



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2008년12월16일
 (11) 등록번호 10-0874226
 (24) 등록일자 2008년12월09일

(51) Int. Cl.
 H04N 7/12 (2006.01) H04N 13/00 (2006.01)
 H04N 7/24 (2006.01)
 (21) 출원번호 10-2007-0002824
 (22) 출원일자 2007년01월10일
 심사청구일자 2007년01월10일
 (65) 공개번호 10-2008-0065766
 (43) 공개일자 2008년07월15일
 (56) 선행기술조사문헌
 JP2005159592 A
 JP2000152216 A
 JP2002171529 A

(73) 특허권자
광주과학기술원
 광주 북구 오룡동 1번지
 (72) 발명자
김만배
 강원 춘천시 퇴계동 쌍용아파트 101-901
호요성
 광주 북구 오룡동 광주과기원 교원아파트 35-103
 (뒷면에 계속)
 (74) 대리인
특허법인우인

전체 청구항 수 : 총 14 항

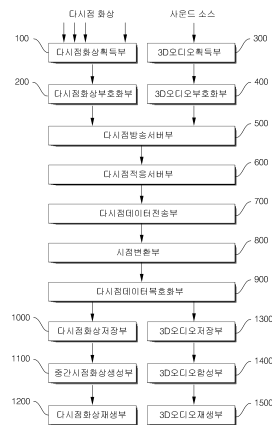
심사관 : 최성진

(54) 다시점 화상 및 3차원 오디오 송수신 장치 및 이를 이용한 송수신 방법

(57) 요약

본 발명은 다시점 화상 및 3차원 오디오 송수신 장치 및 이를 이용한 송수신 방법에 관한 것이다. 보다 상세하게는 다시점 화상과 3차원 오디오 데이터를 결합하되 시청자가 보는 시점화상에 따라 다른 오디오를 재생함으로써 보다 실감나는 미디어를 공급할 수 있고, 수신측에서 원하는 시점의 화상만을 선택적으로 전송할 수 있어 좁은 대역폭으로도 가능하며, 중간시점의 화상을 생성하여 보다 모니터 시차에 맞는 화상을 제공할 수 있는 다시점 화상 및 3차원 오디오 송수신 장치 및 이를 이용한 송수신 방법에 관한 것이다.

대표도 - 도1



(72) 발명자

김종원

광주 광산구 월계동 757-3 동부아파트 104동601호

변혜란

서울 서대문구 남가좌2동 삼성아파트(1차) 109동
602호

유지상

서울 강남구 대치동 미도아파트 211-105

김홍국

경기 수원시 영통구 영통동 신나무실 동보아파트
623-504

이관행

광주 북구 오룡동 1번지 교수아파트 C-302

류제하

광주 광산구 월계동 757-7 (14/1) 첨단금광아파트
101-706

이승현

경기 구리시 인창동 삼보아파트 308-1404

특허청구의 범위

청구항 1

다시점 화상 데이터와 3D 오디오 데이터를 송수신 처리하는 장치에 있어서,

부호화된 다시점 화상 데이터 일부를 다중화하는 다시점 화상 다중화부와, 부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화부를 포함하는 다시점 방송 서버부;

수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 다시점 적응 서버부;

카메라로부터 각 시점 화상의 데이터를 가져 오는 획득 서버부와, 수신측에서 요구된 시점을 선택적으로 전송하도록 제어 명령을 내리는 제어 서버부와, 상기 제어 서버부로부터 명령받은 시점 화상의 데이터를 상기 획득 서버부로부터 가져와서 비압축 데이터 형태로 전송하는 IP(Internet Protocol)멀티캐스트 처리부를 구비하여 비압축 데이터 형태로 전송하는 다시점 데이터 전송부;

시청자가 요구한 시점에 따라, 시점 화상의 기술자(descriptor)와 3D 오디오의 기술자를 생성하는 시점 변환부; 및

복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 중간시점화상 생성부를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 2

제1항에 있어서, 상기 다시점 적응 서버부는

상기 시점 변환부로부터 전송된 시점 기술자(view descriptor)를 파싱(parsing)하는 시점 기술자 파싱부;

상기 시점 기술자 파싱부를 통해 파싱된 다시점 리소스(resource)를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 리소스 Rv를 생성하는 화상 리소스 적응부;

