



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
20.08.21

Reference N411553EP	Application No./Patent No. 18844651.2 - 1230 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication pursuant to Rules 70(2) and 70a(2) EPC

A supplementary European search report has been drawn up concerning the above-identified European patent application (publication number: 3665553).

Since the request for examination has been filed (R. 70(1), 159(1)(f), Art. 94(1) EPC) prior to the transmission of the supplementary European search report, you are hereby invited to indicate within

six months

of notification of this communication whether you wish to proceed further with the European patent application.

If you do not indicate in due time that you wish to proceed further with the European patent application, it **will be deemed to be withdrawn** (R. 70(3) EPC).

You are invited, within the above-mentioned six-month period, to comment on the objections raised in the opinion accompanying the European search report and/or to file any amendments to the description, claims and drawings correcting any deficiencies noted in the opinion (R. 70a(2), R. 137(2) EPC; Guidelines for Examination in the EPO, B-XI, 8).

If filing amendments, you must identify them and indicate the basis for them in the application as filed. Failure to meet either requirement may lead to a communication from the Examining Division requesting that you correct this deficiency (R. 137(4) EPC).

Should the reply to the invitation pursuant to Rule 70a(2) EPC be filed in an admissible non-EPO language, a translation is to be submitted within one month of its filing (R. 6(2) EPC).

Should you not comply with this invitation within the time limit, the application **will be deemed to be withdrawn** in accordance with Rule 70a(3) EPC.

Receiving Section



Registered letter

EPO Form 1224 11.12 (RTB/ESOP01=N) DMEX

page 1 of 1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization

International Bureau

(43) International Publication Date
14 February 2019 (14.02.2019)



(10) International Publication Number
WO 2019/032604 A1

(51) International Patent Classification:

G06F 3/01 (2006.01) H04N 7/14 (2006.01)
G06K 9/00 (2006.01) H04N 13/00 (2018.01)
H04N 5/228 (2006.01)

(21) International Application Number:

PCT/US2018/045648

(22) International Filing Date:

07 August 2018 (07.08.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/542,661 08 August 2017 (08.08.2017) US
62/543,587 10 August 2017 (10.08.2017) US

(71) Applicant: **REALD SPARK, LLC** [US/US]; Suite 200,
100 N. Crescent Drive, Beverly Hills, California 90210
(US).

(72) Inventors: **SOMMERLADE, Eric**; Suite 200, 100 North
Crescent Drive, Beverly Hills, California 90210 (US).
NEOPHYTOU, Alexandros; 100 N. Crescent Drive, Suite
200, Beverly Hills, California 90210 (US).

(74) Agent: **LOWRY, Penny L.** et al.; 5700 Flatiron Parkway,
Boulder, Colorado 80301 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KR,
KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

(54) Title: ADJUSTING A DIGITAL REPRESENTATION OF A HEAD REGION

(57) Abstract: Methods and devices for generating reference data for adjusting a digital representation of a head region, and methods and devices for adjusting the digital representation of a head region are disclosed. In some arrangements, training data are received. A first machine learning algorithm generates first reference data using the training data. A second machine learning algorithm generates second reference data using the same training data and the first reference data generated by the first machine learning algorithm.

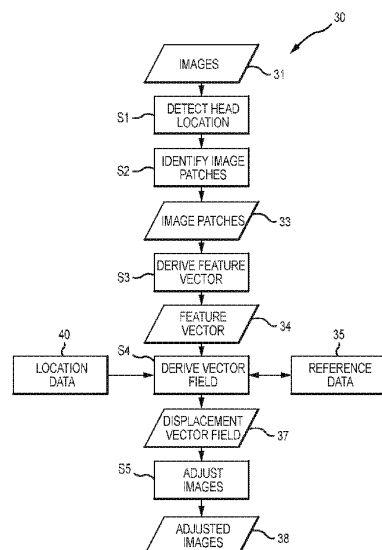


FIG.3

[Continued on next page]

Declarations under Rule 4.17:

- *as to the identity of the inventor (Rule 4.17(i))*
- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*

Adjusting a Digital representation of a Head Region

TECHNICAL FIELD

[01] This application relates to adjusting a digital representation, such as an image or a three-dimensional geometrical representation, of a head region, particularly a facial and/or neck region. The application relates particularly to adjusting target features of the digital representation of the head region, for example to correct a perceived gaze direction of eyes, or to modify the texture and/or shape of features such as the nose, mouth, chin or neck.

BACKGROUND

[02] In many systems, images of a head, which may comprise single images at different times, or stereoscopic pairs of images or other multi-view images, may be captured in one device and displayed on a different device for viewing by an observer. One non-limitative example is a system for performing teleconferencing between two telecommunications devices. In that case, each device may capture images of the observer of that device and transmit them to the other device over a telecommunications network for display and viewing by the observer of the other device. Digital representations other than images and/or complementary to images may also be captured, for example using depth measurements (e.g. using a time-of-flight camera).

[03] When an image or other digital representation of a head is captured and displayed, the gaze of the head may not be directed at the observer. This may be caused for example by the gaze of the head not being directed at the sensing system (e.g. camera system) used to capture the digital representation (e.g. image), for example because a user whose head is imaged is observing a display in the same device as a camera system and the camera system is offset above (or below) that display. In that case, the gaze in the displayed images will be perceived to be downwards (or upwards). The human visual system has evolved high sensitivity to gaze during social interaction, using cues gained from the relative position of the iris and white sclera of other observers. Errors in the perceived gaze are disconcerting. For example in a system for performing teleconferencing, errors in the perceived gaze can create unnatural interactions between the users.

[04] The present disclosure is concerned with processing techniques (e.g. image processing techniques) for adjusting digital representations (e.g. images) of a head region to

correct the perceived gaze and/or to improve other aspects of a computer-generated display of the head region. The present disclosure is particularly concerned with implementing such processing techniques with minimal demands on computer hardware and/or power such that they provide results at or near input data frame rate or user feedback requirements.

BRIEF SUMMARY

[05] According to a first aspect of the present disclosure, there is provided a method of generating reference data for adjusting a digital representation of a head region, the method comprising: receiving training data comprising: a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region; using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

[06] The described use of two machine learning algorithms allows an improved balance to be achieved between accuracy of the adjustment process and speed of execution. The first machine learning algorithm can be configured to provide highly detailed first reference data, which provides high accuracy. Use of this first reference data directly in a processing technique (e.g. image processing technique) to adjust a digital representation (e.g. image or three-dimensional geometrical representation) of a head region would be relatively expensive in terms of computational resources because of the high level of detail. By arranging instead for the first reference data to be provided to a second machine learning algorithm, which provides, based on the first reference data and the original training data, the reference data (the second reference data) that is to be used for the adjustment of the digital representation of the head region, it is possible to benefit to an extent from the high accuracy of the first

machine learning algorithm whilst also providing reference data that is less detailed and thus easier to process efficiently when performing the adjustment of the digital representation of the head region. The quality of the reference data provided by the second machine learning algorithm is found to be significantly improved when the second machine learning algorithm is provided with both the first reference data and the training data in comparison to when the second machine learning algorithm is provided only with the training data.

[07] In an embodiment, efficiency is further improved by providing editing instructions (e.g. image editing instructions) in the second reference data in a compressed representation. The use of a compressed representation reduces data storage and bandwidth requirements during use of the editing instructions to perform adjustment of a digital representation of a head region.

[08] According to a second aspect of the present disclosure, there is provided a device configured to perform a similar method to the first aspect of the invention.

[09] According to a third aspect of the present disclosure, there is provided a method of adjusting a digital representation of a head region, the method comprising: identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region; deriving a feature vector from plural local descriptors of the target patch; using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein the editing instructions in the reference data are provided in a compressed representation.

[10] According to a fourth aspect of the present disclosure, there is provided a device configured to perform a similar method of the third aspect of the invention.

[11] The use of a compressed representation reduces data storage and bandwidth requirements.

[12] According to a fifth aspect of the present disclosure, there is provided a method of training a machine learning algorithm to adjust a digital representation of a head region, comprising: receiving training data comprising: a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and first reference data generated by a pre-trained first machine learning algorithm, the first reference data comprising a set of editing instructions in one-to-one

correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region; updating a pre-trained second machine learning algorithm trained to generate second reference data, where the input for the updating comprises the training data and the generated first reference data, the second reference data comprising editing instructions for adjusting the digital representation of the head region.

[13] According to a sixth aspect of the present disclosure, there is provided a method of training a machine learning algorithm to adjust a digital representation of a head region, the method comprising: receiving training data comprising a set of input digital representations of a head region; training a first machine learning algorithm using the training data to perform an adjustment of a digital representation of a head region; using the trained first machine learning algorithm to generate first reference data, the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations, each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform; and training a second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm.

BRIEF DESCRIPTION OF THE DRAWINGS

[14] Non-limitative embodiments are illustrated by way of example in the accompanying figures, in which like reference numbers indicate similar parts, and in which:

[15] **FIGURE 1** is a schematic perspective view of a device that captures a stereoscopic pair of images;

[16] **FIGURE 2** is a schematic perspective view of a device that displays the stereoscopic pair of images;

[17] **FIGURE 3** is a flow chart of a method of adjusting a stereoscopic pair of images;

[18] **FIGURE 4** is a diagram illustrating the processing of the stereoscopic pair of images in the method of **FIGURE 3**;

[19] **FIGURE 5** is a flow chart of a step of extracting an image patch;

[20] **FIGURE 6** and **FIGURE 7** are flow charts of two alternatives for a step of adjusting an image;

[21] **FIGURE 8** is flow chart of a method of generating reference data;

- [22] **FIGURE 9** schematically depicts data flow in an example method of generating reference data;
- [23] **FIGURE 10** depicts a device for generating reference data;
- [24] **FIGURE 11** is a flow chart of a method of adjusting an image of a head region;
- [25] **FIGURE 12** schematically depicts data flow in an example of a method of adjusting an image of a head region;
- [26] **FIGURE 13** is a diagram of a telecommunications system in which the method may be implemented;
- [27] **FIGURE 14** schematically depicts data flow in an example method of generating reference data for converting a two-dimensional digital representation of a head region to a three-dimensional digital representation of a head region; and
- [28] **FIGURE 15** schematically depicts data flow in an example of a method of adjusting a digital representation of a head region in which a second machine learning algorithm is updated online.

DETAILED DESCRIPTION

[29] **FIGURE 1** and **FIGURE 2** illustrate how incorrect gaze is perceived when a stereoscopic pair of images of a head is captured by the device 10 shown in **FIGURE 1** which will be referred to as the source device 10, and displayed on a different device 20 shown in **FIGURE 2** which will be referred to as the destination device 20. Capturing of a stereoscopic pair of images is shown as an example. A similar effect can occur when a monocular image is captured and when more than two views of the head are captured. A similar effect can also occur when alternative or additional sensing techniques are used to build a digital representation of the head (e.g. where a depth sensor such as a time-of-flight camera is used to obtain three-dimensional geometrical information about positions on the surface of the head).

[30] In the embodiment shown, the capture device 10 includes a display 11 and a camera system 12. In this particular example, the camera system comprises two cameras 13 in order to capture the stereoscopic pair of images of the head of a source observer 14. In a monocular implementation a single camera may be provided instead of the two cameras 13. In other implementations a depth sensor is alternatively or additionally provided. The source observer 14 views the display 11, along line 15. The cameras 13 (optionally including one or more depth sensors) of the camera system 12 are offset from the display 11, in this case being

above the display 11. Thus, the cameras 13 effectively look down on the source observer 14 along line 16.

[31] The display device 20 includes a display 21, which in this example can be a stereoscopic display of any known type, for example an autostereoscopic display of any known type. The display 21 displays the stereoscopic pair of images as captured by the capture device 10. A destination observer 24 views the display 21. If the destination observer 24 is located in a normal viewing position perpendicular to the center of the display 21, as shown by the hard outline of the destination observer 24, then the gaze of the source observer 14 is perceived by the destination observer 24 to be downwards, rather than looking at the destination observer 24, because the cameras 13 of the source device 10 look down on the source observer 14.

[32] Although the cameras 13 are above the display 11 in this example, the cameras 13 could in general could be in any location adjacent the display 11, and the gaze of the source observer 14 perceived by the destination observer 24 would be correspondingly incorrect.

[33] If the destination observer 24 is located in an offset viewing position, as shown by the dotted outline of the destination observer 24 so that the destination observer 24 views the display 21 along line 26, then the offset of the destination observer 24 creates an additional error in the gaze of the source observer 14 perceived by the destination observer 24. A similar additional error in the perceived gaze of the source observer 14 occurs if the destination observer 24 is located in the normal viewing position along line 25, but the displayed image (or stereoscopic pair of images in this example) is displayed on the display 25 in a position offset from the center of the display 25.

[34] A stereoscopic pair of images is an example of multi-view images where there are two images. Although FIGURE 1 illustrates an example where the camera system 12 includes two cameras 13 that capture a stereoscopic pair of images, alternatively the camera system may include more than two cameras 13 that capture more than two multi-view images, in which case similar issues of incorrect perceived gaze exist on display. The camera system may alternatively include only one camera and/or one or more depth sensors.

[35] FIGURE 3 illustrates a method of adjusting multi-view images to correct such errors in the perceived gaze. The method of FIGURE 3 is a specific example of a method of adjusting digital representations of a head region in a case where the digital representations comprise images of the head region and where the images comprise one or more multi-view images of the head region. For simplicity, this method will be described with respect to the

adjustment of multi-view images comprising a stereoscopic pair of images. The method may be generalized to multi-view images comprising more than two images, simply by performing similar processing on a larger number of images. The method may also be generalized to the case where single view (monocular) images of the head region are used and to the case where information from other sensing modalities, such as depth measurements, is included within the digital representations of the head region.

[36] The method may be performed in an image processor 30 (or other processor). The image processor 30 may be implemented by a processor executing a suitable computer program or by dedicated hardware or by some combination of software and hardware. Where a computer program is used, the computer program may comprise instructions in any suitable language and may be stored on a computer readable storage medium, which may be of any type, for example: a recording medium which is insertable into a drive of the computing system and which may store information magnetically, optically or opto-magnetically; a fixed recording medium of the computer system such as a hard drive; or a computer memory.

[37] The image processor 30 (or other processor) may be provided in the source device 10, the destination device 10 or in any other device, for example a server on a telecommunications network, which may be suitable in the case that the source device 10 and the destination device 10 communicate over such a telecommunications network.

[38] In this example, a stereoscopic pair of images 31 are captured by the camera system 12. Although the camera systems 12 is illustrated in FIGURE 1 as including two cameras 13, this is not limitative and more generally the camera system 13 may have the following properties.

[39] The camera system comprises a set of one or more cameras 13, with at least two cameras 13 in the case where multi-view images are processed. Where two cameras are provided, the cameras are typically spaced apart by a distance less than the average human intrapupilar distance. In the alternative that the method is applied to more than two multi-view images, then there are more than two cameras 13, that is one camera 13 per image. In some embodiments, a depth sensor is provided for obtaining three-dimensional geometrical information about a surface of the head region, optionally in addition to one or more other cameras (e.g. optical cameras). The depth sensor may comprise a time-of-flight camera.

[40] Where plural cameras 13 are provided, the cameras 13 may be spatially related to each other and the display 11. The spatial relationship between the cameras 13 themselves and between the cameras 13 and the display 11 is known in advance. Known methods for

finding the spatial relationship may be applied, for example a calibration method using a reference image, or specification *a priori*.

[41] The camera or cameras 13 face in the same direction as the display 11. Thus, when the source observer 14 is viewing the display 11, then the camera or cameras 13 face the source observer 14 and the captured information, such as depth information, image or images (e.g. stereoscopic pair of images) are digital representations (e.g. images and/or three-dimensional geometrical representations) of the head of the source observer 14. Different cameras in the camera system can have different fields of view.

[42] The camera system 12 may include cameras 13 having different sensing modalities, including but not limited to visible light, infrared, and time-of-flight (depth).

[43] In some embodiments, the main output of the camera system 13 is images 31 which are typically video images output at a video rate. The output of the camera system 13 may also include data representing the spatial relationship between the cameras 13 and the display 11, the nature of the sensing modalities and internal parameters of the cameras 13 (for example focal length, optical axis) which may be used for angular localization, as well as three-dimensional geometrical information, for example from depth measurements.

[44] An example of the method performed on a digital representation of a head region comprising a stereoscopic pair of images 31, for the case of adjustment of eye regions, is as follows. To illustrate this example method, reference is also made to **FIGURE 4** which shows an example of the stereoscopic pair of images 31 at various stages of the method.

[45] In step S1, the stereoscopic pair of images 31 are analyzed to detect the location of the head and in particular the eyes of the source observer 14 within the stereoscopic pair of images 31. This is performed by detecting presence of a head, tracking the head, and localizing the eyes of the head. Step S1 may be performed using a variety of techniques that are known in the art.

[46] One possible technique for detecting the presence of the head is to use Haar feature cascades, for example as disclosed in Viola and Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", CVPR 2001, pp 1-9, which is herein incorporated by reference in its entirety.

[47] One possible technique for tracking the head is to use the approach of Active Appearance Models to provide the position of the head of the subject, as well as the location of the eyes, for example as disclosed in Cootes et al., "Active shape models - their training and application", Computer Vision and Image Understanding, 61(1):38-59, Jan. 1995 and in

Cootes et al. "Active appearance models", IEEE Trans. Pattern Analysis and Machine Intelligence, 23(6):681-685, 2001, both of which are herein incorporated by reference in their entireties.

[48] In step S1, typically, a set of individual points ("landmarks") are set to regions of the face, typically the eyes, for example corners of the eye, upper and lower lid locations, etc., thereby localizing the eyes.

[49] In step S2, patches representing portions of a digital representation of the head region, which in this example may be referred to image patches, containing the left and right eyes of the head, respectively, are identified in each image 31 of the stereoscopic pair. FIGURE 4 shows the identified image patches 32 of the right eye in each image 31 (the image patches for the left eye being omitted in FIGURE 4 for clarity).

[50] Step S2 may be performed as shown in FIGURE 5, as follows.

[51] In step S2-1, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, as shown for example in FIGURE 4.

[52] In step S2-2, the image patches 32 identified in step S2-1 are transformed into a normalized coordinate system, being the same normalized coordinate system as used in the machine learning process which is described further below. The transformation is chosen to align the points ("landmarks") of the eye within the image patch that were identified in step S1, with predetermined locations in the normalized coordinate system. The transformation may include translation, rotation and scaling, to appropriate extents to achieve that alignment. The output of step S2-2 is identified image patches 33 of the right eye in each image in the normalized coordinate system as shown for example in FIGURE 4.

[53] The following steps may be performed separately (a) in respect of the image patches containing the left eyes of the head in each image 31 of the stereoscopic pair, and (b) in respect of the image patches containing the right eyes of the head in each image 31 of the stereoscopic pair (in this example). For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes.

[54] In step S3, a feature vector 34 is derived from plural local descriptors (representing information about a local region in a patch), which in this example may be referred to as local image descriptors, of an image patch 33 in at least one image 31 of the stereoscopic pair (in

this example). Depending on the approach and as described further below, this may be an image patch in a single image 31 of the stereoscopic pair or may be both images 31 of the stereoscopic pair. Thus, the local image descriptors are local image descriptors derived in the normalized coordinate system.

[55] The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 to be used for adjusting the image patches. The reference data 35 may comprise reference displacement vector fields that represent transformations of the image patch, or other representations of transformations of the image patch, including compressed representations as described below, and are associated with possible values of the feature vector.

[56] The reference data 35 is obtained and analyzed in advance using a machine learning technique. The machine learning technique may derive the form of the feature vectors 34 and associate transformations such as the reference displacement vector fields with the possible values of the feature vector. A specific example of a machine learning technique applied in the case where it is desired to correct gaze using digital representations of a head region comprising images of the head region, will now be described before reverting to the method of FIGURE 3.

[57] The training input to the machine learning technique is two sets of images (or image patches), which may be stereoscopic pairs of images or monoscopic images, as discussed further below. Each set may comprise images of the head of the same group of individuals but captured from cameras in different locations relative to the gaze so that the perceived gaze differs as between them (in the case where gaze is to be corrected).

[58] The first set are input images, being images of each individual with an incorrect gaze where the error is known *a priori*. In particular, the images in the first set may be captured by at least one cameras in a known camera location where the gaze of the individual which is in a different known direction. For example in the case of the source device of FIGURE 1, the camera location may be the location of a camera 13 and while the gaze of the imaged individual is towards the center of the display 11.

[59] The second set are output images, being images of each individual with correct gaze for a predetermined observer location relative to a display location in which the image is to be displayed. In the simplest case, the observer location is a normal viewing position perpendicular to the center of the display location, for example as shown by the hard outline of the destination observer 24 in the case of the destination device 20 of FIGURE 2.

[60] For each image in the two sets, the image is analyzed to detect the location of the head and in particular the eyes using the same technique as used in step S1 described above, and then image patches containing the left and right eyes of the head, respectively, are identified using the same technique as used in step S2 described above. The following steps may then be performed separately (a) in respect of the image patches containing the left eyes of the head in each image, and (b) in respect of the image patches containing the right eyes of the head in each image. For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes in this embodiment.

[61] Each image patch is transformed into the same normalized coordinate system as used in step S2 described above. As described above, the transformation is chosen to align points ("landmarks") of the eye with predetermined locations in the normalized coordinate system. The transformation may include, for example, translation, rotation and/or scaling, to appropriate extents to achieve that alignment.

[62] Thus, the image patches input and output images of each individual are aligned in the normalized coordinate system.

[63] From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image, for example as follows. Defining positions in the image patches by (x,y) , the displacement vector field $\mathbf{F}$ is given by

$$\mathbf{F}=\{u(x, y), v(x, y)\}$$

where u and v define the horizontal and vertical components of the vector at each position (x,y) .

[64] The displacement vector field $\mathbf{F}$ is chosen so that the image patch of the output image $O(x,y)$ is derived from the image patch of the input image $I(x,y)$ as

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[65] For image data from more than one camera, the system delivers a displacement vector field for the input image from each camera.

[66] The displacement vector field $\mathbf{F}$ for an input and output image of an individual may be derived using a process in which a trial feature vector $\mathbf{F}'=\{u',v'\}$ is modified to minimize error, optionally in an iterative process, for example in accordance with:

$$\sum |O(x, y) - I(x + u'(x, y), y + v'(x, y))| = \min!$$

[67] By way of non-limitative example, the displacement vector field $\mathbf{F}$ may be derived as

disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675, which is herein incorporated by reference in its entirety, wherein the displacement vector field $\mathbf{F}$ is referred to as a “flow field”.

[68] Another example for editing instructions, which may be referred to as image editing instructions in cases where the digital representation of the head region consists of images, which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by filter field $L = \{k(P, x, y)\}$, which defines a filter kernel for a given location (x, y) . This filter field L is chosen so that the image patch of the output image $O(x, y)$ is derived from the image patch of the input image $I(x, y)$ as $O(x, y) = k(P(I, x, y), x, y)$, where $P(I, x, y)$ is a local region around the point (x, y) in the image I , and $k(P, x, y)$ operates on the patch P with coefficients depending on the position (x, y) . For example, it could be the output of a convolution of the patch with a Gaussian filter with width depending on the position x in the image, or a brightness increase of a local pixel depending on the vertical position y .

[69] Another example for editing instructions (e.g. image editing instructions) which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by a set of typical image components that make up the edited image area, e.g. texture showing beard stubbles. These are then blended with a factor depending on the image coordinates and local image content (i.e. a texture blending field). Other transformation fields can be used, such as a brightness adjustment field.

[70] A machine learning technique is used to obtain a map from the displacement vector field $\mathbf{F}$ (or other editing instructions such as image editing instructions) of each individual to respective feature vectors derived from plural local (e.g. image) descriptors of a target patch of an input image.

[71] The local descriptors (e.g. local image descriptors) capture relevant information of a local part of a patch (e.g. image patch) of an input digital representation of the head region (e.g. an input image) and the set of local descriptors usually form a continuous vectorial output.

[72] The local descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific image descriptors are selected and optimized by the machine learning process itself. In general, the local descriptors may be of any suitable type, some non-limitative examples which may be applied

in any combination being as follows.

[73] The local descriptors may include values of individual pixels or a linear combination thereof. Such a linear combination may be, for example, a difference between the pixels at two points, a kernel derived within a mask at an arbitrary location, or a difference between two kernels at different locations.

[74] The local descriptors may include distances of a pixel location from the position of an eye point (“landmark”).

[75] The local descriptors may include SIFT features (Scale-invariant feature transform features), for example as disclosed in Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision 60 (2), pp 91–110, which is herein incorporated by reference in its entirety.

[76] The local descriptors may include HOG features (Histogram of Oriented Gradients features), for example as disclosed in Dalal et al. “Histograms of Oriented Gradients for Human Detection”, Computer Vision and Pattern Recognition, 2005, pp. 886-893, which is herein incorporated by reference in its entirety.

[77] The local descriptors may include “low level representations” from pre-classification stages in deep learning neural networks, for example as disclosed in Yang and Ramanan, “Multi-scale recognition with DAG-CNNs”, ICCV 2015, which is herein incorporated by reference in its entirety. In a classifying deep learning neural network with multiple layers applied to an input digital representation (e.g. image), for example, such low level features could be taken from a layer before the final classification layer of the network.

[78] The derivation of the feature vector from plural local descriptors depends on the type of machine learning applied.

[79] In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local descriptors (e.g. local image descriptors) in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field $\mathbf{F}$ derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 may provide similar functionality to a look-up table, with the machine learning generating a machine learning parameter set that can be used to generate corresponding editing instructions. This allows a reference displacement vector field $\mathbf{F}$ to be simply selected from the reference data 35 on the basis of the feature vector 34 derived in step S3, as described below.

[80] In the case that the feature vector comprises features that are binary values derived from the local descriptors, the feature vector has a binary representation. Such binary values may be derived in various ways from the values of descriptors, for example by comparing the value of a descriptor with a threshold, comparing the value of two descriptors, or by comparing the distance of a pixel location from the position of an eye point (“landmark”).

[81] Alternatively, the feature vector may comprise features that are discretized values of the local descriptors. In this case, more than two discrete values of each feature are possible.

[82] Any suitable machine learning technique may be applied, for example using a decision tree, a decision forest, a decision fern or an ensemble or combination thereof, or a neural network.

[83] By way of example, a suitable machine learning technique using a feature vector comprising features that are binary values derived by comparing a set of individual pixels or a linear combination thereof against a threshold, is disclosed in Ozuysal et al. “Fast Keypoint Recognition in Ten Lines of Code”, Computer Vision and Pattern Recognition, 2007, pp. 1-8, which is herein incorporated by reference in its entirety.

[84] By way of further example, a suitable machine learning technique using a distance of a pixel location with the position of an eye landmark is disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675, which is herein incorporated by reference in its entirety.

[85] By way of further example, a suitable machine learning technique using a random decision forest is disclosed in Ho, “Random Decision Forests”, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995, pp. 278-282, which is herein incorporated by reference in its entirety.

[86] In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local descriptors (e.g. local image descriptors) in a continuous space. In this case, the machine learning technique associates a reference displacement vector field $\mathbf{F}$ (in this example, but other editing instructions could be used) derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field $\mathbf{F}$ to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived in step S3 and the values of the feature vector associated with the reference displacement vector fields.

[87] Any suitable machine learning technique may be applied, for example using support vector regression.

[88] By way of example, a suitable machine learning technique using support vector regression is disclosed in Drucker et al. "Support Vector Regression Machines", Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, which is herein incorporated by reference in its entirety. The output of the technique is a continuously varying set of interpolation directions that form part of the reference data 35 and are used in the interpolation.

[89] The machine learning technique, regardless of its type, inherently also derives the form of the feature vectors 34 that is used to derive the reference displacement vector fields F (or other image editing instructions). This is the form of the feature vectors 34 that is derived in step S3.

[90] The description now reverts to the method of FIGURE 3.

[91] In step S4, at least one displacement vector field 37 representing a transformation of an image patch is derived by using the feature vector 34 derived in step S3 to look up the reference data 35. Due to the derivation of the displacement vector field 37 from the reference data 35, the transformation represented thereby corrects the gaze that will be perceived when the stereoscopic pair of images 31 are displayed.

[92] In the case that the feature vector 34 comprises features that are values in a discrete space and the reference displacement vector fields of the reference data 35 comprise a reference displacement vector field associated with each possible value of the feature vector in the discrete space, then the displacement vector field for the image patch is derived by selecting the reference displacement field associated with the actual value of the derived feature vector 34.

[93] In the case that the feature vector 34 comprises features that are discrete values of the local descriptors in a continuous space, then then the displacement vector field for the image patch is derived by interpolating a displacement vector field from the reference displacement vector fields based on the relationship between the actual value of the derived feature vector 34 and the values of the feature vectors associated with the reference displacement vector fields. In the case that the machine learning technique was support vector regression, this may be done using the interpolation directions that form part of the reference data 35.

[94] In step S5, each image 31 of the stereoscopic pair is adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived

displacement vector fields 37. This produces an adjusted stereoscopic pair of images 38 as shown in FIGURE 4, in which the gaze has been corrected. In particular, the adjustment may be performed using two alternative methods, as follows.

[95] A first method for performing step S5 is shown in **FIGURE 6** and performed as follows.

[96] In step S5-1, the image patch is transformed in the normalised coordinate system in accordance with the corresponding displacement vector field 37 in respect of the same image, thereby correcting the gaze. As described above, for a displacement vector field $\mathbf{F}$ the transformation of the image patch of the input image $I(x,y)$ provides the output image $O(x,y)$ in accordance with

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[97] In step S5-2, the transformed image patch output from step S5-1 is transformed out of the normalised coordinate system, back into the original coordinate system of the corresponding image 31. This is done using the inverse transformation from that applied in step S2-2.

[98] In step S5-3, the transformed image patch output from step S5-2 is superimposed on the corresponding image 31. This may be done with a full replacement within an eye region corresponding to the eye itself, and a smoothed transition between the transformed image patch and the original image 31 over a boundary region around the eye region. The width of the boundary region may be of fixed size or a percentage of the size of the image patch in the original image 31.

[99] A second method for performing step S5 is shown in **FIGURE 7** and performed as follows.

[100] In this second, alternative method, the transformation back into the coordinate system of the corresponding image 31 occurs before the transformation of the image patch in accordance with the transformed displacement vector field $\mathbf{F}$.

[101] In step S5-4, the displacement vector field $\mathbf{F}$ is transformed out of the normalised coordinate system, back into the original coordinate system of the corresponding image 31. This is done using the inverse transformation from that applied in step S2-2.

[102] In step S5-5, the image patch 32 in the coordinate system of the image 31 is transformed in accordance with the displacement vector field $\mathbf{F}$ that has been transformed into the same coordinate system in step S5-4. As described above, for a displacement vector

field $\mathbf{F}$ the transformation of the image patch of the input image $I(x,y)$ provides the output image $O(x,y)$ in accordance with

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

but this is now performed in the coordinate system of the original image 31.

[103] Step S5-6 is the same as S5-3. Thus, in step S5-6, the transformed image patch output from step S5-5 is superimposed on the corresponding image 31. This may be done with a full replacement within an eye region corresponding to the eye itself, and a smoothed transition between the transformed image patch and the original image 31 over a boundary region around the eye region. The width of the boundary region may be of fixed size or a percentage of the size of the image patch in the original image 31.

[104] **FIGURE 8** depicts a method of generating reference data (including reference data 37 of the type described in the specific examples discussed above) for adjusting a digital representation of a head region. In some embodiments, the digital representation of the head region comprises or consists of an image of the head region. In some embodiments, the digital representation of the head region comprises or consists of a three-dimensional digital representation (representing, for example, three-dimensional geometrical information). The three-dimensional digital representation may be obtained from depth measurements, using for example a time-of-flight camera. In an embodiment, the digital representation of the head region is usable to provide a computer generated display of the head region. In an embodiment, the adjustment of the digital representation comprises converting a two-dimensional digital representation of the head region to a three-dimensional digital representation of the head region.

[105] In a case where the digital representation of the head region (prior to adjustment, after adjustment, or both) comprises a three-dimensional digital representation, this may be provided in any of various known ways. For example, the three-dimensional digital representation may comprise a point cloud, a particle system, or a mesh representation. The mesh representation may comprise one or more of: a polygonal surface, a multi-resolution surface, a subdivision surface. The digital representation may comprise a three-dimensional digital representation and texture information associated with the three-dimensional digital representation, e.g. via a texture map. The digital representation may comprise a three-dimensional geometrical representation and color information (e.g. obtained from an optical camera) aligned with the three-dimensional geometrical representation. Other volumetric representations such as particle system or implicit definitions such as signed distance

functions may also be used.

[106] The method comprises a step S100 of receiving training data 100. The training data 100 may be provided for example via a communications interface 112 (e.g. connecting to an external data connection or storage device) connected to a data processing unit 120 that is to perform the method (as depicted schematically in FIGURE 10). The training data 100 comprises a set of input patches. A patch consists of a target portion of a digital representation of a head region. In a case where the digital representation comprises an image, the patch may be referred to as an image patch. Each input patch (e.g. input image patch) comprises a target feature of the digital representation (e.g. image) of the head region prior to adjustment of the digital representation (e.g. adjustment of the image to be displayed) of the head region. The target feature is the same for each input patch. The target feature may comprise one or more of the following: an eye region comprising at least part of an eye (as in the specific examples discussed above with reference to FIGURE 4 for example), a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, and a neck region comprising at least part of a neck. In an embodiment each input image patch comprises a portion of the image of the head region corresponding to the target feature, such as the above-mentioned eye region, nose region, mouth region, chin region, or neck region.

[107] In an embodiment, the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation (e.g. image) of the head region comprises adjusting a gaze direction.

[108] In an embodiment, the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation (e.g. image) of the head region comprises adjusting a shape and/or texture of the nose (e.g. to make nose look smaller and/or slimmer by a fixed proportion, similar to the effect of a “tele lens”).

[109] In an embodiment, the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation (e.g. image) of the head region comprises adjusting a shape and/or texture of the chin (e.g. to reduce or remove double chin appearance).

[110] In an embodiment, the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation (e.g. image) of the head region comprises adjusting a shape and/or texture of the neck (e.g. to reduce or remove wrinkles).

[111] In an embodiment, the target feature comprises a hair region comprising hair and the

adjustment of the digital representation (e.g. image) of the head region comprises adjusting a color of the hair (e.g. by a fixed hue).

[112] In the case where the target feature comprises an eye region, the input patches may be obtained using the methodology described above with reference to steps S2-1 and S2-2, except that stereoscopic pairs of images are not necessarily required. Thus, the input patches may be obtained by using identified points (“landmarks”) corresponding to features of an eye to locate the relevant region of the digital representation (e.g. image) of the head region (e.g. surrounding the eye) and/or transformation of the input patch into a normalized coordinate system, including alignment of the landmarks with predetermined locations in the normalized coordinate system using translation, rotation and/or scaling.

[113] In some embodiments, the training data 100 further comprises a set of output patches. As described above, each patch consists of a target portion of a digital representation of a head region. In a case where the digital representation comprises an image, the patch may be referred to as an image patch. The output patches (e.g. output images patches) are in one-to-one correspondence with the input patches. Each output patch comprises the target feature of the digital representation (e.g. image) of the head region after adjustment of the digital representation (e.g. adjustment of the image to be displayed) of the head region. Thus, in the case where an image of the head region is to be adjusted to correct a gaze direction, each output patch comprises an eye region that has been adjusted so that the gaze appears to be in the desired direction (e.g. directly towards a destination observer 24).

[114] The method further comprises a step S101 comprising using a first machine learning algorithm to generate first reference data 101 using the training data 100. The first reference data 101 comprises editing instructions (e.g. image editing instructions) for adjusting the digital representation (e.g. image) of the head region for a range of possible digital representations (e.g. digital representations representing different states of the head, such as different positions and/or orientations, optionally represented as different images) of the head region.

[115] The method further comprises a step S102 comprising using a second machine learning algorithm to generate second reference data 102. The second machine learning algorithm uses the same training data 100 as the first machine learning algorithm in step S101. The second machine learning algorithm further uses the first reference data 101 output by the first machine learning algorithm in step S101. The second reference data 102 comprises editing instructions (e.g. image editing instructions) for adjusting the digital

representation (e.g. image) of the head region for a range of possible digital representations (e.g. different positions and/or orientations, optionally represented as different images) of the head region.

[116] In an embodiment, the first reference data 101 comprises first editing instructions (e.g. first image editing instructions) for a range of possible configurations of the target feature (e.g. different gaze directions and/or eye morphologies in the case where the target feature comprises an eye region) and first selection instructions for selecting editing instructions (from the first editing instructions) for a particular input patch (e.g. input image patch) based on the configuration of the target feature of the input patch (e.g. the particular gaze direction and/or particular eye morphology of that input patch).

[117] In an embodiment, the second reference data comprises second editing instructions (e.g. second image editing instructions) for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions (from the second editing instructions) for a particular input patch (e.g. input image patch) based on the configuration of the target feature of the input patch.

[118] The configuration of the target feature of each input patch may be represented by a feature vector derived from plural local descriptors (e.g. local image descriptors) of the input patch, as described above with reference to step S3 of FIGURE 3 for the particular case where stereoscopic pairs of images containing eye regions are processed (but the method is applicable more generally than this particular case). As described above, the feature vector may take various forms but is generally adapted to be suitable for looking up editing instructions for performing adjustment of the digital representation (e.g. image) of the head region. In the present embodiment, the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch. In an embodiment, the editing instructions comprise a displacement vector field defining how the input patch is to be transformed to perform the adjustment. The displacement vector field may take any of the forms discussed above. The editing instructions are not limited to displacement vector fields, however. Other editing operations may additionally or alternatively be associated with the features vectors to perform other desired adjustments of the digital representation (e.g. image) of the head region, including for example adjustments to pixel colors or intensities, or changes to underlying geometries (e.g. via a filter field, brightness adjustment field, or texture blending field).

[119] In an embodiment, a first editing algorithm (e.g. first image editing algorithm) is used

by the first machine learning algorithm to define how the first editing instructions are to be applied to an input patch to derive an output patch. A second editing algorithm (e.g. second image editing algorithm) is used by the second machine learning algorithm to define how the second editing instructions are to be applied to an input patch to derive an output patch. The first and second editing algorithms may comprise any of the methods described above for implemented step S5 of FIGURE 3, described with reference to FIGURES 6 and 7.

[120] In an embodiment, the second editing instructions in the second reference data are provided in a compressed representation. The compressed representation may comprise a principle component analysis representation or a wavelet representation for example. In this case, the first and second editing algorithms may be adapted to define how to operate efficiently in this context.

[121] In an embodiment, the second editing instructions are principle component analysis components of a principle component analysis of the first editing instructions. The second editing algorithm in this case transforms the second editing instructions into the first editing instructions by inverse principle component analysis transform.

[122] In an alternative approach, the second editing instructions are wavelet components of the first editing instructions. The second editing algorithm in this case transforms the second editing instructions into the first editing instructions by inverse wavelet transform.

[123] In an embodiment, the first selection instructions for the first reference data are able to select between a larger number of alternative editing instructions (e.g. image editing instructions) than the second selection instructions for the second reference data. The first machine learning algorithm may thus be described as having more input parameters than the second machine learning algorithm. The first machine learning algorithm may provide higher accuracy than the second machine learning algorithm but will typically operate considerably slower. Additionally, the first selection instructions may be significantly more complex (e.g. involving linear algebra or other relatively computer intensive operations) than the second selection instructions (where the selection instructions may resemble a computationally straightforward look-up table, a combination of a look-up table and a tree structure, or similar).

[124] In one particular embodiment, the first machine learning algorithm comprises a neural network (known to provide relatively high accuracy, but at the expense of relatively high computational demands). In such an embodiment, and others, the second machine learning algorithm may comprises a regression forest (known to provide higher computational

efficiency, but at the expense of reduced accuracy). The inventors have found that the combination of the two different machine learning algorithms provides reference data that can be used in an adjustment method with high efficiency while still achieving high adjustment accuracy. The first machine learning algorithm may alternatively comprise a support vector machine or a generative adversarial network (GAN). The second machine learning algorithm may alternatively comprise regression ferns, cluster centres, a lookup table, or separable filter banks. In one embodiment, the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.

[125] **FIGURE 9** schematically depicts data flow in a detailed example of the method of generating reference data of **FIGURE 8**. The first and second machine learning algorithms are respectively labelled MLA1 and MLA2. The first machine learning algorithm MLA1 receives the training data (labelled TD) and, optionally, the first editing algorithm EA1, and a similarity metric SM. The similarity metric SM provides a numerical value to measure similarity between an adjusted image and a desired image and can be used to control the first machine learning algorithm MLA1 and the second machine learning algorithm MLA2 to vary the extent to which differences are penalized according to the nature of the differences. For example, the similarity metric SM may be configured to penalize reductions in portions of images that it is desired to maintain (e.g. sclera in the case where eye regions are being adjusted) or deviations of adjusted features from a known form (e.g. deviations in the shape of an iris from an elliptical form) or from a form that is observed in training data. Alternatively or additionally, the training data may comprise an average absolute or square difference between the adjusted image and a target, or average absolute or square difference between low level representations of the adjusted image and target, such as low level features from a deep learning network (as discussed above). In the example shown, the first editing algorithm EA1 receives auxiliary data AuxD, which defines a basis set used for providing a compressed representation of image editing instructions. The first machine learning algorithm MLA1 outputs first selection instructions LA1 and first editing instructions ED1. The second machine learning algorithm MLA2 receives the same training data TD and, optionally, the second editing algorithm EA2, and the similarity metric SM. The second machine learning algorithm MLA2 additionally receives the first editing instructions ED1. The second machine learning algorithm outputs second selection instructions LA2 and

second editing instructions ED2.

[126] The second machine learning algorithm MLA2 thus gets the editing instructions to match or to approximate, and does not have to infer these from the matched input images as MLA1.

[127] **FIGURE 10** depicts a device 110 for generating the reference data. The device 110 comprises a data processing unit 120 configured to perform the method of generating the reference data according to any of the disclosed embodiments. The data processing unit 110 may be implemented by a processor executing a suitable computer program or by dedicated hardware or by some combination of software and hardware. Data input/output may be provided via a communications interface 112. Where a computer program is used, the computer program may comprise instructions in any suitable language and may be stored on a computer readable storage medium, which may be of any type, for example: a recording medium which is insertable into a drive of the computing system and which may store information magnetically, optically or opto-magnetically; a fixed recording medium of the computer system such as a hard drive; or a computer memory.

[128] **FIGURE 11** depicts a method of adjusting a digital representation of a head region. In the example shown, the digital representation comprises an image of a head region, but the method can be adapted to use any of the digital representations discussed above (e.g. to additionally or alternatively process three-dimensional digital representations). The method may use reference data generated using any of the methods of generating reference data disclosed herein or may use reference data generated using other methods.

[129] The method comprises a step S200 in which a target patch (in this case an image patch) is identified in an image 200 of a head region that is to be adjusted. The target patch comprises a target feature of the head region. The target feature may take any of the forms discussed above. The step S200 may optionally comprise detecting a head and/or eye location as described above with reference to step S1 of **FIGURE 3**. The step S200 may optionally further comprise identifying image patches using the methodology described above with reference to steps S2-1 and S2-2 of **FIGURE 5**, except that the image patches do not necessarily need to be identified as stereoscopic pairs (although they may be if desired).

[130] In step S201, a feature vector is derived from plural local descriptors (e.g. local image descriptors) of the target (e.g. image) patch. The feature vector may be derived using the methodology described above with reference to step S3 of **FIGURE 3**. The feature vector may take any of the forms discussed above.

[131] In step S202, the feature vector is used to select editing instructions (e.g. image editing instructions) from reference data 102. The reference data 102 comprising editing instructions for a range of possible values of the feature vector (representing for example different gaze directions and/or eye morphologies in the case where the target feature comprises an eye region).

[132] In step S203, the selected editing instructions are applied to the target patch to adjust the image of the head region (e.g. to correct a gaze direction).

[133] In an embodiment, the editing instructions (e.g. image editing instructions) are provided in a compressed representation, comprising for example one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers. The use of a compressed representation reduces data storage and bandwidth requirements during use of the image editing instructions to perform adjustment of the digital representation (e.g. image) of the head region. Alternatively or additionally, the reference data containing the editing instructions may be generated using any of the embodiments disclosed herein. The reference data may comprise the second reference data discussed above for example.

[134] Reference to editing instructions herein is understood to encompass any data which can be used to define how a digital representation (e.g. image) of a head region should be adjusted to achieve a desired aim (e.g. gaze correction or conversion from a two-dimensional digital representation to a three-dimensional digital representation or both). The editing instructions may comprise data that can be used directly to modify a digital representation (e.g. image), such as a vector field, or intermediate data such as a machine learning parameter set that can be used to generate data that can be used directly to modify the digital representation (e.g. image).

[135] **FIGURE 12** schematically depicts data flow in a detailed example of a method of adjusting a digital representation of a head region in the case where the digital representation comprises an image, using reference data generated according to the detailed example of **FIGURE 9**. Input data ID is provided from a sensor system SS (e.g. comprising one or more cameras). The input data ID is input to the second selection instructions LA2 to select editing instructions appropriate to the input data ID from the second editing instructions ED2. The selected editing instructions, which in this example are provided in a compressed representation (e.g. principle component analysis components or wavelet components) from the compressed second editing instructions ED2, are then used by the second editing

algorithm EA2, in combination with the auxiliary data AuxD providing the basis set for the compressed representation, to provide output data OD. The output data OD comprises an adjusted image of the head region and is displayed via a display DS.

[136] As described above, the method of adjusting a digital representation of a head region (e.g. an image of a head region) may be implemented in an image processor 30 provided in various different devices. By way of non-limitative example, there will now be described a particular implementation in a telecommunications system which is shown in **FIGURE 13** and arranged as follows.

[137] In this implementation, the source device 10 and the destination device 10 communicate over such a telecommunications network 50. For communication over the telecommunications network 50, the source device 10 includes a telecommunications interface 17 and the destination device 20 includes a telecommunications interface 27.

[138] In this implementation, the image processor 30 is provided in the source device 10 and is provided with an image of a head region directly from a camera system 12 (in this example, a stereoscopic pair of images). The telecommunications interface 17 is arranged to transmit the adjusted images 38 over the telecommunications network 50 to the destination device 20 for display thereon.

[139] The destination device 20 includes an image display module 28 that controls the display 26. The adjusted images 38 are received in the destination device 20 by the telecommunications interface 27 and supplied to the image display module 28 which causes them to be displayed on the display 26.

[140] The following elements of the destination device 20 are optionally included in the case that the method corrects gaze for a destination observer 24 in an observer location other than a normal viewing position perpendicular to the center of the display location. In this case, the destination device 20 includes a camera system 23 and an observer location module 29. The camera system 23 captures an image of the destination observer 24. The observer location module 29 derives the location data 40. The observer location module 29 includes a head tracking module that uses the output of the camera system 23 to detect the location of the destination observer 24. Where the relative observer location also takes into account the location of the image displayed on the display 21, the observer location module 29 obtains the location of the image displayed on the display 21 from the image display module 28. The telecommunications interface 17 is arranged to transmit the location data 40 over the telecommunications network 50 to the source device 10 for use thereby.

