



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년03월03일

(11) 등록번호 10-1497168

(24) 등록일자 2015년02월23일

(51) 국제특허분류(Int. Cl.)

G06T 7/00 (2006.01) H04N 7/15 (2006.01)

(21) 출원번호 10-2009-7027254

(22) 출원일자(국제) 2008년06월02일

심사청구일자 2013년05월21일

(85) 번역문제출일자 2009년12월28일

(65) 공개번호 10-2010-0028060

(43) 공개일자 2010년03월11일

(86) 국제출원번호 PCT/US2008/065579

(87) 국제공개번호 WO 2009/005928

국제공개일자 2009년01월08일

(30) 우선권주장

11/824,412 2007년06월29일 미국(US)

(56) 선행기술조사문헌

US20070127819 A1

WO1996007979 A1

WO2001037552 A1

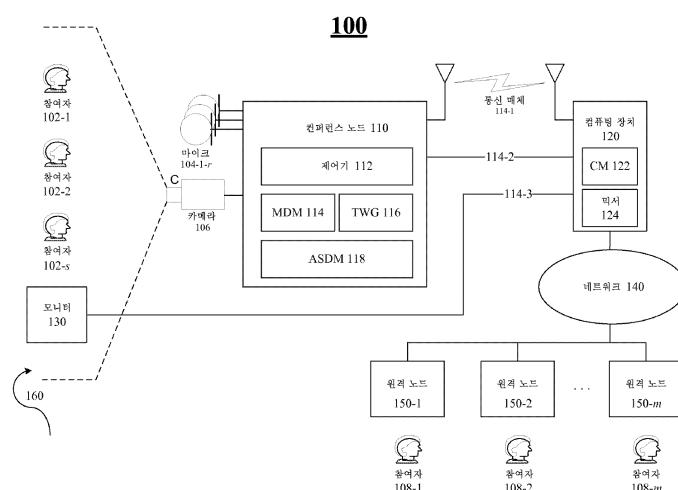
전체 청구항 수 : 총 19 항

심사관 : 신재철

(54) 발명의 명칭 디스플레이 장치 검출 기법

**(57) 요약**

디스플레이 장치를 검출하는 기법들에 대해 기술되어 있다. 장치는 이미지의 비디오 정보를 수신하는 동작을 하거나 비디오 카메라, 및 이미지에 대한 오디오 정보를 수신하는 동작을 하는 마이크를 포함할 수 있다. 상기 장치는 상기 비디오 카메라 및 상기 마이크에 통신 연결된 모니터 검출 모듈을 더 포함할 수 있고, 상기 모니터 검출 모듈은 상기 이미지 내에 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출하고, 이 검출에 기초하여 상기 이미지 내에서 상기 모니터의 위치를 결정하는 동작을 한다. 상기 장치는 또한 상기 모니터 검출 모듈에 통신 연결된 활발한 화자 검출기 모듈을 포함할 수 있고, 상기 활발한 화자 검출기 모듈은 상기 모니터에 의해 야기된 거짓 양성을 배제하는 동작을 한다. 다른 실시예들이 기술되고 특허 청구되어 있다.

**대 표 도**

## 특허청구의 범위

### 청구항 1

이미지 내에 모니터를 갖는 이미지 시퀀스를 수신하는 단계;  
상기 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출하는 단계;  
상기 검출에 기초하여 상기 이미지 내의 상기 모니터의 위치를 결정하는 단계; 및  
전화 회의에서 활발하게 말하는 참여자(an actively speaking participant)가 상기 모니터에 의해 디스플레이되  
는지 여부를 판정하는 단계를 포함하는  
방법.

### 청구항 2

제1항에 있어서,  
상기 모니터에 의해 디스플레이되지 않는 것으로 판정된 전화 회의의 활발하게 말하는 참여자를 선택하는 단계  
를 포함하는  
방법.

### 청구항 3

제1항에 있어서,  
상기 모니터의 상기 위치에 기초한 활발한 화자 검출(active speaker detection)로부터 상기 모니터를 배제시키  
는 단계를 포함하는  
방법.

### 청구항 4

제1항에 있어서,  
상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호를 생성하는 단계를 포함하는  
방법.

### 청구항 5

제1항에 있어서,  
상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 저주파수 패턴을 생성하는 단계를 포함하는  
방법.

### 청구항 6

제1항에 있어서,  
상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 주기 함수를 생성하는 단계를 포함하는  
방법.

### 청구항 7

제1항에 있어서,  
상기 모니터에 의해 디스플레이되는 모니터 신호에 상기 시간 워터마크 신호를 추가하는 단계를 포함하는  
방법.

#### 청구항 8

명령어들을 포함하는 컴퓨터 판독가능 저장 매체로서,  
상기 명령어들은 실행되는 경우에 시스템으로 하여금,  
이미지 내에 모니터를 갖는 이미지 시퀀스를 수신하는 단계;  
상기 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출하는 단계;  
상기 이미지 내의 상기 모니터의 위치를 결정하는 단계; 및  
전화 회의에서 활발하게 말하는 참여자가 상기 모니터에 의해 디스플레이되는지 여부를 판정하는 단계를 수행할 수 있게 하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 9

제8항에 있어서,  
상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터에 의해 디스플레이되지 않는 것으로 판정된, 전화 회의의 활발하게 말하는 참여자를 선택하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 10

제8항에 있어서,  
상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터의 상기 위치에 기초하여 전화 회의의 참여자에 대해 활발한 화자 검출을 수행하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 11

제8항에 있어서,  
상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호를 생성하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 12

제8항에 있어서,  
상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 저주파수 패턴을 생성하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 13

제8항에 있어서,  
상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 주기 함수 직류 오프셋 패턴(periodic function direct current offset pattern)을 생성하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는  
컴퓨터 판독가능 저장 매체.

#### 청구항 14

제8항에 있어서,

상기 명령어들은, 실행되는 경우에 상기 시스템으로 하여금 상기 모니터에 의해 디스플레이되는 모니터 신호에 상기 시간 워터마크 신호를 추가하는 단계를 수행할 수 있게 하는 명령어들을 더 포함하는 컴퓨터 관독가능 저장 매체.

### 청구항 15

이미지의 비디오 정보를 수신하는 동작을 하는 비디오 카메라;

이미지의 오디오 정보를 수신하는 동작을 하는 마이크 어레이;

상기 비디오 카메라 및 마이크에 통신 연결된 모니터 검출 모듈 - 상기 모니터 검출 모듈은 상기 이미지 내에 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출하고, 이 검출에 기초하여 상기 이미지 내에서의 상기 모니터의 위치를 결정하는 동작을 함 -; 및

상기 모니터 검출 모듈에 통신 연결된 활발한 화자 검출기 모듈 - 상기 활발한 화자 검출기 모듈은 상기 모니터에 의해 야기된 거짓 양성(false positive)을 배제하는 동작을 함 - 을 포함하는 장치.

### 청구항 16

제15항에 있어서,

상기 활발한 화자 검출기 모듈은 상기 모니터에 의해 디스플레이되지 않는, 전화 회의의 활발하게 말하는 참여자를 선택하는 동작을 하는

장치.

### 청구항 17

제15항에 있어서,

상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호를 생성하는 동작을 하는 시간 워터마크 발생기를 포함하는

장치.

### 청구항 18

제15항에 있어서,

상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 저주파수 패턴을 생성하는 동작을 하는 시간 워터마크 발생기를 포함하는

장치.

### 청구항 19

제15항에 있어서,

상기 모니터에 의해 디스플레이되는 상기 시간 워터마크 신호로서 주기 함수 직류 오프셋 패턴을 생성하는 동작을 하는 시간 워터마크 발생기를 포함하는

장치.

### 청구항 20

삭제

## 명세서

## 배경 기술

[0001] 전화 회의는 통상적으로 회의실에 있는 다양한 참가자들의 이미지를 캡처하는 비디오 장비, 및 참가자들의 음성을 녹음하는 오디오 장비를 포함한다. 전화 회의 동안, 비디오 카메라의 초점을 주어진 참가자에게 맞추는 것이 바람직할 수 있다. 예를 들어, 비디오 카메라의 초점을 활발한 화자에 맞추기 위해 ASD(active speaker detection, 활발한 화자 검출) 기법이 사용될 수 있다. 이것은 이미지 내에서 사람 음성의 소스를 식별하고 자동으로 비디오 카메라를 이동시키거나 식별된 소스에 초점을 맞춤으로써 달성될 수 있다. 그러나, 일부 경우들에, 회의실 내에 ASD 동작을 방해할 가능성이 있을 수 있는 추가의 물체들이 있다. 이 결과 주어진 화자의 식별, 따라서 비디오 카메라의 초점의 정확도가 감소될 수 있다. 따라서, 이들 및 기타 문제점들을 해결하기 위해 ASD 기법의 개선 필요성이 상당히 있을 수 있다.

[0002] <발명의 요약>

[0003] 이 요약은 이하에서 상세한 설명에 더 기술되는 일련의 개념들을 간략화된 형태로 소개하기 위해 제공된 것이다. 이 요약은 청구된 발명 대상의 중요한 특징들 또는 필수적인 특징들을 확인하기 위한 것이 아니며, 청구된 발명 대상의 범위를 정하는 데 사용되기 위한 것도 아니다.

