(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년05월29일
(11) 등록번호 10-1862128
(24) 등록일자 2018년05월23일

(51) 국제특허분류(Int. Cl.)
H04N 5/262 (2006.01) *G06K 9/20* (2006.01)
(21) 출원번호 10-2012-0018636
(22) 출원일자 2012년02월23일
심사청구일자 2017년02월10일
(65) 공개번호 10-2013-0096983
(43) 공개일자 2013년09월02일
(56) 선행기술조사문헌
US20080218603 A1*
EP01347646 A2*
KR1020080109519 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
삼성전자 주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
(72) 발명자
박미정
경기 수원시 영통구 인계로 219, 2동 1305호 (매탄동, 삼성1차아파트)
강지영
경기 수원시 영통구 영통로90번길 4-27, 108동 1303호 (망포동, 늘푸른벽산아파트)
(뒷면에 계속)
(74) 대리인
윤동열

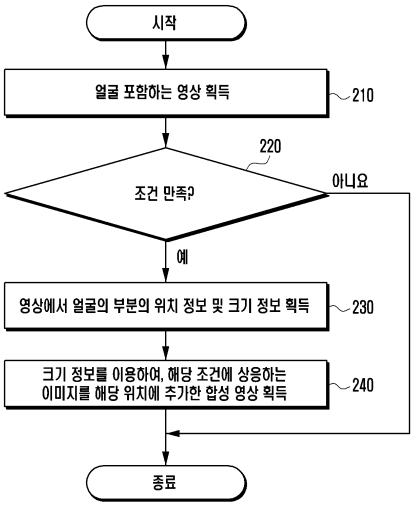
전체 청구항 수 : 총 15 항 심사관 : 진민숙

(54) 발명의 명칭 **얼굴을 포함하는 영상 처리 방법 및 장치**

(57) 요 약

본 발명은 영상 처리 방법 및 장치에 관한 것으로, 본 발명의 일 실시 예에 따르는 영상 처리 방법은, 얼굴을 포함하는 영상을 획득하는 단계, 상기 영상을 획득한 후, 미리 설정된 조건이 만족되는지 여부에 대한 정보를 획득하는 조건 처리 단계, 미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하는 단계 및 상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 영상 획득 단계를 포함할 수 있다. 본 발명의 일 실시 예에 따르면 영상 내의 얼굴과 조화를 이루도록 이미지를 합성하여 제공할 수 있는 영상 처리 방법 및 장치를 제공할 수 있는 효과가 있다.

대 표 도 - 도2



(72) 발명자

이치훈

서울 서대문구 경기대로 26-7, 101동 1401호 (충정로3가, 우리유앤미아파트)

오세기

경기 고양시 일산서구 주엽로 98, 1709동 802호 (주엽동, 문촌마을17단지아파트)

명 세 서

청구범위

청구항 1

영상 처리 방법에 있어서,

얼굴을 포함하는 영상을 획득하는 단계;

상기 영상을 획득한 후, 미리 설정된 조건이 설정된 시간 이상 만족되는지 여부에 대한 정보를 획득하는 조건 처리 단계;

상기 미리 설정된 조건이 상기 설정된 시간 이상 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하는 단계; 및

상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 영상 획득 단계를 포함하는 영상 처리 방법.

청구항 2

제1항에 있어서,

미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 크기 정보를 획득하는 단계를 더 포함하고,

상기 영상 획득 단계는, 상기 영상에 대해, 상기 얼굴의 부분의 크기 정보를 이용하여 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 단계를 포함하는 영상 처리 방법.

청구항 3

제2항에 있어서,

음성을 획득하는 단계를 더 포함하고,

상기 조건 처리 단계는, 상기 음성이 상기 미리 설정된 조건에 부합하는지 여부에 대한 정보를 획득하는 단계를 포함하는 영상 처리 방법.

청구항 4

제3항에 있어서,

상기 조건 처리 단계는,

상기 음성에서 상기 미리 설정된 조건에 부합하는 어구가 추출되는지 여부에 대한 정보를 획득하는 단계를 포함하는 영상 처리 방법.

청구항 5

제3항에 있어서,

상기 조건 처리 단계는,

상기 음성을 음성 분석 장치로 송신하는 단계; 및

상기 음성 분석 장치로부터 상기 음성에서 상기 미리 설정된 조건에 부합하는 어구가 추출되는지 여부에 대한 정보를 획득하는 단계를 포함하는 영상 처리 방법.

청구항 6

제2항에 있어서,

상기 조건 처리 단계는,

상기 영상과 미리 설정된 영상 패턴의 유사도를 획득하는 단계;

상기 영상과 미리 설정된 영상 패턴이 미리 설정된 값 이상의 유사도를 가지는 경우, 상기 미리 설정된 조건이 만족된 것으로 판단하는 단계를 포함하는 영상 처리 방법.

청구항 7

제1항에 있어서,

상기 조건 처리 단계는,

상기 영상으로부터 미리 설정된 표정 정보를 획득하면, 상기 미리 설정된 조건이 만족된 것으로 판단하는 단계를 포함하는 영상 처리 방법.

청구항 8

영상 처리 장치에 있어서,

얼굴을 포함하는 영상을 획득하고, 상기 영상을 획득한 후, 미리 설정된 조건이 설정된 시간 이상 만족되는지 여부에 대한 정보를 획득하고, 상기 미리 설정된 조건이 상기 설정된 시간 이상 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하고, 상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 제어부를 포함하는 영상 처리 장치.

청구항 9

제8항에 있어서,

상기 제어부는, 미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 크기 정보를 획득하고, 상기 영상에 대해, 상기 얼굴의 부분의 크기 정보를 이용하여 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 것을 특징으로 하는 영상 처리 장치.

청구항 10

제9항에 있어서,

상기 제어부는, 음성을 획득하고, 상기 음성이 상기 미리 설정된 조건에 부합하는지 여부에 대한 정보를 획득하여 상기 미리 설정된 조건이 만족되는지 여부에 대한 정보를 획득하는 것을 특징으로 하는 영상 처리 장치

청구항 11

제10항에 있어서,

상기 제어부는, 상기 음성에서 상기 미리 설정된 조건에 부합하는 어구가 추출되는지 여부에 대한 정보를 획득하는 것을 특징으로 하는 영상 처리 장치.

청구항 12

제10항에 있어서,

상기 음성을 음성 분석 장치로 송신하고, 상기 음성 분석 장치로부터 상기 음성에서 상기 미리 설정된 조건에 부합하는 어구가 추출되는지 여부에 대한 정보를 수신하는 통신부를 더 포함하는 영상 처리 장치.

