UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT TRIAL AND APPEAL BOARD

MICROSOFT CORPORATION
Petitioners

v.

EYESMATCH LTD.
Patent Owner

Case No. IPR2025-00296
U.S. Patent No. 8,982,110

Declaration of Dr. Lina Karam, Ph.D.

TABLE OF CONTENTS

I. EXPERIENCE AND QUALIFICATIONS 11

II. TASK SUMMARY AND MATERIALS REVIEWED 20

III. LIST OF EXHIBITS 20

IV. SUMMARY OF CONCLUSIONS AND GROUNDS FOR CHALLENGE 23

V. LEGAL STANDARDS 23

A. Level of Ordinary Skill in the Art 23

B. Legal Standard for Claim Construction 24

C. Anticipation 26

D. Obviousness 27

VI. TECHNOLOGY BACKGROUND 30

A. Exemplary References Corroborating General Engineering Knowledge 31

a. Hartley Text (Ex. 1010) 31

b. Foley Text (Ex. 1027) 31

c. Watt Text Excerpts (Ex. 1028) 31

B. Basic Image Transformations 32

a. Coordinate Representation of 2D and 3D points, World and Camera Coordinate Systems, Camera Image Plane, Camera Projection Matrix and Projective Transformation, Homography Transformation 32

b. Image Flipping 37

c. Synthesis of Virtual Camera Views from Actual Camera Views 38

d. Geometric Transformation 40

e. Image Rotation 42

f. Image Resizing and Field of View Transformation 44

VII. OVERVIEW OF THE CHALLENGED PATENT 47

A. The Subject Matter of the Challenged Patent 48

B. The Challenged Claims of the ’110 Patent 57

a. Claim 1 57
b. Claim 2 58
c. Claim 3 58
d. Claim 4 58
e. Claim 5 59
f. Claim 6 59
g. Claim 7 59
h. Claim 8 59
i. Claim 9 59
j. Claim 10 60
k. Claim 11 60
l. Claim 12 60
m. Claim 13 60
n. Claim 14 60
o. Claim 15 60
p. Claim 16 61
q. Claim 17 61
r. Claim 18 61
C. Prosecution History of the ’110 Patent 61
VIII. PRIORITY DATE AND LEVEL OF ORDINARY SKILL IN THE ART 63
IX. CLAIM CONSTRUCTION 65
X. OVERVIEW OF THE PRIOR ART 65
A. Haan (Ex. 1005) 65
a. Status of Haan as Prior Art 65
b. Overview of the Disclosure of Haan 66
B. Francois (Ex. 1009) 70
a. Status of Francois as Prior Art 70
b. Overview of the Disclosure of Francois 71
C. Rosenberg (Ex. 1006) 74

a. Status of Rosenberg as Prior Art 74
b. Overview of the Disclosure of Rosenberg 75
D. Park (Ex. 1011) 79
a. Status of Park as Prior Art 79
b. Overview of the Disclosure of Park 80
E. Geisner (Ex. 1007) 84
a. Status of Geisner as Prior Art 84
b. Overview of the Disclosure of Geisner 84
F. Lizee (Ex. 1008) 89
a. Status of Lizee as Prior Art 89
b. Overview of the Disclosure of Lizee 89
XI. GROUND 1: CLAIMS 1-6, 10-12, and 16-17 ARE OBVIOUS OVER HAAN IN VIEW OF FRANCOIS 92
A. Motivation to combine Haan and Francois 92
B. Claim 1: 104
a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising: 104
b. 1[a]: operating the camera to generate a stream of images of the user; 106
c. 1[b]: operating the processor to detect the presence of the user in the stream of images; 112
d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view 115
e. 1[d]: displaying the modified images on the monitor; 125
f. 1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and 127
g. 1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies

parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.....135

C. Claim **2.** The method of claim 1, wherein the mapping transformation comprises a combination of at least two of the following transformations: vertical flips about a vertical axis, a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image**.**.....138

D. Claim 3. The method of claim 1, wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images. 140

E. Claim 4. The method of claim 3, wherein the body part comprises one of legs, face, eyes, and hands. 144

F. Claim 5. The method of claim 1, further comprising receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input.....145

G. Claim 6. The method of claim 5, wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images. 146

H. Claim 10. The method of claim 1, further comprising operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images. 147

I. Claim 11. The method of claim 1, further comprising operating the processor to transmit the modified image over a network.....148

J. Claim 12. The method of claim 1, further comprising displaying the modified image on a remote screen for video telepresence.....151

K. Claim 16. The method of claim 1, wherein the processor operates on the images from the camera to generate background segmentation, and replaces segments on the images by replacement segments. 152

L. Claim 17. The method of claim 16, wherein the replacement segments comprise virtual background or background that was captured in a different setting.....155

XII. GROUND 2: CLAIMS 1, 2, 5, 6 AND 10 ARE OBVIOUS OVER ROSENBERG IN VIEW OF PARK 157

A. Motivation to Combine Rosenberg with Park 157

B. Claim 1: 164

a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising 164

b. 1[a]: operating the camera to generate a stream of images of the user; 165

c. 1[b]: operating the processor to detect the presence of the user in the stream of images; 166

d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view; and....167

e. 1[d]: displaying the modified images on the monitor; 171

f. 1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and 173

g. 1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.....176

C. Claim 2: The method of claim 1, wherein the mapping transformation comprises a combination of at least two of the following transformations: vertical flips about a vertical axis,a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image. 182

D. Claims 3-4. Claim 3. The method of claim 1, wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images. Claim 4. The method of claim 3,wherein the body part comprises one of legs, face, eyes, and hands.....187

E. Claim 5: The method of claim 1, further comprising receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input.....189

F. Claim 6: The method of claim 5, wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images. 190

G. Claim 10: The method of claim 1, further comprising operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images. 192

H. Claims 11-12. Claim 11. The method of claim 1, further comprising operating the processor to transmit the modified image over a network. Claim 12. The method of claim 1, further comprising displaying the modified image on a remote screen for video teleprsence. 192

XIII. GROUND 3: CLAIMS 1-18 ARE OBVIOUS IN VIEW OF GEISNER IN VIEW OF LIZEE 194

A. Motivation to Combine Geisner with Lizee.....194

B. Claim 1: 200

a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising: 200

b. 1[a]: operating the camera to generate a stream of images of the user; 202

c. 1[b]: operating the processor to detect the presence of the user in the stream of images; 205

d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view.....209

e. 1[d]: displaying the modified images on the monitor;.....219

f. 1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and.....220

g. 1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.....227

C. Claim 2. The method of claim 1, wherein the mapping transformation comprises a combination of at least two of the following transformations: vertical flips about a vertical axis,a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image. 230

D. Claim 3. The method of claim 1, wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images. 231

E. Claim 4. The method of claim 3, wherein the body part comprises one of legs, face, eyes, and hands. 232

F. Claim 5. The method of claim 1, further comprising receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input.....233

G. Claim 6. The method of claim 5, wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images. 234

H. Claim 7: The method of claim 1, further comprising operating the processor to generate an avatar having body characteristics corresponding to body characteristics of a user appearing in the images. 236

I. Claim 8: The method of claim 7, further comprising operating the processor to display on the monitor the modified images and an image of the avatar. 238

J. Claim 9: The method of claim 7, further comprising operating the processor to augment the modified images by overlaying virtual objects on the avatar. 238

K. Claim 10. The method of claim 1, further comprising operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images. 239

L. Claim 11. The method of claim 1, further comprising operating the processor to transmit the modified image over a network. 240

M. Claim 12. The method of claim 1, further comprising displaying the modified image on a remote screen for video telepresence. 242

N. Claim 13. The method of claim 1, further comprising providing data from a depth sensor to the processor, and operating the processor to measure a texture of the image from the data. 242

O. Claim 14. The method of claim 13, further comprising operating the processor to generate an avatar using the data from the depth sensor.....243

P. Claim 15. The method of claim 13, wherein the processor generate the avatar by analyzing registration pointers on the images of the camera. 245

Q. Claim 16. The method of claim 1, wherein the processor operates on the images from the camera to generate background segmentation, and replaces segments on the images by replacement segments. 246

R. Claim 17. The method of claim 16, wherein the replacement segments comprise virtual background or background that was captured in a different setting.....249

S. Claim 18: The method of claim 1, wherein the processor resides in the remote location, and the method proceeds by transmitting the images from the camera over a network to the processor, and transmitting the modified images back from the processor to themonitor over the network. 252

XIV. GROUND 4: CLAIMS 7-9, 13-15, AND 18 ARE OBVIOUS OVER HAAN IN VIEW OF FRANCOIS AND GEISNER.....255

XV. GROUND 5: CLAIMS 3-4, 7-9, AND 11-18 ARE OBVIOUS OVER ROSENBERG IN VIEW OF PARK AND GEISNER.....258

A. Claims 3-4 261

B. Claims 7-9 and 13-15 264

C. Claims 11-12, 18 265

D. Claims 16-17 268

XVI. SECONDARY CONSIDERATIONS 268

XVII. CONCLUSION 270

INTRODUCTION

1. I, Dr. Lina Karam, Ph.D., submit this declaration to state my opinions on the matter described below.

2. I have been retained as an independent expert in this proceeding before the United States Patent and Trademark Office. Although I am being compensated at my usual and customary rate, no part of my compensation depends on the outcome of this proceeding, and I have no other interest in this proceeding.

3. I understand that this proceeding involves U.S. Patent No. 8,982,110 (the "'110 Patent" or "Challenged Patent"), and I have been asked to provide my opinions as to the unpatentability of the claims of the '110 Patent. I understand that the application for the '110 Patent was filed on April 15, 2014, and claims priority to U.S. Provisional Application No. 61/738,957, filed on December 18, 2012; U.S. Provisional Application No. 61/862,025, filed on Aug. 4, 2013; and U.S. Provisional Application No. 61/892,368, filed on Oct. 17, 2013. I also understand that the '110 Patent is a continuation-in-part of U.S. Application No. 13/088,369, filed on April 17, 2011, which is a continuation of U.S. Application No. 11/817,411, which was a National Phase Application of PCT International Application No. PCT/IL06/000281 having an International Filing Date of March 1, 2006 and which issued as U.S. Patent No. 7,948,481 on May 24, 2011, which claims priority to U.S. Provisional

Application No. 60/656,884 and U.S. Provisional Application No. 60/656,885, both filed on March 1, 2005.

4. I have been asked to consider the unpatentability of the claims of the ’110 Patent based on certain prior art references. I have also been asked to consider the state of the art and the prior art available as of December 18, 2012. Based on the prior art discussed in this declaration, it is my opinion that claims 1-18 of the ’110 Patent are unpatentable for the reasons provided below.

I. EXPERIENCE AND QUALIFICATIONS

5. In formulating my opinions, I have relied upon my knowledge, training, and experience in the relevant art. I believe that I am well qualified to serve as a technical expert in this matter based upon my educational and work experience in image and video processing, computer vision, and multidimensional signal processing, including but not limited to transformations, 3D data processing and visualization, camera calibration and modeling, single-view and multi-view geometry, appearance modification and adjustments, as well as augmented reality.

6. My qualifications are stated more fully in my curriculum vitae, which has been provided as Exhibit 1004. My CV also lists publications on which I am a named author and patents on which I am a named inventor. Here, I provide a brief summary of my qualifications. As set forth in my curriculum vitae:

7. I received a Bachelor of Engineering (B.E.) from the American University of Beirut in Computer and Communications Engineering in 1989, and a M.S. and a Ph.D. from the Georgia Institute of Technology in Electrical Engineering in 1992 and 1995, respectively.

8. I am currently an Emerita Professor in the School of Electrical, Computer, and Energy Engineering at Arizona State University (ASU) and the Director of the R&D Image, Video, and Usability (IVU) Laboratory at ASU. Prior to becoming Emerita Professor, I was a tenured Full Professor, Computer Engineering Program Chair and Computer Engineering Director for Industry Engagement at ASU. From 2020 to 2022, I served as Dean of the School of Engineering at the Lebanese American University. From 2019 to 2021, I served as the Editor-In-Chief of the high-impact IEEE Journal of Special Topics in Signal Processing. I am also the President of PICARIS, LLC, a consulting company that provides expertise on media processing, compression, computer vision, and machine learning. I am also Chief Technical Advisor at AIAEC, LLC.

9. My research includes topics such as Signal, Image, and Video Processing, Compression, and Transmission; Computer Vision; Machine Learning; Perceptual-based Processing; Visual Attention Models; Automated Quality Assessment and Monitoring; Multidimensional Signal Processing; and Digital Filter Design. I have various awards in recognition of my research, including U.S. National

Science Foundation CAREER Award, Intel Outstanding Researcher Award, U.S. National Aeronautics and Space Administration (NASA) Technical Innovation Award, the IEEE Signal Processing Society (SPS) Best Journal Paper Award, IEEE Phoenix Section Outstanding Faculty Award, and IEEE Region 6 Award. I am featured in the 8th and 9th editions of Research.com (2022 and 2023) among World Top Computer Science Scientists.

10. In addition, I am an Institute of Electrical and Electronics Engineers (IEEE) Fellow, which is the highest grade level conferred by IEEE each year to no more than one-tenth of 1% of all IEEE voting members, for my contributions in the areas of image and video processing, visual media compression and transmission, and digital filtering. In addition to serving as Editor-In-Chief of a prestigious IEEE journal, I have also served on several IEEE Boards and chaired multiple IEEE committees. For example, I served on the IEEE Signal Processing Society (SPS) Board of Governors (2016-2018), IEEE SPS Publications Board (2019-2021), IEEE SPS Award Board (2019-2020), IEEE SPS Conference Board (2003-2005 & 2017-2018), IEEE Publication Services and Products Board (PSPB) Strategic Planning Committee (2014-2016), IEEE Technical Activities Board (TAB)/PSPB Product and Services Committee, IEEE Educational Activities Board (EAB) Faculty Resources Committee, and IEEE TechRxiv Advisory and Editorial Boards. I am also an expert delegate of the ISO/IEC JTC1/SC29 Committee (Coding of audio, picture,

multimedia, and hypermedia information) and participating in JPEG/MPEG standardization activities.

11. As part of my various roles at ASU, PICARIS, AIAEC, ISO and IEEE, I am actively involved in industry. For example, as an intern I contributed to the development of image and video processing and compression at AT&T Bell Labs (Murray Hill), and multi-dimensional data processing and visualization at Schlumberger. Later in my career, I collaborated in the research and development of computer vision, machine learning, image/video processing, compression, and transmission projects with various industry leaders, including Intel, Qualcomm, Google, NTT, Motorola, Freescale, NXP, General Dynamics, and NASA. More recently, I took on the role of Chief Technical Advisor at AIAEC, LLC, supervising work and product development in the field of AI and machine learning.

12. In addition, I directed the development of image/video processing, compression, computer vision and machine learning technologies. The developed image/video processing, computer vision and machine learning systems for automated defect detection, 3D characterization, camera calibration and distortion correction, gender/age prediction, and autonomous vehicles were adopted at Intel. The developed image-based non-wet solder joints detection system was granted a Divisional Recognition Award by Intel. More details can be found in Said, Bennett, Karam, and Pettinato, “Automated Detection and Classification of Non-Wet Solder

Joints," IEEE Transactions on Automation Science and Engineering, Jan 2011. The developed image-based automated void detection system helped in enabling two industry standards, JEDEC JC 14-1 void guideline and IPC-7095C. More details can be found in Said, Bennett, Karam, Siah, Goodman, and Pettinato, "Automated Void Detection in Solder Balls in the Presence of Vias and Other Artifacts," IEEE Transactions on Electronics Packaging Manufacturing, November 2012. More details about the developed 3D characterization system including image-based solder ball height and warpage measurements can be found in Li, Bennett, Karam, and Pettinato, "Stereo Vision Based Automated Solder Ball Height and Substrate Coplanarity Inspection," IEEE Transactions on Automation Science and Engineering, April 2016, and in Karam and Li, "Stereo Vision Measurement System and Method," US Patent 9,704,232 B2, issued July 2017. The developed object identification and recognition system was granted a US patent, "Systems, Methods, and Media for Identifying Object Characteristics Based on Fixation Points," US Patent 9,501,710, issued November 22, 2016.

13. I directed the development of camera calibration and distortion correction systems. The developed Fisheye Calibration and Distortion-Correction (FisheyeCDC) Toolbox (https://ivulab.asu.edu/software/fisheyecdc-toolbox/) has been adopted by industry as well as academia. Details about the developed camera calibration, camera modeling, and distortion correction systems can be found in

Prakash and Karam, "Camera Calibration using Adaptive Segmentation and Ellipse Fitting for Localizing Control Points," IEEE International Conference on Image Processing (ICIP), September 2012, and in Li and Karam, "Depth Estimation from Multi-View Images," IEEE International Conference on Emerging Signal Processing Applications, January 2012, and in Andrade and Karam, "Robust Radial Distortion Correction Based on Alternate Optimization," IEEE International Conference on Image Processing, September 2016.

14. The developed machine learning and computer vision systems for autonomous vehicles include Advanced Driver Assist Systems (ADAS) technologies. The developed real-time forward collision warning system prototype included 3D geometry and camera modeling and was demonstrated by our Intel industry collaborators at the 2015 Consumer Electronics Show (CES) in Las Vegas. More details can be found in Prakash, Li, Akhbari, and Karam, "Sparse Depth Calculation using Real-time Key-point Detection and Structure from Motion for Advanced Driver Assist Systems," Springer Lecture Notes, ISVC, December 2014, and in Prakash, Akhbari, and Karam, "Robust Obstacle Detection for Advanced Driver Assistance Systems using Distortions of Inverse Perspective Mapping of a Monocular Camera," Robotics and Autonomous Systems Journal, 2019.

15. I directed the development of scalable visual compression technologies. The developed codecs, utilizing the scalable visual compression technologies, were

commercialized by General Dynamics. More details can be found in Chien, Sadaka, Abousleman, and Karam, "Region-of-Interest-Based Ultra Low-Bit-Rate Video Coding," SPIE Symposium on Defense & Security, March 2008. I also directed the development of perceptual-based visual compression methods and algorithms. The work on JPEG2000 Encoding with Perceptual Distortion Control enabled the integration of adaptive perceptual-based visual processing and compression in the JPEG 2000 image coding standard and demonstrated improved performance in terms of visual quality and compression, while maintaining full compatibility with the JPEG 2000 standard. For this perceptual-based image compression work, I received a Technical Innovation Award from NASA.

16. In addition, I directed the development of automated biomedical image processing and computer vision technologies that enable high-throughput cancer diagnostics and drug discovery. The developed automated image analysis technologies were licensed for cancer research at different industries and institutions, including the Translational Genomics Institute (TGEN) and the New York School of Medicine. More details can be found in Karam and Said, "Automatic Cell Migration and Proliferation Analysis," U.S. Patent 9,082,164, issued July 2015.

17. My work has been published numerous times in prestigious journals, including the IEEE Transactions on Signal Processing, IEEE Transactions on Image Processing, IEEE Transactions on Pattern Analysis and Machine Intelligence

(PAMI), IEEE Transactions on Broadcasting, and IEEE/CVF Computer Vision and Pattern Recognition (CVPR) Conference. Some selected publications include "Morphological Reversible Contour Representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, March 2000, "Morphological Text Extraction from Images," IEEE Transactions on Image Processing," Nov. 2000, "A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB)," IEEE Transactions on Image Processing, April 2009, "Automated Detection and Classification of Non-Wet Solder Joints," IEEE Transactions on Automation Science and Engineering, Jan. 2011, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," IEEE Transactions on Broadcasting, June 2011, "An Efficient Selective Perceptual-Based Super-Resolution Estimator," IEEE Transactions on Image Processing, Dec. 2011, "Select Trends in Image, Video, and Multidimensional Signal Processing," IEEE Signal Processing Magazine, Jan. 2012, "A Distributed Canny Edge Detector: Algorithm and FPGA Implementation," IEEE Transactions on Image Processing, July 2014, "3D Blur Discrimination," ACM Transactions on Applied Perception, May 2016, "Spatially-Varying Blur Detection Based on Multiscale Fused and Sorted Transform Coefficients of Gradient Magnitudes," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017, "Unconstrained Ear Recognition using Deep Neural Networks," IET Biometrics, January 2018, "Visual

Saliency Prediction Using a Mixture of Deep Neural Networks," IEEE Transactions on Image Processing, August 2018, "Quality Robust Mixtures of Deep Neural Networks," IEEE Transactions on Image Processing, November 2018, "Augmented Sparse Representation Classifier (ASRC) for Face Recognition under Quality Distortions," IET Biometrics Journal, November 2019, "DeepCorrect: Correcting DNN Models against Image Distortions," IEEE Transactions on Image Processing, December 2019, "It GAN DO Better: GAN-based Detection of Objects on Images with Varying Quality," IEEE Transactions on Image Processing, November 2021.

18. I am an inventor, and I have been awarded eight patents in the field of image and video processing, compression, and transmission: U.S. Patent Nos. 6,154,493 ("The Compression of Color Images Based on a 2-Dimensional Discrete Wavelet Transform Yielding a Perceptually Lossless Image"), 6,124,811 ("Realtime Algorithms and Architectures for Coding Images Compressed by DWT-Based Techniques"), 6,717,990 ("Communication System and Method for Multi-Rate, Channel-Optimized Trellis-Coded Quantization"), 7,551,671 ("System and Method for Transmission of Video Signals using Multiple Channels"), 9,082,164 ("Automatic Cell Migration and Proliferation Analysis"), 9,501,710 ("Systems, Methods, and Media for Identifying Object Characteristics Based on Fixation Points"), 9,704,232 ("Stereo Vision Measurement System and Method"), and 11,0304,85 ("Systems and Methods for Feature Corrections and Regeneration for

Robust Sensing, Computer Vision, and Classification,"). I was also a named inventor on an additional patent application, published as U.S. 2008/0259796 ("Method and Apparatus for Network-Adaptive Video Coding").

19. Additional information regarding my background, qualifications, publications, and presentations is provided in my CV, which is provided as Exhibit 1004.

II. TASK SUMMARY AND MATERIALS REVIEWED

20. I have been asked to review the Challenged Patent and its prosecution history, to provide an understanding of the technology relevant to the Challenged Patent, to review certain prior-art references, and analyze whether those references disclose or teach the limitations of the claims of the Challenged Patent. The opinions stated in this declaration are from the perspective of a person of ordinary skill in the art (POSITA).

III. LIST OF EXHIBITS

21. In forming my opinions, I have reviewed the Challenged Patent, its prosecution history, the materials cited in the List of Exhibits, and the materials cited throughout my declaration. Those materials include the following:

EX	Description
1001	U.S. Patent No. 8,982,110 to Ofer Saban and Nissi Vilcovsky (filed April 15, 2014 and issued March 17, 2015) (the "'110 Patent" or "Challenged Patent")

EX	Description
1002	File History of U.S Patent No. 8,982,110 (App. No. 14/253,831) ("'110 Patent File History")
1003	[Reserved]
1004	Curriculum Vitae of Dr. Lina Karam, Ph.D.
1005	U.S. Patent Pub. No. 2011/0298910 to De Haan et al. (filed February 15, 2010 and published December 8, 2011) ("Haan")
1006	U.S. Patent Pub. No. US 2007/0040033 to Rosenberg (filed September 26, 2006 and published February 22, 2007) ("Rosenberg")
1007	U.S. Patent Pub. No. 2011/0246329 to Geisner et al. (filed April 1, 2010 and published October 6, 2011) ("Geisner")
1008	WO Patent Pub. No. 2006/108279 to Lizee et al. (filed April 11, 2006 and published October 19, 2006) ("Lizee")
1009	Institute for Robotics and Intelligent Systems, Alexandre R.J. Francois, *Software Architecture for Computer Vision: Beyond Pipes and Filters* (July 2003) ("Francois")
1010	Hartley et al., *Multiple View Geometry in Computer Vision* (2nd ed., published 2004) ("Hartley")
1011	U.S. Patent Pub. No. 2007/0065039 to Park et al. (filed September 22, 2006 and published March 22, 2007) ("Park")
1012-1017	[Reserved]
1018	Excerpts from EyesMatch Ltd.'s and Memomi Labs Inc.'s Objections and Responses to First Set of Interrogatories of Meta Platforms, Inc., Instagram, LLC, and WhatsApp Inc. (Nos. 1-4) in *EyesMatch Ltd. v. Facebook, Inc.*, No. 1:21-cv-00111-RGA-JLH (D. Del.) ("Excerpts from Patent Owner's Infringement Contentions in the Meta Litigation")

EX	Description
1019	Excerpts from Plaintiffs EyesMatch Ltd.'s Disclosure of Asserted Claims and Infringement Contentions in *EyesMatch Ltd. v. Samsung Electronics America, Inc.*, No. 2:23-cv-00363-JRG-RSP (E.D. Tex.)
1020	*EyesMatch Ltd. v. Samsung Electronics America, Inc.*, No. 2:23-cv-00363-JRG-RSP (E.D. Tex.), Docket Control Order, Dkt. No. 27
1021	U.S. Patent No. 8,982,109 to Nissi Vilcovsky and Ofer Saban (filed March 15, 2013 and issued March 17, 2015) (the "'109 Patent")
1022	File History of U.S Patent No. 8,982,109 (App. No. 13/843,001) ("'109 Patent File History")
1023	U.S. Provisional Patent Application No. 61/738,957 (filed December 18, 2012)
1024	File History of U.S. Provisional Patent Application No. 13/088,369 (filed April 17, 2011)
1025	U.S. Provisional Patent Application No. 60/656,884 (filed March 1, 2005)
1026	U.S. Provisional Patent Application No. 60/656,885 (filed March 1, 2005)
1027	James D. Foley et al., *Computer Graphics: Principles and Practice* (2nd ed. in C, published 1995) ("Foley")
1028	Excerpts from Alan Watt, *3D Computer Graphics* (3rd ed., published 2000) ("Watt")
1029	Declaration of Sylvia D. Hall-Ellis, Ph. D. ("Hall-Ellis Declaration")
1030	Patent Owners EyesMatch Ltd., Memomi Labs Inc.'s Preliminary Response in *Meta Platforms, Inc., Instagram, LLC, and WhatsApp Inc. v. EyesMatch Ltd., Memomi Labs Inc.*, IPR2022-00490, Paper 9 ("Patent Owner's Preliminary Response in the Meta IPR2022-00490")

Other than Ex. 1002 ('110 Patent File History), Ex. 1018 (Excerpts from Patent Owner's Infringement Contentions in the Meta Litigation), Ex. 1028 (Watt), and Ex.

1030 (Patent Owner's Preliminary Response in the Meta IPR2022-00490) all citations to exhibits reference original page numbers found in the underlying documents, and all emphases have been added. Citations to Exs. 1002, 1018, 1028, 1030 reference the stamped page numbers added in the lower right corner of each exhibit page.

IV. SUMMARY OF CONCLUSIONS AND GROUNDS FOR CHALLENGE

22. After reviewing the materials identified in the List of Exhibits, I have concluded that each of claims 1-18 (the "Challenged Claims") of the '110 Patent would have been obvious to a person of ordinary skill in the art, based on the following grounds:

Ground	'110 Patent Claims	Combination of Prior-Art References
1	1-6, 10-12, 16-17	Obvious over Haan in view of Francois
2	1-6, 10-12	Obvious over Rosenberg in view of Park
3	1-18	Obvious over Geisner in view of Lizee
4	7-9, 13-15, 18	Obvious over Haan in view of Francois and Geisner
5	3-4, 7-9, 11-18	Obvious over Rosenberg in view of Park and Geisner

V. LEGAL STANDARDS

A. LEVEL OF ORDINARY SKILL IN THE ART

23. When interpreting a patent, I understand that it is important to identify the relevant art pertaining to the patent as well as the level of ordinary skill in that

art at the time of the claimed invention. The "art" is the field of technology to which the patent is related.

24. I have been instructed by counsel that the person of ordinary skill in the art ("POSITA") is a hypothetical person who is presumed to know the relevant prior art. I understand that the actual inventor's skill is not determinative of the level of ordinary skill. I further understand that the factors that may be considered in determining the level of skill include: the types of problems encountered in the art; prior art solutions to those problems; rapidity with which innovations are made; sophistication of the technology; and educational level of active workers in the field. I understand that not all such factors may be present in every case, and one or more of them may predominate.

B. LEGAL STANDARD FOR CLAIM CONSTRUCTION

25. I understand that the first step in determining whether a patent claim would have been anticipated or obvious is to ascertain how a POSITA would have understood the claim terms.

26. I have been instructed by counsel on the law regarding claim construction and patent claims, and I understand that a patent may include two types of claims: independent claims and dependent claims. An independent claim stands alone and includes only the limitations it recites. A dependent claim can depend from an independent claim, or it can further depend from another dependent claim. I

understand that a dependent claim includes all the limitations that it recites, in addition to all the limitations recited in the claim(s) from which it depends.

27. I understand that the United States Patent and Trademark Office interprets claim terms in an *inter partes* review proceeding under the same claim construction standard that is used in a United States federal court. I understand that under this standard, the meaning of claim terms is considered from the viewpoint of one of ordinary skill in the art at the time of the alleged invention.

28. I have been informed that claim terms are generally given their ordinary and customary meaning as understood by one of ordinary skill in the art in light of the specification and the prosecution history pertaining to the patent. I understand, however, that claim terms are generally not limited by the embodiments described in the specification.

29. I understand that in addition to the claims, specification, and prosecution history, other evidence may be considered to ascertain the meaning of claim terms, including textbooks, encyclopedias, articles, and dictionaries. I have been informed that this other evidence is often less significant and less reliable than the claims, specification, and prosecution history.

C. ANTICIPATION

30. I have been told that, in an *inter partes* review proceeding, a patent claim may be deemed unpatentable if it is shown by a preponderance of the evidence that the claim was anticipated by a prior art reference under 35 U.S.C. § 102.

31. I have been informed by counsel that for a claim to be anticipated under 35 U.S.C. § 102, every limitation of the claimed invention must be found either explicitly or inherently in a single prior art reference. I have also been informed that there is a set process as follows: a) the claims of a patent are properly construed, b) then, one must compare the claim language to the prior art on a limitation-by-limitation basis. I understand that an anticipatory reference does not have to recite word for word what is in the anticipated claims. Rather, anticipation can occur when a claimed limitation is "inherent" in the relevant reference. I have been advised that if the prior art necessarily functions in accordance with, or includes, the claim limitations, it can anticipate even though a limitation is not expressly recited. I understand that inherency requires a showing that the missing descriptive matter in the claim is necessarily or implicitly present in the allegedly anticipating reference, and that it would have been so recognized by a POSITA. In addition, I understand that an enabling disclosure is a disclosure that allows a POSITA to make the invention without undue experimentation. Although anticipation typically involves the analysis of a single prior art reference, I understand that additional references

may be used to show that the primary reference has enabling disclosure, to explain the meaning of a term used in the primary reference, and/or to show that a characteristic is inherent in the primary reference.

32. I have been informed that in *inter partes* review proceedings, such as this one, the party challenging the patent bears the burden of proving unpatentability by a preponderance of the evidence. I understand that a preponderance of the evidence means "more likely than not."

D. OBVIOUSNESS

33. I understand that a patent claim is invalid as obvious if the claimed invention would have been obvious to a person of ordinary skill in the art at the time the claimed invention was made. This means that even if all of the elements of the claim cannot be found in a single prior art reference that would anticipate the claim, a person of ordinary skill in the field who knew about the prior art would have come up with the claimed invention. I understand that in an obviousness determination, the person of ordinary skill in the art is presumed to have knowledge of all prior art material. I understand that whether a claim is obvious is based upon the determination of several factual issues.

34. I understand that obviousness is a question of law based on underlying factual issues including the scope and content of the prior art, the level of ordinary

skill in the art, the differences between the claimed invention and the prior art, and secondary considerations of non-obviousness.

35. In considering obviousness, I understand that one must determine the scope and content of the prior art. I understand that, in order to be considered as prior art to a patent being considered, a prior art reference must be reasonably related to the claimed invention of that patent. A reference is reasonably related if it is in the same field as the claimed invention or is from another field to which a person of ordinary skill in the art would look to solve a known problem.

36. I understand that one must determine what differences, if any, existed between the claimed invention and the prior art.

37. I understand that a patent claim composed of several elements is not proved obvious merely by demonstrating that each of its elements was independently known in the prior art. In evaluating whether such a claim would have been obvious, one may consider whether a reason has been identified that would have prompted a person of ordinary skill in the art to combine the elements or concepts from the prior art in the same way as in the claimed invention. There is no single way to define the line between true inventiveness on the one hand (which is patentable) and the application of common sense and ordinary skill to solve a problem on the other hand (which is not patentable). For example, market forces or other design incentives may be what produced a change, rather than true inventiveness.

38. I understand that whether a prior art reference, alone or in combination with one or more other prior art references, renders a patent claim unpatentable as obvious is determined from the perspective of a person of ordinary skill in the art at the time of the alleged invention. I have been told that there is no requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, but a suggestion to combine known elements to achieve the claimed invention may come from the prior art, as filtered through the knowledge of one skilled in the art. In addition, I have been told that the inferences and creative steps a person of ordinary skill in the art would employ are also relevant to the determination of obviousness.

39. I understand that there is no rigid rule that a reference or combination of references must contain a "teaching, suggestion, or motivation" to combine references. But I also understand that the "teaching, suggestion, or motivation" test can be a useful guide in establishing a rationale for combining elements of the prior art. I have been told that this test poses the question as to whether there is an express or implied teaching, suggestion, or motivation to combine prior art elements in a way that realizes the claimed invention, and that the test seeks to counter impermissible hindsight analysis.

40. I understand that one may consider, e.g., whether (1) the change was merely the predictable result of using prior art elements according to their known

functions, or whether it was the result of true inventiveness; (2) there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent; (3) the claimed innovation applies a known technique that had been used to improve a similar device or method in a similar way; (4) the claimed invention would have been obvious to try, meaning that the claimed invention was one of a relatively small number of possible approaches to the problem with a reasonable expectation of success by those skilled in the art; (5) the invention merely substituted one known element for another known element in order to obtain predictable results; (6) the invention merely applies a known technique to a known device, method, or product to yield predictable results; or (7) known work in the field may have prompted variations of use of the same inventions in the same or different fields due to market forces or design incentives that would have been predictable to a person of ordinary skill in the art.

41. I understand that any assertion of secondary considerations of non-obviousness must be accompanied by a nexus between the merits of the invention and the evidence offered.

VI. TECHNOLOGY BACKGROUND

42. In the subsections below, I describe some of the basic principles that would have been well-known to a POSITA as of the priority date of the ’110 Patent, which I understand to be December 18, 2012 (*see infra* §VIII, ¶90), and long before.

A. EXEMPLARY REFERENCES CORROBORATING GENERAL ENGINEERING KNOWLEDGE

a. Hartley Text (Ex. 1010)

43. As an example of engineering knowledge pertaining to multiple-view geometry modeling, Exhibit 1010 is a textbook entitled "Multiple View Geometry in Computer Vision, Second Edition," by Richard Hartley and Andrew Zisserman. This book memorializes the state of the art of multiple-view geometry modeling as of September 2004, when the book was published (Ex. 1029 (Hall-Ellis Declaration) at ¶49), and provides a perspective of the knowledge of one of ordinary skill in the art.

b. Foley Text (Ex. 1027)

44. As an example of engineering knowledge pertaining to 2D and 3D computer graphics and modeling, Exhibit 1027 is a textbook entitled "Computer Graphics Principles and Practice (Second Edition in C)" by James D. Foley et al. This book memorizes the state of the art of computer graphics and representing and modeling 2D and 3D objects as of October 1995, when the book was published (Ex. 1029 (Hall-Ellis Declaration) at ¶58), and provides a perspective of the knowledge of one of ordinary skill in the art.

c. Watt Text Excerpts (Ex. 1028)

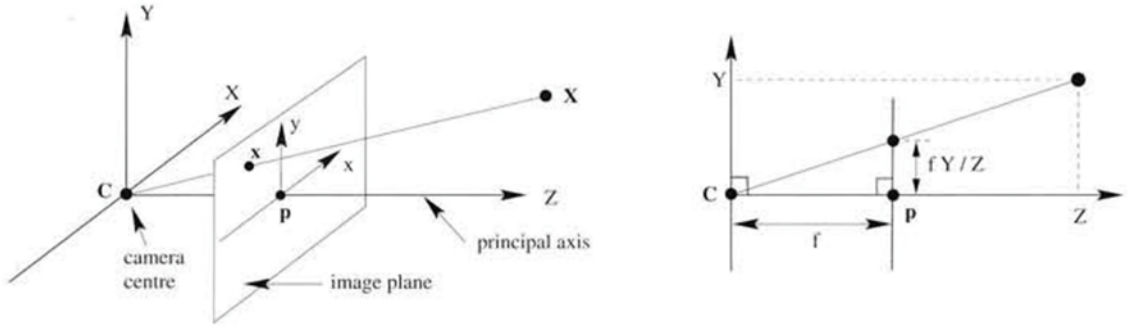
45. As an example of engineering knowledge pertaining to 2D/3D graphics modeling and image transformations to simulate the appearance of a real object,

Exhibit 1028 is a copy of excerpts from the textbook entitled "3D Computer Graphics, Third Edition" by Alan Watt. This book memorizes the state of the art of computer graphics and representing and modeling 3D objects as of December 2000, when the book was published (Ex. 1029 (Hall-Ellis Declaration) at ¶66), and provides a perspective of the knowledge of one of ordinary skill in the art.

B. BASIC IMAGE TRANSFORMATIONS

a. Coordinate Representation of 2D and 3D points, World and Camera Coordinate Systems, Camera Image Plane, Camera Projection Matrix and Projective Transformation, Homography Transformation

46. The figure below illustrates a camera geometry model with a camera center **C** and a camera focal length *f*, where a 3D point **X** in a world coordinate frame is mapped to a 2D point **x** on the camera's image plane by means of a projective transformation. The 2D point **x** is the point at which the line joining the 3D point **X** to the camera center **C** intersects the camera's image plane ("image plane") and its coordinates are expressed in terms of the image plane coordinates. Ex. 1010 (Hartley) at 153-155, Fig. 6.1.



Ex. 1010 (Hartley) at 154, Fig. 6.1.

47. In the figure above, the line from the camera center perpendicular to the image plane is called the principal axis. The point **p** at which the principal axis intersects with the image plane is called the principal point. This projective transformation mapping is represented by the camera's projection matrix P and is given by **x** = P**X,** where **x** and **X** are represented using homogeneous coordinates and P is a 3x4 matrix (i.e., a matrix having 3 rows and 4 columns). Ex. 1010 (Hartley) at 153-155, Fig. 6.1; *see also infra* §§VI.B.c VI.B.d, VI.B.e.

48. In the fields of computer vision and computer graphics, geometric modeling and geometric transformations are commonly expressed by representing 2D and 3D points using homogeneous coordinates. Let $\tilde{\mathbf{x}}$ = (x,y) and $\tilde{\mathbf{X}} = (X, Y, Z)$ be, respectively, a 2D point in a camera's image plane coordinate frame and a 3D point in a world coordinate frame. The homogeneous representation, also called homogeneous coordinates, of $\tilde{\mathbf{x}}$ = (x,y) and $\tilde{\mathbf{X}} = (X, Y, Z)$ is obtained by representing $\tilde{\mathbf{x}}$ using a three-dimensional vector **x** = (x, y, 1) and by representing $\tilde{\mathbf{X}}$ using a four-dimensional vector **X** = (X, Y, Z, 1). More generally, the last element in a homogeneous representation serves as a scale factor. It follows that the homogenous representation (x, y, w) represents the 2D point (x/w, y/w). Similarly, the homogeneous representation (X, Y, Z, W) represents a 3D point (X/W. Y/W, Z/W). Conversely, $\tilde{\mathbf{x}}$ = (x,y) and $\tilde{\mathbf{X}} = (X, Y, Z)$ are referred to as the inhomogeneous

representation or inhomogeneous coordinates. Ex. 1010 (Hartley) at 27; *see also* Ex. 1027 (Foley) at 204.

49. In the figure above, the camera coordinate system (i.e., camera's center **C** and axes X, Y, Z) coincides with the world coordinate system (i.e., world coordinate system's origin and axes). In such case, the camera projection matrix is expressed as P = K[I|**0**] where I is the 3x3 identity matrix, **0** is a 3x1 column vector whose elements are all zero, and K is the intrinsic camera matrix (also known as the camera calibration matrix) and is represented by $K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$ "where $\alpha_x = fm_x$ and $\alpha_y = fm_y$ represent the focal length of the camera in terms of pixel dimensions in the x and y directions respectively[,]" $x_0 = p_x m_x$ and $y_0 = p_y m_x$ represent the coordinates of the camera's principle point in terms of pixel dimensions, and s is a skew parameter which is typically zero for most normal cameras (e.g., CCD cameras).[1] Ex. 1010 (Hartley) at 156-158. K contains the camera's intrinsic parameters (e.g., focal length, principal point, pixel dimensions, skew).

50. In general, the world coordinate frame does not coincide with and can be different from the camera coordinate frame, as shown in the figure below.

---

[1] In the case of square pixels, m_x and m_y can be set to be equal to 1. Ex. 1010 (Hartley) at 155-157.

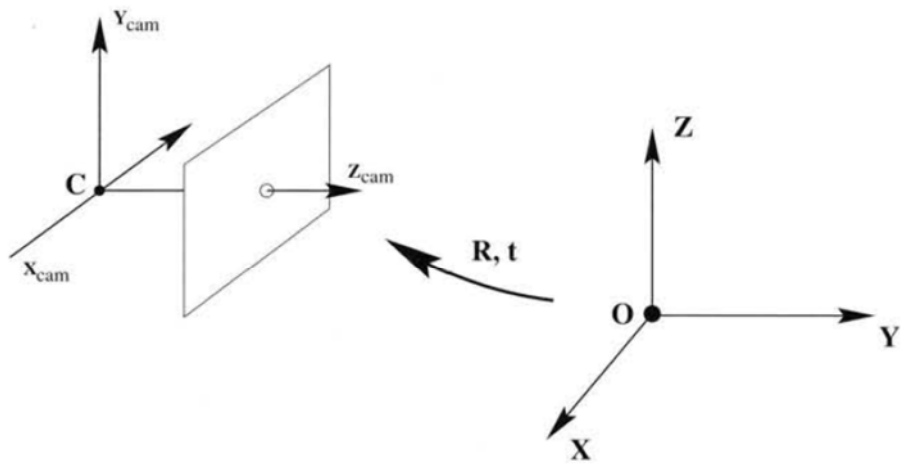


Fig. 6.3. *The Euclidean transformation between the world and camera coordinate frames.*

Ex. 1010 (Hartley) at 155-156, Fig. 6.3.

51. As shown in the figure above, the world coordinate frame and the camera coordinate frame are related via a rotation, i.e., a 3x3 rotation matrix R, and a translation, i.e., a 3x1 translation vector **t**. In this case, the camera projection matrix is expressed as P = as K[R|**t**], where **t**=-R$\tilde{\mathbf{C}}$ and where $\tilde{\mathbf{C}}$ is the inhomogeneous representation of the camera center **C**. It follows that P = KR[I|-$\tilde{\mathbf{C}}$]. Ex. 1010 (Hartley) at 155-157; *see also* Ex. 1028 (Watt) at 23-26.

52. The point imaging equation **x** = P**X** is a map from a point in a world coordinate frame, to a point in image coordinates. Ex. 1010 at 196. This allows the world coordinate frame to be chosen. Ex. 1010 (Hartley) at 196. If, hypothetically, the world coordinate frame is chosen such that the XY-plane corresponds to a plane π in the scene, points on the plane π (also referred to as scene plane) have zero Z-coordinate, as shown in in the figure below. Ex. 1010 (Hartley) at 196, Fig. 8.1.

Then, if the columns of the 3x4 camera projection matrix P are denoted as $\mathbf{p}_i$, the image **x** of a point **X** on the plane π can be given by:

$$\mathbf{x} = \mathrm{P}\mathbf{X} = [\mathrm{p}_1\ \mathrm{p}_2\ \mathrm{p}_3\ \mathrm{p}_4]\begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = [\mathrm{p}_1\ \mathrm{p}_2\ \mathrm{p}_4\ ]\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = [\mathrm{p}_1\ \mathrm{p}_2\ \mathrm{p}_4\ ]\mathbf{x}_{\boldsymbol{\pi}}$$

Ex. 1010 (Hartley) at 196. Thus, the map between points $\mathbf{x}_{\boldsymbol{\pi}} = (X, Y, 1)^T$ on the plane π and their image **x** takes the form of a general planar homography (i.e., a plane-to-plane projective transformation): $\mathbf{x} = \mathrm{H}\mathbf{x}_{\boldsymbol{\pi}}$, with H being a 3x3 matrix of rank 3. This shows that the most general transformation that can occur between a scene plane (e.g., plane π) and an image plane or between two image views under perspective imaging is a plane projective transformation.

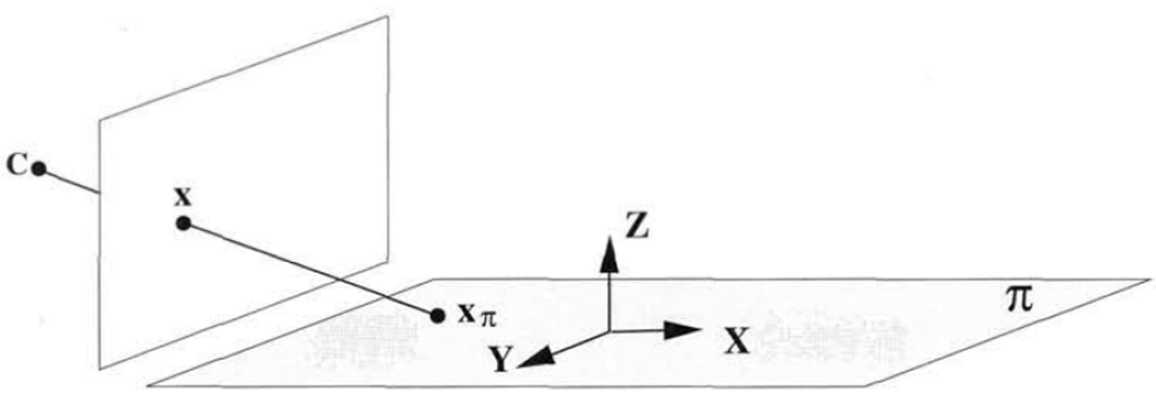


Fig. 8.1. **Perspective image of points on a plane.** *The XY-plane of the world coordinate frame is aligned with the plane π. Points on the image and scene planes are related by a plane projective transformation.*

Ex. 1010 (Hartley) at 196, Fig. 8.1.

b. <u>Image Flipping</u>

53. Flipping an image captured by a camera about a theoretical vertical axis was a well-known technique in the field of image formation and computer vision that allowed for the image to mimic a mirror's reflection. *See, e.g.*, Ex. 1010 (Hartley) at 38. This image flipping transformation falls under the group of isometry transformations, which is mathematically represented by a group of matrices. Ex. 1010 (Hartley) at 38; *see also* Ex. 1027 (Foley) at 203-206 (Eqs. 5.6, 5.7, and 5.24); Ex. 1028 (Watt) at 13-21, 23-26, 37-40. One such exemplary matrix representing an isometry transformation is shown below, where ϵ can either equal +1 or -1. Ex. 1010 (Hartley) at 38:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \epsilon\, cos\theta & -sin\theta & t_x \\ \epsilon\, sin\theta & cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

54. When ϵ is equal to -1, then the orientation of the captured image is reversed, which results in the image being reflected. Ex. 1010 (Hartley) at 38. Flipping of the image can be obtained by setting the rotation angle θ and the translation vector t = (t_x, t_y) to zero and by setting ϵ = -1. Ex. 1010 (Hartley) at 38; *see also* Ex. 1027 (Foley) at 203-206 (Eqs. 5.6, 5.7 and 5.24); Ex. 1028 (Watt) at 16-21, 23-26.

c. <u>Synthesis of Virtual Camera Views from Actual Camera Views</u>

55. It has long been known that modified images (i.e., virtual views) that appear to be captured from different point of views than the actual camera's point of view can be generated from an actual image that was captured by the actual camera (i.e., actual view). Ex. 1028 (Watt) at 13, 23-24, 30-31, 37-38, 42-46; *see also* Ex. 1027 (Foley) at 822-32. The modified images are generated from the actual images using transformation mapping. Ex. 1028 (Watt) at 13, 15-21, 23-31, 37-38; *see also* Ex. 1027 (Foley) at 820-26. The transformation mapping can be derived based on both the actual camera's projection matrix and the virtual camera's projection matrix. Ex. 1010 (Hartley) at 202-208, 228, Figs. 8.6, 8.7, 8.25. *See also generally*, Ex. 1010 (Hartley) at 153-165; *see also* Ex. 1027 (Foley) at 820-826; Ex. 1028 (Watt) at 16-21, 23-26, 37-40.

56. As described in §VI.B.a *supra*, a camera projection matrix P can be represented as $P=K[R|\mathbf{t}]$, where R is a 3×3 rotation matrix representing the orientation of camera coordinates with respect to the world coordinates (also referred to as global coordinates); **t** is a 3×1 translation vector which shifts a camera center C with respect to the world coordinate system; **t** is represented by $\mathbf{t}= -R\tilde{\mathbf{C}}$, where $\tilde{\mathbf{C}}$ is the inhomogeneous representation of the camera center C; K is the intrinsic camera matrix (also known as camera calibration matrix) as discussed in §VI.B.a *supra*, ¶49.

K contains the camera's intrinsic parameters (e.g., focal length, principal point, pixel dimensions, skew); R and **t** correspond to the camera's extrinsic parameters. The transformation between the camera and world coordinates can be seen in the figure below:

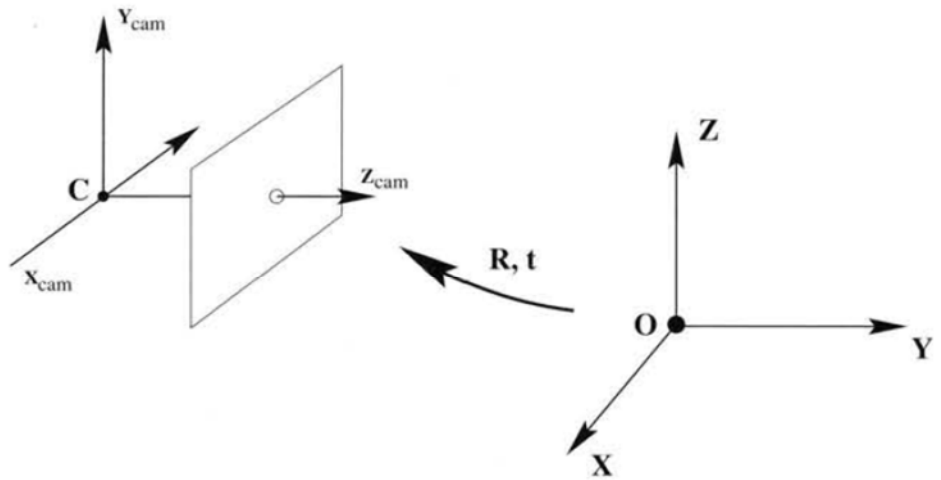


Fig. 6.3. *The Euclidean transformation between the world and camera coordinate frames.*

Ex. 1010 (Hartley) at 156, Fig. 6.3.

57. For example, let the actual camera coordinates be taken to be the world coordinates (i.e., the camera center is the center of the world coordinate system and the camera coordinate system's axes are aligned with the world coordinate system's axes.) Ex. 1010 (Hartley) at 155-156; *see also* Ex. 1028 (Watt) at 16-17, 23-26, 28. In this case, the camera rotation matrix is the 3x3 identity matrix I and the camera translation vector **t** is **0**. Ex. 1010 (Hartley) at 155; *see also* Ex. 1028 (Watt) at 23-26. The actual camera's projection matrix is then given by P_{actual} =K[I|0]. Ex. 1010 (Hartley) at 156. A new synthesized image corresponding to a desired virtual point

of view (i.e., virtual camera's point of view), that is different from the actual camera's point of view, can be generated from the captured image (i.e., image captured by the actual camera) by providing the desired virtual camera's projection matrix, $P_{virtual}$=K'[R|t], and thus providing the desired virtual camera's position and orientation along with its intrinsic parameters. Ex. 1010 (Hartley) at 156-157. A transformation mapping can then be determined based on the provided virtual camera's projection matrix and used to transform the actual image to a virtual image that appears to be captured from the desired virtual point of view. Ex. 1010 (Hartley) at 202-208, 228, Figs. 8.6, 8.7, 8.25. Alternatively, a homography transformation (i.e., plane-to-plane projective transformation) can be determined using the provided actual and virtual camera projection matrices and used to transform the actual camera's image plane (i.e., captured image by actual camera) into the virtual camera's image plane (i.e., image that appears to be captured from the virtual camera's point of view). *See supra* §VI.B.a, ¶52.

d. <u>Geometric Transformation</u>

58. Geometric image transformation was also known in the relevant art as of December 18, 2012, the priority date of the '110 Patent (*see infra* §VIII, ¶90). Given a camera positioned above or to the side of a display screen, simply flipping an image about a vertical axis is insufficient to properly replicate a traditional mirror image given the offset position and/or offset angle of the camera. Therefore,

geometric transformation(s) can be applied to the image to counteract any image distortion or warping caused by the offset position and/or angle.

59. An object in 3D-space and a camera center can define a set of rays (called a "cone"), and an image is obtained by intersecting these rays with a plane. Ex. 1010 (Hartley) at 202-203; *see also* Ex. 1028 (Watt) at 23-26. As shown below, the cone of rays can intersect two planes. Ex. 1010 (Hartley) at 202-203. The two images, each of which are formed where the cone intersects the respective planes, are thus related by a perspective map. This means that images obtained with the same camera center can be mapped to one another by a plane projective transformation, i.e., they are projectively equivalent and so have the same projective properties. Ex. 1010 (Hartley) at 202-203.

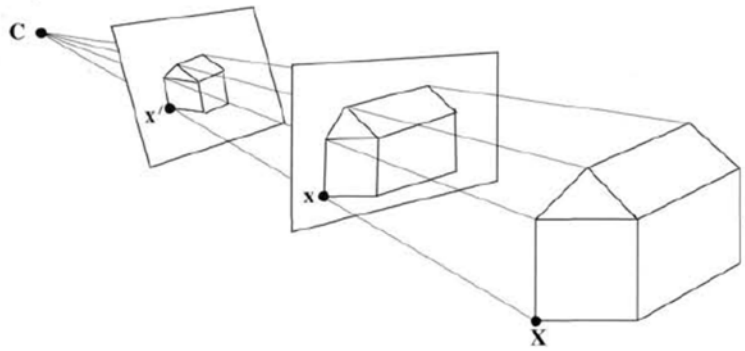


Fig. 8.5. **The cone of rays with vertex the camera centre.** *An image is obtained by intersecting this cone with a plane. A ray between a 3-space point* **X** *and the camera centre* **C** *pierces the planes in the image points* **x** *and* **x**′. *All such image points are related by a planar homography.* $\mathbf{x}' = \mathtt{H}\mathbf{x}$.

Ex. 1010 (Hartley) at 203, Fig. 8.5.

60. The image plane can be moved by first considering an increase in the focal length of the camera. Ex. 1010 (Hartley) at 203. To a first approximation, this

corresponds to a displacement of the image plane along the principal axis, i.e., a simple magnification. Ex. 1010 (Hartley) at 203. However, a compound lens zooming will perturb both the principal point and the effective camera center. Ex. 1010 (Hartley) at 203. Let f be the original focal length and let f^* be the increased focal length. The effect of zooming by a factor $k = f / f^*$ is to multiply the calibration matrix K on the right by diag(k, k, 1). Ex. 1010 (Hartley) at 204.

e. <u>Image Rotation</u>

61. It has long been known that an image can be rotated and/or translated (i.e., shifted) via a transformation mapping with the use of a projection matrix. *See generally* Ex. 1010 (Hartley) at 153-65; *see also* Ex. 1027 (Foley) at 203-206 (Eqs. 5.6, 5.7 and 5.24); Ex. 1028 (Watt) at 13-14, 16-21, 23-26, 31-40; *see supra* §VI.B.a, ¶¶47-52; §VI.B.c.

Fig. 8.6. *Between images (a) and (b) the camera rotates about the camera centre. Corresponding points (that is images of the same 3D point) are related by a plane projective transformation. Note that 3D points at different depths which are coincident in image (a), such as the mug lip and cat body, are also coincident in (b), so there is no motion parallax in this case. However, between images (a) and (c) the camera rotates about the camera centre* and *translates. Under this general motion coincident points of differing depth in (a) are imaged at different points in (c), so there* is *motion parallax in this case due to the camera translation.*

Ex. 10 (Hartley) at 204, Fig. 8.6.