[0004] 다양한 실시예들이 일반적으로 회의 시스템에 관한 것일 수 있다. 일부 실시예들은 특히 회의 시스템의 ASD 동작을 향상시키는 다양한 모니터 검출 기법에 관한 것일 수 있다. 일부 실시예에서, 예를 들어, 시스템 또는 장치는 이미지의 비디오 정보를 수신하는 동작을 하는 비디오 카메라, 및 이미지에 대한 오디오 정보를 수신하는 동작을 하는 마이크를 포함할 수 있다. 본 시스템은 비디오 카메라 및 마이크에 통신 연결되어 있는 모니터 검출 모듈을 더 포함할 수 있다. 모니터 검출 모듈은 이미지 내에 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출하고, 이 검출에 기초하여 이미지 내에서 모니터의 위치를 결정하는 동작을 할 수 있다. 본 시스템은 또한 모니터 검출 모듈에 통신 연결되어 있는 ASD(active speaker detector) 모듈을 포함할 수 있다. ASD 모듈은 모니터에 의해 거짓 양성이 야기되는지 여부를 판정하는 동작을 할 수 있다. 예를 들어, ASD 모듈은 전화 회의에서 활발하게 말하는 참여자가 모니터에 의해 디스플레이되는지 여부를 판정할 수 있다. ASD 모듈은 그 다음에 모니터에 의해 디스플레이되지 않은 전화 회의의 활발하게 말하는 참여자를 선택할 수 있다. 이와 같이, 비디오 카메라가 발화중인 사람 참가자 또는 기타 외란 비디오를 디스플레이하는 모니터가 아니라 발화중인 사람 참가자에게 초점을 맞출 가능성을 증가시킴으로써 ASD 모듈의 정확도가 향상될 수 있다. 다른 실시예들이 기술되고 특히 청구되어 있다.

## 발명의 상세한 설명

[0011] 다양한 실시예들이 하나 이상의 요소들을 포함하고 있을 수 있다. 요소는 실시예와 관련하여 기술된 임의의 특징, 특성, 구조 또는 동작을 포함할 수 있다. 요소의 일례로는 하드웨어 요소, 소프트웨어 요소, 물리 요소, 또는 이들의 임의의 조합이 있을 수 있다. 실시예가 예로서 소정의 구성에서 제한된 수의 요소들로 기술될 수 있지만, 실시예가 주어진 구현에 대해 원하는 바에 따라 대안의 구성에서 더 많은 또는 더 적은 요소들을 포함할 수 있다. 유의해야 할 점은 "일 실시예" 또는 "실시예" 또는 유사한 문구가 꼭 동일한 실시예를 가리키는 것은 아니라는 것이다.

[0012] 다양한 실시예들이 일반적으로 모니터 등의 디스플레이 장치를 검출하는 기법들에 관한 것일 수 있다. 일부 실시예들은 특히 회의 전화 회의용 비디오 카메라에 의해 캡처된 비디오 신호 또는 이미지 내에 들어 있는 회의실 내의 모니터를 검출하는 것에 관한 것일 수 있다. 예를 들어, 일부 실시예들은 전화 회의 동안 회의실 내의 모니터의 존재를 식별하고 찾아내기 위해 모니터 검출 기법을 구현할 수 있다. 일부 실시예들은 또한 오디오(사운드 소스 위치파악) 및 비디오(움직임 및 공간 패턴) 특징들을 사용하여 회의실 내의 화자를 검출하는 ASD 기법을 구현할 수 있다. 그러나, 일부 경우들에, ASD가 모니터를 선택하는 거짓 양성(false positive)이 일어날 수 있다. 거짓 양성은 ASD가 활발하게 발화하지 않는 또는 주도적인 화자가 아닌 어떤 사람 또는 물체를 선택하는 경우를 말한다. 예를 들어, 이것은 모니터가 발화중인 로컬 참여자 또는 모니터로부터 나오는 어떤 다른 노이즈를 디스플레이할 때마다 일어날 수 있다. 그 결과, 모니터 검출 기법은 ASD가 그에 따라 이러한 종류의 거짓 양성을 감소시킬 수 있도록 회의실 내에 모니터가 존재하는지를 검출하는 데 사용될 수 있다.

[0013] 일 실시예에서, 예를 들어, 회의 시스템의 컨퍼런스 노드(conference node)는 회의실의 이미지를 나타내는 비디오 정보를 수신하는 동작을 하는 하나 이상의 비디오 카메라, 및 회의실로부터 나오는 오디오 정보를 수신하는 동작을 하는 하나 이상의 마이크를 포함할 수 있다. 본 시스템은 비디오 카메라 및 마이크에 통신 연결되어 있는 모니터 검출 모듈을 더 포함할 수 있다. 모니터 검출 모듈은 회의실 이미지 내에 모니터에 의해 디스

플레이되는 시간 워터마크 신호를 검출하고, 이 검출에 기초하여 이미지 내에서 모니터의 위치를 결정하고, 모니터의 존재 및 위치를 나타내는 모니터 검출 신호를 출력하는 동작을 할 수 있다. 본 시스템은 또한 모니터 검출 모듈에 통신 연결되어 있는 ASD 모듈을 포함할 수 있다. ASD 모듈은 모니터 검출 신호를 수신하고 이 모니터 검출 신호를 사용하여 전화 회의에서의 주도적 화자가 모니터에 의해 디스플레이되는지 여부 또는 주도적 화자가 모니터로부터 주어진 근접 거리 내에 있는지 여부를 판정하는 동작을 할 수 있다. ASD 모듈은 그 다음에 모니터에 의해 디스플레이되지 않은 전화 회의의 주도적 화자를 선택할 수 있다. 이와 같이, 비디오 카메라가 발화중인 사람 참가자를 디스플레이하는 모니터가 아니라 발화중인 사람 참가자에게 초점을 맞출 가능성을 증가시키고 그에 의해 화상 회의 서비스 및 전반적인 사용자 만족도를 향상시킴으로써 ASD 모듈의 정확도가 향상될 수 있다.

[0014] 도 1은 화상 회의 시스템(100)의 블록도를 나타낸 것이다. 화상 회의 시스템(100)은 화상 회의 서비스를 제공하도록 구성되어 있을 수 있다. 화상 회의는 2개 이상의 장소가 양방향 비디오 및 오디오 전송을 통해 동시에 대화할 수 있게 하는 일련의 대화형 전기 통신 기술이다. 화상 회의는 오디오 및 비디오의 전기 통신을 사용하여 서로 다른 위치에 있는 사람들을 회의에 모아게 한다. 이것은 개인 사무실 내의 2명의 사람 간의 대화만큼 간단할 수 있거나(포인트-투-포인트), 2명 이상의 사람이 서로 다른 장소에 있는 대형 회의실에 있는 경우 몇 개의 장소를 수반할 수 있다(멀티-포인트). 사람들의 오디오 및 비디오 전송 이외에, 화상 회의는 문서, 컴퓨터-디스플레이 정보 및 화이트보드를 공유하는 데 사용될 수 있다.

[0015] 도 1에 도시되어 있는 예시된 실시예에서, 화상 회의 시스템(100)은 컨퍼런스 노드(110)를 포함할 수 있다. 컨퍼런스 노드(110)는 화상 회의 장비를 갖는 이동형 또는 고정형 전자 장치를 포함할 수 있다. 예를 들어, 컨퍼런스 노드(110)는 제어기(112), 모니터 검출 모듈(114), 시간 워터마크 발생기(116) 및 활발한 화자 검출기 모듈(118)을 포함할 수 있다. 컨퍼런스 노드는 하나 이상의 마이크(104-1-r)를 포함하는 마이크 어레이(104), 및 하나 이상의 비디오 카메라(106-1-p)를 더 포함할 수 있으며, 여기서 r 및 p는 양의 정수로서 반드시 똑같을 필요는 없다. 일반적인 동작을 설명하면, 컨퍼런스 노드(110)는, 회의실(160) 내의 일군의 화상 회의 참가자(102-1-s)로부터의 오디오 정보 및 비디오 정보를 캡처 또는 레코딩하는 것 및 회의실(160)로부터 원격지에 있는 다른 참가자들로, 컴퓨팅 장치(120) 및 네트워크(140)를 거쳐 일련의 원격 노드(150-1-m)로 레코딩된 신호를 전달하는 것 등의, 일반적인 화상 회의 서비스를 제공하는 동작을 할 수 있다.

[0016] 일 실시예에서, 예를 들어, 컨퍼런스 노드(110)는 미국 워싱턴주 레드몬드 소재의 마이크로소프트사에 의해 제작된 MICROSOFT® ROUNDTABLE 호환 장치로서 구현될 수 있다. MICROSOFT ROUNDTABLE 제품은 테이블톱 장치로서, 기저부에서 대략 종래의 스피커폰의 크기이다. 이 제품은 동기화된 음성 및 화상 회의를 제공하기 위해 컴퓨팅 장치(120) 등의 표준의 퍼스널 컴퓨터(PC)에 연결될 수 있다. MICROSOFT ROUNDTABLE 장치는 회의에 참여하고 있는 모든 사람의 연이은 이미지들의 360도 파노라마 비디오를 생성한다. 이 장치는 대화의 흐름을 추적하며, 그에 따라 발화 중인 사람의 이미지 및 음성이 주목을 받는다. 많은 장소들에 걸친 사람들이 가상적으로 함께 회의에 참석할 수 있다. MICROSOFT ROUNDTABLE 장치는, 예를 들어, 컴퓨팅 장치(120)를 통해 구현되는, MICROSOFT OFFICE COMMUNICATIONS SERVER 및 MICROSOFT OFFICE LIVE MEETING 제품을 비롯한 다른 화상 회의 장비 및 통신 장비와 상호연동하도록 설계되어 있다. 컨퍼런스 노드(110)가 MICROSOFT ROUNDTABLE 제품과 관련하여 설명될 수 있지만, 컨퍼런스 노드(110)가 동일한 또는 유사한 특징들을 갖는 임의의 적당한 오디오 및 비디오 장비로 구현될 수 있다는 것을 잘 알 것이다. 실시예들이 이와 관련하여 제한되지 않는다.