청구항 13

제9항에 있어서,

상기 제어부는, 상기 영상과 미리 설정된 영상 패턴의 유사도를 획득하고, 상기 영상과 미리 설정된 영상 패턴이 미리 설정된 값 이상의 유사도를 가지는 경우, 상기 미리 설정된 조건이 만족된 것으로 판단하는 것을 특징으로 하는 영상 처리 장치.

청구항 14

제8항에 있어서,

상기 제어부는, 상기 영상으로부터 미리 설정된 표정 정보를 획득하면, 상기 미리 설정된 조건이 만족된 것으로 판단하는 것을 특징으로 하는 영상 처리 장치.

청구항 15

영상 처리 방법을 구현하기 위해 컴퓨터가 읽을 수 있는 실행코드를 포함하는 저장 매체에 있어서,

상기 영상 처리 방법은,

얼굴을 포함하는 영상을 획득하는 단계;

상기 영상을 획득한 후, 미리 설정된 조건이 설정된 시간 이상 만족되는지 여부에 대한 정보를 획득하는 조건 처리 단계;

상기 미리 설정된 조건이 상기 설정된 시간 이상 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하는 단계; 및

상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 영상 획득 단계를 포함하는 저장 매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상 처리 방법 및 장치에 관한 것으로, 더욱 상세하게는 얼굴을 포함하는 영상을 처리하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 과거에는 영상 녹화를 위해 고가의 장비가 요구되었으며, 영상 통화는 일반적인 통화방식이 아니었다. 하지만 근래의 영상 처리 기술의 발전으로 휴대용 단말기를 통해 영상 통화를 수행하거나 영상을 녹화하는 일이 빈번히 일어나고 있다.

[0003] 영상 처리 기술 중 특히 사람의 얼굴 부분을 촬영하거나, 사람의 얼굴 부분을 촬영한 영상을 처리하는 기술이 크게 관심을 받고 있다. 일반적으로 영상 통화를 수행할 때에는 통화자의 얼굴 부분을 촬영하여 상대 통화자에게 송신해 준다. 이와 같은 영상 통화를 제공할 때 통화 영상에 특정 이미지나 애니메이션(움직이는 이미지)을 합성하는 방식이 알려져 있다. 다만, 종래의 방식에 따르면, 사용자가 이미지(움직이는 이미지, 즉 애니메이션을 포함한다. 이하 같다.) 표시를 위한 메뉴를 수동으로 선택하면, 휴대 단말기는 선택된 메뉴에 해당하는 이미지를 통화 영상에 합성하여 제공 및/또는 송신한다. 하지만, 이러한 방식에 따르면 영상 통화 도중 사용자가 이미지를 수동 선택하여야 하고, 그에 따라 대화의 흐름이 끊어지는 문제점이 있다. 또한, 영상 내에 표현된 얼굴 위치나 얼굴 크기에 대한 고려 없이 미리 설정된 위치 또는 임의의 위치에 임의의 크기로 이미지를 합성하여 영상 내의 얼굴과 조화를 이루지 못하는 문제가 있다.

발명의 내용

해결하려는 과제

[0004] 본 발명의 일 실시 예는 상술한 문제점을 해결하기 위하여 제안된 것으로, 영상 내의 얼굴과 조화를 이루도록 이미지를 합성하여 제공할 수 있는 영상 처리 방법 및 장치를 제공하는 데 그 목적이 있다.

[0005] 또한, 본 발명의 일 실시 예는 대화의 흐름을 끊지 않고 적절한 이미지를 영상에 합성하여 제공할 수 있는 영상 처리 방법 및 장치를 제공하는 데 그 목적이 있다.

과제의 해결 수단

[0006] 상술한 과제를 달성하기 위하여, 본 발명의 일 실시 예에 따르는 영상 처리 방법은, 얼굴을 포함하는 영상을 획

득하는 단계, 상기 영상을 획득한 후, 미리 설정된 조건이 만족되는지 여부에 대한 정보를 획득하는 조건 처리 단계, 미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하는 단계 및 상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 영상 획득 단계를 포함할 수 있다.

[0007] 상술한 과제를 달성하기 위하여, 본 발명의 일 실시 예에 따르는 영상 처리 장치는, 얼굴을 포함하는 영상을 획득하고, 상기 영상을 획득한 후, 미리 설정된 조건이 만족되는지 여부에 대한 정보를 획득하고, 미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하고, 상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 제어부를 포함할 수 있다.

[0008] 영상 처리 방법을 구현하기 위해 컴퓨터가 읽을 수 있는 실행코드를 포함하는 저장 매체가 제공된다. 상기 영상 처리 방법은, 얼굴을 포함하는 영상을 획득하는 단계, 상기 영상을 획득한 후, 미리 설정된 조건이 만족되는지 여부에 대한 정보를 획득하는 조건 처리 단계, 미리 설정된 조건이 만족되는 경우 상기 영상에서 상기 얼굴의 부분의 위치 정보를 획득하는 단계 및 상기 영상에 대해, 상기 얼굴의 부분의 위치에 상기 만족된 조건에 상응하는 이미지를 추가한 합성 영상을 획득하는 영상 획득 단계를 포함할 수 있다.

발명의 효과

[0009] 본 발명의 일 실시 예에 따르면 영상 내의 얼굴과 조화를 이루도록 이미지를 합성하여 제공할 수 있는 영상 처리 방법 및 장치를 제공할 수 있는 효과가 있다.

[0010] 또한, 본 발명의 일 실시 예에 따르면, 대화의 흐름을 끊지 않고 적절한 이미지를 영상에 합성하여 제공할 수 있는 영상 처리 방법 및 장치를 제공할 수 있는 효과가 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 일 실시 예에 따르는 영상 처리 장치(100)의 블록구성도이다.

도 2는 본 발명의 제1 실시 예에 따르는 영상 처리 과정의 순서도이다.

도 3은 본 발명의 제2 실시 예에 따르는 영상 처리 과정의 순서도이다.

도 4는 본 발명의 제2 실시 예에 따르는 영상 처리 과정에서 영상의 합성을 나타낸다.

도 5는 본 발명의 제3 실시 예에 따르는 영상 처리 과정의 순서도이다.