62. As discussed in §VI.B.a, *supra* ¶¶47-51, a projection matrix P can be represented as P=K[R|**t**], where R is a 3×3 rotation matrix representing the orientation of camera coordinates with respect to the world coordinates; **t** is a 3×1 translation vector that shifts a camera center $\tilde{\mathbf{C}}$ with respect to the world coordinate system; **t** is represented by **t**= -R$\tilde{\mathbf{C}}$, where $\tilde{\mathbf{C}}$ is the inhomogeneous representation of the camera center C; and K is an intrinsic camera matrix (also known as a camera calibration matrix) and is represented by $K = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$[2]. Ex. 1010 (Hartley) at 157-58; *see also supra* §VI.B.a, ¶49, §VI.B.c, ¶56; Ex. 1028 (Watt) at 16-21. The transformation between different coordinates can be seen in the figure below:

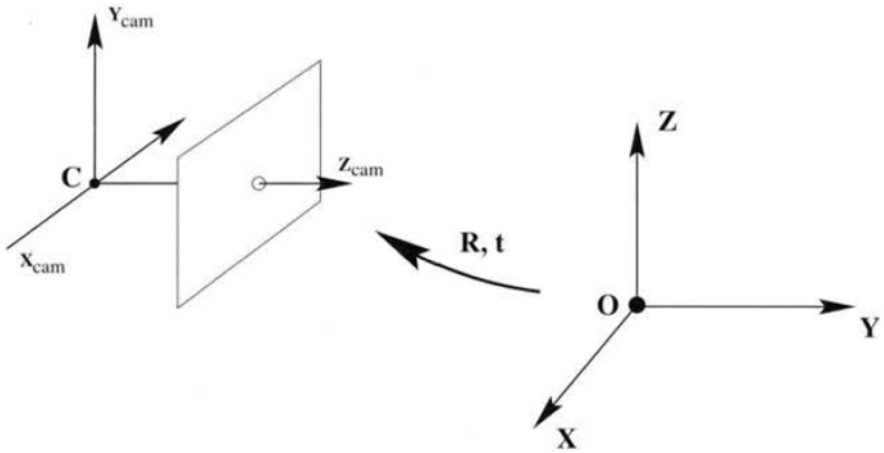


Fig. 6.3. *The Euclidean transformation between the world and camera coordinate frames.*

[2] α_x and α_y represent the focal length of the camera in terms of pixel dimensions in the *x* and *y* directions, respectively. Ex. 1010 at 157; *see also supra* §VI.B.a, ¶49.

Ex. 1010 (Hartley) at 156, Fig. 6.3.

63. For example, let the world coordinates be taken to be the same as the actual camera coordinates. *See supra* §VI.B.c, ¶57; *see also* Ex. 1028 (Watt) at 16-17, 23-26. As it was commonly known, a virtual camera that simulates a rotation of the actual camera around its center with no change in the camera internal parameters can be obtained by setting the virtual camera's projection matrix to be P_{vitual} = K[R|**0**], where the rotation matrix R provides the desired rotation. Ex. 1010 (Hartley) at 204.

64. The virtual view (i.e., modified image) corresponding to the point of view of the virtual camera can also be generated from the captured image (i.e., image captured by the actual camera) by a transformation mapping that takes the form of a planar homography transformation that can be expressed in terms of the virtual camera's intrinsic and extrinsic parameters as follows: $H = KRK^{-1}$, where K^{-1} corresponds to the inverse of the camera intrinsic matrix K. Ex. 1010 (Hartley) at 204.

f. <u>Image Resizing and Field of View Transformation</u>

65. As is also commonly known, a reflection of a user in a traditional mirror appears, to the user, to be the same size regardless of the user's point of view. By contrast, the size of a user image captured by a camera and displayed on a display generally changes depending on the position and orientation of the camera relative

to the user. It is therefore common in the field of multiple-view geometry modelling to transform an image to alleviate this problem via a scaling or perspective transformation. Ex. 1010 (Hartley) at 166-67, 170, Fig. 6.8. For example, a well-known cinematographic technique involves operating a camera to optically zoom in on an object of interest while simultaneously physically tracking the camera back, such that the object stays the same size irrespective of changes in distance between the object and the camera. Ex. 1010 (Hartley) at 166; *see also* Ex. 1010 (Hartley) at Figs. 6.7, 6.8.

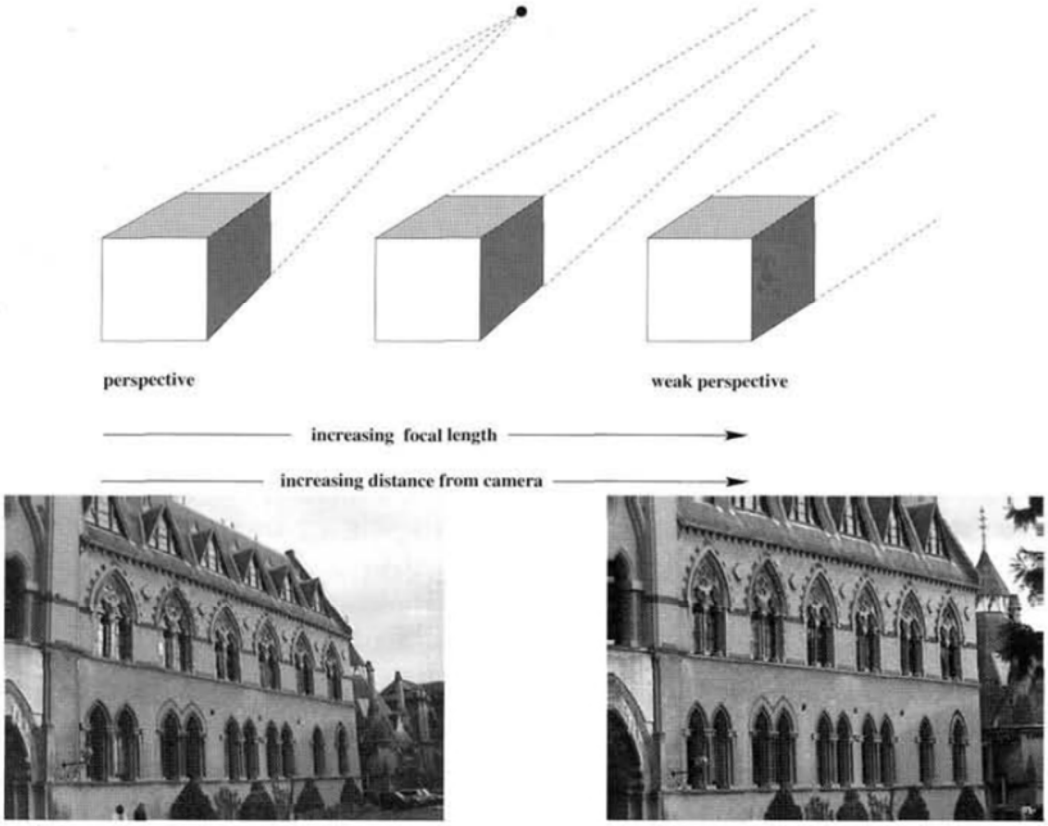


Fig. 6.7. *As the focal length increases and the distance between the camera and object also increases, the image remains the same size but perspective effects diminish.*

Ex. 1010 (Hartley) at 166, Fig. 6.7.

66. Generally, as a user moves away from the camera, the camera can be optically zoomed in and refocused on the user, such that the focal length is increased, and the image of the user is magnified. Ex. 1010 (Hartley) at 167. However, images captured using a lens with a longer focal length generally correspond to a smaller field of view. Ex. 1010 (Hartley) at 167, 169. The "tracking and zooming" approach described above is well-known and alleviates this problem, such that the size of the user remains fixed within a displayed camera image. Ex. 1010 (Hartley) at 167. This approach is well-known in the field of multiple-view geometry modeling and image formation and can be applied digitally (Ex. 1010 (Hartley) at 7, 166-170, 202-204, Figs. 6.7, 6.8, 8.5) to correct the user image based on the camera's distance to the user to match the size of the image that the user would see in a mirror.

67. Further, the figure below illustrates an example of one mapping transformation (e.g., weak perspective) that could be applied such that the size of the imaged user remains fixed within a displayed camera image, with the fixed size corresponding to the size of the user at distance d_o from the camera. Ex. 1010 (Hartley) at Figure. 6.8. Even if the user moves farther away from the camera, beyond distance d_o, the weak perspective mapping ensures that the size of the user is kept the same as the one at distance d_o. Alternatively, as the user moves away from the camera, the size of the user can be kept fixed by increasing the focal length *f*, which corresponds to moving the focal plane (i.e., plane at distance f in figure below)

towards the user. Ex. 1010 (Hartley) at Figure. 6.8. Such zooming in operation can be performed through a transformation mapping where the camera intrinsic matrix K (of the camera projection matrix) is modified by replacing the focal length of the original camera with a larger focal length. Ex. 1010 (Hartley) at Figure. 6.8; *see also supra* §VI.B.d, ¶60.

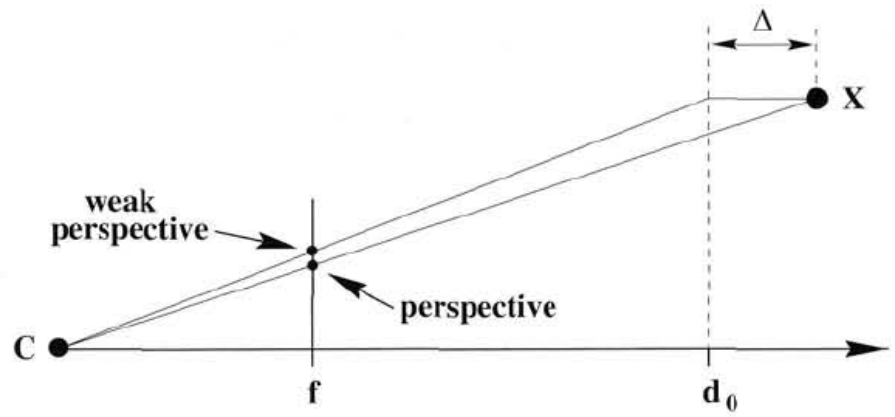


Fig. 6.8. **Perspective vs weak perspective projection.** *The action of the weak perspective camera is equivalent to orthographic projection onto a plane (at $Z = d_0$), followed by perspective projection from the plane. The difference between the perspective and weak perspective image point depends both on the distance Δ of the point* **X** *from the plane, and the distance of the point from the principal ray.*

Ex. 1010 (Hartley) at 170, Figure. 6.8.

VII. <u>OVERVIEW OF THE CHALLENGED PATENT</u>

68. I have reviewed U.S. Patent No. 8,982,110 (the "'110 Patent"), which is entitled "Method for image transformation, augmented reality, and teleperence [sic]." Ex. 1001 ('110 Patent). I understand that the '110 Patent was filed on April 15, 2014, and that it issued on March 17, 2015. *See infra* §VII.C. The '110 Patent claims priority to U.S. Provisional Application No. 61/738,957, filed December 18,

2012 (*see infra* § VII.C), but as I discuss below, the references I rely upon are dated even earlier than this provisional application.

A. THE SUBJECT MATTER OF THE CHALLENGED PATENT

69. The ’110 Patent “relates generally to imaging and display systems and, more particularly, to monitors, and interactive displays, e.g., in retail and/or service environments, medical or home situations, video conferencing, gaming, and the like.” Ex. 1001 (’110 Patent) at 1:43-46, Technical Field. Exemplary implementations relate to “virtualizing a mirror in situations where users expect to see a mirror, e.g., in trying on apparel” and in “situations where a natural appearance is preferable over standard video image, such as in, e.g., video conferencing.” Ex. 1001 (’110 Patent) at 1:47-51, Technical Field.

70. The Background section of the ’110 Patent states that “[t]he conventional mirror (i.e., reflective surface) is the common and most reliable tool for an individual to explore actual self-appearance, in real time.” Ex. 1001 (’110 Patent) at 1:55-57. According to the ’110 Patent, while “[a] few alternatives have been proposed around the combination of a camera and a screen to replace the conventional mirror,” these “are not convincing and are not yet accepted as a reliable image of the individual as if he was looking at himself in a conventional mirror.” Ex. 1001 (’110 Patent) at 1:57-62. The ’110 Patent asserts that “[t]his is mainly

because the image generated by a camera is very different from an image generated by a mirror." Ex. 1001 ('110 Patent) at 1:62-64.

71. The Background section of the '110 Patent further explains that the '110 Patent incorporates by reference U.S. Patent App. Ser. No. 13/843,001 to the same inventors (which issued as U.S. Patent No. 8,982,109 (the "'109 Patent"), Ex. 1021), which purportedly "address[es] the challenges of converting and transforming a still image or 2D/3D video created by one or more cameras, with or without other sensors, into a mirror or video conference experience." Ex. 1001 ('110 Patent) at 1:65-2:2.

72. The '110 Patent states that it describes "a computerized technique" that takes a video or one or more still images "before or after the transformation described in U.S. Patent App. Ser. No. 13/843,001, and continue[s] the computerized method to address additional functionalities, such as user interface, augmented reality, color change, texture change, shape manipulation of an object, body or background, and more." Ex. 1001 ('110 Patent) at 2:15-21. The '110 Patent describes methods to correct projection distortions resulting from the offset angle and height of the camera on top of the mirror screen. Ex. 1001 ('110 Patent) at 3:21-29, 18:58-63. Figure 8 below, from the '109 Patent, which is incorporated by reference into the '110 Patent (Ex. 1001 ('110 Patent) at 1:65-2:2, 2:17-18), shows such distortion. *See* Ex. 1021 ('109 Patent) at Fig. 8, 1:61-2:48, 15:30-40.

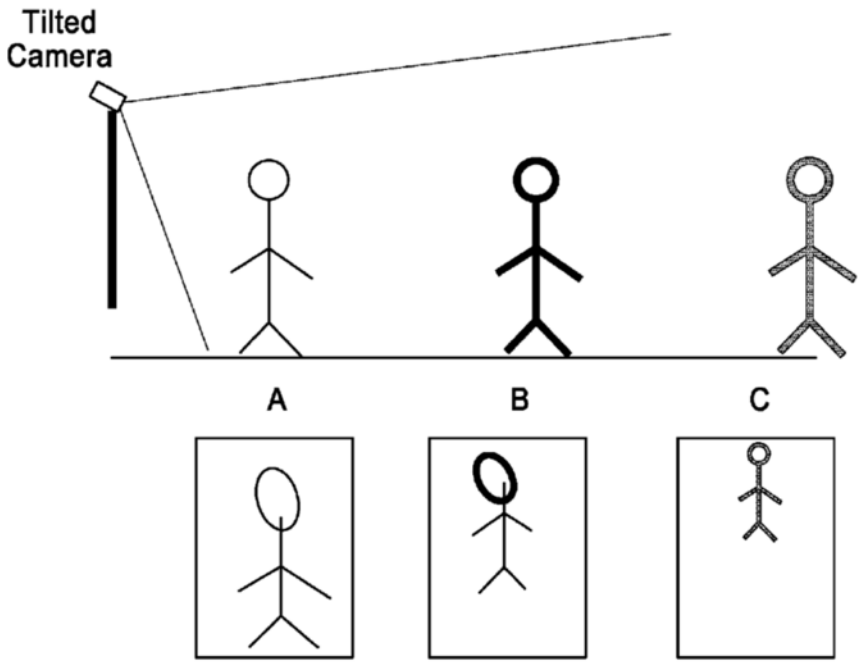


FIG. 8

Ex. 1021 ('109 Patent) at Fig. 8.

73. The virtual mirror system described in the '110 Patent includes one or more cameras, a display, a computing device, as well as "one or more machine vision augmented reality modules," and a user interface module that "can be adapted to send/receive information to/from one or more of a cloud, a web/store or a user device, e.g., smart phone or tablet." Ex. 1001 ('110 Patent) at 8:28-43; *see also* Ex. 1001 ('110 Patent) at Fig. 1.

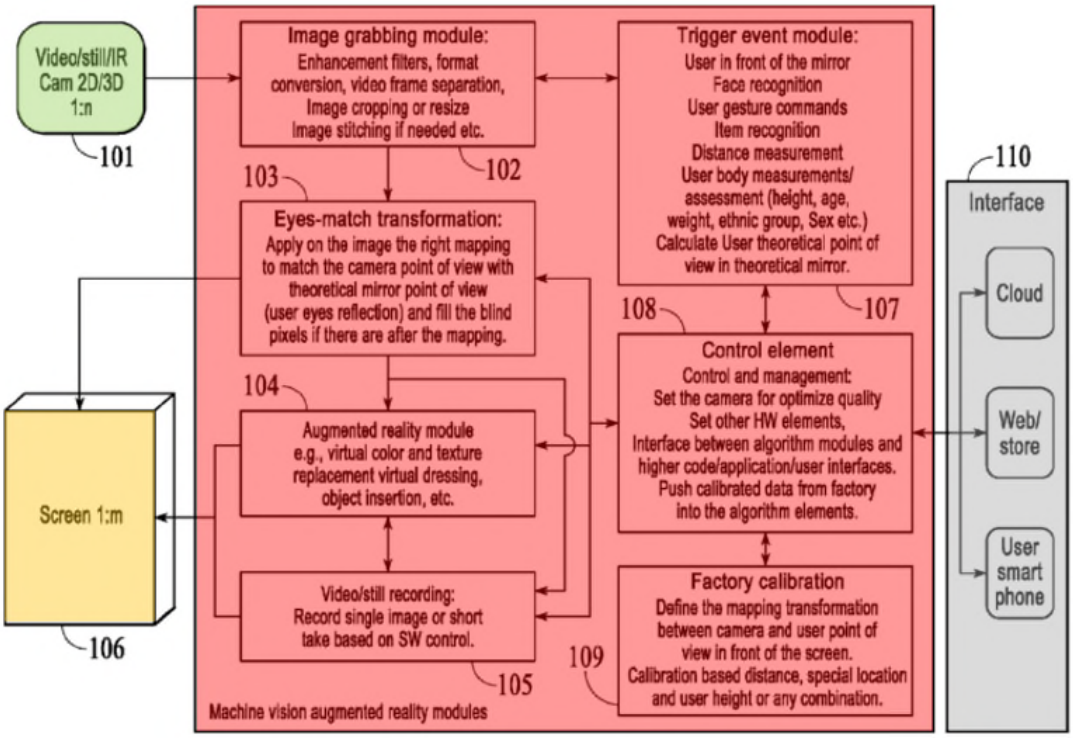
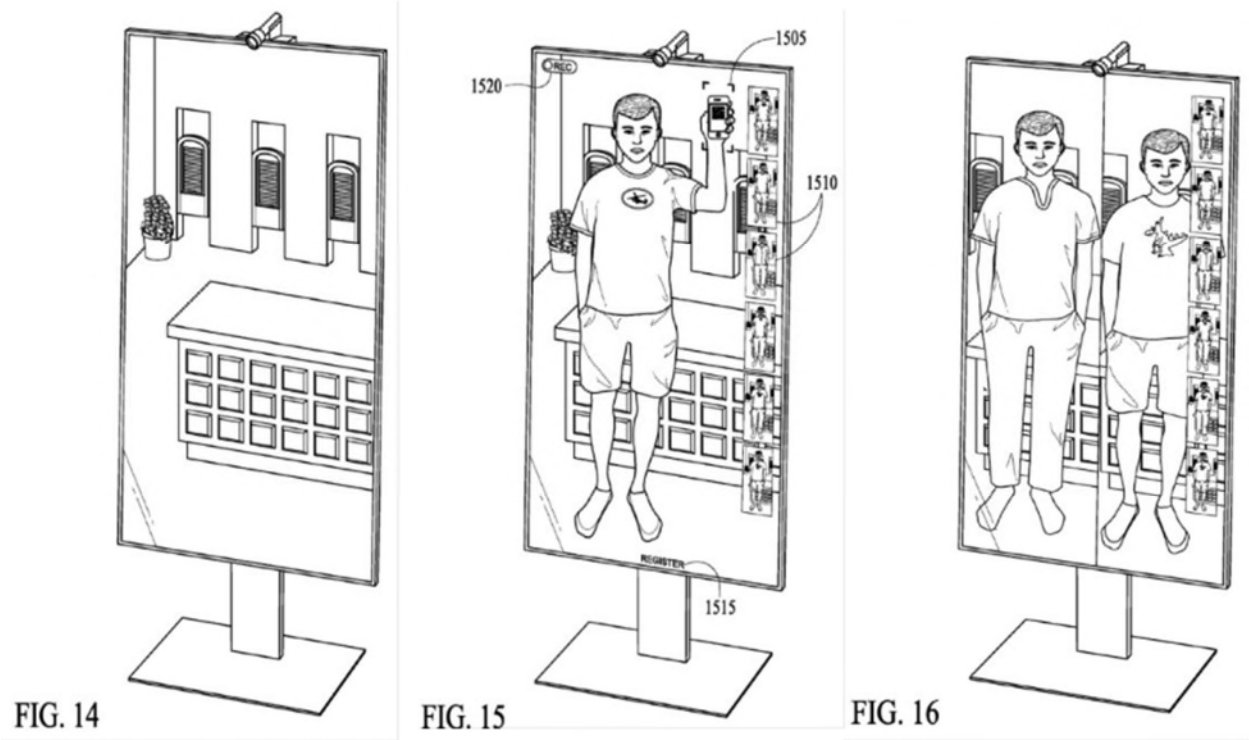


FIG. 1

Ex. 1001 ('110 Patent) at Fig. 1 (annotated).

74. The '110 Patent further discloses that "various embodiments may include single or multiple cameras from all types of sensors, including but not limited to, CCD, CMOS, IR CMOS and the like" and that "[t]he display can include single or multiple screens or virtual screen, and regular screens in all sizes, shapes and proportions," as well as "mirror like screens, smartphones or tablets, projectors, holograms, 3D screens, Google glass, projector head mount display (HMD), glass (with switchable obscure layer) and the like." Ex. 1001 ('110 Patent) at 5:16-27.

75. The ’110 Patent further discloses that one or more cameras “can be located anywhere” (Ex. 1001 (’110 Patent) at 5:28-29), and an exemplary configuration in the ’110 Patent shows a display acting as a mirror and one or more cameras fixed to the perimeter of the display. Ex. 1001 (’110 Patent) at Figs. 14-16, 20:56-60, 5:16-21; *see also* Ex. 1021 (’109 Patent) at 15:6-9.



Ex. 1001 (’110 Patent) at Figs. 14-16.

76. The ’109 Patent, which is incorporated by reference into the ’110 Patent (*see* Ex. 1001 (’110 Patent) at 1:65-2:2, 2:17-18), describes the “transformation” of a video or still images upon which the ’110 Patent purportedly expands. The ’109 Patent discloses methods for causing a computerized flat panel display system to behave similarly to a mirror. Ex. 1021 (’109 Patent) at 1:26-31. The ’109 Patent

describes image processing techniques that use a camera-generated image (captured by, for example, a digital still or video camera) and "manipulat[es] the image such that when it is projected on a screen, it resembles a mirror image, i.e., an image the user would have seen if the screen was actually a standard mirror." Ex. 1021 ('109 Patent) at 14:7-11.

77. In some embodiments, as displayed in Figure 6 of the '109 Patent below, a camera 630 captures live images of a user, a processor 645 processes said images, and a digital screen 640 displays the processed images, such that the digital screen replaces a conventional mirror. Ex. 1021 ('109 Patent) at 14:12-15. The '109 Patent explains "it may be too complex or expensive to place the camera behind the display at the eye level," and indicates an "ideal system" has "a fixed-mounted camera/cameras located at the perimeter of the screen and facing the user." Ex. 1021 ('109 Patent) at 15:4-9.

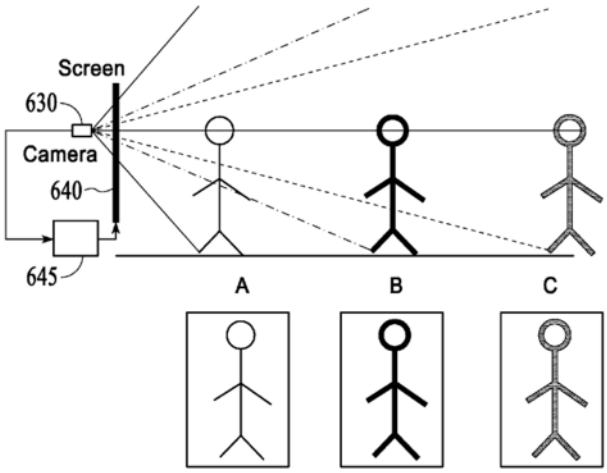


FIG. 6

Ex. 1021 (’109 Patent) at Fig. 6.

The ’109 Patent also explains that transformation mapping “essentially reassigns new addresses to each pixel of the original image.” Ex. 1021 (’109 Patent) at 29:28-30. The ’109 Patent discloses several well-known exemplary virtual image transformations, including “angle, tilt, azimuth, scale, [and] spatial translation” (Ex. 1021 (’109 Patent) at 29:38-42) that “effectively mimic a mirror reflection” (Ex. 1021 (’109 Patent) at 24:4-7). *See also* Ex. 1021 (’109 Patent) at 23:41-24:3. The ’109 Patent further states that a “final transformation is a matrix multiplication of the individual transformation.” Ex. 1021 (’109 Patent) at 29:40-42.

78. According to the ’109 Patent, a tilt transformation corrects distortions resulting from a downwards-facing or downwards-tilted camera by “modif[ying] an input image taken with the camera looking downwards and transform[ing] the image

to appear as if the camera was pointed straight horizontally." Ex. 1021 ('109 Patent) at 23:45-51. Additionally, as described in the '109 Patent, an elevation transformation corrects distortions from a camera positioned externally to a display screen and modifies the image such that it appears to have been taken from the display screen at a user's eye-level. Ex. 1021 ('109 Patent) at 23:53-58. The '109 Patent further describes that a scale transformation "reduces the effect of size change with distance to the screen, so that the user sees almost constant size of himself in the screen, regardless of distance to the screen." Ex. 1021 ('109 Patent) at 23:65-24:3.

79. Additional exemplary image processing manipulations disclosed in the '109 Patent include an adaptive field of view ("FOV") used "to compensate for the changes in the distances between the user and the screen, so that the user will see his image at the same size, just like in a mirror." Ex. 1021 ('109 Patent) at 14:29-32. According to the '109 Patent, a system creates "an adaptive, i.e., continuously variable, FOV" by manipulating an image resolution or controlling in real-time a camera's zoom and focus (based on tracked movements of the user). Ex. 1021 ('109 Patent) at 14:34-39.

80. Another exemplary image manipulation disclosed in the '109 Patent is flipping, via the image processor, a camera-generated image across the vertical access, to mimic the "right-to-left behavior" of a mirror. Ex. 1021 ('109 Patent) at

14:40-41. One exemplary mechanism disclosed in the ’109 Patent that may be used to achieve this image flipping is manipulation of the image’s pixels: “[t]his image transformation may be performed relatively easily by manipulating the pixels addresses of a digital image.” Ex. 1021 (’109 Patent) at 14:41-43. The ’109 Patent additionally discloses an adaptive point of view (“POV”) “to fit the view point that would occur with a conventional mirror.” Ex. 1021 (’109 Patent) at 15:48-52.

81. The ’110 Patent discloses a range of use cases, including clothes fitting, virtual dressing, a video conference device over a network, remotely sharing outfits while shopping, and a gaming system. Ex. 1001 (’110 Patent) at 7:31-55 (“can be used to create reliable virtual dressing”), 13:55-14:16 (“can create a more convincing fitting”), 23:8-9, 23:50-57 (“main experience is to see the user who is standing in front of the mirror and is doing something for review”), 29:56-63, 1:44-51, 8:16-26. The ’109 Patent discuses that embodiments of the disclosed system may include a variety of locations, such as “clothing stores, shoe stores, points of sale, concept showrooms, exhibitions, shopping malls, eyewear stores, cosmetic stores, sport clubs, health institutes, fitness centers, airports, train stations, coffee shops, restaurants, hotels, homes, and the like.” Ex. 1021 (’109 Patent) at 8:33-40.

82. Further, the ’110 Patent discloses “further enhancement features” for a “virtual mirror” image captured from a camera, including image manipulations to change colors, enable multi-level user identification and authentication, and provide

body measurements. Ex 1001 ('110 Patent) at 5:4-9. The '110 Patent also discloses that these enhancement features include a calibrated camera "capable of providing augmented reality features, for example, color, texture, shape manipulation of body items or background in, e.g., a virtual mirror or video conference applications." Ex ('110 Patent) at 5:9-15. The '110 Patent further discusses applying a mapping transformation to data from a camera to generate an avatar of a user. Ex ('110 Patent) at Abstract, 21:5-59. A 3D sensor is contemplated to generate "depth measurement" and "to control a parametric model of the user's avatar." Ex ('110 Patent) at 21:13-15.

83. In my analysis below, I will discuss additional aspects of the '110 Patent, as well as the '109 Patent, which is incorporated by reference into the '110 Patent.

B. THE CHALLENGED CLAIMS OF THE '110 PATENT

84. The Challenged Claims (claims 1-18) are set forth below. Some of the claim elements have been further sub-divided for ease of reference.

a. Claim 1

1[pre]	A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising:
1[a]	operating the camera to generate a stream of images of the user;
1[b]	operating the processor to detect the presence of the user in the stream of images;
1[c]	operating the processor to apply adaptive transformation mapping to the

	stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view; and,
1[d]	displaying the modified images on the monitor;
1[e]	wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and
1[f]	wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.

b. <u>Claim 2</u>

2[pre]	The method of claim 1,
2[a]	wherein the mapping transformation comprises a combination of at least two of the following transformations:
2[b]	vertical flips about a vertical axis, a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image.

c. <u>Claim 3</u>

3[pre]	The method of claim 1,
3[a]	wherein during idle mode the camera continuously sends captured images to the processor, and
3[b]	the user presence is identified by the processor detecting a body part in the images.

d. <u>Claim 4</u>

4[pre]	The method of claim 3,
4[a]	wherein the body part comprises one of legs, face, eyes, and hands.

e. Claim 5

5[pre]	The method of claim 1,
5[a]	further comprising receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input.

f. Claim 6

6[pre]	The method of claim 5,
6[a]	wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images.

g. Claim 7

7[pre]	The method of claim 1,
7[a]	further comprising operating the processor to generate an avatar having body characteristics corresponding to body characteristics of a user appearing in the images.

h. Claim 8

8[pre]	The method of claim 7,
8[a]	further comprising operating the processor to display on the monitor the modified images and an image of the avatar.

i. Claim 9

9[pre]	The method of claim 7,
9[a]	further comprising operating the processor to augment the modified images by overlaying virtual objects on the avatar.

j. Claim 10

10[pre]	The method of claim 1,
10[a]	further comprising operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images.

k. Claim 11

11[pre]	The method of claim 1,
11[a]	further comprising operating the processor to transmit the modified image over a network.

l. Claim 12

12[pre]	The method of claim 1,
12[a]	further comprising displaying the modified image on a remote screen for video telepresence.

m. Claim 13

13[pre]	The method of claim 1,
13[a]	further comprising providing data from a depth sensor to the processor, and operating the processor to measure a texture of the image from the data.

n. Claim 14

14[pre]	The method of claim 13,
14[a]	further comprising operating the processor to generate an avatar using the data from the depth sensor.

o. Claim 15

15[pre]	The method of claim 13,

15[a]	wherein the processor generate the avatar by analyzing registration pointers on the images of the camera.

p. <u>Claim 16</u>

16[pre]	The method of claim 1,
16[a]	wherein the processor operates on the images from the camera to generate background segmentation, and
16[b]	replaces segments on the images by replacement segments.

q. <u>Claim 17</u>

17[pre]	The method of claim 16,
17[a]	wherein the replacement segments comprise virtual background or background that was captured in a different setting.

r. <u>Claim 18</u>

18[pre]	The method of claim 1,
18[a]	wherein the processor resides in the remote location, and the method proceeds by transmitting the images from the camera over a network to the processor, and transmitting the modified images back from the processor to the monitor over the network.

C. <u>PROSECUTION HISTORY OF THE '110 PATENT</u>

85. The application of the '110 Patent was filed on April 15, 2014, and claims priority to a chain of applications. Specifically, the '110 Patent claims priority to three provisional patent applications: U.S. Provisional Application No. 61/738,957, filed on December 18, 2012; U.S. Provisional Application No.

61/862,025, filed on Aug. 4, 2013; and U.S. Provisional Application No. 61/892,368, filed on Oct. 17, 2013.

86. Further, the application of the ’110 Patent is a continuation-in-part of U.S. Application No. 13/843,001, filed on March 15, 2013, which is the application that led to the ’109 Patent, and which is a continuation-in-part of U.S. Application No. 13/088,369, filed on April 17, 2011, which is a continuation of U.S. Application No. 11/817,411, which was a National Phase Application of PCT International Application No. PCT/IL06/000281 having an International Filing Date of March 1, 2006, and which issued as U.S. Patent No. 7,948,481 on May 24, 2011, which claims priority to U.S. Provisional Application No. 60/656,884 and U.S. Provisional Application No. 60/656,885, both filed on March 1, 2005.

87. The application for the ’110 Patent did not receive any rejections and was allowed on December 3, 2014. In the Notice of Allowability of the Notice of Allowance, the Examiner noted that the Examiner had entered an Examiner’s amendment in which the subject matter of two now-cancelled dependent claims was incorporated into claim 1. Ex. 1002 (’110 Patent File History) at 333. The limitations added via the Examiner’s amendment require the processor of the method recited in claim 1 to apply adaptive transformation mapping to a stream of captured images to “generate modified images that appear to be captured from a different point of view of the camera’s actual point of view” such that the “different point of view

corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera." Ex. 1002 ('110 Patent File History) at 333. Additionally, the added limitations require that during a "mirror tracking mode," the processor "utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera." Ex. 1002 ('110 Patent File History) at 333.

88. The '110 Patent issued on March 17, 2015.

VIII. PRIORITY DATE AND LEVEL OF ORDINARY SKILL IN THE ART

89. As noted above, I understand that the '110 Patent issued from U.S. Patent Application No. 14/253,831, filed on April 15, 2014, which claims priority to U.S. Provisional Application No. 61/862,025 (filed Aug. 4, 2013) and U.S. Provisional Application No. 61/892,368 (filed Oct. 17, 2013) and is a continuation-in-part of U.S. Application No. 13/843,001, filed March 15, 2013, which led to the '109 Patent. The application of the '109 Patent claims priority to U.S. Provisional Application No. 61/738,957, filed December 18, 2012 and is a continuation-in-part of U.S. Application No. 13/088,369, filed April 17, 2011, which claims priority, through a chain of continuation and continuation-in-part, to U.S. Provisional

Application No. 60/656,884 and U.S. Provisional Application No. 60/656,885, both filed on March 1, 2005.

90. I have been informed by counsel that Patent Owner ("PO") has only claimed priority based on U.S. Provisional Application No. 61/738,957, filed December 18, 2012 (the "2012 Provisional Application") in a prior district court litigation. (*see* Ex. 1018 (Excerpts from Patent Owner's Infringement Contentions in the Meta Litigation) at p. 3), and when determining the level of ordinary skill in the art in a prior IPR (Ex. 1030 (Patent Owner's Preliminary Response in the Meta IPR2022-00490) at 8). Further, I have reviewed the 2012 Provisional Application (Ex. 1023), as well as the above-listed provisional patent applications predating the 2012 Provisional Application (Exs. 1024-1026) and confirm that at least the limitations "to generate modified images that appear to be captured from a different point of view" and "wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera" do not have supporting disclosure prior to the filing date of the 2012 Provisional Application. Therefore, for purposes of my analysis, I have considered the priority date of the '110 Patent to be December 18, 2012, and have considered the relevant level of ordinary skill in the art at that time.

91. In my opinion, a POSITA as of December 18, 2012, would have had a bachelor's degree in electrical or computer engineering, or a closely related scientific

field such as computer science, and two years of work experience with image capture and processing systems or related fields. A person with less and/or different education but more relevant practical experience, or vice versa, may also meet this standard. The prior art also evidences the level of skill in the art.

IX. CLAIM CONSTRUCTION

92. In my opinion, and for the purposes of this declaration, the claim terms of the ’110 Patent do not require a special construction. I have applied the ordinary and customary meaning of the claims.

X. OVERVIEW OF THE PRIOR ART

A. HAAN (EX. 1005)

a. Status of Haan as Prior Art

93. “Haan” refers to United States Patent Application Publication No. 2011/0298910, entitled “Mirror device.” Haan names Gerard De Haan, Frederik Jan De Bruijn, and Wilhelmus PETRUS Van Der Meer as the inventors and was assigned to Koninklijke Philips Electronics, N.V. Haan was filed in the United States on February 15, 2010, and was published December 8, 2011. Thus, Haan was published more than one year prior to the priority date of the ’110 Patent, which I understand to be December 18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Haan qualifies as prior art to the ’110 Patent.

b. <u>Overview of the Disclosure of Haan</u>

94. Haan is directed to a "mirror device" that "(1) comprises a display device (3) and a camera (2) to record an image" of a user. Ex. 1005 (Haan) at Abstract. Haan further describes this mirror device as comprising a means for determining the position of a user's body part (e.g., the user's head) with respect to the display screen, and based on the user's position, the mirror device modifies "the view point from which the to be displayed image is recorded." Ex. 1005 (Haan) at Abstract; *see also* Ex. 1005 (Haan) at [0006]-[0011], [0013], [0024], [0056]-[0058], [0071].

95. Haan recognizes that "[m]irror devices have been know [sic] for a very long time, in fact since Roman times and before" (Ex. 1005 (Haan) at [0002]) and that "[a] frequently occurring problem with mirrors is that the area one wants to see is just out of sight." Ex. 1005 (Haan) at [0004]. According to Haan, this problem can occur when, for example, a user gets closer to a mirror and "the particular [body] part gets out of sight and/or one cannot get the part in focus" or when a user "can turn the head to get a better view, but often, to the aggravation of the viewer, as the head is turned to get a good look at the part, the particular part moves out of sight." Ex. 1005 (Haan) at [0004].

96. Haan therefore discloses techniques for determining the orientation of a user in relation to the mirror device and identifying specific body parts of the user

to enable a user to modify the point of view of displayed images and display those areas that are "often just out of sight." Ex. 1005 (Haan) at [0004]; *see also* Ex. 1005 (Haan) at [0009]-[0010]. For example, Figure 3 of Haan shows a modified image 32 of a user's side/profile, while the user looks towards the mirror in image 31. Ex. 1005 (Haan) at [0004], Fig. 3.

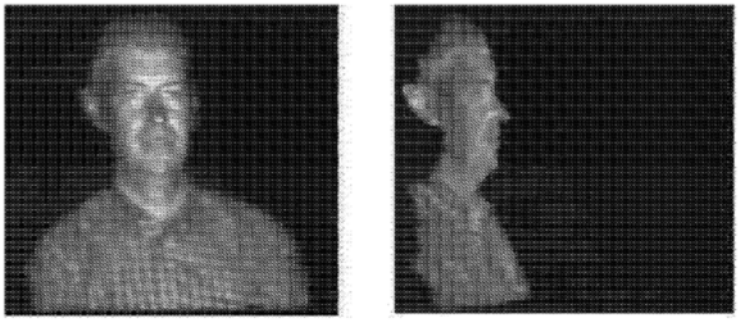


FIG. 3

Ex. 1005 (Haan) at Fig. 3.

97. As shown in Fig. 2, Haan discloses a display device, one or more cameras, and an analysis means and a synthesis means "embodied in a computer program." Ex. 1005 (Haan) at Abstract, Fig. 2, [0058], [0087].

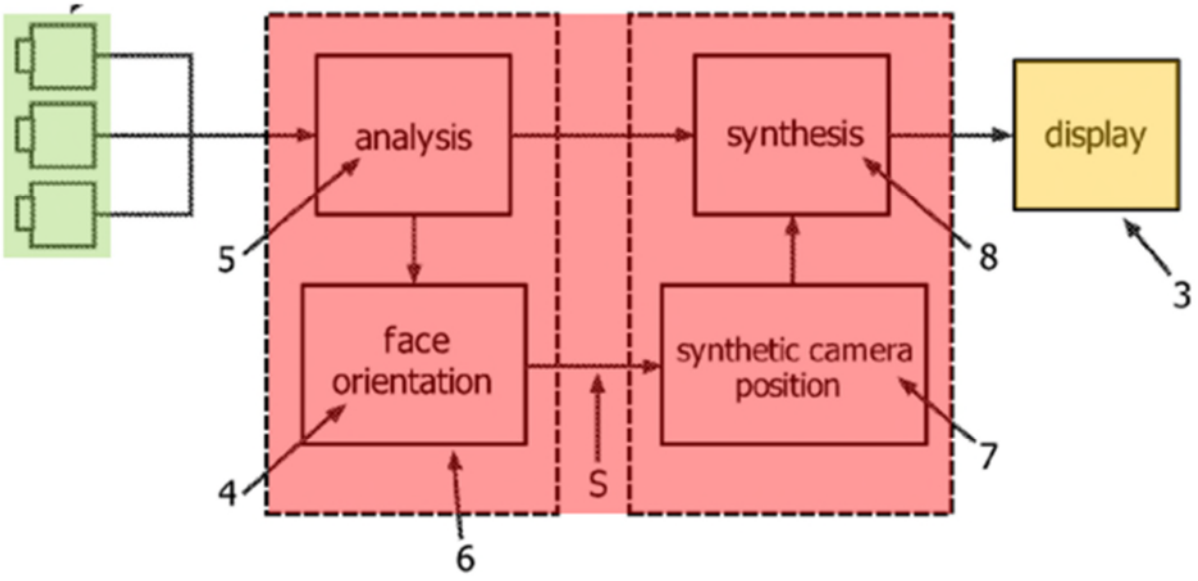


FIG. 2

Ex. 1005 (Haan) at Fig. 2 (partial, annotated).

98. Like the ’110 Patent, Haan’s analysis module determines the position of a user’s body part relative to the display screen. The analysis module disclosed in Haan includes a “position determination means,” into which the images recorded by one or more cameras are fed to determine the position of a user’s body part (e.g., a head or torso) with respect to the mirror device’s display screen, and the positions of the user’s body parts with respect to each other. Ex. 1005 (Haan) at [0058], Fig. 2, claim 2. Specifically, the position determination means performs a 3D analysis of the recorded images “to determine the angle of rotation of the head in respect of the mirror device and to determine the distance to the display screen.” Ex. 1005 (Haan) at [0058]. Further, Haan discloses that the position determination means can analyze

a recorded image to detect specific body parts (e.g., eyes, mouth, and ears) and can calculate the distances between these body parts to determine the orientation of a user's head. Ex. 1005 (Haan) at [0058]. Haan also discloses that an ultrasound means may be used for sensing orientation and determining distances. Ex. 1005 (Haan) at [0067], [0074]; *see also* Ex. 1005 (Haan) at Fig 2.

99. Similar to the '110 Patent, Haan discloses that the mirror device performs image processing to modify and display camera-recorded images, based on a user's position relative to the mirror device. Haan discloses that "by synthesis of the various images," the mirror device, via its synthesis module, can "construct a virtual camera position" (i.e., the "virtual viewpoint") and/or the magnification of the user to be displayed. Ex. 1005 (Haan) at [0058]; *see also* Ex. 1005 (Haan) at [0006]-[0011], [0024], [0056]-[0057], [0071], [0073]. Haan recognizes that "magnification is…equivalent to moving a camera with fixed magnification closer to the object, so all applications can be seen, in fact as changing the point of view" of a real or virtual camera. Ex. 1005 (Haan) at [0073]; *see also* Ex. 1005 (Haan) at [0074]-[0076]. Haan similarly discloses that a camera's field of view changes depending on the user's distance from the mirror device. Ex. 1005 (Haan) at [0097]. Therefore, Haan discloses that the synthesis module changes the view point of captured images based on a user's determined position.

100. In addition to providing mirror images at modified points of view and magnifications, Haan discloses that the mirror device can further augment recorded images. For example, Haan can overlay the recorded images with makeup, hairstyles, and clothing of various colors and styles. The user therefore can view the reflection at various angles and magnifications. Ex. 1005 (Haan) at [0113]-[0118].

101. Therefore, Haan teaches a mirror device that provides users a reliable virtual reflection of themselves, along with enhanced viewing features including, for example, modified points of view and magnifications depending on the user's distance and motions in relation to the mirror device, as well as augmented features (e.g., overlaid makeup, hairstyles, and clothing), solving the known problem identified in the '110 Patent and offering a similar solution.

B. FRANCOIS (EX. 1009)

a. Status of Francois as Prior Art

102. "Francois" refers to a conference paper publication by Alexandre R.J. Francois, entitled "Software Architecture for Computer Vision: Beyond Pipes and Filters." Francois was published in the volume titled *IRIS: Selected Papers of the Information Systems Research Seminar in Scandinavia*. Ex. 1029 ("Hall-Ellis Declaration") at ¶40. I understand that Francois was published and publicly available on July 1, 2003. Ex. 1029 ("Hall-Ellis Declaration") at ¶43. This was well over one year before the priority date of the '110 Patent, which I understand to be December

18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Francois qualifies as prior art to the '110 Patent.

b. <u>Overview of the Disclosure of Francois</u>

103. Francois discloses a software architecture model called Software Architecture for Immersipresence ("SAI") for performing distributed, asynchronous parallel processing of data streams, with a focus on computer vision applications. Ex. 1009 (Francois) at §1.2-1.3. Francois recognizes the "development of real-time video analysis applications" as a "steering concern during the development" of SAI. Ex. 1009 (Francois) at §4.1. Francois describes four real-time functions performed by SAI for video analysis: "capture, segmentation, tracking, visualization." Ex. 1009 (Francois) at §4.1, Fig. 7:

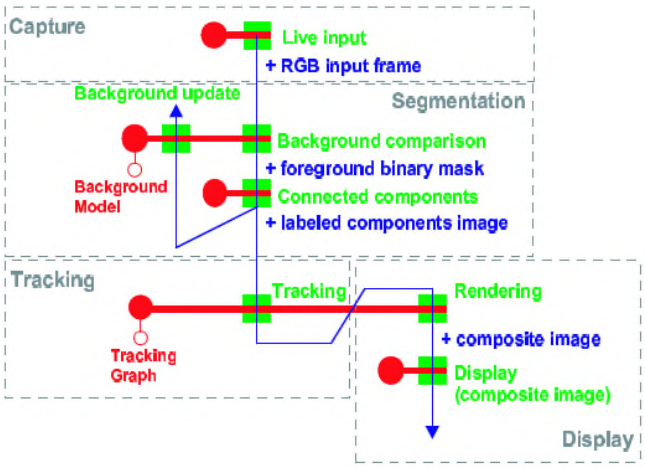


Figure 7: Conceptual graph for real-time color background model-based segmentation and multi-resolution graph tracking application.

Ex. 1009 (Francois) at Fig. 7.

104. Francois discloses a virtual mirror prototype implementing SAI: "[i]n order to simulate a real mirror on a computer screen, images of the observed world, consistent with the viewer's position, must be synthesized and displayed in real-time." Ex. 1009 (Francois) at §4.3. Francois notes that the virtual mirror prototype comprises a user-manipulated LCD screen, camera capturing a continuous video stream, and a tracking device, wherein the "synthesized video stream is a close approximation of what the user would see on the screen surface if it were a real mirror." Ex. 1009 (Francois) at §4.3, Fig. 12:

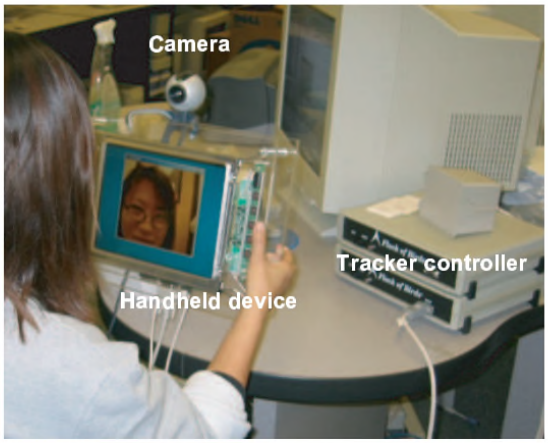


Figure 12: The handheld mirror simulation system ("Virtual Mirror").

Ex. 1009 (Francois) at Fig. 12.

105. Francois describes using a "tracker input cell" to "add[] position and orientation data to the stream" and a "mirror transform cell" to "use the synchronous image and tracking data to synthesize the mirror simulation image." Ex. 1009 (Francois) at §4.3, Fig. 13:

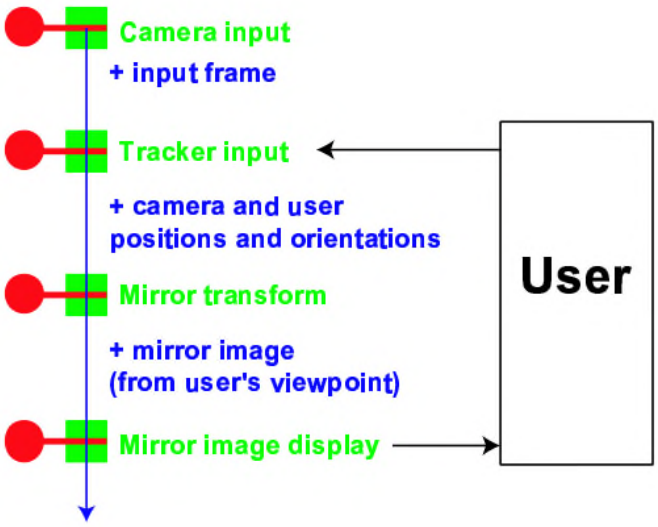


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 13.

106. Further, Francois describes using the SAI model to augment the video streams, enabling real-time video manipulation based on user input: "[t]he (virtual) camera used for rendering can be manipulated in real-time (e.g. with a gamepad or through a head tracking device), making the system interactive." Ex. 1009 (Francois) at §4.5. As Francois details, these manipulations may include adding animated 3D graphics to video streams, or performing background segmentation and substitution. Ex. 1009 (Francois) at §§4.4-4.5, Fig. 15:

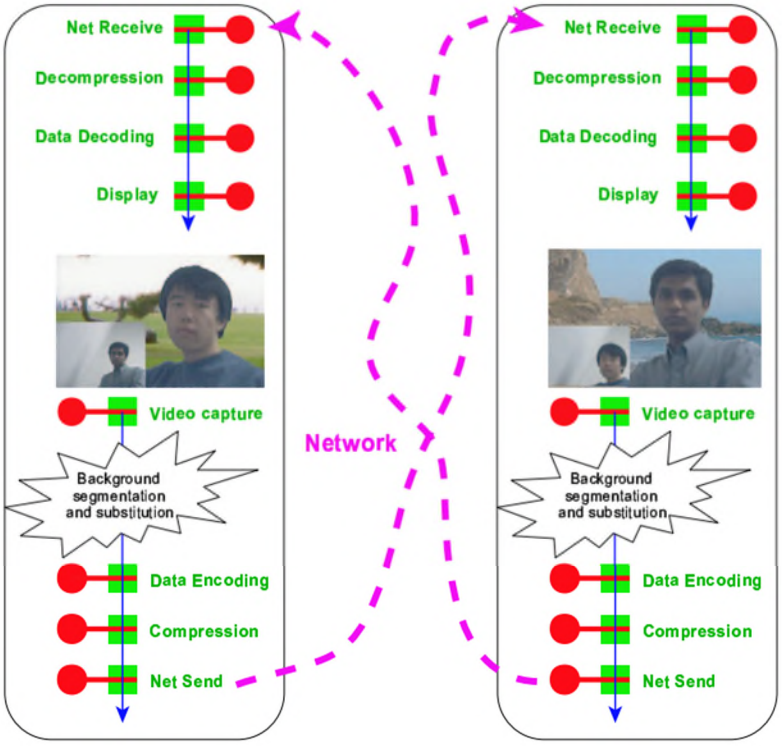


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15.

107. Therefore, Francois teaches a software architecture framework enabling the creation of a virtual mirror system, solving the known problem identified in the '110 Patent and offering a similar solution.

C. ROSENBERG (EX. 1006)

a. Status of Rosenberg as Prior Art

108. "Rosenberg" refers to United States Patent Application Publication No. 2007/0040033, entitled "Digital mirror system with advanced imaging features and hands-free control." Rosenberg names Louis B. Rosenberg as the inventor and was assigned to Outland Research. Rosenberg was filed in the United States on

September 26, 2006, and was published on February 22, 2007. Thus, Rosenberg was published more than one year prior to the priority date of the ’110 Patent, which I understand to be December 18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Rosenberg qualifies as prior art to the ’110 Patent.

b. <u>Overview of the Disclosure of Rosenberg</u>

109. Rosenberg is directed to a “digital mirror system [] adapted to provide real-time captured images that approximate what a user would see when viewing an ordinary mirror, with the addition of various other image processing features.” Ex. 1006 (Rosenberg) at [0006]. Similar to the ’110 Patent, Rosenberg describes techniques for adapting a flat panel display to display images that mimic the behavior of a conventional mirror. Ex. 1006 (Rosenberg) at [0006], [0077]; *see also* Ex. 1006 (Rosenberg) at [0062], Figs. 3, 5.

110. The digital mirror system disclosed in Rosenberg includes a camera, a display, and “a processor for adjusting and/or modifying the images that are displayed upon the screen to achieve unique mirroring features and functions not provided by a standard household wall mirror” as well as “access to one or more communication networks via a network connection.” Ex. 1006 (Rosenberg) at [0074]; *see also* Ex. 1006 (Rosenberg) at Fig. 13:

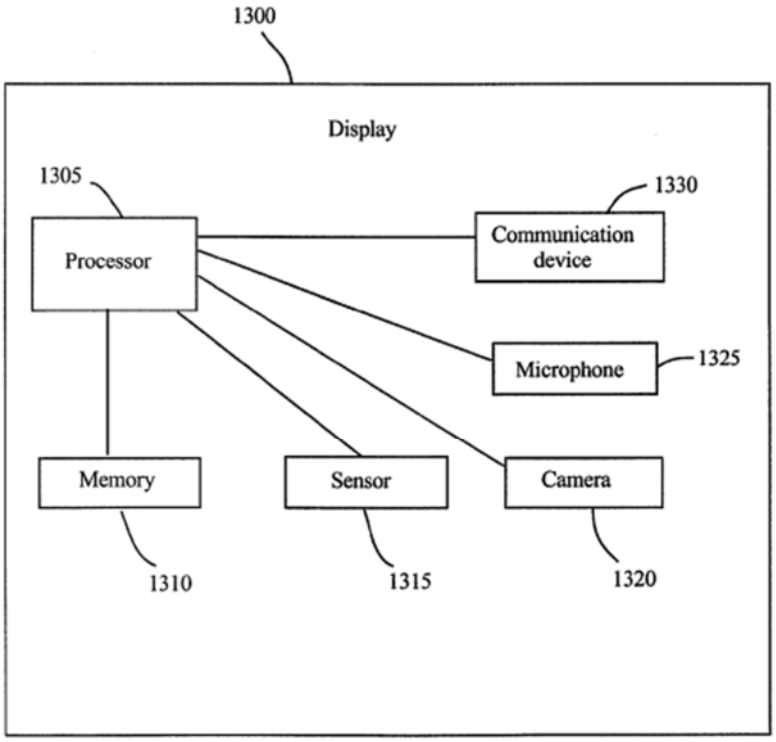


Fig. 13

Ex. 1006 (Rosenberg) at Fig. 13.

Similar to the ’110 Patent, Rosenberg discloses that the “camera is mounted onto the digital mirror display.” Ex. 1006 (Rosenberg) at [0006]. Rosenberg also discloses a hands-free sensor that detects user input and that a “processor controls the camera to capture the at least one image in response to the user input.” Ex. 1006 (Rosenberg) at [0006]; *see also* Ex. 1006 (Rosenberg) at [0007], [0035], [0074], Figs. 5, 8.

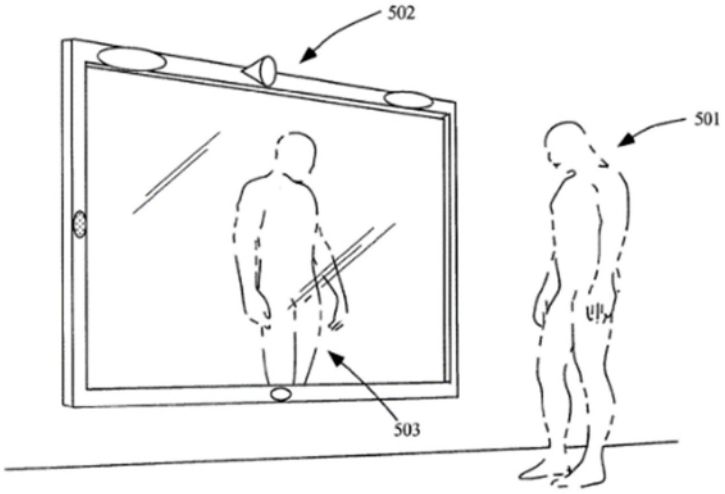


Fig. 5

Ex. 1006 (Rosenberg) at Fig. 5.