[0017] 일 실시예에서, 예를 들어, 마이크(104-1-r)는 참가자(102-1-s)에 의해 발생된 오디오 정보를 수신할 수 있는 지향성 또는 무지향성 마이크를 포함할 수 있다. 오디오 정보의 예는 사람의 음성, 회의실(160) 내의 참가자(102-1-s)에 의한 음성 세그먼트(speech segment) 또는 발화를 포함할 수 있다. 오디오 정보는, 도 6을 참조하여 보다 상세히 설명하게 되는 바와 같이, 원격 노드(150-1-m)로 전달될 수 있고 및/또는 다양한 유형의 메모리 요소에 레코딩될 수 있다. 마이크(104-1-r)는 또한, ASD 모듈(118)을 참조하여 더 상세히 설명하는 바와 같이, 회의실 내의 주도적 화자를 식별하여 분리시키는 데 도움을 주기 위해 사운드 빔형성(sound beamforming) 동작에 사용될 수 있다.

[0018] 일 실시예에서, 예를 들어, 비디오 카메라(106-1-p)는 참가자(102-1-s)의 비디오 정보를 수신 또는 캡처할 수 있는 지향성 또는 무지향성 카메라를 포함할 수 있다. 예를 들어, 하나 이상의 비디오 카메라(106-1-p)가 링 카메라(ring camera) 또는 "링캠(ringcam)"으로서 구현될 수 있다. 링캠은 360도의 비디오를 캡처하는 고해상도 무지향성 카메라이다. 링캠은 통상적으로 각각이, 예를 들어, 640 x 480 또는 800 x 600 등의 몇몇 원하는 수준의 해상도를 갖는 개별적인 또는 서로 다른 디지털 카메라들로 구성되어 있다. 각각의 카메라로부터의 이미지가 실시간으로 짜깁기되어 고해상도 파노라마를 형성한다. 이미지를 디지털적으로 짜깁기한 다음에 처리

장치(예를 들어, 제어기(112) 또는 컴퓨팅 장치(120))를 통해 피드함으로써 회의실(160) 전체로부터 이미지를 캡처할 수 있다.

[0019] 일 실시예에서, 컨퍼런스 노드(110)는 ASD 모듈(118)을 포함하고 있을 수 있다. ASD 모듈(118)은 오디오(사운드 소스 위치파악) 및 비디오(움직임 및 공간 패턴) 특징들을 사용하여 회의실(160) 내의 주도적 화자 또는 활발한 화자를 검출하도록 구성되어 있을 수 있다. ASD 모듈(118)은 몇 명의 사람들이 동시에 말하고 있을 때 회의실 내의 주도적 화자를 결정할 수 있다. ASD 모듈은 또한 배경 노이즈 및 소리를 반사하는 경표면(hard surface)을 보상한다. 예를 들어, ASD 모듈(118)은 서로 다른 소리를 구분하고 범형성이라고 하는 프로세스를 통해 우세한 소리를 분리하기 위해 6개의 개별 마이크(104-1-r)로부터 입력을 수신할 수 있다. 각각의 마이크(104-1-r)가 컨퍼런스 노드(110)의 서로 다른 부분에 내장되어 있다. 음속에도 불구하고, 마이크(104-1-r)는 서로에 대해 서로 다른 시간 간격으로 참여자(102-1-s)로부터 음성 정보를 수신할 수 있다. ASD 모듈(118)은 이 시간차를 사용하여 음성 정보의 소스를 식별할 수 있다. 음성 정보의 소스가 식별되면, 제어기(112)는 비디오 카메라(106-1-p)로부터의 시작적 단서를 사용하여 주도적 화자의 얼굴을 정확하게 찾아내고, 확대하여 강조할 수 있다.

[0020] 그러나, 일부 경우들에, 회의실(160) 내의 다양한 물체들이 ASD 모듈(118)의 정확도에 영향을 줄 수 있다. 예를 들어, 회의실(160)은 모니터(130) 등의 디스플레이 장치를 포함하고 있을 수 있다. 모니터(130)는 CRT(cathode ray tube) 모니터, LCD(liquid crystal display) 모니터, TFT(thin-film transistor) LCD 모니터, 프로젝터 스크린, 텔레비전, 디지털 텔레비전 등의 오디오 및 비디오 정보를 재현할 수 있는 임의의 디스플레이 장치를 나타낼 수 있다. 게다가, 모니터(130)는 참여자(102-1-s) 및/또는 참여자(108-1-m)을 비롯한 다양한 참여자들을 디스플레이할 수 있다. 예를 들어, 컴퓨팅 장치(120)는 모니터(130)를 사용하여 화상 전화 회의의 각각의 참여자를 디스플레이하는 MICROSOFT LIVE MEETING CONSOLE을 실행하는 클라이언트 모듈(122)을 포함하고 있을 수 있다. 사람들의 오디오 및 비디오 전송 이외에, 클라이언트 모듈(122)은 또한 문서, 컴퓨터-디스플레이 정보 및 화이트보드를 공유하는 데도 사용될 수 있다. 이러한 정보는 정리되어 모니터(130)에 의해 다양한 구성 및 배열의 다수의 창에 디스플레이될 수 있다.

[0021] 모니터(130)는 ASD 모듈(118)로 하여금 거짓 양성을 선택하게 할 수 있는 다양한 유형의 오디오 및 비디오 정보를 디스플레이한다. 예를 들어, 모니터(130)가 화상 전화 회의의 다양한 참여자들의 오디오 및 비디오 재현을 디스플레이하기 때문에, 모니터(130)가 주도적 화자의 오디오 정보 및 비디오 정보를 재현할 때, ASD 모듈(118)은 모니터(130)를 주도적 화자로서 잘못 식별할 수 있다. 이 결과 제어기(112)가 잘못하여 비디오 카메라(106)의 초점을 모니터(130)에 맞출 수도 있다. 실제로, 일부 경우들에, 로컬 참여자(102-1-s)가 회의실(160) 내의 주도적 화자일 수 있고 또한 동일한 회의실(160) 내의 모니터(130)를 통해 재현될 수 있으며, 그에 의해 ASD 모듈(118)로 하여금 실제의 주도적 화자(102-1-s) 대신에 모니터(130)를 선택하게 할 수 있다. 이를 및 기타 시나리오에서, 모니터(130)의 존재로 인해 ASD 모듈(118)의 정확도 및 성능이 저하될 수 있다.

[0022] 이를 및 기타 문제점을 해결하기 위해, 컨퍼런스 노드(110)는 모니터 검출 모듈(114)을 포함할 수 있다. 모니터 검출 모듈(114)은 비디오 카메라(106) 및 마이크(104)에 통신 연결될 수 있다. 모니터 검출 모듈(114)은 비디오 카메라(106)에 의해 회의실(160)을 찍은 이미지, 비디오 프레임, 또는 비디오 프레임 블록 내에 모니터(130)에 의해 디스플레이되는 시간 워터마크 신호를 검출하는 동작을 할 수 있다. 시간 워터마크 신호가 컨퍼런스 노드(110)의 시간 워터마크 발생기(116) 및 컴퓨팅 장치(120)의 믹서(124)를 통해 모니터(130)에 의해 디스플레이되는 비디오 신호에 추가될 수 있다. 모니터 검출 모듈(114)은 모니터 검출 동작에 기초하여 회의실(160) 내에서 모니터(130)의 위치를 결정할 수 있다. 모니터 검출 모듈(114)은 ASD 선택 동작들로부터 모니터(130)를 필터링 제거하기 위해 ASD 모듈(118)에 의해 사용될 수 있는 모니터 검출 신호를 ASD 모듈(118)로 출력할 수 있다.

[0023] 일 실시예에서, 컨퍼런스 노드(110)는 시간 워터마크 발생기(116)를 포함하고 있을 수 있다. 시간 워터마크 발생기(116)는 모니터(130)에 의해 디스플레이되는 시간 워터마크 신호를 발생하는 동작을 할 수 있다. 일반적으로, 시간 워터마크 신호는 실질적으로 사람에 의해 검출될 수 없어야만 하지만, 비디오 분석 기술을 사용하여 자동화된 시스템에 의해 안정적으로 검출될 수 있어야만 한다. 시간 워터마크 발생기(116)는 시간 워터마크 신호를 컴퓨팅 장치(120)의 믹서(124) 및 연결(114-2, 114-3)을 통해 모니터(130)로 출력할 수 있다. 도 1의 예시된 실시예들이 시간 워터마크 발생기를 컨퍼런스 노드(110)에 의해 구현되는 것으로 도시하고 있지만, 시간 워터마크 발생기(116)가 주어진 일련의 성능 또는 설계 제약조건을 위해 원하는 바에 따라 화상 회의 시스템(100)의 다른 요소들에서 구현될 수 있다는 것을 잘 알 것이다. 예를 들어, 시간 워터마크 발생기(116)는 컴퓨

텅 장치(120)의 일부로서 구현될 수 있고 여전히 본 실시예들의 의도된 범위 내에 속한다.

[0024] 일부 실시예들에서, 시간 워터마크 발생기(116)는 모니터(130)에 의해 디스플레이되는 시간 워터마크 신호로서 저주파수 패턴을 발생할 수 있다. 예를 들어, 시간 워터마크 발생기(116)는 모니터(130)에 의해 디스플레이되는 시간 워터마크 신호로서 주기함수 직류(DC) 오프셋 패턴을 발생할 수 있다. 램프, 사인파, 구형파 등의 임의의 주기적인 신호라도 사용될 수 있다. 일부 실시예들이 시간 워터마크 신호를 주기적인 램프 함수 직류 오프셋 패턴으로서 기술하고 있지만, 사인파 또는 기타 주기적인 신호가 램프 신호 대신에 사용될 수 있다. 시간 워터마크 신호가 실질적으로 사람에 의해서는 검출될 수 없고 자동화된 시스템에 의해서는 검출될 수 있는 한, 시간 워터마크 발생기(116)에 의해 발생되는 시간 워터마크 신호들의 유형이 원하는 구현에 따라 다를 수 있다.