도 6은 본 발명의 제3 실시 예에 따르는 영상 합성을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0012] 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0013] 실시 예를 설명함에 있어서 본 발명이 속하는 기술 분야에 익히 알려져 있고 본 발명과 직접적으로 관련이 없는 기술 내용에 대해서는 설명을 생략한다. 이는 불필요한 설명을 생략함으로써 본 발명의 요지를 흐리지 않고 더욱 명확히 전달하기 위함이다.

[0014] 마찬가지 이유로 첨부 도면에 있어서 일부 구성요소는 과장되거나 생략되거나 개략적으로 도시되었다. 또한, 각 구성요소의 크기는 실제 크기를 전적으로 반영하는 것이 아니다. 각 도면에서 동일한 또는 대응하는 구성요소에는 동일한 참조 번호를 부여하였다.

[0015] 이하, 본 발명의 실시 예들에 의하여 영상 처리 방법 및 장치를 설명하기 위한 도면들을 참고하여 본 발명에 대해 설명하도록 한다.

[0016] 도 1은 본 발명의 일 실시 예에 따르는 영상 처리 장치(100)의 블록구성도이다.

[0017] 도 1을 참조하면, 본 발명의 일 실시 예에 따르는 영상 처리 장치(100)는 통신부(110), 오디오처리부(120), 표시부(130), 입력부(140), 저장부(150), 제어부(160) 및 카메라부(170)를 포함한다.

[0018] 통신부(110)는 영상 처리 장치(100)의 무선 통신 또는 유선 통신을 위한 해당 데이터의 송수신 기능을 수행한다. 통신부(110)는 예를 들어 송신되는 신호의 주파수를 상승 변환 및 증폭하는 RF송신기와, 수신되는 신

호를 저 잡음 증폭하고 주파수를 하강 변환하는 RF수신기 등으로 구성될 수 있다. 또한, 통신부(110)는 무선 채널을 통해 데이터를 수신하여 제어부(160)로 출력하고, 제어부(160)로부터 출력된 데이터를 무선 채널을 통해 전송할 수 있다.

[0019] 특히 본 발명의 일 실시 예에 따르는 통신부(110)는 얼굴을 포함하는 영상을 송수신하거나, 이미지를 합성한 영상을 송수신할 수 있다. 또한 통신부(110)는 음성을 송수신할 수 있다. 통신부(110)는 음성을 외부의 음성 분석 장치로 송신하고, 음성 분석 장치로부터 해당 음성에 특정 어구가 포함돼 있는지, 또는 기타 조건을 만족하는지에 대한 정보를 수신할 수 있다. 또한 실시 예에 따라 통신부(110)는 외부의 영상 분석 장치에 영상을 송신하고, 영상 분석 장치로부터 해당 영상에 포함된 얼굴의 부분의 위치 및/또는 크기 정보를 수신할 수 있다. 또한 통신부(110)는 해당 영상이 미리 설정된 조건을 만족하는지, 예를 들어 특정 표정 정보가 발견되는지에 대한 정보를 영상 분석 장치로부터 수신할 수 있다. 외부의 음성 분석 장치 또는 영상 분석장치에 의존하지 않고, 제어부(160)가 음성 분석 및/또는 영상 분석을 수행하는 경우 상술한 통신부(110)의 기능들 중 외부의 음성 분석 장치 및 영상 분석 장치와의 통신 기능은 일부 또는 전부가 생략될 수 있다. 또한 영상 통화와 같은 영상/음성 송수신이 불필요한 경우에는 통신부가 생략될 수도 있다.

[0020] 오디오처리부(120)는 디지털 오디오 신호를 오디오 코덱을 통해 아날로그 오디오 신호로 변환하여 스피커(SPK)를 통해 재생하고, 마이크(MIC)로부터 입력되는 아날로그 오디오 신호를 오디오 코덱을 통해 디지털 오디오 신호로 변환한다. 오디오처리부(120)는 코덱(CODEC)으로 구성될 수 있으며, 코덱은 패킷 데이터 등을 처리하는 데이터 코덱과 음성 등의 오디오 신호를 처리하는 오디오 코덱으로 구성될 수 있다. 오디오에 대한 별도의 처리가 필요하지 않은 경우 오디오처리부(120)는 생략될 수 있다.

[0021] 표시부(130)는 음성 처리 장치(100)의 메뉴, 입력된 데이터, 기능 설정 정보 및 기타 다양한 정보를 사용자에게 시각적으로 제공한다. 표시부(130)는 음성 처리 장치(100)의 부팅 화면, 대기 화면, 메뉴 화면, 통화 화면, 기타 어플리케이션 화면을 출력하는 기능을 수행한다. 실시 예에 따라서 표시부(130)는 후술하는 실시 예에 따라 합성된 이미지를 표시할 수 있다. 또한 음성 처리 장치(100)가 영상 통화를 지원하는 경우 표시부(130)는 영상 통화 상대방으로부터 수신한 영상을 표시할 수 있다.

[0022] 표시부(130)는 액정표시장치(LCD, Liquid Crystal Display), 유기 발광 다이오드(OLED, Organic Light Emitting Diodes), 능동형 유기 발광 다이오드(AMOLED, Active Matrix Organic Light Emitting Diodes) 등으로 형성될 수 있다.

[0023] 영상을 직접 표시하지 않고 영상을 다른 장치로 송신하는 실시 예에 따르면 표시부(130)는 생략될 수 있다. 이 경우 영상 처리 장치(100)로부터 영상을 수신한 다른 장치라 해당 영상을 표시할 수 있다.

[0024] 입력부(140)는 사용자의 제어 입력을 수신하여 제어부(160)로 전달한다. 입력부(140)는 터치센서 및/또는 키패드의 형태로 구현될 수 있다.

[0025] 터치센서는 사용자의 터치 입력을 감지한다. 터치센서는 정전용량 방식(capacitive overlay), 압력식 저항막 방식(resistive overlay), 적외선 감지 방식(infrared beam) 등의 터치 감지 센서로 구성되거나, 압력 감지 센서(pressure sensor)로 구성될 수도 있다. 상기 센서들 이외에도 물체의 접촉 또는 압력을 감지할 수 있는 모든 종류의 센서 기기가 본 발명의 터치센서로 구성될 수 있다. 터치센서는 사용자의 터치 입력을 감지하고, 감지 신호를 발생시켜 제어부(160)로 전송한다. 상기 감지 신호에는 사용자가 터치를 입력한 좌표 데이터가 포함된다. 사용자가 터치 위치 이동 동작을 입력한 경우에 터치센서는 터치 위치 이동 경로의 좌표 데이터를 포함한 감지 신호를 발생시켜 제어부(160)로 전송한다.