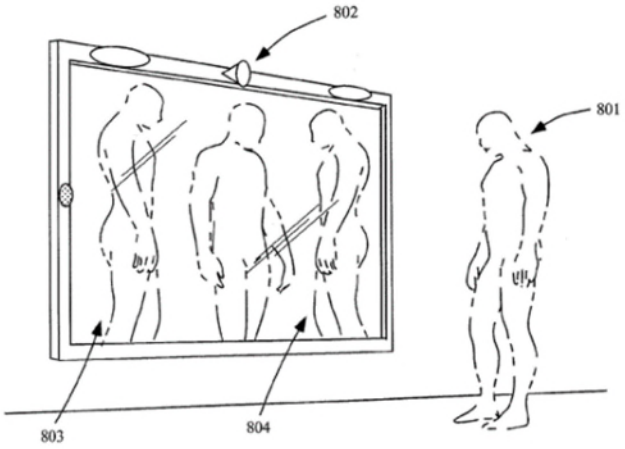


Fig. 8

Ex. 1006 (Rosenberg) at Fig. 8.

111. Additionally, similar to the disclosure of the ’110 Patent, Rosenberg discloses an embodiment in which “the camera is mounted directly above the screen area of the display” and notes that this configuration may be “non-ideal because it

provides the user with a view of himself or herself from a somewhat unnatural perspective as compared to a traditional mirror." Ex. 1006 (Rosenberg) at [0027]. Rosenberg also discloses another embodiment, wherein "the camera is mounted within the screen area of the display itself, approximately at eye-level of users." Ex. 1006 (Rosenberg) at [0027]; *see also* Ex. 1006 (Rosenberg) at [0061]. Rosenberg notes that "an image that is from a vantage point perfectly straight ahead" is often preferred "because it yields a more realistic replication of a true mirror image." Ex. 1006 (Rosenberg) at [0028], *see also* Ex. 1006 (Rosenberg) at [0075].

112. Rosenberg additionally describes that the digital mirror detects the presence of a user, as well as the distance between the user and the mirror system via a "proximity and/or motion sensor for detecting when a user is physically present before and/or approaches and/or stands before the display screen of the mirror within some proximity range." Ex. 1006 (Rosenberg) at [0025].

113. Additionally, similar to the '110 Patent, the digital mirror system described in Rosenberg enables users to view "a mirror image of themselves (i.e., a left-right inverted representation) as is displayed by a traditional mirror." Ex. 1006 (Rosenberg) at [0025]. As described in Rosenberg, the digital mirror system outputs "real-time streamed video images being a left-right inverted representation of the video signal captured by the camera" that is generated "by flipping the pixels of the image around the vertical centerline such that right side of the user, from the user's

perspective, appears upon the right side of the display, and the left side of the user, from the user's perspective, appears upon the left side of the display." Ex. 1006 (Rosenberg) at [0026].

114. Furthermore, similar to the '110 Patent, the digital mirror system described in Rosenberg transforms and/or resizes the real-time images captured by the camera, such that a real-time video image "is then displayed upon the wall mounted screen of the digital mirror 502 with a displayed size and scaling such that it appears like a traditional life-size mirror image 503." Ex. 1006 (Rosenberg) at [0077].

115. Therefore, Rosenberg teaches a digital mirror system adapted to provide inverted and scaled real-time mirror images of a user and offers similar technical solutions as those described in the '110 Patent.

D. PARK (EX. 1011)

a. Status of Park as Prior Art

116. "Park" refers to United States Patent Application Publication No. 2007/0065039, entitled "Image capturing apparatus with image compensation and method thereof." Park names Tae-suh Park, Yeun-bae Kim, and Min-kyu Park as the inventors and was assigned to Samsung Electronics Co., Ltd. Park was filed in the United States on September 22, 2006, and was published on March 22, 2007. Thus, Park was published more than one year prior to the priority date of the '110

Patent, which I understand to be December 18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Park qualifies as prior art to the ’110 Patent.

b. <u>Overview of the Disclosure of Park</u>

117. Park is directed to “an image capturing apparatus with image compensation and method thereof, which can perform compensation of a user’s facial image that is distorted in accordance with the position of a camera relative to an imaged object.” Ex. 1011 (Park) at [0003]. Similar to the ’110 Patent, the image capturing apparatus disclosed in Park includes an image capturing unit (i.e., a camera), a compensation unit (i.e., a processor), and a display screen. *See* Ex. 1011 (Park) at Figs. 2, 9.

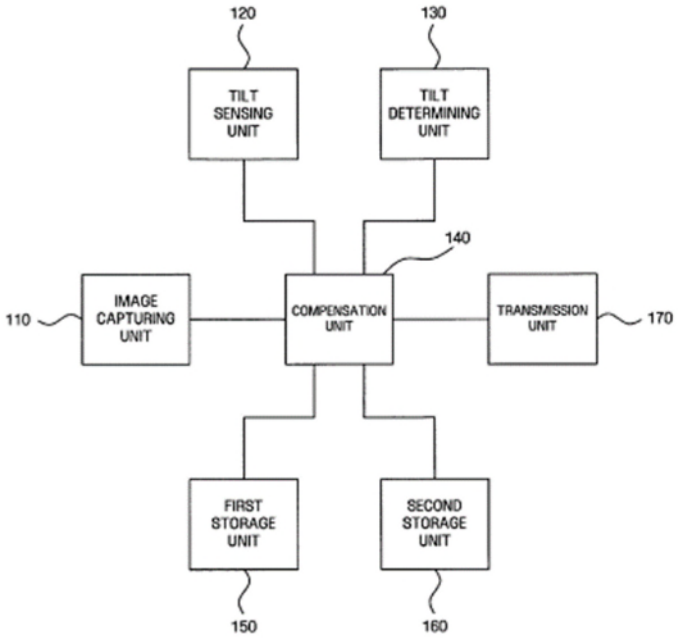


FIG. 2

Ex. 1011 (Park) at Fig 2.

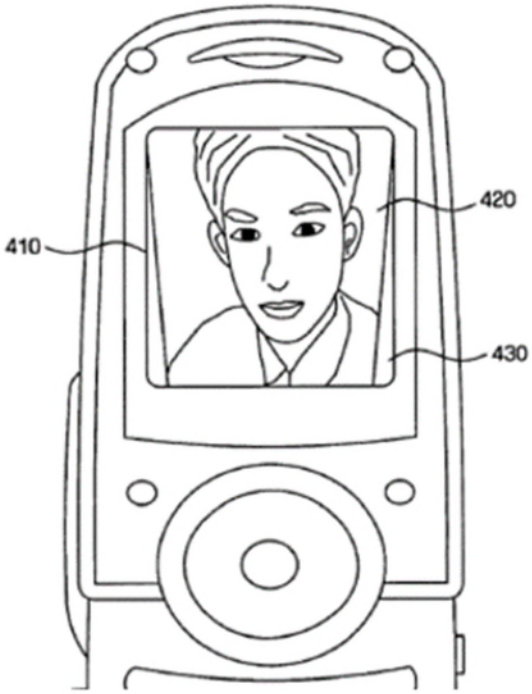


FIG. 9

Ex. 1011 (Park) at Fig. 9 (cropped).

118. Similar to the ’110 Patent, the image capturing apparatus described in Park also includes a transmission unit that “may transmit the image through a wired network” or “a wireless network.” *See* Ex. 1011 (Park) at [0063], Fig. 2.

119. Further, in one embodiment, Park describes that an image capturing apparatus with image compensation “can obtain an image desired by a user by determining the tilt status of a mobile communication device that captures the image of an object and performing compensation of the captured image in accordance with the determined tilt status of the device during a video telephony using the mobile communication device.” Ex. 1011 (Park) at [0010]. As Park describes, “[i]f the

calculated tilt is determined to be an angle that causes a perspective distortion of the facial image that is not desired by the user, the compensation unit 140 compensates for the generated distortion." Ex. 1011 (Park) at [0050]; *see also* Ex. 1011 (Park) at Fig. 5.

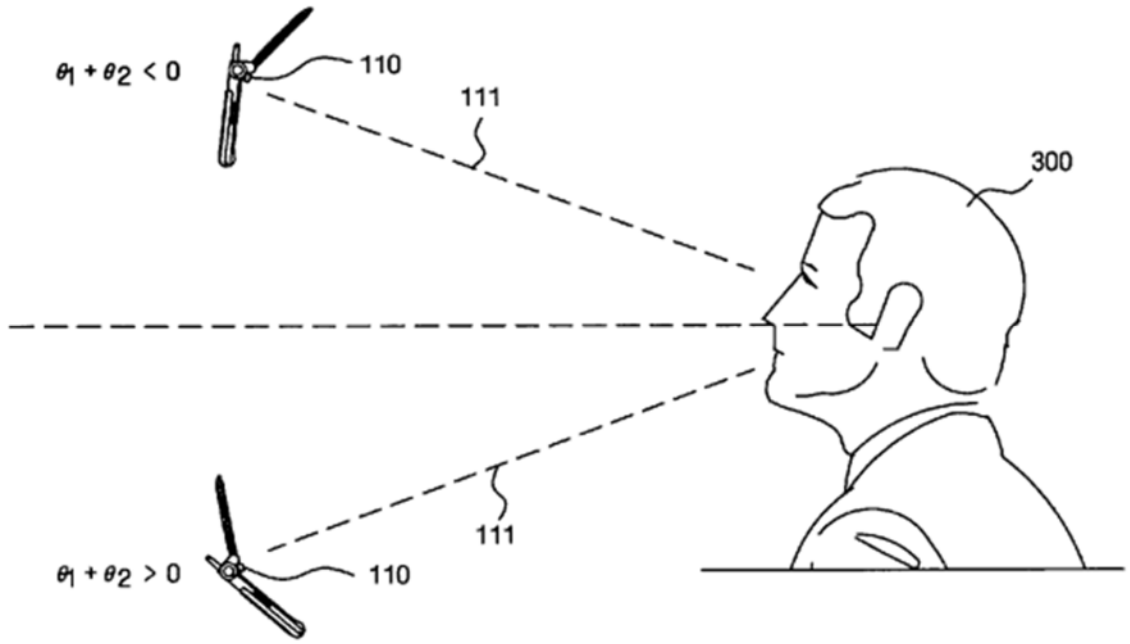
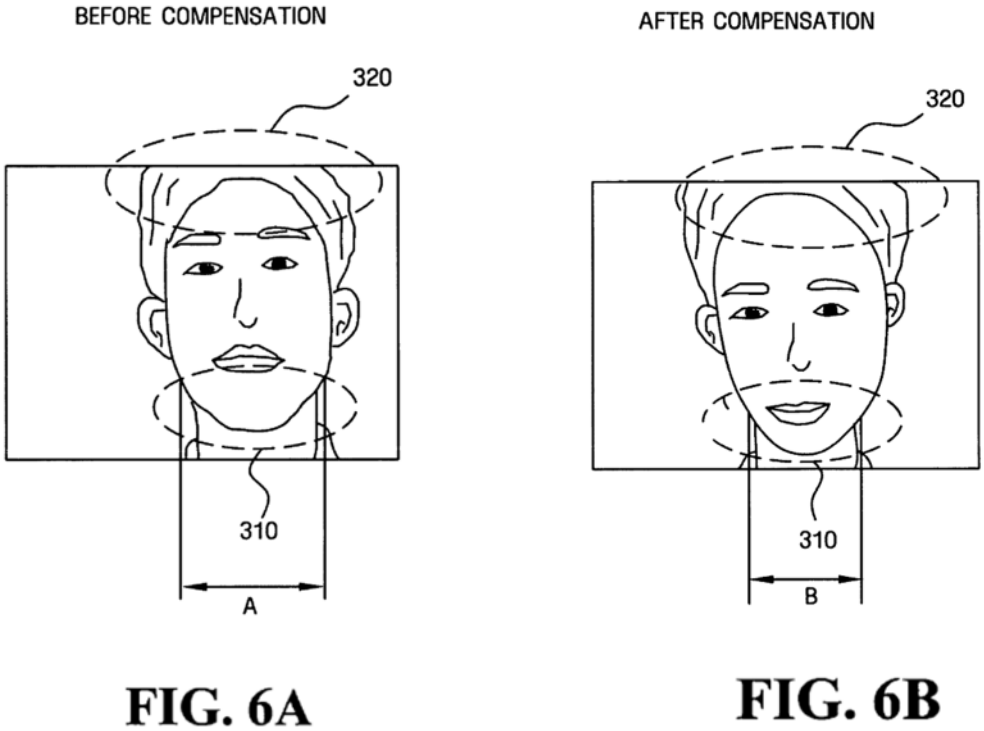


FIG. 5

Ex. 1011 (Park) at Fig. 5.

120. Park teaches performing compensation based on the location of the camera in relation to the user. Ex. 1011 (Park) at [0051]. For example, Park discloses that "if the image capturing unit 110 that is located below the user's face has captured the user's facial image, as illustrated in Fig. 6A, the captured image may be a distorted image in which the chin part 310 appears to be emphasized and the

forehead part 320 appears to be narrowed." Ex. 1011 (Park) at [0051]. To correct this distortion, the processor "performs perspective compensation on the image as illustrated in FIG. 6A to provide a compensated image as illustrated in FIG. 6B." Ex. 1011 (Park) at [0051]; *see also* Ex. 1011 (Park) at Figs. 6A, 6B.



Ex. 1011 (Park) at Fig. 6A (left), Fig. 6B (right).

121. Park also teaches that although "in the description of an embodiment [], an image capturing apparatus with image compensation is descried [*sic*] as being applied to devices such as a camera and mobile communication device (e.g., a mobile phone or a personal digital assistant (PDA)) having a camera attached thereto […] it is to be understood that these are merely non-limiting examples, and that the

apparatus of the present embodiment may also be applied to, for example, a camera connected to a personal computer." Ex. 1011 (Park) at [0034].

122. In describing methods of image correction, or image compensation, Park further describes a pixel filling process. *See, e.g.*, Ex. 1011 (Park) at [0056] ("the empty region is filled with pixel values of the compensated image existing at the closest position…").

E. GEISNER (EX. 1007)

a. Status of Geisner as Prior Art

123. "Geisner" refers to United States Patent Application Publication No. 2011/0246329, entitled "Motion-based interactive shopping environment." Geisner names Kevin A. Geisner, Kudo Tsunoda, Darren Bennett, Brian S. Murphy, Stephen G. Latta, Relja Markovic, and Alex Kipman as the inventors and was assigned to Microsoft Corporation. Geisner was filed in the United States on April 1, 2010, and was published October 6, 2011. Thus, Geisner was published more than one year prior to the priority date of the '110 Patent, which I understand to be December 18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Geisner qualifies as prior art to the '110 Patent.

b. Overview of the Disclosure of Geisner

124. Geisner is directed to an "on-screen shopping application which reacts to a human target user's motions to provide a shopping experience to the user,"

including a "tracking system [that] captures user motions and executes a shopping application allowing a user to manipulate an on-screen representation the user." Ex. 1007 (Geisner) at Abstract. Geisner discloses that the tracking system may be connected to a display (Ex. 1007 (Geisner) at [0029]) and may include a processor, and a capture device (Ex. 1007 (Geisner) at [0027], [0028]) to monitor one or more users "such that gestures and/or movements performed by the one or more users may be captured, analyzed, and tracked to perform one or more controls or actions within the application and/or animate an avatar or on-screen character" (Ex. 1007 (Geisner) at [0028]). Geisner further discloses that the capture device can be a set of two or more cameras to capture "a scene from different angles" and "generate depth information" (Ex. 1007 (Geisner) at [0039]). *See also* Ex. 1007 (Geisner) at Figs. 1, 4.

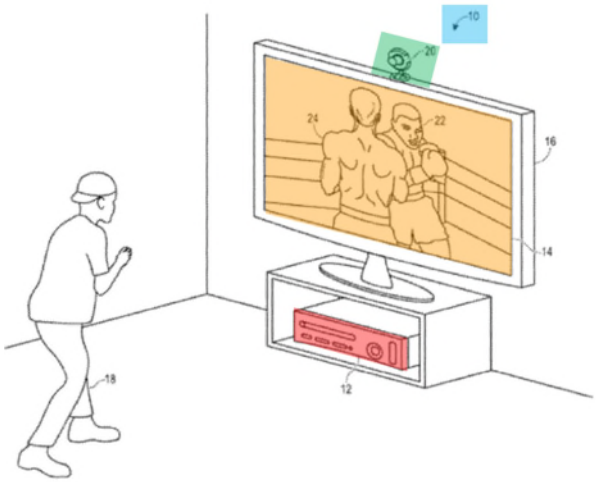


FIG. 1

Ex. 1007 (Geisner) at Fig. 1 (annotated).

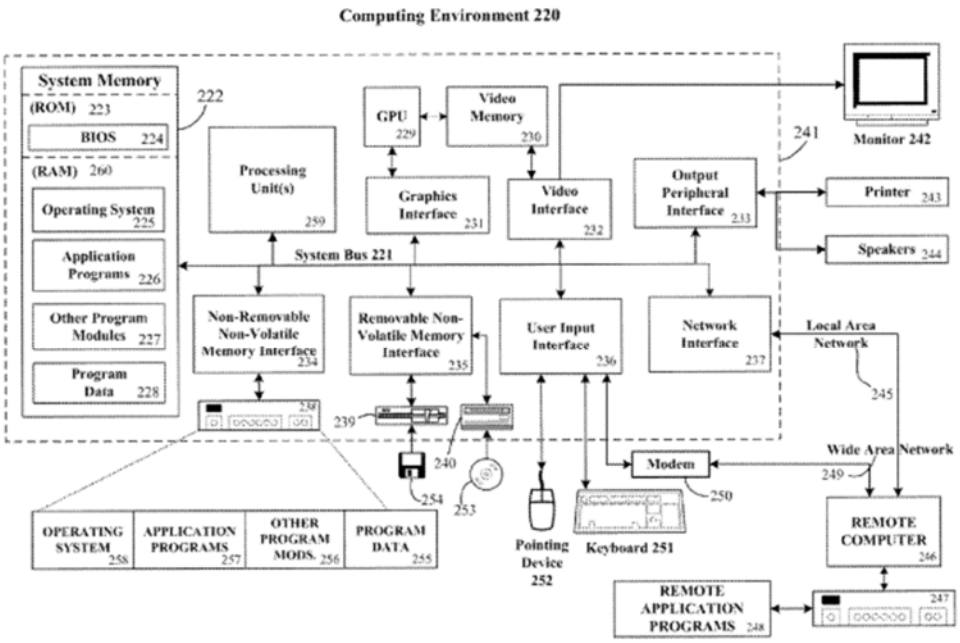


FIG. 4

Ex. 1007 (Geisner) at Fig. 4.

125. Geisner further discloses that the "visual appearance of an on-screen character may then be changed in response to changes to the model being tracked." Ex. 1007 (Geisner) at [0122]. According to Geisner, the system creates a "scanned human target," which may be represented by "one or more data structures" as a 3D model, wherein "[e]ach body part may be characterized as a mathematical vector defining joints and bones." Ex. 1007 (Geisner) at [0123]. This enables image transformations in various angles and directions, as each of the defined joints represents a place where the 3D model "can pivot in the x, y, z directions or a place of interest on the body." Ex. 1007 (Geisner) at [0125], Fig. 8:

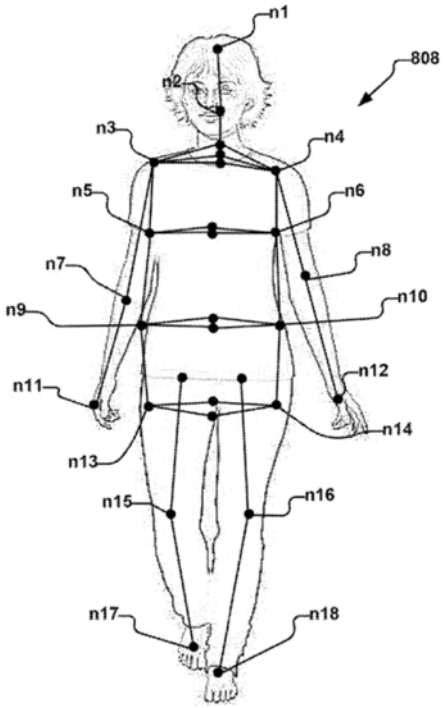


FIG. 8

Ex. 1007 (Geisner) at Fig. 8.

126. Geisner discloses that the tracking system may be deployed for a range of use cases, including an on-screen garment shopping application, a gaming application, and a room modeling application. Ex. 1007 (Geisner) at [0104]-[0111], [0027], [0092]-[0094], and [0163]-[0165]. Similar to the ’110 Patent, Geisner also teaches mirroring the captured image such that it is inverted across a vertical axis. Ex. 1007 (Geisner) at Fig 9:

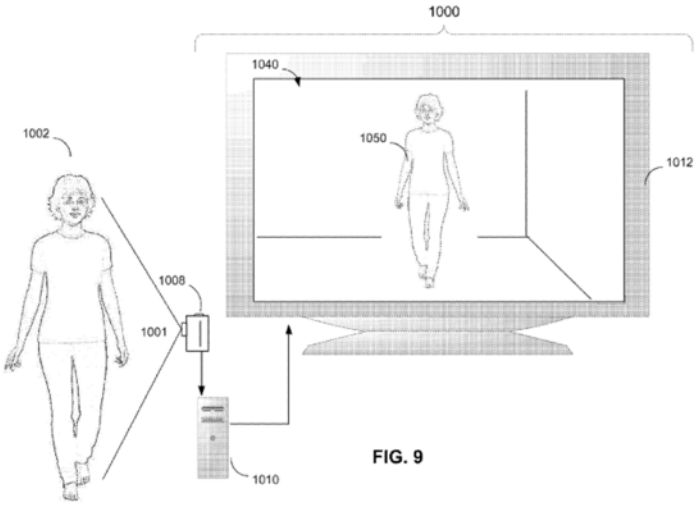


FIG. 9

Ex. 1007 (Geisner) at Fig. 9.

127. Further, Geisner discloses combining and altering a user's features, as tracked and displayed by the system (e.g., eye color, eye size, eye shape, hair color, hair type) by "applying a color, positioning, or scaling to the target" (Ex. 1007 (Geisner) at [0062]), or including sizing controls on a user interface, such that the users can manipulate their virtual representations "with a gain or loss of weight" (Ex. 1007 (Geisner) at [0161]). Geisner teaches additional augmentation features, including allowing users to dress their virtual representations in various articles of clothing and to apply virtual backgrounds, including scanned stores and stock photos. Ex. 1007 (Geisner) at [0153]-[0159].

128. Therefore, Geisner teaches an interactive shopping application that enables users to modify and augment virtual reflections of themselves, solving the known problem identified in the '110 Patent and offering a similar solution.

F. LIZEE (EX. 1008)

a. Status of Lizee as Prior Art

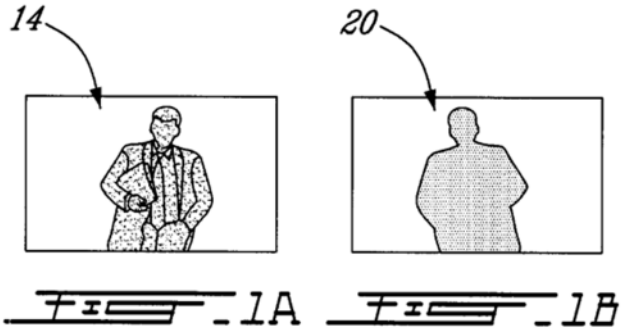
129. "Lizee" refers to International Patent Application Publication No. WO 2006/108279, entitled "Method and Apparatus for Virtual Presence." Lizee names Martin Lizee and Robert Lizee as the inventors. Lizee was filed on April 11, 2006, and was published October 19, 2006. Thus, Lizee was published more than one year prior to the priority date of the '110 Patent, which I understand to be December 18, 2012. *See supra* §VIII, ¶90. Accordingly, I understand that Lizee qualifies as prior art to the '110 Patent.

b. Overview of the Disclosure of Lizee

130. Lizee is directed to a "virtual presence system" comprising an image capture device used to capture user images "from which the virtual representation comprises a mirrored silhouette transformation of the user image." Ex. 1008 (Lizee) at Abstract. Lizee further discloses that the system may calculate the user's position and motion and that the perspective of the captured image can be "altered in response to a given user motion or displacement" based on "evaluating, for example, variations in the size and position of the user's image." Ex. 1008 (Lizee) at Abstract. Lizee teaches that this system can be implemented in "virtual environment systems" that use "real-time motion capture devices," including virtual reality systems and video games. Ex. 1008 (Lizee) at [0004].

131. Similar to the ’110 Patent, the virtual presence system described in Lizee detects the presence of a user in the stream of captured images. Specifically, Lizee discloses the processor using image processing and isolation techniques to “detect[] color differences of the pixels respectively corresponding to the background image and the captured image,” resulting in “isolating the image of the user from the background,” which allows for generation of the virtual silhouette. Ex. 1008 (Lizee) at [0037].

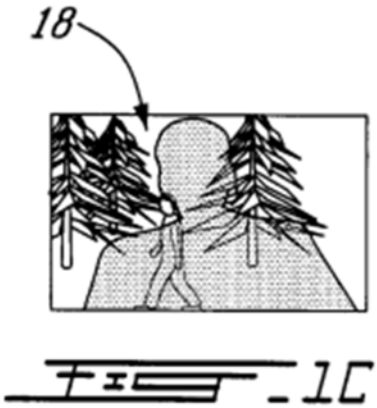
132. Additionally, similar to the ’110 Patent, Lizee discloses various image transformations based on the distance from the user to the image capture device. Lizee teaches that calculating user distances and motion are achieved by methods known to persons skilled in the art (e.g., via linear interpolation based on the size in pixels of “a predefined part or portion” of the imaged user). Ex. 1008 (Lizee) at [0046]. One exemplary image transformation described in Lizee is left-to-right inversion of the user’s image across a vertical axis. Ex. 1008 (Lizee) at [0003], [0034]; *see also* Ex. 1008 (Lizee) at Figs. 1A, 1B:



Ex. 1008 (Lizee) at Figs. 1A, 1B.

133. Further, similar to the ’110 Patent, Lizee discloses that “a position or viewpoint of the virtual camera 28 could be moved to reflect a change in the viewing direction of the virtual silhouette 24.” Ex. 1008 (Lizee) at [0057]. For example, Lizee discloses that “user motions may be detected and used to control various virtual camera perspective modifications, which may include moving, rotating, panning, zooming, orbiting and/or dollying the virtual camera relative to the silhouette model 24 and/or the virtual environment.” Ex. 1008 (Lizee) at [0063].

134. Lizee additionally teaches use of augmented graphical features, including “3D mesh mappings, scaled graphical representations, image and graphical overlays,” (Ex. 1008 (Lizee) at [0075]) as well as processor-generated virtual backgrounds. Ex. 1008 (Lizee) at Fig. 1C:



Ex. 1008 (Lizee) at Fig. 1C.

135. Therefore, similar to the ’110 Patent, Lizee teaches a virtual presence system that modifies and augments video images of a user in real time, solving the known problem identified in the ’110 Patent and offering a similar solution.

XI. GROUND 1: CLAIMS 1-6, 10-12, AND 16-17 ARE OBVIOUS OVER HAAN IN VIEW OF FRANCOIS

136. In the following sections, I demonstrate how Haan in view of Francois renders obvious claims 1-6, 10-12, and 16-17 of the ’110 Patent. All emphases are added unless otherwise indicated.

A. MOTIVATION TO COMBINE HAAN AND FRANCOIS

137. A POSITA would have found it obvious to combine Haan and Francois to achieve an enhanced virtual mirror experience.

138. First, Haan and Francois are analogous to the ’110 Patent and to each other. The ’110 Patent, Haan, and Francois are in the same field of endeavor. The ’110 Patent concerns a system for “virtualizing a mirror.” Ex. 1001 (’110 Patent) at 1:43-51 (Technical Field). It “relates generally to imaging and display systems and, more particularly, to monitors, and interactive displays, e.g., in retail and/or service environments, medical or home situations, video conferencing, [and] gaming.” Ex. 1001 (’110 Patent) at 1:43-46 (Technical Field).

139. Haan is in the same field of endeavor. Haan discloses a mirror device comprising a display device and a camera that converts conventional monitors into

interactive mirroring systems that display images with adjusted (virtual) viewpoints based on the position of a user, including that user's body parts. Ex. 1005 (Haan) at Abstract ("view point from which the to be displayed image is recorded is changed"), (noting "a signal (S) indicative of the position of the body part with respect to the a [sic] display screen"), [0037] ("upgrading a normal TV and/or a computer monitor of a PC to a device with the added function of a mirror device"), [0003] (indicating such devices may be used "at barber shops and in dressing rooms").

140. Similarly, Francois is in the same field of endeavor. Francois discloses computer-based architectural patterns for numerous "projects ranging from single stream, automatic real-time video processing to fully integrated distributed interactive systems mixing live video, graphics and sound." Ex. 1009 (Francois) at Abstract. For example, Francois describes a handheld virtual mirror that synthesizes and displays captured images according to a viewer's position. Ex. 1009 (Francois) at §4.3, Fig. 12.

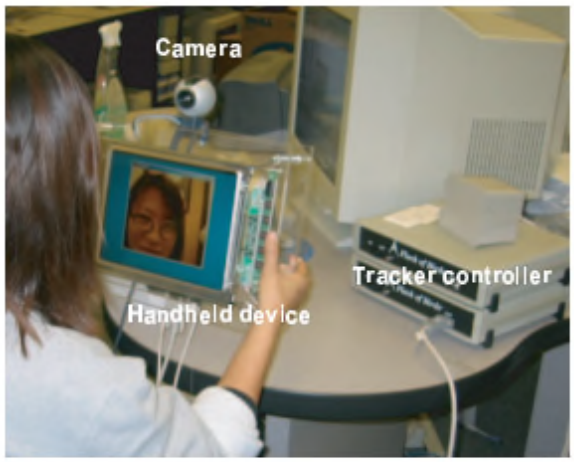


Figure 12: The handheld mirror simulation system ("Virtual Mirror").

Ex. 1009 (Francois) at Fig. 12.

141. Furthermore, Francois concerns adding communication components to support interactive live telecommunications. Ex. 1009 (Francois) at §§4.4-4.5, Fig. 15.

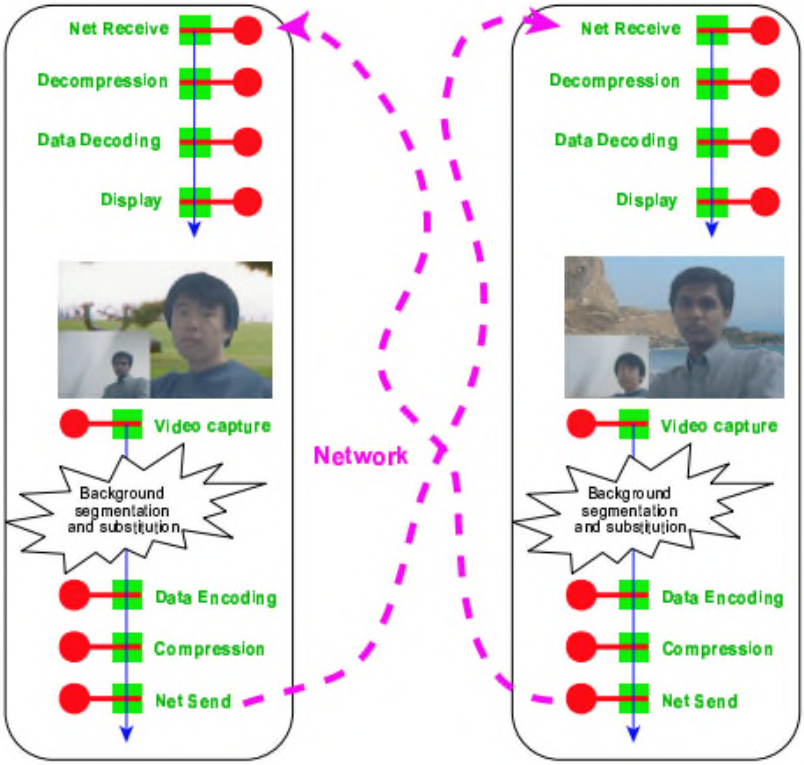


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15.

142. A POSITA would have been motivated to combine Haan and Francois to achieve an enhanced virtual mirror experience. For example, Haan discloses that its virtual mirror system provides "greater freedom in the area of view" allowing the "mirror device to display those areas that are normally just out of sight." Ex. 1005 (Haan) at [0012]. And Haan discloses its system can be implemented on computers and mobile devices. *Id.* at [0037]. A POSITA would have been motivated to look at references improving upon virtual mirror systems for known computers and mobile devices, like Francois, to similarly improve Haan and offer additional enhanced user

experiences. A POSITA would have been motivated to add network capabilities to Haan's computers and mobile devices, such as the network communication module (IMSC Communicator) in Francois (Ex. 1009 (Francois) at §4.4), to enable a user to transmit modified images made locally (like in Haan) to a remote device for display and to provide the benefit of a robust communication functionality in Haan's system. A POSITA would also have understood that making this combination would provide a user with additional flexibility and functionality and thus with an enhanced user experience, for it would allow users to share the side/profile views with other individuals during a video call. Ex. 1005 (Haan) at [0037]; Ex. 1009 (Francois) at Fig. 15.

143. For example, as shown in the below modified figure, combining figures from Francois and Haan, a POSITA would have been motivated to (and would have found it obvious to) display the modified image on both a local display and a remote display to expand upon the benefits of Haan and offer an enhanced user experience:

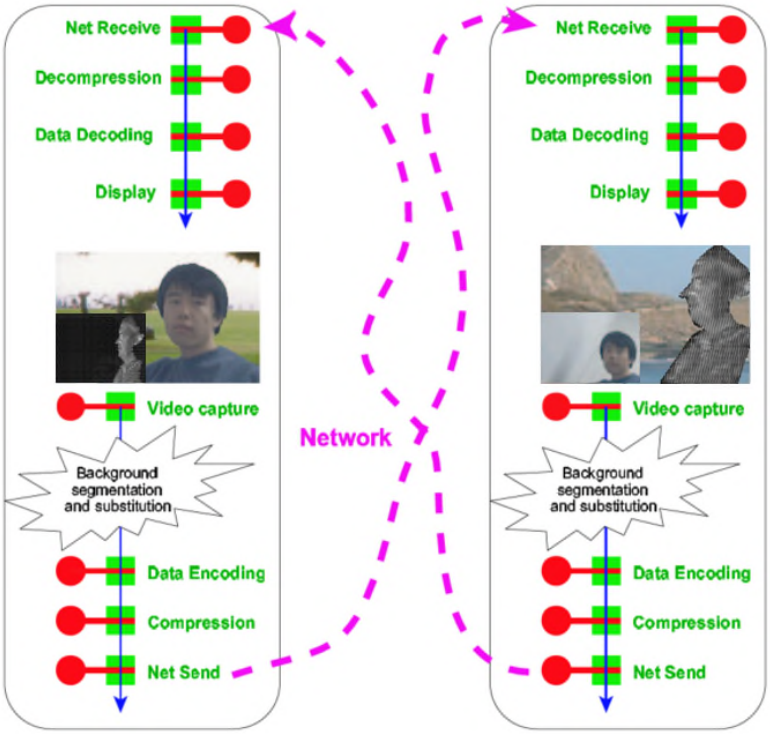


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15, combined with Ex. 1005 (Haan) at Fig. 3.

144. Further, the teachings of Haan and Francois would have led a POSITA to combine them. For example, as noted above, Haan teaches that its system can be implemented on both computers and mobile devices. Ex. 1005 (Haan) at [0037]. A POSITA would have been motivated to seek out references, such as Francois, that fully utilize the capabilities of known computer and mobile devices, including their network capabilities, to improve upon and ultimately expand the usability of Haan's mirror device system.

145. In particular, Francois teaches the disclosed software architecture can be used to implement a virtual mirror, identifying the virtual mirror as an exemplary

project. Ex. 1009 (Francois) at §4.3 ("Handheld mirror simulation"). A POSITA would have recognized the teachings regarding the virtual mirror in Francois would be relevant to Haan's virtual mirror and would offer an improved way to implement the virtual mirror in Haan to provide an enhanced user experience.

146. Francois teaches a software architectural framework, that is called the Software Architecture for Immersipresence ("SAI"), which efficiently streamlines image processing applications and permits flexible system expansion and integration (e.g., enabling new applications to be added). Ex. 1009 (Francois) at §§1.2-1.3. The goal of SAI is to provide a universal framework for numerous projects "ranging from single stream automatic real-time video processing to fully integrated distributed interactive systems mixing live video, graphics and sound." Ex. 1009 (Francois) at §1.3. The virtual mirror in Haan is within the scope of the projects contemplated by Francois; indeed, an exemplary implementation in Francois is a virtual mirror similar to the one described in Haan. For example, as shown below in Figure 13 of Francois, Francois discloses an exemplary processing schematic for virtual mirrors, similar to the virtual mirror in Haan. *See* Ex. 1009 (Francois) at §4.3; Ex. 1009 (Francois) at Figs. 12-13. Thus, the teachings of Haan and Francois underscore their compatibility and would have led a POSITA to combine them to again improve upon and expand the usability of Haan's mirror device system.

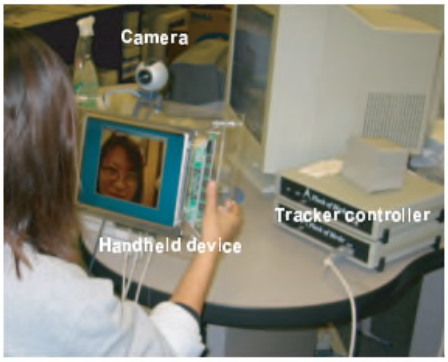


Figure 12: The handheld mirror simulation system ("Virtual Mirror")

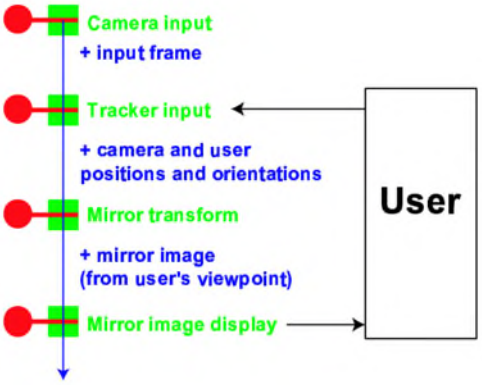


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 12 (left) and Fig. 13 (right).

147. Indeed, Francois teaches that a mirror device can be seamlessly integrated into an extensible platform for remote data sharing. Ex. 1009 (Francois) at §4.4. This is shown, for example, below in Figure 15. Ex. 1009 (Francois) at Fig. 15.

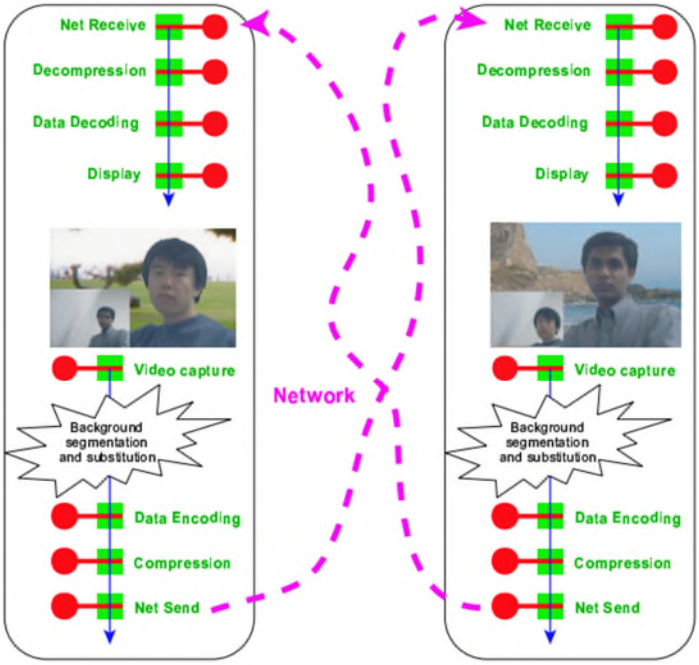


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15.

148. A POSITA would have understood the exemplary system described in Figure 15 can incorporate input tracking and transformation features associated with a mirror, at least as illustrated in the modified figure below:

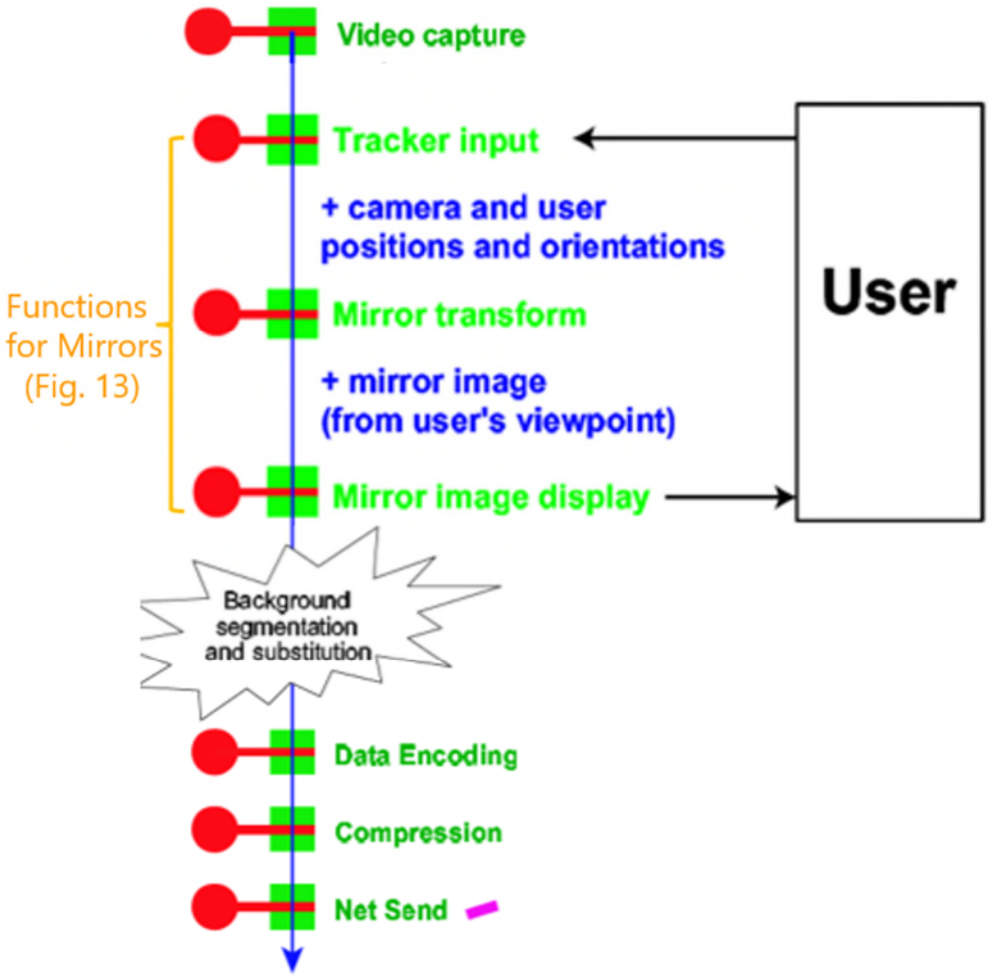


Figure 13: The Virtual Mirror application graph.

Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 13 (combined with Fig. 15 of Ex. 1009 (Francois)). A POSITA would have understood Francois's teachings were relevant to (and compatible with) Haan's device and would have been motivated to modify Haan in view of those teachings to have improved communication capabilities.

149. A POSITA would have had a reasonable expectation of success making such a combination because Haan and Francois apply known techniques to improve similar devices. As shown below, Haan and Francois aim to improve a virtual mirror comprising a camera, a known computing device, and a display. *See, e.g.*, Ex. 1005 (Haan) at Abstract, [0023]-[0027], Fig. 2; Ex. 1009 (Francois) at §4.3, Figs. 12-13.

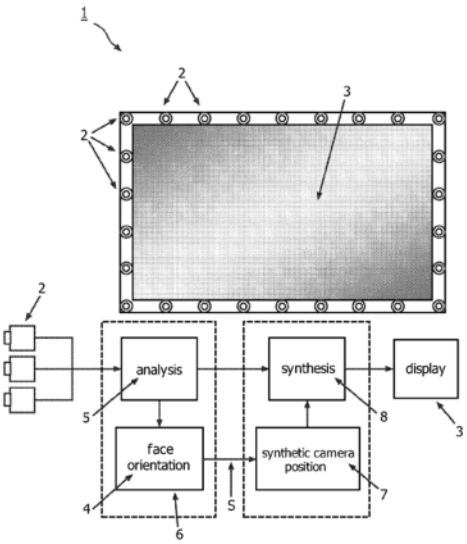


FIG. 2

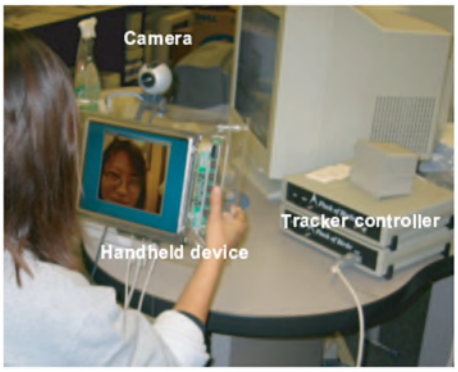


Figure 12: The handheld mirror simulation system ("Virtual Mirror").

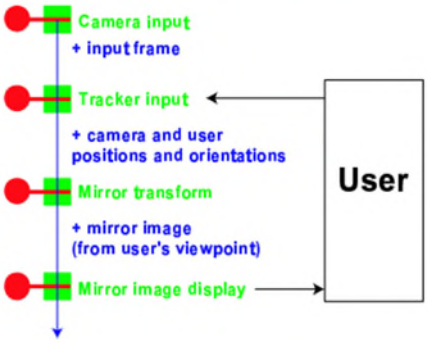


Figure 13: The Virtual Mirror application graph.

Ex. 1005 (Haan) at Fig. 2 (left); Ex. 1009 (Francois) at Figs. 12, 13 (right).

150. A POSITA would have been motivated to combine the teachings of Francois and Haan with a reasonable expectation of success, since both references disclose techniques to synthesize and display captured images based on a user's

position in relation to a display, and thus both references apply similar techniques to enhance a user's experience. *See, e.g.*, Ex. 1005 (Haan) at [0058]-[0059]; Ex. 1009 (Francois) at Fig. 13 (depicting exemplary processing schematic for virtual mirrors, similar to the virtual mirror in Haan). By applying synthesis algorithms, both references create a virtual camera with a modified point of view. Ex. 1005 (Haan) at [0014], [0075]-[0076]; Ex. 1009 (Francois) at §4.3.

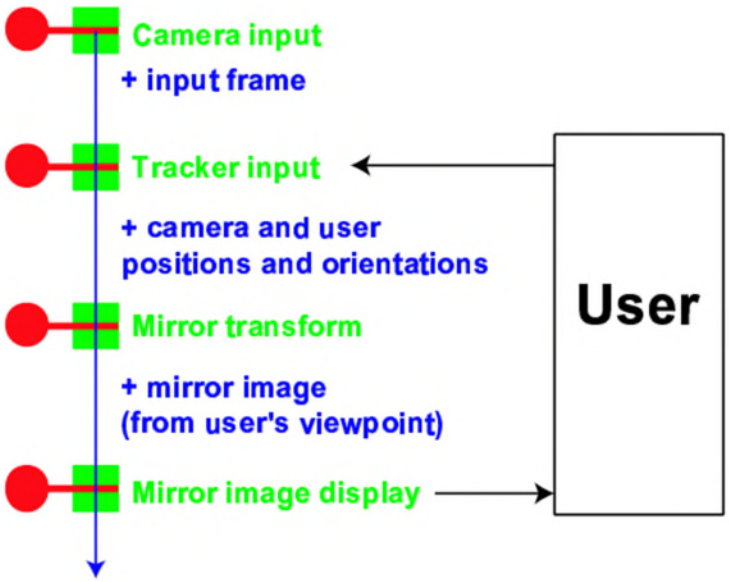


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 13.

151. Combing Haan and Francois would be no more than a simple substitution of known elements to obtain predictable results. For example, a POSITA would have been motivated to modify the virtual mirror in Haan with the software architecture and IMSC communicator in Francois for remote interaction to obtain

the benefit of improved and streamlined telepresence using virtual mirrors that Francois describes (collaboration over network with enhanced image processing). Ex. 1009 (Francois) at §§4.4-4.5. A POSITA would have understood that combining Francois and Haan would yield predictable results, as Francois provides specific implementation details for how its software architecture can be used for the implementation of devices such as a virtual mirror in a computer or mobile device. Ex. 1009 (Francois) at §§4.3-4.5.

B. CLAIM 1:

a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising:

152. I understand that this limitation forms part of the preamble of this claim and therefore may not be limiting. To the extent the Board treats element 1[pre] as limiting, Haan in view of Francois renders obvious this limitation. Specifically, as shown in Fig. 2 of Haan below, Haan teaches a computer program for operating a mirror display system having a camera (an image recording device), a monitor (a display device), and processor (processing means). Ex. 1005 (Haan) at claim 1, [0058]; *see also* Ex. 1005 (Haan) at Abstract, Fig. 2, [0006]-[0011], [0087], [0109]-[0118]. Haan describes said mirror device as having "at least one **image recording device** (2) to record an image." Ex. 1005 (Haan) at claim 1; *see also* Ex. 1005 (Haan)

at Abstract, Fig. 2, [0006]-[0011]. Haan also discloses a mirror device having "a **display device** (3) with a display screen." Ex. 1005 (Haan) at claim 1.

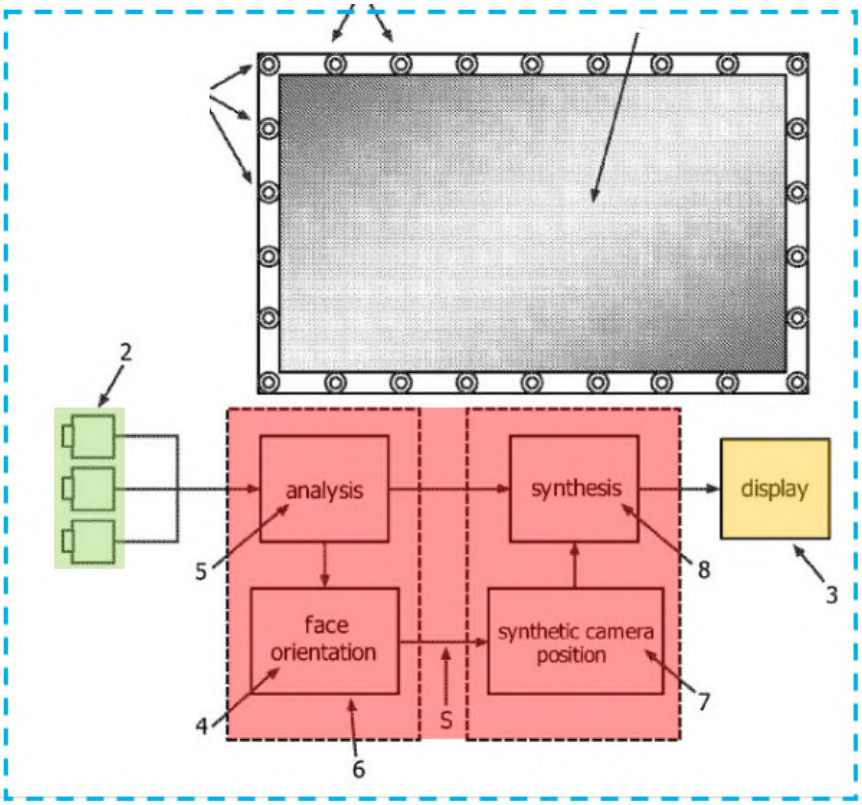


FIG. 2

Ex. 1005 (Haan) at Fig. 2 (annotated).

153. Haan describes that the mirror device including the display device (3) and image recording device (2) includes "a **computer program** comprising program code means for performing a method according to the present invention. Ex. 1005 (Haan) at [0087]. Haan also discloses a "means . . . to determine a position . . . ; a means to change (7) the view point. . . [and] a means to display the recorded image on **the display screen of the display device** (3)." Ex. 1005 (Haan) at claim 1; *see*

also Ex. 1005 (Haan) at Abstract, Fig. 2, [0006]-[0011]. A POSITA would have understood that the computer program described in Haan operates on a processor.

154. Additionally, Haan discloses a processor. As shown in Figure 2 above, Haan discloses an "analysis means" and a "synthesis means." Ex. 1005 (Haan) at Fig. 2, [0058]. Additionally, Haan discloses that "[t]he data [] is … combined by the **processing means** to arrive at a single image signal from a virtual point of view." Ex. 1005 (Haan) at [0058]. Haan further discloses that the invention is "embodied in a computer program comprising program code means for performing a method according to the present invention." Ex. 1005 (Haan) at [0087]. A POSITA would have understood that a computer program operates on a processor. Thus, A POSITA would recognize that Haan discloses a processor, and this limitation would also have been obvious to a POSITA in view of Haan, as a processor would have been an obvious way to implement Haan's disclosure.

b. <u>1[a]: operating the camera to generate a stream of images of the user;</u>

155. Haan in view of Francois renders obvious "operating the camera to generate a stream of images of the user." Haan discloses one or more cameras that generate images of a user. For example, Haan discloses a "[m]irror device comprises . . . a camera (2) to record an image." Ex. 1005 (Haan) at Abstract, [0008], [0057]-[0058], claim 1. Further, Haan describes "[t]he generated image is the image of the

viewer" and that the recorded images from the one or more cameras "are fed to a position determination means" for further processing. Ex. 1005 (Haan) at [0058], [0113]-[0114]; *see also* [0117]-[0118], Fig. 2.

156. Haan describes a display system that functions as a mirror. Haan also discloses "a means for determining a position of a body part with respect to the display screen of the display device [] of a viewer in front of the display screen." Ex. 1005 (Haan) at [0009]. The determined position may comprise "the distance to the display screen," "angle of turn of the head [or] torso in respect of the display screen," and "movements to and/or from the display screen." Ex. 1005 (Haan) at [0014], Fig. 9.

FIG. 9

Ex. 1005 (Haan) at Fig. 9 (partial).

157. Haan discloses a stream of images because it discloses a display system that functions as a conventional mirror that captures movements (e.g., turning of a user's face or body). As Haan describes, "[t]he recorded images by cameras 2 are fed to a position determination means 4." Ex. 1005 (Haan) at [0058]; *see also* Ex. 1005 (Haan) at [0017] ("A nod to the right, followed by a slow returning head to a normal position will enable the viewer to shift the view …"), [0015], Fig. 9. Because movements and the speed of the movements are captured, the camera in Haan would necessarily generate a stream of images in order to capture these movements and in turn behave as a mirror. Ex. 1005 (Haan) at [0017].

158. Haan discloses an enhanced mirror wherein the image(s) from the one or more cameras is (are) magnified or enhanced to view some parts better. When the user is looking straight or is turning by a small angle, the image is produced using a reflective surface rather than a synthesized mirror image. Ex. 1005 (Haan) at [0030], [0076], Fig. 9. Further, Haan discloses that the mirror device shows the turning and moving of the user's face or body. Ex. 1005 (Haan) at [0015],[0018], [0061]-[0063], Figs. 9-10. The mirror device is capable of providing a smooth transition between images where magnification is applied and where magnification is not applied. Ex. 1005 (Haan) at [0063], [0076].

159. Thus, Haan discloses that the camera capturing the image of the user generates a stream of images in order for the display system to function as a mirror

and captures multiple images or movements as the user moves. Ex 1005 (Haan) at [0009], [0014], [0030], [0076], [0063]. In other words, a stream of images of the user is generated by the camera in Haan in order for Haan's display system to behave as a mirror. Ex. 1005 (Haan) at claim 1; *see also* Ex. 1005 (Haan) at Abstract, Fig. 2, [0006]-[0011], [0063].

160. However, if the Patent Owner argues that this limitation is not expressly disclosed in Haan, it is necessarily present in Haan's disclosure. A POSITA would have also found it obvious to operate Haan's one or more cameras to generate video or a stream of images to capture movement if it was intended for the display connected to the one or more cameras to function as a mirror and mimic movements of the user.

161. Additionally, Haan in view of Francois renders obvious this limitation. As discussed, Francois teaches operating a camera to generate a stream of images of a user. For example, Francois discloses that "video frames [are] captured and processed" (Ex. 1009 (Francois) at §2.2.3), and "images of the observed world, consistent with the viewer's position, must be synthesized and displayed in real-time" to that user (Ex. 1009 (Francois) at §4.3). Francois further discloses "[t]he continuous input video stream and tracker data is used to synthesize, in real-time, a continuous video stream displayed on the LCD screen. The synthesized video stream

is a close approximation to what the user would see on the screen surface if it were a real mirror." Ex. 1009 (Francois) at §4.3.