[0025] 일 실시예에서, 예를 들어, 시간 워터마크 발생기(116)는 발생된 시간 워터마크 신호를 이하의 의사 코드에 따라 연결(114-2, 114-3)을 통해 컴퓨팅 장치(120)로 출력할 수 있다.

```
function oscillate_image()
im=imread('client.png');
max_offset=5;
for i=1:100
    for j=i:max_offset,max_offset-1:-1:1];
        im2=im+j;
        imshow(im2,'InitialMagnification','fit');
        title(sprintf('offset=%d',j));
        pause(.2);
    end
end
```

[0026]

[0027] 도 2는 그래프(200)의 일 실시예를 나타낸 것이다. 그래프(200)는 X-축에 대해 0초부터 10초까지 시간 간격을 갖고 Y-축에 대해 0값부터 5 값까지 루미넌스 오프셋 값을 갖는 그래프를 나타낸 것이다. 그래프(200)는 또한 모니터(130)에 대한 시간 워터마크 신호로서 사용하기에 적합한 주기적인 램프 함수 DC 오프셋 패턴 또는 신호(202)도 나타내고 있다. 모니터(130)는 통상적으로 다수의 픽셀을 디스플레이하고, 각각의 픽셀은 일정 범위 내의 픽셀값을 갖는다. 예를 들어, 8-비트 컬러 이미지인 경우, 각각의 픽셀은 0 내지 255의 값을 가질 수 있고, 각각의 픽셀값은 일정 범위의 컬러를 나타낸다. 주기적인 램프 함수 DC 오프셋 패턴은 값을 증가 또는 감소시키기 위해 하나 이상의 픽셀값을 수정함으로써 비디오 신호에 삽입될 수 있다. 예를 들어, 비디오 신호를 가정하고 모니터(130)는 YUV 모델을 사용하여 하나의 루마(luma) 성분 및 2개의 크로미넌스(chrominance) 성분으로 컬러 공간을 정의한다. YUV 모델에서, Y 성분은 루마 성분(회도)을 나타내고, U 및 V 성분은 크로미넌스 또는 컬러 성분을 나타낸다. 이 경우에, 그래프(500)의 각자의 X-축, Y-축에 도시된 바와 같이, 하나 이상의 픽셀값에 대한 Y 성분은 0초부터 10초까지 0부터 5까지 루미넌스 오프셋 값만큼 증가될 수 있다. 비디오 신호의 루미넌스 값이 주기적인 램프 함수 DC 오프셋 신호(202)에 따라 시간에 따라 변화될 때, 모니터(130)는 시간에 따라 회도 또는 Y 성분을 서서히 증가 및 감소시키게 된다. Y 성분의 증가 및 감소 속도는 사람의 인지를 감소 또는 예방하도록 제어되지만, 발진 신호로 모니터(130)를 갖는 비디오 이미지를 분석하는 모니터 검출 모듈(114)에 의해 안정적으로 검출될 수 있다. 화상 회의 동안, 비디오 카메라(106)는 시간 워터마크 신호를 갖는 이미지 또는 일련의 이미지를 레코딩하고, 그 이미지를 모니터 검출 모듈(114)로 출력한다. 모니터 검출 모듈(114)은 그 다음에 시간 워터마크 신호의 존재를 판정하고 그에 따라 모니터(130)의 위치를 결정하기 위해 픽셀 블록 단위로 이미지를 분석할 수 있다.

[0028] 도 3는 그래프(300-A, 300-B)의 일 실시예를 나타낸 것이다. 그래프(300-A, 300-B)는 각각 X-축에 대해 0개부터 300개 프레임까지 프레임 이력을 갖고 Y-축에 대해 231부터 236까지 루미넌스 값(Y)을 갖는 그래프를 나타낸 것이다. 그래프(300-A, 300-B)는 또한 도 2를 참조하여 기술한 것과 같은 내장된 신호(202)를 갖는 이미지를 디스플레이하는 모니터(130)에 대한 각자의 픽셀 이력 엔밸로프(302, 304)도 나타내고 있다. 픽셀 이력 엔밸로프(302, 304)는 주어진 AVI 파일에서 선택된 점의 픽셀 이력을 그리도록 디자인되어 있는 이하의 의사 코드를 비롯하여 다양한 기법을 사용하여 그려질 수 있다.

```
function pixel_slice(avi_file)
fileinfo=aviinfo(avi_file);
%n=fileinfo.NumFrames;
n=300;
mov=aviread(avi_file,1);
```

```

imshow(mov.cdata);
[x,y]=ginput(1);
x=round(x);
y=round(y);
for i=1:n
    mov=aviread(avi_file,i);
    Y=rgb2gray(mov.cdata);
    d=5;
    s(i)=mean(mean(Y(y-d:y+d,x-d:x+d)));
    sprintf('frame %d\n',i);
end
a=1;
m=5;
b=ones(1,m)/m;
s2=filter(b,a,s);
subplot(2,1,1);
plot(s(m:n));
title('pixel history');
ylabel('Y value');
subplot(2,1,2);
plot(s2(m:n));
title('5 point moving average');
xlabel('Frame #');
ylabel('Y value');
figure;
max_lag=60;
s3=s2(m:n)-mean(s2(m:n));
plot(-max_lag:max_lag,xcorr(s3,max_lag,'coeff'));
title('xcorr');
xlabel('Lag');
ylabel('Correlation');

```

[0030]

[0031]

도 3에 도시되어 있는 바와 같이, 주어진 픽셀 블록(예를 들어, 11 x 11)에 대해, 픽셀 이력 엔벨로프(302)는 다수의 주기적인 사이클로 231과 236 사이에서 서서히 발진하며, 각각의 주기적인 사이클은 대략 50 프레임마다 반복한다. 일부 경우들에, 픽셀 이력 엔벨로프(302)는, 예를 들어, 평탄화된 픽셀 이력 엔벨로프(304)를 생성하기 위해 5개 점의 이동 평균을 사용하여 평탄화될 수 있다. 평탄화된 픽셀 이력 엔벨로프(304)는 시간 워터마크 신호의 검출을 용이하게 하기 위해 사용될 수 있다.

[0032]

도 4는 그레프(400)의 일 실시예를 나타낸 것이다. 그레프(400)는 평탄화된 픽셀 이력 엔벨로프(304)에 대한 상관의 일례를 나타낸 것이다. 다양한 실시예들에서, 모니터 검출 모듈(114)은, 비디오 카메라(106)의 캡처된 이미지에서 0.5 Hz 신호(202)의 존재를 검출하기 위해, 평탄화된 픽셀 이력 엔벨로프(304)를 수신하고 수신된 신호를 교차-상관 기법을 사용하여 분석하도록 구성되어 있을 수 있다. 통상적으로, 검출 동작을 정확하게 수행하는 데 단지 몇 주기만 필요하다. 도 4는 모니터 검출 모듈(114)에 의해 수행되는 교차 상관 분석의 결과를 나타내는 상관 신호(402)를 나타낸 것이다. 교차 상관의 대안으로서, 내장된 시간 워터마크 신호를 검출하는 테 정합 필터링(matched filtering) 기법도 사용될 수 있다. 특정의 시간 워터마크 신호가 모니터 검출 기법에 맞게 조정되는 한, 주어진 구현에 사용하기에 적합한 워터마크 검출 기법의 수 및 유형이 상당히 다양할 수 있다. 실시예들이 이와 관련하여 제한되지 않는다.

[0033]

포화된(예를 들어, 백색) 또는 불포화(예를 들어, 흑색) 픽셀 블록이 검출하기 어려울 수 있지만 개선된 모폴로지(morphology) 및/또는 이미지 처리 기법들이 이를 특정의 영역을 채우는 데 사용될 수 있다는 것에 유의할 만한 가치가 있다. 또한, 모니터 검출 기법은 정확도를 향상시키기 위해 많은 시간(예를 들어, 60초)에 걸쳐 시간 필터링될 수 있다. 게다가, 모니터 검출 모듈(114)은 시간 필터링에서 거짓 음성(false negative)을 배제시키기 위해 AEC(automatic exposure control) 및/또는 AGC(automatic gain control)가 수렴될 때에만 모니터 검출을 수행할 수 있어야만 한다. 시간 필터링 기간이 비교적 제한된 기간이면 장치 이동이 꼭 특별한 취급을 필요

요로 하는 것은 아니다.