[0026] 키패드는 영상 처리 장치(100)를 제어하기 위한 사용자의 키 조작을 입력받고 입력 신호를 생성하여 제어부(160)에 전달한다. 키패드는 숫자 키, 방향키를 포함할 수 있으며, 영상 처리 장치(100)의 일면에 소정의 기능키로 형성될 수 있다.

[0027] 입력부(140)는 실시 예에 따라 터치 센서만을 포함하거나 키패드만을 포함할 수도 있고, 별도의 제어 입력이 불필요한 경우 생략될 수도 있다.

[0028] 저장부(150)는 휴대 단말기(100)의 동작에 필요한 프로그램 및 데이터를 저장한다. 저장부(150)는 프로그램 영역과 데이터 영역으로 구분될 수 있다. 프로그램 영역은 영상 처리 장치(100)의 전반적인 동작을 제어하는 프로그램 및 영상 처리 장치(100)를 부팅시키는 운영체제(OS, Operating System), 멀티미디어 컨텐츠 재생 등에 필요한 응용 프로그램, 영상 처리 장치(100)의 기타 옵션 기능, 예컨대, 카메라 기능, 소리 재생 기능, 이미지 또는 동영상 재생 기능에 필요한 응용 프로그램 등을 저장할 수 있다. 데이터 영역은 영상 처리 장치(100)의 사용

에 따라 발생하는 데이터가 저장되는 영역으로서, 이미지, 동영상, 폰 북, 오디오 데이터 등을 저장할 수 있다. 특히 저장부(150)는 카메라부(170)가 촬영한 영상을 저장하거나, 제어부(160)가 합성한 영상을 저장할 수 있다.

[0029] 카메라부(170)는 영상을 촬영한다. 촬영된 영상은 제어부(160)로 전달되거나 저장부(150)에 저장된다. 영상 촬영을 위한 구성은 널리 알려진 구성이므로 여기서는 상세한 설명을 생략한다. 영상을 직접 촬영하지 않고 외부의 다른 장치로부터 수신하는 실시 예에 따르면 카메라부(170)는 생략될 수 있다.

[0030] 제어부(160)는 영상 처리 장치(100)의 각 구성 요소에 대한 전반적인 동작을 제어한다.

[0031] 특히 제어부(160)는 도 2 내지 도 6을 참조하여 후술하는 바와 같이 영상으로부터 해당 영상에 포함된 얼굴의 부분의 위치 및/또는 크기 정보를 획득하고, 그에 따라 이미지를 합성할 수 있다. 또한 제어부(160)는 영상 또는 음성이 특정한 조건을 만족하는지 판단하여 그에 따라 영상에 이미지를 합성할 수 있다. 상세한 사항은 도 2 내지 도 6을 참조하여 후술한다.

[0032] 도 2는 본 발명의 제1 실시 예에 따르는 영상 처리 과정의 순서도이다.

[0033] 단계 210에서 영상 처리 장치(100)의 제어부(160)는 얼굴을 포함하는 영상을 획득한다. 예를 들어 영상 처리 장치(100)의 카메라부(170)가 영상을 촬영하여 제어부(160)가 영상을 획득할 수 있다. 또한 통신부(110)가 영상 통화의 상대방으로부터 영상을 수신하여 제어부(160)가 영상을 획득할 수도 있다. 상술한 방식 또는 기타 다른 방식으로 획득된 영상이 이하의 과정에서 활용될 수 있다.

[0034] 이후, 단계 220에서 제어부(160)는 미리 설정된 조건이 만족되는지 판단한다. 예를 들어 제어부(160)는 영상을 분석하여 조건 만족 여부를 판단할 수 있다. 또한 제어부(160)는 영상과 함께 음성을 획득하고, 해당 음성이 미리 설정된 조건을 만족시키는지 판단할 수 있다. 또한 영상과 음성을 모두 분석하여 조건 만족 여부를 판단할 수도 있다. 또한 다른 예에 따르면, 제어부는 획득한 영상 및/또는 음성을 외부의 분석 장치로 송신하고, 그로부터 수신한 분석 결과를 이용하여 조건 만족 여부를 판단할 수 있다. 구체적인 조건 만족 여부 판단에 대해서는 도 3 내지 도 6을 참조하여 후술한다.

[0035] 해당 조건이 만족되지 않는 경우 과정은 종료된다.

[0036] 해당 조건이 만족되는 경우 과정은 단계 230으로 진행한다.

[0037] 단계 230에서 제어부(160)는 영상에서 얼굴의 부분의 위치 정보 및 크기 정보를 획득한다. 예를 들어 제어부(160)는 얼굴의 눈 및/또는 입의 위치 정보 및 크기 정보를 획득할 수 있다. 예를 들어 제어부(160)는 직접 영상을 분석하여 얼굴의 부분의 위치 정보 및 크기 정보를 획득할 수 있다. 다른 예에 따르면, 제어부(160)는 영상 처리 장치(100) 외부의 영상 분석 장치에 영상을 송신하고 그 영상 분석 장치로부터 분석 결과를 수신하여 얼굴의 부분의 위치 정보 및 크기 정보를 획득할 수도 있다. 영상에서 얼굴의 눈, 코의 위치/크기를 인식하는 기술은 종래에 알려진 기술이므로 여기서는 상세한 설명을 생략한다.

[0038] 여기서 위치 정보는 해당 얼굴 부분의 무게중심이 되거나 또는 해당 얼굴 부분을 포함하는 최소 크기의 직사각형의 어느 한 꼭지점이 될 수 있다. 또한 크기 정보는 가로 길이, 세로 길이 정보를 포함하거나 지름/반지름 길이, 대각선 길이 등의 정보를 포함할 수 있다. 크기 정보는 해당 얼굴 부분을 포함하는 최소 크기의 직사각형의 마주보는 꼭지점의 위치 정보의 형태로 표현될 수도 있다. 기타 다양한 방식이 위치/크기 정보를 나타내기 위해 활용될 수 있으며, 이는 단순히 표현 방식의 차이에 불과한 것이며 본질적으로 위치/크기 정보를 나타내기 위해 사용될 수 있는 정보는 모두 위치/크기 정보가 될 수 있다.