[141] Although the above description refers to a method applied to images supplied from a source device 10 to a destination device 20, the method may equally be applied to images supplied in the opposite direction from the destination device 20 to the source device 10, in which case the destination device 20 effectively becomes the “source device” and the source device 10 effectively becomes the “destination device”. Where images are supplied bi-directionally, the labels “source” and “destination” may be applied to both devices, depending on the direction of communication being considered.

[142] **FIGURE 14** depicts a further embodiment of a method of generating reference data for adjusting a digital representation of a head region using a framework of the type depicted in **FIGURE 8**. In this embodiment, the training data 100 received in step S100 (**FIGURE 8**) comprises a set of input digital representations of a head region (e.g. input patches each consisting of a target portion of a two-dimensional digital representation of the head region, such as a captured image).

[143] In step S101, a first machine learning algorithm MLA1 is trained using the training data 100, the training causing the first machine learning algorithm MLA1 to become capable of performing an adjustment of a digital representation of a head region. In an embodiment, the adjustment of the digital representation comprises converting from a two-dimensional digital representation to a three-dimensional digital representation (e.g. converting from a 2D image of a portion of a head region to a 3D mesh of a portion of the head region). The trained first machine learning algorithm MLA1 is then used to generate first reference data 101. The first reference data 101 comprises an adjusted digital representation of the head region for each of at least a subset of the input digital representations in the training data 100. Each adjusted digital representation is obtained by performing the adjustment that the first machine learning algorithm MLA1 was trained to perform.

[144] In step S102, a second machine learning algorithm MLA2 is trained using at least a subset of the training data 100 used to train the first machine learning algorithm MLA1 and the first reference data 101. The training causes the second machine learning algorithm MLA2 to become capable of performing the same adjustment of a digital representation of a head region as the first machine learning algorithm MLA1.

[145] In the particular embodiment of **FIGURE 14**, the first machine learning algorithm MLA1 comprises a first encoder 306A and a first predictor 308A. The training data 100 is input to the first encoder 306A. In this embodiment, the first encoder 306A comprises a feature extractor algorithm. The feature extractor algorithm derives informative and non-

redundant values from the training data 100 (i.e. extracts meaningful features from the training data 100). Examples of feature extractor algorithms include Convolutional Neural Networks, Principal Component Analysis, SIFT (Scale Invariant Feature Transform). An output from the first encoder 306A is input to a first predictor 308A. The first predictor 308A generates an adjusted version of each input digital representation in the training data 100 based on the output of the first encoder 306A (e.g. features extracted by the first encoder 306A). In this embodiment, each input digital representation comprises a two-dimensional digital representation (e.g. a two-dimensional image) and the adjusted version of each input digital representation comprises a three-dimensional digital representation (e.g. a mesh). Each three-dimensional digital representation is input to a renderer 310. The renderer 310 synthesizes one or more two-dimensional digital representations corresponding to each input three-dimensional digital representation (e.g. one or more photorealistic images of the head region defined by the three-dimensional digital representation). The output from the renderer 310 is then input to a second encoder 306B. The second encoder 306B may be identical to the first encoder 306A. An output from the second encoder 306B is input to a second predictor 308B. The second predictor 308B may be identical to the first predictor 308A. A first regularizer 314 is provided that compares the output from the first encoder 306A with the output from the second encoder 306B and imposes one or more predetermined first constraints. A second regularizer 316 is provided that compares the output from the first predictor 308A with the output from the second predictor 308B and imposes one or more predetermined second constraints. The first regularizer 314 and the second regularizer 316 may use a set of semantically meaningful constraints (examples of the first constraints and second constraints) or additional information to help reach a desirable solution and to prevent overfitting. The constraints may help to ensure that generated three-dimensional digital representations are natural looking for example, by requiring high levels of natural levels of smoothness. Using this architecture, the first machine learning algorithm MLAI iteratively updates properties of the first predictor 308A and the second predictor 308B (e.g. by adjusting parameters such as weights defining the operation of the predictor) to maximize matching between the outputs from the first and second encoders (as measured by the first regularizer 314) subject to the first constraints and to maximize matching between the outputs from the first and second predictors (as measured by the second regularizer 316) subject to the second constraints. In other embodiments, the first encoder 306A and second encoder 306B may also be iteratively updated. The training of the first machine learning algorithm

MLA1 thus comprises iteratively using a rendering process to generate a two-dimensional digital representation from a three-dimensional digital representation generated by the first machine learning algorithm MLA1 and comparing the generated digital representation with a corresponding digital representation in the training data. Once matching has been achieved or a predetermined number of iterations have been performed, the first machine learning algorithm MLA1 is considered trained and the resulting output from the second predictor 308B can be used to provide the first reference data 101 (which in this embodiment comprises three-dimensional digital representations corresponding to the input two-dimensional representations in the training data 100).

[146] As an extension, new two-dimensional and/or three-dimensional digital representations can be generated by applying editing instructions either to two-dimensional digital representations (which are converted to three-dimensional digital representations by the trained first machine learning algorithm MLA1) or to three-dimensional digital representations output by the trained first machine learning algorithm MLA1). In this case, the first reference data 101 may include the two-dimensional digital representations and/or three-dimensional representations after modification by the editing instructions.

[147] A set of thus generated three-dimensional digital representations output as first reference data 101 and corresponding two-dimensional digital representations (e.g. directly from the training data 100 as shown in FIGURE 14) are used to train the second machine learning algorithm MLA2 (input at block 304). In this embodiment, the second machine learning algorithm MLA2 also comprises an encoder 306C and a predictor 308C, which may be configured as described above for the first machine learning algorithm MLA1. The predictor 308C (and, optionally, the encoder 308C) may be iteratively updated (trained) using a regularizer 318 based on the input training data 100 and first reference data 101. Thus, the second machine learning algorithm MLA2 learns to convert between a two-dimensional digital representation and a three-dimensional digital representation based on the mapping between two-dimensional digital representations and three-dimensional digital representations derived using the renderer 310 in the first machine learning algorithm MLA1. Provided the mapping derived by the first machine learning algorithm MLA1 is reliable, the second machine learning algorithm MLA2 will be able to provide accurate conversion between two-dimensional digital representations and three-dimensional digital representations using a simpler trained machine learning model (which can be stored and operated using fewer computing resources than the first machine learning algorithm MLA1).

[148] **FIGURE 15** depicts data flow in an example implementation of an embodiment in which a pre-trained second machine learning algorithm MLA2 is updated (i.e. trained further) in at or near input data frame rate (i.e. online). In embodiments of this type, a plurality of the digital representations (e.g. images) of the head region may be received as input data ID (e.g. from a sensor system SS as described above) and adjusted (e.g. as described above with reference to FIGURES 11, 12 and 14) to provide output data OD that is subsequently displayed via display DS (as described above with reference to FIGURE 12). The input data ID may comprise a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch. The adjustment may be performed using a pre-trained version of the second machine learning algorithm MLA2, optionally in combination with a geometrical model. The input data ID may comprise digital representations captured at different points in times, such as different frames in a sequence of images obtained by the sensor system SS in a video capture mode. In an embodiment, the method comprises updating the pre-trained second machine learning algorithm MLA2 using first reference data (e.g. editing instructions) generated by the first machine learning algorithm MLA1 using one or more of the received digital representations (from the input data ID). The first reference data may comprise a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region. In the example depicted in FIGURE 15, the first reference data for the updating is generated in a background process and the second machine learning algorithm MLA2 is updated at or near input data frame rate (i.e. online).

[149] The updating thus uses a pre-trained version of the first machine learning algorithm MLA1 (which, as described above, is configured to provide more highly detailed reference data at the expense of higher computational demands, relative to the second machine learning algorithm MLA2) to generate first reference data that is used to update the second machine learning algorithm MLA2. Where the first machine learning algorithm MLA1 is slower than the second machine learning algorithm MLA2 (which will normally be the case where the first machine learning algorithm MLA1 is configured to provide more highly detailed reference data than the second machine learning algorithm MLA2), only a subset of the input data ID (comprising the received and adjusted digital representations) are used to update the second machine learning algorithm MLA2, thereby allowing the updating process to keep up with the adjusting of the input data ID by the second machine learning algorithm MLA2 to

provide the output data OD. Any of various known techniques may be used to perform the updating of the second machine learning algorithm MLA2. For example, the online updating may be performed as described in Amir Saffari, Christian Leistner, Jakob Santner, Martin Godec, and Horst Bischof, "On-line Random Forests," in 3rd IEEE ICCV Workshop on On-line Computer Vision, 2009, which is herein incorporated by reference in its entirety, or as described in Online Deep Learning: Learning Deep Neural Networks on the Fly, Doyen Sahoo, Quang Pham, Jing Lu, Steven C.H., Hoi School of Information Systems, Singapore Management University (<https://arxiv.org/pdf/1711.03705.pdf>), which is herein incorporated by reference in its entirety. The second machine learning algorithm MLA2 can therefore be gradually improved during use, as more input data ID are encountered and processed by the system.

[150] Additional embodiments of the disclosure are described in the following numbered clauses.

1. A method of generating reference data for adjusting an image of a head region, the method comprising:
 - receiving training data comprising:
 - a set of input image patches, each input image patch comprising a target feature of an image of a head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and
 - a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target feature of the image of the head region after adjustment of the image of the head region;
 - using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region; and
 - using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region.
2. The method of clause 1, wherein:
 - the first reference data comprise first image editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting image editing

instructions for a particular input image patch from the first image editing instructions based on the configuration of the target feature of the input image patch; and

the second reference data comprise second image editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting image editing instructions for a particular input image patch from the second image editing instructions based on the configuration of the target feature of the input image patch.

3. The method of clause 2, wherein the configuration of the target feature of each input image patch is represented by a feature vector derived from plural local image descriptors of the input image patch, and the first and second selection instructions define how the feature vector is used to select image editing instructions for the input image patch.

4. The method of clause 2 or 3, wherein the image editing instructions comprise a displacement vector field defining how the input image patch is to be transformed.

5. The method of clause 2 or 3, wherein the image editing instructions comprise a filter field, a brightness adjustment field, or a texture blending field.

6. The method of any of clauses 2-5, wherein:

a first image editing algorithm is used by the first machine learning algorithm to define how the first image editing instructions are to be applied to an input image patch to derive an output image patch; and

a second image editing algorithm is used by the second machine learning algorithm to define how the second image editing instructions are to be applied to an input image patch to derive an output image patch.

7. The method of clause 6, wherein:

the second image editing instructions are principle component analysis components of a principle component analysis of the first image editing instructions; and

the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse principle component analysis transform.

8. The method of clause 6, wherein:

the second image editing instructions are wavelet components of the first image editing instructions; and

the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse wavelet transform.

9. The method of any of clauses 2-8, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative image editing instructions than the second selection instructions for the second reference data.
10. The method of any preceding clause, wherein the first machine learning algorithm is of a different machine learning algorithm type than the second machine learning algorithm.
11. The method of any preceding clause, wherein the first machine learning algorithm comprises one or more of the following: a neural network; a support vector machine; a generative adversarial network, GAN.
12. The method of any preceding clause, wherein the second machine learning algorithm comprises one or more of the following: a regression forest; regression ferns, cluster centres, a lookup table, separable filter banks.
13. The method of any of clauses 1-10, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.
14. The method of any preceding clause, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair.
15. The method of any of clauses 1-13, wherein:
- the target feature comprises an eye region comprising at least part of an eye and the adjustment of the image of the head region comprises adjusting a gaze direction;
 - the target feature comprises a nose region comprising at least part of a nose and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the nose;
 - the target feature comprises a chin region comprising at least part of a chin and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the chin;
 - the target feature comprises a neck region comprising at least part of a neck and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the image of the head region comprises adjusting a color of the hair.

16. The method of any preceding clause, wherein the second image editing instructions in the second reference data are provided in a compressed representation.

17. The method of clause 16, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.

18. A method of adjusting an image of a head region, the method comprising:
identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;
deriving a feature vector from plural local image descriptors of the image patch;
using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and

applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:

the reference data comprises the second reference data generated by the method of any of clauses 1-17.

19. A method of adjusting an image of a head region, the method comprising:
identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;
deriving a feature vector from plural local image descriptors of the image patch;
using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and

applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:

the image editing instructions in the reference data are provided in a compressed representation.

20. The method of clause 19, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.

21. A computer program capable of execution by a processor and arranged on execution

to cause the processor to perform a method according to any of the preceding clauses.

22. A computer readable storage medium storing a computer program according to clause 21.

23. A device for generating reference data for adjusting an image of a head region, the device comprising a data processing unit arranged to:

receive training data comprising:

a set of input image patches, each input image patch comprising information about a target feature of an image of the head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and

a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target portion of the image of the head region after adjustment of the image of the head region;

use a first machine learning algorithm to generate first reference data using the training data; and

use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

24. A device for adjusting an image of a head region, the device comprising an image processor arranged to process the image of the head region by:

identifying an image patch in the image of the head region, the image patch comprising information about a target feature of the image of the head region;

deriving a feature vector from plural local image descriptors of the image patch;

using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and

applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:

the image editing instructions in the reference data are provided in a compressed representation.

25. The device of clause 24, further comprising a telecommunications interface arranged to transmit the adjusted image over a telecommunications network to a destination device for display thereon.

[151] While various embodiments in accordance with the principles disclosed herein have been described above, it should be understood that they have been presented by way of example only, and not limitation. Thus, the breadth and scope of this disclosure should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with any claims and their equivalents issuing from this disclosure. Furthermore, the above advantages and features are provided in described embodiments, but shall not limit the application of such issued claims to processes and structures accomplishing any or all of the above advantages.

[152] Additionally, the section headings herein are provided for consistency with the suggestions under 37 CFR 1.77 or otherwise to provide organizational cues. These headings shall not limit or characterize the embodiment(s) set out in any claims that may issue from this disclosure. Specifically and by way of example, although the headings refer to a "Technical Field," the claims should not be limited by the language chosen under this heading to describe the so-called field. Further, a description of a technology in the "Background" is not to be construed as an admission that certain technology is prior art to any embodiment(s) in this disclosure. Neither is the "Summary" to be considered as a characterization of the embodiment(s) set forth in issued claims. Furthermore, any reference in this disclosure to "invention" in the singular should not be used to argue that there is only a single point of novelty in this disclosure. Multiple embodiments may be set forth according to the limitations of the multiple claims issuing from this disclosure, and such claims accordingly define the embodiment(s), and their equivalents, that are protected thereby. In all instances, the scope of such claims shall be considered on their own merits in light of this disclosure, but should not be constrained by the headings set forth herein.

Claims:

1. A method of generating reference data for adjusting a digital representation of a head region, the method comprising:

receiving training data comprising:

a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and

a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;

using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and

using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

2. The method of claim 1, wherein:

the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch; and

the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch.

3. The method of claim 2, wherein the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input

patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch.

4. The method of claim 2 or 3, wherein the editing instructions comprise a displacement vector field defining how the input patch is to be transformed.

5. The method of claim 2 or 3, wherein the editing instructions comprise a filter field, a brightness adjustment field, or a texture blending field.

6. The method of any of claims 2-5, wherein:
a first editing algorithm is used by the first machine learning algorithm to define how the first editing instructions are to be applied to an input patch to derive an output patch; and
a second editing algorithm is used by the second machine learning algorithm to define how the second editing instructions are to be applied to an input patch to derive an output patch.

7. The method of claim 6, wherein:
the second editing instructions are principle component analysis components of a principle component analysis of the first editing instructions; and
the second image editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse principle component analysis transform.

8. The method of claim 6, wherein:
the second editing instructions are wavelet components of the first editing instructions; and
the second editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse wavelet transform.

9. The method of any of claims 2-8, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative editing instructions than the second selection instructions for the second reference data.

10. The method of any preceding claim, wherein the first machine learning algorithm is of a different machine learning algorithm type than the second machine learning algorithm.
11. The method of any preceding claim, wherein the first machine learning algorithm comprises one or more of the following: a neural network; a support vector machine; a generative adversarial network, GAN.
12. The method of any preceding claim, wherein the second machine learning algorithm comprises one or more of the following: a regression forest; regression ferns, cluster centres, a lookup table, separable filter banks.
13. The method of any of claims 1-10, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.
14. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair.
15. The method of any of claims 1-13, wherein:
- the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation of the head region comprises adjusting a gaze direction;
 - the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the nose;
 - the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the chin;

the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the digital representation of the head region comprises adjusting a color of the hair.

16. The method of any preceding claim, wherein the second editing instructions in the second reference data are provided in a compressed representation.

17. The method of claim 16, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers.

18. A method of adjusting a digital representation of a head region, the method comprising
receiving a digital representation of a head region; and
using reference data comprising editing instructions to adjust the digital representation of the head region, wherein the reference data comprises the second reference data generated by the method of any of claims 1-17.

19. A method of adjusting a digital representation of a head region, the method comprising:
identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;
deriving a feature vector from plural local descriptors of the target patch;
using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and
applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:
the reference data comprises the second reference data generated by the method of any of claims 1-17.

20. A method of adjusting a digital representation of a head region, the method comprising:
- identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;
 - deriving a feature vector from plural local descriptors of the target patch;
 - using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and
 - applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:
 - the editing instructions in the reference data are provided in a compressed representation.
21. The method of claim 20, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.
22. The method of any of claims 18-21, wherein a plurality of the digital representations of the head region are received and adjusted, each digital representation being captured from the head region at a different point in time, and the method further comprises updating the second machine learning algorithm using editing instructions generated by the first machine learning algorithm using one or more of the received and adjusted digital representations.
23. The method of claim 22, wherein the updating of the second machine learning algorithm is performed at or near input data frame rate.
24. The method of claim 23, wherein only a subset of the received and adjusted digital representations are used to update the second machine learning algorithm.
25. A method of training a machine learning algorithm to adjust a digital representation of a head region, comprising:
- receiving training data comprising:

a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and

first reference data generated by a pre-trained first machine learning algorithm, the first reference data comprising a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region;

updating a pre-trained second machine learning algorithm trained to generate second reference data, where the input for the updating comprises the training data and the generated first reference data, the second reference data comprising editing instructions for adjusting the digital representation of the head region.

26. The method of claim 25, wherein the updating of the second machine learning algorithm is performed at or near input data frame rate.

27. A method of training a machine learning algorithm to adjust a digital representation of a head region, the method comprising:

receiving training data comprising a set of input digital representations of a head region;

training a first machine learning algorithm using the training data to perform an adjustment of a digital representation of a head region;

using the trained first machine learning algorithm to generate first reference data, the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations, each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform; and

training a second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm.

28. The method of claim 27, wherein the adjustment of the digital representation comprises converting from a two-dimensional digital representation to a three-dimensional digital representation.
29. The method of claim 28, wherein the training of the first machine learning algorithm comprises iteratively using a rendering process to generate a two-dimensional digital representation from a three-dimensional digital representation generated by the first machine learning algorithm and comparing the generated digital representation with a corresponding digital representation in the training data.
30. The method of any of claims 25-29, further comprising receiving a further digital representation of the head region and using the updated or trained second machine learning algorithm to adjust the further digital representation.
31. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.
32. A computer readable storage medium storing a computer program according to claim 31.
33. A device for generating reference data for adjusting a digital representation of a head region, the device comprising a data processing unit arranged to:
- receive training data comprising:
 - a set of input patches, each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
 - a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target portion of the digital representation of the head region after adjustment of the digital representation of the head region;
 - use a first machine learning algorithm to generate first reference data using the training data; and

use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

34. The device of claim 33, wherein:

the first reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;

the second reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;

the second editing instructions in the second reference data are provided in a compressed representation.

35. A device for adjusting a digital representation of a head region, the device comprising a data processing unit arranged to process the digital representation of the head region by:

identifying a target patch in the digital representation of the head region, the target patch comprising information about a target feature of the digital representation of the head region;

deriving a feature vector from plural local descriptors of the target patch;

using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and

applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:

the editing instructions in the reference data are provided in a compressed representation.

36. The device of any of claims 33-35, further comprising a telecommunications interface arranged to transmit the adjusted digital representation over a telecommunications network to a destination device to allow display of the head region thereon.

1/12

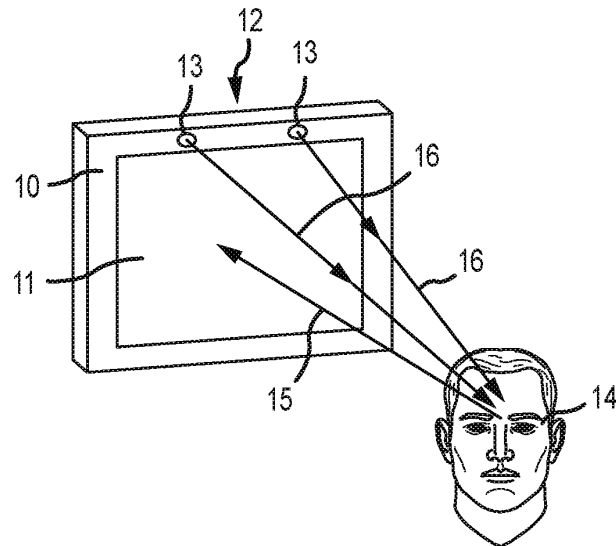


FIG. 1

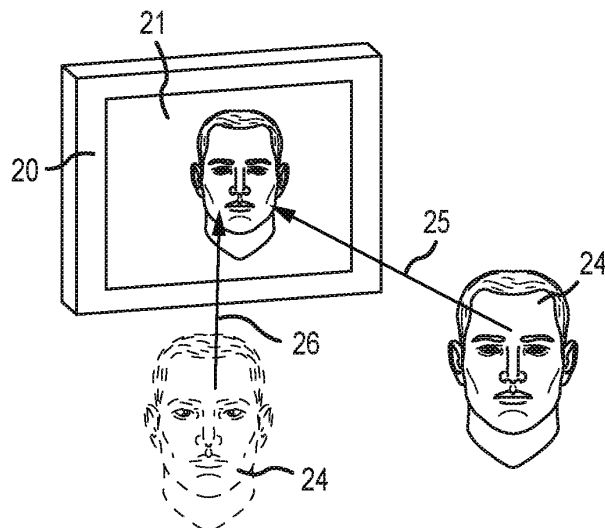


FIG. 2

2/12

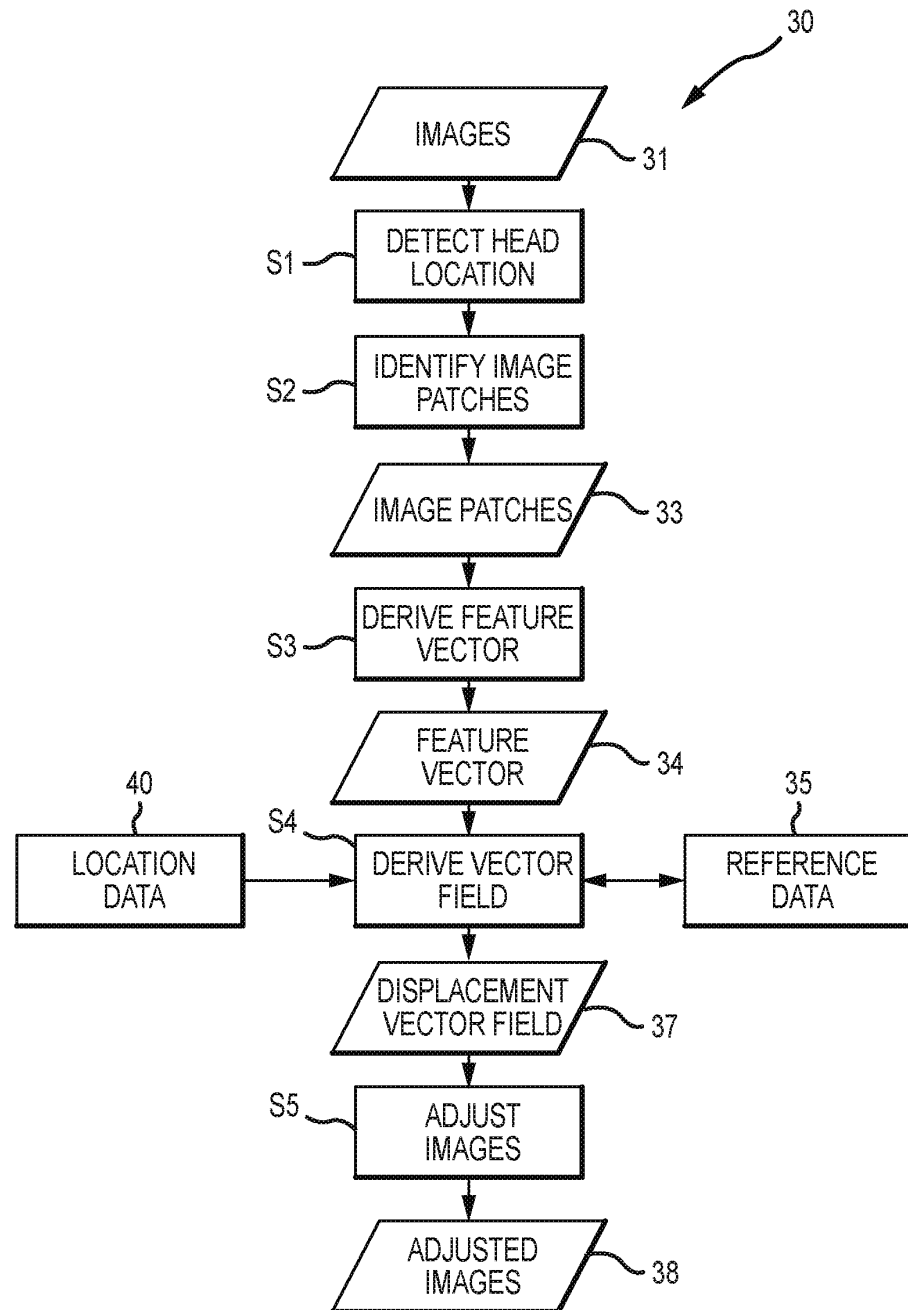


FIG.3

3/12

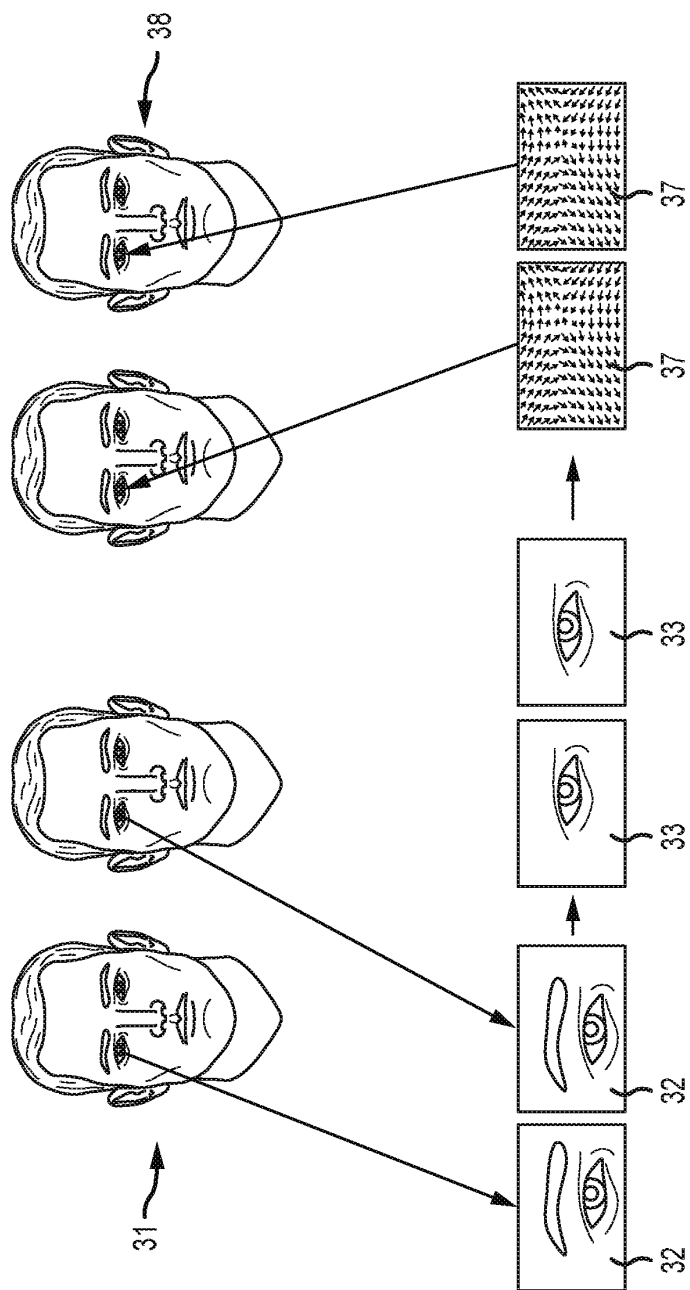


FIG.4

4/12

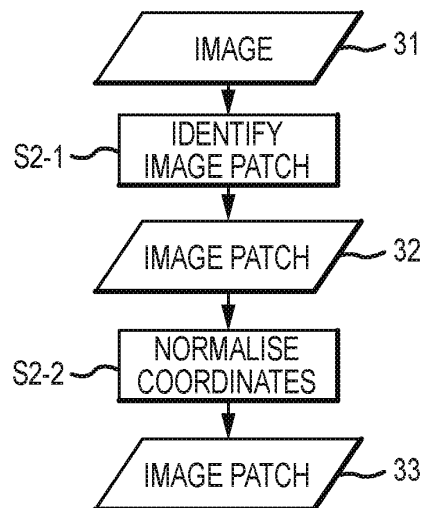


FIG.5

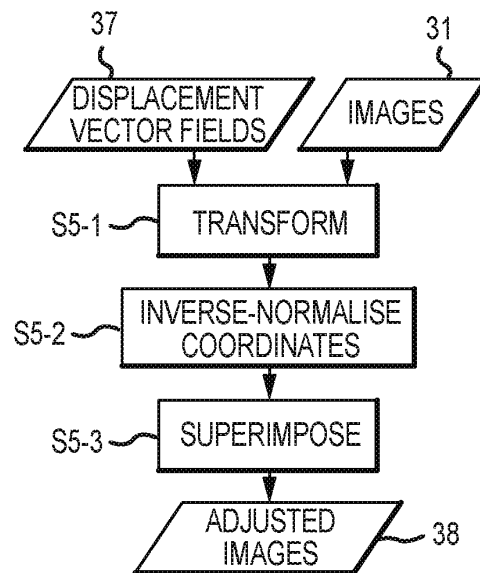


FIG.6

5/12

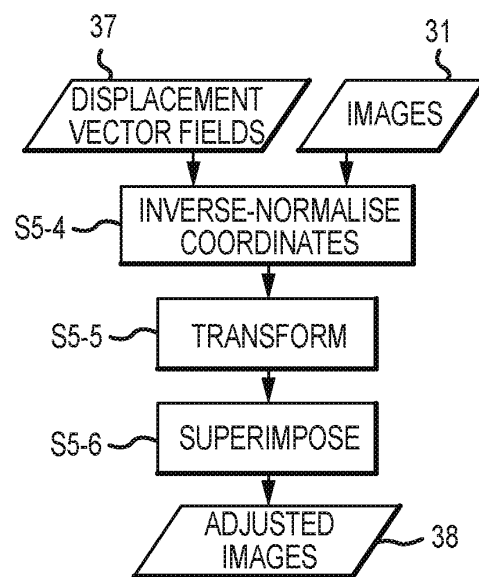


FIG.7

6/12

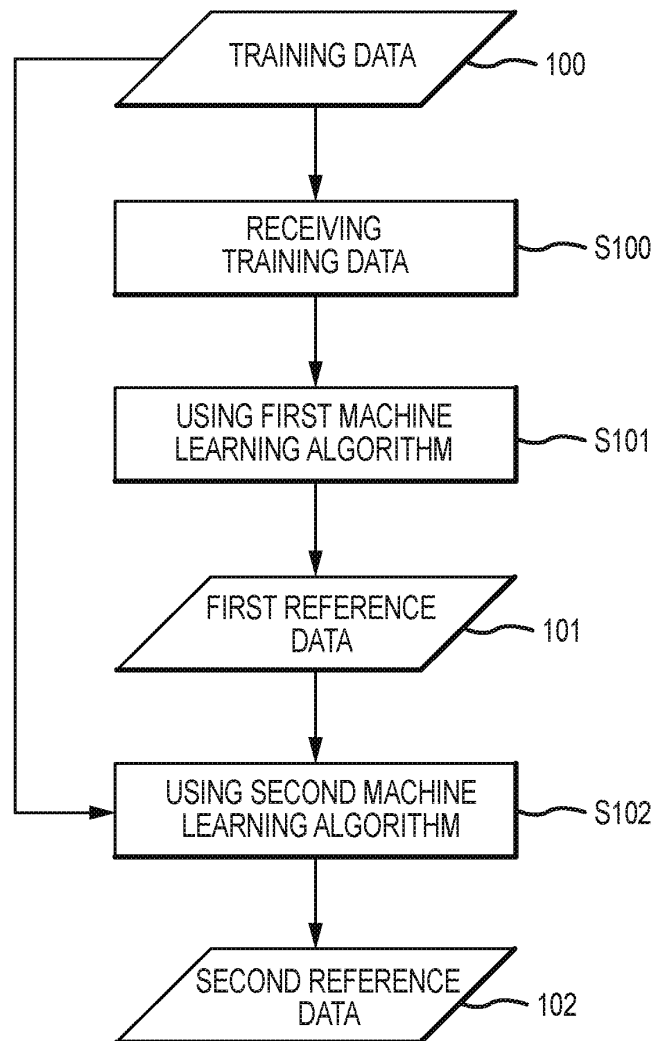


FIG.8

7/12

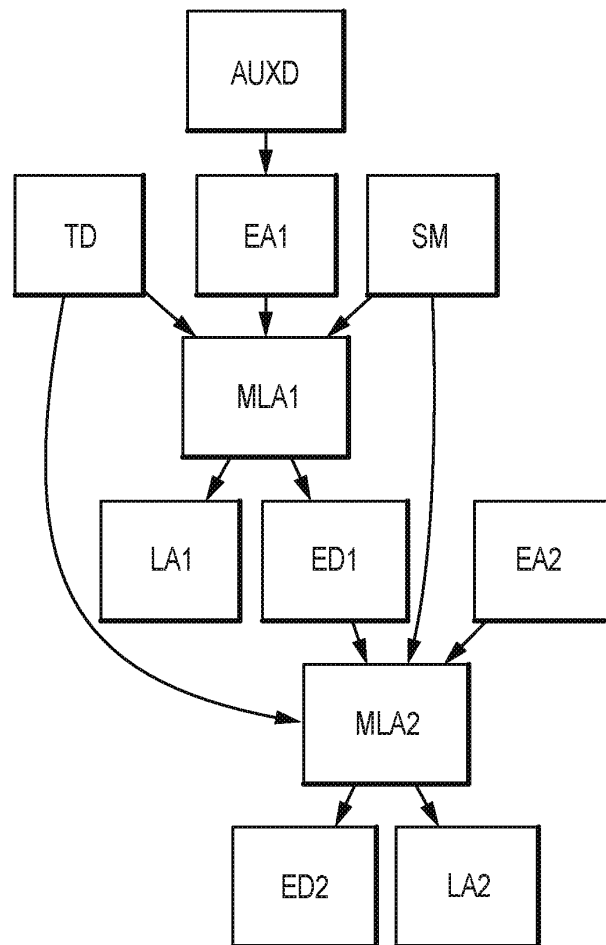


FIG.9

8/12

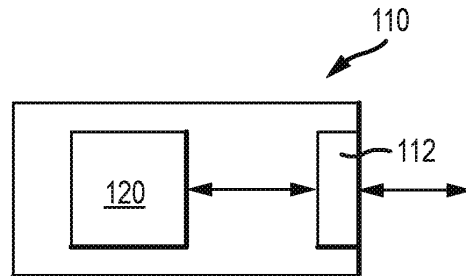


FIG. 10

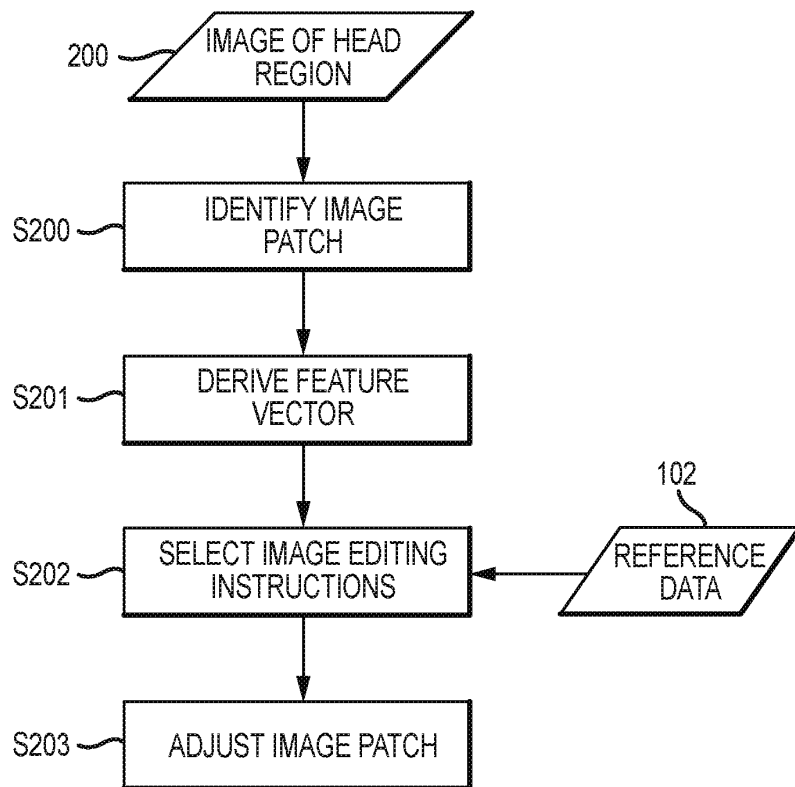


FIG. 11

9/12

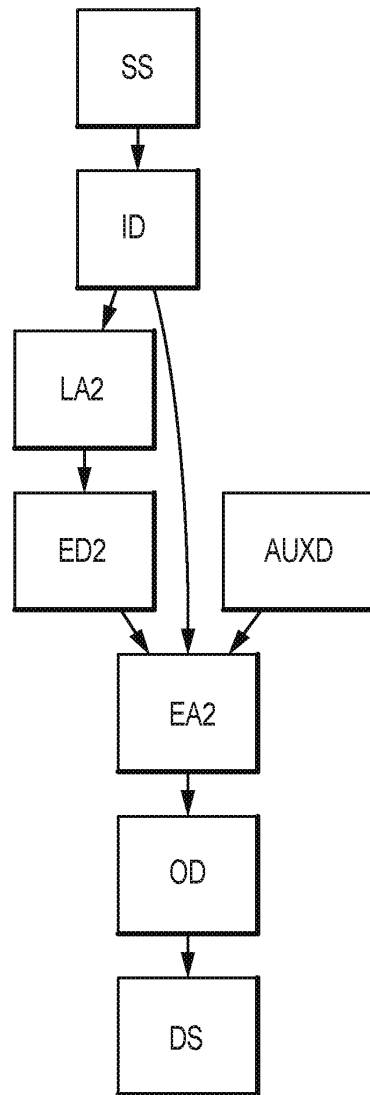


FIG.12

10/12

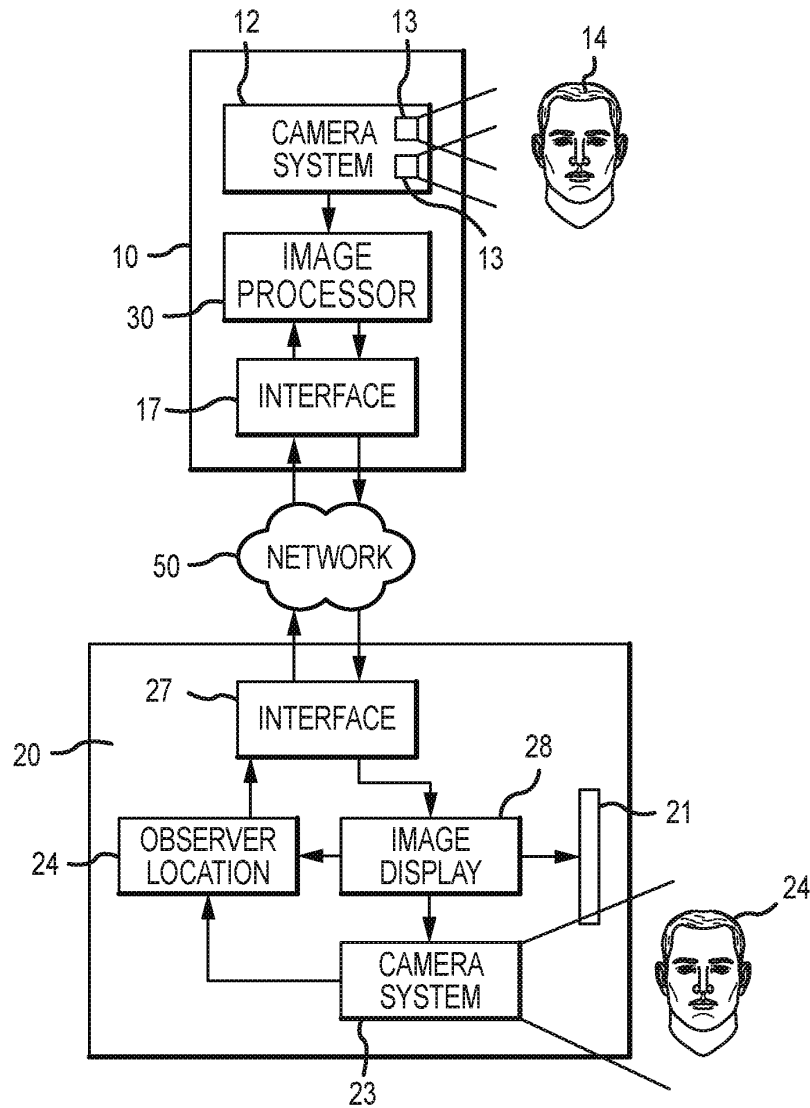


FIG.13

11/12

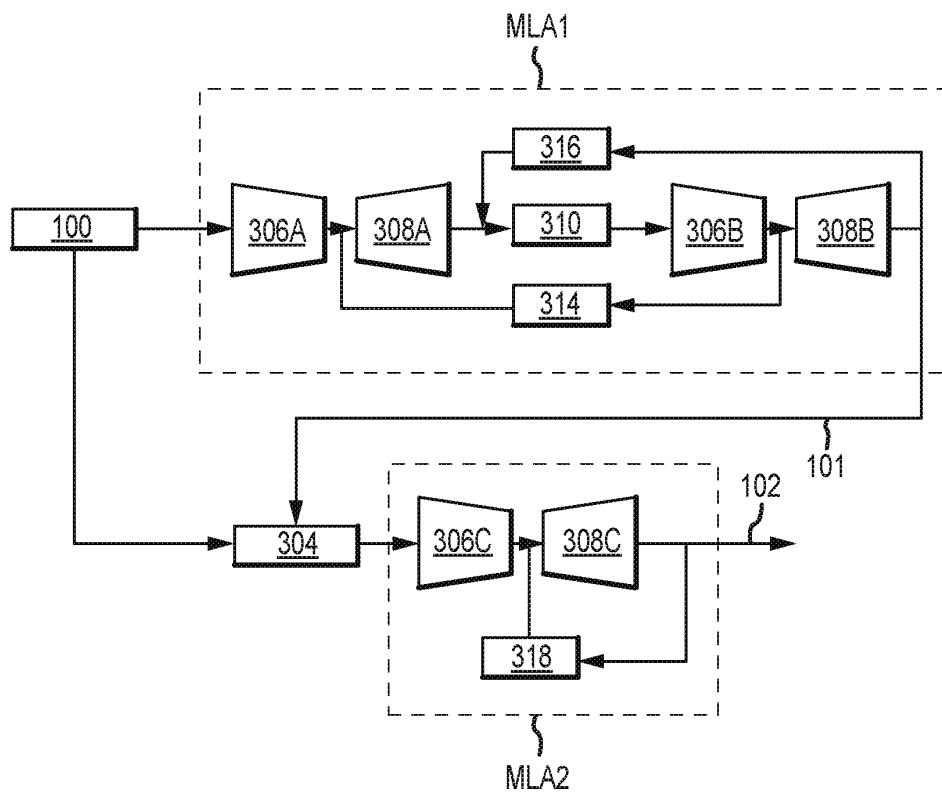


FIG.14

12/12

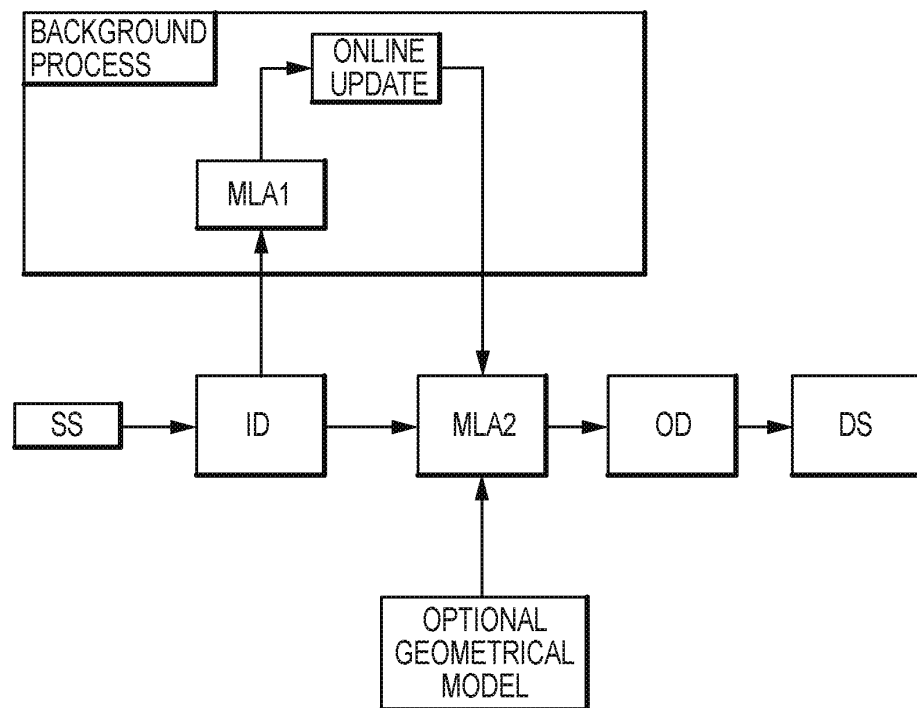


FIG.15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/045648

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 6-19, 22-24, 31, 32
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (2)) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/045648

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06F 3/01; G06K 9/00; H04N 5/228; H04N 7/14; H04N 13/00 (2018.01)
CPC - G06F 3/011; G06F 3/013; G06T 5/005; G06T 5/008; G06T 7/11 (2018.08)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History documentDocumentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 348/222.1; 382/167; 348/E5.024 (keyword delimited)Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2017/0195662 A1 (REALD SPARK, LLC) 06 July 2017 (06.07.2017) entire document	1-5, 20, 21, 25-30, 33-36
Y	WO 2016/132148 A1 (MAGIC PONY TECHNOLOGY LIMITED et al) 25 August 2016 (25.08.2016) entire document	1-5, 25-30, 33, 34, 36
Y	US 2014/0002586 A1 (NOURBAKSH) 02 January 2014 (02.01.2014) entire document	5
Y	US 2012/0319928 A1 (RHODES) 20 December 2012 (20.12.2012) entire document	20, 21, 34, 35
Y	US 2005/0104878 A1 (KAYE et al) 19 May 2005 (19.05.2005) entire document	28, 29
A	US 2012/0114201 A1 (LUI SI et al) 10 May 2012 (10.05.2012) entire document	1-5, 20, 21, 25-30, 33-36
A	US 2011/0063465 A1 (NANU et al) 17 March 2011 (17.03.2011) entire document	1-5, 20, 21, 25-30, 33-36
A	US 2016/0219258 A1 (REALD INC.) 28 July 2016 (28.07.2016) entire document	1-5, 20, 21, 25-30, 33-36

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

01 October 2018

Date of mailing of the international search report

16 OCT 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

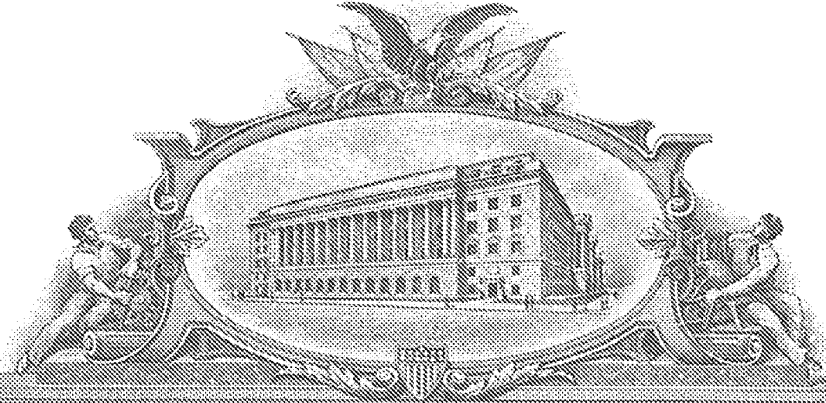
Form PCT/ISA/210 (second sheet) (January 2015)

**DOCUMENT MADE AVAILABLE UNDER THE
PATENT COOPERATION TREATY (PCT)**

International application number:	PCT/US2018/045648
International filing date:	07 August 2018 (07.08.2018)
Document type:	Certified copy of priority document
Document details:	Country/Office: US
	Number: 62/542,661
	Filing date: 08 August 2017 (08.08.2017)
Date of receipt at the International Bureau:	22 August 2018 (22.08.2018)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-bis)

944499



THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 21, 2018

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 62/542,661

FILING DATE: *August 08, 2017*

RELATED PCT APPLICATION NUMBER: *PCT/US18/45648*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US62/542,661*



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2:

☐	Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)
--------------------------	---

Inventor Information:

Inventor 1				
Legal Name				
Prefix	Given Name	Middle Name	Family Name	Suffix
	Eric		Sommerlade	
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service				
City	Oxford	Country of Residence ⁱ	GB	
Mailing Address of Inventor:				
Address 1	100 N. Crescent Drive			
Address 2	Suite 200			
City	Beverly Hills	State/Province	CA	
Postal Code	90210	Country ⁱ	US	
All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the Add button. Add				

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).			
☐ An Address is being provided for the correspondence information of this application.			
Customer Number	145234		
Email Address	patpros@reald.com	Add Email	Remove Email

Application Information:

Title of the Invention	Method and system for improved facial visual data editing using machine learning		
Attorney Docket Number	409000	Small Entity Status Claimed	☒
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	3	Suggested Figure for Publication (if any)	

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:		☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)	
Customer Number					
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Additional Representative Information blocks may be generated within this form by selecting the Add button.					

