



US 20020007718A1

(19) **United States**

(12) **Patent Application Publication**
Corset

(10) **Pub. No.: US 2002/0007718 A1**

(43) **Pub. Date: Jan. 24, 2002**

(54) **KARAOKE SYSTEM**

Publication Classification

(76) Inventor: **Isabelle Corset, San Mateo, CA (US)**

(51) **Int. Cl.⁷ A63H 5/00**

(52) **U.S. Cl. 84/609**

Correspondence Address:

Corporate Patent Counsel

U.S. Philips Corporation

580 White Plains Road

Tarrytown, NY 10591 (US)

(21) Appl. No.: **09/883,442**

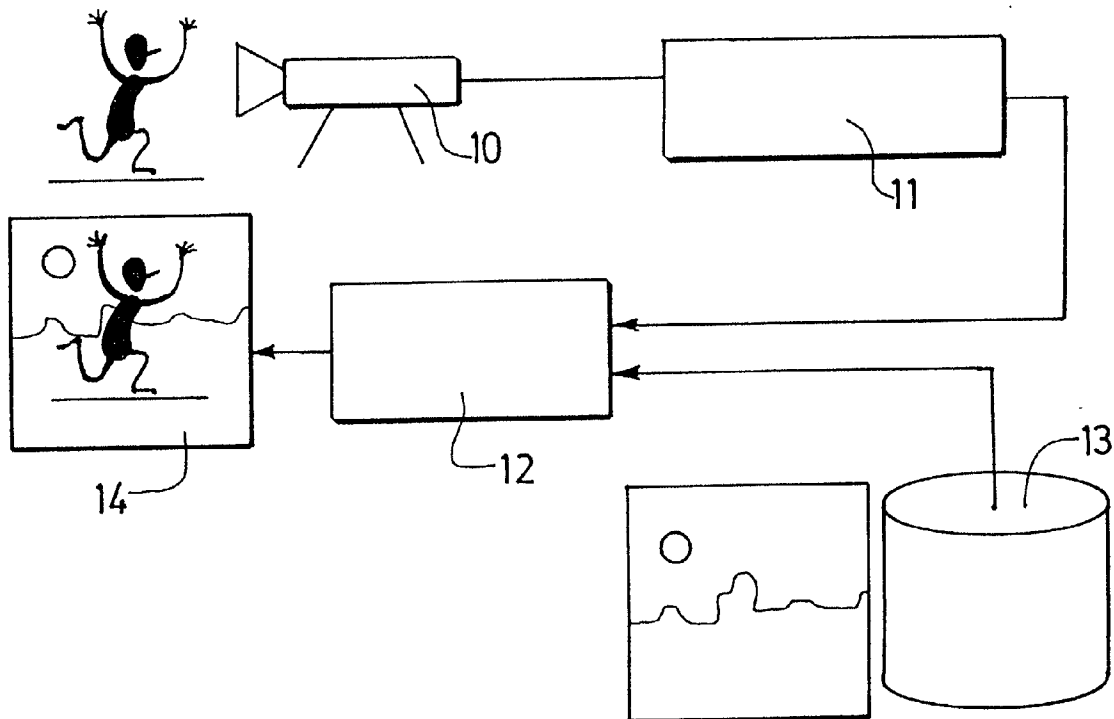
(22) Filed: **Jun. 18, 2001**

(30) **Foreign Application Priority Data**

Jun. 20, 2000 (EP)..... 00401758.8

(57) **ABSTRACT**

The concept of singing karaoke exclusively relies on audio based techniques. The invention consists of developing the concept of video immersion: the user will see his (her) image inserted into the video clip or movie at the place of his (her) favorite dancer, singer or player, and will therefore be able to play clip/song on video tape, and replace any famous star. More precisely, the invention relates to a karaoke system in which are provided successive means for picking up an image of the user and his/her voice, analyzing and processing the obtained signals, mixing the audio/video signals thus analyzed and processed and a pre-recorded material and displaying the combination signals thus obtained.