[0039] 단계 240에서 제어부(160)는 해당 영상에 상기 위치 정보 및 크기 정보를 이용하여 단계 220에서 만족된 조건에 상응하는 이미지(움직이는 이미지를 포함한다, 이하 같다)를 영상에 추가하여 합성한다. 예를 들어 음성에서 "happy"라는 어구가 발견된 경우, 해당 어구의 발견 조건에 상응하는 이미지인 웃는 형상의 이미지를 영상의 눈 및 입 위치에 추가하여 합성할 수 있다. 이 때 눈 및 입의 크기 정보에 따라 웃는 형상의 이미지의 크기가 조정될 수 있다. 실시 예에 따라서 얼굴의 부분의 크기 정보의 획득 및 이용은 생략될 수 있다. 즉, 제어부(160)는 미리 설정된 크기대로 웃는 형상의 이미지를 추가하여 합성할 수도 있다.

[0040] 도 3은 본 발명의 제2 실시 예에 따르는 영상 처리 과정의 순서도이다.

[0041] 도 4는 본 발명의 제2 실시 예에 따르는 영상 처리 과정에서 영상의 합성을 나타낸다.

[0042] 단계 310에서 영상 처리 장치(100)의 제어부(160)는 영상 및 음성을 획득한다. 예를 들어 제어부(160)는 카메라부(170)및 오디오처리부(120)를 통해 사용자의 촬영된 영상 및 음성을 획득할 수 있다. 다른 예에 따르면 제어부(160)는 통신부(110)를 통해 다른 장치로부터 영상 및 음성을 획득할 수 있다. 도 4를 참조하면 영상 처리 장치(100)는 첫 번째 화면(410)과 같이 영상(415)을 획득한다. 영상(415)는 영상 처리 장치(100)의 카메라부(170)를 통해 촬영된 것이라고 가정한다. 하지만 다른 방식으로 획득된 영상도 마찬가지로 처리될 수 있음은 자명하다.

[0043] 단계 320에서 제어부(160)는 통신부(110)를 통해 획득된 음성을 영상 처리 장치(100) 외부의 음성 분석 장치에 송신한다. 단계 330에서 제어부(160)는 통신부(110)를 통해 해당 음성의 분석 결과를 외부의 음성 분석 장치로부터 수신한다. 음성 분석 장치는 예를 들어 음성을 분석하여 미리 설정된 어구가 발견되는 경우 이를 영상 처리 장치(100)에게 알려준다. 음성을 분석하여 특정 어구를 발견하는 기술은 널리 공지돼 있으므로 여기서는 상세한 설명을 생략한다. 현재까지 알려진 음성 분석 기술 및/또는 미래에 알려질 음성 분석 기술들이 활용될 수 있다.

[0044] 단계 340에서 제어부(160)는 해당 음성에서 미리 설정된 어구가 발견되는지 판단한다. 예를 들어 "happy"라는 어구가 발견되는 경우 영상 처리 장치(100)가 웃는 표정의 이미지를 합성하는 것으로 가정한다. 사용자가 마이크를 통해 "I'm so happy"라는 음성을 영상 처리 장치(100)에 전달한다면 제어부(160)는 "happy"라는 어구를 발견할 수 있다. 미리 설정된 어구는 사용자에 의해 설정되거나 서비스 제공자, 단말 생산자, 소프트웨어 제공자 중 어느 하나 이상에 의하여 설정될 수도 있다.

[0045] 본 실시 예에 따르면 음성 분석을 통해 특정 어구가 발견되는지에 따라 영상 합성이 진행된다. 하지만 변형 예에 따르면 음성의 음색(tone) 및 음높이(pitch), 크기(loudness) 중 어느 하나 이상을 이용하여 특정한 이미지를 영상에 합성할지 여부를 결정할 수도 있다. 또한 특정 어구의 발견된 경우 해당 어구 부분의 음성의 음색 및 음높이, 크기 중 어느 하나 이상을 고려하여 특정한 이미지를 영상에 합성할지 여부를 결정하는 실시 예도 가능하다. 여기서는 특정 어구가 발견되는지 여부만을 고려하는 것으로 가정한다.

[0046] 해당 어구가 발견되지 않는 경우 과정은 종료된다.

[0047] 해당 어구가 발견되는 경우 과정은 단계 350으로 진행한다.

[0048] 단계 350에서, 제어부(160)는 영상에서 얼굴의 부분의 위치 정보 및 크기 정보를 획득한다. 도 4를 참조하면, 두 눈의 위치(421, 422) 및 입의 위치(423)가 표시된다. 단계 360에서 제어부(160)는 해당 영상에 상기 위치 정보 및 크기 정보를 이용하여 단계 340에서 발견된 어구에 상응하는 이미지를 영상에 추가하여 합성한다. "happy"에 상응하는 이미지는 두 번째 화면(440)의 이미지들(451, 452, 453)이라고 가정한다. 이미지들(451, 452, 453)은 눈 및 입의 위치들(421, 422, 423)에 각각 합성된다. 또한 눈의 크기 및 입의 크기에 따라 이미지들(451, 452, 453)의 크기가 조정될 수 있다. 예를 들어 이미지(451)의 좌우 길이는 눈의 왼쪽 끝으로부터 오른쪽 끝까지의 길이의 두 배 크기가 되도록 조정될 수 있다.

[0049] 예를 들어 이미지(451)는 좌우/상하 크기 비율이 유지되면서 크기가 조정될 수 있다. 즉 눈의 상-하 높이에 관계 없이 이미지(451)의 좌-우 길이가 결정되고, 그에 비례하여 이미지(451)의 상-하 높이도 결정될 수 있다. 반대로 이미지(451)는 좌우/상하 크기 비율이 유지되지 않으면서 크기가 조정될 수 있다. 예를 들어 가늘고 긴 모양의 눈에 대해서는 마찬가지로 이미지(451)를 상대적으로 가늘고 긴 모양으로 변형하여 합성할 수 있다.

[0050] 이렇게 합성된 영상은 두 번째 화면(440)과 같이 영상 처리 장치(100)의 표시부(130)에 표시될 수 있다. 또한 합성된 영상은 영상 통화의 상대방에게 전달될 수도 있다. 영상 통화의 상대방은 세 번째 화면(470)과 같이 이미지들(481, 482, 483)이 합성된 상대방의 얼굴이 포함된 영상을 제공받을 수 있다. 다른 실시 예에 따르면 합성된 영상은 저장부(150)에 반 영구적 또는 임시적으로 저장될 수도 있다. 저장부(150)에 저장된 영상은 추후 재생되거나 추가적 가공에 활용될 수도 있다. 합성된 영상은 저작 툴 기타 애플리케이션을 통해 소셜 네트워크 서비스(SNS)에 활용되거나 웹 페이지에 게시될 수도 있다.