[0034] 일 실시예에서, 컨퍼런스 노드(110)는 ASD 모듈(118)을 포함하고 있을 수 있다. ASD 모듈(118)은 모니터 검출 모듈(114)에 통신 연결될 수 있다. ASD 모듈(118)은 화상 전화 회의의 참여자(102-1-s) 또는 참여자(108-1-m) 중에서 주도적 또는 활발하게 말하는 참여자가 모니터(130)에 의해 디스플레이되는지 여부를 판정하는 동작을 할 수 있다. ASD 모듈(118)은 모니터 검출 모듈(114)로부터 모니터 검출 신호를 수신할 수 있다. 모니터 검출 신호는 회의실(160) 내의 모니터(130)의 존재를 나타낼 수 있다. 게다가, 모니터 검출 신호는 모니터(130)의 근사 위치를 나타낼 수 있다. 모니터(130)의 위치가 주어지면, ASD 모듈(118)은 거짓 양성을 감소시키기 위해 모니터(130)의 위치에 있는 픽셀 블록을 사용할 수 있다. ASD 모듈(118)은 모니터 검출 결과를, 훈련 알고리즘 또는 하드코딩된 규칙 집합에의 입력으로서 사용할 수 있다. 하드코딩된 규칙의 일례는 검출된 화자가, 대략 25%보다 적은 모니터 픽셀과 같은, 몇몇 미리 정해진 문턱값보다 적은 모니터 픽셀을 가져야만 한다는 것일 수 있다. 주도적 화자가 어쩌면 모니터(130)의 전방에 서있을 수 있기 때문에, 전체 모니터 영역을 ASD 모듈(118) 선택 알고리즘으로부터 제거하는 것이 바람직하지 않을 수 있지만 가능할 수 있다. 전체 모니터 영역을 제거할 필요가 없도록 하기 위해, 모니터 검출 모듈(114) 및/또는 ASD 모듈(118)은 블록-기반 처리를 수행하도록 구성될 수 있다. 블록-기반 처리는 또한 계산 부하도 감소시킨다. 그 결과, ASD 모듈(118)은 모니터(130)에 의해 디스플레이되지 않은 전화 회의의 활발하게 말하는 참여자를 선택할 수 있다.

[0035] 화상 회의 시스템(100)의 동작들이 하나 이상의 논리 흐름을 참조하여 더 설명될 수 있다. 달리 언급하지 않는 한, 대표적인 논리 흐름들이 꼭 제공된 순서로 또는 임의의 특정의 순서로 실행되어야만 하는 것이 아님을 잘 알 것이다. 게다가, 논리 흐름들과 관련하여 기술된 다양한 활동들이 직렬로 또는 병렬로 실행될 수 있다. 논리 흐름들은, 주어진 일련의 설계 및 성능 제약조건에서 원하는 바에 따라, 화상 회의 시스템(100)의 하나 이상의 요소들 또는 대안의 요소들을 사용하여 구현될 수 있다.

[0036] 도 5는 논리 흐름(500)을 나타낸 것이다. 논리 흐름(500)은 본 명세서에 기술되어 있는 하나 이상의 실시예들에 의해 실행되는 동작들을 나타낼 수 있다. 논리 흐름(500)에 도시한 바와 같이, 논리 흐름(500)은 블록(502)에서 이미지 내에 모니터를 갖는 이미지 시퀀스를 수신할 수 있다. 논리 흐름(500)은 블록(504)에서 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출할 수 있다. 논리 흐름(500)은 블록(506)에서 검출에 기초하여 이미지 내에서의 모니터의 위치를 결정할 수 있다. 실시예들이 이와 관련하여 제한되지 않는다.

[0037] 일 실시예에서, 논리 흐름(500)은 블록(502)에서 이미지 내에 모니터를 갖는 이미지를 수신할 수 있다. 예를 들어, 비디오 카메라(106)는 회의실(160) 내의 참여자(102-1-s) 및 모니터(130)의 실시간 비디오 이미지를 캡처할 수 있다. 스트리밍 비디오가 무선 공유 매체(114-1) 또는 유선 통신 매체(114-2)를 거쳐 컨퍼런스 노드(110)로부터 컴퓨팅 장치(120)로 전달될 수 있다. 컴퓨팅 장치(120)는 로컬 참여자(102-1-s) 및 원격 참여자(108-1-m)를 비롯한, 화상 전화 회의의 참여자들의 이미지를 디스플레이하도록 화상 회의 소프트웨어를 구현하는 클라이언트 모듈(122)을 포함할 수 있다. 컴퓨팅 장치(120)는 또한 비디오 카메라(106)에 의해 캡처된 스트리밍 비디오를 네트워크(140)를 거쳐 원격 노드(150-1-m)로 전달할 수 있다. 네트워크(140)는 적절한 인터페이스 및 장비를 갖는 패킷-교환 네트워크, 회선-교환 네트워크, 또는 이 둘의 조합을 비롯한 임의의 유형의 적당한 네트워크를 포함할 수 있다.

[0038] 참여자(102-1-s) 각각이 다양한 때에 차례로 말을 할 때, 마이크(104) 및 ASD 모듈(118)이 다양한 사운스 소스 위치 파악 기법들을 사용하여 참여자들(102-1-s) 중에서 주도적 또는 활발한 화자를 결정하는 데 사용될 수 있다. 주도적 화자가 식별되고 위치파악되면, 제어기(112)는 비디오 카메라(106)로 하여금 자동으로 주도적 화자에 초점을 맞추게 할 수 있거나, 클라이언트 모듈(122)이 다른 방식으로 주도적 화자를 디스플레이하도록 그의 그래픽 사용자 인터페이스(GUI)를 수정할 수 있다. 후자의 경우의 일례는 다른 GUI 창들에 대해 주도적 화자를 디스플레이하는 GUI 창을 확대시키는 것, 주도적 화자를 디스플레이하는 GUI 창을 모니터(130) 상의 다른 위치로 이동시키는 것 등일 수 있다.

[0039] 일 실시예에서, 논리 흐름(500)은 블록(504)에서 모니터에 의해 디스플레이되는 시간 워터마크 신호를 검출할 수 있다. ASD 모듈(118)이 모니터(130)를 활발한 화자의 소스로서 식별하고 위치 파악하는 것을 감소시키거나 방지하기 위해, 모니터 검출 모듈(114)은 시간 워터마크 발생기(116)를 통해 모니터(130)에 의해 디스플레이되는 비디오 신호에 삽입되는 시간 워터마크 신호를 검출하기 위해 다양한 비디오 분석 기법들을 구현할 수 있다. 예를 들어, 시간 워터마크 발생기(116)는 발생된 시간 워터마크 신호를 연결(114-2, 114-3)을 통해 컴퓨팅 장치(120)로 출력할 수 있다. 컴퓨팅 장치(120)는 맥서(124)을 포함할 수 있다. 맥서(124)는 모니터(130)에 대한 시간 워터마크 신호 및 비디오 신호를 입력으로서 수신할 수 있다. 맥서(124)는 시간 워터마크 신호를 비디오

신호에 내장하기 위해 시간 워터마크 신호를 비디오 신호에 믹싱 또는 추가하도록 구성되어 있을 수 있다. 믹서(124)는 모니터(130)에 의한 차후의 디스플레이를 위해 믹싱된 신호를 모니터(130)의 비디오 카드로 출력할 수 있다.

[0040]

일 실시예에서, 예를 들어, 모니터 검출 모듈(114)은 이하의 의사 코드에 따라 모니터(130)에 의해 디스플레이 되는 시간 워터마크 신호를 검출할 수 있다.

```
function detect_monitor(avi_file)
n=300;
width=80;
height=60;
for i=1:n
    mov=aviread(avi_file,i);
    Y=rgb2gray(mov.cdata);
    s(i,:,:)=imresize(Y,[height width],'bilinear');
```

[0041]

```
fprintf('frame %d\n',i);
end
r=zeros(height,width);
for x=1:width
    for y=1:height
        S=detrend(double(s(:,y,x)));
        a=1;
        m=5;
        b=ones(l,m)/m;
        S2=filter(b,a,S);
        S2=S2(m:length(S2));
        max_lag=60;
        c=xcorr(S2,max_lag,'coeff');
        [max_val,max_i]=local_max(c);
        if length(max_i)~=3
            continue;
        end
        lag=max_i(3)-max_i(2);
        target_lag=39;
        lag_tol=4;
        if max_val(3)>0.5 && lag>=target_lag-lag_tol &&
        lag<=target_lag+lag_tol
            r(y,x)=1;
        end
    end
end
se=strel('square',3);
r=imerode(r,se);
r=imdilate(r,se);
width2=800;
height2=600;
r2=imresize(r,[height2 width2]);
mov=aviread(avi_file,l);
im=mov.cdata;
imwrite(im,'monitor.png');
```

[0042]

```

for x=1:width2
    for y=1:height2
        if r2(y,x)==1
            im(y,x,:)=[255,0,0];
        end
    end
end
imshow(r);
figure;
imshow(im);
imwrite(im,'monitor results.png');

function [max_val,max_i]=local_max(x)
% return the local maximums
max_val=x(1);
max_i=1;
n=length(x);
j=1;
for i=2:n-1
    if x(i)>x(i-1) && x(i)>x(i+1)
        max_i(j)=i;
        max_val(j)=x(i);
        j=j+1;
    end
end

```

[0043]

[0044]

일 실시예에서, 논리 호름(500)은 블록(506)에서 검출에 기초하여 이미지 내에서의 모니터의 위치를 결정할 수 있다. 예를 들어, 모니터(130)에 의해 디스플레이되는 시간 워터마크 신호가 모니터 검출 모듈(114)에 의해 검출되면, 모니터 검출 모듈(114) 및/또는 ASD 모듈(118)은 비디오 프레임을 분석하여 모니터(130)의 위치를 분리 시킬 수 있다. 모듈(114 및/또는 118)은 비디오 프레임 또는 일군의 비디오 프레임 내에서 발진이 일어나는 근사 영역을 결정하기 위해 블록마다 비디오 분석을 수행할 수 있다. 식별된 영역에 대응하는 픽셀 블록은 그 다음에 모니터 픽셀 블록으로서 태깅 또는 식별될 수 있다. ASD 모듈(118)은 그 다음에 모니터 픽셀 블록을 사용하여 ASD 선택 동작으로부터 모니터(130)를 필터링 제거할 수 있다. 예를 들어, ASD 모듈(118)은 전화 회의에서 활발하게 말하는 참여자가 모니터(130)에 의해 디스플레이되는지 여부를 판정하기 위해 모니터 픽셀 블록을 사용할 수 있다. ASD 모듈(118)은 그 다음에 모니터(130)에 의해 디스플레이되지 않은 화상 전화 회의의 활발하게 말하는 참여자를 선택할 수 있다.