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

When referring to the current application, please leave the application number blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button.

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country ¹	Filing Date (YYYY-MM-DD)	Access Code ¹ (if applicable)

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.

☐ NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

☐ Assignee	☐ Legal Representative under 35 U.S.C. 117	☐ Joint Inventor
☐ Person to whom the inventor is obligated to assign.	☐ Person who shows sufficient proprietary interest	

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

--	--

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. ☐

Prefix	Given Name	Middle Name	Family Name	Suffix

Mailing Address Information For Applicant:

Address 1			
Address 2			
City		State/Province	
Country [†]		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Assignee 1									
Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.									
If the Assignee or Non-Applicant Assignee is an Organization check here. ☐									
Prefix	Given Name	Middle Name	Family Name	Suffix					
Mailing Address Information For Assignee including Non-Applicant Assignee:									
Address 1									
Address 2									
City		State/Province							
Country i			Postal Code						
Phone Number		Fax Number							
Email Address									
Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.									

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000
		Application Number	
Title of Invention	Method and system for improved facial visual data editing using machine learning		

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Penny L. Lowry/			Date (YYYY-MM-DD)	
First Name	Penny L.	Last Name	Lowry	Registration Number	57186
Additional Signature may be generated within this form by selecting the Add button.					

Doc Code: **TR.PROV**

Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (11-08)

Approved for use through 05/31/2015. OMB 0651-0032

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Provisional Application for Patent Cover Sheet					
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)					
Inventor(s)					
Inventor 1					Remove
Given Name	Middle Name	Family Name	City	State	Country j
Eric		Sommerlade	Oxford		GB
All Inventors Must Be Listed – Additional Inventor Information blocks may be generated within this form by selecting the Add button.					Add
Title of Invention		Method and system for improved facial visual data editing using machine learning			
Attorney Docket Number (if applicable)		409000			
Correspondence Address					
Direct all correspondence to (select one):					
☒ The address corresponding to Customer Number			☐ Firm or Individual Name		
Customer Number			145234		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.	
☒ No.	
Yes, the invention was made by an agency of the United States Government. The U.S. Government agency name is:	
Yes, the invention was under a contract with an agency of the United States Government. The name of the U.S. Government agency and Government contract number are:	

Entity Status**Applicant asserts small entity status under 37 CFR 1.27 or applicant certifies micro entity status under 37 CFR 1.29**

- ☒ Applicant asserts small entity status under 37 CFR 1.27
- ☐ Applicant certifies micro entity status under 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
- ☐ No

Warning

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

Signature

Please see 37 CFR 1.4(d) for the form of the signature.

Signature	Penny L. Lowry/		Date (YYYY-MM-DD)	2017-08-08
First Name	Penny L.	Last Name	Lowry	Registration Number (If appropriate)
				57186

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **This form can only be used when in conjunction with EFS-Web. If this form is mailed to the USPTO, it may cause delays in handling the provisional application.**

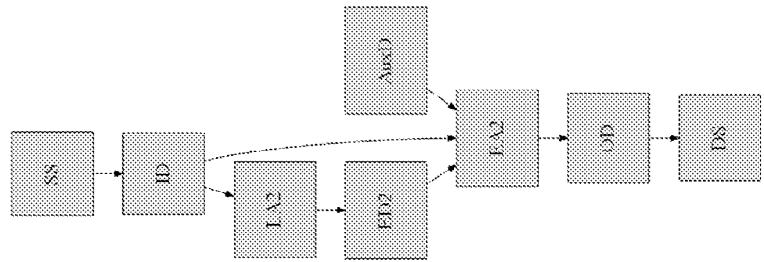
Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that : (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

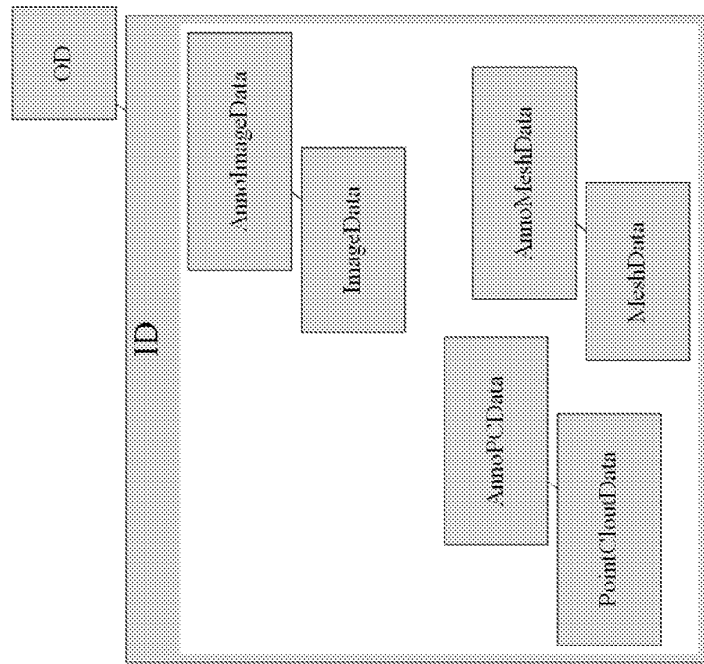
The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

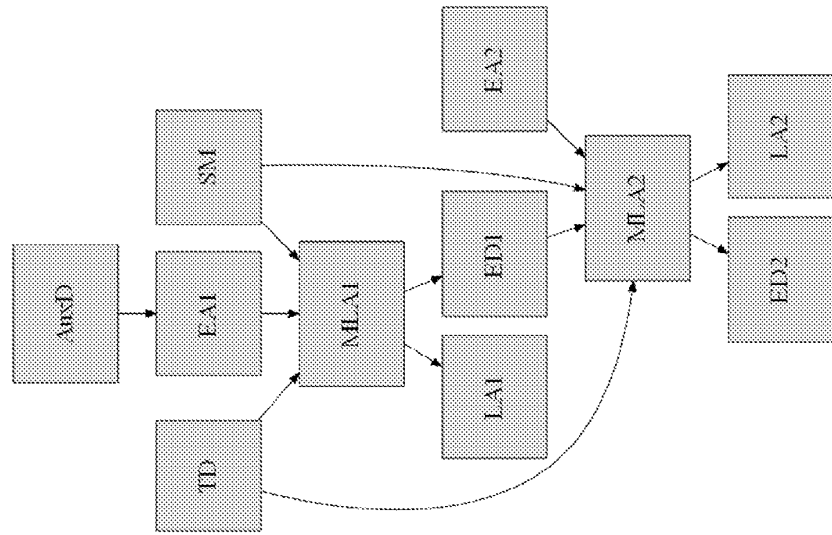
Example Deployed System



Example Training Data TD



Example Training System TS



409000 Patent description

Method and system for improved facial visual data editing using machine learning

Abstract:

Facial image features are among the most critically perceived elements by observers. It is desirable to bridge the “uncanny valley”. Changes to facial features in a communication setting are often desired by participants: eye gaze direction, nose size and double necks. This patent describes embodiments of a method to apply changes to input data of varying modality by using the input data as guideline. Relevant image patches may be identified, and a feature vector may be derived from plural local image descriptors of the image patches. A displacement vector field representing a transformation of an image patch and geometry may be derived, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector produced by machine learning. The input data may be adjusted by transforming the relevant image patches and geometry data in accordance with the derived displacement vector field.

Example purposes of system:

- The system learns editing instructions to transform input data to a desired target output.
- The system works on data from a given category, in particular human heads, faces, eyes, necks, ears. The category is fixed per embodiment.
- The editing instructions are evaluated by the same implementation per embodiment. One embodiment of the editing instructions is to shift a rectangular set of pixels depending on the underlying color information, achieving the adjustment of the gaze of the individual depicted
- The achieved effect of the editing instruction is the same for all input data. Example effects are
 - o Adjust the gaze of an individual in the input data to look upward by a certain angle
 - o Adjust the nasal region of an individual in the input data to be smaller/slimmer by a fixed proportion (this has the effect of a “tele lens” more suitable for profile photos)
 - o Change the hair colour of the subject by a fixed hue
- Output of the system are data editing instructions. These are subsequently applied to any new input not part of the image database or training set
- In one embodiment of the system, the system takes a previously unseen input image of a face, and provides a transformation that adjusts the output image. The output image has the gaze of the observed face in a different direction than in the input image. For all input images, the relative adjustment of the angle is the same (e.g. 15 degrees up), depending on the input conditions (e.g. position of head in input data).
- Example Input domains:
 - o 1a at least one image with at least one colour channel, e.g. a gray, RGB, YCrCb image
 - o 1b multiple images with different colour channels or sensing modalities, e.g. one RGB, one near-infrared image, or one RGB, one image captured with structured light

- structured light can be a projected pattern or background flickering of ambient lighting (temporal structure)
 - 1c a mesh or point cloud structure with coloured vertices
 - 1d a mesh with texture index per vertex and associated texture map
 - 1e any other volumetric representation of data which has image data and geometric data
- Output Domain: any of the input domains

Example Components:

- Training system **TS** and evaluation system **ES**
- Editing algorithm **EA**
 - The algorithm takes a given editing instruction (derived from the input data and auxiliary), applies these to the input data to provide the output data
- Lookup instructions **LA**
 - These instructions provide the editing instructions appropriate for the given input data. Embodiments are:
 - A lookup table
 - One or a set of linear regressors
 - A neural network
 - Embodiments for Inputs for to these instructions are:
 - A set of input pixels or descriptions thereof
 - A set of geometrical descriptions of the input 3d data

Example Editing instructions **ED**

- These instructions change the input object from the input domain into the output domain representation. This transformation can be supported by auxiliary data, e.g. knowledge of the mesh representation, which can be added to the machine learning algorithm
 - Shifts of pixel colours and underlying geometry (e.g. if the texture of the eye is adjusted to move the eyelid, the corresponding 3d geometry also is adjusted)
 - Intensity changes to the pixel colours
 - Representations of the adjustments using e.g. PCA or a Fourier Transform
- **TS** may be used offline to learn the editing instructions. These instructions may then be deployed to a multitude of **ES**

- Example Input to system **TS**:
 - Machine Learning Algorithm 1 (**MLA1**)
 - Machine Learning Algorithm 2 (**MLA2**), where the accuracy of ML algorithm 2 is not necessarily as good as using **MLA1**. However, **MLA2** performs faster, and can approximate the accuracy of **MLA1** with more training data
 - Editing algorithm **EA1**
 - Editing algorithm **EA2**

- Example Input data to **MLA1**:
 - **TD1**: a set of objects from the input domain (Source), and a set of objects from the output domain with the desired effect, whereas input and desired output object are in one-to-one correspondence (Target)
 MLA1 learns the editing instructions indirectly, through a latent representation.
 - Similarity metric **SM**: a numerical value to measure similarity between the modified input object and the desired output object. .
 - embodiments of similarity metrics may be defined by
 - Average absolute or square difference between result and target
 - Average absolute or square difference between low level representations of result and target (e.g. low level features from deep learning networks, such as VGG16)
 - Penalize “taboo zones” (e.g. sclera) that need to be maintained
 - Penalize deviation of an adjusted feature from known form, (e.g. the iris must stay an elliptical form) – or difference to said form in training data

 - Example Output of **MLA1**:
 - Lookup instruction **LA1** to provide appropriate editing instructions **ED1** for the given data
 - editing instructions **ED1** to modify input object, with the aim to make input object similar to the output object.

 - Example Input to **MLA2**:
 - Training data **TD1**
 - Editing instructions **ED1**, the set of editing information obtained by MLA1.
 - Similarity metric **SM**
 - MLA2 learns the editing information directly. By making the desired output explicit, MLA2 has a better chance at fitting the output,
 - Example Output of **MLA2**:
 - editing instructions **ED2** to modify input object, with the aim to make input object similar to the output object.
 - Lookup instruction **LA2** to provide appropriate editing instructions **ED2** for the given data

- Example Input to system **ES**:
 - Sensor system or network connection **SS**
 - Input data **ID**
 - editing instructions **ED2**
 - lookup instruction **LA2**
 - editing algorithm **EA2**
- Example Output of system **ES**:
 - Output Data **OD**
 - Display system or network connection **DS**

Example Claims:

- 1) System and method to apply editing information on input data relating to facial imagery and geometry, whereas the input geometry is aligned with the input image, comprising
 - a. A method to locate relevant parts of input data, in particular eye regions, nose or mouth
 - b. A method deriving a feature vectors from relevant parts
 - c. Applying said feature vectors to LA2 to obtain editing instructions ED2
 - d. Applying editing instructions to the input images and input geometry, making consistent adjustments
- 2) A system according to 1), adding an input camera system comprising at least one camera and a method to match input data to 3d geometry
- 3) A system according to 1), adding an input camera system comprising a plurality of cameras and a method to estimate 3d geometry
- 4) A system according to 1), adding an input camera system comprising at least one camera and a depth sensing camera, providing 3d geometry
- 5) A system according to 1), whereas the facial image region is constrained to be the region of the eyes, and the editing information is adjusting the gaze in the image as well as the eyelids geometry information
 - a. Same for nose, (double) chin and neck region
- 6) A system according to 1), where MLA1 is a neural network and MLA2 is a regression forest
- 7) A system according to 1), where the output ED2 is are PCA components of the PCA of ED1 data, and EA2 first transforms ED2 output into ED1 by inverse PCA transform
- 8) A system according to 1), where the output ED2 is are wavelet components of ED1 data, and EA2 first transforms ED2 output into ED1 by inverse wavelet transform
- 9) A system according to 1) where the input ID is a single image and output OD is single image
- 10) A system according to 1) where the input ID is a single image and output OD is a pair of stereoscopic images
- 11) A system according to 1) where the input ID and output OD is a pair of stereoscopic images each

Electronic Acknowledgement Receipt	
EFS ID:	30022345
Application Number:	62542661
International Application Number:	
Confirmation Number:	7825
Title of Invention:	Method and system for improved facial visual data editing using machine learning
First Named Inventor/Applicant Name:	Eric Sommerlade
Customer Number:	145234
Filer:	Penny Lea Lowry/Stefan Bump
Filer Authorized By:	Penny Lea Lowry
Attorney Docket Number:	409000
Receipt Date:	08-AUG-2017
Filing Date:	
Time Stamp:	18:10:53
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	\$130
RAM confirmation Number	080917INTEFSW18134600
Deposit Account	506262
Authorized User	Stefan Bump
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: 37 CFR 1.16 (National application filing, search, and examination fees) 37 CFR 1.17 (Patent application and reexamination processing fees)	

37 CFR 1.19 (Document supply fees)					
37 CFR 1.20 (Post Issuance fees)					
37 CFR 1.21 (Miscellaneous fees and charges)					
File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	WebADS.pdf	114911	no	7
			6e5e2657678790ebea50faad15fe0931af4781ff		
Warnings:					
Information:					
2	Provisional Cover Sheet (SB16)	409000_ProvisionalSB.pdf	1477153	no	3
			cd3a9fbd9dd79fd34b1b745613f949410f919445		
Warnings:					
Information:					
3		409000_PRV_Application.pdf	205423	yes	7
			ec7b7e5725081528e9dc7794275dff6c6c0d2a5e		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Drawings-only black and white line drawings		5	7	
	Specification		1	4	
Warnings:					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	29921	no	2
			2782fefa6d86af5af109b722b77ca7233a86f3c6		
Warnings:					
Information:					
Total Files Size (in bytes):			1827408		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

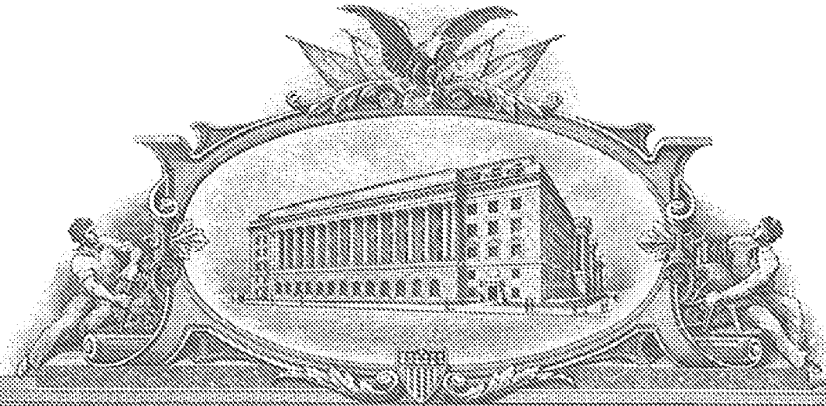
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

**DOCUMENT MADE AVAILABLE UNDER THE
PATENT COOPERATION TREATY (PCT)**

International application number:	PCT/US2018/045648
International filing date:	07 August 2018 (07.08.2018)
Document type:	Certified copy of priority document
Document details:	Country/Office: US
	Number: 62/543,587
	Filing date: 10 August 2017 (10.08.2017)
Date of receipt at the International Bureau:	22 August 2018 (22.08.2018)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-bis)

944500



THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 21, 2018

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 62/543,587

FILING DATE: *August 10, 2017*

RELATED PCT APPLICATION NUMBER: *PCT/US18/45648*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US62/543,587*



Certified by

Michelle K. Lee

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2:

☐	Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)
--------------------------	---

Inventor Information:

Inventor 1				
Legal Name				
Prefix	Given Name	Middle Name	Family Name	Suffix
	Eric		Sommerlade	
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service				
City	Oxford	Country of Residence ⁱ	GB	
Mailing Address of Inventor:				
Address 1	100 N. Crescent Drive			
Address 2	Suite 200			
City	Beverly Hills	State/Province	CA	
Postal Code	90210	Country ⁱ	US	
All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the Add button. Add				

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).			
☐ An Address is being provided for the correspondence information of this application.			
Customer Number	145234		
Email Address	patpros@reald.com	Add Email	Remove Email

Application Information:

Title of the Invention	Generating reference data for adjusting an image of a head region		
Attorney Docket Number	409000A	Small Entity Status Claimed	☒
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	10	Suggested Figure for Publication (if any)	

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)		
Customer Number	145234				
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Additional Representative Information blocks may be generated within this form by selecting the Add button.					

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

When referring to the current application, please leave the application number blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button.

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country ¹	Filing Date (YYYY-MM-DD)	Access Code ¹ (if applicable)

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013. ☐ NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.
--

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

☐ Assignee	☐ Legal Representative under 35 U.S.C. 117	☐ Joint Inventor
☐ Person to whom the inventor is obligated to assign.	☐ Person who shows sufficient proprietary interest	

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

--	--

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. ☐

Prefix	Given Name	Middle Name	Family Name	Suffix

Mailing Address Information For Applicant:

Address 1			
Address 2			
City		State/Province	
Country ¹		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Assignee 1									
Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.									
If the Assignee or Non-Applicant Assignee is an Organization check here. ☐									
Prefix	Given Name	Middle Name	Family Name	Suffix					
Mailing Address Information For Assignee including Non-Applicant Assignee:									
Address 1									
Address 2									
City		State/Province							
Country i			Postal Code						
Phone Number		Fax Number							
Email Address									
Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.									

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	409000A
		Application Number	
Title of Invention	Generating reference data for adjusting an image of a head region		

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Penny L. Lowry/			Date (YYYY-MM-DD)	
First Name	Penny L.	Last Name	Lowry	Registration Number	57186
Additional Signature may be generated within this form by selecting the Add button.					

Doc Code: **TR.PROV**

Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (11-08)

Approved for use through 05/31/2015. OMB 0651-0032

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Provisional Application for Patent Cover Sheet					
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)					
Inventor(s)					
Inventor 1					Remove
Given Name	Middle Name	Family Name	City	State	Country j
Eric		Sommerlade	Oxford		GB
All Inventors Must Be Listed – Additional Inventor Information blocks may be generated within this form by selecting the Add button.					Add
Title of Invention		Method and system for improved facial visual data editing using machine learning			
Attorney Docket Number (if applicable)		409000A			
Correspondence Address					
Direct all correspondence to (select one):					
☒ The address corresponding to Customer Number			☐ Firm or Individual Name		
Customer Number			145234		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.	
☒ No.	
Yes, the invention was made by an agency of the United States Government. The U.S. Government agency name is:	
Yes, the invention was under a contract with an agency of the United States Government. The name of the U.S. Government agency and Government contract number are:	

Entity Status**Applicant asserts small entity status under 37 CFR 1.27 or applicant certifies micro entity status under 37 CFR 1.29**

- ☒ Applicant asserts small entity status under 37 CFR 1.27
- ☐ Applicant certifies micro entity status under 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
- ☐ No

Warning

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

Signature

Please see 37 CFR 1.4(d) for the form of the signature.

Signature	Penny L. Lowry/		Date (YYYY-MM-DD)	2017-08-10
First Name	Penny L.	Last Name	Lowry	Registration Number (If appropriate)
				57186

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **This form can only be used when in conjunction with EFS-Web. If this form is mailed to the USPTO, it may cause delays in handling the provisional application.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that : (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Fig.1

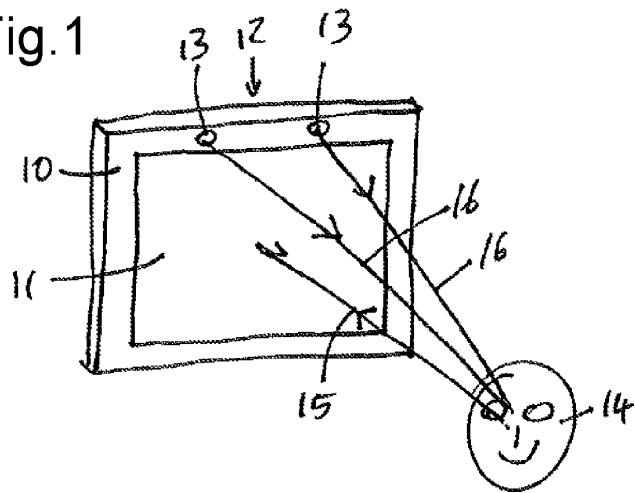


Fig.2

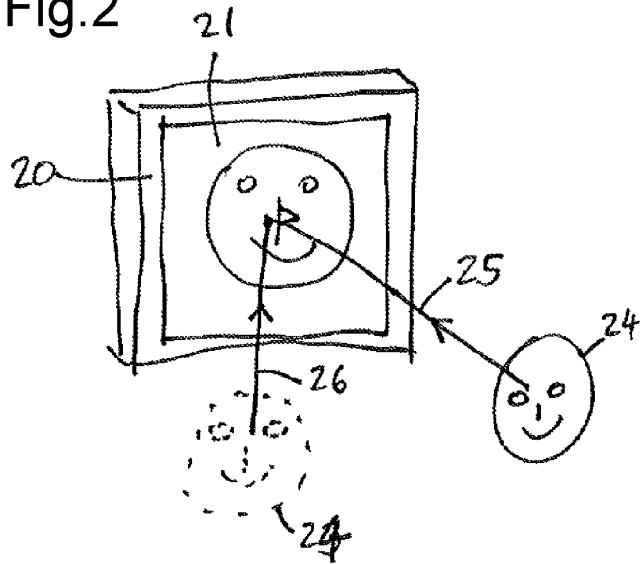


Fig.3

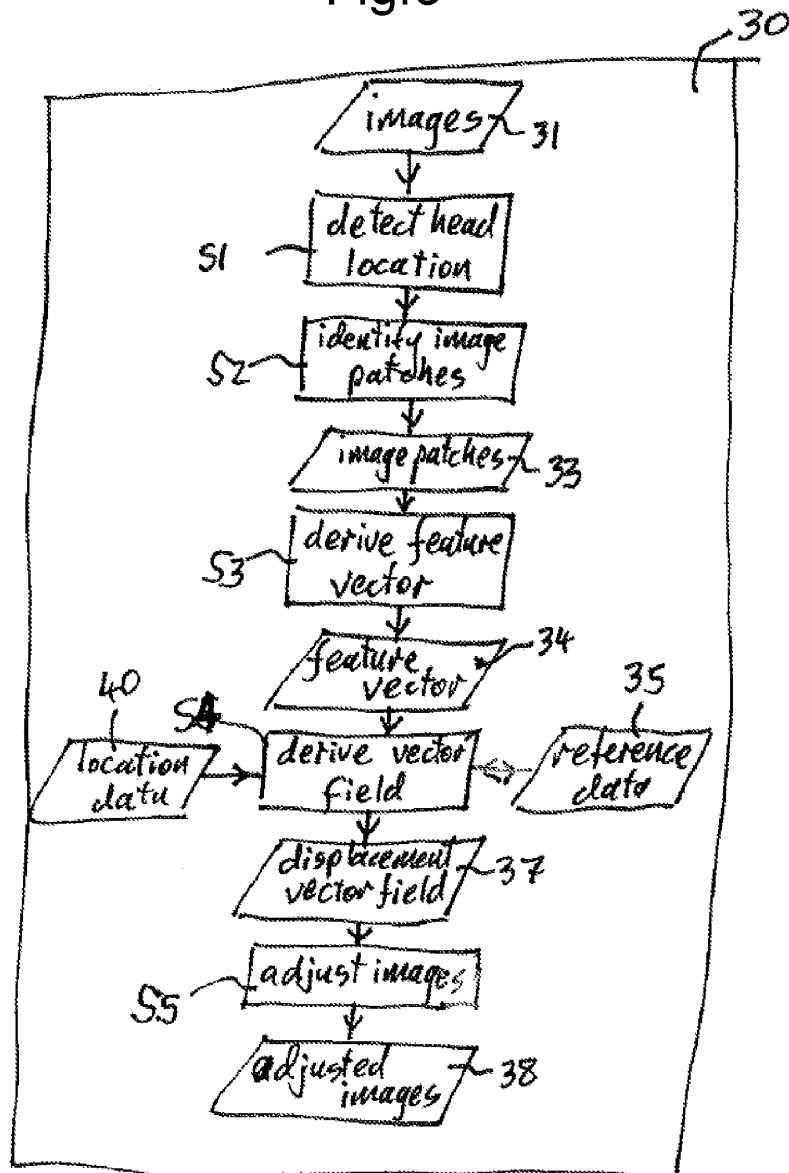


Fig.4

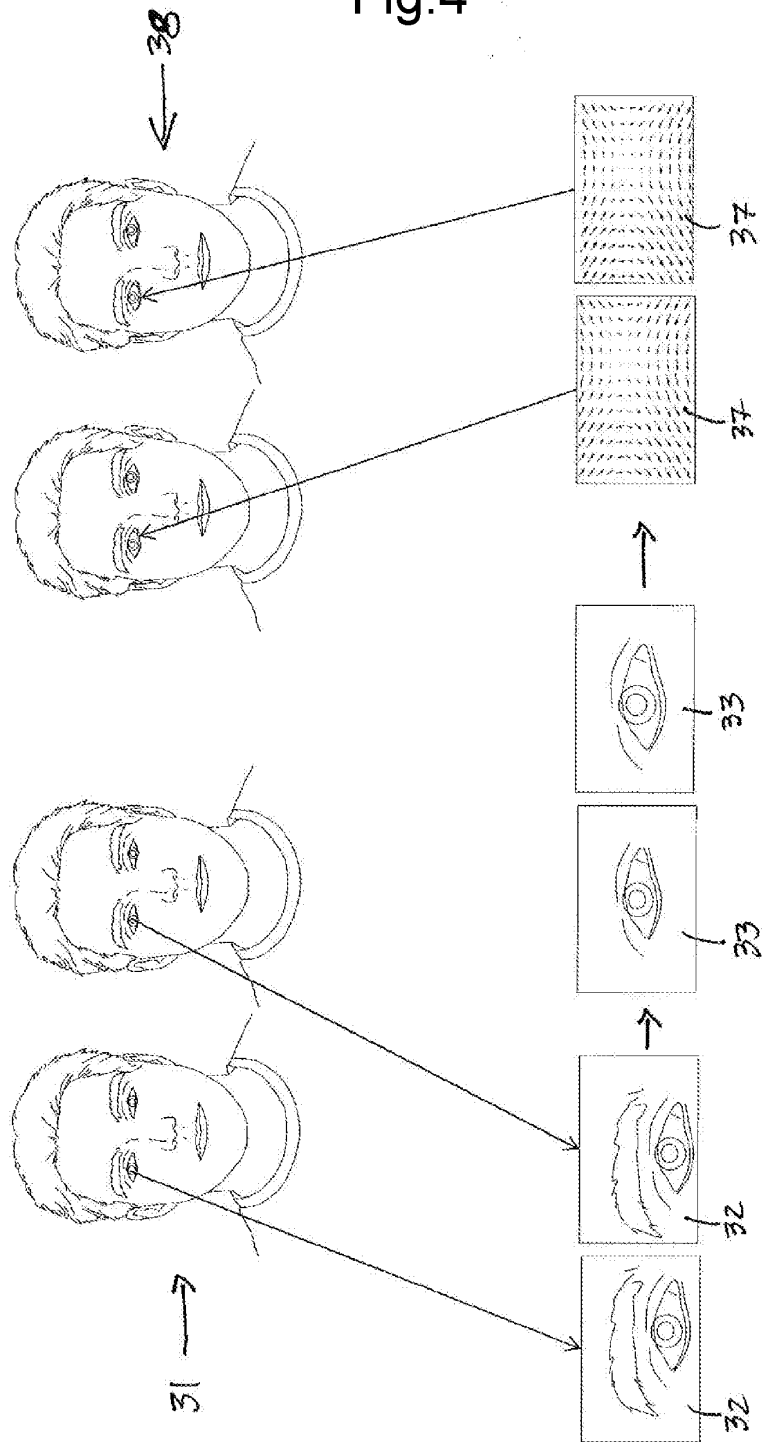


Fig.5

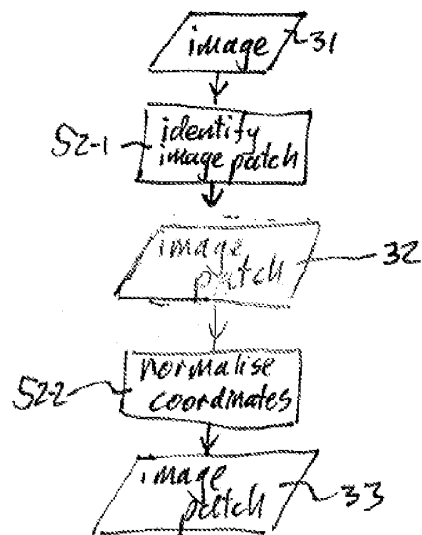


Fig.6

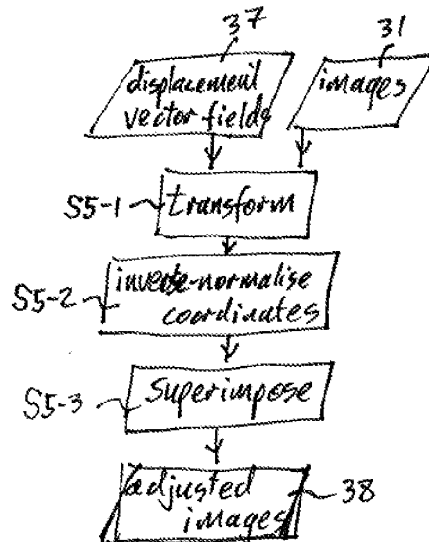


Fig.7

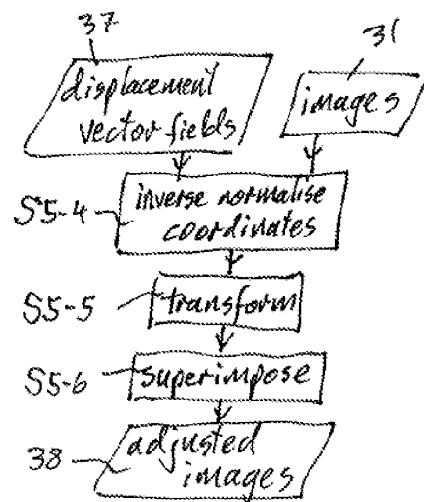


Fig.8

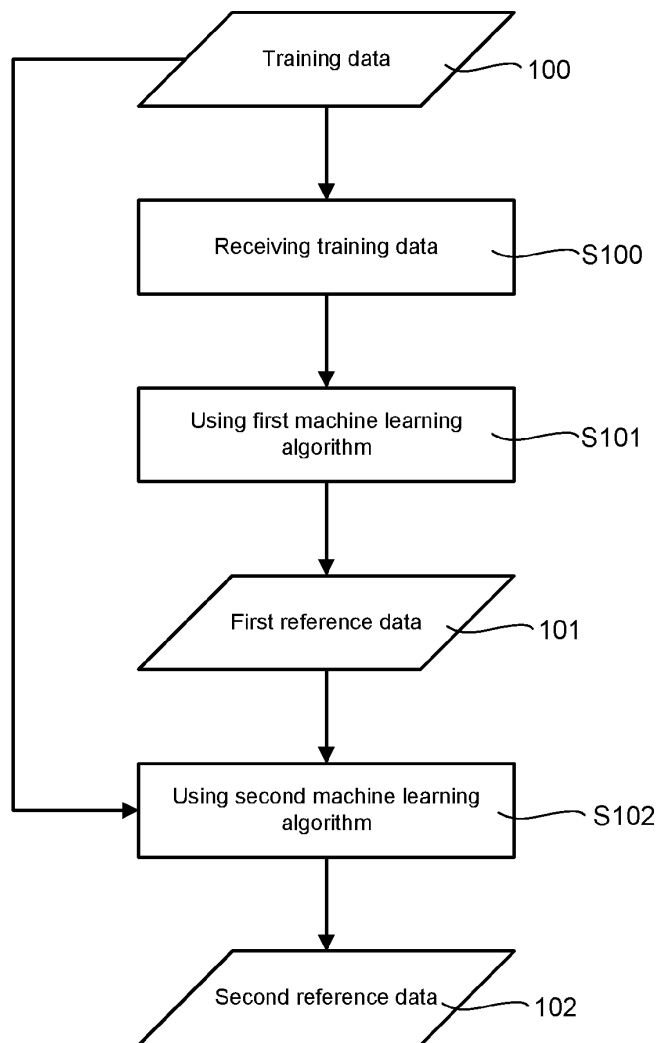


Fig.9

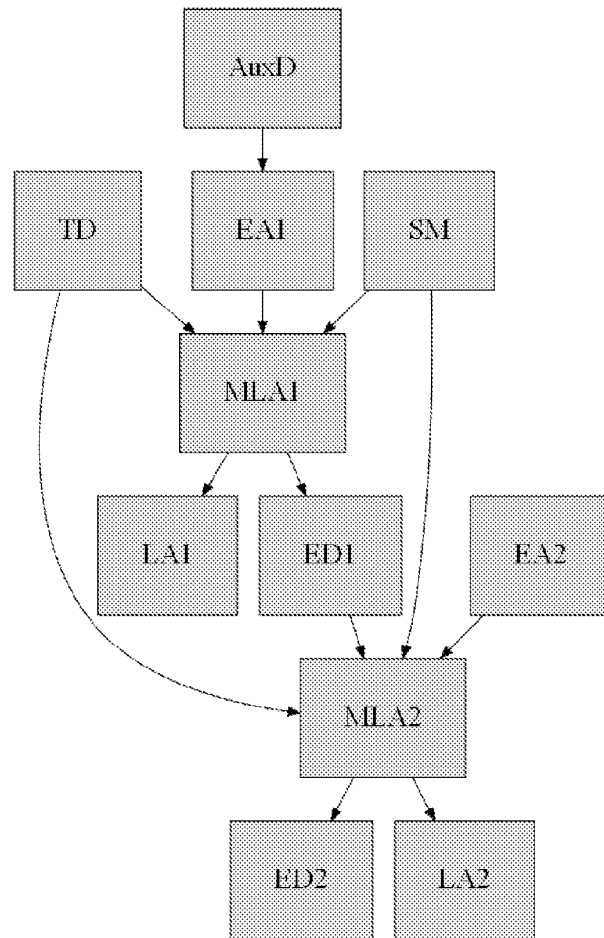


Fig.10

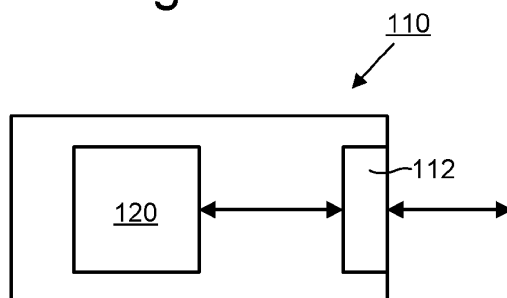


Fig.11

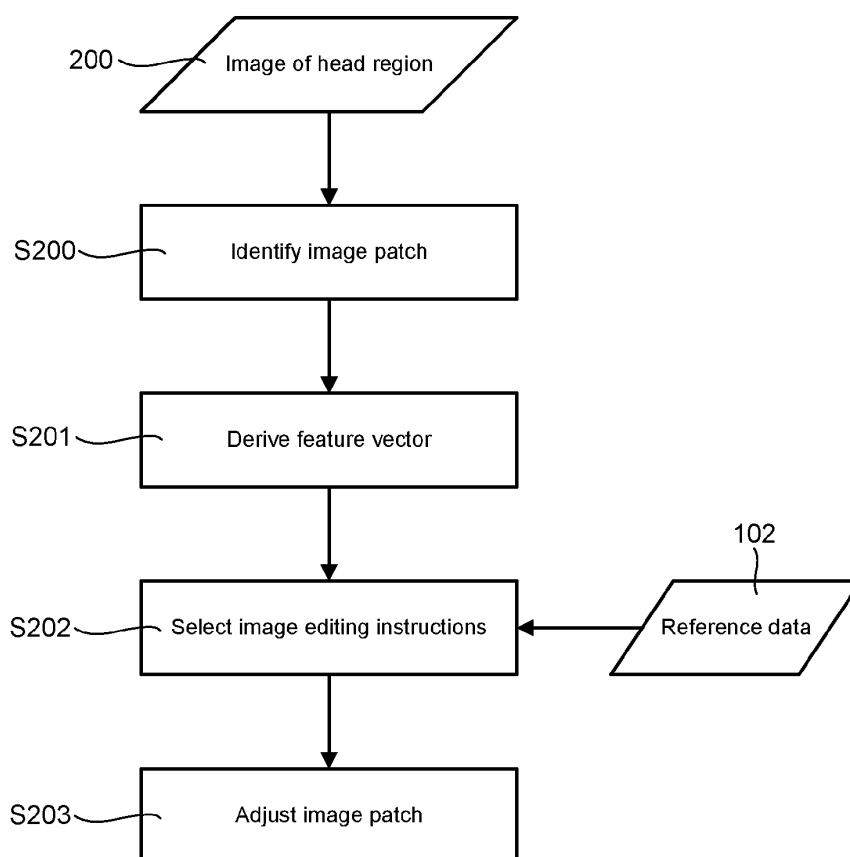


Fig.12

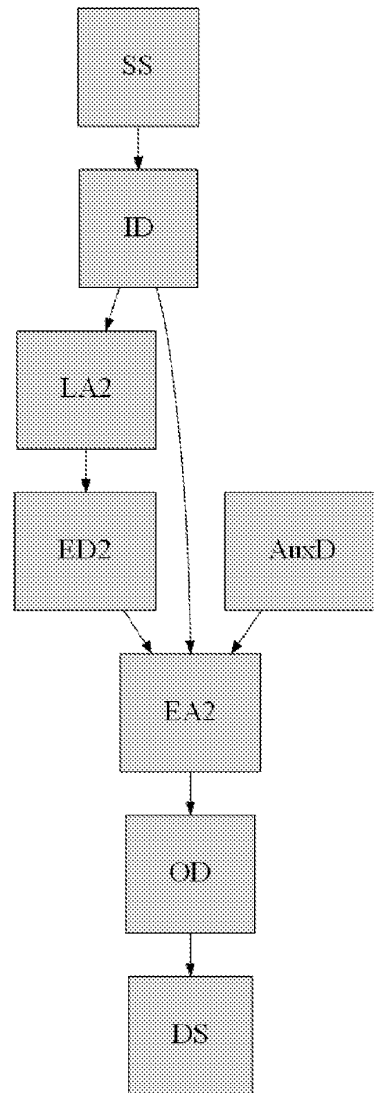
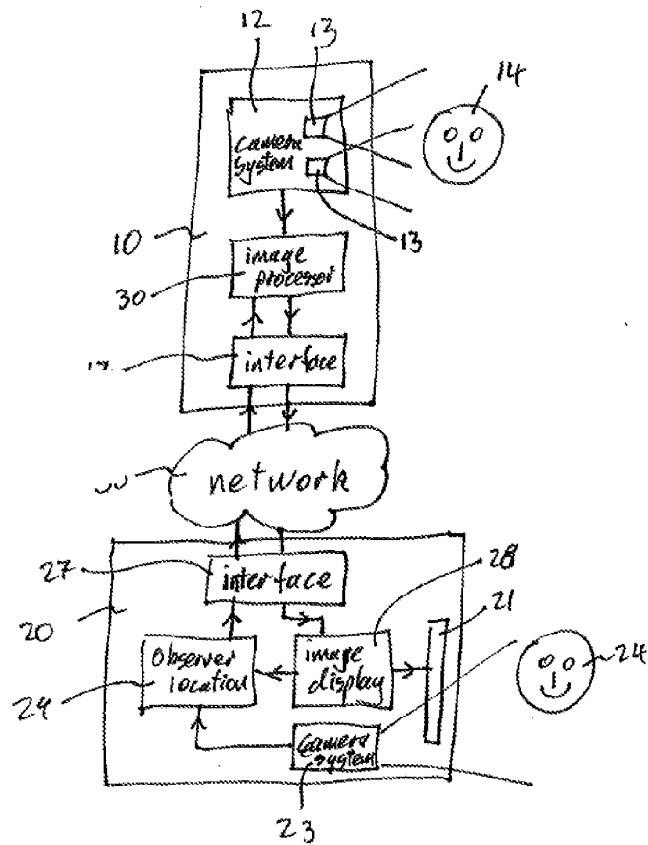


Fig. 13



Generating Reference Data for Adjusting an Image of a Head Region

TECHNICAL FIELD

[01] This application relates to generating reference data for adjusting an image of a head region, particularly a facial and/or neck region, and to adjusting an image of a head region using the reference data or other reference data. The application relates particularly to adjusting target features of the image of the head region, for example to correct a perceived gaze direction of eyes, or to modify the texture and/or shape of features such as the nose, mouth, chin or neck.

BACKGROUND

[02] In many systems, images of a head, which may comprise single images at different times, or stereoscopic pairs of images or other multi-view images, may be captured in one device and displayed on a different device for viewing by an observer. One non-limitative example is a system for performing teleconferencing between two telecommunications devices. In that case, each device may capture images of the observer of that device and transmit them to the other device over a telecommunications network for display and viewing by the observer of the other device.

[03] When an image of a head is captured and displayed, the gaze of the head may not be directed at the observer. This may be caused for example by the gaze of the head not being directed at the camera system used to capture the image, for example because the user whose head is imaged is observing a display in the same device as the camera system and the camera system is offset above (or below) that display. In that case, the gaze in the displayed images will be perceived to be downwards (or upwards). The human visual system has evolved high sensitivity to gaze during social interaction, using cues gained from the relative position of the iris and white sclera of other observers. Errors in the perceived gaze are disconcerting. For example in a system for performing teleconferencing, errors in the perceived gaze can create unnatural interactions between the users.