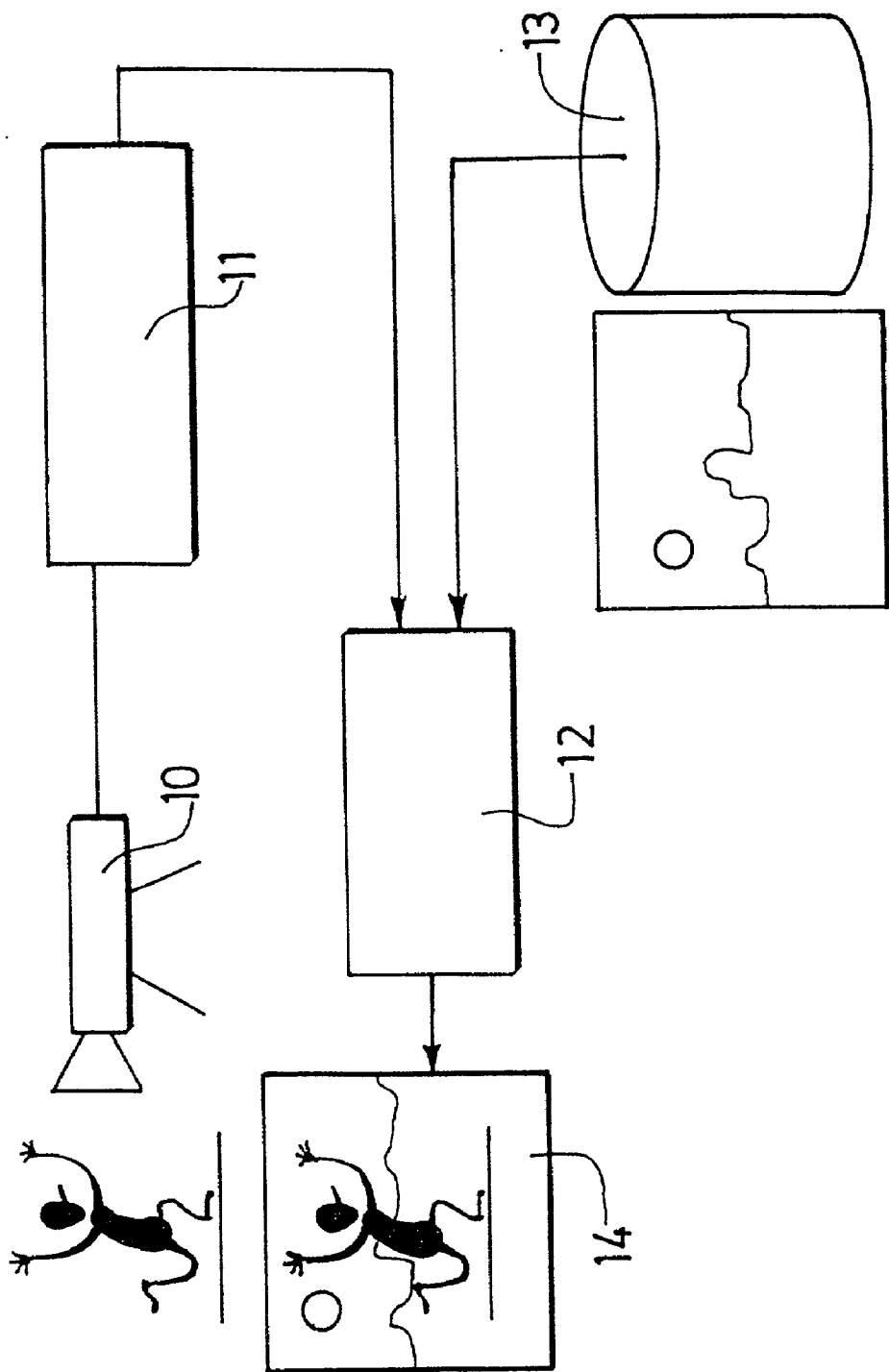


FIG.1

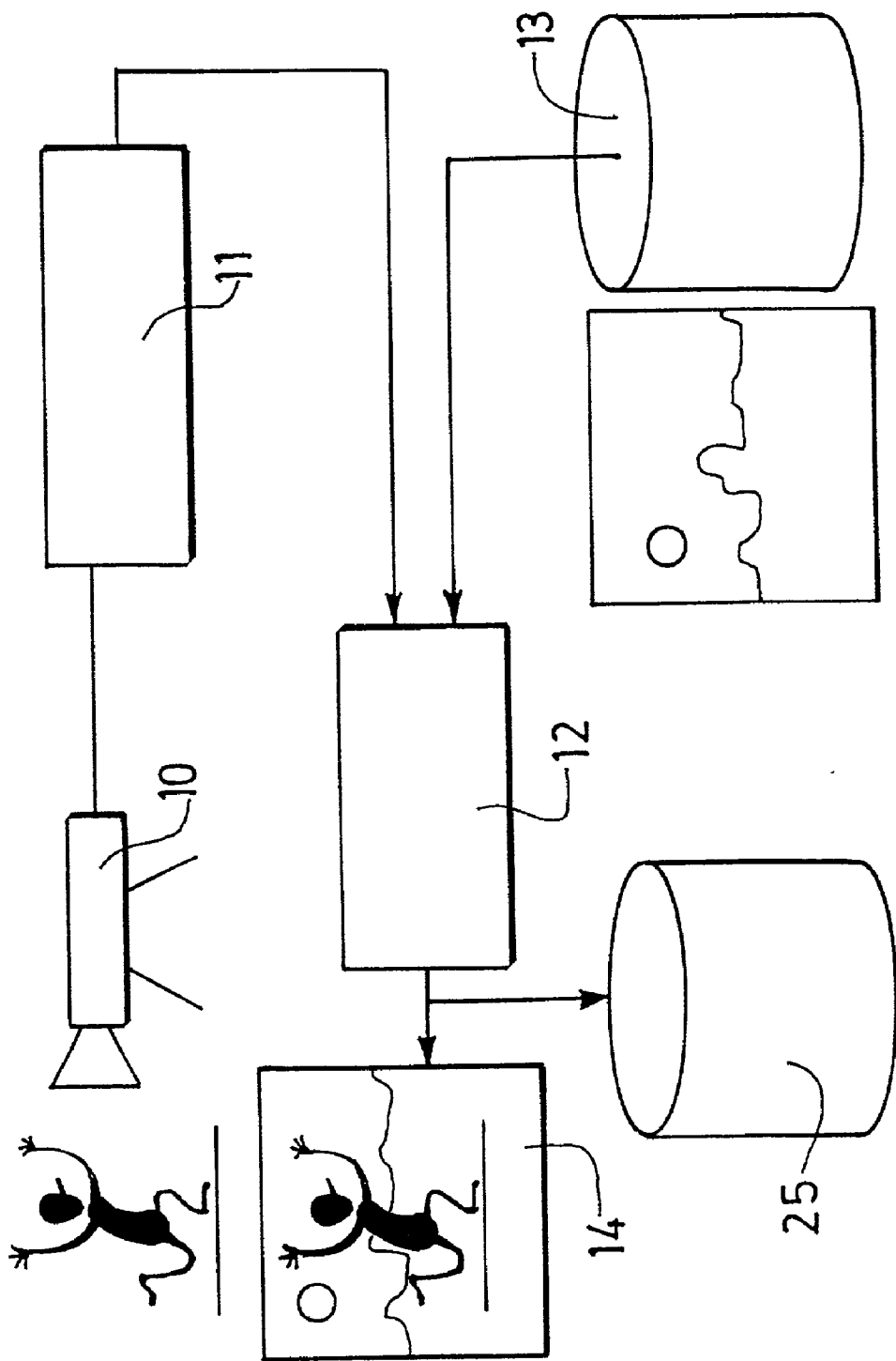


FIG.2

KARAOKE SYSTEM

[0001] The present invention relates to a karaoke system for singing during a sequence such as a video clip or a movie.

[0002] In a karaoke system such as described for instance in the european patent application EP 0782338, music, lyrics or any kind of audio data are transmitted from a transmission station to a distribution station. Music control means of a main module of the system airs music through a built-in speaker of a monitor and voice from an unshown microphone through said speaker. Image control means display a background image (such as a video image or a still image extracted from background image storage means) on the monitor, and lyrics control means display lyrics by superimposing them upon the background image. An image pickup device such as a CCD camera picks up an image of a singer and superimposes it on the screen of the monitor through video image control means as a superimposed image. Such a system defines what may be called "video mixing" in the concept of karaoke.

[0003] It is an object of the invention to propose another type of karaoke system, equipped with an additional functionality.