상기 시점 변환부로부터 전송된 다시점 화상 기술자를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 기술자 Dv를 생성하는 화상 기술자 적응부;

상기 시점 기술자 파싱부를 통해 파싱된 오디오 리소스를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 리소스 Ra를 생성하는 오디오 리소스 적응부; 및

상기 시점 변환부로부터 전송된 오디오 기술자를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 기술자 Da를 생성하는 오디오 기술자 적응부

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 3

제1항에 있어서, 상기 시점 변환부는

전송받은 복수개의 시점 화상으로부터 시청자가 원하는 적어도 하나의 시점을 선택하는 시점 결정부; 및

상기 시점 결정부로부터 선택된 시점의 기술자를 생성하여 상기 다시점 적응 서버부로 전송하는 시점 기술자 생성부

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 4

제1항에 있어서, 상기 중간시점화상 생성부는

카메라의 외부 및 내부 파라미터를 계산하는 카메라 파라메타 계산부;

이웃하는 시점 화상간의 매칭점을 찾아 초기 깊이정보를 찾아내는 깊이정보 계산부;

이웃하는 시점 화상간에 기준 화상을 정하고, 상기 기준 화상에 대한 색상 정보를 바탕으로 서로 다른 색상정보를 가지는 영역으로 상기 이웃하는 시점 화상들의 영역을 분할하는 영역 분할부;

상기 영역 분할부를 통해 색상별로 분할된 영역에 상기 깊이정보 계산부를 통해 계산된 초기 깊이 정보를 반영하여 깊이 계층을 생성하되, 동일한 영역에 대해 대표 깊이정보를 할당하는 작업을 모든 영역에 반복하고 상기 대표 깊이정보를 그룹화한 후 그룹 내에서 대표 깊이값 하나를 추출하여 깊이 계층을 생성하는 깊이계층 예측부;

상기 깊이계층 예측부로부터 생성된 깊이 계층에 반복적으로 깊이할당을 수행하는 깊이정보 개선부; 및

상기 깊이정보 개선부를 통해 최적으로 생성된 깊이 정보와, 상기 카메라 파라메타 계산부를 통해 계산된 카메라 파라메타를 이용하여 임의 시점의 화상을 생성하는 중간시점 제작부

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 5

제1항에 있어서,

복수개의 카메라로부터 다시점 화상을 획득하는 다시점 화상 획득부;

상기 다시점 화상 획득부를 통해 얻은 다시점 화상 데이터를 부호화하는 다시점 화상 부호화부;

복수개의 마이크로폰으로부터 다채널 3D 오디오를 획득하는 3D 오디오 획득부;

상기 3D 오디오 획득부를 통해 얻은 3D 오디오 데이터를 부호화하는 3D 오디오 부호화부

를 더 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 6

삭제

청구항 7

제1항에 있어서, 상기 다시점 데이터 전송부는

상기 획득 서버부로부터 각 시점 화상 및 3D 오디오의 생성 시간 정보를 포함한 모든 시점 데이터를 받아 오는 시점 데이터 모음부; 및

상기 생성 시간 정보를 이용하여 상기 시점 데이터를 동기 다중화하는 동기 다중화부

를 구비하여 압축 데이터 형태로 더 전송하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 8

제1항에 있어서,

수신된 n 개의 다시점 부호화 데이터 중에서 j 개의 시점을 복호화하는 j 시점 복호화부;

상기 j 시점 복호화부를 통해 복호화된 데이터를 하나로 모으는 시점 데이터 모음부; 및

수신된 다채널 오디오 데이터를 복호화하는 3D 오디오 복호화부

를 구비하는 다시점 데이터 복호화부를 더 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치 (단, $n \geq j$).

청구항 9

제8항에 있어서,

상기 시점 데이터 모음부를 통해 모인 데이터를 각각의 시점별로 프레임 단위의 시점화상을 생성하기 위해 저장하는 다시점 화상 저장부;

다시점 화상과 동기화하고 3D 오디오 합성을 위해 복호화된 3D 오디오 데이터를 저장하는 3D 오디오 저장부;

시청자에 의해 선택된 시점 화상에 일치하는 3D 오디오를 상기 3D 오디오 저장부로부터 가져와서 합성하는 3D 오디오 합성부; 및

상기 3D 오디오 합성부를 통해 합성된 3D 오디오 데이터를 다채널 스피커 또는 스테레오 스피커를 통해 재생하는 3D 오디오 재생부

를 더 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 10

제9항에 있어서, 상기 3D 오디오 합성부는

시청자의 시점 선택에 따라 서라운드 패닝(surround panning)을 통해 선택된 시점 화상에 일치하는 3D 오디오를 합성하는 서라운드 패닝부;