[0051] 상술한 예에서 특정 어구가 발견된 경우 해당 어구가 발견된 시점에 상응하는 영상 부분에 대해 이미지 합성을 적용할 수 있다. 예를 들어 전체 분량 중 2분 00초 위치의 음성에서 "happy"가 발견된 경우 2분 00초~2분 03초 사이의 영상에 웃는 표정의 이미지를 합성할 수 있다. 다만 구체적인 합성 적용의 시작/종료 시점은 실시 예에 따라 다양하게 변형될 수 있다.

[0052] 상술한 어구 외에 복수의 어구가 미리 설정될 수 있으며, 각 어구에 대해 서로 다른 이미지가 할당될 수 있다. 예를 들어 "sad"라는 어구에 대해서는 우는 표정의 이미지가 할당될 수 있다.

[0053] 도 5는 본 발명의 제3 실시 예에 따르는 영상 처리 과정의 순서도이다.

[0054] 도 6은 본 발명의 제3 실시 예에 따르는 영상 합성을 나타낸 도면이다.

[0055] 단계 510에서 제어부(160)는 영상을 획득한다. 제1, 제2 실시 예와 마찬가지로 제어부(160)는 카메라부(170) 또는 통신부(110)를 통해 영상을 획득할 수 있다. 단계 520에서 제어부(520)는 표정 정보를 획득한다. 영상 처리 장치(110)는 미리 설정된 영상 패턴을 영상의 얼굴 부분과 비교하여 표정 정보를 획득할 수 있다. 예를 들어 미리 설정된 웃는 표정의 영상 패턴과 영상을 비교한 결과 미리 설정된 값 이상의 유사도를 얻은 경우 해당 영상의 얼굴은 웃는 것으로 판단할 수 있다. 이 경우 제어부(160)는 웃는 표정에 해당하는 표정 정보를 획득할 수 있다. 영상 패턴은 웃는 표정과 같은 표정의 특징적 부분을 나타낸 패턴이다. 예를 들어 웃는 표정의 경우 눈이 가늘어지거나 보조개가 생기는 특징이 있으므로, 그 부분을 판단할 수 있도록 영상 패턴이 형성될 수 있다. 뽀뽀하는 동작의 경우 입술을 오므렸다가 펴는 특징을 판단할 수 있는 영상 패턴을 이용하여 검출할 수 있다. 해당 부분에 대해 영상과 비교한 결과 보조개가 있고 눈이 가늘어졌음을 판단할 수 있다면, 웃는 표정으로 인식할 수 있을 것이다. 기타 표정의 인식을 위한 현재까지 알려지거나 미래에 알려질 기술이 활용될 수 있다.

[0056] 제어부(160)는 영상을 직접 분석하여 표정을 검출할 수도 있고, 영상 처리 장치(100) 외부의 영상 분석 장치를 이용하여 표정을 검출할 수도 있다.

[0057] 도 6의 첫 번째 화면(610)에는 윙크하는 표정의 얼굴(615)이 도시된다.

[0058] 530에서 제어부(160)는 특정 표정이 검출되는지 판단한다. 즉, 예를 들어 윙크하는 표정의 표정 정보가 검출되는지, 또는 윙크하는 표정에 해당하는 영상 패턴과 영상을 비교하였을 때 미리 설정된 값 이상의 유사도가 도출되는지 확인한다. 도 6의 첫 번째 화면(610)과 같이 윙크하는 표정 등 미리 설정된 표정이 검출되는 경우 과정은 단계 540으로 진행한다. 반대로 특정 표정이 검출되지 않으면 과정은 종료된다. 오동작을 막기 위해 특정 표정의 검출이 미리 설정된 시간, 예를 들어 2초 이상 지속된 경우에만 단계 540 이후의 합성을 진행할 수도 있다.

[0059] 단계 540에서 제어부(160)는 영상에서 얼굴의 부분의 위치/크기 정보를 획득한다. 위치/크기 정보의 획득의 구체적 방식은 도 2 내지 도 4를 참조하여 상술한 바와 같다. 윙크하는 표정에 대응하여 눈초리 부분에 하트 모양의 이미지를 표시하는 예를 가정한다. 제어부(160)는 도 6의 첫 번째 화면(610)의 영상(615)에서 윙크하는 눈의 눈초리의 위치(621) 정보 및 눈의 크기 정보를 획득한다.

[0060] 단계 550에서 제어부(160)는 해당 위치에, 검출된 표정에 대응하는 이미지를 눈의 크기에 따라 조정하여 표시한다. 도 2 내지 도 4의 실시 예와 마찬가지로 크기 정보의 획득 및 이용은 생략될 수도 있다. 도 6의 두 번째 화면(640)을 참조하면 윙크하는 표정에 따라 하트 이미지(651)가 눈초리 위치(621)에 합성되어 표시된다.

[0061] 이렇게 합성된 영상은 두 번째 화면(640)과 같이 영상 처리 장치(100)의 표시부(130)에 표시될 수 있다. 또한 합성된 영상은 영상 통화의 상대방에게 전달될 수도 있다. 영상 통화의 상대방은 세 번째 화면(670)과 같이 이미지(681)가 합성된 상대방의 얼굴이 포함된 영상을 제공받을 수 있다. 다른 실시 예에 따르면 합성된 영상은 저장부(150)에 반 영구적 또는 임시적으로 저장될 수도 있다. 저장부(150)에 저장된 영상은 추후 재생되거나 추가적 가공에 활용될 수도 있다. 합성된 영상은 저작 툴 기타 애플리케이션을 통해 소셜 네트워크 서비스(SNS)에 활용되거나 웹 페이지에 게시될 수도 있다.

[0062] 이 때, 처리 흐름도 도면들의 각 블록과 흐름도 도면들의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수 있음을 이해할 수 있을 것이다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 흐름도 블록(들)에서 설명된 기능들을 수행하는 수단을 생성하게 된다. 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독

가능 메모리에 저장되는 것도 가능하므로, 그 컴퓨터 이용가능 또는 컴퓨터 판독 가능 메모리에 저장된 인스트럭션들은 흐름도 블록(들)에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다. 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 흐름도 블록(들)에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.

[0063] 또한, 각 블록은 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 실행 예들에서는 블록들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 블록들은 사실 실질적으로 동시에 수행되는 것도 가능하고 또는 그 블록들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.