[0004] To this end, the system relates to a karaoke system for singing during a sequence such a video clip or movie and comprising in series a pickup device, for picking up an image and the voice of an user, an analysis and processing device, for separating at least a part of the user from the background, a mixing and rendering device, for combining the output signal of said analysis and processing device and a pre-recorded material, and a displaying device, for displaying the combination signal.

[0005] Up to now, the concept of singing karaoke was exclusively relying on audio based techniques giving limited functionality and no possibility of real insertion of the user into the video imaginary world. The proposed solution, that introduces the idea of video mixing, allows to extend this karaoke concept to video and, more generally, to develop the concept of a complete audio-video immersion: according to said concept, the voice and the face in the video clip of the song can be replaced by the voice and the face of the fortuitous singer (hereinafter also called user since he/she may be in fact a singer, a player, a dancer, and so on . . .). The same proposed technique may find similar applications in other contexts, for instance in the field of e-commerce or for video editing of pre-recorded content.

[0006] The present invention will now be described, by way of example, with reference to the accompanying drawing in which:

[0007] **FIG. 1** is a block diagram of a karaoke system according to the invention;

[0008] **FIG. 2** illustrates another implementation of a karaoke system according to the invention.

[0009] As shown in **FIG. 1**, the different sub-systems necessary to implement the karaoke system according to the invention are mainly an analysis and processing device **11** and a mixing and rendering device **12**.

[0010] The analysis and processing device **11**, that receives the image and voice of the user (the person shown

in black) picked up by a pickup device **10**, consists of a segmentation circuit, provided for separating for instance the face of the user from the background and thus defining an alpha plane (such a circuit can be based for example on the chroma key technique, if the user is put on stage, in front of a blue screen). The mixing and rendering device **12** is a circuit using the shape information analyzed in the device **11** for composing the user with a background pre-recorded video or audio-video delivered by a medium **13** (said pre-recorded material is shown at the left side of the medium). This composition completes the audio one, provided for mixing the voice of said user with the background pre-recorded music from the song. Using the alpha plane defined thanks to the device **11**, it is then easy to combine the two sources, according to a relation of the following type: $[(\text{video } 1 \times \alpha) + (\text{video } 2 \times (255 - \alpha))] / 255 = \text{final video}$. A displaying device **14** such as a monitor is finally used for displaying the final result (i.e. the combination of the pre-recorded material and of what belongs specifically to the user).

[0011] Obviously, for an improved quality, the analysis implemented in the device **11** can produce 8-bit alpha plane, which enables a better mixing on the frontier of the incrustated object. It may also be noted that the system can replace either only the head of the user or his/her whole body.

[0012] Different cases may be considered, with respect to the type of audio/video sources:

[0013] (a) the two audio/video sources are not compressed: this option may be used for instance in a karaoke restaurant, when the whole body of a singer is incrustated in the clip/movie (the pre-recorded data can be stored on a tape, and the fortuitous singer video can be analyzed and transmitted directly to the video mixer);

[0014] (b) one or the two sources are compressed: one adapted framework for such a case is then the newly developed MPEG-4 standard, that enables to encode the shape and alpha plane of an object—here the face of the fortuitous user—(MPEG-4 has defined a whole system framework enabling the composition of audio and video objects).

[0015] Different cases of application of the invention may also be considered:

[0016] (a) the user may want to record the result of the mixing operation: this is illustrated in **FIG. 2** that shows a system similar to the implementation of **FIG. 1** but including an additional recording device **25**;

[0017] (b) in some cases, the karaoke system will work on line: the pre-recorded clips are then stored on a database (on the Internet for instance), and the user is recording his/her performance at home and wants to produce the combination karaoke clip and put it on his/her homepage (the use of compression techniques is particularly useful in such a case and, more generally, in all the applications that are run in a bandwidth constrained environment);

[0018] (c) in some cases also, the user may want to only put his/her head in place of the original singer's head, which induces further processing in the mixing and rendering device 12, since it is needed that the position of the head of the user and the orientation and posture of the body of the original singer match up.

1. A karaoke system for singing during a sequence such a video clip or movie and comprising in series a pickup

device, for picking up an image and the voice of an user, an analysis and processing device, for separating at least a part of the user from the background, a mixing and rendering device, for combining the output signal of said analysis and processing device and a pre-recorded material, and a displaying device, for displaying the combination signal.

* * * * *