162. Additionally, Francois discloses an architectural style, called SAI, and teaches that "[a]n appropriate style should support the design and implementation of software systems capable of handling images, 2-D and 3-D geometric models, video streams, various data structures, in a variety of algorithms and computational frameworks." Ex. 1009 (Francois) at §1.1. Francois further discloses that the supported "applications may be interactive and/or have real-time constraints" and that "[a] fundamental underlying computational invariant across such systems is distributed parallel processing of generic data-streams" Ex. 1009 (Francois) at §1.1. Francois also teaches that the user reaction can be collected using "various sensory devices" generating "data streams." Ex. 1009 (Francois) at §§4.3, 4.5. Francois further describes that the SAI architecture handles streams of data coming from various sources. Ex. 1009 (Francois) at §§2.2, 4.3, 4.5. Figs. 5, 13.

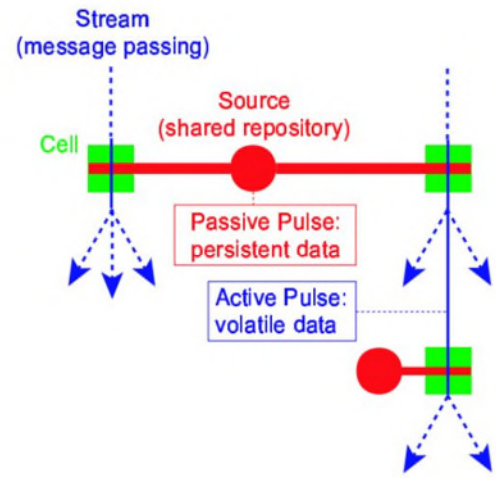


Figure 5: Summary of notation for SAI designs.

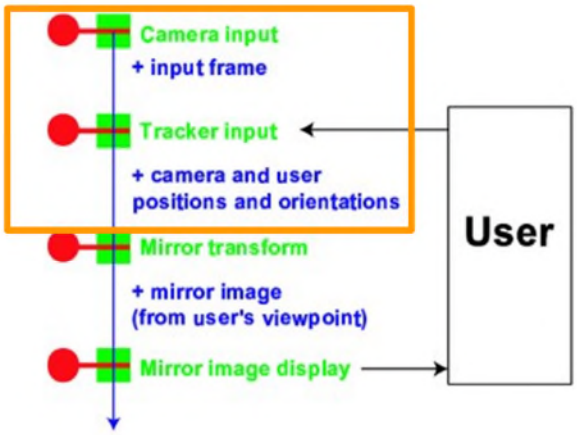


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 5 (left), Fig. 13 (right) (annotated).

163. A POSITA would have understood that the data streams disclosed by Francois, including images, captured video frames, and/or video streams consist of a stream of images. A POSITA would have been motivated to combine Francois's teachings of video or streams of images with Haan to achieve an enhanced virtual mirror experience, and a POSITA would have had a reasonable expectation of success in doing so. *See supra* §XI.A. Moreover, a POSITA would have found it obvious to use Francois' SAI architecture for implementing Haan's virtual mirror functionality. *See supra* §XI.A, ¶¶145-150. For example, a POSITA would have recognized that Haan seeks to mimic a mirror by using a virtual display system (*see supra* §X.A.b, ¶94) and, to the extent not explicitly disclosed in Haan, a POSITA would have understood this can be more accurately and readily achieved with a

stream of images that allow for real-time processing, as supported by Francois' SAI architecture.

c. **1[b]: operating the processor to detect the presence of the user in the stream of images;**

164. Haan in view of Francois renders obvious "operating the processor to detect the presence of the user in the stream of images."

165. Haan discloses the mirror device "comprises [a] means for determining . . . the position of a body part." Ex. 1005 (Haan) at Abstract. Further, Haan discloses the mirror device "determine[s] a position of a body part with respect to the display screen of the display device." Ex. 1005 (Haan) at [0009]. Haan describes the body part may be "the head or face of a person, [and/or] in other embodiments the body part is the torso." Ex. 1005 (Haan) at [0013].

166. Haan discloses detecting various body parts, including a user's face, eyes, and mouth, to determine face orientation. Ex. 1005 (Haan) at [0058]. In particular, Haan explains the recorded images "are fed to a position determination means," which analyzes face orientation of a user by detecting "the ratio between the size of the head and the distance between the eyes, or the ratio between the distance between the eyes and the distance between the eyes and the mouth." Ex. 1005 (Haan) at [0058]. Given that Haan describes detecting various body parts (e.g., face, eyes, and mouth) and performing various calculations based on that detection

to determine face orientation (Ex. 1005 (Haan) at [0058]), Haan discloses detecting a user's presence, as claim 3 of the '110 Patent states that "user presence is identified by … detecting a body part." Ex. 1001 ('110 Patent) at claim 3.

167. Further, Haan discloses operating a processor to perform this step. For example, Haan teaches that a "[t]he invention is [] embodied in a computer program comprising program code means for performing a method according to the present invention" and that "face orientation is determined in means 6." Ex. 1005 (Haan) at [0087]. A POSITA would have understood that a computer program, such as the one used for the determination of a user's face orientation and position, would operate on a processor. *See supra* §XI.B.a, ¶¶153-154. Haan also teaches that "[t]he data from a number of cameras is [] combined by the processing means to arrive at a single image signal from a virtual view point." Ex. 1005 (Haan) at [0058]. Thus, Haan discloses operating a processor to perform its method, including detecting the presence of the user in the stream of images from the one or more cameras.

168. Further, a POSITA would have understood that to operate the processor to detect the presence of a *body part of the user*, the processor must also detect *the presence of the user* itself. Or, at the very least, it would have been obvious to detect the presence of the user if the processor is detecting the presence of a body part of the user. In other words, a POSITA would have understood that an identified body part is most likely to be associated with a user.

169. Moreover, this limitation would have been obvious in view of Francois. As shown below in Figure 13 of Francois, the tracking step identifies user positions and orientations. Ex. 1009 (Francois) at §4.3, Fig. 13. Francois teaches additional video analysis techniques may be used, "including face detection and tracking, and gaze tracking" and provides an expandible software framework for implementing functions such as those disclosed in Haan, including detecting the presence of a user in a stream of images. Ex. 1009 (Francois) at Fig. 13, §4.3.

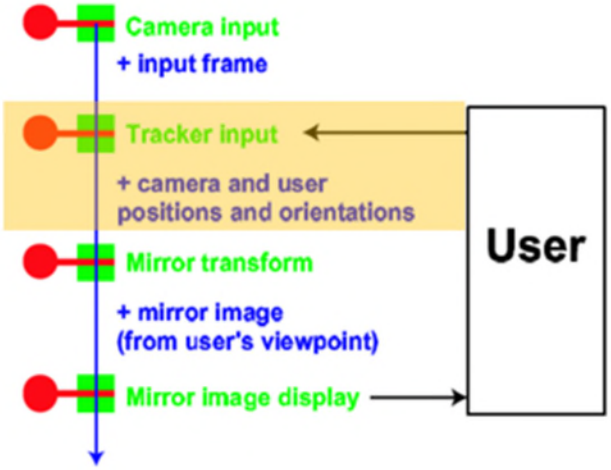


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 13 (annotated).

170. Francois teaches that such a system may be implemented in a system with a dual processor such as Pentium 4@1.7GHz. Francois at §5. A POSITA would have found it obvious to apply the SAI architecture in Francois and deploy the virtual mirror in Haan with Francois's recommended dual processor to accurately mimic a

mirror using a virtual display system (*see supra* §XI.A, ¶¶145-150) given that Haan discloses its analysis and synthesis means are in the form of computer programs. A POSITA would have done so with a reasonable expectation of success, as Francois discloses a known processor that is compatible with virtual mirror systems like those disclosed in Haan. *See supra* §XI.A. Thus, Haan in view of Francois at least renders the limitation obvious.

d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view

171. Haan in view of Francois renders obvious this limitation.

i. Haan in view of Francois discloses "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera. . . "

172. Claim 2 of the '110 Patent clarifies that transformation mapping may include "vertical flips about a vertical axis,[3] a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image." Ex. 1001 ('110 Patent) at claim 2. Haan in view

[3] *See* below footnote No. 4 discussing claim 2 (*see* §XI.C)

of Francois discloses such "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera."

173. The '109 Patent explains transformation mapping "essentially reassigns new addresses to each pixel of the original image." Ex. 1021 ('109 Patent) at 29:28-30; *see also* Ex. 1021 ('109 Patent) at 14:40-43. The '109 Patent further discloses exemplary transformation mappings, including "angle, tilt, azimuth, scale, [and] spatial translation." Ex. 1021 ('109 Patent) at 29:38-42. Additionally, the '109 Patent states that a "final transformation is a matrix multiplication of the individual transformation." Ex. 1021 ('109 Patent) at 29:40-42. *See supra* §VII.A, ¶77. A POSITA would have understood a transformation mapping may be represented by a matrix. *See, e.g.*, §VI.B.a, ¶47; §VI.B.c, ¶¶55-57; VI.B.e, ¶¶61-64; VI.B.f, ¶67.

174. Haan discloses at least two exemplary types of adaptive transformation mapping: (1) an adaptive rotation based on the orientation of the head or torso of a viewer (Ex. 1005 (Haan) at [0060]-[0071]) and (2) an adaptive resizing based on the distance from a viewer to the screen or camera (Ex. 1005 (Haan) at [0072]-[0076]). For example, Haan discloses "the turning of the head can be exaggerated so that . . . the turning angle (and thereby the camera position, i.e.[,] the point of view from the image is recorded) is increased." Ex. 1005 (Haan) at [0061]. Figure 6 below shows two exemplary correlations (linear or non-linear) between the actual angle of the viewer's head and the displayed angle. Ex. 1005 (Haan) at [0070]-[0071].

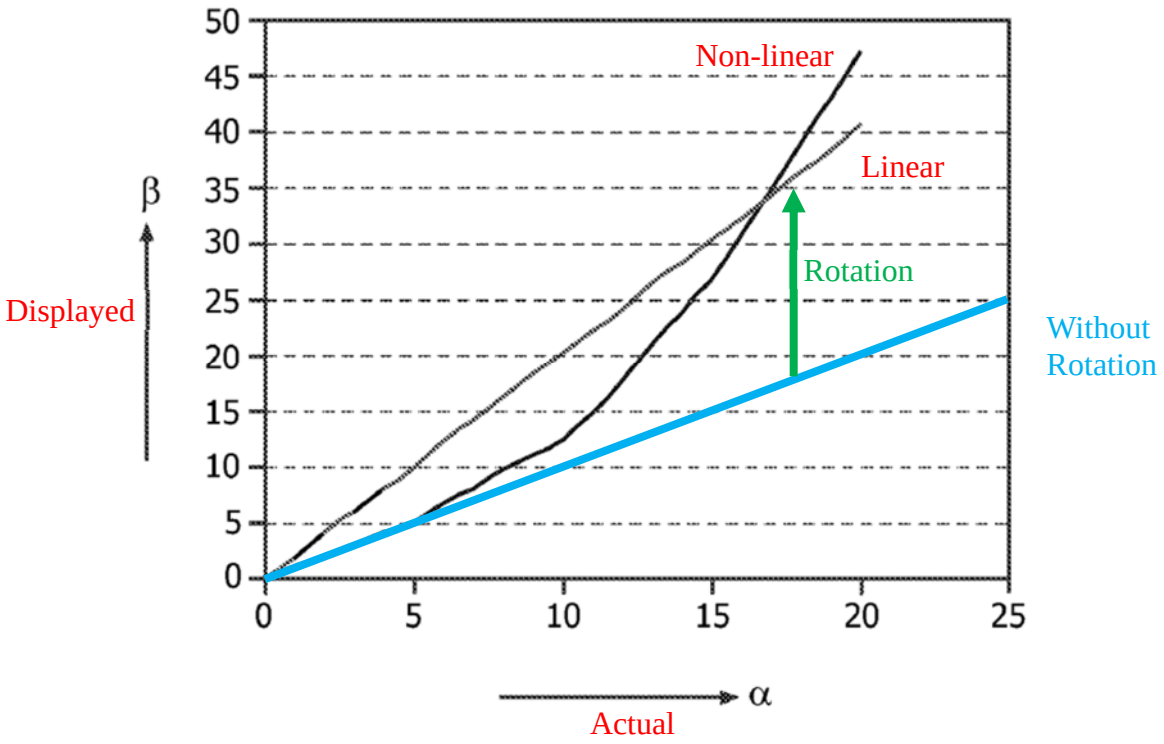


FIG. 6

Ex. 1005 at Fig. 6 (annotated).

175. If there was no rotation transformation applied to the image in Haan, the displayed angle would equal the actual angle (as shown with the blue line above in annotated Fig. 6). However, Haan discloses that there is a rotation transformation applied to the images. Thus, the displayed angle becomes larger than the actual angle (as shown above by the green annotation in annotated Figure 6).

176. In addition to disclosing adaptive rotation based on the orientation of the head or torso of a viewer, Haan discloses that "the displayed image is magnified [(i.e., applying a resizing transformation)] as a function of the distance to the mirror

device." Ex. 1005 (Haan) at [0072], Figs. 7-8. As explained in Haan, the system exaggerates the movement of a user by applying the resizing transformation, such as to enlarge the image to make the user appear closer than they actually are. Ex. 1005 (Haan) at [0022]-[0023], [0072], [0076]. This transformation is described in Haan as allowing "a viewer to change the magnification by for instance getting closer to the display screen." Ex. 1005 (Haan) at [0023]; *see also* [0022], [0072], [0076]. A POSITA would have known that rotation is an angle transformation while resizing is a scaling transformation. The two transformations (rotation and resizing) are both adaptive transformations, as claimed, because they change according to the user's orientation or distance.

177. To the extent Patent Owner argues that the exact terms "transformation mapping" and "transformation" were not used in Haan, a POSITA would have understood that the synthesis in Haan corresponds to the application of transformation mappings, as the synthesis of virtual camera views from actual camera views was well-known in the art to be a transformation. *See, e.g.*, §VI.B.c, ¶¶55-57. Moreover, as discussed, the '109 Patent states that a transformation mapping "essentially reassigns new addresses to each pixel of the original image" (Ex. 1021 ('109 Patent) at 29:28-30; *see also* Ex. 1021 ('109 Patent) at 14:40-43) and discloses exemplary transformation mappings (Ex. 1021 ('109 Patent) at 29:38-42). *See supra* §XI.B.d.i, ¶173. Further, as discussed, the '109 Patent states that a

"final transformation is a matrix multiplication of the individual transformation." Ex. 1021 ('109 Patent) at 29:40-42. *See supra* §VII.A, ¶77; §XI.B.d.i, ¶173. A POSITA would have understood that at least the exemplary adaptive transformation mappings disclosed in Haan, such as the adaptive rotation based on the orientation of the head or torso of a viewer and the adaptive resizing based on the distance from a viewer to the screen or camera (Ex. 1005 (Haan) at [0061]-[0076]; *see supra* §XI.B.d.i, ¶174), would be carried out by assigning new addresses to individual pixels of the actual captured images. A POSITA would have found it obvious to apply transformation mappings to the captured images to carry out Haan's synthesizing steps, and would have understood synthesis is performed by using one or more transformation matrices and giving new locations to each pixel of the actual images. *See, e.g., supra* §VI.B.a, ¶47; §VI.B.c, ¶¶55-57; VI.B.e, ¶¶61-64; VI.B.f, ¶67.

178. Further, as I discussed above, Haan discloses operating a processor to perform this step. Haan also discusses "[t]he data from a number of cameras is . . . combined by the processing means to arrive at a single image signal from a virtual view point." *See* §XI.B.c, ¶167; Ex. 1005 (Haan) at [0058]. Thus, Haan discloses "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera," as claimed.

179. Additionally, as discussed above, Francois teaches a processor applying adaptive transformation mapping to the stream of images captured by the camera. For example, Francois discloses a processor applying a mirror transformation. Ex. 1009 (Francois) at §4.3, Fig. 13.

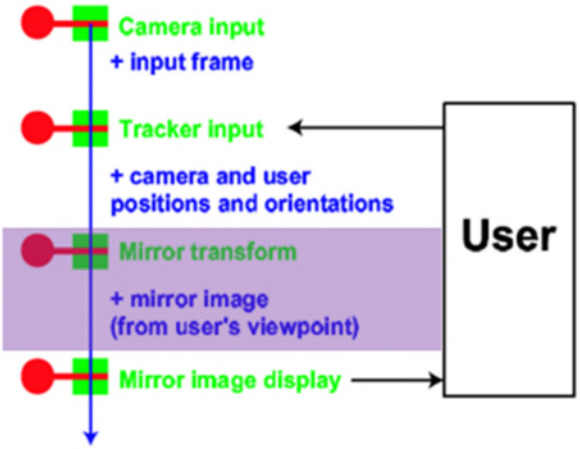


Figure 13: The Virtual Mirror application graph.

Ex. 1009 (Francois) at Fig. 13 (annotated).

A POSITA would have been motivated to perform this step on a processor based on Haan in view of Francois. As discussed, a POSITA would have found it obvious to use Francois' SAI architecture for implementing Haan's virtual mirror functionality. *See supra* §XI.A, ¶¶145-150. Further, Haan discloses that the means for the rotation and resizing are in the form of computer programs. *See supra* §XI.A; Ex. 1005 (Haan) at [0058], [0087], [0099], [0105]. A POSITA would understand that computer programs operate on a processor and a POSITA would have done so with

a reasonable expectation of success, as Francois discloses a known processor that is compatible with virtual mirror systems like those disclosed in Haan. *See supra* §XI.A; §XI.B.a, ¶¶153-154; §XI.B.c, ¶167.

ii. Haan in view of Francois discloses "generate modified images that appear to be captured from a different point of view of the camera's actual point of view"

180. Haan in view of Francois discloses "generat[ing] modified images that appear to be captured from a different point of view of the camera's actual point of view." Haan discloses "a means to change [] the view point from which the image is recorded." Ex. 1005 (Haan) at claim 1, *see also* [0027], [0058].

181. Further, Haan describes the "camera position" as "the point of view from [which] the image is recorded." Ex. 1005 (Haan) at [0061]. In other words, Haan describes that a point of view is the same as a camera position. Therefore, an image with a virtual point of view appears to be taken from a virtual camera position. Haan further teaches applying a rotation transformation to change the camera angle and a resizing transformation to change camera distance. Ex. 1005 (Haan) at [0027], [0058], [0060]-[0064], [0070]-[0076]; *see also infra* §XI.B.f.

182. Haan discloses the mirror device performs image processing to modify and display camera-recorded images, based on a user's position relative to the mirror device. Haan discloses the mirror device can "construct a virtual camera position" (i.e., the "virtual viewpoint") and/or can synthesize an image with a magnification

of the user to be displayed. Ex. 1005 (Haan) at [0058], [0073]. Haan recognizes that "magnification is…equivalent to moving a camera with fixed magnification closer to the object, so all applications can be seen, in fact as changing the point of view" of a real or virtual camera. Ex. 1005 (Haan) at [0073]; *see also* Ex. 1005 (Haan) at [0074]-[0076].

183. Haan similarly discloses that the magnification and thus the camera's point of view changes depending on the user's distance from the mirror device. Ex. 1005 (Haan) at [0073]-[0076],; Fig. 8. Haan discloses that the system exaggerates a user's movement by applying the resizing transformation, such as to enlarge the image to make the user appear closer. *See* Ex. 1005 (Haan) at [0022]-[0023], [0072], [0076]. Because Haan teaches that the magnification is obtained by synthesizing an image, a POSITA would understand the camera's point of view corresponding to the synthesized image is a virtual camera's point of view that is different from the actual camera's point of view. Ex. 1005 (Haan) at [0073]-[0076], Fig. 8.

184. Additionally, Haan describes that "[u]sing several cameras one can, by synthesis of the various images, construct a virtual camera position and/or magnification." Ex. 1005 (Haan) at [0058]; *see also supra* VI.B.c. Thus, in Haan, a final ("modified") image with a virtual camera position is generated by applying the synthesis means. Because a virtual point of view is formed by synthesizing images from multiple cameras, the virtual point of view would be different from at least one,

if not all, of the actual camera points of view. As shown below, the modified image (right) appears to be captured from a different point of view than the original (left).

31
Without rotation
(normal mirror)

FIG. 3

32
With rotation
(Virtual point of view)

Ex. 1005 (Haan) at Fig. 3 (annotated).

185. Synthesizing requires at least two cameras, and the synthesized virtual point of view must be different from at least one of the cameras. The synthesized virtual point of view cannot be the same as both cameras. Ex. 1005 (Haan) at [0058]. Thus, the virtual point of view is different from at least one actual point of view of at least one camera. Thus, Haan discloses a means to "generate modified images that appear to be captured from a different point of view of the camera's actual point of view."

186. A POSITA would have understood a synthesis process for adjusting the camera point of view is a well-known type of transformation (*see, e.g.,* §VI.B.c) and therefore, a POSITA would have found it obvious to implement a synthesis by applying a transformation mapping to the captured images. Thus, a POSITA would have understood that the disclosure in Haan relating to synthesizing images to "construct a virtual camera position and/or magnification" discloses or at least renders obvious applying "transformation mapping.". Ex. 1005 (Haan) at [0058].

187. Further, a POSITA would have found it obvious to combine the teachings of Haan and Francois to achieve an enhanced virtual mirror experience. For example, Francois discloses adding communication components to support interactive live telecommunications (Ex. 1009 (Francois) at §§4.4-4.5, Fig. 15; *see, e.g., supra* §XI.A, ¶141) and discloses applying transformations that change the viewpoint of the image. Ex. 1009 (Francois) at Fig. 15. *See supra* §XI.A, ¶142. As shown in Figure 15 of Francois below, the right image has a different (i.e., virtual) point of view with respect to the user than the point of view in the left image. Additionally, because the user in the right image appears larger than in the left image, the distances between the user and the camera position in the two images are different. Ex. 1009 (Francois) at Fig. 15 (partial, annotated). A POSITA would have found it obvious to combine the features relating to Francois' communication functionality (including applying transformations that change the viewpoint of the

image during interactive live telecommunications) with Haan's mirror system to enable a user to generate and transmit modified images (i.e., images captured from a different point of view) to a remote display in order to achieve a robust communication functionality.

Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15 (partial, annotated).

e. <u>1[d]: displaying the modified images on the monitor;</u>

188. Haan in view of Francois renders obvious this limitation. As discussed above, Haan discloses various types of display (monitor), including a conventional computer monitor, a TV, or a display with semi-transparent layer. Ex. 1005 (Haan) at [0033], [0037], [0099]-[0101]. Haan also claims "a means to display the recorded image on the display device" after "changing the view point from which the image is recorded." Ex. 1005 (Haan) at [0010]-[0011], claim 1.

189. Based on the description in Haan, the modified image is displayed on the monitor. As shown in multiple figures, the synthesized image (an example of the claimed modified image) is sent to a display for displaying. Ex. 1005 (Haan) at Figs. 2, 7, 14, [0058], [0070]-[0073], [0099]-[0101].

190. Further, Haan discloses "the displayed image may be an image derived from or associated with the captured image." Ex. 1005 (Haan) at [0110], Figs. 2, 7, 14.

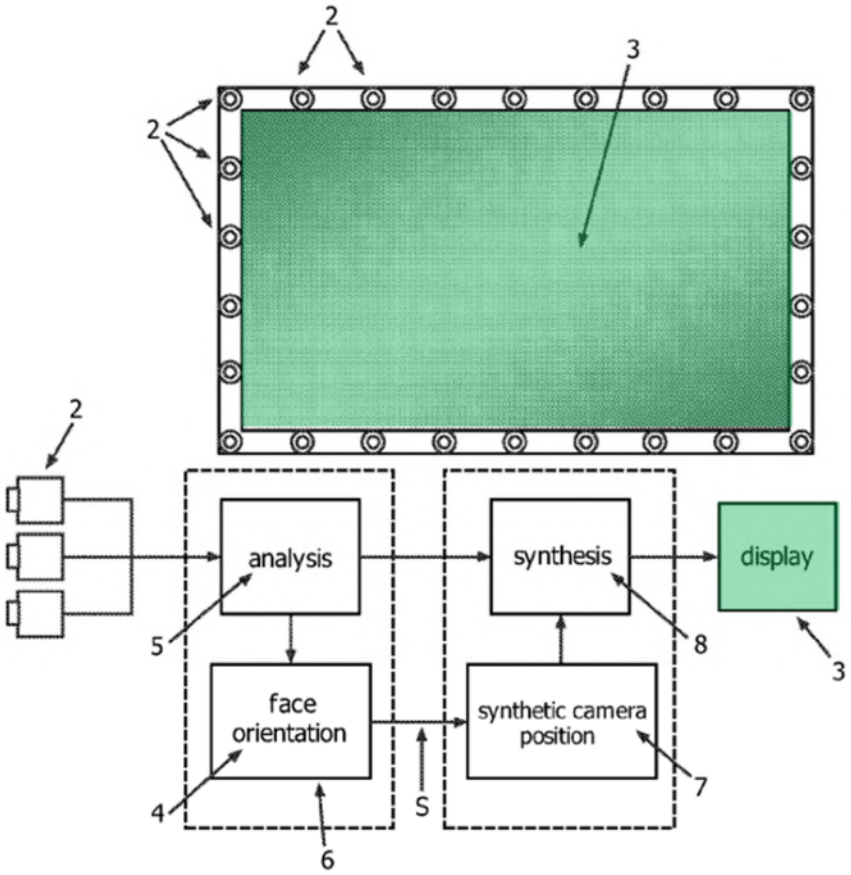


FIG. 2

Ex. 1005 (Haan) at Fig. 2 (annotated).

f. 1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and

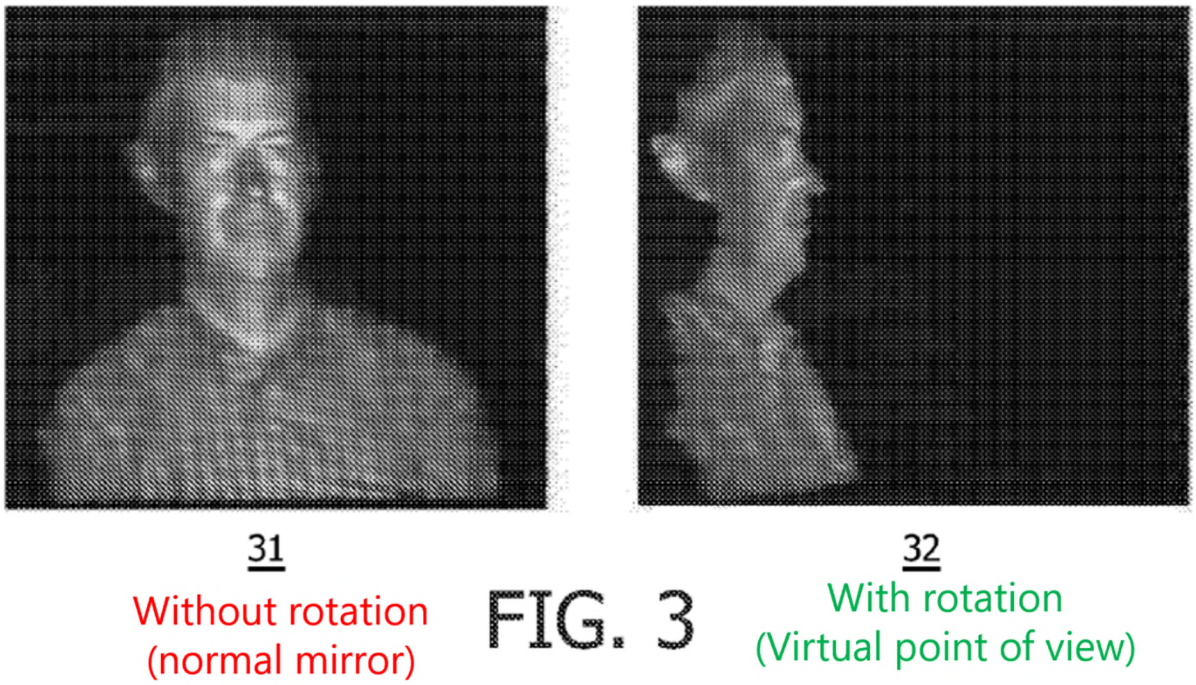
191. Haan in view of Francois renders obvious this limitation. As discussed above, Haan discloses a virtual point of view (i.e., a different point of view) which corresponds to an image "taken from a different angle and a different distance than the actual images captured by the camera." As discussed above, the means to change the point of view includes at least two adaptive mapping transformations—rotation transformation and resizing transformation—making the virtual point of view different than the actual images captured by the camera. *See* §XI.B.d.i, ¶173. As discussed, a POSITA would have understood the synthesis disclosed in Haan is a transformation mapping, as claimed. *See supra* §XI.B.d.i, ¶¶177-178.

i. "wherein the different point of view corresponds to an image taken from a different angle… than the actual images captured by the camera"

192. Haan describes the "different point of view corresponds to an image taken from a different angle." For example, as discussed above, Haan describes that the adaptive rotation transformation rotates the viewpoint so that the display image appears to be taken from a different angle than the actual camera position. Ex. 1005 (Haan) at [0060]-[0063], [0073]. As discussed above, Haan explains a point of view is the same as a camera position (Ex. 1005 (Haan) at [0061]: "the camera position, i.e. the point of view from [which] the image is recorded") and further explains that

rotation transformation, e.g., increasing the turning angle, "means in fact that one shifts the point of view from which the image is recorded." Ex. 1005 (Haan) at [0073]; *see also supra* §XI.B.d.i, ¶¶173-175.

193. As shown in Figure 3 of Haan below, the rotation transformation changes the point of view. Ex. 1005 (Haan) at [0061]. For example, as shown in Figure 3, "a 12 degree turn of the head causes the head, as displayed on the display screen, to be turned 84 degrees." Ex. 1005 (Haan) at Fig. 3, [0061].



Ex. 1005 (Haan) at Fig. 3 (annotated).

194. Haan describes at least two algorithms to increase the turning angle—one linear and the other non-linear. For the linear algorithm, Haan describes that "a turn of x degrees of the head will turn the head in the display by x*y degrees where

y is constant." Ex. 1005 (Haan) at [0020]. Further, with respect to the non-linear algorithm, Haan describes that "the increase in angle increases as the head is turned more." Ex. 1005 (Haan) at [0020].

195. Figure 6 (annotated) below shows the linear and non-linear exemplary corrections contemplated by Haan for the actual turning of the head and displayed turning of the head.

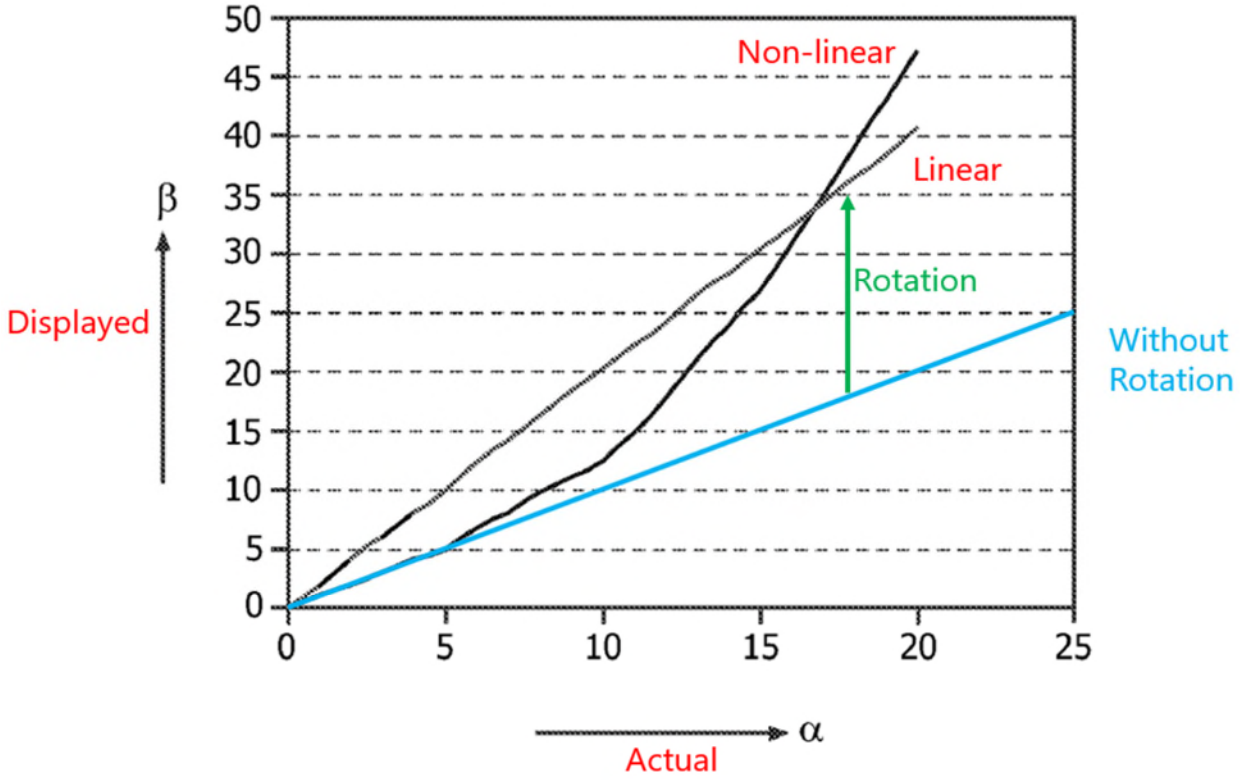


FIG. 6

Ex. 1005 (Haan) at Fig. 6 (annotated).

196. Further, Figures 2, 7, and 14 of Haan show additional configurations that may implement rotation transformations. Ex. 1005 (Haan) at Figs. 2, 7, 14.

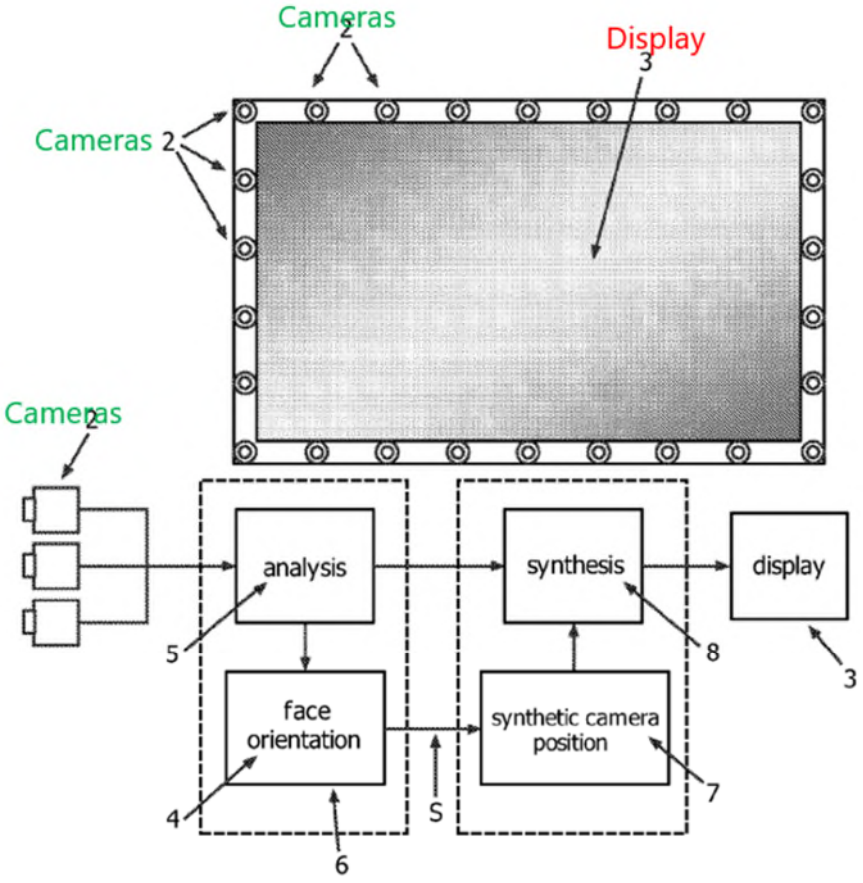


FIG. 2

Ex. 1005 (Haan) at Fig. 2 (annotated).

197. For example, as shown in Figure 2, the images taken from the cameras are analyzed to determine face orientation 4, which is then processed to generate a synthetic camera position 7, that is a virtual camera position, or a "different point of view," as claimed in the '110 Patent. Ex. 1005 (Haan) at [0058]. Based on the virtual camera position, the mirror device in Haan synthesizes, or combines, actual images

to form "a single image signal from a virtual point of view," which is then displayed on a screen. Ex. 1005 (Haan) at [0058]. A POSITA would have understood that the above rotation transformation is a transformation mapping that takes the form of a planar homography transformation ($H= KRK^{-1}$). *See supra* §VI.B.e, ¶64.

ii. "wherein the different point of view corresponds to an image taken from … a different distance than the actual images captured by the camera"

198. Haan discloses "wherein the different point of view corresponds to an image taken from … a different distance than the actual images captured by the camera." The adaptive resizing transformation described in the '110 Patent changes magnification of an image so that the virtual viewpoint of the image appears to have a different distance than the actual distance from the camera to the viewer. Ex. 1001 ('110 Patent) at 3:21-29. The system in Haan exaggerates the movement of a viewer by applying resizing transformation, such as to enlarge the image to make the user appear closer. *See* Ex. 1005 (Haan) at [0022]-[0023], [0072], [0076].

199. Haan explains its "magnification" (i.e. amplification) is "equivalent to moving a camera" and "changing the point of view." Ex. 1005 (Haan) at [0073]. Haan further explains "magnification is in fact equivalent to moving a camera with fixed magnification closer to the object, so all applications can be seen, in fact as changing the point of view, or in other words the position of the camera." Ex. 1005 (Haan) at [0073].

200. In Haan, the magnification of a displayed image adapts to the distance to the mirror. For example, as shown in Figure 8, there is "an exemplary relation between the magnification M of the image and the distance d in centimeter to the mirror." Ex. 1005 (Haan) at [0075], Fig. 8. Figure 8 depicts how a user moving towards the mirror results in the system exaggerating the movement, causing the user to appear closer to the camera than he/she is. Figure 8 shows the relationship "between the magnification *M* of the image and the distance *d* in centimeter to the mirror." Ex. 1005 (Haan) at [0075], Fig. 8.

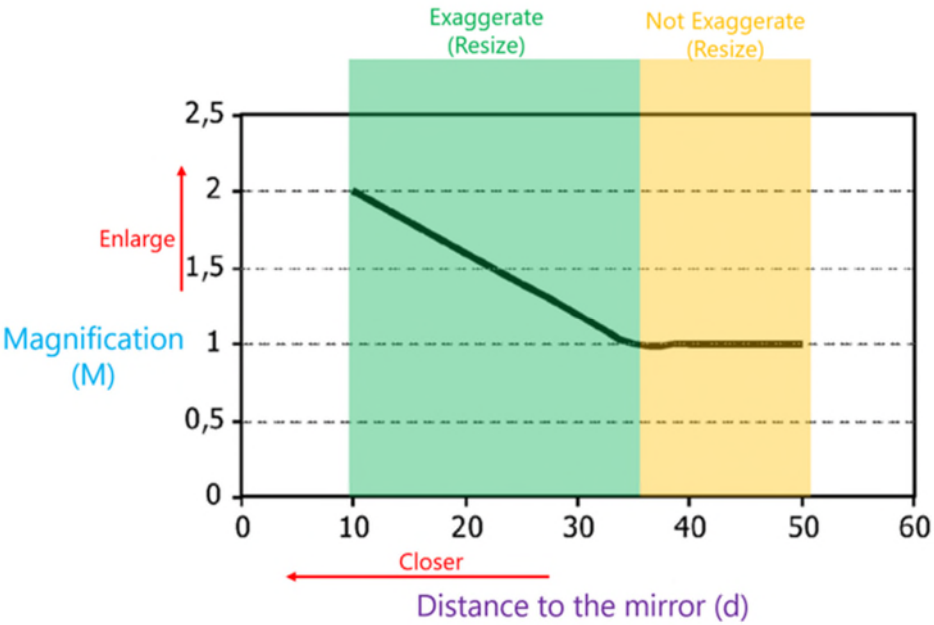
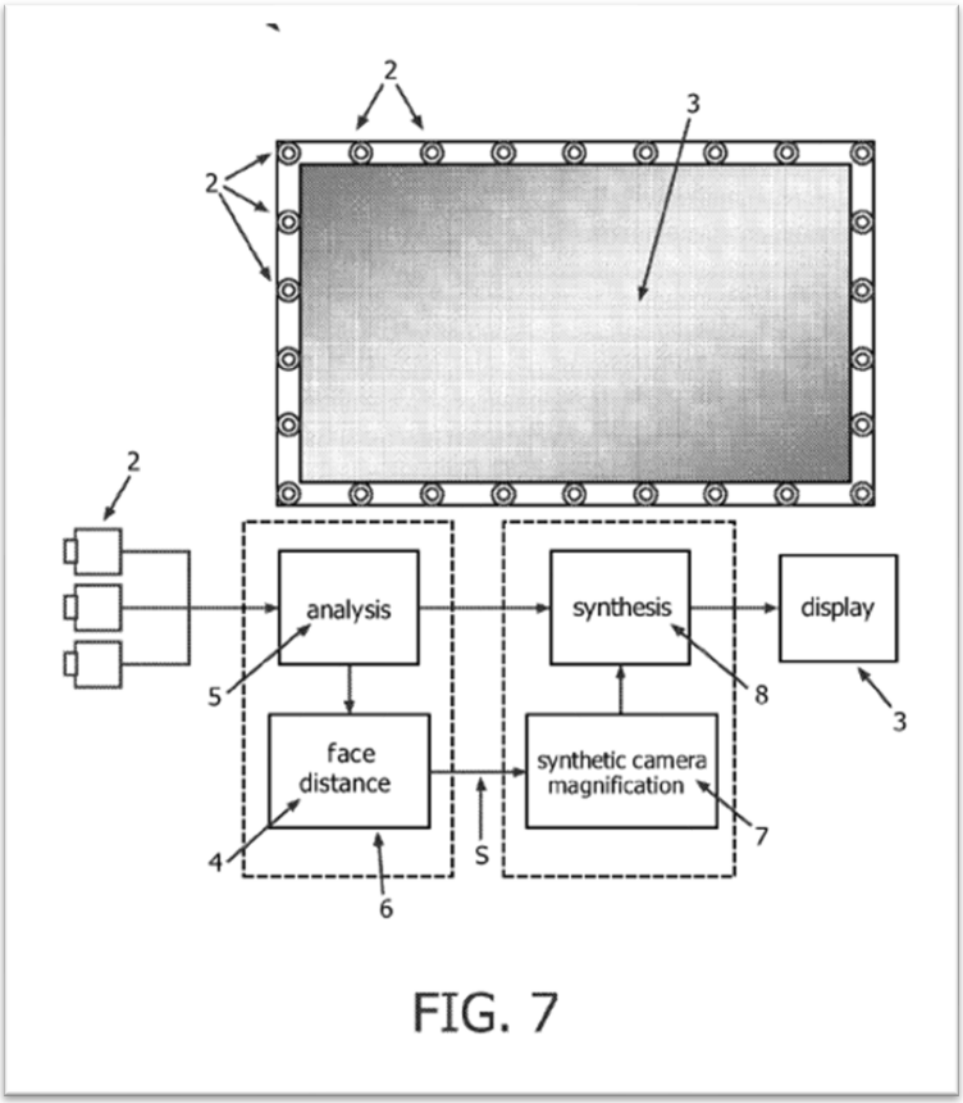


FIG. 8

Ex. 1005 (Haan) at Fig. 8 (annotated).

201. To apply this resizing transformation, the mirror may comprise a face distance determining means and a synthetic camera magnification means. Ex. 1005 (Haan) at [0073], Fig. 7.



Ex. 1005 (Haan) at Fig. 7.

202. Further, as shown in Figure 7 above, the images taken from the cameras are analyzed to determine a distance from the face of a viewer to the mirror device.

Ex. 1005 (Haan) at [0073]. Haan discloses ways to determine the face distance from the mirror device, including using the size of the object, parallax, or ultrasound. Ex. (Haan) 1005 at [0074]. The determined distance is processed to generate a synthetic camera magnification, which is fed to the synthesis means to adjust a magnification of the image for display. Ex. 1005 (Haan) at Fig. 7, [0073]. Thus, Haan discloses "wherein the different point of view corresponds to an image taken from … a different distance than the actual images captured by the camera."

iii. Haan discloses both angle and distance transformations

203. Haan discloses image transformations wherein the angle and distance are modified simultaneously. Haan states that "rotation as well as magnification of the image can be done simultaneous[ly]." Ex. 1005 (Haan) at [0088]. Further, Haan discloses that "using several cameras[,] one can, by synthesis of the various images, construct a virtual camera position and/or magnification." *See, e.g.*, Ex. 1005 (Haan) at [0058]; *see also* Ex. 1005 (Haan) at [0095], [0014].

204. A POSITA would have further understood that the adaptive rotation transformation and resizing transformation can both be performed in the synthesis means 8, including simultaneously, to generate a modified image, as was well known in the art as of the '110 Patent's December 18, 2012 priority date. Ex. 1005 (Haan) at [0027], [0058]; *see supra* §§ VI.B.c-VI.B.f.

g. <u>1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.</u>

205. Haan in view of Francois renders obvious "during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera."

206. The '110 Patent describes a mirror tracking mode when "the user enters a specified zone in front of the field of view of the camera." Ex. 1001 ('110 Patent) at 25:36-43. When the user enters this specified zone (which is within a certain distance between the mirror and the user), the mirror goes into mirror tracking mode such that the system may apply transformations to the image and generate modified images with virtual points of view. Ex. 1001 ('110 Patent) at 25:41-43, 26:22-23. In contrast, the '110 Patent describes that the system operates in an idle mode when the user is not within the specified zone or outside of a certain distance from the mirror. Ex. 1001 ('110 Patent) at 25:36-39, 25:44-56. In the idle mode, the system may not apply a mapping transformation to generate modified images. Ex. 1001 ('110 Patent) at 25:36-39.

207. Similarly, Haan discloses a camera's field of view changes depending on the user's distance from the mirror device (i.e., the user's position/orientation).

Ex. 1005 (Haan) at [0097]. Haan describes that its mirror device may switch between a display mode (or a transparent mode if a semi-transparent layer is used) and a reflective mode. Ex. 1005 (Haan) at [0080]-[0081], [0090]-[0092]. Haan explains that "x stands for the angle of orientation or the distance at which the mirror device changes from reflective mode to display mode." Ex. 1005 (Haan) at [0082]. If a viewer is "within a range of angles and/or positions," the mirror operates in a mirror mode wherein transformations are performed, as described above. Ex. 1005 (Haan) at [0093]; *see also supra* §§XI.B.c, XI.B.d, XI.B.e (1[b], 1[c], 1[e]). For example, if the distance between the face of a viewer and the camera is within 35 cm, the system may operate in a mirror mode, applying resizing transformations. Ex. 1005 (Haan) at [0075]-[0076]. However, if a user is not within the range of angles or positions, the system is in idle mode, acting as an ordinary display and/or as a regular reflective mirror. Ex. 1005 (Haan) at [0091]-[0094].

208. Further, Haan discloses that the processor of the mirror device utilizes the angle of the images captured by the camera and applies parametric input (e.g., signal S, face orientation, face distance) for a mapping transformation that changes imaging perspective (e.g., point of view and field of view) of the images captured by the camera. Ex. 1005 (Haan) at [0058], [0097], Figs. 2, 7. As discussed above, the change of the angular aspect of the point of view requires applying a rotation transformation, which utilizes an angle of the images (e.g., turning angle of viewer's

head or the angle of a camera with respect to the viewer). Ex. 1005 (Haan) at [0058], [0097]; *see also supra* §XI.B.d.i, ¶¶173-175; §XI.B.f.i, ¶¶192-195.

209. Haan also teaches applying parametric input. The ’110 Patent does not provide a definition for “parametric input.” However, the specification suggests that parametric input may include variables associated with a user feature (“e.g., weight, height, etc.”). Ex. 1001 (’110 Patent) at 29:64-30:1. Further, the ’109 Patent suggests input parameters may include “position of [the] user,” “camera position,” “capturing angle,” or the like. Ex. 1021 (’109 Patent) at 9:26-52.

210. Haan describes that “the view point from which the image is recorded is in dependence on said indication signal (S).” Ex. 1005 (Haan) at [0099],[0058] (“The position of the virtual view point is dependent on the determined face orientation, i.e., signal S.”), claim 1. Signal S is the output of the determining means 4, indictive of the position of a viewer. Ex. 1005 (Haan) at [0099] (“said means outputting [an] indication signal (S) to indicate the position of the body part of a viewer with respect to the frame”), Figs. 2, 7, 11, 14. As explained above, the position information, including angular or distance data, is applied in a rotation or resizing transformation in the synthesis unit, which takes input from the means 4. Ex. 1005 (Haan) at Figs. 2, 7, 11, 14. Thus, Haan discloses or at least renders obvious “during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that

changes imaging perspective of the images captured by the camera." Moreover, a POSITA would have been motivated to perform this step on a processor in view of Francois, given that Haan discloses the means for rotation and resizing are computer programs, to ensure compatibility and implement the virtual mirror system in Haan. *See supra* §XI.A; Ex. 1005 (Haan) at [0058], [0087], [0099], [0105]. A POSITA would have done so with a reasonable expectation of success, as Francois discloses a known processor that is compatible with virtual mirror systems like those disclosed in Haan. *See supra* §XI.A.

C. CLAIM 2. THE METHOD OF CLAIM 1, WHEREIN THE MAPPING TRANSFORMATION COMPRISES A COMBINATION OF AT LEAST TWO OF THE FOLLOWING TRANSFORMATIONS: VERTICAL FLIPS ABOUT A VERTICAL AXIS,[4] A PROJECTION TRANSFORMATION, ROTATION, RESIZE, TWO-DIMENSIONAL OFFSET, OPTIC DISTORTION

[4] To the extent that "vertical flips about a vertical axis" has discernible meaning, it should be understood to include at least the '110 Patent's related disclosure in the specification, which includes "flipping the image about a vertical axis to reverse right and left sides of the image." Ex. 1001 ('110 Patent) at 3:21-24; *see also, e.g.*, describing transformations including flipping, translation, and resizing for zoom in/out, EX1021 ('109 Patent), 16:46-17:28, 22:1-13, 23:41-24:9, 24:51-53, 25:48-51, 29:22-58. And the prior art discloses such flipping as discussed in this section and others. *See, e.g., supra* §§VI.B.b, XII.C, XIII.C.

CORRECTION, AND POLYNOMIAL MAPPING ON PART OR ENTIRE CAPTURED IMAGE.

211. Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders obvious wherein the mapping transformation comprises a combination of at least two of the following transformations: flips about a vertical axis, a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image. *See, e.g., supra* §§XI.B.c, XI.B.d, XI.B.e (1[b], 1[c], 1[e]).

212. Haan primarily discloses two transformations for changing the point of view: rotation transformation, which changes the angle of the point of view, and resizing transformation, which changes the distance of the point of view. Ex. 1005 (Haan) at [0073], [0070]-[0072], [0074]-[0076]. Thus, the mapping transformation in Haan comprises a combination of at least the resizing and rotation transformations. A POSITA would also have found it obvious to apply transformation mapping to the captured images to carry out the synthesizing of images from multiple cameras. Further, a POSITA would have understood that a synthesis process adjusting the point of view is a type of transformation. *See supra* §VI.B.c.

213. Further, to the extent Patent Ower argues the claim requires two transformations to be implemented, Haan does so. For example, Haan discloses an embodiment having "[c]ombinations of the two, in which both orientation and distance are measured." Ex. 1005 (Haan) at [0095].

214. Moreover, Haan teaches flips about a vertical axis and that the display may act as a normal mirror. Ex. 1005 (Haan) at Title, Abstract, [0037]. A POSITA would have understood the images in a virtual mirror are left/right flipped about a vertical axis. *See supra* §VI.B.b.

215. Additionally, Haan renders obvious this claim in view of the knowledge of a POSITA because a POSITA would have understood the claimed mapping transformation is disclosed in Haan. *See supra* ¶¶55-57, 61-64, 67, 177, 186.

D. CLAIM 3. THE METHOD OF CLAIM 1, WHEREIN DURING IDLE MODE THE CAMERA CONTINUOUSLY SENDS CAPTURED IMAGES TO THE PROCESSOR, AND THE USER PRESENCE IS IDENTIFIED BY THE PROCESSOR DETECTING A BODY PART IN THE IMAGES.

216. Haan in view of Francois, renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders obvious "wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images."

217. Haan discloses a mirror tracking mode (display mode, as explained with regard to limitation 1[f]), and an idle mode (reflective mode). *See supra* XI.B.g. In

idle mode, the mirror behaves like an ordinary mirror, where no resizing is applied. Ex. 1005 (Haan) at [0076], Fig. 8.

218. While the rotation and resizing transformations are not applied in idle mode, as shown in Figure 8, Haan explains the mirror nevertheless continuously monitors the distance and orientation of the viewer. Indeed, Haan discloses the idle mode is activated when the detected distance of the user exceeds a certain threshold and is deactivated when the detected distance falls below that certain threshold. Ex. 1005 (Haan) at [0076], Fig. 8. For example, if the distance from the user to the mirror is larger than a threshold value (e.g., 35 cm), the magnification of the mirror device may be set to 1, and no resizing is applied. Ex. 1005 (Haan) at [0076]. Haan describes that no resizing is applied "by using the device as a simple mirror, i.e. not displaying an image on the display device, but [rather] using the device in the reflective mode." Ex. 1005 (Haan) at [0076].

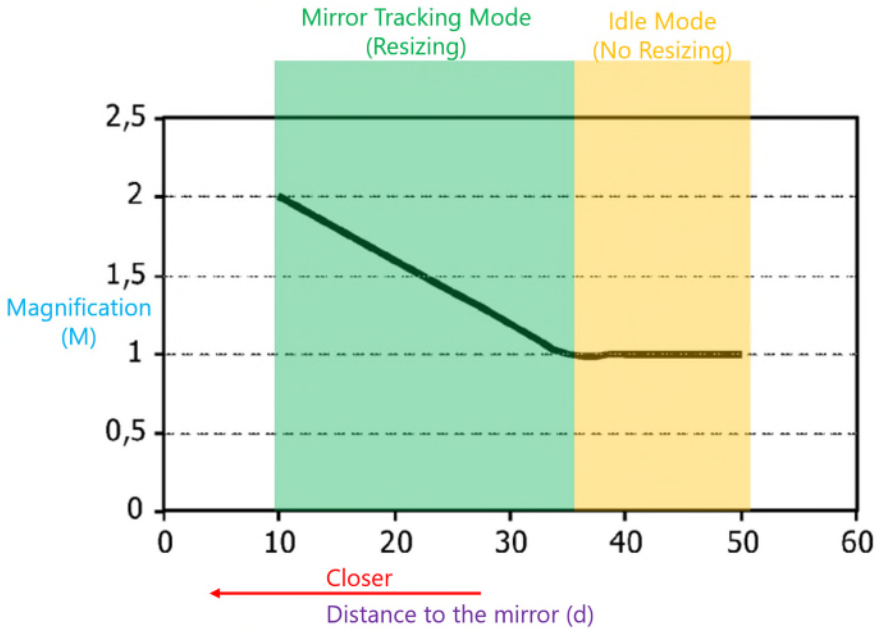


FIG. 8

Ex. 1005 (Haan) at Fig. 8 (annotated).

219. As shown in Fig. 8, for example, the distance between the user and the mirror device is continuously measured even when the distance is larger than 35 cm. This allows the mirror device to automatically switch back to the interactive mirror display mode when the distance is smaller than the threshold value of 35 cm, for example. Fig. 8 shows the distance is continuously measured when the distance is larger than 35 cm, where the resizing transformation is not performed. Further, Fig. 6 shows that the face orientation is continuously measured when the orientation is lower than 5 cm, where the rotation transformation is not performed in the non-linear configuration. Therefore, the mirror device of Haan may enter into idle mode (e.g., reflective mode) when face orientation is lower than 5, and rotation transformations

are not performed. Ex. 1005 (Haan) at Fig. 6. Thus, even in the idle mode, when the mirror device is just acting like a regular reflective mirror, the camera continuously sends captured images to the processor.