[04] The present disclosure is concerned with image processing techniques for adjusting images of a head region to correct the perceived gaze and/or to improve other aspects of the images. The present disclosure is particularly concerned with implementing such image processing techniques in real time with minimal demands on computer hardware and/or power.

BRIEF SUMMARY

[05] According to a first aspect of the present disclosure, there is provided a method of generating reference data for adjusting an image of a head region, the method comprising: receiving training data comprising: a set of input image patches, each input image patch comprising a target feature of an image of a head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target feature of the image of the head region after adjustment of the image of the head region; using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region; and using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region.

[06] The described use of two machine learning algorithms allows an improved balance to be achieved between accuracy of the adjustment process and speed of execution. The first machine learning algorithm can be configured to provide highly detailed first reference data, which provides high accuracy. Use of this first reference data directly in an image processing technique to adjust an image of a head region would be relatively expensive in terms of computational resources because of the high level of detail. By arranging instead for the first reference data to be provided to a second machine learning algorithm, which provides, based on the first reference data and the original training data, the reference data (the second reference data) that is to be used for the adjustment of the image of the head region, it is possible to benefit to an extent from the high accuracy of the first machine learning algorithm whilst also providing reference data that is less detailed and thus easier to process efficiently when performing the adjustment of the image of the head region. The quality of the reference data provided by the second machine learning algorithm is found to be significantly improved when the second machine learning algorithm is provided with both the first reference data and the training data in comparison to when the second machine learning algorithm is provided only with the training data.

[07] In an embodiment, efficiency is further improved by providing image editing instructions in the second reference data in a compressed representation. The use of a compressed representation reduces data storage and bandwidth requirements during use of the image editing instructions to perform adjustment of an image of a head region.

[08] According to a second aspect of the present disclosure, there is provided a device configured to perform a similar method to the first aspect of the invention.

[09] According to a third aspect of the present disclosure, there is provided a method of adjusting an image of a head region, the method comprising: identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region; deriving a feature vector from plural local image descriptors of the image patch; using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein: the image editing instructions in the reference data are provided in a compressed representation.

[10] According to a fourth aspect of the present disclosure, there is provided a device configured to perform a similar method of the third aspect of the invention.

[11] The use of a compressed representation reduces data storage and bandwidth requirements.

BRIEF DESCRIPTION OF THE DRAWINGS

[12] Non-limitative embodiments are illustrated by way of example in the accompanying figures, in which like reference numbers indicate similar parts, and in which:

[13] **FIGURE 1** is a schematic perspective view of a device that captures a stereoscopic pair of images;

[14] **FIGURE 2** is a schematic perspective view of a device that displays the stereoscopic pair of images;

[15] **FIGURE 3** is a flow chart of a method of adjusting a stereoscopic pair of images;

[16] **FIGURE 4** is a diagram illustrating the processing of the stereoscopic pair of images in the method of FIGURE 3;

[17] **FIGURE 5** is a flow chart of a step of extracting an image patch;

[18] **FIGURE 6** and **FIGURE 7** are flow charts of two alternatives for a step of adjusting an image;

[19] **FIGURE 8** is flow chart of a method of generating reference data;

[20] **FIGURE 9** schematically depicts data flow in an example method of generating reference data;

[21] **FIGURE 10** depicts a device for generating reference data;

[22] **FIGURE 11** is a flow chart of a method of adjusting an image of a head region;

[23] **FIGURE 12** schematically depicts data flow in an example of a method of adjusting an image of a head region;

[24] **FIGURE 13** is a diagram of a telecommunications system in which the method may be implemented.

DETAILED DESCRIPTION

[25] **FIGURE 1** and **FIGURE 2** illustrate how incorrect gaze is perceived when a stereoscopic pair of images of a head is captured by the device 10 shown in **FIGURE 1** which will be referred to as the source device 10 and displayed on a different device 20 shown in **FIGURE 2** which will be referred to as the destination device 20. Capturing of a stereoscopic pair of images is shown as an example. A similar effect can occur when a monocular image is captured and when more than two views of the head are captured.

[26] The capture device 10 includes a display 11 and a camera system 12. In this particular example, the camera system comprises two cameras 13 in order to capture the stereoscopic pair of images of the head of a source observer 14. In a monocular implementation a single camera may be provided instead of the two cameras 13. The source observer 14 views the display 11, along line 15. The cameras 13 of the camera system 12 are offset from the display 11, in this case being above the display 11. Thus, the cameras 13 effectively look down on the source observer 14 along line 16.

[27] The display device 20 includes a display 21, which in this example can be a stereoscopic display of any known type, for example an autostereoscopic display of any known type. The display 21 displays the stereoscopic pair of images as captured by the capture device 10. A destination observer 24 views the display 21. If the destination observer 24 is located in a normal viewing position perpendicular to the center of the display 21, as shown by the hard outline of the destination observer 24, then the gaze of the source observer 14 is perceived by the

destination observer 24 to be downwards, rather than looking at the destination observer 24, because the cameras 13 of the source device 10 look down on the source observer 14.

[28] Although the cameras 13 are above the display 11 in this example, the cameras 13 could in general could be in any location adjacent the display 11, and the gaze of the source observer 14 perceived by the destination observer 24 would be correspondingly incorrect.

[29] If the destination observer 24 is located in an offset viewing position, as shown by the dotted outline of the destination observer 24 so that the destination observer 24 views the display 21 along line 26, then the offset of the destination observer 24 creates an additional error in the gaze of the source observer 14 perceived by the destination observer 24. A similar additional error in the perceived gaze of the source observer 14 occurs if the destination observer 24 is located in the normal viewing position along line 25, but the displayed image (or stereoscopic pair of images in this example) is displayed on the display 25 in a position offset from the center of the display 25.

[30] A stereoscopic pair of images is an example of multi-view images where there are two images. Although FIGURE 1 illustrates an example where the camera system 12 includes two cameras 13 that capture of a stereoscopic pair of images, alternatively the camera system may include more than two cameras 13 that capture more than two multi-view images, in which case similar issues of incorrect perceived gaze exist on display. The camera system may alternatively include only one camera.

[31] **FIGURE 3** illustrates a method of adjusting multi-view images to correct such errors in the perceived gaze. The method of a FIGURE 3 is a specific example of a method of adjusting images of a head region in the case where the images comprise one or more multi-view images of the head region. For simplicity, this method will be described with respect to the adjustment of multi-view images comprising a stereoscopic pair of images. The method may be generalized to multi-view images comprising more than two images, simply by performing the similar processing on a larger number of images. The method may also be generalized to the case where single view (monocular) images of the head region are used.

[32] The method may be performed in an image processor 30. The image processor 30 may be implemented by a processor executing a suitable computer program or by dedicated hardware or by some combination of software and hardware. Where a computer program is used, the computer program may comprise instructions in any suitable language and may be stored on a

computer readable storage medium, which may be of any type, for example: a recording medium which is insertable into a drive of the computing system and which may store information magnetically, optically or opto-magnetically; a fixed recording medium of the computer system such as a hard drive; or a computer memory.

[33] The image processor 30 may be provided in the source device 10, the destination device 10 or in any other device, for example a server on a telecommunications network, which may be suitable in the case that the source device 10 and the destination device 10 communicate over such a telecommunications network.

[34] In this example, a stereoscopic pair of images 31 are captured by the camera system 12. Although the camera systems 12 is illustrated in FIGURE 1 as including two cameras 13, this is not limitative and more generally the camera system 13 may have the following properties.

[35] The camera system comprises a set of one or more cameras 13, with at least two cameras 13 in the case where multi-view images are processed. Where two cameras are provided, the cameras are typically spaced apart by a distance less than the average human intrapupilar distance. In the alternative that the method is applied to more than two multi-view images, then there are more than two cameras 13, that is one camera 13 per image.

[36] Where plural cameras 13 are provided, the cameras 13 are spatially related to each other and the display 11. The spatial relationship between the cameras 13 themselves and between the cameras 13 and the display 11 is known in advance. Known methods for finding the spatial relationship may be applied, for example a calibration method using a reference image, or specification *a priori*.

[37] The camera or cameras 13 face in the same direction as the display 11. Thus, when the source observer 14 is viewing the display 11, then the camera or cameras 13 face the source observer 14 and the captured image or images (e.g. stereoscopic pair of images) are images of the head of the source observer 14. Different cameras in the camera system can have different fields of view.

[38] The camera system 12 may include cameras 13 having different sensing modalities, including visible light and infrared.

[39] The main output of the camera system 13 is images 31 which are typically video images output at a video rate. The output of the camera system 13 may also include data representing the spatial relationship between the cameras 13 and the display 11, the nature of the sensing

modalities and internal parameters of the cameras 13 (for example focal length, optical axis) which may be used for angular localization.

[40] An example of the method performed on a stereoscopic pair of images 31, for the case of adjustment of eye regions, is as follows. To illustrate this example method, reference is also made to **FIGURE 4** which shows an example of the stereoscopic pair of images 31 at various stages of the method.

[41] In step S1, the stereoscopic pair of images 31 are analyzed to detect the location of the head and in particular the eyes of the source observer 14 within the stereoscopic pair of images 31. This is performed by detecting presence of a head, tracking the head, and localizing the eyes of the head. Step S1 may be performed using a variety of techniques that are known in the art.

[42] One possible technique for detecting the presence of the head is to use Haar feature cascades, for example as disclosed in Viola and Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", CVPR 2001, pp 1-9 (incorporated herein by reference).

[43] One possible technique for tracking the head is to use the approach of Active Appearance Models to provide the position of the head of the subject, as well as the location of the eyes, for example as disclosed in Cootes et al., "Active shape models - their training and application", Computer Vision and Image Understanding, 61(1):38-59, Jan. 1995 and in Cootes et al. "Active appearance models", IEEE Trans. Pattern Analysis and Machine Intelligence, 23(6):681-685, 2001 (incorporated herein by reference).

[44] In step S1, typically, a set of individual points ("landmarks") are set to regions of the face, typically the eyes, for example corners of the eye, upper and lower lid locations, etc, thereby localizing the eyes.

[45] In step S2, image patches containing the left and right eyes of the head, respectively are identified in each image 31 of the stereoscopic pair. **FIGURE 4** shows the identified image patches 32 of the right eye in each image 31 (the image patches for the left eye being omitted in **FIGURE 4** for clarity).

[46] Step S2 may be performed as shown in **FIGURE 5**, as follows.

[47] In step S2-1, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, as shown for example in **FIGURE 4**.

[48] In step S2-2, the image patches 32 identified in step S2-1 are transformed into a normalized coordinate system, being the same normalized coordinate system as used in the machine learning process which is described further below. The transformation is chosen to align the points (“landmarks”) of the eye within the image patch that were identified in step S1, with predetermined locations in the normalized coordinate system. The transformation may include translation, rotation and scaling, to appropriate extents to achieve that alignment. The output of step S2-2 is identified image patches 33 of the right eye in each image in the normalized coordinate system as shown for example in FIGURE 4.

[49] The following steps are performed separately (a) in respect of the image patches containing the left eyes of the head in each image 31 of the stereoscopic pair, and (b) in respect of the image patches containing the right eyes of the head in each image 31 of the stereoscopic pair (in this example). For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes.

[50] In step S3, a feature vector 34 is derived from plural local image descriptors of an image patch 33 in at least one image 31 of the stereoscopic pair (in this example). Depending on the approach and as described further below, this may be an image patch in a single image 31 of the stereoscopic pair or may be both images 31 of the stereoscopic pair. Thus, the local image descriptors are local image descriptors derived in the normalized coordinate system.

[51] The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 to be used for adjusting the image patches. The reference data 35 may comprise reference displacement vector fields that represent transformations of the image patch, or other representations of transformations of the image patch, including compressed representations as described below, and are associated with possible values of the feature vector.

[52] The reference data 35 is obtained and analyzed in advance using a machine learning technique. The machine learning technique may derive the form of the feature vectors 34 and associate transformations such as the reference displacement vector fields with the possible values of the feature vector. A specific example of a machine learning technique applied in the case where it is desired to correct gaze will now be described before reverting to the method of FIGURE 3.

[53] The training input to the machine learning technique is two sets of images (or image patches), which may be stereoscopic pairs of images or monoscopic images, as discussed further below. Each set may comprise images of the head of the same group of individuals but captured from cameras in different locations relative to the gaze so that the perceived gaze differs as between them (in the case where gaze is to be corrected).

[54] The first set are input images, being images of each individual with an incorrect gaze where the error is known *a priori*. In particular, the images in the first set may be captured by at least one cameras in a known camera location where the gaze of the individual which is in a different known direction. For example in the case of the source device of FIGURE 1, the camera location may be the location of a camera 13 and while the gaze of the imaged individual is towards the center of the display 11.

[55] The second set are output images, being images of each individual with correct gaze for a predetermined observer location relative to a display location in which the image is to be displayed. In the simplest case, the observer location is a normal viewing position perpendicular to the center of the display location, for example as shown by the hard outline of the destination observer 24 in the case of the destination device 20 of FIGURE 2.

[56] For each image in the two sets, the image is analyzed to detect the location of the head and in particular the eyes using the same technique as used in step S1 described above, and then image patches containing the left and right eyes of the head, respectively, are identified using the same technique as used in step S2 described above. The following steps are then performed separately (a) in respect of the image patches containing the left eyes of the head in each image, and (b) in respect of the image patches containing the right eyes of the head in each image. For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes in this embodiment.

[57] Each image patch is transformed into the same normalized coordinate system as used in step S2 described above. As described above, the transformation is chosen to align points (“landmarks”) of the eye with predetermined locations in the normalized coordinate system. The transformation may include translation, rotation and scaling, to appropriate extents to achieve that alignment.

[58] Thus, the image patches input and output images of each individual are aligned in the

normalized coordinate system.

[59] From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image, for example as follows. Defining positions in the image patches by (x,y) , the displacement vector field $\mathbf{F}$ is given by

$$\mathbf{F}=\{u(x, y), v(x, y)\}$$

where u and v define the horizontal and vertical components of the vector at each position (x,y) .

[60] The displacement vector field $\mathbf{F}$ is chosen so that the image patch of the output image $O(x,y)$ is derived from the image patch of the input image $I(x,y)$ as

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[61] For image data from more than one camera, the system delivers a displacement vector field for the input image from each camera.

[62] The displacement vector field $\mathbf{F}$ for an input and output image of an individual may be derived using a process in which a trial feature vector $\mathbf{F}'=\{u',v'\}$ is modified to minimize error, optionally in an iterative process, for example in accordance with:

$$\sum |O(x, y) - I(x + u'(x, y), y + v'(x, y))| = \min!$$

[63] By way of non-limitative example, the displacement vector field $\mathbf{F}$ may be derived as disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675 (incorporated herein by reference), wherein the displacement vector field $\mathbf{F}$ is referred to as a “flow field”.

[64] Another example for image editing instructions, which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by filter field $L = \{k(P, x, y)\}$, which defines a filter kernel for a given location (x, y) . This filter field L is chosen so that the image patch of the output image $O(x, y)$ is derived from the image patch of the input image $I(x, y)$ as $O(x, y) = k(P(I, x, y), x, y)$, where $P(I, x, y)$ is a local region around the point (x,y) in the image I , and $k(P, x, y)$ operates on the patch P with coefficients depending on the position (x, y) . For example, it could be the output of a convolution of the patch with a Gaussian filter with width depending on the position x in the image, or a brightness increase of a local pixel depending on the vertical position y .

[65] Another example for image editing instructions which, can be used additionally or as an

alternative to the displacement vector field in any of the arrangements disclosed herein, is given by a set of typical image components that make up the edited image area, e.g. texture showing beard stubbles. These are then blended with a factor depending on the image coordinates and local image content (i.e. a texture blending field). Other transformation fields can be used, such as a brightness adjustment field.

[66] A machine learning technique is used to obtain a map from the displacement vector field F (or other image editing instructions) of each individual to respective feature vectors derived from plural local image descriptors of the image patch of the input image.

[67] The local descriptors capture relevant information of a local part of the image patch of the input image and the set of descriptors usually form a continuous vectorial output.

[68] The local image descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific local image descriptors are selected and optimized by the machine learning process itself. In general, the local image descriptors may be of any suitable type, some non-limitative examples which may be applied in any combination being as follows.

[69] The local image descriptors may include values of individual pixels or a linear combination thereof. Such a linear combination may be, for example, a difference between the pixels at two points, a kernel derived within a mask at an arbitrary location, or a difference between two kernels at different locations.

[70] The local image descriptors may include distances of a pixel location from the position of an eye point (“landmark”).

[71] The local image descriptors may include SIFT features (Scale-invariant feature transform features), for example as disclosed in Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision 60 (2), pp 91–110 (incorporated herein by reference).

[72] The local image descriptors may include HOG features (Histogram of Oriented Gradients features), for example as disclosed in Dalal et al. “Histograms of Oriented Gradients for Human Detection”, Computer Vision and Pattern Recognition, 2005, pp. 886-893 (incorporated herein by reference).

[73] The local image descriptors may include “low level representations” from pre-classification stages in deep learning neural networks, for example as disclosed in Yang and

Ramanan, “Multi-scale recognition with DAG-CNNs”, ICCV 2015. In a classifying deep learning neural network with multiple layers applied to an input image, for example, such low level features could be taken from a layer before the final classification layer of the network.

[74] The derivation of the feature vector from plural local image descriptors depends on the type of machine learning applied.

[75] In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field **F** derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field **F** to be simply selected from the reference data 35 on the basis of the feature vector 34 derived in step S3, as described below.

[76] In the case that the feature vector comprises features that are binary values derived from the local image descriptors, the feature vector has a binary representation. Such binary values may be derived in various ways from the values of descriptors, for example by comparing the value of a descriptor with a threshold, comparing the value of two descriptors, or by comparing the distance of a pixel location from the position of an eye point (“landmark”).

[77] Alternatively, the feature vector may comprise features that are discretized values of the local image descriptors. In this case, more than two discrete values of each feature are possible.

[78] Any suitable machine learning technique may be applied, for example using a decision tree, a decision forest, a decision fern or an ensemble or combination thereof, or a neural network.

[79] By way of example, a suitable machine learning technique using a feature vector comprising features that are binary values derived by comparing a set of individual pixels or a linear combination thereof against a threshold, is disclosed in Ozuysal et al. “Fast Keypoint Recognition in Ten Lines of Code”, Computer Vision and Pattern Recognition, 2007, pp. 1-8 (incorporated herein by reference).

[80] By way of further example, a suitable machine learning technique using a distance of a pixel location with the position of an eye landmark is disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision

and Pattern Recognition, 2015, pp. 4667-4675 (incorporated herein by reference).

[81] By way of further example, a suitable machine learning technique using a random decision forest is disclosed in Ho, “Random Decision Forests”, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995, pp. 278-282 (incorporated herein by reference).

[82] In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field **F** (in this example, but other image editing instructions could be used) derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field **F** to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived in step S3 and the values of the feature vector associated with the reference displacement vector fields.

[83] Any suitable machine learning technique may be applied, for example using support vector regression.

[84] By way of example, a suitable machine learning technique using support vector regression is disclosed in Drucker et al. “Support Vector Regression Machines”, Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, (incorporated herein by reference). The output of the technique is a continuously varying set of interpolation directions that form part of the reference data 35 and are used in the interpolation.

[85] The machine learning technique, regardless of its type, inherently also derives the form of the feature vectors 34 that is used to derive the reference displacement vector fields **F** (or other image editing instructions). This is the form of the feature vectors 34 that is derived in step S3.

[86] The description now reverts to the method of FIGURE 3.

[87] In step S4, at least one displacement vector field 37 representing a transformation of an image patch is derived by using the feature vector 34 derived in step S3 is to look up the reference data 35. Due to the derivation of the displacement vector field 37 from the reference data 35, the transformation represented thereby corrects the gaze that will be perceived when the stereoscopic pair of images 31 are displayed.

[88] In the case that the feature vector 34 comprises features that are values in a discrete space and the reference displacement vector fields of the reference data 35 comprise a reference

displacement vector field associated with each possible value of the feature vector in the discrete space, then the displacement vector field for the image patch is derived by selecting the reference displacement field associated with the actual value of the derived feature vector 34.

[89] In the case that the feature vector 34 comprises features that are discrete values of the local image descriptors in a continuous space, then the displacement vector field for the image patch is derived by interpolating a displacement vector field from the reference displacement vector fields based on the relationship between the actual value of the derived feature vector 34 and the values of the feature vectors associated with the reference displacement vector fields. In the case that the machine learning technique was support vector regression, This may be done using the interpolation directions that form part of the reference data 35.

[90] In step S5, each image 31 of the stereoscopic pair is adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector fields 37. This produces an adjusted stereoscopic pair of images 38 as shown in FIGURE 4, in which the gaze has been corrected. In particular, the adjustment may be performed using two alternative methods, as follows.

[91] A first method for performing step S5 is shown in **FIGURE 6** and performed as follows.

[92] In step S5-1, the image patch is transformed in the normalised coordinate system in accordance with the corresponding displacement vector field 37 in respect of the same image, thereby correcting the gaze. As described above, for a displacement vector field **F** the transformation of the image patch of the input image $I(x,y)$ provides the output image $O(x,y)$ in accordance with

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[93] In step S5-2, the transformed image patch output from step S5-1 is transformed out of the normalised coordinate system, back into the original coordinate system of the corresponding image 31. This is done using the inverse transformation from that applied in step S2-2.

[94] In step S5-3, the transformed image patch output from step S5-2 is superimposed on the corresponding image 31. This may be done with a full replacement within an eye region corresponding to the eye itself, and a smoothed transition between the transformed image patch and the original image 31 over a boundary region around the eye region. The width of the boundary region may be of fixed size or a percentage of the size of the image patch in the original image 31.

[95] A second method for performing step S5 is shown in **FIGURE 7** and performed as follows.

[96] In this second, alternative method, the transformation back into the coordinate system of the corresponding image 31 occurs before the transformation of the image patch in accordance with the transformed displacement vector field **F**.

[97] In step S5-4, the displacement vector field **F** is transformed out of the normalised coordinate system, back into the original coordinate system of the corresponding image 31. This is done using the inverse transformation from that applied in step S2-2.

[98] In step S5-5, the image patch 32 in the coordinate system of the image 31 is transformed in accordance with the displacement vector field **F** that has been transformed into the same coordinate system in step S5-4. As described above, for a displacement vector field **F** the transformation of the image patch of the input image $I(x,y)$ provides the output image $O(x,y)$ in accordance with

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

but this is now performed in the coordinate system of the original image 31.

[99] Step S5-6 is the same as S5-3. Thus, in step S5-6, the transformed image patch output from step S5-5 is superimposed on the corresponding image 31. This may be done with a full replacement within an eye region corresponding to the eye itself, and a smoothed transition between the transformed image patch and the original image 31 over a boundary region around the eye region. The width of the boundary region may be of fixed size or a percentage of the size of the image patch in the original image 31.

[100] **FIGURE 8** depicts a method of generating reference data (including reference data 37 of the type described in the specific examples discussed above) for adjusting an image of a head region. The method comprises a step S100 of receiving training data 100. The training data 100 may be provided for example via a communications interface 112 (e.g. connecting to an external data connection or storage device) connected to a data processing unit 120 that is to perform the method (as depicted schematically in **FIGURE 10**). The training data 100 comprises a set of input image patches. Each input image patch comprises a target feature of an image of a head region prior to adjustment of the image of the head region. The target feature is the same for each input image patch. The target feature may comprise one or more of the following: an eye region comprising at least part of an eye (as in the specific examples discussed above with

reference to FIGURE 4 for example), a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, and a neck region comprising at least part of a neck. In an embodiment each input image patch comprises a portion of the image of the head region corresponding to the target feature, such as the above-mentioned eye region, nose region, mouth region, chin region, or neck region.

[101] In an embodiment, the target feature comprises an eye region comprising at least part of an eye and the adjustment of the image of the head region comprises adjusting a gaze direction.

[102] In an embodiment, the target feature comprises a nose region comprising at least part of a nose and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the nose (e.g. to make nose look smaller and/or slimmer by a fixed proportion, similar to the effect of a “tele lens”).

[103] In an embodiment, the target feature comprises a chin region comprising at least part of a chin and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the chin (e.g. to reduce or remove double chin appearance).

[104] In an embodiment, the target feature comprises a neck region comprising at least part of a neck and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the neck (e.g. to reduce or remove wrinkles).

[105] In an embodiment, the target feature comprises a hair region comprising hair and the adjustment of the image of the head region comprises adjusting a color of the hair (e.g. by a fixed hue).

[106] In the case where the target feature comprises an eye region, the input image patches may be obtained using the methodology described above with reference to steps S2-1 and S2-2, except that stereoscopic pairs of images are not necessarily required. Thus, the input image patches may be obtained by using identified points (“landmarks”) corresponding to features of an eye to locate the relevant region of the image of the head region (e.g. surrounding the eye) and/or transformation of the input image patch into a normalized coordinate system, including alignment of the landmarks with predetermined locations in the normalized coordinate system using translation, rotation and/or scaling.

[107] The training data 100 further comprises a set of output image patches. The output image patches are in one-to-one correspondence with the input image patches. Each output image patch comprises the target feature of the image of the head region after adjustment of the image

of the head region. Thus, in the case where an image of the head region is to be adjusted to correct a gaze direction, each output image patch comprises an eye region that has been adjusted so that the gaze appears to be in the desired direction (e.g. directly towards a destination observer 24).

[108] The method further comprises a step S101 comprising using a first machine learning algorithm to generate first reference data 101 using the training data 100. The first reference data 101 comprises image editing instructions for adjusting the image of the head region for a range of possible images of the head region.

[109] The method further comprises a step S102 comprising using a second machine learning algorithm to generate second reference data 102. The second machine learning algorithm uses the same training data 100 as the first machine learning algorithm in step S101. The second machine learning algorithm further uses the first reference data 101 output by the first machine learning algorithm in step S101. The second reference data 102 comprises image editing instructions for adjusting the image of the head region for a range of possible images of the head region.

[110] In an embodiment, the first reference data 101 comprises first image editing instructions for a range of possible configurations of the target feature (e.g. different gaze directions and/or eye morphologies in the case where the target feature comprises an eye region) and first selection instructions for selecting image editing instructions (from the first image editing instructions) for a particular input image patch based on the configuration of the target feature of the input image patch (e.g. the particular gaze direction and/or particular eye morphology of that input image patch).

[111] In an embodiment, the second reference data comprises second image editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting image editing instructions (from the second image editing instructions) for a particular input image patch based on the configuration of the target feature of the input image patch.

[112] The configuration of the target feature of each input image patch may be represented by a feature vector derived from plural local image descriptors of the input image patch, as described above with reference to step S3 of FIGURE 3 for the particular case where stereoscopic pairs of images containing eye regions are processed (but the method is applicable more generally than this particular case). As described above, the feature vector may take various forms but is

generally adapted to be suitable for looking up image editing instructions for performing adjustment of the image of the head region. In the present embodiment, the first and second selection instructions define how the feature vector is used to select image editing instructions for the input image patch. In an embodiment, the image editing instructions comprise a displacement vector field defining how the input image patch is to be transformed to perform the adjustment. The displacement vector field may take any of the forms discussed above. The image editing instructions are not limited to displacement vector fields, however. Other image editing operations may additionally or alternatively be associated with the features vectors to perform other desired adjustments of the image of the head region, including for example adjustments to pixel colors or intensities, or changes to underlying geometries (e.g. via a filter field, brightness adjustment field, or texture blending field).

[113] In an embodiment, a first image editing algorithm is used by the first machine learning algorithm to define how the first image editing instructions are to be applied to an input image patch to derive an output image patch. A second image editing algorithm is used by the second machine learning algorithm to define how the second image editing instructions are to be applied to an input image patch to derive an output image patch. The first and second image editing algorithms may comprise any of the methods described above for implemented step S5 of FIGURE 3, described with reference to FIGURES 6 and 7.

[114] In an embodiment, the second image editing instructions in the second reference data are provided in a compressed representation. The compressed representation may comprise a principle component analysis representation or a wavelet representation for example. In this case, the first and second image editing algorithms may be adapted to define how to operate efficiently in this context.

[115] In an embodiment, the second image editing instructions are principle component analysis components of a principle component analysis of the first image editing instructions. The second image editing algorithm in this case transforms the second image editing instructions into the first image editing instructions by inverse principle component analysis transform.

[116] In an alternative approach, the second image editing instructions are wavelet components of the first image editing instructions. The second image editing algorithm in this case transforms the second image editing instructions into the first image editing instructions by

inverse wavelet transform.

[117] In an embodiment, the first selection instructions for the first reference data are able to select between a larger number of alternative image editing instructions than the second selection instructions for the second reference data. The first machine learning algorithm may thus be described as having more input parameters than the second machine learning algorithm. The first machine learning algorithm may provide higher accuracy than the second machine learning but will typically operate considerably slower. Additionally, the first selection instructions may be significantly more complex (e.g. involving linear algebra or other relatively computer intensive operations) than the second selection instructions (where the selection instructions may resemble a computationally straightforward look-up table, a combination of a look-up table and a tree structure, or similar).

[118] In one particular embodiment, the first machine learning algorithm comprises a neural network (known to provide relatively high accuracy, but at the expense of relatively high computational demands). In such an embodiment, and others, the second machine learning algorithm may comprises a regression forest (known to provide higher computational efficiency, but at the expense of reduced accuracy). The inventors have found that the combination of the two different machine learning algorithms provides reference data that can be used in an image adjustment method with high efficiency while still achieving high adjustment accuracy. The first machine learning algorithm may alternatively comprise a support vector machine or a generative adversarial network (GAN). The second machine learning algorithm may alternatively comprise regression ferns, cluster centres, a lookup table, or separable filter banks. In one embodiment, the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.

[119] **FIGURE 9** schematically depicts data flow in a detailed example of the method of generating reference data of **FIGURE 8**. The first and second machine learning algorithms are respectively labelled MLA1 and MLA2. The first machine learning algorithm MLA1 receives the training data (labelled TD) and, optionally, the first machine learning algorithm EA1, and a similarity metric SM. The similarity metric SM provides a numerical value to measure similarity between an adjusted image and a desired image and can be used to control the first machine learning algorithm MLA1 and the second machine learning algorithm MLA2 to vary the extent

to which differences are penalized according to the nature of the differences. For example, the similarity metric SM may be configured to penalize reductions in portions of images that it is desired to maintain (e.g. sclera in the case where eye regions are being adjusted) or deviations of adjusted features from a known form (e.g. deviations in the shape of an iris from an elliptical form) or from a form that is observed in training data. Alternatively or additionally, the training data may comprise an average absolute or square difference between the adjusted image and a target, or average absolute or square difference between low level representations of the adjusted image and target, such as low level features from a deep learning network (as discussed above). In the example shown, the first machine learning algorithm EA1 receives auxiliary data AuxD, which defines a basis set used for providing a compressed representation of image editing instructions. The first machine learning algorithm outputs first selection instructions LA1 and first image editing instructions ED1. The second machine learning algorithm MLA2 receives the same training data TD and, optionally, the second machine learning algorithm EA2, and the similarity metric SM. The second machine learning algorithm MLA2 additionally receives the first image editing instructions ED1. The second machine learning algorithm outputs second selection instructions LA2 and second image editing instructions ED2.

[120] The second machine learning algorithm MLA2 thus gets the editing instructions to match or to approximate, and does not have to infer these from the matched input images as MLA1.

[121] **FIGURE 10** depicts a device 110 for generating the reference data. The device 110 comprises a data processing unit 120 configured to perform the method of generating the reference data according to any of the disclosed embodiments. The data processing unit 110 may be implemented by a processor executing a suitable computer program or by dedicated hardware or by some combination of software and hardware. Data input/output may be provided via a communications interface 112. Where a computer program is used, the computer program may comprise instructions in any suitable language and may be stored on a computer readable storage medium, which may be of any type, for example: a recording medium which is insertable into a drive of the computing system and which may store information magnetically, optically or opto-magnetically; a fixed recording medium of the computer system such as a hard drive; or a computer memory.

[122] **FIGURE 11** depicts a method of adjusting an image of a head region. The method may use reference data generated using any of the methods of generating reference data disclosed

herein or may use reference data generated using other methods.

[123] The method comprises a step S200 in which an image patch is identified in an image 200 of a head region that is to be adjusted. The image patch comprises a target feature of the head region. The target feature may take any of the forms discussed above. The step S200 may optionally comprise detecting a head and/or eye location as described above with reference to step S1 of FIGURE 3. The step S200 may optionally further comprise identifying image patches using the methodology described above with reference to steps S2-1 and S2-2 of FIGURE 5, except that the image patches do not necessarily need to be identified as stereoscopic pairs (although they may be if desired).

[124] In step S201, a feature vector is derived from plural local image descriptors of the image patch. The feature vector may be derived using the methodology described above with reference to step S3 of FIGURE 3. The feature vector may take any of the forms discussed above.

[125] In step S202, the feature vector is used to select image editing instructions from reference data 102. The reference data 102 comprising image editing instructions for a range of possible values of the feature vector (representing for example different gaze directions and/or eye morphologies in the case where the target feature comprises an eye region).

[126] In step S203, the selected image editing instructions are applied to the image patch to adjust the image of the head region (e.g. to correct a gaze direction).

[127] In an embodiment, the image editing instructions are provided in a compressed representation, comprising for example one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres. The use of a compressed representation reduces data storage and bandwidth requirements during use of the image editing instructions to perform adjustment of an image of a head region. Alternatively or additionally, the reference data containing the image editing instructions may be generated using any of the embodiments disclosed herein. The reference data may comprise the second reference data discussed above for example.

[128] **FIGURE 12** schematically depicts data flow in a detailed example of a method of adjusting an image of a head region using reference data generated according to the detailed example of FIGURE 9. Input data ID is provided from a sensor system SS (e.g. comprising one or more cameras). The input data ID is input to the second selection instructions LA2 to select image editing instructions appropriate to the input data ID from the second image editing

instructions ED2. The selected image editing instructions, which in this example are provided in a compressed representation (e.g. principle component analysis components or wavelet components) from the compressed second image editing instructions ED2, are then used by the second image editing algorithm EA2, in combination with the auxiliary data AuxD providing the basis set for the compressed representation, to provide output data OD. The output data OD comprises an adjusted image of the head region and is displayed via a display DS.

[129] As described above, the method of adjusting an image of a head region may be implemented in an image processor 30 provided in various different devices. By way of non-limitative example, there will now be described a particular implementation in a telecommunications system which is shown in **FIGURE 13** and arranged as follows.

[130] In this implementation, the source device 10 and the destination device 10 communicate over such a telecommunications network 50. For communication over the telecommunications network 50, the source device 10 includes a telecommunications interface 17 and the destination device 20 includes a telecommunications interface 27.

[131] In this implementation, the image processor 30 is provided in the source device 10 and is provided with an image of a head region directly from a camera system 12 (in this example, a stereoscopic pair of images). The telecommunications interface 17 is arranged to transmit the adjusted images 38 over the telecommunications network 50 to the destination device 20 for display thereon.

[132] The destination device 20 includes an image display module 28 that controls the display 26. The adjusted images 38 are received in the destination device 20 by the telecommunications interface 27 and supplied to the image display module 28 which causes them to be displayed on the display 26.

[133] The following elements of the destination device 20 are optionally included in the case that the method corrects gaze for a destination observer 24 in an observer location other than a normal viewing position perpendicular to the center of the display location. In this case, the destination device 20 includes a camera system 23 and an observer location module 29. The camera system 23 captures an image of the destination observer 24. The observer location module 29 derives the location data 40. The observer location module 29 includes a head tracking module that uses the output of the camera system 23 to detect the location of the destination observer 24. The observer location module 29. Where the relative observer location

also takes into account the location of the image displayed on the display 21, the observer location module 29 obtains the location of the image displayed on the display 21 from the image display module 28. The telecommunications interface 17 is arranged to transmit the location data 40 over the telecommunications network 50 to the source device 10 for use thereby.

[134] Although the above description refers to a method applied to images supplied from a source device 10 to a destination device 20, the method may equally be applied to images supplied from in the opposite direction from the destination device 20 to the source device 10, in which case the destination device 20 effectively becomes the “source device” and the source device 10 effectively becomes the “destination device”. Where images are supplied bi-directionally, the labels “source” and “destination” may be applied to both devices, depending on the direction of communication being considered.

[135] While various embodiments in accordance with the principles disclosed herein have been described above, it should be understood that they have been presented by way of example only, and not limitation. Thus, the breadth and scope of this disclosure should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with any claims and their equivalents issuing from this disclosure. Furthermore, the above advantages and features are provided in described embodiments, but shall not limit the application of such issued claims to processes and structures accomplishing any or all of the above advantages.

[136] Additionally, the section headings herein are provided for consistency with the suggestions under 37 CFR 1.77 or otherwise to provide organizational cues. These headings shall not limit or characterize the embodiment(s) set out in any claims that may issue from this disclosure. Specifically and by way of example, although the headings refer to a “Technical Field,” the claims should not be limited by the language chosen under this heading to describe the so-called field. Further, a description of a technology in the “Background” is not to be construed as an admission that certain technology is prior art to any embodiment(s) in this disclosure. Neither is the “Summary” to be considered as a characterization of the embodiment(s) set forth in issued claims. Furthermore, any reference in this disclosure to “invention” in the singular should not be used to argue that there is only a single point of novelty in this disclosure. Multiple embodiments may be set forth according to the limitations of the multiple claims issuing from this disclosure, and such claims accordingly define the embodiment(s), and their equivalents, that are protected thereby. In all instances, the scope of such claims shall be

considered on their own merits in light of this disclosure, but should not be constrained by the headings set forth herein.

Claims:

1. A method of generating reference data for adjusting an image of a head region, the method comprising:

receiving training data comprising:

a set of input image patches, each input image patch comprising a target feature of an image of a head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and

a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target feature of the image of the head region after adjustment of the image of the head region;

using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region; and

using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region.

2. The method of claim 1, wherein:

the first reference data comprise first image editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting image editing instructions for a particular input image patch from the first image editing instructions based on the configuration of the target feature of the input image patch; and

the second reference data comprise second image editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting image editing instructions for a particular input image patch from the second image editing instructions based on the configuration of the target feature of the input image patch.

3. The method of claim 2, wherein the configuration of the target feature of each input image patch is represented by a feature vector derived from plural local image descriptors of the input image patch, and the first and second selection instructions define how the feature vector is used to select image editing instructions for the input image patch.
4. The method of claim 2 or 3, wherein the image editing instructions comprise a displacement vector field defining how the input image patch is to be transformed.
5. The method of claim 2 or 3, wherein the image editing instructions comprise a filter field, a brightness adjustment field, or a texture blending field.
6. The method of any of claims 2-5, wherein:
 - a first image editing algorithm is used by the first machine learning algorithm to define how the first image editing instructions are to be applied to an input image patch to derive an output image patch; and
 - a second image editing algorithm is used by the second machine learning algorithm to define how the second image editing instructions are to be applied to an input image patch to derive an output image patch.
7. The method of claim 6, wherein:
 - the second image editing instructions are principle component analysis components of a principle component analysis of the first image editing instructions; and
 - the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse principle component analysis transform.
8. The method of claim 6, wherein:
 - the second image editing instructions are wavelet components of the first image editing instructions; and
 - the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse wavelet transform.

9. The method of any of claims 2-8, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative image editing instructions than the second selection instructions for the second reference data.
10. The method of any preceding claim, wherein the first machine learning algorithm is of a different machine learning algorithm type than the second machine learning algorithm.
11. The method of any preceding claim, wherein the first machine learning algorithm comprises one or more of the following: a neural network; a support vector machine; a generative adversarial network, GAN.
12. The method of any preceding claim, wherein the second machine learning algorithm comprises one or more of the following: a regression forest; regression ferns, cluster centres, a lookup table, separable filter banks.
13. The method of any of claims 1-10, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.
14. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair.
15. The method of any of claims 1-13, wherein:
the target feature comprises an eye region comprising at least part of an eye and the adjustment of the image of the head region comprises adjusting a gaze direction;

the target feature comprises a nose region comprising at least part of a nose and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the nose;

the target feature comprises a chin region comprising at least part of a chin and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the chin;

the target feature comprises a neck region comprising at least part of a neck and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the image of the head region comprises adjusting a color of the hair.

16. The method of any preceding claim, wherein the second image editing instructions in the second reference data are provided in a compressed representation.

17. The method of claim 16, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.

18. A method of adjusting an image of a head region, the method comprising:

identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;

deriving a feature vector from plural local image descriptors of the image patch;

using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and

applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:

the reference data comprises the second reference data generated by the method of any of claims 1-17.

19. A method of adjusting an image of a head region, the method comprising:
 - identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;
 - deriving a feature vector from plural local image descriptors of the image patch;
 - using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and
 - applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:
 - the image editing instructions in the reference data are provided in a compressed representation.
20. The method of claim 19, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.
21. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.
22. A computer readable storage medium storing a computer program according to claim 21.
23. A device for generating reference data for adjusting an image of a head region, the device comprising a data processing unit arranged to:
 - receive training data comprising:
 - a set of input image patches, each input image patch comprising information about a target feature of an image of the head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and
 - a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target portion of the image of the head region after adjustment of the image of the head region;

use a first machine learning algorithm to generate first reference data using the training data; and

use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

24. A device for adjusting an image of a head region, the device comprising an image processor arranged to process the image of the head region by:

identifying an image patch in the image of the head region, the image patch comprising information about a target feature of the image of the head region;

deriving a feature vector from plural local image descriptors of the image patch;

using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and

applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:

the image editing instructions in the reference data are provided in a compressed representation.

25. The device of claim 24, further comprising a telecommunications interface arranged to transmit the adjusted image over a telecommunications network to a destination device for display thereon.

Abstract

Methods and devices for generating reference data for adjusting an image of a head region, and methods and devices for adjusting the image of a head region are disclosed. In one arrangement, training data are received. A first machine learning algorithm generates first reference data using the training data. A second machine learning algorithm generates second reference data using the same training data and the first reference data generated by the first machine learning algorithm.

Electronic Acknowledgement Receipt	
EFS ID:	30044885
Application Number:	62543587
International Application Number:	
Confirmation Number:	7127
Title of Invention:	Generating reference data for adjusting an image of a head region
First Named Inventor/Applicant Name:	Eric Sommerlade
Customer Number:	145234
Filer:	Penny Lea Lowry/Stefan Bump
Filer Authorized By:	Penny Lea Lowry
Attorney Docket Number:	409000A
Receipt Date:	10-AUG-2017
Filing Date:	
Time Stamp:	15:22:24
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	\$130
RAM confirmation Number	081117INTEFSW15280600
Deposit Account	506262
Authorized User	Stefan Bump
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: 37 CFR 1.16 (National application filing, search, and examination fees) 37 CFR 1.17 (Patent application and reexamination processing fees)	

37 CFR 1.19 (Document supply fees)					
37 CFR 1.20 (Post Issuance fees)					
37 CFR 1.21 (Miscellaneous fees and charges)					
File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	WebADS.pdf	114937	no	7
			e3d188152fa2a5e5f27463cebd8ad071e3156853		
Warnings:					
Information:					
2	Provisional Cover Sheet (SB16)	409000A_ProvisionalSB.pdf	1477146	no	3
			ac68da89a090c8c96badb92700a011ead95b62e2		
Warnings:					
Information:					
3		409000A_PRV_Application.pdf	547583	yes	41
			8ffac0fd9e1da991ecee52dce679027811375569		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Drawings-only black and white line drawings		32	41	
	Specification		1	31	
Warnings:					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	29808	no	2
			550917395b99959f8ea43d41da6e80779aa53676		
Warnings:					
Information:					
Total Files Size (in bytes):			2169474		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



LOWRY, Penny L. et al.
5700 Flatiron Parkway
Boulder, Colorado 80301
ETATS-UNIS D'AMERIQUE

Date
20.12.19

Reference	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Entry into the European phase before the European Patent Office

The following information describes the procedural steps required for entry into the European phase before the European Patent Office (EPO) as designated/elected Office. You are advised to read it carefully because failure to take the necessary action in due time can lead to a loss of rights.

1. The above mentioned international patent application has been given the **European application No. 18844651.2**. This application number is to be used in correspondence with the EPO.
2. Applicants **without a residence or their principal place of business** in an EPC Contracting State may themselves initiate European processing of their international applications, provided they do so before expiry of the 31st month from the priority date.

During the European phase before the EPO as designated or elected Office, however, such applicants **must** be represented by a professional representative (Art. 133(2) and Art. 134(1) and (8) EPC).

Where, at the expiry of the time period laid down in Rule 163(5) EPC, the requirements of Article 133(2) EPC have not been complied with, the European patent application will be **refused**, pursuant to Rule 163(6) EPC.

Please note that a professional representative authorised to act before the EPO and who acted for the applicant during the international phase does not automatically become the representative for the European phase. Applicants are therefore strongly advised to appoint in good time any representative they wish to initiate the European phase for them; otherwise the EPO has to send all communications directly to the applicant.

3. Applicants **with a residence or their principal place of business** in an EPC Contracting State are not obliged to appoint for the European phase a professional representative authorised to act before the EPO. However, in view of the complexity of the procedure it is recommended that they do so.
4. Applicants and professional representatives are also strongly advised to initiate the European phase using **EPO Form 1200**. It is available free of charge from the EPO or via the EPO website at **www.epo.org**. Similarly, it can be generated using the Online Filing software, obtainable free of charge from the EPO, or the new online filing (CMS). The use of the form is not compulsory.

-
5. Applicants wishing to use an address for correspondence which is different from the address of their seat or residence are advised to indicate this address in section 1 of EPO Form 1200. When acting as designated or elected Office, the EPO will accept an address for correspondence only if no professional representative must be or is appointed (Art. 133 and 134 EPC). This address must be the applicant's own address and must be located in an EPC Contracting State. Any address for correspondence accepted for the procedure in the international phase which does not fulfil those conditions will not be accepted in the procedure before the EPO in the European phase.
6. To enter the European phase before the EPO as designated or elected Office (Art. 22(1) and (3) and 39(1) PCT), the **following acts** must be performed by the applicant within **31 months** from the date of filing of the international application or (where applicable) the earliest priority date:
- a) Supply a translation of the international application into an EPO official language, if the International Bureau did not publish the application in such language (Art. 22(1) PCT and R. 159(1)(a) EPC);
 - b) Specify the application documents, as originally filed or as amended, on which the European grant procedure is to be based (R. 159(1)(b) EPC);
 - c) Pay the filing fee and, where applicable, the additional fee for a European patent application comprising more than 35 pages (R. 159(1)(c) EPC, Art. 2(1), items 1, 1a Rules relating to Fees (RFees));
 - d) Pay the search fee where a supplementary European search report has to be drawn up (R. 159(1)(e) EPC);
 - e) Pay the designation fee if the time limit laid down in Rule 39(1) EPC (i.e. six months after publication of the international search report) has expired before the 31-month period pursuant to Rule 159(1) EPC (R. 159(1)(d) EPC);
 - f) File the written request for examination and pay the examination fee if the time limit laid down in Rule 70(1) EPC has expired before the 31-month period pursuant to Rule 159(1) EPC (R. 159(1)(f) EPC);
 - g) Pay the renewal fee in respect of the third year, if the fee has fallen due (see Rule 51(1) EPC) before expiry of the 31-month period pursuant to Rule 159(1) EPC (R. 159(1)(g) EPC);
 - h) File, where applicable, the certificate of exhibition referred to in Article 55(2) and Rule 25 EPC (R. 159(1)(h) EPC);
 - i) Pay the claims fees for the sixteenth and each subsequent claim when the application documents on which the European grant procedure is to be based comprise more than fifteen claims (R. 162(1) EPC). A higher amount is payable for the 51st and each subsequent claim (Art.2(1), item 15 RFees).