[0045]

도 6은 화상 회의 시스템(100)을 비롯한 다양한 실시예들을 구현하는 데 적합한 컴퓨팅 시스템 아키텍처(600)의 블록도를 나타낸 것이다. 컴퓨팅 시스템 아키텍처(600)는 적합한 컴퓨팅 환경의 일례에 불과하며, 실시예들의 용도 또는 기능성의 범위에 관해 어떤 제한을 암시하고자 하는 것이 아님을 잘 알 것이다. 컴퓨팅 시스템 아키텍처(600)가 예시적인 컴퓨팅 시스템 아키텍처(600)에 도시된 컴포넌트들 중 임의의 하나 또는 그 컴포넌트들의 임의의 조합과 관련하여 어떤 의존성 또는 요구사항을 갖는 것으로 해석되어서도 안 된다.

[0046]

다양한 실시예들이 일반적으로 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터 실행 가능 명령어와 관련하여 기술될 수 있다. 일반적으로, 프로그램 모듈은 특정 동작을 수행하거나 특정 추상 데이터 유형을 구현하는 소프트웨어 요소를 포함한다. 일부 실시예들은 또한 통신 네트워크를 통해 연결되어 있는 하나 이상의 원격 처리 장치들에 의해 동작이 수행되는 분산 컴퓨팅 환경에서 실시될 수 있다. 분산 컴퓨팅 환경에서, 프로그램 모듈은 메모리 저장 장치를 비롯한 로컬 및 원격 컴퓨터 저장 매체 둘다에 위치할 수 있다.

[0047]

도 6에 도시된 바와 같이, 컴퓨팅 시스템 아키텍처(600)는 컴퓨터(610) 형태의 범용 컴퓨팅 장치를 포함한다. 컴퓨터(610)는 컴퓨터 또는 처리 시스템에서 통상적으로 발견되는 다양한 컴포넌트들을 포함할 수 있다. 컴퓨터(610)의 일부 예시적인 컴포넌트들은 처리 장치(620) 및 메모리 장치(630)를 포함할 수 있지만 이에 제한되는 것은 아니다.

[0048]

일 실시예에서, 예를 들어, 컴퓨터(610)는 하나 이상의 처리 장치(620)를 포함할 수 있다. 처리 장치(620)는 정보 또는 데이터를 처리하도록 구성되어 있는 임의의 하드웨어 요소 또는 소프트웨어 요소를 포함할 수 있다. 처리 장치(620)의 일부 예는 CISC(complex instruction set computer) 마이크로프로세서, RISC(reduced instruction set computing) 마이크로프로세서, VLIW(very long instruction word) 마이크로프로세서, 명령 세트들의 조합을 구현하는 프로세서, 또는 기타 프로세서 장치를 포함할 수 있지만 이에 제한되는 것은 아니다. 일 실시예에서, 예를 들어, 처리 장치(620)는 범용 프로세서로서 구현될 수 있다. 대안으로서, 처리 장치(620)는 컨트롤러, 마이크로컨트롤러, 임베디드 프로세서, DSP(digital signal processor), 네트워크 프로세서, 미디어 프로세서, 입/출력(I/O) 프로세서, MAC(media access control) 프로세서, 무선 기지대역 프로세서, FPGA(field programmable gate array), PLD(programmable logic device), ASIC(application specific integrated circuit) 등의 전용 프로세서로서 구현될 수 있다. 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0049]

일 실시예에서, 예를 들어, 컴퓨터(610)는 처리 장치(620)에 결합된 하나 이상의 메모리 장치(630)를 포함할 수 있다. 메모리 장치(630)는 정보 또는 데이터를 저장하도록 구성되어 있는 임의의 하드웨어 요소일 수 있다. 메모리 장치의 일부 예는 랜덤 액세스 메모리(RAM), DRAM(dynamic RAM), DDRAM(Double-Data-Rate DRAM), SDRAM(synchronous DRAM), SRAM(static RAM), 판독 전용 메모리(ROM), PROM(programmable ROM), EPROM(erasable programmable ROM), EEPROM, CD-ROM(Compact Disk ROM), CD-R(Compact Disk Recordable), CD-RW(Compact Disk Rewriteable), 플래쉬 메모리(예를 들어, NOR 또는 NAND 플래쉬 메모리), CAM(content addressable memory), 폴리머 메모리(polymer memory)(예를 들어, 강유전성 폴리머 메모리), 상변화 메모리(phase-change memory)(예를 들어, 오보닉 메모리(ovonic memory)), 강유전성 메모리(ferroelectric memory), SONOS(silicon-oxide-nitride-oxide-silicon) 메모리, 디스크(예를 들어, 플로피 디스크, 하드 디스크, 광 디스크, 자기 디스크, 광자기 디스크), 또는 카드(예를 들어, 자기 카드, 광 카드), 테이프, 카세트, 또는 컴퓨터(610)에 의해 액세스되고 원하는 정보를 저장하는 데 사용될 수 있는 임의의 기타 매체를 포함할 수 있지만 이에 제한되는 것은 아니다. 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0050]

일 실시예에서, 예를 들어, 컴퓨터(610)는 메모리 장치(630)를 비롯한 각종 시스템 컴포넌트들을 처리 장치(620)에 연결시키는 시스템 버스(621)를 포함할 수 있다. 시스템 버스(621)는 메모리 버스 또는 메모리 컨트롤러, 주변 장치 버스, 및 각종 버스 아키텍처 중 임의의 것을 이용하는 로컬 버스를 비롯한 몇몇 유형의 버스 구조 중 어느 것이라도 될 수 있다. 예로서, 이러한 아키텍처는 ISA(Industry Standard Architecture) 버스, MCA(Micro Channel Architecture) 버스, EISA(Enhanced ISA) 버스, VESA(Video Electronics Standards Association) 로컬 버스, 및 메자닌(Mezzanine) 버스라고도 하는 PCI(Peripheral Component Interconnect) 버스 등을 포함하지만 이에 제한되는 것은 아니다. 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0051]

다양한 실시예들에서, 컴퓨터(610)는 다양한 유형의 저장 매체를 포함할 수 있다. 저장 매체는 휘발성 또는 비휘발성 메모리, 이동식 또는 비이동식 메모리, 소거가능 또는 비소거가능 메모리, 기록가능 또는 재기록가능 메모리 등과 같이 데이터 또는 정보를 저장할 수 있는 저장 매체라면 어느 것이라도 될 수 있다. 저장 매체로는 컴퓨터 판독가능 매체 또는 통신 매체를 비롯한, 2가지 일반 유형이 있을 수 있다. 컴퓨터 판독가능 매체는 컴퓨팅 시스템 아키텍처(600)와 같은 컴퓨팅 시스템에 기록을 하거나 그로부터 판독을 하도록 구성되어 있는 저장 매체를 포함할 수 있다. 컴퓨팅 시스템 아키텍처(600)의 컴퓨터 판독가능 매체의 예는 ROM(631) 및 RAM(632)과 같은 휘발성 및/또는 비휘발성 메모리를 포함할 수 있지만 이에 제한되는 것은 아니다. 통신 매체는 통상적으로 반송파(carrier wave) 또는 기타 전송 메커니즘(transport mechanism)과 같은 피변조 데이터 신호(modulated data signal)에 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터 등을 구현하고 모든 정보 전달 매체를 포함한다. "피변조 데이터 신호"라는 용어는, 신호 내에 정보를 인코딩하도록 그 신호의 특성을 중 하나 이상을 설정 또는 변경시킨 신호를 의미한다. 예로서, 통신 매체는 유선 네트워크 또는 직접 배선 접속(direct-wired connection)과 같은 유선 매체, 그리고 음향, RF 스펙트럼, 적외선, 기타 무선 매체와 같은 무선 매체를 포함한다. 상술된 매체들의 모든 조합이 또한 컴퓨터 판독가능 매체의 영역 안에 포함되는 것으로 한다.

[0052]

다양한 실시예들에서, 메모리 장치(630)는 ROM(631) 및 RAM(632)과 같은 휘발성 및/또는 비휘발성 메모리 형태의 컴퓨터 저장 매체를 포함한다. 시동 중과 같은 때에, 컴퓨터(610) 내의 구성요소들 사이의 정보 전송을 돋는 기본 루틴을 포함하는 기본 입/출력 시스템(BIOS)(633)이 통상적으로 ROM(631)에 저장되어 있다. RAM(632)은 통상적으로 처리 장치(620)가 즉시 액세스할 수 있고 및/또는 현재 동작시키고 있는 데이터 및/또는 프로그램 모듈을 포함한다. 예로서, 도 6는 운영 체제(634), 애플리케이션 프로그램(635), 기타 프로그램 모듈(636),

및 프로그램 데이터(637)를 도시하고 있지만 이에 제한되는 것은 아니다.