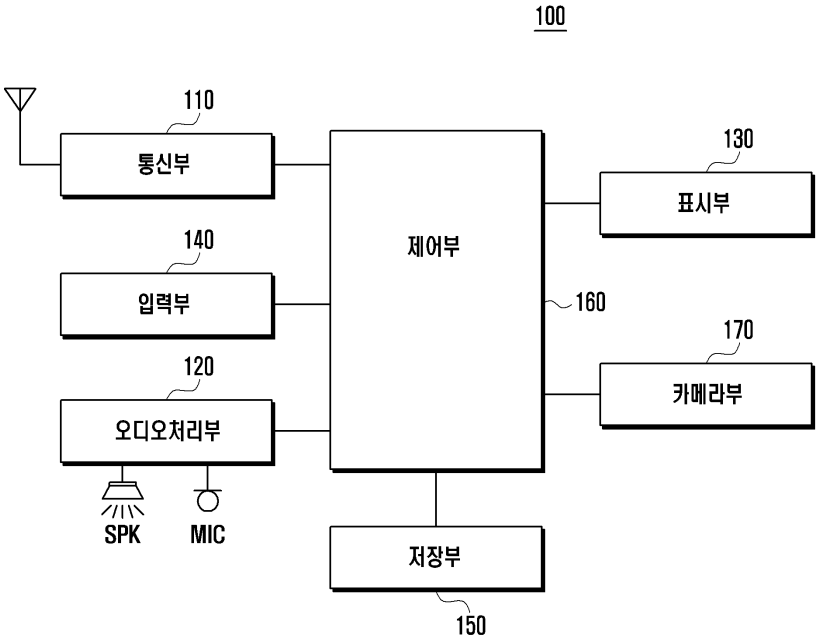
[0064] 이 때, 본 실시 예에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생시키도록 구현될 수도 있다.

[0065] 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명이 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시 예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구의 범위에 의하여 나타내어지며, 특허청구의 범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

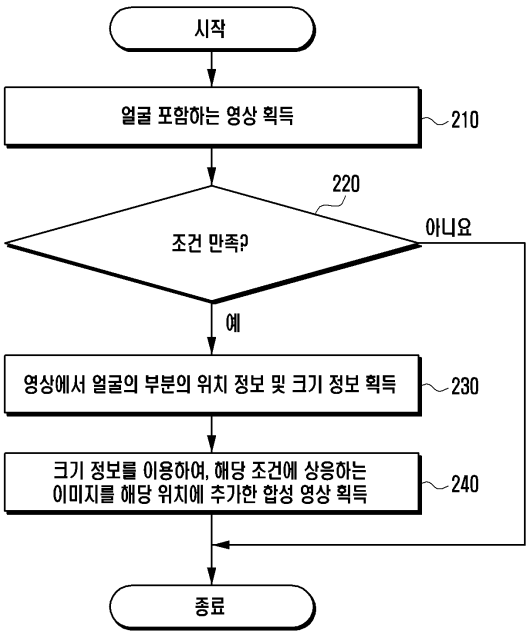
[0066] 한편, 본 명세서와 도면에는 본 발명의 바람직한 실시 예에 대하여 개시하였으며, 비록 특정 용어들이 사용되었으나, 이는 단지 본 발명의 기술 내용을 쉽게 설명하고 발명의 이해를 돕기 위한 일반적인 의미에서 사용된 것이지, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시 예 외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형 예들이 실시 가능하다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다.