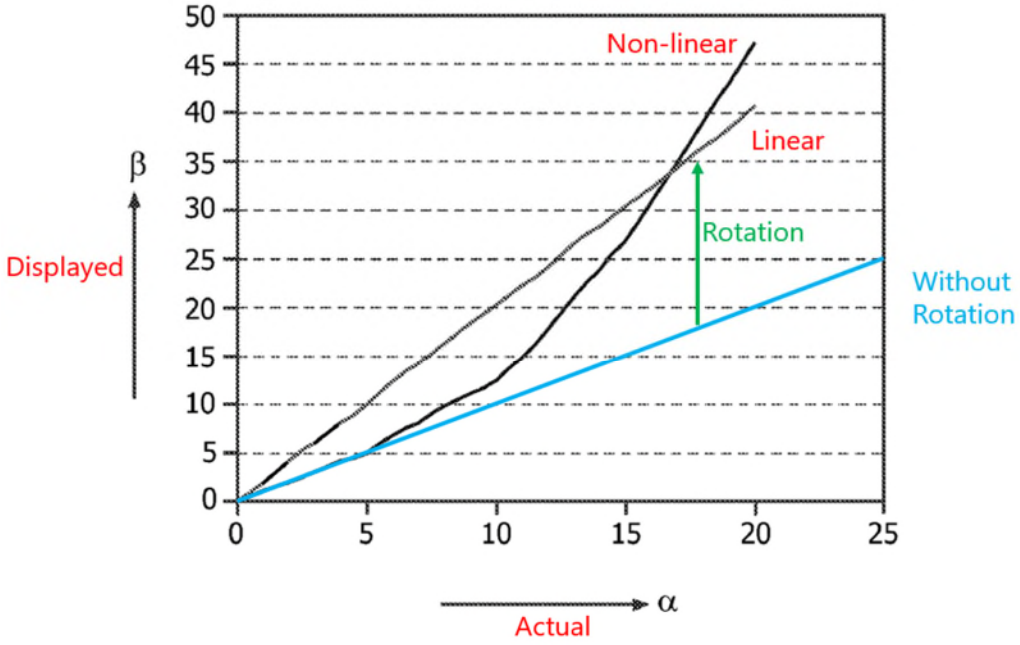


FIG. 6

Ex. 1005 (Haan) at Fig. 6 (annotated).

220. Because the user orientation and/or distance need to be measured to exit idle mode, a POSITA would have found it obvious to detect the presence of the user during idle mode. Further, as discussed above, when the distance or orientation of a body part is detected, the presence of the user is also detected. *See supra* §XI.B.c. Thus, Haan in view of Francois discloses that the user presence is identified by the processor detecting a body part in the images, and further discloses or renders

obvious "wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images."

221. Additionally, Haan renders obvious this claim in view of the knowledge of a POSITA because a POSITA would have understood that a camera would need to continuously send captured images to the processor so the camera can determine when to switch modes. *See supra* §XI.B.g, ¶¶205-207.

E. CLAIM 4. THE METHOD OF CLAIM 3, WHEREIN THE BODY PART COMPRISES ONE OF LEGS, FACE, EYES, AND HANDS.

222. Haan in view of Francois renders obvious the method of claim 3. *See supra* §XI.D. Further, Haan in view of Francois renders obvious "wherein the body part comprises one of legs, face, eyes, and hands." As discussed, Haan discloses detecting a body part like the "head or face of a person" in the images. Exhibit 1005 (Haan) at [0013]. *See supra* §XI.B.c, ¶¶165-166.

223. Further, Haan renders obvious this claim in view of the knowledge of a POSITA because a POSITA would have understood a body part may comprise a leg, face, eyes, or hands.

F. CLAIM 5. THE METHOD OF CLAIM 1, FURTHER COMPRISING RECEIVING USER INPUT AND OPERATING THE PROCESSOR TO VARY IMAGE MAGNITUDE OR FIELD-OF-VIEW OF THE MODIFIED IMAGES ACCORDING TO THE USER INPUT.

224. Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, as explained below, Haan in view of Francois renders obvious "receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input." Patent Owner alleges in its infringement contentions "[u]ser input includes the user's movements captured by the camera." Ex. 1019 (Excerpts from EyesMatch Ltd.'s Disclosure of Asserted Claims and Infringement Contentions) at 63. The captured user movement is one type of user input under at least Patent Owner's interpretation.

225. Haan describes a display system that functions as a mirror which has "a means for determining a position of a body part with respect to the display screen of the display device [] of a viewer in front of the display screen." Ex, 1005 (Haan) at [0009]. The determined position may comprise "the distance to the display screen," "angle of turn of the head [or] torso in respect of the display screen" and "movements to and/or from the display screen" (i.e., user's movements captured by the camera, or "user input" as described by PO). Ex. 1005 (Haan) at [0014]. Further, Haan discloses an enhanced mirror wherein an image is magnified or enhanced to view some parts better. For example, when a user is looking straight or turning by a small

angle, the image is produced using a reflective surface rather than a synthesized mirror image. Ex. 1005 (Haan) at [0030], [0073], [0076], Fig. 9. Haan discloses that the processor of the mirror device utilizes the position of the user (e.g., "angle of turn of the head [or] torso," "distance to the display screen") in the images captured by the camera and applies parametric input (e.g., signal S, face orientation, face distance), further corresponding to user input, for a mapping transformation that changes imaging perspective (e.g., image magnification, field of view, point of view) of the images captured by the camera. Ex. 1005 (Haan) at [0058], [0097]; *see also* Ex. 1005 (Haan) at [0015]-[0018], [0022]-[0023], Figs. 2, 7.

G. CLAIM 6. THE METHOD OF CLAIM 5, WHEREIN THE PROCESSOR VARIES THE IMAGE MAGNITUDE OR FIELD-OF-VIEW MAPPING TO DISPLAY A PARTIAL OR FULL BODY OF THE USER IN THE MODIFIED IMAGES.

226. Haan in view of Francois renders obvious the method of claim 5. *See supra* §XI.F. Further, Haan in view of Francois renders obvious "wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images." For example, as discussed above, Haan discloses a processing means to increase the image magnification or vary the field of view to display a partial body of the user in the modified images. *See supra* §§XI.B.c-XI.B.e, XI.B.g (1[b]-1[d], 1[f]); Ex. 1005 (Haan) at [0072]-[0073], [0097], Figs. 1, 3, 8, 9, 10.

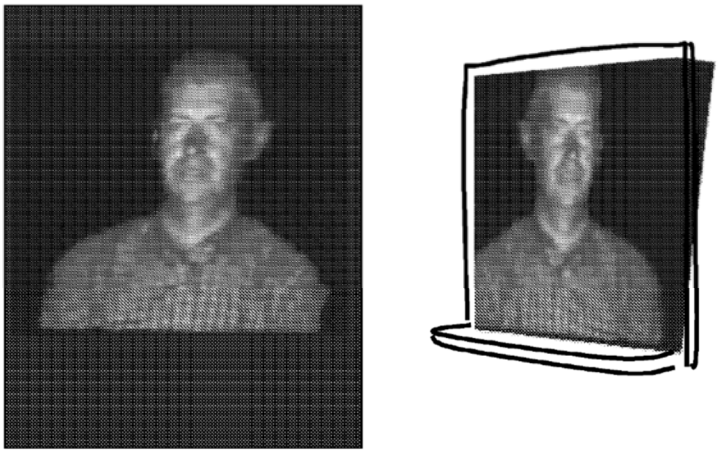


FIG. 1

Ex. 1005 (Haan) at Fig. 1; *see also* Ex. 1005 (Haan) at Figs. 3, 8, 9, 10.

227. Haan discloses a zoom out effect achieved "[b]y using a magnification smaller than 1." Ex. 1005 (Haan) at [0076]. Therefore, a POSITA would have found it obvious to display a full body of a user when the user zooms out on the images to show the areas of the user's image that is just out of sight. *See* Ex. 1005 (Haan) [0004] ("[a] frequently occurring problem with mirrors is that the area one wants to see is just out of sight."); *see also* Ex. 1006 (Rosenberg), [0050], Fig. 11.

H. CLAIM 10. THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO AUGMENT THE MODIFIED IMAGES BY OVERLAYING VIRTUAL OBJECTS ON A USER'S IMAGE APPEARING WITHIN THE MODIFIED IMAGES.

228. Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders obvious "operating the processor to augment the modified images by overlaying virtual objects on a user's

image appearing within the modified images." For example, Haan discloses that the "generated image is the image of the viewer" and "[a] computer program would overlay and adjust the recorded image to simulate applied make-up." Ex. 1005 (Haan) at [0113]. For example, Haan describes that "the viewer can [then] take a good look, and at various angles . . . or at a magnification to judge the effect of the application of make-up." Ex. 1005 (Haan) at [0113]. Thus, Haan discloses overlaying make-up to user images.

229. Further, Haan discloses "a shown mirror image is the image of the viewer as taken, but computer altered to provide various different hair styles and color. The viewer can then judge various hairstyles and colors, at various angels [sic]." Ex. 1005 (Haan) at [0114]. Moreover, Haan explains that "the mirror image shown is a composite of the recorded image with some artificial compounds." Ex. 1005 (Haan) at [0118]; *see also* Ex. 1005 (Haan) at [0115]-[0117]. Thus, Haan in view of Francois renders obvious "operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images."

I. CLAIM 11. THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO TRANSMIT THE MODIFIED IMAGE OVER A NETWORK.

230. As discussed above, Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders

obvious "operating the processor to transmit the modified image over a network." Francois discloses transmitting the modified image over a network. Specifically, Francois discloses "encoding, compression and networking (send) on the emitting side, [and] networking (receive), decompression and decoding on the receiving side." Ex. 1009 (Francois) at §4.4. Figure 15 below illustrates an exemplary implementation.

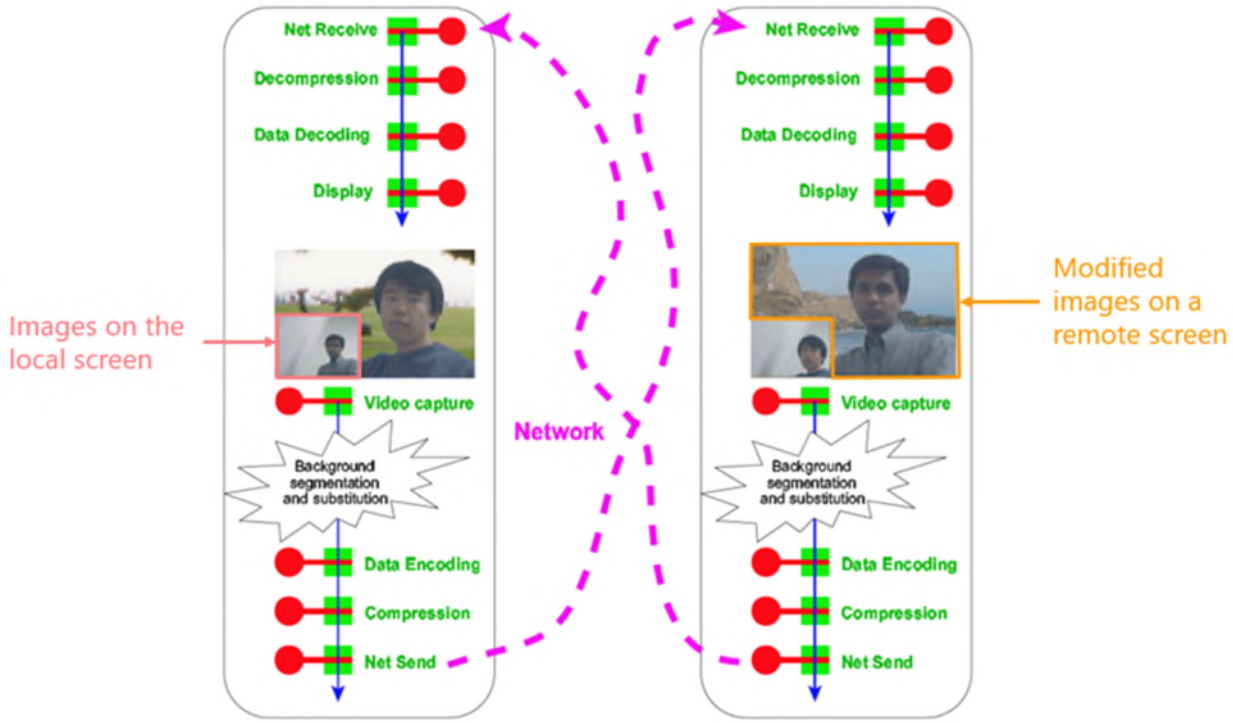


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15 (annotated).

231. As shown in the exemplary implementation above, the Haan-Francois system transmits images captured by the camera to a remote screen over a network.

Ex. 1009 (Francois) at Fig. 15, §4.4. The received image is then decompressed, decoded, and displayed on the screen at the receiving side. Ex. 1009 (Francois) at Fig. 15, §4.4.

232. Further, Francois explains in detail using an IMSC Communicator for remote collaborative data sharing. Ex. 1009 (Francois) at §4.4. Such communication system in Francois may use popular architectures for communication applications including "Client/Server and Peer To Peer." Ex. 1009 (Francois) at §4.4.

233. As discussed, a POSITA would have recognized that adding Francois' network communication module (e.g., IMSC Communicator) (Ex. 1009 (Francois) at §4.4) with Haan's virtual mirror system would enable a user to transmit local modified images to a remote device for display, enabling a robust communication functionality for a user, including sharing modified user images during video calls (*see, e.g.*, Ex. 1005 (Haan) at [0037]; Ex. 1009 (Francois) at Fig. 15). *See supra* §XI.A, ¶¶142-143. As explained above, a POSITA would have found it obvious to use the network feature from Francois with the virtual display system from Haan to create a real-time interactive system capable of transmitting the modified images to a remote screen via a network for enhanced video telepresence to expand on known benefits of Haan and offer an enhanced user experience. *See supra* §XI.A. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XI.A.

J. CLAIM 12. THE METHOD OF CLAIM 1, FURTHER COMPRISING DISPLAYING THE MODIFIED IMAGE ON A REMOTE SCREEN FOR VIDEO TELEPRESENCE.

234. As discussed above, Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders obvious displaying the modified image on a remote screen for video telepresence as similarly discussed above for claim 11 and as shown below in Figure 15. *See also, e.g., supra* §XI.A, ¶¶142-143.

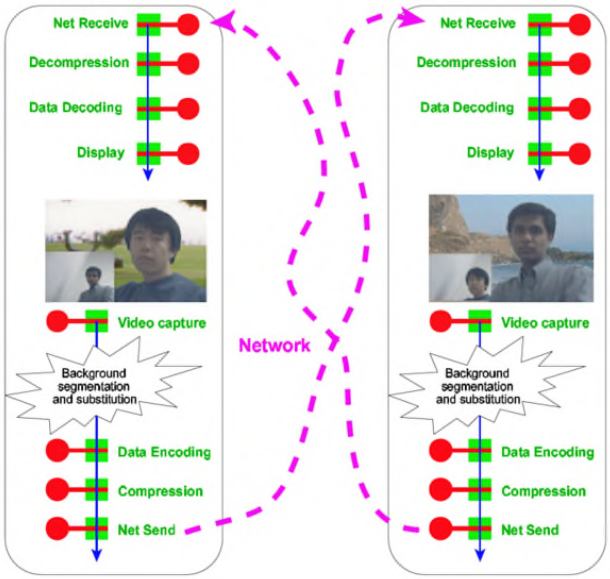


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15.

As explained above, a POSITA would have found it obvious to use the network feature from Francois with the virtual mirror system in Haan to create a real-time interactive system capable of transmitting the modified images to a remote screen

via a network for enhanced video telepresence to expand on known benefits of Haan and offer an enhanced user experience. *See supra* §XI.A. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XI.A.

K. CLAIM 16. THE METHOD OF CLAIM 1, WHEREIN THE PROCESSOR OPERATES ON THE IMAGES FROM THE CAMERA TO GENERATE BACKGROUND SEGMENTATION, AND REPLACES SEGMENTS ON THE IMAGES BY REPLACEMENT SEGMENTS.

235. As discussed above, Haan in view of Francois renders obvious the method of claim 1. *See supra* §XI.B. Further, Haan in view of Francois renders obvious "wherein the processor operates on the images from the camera to generate background segmentation, and replaces segments on the images by replacement segments," as claimed.

236. Francois discloses "operat[ing] on the images from the camera to generate background segmentation," as claimed. For example, Francois discloses real-time video segmentation and tracking. Ex. 1009 (Francois) at §4.1 ("Real-time video segmentation and tracking"), Figs. 7-9. Specifically, Francois discloses applying "an asynchronous parallel implementation of the segmentation algorithm" to generate a foreground binary mask. Ex. 1009 (Francois) at §4.1, Fig. 7. As shown in Figure 7, the system performs a background comparison using a background model and generates a foreground binary mask.

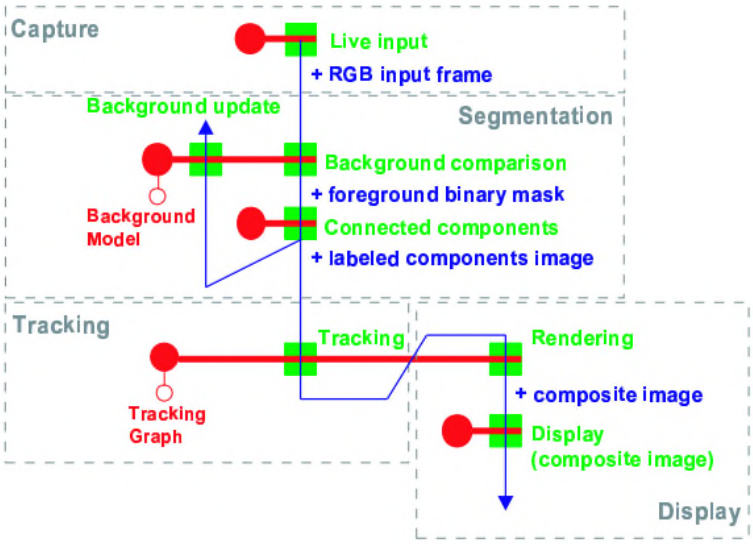


Figure 7: Conceptual graph for real-time color background model-based segmentation and multi-resolution graph tracking application.

Ex. 1009 (Francois) at Fig. 7.

237. Francois discloses "replac[ing] segments on the images by replacement segments," as claimed.

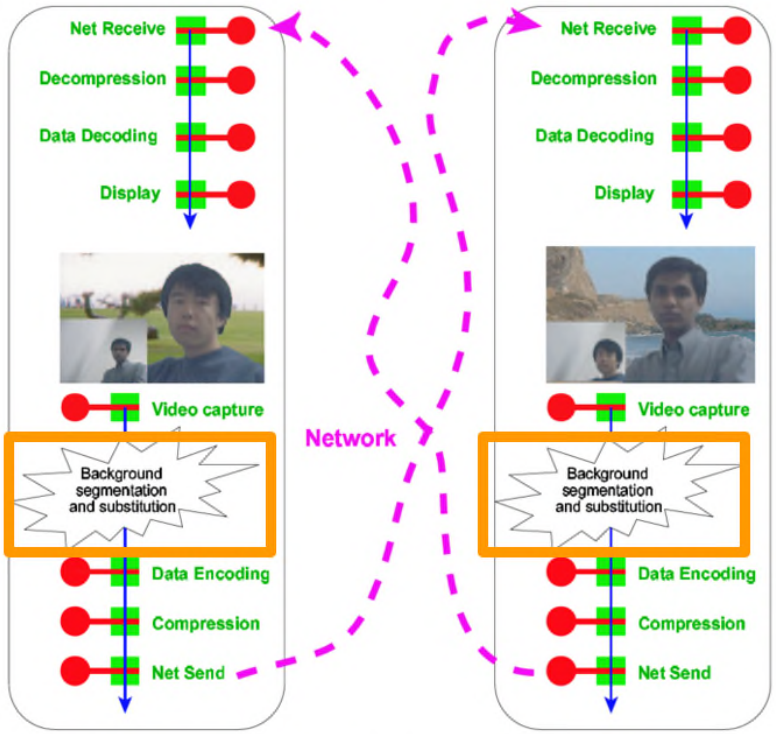


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15 (annotated).

238. Francois discloses rendering the image using the output segmentation to create composite images. Ex. 1009 (Francois) at §4.1. Francois discloses adding "[a] background replacement unit, based on the [background] segmentation by change detection" to the video capture. Ex. 1009 (Francois) at §4.1, §4.4, Fig. 15. As depicted in Figure 15, Francois discloses that this background segmentation can be integrated into a network communication module (e.g., IMSC Communicator) for use during video calls.

239. A POSITA would have found it obvious to combine Francois' network communication module (e.g., IMSC Communicator) (Ex. 1009 (Francois) at §4.4,

Fig. 15) with Haan's virtual mirror system (Ex. 1005 (Haan) at Abstract, [0037]) to enable a robust communication functionality for a user, including enabling enhanced video telepresence. *See supra* §XI.A, ¶¶142-143; *supra* §XI.I. Further, a POSITA would have found it obvious to combine the teachings from Francois with Haan to expand on known benefits of Haan and offer an enhanced user experience. *See supra* §XI.A. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XI.A.

L. CLAIM 17. THE METHOD OF CLAIM 16, WHEREIN THE REPLACEMENT SEGMENTS COMPRISE VIRTUAL BACKGROUND OR BACKGROUND THAT WAS CAPTURED IN A DIFFERENT SETTING.

240. As explained above, Haan in view of Francois renders obvious the method of claim 16. *See supra* §XI.K. Further, Haan in view of Francois renders obvious "wherein the replacement segments comprise virtual background or background that was captured in a different setting." As shown in Figure 15 below, Francois discloses "background segmentation and substitution" and adding "[a] background replacement unit, based on the [background] segmentation by change detection" to the video capture. Ex. 1009 (Francois) at §4.1, §4.4, Fig. 15.

241. As shown in Figure 15, the background segments in the left image are replaced with segments from another stream, resulting in the right image with a background of a beach. Ex. 1009 (Francois) at Fig. 15. A POSITA would have understood that the background substitution in Figure 15 is a virtual background or

background that was captured in a different setting. A POSITA would have found it obvious to use a virtual background when background substitution is applied. For example, a POSITA would have found it obvious to use a virtual background to minimize distraction from the physical background and thereby enable an enhanced video telepresence experience for a user. *See supra* §XI.A, ¶¶142-143; *supra* §XI.I.

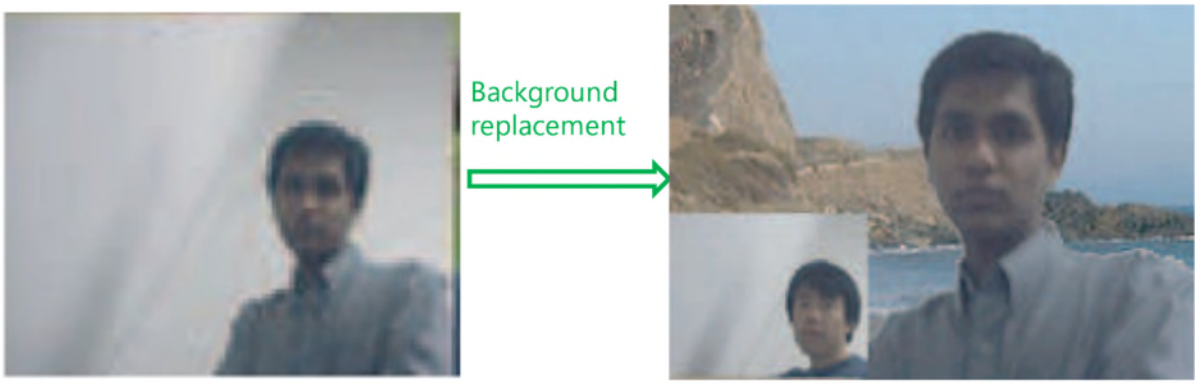


Figure 15: Conceptual graph for an early version of the IMSC Communicator, with support for video (synchronized image and sound).

Ex. 1009 (Francois) at Fig. 15 (partial, annotated).

A POSITA would have found it obvious to combine the teachings from Francois with Haan to expand on the known benefits of Haan and offer an enhanced user experience. *See* §XI.A. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XI.A.

XII. GROUND 2: CLAIMS 1, 2, 5, 6 AND 10 ARE OBVIOUS OVER ROSENBERG IN VIEW OF PARK

A. MOTIVATION TO COMBINE ROSENBERG WITH PARK

242. A POSITA would have found it obvious to combine Rosenberg with Park, as both references are analogous art to the '110 Patent and to each other. Rosenberg, Park, and the '110 Patent all concern the same field of endeavor and are therefore analogous.

243. The '110 Patent discloses a system for "virtualizing a mirror." Ex. 1001 ('110 Patent) at 1:47-49, Technical Field. The virtual mirror system described in the '110 Patent may be deployed in a fitting room setting or on a user smart phone to display on screens images that were captured by cameras and modified. Ex. 1001 ('110 Patent) at Fig. 1; *see also* Ex. 1001 ('110 Patent) at Fig. 5, 2:45-57, 8:41-44.

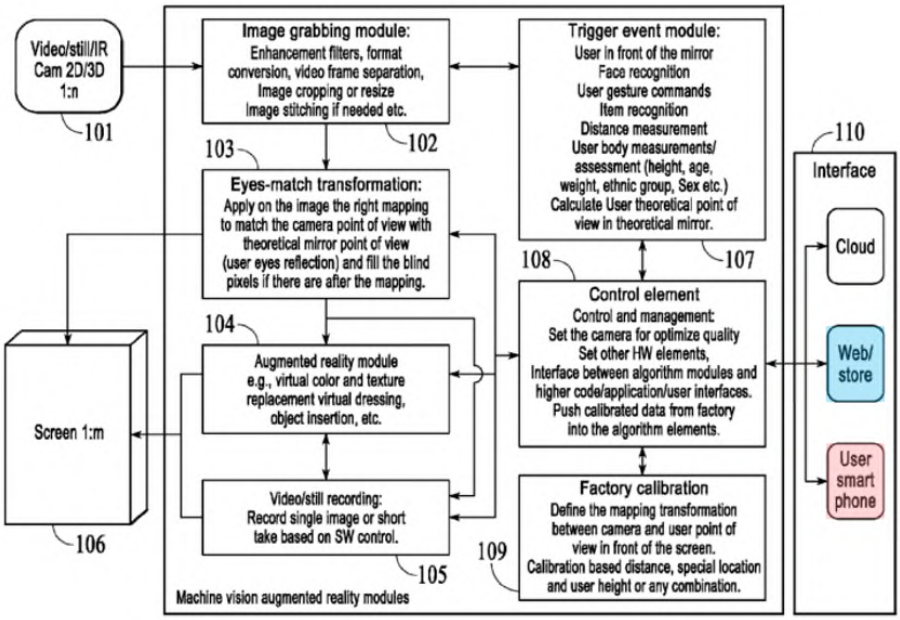


FIG. 1

Ex. 1001 ('110 Patent) at Fig. 1 (annotated).

244. Rosenberg also describes techniques for adapting a flat panel display to display images that mimic the behavior of a conventional mirror. *See, e.g.*, Ex. 1006 (Rosenberg) at [0006], [0077], Figs. 3, 5. As shown in the figure below, Rosenberg's system includes a digital video camera mounted above the screen. *See, e.g.*, Ex. 1006 (Rosenberg) at [0059]-[0060], [0062], [0077]. The camera is centered horizontally across the display and aimed slightly downward such that it will be pointed at any typical user in front of the display. Ex. 1006 (Rosenberg) at [0062]. Rosenberg discloses that, in a digital mirror state, the camera captures a real-time video image of the user, which is "left-right inverted" and scaled to a display size to emulate "a traditional life-size mirror image." Ex. 1006 (Rosenberg) at [0077]. Further, Rosenberg notes that a configuration where the camera is positioned above the screen may be "non-ideal because it provides the user with a view of himself or herself from a somewhat unnatural perspective as compared to a traditional mirror." Ex. 1006 (Rosenberg) at [0027]; *see supra* §X.C.b, ¶111.

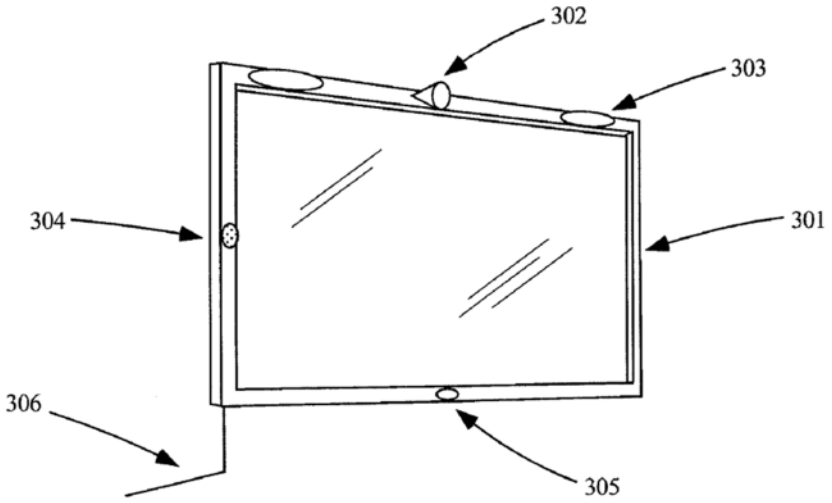


Fig. 3

Ex. 1006 (Rosenberg) at Fig. 3.

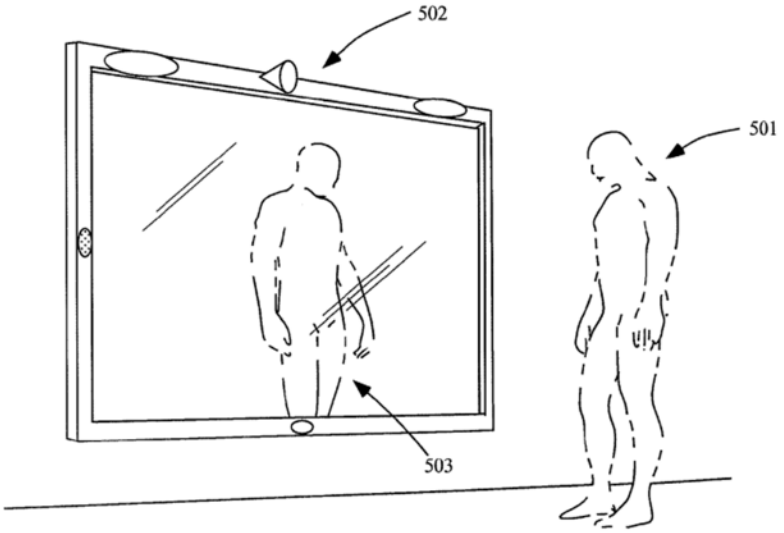
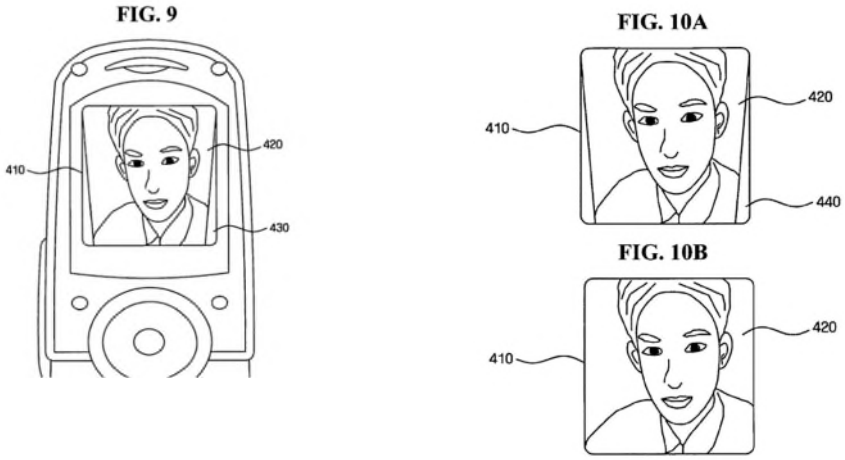


Fig. 5

Ex. 1006 (Rosenberg) at Fig. 5.

245. Similarly, Park discloses techniques for transforming images, including modified images of a user captured by a tilt camera positioned above the display, and more specifically relates to displaying images of a user looking into a smart phone screen. Ex. 1011 (Park) at [0010], Figs. 10A-10C. Park also falls within the field of the ’110 Patent because it concerns displaying camera-captured images of a user looking into a screen. *See, e.g.*, Ex 1011 (Park) at [0010], Figs. 10A-10C.



Ex. 1011 (Park) at Figs. 9, 10A, 10B.

246. Park describes a method for applying transformations to (i.e., compensating) the “captured image in accordance with the position of a module that captures the image of an object.” Ex. 1011 (Park) at [0034]. Specifically, Park teaches that if the camera has an angle “that causes a perspective distortion of the facial image that is not desired by the user, the compensation unit 140 compensates for the generated distortion.” Ex. 1011 (Park) at [0050].

247. A POSITA would have been motivated to combine Rosenberg and Park based on the teachings of both references. For example, as previously explained, Rosenberg recognizes that a configuration in which "the camera is mounted directly above the screen area of the display" may be "non-ideal because it provides the user with a view of himself or herself from a somewhat unnatural perspective as compared to a traditional mirror." Ex. 1006 (Rosenberg) at [0027]; *see supra* §X.C.b, ¶111. A POSITA would have recognized Park discloses a relevant known solution to this problem when describing applying transformations (compensations) to the "captured image[s]" in accordance with camera angles determined by a tilt sensor. Ex. 1011 (Park) at [0034]; *see also* Ex. 1011 (Park) at [0036]. Park discloses that if the camera's angle "causes a perspective distortion of the facial image that is not desired by the user, the compensation unit 140 compensates for the generated distortion." Ex. 1011 (Park) at [0050]; *see also* Ex. 1011 (Park) at [0051]. Further, Park teaches that said compensation may be performed when "the image capturing unit is determined to be above or below the object." Ex. 1011 (Park) at claim 44. Therefore, a POSITA would have been motivated to implement Park's image compensation into Rosenberg to correct image distortions and provide a more realistic mirror image.

248. While Rosenberg discloses a hardware-based solution to prevent image distortion when disclosing an embodiment in which the camera is mounted behind

the display at approximately the user's eye-level (Ex. 1006 (Rosenberg) at [0027]; *see supra* §X.C.b, ¶111), a POSITA would have been motivated to implement Park's software-based image compensation solution because it avoids the need for Rosenberg's optical components and specialized displays, which would add cost to the system. *See* Ex. 1006 (Rosenberg at [0028] (teaching using lenses, mirrors, and/or fiber optics for positioning a camera behind a screen having a removed non-active area). Even in the case in which the camera is located behind the display screen, said behind-screen camera may not match the eye-level of each user, and therefore, a POSITA would still have been motivated to apply Park's transformations to correct any potential distortions from mismatched eye-levels. First, Rosenberg notes "an image that is from a vantage point perfectly straight ahead" is often preferred "because it yields a more realistic replication of a true mirror image." Rosenberg, [0028], [0075]. A POSITA would have been motivated to use Park's image compensation in Rosenberg's first embodiment, where the camera is placed above the display, to correct unnatural perspectives so that images appear to be taken from a position directly in front of the user, "yield[ing] a more realistic replication of a true mirror image." *Id*. Second, a POSITA would have also been motivated to use Park's image compensation in Rosenberg's second embodiment, where the camera is placed behind the display. In such embodiments, Rosenberg's camera "is centered horizontally with respect to the display screen[,]

and at an elevation vertically that generally corresponds with a user's typical eye level." *Id.*, [0061], [0027]. A POSITA would have known that eye level varies from person to person—an image taken at one person's eye level could be below or above an image taken at another person's eye level—causing unnatural distortions in the captured images. Park, [0006], [0014] (teaching camera above or below user causes image distortion). A POSITA would have been motivated to implement Park's image compensation for these embodiments in Rosenberg to also correct the distortions and improve user experience of using the digital mirror like a real mirror. *Rosenberg*, [0027]-[0028], [0061], [0075].

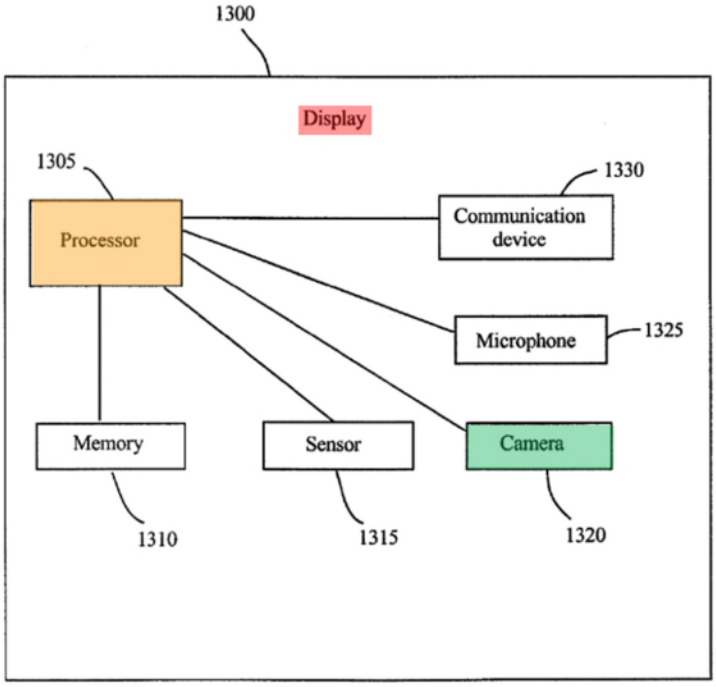
249. A POSITA would have understood implementing Park's image compensation in Rosenberg would simply require adding to Rosenberg's image display system 1) a tilt determining module that detects the camera angle, and 2) a compensation module that compensates the captured images when the camera is positioned at an undesirable angle as disclosed in Park. A POSITA would have understood that these modules would be readily implemented in Rosenberg's system including a programmable processor and memory storing program code, and would have had a reasonable expectation of success in such implementation. Ex. 1006 (Rosenberg), [0035], [0109]. Further, modifying the system disclosed in Rosenberg in view of Park is nothing more than applying a known technique to improve a similar device, as both Rosenberg and Park relate to improving a virtual display

system that includes a camera mounted on an edge of the display above the user. Ex. 1006 (Rosenberg) at Figs. 3, 5, Abstract; Ex. 1011 (Park) at Fig. 9, Abstract. Both Rosenberg and Park disclose devices used for taking images of a user and displaying the captured images to the user, and the image modification techniques disclosed in Park would have resulted in known, predictable results of a mirror image with a natural and accurate visual perspective when applied to Rosenberg. Ex. 1006 (Rosenberg) at Abstract, [0027]-[0028]; Ex. 1011 (Park) at Abstract.

B. CLAIM 1:

a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising

250. To the extent the preamble is limiting, Rosenberg in view of Park renders obvious this limitation. For example, Rosenberg discloses a computerized method for operating a display system, the display system having a camera, a monitor (a display), and a processor. Ex. 1006 (Rosenberg) at Fig. 13. For example, Rosenberg discloses "the display 1300 includes a processor 1305, a memory 1310, a sensor 1315, a camera 1320, a microphone 1325, and a communication device 1330." Ex. 1006 (Rosenberg) at Fig. 13, [0109]; *see also* Ex. 1006 (Rosenberg) at Figs. 3, 5-12, [0006], [0025].



Ex. 1006 (Rosenberg) at Fig. 13.

b. <u>1[a]: operating the camera to generate a stream of images of the user;</u>

251. Rosenberg in view of Park renders obvious operating the camera to generate a stream of images of the user. For example, Rosenberg discloses "one or more digital camera(s) configured to provide real-time images of a user who stands before the display." Ex. 1006 (Rosenberg) at [0026]; *see also* Ex. 1006 (Rosenberg) at [0077] ("The user's image is captured as a real-time video image by a camera upon or within the digital mirror 502 as described previously."); *see also* Ex. 1006 (Rosenberg) at Fig. 5. A POSITA would have recognized that this real-time video image is a stream of images, as claimed.

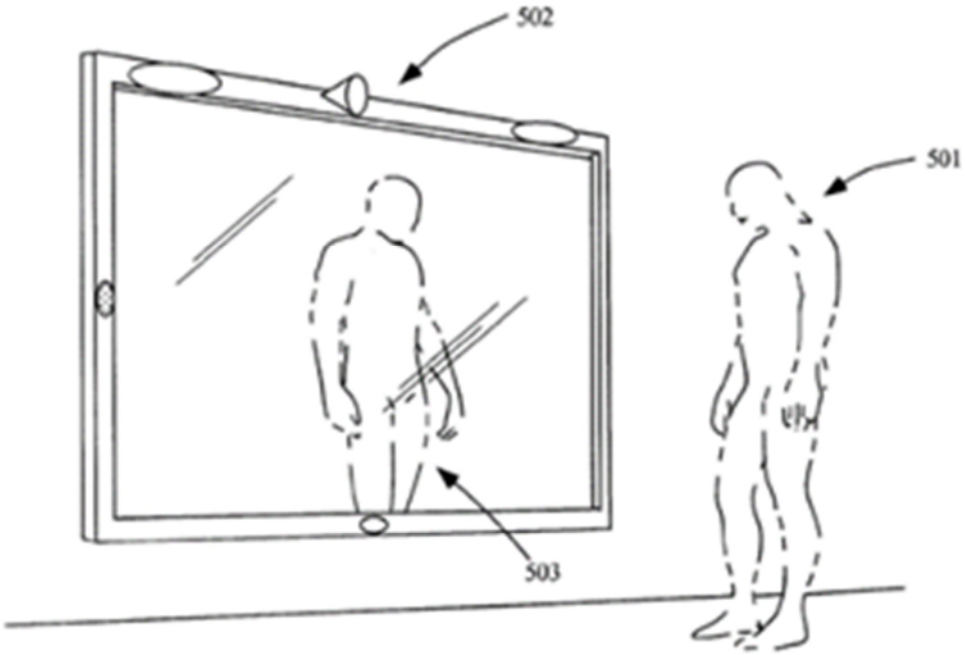


Fig. 5

Ex. 1006 (Rosenberg) at Fig. 5.

c. <u>1[b]: operating the processor to detect the presence of the user in the stream of images;</u>

252. Rosenberg in view of Park renders obvious this limitation. For example, Rosenberg discloses embodiments where the processor works with "one or more position and/or motion sensors for **<u>detecting the presence of a person</u>** standing before the digital mirror." Ex. 1006 (Rosenberg) at [0058]; *see also* at [0071]. Further, Rosenberg states that a digital camera may function as a proximity/motion sensor. Ex. 1006 (Rosenberg) at [0034], [0069], [0073], [0109], Fig. 13. A POSITA would have understood that in such instances, the video images captured by the

digital camera would be used to detect the user's presence, and the detected user would be the user appearing in the videos (i.e., "the stream of images") disclosed in Rosenberg.

d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view; and

253. Rosenberg in view of Park renders obvious this limitation. For example, Rosenberg teaches "provid[ing] the user with a natural and accurate **visual perspective** of himself or herself as compared to a traditional mirror." Ex. 1006 (Rosenberg) at [0027]. Additionally, Rosenberg teaches that a camera mounted above the display would generate an "unnatural perspective as compared to a traditional mirror." Ex. 1006 (Rosenberg) at [0027]. Rosenberg discloses the need of providing images that appear to be captured "approximately at eye-level." Ex. 1006 (Rosenberg) at [0027]. Rosenberg teaches that "an eye-level viewing angle [] provides a highly realistic mirror replication image." Ex. 1006 (Rosenberg) at [0061]; *see also* Ex. 1006 (Rosenberg) at [0028], [0075].

254. Further, Park teaches a method of generating images that appear to be taken at a different angle. For example, Park discloses modifying images having "a perspective distortion of the facial image that is not desired by the user." Ex. 1011 (Park) at [0050], Figs. 6A-6B.

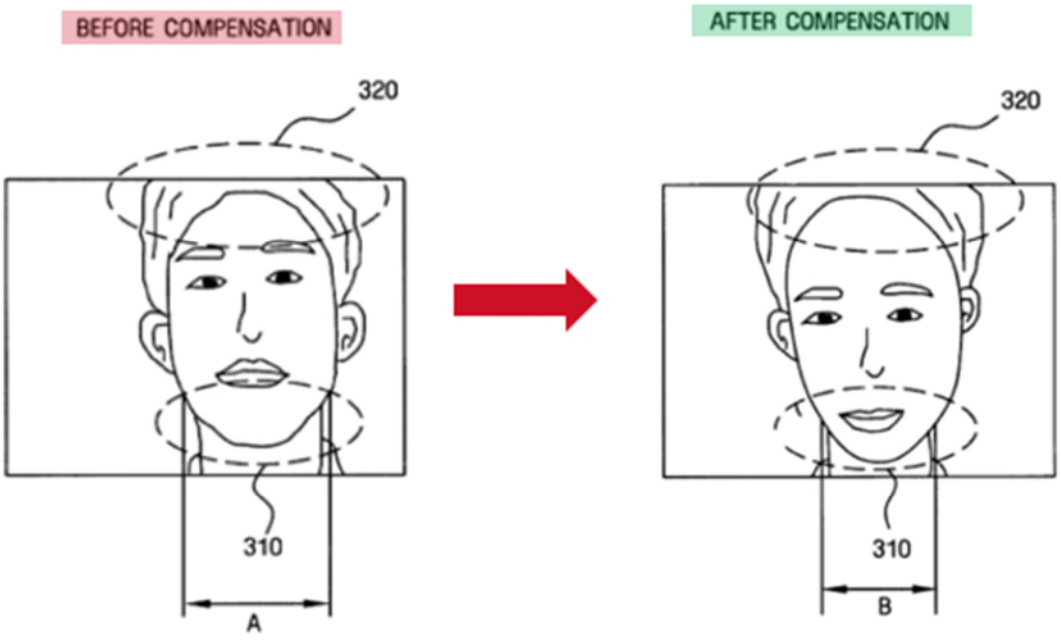


FIG. 6A FIG. 6B

Ex. 1011 (Park) at Figs. 6A and 6B (annotated).

255. Park specifically teaches that when a camera is positioned below the user's face, the captured images may be distorted with an "emphasized" chin part and a "narrowed" forehead part. Ex. 1011 (Park) at [0051]. Park further teaches that to adjust the user's perspective, "image compensation using a geometric transform technique and a perspective transform technique" may be used, as depicted in Figs. 7-8 of Park. Ex. 1011 (Park) at [0059]. Park discloses that in the compensated image, the chin is reduced and the forehead is enlarged as compared to the distorted image, as shown in Fig. 6B (above). Ex. 1011 (Park) at [0051], Fig. 6B.

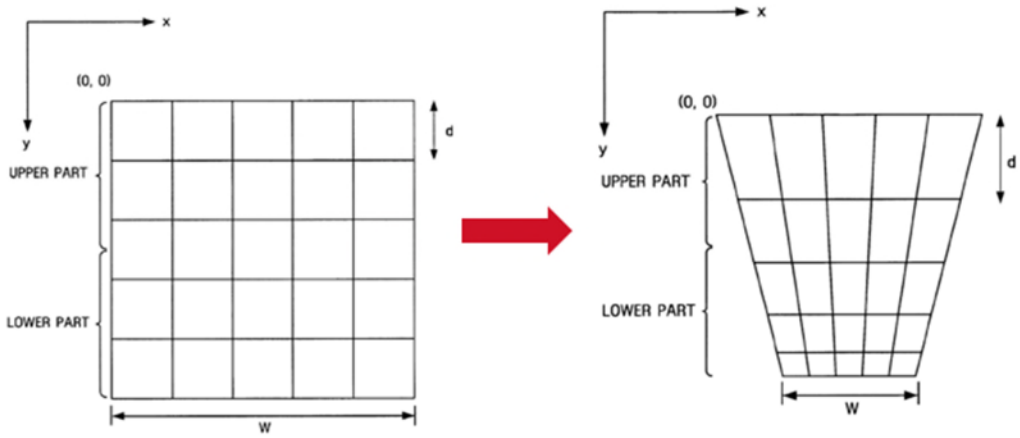


FIG. 7 FIG. 8

Ex. 1011 (Park) at Fig. 7 (left) and Fig. 8 (right) (annotated).

256. Park further discloses "compensating for image distortion of the captured image when the image capturing unit is determined to be above or below the object." Ex. 1011 (Park) at claim 44; *see also* Ex. 1011 (Park) at [0014]. A POSITA would have understood that Park's perspective transformation is an example of a transformation mapping, because said perspective transformation maps the positions of the pixels from the captured image to those of the compensated image. Ex. 1011 (Park) at [0056], [0059]-[0062]. More specifically, Park discloses a matrix-based image transformation that maps pixels of an image before transformation to pixels of a compensated image. Ex. 1011 (Park), [0060]-[0061]. Likewise, the '109 Patent describes the transformation mapping is performed using a matrix. Ex. 1021 ('109 Patent), 29:38-42. A POSITA would have understood

Park's transformation to be "adaptive" because it adapts to the tilt angle of the camera. Ex. 1011 [Park] at [0060]-[0061].

257. As discussed, a POSITA would have been motivated to implement Park's image compensation into Rosenberg's camera/display configuration to correct image distortions and provide a more realistic mirror image for display to a user. *See supra* §XII.A, ¶¶244-249. A POSITA therefore would have found it obvious to apply the transformation mapping disclosed in Park to the images captured by the cameras in Rosenberg to correct the unnatural perspective and provide an ideal user perspective ("i.e., an image that is from a vantage point perfectly straight ahead" (Ex. 1006 (Rosenberg) at [0028])). *See supra* §XII.A; ¶¶242-249. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XII.A; ¶249. A POSITA would have understood Rosenberg's camera-captured images modified by Park's transformation mapping would appear to be captured from a point of view different from the camera's actual point of view. A POSITA would have recognized the transformed images would appear to be captured from a point of view "straight ahead" of the user instead of the camera's actual point of view above and tilted at the user, which yields more realistic replications of true mirror images, as desired by Rosenberg. Ex. 1006 (Rosenberg) at [0028]; *see also* Ex. 1006 (Rosenberg) at [0075].

e. <u>1[d]: displaying the modified images on the monitor;</u>

258. Rosenberg in view of Park renders obvious displaying the modified images on the monitor. For example, Rosenberg discloses displaying the real-time video images captured by the camera after performing modifications, such as left-right inversion, scaling, and/or zooming. *See, e.g.*, Ex. 1006 (Rosenberg) at [0026], [0035], [0038], [0050]-[0052], [0077], Fig. 5. Rosenberg also discloses displaying "composite images that include a concurrent display of current and previous reflective images, thereby allowing users to view changes over time in their appearance." Ex. 1006 (Rosenberg) at [0025]. Rosenberg discloses "adjusting and/or modifying the images that are displayed upon the screen to achieve unique mirroring features and functions not provided by a standard household wall mirror." Ex. 1006 (Rosenberg) at [0035].

259. Park also discloses displaying the modified images on the display. For example, Park discloses "a display unit displaying the image compensated by the compensation unit." Ex. 1011 (Park) at claim 16; *see also* Ex. 1011 (Park) at Fig. 9, [0055]-[0056].

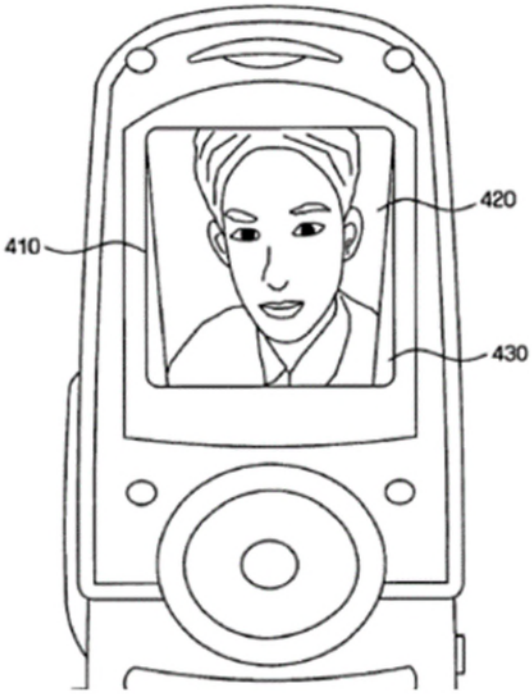


FIG. 9

Ex. 1011 (Park) at Fig. 9 (cropped).

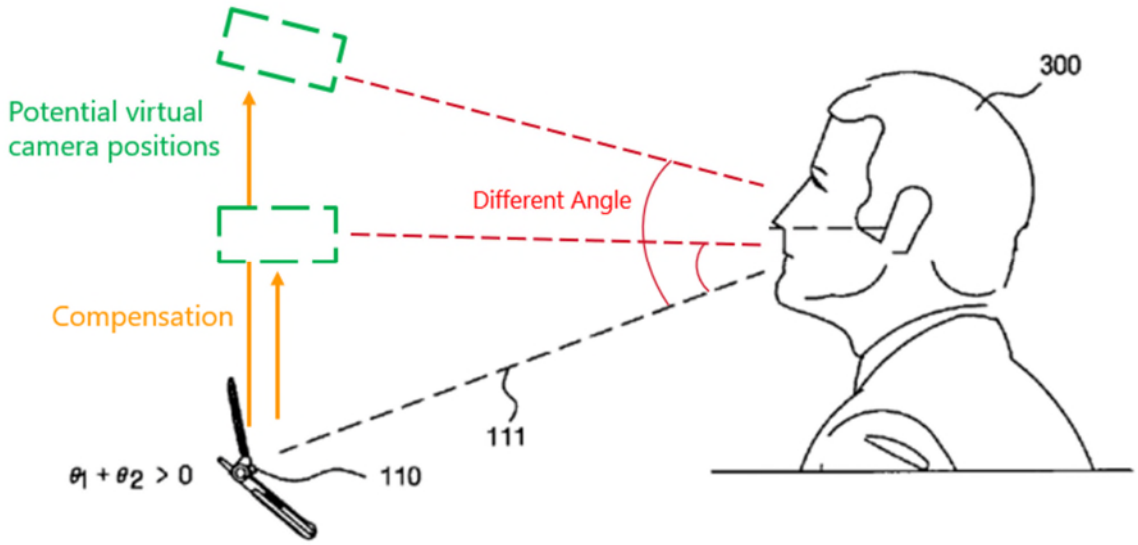
260. As discussed, a POSITA would have found it obvious to implement Park's image compensation into Rosenberg's camera/display configuration to correct image distortions and display a more realistic mirror image to a user. *See supra* §XII.A, ¶¶244-249. A POSITA therefore would have found it obvious to display the modified images on the monitor to show the user images with a natural perspective mimicking reflections of a traditional mirror as desired by Rosenberg. Ex. 1006 (Rosenberg, [0027]). *See supra* §XII.A; ¶¶242-249. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XII.A; ¶249.

f. <u>1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and</u>

261. Rosenberg in view of Park renders obvious this limitation.

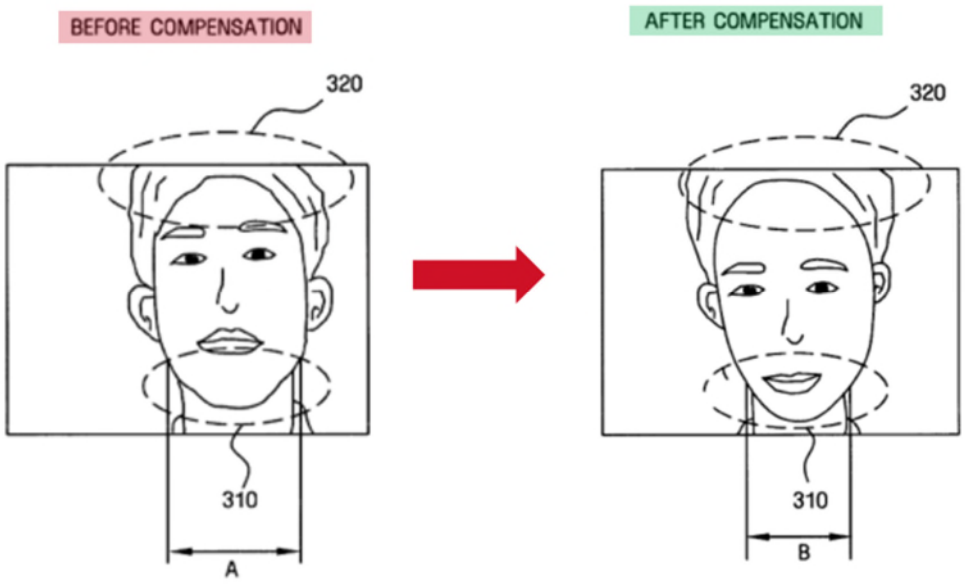
i. "wherein the different point of view corresponds to an image taken from a different angle"

262. A POSITA would understand that implementing Park's compensation in Rosenberg would result in a different point of view that corresponds to an image taken from a different angle. For example, Park discloses generating images having a different point of view corresponds to an image taken from a different angle, and as shown below, Park discloses applying an image compensation to generate images with different user perspectives. *See, e.g.*, Ex. 1011 (Park) at Fig. 5.



Ex. 1011 (Park) at Fig. 5 (partial, annotated).