If either the translation of the international application or the request for examination is not filed in time, or if the filing fee, the additional fee, the search fee, the designation fee or the examination fee is not paid in due time, the application shall be deemed to be withdrawn (R. 160(1) EPC).

7. Making payments

For payments made via deposit account, please note that as from 1 December 2017 debit orders will only be carried out if filed in an electronically processable format (xml), using an accepted means of filing as laid down in the Arrangements for deposit accounts (ADA), published in a Supplementary publication to the Official Journal.

All relevant information related to the modes of payment of fees to the EPO can be retrieved from the EPO website at "**Making Payments**".

Information concerning fee amounts

Procedural fees are usually adjusted every two years, on even years, with effect from 1 April. Therefore, before making a payment, parties should verify the amounts actually due on the date of payment using the applicable version of the Schedule of fees and expenses, published as a Supplement to the Official Journal of the EPO, available on the EPO website (www.epo.org) at www.epo.org/schedule-of-fees. The "Schedule of fees" table allows the viewing, downloading and searching of individual fee amounts, both current and previous.

8. Restoration of priority right

Where the international application contains a priority claim to an earlier application and it has been filed within two months from the expiration of the 12-month priority period, a request for restoration before the EPO as designated Office (R. 49ter.2 PCT) applies under the following circumstances:

- a) No request for restoration filed before the receiving Office (RO) during the international phase (R. 26bis.3 PCT)
- b) Negative decision by the RO irrespective of the criterion applied (due care/unintentionality)
- c) Positive decision by the RO based on the unintentionality criterion.

For a request to be admissible, it must be filed and the requisite fee must be paid (R. 49ter.2(b)(iii) and R. 49ter.2(d) PCT, Art.2(1), item 13 RFees) within the one-month time limit under R. 49ter.2(b)(i) PCT for filing the request, which starts from the expiry of the 31-month time limit pursuant to Rule 159(1)EPC and Article 22(3) or 39(1)(b) PCT or, where applicable, from the date on which an effective request for early processing (Art. 23(2) or 40(2) PCT) is filed ('Notice from the European Patent Office dated 21 February 2013 concerning the request for early processing', OJ EPO 2013, 158). The request for restoration also needs to state the reasons for the failure to file the international application within the priority period (Rule 49ter.2(b)(ii) PCT).

9. If the applicant had appointed a representative during the application's international phase, the present Form will be sent to the representative, asking him to inform the applicant accordingly. All subsequent communications will be sent to the applicant, or - if the EPO is informed of his appointment in time - to the applicant's European representative.
10. For more details about time limits and procedural acts before the EPO as designated or elected Office, see the EPO brochure Euro-PCT Guide: PCT procedure at the EPO. This brochure, the list of professional representatives before the EPO, Form 1200 and details of the latest fees are available on the EPO website (www.epo.org).

11. Programme for accelerated prosecution of European patent applications - "PACE"

Applicants wishing to accelerate proceedings before the EPO as designated or elected Office may request accelerated prosecution under the PACE programme (Notice from the European Patent Office dated 30 November 2015 concerning the programme for accelerated prosecution of European patent applications - "PACE", OJ EPO 2015, A93).

Receiving Section

Copy for (DO-EP) 31
PATENT COOPERATION TREATY

PCT/US2018/045648

ADVANCE E-MAIL

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING
OF A CHANGE

(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

LOWRY, Penny L.
5700 Flatiron Parkway
Boulder, Colorado 80301
ÉTATS-UNIS D'AMÉRIQUE

Date of mailing (day/month/year) 12 February 2020 (12.02.2020)	
Applicant's or agent's file reference 409002	IMPORTANT NOTIFICATION
International application No. PCT/US2018/045648	International filing date (day/month/year) 07 August 2018 (07.08.2018)

1. The following indications appeared on record concerning:

☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address REALD SPARK, LLC Suite 200 100 N. Crescent Drive Beverly Hills, California 90210 United States of America	State of Nationality US	State of Residence US
	Telephone No. 303-500-6151	
	Facsimile No. 303-440-0206	
	E-mail address patpros@reald.com	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☐ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address REALD SPARK, LLC Suite 400 9777 Wilshire Blvd., Beverly Hills, California 90212 United States of America	State of Nationality US	State of Residence US
	Telephone No. 303-500-6151	
	Facsimile No. 303-440-0206	
	E-mail address patpros@reald.com	
	☒ Notifications by e-mail authorized	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office ☒ the International Preliminary Examining Authority
☐ the International Searching Authority ☒ the designated Offices concerned
☐ the Authority(ies) specified for supplementary search ☐ the elected Offices concerned
☐ other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer N'Diaye Anta e-mail pct.team6@wipo.int Telephone No. +41 22 338 74 06
---	---

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(Chapter I of the Patent Cooperation Treaty)

(PCT Rule 44bis)

Applicant's or agent's file reference 409002	FOR FURTHER ACTION		See item 4 below
International application No. PCT/US2018/045648	International filing date (day/month/year) 07 August 2018 (07.08.2018)	Priority date (day/month/year) 08 August 2017 (08.08.2017)	
International Patent Classification (8th edition unless older edition indicated) See relevant information in Form PCT/ISA/237			
Applicant REALD SPARK, LLC			

1.	This international preliminary report on patentability (Chapter I) is issued by the International Bureau on behalf of the International Searching Authority under Rule 44 bis.1(a).																										
2.	This REPORT consists of a total of 16 sheets, including this cover sheet. In the attached sheets, any reference to the written opinion of the International Searching Authority should be read as a reference to the international preliminary report on patentability (Chapter I) instead.																										
3.	This report contains indications relating to the following items: <table><tr><td>☒</td><td>Box No. I</td><td>Basis of the report</td></tr><tr><td>☐</td><td>Box No. II</td><td>Priority</td></tr><tr><td>☒</td><td>Box No. III</td><td>Non-establishment of opinion with regard to novelty, inventive step and industrial applicability</td></tr><tr><td>☐</td><td>Box No. IV</td><td>Lack of unity of invention</td></tr><tr><td>☒</td><td>Box No. V</td><td>Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</td></tr><tr><td>☐</td><td>Box No. VI</td><td>Certain documents cited</td></tr><tr><td>☐</td><td>Box No. VII</td><td>Certain defects in the international application</td></tr><tr><td>☐</td><td>Box No. VIII</td><td>Certain observations on the international application</td></tr></table>			☒	Box No. I	Basis of the report	☐	Box No. II	Priority	☒	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability	☐	Box No. IV	Lack of unity of invention	☒	Box No. V	Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement	☐	Box No. VI	Certain documents cited	☐	Box No. VII	Certain defects in the international application	☐	Box No. VIII	Certain observations on the international application
☒	Box No. I	Basis of the report																									
☐	Box No. II	Priority																									
☒	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability																									
☐	Box No. IV	Lack of unity of invention																									
☒	Box No. V	Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement																									
☐	Box No. VI	Certain documents cited																									
☐	Box No. VII	Certain defects in the international application																									
☐	Box No. VIII	Certain observations on the international application																									
4.	The International Bureau will communicate this report to designated Offices in accordance with Rules 44bis.3(c) and 93bis.1 but not, except where the applicant makes an express request under Article 23(2), before the expiration of 30 months from the priority date (Rule 44bis .2).																										

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland		Date of issuance of this report 11 February 2020 (11.02.2020)
Facsimile No. +41 22 338 82 70		Authorized officer Agnès Wittmann-Regis e-mail: pct.team6@wipo.int

Form PCT/IB/373 (January 2004)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43bis.1)

To: PENNY L. LOWRY
5700 FLATIRON PARKWAY
BOULDER, CO 80301

Date of mailing
(day/month/year)

16 OCT 2018

Applicant's or agent's file reference
409002

FOR FURTHER ACTION

See paragraph 2 below

International application No.

PCT/US2018/045648

International filing date (day/month/year)

07 August 2018

Priority date (day/month/year)

08 August 2017

International Patent Classification (IPC) or both national classification and IPC

IPC(8) - G06F 3/01; G06K 9/00; H04N 5/228; H04N 7/14; H04N 13/00 (2018.01)

CPC - G06F 3/011; G06F 3/013; G06T 5/005; G06T 5/008; G06T 7/11 (2018.08)

Applicant REALD SPARK, LLC

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☒ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. FURTHER ACTION

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450
Facsimile No. 571-273-8300

Date of completion of this opinion

01 October 2018

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (January 2015)

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Box No. 1 Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:
 - ☒ the international application in the language in which it was filed.
 - ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).
2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a)).
3. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).
4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
5. Additional comments:

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application.

☒ claims Nos. 6-19, 22-24, 31, 32

because:

☐ the said international application, or the said claims Nos. _____ relate to the following subject matter which does not require an international search (*specify*):

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 6-19, 22-24, 31, 32 are so unclear that no meaningful opinion could be formed (*specify*):

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons: Claims 6-19, 22-24, 31, 32 are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

☐ the claims, or said claims Nos. _____ are so inadequately supported by the description that no meaningful opinion could be formed (*specify*):

☒ no international search report has been established for said claims Nos. 6-19, 22-24, 31, 32

☐ a meaningful opinion could not be formed without the sequence listing; the applicant did not, within the prescribed time limit:

☐ furnish a sequence listing in the form of an Annex C/ST.25 text file, and such listing was not available to the International Searching Authority in the form and manner acceptable to it; or the sequence listing furnished did not comply with the standard provided for in Annex C of the Administrative Instructions.

☐ furnish a sequence listing on paper or in the form of an image file complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in the form and manner acceptable to it; or the sequence listing furnished did not comply with the standard provided for in Annex C of the Administrative Instructions.

☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13ter.1(a) or (b).

☐ See Supplemental Box for further details.

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US2018/045648

Box No. V	Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement			
1. Statement	Novelty (N)	Claims	1-5, 20, 21, 25-30, 33-36	YES
		Claims	None	NO
	Inventive step (IS)	Claims	None	YES
		Claims	1-5, 20, 21, 25-30, 33-36	NO
	Industrial applicability (IA)	Claims	1-5, 20, 21, 25-30, 33-36	YES
		Claims	None	NO
2. Citations and explanations: Claims 1-4, 25-27, 30, 33, and 36 lack an inventive step under PCT Article 33(3) as being obvious over RealD Spark, LLC (hereinafter RealD) in view of Magic Pony Technology Limited et al. (hereinafter Magic). Regarding claim 1, RealD discloses a method of generating reference data (The reference data 35 is obtained and analyzed in advance [i.e. generating reference data] using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056) for adjusting a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), the method comprising: receiving training data (The capture device 10 includes a display 11 and a camera system 12 comprises two cameras 13 used to capture the stereoscopic pair of images of the head of a source observer 14, para 0030, The training input to the machine learning technique is two sets of images [i.e. receiving training data], which may be stereoscopic pairs of images or monoscopic images, para 0057) comprising: a set of input patches (in each image of the multi-view images, identifying image patches [i.e. a set of input patches] containing the left and right eyes of the head, respectively, para 0005), each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region (in each image of the multi-view images, identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region], para 0005, the image patches input and output images of each individual are aligned in the normalized coordinate system, para 0062), wherein the target feature is the same for each input patch (image patches containing the left and right eyes of the head, respectively are identified in each image 31 of the stereoscopic pair. FIG. 4 shows the identified image patches 32 of the right eye in each image 31 [i.e. the target feature is the same for each input patch], para 0049, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, para 0051); and a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region (adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, para 0005, image patches containing the left and right eyes of the head are identified and transformed, para 0006, the image patches input and output images of each individual are aligned in the normalized coordinate system [i.e. a set of output patches in one-to-one correspondence with the input patches], para 0062, From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image [i.e. each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region], para 0063, The displacement vector field F is chosen so that the image patch of the output image O(x,y) is derived from the image patch of the input image I(x,y) [i.e. a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region], para 0064); using a first machine learning algorithm to generate first reference data using the training data (To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning, para 0006, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The training input to the machine learning technique is two sets of images, which may be stereoscopic pairs of images or monoscopic images [i.e. using a first machine learning algorithm to generate first reference data using the training data], para 0057, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F [i.e. first reference data] to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076), the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each				

Form PCT/ISA/237 (Box No. V) (January 2015)

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006. The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055. The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056); and

using a second machine learning algorithm to generate second reference data (In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field F [i.e. second reference data] to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived and the values of the feature vector associated with the reference displacement vector fields [i.e. using a second machine learning algorithm to generate second reference data, para 0083], the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006. The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055. The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056).

RealD does not specifically disclose the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm. However, Magic is in the field of generating hierarchical algorithms using machine learning techniques to enable high quality reconstruction of lower-quality visual data (pg 1 In 6-8) and discloses the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm (receiving one or more sections of higher-quality visual data; training a hierarchical algorithm, wherein the hierarchical algorithm is operable to increase the quality of one or more sections of lower-quality visual data to substantially reproduce the one or more sections of higher-quality visual data; and outputting the hierarchical algorithm, pg 12 In 13-17. Training hierarchical algorithms can allow enhancement and/or reconstruction models to be developed for enhancing visual data in at least one embodiment, visual data can be image data and/or video data, enhancement models can be developed to increase the accuracy of higher-quality visual data that can be reproduced from lower-quality visual data in comparison to original higher-quality visual data, knowledge of the original visual data can allow the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data [i.e. using the same training data as the first machine learning algorithm] and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data, pg 12 In 18-27, a trained hierarchical algorithm can be developed for input visual data, wherein the trained hierarchical algorithm is developed for that input data based on the selected most similar pre-trained algorithm, the selection of the one or more most similar pre-trained algorithm(s) can be made based on one or more metrics associated with the pre-trained models [i.e. using the first reference data generated by the first machine learning algorithm] when compared and/or applied to the input data [i.e. the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm], pg 12 In 29 to pg 13 In 10). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of allowing the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data (Magic pg 12 In 18-27).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

Regarding claim 2, RealD further discloses that the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, para 0005, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The local image descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific local image descriptors are selected and optimized by the machine learning process itself, para 0070, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076); and the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, para 0005, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The local image descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific local image descriptors are selected and optimized by the machine learning process itself, para 0070, In the case that the feature vector 34 comprises features that are values in a discrete space and the reference displacement vector fields of the reference data 35 comprise a reference displacement vector field associated with each possible value of the feature vector in the discrete space, then the displacement vector field for the image patch is derived by selecting the reference displacement field associated with the actual value of the derived feature vector 34, para 0095).

Regarding claim 3, RealD further discloses that the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch (deriving a feature vector from plural local image descriptors of the image patch in at least one image of the multi-view images; and deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector, para 0005, The local image descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific local image descriptors are selected and optimized by the machine learning process itself, para 0070, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076).

Regarding claim 4, RealD further discloses that the editing instructions comprise a displacement vector field defining how the input patch is to be transformed (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, para 0005).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

Regarding claim 25, RealD discloses a method of training a machine learning algorithm (The training input to the machine learning technique [i.e. training a machine learning algorithm] is two sets of images, which may be stereoscopic pairs of images or monoscopic images, para 0057, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table, para 0076) to adjust a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), comprising:

receiving training data (The capture device 10 includes a display 11 and a camera system 12 comprises two cameras 13 used to capture the stereoscopic pair of images of the head of a source observer 14, para 0030, The training input to the machine learning technique is two sets of images [i.e. receiving training data], which may be stereoscopic pairs of images or monoscopic images, para 0057) comprising:

a set of input patches (in each image of the multi-view images, identifying image patches [i.e. a set of input patches] containing the left and right eyes of the head, respectively, para 0005). each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region (in each image of the multi-view images, identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region], para 0005, the image patches input and output images of each individual are aligned in the normalized coordinate system, para 0062), wherein the target feature is the same for each input patch (image patches containing the left and right eyes of the head, respectively are identified in each image 31 of the stereoscopic pair. FIG. 4 shows the identified image patches 32 of the right eye in each image 31 [i.e. the target feature is the same for each input patch], para 0049, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, para 0051); and first reference data generated by a pre-trained first machine learning algorithm (To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning, para 0006, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The training input to the machine learning technique is two sets of images, which may be stereoscopic pairs of images or monoscopic images [i.e. first reference data generated by a pre-trained first machine learning algorithm], para 0057, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076), the first reference data comprising a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. each editing instruction being for adjusting the digital representation of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, the image patches input and output images of each individual are aligned in the normalized coordinate system, para 0062, From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image [i.e. a set of editing instructions in one-to-one correspondence with the input patches], para 0063);

a pre-trained second machine learning algorithm trained to generate second reference data (In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field F to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived and the values of the feature vector associated with the reference displacement vector fields [i.e. updating a pre-trained second machine learning algorithm trained to generate second reference data], para 0083), the second reference data comprising editing instructions for adjusting the digital representation of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the second reference data comprising editing instructions for adjusting the digital representation of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected,

Form PCT/ISA/237 (Supplemental Box) (January 2015)

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056).

RealD does not specifically disclose that the input for the updating comprises the training data and the generated first reference data. However, Magic discloses that the input for the updating comprises the training data and the generated first reference data (receiving one or more sections of higher-quality visual data; training a hierarchical algorithm, wherein the hierarchical algorithm is operable to increase the quality of one or more sections of lower-quality visual data to substantially reproduce the one or more sections of higher-quality visual data; and outputting the hierarchical algorithm, pg 12 ln 13-17, Training hierarchical algorithms can allow enhancement and/or reconstruction models to be developed for enhancing visual data in at least one embodiment, visual data can be image data and/or video data, enhancement models can be developed to increase the accuracy of higher-quality visual data that can be reproduced from lower-quality visual data in comparison to original higher-quality visual data, knowledge of the original visual data can allow the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data [i.e. the input for the updating comprises the training data and the generated first reference data], pg 12 ln 18-27). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of allowing the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data (Magic pg 12 ln 18-27).

Regarding claim 26, modified RealD discloses the invention above, but RealD did not disclose that the updating of the second machine learning algorithm is performed at or near input data frame rate. However, Magic discloses that the updating of the second machine learning algorithm is performed at or near input data frame rate (the network is adapted for video super-resolution with a modification to the input layer. Instead of handling single frames independently, consecutive frames are jointly processed by feeding the consecutive multi-frame data such that the input can be expressed or converted into a three-dimensional vector of corresponding patches from multiple frames such that the temporal changes between consecutive frames are captured to the network. In an alternative embodiment, fixed temporal sampling can be used where videos have varying or different framerates, optionally selecting the lowest frame rate when the network is trained, pg 68 ln 29 to pg 69 ln 5). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of training the network (Magic pg 68 ln 29 to pg 69 ln 5).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

Regarding claim 27, RealD discloses a method of training a machine learning algorithm (The training input to the machine learning technique [i.e. training a machine learning algorithm] is two sets of images, which may be stereoscopic pairs of images or monoscopic images, para 0057, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table, para 0076) to adjust a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), the method comprising:

receiving training data (The capture device 10 includes a display 11 and a camera system 12 comprises two cameras 13 used to capture the stereoscopic pair of images of the head of a source observer 14, para 0030, The training input to the machine learning technique is two sets of images [i.e. receiving training data], which may be stereoscopic pairs of images or monoscopic images, para 0057) comprising a set of input digital representations of a head region (in each image of the multi-view images, identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. a set of input digital representations of a head region], para 0005);

training a first machine learning algorithm using the training data (To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning, para 0006, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The training input to the machine learning technique is two sets of images, which may be stereoscopic pairs of images or monoscopic images [i.e. training a first machine learning algorithm using the training data], para 0057, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076) to perform an adjustment of a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. perform an adjustment of a digital representation of a head region], para 0005);

using the trained first machine learning algorithm to generate first reference data (To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning, para 0006, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The training input to the machine learning technique is two sets of images, which may be stereoscopic pairs of images or monoscopic images [i.e. using the trained first machine learning algorithm to generate first reference data], para 0057, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076), the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056), each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform (A displacement vector field representing a transformation of an image patch is derived, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector produced by machine learning. The multi-view images are adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, abstract, each image 31 of the stereoscopic pair is adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector fields 37. This produces an adjusted stereoscopic pair of images 38, in which the gaze has been corrected [i.e. each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform], para 0116); and

training a second machine learning algorithm (In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field F to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived and the values of the feature vector

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

associated with the reference displacement vector fields [i.e. training a second machine learning algorithm) to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. perform the same adjustment of a digital representation of a head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056).

RealD does not specifically disclose the second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data.

However, Magic discloses the second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data (receiving one or more sections of higher-quality visual data; training a hierarchical algorithm, wherein the hierarchical algorithm is operable to increase the quality of one or more sections of lower-quality visual data to substantially reproduce the one or more sections of higher-quality visual data; and outputting the hierarchical algorithm, pg 12 In 13-17, Training hierarchical algorithms can allow enhancement and/or reconstruction models to be developed for enhancing visual data in at least one embodiment, visual data can be image data and/or video data, enhancement models can be developed to increase the accuracy of higher-quality visual data that can be reproduced from lower-quality visual data in comparison to original higher-quality visual data, knowledge of the original visual data can allow the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data [i.e. the second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data], pg 12 In 18-27). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of allowing the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data (Magic pg 12 In 18-27).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

Regarding Claim 30, RealD fails to explicitly disclose further comprising receiving a further digital representation of the head region and using the updated or trained second machine learning algorithm to adjust the further digital representation.

However, RealD teaches a pre-trained second machine learning algorithm trained to generate second reference data (In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field F to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived and the values of the feature vector associated with the reference displacement vector fields [i.e. updating a pre-trained second machine learning algorithm trained to generate second reference data], para 0083) and adjust the digital representation of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the second reference data comprising editing instructions for adjusting the digital representation of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006. The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055. The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056).

Further, Magic teaches updated algorithm (receiving one or more sections of higher-quality visual data; training a hierarchical algorithm, wherein the hierarchical algorithm is operable to increase the quality of one or more sections of lower-quality visual data to substantially reproduce the one or more sections of higher-quality visual data; and outputting the hierarchical algorithm, pg 12 ln 13-17, Training hierarchical algorithms can allow enhancement and/or reconstruction models to be developed for enhancing visual data in at least one embodiment, visual data can be image data and/or video data, enhancement models can be developed to increase the accuracy of higher-quality visual data that can be reproduced from lower-quality visual data in comparison to original higher-quality visual data, knowledge of the original visual data can allow the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data [i.e. the input for the updating comprises the training data and the generated first reference data], pg 12 ln 18-27).

Furthermore, defining an analysis and training of a learning algorithm to update the representation as taught by RealD and Magic and the specifics [update the second algorithm for head representation] would involve routine skill in the art.

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of allowing the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data (Magic pg 12 ln 18-27).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

Regarding claim 33, RealD discloses a device for generating reference data (The reference data 35 is obtained and analyzed in advance [i.e. generating reference data] using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056) for adjusting a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), the device comprising a data processing unit (an image processor 30, para 0036) arranged to:

receive training data (The capture device 10 includes a display 11 and a camera system 12 comprises two cameras 13 used to capture the stereoscopic pair of images of the head of a source observer 14, para 0030, The training input to the machine learning technique is two sets of images [i.e. receiving training data], which may be stereoscopic pairs of images or monoscopic images, para 0057) comprising:

a set of input patches (in each image of the multi-view images, identifying image patches [i.e. a set of input patches] containing the left and right eyes of the head, respectively, para 0005), each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region (in each image of the multi-view images, identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region], para 0005, the image patches input and output images of each individual are aligned in the normalized coordinate system, para 0062), wherein the target feature is the same for each input patch (image patches containing the left and right eyes of the head, respectively are identified in each image 31 of the stereoscopic pair. FIG. 4 shows the identified image patches 32 of the right eye in each image 31 [i.e. the target feature is the same for each input patch], para 0049, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, para 0051); and

a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target portion of the digital representation of the head region after adjustment of the digital representation of the head region (adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field, para 0005, image patches containing the left and right eyes of the head are identified and transformed, para 0006, the image patches input and output images of each individual are aligned in the normalized coordinate system [i.e. a set of output patches in one-to-one correspondence with the input patches], para 0062, From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image [i.e. each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region], para 0063, The displacement vector field F is chosen so that the image patch of the output image $O(x,y)$ is derived from the image patch of the input image $I(x,y)$ [i.e. a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region], para 0064);

use a first machine learning algorithm to generate first reference data using the training data (To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning, para 0006, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056, The training input to the machine learning technique is two sets of images, which may be stereoscopic pairs of images or monoscopic images [i.e. using a first machine learning algorithm to generate first reference data using the training data], para 0057, In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local image descriptors in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076); and

use a second machine learning algorithm to generate second reference data (In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local image descriptors in a continuous space. In this case, the machine learning technique associates a reference displacement vector field F derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field F to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived and the values of the feature vector associated with the reference displacement vector fields [i.e. using a second machine learning algorithm to generate second reference data]).

RealD does not specifically disclose the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

However, Magic discloses the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm (receiving one or more sections of higher-quality visual data; training a hierarchical algorithm, wherein the hierarchical algorithm is operable to increase the quality of one or more sections of lower-quality visual data to substantially reproduce the one or more sections of higher-quality visual data; and outputting the hierarchical algorithm, pg 12 In 13-17, Training hierarchical algorithms can allow enhancement and/or reconstruction models to be developed for enhancing visual data in at least one embodiment, visual data can be image data and/or video data, enhancement models can be developed to increase the accuracy of higher-quality visual data that can be reproduced from lower-quality visual data in comparison to original higher-quality visual data, knowledge of the original visual data can allow the hierarchical algorithm to be trained (and/or developed) based on knowledge of both the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data [i.e. the second machine learning algorithm using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm], pg 12 In 18-27). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Magic for the purpose of allowing the hierarchical algorithm to be trained (and/or developed) based on knowledge of both

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

the original visual data and the low-quality visual data in order to train a hierarchical algorithm to substantially reproduce the original visual data from the low-quality visual data (Magic pg 12 ln 18-27).

Regarding claim 36, RealD further discloses further comprising a telecommunications interface arranged to transmit the adjusted digital representation over a telecommunications network to a destination device to allow display of the head region thereon (The image processor 30 may be provided in the source device 10, the destination device 10 or in any other device, for example a server on a telecommunications network, which may be suitable in the case that the source device 10 and the destination device 10 communicate over such a telecommunications network, para 0037, the source device 10 and the destination device 10 communicate over such a telecommunications network 50, para 0141, the image processor 30 is provided in the source device 10 and is provided with the stereoscopic pair of images directly from the camera system 12. The telecommunications interface 17 is arranged to transmit the adjusted stereoscopic pair of images 38 over the telecommunications network 50 to the destination device 20 for display thereon, para 0142).

Claim 5 lacks an inventive step under PCT Article 33(3) as being obvious over RealD Spark, LLC (hereinafter RealD) in view of Magic Pony Technology Limited et al. (hereinafter Magic) and Nourbakhsh.

Regarding claim 5, modified RealD discloses the invention above, but RealD did not disclose that the editing instructions comprise a filter field, a brightness adjustment field, or a texture blending field. However, Nourbakhsh is in the field of video calls and meetings (para 0001) and discloses that the editing instructions comprise a filter field (analyzes video frames for the orientations of B's head and gaze direction, and uses image editing software to virtually correct the gaze direction of B's eye upward, toward B's camera, para 0002, The logic calculates 670 the angular correction for the eyes. The logic may then determine 680 filter parameters that set bounds on when eye gaze direction correction will be applied at all, para 0046), a brightness adjustment field, or a texture blending field. It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Nourbakhsh for the purpose of automatically adjusting eye gaze direction in a set of images to simulate eye-contact (Nourbakhsh para 0001).

Claims 20, 21, and 35 lack an inventive step under PCT Article 33(3) as being obvious over RealD Spark, LLC (hereinafter RealD) in view of Rhodes.

Regarding claim 20, RealD discloses a method of adjusting a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), the method comprising: identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region (identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. identifying a target patch in the digital representation of the head region], the target patch comprising a target feature of the digital representation of the head region), para 0005); deriving a feature vector from plural local descriptors of the target patch (deriving a feature vector from plural local image descriptors of the image patch [i.e. deriving a feature vector from plural local descriptors of the target patch] in at least one image of the multi-view images, para 0005); using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector, para 0005, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived [i.e. using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector], para 0076); and applying the selected editing instructions to the target patch to adjust the digital representation of the head region (adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. applying the selected editing instructions to the target patch to adjust the digital representation of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed, para 0006, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076). RealD does not specifically disclose that the editing instructions in the reference data are provided in a compressed representation. However, Rhodes is in the field of transmitting, receiving, and displaying data (abstract) and discloses that the editing instructions in the reference data are provided in a compressed representation (After computing the discrete wavelet transform (DWT), the data can be compressed according to any type of compression technique, such as, a variable thresholding mechanism governed by the embedded zerotree wavelet (EZW) in which coefficients that are larger than a threshold may be selected for transmission, para 0050, The additional data subsequently transmitted can be data compressed using a lower amount of compression, if prioritization instructions (e.g., information indicative of a gaze direction) are provided to the transmitter, the transmitter may encode data representing information about the gaze direction in better quality and less distortion than data representing other areas [i.e. the editing instructions in the reference data are provided in a compressed representation], para 0052). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Rhodes for the purpose of lowering bandwidth (Rhodes abstract).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

Regarding claim 21, modified RealD discloses the invention above, but RealD did not disclose that the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres. However, Rhodes discloses that the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation (Any type of encoding or compression techniques may be used, such as for example, a wavelet based technique (e.g., such as discrete wavelet transform (DWT) or embedded zerotree wavelet (EZW)), para 0050); Fourier and/or discrete cosine transform components; cluster centres. It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Rhodes for the purpose of lowering bandwidth (Rhodes abstract).

Regarding claim 35, RealD discloses a device for adjusting a digital representation of a head region (adjusting the stereoscopic pair of images, or more generally multi-view images, of a head to correct the perceived gaze, para 0004, adjusting multi-view images of a head to correct gaze [i.e. adjusting a digital representation of a head region], para 0005), the device comprising a data processing unit arranged to process the digital representation of the head region (an image processor 30, para 0036, the image processor 30 is provided in the source device 10 and is provided with the stereoscopic pair of images directly from the camera system 12. The telecommunications interface 17 is arranged to transmit the adjusted stereoscopic pair of images 38 over the telecommunications network 50 to the destination device 20 for display thereon [i.e. a data processing unit arranged to process the digital representation of the head region], para 0142) by:

identifying a target patch in the digital representation of the head region, the target patch comprising information about a target feature of the digital representation of the head region (identifying image patches containing the left and right eyes of the head, respectively; in respect of the image patches containing the left eyes of the head in each image of the multi-view images, and also in respect of the image patches containing the right eyes of the head in each image of the multi-view images [i.e. identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region], para 0005);

deriving a feature vector from plural local descriptors of the target patch (deriving a feature vector from plural local image descriptors of the image patch [i.e. deriving a feature vector from plural local descriptors of the target patch] in at least one image of the multi-view images, para 0005);

using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector, para 0005, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived [i.e. using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector], para 0076); and

applying the selected editing instructions to the target patch to adjust the digital representation of the head region (adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. applying the selected editing instructions to the target patch to adjust the digital representation of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed, para 0006, the machine learning technique associates a reference displacement vector field F derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 is essentially a look-up table. This allows a reference displacement vector field F to be simply selected from the reference data 35 on the basis of the feature vector 34 derived, para 0076).

RealD does not specifically disclose that the editing instructions in the reference data are provided in a compressed representation. However, Rhodes discloses that the editing instructions in the reference data are provided in a compressed representation (After computing the discrete wavelet transform (DWT), the data can be compressed according to any type of compression technique, such as, a variable thresholding mechanism governed by the embedded zerotree wavelet (EZW) in which coefficients that are larger than a threshold may be selected for transmission, para 0050. The additional data subsequently transmitted can be data compressed using a lower amount of compression, if prioritization instructions (e.g., information indicative of a gaze direction) are provided to the transmitter, the transmitter may encode data representing information about the gaze direction in better quality and less distortion than data representing other areas [i.e. the editing instructions in the reference data are provided in a compressed representation], para 0052). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Rhodes for the purpose of lowering bandwidth (Rhodes abstract).

Claims 28 and 29 lack an inventive step under PCT Article 33(3) as being obvious over RealD Spark, LLC (hereinafter RealD) in view of Magic Pony Technology Limited et al. (hereinafter Magic) and Kaye et al. (hereinafter Kaye).

Regarding claim 28, modified RealD discloses the invention above, but RealD did not disclose that the adjustment of the digital representation comprises converting from a two-dimensional digital representation to a three-dimensional digital representation. However, Kaye is in the field of reconstructing a hidden surface area in an image (abstract) and discloses that the adjustment of the digital representation comprises converting from a two-dimensional digital representation to a three-dimensional digital representation (converting two-dimensional images into three-dimensional images, para 0037). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Kaye for the purpose of converting two-dimensional (2D) motion picture images into three-dimensional (3D) motion picture images and to be able reduce or eliminate image artifacts that occur in the process of applying depth to objects during the conversion of 2D images into 3D images (Kaye para 0002, 0008).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2018/045648

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

Regarding claim 29, modified RealD discloses the invention above, but RealD did not disclose that the training of the first machine learning algorithm comprises iteratively using a rendering process to generate a two-dimensional digital representation from a three-dimensional digital representation generated by the first machine learning algorithm and comparing the generated digital representation with a corresponding digital representation in the training data. However, Kaye discloses that the training of the first machine learning algorithm comprises iteratively using a rendering process to generate a two-dimensional digital representation from a three-dimensional digital representation generated by the first machine learning algorithm and comparing the generated digital representation with a corresponding digital representation in the training data (An image region comparison technique is used to determine if an image region in an alternate frame contains missing image information hidden in the original frame, an image region is represented as a 2-dimensional topological set. Basic point-set topological operations then facilitate the pertinent region comparisons in the reconstruction process, para 0045, a region of image discrepancy caused by the shift of the object is determined quantitatively. This region is compared with the object's original region of definition in alternate frames to determine if any portion of the area of discrepancy is uncovered in those frames, para 0046, comparing image regions to determine if an image region in an alternate image contains information missing from the image due to the separation pixel gap. In one embodiment, the step of comparing image regions further includes the steps of: representing the image regions as a two-dimensional topological sets; and employing point-set topological operations for pertinent image region comparisons, para 0051). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Kaye for the purpose of determining if an image region in an alternate frame contains missing image information hidden in the original frame (Kaye para 0045).

Claim 34 lacks an inventive step under PCT Article 33(3) as being obvious over RealD Spark, LLC (hereinafter RealD) in view of Magic Pony Technology Limited et al. (hereinafter Magic) and Rhodes.

Regarding claim 34, modified RealD discloses the invention above, and RealD further discloses that the first reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056); and the second reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region (deriving a displacement vector field representing a transformation of an image patch, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector; and adjusting each image of the multi-view images by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field [i.e. the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region], para 0005, image patches containing the left and right eyes of the head are identified and transformed. To derive a displacement vector field that represents the transformation, a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images and used to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector. The form of the feature vector may be derived in advance from the reference data using machine learning. This method allows the gaze to be corrected, thereby reducing the disconcerting effect of incorrect gaze when the multi-view images are subsequently displayed, para 0006, The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 comprising reference displacement vector fields that represent transformations of the image patch and are associated with possible values of the feature vector, para 0055, The reference data 35 is obtained and analyzed in advance using a machine learning technique which derives the form of the feature vectors 34 and associates the reference displacement vector fields with the possible values of the feature vector, para 0056), but RealD did not disclose that the second editing instructions in the second reference data are provided in a compressed representation. However, Rhodes discloses that the second editing instructions in the second reference data are provided in a compressed representation (After computing the discrete wavelet transform (DWT), the data can be compressed according to any type of compression technique, such as, a variable thresholding mechanism governed by the embedded zerotree wavelet (EZW) in which coefficients that are larger than a threshold may be selected for transmission, para 0050, The additional data subsequently transmitted can be data compressed using a lower amount of compression, if prioritization instructions (e.g., information indicative of a gaze direction) are provided to the transmitter, the transmitter may encode data representing information about the gaze direction in better quality and less distortion than data representing other areas [i.e. the editing instructions in the reference data are provided in a compressed representation], para 0052). It would have been obvious to one of ordinary skill in the art at the time of the invention to modify RealD with the teachings of Rhodes for the purpose of lowering bandwidth (Rhodes abstract).

Claims 1-5, 20, 21, 25-30, and 33-36 meet the criteria set out in PCT Article 33(4), and thus have industrial applicability because the subject matter claimed can be made or used in industry.



**Entry into the European phase
(EPO as designated or elected Office)**
To the European Patent Office

European application number	EP18844651.2
PCT application number	PCT/US2018/045648
PCT publication number	WO2019032604
Applicant's or representative's reference	N411553EP
International Filing Date	07.08.2018
International Searching Authority (ISA)	US
International Preliminary Examining Authority (IPEA)	not applicable
1. Applicant Indications concerning the applicant(s) are contained in the international publication or were recorded by the International Bureau after the international publication. ☒ Changes which have not yet been recorded by the International Bureau are set out here: ☐	
2. Representative Representative 1 Representative or association of representatives to be listed in the Register of European Patents and to whom communications are to be notified Name: J A Kemp LLP Association No.: 811 Address of place of business: 14 South Square Gray's Inn London, Greater London, WC1R 5JJ United Kingdom Telephone: +44 20 3077 8600 Fax: +44 20 7430 1000 E-mail: mail@jakemp.com	
3. Authorisation Representative 1 An individual authorisation is attached. ☐ A general authorisation has been registered under No: ☐ A general authorisation has been filed, but not yet registered. ☐ The authorisation filed with the EPO as PCT receiving Office expressly includes the European phase. ☐	
4. Request for examination Examination of the application under Art. 94 EPC is hereby requested. The examination fee is being (has been, will be) paid. ☒ Request for examination in an admissible non-EPO language: ☐ The/Each applicant hereby declares that he is an entity or a natural person under Rule 6(4) EPC. ☐	
5. Copies Additional copies of the documents cited in the supplementary European search report are requested. ☐ Number of additional sets of copies	
6. Documents intended for proceedings before the EPO Number of claims on entry into the European phase: 36 6.1 Proceedings before the EPO as designated Office (PCT I) are to be based on the following documents the application documents published by the International Bureau; where the publication includes a set of claims amended under Article 19 PCT, the latter ☒	

N411553EP

replaces the originally filed claims unless replaced by the amendments attached.	☐
Comments on the international preliminary examination report established by the EPO as the International Preliminary Examining Authority and/or observations are attached.	☐
<i>Where necessary, further details should be attached as "Other documents"</i>	
6.2 Proceedings before the EPO as elected Office (PCT II) are to be based on the following documents:	
the documents on which the international preliminary examination report is based, including any annexes	☐
unless replaced by the amendments attached.	☐
comments on the international preliminary examination report established by the EPO as the International Preliminary Examining Authority and/or observations are enclosed.	☐
<i>Where necessary, further details should be attached as "Other documents"</i>	
If the EPO as International Preliminary Examining Authority has been supplied with test reports, these may be used as the basis of proceedings before the EPO.	☐
6.3 A copy of the results of the search carried out by the authority with which the previous application(s) whose priority is claimed was (were) filed is attached (R. 141(1) EPC).	☐
7. Translations	
Translations in one of the official languages of the EPO (English, French, German) are attached as crossed below:	☐
<i>* In proceedings before the EPO as designated or elected Office (PCT I + II):</i>	
7.1 Translation of the international application (description, claims, any text in the drawings) as originally filed, of the abstract as published and of any indication under Rule 13bis.3 and 13bis.4 PCT regarding biological material	☐
7.2 Translation of the priority application(s) (to be filed only at the EPO's request, Rule 53(3) EPC)	☐
7.3 It is hereby declared that the international application as originally filed is a complete translation of the previous application (Rule 53(3) EPC)	☐
<i>* In addition, in proceedings before the EPO as designated Office (PCT I):</i>	
7.4 Translation of amended claims and any statement under Art. 19 PCT, if the claims as amended are to form the basis for the proceedings before the EPO (see Section 6).	☐
<i>* In addition, in proceedings before the EPO as elected Office (PCT II):</i>	
7.5 Translation of any annexes to the international preliminary examination report	☐
8. Biological material	
The invention uses and/or relates to biological material deposited under Rule 31 EPC.	☐
The particulars referred to in Rule 31(1)(c) EPC (if not yet known, the depositary institution and the identification reference(s)) [number, symbols, etc.] of the depositor) are given in the international publication or in the translation submitted in Section 7 on:	☐
page(s) / line(s)	
The receipt(s) of deposit issued by the depositary institution	
is (are) enclosed.	☐
will be filed later.	☐
Waiver of the right to an undertaking from the requester pursuant to Rule 33(2) EPC attached.	☐
9. Nucleotide and amino acid sequences	
The international application discloses nucleotide and/or amino acid sequences.	☐
9.1 The sequence listing was filed under Rule 5.2(a) PCT, or furnished to the EPO as ISA under Rule 13ter.1(a) PCT, or is otherwise available to the EPO, in computer-readable format in accordance with WIPO ST.25.	☐
9.2	
The sequence listing is attached in computer-readable format in accordance with	☐

N411553EP

17. Annotations

18. Signature(s) of applicant(s) or representative

Place:

Date:

04 March 2020

Signed by:

Colin Merryweather 23462

Association:

J A Kemp LLP

Representative name:

/Colin Merryweather/

Capacity:

(Representative)

N411553EP

Table for section 6 of Form 1200.3

In accordance with the Notice from the European Patent Office dated 26 January 2009 concerning the 2009 fee structure (OJ EPO 2009, 118, and Guidelines for Examination in the EPO, April 2009, A-III, 13.2), the amount of the additional fee (Art. 2, item 1a, Rules relating to Fees) for the pages of this European patent application is calculated as follows:

Documents intended for proceedings before the EPO (R. 159 (1) (b) EPC) and for calculating the additional fee (Art. 2, item 1a, RFees):		
	Page(s) from ... to ...	Number of pages
Description: International application as published	1-35	35
Claims: International application as published	36-43	8
Drawings: International application as published	1/12-12/12	12
Abstract: Default count: one page		1
Total number of pages		56
Fee-exempt pages (Art. 2, item 1a, RFees)		-35
Number of pages to be paid for		21
(x 15 EUR per page)		
Total amount payable	EUR	315

N411553EP



Acknowledgement of receipt

We hereby acknowledge receipt of the form for entry into the European phase (EPO as designated or elected Office) as follows:

Submission number	8397264	
PCT application number	PCT/US2018/045648	
EP application number	18844651.2	
Date of receipt	04 March 2020	
Receiving Office	European Patent Office, The Hague	
Your reference	N411553EP	
Applicant		
Country		
Documents submitted	package-data.xml application-body.xml	ep-euro-pct.xml epf1200.pdf (5 p.)
Submitted by	CN=Colin Merryweather 23462	
Method of submission	Online	
Date and time receipt generated	04 March 2020, 16:51 (CET)	
Message Digest	0D:C2:AD:A1:A3:74:D4:41:4C:6D:81:91:CA:EA:86:39:83:98:37:F5	

Correction by the EPO of errors in debit instructions filed by eOLF

Errors in debit instructions filed by eOLF that are caused by the editing of Form 1038E entries or the continued use of outdated software (all forms) may be corrected automatically by the EPO, leaving the payment date unchanged (see decision T 152/82, OJ EPO 1984, 301 and point 6.3 ff ADA, Supplement to OJ EPO 10/2007).

/European Patent Office/



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



SOMMERLADE, Eric
Suite 200
100 North Crescent Drive
Beverly Hills, California 90210
ETATS-UNIS D'AMERIQUE

Date
13.03.20

Reference	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Notification of the data mentioned in Rule 19(3) EPC

In the above-identified patent application you are designated as inventor/co-inventor.
Pursuant to Rule 19(3) EPC the following data are notified herewith:

DATE OF FILING : 07.08.18

PRIORITY : US/08.08.17/ USP201762542661
: US/10.08.17/ USP201762543587

TITLE : ADJUSTING A DIGITAL REPRESENTATION OF A HEAD
REGION

DESIGNATED STATES : AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE
IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM
TR

Receiving Section





European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?

Contact Customer Services at www.epo.org/contact



NEOPHYTOU, Alexandros
100 N. Crescent Drive
Suite 200
Beverly Hills, California 90210
ETATS-UNIS D'AMERIQUE

Date
13.03.20

Reference	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Notification of the data mentioned in Rule 19(3) EPC

In the above-identified patent application you are designated as inventor/co-inventor.
Pursuant to Rule 19(3) EPC the following data are notified herewith:

DATE OF FILING : 07.08.18

PRIORITY : US/08.08.17/ USP201762542661
: US/10.08.17/ USP201762543587

TITLE : ADJUSTING A DIGITAL REPRESENTATION OF A HEAD
REGION

DESIGNATED STATES : AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE
IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM
TR

Receiving Section





European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?

Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
17.03.20

Reference N411553EP	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication pursuant to Rules 161(2) and 162 EPC

1. Amendment of the application (R. 161(2) EPC)

The above-mentioned international (Euro-PCT) application has entered the European phase.

Under Articles 28, 41 PCT and Rules 52, 78 PCT the application may be amended before a designated or elected Office.

In accordance with Rule 161(2) EPC, you may amend your application once within a **non-extendable period of six months** after notification of the present communication.

If filing amendments, you must identify them and indicate the basis for them in the application as filed. Failure to meet either requirement may lead to a communication from the Examining Division requesting that you correct this deficiency (R. 137(4) EPC).

The claims applicable on expiry of this period, i.e. those filed on entry into the European phase or in response to the present communication, will form the basis for the calculation of any claims fee to be paid (see page 2).

Registered letter

EPO Form 1226CC 12.17 (NFS) EUCL (09/03/20)

page 1 of 3

2 Claims fees under Rule 162 EPC

If the application documents on which the European grant procedure is to be based comprise more than fifteen claims, a claims fee shall be payable for the sixteenth and each subsequent claim within the period provided for in Rule 159(1) EPC.

- ☐ Based on the application documents currently on file, all necessary claims fees have already been paid (or the documents do not comprise more than 15 claims).
- ☐ All necessary fees will be debited automatically according to the automatic debit order.
- ☒ The claims fees due for the claims 16 to 36 were not paid within the above-mentioned period.

Any outstanding claims fee, either based on the current set of claims or on any amended claims to be filed pursuant to Rule 161 EPC (see page 1), may still be validly paid within a **non-extendable period of six months** after notification of this communication (R. 162(2) EPC).

If a payment is made for only some of the claims, you must indicate for which claims it is intended. If a claims fee is not paid in due time, the claim concerned is deemed to be abandoned (R. 162(4) EPC).

If claims fees have already been paid, but on expiry of the above-mentioned period there is a new set of claims containing fewer fee-incurring claims than before, the claims fees in excess of those due under Rule 162(2), second sentence EPC will be refunded (R. 162(3) EPC).