[0053] 컴퓨터(610)는 또한 기타 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체를 포함할 수 있다. 단지 예로서, 도 6은 비이동식·비휘발성 자기 매체에 기록을 하거나 그로부터 판독을 하는 하드 디스크 드라이브(640), 이동식·비휘발성 자기 디스크(652)에 기록을 하거나 그로부터 판독을 하는 자기 디스크 드라이브(651), 및 CD-ROM 또는 기타 광 매체 등의 이동식·비휘발성 광 디스크(656)에 기록을 하거나 그로부터 판독을 하는 광 디스크 드라이브(655)를 도시하고 있다. 예시적인 운영 환경에서 사용될 수 있는 기타 이동식/비이동식, 휘발성/비휘발성 컴퓨터 기억 매체로는 자기 테이프 카세트, 플래쉬 메모리 카드, DVD(digital versatile disk), 디지털 비디오 테이프, 고상(solid state) RAM, 고상 ROM 등이 있지만 이에 제한되는 것은 아니다. 하드 디스크 드라이브(641)는 통상적으로 인터페이스(640)와 같은 비이동식 메모리 인터페이스를 통해 시스템 버스(621)에 접속되고, 자기 디스크 드라이브(651) 및 광 디스크 드라이브(655)는 통상적으로 인터페이스(650)와 같은 이동식 메모리 인터페이스에 의해 시스템 버스(621)에 접속된다.

[0054] 위에서 설명되고 도 6에 도시된 드라이브들 및 이들과 관련된 컴퓨터 저장 매체는 컴퓨터(610)에 대한 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 및 기타 데이터의 저장을 제공한다. 도 6에서, 예를 들어, 하드 디스크 드라이브(641)는 운영 체제(644), 애플리케이션 프로그램(645), 기타 프로그램 모듈(646), 및 프로그램 데이터(647)를 저장하는 것으로 도시되어 있다. 여기서 유의할 점은 이들 컴포넌트가 운영 체제(634), 애플리케이션 프로그램(635), 기타 프로그램 모듈(636), 및 프로그램 데이터(637)와 동일하거나 그와 다를 수 있다는 것이다. 이에 관해, 운영 체제(644), 애플리케이션 프로그램(645), 기타 프로그램 모듈(646), 및 프로그램 데이터(647)에 다른 번호가 부여되어 있는 것은 적어도 이들이 다른 사본(copy)이라는 것을 나타내기 위한 것이다. 사용자는 키보드(662) 및 통상 마우스, 트랙볼(trackball) 또는 터치 패드로 지침되는 포인팅 장치(661) 등의 입력 장치를 통해 명령 및 정보를 컴퓨터(610)에 입력할 수 있다. 다른 입력 장치(도시 생략)로는 마이크, 조이스틱, 게임 패드, 위성 안테나, 스캐너 등을 포함할 수 있다. 이들 및 기타 입력 장치는 종종 시스템 버스에 결합된 사용자 입력 인터페이스(660)를 통해 처리 장치(620)에 접속되지만, 병렬 포트, 게임 포트 또는 USB(universal serial bus) 등의 다른 인터페이스 및 버스 구조에 의해 접속될 수도 있다. 모니터(684) 또는 다른 유형의 디스플레이 장치도 비디오 인터페이스(682) 등의 인터페이스를 통해 시스템 버스(621)에 접속되어 있다. 모니터(684) 외에, 컴퓨터는 또한 스피커(687) 및 프린터(686) 등의 기타 주변 출력 장치를 포함할 수 있고, 이들은 출력 주변장치 인터페이스(683)를 통해 접속될 수 있다.

[0055] 컴퓨터(610)는 원격 컴퓨터(680)와 같은 하나 이상의 원격 컴퓨터로의 논리적 접속을 사용하여 네트워크화된 환경에서 동작할 수 있다. 원격 컴퓨터(680)는 퍼스널 컴퓨터(PC), 서버, 라우터, 네트워크 PC, 피어 장치 또는 기타 통상의 네트워크 노드일 수 있고, 통상적으로 컴퓨터(610)와 관련하여 상술된 구성요소들의 대부분 또는 그 전부를 포함하지만, 명확함을 위해 도 6에는 메모리 저장 장치(681)만이 도시되어 있다. 도 6에 도시된 논리적 접속으로는 LAN(671) 및 WAN(673)이 있지만, 기타 네트워크를 포함할 수도 있다. 이러한 네트워킹 환경은 사무실, 전사적 컴퓨터 네트워크(enterprise-wide computer network), 인트라넷 및 인터넷에서 일반적인 것이다.

[0056] LAN 네트워킹 환경에서 사용될 때, 컴퓨터(610)는 네트워크 인터페이스 또는 어댑터(670)를 통해 LAN(671)에 접속된다. WAN 네트워킹 환경에서 사용될 때, 컴퓨터(610)는 통상적으로 인터넷 등의 WAN(673)을 통해 통신을 설정하는 데 적합한 모뎀(672) 또는 기타 수단을 포함한다. 내장형 또는 외장형일 수 있는 모뎀(672)은 사용자 입력 인터페이스(660) 또는 기타 적절한 메커니즘을 통해 시스템 버스(621)에 접속될 수 있다. 네트워크화된 환경에서, 컴퓨터(610) 또는 그의 일부와 관련하여 기술된 프로그램 모듈은 원격 메모리 저장 장치에 저장될 수 있다. 예로서, 도 6는 원격 애플리케이션 프로그램(685)이 메모리 장치(681)에 있는 것으로 도시하고 있지만 이에 제한되는 것은 아니다. 도시된 네트워크 접속이 예시적인 것이며 이 컴퓨터들 사이에 통신 링크를 설정하는 기타 수단이 사용될 수 있다는 것을 이해할 것이다. 게다가, 네트워크 연결이 유선 또는 무선 연결로서 구현될 수 있다. 후자의 경우에, 컴퓨팅 시스템 아키텍처(600)는 하나 이상의 안테나, 송신기, 수신기, 송수신기, 무선기, 증폭기, 필터, 통신 인터페이스, 및 기타 무선 요소와 같은 무선 통신에 적합한 다양한 요소들로 수정될 수 있다. 무선 통신 시스템은, 예를 들어, 하나 이상의 RF 스펙트럼의 부분들 또는 대역들과 같은 무선 통신 매체를 통해 정보 또는 데이터를 주고받는다. 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0057] 화상 회의 시스템(100) 및/또는 컴퓨팅 시스템 아키텍처(600)의 일부 또는 전부가 전자 장치의 부품, 컴포넌트 또는 서브시스템으로서 구현될 수 있다. 전자 장치의 일례로는 처리 시스템, 컴퓨터, 서버, 워크스테이션, 가전기기, 단말기, 퍼스널 컴퓨터, 랩톱, 올트라-랩톱, 핸드헬드 컴퓨터, 미니컴퓨터, 메인프레임 컴퓨터, 분산 컴퓨팅 시스템, 멀티프로세서 시스템, 프로세서-기반 시스템, 가전제품, 프로그램가능 가전 제품, PDA, 텔레비

전, 디지털 텔레비전, 셋톱박스, 전화, 이동 전화, 셀룰러 전화, 핸드셋, 무선 액세스 포인트, 기지국, 가입자국, 이동전화 가입자 센터, 무선 네트워크 제어기, 라우터, 허브, 게이트웨이, 브리지, 스위치, 기계, 또는 이들의 조합이 있을 수 있지만, 이에 제한되는 것은 아니다. 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0058] 일부 경우들에, 다양한 실시예들이 제조 물품으로서 구현될 수 있다. 제조 물품은 하나 이상의 실시예들의 다양한 동작들을 수행하는 논리 및/또는 데이터를 저장하도록 구성되어 있는 저장 매체를 포함할 수 있다. 저장 매체의 일례로는 이전에 설명한 예들이 있을 수 있지만, 이에 제한되는 것은 아니다. 다양한 실시예들에서, 예를 들어, 제조 물품은 범용 프로세서 또는 전용 프로세서에서 실행하기에 적합한 컴퓨터 프로그램 명령어들이 들어 있는 자기 디스크, 광 디스크, 플래쉬 메모리 또는 펌웨어를 포함할 수 있다. 그러나, 실시예들이 이와 관련하여 제한되는 것은 아니다.

[0059] 다양한 실시예들이 하드웨어 요소, 소프트웨어 요소 또는 이 둘의 조합을 사용하여 구현될 수 있다. 하드웨어 요소의 일례로는 논리 장치에 대해 이전에 제공된 일례들이 있을 수 있으며, 또한 마이크로프로세서, 회로, 회로 요소(예를 들어, 트랜지스터, 저항기, 커패시터, 인덕터, 기타), 집적 회로, 논리 게이트, 레지스터, 반도체 장치, 칩, 마이크로칩, 칩셋 등이 있을 수 있다. 소프트웨어 요소의 일례로는 소프트웨어 컴포넌트, 프로그램, 애플리케이션, 컴퓨터 프로그램, 애플리케이션 프로그램, 시스템 프로그램, 기계 프로그램, 운영 체제 소프트웨어, 미들웨어, 펌웨어, 소프트웨어 모듈, 루틴, 서브루틴, 함수, 메서드, 프로시저, 소프트웨어 인터페이스, API(application program interface), 명령어 세트, 컴퓨팅 코드, 컴퓨터 코드, 코드 세그먼트, 컴퓨터 코드 세그먼트, 워드, 값, 심볼, 또는 이들의 임의의 조합이 있을 수 있다. 실시예를 하드웨어 요소 및/또는 소프트웨어 요소를 사용하여 구현할지를 결정하는 것은, 주어진 구현에 대해 원하는 바에 따라, 원하는 계산 속도, 전력 레벨, 내열성, 처리 사이클 비용(processing cycle budget), 입력 데이터 속도, 출력 데이터 속도, 메모리 자원, 데이터 버스 속도 및 기타 설계 또는 성능 제약조건 등의 다수의 인자들에 따라 변할 수 있다.