263. A POSITA would have understood that after applying the compensation as taught by Park, the system would create virtual camera positions having a point of view with different angles. For example, prior to implementing the transformation (i.e., compensation) disclosed in Park, the user's image is captured at a point of view below the user's face and contains perspective distortion. *See* Ex. 1011 (Park) at Fig. 6A. After the transformation is applied, and the distortion is compensated, the user's image appears to be captured at a point of view straight ahead of the user's face. *See* Ex. 1011 (Park) at Fig. 6B. *See also* Ex. 1011 (Park) at [0050]-[0051], Figs. 14-15.



See Ex. 1011 (Park) at Figs. 6A-6B (annotated).

ii. "…a different distance than the actual images captured by the camera"

264. Additionally, a POSITA would understand that combining the compensation disclosed in Park with the mirror image processing disclosed in Rosenberg would result in a different point of view corresponding to an image taken from a different distance.

265. The '110 Patent describes "resizing the image to reduce variations caused by changes in the object's distance to the camera." *See* Ex. 1001 ('110 Patent) at 3:26-28. Similarly, the '109 Patent discloses resizing the image based on distance to the user to match the size that the user would see in a mirror. *See* Ex. 1021 ('109 Patent) at 2:19-27 (teaching unlike a mirror image, the user appears larger in a camera image as the user approaches the camera), 17:19-23, 21:57-60, 23:65-24:9, 25:46-48, 30:4-11, Fig. 11.

266. Rosenberg, similar to the '110 and '109 Patents, discloses scaling (i.e., resizing) user's images that are captured as real-time video images by the camera so that each of the scaled video images appear to emulate "a traditional life-size mirror image." Ex. 1006 (Rosenberg) at [0077]. A POSITA would have understood that resizing an image would make the image appear to be captured by the camera at a different distance, thereby creating a virtual point of view corresponding to the different distance.

267. Rosenberg also discloses that "the digital mirror also enables the user to zoom-in and/or zoom-out the real-time image." Ex. 1006 (Rosenberg) at [0050]. In Rosenberg, a digital zoom may be performed by enlarging the displayed images. Ex. 1006 (Rosenberg) at [0100]. A POSITA would have understood that zooming in/out on an image captured by a fixed-focus camera, such as the ones used in Rosenberg, creates a modified image with a virtual point of view corresponding to a different distance from the user to the camera. *See supra* §VI.B.f.

268. As discussed above, a POSITA would have understood that resizing an image would create a virtual point of view corresponding to a different distance. *See supra* §VI.B.f.

g. <u>1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.</u>

269. Rosenberg in view of Park renders obvious this limitation.

270. Rosenberg teaches a mirror tracking mode. For example, Rosenberg discloses a "sensor for detecting when a user is physically present before and/or approaches and/or stands before the display screen of the mirror within some proximity range." Ex. 1006 (Rosenberg) at [0025] (also teaching "the system automatically enters a mirroring mode when a user steps in front of the display screen for more than a threshold amount of time"); *see also* Ex. 1006 (Rosenberg)

at [0032]-[0033], [0056], [0069]-[0073], [0109], [0114], Fig. 13. When the sensor disclosed in Rosenberg detects that the user is within a specified proximity range, the virtual mirror system "automatically engages the digital mirror functionality of the present invention." Ex. 1006 (Rosenberg) at [0025]. In other words, Rosenberg's system detects the presence of a user to automatically engage a mirror mode. *See* Ex. 1006 (Rosenberg) at [0026], [0028], [0035], [0077]. A POSITA would have understood this discloses a "mirror tracking mode" because during the mirror mode, real-time images of the user captured by the camera are modified (e.g., left-right inversion, scaling) to generate and display real-time mirror images to the user.

271. When the sensor disclosed in Rosenberg detects the user is no longer before the mirror for more than a threshold amount of time, then the system automatically switches from the digital mirror state to "a non-mirroring state," which may include a digital picture frame state or a power conservation state. Ex. 1006 (Rosenberg) at [0083], [0114]; *see also* Ex. 1006 (Rosenberg) at [0025], [0032]-[0033], [0056], [0069]-[0073], [0109]. Rosenberg also discloses that the digital camera can be used, either alone or in combination with other sensors, to provide this proximity and/or motion sensing functionality. Ex. 1006 (Rosenberg) at [0034], [0069], [0073], [0109], Fig. 13.

272. When Rosenberg is modified with Park's transformation mapping (*see supra* §XII.B.d), the processor would utilize the angle of the images that are captured

by the camera of Rosenberg and apply parametric input for a mapping transformation that changes imaging perspective of the images as taught by Park. More specifically, Park teaches utilizing the angle of the images captured by the camera (the calculated tilt $\Theta_1+\Theta_2$) to apply parametric input for a transformation mapping that changes imaging perspective of the images captured by the camera. *See, e.g.*, Ex. 1011 (Park) at Fig. 3. The system disclosed in Park measures the tilt of the main body against a horizontal plane (Θ_1) and the angle between the camera and the main body (Θ_2). At a position shown in Figure 3, Θ_1 has a negative value. *See, e.g.*, Ex. 1011 (Park) at Fig. 3.

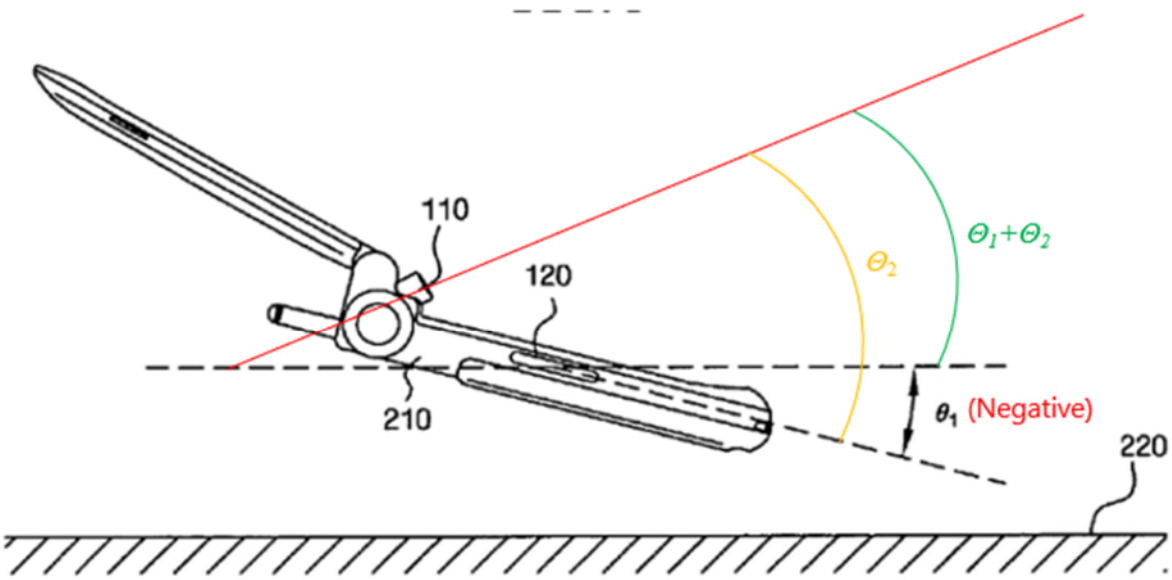


FIG. 3

Ex. 1011 (Park) at Fig. 3 (annotated).

273. As shown in the figure below, the calculated tilt $\Theta_1+\Theta_2$ represents an angle of the image captured by the camera.

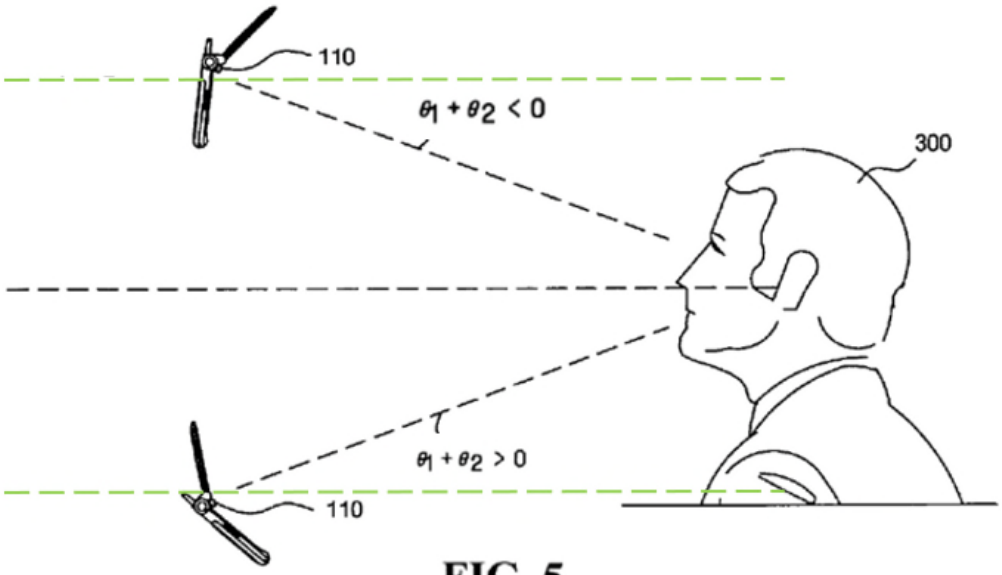


FIG. 5

Ex. 1011 (Park) at Fig. 5 (annotated).

274. The device in Park utilizes the calculated tilt for image processing and applies parametric input for a mapping transformation. Ex. 1011 [Park] at [0060]-[0061]. For example, Park discloses a transformation mapping that applies parametric input, as illustrated in the equation below:

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} m00 & m01 & m02 \\ m10 & m11 & m12 \\ m20 & m21 & m22 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$x' = \frac{m00x + m01y + m02}{m20x + m21y + m22}$$

$$y' = \frac{m10x + m11y + m12}{m20x + m21y + m22}$$

$$X = \frac{x'}{w}, \; Y = \frac{y'}{w}$$

Ex. 1011 (Park) at Equation (1), [0061].

275. In the equation above, m00-m22 are matrix coefficients adjusted on the basis of the calculated tilt. Ex. 1011 (Park) at [0061]. A POSITA would have understood Park's coefficients are a "parametric input" of the transformation because changes of these coefficients would result in changes of the performed transformation. Ex. 1011 (Park) at [0061]. Therefore, Park teaches utilizing the angle of the images captured by the camera (the calculated tilt Θ1+Θ2) and applies parametric input for a mapping transformation. A POSITA would have understood that Rosenberg's camera angle, in relation to the horizontal, is equivalent to the calculated tilt $\Theta_1+\Theta_2$. In other words, a POSITA would understand tilt angle Θ_1 (angle of the device relative to the horizontal plane) in Rosenberg to be 90-degrees

because Rosenberg teaches a wall-mounted display (Ex. 1006 (Rosenberg) at [0025]-[0026]), and tilt angle Θ2 would be the angle of the camera relative to the display. Ex. 1011 (Park) at Fig. 5.

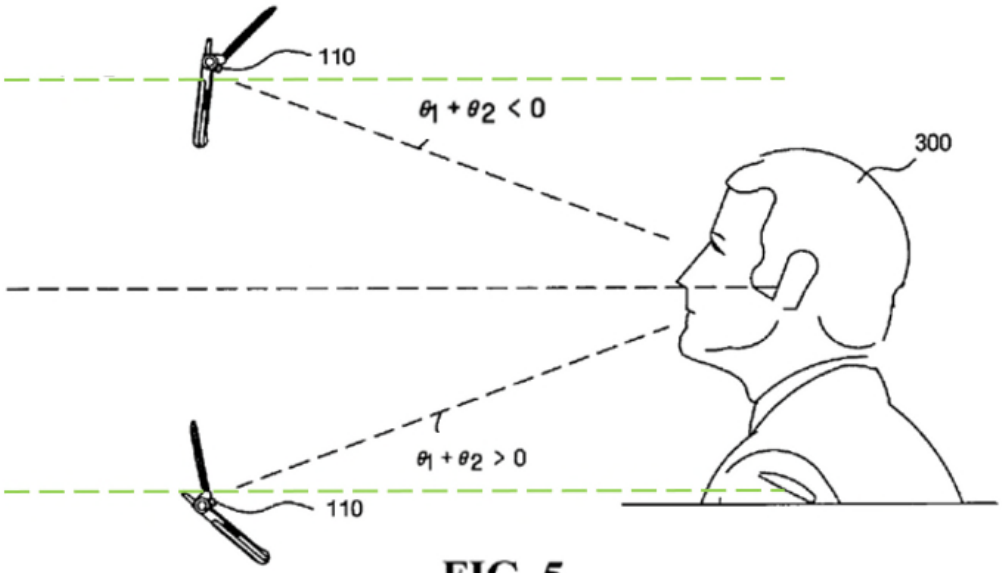


FIG. 5

Ex. 1011 (Park) at Fig. 5 (annotated).

276. A POSITA would be motivated to implement Park's transformation mapping into the mirror tracking mode disclosed in Rosenberg, which would enable Rosenberg's processor to perform Park's compensation method (i.e., "a mapping transformation") that utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera. Additionally, as discussed above, a POSITA would have been motivated to implement the transformation mapping disclosed in Park in Rosenberg as an efficient solution to correcting image

distortions based on the angle of a camera, as described in Rosenberg. *See* Ex. 1006 (Rosenberg) at [0027]-[0028], [0061]; *see supra* §XII.A, ¶247. A POSITA would have also done so with a reasonable expectation of success. *See supra* §XII.A.

C. CLAIM 2: THE METHOD OF CLAIM 1, WHEREIN THE MAPPING TRANSFORMATION COMPRISES A COMBINATION OF AT LEAST TWO OF THE FOLLOWING TRANSFORMATIONS: VERTICAL FLIPS ABOUT A VERTICAL AXIS, A PROJECTION TRANSFORMATION, ROTATION, RESIZE, TWO-DIMENSIONAL OFFSET, OPTIC DISTORTION CORRECTION, AND POLYNOMIAL MAPPING ON PART OR ENTIRE CAPTURED IMAGE.

277. Rosenberg in view of Park renders obvious the method of claim 1. *See supra* §XII.B. Further, Rosenberg in view of Park renders obvious this claim. Rosenberg in view of Park teaches at least a combination of two of the following transformations: flips about a vertical axis, a projection transformation, rotation, resizing, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image. For example, Rosenberg and Park disclose applying transformations including at least flips about a vertical axis, a projection transformation, two-dimensional offset, and resizing on part of or on the entire captured image. For example, as described below, Rosenberg discloses applying transformations including at least flips about a vertical axis, as well as resizing part or the entire captured image, moving part or the entire captured image, and resizing by zooming in/out. See, e.g., Ex. 1006 (Rosenberg) at [0026] ("the real-time streamed video images being a **left-right inverted** representation of the video signal

captured by the camera" and "[t]he left-right inverted representation is generally produced by **flipping the pixels of the image around the vertical centerline**."); *see also* Ex. 1006 (Rosenberg) at [0035], [0038], [0048]-[0052], [0077], [0093]-[0111], Fig. 5 (showing a modified image with "a displayed size and scaling such that it appears like a traditional life-size mirror image," *see* Ex. 1006 (Rosenberg) at [0077]), Fig. 8 (showing moving the user left or right, resulting in a two-dimensional offset), Figs. 10-11 (showing zooming in and zooming out).

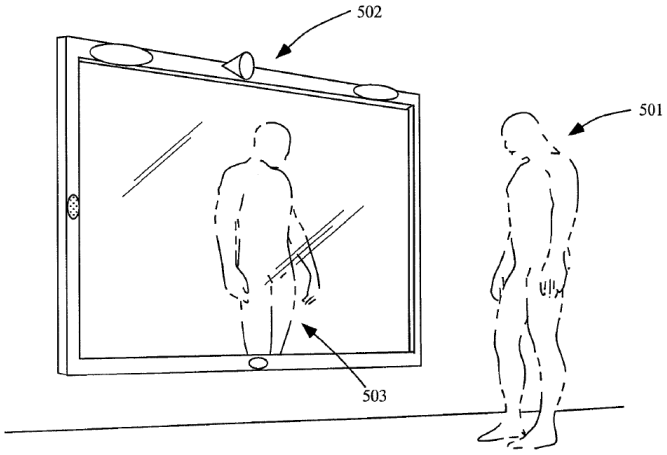


Fig. 5

Ex. 1006 (Rosenberg) at Fig. 5; *see also* Ex. 1006 (Rosenberg) at [0077].

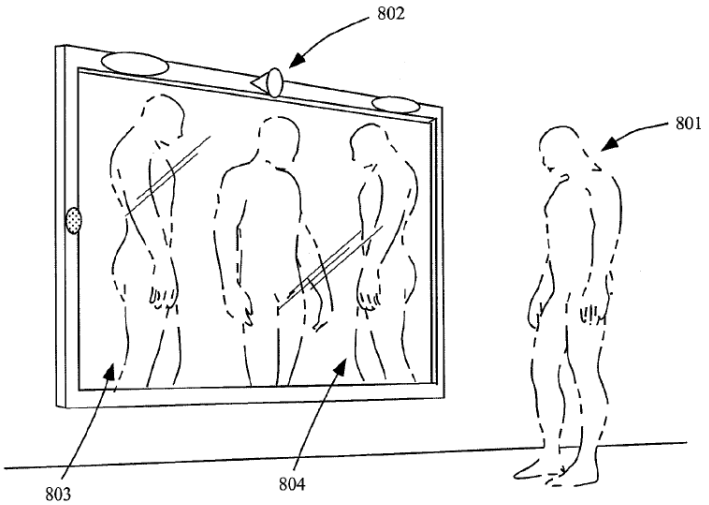


Fig. 8

Ex. 1006 (Rosenberg) at Fig. 8.

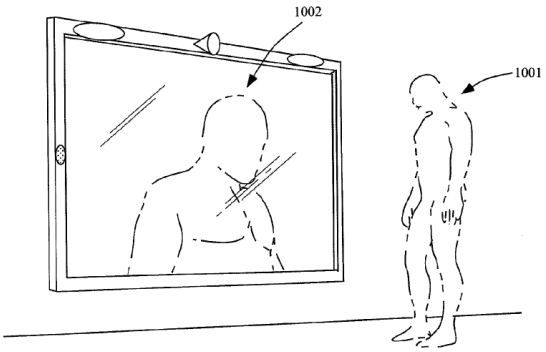


Fig. 10

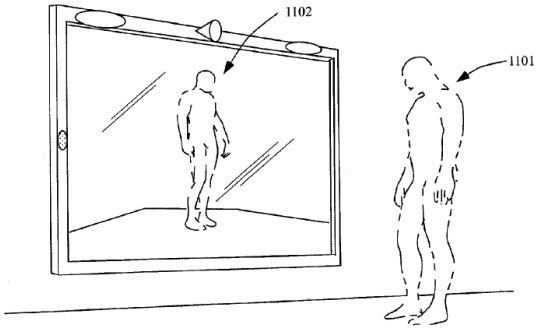


Fig. 11

Ex. 1006 (Rosenberg) at Fig. 10 (right) and Fig. 11 (left).

Additionally, Park discloses a projection transformation. For example, Park discloses correcting a perspective distortion using a perspective transformation technique. Ex. 1011 (Park) at [0050]-[0051], [0059]-[0062], Figs. 6-8; *see also supra* §XII.B.d, ¶¶254-256.

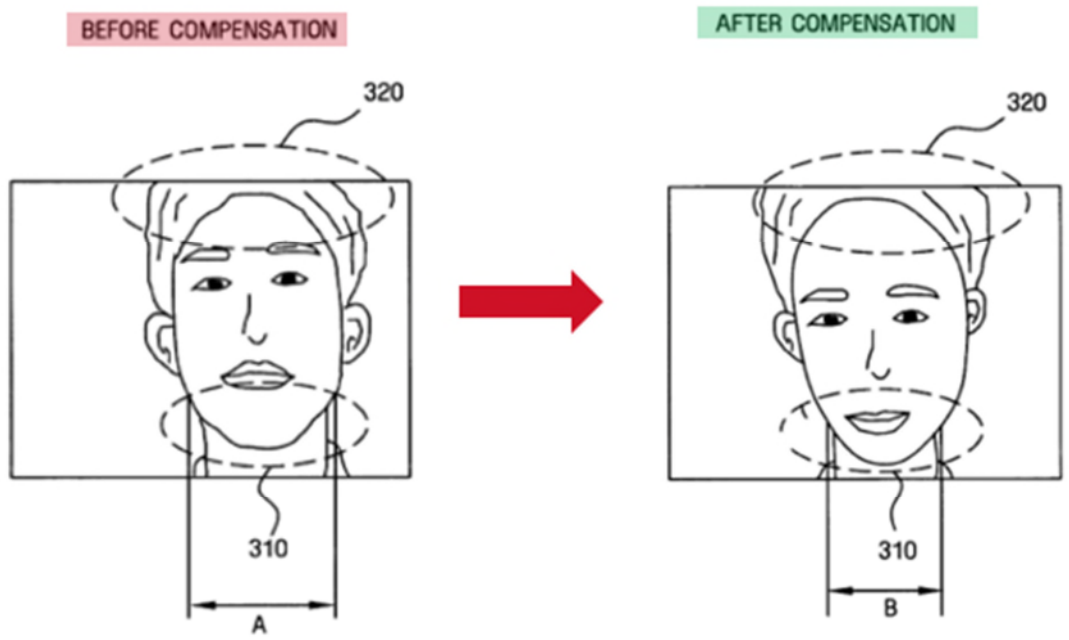


FIG. 6A FIG. 6B

Ex. 1011 (Park) at Fig. 6 (annotated).

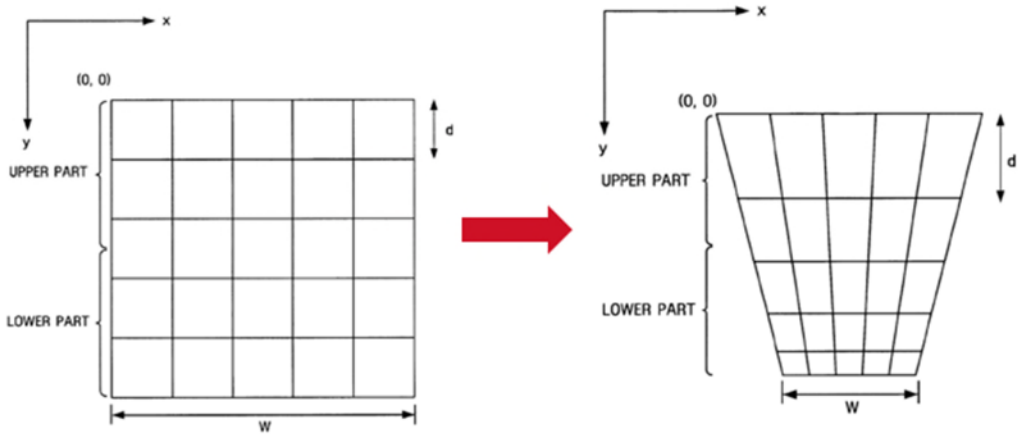


FIG. 7 FIG. 8

Ex. 1011 (Park) at Fig. 7 (left) and Fig. 8 (right) (annotated).

278. A POSITA would have understood that Park's perspective transformation technique is a projection transformation because, as discussed above,

Park's perspective transformation maps the positions of the pixels from the captured image to those of the compensated image (Ex. 1011 (Park) at [0056]; *see supra* §XII.B.d, ¶256), and it changes the point of view of the image from above or below the user to a point of view straight ahead of the user (*see supra* §XII.B.d, ¶257). Further, as explained above, Rosenberg teaches a resizing transformation. *See* supra §XII.B.f.ii. Park also teaches a resizing transformation. Park discloses "the compensation unit 140 performs perspective compensation on the image as illustrated in FIG. 6A to provide a compensated image as illustrated in FIG. 6B. In the compensated image of FIG. 6B, **the chin part 310 is reduced and the forehead part 320 is enlarged**." Ex. 1011 (Park) at [0051], Fig. 6. Park further discloses "a method of cutting out an unimpaired square region of the compensated image 420," and "enlarging cut region to a size enough to fill the display region." Ex. 1011 (Park) at [0057]; *see also* Ex. 1011 (Park) at claims 21, 41. A POSITA would have understood the transformations disclosed in Rosenberg and Park are similar to the flipping, translation, and resizing for zoom in/out described in the '109 Patent. Ex. 1021 ('109 Patent) at 16:46-17:28, 22:1-13, 23:41-24:9, 24:51-53, 25:48-51, 29:22-58.

D. CLAIMS 3-4. CLAIM 3. THE METHOD OF CLAIM 1, WHEREIN DURING IDLE MODE THE CAMERA CONTINUOUSLY SENDS CAPTURED IMAGES TO THE PROCESSOR, AND THE USER PRESENCE IS IDENTIFIED BY THE PROCESSOR DETECTING A BODY PART IN THE IMAGES. CLAIM 4. THE METHOD OF CLAIM 3, WHEREIN THE BODY PART COMPRISES ONE OF LEGS, FACE, EYES, AND HANDS.

279. Rosenberg in view of Park renders obvious the method of claim 1. *See supra* §XII.B. Further, Rosenberg in view of Park renders obvious claims 3 and 4 of the '110 Patent. For example, Rosenberg discloses an idle mode (a low-power idle state) with an auto-activation feature, in which the proximity and/or motion sensors continuously operating to detect whether or not a person is located before the display screen. *See* Ex. 1006 (Rosenberg) at [0032]. In a low-power idle mode, while the system may turn off or dim the display screen, the sensor would continuously send data to the processor for identifying the presence of a user to automatically engage the mirror mode. *See* Ex. 1006 (Rosenberg) at [0032], [0070]. Rosenberg discloses that the proximity and motion detection may be accomplished using a dedicated sensor, such as an infra-red sensor, or using the camera of the system. Ex. 1006 (Rosenberg) at [0069], [0073], [0109], Fig. 13. A POSITA would have understood that for those embodiments using the camera, the data sent to the processor are captured real-time video images and would be continuously sent to the processor to detect the user presence because it would allow the system to engage the mirror functionality "when a user steps in front of the display screen for more than a

threshold amount of time". Ex. 1006 (Rosenberg) at [0025]; *see also* Ex. 1006 (Rosenberg) at [0056], [0061]-[0062], [0072], [0114], claim 17.

280. Further, a POSITA would understand that methods for detecting a user's presence using image data from a digital camera were well-known in the art and would have been aware of various known and suitable image processing methods for detecting a user's presence using the digital camera and processor disclosed in Rosenberg. For example, Lizee discloses estimating the distance between a user and the video capture device based on the number of pixels corresponding to features of the user. *See* Ex. 1008 (Lizee) at [0043]-[0046]. Further, Haan discloses estimating distance based on the size of the object or by parallax. *See* Ex. 1005 (Haan) at [0074]. Moreover, a POSITA would have understood that such methods would detect one of a user's legs, face, eyes, and hands before processing the image to determine the size of such body features and thus the proximity/presence of the user. Indeed, it was well-known in the art that when the features of a user, such as eyes, are used for further image processing, the eyes would need to be detected first. For example, in its background section, Park discloses that a Korean patent publication (Korean Patent Unexamined Publication No. 2004-0107890) "discloses an image display compensation method for a mobile phone that adjusts a captured image parallel to the user's eyes." Ex. 1011 (Park) at [0008]. A POSITA would understand that the disclosed system must necessarily detect at least

the user's eyes. Therefore, a POSITA would have understood that known and suitable methods for detecting a user's presence would include detecting one of a user's legs, face, eyes, and hands. Further, a POSITA would have understood that the methods would include detecting one of a user's legs, face, eyes, and hands because these are among finite options for detecting a person.

E. CLAIM 5: THE METHOD OF CLAIM 1, FURTHER COMPRISING RECEIVING USER INPUT AND OPERATING THE PROCESSOR TO VARY IMAGE MAGNITUDE OR FIELD-OF-VIEW OF THE MODIFIED IMAGES ACCORDING TO THE USER INPUT.

281. Rosenberg in view of Park renders obvious the method of claim 1. *See supra* §XII.B. Further, Rosenberg in view of Park renders obvious this claim. For example, Rosenberg discloses receiving user verbal input, such as commands "zoom-in" and "zoom-out," to operate the processor to vary image magnitude (perform zooming function). Ex. 1006 (Rosenberg) at [0100]. Rosenberg teaches using a microphone to receive a user input and a speech processor to recognize the commands. Ex. 1006 (Rosenberg) at [0031], [0100], [0109].

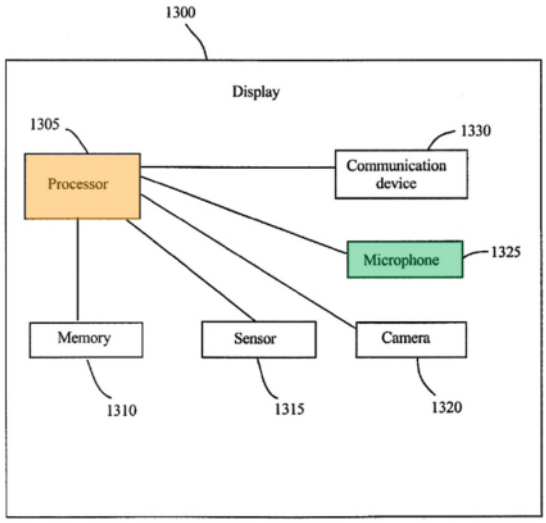


Fig. 13

Ex. 1006 (Rosenberg) at Fig. 13 (annotated). Rosenberg discloses zooming functions that "may be performed by digital zoom," which is "performed by enlarging the display graphically without changing the optics of the camera in the image capture process." Ex. 1006 (Rosenberg) at [0100].

F. CLAIM 6: THE METHOD OF CLAIM 5, WHEREIN THE PROCESSOR VARIES THE IMAGE MAGNITUDE OR FIELD-OF-VIEW MAPPING TO DISPLAY A PARTIAL OR FULL BODY OF THE USER IN THE MODIFIED IMAGES.

282. Rosenberg in view of Park renders obvious the method of claim 5. *See supra* §XII.E. Further, Rosenberg in view of Park renders obvious this claim. As discussed above, Rosenberg discloses varying the image magnitude. *See* supra §XII.B.f.ii. Rosenberg additionally discloses increasing magnitude (zooming in) to display a partial body of the user. *See* Ex. 1006 (Rosenberg) at [0100] ("Zooming in is particularly useful if a user is putting on makeup or performing some other action

that requires close attention to small details."). Additionally, Rosenberg discloses decreasing the image magnitude (zooming out) to display a full body of the user when "a user wants to see a full length image of himself or herself on a mirror screen." Ex. 1006 (Rosenberg) at [0100]. Rosenberg also teaches that "[s]uch zooming in and zooming out functions make the current digital mirror system substantially improved over traditional wall mirrors for it allows a single mirror to be used for close-up actions that require magnification" and "full length views that are often not possible in small spaces." Ex. 1006 (Rosenberg) at [0100]. *See also* Ex. 1006 (Rosenberg) at [0050]-[0051], [0099]-[0105], Figs. 10-11.

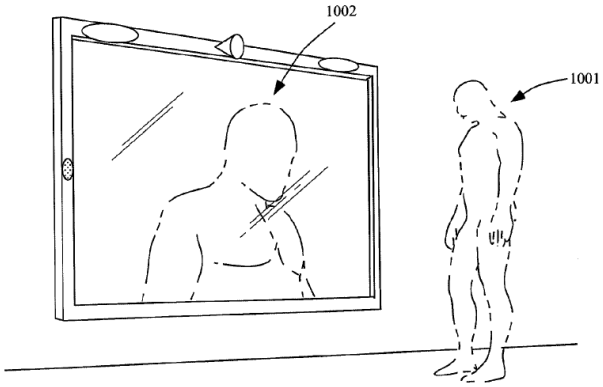


Fig. 10

Ex. 1006 (Rosenberg) at Fig. 10.

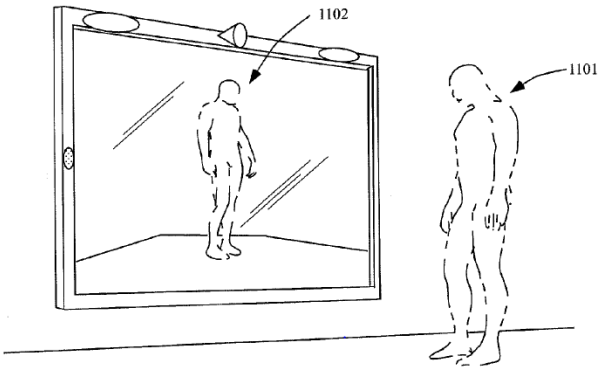


Fig. 11

Ex. 1006 (Rosenberg) at Fig. 11.

G. CLAIM 10: THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO AUGMENT THE MODIFIED IMAGES BY OVERLAYING VIRTUAL OBJECTS ON A USER'S IMAGE APPEARING WITHIN THE MODIFIED IMAGES.

283. Rosenberg in view of Park renders obvious the method of claim 1. *See supra* §XII.B. Further, Rosenberg in view of Park renders obvious this claim. For example, Rosenberg discloses overlaying stock hairstyle images "as an onionskin transparency upon the real-time image of the user himself or herself." Ex. 1006 (Rosenberg) at [0111]. A POSITA would have understood Rosenberg's image overlays augment the modified images because it combines the modified images of real objects with virtual objects that enhances user's experience.

H. CLAIMS 11-12. CLAIM 11. THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO TRANSMIT THE MODIFIED IMAGE OVER A NETWORK. CLAIM 12. THE METHOD OF

CLAIM 1, FURTHER COMPRISING DISPLAYING THE MODIFIED IMAGE ON A REMOTE SCREEN FOR VIDEO TELEPRESENCE.

284. Rosenberg in view of Park renders obvious the method of claim 1. *See supra* §XII.B. Further, Rosenberg in view of Park renders obvious claims 11 and 12 of the '110 Patent. Rosenberg, for example, discloses that "the processor within the advanced digital mirror may have communication access to one or more communication networks via a network connection 306." Ex. 1006 (Rosenberg) at [0074]. Rosenberg further discloses that the "communication device 1330 is utilized to receive and transmit images across a network." Ex. 1006 (Rosenberg) at [0109], Fig. 13. Park also discloses a transmission unit for transmitting compensated images to another user over a network for video telephony. Ex. 1011 (Park) at [0010], [0013], [0037], [0063], [0082], [0092], Fig. 2. As discussed above, Rosenberg in view of Park discloses generating and displaying a modified image. *See supra* §§XII.A, XII.B.c-XII.B.g. A POSITA would have found it obvious to utilize Rosenberg's system for transmitting and displaying the modified image on the screen of another user at a remote location for video telepresence. For example, a POSITA would have understood Rosenberg's system would be suitable for showing the modified image to another user to collaborate on the user's look, such as a hairstyle, makeup, or clothing. *See* Ex. 1006 (Rosenberg) at [0053], [0111], Ex. 1005 (Haan) at [0113]-[0118], Ex. 1007 (Geisner) at [0154], [0156], [0171]-[0172]. A

POSITA would have understood such functionality improves Rosenberg's user experience by improving the versatility of Rosenberg's digital mirror system. *See* Ex. 1006 (Rosenberg) at [0005] (teaching "a substantial need for a significantly more versatile digital mirror system with a variety of unique and useful digital mirroring features and functions."). Rosenberg also teaches a "third person mirror mode" that allows the user to see what he/she looks like from the perspective of a third person. Ex. 1006 (Rosenberg) at [0036]. A POSITA also would have utilized Rosenberg's system (as modified by Park) for transmitting and displaying a modified image from the third person perspective to a remote user for telepresence, such as in a video call (*see* Ex. 1011 (Park), [0010], [0013]).

XIII. GROUND 3: CLAIMS 1-18 ARE OBVIOUS IN VIEW OF GEISNER IN VIEW OF LIZEE

A. MOTIVATION TO COMBINE GEISNER WITH LIZEE

285. A POSITA would have found it obvious to combine Geisner with Lizee. For example, both Geisner and Lizee are analogous art to the '110 Patent and to each other.

286. As mentioned above, the '110 Patent "relates generally to imaging and display systems and, more particularly, to monitors, and interactive displays, e.g., in retail and/or service environments, medical or home situations, video conferencing,

gaming, and the like." Ex. 1001 ('110 Patent) at 1:43-51 (Technical Field). It specifically relates to a virtual mirror for trying on apparels.

287. Geisner is in the same field. Geisner discloses an interactive shopping application deployed on a display screen. Ex. 1007 (Geisner) at Title ("Motion-based interactive shopping environment"); Abstract ("An on-screen shopping application … provid[ing] a shopping experience"), ("allow[ing] the user to interact with virtual articles"). The implementation of the system includes trying on virtual garments in a retail setting or playing games in commercial or home situations. Ex. 1007 (Geisner) at [0001], [0026].

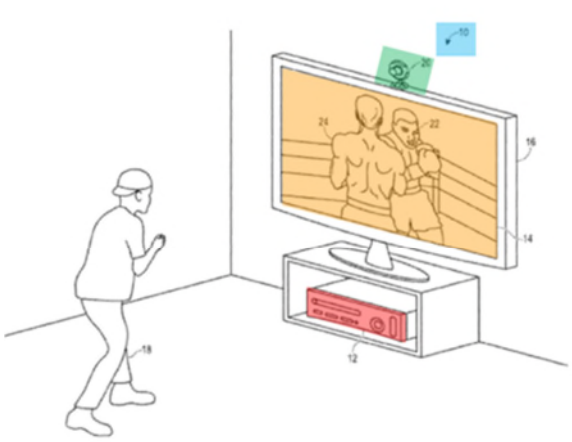
288. Lizee is also in the same field. Lizee discloses an "interactive immersion of a silhouette image on a display screen of a user interacting therewith" Ex. 1008 (Lizee) at Field of Invention. In particular, Lizee teaches a video system with "an image capture device" enabling a user to interact with "a given virtual environment" with "a virtual representation of the user." Ex. 1008 (Lizee) at Abstract; Field of Invention. A primary implementation discussed in Lizee is for gaming, such as a virtual boxing game. Ex. 1008 (Lizee) at [0079], [0004].

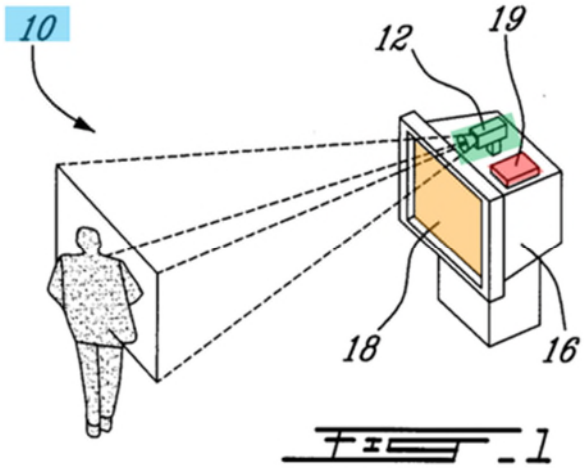
289. The '110 Patent, Geisner, and Lizee aim to solve a similar problem: enhancing techniques that create a user representation on a screen, track user movements, and enable interactions of the user with the virtual environment. For example, the '110 Patent describes a method to "generate an avatar having body

characteristics corresponding to body characteristics of a user." Ex. 1001 ('110 Patent) at Abstract. Further, the '110 Patent discuses "calibrating the image or video based on a user tracking system" and "interactive displays." Ex. 1001 ('110 Patent) at 1:44-45, 3:2-4. Similarly, Geisner discloses "[t]he on-screen representation has a likeness of the user or another individual" and "[a] tracking system captures user motions." Ex. 1007 (Geisner) at [0003], Abstract. Geisner describes that its system "allows the user to interact with virtual articles that represent real-world articles." Ex, 1007 at Abstract.

290. Similarly, Lizee describes a system that can "capture a user image from which a virtual representation of the user may be generated." Ex. 1008 (Lizee) at Abstract. Lizee also describes a "position and motion of the user may also be calculated to position and move the virtual representation" and "a method and apparatus for the interactive immersion of a silhouette image on a display screen of a user interacting therewith." Ex. 1008 (Lizee) at Abstract, Field of the Invention.

291. Accordingly, a POSITA would be motivated to combine Geisner and Lizee because they use known techniques to improve similar devices in the same way. For example, as shown in the figures below, Geisner (left) and Lizee (right) aim to improve similar display systems (shown in blue) comprising a camera (shown in green), a processor (shown in red), and a display (shown in yellow).





Ex. 1007 (Geisner) at Fig. 1 (left); Ex. 1008 (Lizee) at Fig. 1 (right).

292. As shown from the above figure example, Geisner and Lizee use similar techniques in the same way. Further, as shown in the comparison below, both Geisner and Lizee consider a virtual boxing game, and teach extracting user body and movement information and transforming the information into interactions with virtual objects in a virtual environment, such as in gaming. Ex. 1007 (Geisner) at [0026]; Ex. 1008 (Lizee) at [0079].

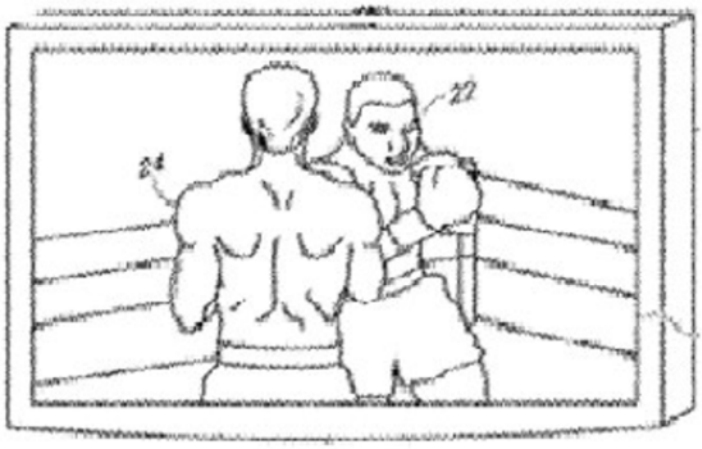


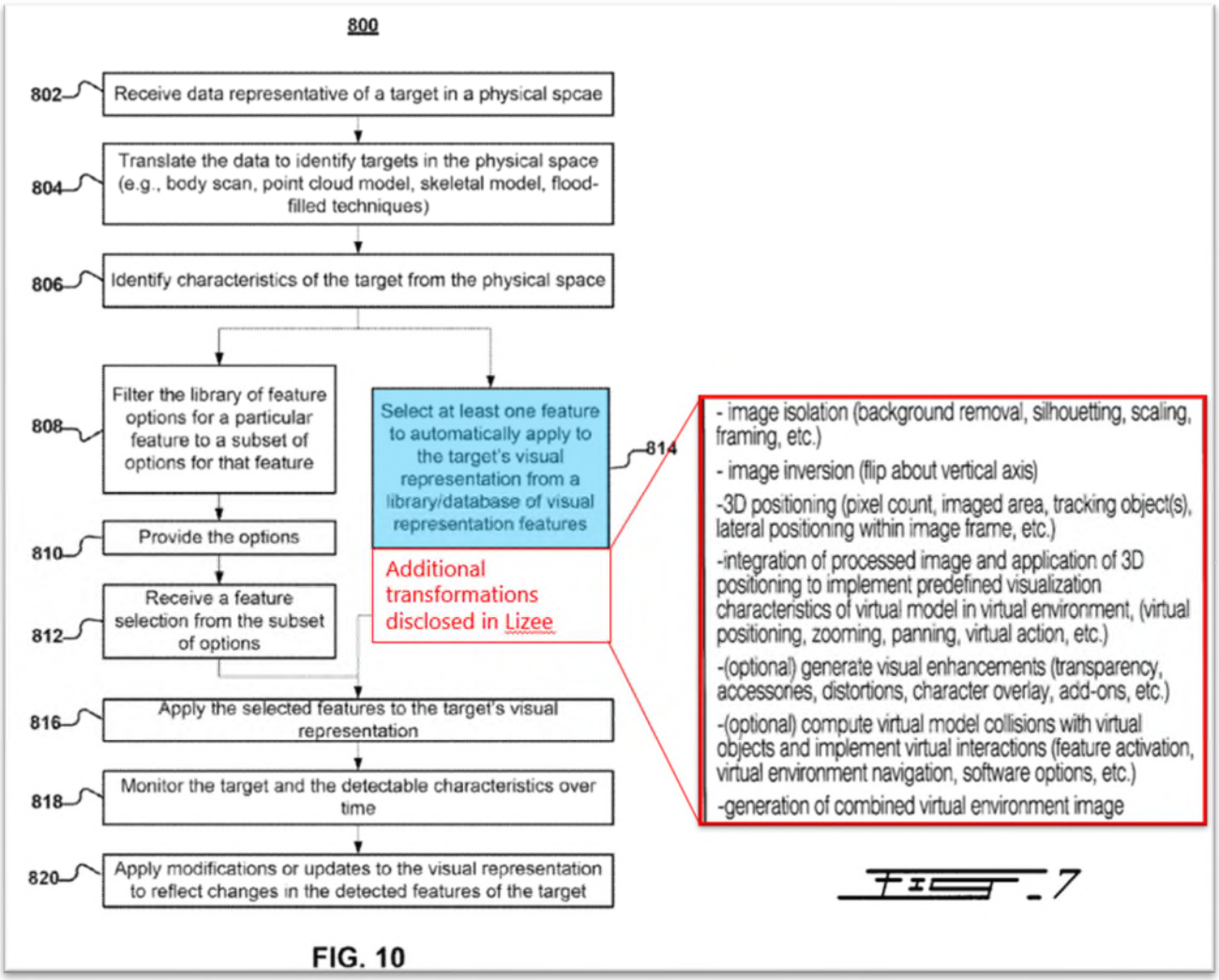
FIG. 1

FIG. 5

Ex. 1007 (Geisner) at Fig. 1 (partial, adjusted) (left); Ex. 1008 (Lizee) at Fig. 5 (partial) (right).

293. Combining Geisner and Lizee according to known methods would yield predicable results. Because of the similarity of hardware configurations in Geisner and Lizee, a POSITA would have understood that combining the teachings of Lizee into Geisner would yield predicable results. For example, a POSITA would have understood that it is feasible to incorporate the transformation algorithms disclosed in Lizee with the device disclosed in Geisner.

294. For example, a POSITA would have found it obvious to combine transformations in Lizee into the library/database of visual representation of features in Geisner as shown in the combined image below.



Ex. 1007 (Geisner) at Fig. 10 (annotated) (left); Ex. 1008 (Lizee) at Fig. 7 (partial) (right).

295. The resulting combination would use the transformations disclosed in Lizee to generate enhanced modified images with improved perceived quality. Ex. 1008 (Lizee) at [0084]-[0085]. The resulting combination would include using the transformations to generate better visual representations of a user with less perceived image distortions, providing the benefits disclosed in Lizee to the system of Geisner.

296. Finally, the teachings in Geisner and Lizee would lead a POSITA to combine them. For example, Geisner discloses a "human body model" (e.g., "skeletal model") that may "be provided to the computing environment for tracking . . . and rendering a visual representation." Ex. 1007 (Geisner) at [0139]. Lizee discloses "image rendering and transfer techniques including, but not limited to, image outlines, graphical overlays, shadows, 3D meshes and mappings and the like, may also be used to produce a similar effect." Ex. 1008 (Lizee) at [0036]. A POSITA considering this would have found it obvious to combine Lizee's transformation methods and algorithms with Geisner to provide image renderings, image outlines, or as an alternative/addition to rendering a visual representation in Geisner to generate enhanced modified images with improved perceived quality.

B. CLAIM 1:

a. 1[pre]: A computerized method for operating a display system, the display system having a camera, a monitor, and a processor, the method comprising:

297. To the extent the preamble is limiting, Geisner in view of Lizee renders this limitation obvious. Geisner discloses "a system 10 which tracks user movements and presents an on-screen representation of a user." Ex. 1007 (Geisner) at [0025], Figs. 1, 1B. The system comprises a camera (a capture device 20), a monitor (a visual device 16), and a processor (a computing system 12). Ex. 1007 (Geisner) at [0031].

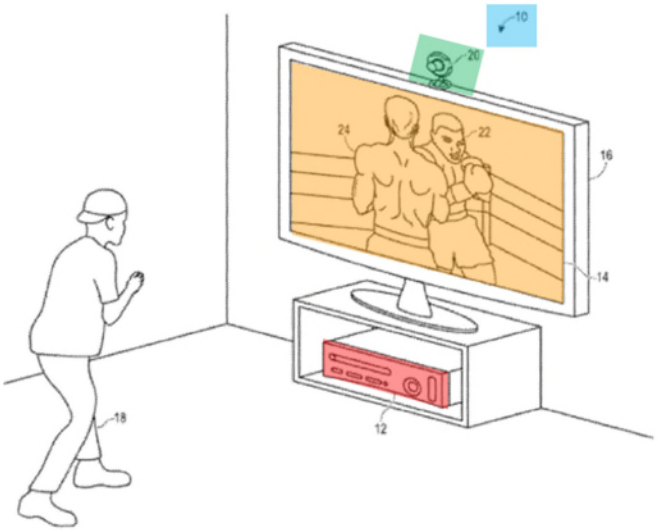


FIG. 1

Ex. 1007 (Geisner) at Fig. 1.

298. Figure 2 provides a more detailed configuration, in which the capture device 20 comprises a 3-D camera 26 and/or RGB camera 28 and a processor 32. Ex. 1007 (Geisner) at [0036]. The capture device communicates with the computing system 12, which comprises a processor 195 and may comprise a display device 194. Ex. 1007 (Geisner) at [0044].

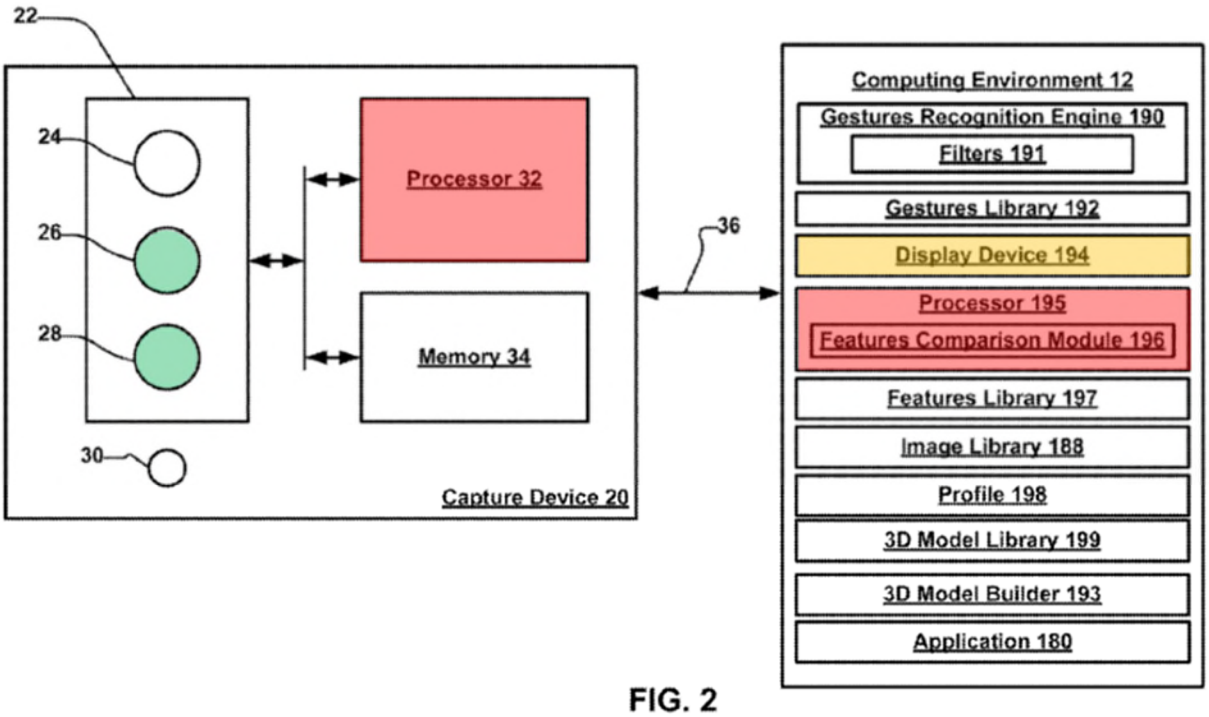


FIG. 2

Ex. 1007 (Geisner) at Fig. 2.

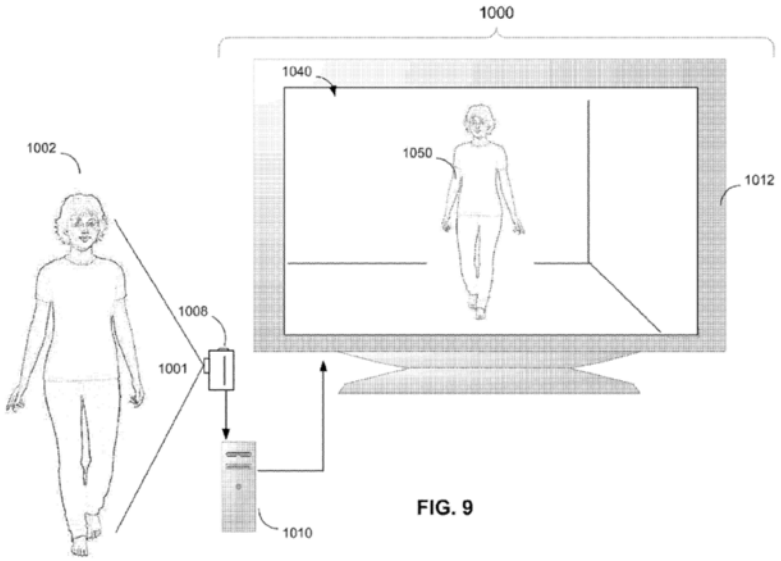
b. <u>1[a]: operating the camera to generate a stream of images of the user;</u>

299. Geisner in view of Lizee renders obvious "operating the camera to generate a stream of images of the user."

300. Geisner discloses operating the camera (capture device 20) to generate a stream of images (video) of the user. For example, Geisner discloses capture device 20 may be "configured to capture video with depth information including a depth image." Ex. 1007 (Geisner) at [0034]. Geisner also discloses "the capture device 20 may provide … images captured to the computing environment." Ex. 1007 (Geisner) at [0045]. Further, Geisner discloses that the device may "be used for target

recognition, where the target can be a user or an object." Ex. 1007 (Geisner) at [0034].

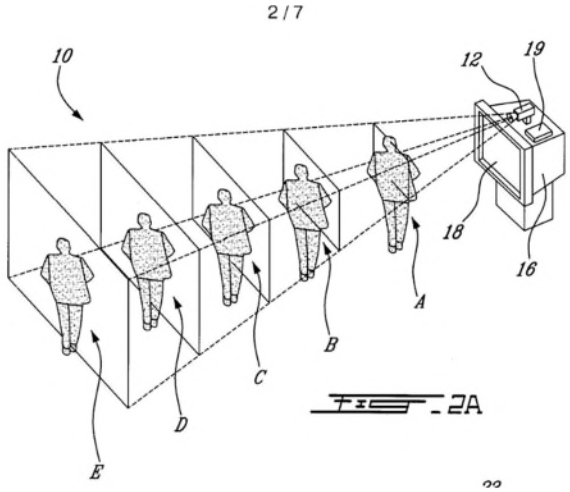
301. As shown in Figure 9 below, Geisner discloses a user and an on-screen mirror-like representation of the user generated by the system. Ex. 1007 (Geisner) at [0013]. As shown in Figure 9, the camera 1008 captures images of a user, and the visual representation 1050 of the user is displayed on a screen 1000. Ex. 1007 (Geisner) at Fig. 9, [0130], [0132], [0137].

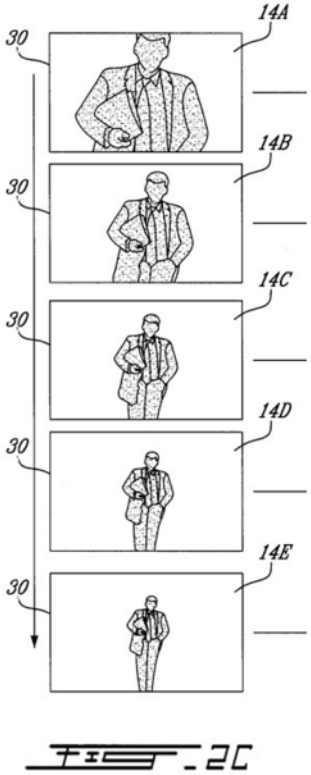


Ex. 1007 (Geisner) at Fig. 9.

302. Further, operating a camera to capture "a stream of images of the user," as claimed, was well known in the art. *See, e.g.*, Ex. 1008 (Lizee) at claims 1, 26, 40; *see also* Abstract, [0012], [0014], [0015], [0029], Figs. 1, 2A, 2B. Specifically, it was well known in the art to use an "image capture device for capturing successive frontal images of the user." *See* Ex. 1008 (Lizee) at claim 4; *see also* Ex. 1008

(Lizee) at claims 12, 42, 50, 68, 71. It was also well known in the art that "the image capture device may comprise a digital video camera" and/or "an analog camera providing a continuous video feed to an analog-to-digital (A/D) converter." Ex. 1008 (Lizee) at [0030]. Further, it was well known in the art to operate a camera to capture images of a user at successive positions. *See, e.g.*, Ex. 1008 (Lizee) at Fig. 2A, [0040]; *see also* Ex. 1008 (Lizee) at Fig. 2C.