You are reminded that the supplementary search carried out according to Article 153(7) EPC will relate only to the last set of claims applicable on expiry of the above period AND will be confined to those fee-incurring claims for which fees have been paid in due time.

The claims fee is currently

EUR 235 for the 16th and each subsequent claim up to the limit of 50
EUR 585 for the 51st and each subsequent claim

Making payments

For payments made via deposit account, please note that as from 1 December 2017 debit orders will only be carried out if filed in an electronically processable format (xml), using an accepted means of filing as laid down in the Arrangements for deposit accounts (ADA), published in the Supplementary publication in the Official Journal.

All relevant information related to the modes of payment of fees to the EPO can be retrieved from the EPO website at "**Making Payments**".

Information concerning fee amounts

Procedural fees are usually adjusted every two years, on even years, with effect from 1 April. Therefore, before making a payment, parties should verify the amounts actually due on the date of payment using the applicable version of the Schedule of fees and expenses, published as a Supplement to the Official Journal of the EPO, available on the EPO website (www.epo.org) at www.epo.org/schedule-of-fees. The "Schedule of fees" table allows the viewing, downloading and searching of individual fee amounts, both current and previous.

Important note to users of the automatic debiting procedure

Unless the EPO receives prior instructions to the contrary, the fees for all claims incurring fees will be debited on the last day of the period for payment.

Receiving Section





European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
20.05.20

Reference N411553EP	Application No./Patent No. 18844651.2 - 1216 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication of European publication number and information on the application of Article 67(3) EPC

Provisional protection under Article 67(1) and (2) EPC in the individual contracting states becomes effective when the conditions referred to in Article 67(3) EPC have been fulfilled. Provisional protection in the individual extension and validation states is governed by the national laws of the countries concerned. For more details, see Chapter III of the EPO brochure "National Law relating to the EPC" (www.epo.org/law-practice/legal-texts/national-law.html), which also contains information about the extension and validation system. Further information can also be found in the Official Journal of the European Patent Office.

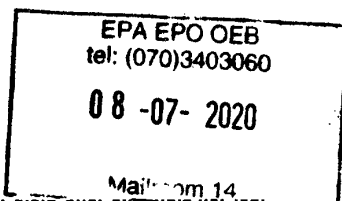
Pursuant to Article 153(3) EPC the publication under Article 21 PCT of an international application for which the European Patent Office is a designated or elected Office takes the place of the publication of a European patent application.

The bibliographic data of the above-mentioned Euro-PCT application will be published on 17.06.20 in Section I.1 of the European Patent Bulletin. The European publication number is 3665553.

In all future communications to the European Patent Office, please quote the application number plus Directorate number.

Receiving Section





European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



SOMMERLADE, Eric
Suite 200
100 North Crescent Drive
Beverly Hills, California 90210
ETATS-UNIS D'AMERIQUE

Date
13.03.20

Reference	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Notification of the data mentioned in Rule 19(3) EPC

In the above-identified patent application you are designated as inventor/co-inventor.
Pursuant to Rule 19(3) EPC the following data are notified herewith:

DATE OF FILING : 07.08.18

PRIORITY : US/08.08.17/ USP201762542661
: US/10.08.17/ USP201762543587

TITLE : ADJUSTING A DIGITAL REPRESENTATION OF A HEAD
REGION

DESIGNATED STATES : AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE
IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM
TR

Receiving Section





Europäisches
Patentamt
European
Patent Office
Office européen
des brevets

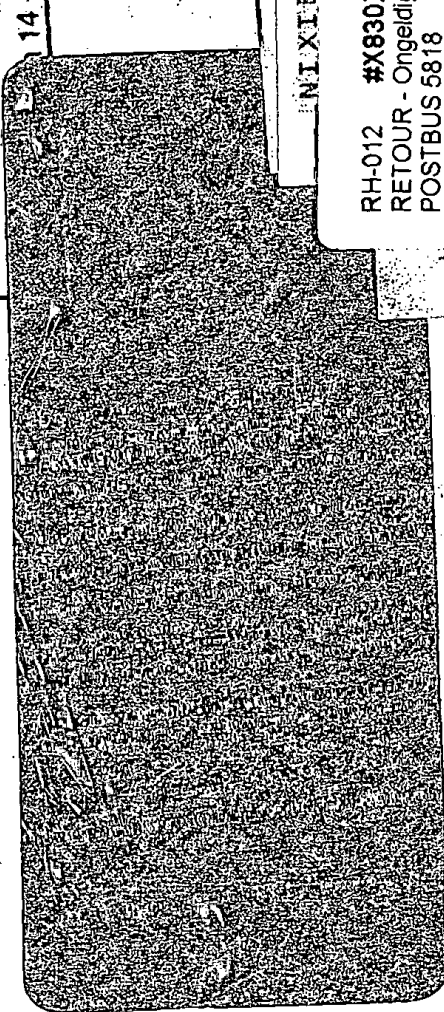


EPA EPO OEB
tel: (070)3403060

08-07-2020

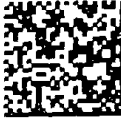


Port Betaald
Port Payé
Pays-Bas



NIXIE 911 CE 1 7204/07/20

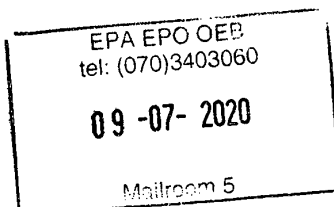
RH-012 #X830X0X#00#0000#
RETOUR - Ongeldig/Incompleet adres
POSTBUS 5818
2280 HV RIJSWIJK ZH



66-24-39

502103E450

4.111



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



NEOPHYTOU, Alexandros
100 N. Crescent Drive
Suite 200
Beverly Hills, California 90210
ETATS-UNIS D'AMERIQUE

Date
13.03.20

Reference	Application No./Patent No. 18844651.2 - 1216 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Notification of the data mentioned in Rule 19(3) EPC

In the above-identified patent application you are designated as inventor/co-inventor.
Pursuant to Rule 19(3) EPC the following data are notified herewith:

DATE OF FILING : 07.08.18

PRIORITY : US/08.08.17/ USP201762542661
: US/10.08.17/ USP201762543587

TITLE : ADJUSTING A DIGITAL REPRESENTATION OF A HEAD
REGION

DESIGNATED STATES : AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE
IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM
TR

Receiving Section



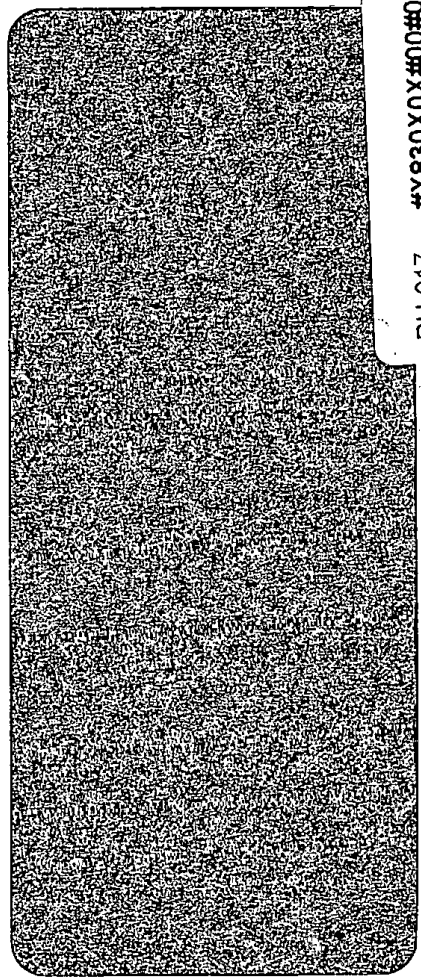


Port Betaald
Port Payé
Pays-Bas

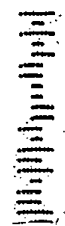


EPA EPO OEB
tel: (070)3403060
09-07-2020
Mailroom 5

8C: 00100
*0852-02770-30-27
RETURN TO SENDER
NOT DELIVERABLE AS ADDRESSED
UNABLE TO FORWARD
02/20/2020 1 59 11
NIXE



RH-017 #X830X0X#00#0000#
RETOUR - Ongeldig/incomplete adres
POSTBUS 5818
2280 HV RIJSWIJK ZH



502103

4.111



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?

Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
16.07.20

Reference N411553EP	Application No./Patent No. 18844651.2 - 1216 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication pursuant to Rule 19(1) EPC

The communication issued pursuant to Rule 19(3) EPC, sent to the inventor designated below, has been returned by the postal services.

You are requested to indicate the correct address of the inventor (R. 19(1) EPC).

Inventor : SOMMERLADE, Eric
Suite 200 100 North Crescent Drive
US / Beverly Hills, California 90210

Receiving Section





European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?

Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
17.07.20

Reference N411553EP	Application No./Patent No. 18844651.2 - 1216 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication pursuant to Rule 19(1) EPC

The communication issued pursuant to Rule 19(3) EPC, sent to the inventor designated below, has been returned by the postal services.

You are requested to indicate the correct address of the inventor (R. 19(1) EPC).

Inventor : NEOPHYTOU, Alexandros
100 N. Crescent Drive Suite 200
US / Beverly Hills, California 90210

Receiving Section





Letter accompanying subsequently filed items

Representative:

J A Kemp LLP
811
14 South Square
Gray's Inn
London Greater London WC1R 5JJ
United Kingdom

Phone: +44 20 3077 8600

Fax: +44 20 7430 1000

E-mail: mail@jakemp.com

80298 Munich

Germany

Tel. +49(0)89 2399-0 | Fax -4465

P.O. Box 5818

2280 HV Rijswijk

Netherlands

Tel. +31(0)70 340-2040 | Fax -3016

10958 Berlin

Germany

Tel. +49(0)30 25901-0 | Fax -840

The document(s) listed below is (are) subsequently filed documents pertaining to the following application:

Application number

18844651.2

Applicant's or representative's reference

N411553EP DQF

	Description of document	Original file name	Assigned file name
1	Reply to search opinion/written opinion/IPER	Response to R161 162 communication - N411553EP.PDF	WOREPLY-1.PDF
2	Amended claims with annotations	Amended claims tracked filed Sept 20 N411553EP DQF.PDF	CLMS-HWA-1.PDF
3	Amended claims (clean copy)	Amended claims clean filed Sept 20 N411553EP DQF.PDF	CLMS-1.PDF

	Payment	
1	Method of payment	Not specified

Signatures

Place: Paris
Date: 17 September 2020
Signed by: Dominic Forsythe 23485
Association: J A Kemp LLP
Representative name: /Dominic Forsythe/
Capacity: (Representative)

N411553EP DQF

J A • K E M P

BY ONLINE FILING

The European Patent Office
80298 Munich
Germany

17 September 2020

Dear Sirs

European Patent Application No. 18844651.2
REALD SPARK, LLC
Our Ref: N411553EP DQF

In reply to the communication of 17 March 2020, we request that amended claims pages 37-41 replace all of the claims pages currently on file. We attach also a tracked version of the amended pages to show where changes have been made.

Amendments

The claims have been amended to delete independent claims other than claims 1 and 33 (now 15). The remaining claims have been amended to improve conciseness by deleting or combining selected dependent claims.

Conclusion

We request that a communication under Article 94(3) be issued if any objections remain. Alternatively, if considered likely to be more efficient (for instance, if only minor remaining issues are identified), then the examiner is invited to contact us on +44 20 3077 8600.

As a precaution, we request that oral proceedings be held before any decision adverse to the applicant is made.

Yours faithfully

Electronically Signed
DOMINIC FORSYTHE
FOR J A KEMP LLP
ASSOCIATION NO. 811

PATENT ATTORNEYS • TRADE MARK ATTORNEYS
LONDON • OXFORD • CAMBRIDGE • PARIS • MUNICH

14 South Square, Gray's Inn, London WC1R 5JJ
T +44 20 3077 8600 F +44 20 7430 1000
mail@jakemp.com www.jakemp.com

J A Kemp LLP is a Limited Liability Partnership, registered in England and Wales (registered number OC427581) and regulated by IPReg.
Our registered office is at 14 South Square, Gray's Inn, London WC1R 5JJ, UK. Our use of the word 'partner' refers to a member of J A Kemp LLP.

Claims:

1. A method of generating reference data for adjusting a digital representation of a head region, the method comprising:
 - receiving training data comprising:
 - a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
 - a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;
 - using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and
 - using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.
2. The method of claim 1, wherein:
 - the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch; and
 - the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch,

wherein, optionally: the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch.

3. The method of claim 2, wherein the editing instructions comprise:
a displacement vector field defining how the input patch is to be transformed or a filter field, a brightness adjustment field, or a texture blending field.
4. The method of claim 2 or 3, wherein:
a first editing algorithm is used by the first machine learning algorithm to define how the first editing instructions are to be applied to an input patch to derive an output patch; and
a second editing algorithm is used by the second machine learning algorithm to define how the second editing instructions are to be applied to an input patch to derive an output patch.
5. The method of claim 4, wherein:
the second editing instructions are principle component analysis components of a principle component analysis of the first editing instructions; and the second image editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse principle component analysis transform; or
the second editing instructions are wavelet components of the first editing instructions; and the second editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse wavelet transform.
6. The method of any of claims 2-5, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative editing instructions than the second selection instructions for the second reference data.
7. The method of any of claims 1-6, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural

network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.

8. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair, wherein, optionally:

the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation of the head region comprises adjusting a gaze direction;

the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the nose;

the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the chin;

the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the digital representation of the head region comprises adjusting a color of the hair.

9. The method of any preceding claim, wherein the second editing instructions in the second reference data are provided in a compressed representation, the compressed representation optionally comprising one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers.

10. A method of adjusting a digital representation of a head region, the method comprising receiving a digital representation of a head region; and

using reference data comprising editing instructions to adjust the digital representation of the head region, wherein the reference data comprises the second reference data generated by the method of any of claims 1-9.

11. A method of adjusting a digital representation of a head region, the method comprising:
identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;
deriving a feature vector from plural local descriptors of the target patch;
using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and
applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:
the reference data comprises the second reference data generated by the method of any of claims 1-9.

12. The method of claim 10 or 11, wherein a plurality of the digital representations of the head region are received and adjusted, each digital representation being captured from the head region at a different point in time, and the method further comprises updating the second machine learning algorithm using editing instructions generated by the first machine learning algorithm using one or more of the received and adjusted digital representations, wherein the updating of the second machine learning algorithm is optionally performed at or near input data frame rate and, optionally, only a subset of the received and adjusted digital representations are used to update the second machine learning algorithm.

13. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.

14. A computer readable storage medium storing a computer program according to claim 13.

15. A device for generating reference data for adjusting a digital representation of a head region, the device comprising a data processing unit arranged to:

receive training data comprising:

- a set of input patches, each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and

- a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target portion of the digital representation of the head region after adjustment of the digital representation of the head region;

- use a first machine learning algorithm to generate first reference data using the training data; and

- use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

Claims:

1. A method of generating reference data for adjusting a digital representation of a head region, the method comprising:
 - receiving training data comprising:
 - a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
 - a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;
 - using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and
 - using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.
2. The method of claim 1, wherein:
 - the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch; and
 - the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch,

wherein, optionally: -

~~3. The method of claim 2, wherein~~ the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch.

34. The method of claim ~~2 or 3~~, wherein the editing instructions comprise:
a displacement vector field defining how the input patch is to be transformed or -

~~5. The method of claim 2 or 3, wherein the editing instructions comprise~~ a filter field, a brightness adjustment field, or a texture blending field.

46. The method of ~~any of claims 2 or 3-5~~, wherein:
a first editing algorithm is used by the first machine learning algorithm to define how the first editing instructions are to be applied to an input patch to derive an output patch; and
a second editing algorithm is used by the second machine learning algorithm to define how the second editing instructions are to be applied to an input patch to derive an output patch.

57. The method of claim 46, wherein:
the second editing instructions are principle component analysis components of a principle component analysis of the first editing instructions; and
_the second image editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse principle component analysis transform; or -

~~8. The method of claim 6, wherein:~~
the second editing instructions are wavelet components of the first editing instructions; and
_the second editing algorithm is configured to transform the second editing instructions into the first editing instructions by inverse wavelet transform.

69. The method of any of claims 2-58, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative editing instructions than the second selection instructions for the second reference data.

~~10. — The method of any preceding claim, wherein the first machine learning algorithm is of a different machine learning algorithm type than the second machine learning algorithm.~~

~~11. — The method of any preceding claim, wherein the first machine learning algorithm comprises one or more of the following: a neural network; a support vector machine; a generative adversarial network, GAN.~~

~~12. — The method of any preceding claim, wherein the second machine learning algorithm comprises one or more of the following: a regression forest; regression ferns, cluster centres, a lookup table, separable filter banks.~~

713. The method of any of claims 1-610, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.

814. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair, wherein, optionally:-

~~15. — The method of any of claims 1-13, wherein:~~

the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation of the head region comprises adjusting a gaze direction;

the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the nose;

the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the chin;

the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the digital representation of the head region comprises adjusting a color of the hair.

946. The method of any preceding claim, wherein the second editing instructions in the second reference data are provided in a compressed representation.

17. ~~The method of claim 16, wherein the compressed representation~~ optionally comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers.

108. A method of adjusting a digital representation of a head region, the method comprising receiving a digital representation of a head region; and using reference data comprising editing instructions to adjust the digital representation of the head region, wherein the reference data comprises the second reference data generated by the method of any of claims 1-947.

119. A method of adjusting a digital representation of a head region, the method comprising: identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region; deriving a feature vector from plural local descriptors of the target patch; using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and

applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:

the reference data comprises the second reference data generated by the method of any of claims 1-9~~17~~.

~~20. — A method of adjusting a digital representation of a head region, the method comprising:~~
~~—— identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;~~
~~—— deriving a feature vector from plural local descriptors of the target patch;~~
~~—— using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and~~
~~—— applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:~~
~~—— the editing instructions in the reference data are provided in a compressed representation.~~

~~21. — The method of claim 20, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.~~

122. The method of ~~any of claim 10 or 11~~~~s 18-21~~, wherein a plurality of the digital representations of the head region are received and adjusted, each digital representation being captured from the head region at a different point in time, and the method further comprises updating the second machine learning algorithm using editing instructions generated by the first machine learning algorithm using one or more of the received and adjusted digital representations, wherein:

~~23. — The method of claim 22, wherein the updating of the second machine learning algorithm is~~ optionally ~~performed at or near input data frame rate and, optionally, -~~

~~24. — The method of claim 23, wherein~~ only a subset of the received and adjusted digital representations are used to update the second machine learning algorithm.

~~25. — A method of training a machine learning algorithm to adjust a digital representation of a head region, comprising:~~

~~receiving training data comprising:~~

~~a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and~~

~~first reference data generated by a pre-trained first machine learning algorithm, the first reference data comprising a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region;~~

~~updating a pre-trained second machine learning algorithm trained to generate second reference data, where the input for the updating comprises the training data and the generated first reference data, the second reference data comprising editing instructions for adjusting the digital representation of the head region.~~

~~26. — The method of claim 25, wherein the updating of the second machine learning algorithm is performed at or near input data frame rate.~~

~~27. — A method of training a machine learning algorithm to adjust a digital representation of a head region, the method comprising:~~

~~— receiving training data comprising a set of input digital representations of a head region;~~

~~— training a first machine learning algorithm using the training data to perform an adjustment of a digital representation of a head region;~~

~~using the trained first machine learning algorithm to generate first reference data, the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations, each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform; and~~

~~—training a second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm.~~

~~28. The method of claim 27, wherein the adjustment of the digital representation comprises converting from a two-dimensional digital representation to a three-dimensional digital representation.~~

~~29. The method of claim 28, wherein the training of the first machine learning algorithm comprises iteratively using a rendering process to generate a two-dimensional digital representation from a three-dimensional digital representation generated by the first machine learning algorithm and comparing the generated digital representation with a corresponding digital representation in the training data.~~

~~30. The method of any of claims 25-29, further comprising receiving a further digital representation of the head region and using the updated or trained second machine learning algorithm to adjust the further digital representation.~~

1334. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.

1432. A computer readable storage medium storing a computer program according to claim 1334.

1533. A device for generating reference data for adjusting a digital representation of a head region, the device comprising a data processing unit arranged to:

receive training data comprising:

a set of input patches, each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and

a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target portion of the digital representation of the head region after adjustment of the digital representation of the head region;
use a first machine learning algorithm to generate first reference data using the training data; and
use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.

34. ~~The device of claim 33, wherein:~~

~~the first reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;~~

~~the second reference data comprises editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;~~

~~the second editing instructions in the second reference data are provided in a compressed representation.~~

35. ~~A device for adjusting a digital representation of a head region, the device comprising a data processing unit arranged to process the digital representation of the head region by:~~

~~—— identifying a target patch in the digital representation of the head region, the target patch comprising information about a target feature of the digital representation of the head region;~~

~~—— deriving a feature vector from plural local descriptors of the target patch;~~

~~—— using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and~~

~~—— applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein:~~

~~—— the editing instructions in the reference data are provided in a compressed representation.~~

36. ~~The device of any of claims 33-35, further comprising a telecommunications interface~~

~~arranged to transmit the adjusted digital representation over a telecommunications network to a destination device to allow display of the head region thereon.~~



Acknowledgement of receipt

We hereby acknowledge receipt of the following subsequently filed document(s):

Submission number	9028253	
Application number	EP18844651.2	
Date of receipt	17 September 2020	
Receiving Office	European Patent Office, The Hague	
Your reference	N411553EP DQF	
Applicant	All applicants as on file	
Documents submitted	package-data.xml epf1038.pdf (1 p.) CLMS-HWA-1.PDFAmended claims tracked filed Sept 20 N411553EP DQF.PDF (9 p.)	ep-sfd-request.xml WOREPLY-1.PDFResponse to R161 162 communication - N411553EP.PDF (1 p.) CLMS-1.PDFAmended claims clean filed Sept 20 N411553EP DQF.PDF (5 p.)
Submitted by	CN=Dominic Forsythe 23485	
Method of submission	Online	
Date and time receipt generated	17 September 2020, 16:59 (CEST)	
Message Digest	BE:06:9D:E9:5F:51:9D:F0:08:C0:D9:93:80:8B:40:A4:26:30:39:6A	

/European Patent Office/



Application No.: 18844651.2

The search has started on 15.04.21 (see Notice from the European Patent Office dated 03/2013, OJ EPO 2013, 153).

Receiving Section

**SUPPLEMENTARY PARTIAL
EUROPEAN SEARCH REPORT**

Application Number
EP 18 84 4651

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (IPC)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
A	SUN XIANGRAN ET AL: "Construction and compression of face models for multi-party videoconferencing with multi-camera", 2013 6TH INTERNATIONAL CONGRESS ON IMAGE AND SIGNAL PROCESSING (CISP), IEEE, vol. 2, 16 December 2013 (2013-12-16), pages 948-952, XP032569425, DOI: 10.1109/CISP.2013.6745301 [retrieved on 2014-02-19] * section II *	5,9	
A	----- KATSIGIANNIS STAMOS ET AL: "A GPU based real-time video compression method for video conferencing", 2013 18TH INTERNATIONAL CONFERENCE ON DIGITAL SIGNAL PROCESSING (DSP), IEEE, July 2013 (2013-07), pages 1-6, XP032498481, ISSN: 1546-1874, DOI: 10.1109/ICDSP.2013.6622719 [retrieved on 2013-10-06] * abstract *	5,9	TECHNICAL FIELDS SEARCHED (IPC)
A	----- GEOFFREY HINTON ET AL: "Distilling the Knowledge in a Neural Network", CORR (ARXIV), vol. 1503.02531v1, 9 March 2015 (2015-03-09), pages 1-9, XP055549014, * abstract *	6,7	
	----- -/--		

2
EPO FORM 1503 09-98 (P04N26)

**SUPPLEMENTARY PARTIAL
EUROPEAN SEARCH REPORT**

Application Number
EP 18 84 4651

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (IPC)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
A	YIM JUNHO ET AL: "A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning", 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), IEEE COMPUTER SOCIETY, US, 21 July 2017 (2017-07-21), pages 7130-7138, XP033250080, ISSN: 1063-6919, DOI: 10.1109/CVPR.2017.754 [retrieved on 2017-11-06] * sections 1, 2 *	6,7	
A	----- CRISTIAN BUCILUA ET AL: "Model compression", PROCEEDINGS OF THE TWELFTH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING : AUGUST 20 - 23, 2006, PHILADELPHIA, PA, USA, NEW YORK, NY : ACM PRESS, 2 PENN PLAZA, SUITE 701 NEW YORK NY 10121-0701 USA, 20 August 2006 (2006-08-20), pages 535-541, XP058107730, DOI: 10.1145/1150402.1150464 ISBN: 978-1-59593-339-3 * abstract *	6,7	TECHNICAL FIELDS SEARCHED (IPC)
A	----- KONONENKO DANIIL ET AL: "Learning to look up: Realtime monocular gaze correction using machine learning", 2015 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), IEEE, 7 June 2015 (2015-06-07), pages 4667-4675, XP032793924, DOI: 10.1109/CVPR.2015.7299098 [retrieved on 2015-10-14] * section 3 *	1-11, 13-15	

2
EPO FORM 1503 09-98 (P04N26)

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. claims: 1-11, 13-15

providing the 2nd editing instructions in the 2nd reference data in a compressed representation

1.1. claims: 6, 7

providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data

2. claim: 12

online updating of the 2nd machine-learning-algorithm, based on the correction of a video stream predicted by the 1st machine-learning-algorithm, in order to improve the machine learning prediction "on the fly"

Please note that all inventions mentioned under item 1, although not necessarily linked by a common inventive concept, could be searched without effort justifying an additional fee.

The search has been limited to the first subject.

* * * *

Reasoning: see European search opinion

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 18 84 4651

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-04-2021

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017195662 A1	06-07-2017	CN 108463787 A	28-08-2018
		EP 3400706 A1	14-11-2018
		US 2017195662 A1	06-07-2017
		US 2020021794 A1	16-01-2020
		US 2021044794 A1	11-02-2021
		WO 2017120247 A1	13-07-2017

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
03.05.21

Reference N411553EP	Application No./Patent No. 18844651.2 - 1230 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication pursuant to Rule 164(1) EPC

The partial supplementary European search report (R. 164 EPC) is enclosed.

The applicant is informed that if the supplementary European search report is also to cover inventions other than the invention(s) already searched and for which a meaningful search can be carried out, a further search fee must be paid for each of these inventions, in the present instance

1 search fee(s)

within **two months** after notification of this communication.

The amount payable for each further search is EUR 1350,00.

Important note to users of the automatic debiting procedure

Unless the EPO receives prior instructions to the contrary, the fee(s) will be debited on the last day of the period for payment. For earlier payment(s) another method of payment must be used.

If applicable, a European search opinion covering those invention(s) for which a search fee has been paid will then be sent together with the supplementary European search report.

Copies of documents cited in the supplementary European search report are attached.

☒ 0 additional set(s) of copies of such documents is (are) enclosed as well.

Making payments

For payments made via deposit account, please note that as from 1 December 2017 debit orders will only be carried out if filed in an electronically processable format (xml), using an accepted means of filing as laid down in the Arrangements for deposit accounts (ADA), published in the Supplementary publication in the Official Journal.

All relevant information related to the modes of payment of fees to the EPO can be retrieved from the EPO website at "**Making Payments**".

Information concerning fee amounts

Procedural fees are usually adjusted every two years, on even years, with effect from 1 April. Therefore, before making a payment, parties should verify the amounts actually due on the date of payment using the applicable version of the Schedule of fees and expenses, published as a Supplement to the Official Journal of the EPO, available on the EPO website (www.epo.org) at www.epo.org/schedule-of-fees. The "Schedule of fees" table allows the viewing, downloading and searching of individual fee amounts, both current and previous.



The examination is being carried out on the **following application documents**

Description, Pages

1-35 as published

Claims, Numbers

1-15 filed in electronic form on 17-09-2020

Drawings, Sheets

1/12-12/12 as published

* * * *

Following documents are cited in the European search report:

- D1 XP055295123
Yaroslav Ganin et al.: "DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation", pages 1-18, 26 July 2016 (2016-07-26), <http://arxiv.org/pdf/1607.07215v2.pdf>
- D2 US 2017/195662 A1 (SOMMERLADE ERIC [GB] ET AL) 6 July 2017 (2017-07-06)
- D3 XP055295123
Jiebo Luo et al.: "Face Location in Wavelet-Based Video Compression for High Perceptual Quality Videoconferencing", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 4, 1 August 1996 (1996-08-01), ISSN: 1051-8215
- D4 XP032569425
Sun Xiangran et al.: "Construction and compression of face models for multi-party videoconferencing with multi-camera", 6th IEEE International Congress on image and Signal Processing (CISP), vol. 2, 16 December 2013 (2013-12-16), pages 948-952, DOI: 10.1109/CISP.2013.6745301

- D5 XP058107730
Katsigiannis Stamos et al.: "A GPU based real-time video compression method for video conferencing", 18th IEEE International Conference on Digital Signal Processing July 2013 (2013-07), pages 1-6, ISSN: 1546-1874 , DOI: 10.1109/ICDSP.2013.6622719
- D6 XP055549014
Geoffrey Hinton et al.: "Distilling the Knowledge in a Neural Network", 9 March 2015 (2015-03-09), pages 1-9, <https://arxiv.org/pdf/1503.02531.pdf>
- D7 XP033250080
Yim Junho et al.: "A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21 July 2017 (2017-07-21), pages 7130-7138, ISSN: 1063-6919, DOI: 10.1109/CVPR.2017.754
- D8 XP058107730
Cristian Bucilă et al.: "Model compression", Proceedings of the 12th ACM International Conference on Knowledge Discovery and Data Mining, 20 August 2006 (2006-08-20), pages 535-541, DOI: 10.1145/1150402.1150464 , ISBN: 978-1-59593-339-3
- D9 XP032793924
Kononenko D. et al.: "Learning to look up: Realtime monocular gaze correction using machine learning", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7 June 2015 (2015-06-07), pages 4667-4675, DOI: 10.1109/CVPR.2015.7299098

* * * *

- 1 An international preliminary report on patentability (IPRP) has already been drawn up for the present application in accordance with the PCT, stating a lack of inventive step of the published PCT-claims 1-5, 20, 21, 25-30, 33-36. No opinion has been established for the remainder of the published PCT-claims, for not being drafted in accordance with Rule 6.4(a) PCT.

- 2 After entering the regional before the EPO, and with letter L0 received on 17-09-2020, the applicant has filed replacement claims 1-15, featuring the merging of PCT-claims, and the deletion of PCT-claims 10-12, 20, 21, 25-30, 34-36. The remaining independent claims were not amended. No arguments with respect to the IPRP's statement of lack of inventive step have been filed, which is concerning the present claims 1-3, 15.

Various deficiencies

- 3 The application exhibits following deficiencies:
- 3.1 Even though claim 11 includes explicitly or implicitly all the features of claim 10, it is not drafted as dependent on claim 10 as required by Rule 43(4) EPC.
- 3.2 The features of the claims should be provided with reference signs placed in parentheses to increase the intelligibility of the claims (Rule 43(7) EPC).
- 3.3 Documents reflecting the described prior art should preferably be cited in the background part of the description (Rule 42(1)(b) EPC). In the present case, documents D1 and D2 should be cited.
- 3.4 In the description paragraphs [47], [48], [67], [75], [76], [77], [83], [84], [85], [88], [149], the cited background art is "incorporated by reference". In case the applicant considers portions of the referenced document as essential for carrying out the claimed invention, the matter in question should be expressly incorporated into the description, as the patent specification must, regarding the essential features of the invention, be self-contained (Article 83 EPC). As to the particular conditions allowing the introduction of features from a cross-referenced document into the claims, reference is made to the Guidelines H, IV, 2.2.1. It is the provisional opinion of the search examiner that no insertion of background art is further necessary to comply with Article 83 EPC. Said incorporation by reference needs however to be deleted from the description before the end of the examination procedure (Guidelines F-III, 8).
- 3.5 Statements implying that the scope of protection may deviate from the claimed invention (sub-combinations of features, "spirit", "equivalents", claim-like clauses) should be deleted or amended in order to clarify that the scope of protection of the invention is exclusively defined by the claims (Article 84 EPC, Guidelines F-IV, 4.4). This appears to apply to paragraph [150] (necessary complete deletion of claim-like clauses), to paragraph [151] (suggested

complete deletion or at least deletion of "equivalents"), and to paragraph [152] (suggested complete deletion or at least deletion of the contradictions to the definitions of Rule 43 EPC, deletion of "equivalents").

Lack of clarity

- 4 The claims do not comply with the requirements of Article 84 EPC:
- 4.1 In independent claim 15, the claimed first and second "reference data" are not clearly defined in their relation to the problem of head pose correction and may therefore be seen as arbitrary processing features at least over a part of their scope of protection. In fact, the training of the claimed machine learning algorithms, may result also in processing data other than only head pose correction instructions.
- 4.2 In order to include at least implicitly a tangible technical effect of head pose correction, the claimed reference data has to include the editing instructions as presently claimed in independent claim 1.
- 4.3 It is to be noted that this lack of clarity of the claimed reference data has the consequence that the subject-matter of claim 15 is obvious over any prior art of head pose correction at least over part of its scope of protection, as the machine learning may presently be construed as an arbitrary mathematical processing not amenable to contribute to an inventive step in the sense of Article 56 EPC (see Guidelines G VII, 5.4).

Lack of unity of invention

- 5 The application lacks unity within the meaning of Article 82 EPC.
The examiner considers that the following separate alleged inventions are not so linked as to form a single general inventive concept:
 - 1. claims: 1-5, 8-11, 13-15
providing the 2nd editing instructions in the 2nd reference data in a compressed representation in order to reduce the bandwidth needs of a gaze correction
 - 1.1. claims: 6, 7

providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data, ultimately a reduction of the bandwidth needs of a gaze correction

2. claim: 12

online updating of the 2nd machine-learning-algorithm, based on the correction of a video stream predicted by the 1st machine-learning-algorithm, in order to improve the machine learning prediction "on the fly"

5.1 An assessment with respect to the prior art is first to be made for the subject-matter in common to the three claim-groups:

5.1.1 D1 represents the document of prior art closest to the originally claimed invention, and anticipates the subject-matter of claim 1, namely

A method of generating reference data for adjusting a digital representation of a head region,

(see D1, abstract: "[...] Our deep architecture performs coarse-to-fine warping with an additional intensity correction of individual pixels. All these operations are performed in a feed-forward manner, and the parameters associated with different operations are learned jointly in the end-to-end fashion. After learning, the resulting neural network can synthesize images with manipulated gaze [...])"

the method comprising:

receiving training data comprising:

a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;

(see D1, section 3 "The model" 2nd sentence: "The model is trained on pairs of images corresponding to eye appearance before and after the redirection. The redirection angle serves as an additional input parameter that is provided both during training and at test time."

see also section 3. 1 "Data preparation"; see also section 4.2 "Training procedure")

using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;

and

using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

(see D1, figure 2 and section 3 entitled "The model", in particular in the 2nd paragraph: " The task of the network is therefore the prediction of the warping field. This field is predicted in two stages in a coarse-to-fine manner, where the decisions at the fine scale are being informed by the result of the coarse stage.";

see also section 3.2 "Warping modules" describing the two concatenated machine learning algorithms, as claimed: "Each of the two warping modules takes as an input the image, the position of the feature points, and the redirection angle. [...]"

Coarse warping. *The last convolutional layer of the first (half-scale) warping module produces a pixel-flow field (a two-channel map), which is then up-sampled $D_{\text{coarse}}(I, \alpha)$ and applied to warp the input image by means of a bilinear sampler S [12,21] that finds the coarse estimate" [...]*

Fine warping. *In the fine warping module, the rough image estimate O_{coarse} and the upsampled low-resolution flow $D_{\text{coarse}}(I, \alpha)$ are concatenated with the input data (the image, the angle encoding, and the feature point encoding) at the original scale and sent to the $1\times$ -*

scale network which predicts another two-channel flow D_{res} that amends the half-scale pixel-flow [...] The amended flow is used to obtain the final output";

Remark: the joint learning of the two CNNs for coarse and fine warping implies the claimed training)

- 5.1.2 The subject-matter of independent method-claim 1, and of the corresponding claims 13-15 in other categories is therefore not new in the sense of Article 54(1)(2) EPC.
- 5.1.3 The passages of D1 recited above disclose also the subject matter of independent claim 10, which regards the adjustment of a head image with the editing instructions resulting from the training of the 2nd machine learning algorithm. The subject-matter of claim 10 is therefore also not new over D1.
- 5.1.4 Moreover, the passages of D1 recited above explicitly or implicitly disclose also the additional features of dependent claims 2, 3, 4, 8, 11. The subject-matter of these claims is therefore also not new over D1.
- 5.2 The contribution of the different claim groups to the known prior art disclosed by D1, i.e. the Special Technical Features (STF) as defined by Rule 44(1) EPC, and their associated objective problems, may therefore be defined as follows:
- 5.2.1 Concerning claim group 1, the STF consist in the additional features of claims 5 or 9 of a compressed representation of the 2nd editing instructions. With regard to D1, the objective problem can be defined as how to reduce the transmission bandwidth of the transform parameters, for example if the gaze correction of D1 is to be used in teleconferencing.
- 5.2.2 Concerning claim group 1.1, the STF consist in the additional features of claim 6 or 7. With regard to D1, the objective problem may be the definition of an alternative architecture, with the purpose of improving the training of a (2nd) machine learning algorithm of poor accuracy, which is ultimately used for solving the objective problem of reducing the transmission bandwidth of the transform parameters necessary for gaze correction.
- 5.2.3 Concerning claim group 2, the STF consist in the additional features of claim 12. With regard to D1, the objective problem can be defined as how to train the 2nd machine learning algorithm "on the fly".
- 5.3 Conclusion

The STF of the different claim-groups are obviously not the same. Moreover, they can be implemented independently from each other, and the solved problems either are not the same, or, when the same (groups 1 and 2), cannot define a common inventive concept ("problem invention") because consisting in a standard problem of the technical field. Therefore, no technical relationship in the sense of Rule 44(1) EPC exists between the different claim groups, which may be regarded as an inventive link.

Hence, the aforementioned claim groups are not so linked as to form a single general inventive concept, and do therefore not comply with the requirements of Article 82 EPC.

Lack of novelty over D1

- 6 The subject-matter of **claims 1-4, 8, 10, 11 ,13-15** is not new in the sense of Article 54(1) and (2) EPC, and therefore the requirements of Article 52(1) EPC are not met, as it is anticipated by D1 (see section 5.1 above).

Lack of inventive step

- 7 Furthermore, the subject-matter of **claims 5 and 9** does not involve an inventive step in the sense of Article 56 EPC, and therefore the requirements of Article 52(1) EPC are not met, for the following reasons:
- 7.1 The additional features of dependent claims 5, 9 concern the claimed second editing instructions being in a compression representation.
- 7.2 In the overall description, the compression is meant as an integral part of the training of the machine learning algorithms. This can be derived in particular from figures 9, 12, where the auxiliary data "Auxd" is defined in paragraph [125] to be input into the 1st machine to refine "a basis set used for providing a compressed representation of image editing instructions".
- 7.3 However, the wording of claim 5 or 9 at present includes also a post-compression of the 2nd editing instructions output by the training of the machine learning algorithms. As video compression including the compression of imaging instructions is a standard measure for video-conferencing (*see for*

example documents D3-D5), the skilled person would straightforwardly implement the additional features of claims 5, 9, when using the method of D1 in the framework of videoconferencing.

- 7.4 The features of PCA or wavelet compression are standard compression methods, which do *prima facie* not add any surprising technical advantage that would prove the involvement of inventive skills.
- 7.5 It is suggested to alleviate this objection with respect to D1 by including that the compression is an integral part of the training of the machine learning algorithms, for example by including in the claimed training data "a basis set used for providing a compressed representation of editing instructions" based on paragraph [25]. This intermediate generalisation from [25] and figure 9, is deemed by the search examiner to fulfil the criteria set out in the Guidelines H-V, 3.2.1 and not to contravene Article 123(2) EPC.
- 8 Finally, the subject-matter of **claims 6 and 7** does not involve an inventive step in the sense of Article 56 EPC, and therefore the requirements of Article 52(1) EPC are not met, for the following reasons:
- 8.1 When contemplated with respect to D1, the additional features of dependent claims 5, 9 are providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data, ultimately a reduction of the bandwidth needs of a gaze correction. Starting from D1, this would mean to reduce the accuracy of the 2nd CNN trained to provide a refined transform. Even though bandwidth requirements could force a skilled person to reduce the accuracy of D1, there seems to be no obvious reason to modify the coarse-to-fine processing of D1 in this way.
- 8.2 In assessing other prior art, it however resorts that the application's concept (underlying these features) of improving the training of a less accurate machine learning algorithm additionally with the result of a pre-trained more accurate machine learning algorithm, is a concept well known in the art of artificial intelligence, in particular under the name of "knowledge distillation" (*see for example documents D6-D8*). When starting from D2, which discloses the gaze redirection prior work of the applicant, featuring a single neural network, a skilled person would obviously think to apply to the single neural network of D2 the aforementioned standard concept of knowledge distillation in order to reduce its size and the number of outputted editing instructions, hence the

bandwidth requirements of D2, for its better use in videoconferencing systems. A skilled person would therefore arrive to the combination of features of claims 6 and 7 without having to involve inventive skills.

Outlook

- 9 In case of entering examination before the EPO, it is suggested to advance towards patentability by filing amended independent claims including the suggestion of point 7.5, accompanied by the alleviation of the further deficiencies of the claims and description recited under points 3 and 4.
- 10 It is to be noted that the aforementioned suggestion of the search examiner is given without prejudice as to the final opinion of a future examining division, and that it is directed to assist the applicant in his decision on how to proceed. It in no way precludes consideration of alternative solutions submitted by the applicant.
- 11 Other inventions than the one chosen for the prosecution of this application may be made the object of divisional applications under Article 76 EPC. The applicant is also invited to thoroughly check the application for necessary excisions of subject-matter that does not fall under the scope of protection of the prosecuted invention.

TITLE: ADJUSTING A DIGITAL REPRESENTATION OF A HEAD REGION

APPLICANT: RealD Spark, LLC

IPC CLASSIFICATION: G06F3/01, G06K9/00, H04N5/228, H04N7/14, H04N13/00

EXAMINER: Bouchaâla, Nicolas

CONSULTED DATABASES: DOSYS, EPODOC, INET, INSPEC, NPL

CLASSIFICATION SYMBOLS DEFINING EXTENT OF THE SEARCH:

IPC:

CPC: G06T19/LOW, G06T99-SK ARTIFICIAL INTELLIGENCE (G06T), G06T7/70/LOW, G06N3/08, G06N3/0454, G06K9/00597/LOW, LEARNING OR TRAINING N3N5N7N20 - TRANSFER LEARNING (G06N)

FI/F-TERMS:

KEYWORDS OR OTHER ELEMENTS FEATURING THE INVENTION:

image correction of head pose / gaze correction; image feature analysis; machine-learning; training of less accurate machine with result of high accurate machine, knowledge distillation, teacher student; training compressed transform

**SUPPLEMENTARY
EUROPEAN SEARCH REPORT**

Application Number
EP 18 84 4651

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	Yaroslav Ganin ET AL: "DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation", 26 July 2016 (2016-07-26), XP055295123, DOI: 10.1007/978-3-319-46475-6_20 Retrieved from the Internet: URL:http://arxiv.org/pdf/1607.07215v2.pdf * sections 3, 4 *	1-5, 8-11, 13-15	INV. G06F3/01 G06K9/00 H04N5/228 H04N7/14 H04N13/00
X	US 2017/195662 A1 (SOMMERLADE ERIC [GB] ET AL) 6 July 2017 (2017-07-06) * the whole document *	1-11, 13-15	
A	JIEBO LUO ET AL: "Face Location in Wavelet-Based Video Compression for High Perceptual Quality Videoconferencing", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, US, vol. 6, no. 4, August 1996 (1996-08), XP011014312, ISSN: 1051-8215 * abstract *	5,9	TECHNICAL FIELDS SEARCHED (IPC) G06T G06N
A	SUN XIANGNAN ET AL: "Construction and compression of face models for multi-party videoconferencing with multi-camera", 2013 6TH INTERNATIONAL CONGRESS ON IMAGE AND SIGNAL PROCESSING (CISP), IEEE, vol. 2, 16 December 2013 (2013-12-16), pages 948-952, XP032569425, DOI: 10.1109/CISP.2013.6745301 * section II *	5,9	
The supplementary search report has been based on the last set of claims valid and available at the start of the search.			
Place of search Berlin		Date of completion of the search 20 April 2021	Examiner Bouchaâla, Nicolas
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

2
EPO FORM 1503 03.82 (P04N04)

**SUPPLEMENTARY
EUROPEAN SEARCH REPORT**

Application Number
EP 18 84 4651

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	KATSIGIANNIS STAMOS ET AL: "A GPU based real-time video compression method for video conferencing", 2013 18TH INTERNATIONAL CONFERENCE ON DIGITAL SIGNAL PROCESSING (DSP), IEEE, July 2013 (2013-07), pages 1-6, XP032498481, ISSN: 1546-1874, DOI: 10.1109/ICDSP.2013.6622719 [retrieved on 2013-10-06] * abstract *	5,9	
A	----- GEOFFREY HINTON ET AL: "Distilling the Knowledge in a Neural Network", CORR (ARXIV), vol. 1503.02531v1, 9 March 2015 (2015-03-09), pages 1-9, XP055549014, * abstract *	6,7	
A	----- YIM JUNHO ET AL: "A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning", 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), IEEE COMPUTER SOCIETY, US, 21 July 2017 (2017-07-21), pages 7130-7138, XP033250080, ISSN: 1063-6919, DOI: 10.1109/CVPR.2017.754 [retrieved on 2017-11-06] * sections 1, 2 *	6,7	
The supplementary search report has been based on the last set of claims valid and available at the start of the search.			TECHNICAL FIELDS SEARCHED (IPC)
Place of search Berlin		Date of completion of the search 20 April 2021	Examiner Bouchaâla, Nicolas
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

2
EPO FORM 1503 03.82 (P04N04)

**SUPPLEMENTARY
EUROPEAN SEARCH REPORT**

Application Number
EP 18 84 4651

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	CRISTIAN BUCILUA ET AL: "Model compression", PROCEEDINGS OF THE TWELFTH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING : AUGUST 20 - 23, 2006, PHILADELPHIA, PA, USA, NEW YORK, NY : ACM PRESS, 2 PENN PLAZA, SUITE 701 NEW YORK NY 10121-0701 USA, 20 August 2006 (2006-08-20), pages 535-541, XP058107730, DOI: 10.1145/1150402.1150464 ISBN: 978-1-59593-339-3 * abstract *	6,7	
A	KONONENKO DANIIL ET AL: "Learning to look up: Realtime monocular gaze correction using machine learning", 2015 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), IEEE, 7 June 2015 (2015-06-07), pages 4667-4675, XP032793924, DOI: 10.1109/CVPR.2015.7299098 * section 3 *	1-11, 13-15	
The supplementary search report has been based on the last set of claims valid and available at the start of the search.			TECHNICAL FIELDS SEARCHED (IPC)
Place of search Berlin		Date of completion of the search 20 April 2021	Examiner Bouchaâla, Nicolas
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

2
EPO FORM 1503 03.82 (P04N04)



Application Number

EP 18 84 4651

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing claims for which payment was due.

☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for those claims for which no payment was due and for those claims for which claims fees have been paid, namely claim(s):

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for those claims for which no payment was due.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

☐ All further search fees have been paid within the fixed time limit. The present (supplementary) European search report has been drawn up for all claims.

☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.

☐ Only part of the further search fees have been paid within the fixed time limit. The present (supplementary) European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:

☒ None of the further search fees have been paid within the fixed time limit. The present (supplementary) European search report has been drawn up for those parts of the European patent application which relate to the first mentioned in the claims, namely claims:

1-11, 13-15

**LACK OF UNITY OF INVENTION
SHEET B**

Application Number
EP 18 84 4651

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. claims: 1-11, 13-15

providing the 2nd editing instructions in the 2nd reference data in a compressed representation

1.1. claims: 6, 7

providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data

2. claim: 12

online updating of the 2nd machine-learning-algorithm, based on the correction of a video stream predicted by the 1st machine-learning-algorithm, in order to improve the machine learning prediction "on the fly"

Please note that all inventions mentioned under item 1, although not necessarily linked by a common inventive concept, could be searched without effort justifying an additional fee.

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 18 84 4651

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-04-2021

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017195662 A1	06-07-2017	CN 108463787 A	28-08-2018
		EP 3400706 A1	14-11-2018
		US 2017195662 A1	06-07-2017
		US 2020021794 A1	16-01-2020
		US 2021044794 A1	11-02-2021
		WO 2017120247 A1	13-07-2017

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication ?
Contact Customer Services at www.epo.org/contact



J A Kemp LLP
14 South Square
Gray's Inn
London WC1R 5JJ
ROYAUME UNI

Date
03.08.21

Reference N411553EP	Application No./Patent No. 18844651.2 - 1230 / 3665553 PCT/US2018045648
Applicant/Proprietor RealD Spark, LLC	

Communication

The extended European search report is enclosed.

The extended European search report includes, pursuant to Rule 62 EPC, the supplementary European search report (Art. 153(7) EPC) and the European search opinion.

Copies of documents cited in the European search report are attached.

☒ 0 additional set(s) of copies of such documents is (are) enclosed as well.

Refund of the search fee

If applicable under Article 9 Rules relating to fees, a separate communication from the Receiving Section on the refund of the search fee will be sent later.

Should you wish to further prosecute this application in the examination phase, your attention is drawn to the provisions of Rule 70a EPC. An invitation to respond to the extended European search report will be issued shortly (R. 70(2) EPC).



Datum
Date cf Form 1507
Date

Blatt
Sheet 1
Feuille

Anmelde-Nr:
Application No: 18 844 651.2
Demande n°:

The examination is being carried out on the **following application documents**

Description, Pages

1-35 as published

Claims, Numbers

1-15 filed in electronic form on 17-09-2020

Drawings, Sheets

1/12-12/12 as published

* * * *

Following documents are cited in the European search report:

- D1 XP055295123
Yaroslav Ganin et al.: "DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation", pages 1-18, 26 July 2016 (2016-07-26), <http://arxiv.org/pdf/1607.07215v2.pdf>
- D2 US 2017/195662 A1 (SOMMERLADE ERIC [GB] ET AL) 6 July 2017 (2017-07-06)
- D3 XP055295123
Jiebo Luo et al.: "Face Location in Wavelet-Based Video Compression for High Perceptual Quality Videoconferencing", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 4, 1 August 1996 (1996-08-01), ISSN: 1051-8215
- D4 XP032569425
Sun Xiangran et al.: "Construction and compression of face models for multi-party videoconferencing with multi-camera", 6th IEEE International Congress on image and Signal Processing (CISP), vol. 2, 16 December 2013 (2013-12-16), pages 948-952, DOI: 10.1109/CISP.2013.6745301

- D5 XP058107730
Katsigiannis Stamos et al.: "A GPU based real-time video compression method for video conferencing", 18th IEEE International Conference on Digital Signal Processing July 2013 (2013-07), pages 1-6, ISSN: 1546-1874 , DOI: 10.1109/ICDSP.2013.6622719
- D6 XP055549014
Geoffrey Hinton et al.: "Distilling the Knowledge in a Neural Network", 9 March 2015 (2015-03-09), pages 1-9, <https://arxiv.org/pdf/1503.02531.pdf>
- D7 XP033250080
Yim Junho et al.: "A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21 July 2017 (2017-07-21), pages 7130-7138, ISSN: 1063-6919, DOI: 10.1109/CVPR.2017.754
- D8 XP058107730
Cristian Bucilă et al.: "Model compression", Proceedings of the 12th ACM International Conference on Knowledge Discovery and Data Mining, 20 August 2006 (2006-08-20), pages 535-541, DOI: 10.1145/1150402.1150464 , ISBN: 978-1-59593-339-3
- D9 XP032793924
Kononenko D. et al.: "Learning to look up: Realtime monocular gaze correction using machine learning", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7 June 2015 (2015-06-07), pages 4667-4675, DOI: 10.1109/CVPR.2015.7299098

* * * *

- 1 An international preliminary report on patentability (IPRP) has already been drawn up for the present application in accordance with the PCT, stating a lack of inventive step of the published PCT-claims 1-5, 20, 21, 25-30, 33-36. No opinion has been established for the remainder of the published PCT-claims, for not being drafted in accordance with Rule 6.4(a) PCT.

- 2 After entering the regional before the EPO, and with letter L0 received on 17-09-2020, the applicant has filed replacement claims 1-15, featuring the merging of PCT-claims, and the deletion of PCT-claims 10-12, 20, 21, 25-30, 34-36. The remaining independent claims were not amended. No arguments with respect to the IPRP's statement of lack of inventive step have been filed, which is concerning the present claims 1-3, 15.

Various deficiencies

- 3 The application exhibits following deficiencies:
- 3.1 Even though claim 11 includes explicitly or implicitly all the features of claim 10, it is not drafted as dependent on claim 10 as required by Rule 43(4) EPC.
- 3.2 The features of the claims should be provided with reference signs placed in parentheses to increase the intelligibility of the claims (Rule 43(7) EPC).
- 3.3 Documents reflecting the described prior art should preferably be cited in the background part of the description (Rule 42(1)(b) EPC). In the present case, documents D1 and D2 should be cited.
- 3.4 In the description paragraphs [47], [48], [67], [75], [76], [77], [83], [84], [85], [88], [149], the cited background art is "incorporated by reference". In case the applicant considers portions of the referenced document as essential for carrying out the claimed invention, the matter in question should be expressly incorporated into the description, as the patent specification must, regarding the essential features of the invention, be self-contained (Article 83 EPC). As to the particular conditions allowing the introduction of features from a cross-referenced document into the claims, reference is made to the Guidelines H, IV, 2.2.1. It is the provisional opinion of the search examiner that no insertion of background art is further necessary to comply with Article 83 EPC. Said incorporation by reference needs however to be deleted from the description before the end of the examination procedure (Guidelines F-III, 8).
- 3.5 Statements implying that the scope of protection may deviate from the claimed invention (sub-combinations of features, "spirit", "equivalents", claim-like clauses) should be deleted or amended in order to clarify that the scope of protection of the invention is exclusively defined by the claims (Article 84 EPC, Guidelines F-IV, 4.4). This appears to apply to paragraph [150] (necessary complete deletion of claim-like clauses), to paragraph [151] (suggested

complete deletion or at least deletion of "equivalents"), and to paragraph [152] (suggested complete deletion or at least deletion of the contradictions to the definitions of Rule 43 EPC, deletion of "equivalents").

Lack of clarity

- 4 The claims do not comply with the requirements of Article 84 EPC:
- 4.1 In independent claim 15, the claimed first and second "reference data" are not clearly defined in their relation to the problem of head pose correction and may therefore be seen as arbitrary processing features at least over a part of their scope of protection. In fact, the training of the claimed machine learning algorithms, may result also in processing data other than only head pose correction instructions.
- 4.2 In order to include at least implicitly a tangible technical effect of head pose correction, the claimed reference data has to include the editing instructions as presently claimed in independent claim 1.
- 4.3 It is to be noted that this lack of clarity of the claimed reference data has the consequence that the subject-matter of claim 15 is obvious over any prior art of head pose correction at least over part of its scope of protection, as the machine learning may presently be construed as an arbitrary mathematical processing not amenable to contribute to an inventive step in the sense of Article 56 EPC (see Guidelines G VII, 5.4).

Lack of unity of invention

- 5 The application lacks unity within the meaning of Article 82 EPC.
- The examiner considers that the following separate alleged inventions are not so linked as to form a single general inventive concept:
1. claims: 1-5, 8-11, 13-15
- providing the 2nd editing instructions in the 2nd reference data in a compressed representation in order to reduce the bandwidth needs of a gaze correction
- 1.1. claims: 6, 7

providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data, ultimately a reduction of the bandwidth needs of a gaze correction

2. claim: 12

online updating of the 2nd machine-learning-algorithm, based on the correction of a video stream predicted by the 1st machine-learning-algorithm, in order to improve the machine learning prediction "on the fly"

5.1 An assessment with respect to the prior art is first to be made for the subject-matter in common to the three claim-groups:

5.1.1 D1 represents the document of prior art closest to the originally claimed invention, and anticipates the subject-matter of claim 1, namely

A method of generating reference data for adjusting a digital representation of a head region,

(see D1, abstract: "[...] Our deep architecture performs coarse-to-fine warping with an additional intensity correction of individual pixels. All these operations are performed in a feed-forward manner, and the parameters associated with different operations are learned jointly in the end-to-end fashion. After learning, the resulting neural network can synthesize images with manipulated gaze [...])"

the method comprising:

receiving training data comprising:

a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;

(see D1, section 3 "The model" 2nd sentence: "The model is trained on pairs of images corresponding to eye appearance before and after the redirection. The redirection angle serves as an additional input parameter that is provided both during training and at test time."

see also section 3. 1 "Data preparation"; see also section 4.2 "Training procedure")

using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;

and

using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

(see D1, figure 2 and section 3 entitled "The model", in particular in the 2nd paragraph: " The task of the network is therefore the prediction of the warping field. This field is predicted in two stages in a coarse-to-fine manner, where the decisions at the fine scale are being informed by the result of the coarse stage.";

see also section 3.2 "Warping modules" describing the two concatenated machine learning algorithms, as claimed: "Each of the two warping modules takes as an input the image, the position of the feature points, and the redirection angle. [...]"

Coarse warping. The last convolutional layer of the first (half-scale) warping module produces a pixel-flow field (a two-channel map), which is then up-sampled $D_{\text{coarse}}(I, \alpha)$ and applied to warp the input image by means of a bilinear sampler S [12,21] that finds the coarse estimate" [...]

Fine warping. In the fine warping module, the rough image estimate O_{coarse} and the upsampled low-resolution flow $D_{\text{coarse}}(I, \alpha)$ are concatenated with the input data (the image, the angle encoding, and the feature point encoding) at the original scale and sent to the $1\times$ -

scale network which predicts another two-channel flow D_{res} that amends the half-scale pixel-flow [...] The amended flow is used to obtain the final output";

Remark: the joint learning of the two CNNs for coarse and fine warping implies the claimed training)

- 5.1.2 The subject-matter of independent method-claim 1, and of the corresponding claims 13-15 in other categories is therefore not new in the sense of Article 54(1)(2) EPC.
- 5.1.3 The passages of D1 recited above disclose also the subject matter of independent claim 10, which regards the adjustment of a head image with the editing instructions resulting from the training of the 2nd machine learning algorithm. The subject-matter of claim 10 is therefore also not new over D1.
- 5.1.4 Moreover, the passages of D1 recited above explicitly or implicitly disclose also the additional features of dependent claims 2, 3, 4, 8, 11. The subject-matter of these claims is therefore also not new over D1.
- 5.2 The contribution of the different claim groups to the known prior art disclosed by D1, i.e. the Special Technical Features (STF) as defined by Rule 44(1) EPC, and their associated objective problems, may therefore be defined as follows:
- 5.2.1 Concerning claim group 1, the STF consist in the additional features of claims 5 or 9 of a compressed representation of the 2nd editing instructions. With regard to D1, the objective problem can be defined as how to reduce the transmission bandwidth of the transform parameters, for example if the gaze correction of D1 is to be used in teleconferencing.
- 5.2.2 Concerning claim group 1.1, the STF consist in the additional features of claim 6 or 7. With regard to D1, the objective problem may be the definition of an alternative architecture, with the purpose of improving the training of a (2nd) machine learning algorithm of poor accuracy, which is ultimately used for solving the objective problem of reducing the transmission bandwidth of the transform parameters necessary for gaze correction.
- 5.2.3 Concerning claim group 2, the STF consist in the additional features of claim 12. With regard to D1, the objective problem can be defined as how to train the 2nd machine learning algorithm "on the fly".
- 5.3 Conclusion

The STF of the different claim-groups are obviously not the same. Moreover, they can be implemented independently from each other, and the solved problems either are not the same, or, when the same (groups 1 and 2), cannot define a common inventive concept ("problem invention") because consisting in a standard problem of the technical field. Therefore, no technical relationship in the sense of Rule 44(1) EPC exists between the different claim groups, which may be regarded as an inventive link.

Hence, the aforementioned claim groups are not so linked as to form a single general inventive concept, and do therefore not comply with the requirements of Article 82 EPC.

Lack of novelty over D1

- 6 The subject-matter of **claims 1-4, 8, 10, 11 ,13-15** is not new in the sense of Article 54(1) and (2) EPC, and therefore the requirements of Article 52(1) EPC are not met, as it is anticipated by D1 (see section 5.1 above).

Lack of inventive step

- 7 Furthermore, the subject-matter of **claims 5 and 9** does not involve an inventive step in the sense of Article 56 EPC, and therefore the requirements of Article 52(1) EPC are not met, for the following reasons:
- 7.1 The additional features of dependent claims 5, 9 concern the claimed second editing instructions being in a compression representation.
- 7.2 In the overall description, the compression is meant as an integral part of the training of the machine learning algorithms. This can be derived in particular from figures 9, 12, where the auxiliary data "Auxd" is defined in paragraph [125] to be input into the 1st machine to refine "a basis set used for providing a compressed representation of image editing instructions".
- 7.3 However, the wording of claim 5 or 9 at present includes also a post-compression of the 2nd editing instructions output by the training of the machine learning algorithms. As video compression including the compression of imaging instructions is a standard measure for video-conferencing (*see for*

example documents D3-D5), the skilled person would straightforwardly implement the additional features of claims 5, 9, when using the method of D1 in the framework of videoconferencing.

- 7.4 The features of PCA or wavelet compression are standard compression methods, which do *prima facie* not add any surprising technical advantage that would prove the involvement of inventive skills.
- 7.5 It is suggested to alleviate this objection with respect to D1 by including that the compression is an integral part of the training of the machine learning algorithms, for example by including in the claimed training data "a basis set used for providing a compressed representation of editing instructions" based on paragraph [25]. This intermediate generalisation from [25] and figure 9, is deemed by the search examiner to fulfil the criteria set out in the Guidelines H-V, 3.2.1 and not to contravene Article 123(2) EPC.
- 8 Finally, the subject-matter of **claims 6 and 7** does not involve an inventive step in the sense of Article 56 EPC, and therefore the requirements of Article 52(1) EPC are not met, for the following reasons:
- 8.1 When contemplated with respect to D1, the additional features of dependent claims 5, 9 are providing an alternative machine learning architecture ensuring a number of 2nd editing instructions lower than the number of 1st editing instructions, i.e a reduction of the 2nd reference data, ultimately a reduction of the bandwidth needs of a gaze correction. Starting from D1, this would mean to reduce the accuracy of the 2nd CNN trained to provide a refined transform. Even though bandwidth requirements could force a skilled person to reduce the accuracy of D1, there seems to be no obvious reason to modify the coarse-to-fine processing of D1 in this way.
- 8.2 In assessing other prior art, it however resorts that the application's concept (underlying these features) of improving the training of a less accurate machine learning algorithm additionally with the result of a pre-trained more accurate machine learning algorithm, is a concept well known in the art of artificial intelligence, in particular under the name of "knowledge distillation" (*see for example documents D6-D8*). When starting from D2, which discloses the gaze redirection prior work of the applicant, featuring a single neural network, a skilled person would obviously think to apply to the single neural network of D2 the aforementioned standard concept of knowledge distillation in order to reduce its size and the number of outputted editing instructions, hence the

bandwidth requirements of D2, for its better use in videoconferencing systems. A skilled person would therefore arrive to the combination of features of claims 6 and 7 without having to involve inventive skills.

Outlook

- 9 In case of entering examination before the EPO, it is suggested to advance towards patentability by filing amended independent claims including the suggestion of point 7.5, accompanied by the alleviation of the further deficiencies of the claims and description recited under points 3 and 4.
- 10 It is to be noted that the aforementioned suggestion of the search examiner is given without prejudice as to the final opinion of a future examining division, and that it is directed to assist the applicant in his decision on how to proceed. It in no way precludes consideration of alternative solutions submitted by the applicant.
- 11 Other inventions than the one chosen for the prosecution of this application may be made the object of divisional applications under Article 76 EPC. The applicant is also invited to thoroughly check the application for necessary excisions of subject-matter that does not fall under the scope of protection of the prosecuted invention.

TITLE: ADJUSTING A DIGITAL REPRESENTATION OF A HEAD REGION

APPLICANT: RealD Spark, LLC

IPC CLASSIFICATION: G06F3/01, G06K9/00, H04N5/228, H04N7/14, H04N13/00

EXAMINER: Bouchaâla, Nicolas

CONSULTED DATABASES: DOSYS, EPODOC, INET, INSPEC, NPL

CLASSIFICATION SYMBOLS DEFINING EXTENT OF THE SEARCH:

IPC:

CPC: G06T19/LOW, G06T99-SK ARTIFICIAL INTELLIGENCE (G06T), G06T7/70/LOW, G06N3/08, G06N3/0454, G06K9/00597/LOW, LEARNING OR TRAINING N3N5N7N20 - TRANSFER LEARNING (G06N)

FI/F-TERMS:

KEYWORDS OR OTHER ELEMENTS FEATURING THE INVENTION:

image correction of head pose / gaze correction; image feature analysis; machine-learning; training of less accurate machine with result of high accurate machine, knowledge distillation, teacher student; training compressed transform



Request for changes relating to an association of professional representatives

European Patent Office
Directorate 5.2.3 – Legal Division
80298 Munich
GERMANY

Name of association

J A Kemp LLP

Registration No.

811

Address

14 South Square
Gray's Inn
GB-LONDON WC1R 5JJ

I. Request for a change in the composition

We request registration of a change in the composition of the association with effect from the date of receipt of the present request at the EPO or from the following subsequent date

Date

Members to be added¹ (see additional sheet where necessary)

Surname	First name	Representative No.	Signature	Date

Members to be deleted¹ (see additional sheet where necessary)

Surname	First name	Representative No.	Signature	Date

☒ Confirmation by at least one member of the association (see additional sheet where necessary)

Surname	First name	Representative No.	Signature	Date

¹ This form is only valid for changes relating to an association. If the address in the list of professional representatives before the European Patent Office (see Article 134(1) EPC) should also be changed, a separate request using EPO Form 52301 (Request for registration of changes relating to an entry in the list of professional representatives before the European Patent Office) should be filed.

II. Request for a change of name¹

We request registration of a change of name of the association with effect from the date of receipt of the present request at the EPO or from the following subsequent date

Date

The new name is

☒ Confirmation by at least one member of the association (see additional sheet where necessary)

Surname	First name	Representative No.	Signature	Date

III. Request for a change of address¹

☒ We request registration of a change of address of the association with effect from the date of receipt of the present request at the EPO or from the following subsequent date

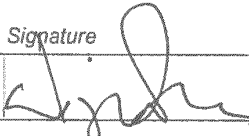
Date

27.9.21

The new address is

80 Turnmill Street
GB-LONDON EC1M 5QU

☒ Confirmation by at least one member of the association (see additional sheet where necessary)

Surname	First name	Representative No.	Signature	Date
Price	Nigel John King	62100		20.9.21

IV. Number of annexes (e.g. additional sheet for further members)

¹ This form is only valid for changes relating to an association. If the address in the list of professional representatives before the European Patent Office (see Article 134(1) EPC) should also be changed, a separate request using EPO Form 52301 (Request for registration of changes relating to an entry in the list of professional representatives before the European Patent Office) should be filed.



Letter accompanying subsequently filed items

Representative:

J A Kemp LLP
811
80 Turnmill Street
London Greater London EC1M 5QU
United Kingdom

Phone: +44 20 3077 8600
Fax: +44 20 7430 1000
E-mail: mail@jakemp.com

80298 Munich
Germany
Tel. +49(0)89 2399-0 | Fax -4465

P.O. Box 5818
2280 HV Rijswijk
Netherlands
Tel. +31(0)70 340-2040 | Fax -3016

10958 Berlin
Germany
Tel. +49(0)30 25901-0 | Fax -840

The document(s) listed below is (are) subsequently filed documents pertaining to the following application:

Application number

18844651.2

Applicant's or representative's reference

N411553EP DQF

	Description of document	Original file name	Assigned file name
1	Reply to examination report	Response to examination report - N411553EP DQF.PDF	EXRE3-1.PDF
2	Amended description (clean copy)	Amended description clean pages Feb 22 N411553EP.pdf	DESC-1.pdf
3	Amended description with annotations	Amended description tracked pages Feb 22 N411553EP.pdf	DESC-HWA-1.pdf
4	Amended claims (clean copy)	Amended claims clean Feb 2022 - N411553EP.PDF	CLMS-1.PDF
5	Amended claims with annotations	Amended claims tracked Feb 2022 - N411553EP.PDF	CLMS-HWA-1.PDF

	Payment	
1	Method of payment	Not specified

Signatures

Place: Paris
Date: 22 February 2022
Signed by: Dominic Forsythe 23485
Association: J A Kemp LLP
Representative name: /Dominic Forsythe/
Capacity: (Representative)

N411553EP DQF

J A • K E M P

BY ONLINE FILING

The European Patent Office
80298 Munich
Germany

22 February 2022

European Patent Application No. 18844651.2
REALD SPARK, LLC
Our Ref: N411553EP DQF/EE15

In response to the communication pursuant to Rules 70(2) and 70a(2) EPC, we hereby confirm that the applicant wishes to proceed further with this application and to have it examined.

We request that examination proceed on the basis of attached amended description pages 1-4, 8-15 and 29-30 replacing description pages 1-4, 8-15 and 29-35, and attached amended claims pages 36-40 replacing all of the claims pages currently on file. We attach also marked up versions of the amended pages to show where changes have been made.

The attached amended pages and the comments provided below respond to the opinion accompanying the European search report.

Amendments and basis

In response to item 3.1, we have amended claim 11 to be dependent on claim 10.

In response to item 3.2, we have added reference numerals throughout the claims.

In response to item 3.3, we have amended the introductory part of the description to acknowledge D1 and D2. We have also made amendments to ensure conformity with the scope of the newly amended claims.

In response to item 3.4, we have deleted “incorporated by reference” statements throughout the description.

In response to item 3.5, we have deleted the claim-like clauses and paragraphs [151] and [152].

In response to items 4.1-4.3, we have amended claim 15 to correspond in scope to claim 1.

PATENT ATTORNEYS • TRADE MARK ATTORNEYS
LONDON • PARIS • OXFORD • CAMBRIDGE • MUNICH

80 Turnmill Street, London EC1M 5QU
T +44 20 3077 8600 F +44 20 7430 1000
mail@jakemp.com www.jakemp.com

J A Kemp LLP is a Limited Liability Partnership, registered in England and Wales (registered number OC427581) and regulated by IPReg.
Our registered office is at 80 Turnmill Street, London EC1M 5QU, UK. Our use of the word ‘partner’ refers to a member of J A Kemp LLP.

In response to items 5-7, we have amended independent claims 1 and 15 in accordance with the helpful suggestion in item 7.5. In particular, we have specified that the received training data comprises “a basis set used for providing a compressed representation of editing instructions” and that “the second editing instructions (ED2) in the second reference data (102) are provided in the compressed representation using the basis set”.

Basis may be found for example in Figures 9 and 12 and originally filed [125], which describe how the compression is meant as an integral part of the training of the machine learning algorithms. The auxiliary data “Auxd” is defined in [125] to be input into the first machine learning algorithm to define “a basis set used for providing a compressed representation of image editing instructions”.

Clarity

The objections in item 4 have been dealt with by the amendments described above.

Unity

The amendments to claims 1 and 15 render these claims novel and inventive relative to the cited prior art. The dependent claims are unified by their dependencies.

Novelty and Inventive Step

The amendments to claims 1 and 15 render these claims novel and inventive. There is no hint or suggestion in the prior art which would have led the skilled person to modify any of the arrangements of D1 in such a way as to incorporate compression as an integral part of the training of the machine learning algorithm.

Concluding remarks

We request that a communication under Article 94(3) be issued if any objections remain. Alternatively, if considered likely to be more efficient (for instance, if only minor remaining issues are identified), then the examiner is invited to contact us on +44 20 3077 8600 or mail@jakemp.com.

As a precaution, we request that oral proceedings be held before any decision adverse to the applicant is made.

Electronically Signed
DOMINIC FORSYTHE
FOR J A KEMP LLP
ASSOCIATION NO. 811

Claims:

1. A method of generating reference data for adjusting a digital representation of a head region, the method comprising:
 - receiving training data (100, TD) comprising:
 - a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch;
 - a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region; and
 - a basis set used for providing a compressed representation of editing instructions;
 - using a first machine learning algorithm (MLA1) to generate first reference data (101) using the training data (100, TD), the first reference data (101) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and
 - using a second machine learning algorithm (MLA2) to generate second reference data (102) using the same training data (100, TD) as the first machine learning algorithm (MLA1) and the first reference data (101) generated by the first machine learning algorithm (MLA1), the second reference data (102) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region,
 - wherein the second editing instructions (ED2) in the second reference data (102) are provided in the compressed representation using the basis set.
2. The method of claim 1, wherein:
 - the first reference data (101) comprise first editing instructions (ED1) for a range of possible configurations of the target feature and first selection instructions (LA1) for selecting editing instructions for a particular input patch from the first editing instructions (ED1) based on the configuration of the target feature of the input patch; and

the second reference data (102) comprise second editing instructions (ED2) for a range of possible configurations of the target feature and second selection instructions (LA2) for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch, wherein, optionally: the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch.

3. The method of claim 2, wherein the editing instructions comprise:
a displacement vector field defining how the input patch is to be transformed or a filter field, a brightness adjustment field, or a texture blending field.
4. The method of claim 2 or 3, wherein:
a first editing algorithm (EA1) is used by the first machine learning algorithm (MLA1) to define how the first editing instructions (ED1) are to be applied to an input patch to derive an output patch; and
a second editing algorithm (EA2) is used by the second machine learning algorithm (MLA2) to define how the second editing instructions (ED2) are to be applied to an input patch to derive an output patch.
5. The method of claim 4, wherein:
the second editing instructions (ED2) are principle component analysis components of a principle component analysis of the first editing instructions (ED1); and the second image editing algorithm (EA2) is configured to transform the second editing instructions (ED2) into the first editing instructions (ED1) by inverse principle component analysis transform; or
the second editing instructions (ED2) are wavelet components of the first editing instructions (ED1); and the second editing algorithm (EA2) is configured to transform the second editing instructions (ED2) into the first editing instructions (ED1) by inverse wavelet transform.

6. The method of any of claims 2-5, wherein the first selection instructions (LA1) for the first reference data (101) are able to select between a larger number of alternative editing instructions than the second selection instructions (LA2) for the second reference data (102).
7. The method of any of claims 1-6, wherein the first machine learning algorithm (MLA1) comprises a first neural network and the second machine learning algorithm (MLA2) comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.
8. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair, wherein, optionally:
- the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation of the head region comprises adjusting a gaze direction;
 - the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the nose;
 - the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the chin;
 - the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the neck; and/or
 - the target feature comprises a hair region comprising hair and the adjustment of the digital representation of the head region comprises adjusting a color of the hair.
9. The method of any preceding claim, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers.

10. A method of adjusting a digital representation of a head region, the method comprising receiving a digital representation of a head region; and using reference data comprising editing instructions to adjust the digital representation of the head region, wherein the reference data comprises the second reference data generated by the method of any of claims 1-9.
11. The method of claim 10, comprising:
identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;
deriving a feature vector from plural local descriptors of the target patch;
using the feature vector to select editing instructions from the reference data, the reference data comprising editing instructions for a range of possible values of the feature vector;
and
applying the selected editing instructions to the target patch to adjust the digital representation of the head region.
12. The method of claim 10 or 11, wherein a plurality of the digital representations of the head region are received and adjusted, each digital representation being captured from the head region at a different point in time, and the method further comprises updating the second machine learning algorithm (MLA2) using editing instructions generated by the first machine learning algorithm (MLA1) using one or more of the received and adjusted digital representations, wherein the updating of the second machine learning algorithm (MLA2) is optionally performed at or near input data frame rate and, optionally, only a subset of the received and adjusted digital representations are used to update the second machine learning algorithm (MLA2).
13. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.
14. A computer readable storage medium storing a computer program according to claim 13.

15. A device for generating reference data for adjusting a digital representation of a head region, the device comprising a data processing unit (120) arranged to:

- receive training data (100, TD) comprising:
 - a set of input patches, each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
 - a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;
 - a basis set used for providing a compressed representation of editing instructions;
- use a first machine learning algorithm (MLA1) to generate first reference data (101) using the training data (100, TD), the first reference data (101) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and
- use a second machine learning algorithm (MLA2) to generate second reference data (102) using the same training data (100, TD) as the first machine learning algorithm (MLA1) and the first reference data (101) generated by the first machine learning algorithm (MLA1), the second reference data (102) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region,
- wherein the second editing instructions (ED2) in the second reference data (102) are provided in the compressed representation using the basis set.

Claims:

1. A method of generating reference data for adjusting a digital representation of a head region, the method comprising:

receiving training data (100, TD) comprising:

a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; ~~and~~

a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region; and

a basis set used for providing a compressed representation of editing instructions;

using a first machine learning algorithm (MLA1) to generate first reference data (101) using the training data (100, TD), the first reference data (101) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and

using a second machine learning algorithm (MLA2) to generate second reference data (102) using the same training data (100, TD) as the first machine learning algorithm (MLA1) and the first reference data (101) generated by the first machine learning algorithm (MLA1), the second reference data (102) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region;-

wherein the second editing instructions (ED2) in the second reference data (102) are provided in the compressed representation using the basis set.

2. The method of claim 1, wherein:

the first reference data (101) comprise first editing instructions (ED1) for a range of possible configurations of the target feature and first selection instructions (LA1) for selecting editing instructions for a particular input patch from the first editing instructions (ED1) based on the configuration of the target feature of the input patch; and

the second reference data (102) comprise second editing instructions (ED2) for a range of possible configurations of the target feature and second selection instructions (LA2) for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch, wherein, optionally: the configuration of the target feature of each input patch is represented by a feature vector derived from plural local descriptors of the input patch, and the first and second selection instructions define how the feature vector is used to select editing instructions for the input patch.

3. The method of claim 2, wherein the editing instructions comprise:

a displacement vector field defining how the input patch is to be transformed or a filter field, a brightness adjustment field, or a texture blending field.

4. The method of claim 2 or 3, wherein:

a first editing algorithm (EA1) is used by the first machine learning algorithm (MLA1) to define how the first editing instructions (ED1) are to be applied to an input patch to derive an output patch; and

a second editing algorithm (EA2) is used by the second machine learning algorithm (MLA2) to define how the second editing instructions (ED2) are to be applied to an input patch to derive an output patch.

5. The method of claim 4, wherein:

the second editing instructions (ED2) are principle component analysis components of a principle component analysis of the first editing instructions (ED1); and the second image editing algorithm (EA2) is configured to transform the second editing instructions (ED2) into the first editing instructions (ED1) by inverse principle component analysis transform; or

the second editing instructions (ED2) are wavelet components of the first editing instructions (ED1); and the second editing algorithm (EA2) is configured to transform the second editing instructions (ED2) into the first editing instructions (ED1) by inverse wavelet transform.

6. The method of any of claims 2-5, wherein the first selection instructions (LA1) for the first reference data (101) are able to select between a larger number of alternative editing instructions than the second selection instructions (LA2) for the second reference data (102).

7. The method of any of claims 1-6, wherein the first machine learning algorithm (MLA1) comprises a first neural network and the second machine learning algorithm (MLA2) comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.

8. The method of any preceding claim, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair, wherein, optionally:

the target feature comprises an eye region comprising at least part of an eye and the adjustment of the digital representation of the head region comprises adjusting a gaze direction;

the target feature comprises a nose region comprising at least part of a nose and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the nose;

the target feature comprises a chin region comprising at least part of a chin and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the chin;

the target feature comprises a neck region comprising at least part of a neck and the adjustment of the digital representation of the head region comprises adjusting a shape and/or texture of the neck; and/or

the target feature comprises a hair region comprising hair and the adjustment of the digital representation of the head region comprises adjusting a color of the hair.

9. The method of any preceding claim, wherein ~~the second editing instructions in the second reference data are provided in a compressed representation, the~~ the compressed representation optionally comprising one or more of the following: a principle component analysis

representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centers.

10. A method of adjusting a digital representation of a head region, the method comprising receiving a digital representation of a head region; and
using reference data comprising editing instructions to adjust the digital representation of the head region, wherein the reference data comprises the second reference data generated by the method of any of claims 1-9.

11. ~~The~~A method of claim 10, ~~adjusting a digital representation of a head region, the method~~ comprising:

identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region;

deriving a feature vector from plural local descriptors of the target patch;

using the feature vector to select editing instructions from the reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and

applying the selected editing instructions to the target patch to adjust the digital representation of the head region, ~~wherein:~~

~~the reference data comprises the second reference data generated by the method of any of claims 1-9.~~

12. The method of claim 10 or 11, wherein a plurality of the digital representations of the head region are received and adjusted, each digital representation being captured from the head region at a different point in time, and the method further comprises updating the second machine learning algorithm (MLA2) using editing instructions generated by the first machine learning algorithm (MLA1) using one or more of the received and adjusted digital representations, wherein the updating of the second machine learning algorithm (MLA2) is optionally performed at or near input data frame rate and, optionally, only a subset of the received and adjusted digital representations are used to update the second machine learning algorithm (MLA2).

13. A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding claims.

14. A computer readable storage medium storing a computer program according to claim 13.

15. A device for generating reference data for adjusting a digital representation of a head region, the device comprising a data processing unit (120) arranged to:

receive training data (100, TD) comprising:

a set of input patches, each input patch comprising information about a target feature of a digital representation of the head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and

a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target ~~feature~~portion of the digital representation of the head region after adjustment of the digital representation of the head region;

a basis set used for providing a compressed representation of editing instructions;

use a first machine learning algorithm (MLA1) to generate first reference data (101) using the training data (100, TD), the first reference data (101) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and

use a second machine learning algorithm (MLA2) to generate second reference data (102) using the same training data (100, TD) as the first machine learning algorithm (MLA1) and the first reference data (101) generated output by the first machine learning algorithm (MLA1), the second reference data (102) comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

wherein the second editing instructions (ED2) in the second reference data (102) are provided in the compressed representation using the basis set.-

Adjusting a Digital representation of a Head Region

TECHNICAL FIELD

[01] This application relates to adjusting a digital representation, such as an image or a three-dimensional geometrical representation, of a head region, particularly a facial and/or neck region. The application relates particularly to adjusting target features of the digital representation of the head region, for example to correct a perceived gaze direction of eyes, or to modify the texture and/or shape of features such as the nose, mouth, chin or neck.

BACKGROUND

[02] In many systems, images of a head, which may comprise single images at different times, or stereoscopic pairs of images or other multi-view images, may be captured in one device and displayed on a different device for viewing by an observer. One non-limitative example is a system for performing teleconferencing between two telecommunications devices. In that case, each device may capture images of the observer of that device and transmit them to the other device over a telecommunications network for display and viewing by the observer of the other device. Digital representations other than images and/or complementary to images may also be captured, for example using depth measurements (e.g. using a time-of-flight camera).

[03] When an image or other digital representation of a head is captured and displayed, the gaze of the head may not be directed at the observer. This may be caused for example by the gaze of the head not being directed at the sensing system (e.g. camera system) used to capture the digital representation (e.g. image), for example because a user whose head is imaged is observing a display in the same device as a camera system and the camera system is offset above (or below) that display. In that case, the gaze in the displayed images will be perceived to be downwards (or upwards). The human visual system has evolved high sensitivity to gaze during social interaction, using cues gained from the relative position of the iris and white sclera of other observers. Errors in the perceived gaze are disconcerting. For example in a system for performing teleconferencing, errors in the perceived gaze can create unnatural interactions between the users.

[04] The present disclosure is concerned with processing techniques (e.g. image processing techniques) for adjusting digital representations (e.g. images) of a head region to

correct the perceived gaze and/or to improve other aspects of a computer-generated display of the head region. The present disclosure is particularly concerned with implementing such processing techniques with minimal demands on computer hardware and/or power such that they provide results at or near input data frame rate or user feedback requirements.

Yaroslav Ganin et al.: "DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation", pages 1-18, 26 July 2016 discloses methods for generating images of a given face with a redirected gaze. The problem is treated as a specific instance of conditional image generation. A deep architecture is disclosed for handling the task. The deep architecture performs coarse-to-fine warping with an additional intensity correction of individual pixels. All these operations are performed in a feed-forward manner, and the parameters associated with different operations are learned jointly in the end-to-end fashion. After learning, the resulting neural network can synthesize images with manipulated gaze, while the redirection angle can be selected arbitrarily from a certain range and provided as an input to the network.

US 2017/195662 A1 discloses a method in which gaze is corrected by adjusting multi-view images of a head. Image patches containing the left and right eyes of the head are identified and a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images. A displacement vector field representing a transformation of an image patch is derived, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector produced by machine learning. The multi-view images are adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field.

BRIEF SUMMARY

[05] The scope of the invention is defined in the appended claims.

[06] According to an aspect of the invention, there is provided the method of claim 1.

[07] The described use of two machine learning algorithms allows an improved balance to be achieved between accuracy of the adjustment process and speed of execution. The first machine learning algorithm can be configured to provide highly detailed first reference data, which provides high accuracy. Use of this first reference data directly in a processing technique (e.g. image processing technique) to adjust a digital representation (e.g. image or three-dimensional geometrical representation) of a head region would be relatively expensive in terms of computational resources because of the high level of detail. By arranging instead

for the first reference data to be provided to a second machine learning algorithm, which provides, based on the first reference data and the original training data, the reference data (the second reference data) that is to be used for the adjustment of the digital representation of the head region, it is possible to benefit to an extent from the high accuracy of the first machine learning algorithm whilst also providing reference data that is less detailed and thus easier to process efficiently when performing the adjustment of the digital representation of the head region. The quality of the reference data provided by the second machine learning algorithm is found to be significantly improved when the second machine learning algorithm is provided with both the first reference data and the training data in comparison to when the second machine learning algorithm is provided only with the training data.

[08] Efficiency is further improved by providing editing instructions (e.g. image editing instructions) in the second reference data in a compressed representation. The use of a compressed representation reduces data storage and bandwidth requirements during use of the editing instructions to perform adjustment of a digital representation of a head region.

[09] According to an alternative aspect of the invention, there is provided the device of claim 15.

[10] The present disclosure also describes a method of adjusting a digital representation of a head region, the method comprising: identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region; deriving a feature vector from plural local descriptors of the target patch; using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein: the editing instructions in the reference data are provided in a compressed representation.

[11] The present disclosure also describes a device configured to perform the above method.

[12] The present disclosure also describes a method of training a machine learning algorithm to adjust a digital representation of a head region, comprising: receiving training data comprising: a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and first reference data generated by a pre-trained first machine learning algorithm, the first reference data comprising

a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region; updating a pre-trained second machine learning algorithm trained to generate second reference data, where the input for the updating comprises the training data and the generated first reference data, the second reference data comprising editing instructions for adjusting the digital representation of the head region.

[13] The present disclosure also describes a method of training a machine learning algorithm to adjust a digital representation of a head region, the method comprising: receiving training data comprising a set of input digital representations of a head region; training a first machine learning algorithm using the training data to perform an adjustment of a digital representation of a head region; using the trained first machine learning algorithm to generate first reference data, the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations, each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform; and training a second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm.

BRIEF DESCRIPTION OF THE DRAWINGS

[14] Non-limitative embodiments are illustrated by way of example in the accompanying figures, in which like reference numbers indicate similar parts, and in which:

[15] **FIGURE 1** is a schematic perspective view of a device that captures a stereoscopic pair of images;

[16] **FIGURE 2** is a schematic perspective view of a device that displays the stereoscopic pair of images;

[17] **FIGURE 3** is a flow chart of a method of adjusting a stereoscopic pair of images;

[18] **FIGURE 4** is a diagram illustrating the processing of the stereoscopic pair of images in the method of FIGURE 3;

[19] **FIGURE 5** is a flow chart of a step of extracting an image patch;

[20] **FIGURE 6** and **FIGURE 7** are flow charts of two alternatives for a step of adjusting an image;

[21] **FIGURE 8** is flow chart of a method of generating reference data;

finding the spatial relationship may be applied, for example a calibration method using a reference image, or specification *a priori*.

[41] The camera or cameras 13 face in the same direction as the display 11. Thus, when the source observer 14 is viewing the display 11, then the camera or cameras 13 face the source observer 14 and the captured information, such as depth information, image or images (e.g. stereoscopic pair of images) are digital representations (e.g. images and/or three-dimensional geometrical representations) of the head of the source observer 14. Different cameras in the camera system can have different fields of view.

[42] The camera system 12 may include cameras 13 having different sensing modalities, including but not limited to visible light, infrared, and time-of-flight (depth).

[43] In some embodiments, the main output of the camera system 13 is images 31 which are typically video images output at a video rate. The output of the camera system 13 may also include data representing the spatial relationship between the cameras 13 and the display 11, the nature of the sensing modalities and internal parameters of the cameras 13 (for example focal length, optical axis) which may be used for angular localization, as well as three-dimensional geometrical information, for example from depth measurements.

[44] An example of the method performed on a digital representation of a head region comprising a stereoscopic pair of images 31, for the case of adjustment of eye regions, is as follows. To illustrate this example method, reference is also made to **FIGURE 4** which shows an example of the stereoscopic pair of images 31 at various stages of the method.

[45] In step S1, the stereoscopic pair of images 31 are analyzed to detect the location of the head and in particular the eyes of the source observer 14 within the stereoscopic pair of images 31. This is performed by detecting presence of a head, tracking the head, and localizing the eyes of the head. Step S1 may be performed using a variety of techniques that are known in the art.

[46] One possible technique for detecting the presence of the head is to use Haar feature cascades, for example as disclosed in Viola and Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, CVPR 2001, pp 1-9.

[47] One possible technique for tracking the head is to use the approach of Active Appearance Models to provide the position of the head of the subject, as well as the location of the eyes, for example as disclosed in Cootes et al., “Active shape models - their training and application”, Computer Vision and Image Understanding, 61(1):38-59, Jan. 1995 and in Cootes et al. “Active appearance models”, IEEE Trans. Pattern Analysis and Machine

Intelligence, 23(6):681-685, 2001.

[48] In step S1, typically, a set of individual points (“landmarks”) are set to regions of the face, typically the eyes, for example corners of the eye, upper and lower lid locations, etc, thereby localizing the eyes.

[49] In step S2, patches representing portions of a digital representation of the head region, which in this example may be referred to image patches, containing the left and right eyes of the head, respectively, are identified in each image 31 of the stereoscopic pair. FIGURE 4 shows the identified image patches 32 of the right eye in each image 31 (the image patches for the left eye being omitted in FIGURE 4 for clarity).

[50] Step S2 may be performed as shown in **FIGURE 5**, as follows.

[51] In step S2-1, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points (“landmarks”) corresponding to features of an eye, as shown for example in FIGURE 4.

[52] In step S2-2, the image patches 32 identified in step S2-1 are transformed into a normalized coordinate system, being the same normalized coordinate system as used in the machine learning process which is described further below. The transformation is chosen to align the points (“landmarks”) of the eye within the image patch that were identified in step S1, with predetermined locations in the normalized coordinate system. The transformation may include translation, rotation and scaling, to appropriate extents to achieve that alignment. The output of step S2-2 is identified image patches 33 of the right eye in each image in the normalized coordinate system as shown for example in FIGURE 4.

[53] The following steps may be performed separately (a) in respect of the image patches containing the left eyes of the head in each image 31 of the stereoscopic pair, and (b) in respect of the image patches containing the right eyes of the head in each image 31 of the stereoscopic pair (in this example). For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes.

[54] In step S3, a feature vector 34 is derived from plural local descriptors (representing information about a local region in a patch), which in this example may be referred to as local image descriptors, of an image patch 33 in at least one image 31 of the stereoscopic pair (in this example). Depending on the approach and as described further below, this may be an image patch in a single image 31 of the stereoscopic pair or may be both images 31 of the

stereoscopic pair. Thus, the local image descriptors are local image descriptors derived in the normalized coordinate system.

[55] The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 to be used for adjusting the image patches. The reference data 35 may comprise reference displacement vector fields that represent transformations of the image patch, or other representations of transformations of the image patch, including compressed representations as described below, and are associated with possible values of the feature vector.

[56] The reference data 35 is obtained and analyzed in advance using a machine learning technique. The machine learning technique may derive the form of the feature vectors 34 and associate transformations such as the reference displacement vector fields with the possible values of the feature vector. A specific example of a machine learning technique applied in the case where it is desired to correct gaze using digital representations of a head region comprising images of the head region, will now be described before reverting to the method of FIGURE 3.

[57] The training input to the machine learning technique is two sets of images (or image patches), which may be stereoscopic pairs of images or monoscopic images, as discussed further below. Each set may comprise images of the head of the same group of individuals but captured from cameras in different locations relative to the gaze so that the perceived gaze differs as between them (in the case where gaze is to be corrected).

[58] The first set are input images, being images of each individual with an incorrect gaze where the error is known *a priori*. In particular, the images in the first set may be captured by at least one cameras in a known camera location where the gaze of the individual which is in a different known direction. For example in the case of the source device of FIGURE 1, the camera location may be the location of a camera 13 and while the gaze of the imaged individual is towards the center of the display 11.

[59] The second set are output images, being images of each individual with correct gaze for a predetermined observer location relative to a display location in which the image is to be displayed. In the simplest case, the observer location is a normal viewing position perpendicular to the center of the display location, for example as shown by the hard outline of the destination observer 24 in the case of the destination device 20 of FIGURE 2.

[60] For each image in the two sets, the image is analyzed to detect the location of the head and in particular the eyes using the same technique as used in step S1 described above, and

then image patches containing the left and right eyes of the head, respectively, are identified using the same technique as used in step S2 described above. The following steps may then be performed separately (a) in respect of the image patches containing the left eyes of the head in each image, and (b) in respect of the image patches containing the right eyes of the head in each image. For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes in this embodiment.