도면

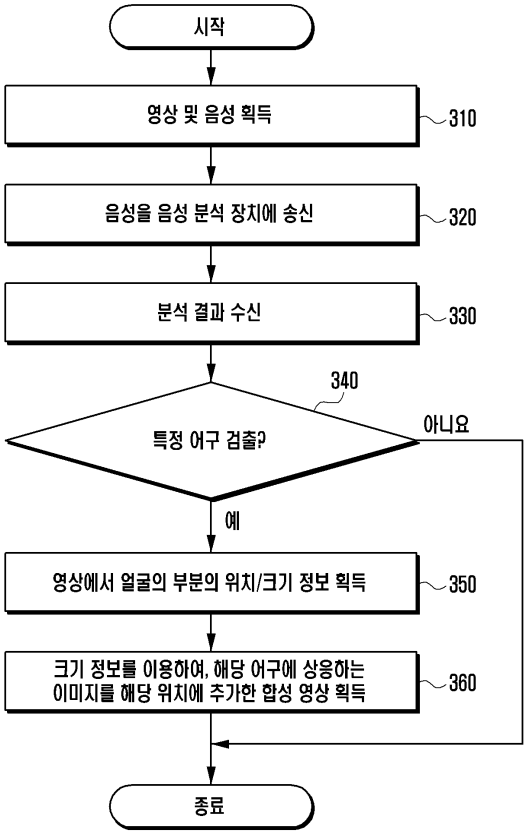
도면1


100
110
통신부
130
표시부
제어부
140
입력부
160
170
카메라부
120
오디오처리부
SPK
MIC
저장부
150

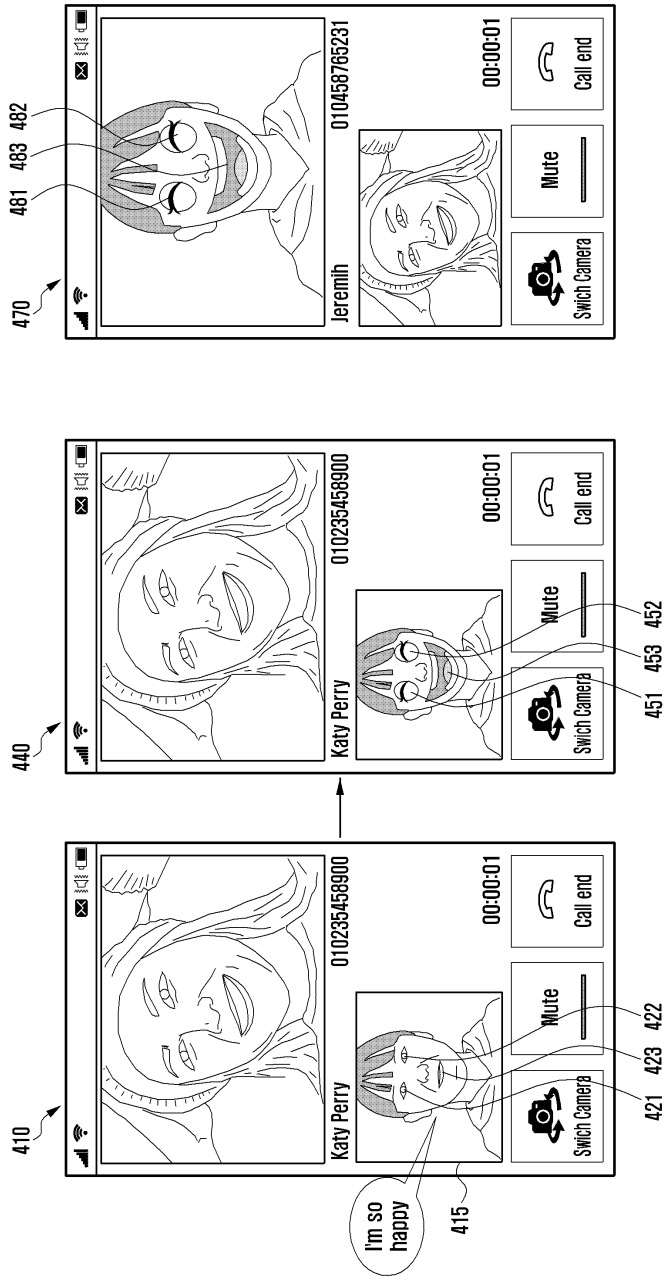
도면2


시작
얼굴 포함하는 영상 획득
210
220
조건 만족?
아니요
예
영상에서 얼굴의 부분의 위치 정보 및 크기 정보 획득
230
크기 정보를 이용하여, 해당 조건에 상응하는
이미지를 해당 위치에 추가한 합성 영상 획득
240
종료

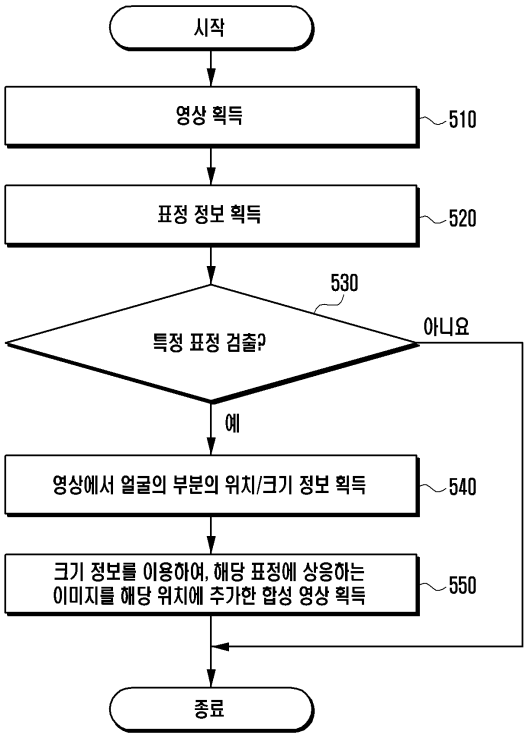
도면3


시작
영상 및 음성 획득
310
음성을 음성 분석 장치에 송신
320
분석 결과 수신
330
340
특정 어구 검출?
아니요
예
영상에서 얼굴의 부분의 위치/크기 정보 획득
350
크기 정보를 이용하여, 해당 어구에 상응하는
이미지를 해당 위치에 추가한 합성 영상 획득
360
종료

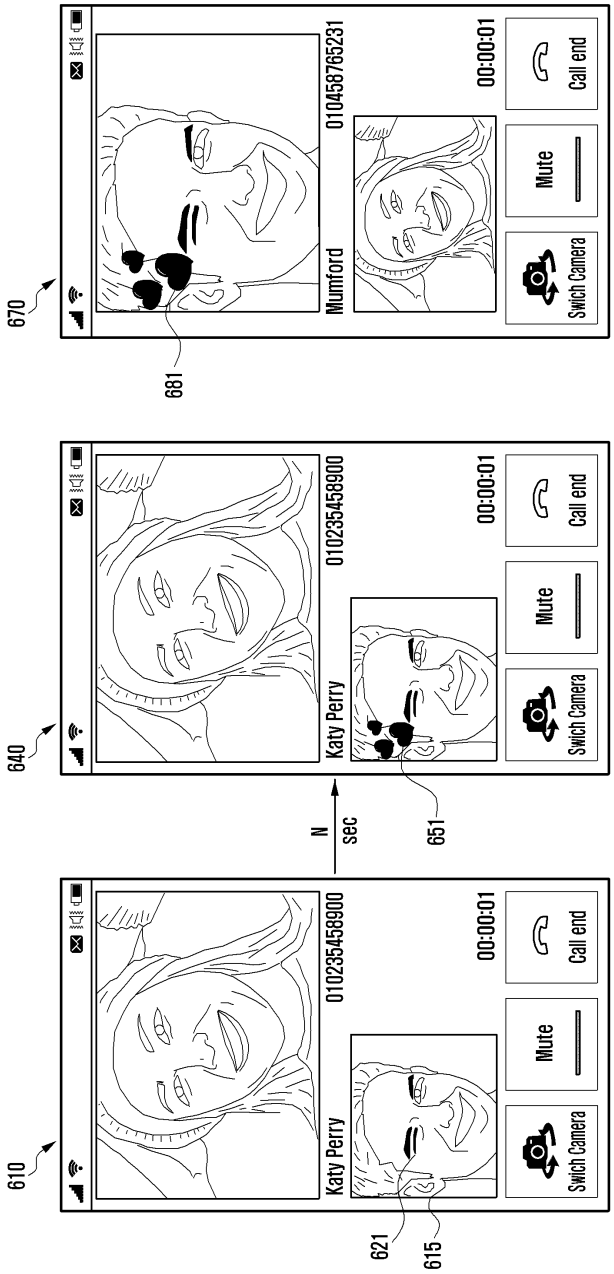
도면4


410
Katy Perry
01023545890O
00:00:01
Call end
Mute
Swich Camera
I'm so happy
415
421
423
422
440
Katy Perry
01023545890O
00:00:01
Call end
Mute
Swich Camera
451
453
452
470
Jeremih
01045876523 1
00:00:01
Call end
Mute
Swich Camera
481
483
482

도면5


시작
영상 획득
510
표정 정보 획득
520
530
특정 표정 검출?
아니요
예
영상에서 얼굴의 부분의 위치/크기 정보 획득
540
크기 정보를 이용하여, 해당 표정에 상응하는
이미지를 해당 위치에 추가한 합성 영상 획득
550
종료

도면6


610
Katy Perry
010235458900
00:00:01
Switch Camera
Mute
Call end
621
615
N
sec
640
Katy Perry
010235458900
00:00:01
Switch Camera
Mute
Call end
651
670
Mumford
010458765231
00:00:01
Switch Camera
Mute
Call end
681