Ex. 1008 (Lizee) at Fig. 2A (left), Fig 2C (right, partial).

303. Thus, Geisner in view of Lizee renders obvious "operating the camera to generate a stream of images of the user."

c. <u>1[b]: operating the processor to detect the presence of the user in the stream of images;</u>

304. Geisner in view of Lizee renders obvious "operating the processor to detect the presence of the user in the stream of images." Geisner discloses a tracking system "used to recognize, analyze, and/or track a human target such as the user." Ex. 1007 (Geisner) at [0030]. As shown in Figure 5, the system determines, at step 404, whether a scene includes a human target. Ex. 1007 (Geisner) at [0105]. If no human target is present, the system will re-scan the scene until a human target is identified, in which case, the system proceeds to generate a model of the human target at step 408. Ex. 1007 (Geisner) at [0105].

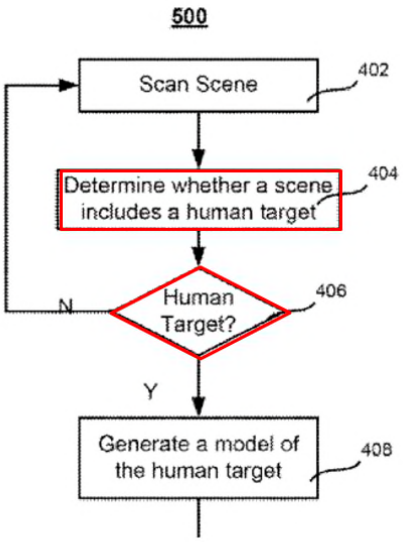


FIG. 5

Ex. 1007 (Geisner) at Fig. 5 (partial, annotated).

305. If a human target is identified, the system generates a model of the human target, an example of which is shown in Figure 8 below. Ex. 1007 (Geisner) at [0105].

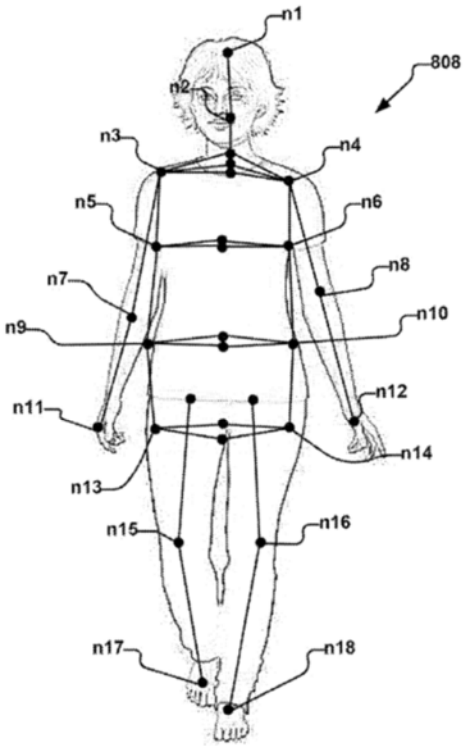


FIG. 8

Ex. 1007 (Geisner) at Fig. 8.

306. Figure 10 provides another example where a user is identified. As shown, the system determines "whether one or more targets in the scene corresponds to a human target such as a user" at step 804. Ex. 1007 (Geisner) at Fig. 10, [0139]. Geisner further provides various techniques to identify the target, including "a body

scan, point cloud models, skeletal models, flood-filled techniques, or the like." Ex. 1007 (Geisner) at [0139].

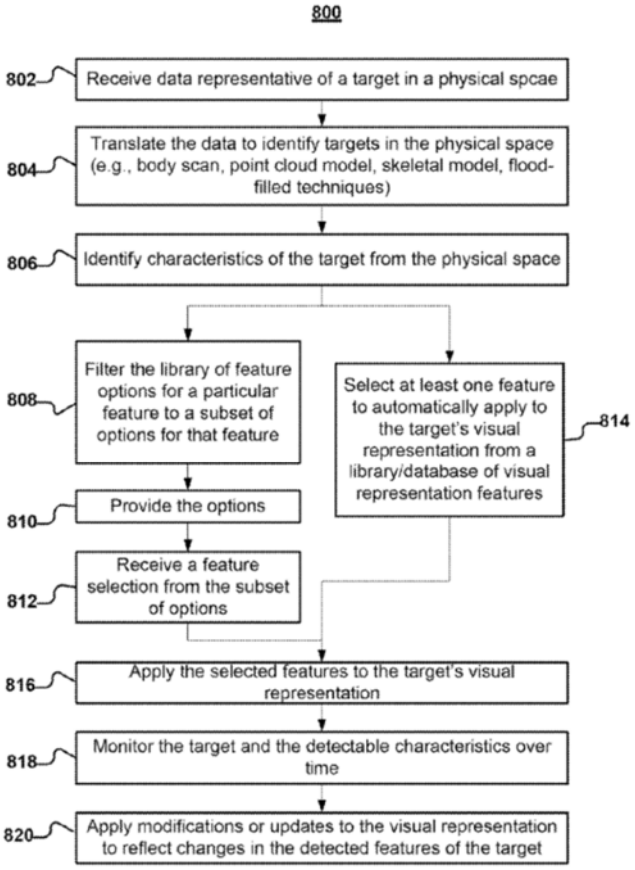


FIG. 10

Ex. 1007 (Geisner) at Fig. 10.

307. Further, Geisner discloses detecting a body part of a user for facial recognition and during gesture identification. For example, Geisner discloses detecting "five key data points (i.e., eyes, corner points of the mouth, and nose)" for facial recognition. Ex. 1007 (Geisner) at [0135]. Geisner further discloses identifying a user's hand in gesture identification. Ex. 1007 (Geisner) at [0051]

("identify that the user's hand is in front of the user's body"), [0127] ("various motions of the hands or other body parts may correspond to common system wide tasks such as navigate up or down in a hierarchical list"). Geisner further discloses identifying hands when generating a skeleton model of the user. Ex. 1007 (Geisner) at [0054] ("joints of the hand, wrists, elbows, knees, nose, ankles, shoulders, and where the pelvis meets the spine"). Moreover, detecting the position of a body part using a processor and tracking algorithms was well known in the art. For example, Lizee discloses detecting the position of a body part, disclosing that "the CPU 66 may detect a position of the user (e.g. pixel count, marker detection, etc.) and implement these position-dependent instructions/effects." Ex. 1008 (Lizee) at [0088], Claims 4, 9, 12, 23, 24, 25, 29, 32, 42, 50, 62-64, 68, 71, Fig. 7. Specifically, Lizee discloses that "conventional or proprietary tracking algorithms may be used to compute the locations of specific user features, for instance the user's face, eyes and/or hands from the image capture device 12 (e.g. tracking on a shape, size and/or color thereof)." Ex. 1008 (Lizee) at [0065].

308. A POSITA would have understood Geisner renders detecting the presence of the user in the stream of images obvious, in order to detect and track the motion of the user.

d. 1[c]: operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view

309. As explained below, Geisner in view of Lizee renders obvious "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera to generate modified images that appear to be captured from a different point of view of the camera's actual point of view." Below, I provide a more detailed analysis of the components of this limitation.

i. "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera"

310. Geisner discloses "operating the processor to apply adaptive transformation mapping to the stream of images captured by the camera." As stated in dependent claim 2 of the '110 patent, the mapping transformation includes "flips about a vertical axis, a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image." Ex. 1001 ('110 Patent) at claim 2.

311. Geisner discloses operating the processor to apply adaptive transformation mapping to the stream of images. For example, Geisner discloses applying a flip transformation. As shown in Figures 9 and 12, the displayed image 1050 is a flipped image of the user 1002 (as shown by the relative front-back position

of the left and right feet). Ex. 1007 (Geisner) at Figs. 9, 12; *see also* Ex. 1007 (Geisner) at Figs. 13, 17.

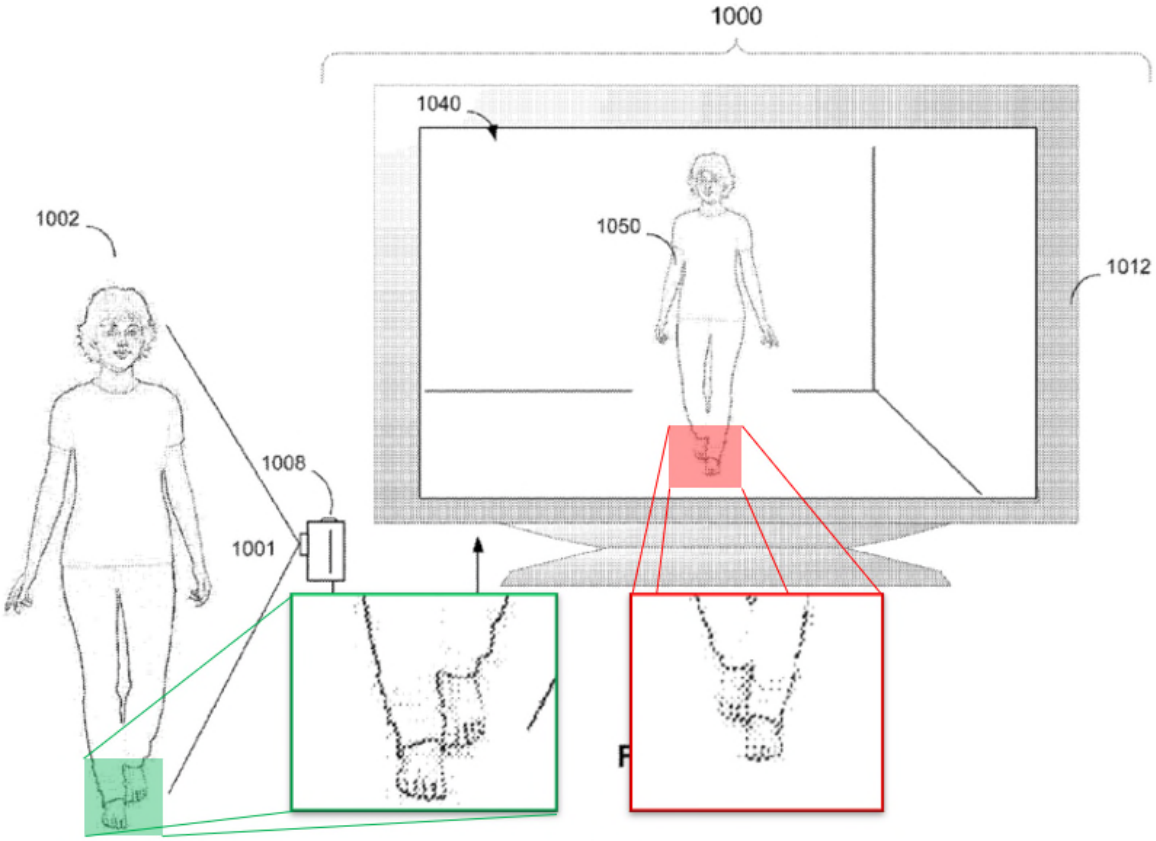


FIG. 9

Ex. 1007 (Geisner) at Fig. 9 (partial, annotated).

312. Geisner discloses other transformations including resizing and positioning. Ex. 1007 (Geisner) at [0062] (When the target is a human, the features of the target "may be altered by applying a color, positioning, or scaling to the target.").

313. Geisner further discloses rotation transformation. Geisner discloses a stereo vision configuration. Ex. 1007 (Geisner) at [0039], [0115]. Geisner also

discloses that "two or more physically separated cameras" may be employed to capture "a scene from different angles" Ex. 1007 (Geisner) at [0039]. A POSITA would have found it obvious to rotate images that are captured from two synchronized angled cameras while performing "stereo vision analysis" and to form a single image with depth information. Ex. 1007 (Geisner) at [0039], [0115]. A POSITA would have known that rotating the images captured from cameras can offset the angular deviation of the camera from the central line (the virtual point of view).

314. Additionally, Lizee discloses methods and algorithms to apply transformations as mentioned above and further transformations.

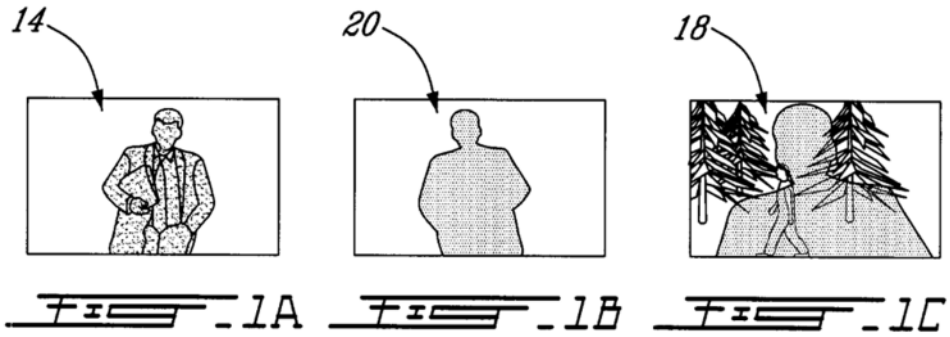
315. If the Patent Owner argues that Geisner does not disclose this limitation, Geisner in view of Lizee renders it obvious. Lizee discloses applying "position-dependent instructions/effects" to the captured images, including "background removal, silhouetting, scaling, framing, mirrored transformation." Ex. 1008 (Lizee) at [0088]. Lizee further discloses applying transformations to create one or more selected displacements of the virtual representation, including "a forward displacement, a backward displacement, an orbital displacement, a rotational displacement and a side displacement." Ex. 1008 (Lizee) at [0063], claim 6; *see also* claims 30, 34, 40, 42-44, 49-51, 53-55, 59, 62-65, 67-71, 76.

316. Further, Lizee discloses operating the processor to applying various transformations. Lizee discloses that "the CPU 66 may detect a position of the user (e.g. pixel count, marker detection, etc.) and implement these position-dependent instructions/effects." Ex. 1008 (Lizee) at [0088].

317. The instructions/effects comprise various transformations. For example, Lizee describes "e.g. background removal, silhouetting, scaling, framing, mirrored transformation." Ex. 1008 (Lizee) at [0088]. Further, Lizee discusses "wherein said virtual camera action comprises at least one of a camera moving action, a camera rotation action, a camera panning action, a camera zooming action, a camera orbiting action and a camera dollying action." Ex. 1008 (Lizee) at claim 13; *see also* claim 6 ("said predetermined motion comprises a displacement of the virtual representation in the virtual environment comprising at least one of a forward displacement, a backward displacement, an orbital displacement, a rotational displacement and a side displacement"), claims 30, 34, 40, 42-44, 49-51, 53-55, 59, 62-65, 67-71, 76. Moreover, Lizee discloses "a mirrored silhouette transformation of said user image such that the virtual representation shares a substantially same orientation as the user." Ex. 1008 (Lizee) at claim 1.

318. Lizee discloses at least three transformations: flipping, resizing, and rotating. Lizee discloses a mirroring transformation, i.e., flips about a vertical axis. Ex. 1008 (Lizee) at Abstract, [0012]-[0015], claim 1. For example, Lizee discloses

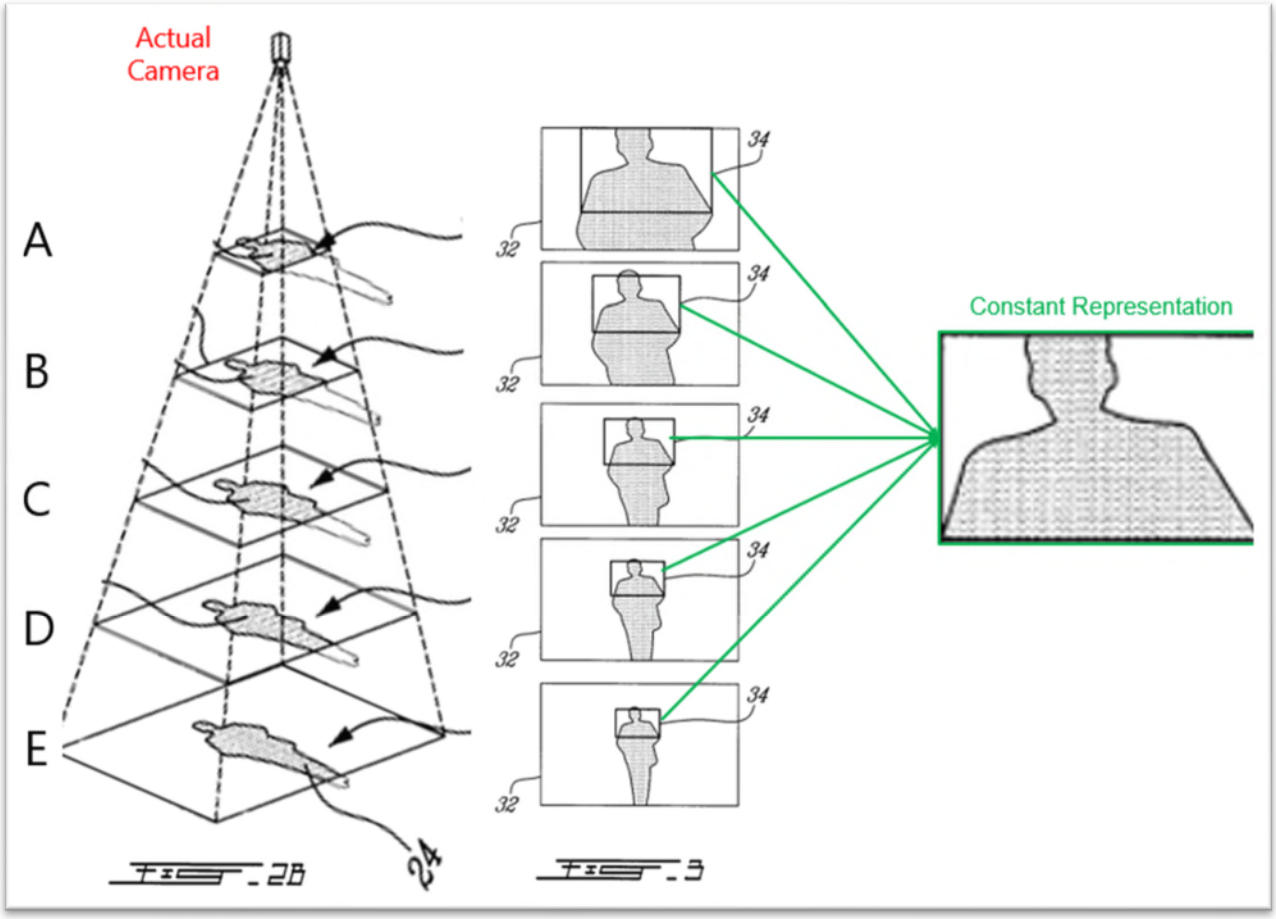
and claims "a mirrored silhouette transformation of the user image such that the virtual representation shares a substantially same orientation as the user." Ex. 1008 (Lizee) at [0012], claim 1. Figure 1B shows mirrored images generated from captured images of Figure 1A. Ex. 1008 (Lizee) at Figs. 1A, 1B, [0019]. Lizee explains, "the captured image 14 is isolated by the computing device 19, inverted from left to right about a vertical axis, and transformed into a translucent silhouette image 20 (Figure 1B) for display on the screen 18 (Figure 1C)." Ex. 1008 (Lizee) at [0034].



Ex. 1008 (Lizee) at Figs. 1A-1C.

319. Second, Lizee discloses an adaptive resizing transformation (e.g., scaling). The resizing transformation is applied for compensating changes in the size of a user caused by changes of the distance from the user to the actual camera. Ex. 1008 (Lizee) at [0051] ("as the user moves from position A to E in the real world (Figure 2A), the virtual camera 28 maintains a constant framing 34").

320. As shown in the combined figure below, the constant framing is maintained by resizing the captured image based on the distance. Figure 2B (left) shows that as the distance varies when the user moves from point A to point E, the size of the user in the captured image is reduced. Ex. 1008 (Lizee) at Fig. 2B. However, as shown in Figure 3 (right), the size of the user in the displayed image remains constant because cropping and a resizing transformation are applied to the captured images.



Ex. 1008 (Lizee) at Fig. 2B (left, partial, rotated, annotated), Fig. 3 (right, partial

annotated).

321. In addition, Lizee discloses that "the framing 34 of the virtual camera 28 may also be altered dynamically to provide a zooming action." Ex. 1008 (Lizee) at [0051]-[0055], Fig. 3. Lizee provides that "a constant representation thereof, limited herein by the virtual camera framing 34, is maintained at every position." Ex. 1008 (Lizee) at [0051]-[0052].

322. Finally, Lizee discloses a rotation transformation. For example, Lizee discloses and claims "a camera rotation action" and "a camera orbiting action." Ex. 1008 (Lizee) at claim 13; *see also* Ex. 1008 (Lizee) at [0063] ("as the system 10 detects a motion of the user, the computing device 19 may be configured to . . . implement various modifications to a visualization thereof from the perspective of the virtual camera 28"), ("user motions may be used to implement virtual displacements of the virtual model such as a forward displacement, a backward displacement, an orbital displacement, a rotational displacement and a side displacement"). Further in claim 13, Lizee discloses "virtual camera action comprises at least one of a camera moving action, a camera rotation action, a camera panning action, a camera zooming action, a camera orbiting action and a camera dollying action." Ex. 1008 (Lizee) at claim 13. As discussed, a POSITA would have been motivated to implement Lizee's adaptive transformation mapping in Geisner to generate enhanced modified images with improved perceived quality, and a

POSITA would have had a reasonable expectation of success in doing so. *See supra* §XIII.A. Further, a POSITA would have understood that it is feasible to incorporate the transformation algorithms disclosed in Lizee with the device disclosed in Geisner, at least because both Geisner and Lizee have similar hardware configurations and teach similar known transformation techniques. *See supra* §XIII.A, ¶¶290-296.

ii. "operating the processor …to generate modified images that appear to be captured from a different point of view of the camera's actual point of view"

323. Geisner discloses "generat[ing] modified images that appear to be captured from a different point of view of the camera's actual point of view." For example, Geisner discloses applying a scaling transformation to the images with targets as discussed above. *See* §XIII.B.d.i, ¶312. As explained above with respect to Haan, a POSITA would have understood that the resizing transformation is equivalent to moving a camera back or forth while a rotation transformation results in changing the angle of point of view. *See supra* §XI.B.f.i, ¶192, §XI.B.f.ii, ¶199.

324. A POSITA would have understood that a scaling transformation results in moving the viewpoint back or forth. Geisner further discloses a stereo vision configuration. Ex. 1007 (Geisner) at [0115]; *see also* Ex. 1007 (Geisner) at [0039]. As explained above, a POSITA would have found it obvious to rotate images that are captured from two synchronized angled cameras while performing "stereo vision

analysis." Ex. 1007 (Geisner) at [0039], [0115]; *see also supra* §XIII.B.d.i, ¶313. A POSITA would have known that rotating the images captured from cameras can offset the angular deviation of the cameras from the central line (the virtual point of view).

325. If the Patent Owner argues that Geisner does not disclose this limitation, Geisner in view of Lizee renders it obvious. For example, Lizee discloses that "a position or viewpoint of the virtual camera 28 could be moved to reflect a change in the viewing direction of the virtual silhouette 24." Ex. 1008 (Lizee) at [0057]. Lizee discloses and claims that "said computing device being further configured to generate a virtual camera such that the virtual representation is displayed from a perspective of said virtual camera." Ex. 1008 (Lizee) at claim 11; *see also* Ex. 1008 (Lizee) at claims 12, 15, 29 ("modifying said perspective as a function of said user motion"). The various "virtual camera perspective modifications . . . may include moving, rotating, panning, zooming, orbiting and/or dollying the virtual camera relative to the silhouette model 24 and/or the virtual environment." Ex. 1008 (Lizee) at [0063]. A POSITA would have understood that modifying perspective results in an image appearing to be captured from a different point of view. *See supra* §VI.B.c

326. As explained above, Lizee discloses at least three transformations: flipping, resizing, and rotating. *See supra* §XIII.B.d.i, ¶318. As explained above with

respect to Haan, the resizing transformation is equivalent to moving a camera closer or farther from the imaged object. *See supra* §XI.B.f.ii, ¶199; *see also supra* §XI.B.d.ii, ¶182. Similarly, Lizee discloses an adaptive resizing transformation (e.g., scaling) and explains that the degree of resizing correlates to the distance from the user to the camera. Ex. 1008 (Lizee) at [0051]-[0052]. Lizee teaches that the resizing transformation is applied for compensating changes in the size of a user that are caused by changes of the distance from the user to the actual camera. Ex. 1008 (Lizee) at [0051] ("as the user moves from position A to E in the real world (Figure 2A), the virtual camera 28 maintains a constant framing 34"). Lizee teaches that constant framing is maintained by resizing the captured image based on the distance. Ex. 1008 (Lizee) at [0051]-[0052], Fig. 2B, Fig. 3; *see also supra* §XIII.B.d.i, ¶¶319-320.

327. Lizee further discloses that "the framing 34 of the virtual camera 28 may also be altered dynamically to provide a zooming action." Ex. 1008 (Lizee) at [0051]-[0055], Fig. 3. Lizee provides that "a constant representation thereof, limited herein by the virtual camera framing 34, is maintained at every position." Ex. 1008 (Lizee) at [0051]-[0052]. Therefore, the displayed images in Lizee appear to be captured from a different point of view than the camera's actual point view since, regardless of the variation of the actual distance between the user and the actual

camera, the displayed images appear to be taken from a virtual camera, whose distance to the user is fixed.

328. Lizee further discloses "control[ling] various virtual camera perspective modifications, which may include moving, rotating, panning, zooming, orbiting and/or dollying the virtual camera relative to the silhouette model 24 and/or the virtual environment." Ex. 1008 (Lizee) at [0063]. A POSITA would have understood that modifying the virtual camera perspective results in an image appearing to be captured from a different point of view. Ex. 1008 (Lizee) at [0063], Fig. 2B, Fig. 3; *see also supra* §VI.B.c. *supra* §XIII.A, ¶¶290-296. A POSITA would have been motivated to implement Lizee's modified image generation in Geisner to generate enhanced modified images with improved perceived quality, and a POSITA would have had a reasonable expectation of success in doing so. *See supra* §XIII.A; ¶¶285-296.

e. <u>1[d]: displaying the modified images on the monitor;</u>

329. Geisner in view of Lizee renders obvious displaying the modified images on the monitor. For example, Geisner discloses and claims "rendering a representation of the scene and the human target on a display device." Ex. 1007 (Geisner) at claim 1, [0004], [0024]. As shown in Figure 12, the system displays modified images on the screen 1200.

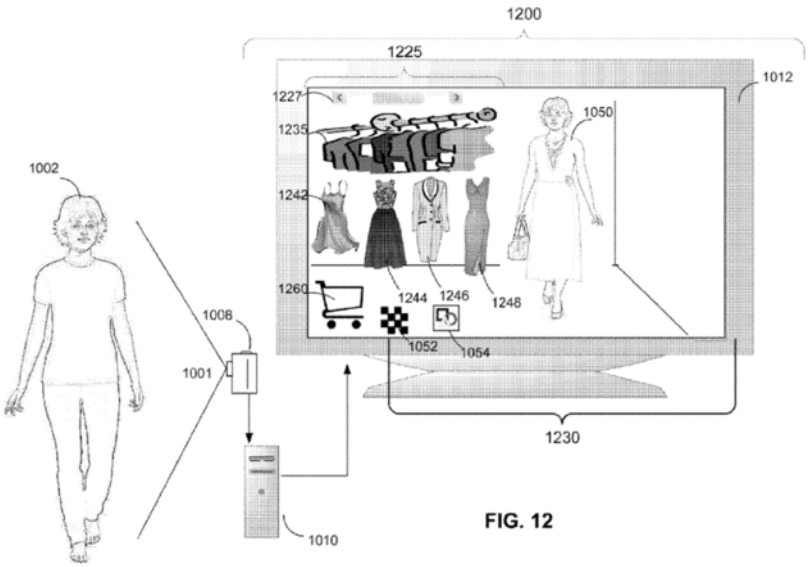


FIG. 12

Ex. 1007 (Geisner) at Fig. 12; *see also* Ex. 1007 (Geisner) at Figs. 1, 1b, 9, 13-15, and 17.

f. <u>1[e]: wherein the different point of view corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera; and</u>

330. Geisner in view of Lizee renders obvious this limitation. As explained above, Geisner in view of Lizee discloses that the modified images appear to be taken from a different point of view by applying various transformations. S*ee supra* §XIII.B.d. Geisner in view of Lizee discloses that the different point of view disclosed in the combination corresponds to an image taken from a different angle and a different distance than the actual images captured by the camera.

i. "wherein the different point of view corresponds to an image taken from a different angle… than the actual images captured by the camera"

331. Lizee discloses that the different point of view corresponds to a different angle than the actual images captured by the camera. Lizee discloses modifying captured images by moving "a position or viewpoint of the virtual camera … to reflect a change in the viewing direction." Ex. 1008 (Lizee) at [0057]. Further, Lizee discloses "implement[ing] virtual displacements of the virtual model 24 such as a forward displacement, a backward displacement, an orbital displacement, a rotational displacement and a side displacement." Ex. 1008 (Lizee) at [0063]. Lizee further discloses "control[ling] various virtual camera perspective modifications, which may include moving, rotating, panning, zooming, orbiting and/or dollying the virtual camera relative to the silhouette model 24 and/or the virtual environment." Ex. 1008 (Lizee) at [0063]. A POSITA would have understood that both an orbital displacement and a rotational displacement would result in a change of the angle of the view point because the displayed silhouette or image needs to be rotated to be properly positioned for a rotating or orbiting placement. Rotating an image or silhouette changes the angle of the virtual camera (i.e., the modified images appear to have been captured from an angle that is different than the actual angle).

332. Additionally, Lizee further discloses that the mirrored silhouette images may be "viewed a virtual camera 28 positioned behind the silhouette model

24 and providing a rear view thereof such that the user and the silhouette model 24 share a substantially same orientation." Ex. 1008 (Lizee) at [0039]. Thus, the mirror (flipping) transformation in Lizee renders a silhouette that appears to be taken from the back of the user while the camera is actually located in front of the user (i.e., a 180- degree rotation of the virtual camera). Ex. 1008 (Lizee) at [0039], Fig. 7.

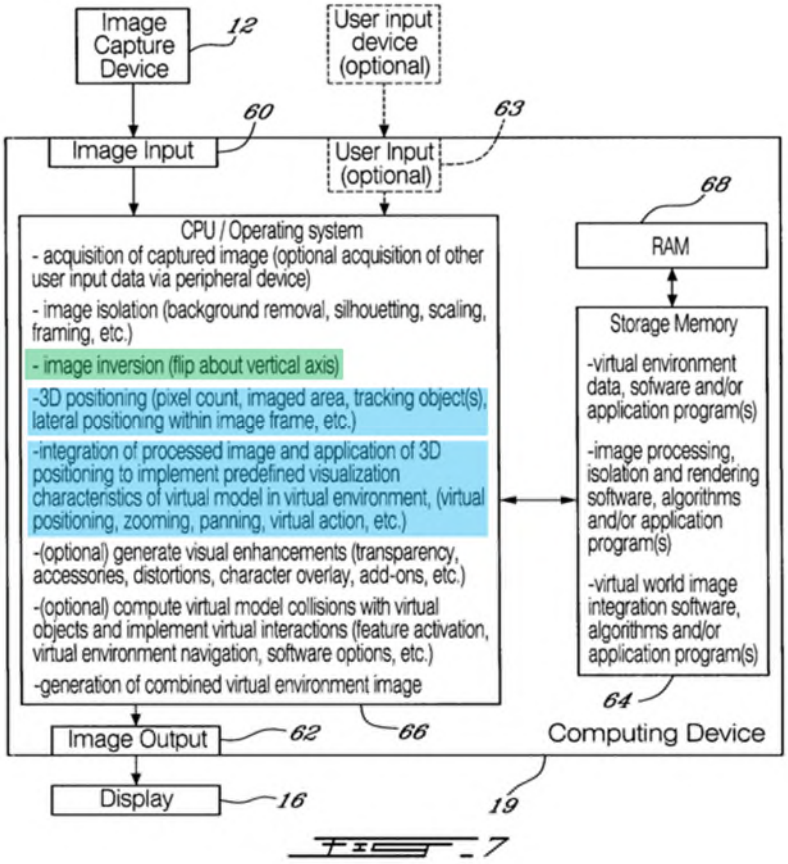


FIG. 7

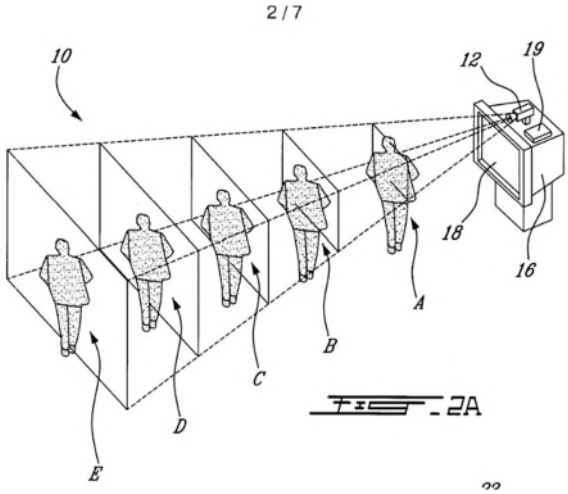
Ex. 1008 (Lizee) at Fig. 7 (annotated).

ii. "…a different distance than the actual images captured by the camera

333. Lizee teaches that the different point of view corresponds to a different distance than the actual images captured by the camera. As explained above with

respect to Haan, the resizing transformation is equivalent to moving a camera closer to or father from the image object. *See supra* §XI.B.f.ii, ¶199; *see also supra* §XI.B.d.ii, ¶198.

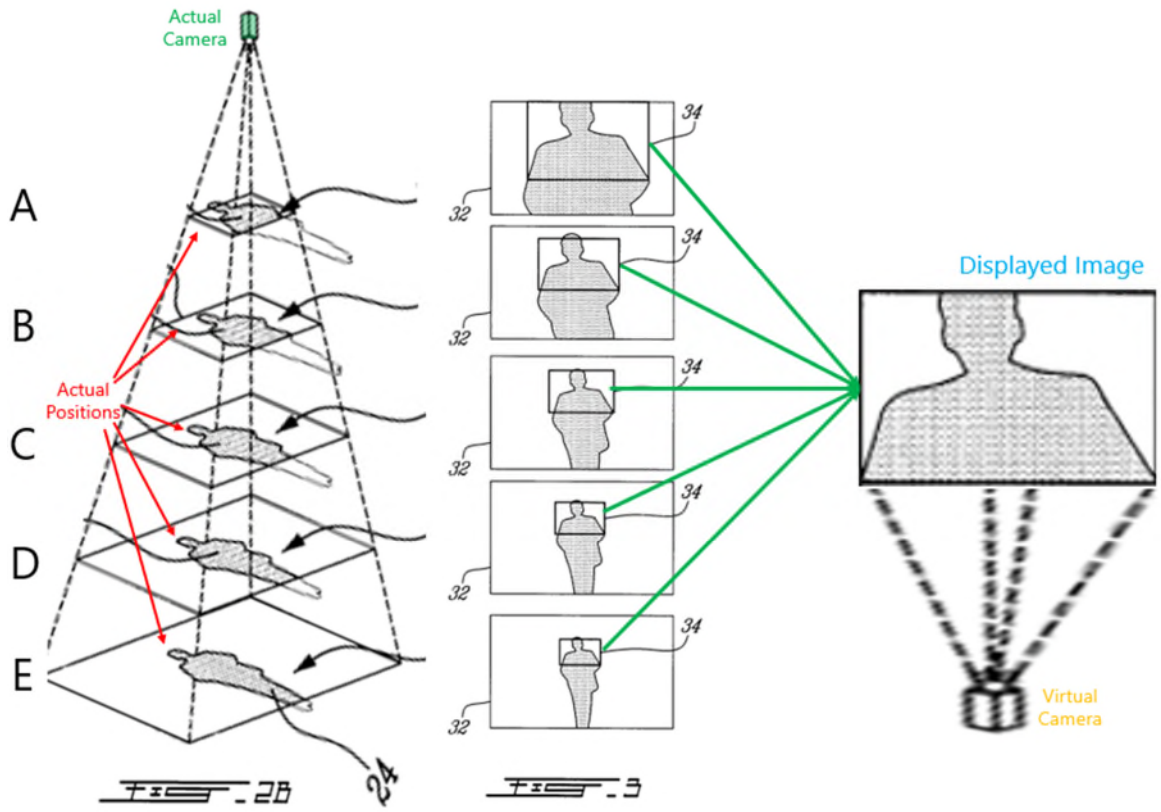
334. As shown in the figure below, Lizee discloses operating the camera to capture images of a user at five successive positions as "the user moves … from positions A to E." Ex. 1008 (Lizee) at [0040]; *see also* Ex. 1008 (Lizee) at Fig. 2A.



Ex. 1008 (Lizee) at Fig. 2A.

335. As shown in figures below, while the actual distance varies, i.e., as the user moves away from the camera from point A to E (Figure 2B), the user representation in the captured images becomes smaller (Figure 3), and the system adjusts the scaling of the displayed images to maintain a constant size of the user representation. Ex. 1008 (Lizee) at [0051]. Therefore, the displayed images appear to be captured from a virtual camera with a fixed distance to the user. A POSITA would have understood that when the user locates at points A through E, the

displayed image has a virtual point of view different from the actual point of view in terms of distance.

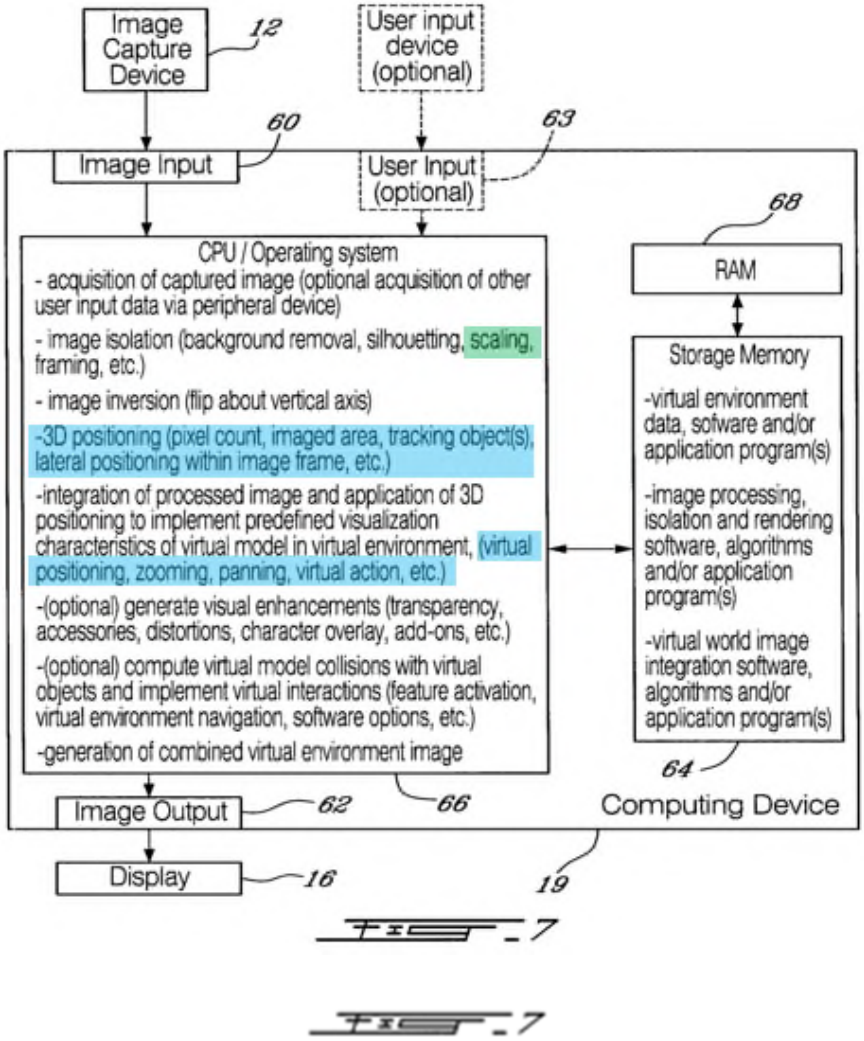


Ex. 1008 (Lizee) at Fig. 2B (left, partial, rotated, annotated); Fig. 3 (right, partial annotated).

336. Additionally, Lizee provides a dynamic zooming (i.e., resizing) transformation. Ex. 1008 (Lizee) at [0056] ("the framing 34 of the virtual camera 28 may also be altered dynamically to provide a zooming action"). Lizee explains "[a]s the distance of the user from the image capture device 12 increases, the virtual camera framing 34 may also increase, thereby providing a zoom-out effect." Ex. 1008 (Lizee) at [0056].

iii. Both angle and distance transformations

337. Further, Lizee teaches that the distance and angle of the virtual view point of a given image may both be changed. As shown in Figure 7 below, Lizee discloses applying to a user image both scaling (changing the distance) and positioning (changing the angle). Ex. 1008 (Lizee) at Fig. 7.



Ex. 1008 (Lizee) at Fig. 7 (annotated).

Additionally, a POSITA would have found it obvious to apply both angle and distance modifications in the Geisner-Lizee combination. As shown below, Geisner (left) and Lizee (right) concern a similar use case, where a user representation needs to be scaled and rotated for proper displacements:

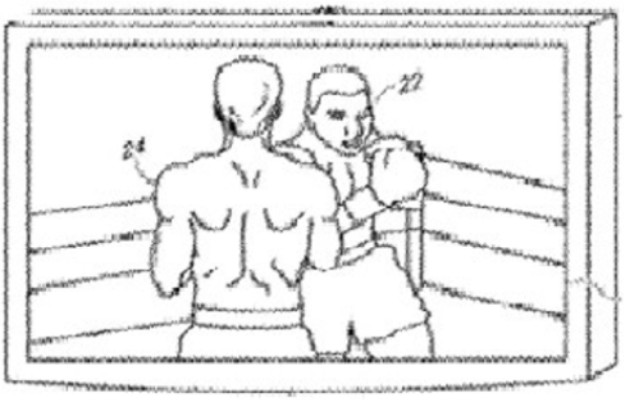
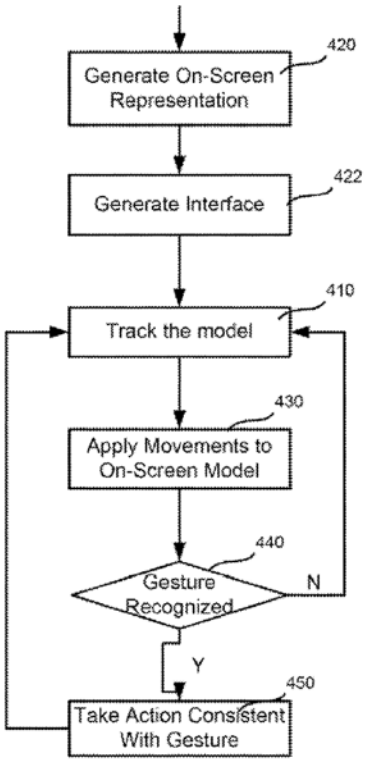


FIG. 1

FIG. 5

Ex. 1007 (Geiser) at Fig. 1 (left) and Ex. 1008 (Lizee) at Fig. 5 (right).

338. As shown in Figure 5 of Geisner, a POSITA would have understood that to enrich the user experience in a virtual environment, Lizee's scaling can be applied at step 420 (Generate On-Screen Representation), and Lizee's positioning can be applied at step 430 (Apply Movements to On-Screen Model).



Ex. 1007 (Geiser) at Fig. 5.

Therefore, a POSITA would have found it obvious that the corresponding virtual point of view has a different distance and a different angle than the actual point of view of the captured image. *See supra* §§ VI.B.c-VI.B.f.

g. <u>1[f]: wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera.</u>

339. As explained with respect to claim elements 1[c] and 1[e], Geisner in view of Lizee renders obvious applying a mapping transformation that changes imaging perspective of the images captured by the camera. *See supra* §§XIII.B.d,

XIII.B.f. Further, Geisner in view of Lizee renders obvious "wherein during mirror tracking mode the processor utilizes the angle of the images captured by the camera and applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera."

i. Geisner teaches that "during mirror tracking mode the processor utilizes the angle of the images captured by the camera"

340. Geisner teaches a tracking system "used to track user motions and execute a shopping application." Ex. 1007 (Geisner) at [0003]; *see also* Abstract. Geisner discloses that, while tracking the user motion, the system obtains image depth information by utilizing two or more cameras viewing a scene from different angles. Ex. 1007 (Geisner) at [0039], [0102]; *see also* Ex. 1007 (Geisner) at Abstract. A POSITA would have found it obvious to use the angles of the separated cameras to register their corresponding captured images in order to properly generate depth information.

ii. Geisner teaches "applies parametric input for a mapping transformation that changes imaging perspective of the images captured by the camera"

341. Geisner further teaches "apply[ing] parametric input for a mapping transformation." With respect to "parametric input," the specification of the '110 Patent suggests that "parametric input" may include user parameters such as weight and height, or depth measurements used to describe and control a user's avatar. Ex.

1001 ('110 Patent) at 21:13-15 ("only a 3D sensor will generate depth measurement that will be used to control a parametric model of the user's avatar."), *see also* 29:64-30:1 ("the system is able to identify the user and also calculate parameters of the user, e.g., weight, height, etc."). The '110 Patent further describes applying the depth data for a mapping transformation. Ex. 1001 ('110 Patent) at 21:24-28 ("The data from a depth sensor or 3D camera can pass a similar mapping transformation engine and the result can be the base of a more accurate 3D avatar of the user.").

342. Geisner teaches a similar configuration and the same use of parametric inputs. For example, Geisner discloses acquiring image depth information (an example of the claimed "parametric input") using two or more stereo cameras (a type of 3D sensor) viewing a scene from different angles. Ex. 1007 (Geisner) at [0039]. The depth information is used to generate a skeleton model of the user, which includes various parameters (another example of the claimed "parametric input"), such as "spots on that user's skeleton, joints of the hand, wrists, elbows, knees, nose, ankles, shoulders." Ex. 1007 (Geisner) at Fig. 8, [0054], [0123], [0054], [0101]. The spots of joints are then processed to identify user's features (yet another example of the claimed "parametric input"), such as a user's gesture or motion. Ex. 1007 (Geisner) at [0053]-[0054], [0062]. Geisner teaches applying the parametric inputs (e.g., skeleton model, depth information, or user's features) for transformations to

change the "positioning, or scaling [of] the target" or "of a selected feature based on the detected target features." Ex. 1007 (Geisner) at [0059]-[0066].

343. Additionally, Geisner discloses applying "exist[ing] skeletal mapping techniques to capture a person with a depth camera" and "[o]ther techniques includ[ing] transforming the image into a body model representation of the person and transforming the image into a mesh model representation of the person." Ex. 1007 (Geisner) at [0054].

C. CLAIM 2. THE METHOD OF CLAIM 1, WHEREIN THE MAPPING TRANSFORMATION COMPRISES A COMBINATION OF AT LEAST TWO OF THE FOLLOWING TRANSFORMATIONS: VERTICAL FLIPS ABOUT A VERTICAL AXIS, A PROJECTION TRANSFORMATION, ROTATION, RESIZE, TWO-DIMENSIONAL OFFSET, OPTIC DISTORTION CORRECTION, AND POLYNOMIAL MAPPING ON PART OR ENTIRE CAPTURED IMAGE.

344. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious wherein "the mapping transformation comprises a combination of at least two of the following transformations: flips about a vertical axis, a projection transformation, rotation, resize, two-dimensional offset, optic distortion correction, and polynomial mapping on part or entire captured image."

345. As explained above with respect to claim limitations 1[c] and 1[e], the mapping transformations in Lizee and Geisner comprise a combination of at least flips about a vertical axis, resize, and rotation. *See supra* §§XIII.B.d, XIII.B.f.

Further, Geisner in view of Lizee discloses additional transformations. For example, Lizee discloses a projection transformation. Ex. 1008 (Lizee) at [0040]. Lizee discloses projecting images of the user on a virtual panel in the virtual environment. Ex. 1008 (Lizee) at [0039] ("the silhouette model 24 . . . is conceptually defined by the projection 26 of the silhouette image 20 . . . on a virtual panel 27"). The projection compensates for the displacement of the user relative to the camera. Ex. 1008 (Lizee) at [0040]. A POSITA would have been motivated to implement Lizee's additional transformations in Geisner to generate enhanced modified images with improved perceived quality, and a POSITA would have had a reasonable expectation of success in doing so. *See supra* §XIII.A; ¶¶285-296. Additionally, Geisner discloses applying a positioning transformation. Ex. 1007 (Geisner) at [0062] (When the target is a human, the features of the target "may be altered by applying a color, positioning, or scaling to the target."). *See supra* §XIII.B.d.

D. CLAIM 3. THE METHOD OF CLAIM 1, WHEREIN DURING IDLE MODE THE CAMERA CONTINUOUSLY SENDS CAPTURED IMAGES TO THE PROCESSOR, AND THE USER PRESENCE IS IDENTIFIED BY THE PROCESSOR DETECTING A BODY PART IN THE IMAGES.

346. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images."

347. Geisner discloses a modification mode that can be triggered by a user gesture. Ex. 1007 (Geisner) at [0052]. The triggering gesture may interrupt "the current state of the application" and put the system into a modification mode. Ex. 1007 (Geisner) at [0052]. The modification mode may pause the application (i.e., enter into an idle mode) and resume the application upon exiting the mode. Ex. 1007 (Geisner) at [0052]. During the modification mode, a user can modify the visual representation while the application is paused. Ex. 1007 (Geisner) at [0052].

348. To modify the visual representation, the system keeps "monitoring the target and detecting a change in the detected target feature" and "updating the visual representation of the target . . . based on the change in the detected target feature." Ex. 1007 (Geisner) at [0048]. As explained above, gesture identification in Geisner requires identifying at least one user body part. *See supra* §XIII.B.c, ¶307. Therefore, a POSITA would understand that during the modification mode, when the application is paused (i.e., during idle mode), the camera continuously sends captured images to the processor for gesture identification, which requires identifying the presence of a user by detecting a user body part.

E. CLAIM 4. THE METHOD OF CLAIM 3, WHEREIN THE BODY PART COMPRISES ONE OF LEGS, FACE, EYES, AND HANDS.

349. Geisner in view of Lizee renders obvious the method of claim 3. *See supra* §XIII.D. Further, Geisner in view of Lizee renders obvious "wherein the body

part comprises one of legs, face, eyes, and hands." As discussed above, during idle mode, Geisner discloses identifying a user's hand in gesture identification. Ex. 1007 (Geisner) at [0051] ("identify that the user's hand is in front of the user's body"), [0127]; *See supra* §XIII.D. Geisner also discloses detecting a user's eyes, corner points of the mouth, and nose for facial recognition. Ex. 1007 (Geisner) at [0135]. Geisner further discloses identifying hands when generating a skeleton model of the user. Ex. 1007 (Geisner) at [0054] ("joints of the hand, wrists, elbows, knees, nose, ankles, shoulders, and where the pelvis meets the spine"). *See supra* §XIII.B.c, ¶307. Further, as discussed above, detecting the position of a body part using a processor and tracking algorithms was well known in the art. *See, e.g.*, Ex. 1008 (Lizee) at [0065]. *See supra* §XIII.B.c, ¶307.

F. CLAIM 5. THE METHOD OF CLAIM 1, FURTHER COMPRISING RECEIVING USER INPUT AND OPERATING THE PROCESSOR TO VARY IMAGE MAGNITUDE OR FIELD-OF-VIEW OF THE MODIFIED IMAGES ACCORDING TO THE USER INPUT.

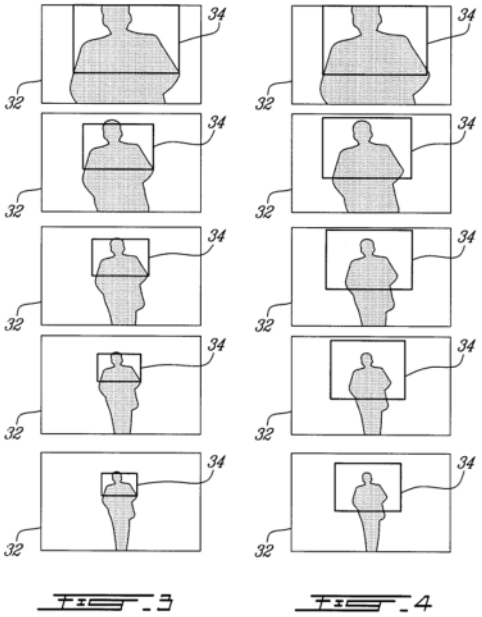
350. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "comprising receiving user input and operating the processor to vary image magnitude or field-of-view of the modified images according to the user input." As previously explained, a captured user movement can be considered a type of user input, at least based on the Patent Owner's interpretation. *See supra* §XI.F, ¶224.

351. Lizee discloses receiving user input and operating the processor to vary image magnitude or field-of-view according to the user input. Specifically, Lizee teaches receiving user motion to apply zoom-in/out function, which varies image magnitude or field-of-view. For example, Lizee discloses that "the user may use a backward motion to implement a zoom-out function." Ex. 1008 (Lizee) at [0056]. Lizee explains that the zoom-out effect can be carried out by increasing the virtual camera framing when the distance of the user from the actual camera increases. Ex. 1008 (Lizee) at [0056]. A POSITA would have been motivated to implement Lizee's receiving user input and operating the processor to vary image magnitude or field-of-view according to the user input in Geisner to generate enhanced modified images with improved perceived quality, and would have had a reasonable expectation of success in doing so. *See supra* §XIII.A; ¶¶285-296.

G. CLAIM 6. THE METHOD OF CLAIM 5, WHEREIN THE PROCESSOR VARIES THE IMAGE MAGNITUDE OR FIELD-OF-VIEW MAPPING TO DISPLAY A PARTIAL OR FULL BODY OF THE USER IN THE MODIFIED IMAGES.

352. Geisner in view of Lizee renders obvious the method of claim 5. *See supra* §XIII.F. Further, Geisner in view of Lizee renders obvious "wherein the processor varies the image magnitude or field-of-view mapping to display a partial or full body of the user in the modified images."

353. Lizee teaches varying the image magnitude or filed-of-view to display a partial or full body of the user. For example, as shown in Figures 3 and 4 in Lizee, the processor varies the virtual camera framing 34 (an example of virtual field-of-view) to display a partial or full body of the user in the modified images. Ex. 1008 (Lizee) at Figs. 3, 4, [0051]-[0056]. A POSITA would have been motivated to implement Lizee's varying the image magnitude or filed-of-view to display a partial or full body of the user in Geisner to generate enhanced modified images with improved perceived quality, and a POSITA would have had a reasonable expectation of success in doing so. *See supra* Section XIII.A; ¶¶285-296.



Ex. 1008 (Lizee) at Figs. 3, 4.

354. A POSITA would have found it obvious to zoom-in/out images to display a partial or full body of the user, similar to the effect depicted in Figure 9 of

Geisner, as applying an optical or digital zoom was a well-known image transformation technique. *See supra* §VI.B.f. Ex. 1007 (Geisner) at Fig. 9; *see also* Figs. 10-11. For example, Geisner discloses the modified user representation (modified image) may display a full body of the user as shown in Figure 9 below, for example.

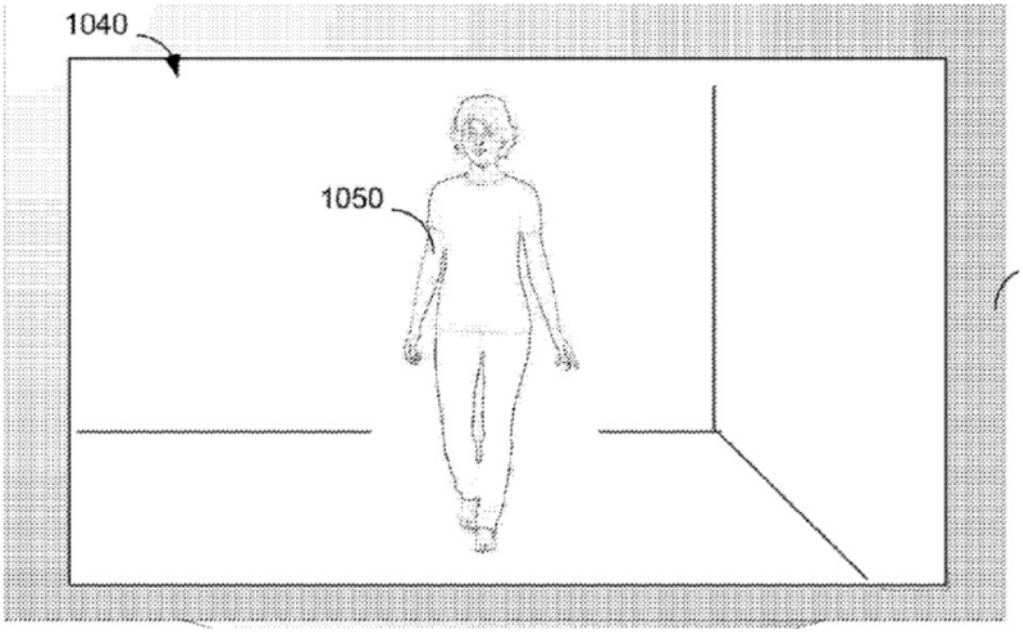


FIG. 9

Ex. 1007 (Geisner) at Fig. 9 (partial).