[61] Each image patch is transformed into the same normalized coordinate system as used in step S2 described above. As described above, the transformation is chosen to align points (“landmarks”) of the eye with predetermined locations in the normalized coordinate system. The transformation may include, for example, translation, rotation and/or scaling, to appropriate extents to achieve that alignment.

[62] Thus, the image patches input and output images of each individual are aligned in the normalized coordinate system.

[63] From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image, for example as follows. Defining positions in the image patches by (x,y) , the displacement vector field $\mathbf{F}$ is given by

$$\mathbf{F}=\{u(x, y), v(x, y)\}$$

where u and v define the horizontal and vertical components of the vector at each position (x,y) .

[64] The displacement vector field $\mathbf{F}$ is chosen so that the image patch of the output image $O(x,y)$ is derived from the image patch of the input image $I(x,y)$ as

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[65] For image data from more than one camera, the system delivers a displacement vector field for the input image from each camera.

[66] The displacement vector field $\mathbf{F}$ for an input and output image of an individual may be derived using a process in which a trial feature vector $\mathbf{F}'=\{u',v'\}$ is modified to minimize error, optionally in an iterative process, for example in accordance with:

$$\sum |O(x, y) - I(x + u'(x, y), y + v'(x, y))| = \min!$$

[67] By way of non-limitative example, the displacement vector field $\mathbf{F}$ may be derived as disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675,

wherein the displacement vector field **F** is referred to as a “flow field”.

[68] Another example for editing instructions, which may be referred to as image editing instructions in cases where the digital representation of the head region consists of images, which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by filter field $L = \{k(P, x, y)\}$, which defines a filter kernel for a given location (x, y) . This filter field L is chosen so that the image patch of the output image $O(x, y)$ is derived from the image patch of the input image $I(x, y)$ as $O(x, y) = k(P(I, x, y), x, y)$, where $P(I, x, y)$ is a local region around the point (x, y) in the image I , and $k(P, x, y)$ operates on the patch P with coefficients depending on the position (x, y) . For example, it could be the output of a convolution of the patch with a Gaussian filter with width depending on the position x in the image, or a brightness increase of a local pixel depending on the vertical position y .

[69] Another example for editing instructions (e.g. image editing instructions) which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by a set of typical image components that make up the edited image area, e.g. texture showing beard stubbles. These are then blended with a factor depending on the image coordinates and local image content (i.e. a texture blending field). Other transformation fields can be used, such as a brightness adjustment field.

[70] A machine learning technique is used to obtain a map from the displacement vector field **F** (or other editing instructions such as image editing instructions) of each individual to respective feature vectors derived from plural local (e.g. image) descriptors of a target patch of an input image.

[71] The local descriptors (e.g. local image descriptors) capture relevant information of a local part of a patch (e.g. image patch) of an input digital representation of the head region (e.g. an input image) and the set of local descriptors usually form a continuous vectorial output.

[72] The local descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific image descriptors are selected and optimized by the machine learning process itself. In general, the local descriptors may be of any suitable type, some non-limitative examples which may be applied in any combination being as follows.

[73] The local descriptors may include values of individual pixels or a linear combination thereof. Such a linear combination may be, for example, a difference between the pixels at

two points, a kernel derived within a mask at an arbitrary location, or a difference between two kernels at different locations.

[74] The local descriptors may include distances of a pixel location from the position of an eye point (“landmark”).

[75] The local descriptors may include SIFT features (Scale-invariant feature transform features), for example as disclosed in Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision* 60 (2), pp 91–110.

[76] The local descriptors may include HOG features (Histogram of Oriented Gradients features), for example as disclosed in Dalal et al. “Histograms of Oriented Gradients for Human Detection”, *Computer Vision and Pattern Recognition*, 2005, pp. 886-893.

[77] The local descriptors may include “low level representations” from pre-classification stages in deep learning neural networks, for example as disclosed in Yang and Ramanan, “Multi-scale recognition with DAG-CNNs”, *ICCV 2015*. In a classifying deep learning neural network with multiple layers applied to an input digital representation (e.g. image), for example, such low level features could be taken from a layer before the final classification layer of the network.

[78] The derivation of the feature vector from plural local descriptors depends on the type of machine learning applied.

[79] In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local descriptors (e.g. local image descriptors) in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field $\mathbf{F}$ derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 may provide similar functionality to a look-up table, with the machine learning generating a machine learning parameter set that can be used to generate corresponding editing instructions. This allows a reference displacement vector field $\mathbf{F}$ to be simply selected from the reference data 35 on the basis of the feature vector 34 derived in step S3, as described below.

[80] In the case that the feature vector comprises features that are binary values derived from the local descriptors, the feature vector has a binary representation. Such binary values may be derived in various ways from the values of descriptors, for example by comparing the value of a descriptor with a threshold, comparing the value of two descriptors, or by comparing the distance of a pixel location from the position of an eye point (“landmark”).

[81] Alternatively, the feature vector may comprise features that are discretized values of the local descriptors. In this case, more than two discrete values of each feature are possible.

[82] Any suitable machine learning technique may be applied, for example using a decision tree, a decision forest, a decision fern or an ensemble or combination thereof, or a neural network.

[83] By way of example, a suitable machine learning technique using a feature vector comprising features that are binary values derived by comparing a set of individual pixels or a linear combination thereof against a threshold, is disclosed in Ozuysal et al. “Fast Keypoint Recognition in Ten Lines of Code”, Computer Vision and Pattern Recognition, 2007, pp. 1-8.

[84] By way of further example, a suitable machine learning technique using a distance of a pixel location with the position of an eye landmark is disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675.

[85] By way of further example, a suitable machine learning technique using a random decision forest is disclosed in Ho, “Random Decision Forests”, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995, pp. 278-282.

[86] In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local descriptors (e.g. local image descriptors) in a continuous space. In this case, the machine learning technique associates a reference displacement vector field $\mathbf{F}$ (in this example, but other editing instructions could be used) derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field $\mathbf{F}$ to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived in step S3 and the values of the feature vector associated with the reference displacement vector fields.

[87] Any suitable machine learning technique may be applied, for example using support vector regression.

[88] By way of example, a suitable machine learning technique using support vector regression is disclosed in Drucker et al. “Support Vector Regression Machines”, Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161. The output of the technique is a continuously varying set of interpolation directions that form part of the reference data 35

and are used in the interpolation.

[89] The machine learning technique, regardless of its type, inherently also derives the form of the feature vectors 34 that is used to derive the reference displacement vector fields F (or other image editing instructions). This is the form of the feature vectors 34 that is derived in step S3.

[90] The description now reverts to the method of FIGURE 3.

[91] In step S4, at least one displacement vector field 37 representing a transformation of an image patch is derived by using the feature vector 34 derived in step S3 to look up the reference data 35. Due to the derivation of the displacement vector field 37 from the reference data 35, the transformation represented thereby corrects the gaze that will be perceived when the stereoscopic pair of images 31 are displayed.

[92] In the case that the feature vector 34 comprises features that are values in a discrete space and the reference displacement vector fields of the reference data 35 comprise a reference displacement vector field associated with each possible value of the feature vector in the discrete space, then the displacement vector field for the image patch is derived by selecting the reference displacement field associated with the actual value of the derived feature vector 34.

[93] In the case that the feature vector 34 comprises features that are discrete values of the local descriptors in a continuous space, then then the displacement vector field for the image patch is derived by interpolating a displacement vector field from the reference displacement vector fields based on the relationship between the actual value of the derived feature vector 34 and the values of the feature vectors associated with the reference displacement vector fields. In the case that the machine learning technique was support vector regression, this may be done using the interpolation directions that form part of the reference data 35.

[94] In step S5, each image 31 of the stereoscopic pair is adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived

[148] **FIGURE 15** depicts data flow in an example implementation of an embodiment in which a pre-trained second machine learning algorithm MLA2 is updated (i.e. trained further) in at or near input data frame rate (i.e. online). In embodiments of this type, a plurality of the digital representations (e.g. images) of the head region may be received as input data ID (e.g. from a sensor system SS as described above) and adjusted (e.g. as described above with reference to FIGURES 11, 12 and 14) to provide output data OD that is subsequently displayed via display DS (as described above with reference to FIGURE 12). The input data ID may comprise a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch. The adjustment may be performed using a pre-trained version of the second machine learning algorithm MLA2, optionally in combination with a geometrical model. The input data ID may comprise digital representations captured at different points in times, such as different frames in a sequence of images obtained by the sensor system SS in a video capture mode. In an embodiment, the method comprises updating the pre-trained second machine learning algorithm MLA2 using first reference data (e.g. editing instructions) generated by the first machine learning algorithm MLA1 using one or more of the received digital representations (from the input data ID). The first reference data may comprise a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region. In the example depicted in FIGURE 15, the first reference data for the updating is generated in a background process and the second machine learning algorithm MLA2 is updated at or near input data frame rate (i.e. online).

[149] The updating thus uses a pre-trained version of the first machine learning algorithm MLA1 (which, as described above, is configured to provide more highly detailed reference data at the expense of higher computational demands, relative to the second machine learning algorithm MLA2) to generate first reference data that is used to update the second machine learning algorithm MLA2. Where the first machine learning algorithm MLA1 is slower than the second machine learning algorithm MLA2 (which will normally be the case where the first machine learning algorithm MLA1 is configured to provide more highly detailed reference data than the second machine learning algorithm MLA2), only a subset of the input

data ID (comprising the received and adjusted digital representations) are used to update the second machine learning algorithm MLA2, thereby allowing the updating process to keep up with the adjusting of the input data ID by the second machine learning algorithm MLA2 to provide the output data OD. Any of various known techniques may be used to perform the updating of the second machine learning algorithm MLA2. For example, the online updating may be performed as described in Amir Saffari, Christian Leistner, Jakob Santner, Martin Godec, and Horst Bischof, "On-line Random Forests," in 3rd IEEE ICCV Workshop on On-line Computer Vision, 2009, or as described in Online Deep Learning: Learning Deep Neural Networks on the Fly, Doyen Sahoo, Quang Pham, Jing Lu, Steven C.H., Hoi School of Information Systems, Singapore Management University, <https://arxiv.org/pdf/1711.03705.pdf>. The second machine learning algorithm MLA2 can therefore be gradually improved during use, as more input data ID are encountered and processed by the system.

Adjusting a Digital representation of a Head Region

TECHNICAL FIELD

[01] This application relates to adjusting a digital representation, such as an image or a three-dimensional geometrical representation, of a head region, particularly a facial and/or neck region. The application relates particularly to adjusting target features of the digital representation of the head region, for example to correct a perceived gaze direction of eyes, or to modify the texture and/or shape of features such as the nose, mouth, chin or neck.

BACKGROUND

[02] In many systems, images of a head, which may comprise single images at different times, or stereoscopic pairs of images or other multi-view images, may be captured in one device and displayed on a different device for viewing by an observer. One non-limitative example is a system for performing teleconferencing between two telecommunications devices. In that case, each device may capture images of the observer of that device and transmit them to the other device over a telecommunications network for display and viewing by the observer of the other device. Digital representations other than images and/or complementary to images may also be captured, for example using depth measurements (e.g. using a time-of-flight camera).

[03] When an image or other digital representation of a head is captured and displayed, the gaze of the head may not be directed at the observer. This may be caused for example by the gaze of the head not being directed at the sensing system (e.g. camera system) used to capture the digital representation (e.g. image), for example because a user whose head is imaged is observing a display in the same device as a camera system and the camera system is offset above (or below) that display. In that case, the gaze in the displayed images will be perceived to be downwards (or upwards). The human visual system has evolved high sensitivity to gaze during social interaction, using cues gained from the relative position of the iris and white sclera of other observers. Errors in the perceived gaze are disconcerting. For example in a system for performing teleconferencing, errors in the perceived gaze can create unnatural interactions between the users.

[04] The present disclosure is concerned with processing techniques (e.g. image processing techniques) for adjusting digital representations (e.g. images) of a head region to

correct the perceived gaze and/or to improve other aspects of a computer-generated display of the head region. The present disclosure is particularly concerned with implementing such processing techniques with minimal demands on computer hardware and/or power such that they provide results at or near input data frame rate or user feedback requirements.

Yaroslav Ganin et al.: "DeepWarp: Photorealistic Image Resynthesis for Gaze Manipulation", pages 1-18, 26 July 2016 discloses methods for generating images of a given face with a redirected gaze. The problem is treated as a specific instance of conditional image generation. A deep architecture is disclosed for handling the task. The deep architecture performs coarse-to-fine warping with an additional intensity correction of individual pixels. All these operations are performed in a feed-forward manner, and the parameters associated with different operations are learned jointly in the end-to-end fashion. After learning, the resulting neural network can synthesize images with manipulated gaze, while the redirection angle can be selected arbitrarily from a certain range and provided as an input to the network.

US 2017/195662 A1 discloses a method in which gaze is corrected by adjusting multi-view images of a head. Image patches containing the left and right eyes of the head are identified and a feature vector is derived from plural local image descriptors of the image patch in at least one image of the multi-view images. A displacement vector field representing a transformation of an image patch is derived, using the derived feature vector to look up reference data comprising reference displacement vector fields associated with possible values of the feature vector produced by machine learning. The multi-view images are adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived displacement vector field.

BRIEF SUMMARY

[05] The scope of the invention is defined in the appended claims.

~~**[05][06]** According to an aspect of the invention first aspect of the present disclosure, there is provided the method of claim 1 a method of generating reference data for adjusting a digital representation of a head region, the method comprising: receiving training data comprising: a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and a set of output patches in one to one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the~~

~~head region; using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.~~

~~[06]~~[07] The described use of two machine learning algorithms allows an improved balance to be achieved between accuracy of the adjustment process and speed of execution. The first machine learning algorithm can be configured to provide highly detailed first reference data, which provides high accuracy. Use of this first reference data directly in a processing technique (e.g. image processing technique) to adjust a digital representation (e.g. image or three-dimensional geometrical representation) of a head region would be relatively expensive in terms of computational resources because of the high level of detail. By arranging instead for the first reference data to be provided to a second machine learning algorithm, which provides, based on the first reference data and the original training data, the reference data (the second reference data) that is to be used for the adjustment of the digital representation of the head region, it is possible to benefit to an extent from the high accuracy of the first machine learning algorithm whilst also providing reference data that is less detailed and thus easier to process efficiently when performing the adjustment of the digital representation of the head region. The quality of the reference data provided by the second machine learning algorithm is found to be significantly improved when the second machine learning algorithm is provided with both the first reference data and the training data in comparison to when the second machine learning algorithm is provided only with the training data.

~~[07]~~[08] ~~In an embodiment, e~~Efficiency is further improved by providing editing instructions (e.g. image editing instructions) in the second reference data in a compressed representation. The use of a compressed representation reduces data storage and bandwidth requirements during use of the editing instructions to perform adjustment of a digital representation of a head region.

~~[08]~~~~[09]~~ According to an alternative second aspect of the invention~~present disclosure~~, there is provided the a device of claim 15~~configured to perform a similar method to the first aspect of the invention~~.

~~[09]~~~~[10]~~ The present disclosure also describes ~~According to a third aspect of the present disclosure, there is provided~~ a method of adjusting a digital representation of a head region, the method comprising: ~~—~~identifying a target patch in the digital representation of the head region, the target patch comprising a target feature of the digital representation of the head region; deriving a feature vector from plural local descriptors of the target patch; using the feature vector to select editing instructions from reference data, the reference data comprising editing instructions for a range of possible values of the feature vector; and applying the selected editing instructions to the target patch to adjust the digital representation of the head region, wherein: the editing instructions in the reference data are provided in a compressed representation.

~~[10]~~~~[11]~~ The present disclosure also describes ~~According to a fourth aspect of the present disclosure, there is provided~~ a device configured to perform the above method~~a similar method of the third aspect of the invention~~.

~~[11]~~ ~~The use of a compressed representation reduces data storage and bandwidth requirements.~~

~~[12]~~ The present disclosure also describes~~According to a fifth aspect of the present disclosure, there is provided~~ a method of training a machine learning algorithm to adjust a digital representation of a head region, comprising: receiving training data comprising: a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and first reference data generated by a pre-trained first machine learning algorithm, the first reference data comprising a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region; updating a pre-trained second machine learning algorithm trained to generate second reference data, where the input for the updating comprises the training data and the generated first reference data, the second reference data comprising editing instructions for adjusting the digital representation of the head region.

~~[13]~~ The present disclosure also describes ~~According to a sixth aspect of the present disclosure, there is provided~~ a method of training a machine learning algorithm to adjust a

digital representation of a head region, the method comprising: receiving training data comprising a set of input digital representations of a head region; training a first machine learning algorithm using the training data to perform an adjustment of a digital representation of a head region; using the trained first machine learning algorithm to generate first reference data, the first reference data comprising an adjusted digital representation of the head region for each of at least a subset of the input digital representations, each adjusted digital representation being obtained by performed the adjustment that the first machine learning algorithm was trained to perform; and training a second machine learning algorithm using at least a subset of the training data used to train the first machine learning algorithm and the first reference data to perform the same adjustment of a digital representation of a head region as the first machine learning algorithm.

BRIEF DESCRIPTION OF THE DRAWINGS

[14] Non-limitative embodiments are illustrated by way of example in the accompanying figures, in which like reference numbers indicate similar parts, and in which:

[15] **FIGURE 1** is a schematic perspective view of a device that captures a stereoscopic pair of images;

[16] **FIGURE 2** is a schematic perspective view of a device that displays the stereoscopic pair of images;

[17] **FIGURE 3** is a flow chart of a method of adjusting a stereoscopic pair of images;

[18] **FIGURE 4** is a diagram illustrating the processing of the stereoscopic pair of images in the method of FIGURE 3;

[19] **FIGURE 5** is a flow chart of a step of extracting an image patch;

[20] **FIGURE 6** and **FIGURE 7** are flow charts of two alternatives for a step of adjusting an image;

[21] **FIGURE 8** is flow chart of a method of generating reference data;

finding the spatial relationship may be applied, for example a calibration method using a reference image, or specification *a priori*.

[41] The camera or cameras 13 face in the same direction as the display 11. Thus, when the source observer 14 is viewing the display 11, then the camera or cameras 13 face the source observer 14 and the captured information, such as depth information, image or images (e.g. stereoscopic pair of images) are digital representations (e.g. images and/or three-dimensional geometrical representations) of the head of the source observer 14. Different cameras in the camera system can have different fields of view.

[42] The camera system 12 may include cameras 13 having different sensing modalities, including but not limited to visible light, infrared, and time-of-flight (depth).

[43] In some embodiments, the main output of the camera system 13 is images 31 which are typically video images output at a video rate. The output of the camera system 13 may also include data representing the spatial relationship between the cameras 13 and the display 11, the nature of the sensing modalities and internal parameters of the cameras 13 (for example focal length, optical axis) which may be used for angular localization, as well as three-dimensional geometrical information, for example from depth measurements.

[44] An example of the method performed on a digital representation of a head region comprising a stereoscopic pair of images 31, for the case of adjustment of eye regions, is as follows. To illustrate this example method, reference is also made to **FIGURE 4** which shows an example of the stereoscopic pair of images 31 at various stages of the method.

[45] In step S1, the stereoscopic pair of images 31 are analyzed to detect the location of the head and in particular the eyes of the source observer 14 within the stereoscopic pair of images 31. This is performed by detecting presence of a head, tracking the head, and localizing the eyes of the head. Step S1 may be performed using a variety of techniques that are known in the art.

[46] One possible technique for detecting the presence of the head is to use Haar feature cascades, for example as disclosed in Viola and Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", CVPR 2001, pp 1-9, ~~which is herein incorporated by reference in its entirety.~~

[47] One possible technique for tracking the head is to use the approach of Active Appearance Models to provide the position of the head of the subject, as well as the location of the eyes, for example as disclosed in Cootes et al., "Active shape models - their training and application", Computer Vision and Image Understanding, 61(1):38-59, Jan. 1995 and in

Cootes et al. "Active appearance models", IEEE Trans. Pattern Analysis and Machine Intelligence, 23(6):681-685, 2001, both of which are herein incorporated by reference in their entireties.

[48] In step S1, typically, a set of individual points ("landmarks") are set to regions of the face, typically the eyes, for example corners of the eye, upper and lower lid locations, etc, thereby localizing the eyes.

[49] In step S2, patches representing portions of a digital representation of the head region, which in this example may be referred to image patches, containing the left and right eyes of the head, respectively, are identified in each image 31 of the stereoscopic pair. FIGURE 4 shows the identified image patches 32 of the right eye in each image 31 (the image patches for the left eye being omitted in FIGURE 4 for clarity).

[50] Step S2 may be performed as shown in **FIGURE 5**, as follows.

[51] In step S2-1, image patches 32 containing the left and right eyes of the head are identified in each image 31 of the stereoscopic pair. This is done by identifying an image patch 39 in each image 31 located around the identified points ("landmarks") corresponding to features of an eye, as shown for example in FIGURE 4.

[52] In step S2-2, the image patches 32 identified in step S2-1 are transformed into a normalized coordinate system, being the same normalized coordinate system as used in the machine learning process which is described further below. The transformation is chosen to align the points ("landmarks") of the eye within the image patch that were identified in step S1, with predetermined locations in the normalized coordinate system. The transformation may include translation, rotation and scaling, to appropriate extents to achieve that alignment. The output of step S2-2 is identified image patches 33 of the right eye in each image in the normalized coordinate system as shown for example in FIGURE 4.

[53] The following steps may be performed separately (a) in respect of the image patches containing the left eyes of the head in each image 31 of the stereoscopic pair, and (b) in respect of the image patches containing the right eyes of the head in each image 31 of the stereoscopic pair (in this example). For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes.

[54] In step S3, a feature vector 34 is derived from plural local descriptors (representing information about a local region in a patch), which in this example may be referred to as local image descriptors, of an image patch 33 in at least one image 31 of the stereoscopic pair (in

this example). Depending on the approach and as described further below, this may be an image patch in a single image 31 of the stereoscopic pair or may be both images 31 of the stereoscopic pair. Thus, the local image descriptors are local image descriptors derived in the normalized coordinate system.

[55] The feature vectors 34 are representations of the image patches 33 that are suitable for use in looking up reference data 35 to be used for adjusting the image patches. The reference data 35 may comprise reference displacement vector fields that represent transformations of the image patch, or other representations of transformations of the image patch, including compressed representations as described below, and are associated with possible values of the feature vector.

[56] The reference data 35 is obtained and analyzed in advance using a machine learning technique. The machine learning technique may derive the form of the feature vectors 34 and associate transformations such as the reference displacement vector fields with the possible values of the feature vector. A specific example of a machine learning technique applied in the case where it is desired to correct gaze using digital representations of a head region comprising images of the head region, will now be described before reverting to the method of FIGURE 3.

[57] The training input to the machine learning technique is two sets of images (or image patches), which may be stereoscopic pairs of images or monoscopic images, as discussed further below. Each set may comprise images of the head of the same group of individuals but captured from cameras in different locations relative to the gaze so that the perceived gaze differs as between them (in the case where gaze is to be corrected).

[58] The first set are input images, being images of each individual with an incorrect gaze where the error is known *a priori*. In particular, the images in the first set may be captured by at least one cameras in a known camera location where the gaze of the individual which is in a different known direction. For example in the case of the source device of FIGURE 1, the camera location may be the location of a camera 13 and while the gaze of the imaged individual is towards the center of the display 11.

[59] The second set are output images, being images of each individual with correct gaze for a predetermined observer location relative to a display location in which the image is to be displayed. In the simplest case, the observer location is a normal viewing position perpendicular to the center of the display location, for example as shown by the hard outline of the destination observer 24 in the case of the destination device 20 of FIGURE 2.

[60] For each image in the two sets, the image is analyzed to detect the location of the head and in particular the eyes using the same technique as used in step S1 described above, and then image patches containing the left and right eyes of the head, respectively, are identified using the same technique as used in step S2 described above. The following steps may then be performed separately (a) in respect of the image patches containing the left eyes of the head in each image, and (b) in respect of the image patches containing the right eyes of the head in each image. For brevity, the following description will refer merely to image patches and eyes without specifying the left or right eye, but noting the same steps are performed for both left and right eyes in this embodiment.

[61] Each image patch is transformed into the same normalized coordinate system as used in step S2 described above. As described above, the transformation is chosen to align points (“landmarks”) of the eye with predetermined locations in the normalized coordinate system. The transformation may include, for example, translation, rotation and/or scaling, to appropriate extents to achieve that alignment.

[62] Thus, the image patches input and output images of each individual are aligned in the normalized coordinate system.

[63] From an input and output image of each individual, there is derived a displacement vector field that represents the transformation of the image patch in the input image required to obtain the image patch of the output image, for example as follows. Defining positions in the image patches by (x,y) , the displacement vector field $\mathbf{F}$ is given by

$$\mathbf{F}=\{u(x, y), v(x, y)\}$$

where u and v define the horizontal and vertical components of the vector at each position (x,y) .

[64] The displacement vector field $\mathbf{F}$ is chosen so that the image patch of the output image $O(x,y)$ is derived from the image patch of the input image $I(x,y)$ as

$$O(x, y) = I(x+u(x, y), y+v(x,y))$$

[65] For image data from more than one camera, the system delivers a displacement vector field for the input image from each camera.

[66] The displacement vector field $\mathbf{F}$ for an input and output image of an individual may be derived using a process in which a trial feature vector $\mathbf{F}'=\{u',v'\}$ is modified to minimize error, optionally in an iterative process, for example in accordance with:

$$\sum |O(x, y) - I(x + u'(x, y), y + v'(x, y))| = \min!$$

[67] By way of non-limitative example, the displacement vector field $\mathbf{F}$ may be derived as

disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675; ~~which is herein incorporated by reference in its entirety~~, wherein the displacement vector field **F** is referred to as a “flow field”.

[68] Another example for editing instructions, which may be referred to as image editing instructions in cases where the digital representation of the head region consists of images, which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by filter field $L = \{k(P, x, y)\}$, which defines a filter kernel for a given location (x, y) . This filter field **L** is chosen so that the image patch of the output image $O(x, y)$ is derived from the image patch of the input image $I(x, y)$ as $O(x, y) = k(P(I, x, y), x, y)$, where $P(I, x, y)$ is a local region around the point (x, y) in the image **I**, and $k(P, x, y)$ operates on the patch **P** with coefficients depending on the position (x, y) . For example, it could be the output of a convolution of the patch with a Gaussian filter with width depending on the position **x** in the image, or a brightness increase of a local pixel depending on the vertical position **y**.

[69] Another example for editing instructions (e.g. image editing instructions) which can be used additionally or as an alternative to the displacement vector field in any of the arrangements disclosed herein, is given by a set of typical image components that make up the edited image area, e.g. texture showing beard stubbles. These are then blended with a factor depending on the image coordinates and local image content (i.e. a texture blending field). Other transformation fields can be used, such as a brightness adjustment field.

[70] A machine learning technique is used to obtain a map from the displacement vector field **F** (or other editing instructions such as image editing instructions) of each individual to respective feature vectors derived from plural local (e.g. image) descriptors of a target patch of an input image.

[71] The local descriptors (e.g. local image descriptors) capture relevant information of a local part of a patch (e.g. image patch) of an input digital representation of the head region (e.g. an input image) and the set of local descriptors usually form a continuous vectorial output.

[72] The local descriptors input into the machine learning process are of types expected to provide discrimination between different individuals, although the specific image descriptors are selected and optimized by the machine learning process itself. In general, the local descriptors may be of any suitable type, some non-limitative examples which may be applied

in any combination being as follows.

[73] The local descriptors may include values of individual pixels or a linear combination thereof. Such a linear combination may be, for example, a difference between the pixels at two points, a kernel derived within a mask at an arbitrary location, or a difference between two kernels at different locations.

[74] The local descriptors may include distances of a pixel location from the position of an eye point (“landmark”).

[75] The local descriptors may include SIFT features (Scale-invariant feature transform features), for example as disclosed in Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision 60 (2), pp 91–110, ~~which is herein incorporated by reference in its entirety.~~

[76] The local descriptors may include HOG features (Histogram of Oriented Gradients features) , for example as disclosed in Dalal et al. “Histograms of Oriented Gradients for Human Detection”, Computer Vision and Pattern Recognition, 2005, pp. 886-893, ~~which is herein incorporated by reference in its entirety.~~

[77] The local descriptors may include “low level representations” from pre-classification stages in deep learning neural networks, for example as disclosed in Yang and Ramanan, “Multi-scale recognition with DAG-CNNs”, ICCV 2015, ~~which is herein incorporated by reference in its entirety.~~ In a classifying deep learning neural network with multiple layers applied to an input digital representation (e.g. image), for example, such low level features could be taken from a layer before the final classification layer of the network.

[78] The derivation of the feature vector from plural local descriptors depends on the type of machine learning applied.

[79] In a first type of machine learning technique, the feature vector may comprise features that are values derived from the local descriptors (e.g. local image descriptors) in a discrete space, being binary values or values discretized into more than two possible values. In this case, the machine learning technique associates a reference displacement vector field $\mathbf{F}$ derived from the training input with each possible value of the feature vector in the discrete space, so the reference data 35 may provide similar functionality to a look-up table, with the machine learning generating a machine learning parameter set that can be used to generate corresponding editing instructions. This allows a reference displacement vector field $\mathbf{F}$ to be simply selected from the reference data 35 on the basis of the feature vector 34 derived in step S3, as described below.

[80] In the case that the feature vector comprises features that are binary values derived from the local descriptors, the feature vector has a binary representation. Such binary values may be derived in various ways from the values of descriptors, for example by comparing the value of a descriptor with a threshold, comparing the value of two descriptors, or by comparing the distance of a pixel location from the position of an eye point (“landmark”).

[81] Alternatively, the feature vector may comprise features that are discretized values of the local descriptors. In this case, more than two discrete values of each feature are possible.

[82] Any suitable machine learning technique may be applied, for example using a decision tree, a decision forest, a decision fern or an ensemble or combination thereof, or a neural network.

[83] By way of example, a suitable machine learning technique using a feature vector comprising features that are binary values derived by comparing a set of individual pixels or a linear combination thereof against a threshold, is disclosed in Ozuysal et al. “Fast Keypoint Recognition in Ten Lines of Code”, Computer Vision and Pattern Recognition, 2007, pp. 1-8, ~~which is herein incorporated by reference in its entirety.~~

[84] By way of further example, a suitable machine learning technique using a distance of a pixel location with the position of an eye landmark is disclosed in Kononenko et al., “Learning To Look Up: Realtime Monocular Gaze Correction Using Machine Learning”, Computer Vision and Pattern Recognition, 2015, pp. 4667-4675, ~~which is herein incorporated by reference in its entirety.~~

[85] By way of further example, a suitable machine learning technique using a random decision forest is disclosed in Ho, “Random Decision Forests”, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995, pp. 278-282, ~~which is herein incorporated by reference in its entirety.~~

[86] In a second type of machine learning technique, the feature vector may comprise features that are discrete values of the local descriptors (e.g. local image descriptors) in a continuous space. In this case, the machine learning technique associates a reference displacement vector field **F** (in this example, but other editing instructions could be used) derived from the training input with possible discrete values of the feature vector in the continuous space. This allows a displacement vector field **F** to be derived from the reference data 35 by interpolation from the reference displacement vector fields based on the relationship between the feature vector 34 derived in step S3 and the values of the feature vector associated with the reference displacement vector fields.

[87] Any suitable machine learning technique may be applied, for example using support vector regression.

[88] By way of example, a suitable machine learning technique using support vector regression is disclosed in Drucker et al. "Support Vector Regression Machines", Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, ~~which is herein incorporated by reference in its entirety~~. The output of the technique is a continuously varying set of interpolation directions that form part of the reference data 35 and are used in the interpolation.

[89] The machine learning technique, regardless of its type, inherently also derives the form of the feature vectors 34 that is used to derive the reference displacement vector fields $\mathbf{F}$ (or other image editing instructions). This is the form of the feature vectors 34 that is derived in step S3.

[90] The description now reverts to the method of FIGURE 3.

[91] In step S4, at least one displacement vector field 37 representing a transformation of an image patch is derived by using the feature vector 34 derived in step S3 to look up the reference data 35. Due to the derivation of the displacement vector field 37 from the reference data 35, the transformation represented thereby corrects the gaze that will be perceived when the stereoscopic pair of images 31 are displayed.

[92] In the case that the feature vector 34 comprises features that are values in a discrete space and the reference displacement vector fields of the reference data 35 comprise a reference displacement vector field associated with each possible value of the feature vector in the discrete space, then the displacement vector field for the image patch is derived by selecting the reference displacement field associated with the actual value of the derived feature vector 34.

[93] In the case that the feature vector 34 comprises features that are discrete values of the local descriptors in a continuous space, then then the displacement vector field for the image patch is derived by interpolating a displacement vector field from the reference displacement vector fields based on the relationship between the actual value of the derived feature vector 34 and the values of the feature vectors associated with the reference displacement vector fields. In the case that the machine learning technique was support vector regression, this may be done using the interpolation directions that form part of the reference data 35.

[94] In step S5, each image 31 of the stereoscopic pair is adjusted by transforming the image patches containing the left and right eyes of the head in accordance with the derived

[148] **FIGURE 15** depicts data flow in an example implementation of an embodiment in which a pre-trained second machine learning algorithm MLA2 is updated (i.e. trained further) in at or near input data frame rate (i.e. online). In embodiments of this type, a plurality of the digital representations (e.g. images) of the head region may be received as input data ID (e.g. from a sensor system SS as described above) and adjusted (e.g. as described above with reference to FIGURES 11, 12 and 14) to provide output data OD that is subsequently displayed via display DS (as described above with reference to FIGURE 12). The input data ID may comprise a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch. The adjustment may be performed using a pre-trained version of the second machine learning algorithm MLA2, optionally in combination with a geometrical model. The input data ID may comprise digital representations captured at different points in times, such as different frames in a sequence of images obtained by the sensor system SS in a video capture mode. In an embodiment, the method comprises updating the pre-trained second machine learning algorithm MLA2 using first reference data (e.g. editing instructions) generated by the first machine learning algorithm MLA1 using one or more of the received digital representations (from the input data ID). The first reference data may comprise a set of editing instructions in one-to-one correspondence with the input patches, each editing instruction being for adjusting the digital representation of the head region. In the example depicted in FIGURE 15, the first reference data for the updating is generated in a background process and the second machine learning algorithm MLA2 is updated at or near input data frame rate (i.e. online).

[149]—The updating thus uses a pre-trained version of the first machine learning algorithm MLA1 (which, as described above, is configured to provide more highly detailed reference data at the expense of higher computational demands, relative to the second machine learning algorithm MLA2) to generate first reference data that is used to update the second machine learning algorithm MLA2. Where the first machine learning algorithm MLA1 is slower than the second machine learning algorithm MLA2 (which will normally be the case where the first machine learning algorithm MLA1 is configured to provide more highly detailed reference data than the second machine learning algorithm MLA2), only a subset of the input

data ID (comprising the received and adjusted digital representations) are used to update the second machine learning algorithm MLA2, thereby allowing the updating process to keep up with the adjusting of the input data ID by the second machine learning algorithm MLA2 to provide the output data OD. Any of various known techniques may be used to perform the updating of the second machine learning algorithm MLA2. For example, the online updating may be performed as described in Amir Saffari, Christian Leistner, Jakob Santner, Martin Godec, and Horst Bischof, "On-line Random Forests," in 3rd IEEE ICCV Workshop on On-line Computer Vision, 2009, ~~which is herein incorporated by reference in its entirety~~, or as described in Online Deep Learning: Learning Deep Neural Networks on the Fly, Doyen Sahoo, Quang Pham, Jing Lu, Steven C.H., Hoi School of Information Systems, Singapore Management University, <https://arxiv.org/pdf/1711.03705.pdf>, ~~which is herein incorporated by reference in its entirety~~. The second machine learning algorithm MLA2 can therefore be gradually improved during use, as more input data ID are encountered and processed by the system.

~~[150] Additional embodiments of the disclosure are described in the following numbered clauses:~~

- ~~1. A method of generating reference data for adjusting an image of a head region, the method comprising:~~
 - ~~receiving training data comprising:~~
 - ~~a set of input image patches, each input image patch comprising a target feature of an image of a head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and~~
 - ~~a set of output image patches in one-to-one correspondence with the input image patches, each output image patch comprising the target feature of the image of the head region after adjustment of the image of the head region;~~
 - ~~using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising image editing instructions for adjusting the image of the head region for a range of possible images of the head region; and~~
 - ~~using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm, the second reference data comprising image~~

editing instructions for adjusting the image of the head region for a range of possible images of the head region.

2. The method of clause 1, wherein:

the first reference data comprise first image editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting image editing instructions for a particular input image patch from the first image editing instructions based on the configuration of the target feature of the input image patch; and

the second reference data comprise second image editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting image editing instructions for a particular input image patch from the second image editing instructions based on the configuration of the target feature of the input image patch.

3. The method of clause 2, wherein the configuration of the target feature of each input image patch is represented by a feature vector derived from plural local image descriptors of the input image patch, and the first and second selection instructions define how the feature vector is used to select image editing instructions for the input image patch.

4. The method of clause 2 or 3, wherein the image editing instructions comprise a displacement vector field defining how the input image patch is to be transformed.

5. The method of clause 2 or 3, wherein the image editing instructions comprise a filter field, a brightness adjustment field, or a texture blending field.

6. The method of any of clauses 2-5, wherein:

a first image editing algorithm is used by the first machine learning algorithm to define how the first image editing instructions are to be applied to an input image patch to derive an output image patch; and

a second image editing algorithm is used by the second machine learning algorithm to define how the second image editing instructions are to be applied to an input image patch to derive an output image patch.

7. The method of clause 6, wherein:

the second image editing instructions are principle component analysis components of a principle component analysis of the first image editing instructions; and

~~the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse principle component analysis transform.~~

~~8. The method of clause 6, wherein:~~

~~the second image editing instructions are wavelet components of the first image editing instructions; and~~

~~the second image editing algorithm is configured to transform the second image editing instructions into the first image editing instructions by inverse wavelet transform.~~

~~9. The method of any of clauses 2-8, wherein the first selection instructions for the first reference data are able to select between a larger number of alternative image editing instructions than the second selection instructions for the second reference data.~~

~~10. The method of any preceding clause, wherein the first machine learning algorithm is of a different machine learning algorithm type than the second machine learning algorithm.~~

~~11. The method of any preceding clause, wherein the first machine learning algorithm comprises one or more of the following: a neural network; a support vector machine; a generative adversarial network, GAN.~~

~~12. The method of any preceding clause, wherein the second machine learning algorithm comprises one or more of the following: a regression forest; regression ferns; cluster centres, a lookup table, separable filter banks.~~

~~13. The method of any of clauses 1-10, wherein the first machine learning algorithm comprises a first neural network and the second machine learning algorithm comprises a second neural network, wherein the second neural network comprises fewer layers and/or smaller convolution fields than the first neural network.~~

~~14. The method of any preceding clause, wherein the target feature comprises one or more of the following: an eye region comprising at least part of an eye, a nose region comprising at least part of a nose, a mouth region comprising at least part of a mouth, a chin region comprising at least part of a chin, a neck region comprising at least part of a neck, and a hair region comprising hair.~~

~~15. The method of any of clauses 1-13, wherein:~~

~~the target feature comprises an eye region comprising at least part of an eye and the adjustment of the image of the head region comprises adjusting a gaze direction;~~

~~the target feature comprises a nose region comprising at least part of a nose and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the nose;~~

~~the target feature comprises a chin region comprising at least part of a chin and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the chin;~~

~~the target feature comprises a neck region comprising at least part of a neck and the adjustment of the image of the head region comprises adjusting a shape and/or texture of the neck; and/or~~

~~the target feature comprises a hair region comprising hair and the adjustment of the image of the head region comprises adjusting a color of the hair.~~

~~16. The method of any preceding clause, wherein the second image editing instructions in the second reference data are provided in a compressed representation.~~

~~17. The method of clause 16, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.~~

~~18. A method of adjusting an image of a head region, the method comprising:~~

~~identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;~~

~~deriving a feature vector from plural local image descriptors of the image patch;~~

~~using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and~~

~~applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:~~

~~the reference data comprises the second reference data generated by the method of any of clauses 1-17.~~

~~19. A method of adjusting an image of a head region, the method comprising:~~

~~————— identifying an image patch in the image of the head region, the image patch comprising a target feature of the image of the head region;~~

~~————— deriving a feature vector from plural local image descriptors of the image patch;~~

~~————— using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and~~

~~————— applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:~~

~~————— the image editing instructions in the reference data are provided in a compressed representation.~~

~~——— 20. — The method of clause 19, wherein the compressed representation comprises one or more of the following: a principle component analysis representation; a wavelet representation; Fourier and/or discrete cosine transform components; cluster centres.~~

~~——— 21. — A computer program capable of execution by a processor and arranged on execution to cause the processor to perform a method according to any of the preceding clauses.~~

~~——— 22. — A computer readable storage medium storing a computer program according to clause 21.~~

~~——— 23. — A device for generating reference data for adjusting an image of a head region, the device comprising a data processing unit arranged to:~~

~~————— receive training data comprising:~~

~~————— a set of input image patches, each input image patch comprising information about a target feature of an image of the head region prior to adjustment of the image of the head region, wherein the target feature is the same for each input image patch; and~~

~~————— a set of output image patches in one to one correspondence with the input image patches, each output image patch comprising the target portion of the image of the head region after adjustment of the image of the head region;~~

~~————— use a first machine learning algorithm to generate first reference data using the training data; and~~

~~_____ use a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data output by the first machine learning algorithm.~~

~~_____ 24. A device for adjusting an image of a head region, the device comprising an image processor arranged to process the image of the head region by:~~

~~_____ identifying an image patch in the image of the head region, the image patch comprising information about a target feature of the image of the head region;~~

~~_____ deriving a feature vector from plural local image descriptors of the image patch;~~

~~_____ using the feature vector to select image editing instructions from reference data, the reference data comprising image editing instructions for a range of possible values of the feature vector; and~~

~~_____ applying the selected image editing instructions to the image patch to adjust the image of the head region, wherein:~~

~~_____ the image editing instructions in the reference data are provided in a compressed representation.~~

~~_____ 25. The device of clause 24, further comprising a telecommunications interface arranged to transmit the adjusted image over a telecommunications network to a destination device for display thereon.~~

~~[151] While various embodiments in accordance with the principles disclosed herein have been described above, it should be understood that they have been presented by way of example only, and not limitation. Thus, the breadth and scope of this disclosure should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with any claims and their equivalents issuing from this disclosure. Furthermore, the above advantages and features are provided in described embodiments, but shall not limit the application of such issued claims to processes and structures accomplishing any or all of the above advantages.~~

~~[152][149] Additionally, the section headings herein are provided for consistency with the suggestions under 37 CFR 1.77 or otherwise to provide organizational cues. These headings shall not limit or characterize the embodiment(s) set out in any claims that may~~

~~issue from this disclosure. Specifically and by way of example, although the headings refer to a "Technical Field," the claims should not be limited by the language chosen under this heading to describe the so-called field. Further, a description of a technology in the "Background" is not to be construed as an admission that certain technology is prior art to any embodiment(s) in this disclosure. Neither is the "Summary" to be considered as a characterization of the embodiment(s) set forth in issued claims. Furthermore, any reference in this disclosure to "invention" in the singular should not be used to argue that there is only a single point of novelty in this disclosure. Multiple embodiments may be set forth according to the limitations of the multiple claims issuing from this disclosure, and such claims accordingly define the embodiment(s), and their equivalents, that are protected thereby. In all instances, the scope of such claims shall be considered on their own merits in light of this disclosure, but should not be constrained by the headings set forth herein.~~



Acknowledgement of receipt

We hereby acknowledge receipt of the following subsequently filed document(s):

Submission number	10691621	
Application number	EP18844651.2	
Date of receipt	22 February 2022	
Receiving Office	European Patent Office, The Hague	
Your reference	N411553EP DQF	
Applicant	All applicants as on file	
Documents submitted	package-data.xml epf1038.pdf (1 p.) DESC-1.pdfAmended description clean pages Feb 22 N411553EP.pdf (14 p.) CLMS-1.PDFAmended claims clean Feb 2022 - N411553EP.PDF (5 p.)	ep-sfd-request.xml EXRE3-1.PDFResponse to examination report - N411553EP DQF.PDF (2 p.) DESC-HWA-1.pdfAmended description tracked pages Feb 22 N411553EP.pdf (21 p.) CLMS-HWA-1.PDFAmended claims tracked Feb 2022 - N411553EP.PDF (5 p.)
Submitted by	CN=Dominic Forsythe 23485	
Method of submission	Online	
Date and time receipt generated	22 February 2022, 15:19 (CET)	
Message Digest	A0:72:87:B8:FD:1B:29:04:C5:0F:E1:DD:5D:F9:19:AB:5B:46:98:D9	

/European Patent Office/



European Patent Office
80298 MUNICH
GERMANY

Questions about this communication?
Contact Customer Services at www.epo.org/contact



J A Kemp LLP
80 Turnmill Street
London EC1M 5QU
ROYAUME UNI

Date	01.02.2023
------	------------

Reference N411553EP	Application No./Patent No. 18844651.2 - 1207 / 3665553
Applicant/Proprietor RealD Spark, LLC	

Communication

concerning the registration of amendments relating to

- ☐ a transfer (R. 22 and 85 EPC)
- ☒ entries pertaining to the applicant / the proprietor (R. 143(1)(f) EPC)

As requested, the entries pertaining to the applicant of the above-mentioned European patent application / to the proprietor of the above-mentioned European patent have been amended to the following:

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV
MC MK MT NL NO PL PT RO RS SE SI SK SM TR
RealD Spark, LLC
1930 Central Ave.
Suite A-2
Boulder, Colorado 80301/US

The registration of the changes has taken effect on 24.01.23.

In the case of a published application / a patent, the change will be recorded in the European Patent Register and published in the European Patent Bulletin (Section I.12 / II.12).

Important note for EPO deposit account holders

If you would like to have fee refunds credited to a deposit account held with the EPO, you should ensure that you provide clear and up-to-date refund instructions to this effect, using one of the permitted electronic means of filing (see Arrangements for deposit accounts (ADA)).

Important note for users of the automatic debiting procedure

If a transfer is registered, any existing automatic debit order will cease to be effective only from the date of its express revocation (see Annex A.1 to the ADA).

Receiving Section / For the Examining Division / For the Opposition Division / For the Legal Division *)



*)See note.

Note

This communication is issued by/for the department with whom responsibility lies. The Legal Division is responsible for the registration of transfers, changes of name (Articles 71, 72 and 74 EPC and Rules 22 and 85 EPC) as well as for the rectification of the designation of the inventor (Rule 21 EPC) (see Decision of the President of the EPO, OJ EPO 2013, 600). In all other cases, the Receiving Section, the Examining Division or the Opposition Division is responsible, depending on the stage in proceedings.



Application No.: 18844651.2

Substantive examination has started on 15.06.23 (see Notice from the European Patent Office dated 29.01.2013, OJ EPO 2013, 153).

For the Examining Division