[0060] 일부 실시예들은 "결합" 및 "접속"이라는 표현을 다른 파생어들과 함께 사용하여 설명될 수 있다. 이를 용어가 서로에 대해 꼭 동의어인 것은 아니다. 예를 들어, 일부 실시예들이, 2개 이상의 구성요소가 서로 물리적 또는 전기적으로 직접 접촉하고 있음을 나타내기 위해, 용어 "접속" 및/또는 "결합"을 사용하여 설명될 수 있다. 그러나, 용어 "결합"은 또한 2개 이상의 구성요소가 서로 직접 접촉하고 있지 않지만 여전히 서로 협력하거나 상호작용하는 것도 의미할 수 있다.

[0061] 읽는 사람이 본 발명의 기술 내용을 신속하게 확인할 수 있게 하는 요약서를 요구하는 37 C.F.R. Section 1.72(b)에 부합하기 위해 요약서가 제공되어 있는 것을 강조해둔다. 제공된 요약서가 청구항들의 범위 또는 의미를 해석하거나 제한하는 데 사용되지 않는다는 것을 이해할 것이다. 게다가, 이상의 상세한 설명에서, 개시 내용을 간소화하기 위해 다양한 특징들이 하나의 실시예에 다 들어가 있다는 것을 잘 알 것이다. 이러한 개시 방법이 청구된 실시예가 각각의 청구항에 명시적으로 기재되어 있는 것보다 더 많은 특징들을 필요로 한다는 의사를 반영하는 것으로 해석되어서는 안 된다. 오히려, 이하의 청구 범위가 반영하고 있는 바와 같이, 발명의 청구 대상이 하나의 개시된 실시예의 모든 특징들보다 적은 것에 있다. 따라서, 이하의 청구 범위는 이로써 상세한 설명에 포함되는 것으로 보아야 하며, 각각의 청구항은 그것만으로 별개의 실시예이다. 첨부된 청구항들에서, 용어 "포함하는(comprising)" 및 "~을 특징으로 하는(wherein)"의 동등 어구로서 보통의 표현에서의 용어 "포함하는(including)" 및 "~을 특징으로 하는(in which)"이 사용된다. 게다가, 용어 "제1", "제2", "제3" 등은 단순히 라벨로서 사용되며, 그의 개체에 대한 수치적 제한을 두기 위한 것이 아니다.

[0062] 발명 대상이 구조적 특징들 및/또는 방법적 동작들과 관련하여 기술되어 있지만, 첨부된 청구항들에 정의된 발명 대상이 상기한 특정의 특징들 또는 동작들로 반드시 제한되는 것은 아니라는 것을 잘 알 것이다. 오히려, 상기한 특정의 특징들 및 동작들은 청구항들을 구현하는 예시적인 형태로서 개시되어 있다.

## 도면의 간단한 설명

[0005] 도 1은 화상 회의 시스템의 일 실시예를 나타낸 도면이다.

[0006] 도 2는 제1 그래프의 일 실시예를 나타낸 도면이다.

[0007] 도 3은 제2 그래프의 일 실시예를 나타낸 도면이다.

[0008] 도 4는 제3 그래프의 일 실시예를 나타낸 도면이다.

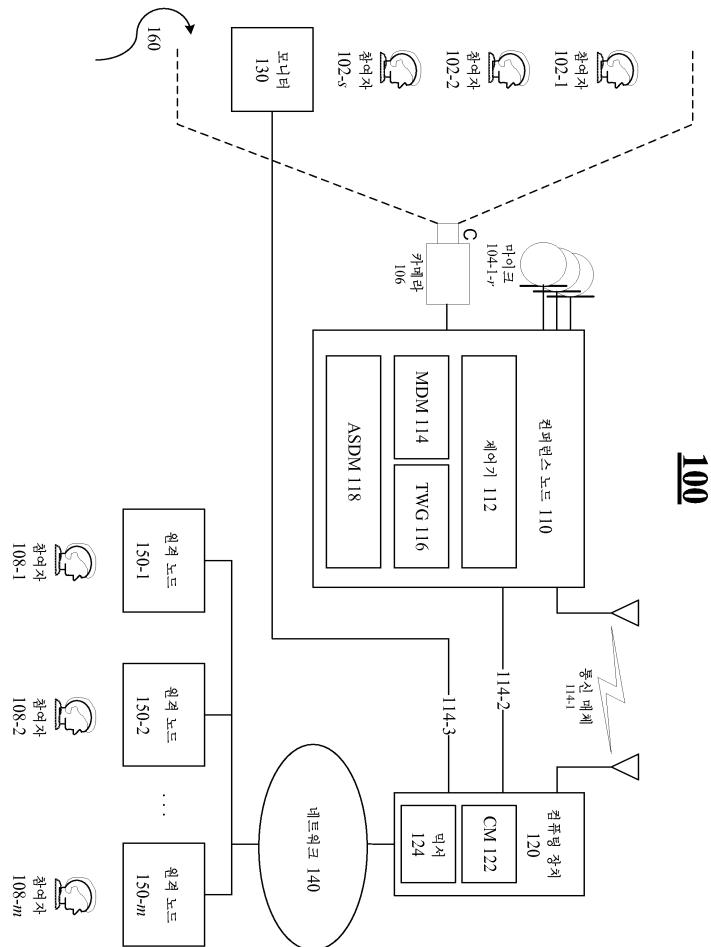
[0009] 도 5는 논리 흐름의 일 실시예를 나타낸 도면이다.

[0010]

도 6은 컴퓨팅 시스템 아키텍처의 일 실시예를 나타낸 도면이다.

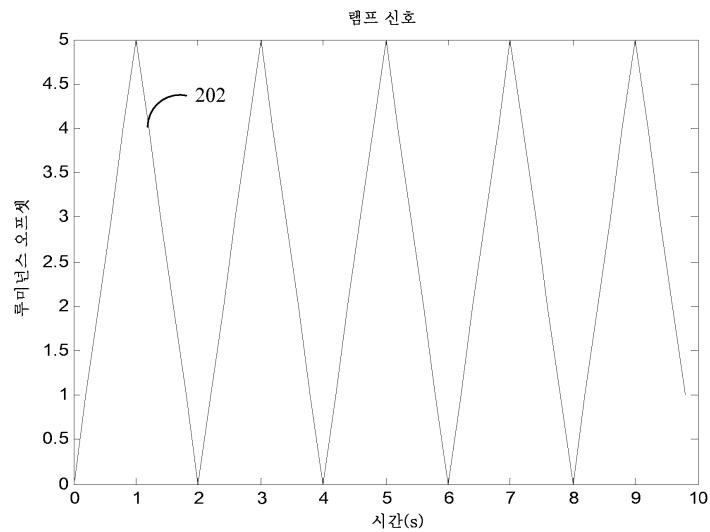
## 도면

### 도면1

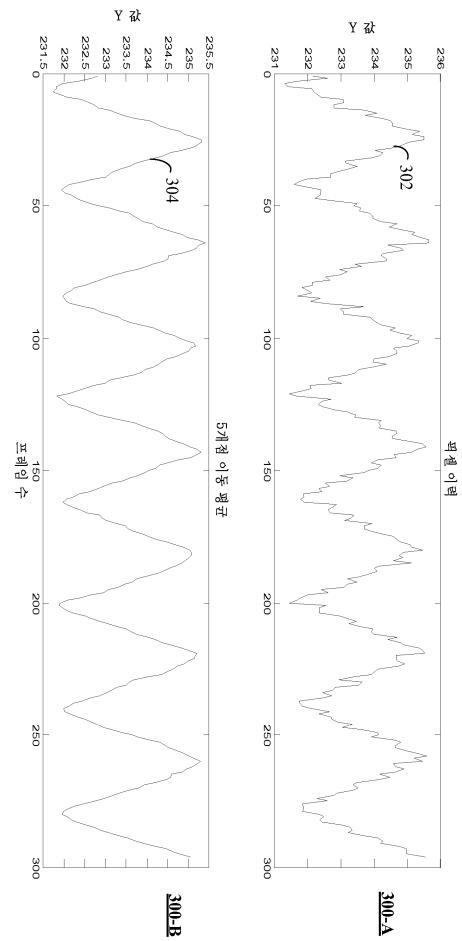


도면2

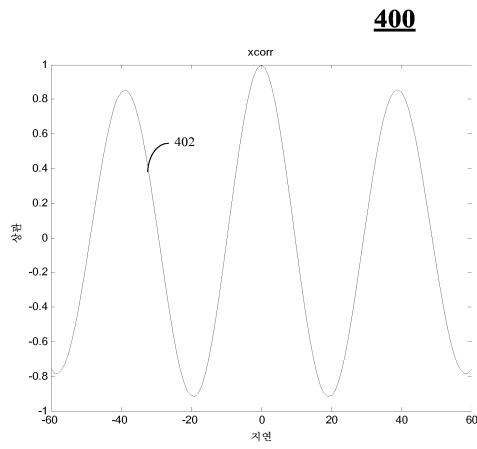
**200**



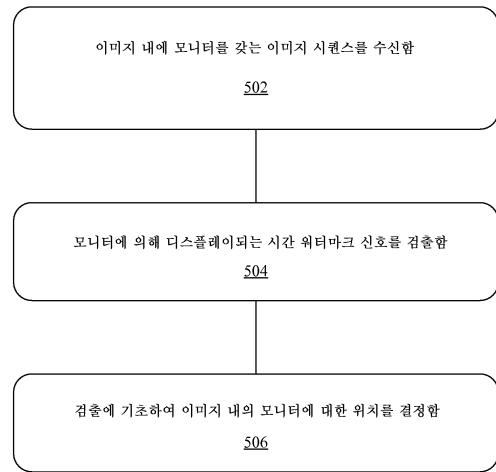
도면3



도면4



## 도면5

**500**

## 도면6