H. CLAIM 7: THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO GENERATE AN AVATAR HAVING BODY CHARACTERISTICS CORRESPONDING TO BODY CHARACTERISTICS OF A USER APPEARING IN THE IMAGES.

355. Geisner in view of Lizee, renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious this claim. For

example, Geisner discloses operating the processor (a computing system) to generate an avatar having body characteristics (user features) corresponding to body characteristics of a user appearing in the displayed images (displayed user features) as claimed. *See, e.g.*, Ex. 1007 (Geisner) at Fig 1B:

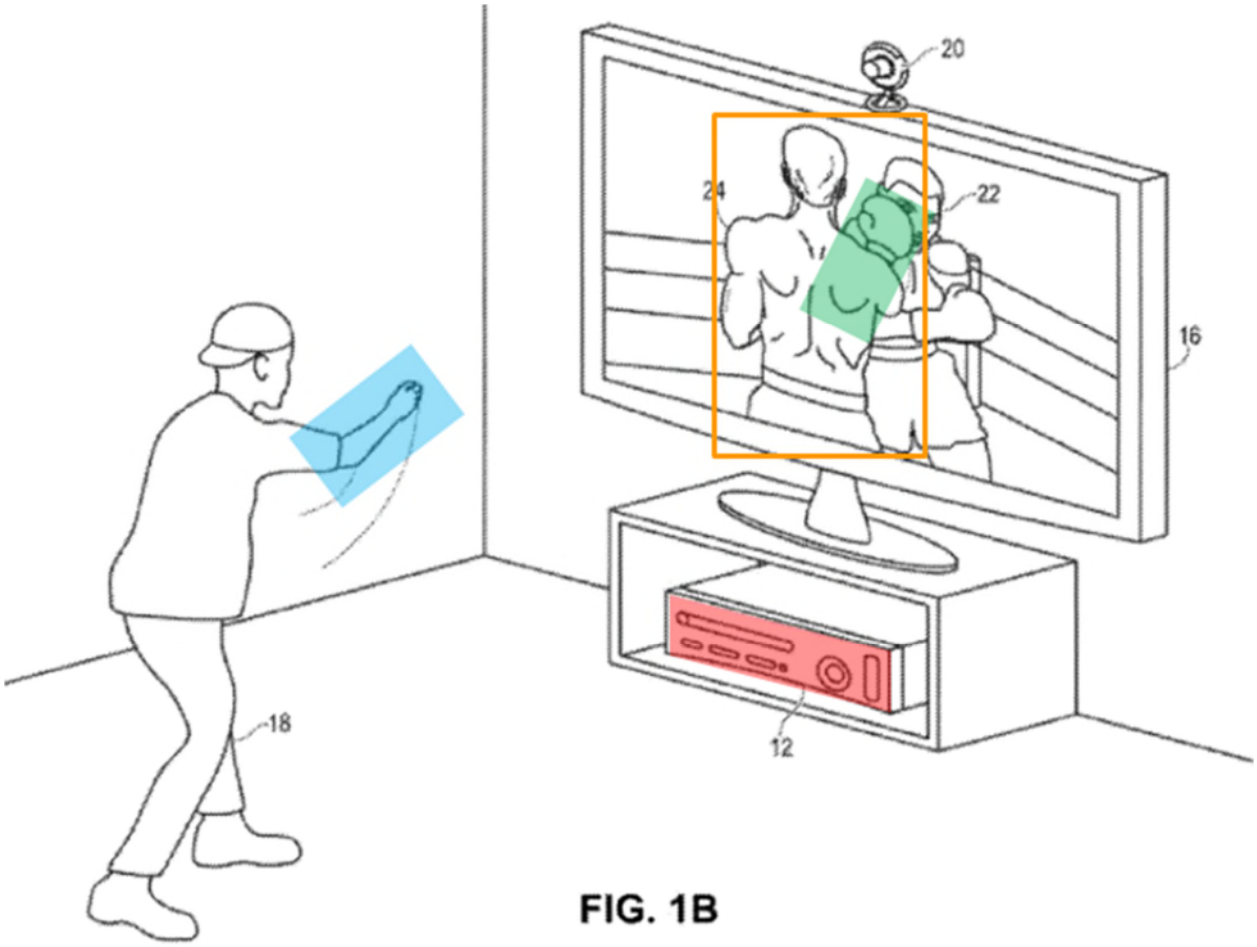


FIG. 1B

Ex. 1007 (Geisner) at Fig. 1B.

356. Geisner, for example, discloses a computing system used to generate a body model, which is "used to render an on-screen player avatar." Ex. 1007 (Geisner) at [0122]; *see also* Ex. 1007 (Geisner) at Fig.1-1B, [0028], [0030], [0072],

[0080]. Furthermore, in Geisner, "the system may detect features specific to the user, such as the user's facial features, body types, etc." and "select the features that resemble the detected features for application to the target's visual representation." Ex. 1007 (Geisner) at [0146]; *see also* Ex. 1007 (Geisner) at [0048],[0056].

I. CLAIM 8: THE METHOD OF CLAIM 7, FURTHER COMPRISING OPERATING THE PROCESSOR TO DISPLAY ON THE MONITOR THE MODIFIED IMAGES AND AN IMAGE OF THE AVATAR.

357. Geisner in view of Lizee renders obvious the method of claim 7 *See supra* §XIII.H. Additionally, Geisner in view of Lizee renders obvious this claim. For example, Geisner discloses displaying an image of the avatar. For example, Geisner discloses "**display[ing]** a representation of a human target user … such as in the form of **an avatar** or a pointer on a display, such as in a display device 193." Ex. 1007 (Geisner) at [0072]; *see also* Ex. 1007 (Geisner) at Figs.1-1B, [0122] ("Upon making such adjustments and applying such constraints, the gaming console may **display** **the adjusted player avatar**.").

J. CLAIM 9: THE METHOD OF CLAIM 7, FURTHER COMPRISING OPERATING THE PROCESSOR TO AUGMENT THE MODIFIED IMAGES BY OVERLAYING VIRTUAL OBJECTS ON THE AVATAR.

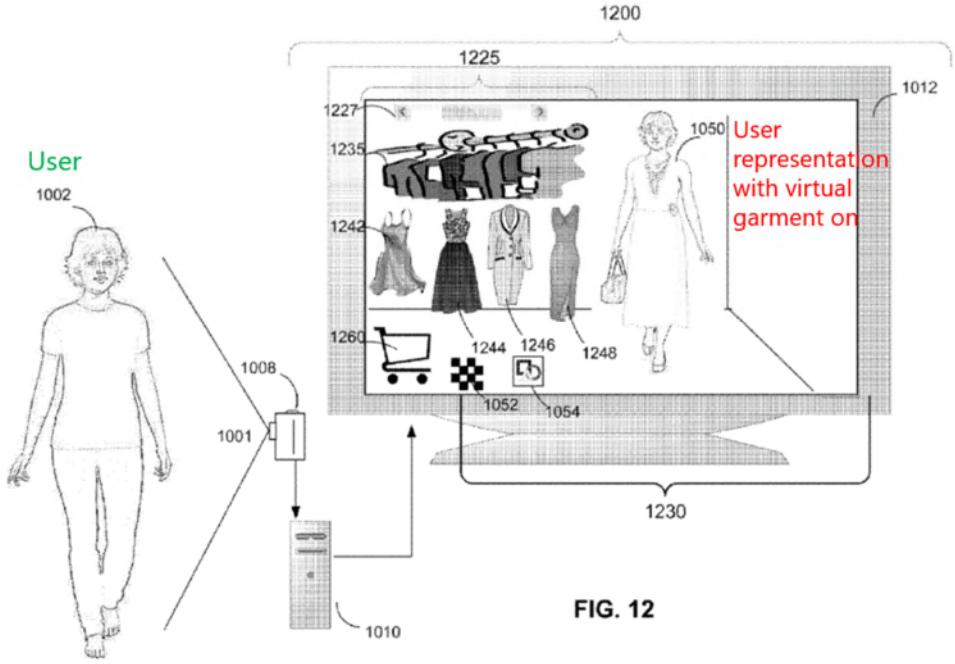
358. As noted above, Geisner in view of Lizee renders obvious the method of claim 7. *See supra* §XIII.H. Further, Geisner in view of Lizee renders obvious "operating the processor to augment the modified images by overlaying virtual objects on the avatar," as claimed. For example, Geisner discloses overlaying a

virtual garment on the avatar in the modified images. Ex. 1007 (Geisner) at [0154] ("A gesture applying an article to an on-screen image may comprise … placing the garment on top of the on-screen representation."); Ex. 1007 (Geisner) at [0060]-[0061],[0147],[0153],[0156]-[0157], Figs. 12, 17.

K. CLAIM 10. THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO AUGMENT THE MODIFIED IMAGES BY OVERLAYING VIRTUAL OBJECTS ON A USER'S IMAGE APPEARING WITHIN THE MODIFIED IMAGES.

359. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "comprising operating the processor to augment the modified images by overlaying virtual objects on a user's image appearing within the modified images."

360. As shown in the figure below, Geisner discloses overlaying a virtual object (e.g., a virtual garment) on a user's image appearing within the modified images (user representation). Ex. 1007 (Geisner) at [0154], [0156], Fig. 12.



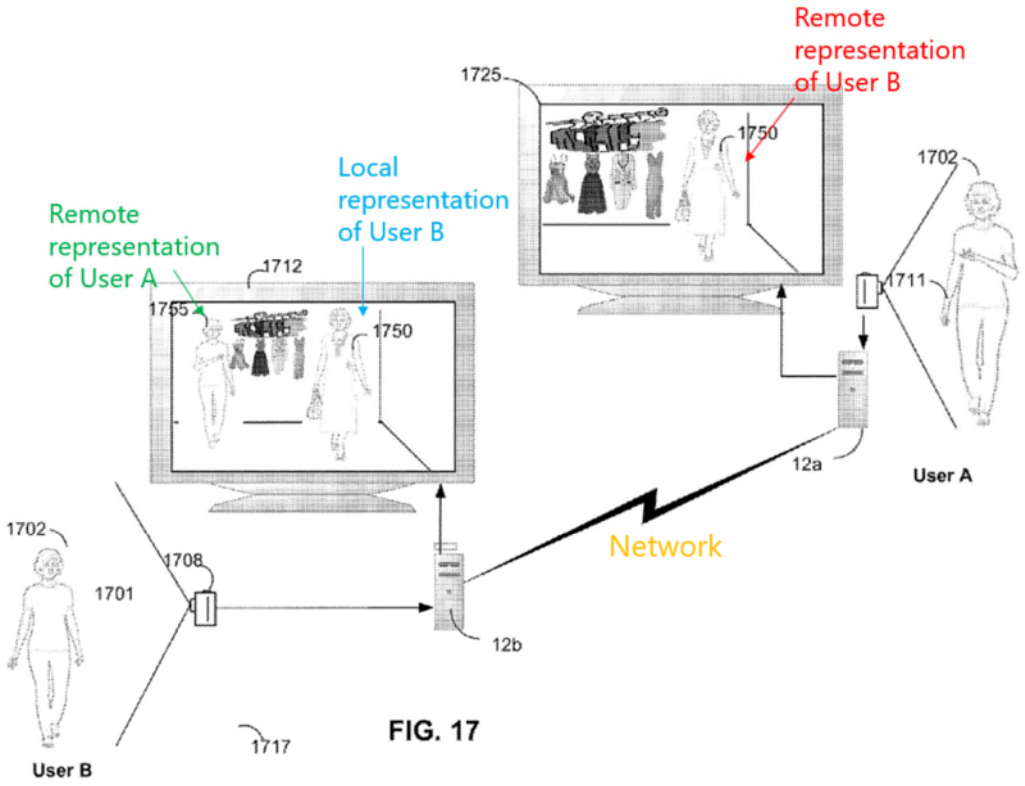
Ex. 1007 (Geisner) at Fig. 12.

361. Geisner discloses "placing the garment on top of the on-screen representation." Ex. 1007 (Geisner) at [0154]. Specifically, Geisner discloses that a user may use a gesture to instruct the system to grab a virtual garment and try that garment on. Ex. 1007 (Geisner) at [0154]; *see also* Ex. 1007 (Geisner) at [0171], Fig. 17.

L. CLAIM 11. THE METHOD OF CLAIM 1, FURTHER COMPRISING OPERATING THE PROCESSOR TO TRANSMIT THE MODIFIED IMAGE OVER A NETWORK.

362. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "operating the processor to transmit the modified image over a network."

363. Geisner discloses operating the processor to transmit the modified image (a user representation) over a network. Geisner discloses that "two users can communicate and collaborate via a networked environment in the shopping application 180." Ex. 1007 (Geisner) at [0171]. As shown in the figure below, the representation of user B may be transmitted over a network to the device at User A's location and is displayed as a remote representation of User B (1750). Similarly, the representation of User A is also transmitted and displayed as the representation 1755. Ex. 1007 (Geisner) at [0171], Fig. 17.



Ex. 1007 (Geisner) at Fig. 17.

M. CLAIM 12. THE METHOD OF CLAIM 1, FURTHER COMPRISING DISPLAYING THE MODIFIED IMAGE ON A REMOTE SCREEN FOR VIDEO TELEPRESENCE.

364. As explained above with respect to claim 11, Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "further comprising displaying the modified image on a remote screen for video telepresence."

365. As mentioned above, Geisner discloses that the user representation can be in the format of a video. *See supra* §XIII.B.b, ¶300. As shown in the annotated version of Figure 17 above (s*ee supra* §XIII.L), the user representation 1750 of user B is displayed on a remote screen 1725. Ex. 1007 (Geisner) at Fig. 17, [0171]. A POSITA would have understood that such mutual display configuration is one type of video telepresence because it captures and displays user movements (Ex. 1007 (Geisner) at [0151]) and allows communications and interactions between User A and User B (Ex. 1007 (Geisner) at [0171]).

N. CLAIM 13. THE METHOD OF CLAIM 1, FURTHER COMPRISING PROVIDING DATA FROM A DEPTH SENSOR TO THE PROCESSOR, AND OPERATING THE PROCESSOR TO MEASURE A TEXTURE OF THE IMAGE FROM THE DATA.

366. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Additionally, Geisner in view of Lizee renders obvious "providing data from a depth sensor to the processor, and operating the processor to measure a

texture of the image from the data," as claimed. For example, Geisner discloses "identifying surfaces, **textures**, and object dimensions from unorganized point clouds derived from a capture device, such as **a depth sensing device**." Ex. 1007 (Geisner) at [0133]; *see also* Ex. 1007 (Geisner) at [0035], [0116], [0132], [0134].

O. CLAIM 14. THE METHOD OF CLAIM 13, FURTHER COMPRISING OPERATING THE PROCESSOR TO GENERATE AN AVATAR USING THE DATA FROM THE DEPTH SENSOR.

367. Geisner in view of Lizee renders obvious the method of claim 13. *See supra* §XIII.N. Geisner in view of Lizee additionally renders obvious operating the processor to generate an avatar using the data from the depth sensor, as claimed, because Geisner teaches generating an avatar using the data from the depth sensor (i.e., depth images). Ex. 1007 (Geisner) at Fig. 6.

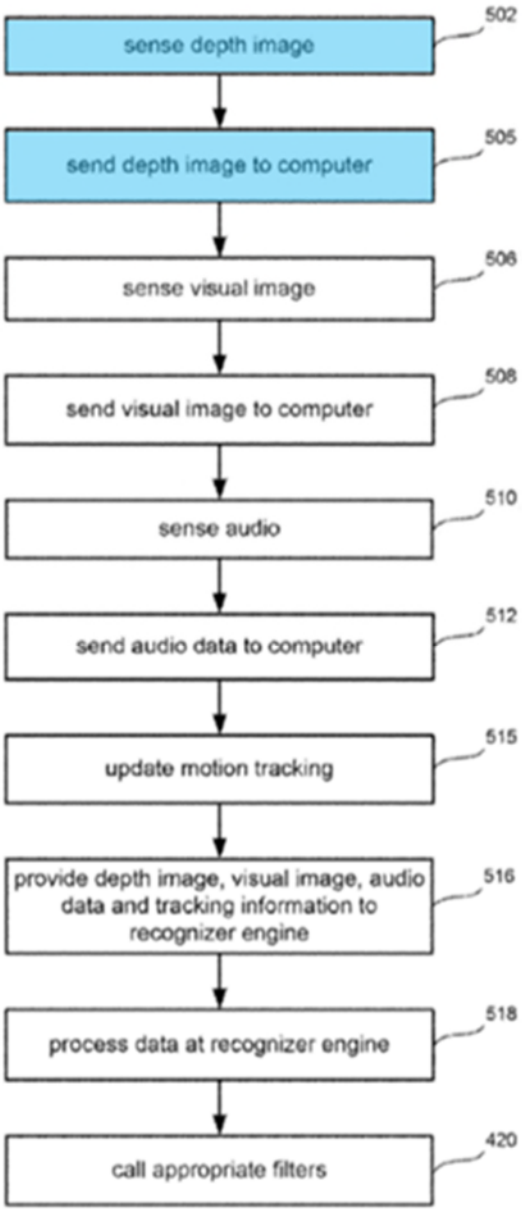


FIG. 6

Ex. 1007 (Geisner) at Fig. 6 (annotated).

368. Geisner discloses generating a skeleton using data from a depth sensor Ex. 1007 (Geisner) at [0102] ("[T]he system will track a skeleton of a person using a[sic] depth images. … Based on those steps, the system will fit a model with the

data and create a skeleton."); *see also*, Ex. 1007 (Geisner) at [0103], [0112]-[0113], [0116], [0120]-[0125]; *see supra* §XIII.B.f. The skeleton is then "used to render an on-screen player avatar." Ex. 1007 (Geisner) at [0122].

P. CLAIM 15. THE METHOD OF CLAIM 13, WHEREIN THE PROCESSOR GENERATE THE AVATAR BY ANALYZING REGISTRATION POINTERS ON THE IMAGES OF THE CAMERA.

369. Geisner in view of Lizee renders obvious the method of claim 13. *See supra* §XIII.N. Further, Geisner in view of Lizee renders obvious "generating the avatar via the processor by analyzing registration pointers on the images of the camera," as claimed. The '110 Patent discloses a "registration technique" that involves "register[ing] the orientation, location and size" of a user, wherein a "[r]egistration point may include, for example, midpoint between the eyes, the edges of the shoulders, the edges of the waist, the knees, etc." *See* Ex. 1001 ('110 Patent) at 14:4-13.

370. In Geisner, a processor uses modeling to generate a skeleton using data from a depth sensor, and this skeleton is used to generate an on-screen avatar. *See supra* §XIII.O, ¶368. The skeleton is generated by analyzing various spots (an example of the claimed registration pointers) including joints and other pointers. Ex. 1007 (Geisner) at [0054] ("skeletal mapping techniques to capture a person with a depth camera and from that **determine various spots** on that user's skeleton, joints of the hand, wrists, elbows, knees, nose, ankles, shoulders, and where the pelvis

meets the spine"); *see also* Ex. 1007 (Geisner) at [0102]("The skeleton will include a set of joints and connections between the joints."), [0103], [0122].

Q. CLAIM 16. THE METHOD OF CLAIM 1, WHEREIN THE PROCESSOR OPERATES ON THE IMAGES FROM THE CAMERA TO GENERATE BACKGROUND SEGMENTATION, AND REPLACES SEGMENTS ON THE IMAGES BY REPLACEMENT SEGMENTS.

371. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "wherein the processor operates on the images from the camera to generate background segmentation, and replaces segments on the images by replacement segments."

372. For example, Geisner discloses generating a background segmentation and replacing segments on the images with replacement segments. Geisner discloses "acquiring a depth image, . . . identifying and removing the background, and assigning each of the foreground pixels to different parts of the body." Ex. 1007 (Geisner) at [0102]. Geisner further incorporates by reference U.S. App. No. 12/603,437 (issued as U.S. Patent No. 8,295,546, Ex. 1012 (Craig)) which describes a detailed process for removing the background.

373. Geisner further discloses applying virtual segments to replace the identified background segments. For example, Geisner discloses that manipulation actions including "changing backgrounds, . . . moving room elements within a scene,

placing items within a scene, placing the on-screen representation with a variety of backgrounds in particular clothing, and the like." Ex. 1007 (Geisner) at [0111].

374. For example, as shown in Figure 12 of Geisner, the interface 1200 comprises a series of background icons 1052, 1054. Ex. 1007 (Geisner) at [0156]. Geisner describes that a user may further "select one or more of the background icons 1052, 1054 to provide a scene specific background behind the virtual representation 1050 of user 1002." Ex. 1007 (Geisner) at Fig. 12, [0158]; *see also* [0153]-[0157], [0159].

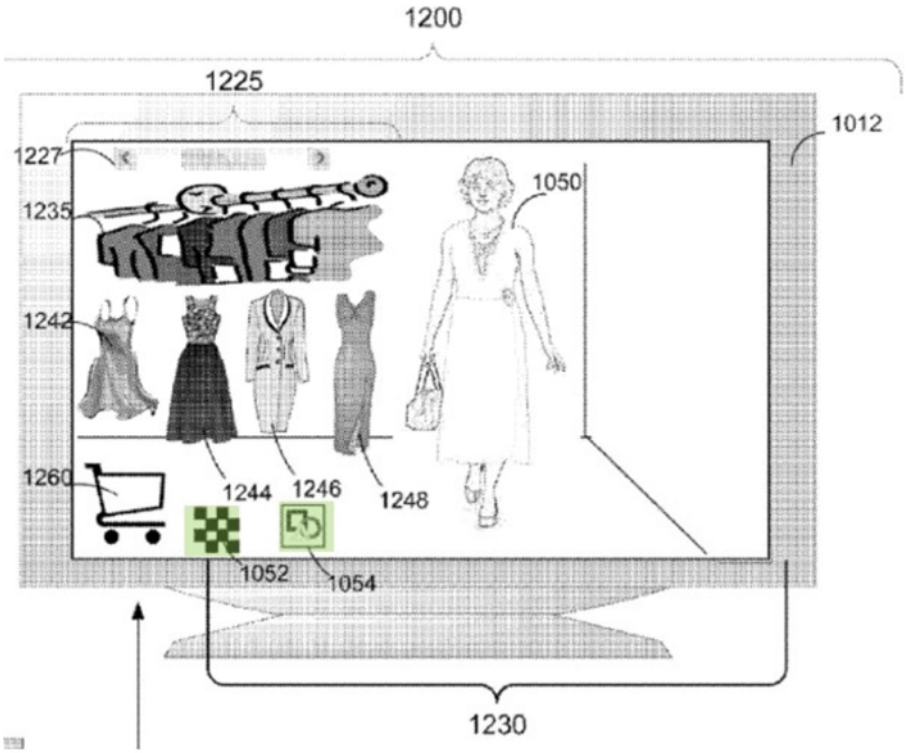
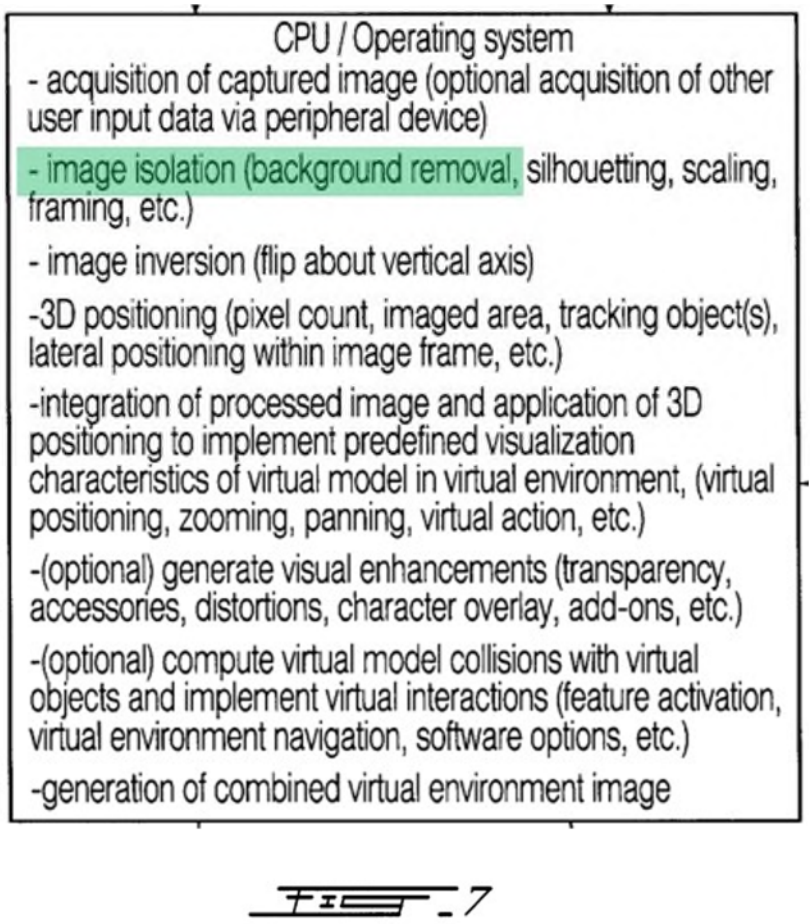


FIG. 12

Ex. 1007 (Geisner) at Fig. 12 (partial, annotated).

375. Additionally, Lizee teaches algorithms that can be used to generate background segmentation. As shown in the figure below, the processor performs the image isolation, which includes background removal. Ex. 1008 (Lizee) at Fig. 7, [0088].



Ex. 1008 (Lizee) at Fig. 7 (partial, annotated).

376. In particular, Lizee discloses a method of image isolation, including (1) capturing a background image, (2) "detecting color differences of the pixels respectively corresponding to the background image and the captured image," and

(3) separating the image of the user from the background by comparing the detected color difference in each pixel with a predefined threshold. Ex. 1008 (Lizee) at [0037]. Lizee further discloses that other algorithms may be used to achieve a similar result. Ex. 1008 (Lizee) at [0037].

377. As discussed, a POSITA would have recognized that both Geisner and Lizee disclose techniques for extracting user body and movement information and transforming the information into interactions with virtual objects in a virtual environment. *See supra* §XIII.A, ¶292. A POSITA would therefore would have been motivated to modify the teachings of Geisner with the teachings of Lizee to isolate the image and provide the capability for the user to replace the actual background with a desired virtual background. A POSITA would have had a reasonable expectation of success in doing so. *See supra* §XIII.A; ¶¶285-296.

R. CLAIM 17. THE METHOD OF CLAIM 16, WHEREIN THE REPLACEMENT SEGMENTS COMPRISE VIRTUAL BACKGROUND OR BACKGROUND THAT WAS CAPTURED IN A DIFFERENT SETTING.

378. Geisner in view of Lizee renders obvious the method of claim 16. *See supra* §XIII.Q. Further, Geisner in view of Lizee renders obvious "wherein the replacement segments comprise virtual background or background that was captured in a different setting."

379. Geisner teaches that a user may change the background of the virtual scene by selecting one or more background icons. Ex. 1007 (Geisner) at [0158]; *see*

also Ex. 1007 (Geisner) at [0153]-[0159]. Geisner describes that the "interface 1200 may include a control or selection portion 1225 and a rendered scene 1230," which is shown in the figure below. Ex. 1007 (Geisner) at [0156], Figure 12. The figure below shows a series of background icons 1052, 1054. A user may select one of the background icons, causing the system to replace the background with the selected background. Ex.1007 (Geisner) at [0158]; *see also* [0153]-[0159].

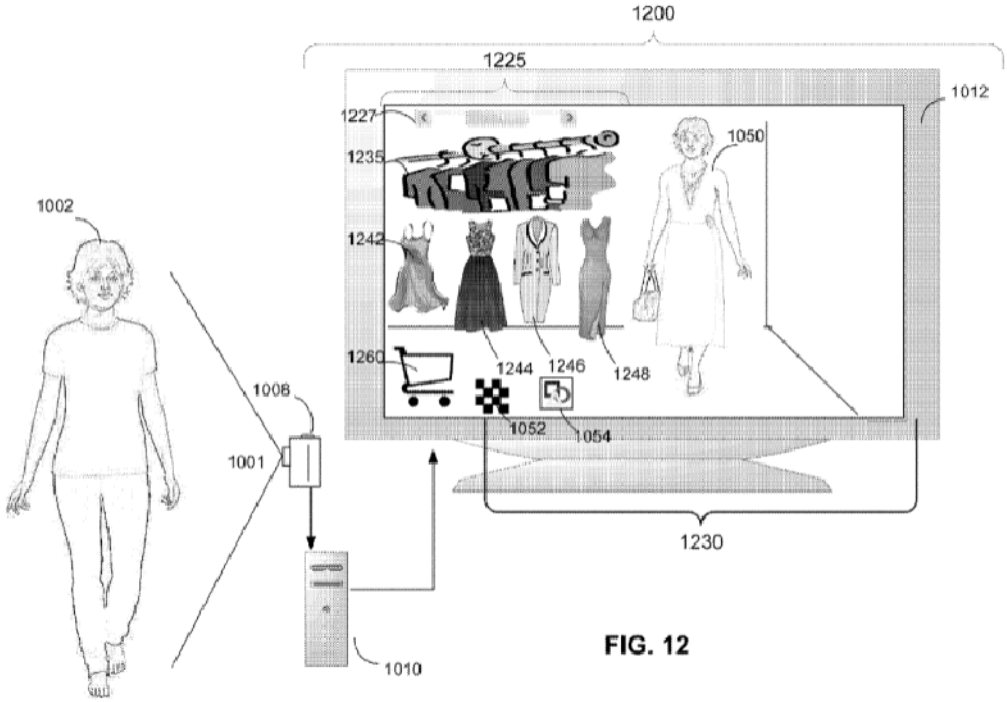


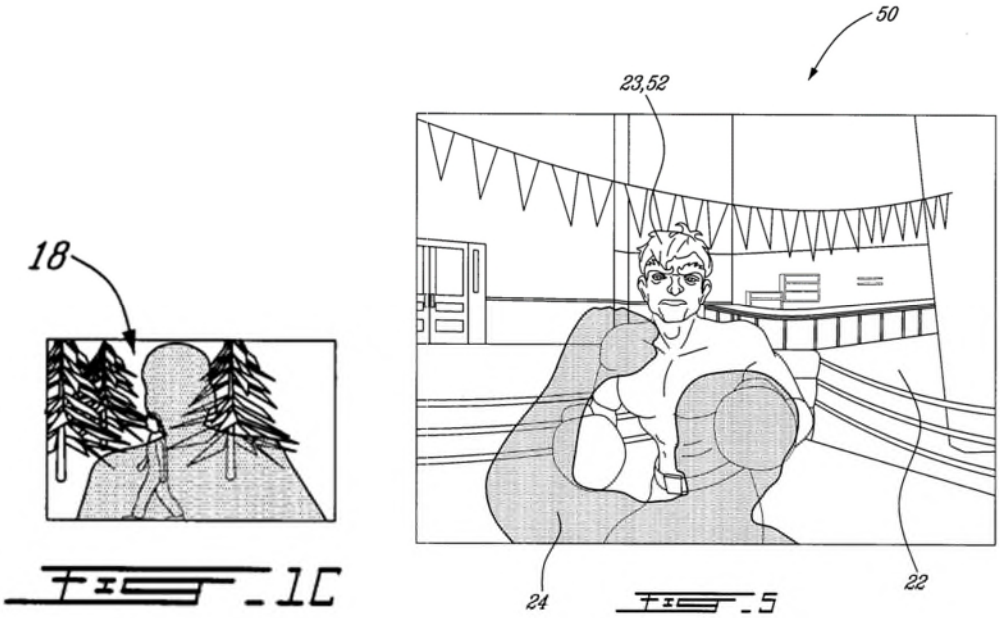
FIG. 12

Ex. 1007 (Geisner) at Fig. 12.

380. The control selection portion and rendered scene described in Geisner are virtual backgrounds. Further, Geisner discloses that a user may change

backgrounds by selecting background icons 1052 or 1054. Ex. 1007 (Geisner) at [0158]; *see also* [0153]-[0159].

381. Additionally, Lizee discloses replacing background segments with replacement segments comprising virtual backgrounds. For example, Lizee discloses "a virtual environment 22 comprising a boxing arena." Ex. 1008 (Lizee) at Fig. 5, [0079].



Ex. 1008 (Lizee) at Fig. 1C (left), Fig. 5 (right).

382. As shown in the figures above, a background is replaced with segments of a virtual background. Ex. 1008 (Lizee) at [0029], Figs. 1C, 5.

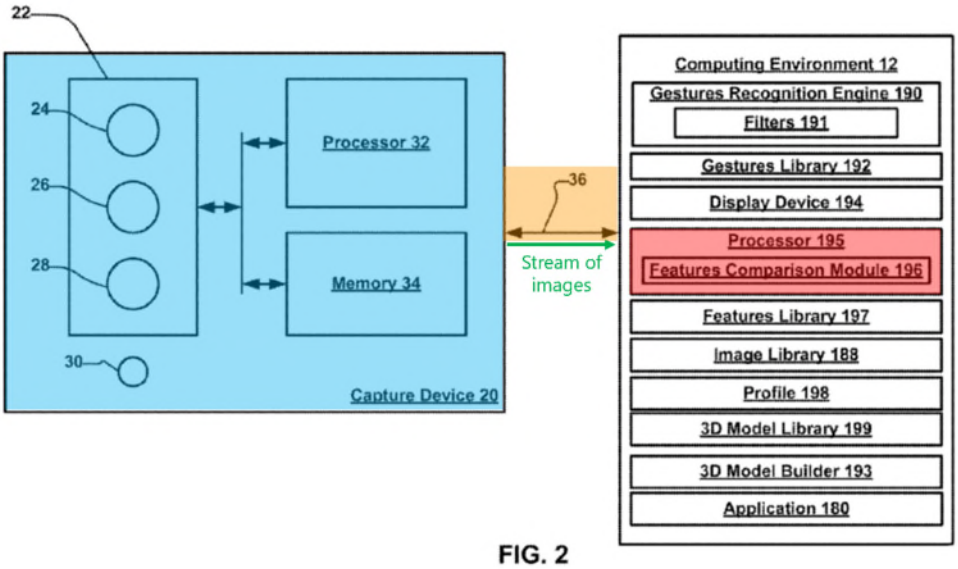
383. Given that both Geisner and Lizee disclose techniques for extracting user body and movement information and transforming the information into interactions with virtual objects in a virtual environment (*see supra* §XIII.A, ¶292),

a POSITA would have been motivated to combine the virtual backgrounds in Lizee with the teachings of Geisner to allow for the user to be more fully immersed in a virtual environment. A POSITA would have had a reasonable expectation of success in doing so. *See supra* §XIII.A; ¶¶285-296. Thus, Geisner in view of Lizee renders obvious "wherein the replacement segments comprise virtual background or background that was captured in a different setting."

S. CLAIM 18: THE METHOD OF CLAIM 1, WHEREIN THE PROCESSOR RESIDES IN THE REMOTE LOCATION, AND THE METHOD PROCEEDS BY TRANSMITTING THE IMAGES FROM THE CAMERA OVER A NETWORK TO THE PROCESSOR, AND TRANSMITTING THE MODIFIED IMAGES BACK FROM THE PROCESSOR TO THE MONITOR OVER THE NETWORK.

384. Geisner in view of Lizee renders obvious the method of claim 1. *See supra* §XIII.B. Further, Geisner in view of Lizee renders obvious "wherein the processor resides in the remote location, and the method proceeds by transmitting the images from the camera over a network to the processor, and transmitting the modified images back from the processor to the monitor over the network."

385. Geisner discloses that images are captured at the Capture Device and transmitted via a communication link to a processor in Computing Environment for processing. Ex. 1007 (Geisner) at Fig. 2.



Ex. 1007 (Geisner) at Fig. 2 (annotated).

386. Additionally, Geisner teaches that the computing environment (and the processor located therein) may reside in a remote location. For example, Geisner discloses that "[i]n a networked environment, program modules depicted relative to the computer 241 … may be stored in the remote memory storage device." Ex. 1007 (Geisner) at [0100]; *see also* [0099]. A POSITA would have found it obvious that the processor resides in the remote location when the program modules are stored remotely and the computing environment is connected via an Ethernet cable or wireless connection.

387. Further, Geisner teaches that images are processed in a computing environment connected via a communication link, which may comprise "an Ethernet cable connection, or the like and/or a wireless connection" through which the images are transmitted to Computing Environment connected to a LAN or WAN. Ex. 1007

(Geisner) at [0044]; *see also* Fig. 2. Geisner further discloses that "[t]he computer 241 may operate in a networked environment using logical connections to one or more remote computers, such as a remote computer 246" and that "[t]he remote computer 246 may be a personal computer, a server, a router, a network PC, a peer device or other common network node, and typically includes many or all of the elements described above relative to the computer 241." Ex. 1007 (Geisner) at [0099], Fig. 4. POSITA would have found it obvious that, for reducing the system requirements of local devices and increasing efficiency, the processor resides in a remote computing environment when the program modules are stored remotely and when the computing environment is connected via an Ethernet cable or wireless connection. *See* Ex. 1007 (Geisner) at [0044].

388. Therefore, Geisner discloses that the modified images may be transmitted back to a local display. Further, the display device disclosed in Geisner may be separate from the computing environment. Ex. 1007 (Geisner) at [0077]. Geisner additionally discloses "transmit[ing] the raw data for processing by another computing environment." Ex. 1007 (Geisner) at [0055]. A POSITA therefore would have found it obvious to position a display and camera(s) in the vicinity of a user (to allow the user to view himself/herself), and would understand that the processor could be located remotely. A POSITA would have found it obvious to transmit the modified images from the computing environment to the user so that the user would

see the modified images on the screen. A POSITA would further understand that in the case in which the processor is located remotely from the user (who is in proximity to the display and camera), it would be necessary to: 1) transmit images from the camera to the processor via a network, and 2) transmit the modified images from the processor back to the display via the network.

XIV. GROUND 4: CLAIMS 7-9, 13-15, AND 18 ARE OBVIOUS OVER HAAN IN VIEW OF FRANCOIS AND GEISNER

389. Haan in view of Geisner and Francois renders obvious claims 7-9, 13-15, and 18. A POSITA would have found it obvious to combine Haan, Francois, and Geisner. As explained above, it would have been obvious to combine Haan and Francois. *See supra* §XI.A. A POSITA would have also found it obvious to further modify Haan and Francois in view of Geisner. First, Haan, Francois, and Geisner are analogous to the ’110 Patent and are in the same field of endeavor, as discussed above with respect to grounds 1 and 3. *See supra* §XI.A, ¶138; *supra* §XIII.A, ¶287.

390. Second, a POSITA would have been motivated to combine Haan and Geisner to enhance the user’s virtual mirror experience adding more comprehensive augmented reality features. For example, Haan describes implementing augmented reality features to the mirror device. Ex. 1005 (Haan) at [0113]. Geisner, however, discloses more comprehensive augmented reality features that would be compatible with Haan’s systems. Ex. 1007 (Geisner) at [0175].

391. A POSITA would have found it obvious to look to Geisner to provide more comprehensive augmented reality features, like changing the user's features including body parts, and providing an avatar, to build upon the virtual mirror system in Haan and offer an enhanced user experience. For example, a user may want to share images of new garments on a remote social platform, but may prefer to use an avatar rather than captured images, either for privacy concerns or to avoid sharing user images that are not ready for sharing. A POSITA would have also been motivated to combine Haan and Geisner to include Geisner's depth sensor with the virtual mirror sensor disclosed in Haan to improve the accuracy of the avatar and/or to enhance a user's augmented reality experience. *See, e.g.,* Ex. 1007 (Geisner) at [0132]-[0134], [0035], [0116].

392. Further, a POSITA would have also been motivated to modify Haan's virtual mirror system with the remote communication/telepresence functionality disclosed in Francois (Ex. 1009 (Francois) at §§4.4-4.5, Fig. 15) and the interactive gesture-identification disclosed in Geisner (Ex. 1007 (Geisner) at Abstract, [0001], [0026]). A POSITA would have found it desirable to provide a user of Haan's virtual mirror the ability to share modified user images remotely, via telepresence. *See supra* §XI.A. Further, as Haan describes, a "frequently occurring problem with mirrors is that the area one wants to see is just out of sight." *See* Ex. 1005 (Haan) at [0004]. A POSITA would have found it obvious to implement the gesture

identification disclosed in Geisner (Ex. 1007 (Geisner) at [0051], [0127]) to allow a user to interact with the Haan-Francois virtual mirror system, even when a user is turned away from the display of the system, by using, for example, certain hand gestures.

393. Combining Geisner's process with Haan and Francois would involve no more than using a known technique to improve a similar device in the same way to achieve predictable results. For example, Haan discloses a virtual mirror system using known computers and mobile devices, and Francois discloses "a universal framework" for "easy integration" into such known systems, like the virtual mirror system in Haan. *See supra* §XI.A. Geisner discloses single or multiple processors executing software instructions stored in storage media that implement improved augmented reality features. *See, e.g.*, Geisner (Ex. 1007), [0027], [0042]-[0043], [0045]-[0047], [0055]-[0056], [0077]-[0079], [0085]-[0086], [0097]-[0098]. And a POSITA would have understood the Haan-Francois system (and consequent disclosure of standard processors for implementing a virtual mirror system) would be suitable for performing Gesiner's processes. *See supra* §§XI.A, XI.B.c. A POSITA would have expected success in implementing Geisner's processes in Haan (as modified by Francois) because this would have involved using a known technique to improve a similar device in a similar way to yield the predictable result of a more versatile digital mirror system with an enhanced user experience.

394. As explained above, Haan in view of Francois renders obvious claim 1. *See supra* §XI.B. The combination of Haan, Francois, and Geisner teaches claims 7-9 and 13-15. As explained, Geisner discloses the additional limitations in claims 7-9 and 13-15. *See supra* §§XIII.H-XIII.J, XIII.N-XIII.P. A POSITA would have been motivated to pursue the combination and would have had a reasonable expectation of success in doing so, as I just discussed. *See supra* §XIV, ¶394.

395. Further, Haan in view of Francois and Geisner renders obvious claim 18. As discussed above, Geisner discloses the additional limitations of claim 18. *See supra* §XIII.S. A POSITA would have found been motivated to modify Haan (as also modified by Francois) with Geisner's remote processor to further reduce the processing capacity requirements and thus cost for a local device at the user's site, thereby improving efficiency and accessibility of the system. *See supra* §XIII.S. As I just discussed above, a POSITA also would have done so with a reasonable expectation of success. *See supra* §XIV, ¶394.

396. Therefore, Haan in view of Francois and Geisner renders claims 7-9, 13-15, and 18 obvious.

XV. GROUND 5: CLAIMS 3-4, 7-9, AND 11-18 ARE OBVIOUS OVER ROSENBERG IN VIEW OF PARK AND GEISNER

397. Rosenberg in view of Park and Geisner renders obvious claims 3-4, 7-9, and 11-18. A POSITA would have found it obvious to combine the teachings of

Geisner with Rosenberg and Park. As described above, it would have been obvious to combine Rosenberg and Park. *See supra* §XII.A. A POSITA would have also found obvious to add Geisner to this combination.

398. Geisner is analogous to Rosenberg, Park, and the ’110 Patent because they are in the same field of endeavor. *See supra* §X.E.b, §XII.A; §XIII.A, ¶¶287, 289. A POSITA would have been motivated to modify Rosenberg (as modified by Park) with Geisner’s gesture identification (*see supra* §§XIII.D, XIII.E) to detect the presence of the user and allow the user to have more control of the mirror operation. A POSITA would have been motivated to modify Rosenberg (as modified by Park) with Geisner’s remote image transfer (*see supra* §§XIII.L-XIII.M) to improve the user experience and versatility of Rosenberg’s system by allowing the user to share how they look with another user to assist the user in choosing their look (e.g., sharing image of the user with “particular hairstyles, clothing choices, and/or makeup work”). Rosenberg (Ex. 1006), [0005], [0053], [0108]). A POSITA would have been motivated to modify Rosenberg (as modified by Park) with Geisner’s avatar generation and display (*see supra* §§XIII.H-XIII.J, XIII.N-XIII.P) to provide the virtual mirror in Rosenberg with an enhanced user experience by having an avatar of the user. For example, a user may want to share images of selected new garments with another user viewing a remote screen (e.g., sharing on a remote social platform) (e.g., for creative expression or entertainment), but may prefer to use an avatar rather

than captured images, either for privacy concerns or to avoid sharing user images that are not ready for sharing. A POSITA also would have been motivated to modify Rosenberg (as modified by Park) with Geisner's remote processor (*see supra* §XIII.S) to reduce the processing capacity requirements and thus cost for the local device at the user's site, thereby improving efficiency and accessibility of the system. A POSITA also would have been motivated to modify Rosenberg (as modified by Park) with Geisner's virtual background (*see supra* §§XIII.Q-XIII.R) to improve the user experience and versatility of Rosenberg's system by allowing the user to, for example, view how a hairstyle, clothing choice, and/or makeup would look in a particular background environment. Geisner (Ex. 1007), [0158].

399. A POSITA would have found it obvious to add Geisner to the Rosenberg-Park combination, as these references each concern improving a display system that displays user images captured by a camera. Both teachings also concern fitting scenarios (*see* Ex. 1006 (Rosenberg) at [0053], [0111], Ex. 1007 (Geisner) at [0154], [0156]), and providing users with options to view their images and to send the images over a communication network. Further, Geisner discloses that a single or multiple processors executing software instructions stored in storage media can perform the processes for implementing these features. Geisner (Ex. 1007), [0027], [0042]-[0043], [0045]-[0047], [0055]-[0056], [0077]-[0079], [0085]-[0086], [0097]-[0098]. A POSITA would have understood that Rosenberg's programmable

processor, memory, and other hardware or software components would be suitable for performing Geisner's processes. Rosenberg (Ex. 1006), [0035], [0068], [0074], [0109], Fig. 13. A POSITA would have expected success in implementing Geisner's processes in Rosenberg (as modified by Park) because this would have involved using a known technique to improve a similar device in the same way and yield the predictable result of a more versatile digital mirror system with enhanced user experience.

A. CLAIMS 3-4

400. Rosenberg in view of Park and Geisner also renders claims 3 and 4 obvious. First, as described above, Rosenberg in view of Park renders claim 1 obvious. *See supra* §XII.B.

401. Further, Rosenberg (as modified by Park) in view of Geisner renders obvious "wherein during idle mode the camera continuously sends captured images to the processor, and the user presence is identified by the processor detecting a body part in the images," as recited in claim 3. *See* Ex. 1001 ('110 Patent) at claim 3. For example, as described previously, Geisner discloses a modification mode that can be triggered by a user gesture, wherein a trigger gesture may interrupt "the current state of the application" and put the system into a "modification mode [that] may cause the application to pause, where the application can be resumed" (i.e., an idle mode) upon exiting the modification mode, wherein a user presence is continuously

identified during the idle mode by detecting a body part. Ex. 1007 (Geisner) at [0052]; *see supra* §§XIII.D, XIII.E.

402. Further, as previously discussed, Rosenberg teaches an idle mode (a low-power idle state) with an auto-activation feature, in which the proximity and/or motion sensors continuously operate to determine whether or not a person is located before the display screen. *See* Ex. 1006 (Rosenberg) at [0032]; *see supra* §XII.D, ¶279. In a low-power idle mode, while the system may turn off display, the sensor would continuously send data to a processor for identifying the presence of a user. *See* Ex. 1006 (Rosenberg) at [0032], *see supra* §XII.D, ¶279. Rosenberg discloses that the proximity and motion detection may be accomplished using a dedicated sensor (e.g., infra-red sensor, camera). Ex. 1006 (Rosenberg) at [0069], [0073], [0109], Fig. 13. As earlier described, a POSITA would have understood that for those embodiments using the camera, the data sent to the processor are captured real-time video images and would be continuously sent to the processor to detect the user presence. *See supra* §XII.D, ¶279.

403. Additionally, a POSITA would have been aware of various known and suitable image processing methods for detecting a user's presence using the digital camera and processor disclosed in Rosenberg. *See supra* §XII.D, ¶280. For example, as discussed, a POSITA would be motivated to seek known and more specific

implementing details on the use of a digital camera to detect a user's presence, as disclosed in Rosenberg. *See supra* §XII.D, ¶280

404. A POSITA would have understood that Geisner provides a suitable known method for detecting a user's presence in Rosenberg's camera-captured images. Geisner discloses gesture identification that requires identifying a user body part and discloses that the body part comprises one of legs, face, eyes, and hands. *See supra* §XIII.E. A POSITA would have found it obvious to modify Rosenberg's idle mode with Geisner's gesture identification method (i.e., identifying a user's hand gesture) for identifying user presence. *See supra* §XIII.E. A POSITA would have understood allowing the user to use a gesture to engage or disengage the mirror would allow the user to have more active control of the operation of the mirror. A POSITA would have understood that using a gesture would be a suitable alternative to Rosenberg's teaching of the user commanding the mirror to turn on and off using a verbal command. *See* Rosenberg (Ex. 1006), [0034], [0068], [0073], [0115]. A POSITA would have had a reasonable expectation of success in doing so. *See* ¶ 399. For example, a POSITA would understand that implementing the identification of a user's body part (i.e., at least one of a user's legs, face, eyes, and hands), as disclosed in Geisner, into the Rosenberg-Park combination, would create a more comprehensive and interactive idle mode by allowing, for example, a user's gesture based on movement of a specific body part to enter or exit the idle mode.

405. Accordingly, Rosenberg in view of Park and Geisner renders these claims obvious.

B. CLAIMS 7-9 AND 13-15

406. Further, Rosenberg in view of Park and Geisner renders claims 7-9 and 13-15 obvious. First, as described above, Rosenberg in view of Park renders claim 1 obvious. *See supra* §XII.B.

407. Additionally, as explained above, Geisner discloses "operating the processor to generate an avatar having body characteristics corresponding to body characteristics of a user appearing in the images," as recited in claim 7. *See* Ex. 1001 ('110 Patent) at claim 7; *see supra* §XIII.H. Geisner further teaches "operating the processor to display on the monitor the modified images and an image of the avatar," as recited in claim 8. *See* Ex. 1001 ('110 Patent) at claim 8; *see supra* §XIII.I. Additionally, Geisner teaches "operating the processor to augment the modified images by overlaying virtual objects on the avatar," as recited in claim 9. *See* Ex. 1001 ('110 Patent) at claim 9; *see supra* §XIII.J.

408. Geisner further discloses "providing data from a depth sensor to the processor, and operating the processor to measure a texture of the image from the data" and "operating the processor to generate an avatar using the data from the depth sensor," as recited in claims 13 and 14. *See* Ex. 1001 ('110 Patent) at claims 13, 14; *see supra* §§XIII.N-XIII.O. Further, Geisner teaches "wherein the processor

generate the avatar by analyzing registration pointers on the images of the camera," as recited in claim 15. *See* Ex. 1001 ('110 Patent) at claim 15; *see supra* §XIII.P.

409. A POSITA would have found it obvious to combine Geisner with the Rosenberg-Park combination by adding avatar related features and a depth sensor for generating the avatar to enhance the user experience. *See supra* ¶¶398-399. A POSITA would have had a reasonable expectation of success in doing so. *See* ¶ 399. Therefore, Rosenberg in view of Park and Geisner would render claims 7-9 and 13-15 obvious.

410. Accordingly, Rosenberg in view of Park and Geisner renders these claims obvious.

C. CLAIMS 11-12, 18

411. Further, Rosenberg in view of Park and Geisner renders claims 11-12 and 18 obvious. First, as described above, Rosenberg in view of Park renders claim 1 obvious. *See supra* §XII.B.

412. Rosenberg (as modified by Park) in view of Geisner teaches "operating the processor to transmit the modified image over a network" and "displaying the modified image on a remote screen for video telepresence," as recited in claims 11 and 12. *See* Ex. 1001 ('110 Patent) at claims 11, 12; *see supra* §§XIII.L-XIII.M. As discussed, Geisner discloses operating the processor to transmit a modified image of a user (a user representation) over a network (a networked environment in a shopping

application). Ex. 1007 (Geisner) at [0171], Fig. 17; *see supra* §XIII.L. Further, Geisner discloses displaying the modified user image on a remote screen. Ex. 1007 (Geisner) at [0171], Fig. 17; *see supra* §XIII.M. Additionally, Rosenberg, for example, discloses that "the processor within the advanced digital mirror may have communication access to one or more communication networks via a network connection 306." Ex. 1006 (Rosenberg) at [0074]. Rosenberg further discloses that the "communication device 1330 is utilized to receive and transmit images across a network." Ex. 1006 (Rosenberg) at [0109], Fig. 13. Park also discloses a transmission unit for transmitting compensated images to another user over a network for video telephony. Ex. 1011 (Park) at [0009]-[0010], [0013], [0037], [0063], [0082], [0092], Fig. 2. As discussed in Ground 3, a POSITA would have understood mutual display configurations, such as in video telepresence, would capture and display user movements and allow communications/interactions between two users. *See supra* §XIII.M; see Ex. 1007 (Geisner), Fig. 17, [0171]; Ex. 1011 (Park) at [0009]-[0010]. A POSITA would have found it obvious to use Rosenberg's system to display the modified images on a screen of another user at a remote location for video telepresence.

413. Additionally, Geisner discloses "wherein the processor resides in the remote location, and the method proceeds by transmitting the images from the camera over a network to the processor, and transmitting the modified images back

from the processor to the monitor over the network," as disclosed in claim 18. *See* Ex. 1001 ('110 Patent) at claim 18; *see supra* §XIII.S. Further, as discussed earlier, Geisner discloses transferring the modified images over the network and displaying the modified images on a remote screen for video telepresence. *See supra* §XIII.L. As discussed above, a POSITA would have been motivated to modify Rosenberg (as modified by Park) with Geisner's remote image transfer to improve the user experience and versatility of Rosenberg's system by allowing the user to share looks with another user. *See* ¶¶398-399. A POSITA would have expected success in applying the teachings from Geisner to transmit modified images over a network using the communication device in Rosenberg and to display the transmitted images on a remote screen because such application requires nothing beyond ordinary skill. *See* ¶ 399.

414. Further, Rosenberg in view of Park and Geisner renders claim 18 obvious. A POSITA would have found it obvious to combine the Rosenberg-Park combination with Geisner's networked computing environment feature to reduce the processing requirements for the local device at the user's end and improve the system's efficiency and accessibility, rendering claim 18 obvious. *See supra* ¶¶398-399.

D. CLAIMS 16-17

415. Rosenberg in view of Park and Geisner also renders obvious claims 16 and 17. First, as described above, Rosenberg in view of Park renders claim 1 obvious. *See supra* §XII.B. Geisner also discloses that "the processor operates on the images from the camera to generate background segmentation, and replaces segments on the images by replacement segments" and that "the replacement segments comprise virtual background or background that was captured in a different setting," as recited in claims 16 and 17. *See* Ex. 1001 ('110 Patent) at claims 16, 17; *see supra* §§XIII.Q-XIII.R.

416. A POSITA would have found it obvious to apply Geisner's background replacement features to the Rosenberg and Park combination to improve the user experience and versatility of Rosenberg's system, by enabling the user to, for example, view how a garment would look in a different background environment. Geisner (Ex. 1007), [0158]. A POSITA would have had a reasonable expectation of success in doing so. *See* ¶ 399. Therefore, Rosenberg in view of Park and Geisner renders claims 16 and 17 obvious. *See supra* ¶¶398-399.

XVI. SECONDARY CONSIDERATIONS

417. During prosecution of the Challenged Patent, the applicants did not discuss any secondary considerations of non-obviousness, and I am not aware of any secondary considerations that would support non-obviousness. Should the Patent

Owner put forth any alleged evidence of secondary considerations in connection with the Challenged Patent, I reserve the right to consider such evidence, to consider whether any such evidence has a nexus to any claims of the Challenged Patent, and to update my analysis.

XVII. CONCLUSION

418. For the reasons set forth above, I believe claims 1-18 of the ’110 patent would have been obvious in view of the prior art.

419. In signing this Declaration, I understand it will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I understand that I may be subject to cross-examination in this case, and that any cross-examination will take place in the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

420. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that those statements were made with the knowledge that willful false statements and the like so made are punishable by fine, imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on May 10, 2024.

Dr. Lina Karam, Ph.D.