스테레오 재생시 3D 효과를 주기위해 다채널 오디오 데이터로부터 3D 파라미터를 추출하는 3D 파라미터 추출부; 및

상기 3D 파라미터 추출부를 통해 추출된 3D 파라미터를 이용하여 스테레오 오디오 데이터를 생성하는 다운 믹싱부

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 장치.

청구항 11

다시점 화상 데이터와 3D 오디오 데이터를 송수신 처리하는 방법에 있어서,

(a) 부호화된 다시점 화상 데이터 일부를 다중화하는 다시점 화상 다중화 단계;

(b) 부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화 단계;

(c) 시청자가 요구한 시점에 따라, 시점 화상의 기술자(descriptor)와 3D 오디오의 기술자를 생성하는 단계;

(d) 수신측에서 요구된 시점 화상을 비압축 데이터 형태로 전송하는 비압축 데이터 전송단계 및 모든 시점 화상과 이에 대응되는 3D 오디오를 하나의 데이터스트림(data stream)으로 다중화하여 압축 데이터 형태로 전송하는 압축 데이터 전송단계를 포함하는 다시점 데이터 전송단계;

(e) 수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 단계; 및

(f) 복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 단계

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 방법.

청구항 12

제11항에 있어서, 상기 (e)단계는

상기 (c)단계를 통해 전송된 시점 기술자(view descriptor)를 파싱(parsing)하는 시점 기술자 파싱 단계;

상기 시점 기술자 파싱 단계를 통해 파싱된 다시점 리소스(resource)를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 리소스 Rv를 생성하는 화상 리소스 적응 단계;

상기 (c)단계를 통해 전송된 다시점 화상 기술자를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 기술자 Dv를 생성하는 화상 기술자 적응 단계;

상기 시점 기술자 파싱 단계를 통해 파싱된 오디오 리소스를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 리소스 Ra를 생성하는 오디오 리소스 적응 단계; 및

상기 (c)단계를 통해 전송된 오디오 기술자를 수신측 단말 특성 또는 시청자의 재생 취향 특성에 적응 변환하여 적응 기술자 Da를 생성하는 오디오 기술자 적응 단계

를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 방법.

청구항 13

제11항에 있어서, 상기 (c)단계는

전송받은 복수개의 시점 화상으로부터 시청자가 원하는 적어도 하나의 시점을 선택하는 단계; 및

선택된 시점의 기술자를 생성하여 전송하는 단계
를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 방법.

청구항 14

제11항에 있어서, 상기 (f)단계는
카메라의 외부 및 내부 파라미터를 계산하는 단계;
이웃하는 시점 화상 간의 매칭점을 찾아 초기 깊이 정보를 찾아내는 단계;
이웃하는 시점 화상 간에 기준 화상을 정하고, 상기 기준 화상에 대한 색상 정보를 바탕으로 서로 다른 색상 정보를 가지는 영역으로 상기 이웃하는 시점 화상들의 영역을 분할하는 단계;
색상별로 분할된 영역에 상기 초기 깊이 정보를 반영하여 깊이 계층을 생성하되, 동일한 영역에 대해 대표 깊이 정보를 할당하는 작업을 모든 영역에 반복하고 상기 대표 깊이정보를 그룹화한 후 그룹 내에서 대표 깊이값 하나를 추출하여 깊이 계층을 생성하는 단계;
상기 깊이 계층에 반복적으로 깊이 할당을 수행하는 단계; 및
상기 깊이 할당을 수행하는 단계를 통해 최적으로 생성된 깊이 정보와, 계산된 카메라 파라미터를 이용하여 임의 시점의 화상을 생성하는 단계
를 포함하는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 방법.

청구항 15

삭제

청구항 16

제11항에 있어서, 상기 압축 데이터 전송단계에서 상기 하나의 데이터 스트림으로의 다중화는
각 시점 화상 및 3D 오디오의 생성 시간 정보를 이용하여 동기화된 형태로 수행되는 것을 특징으로 하는 다시점 화상 및 3D 오디오 송수신 방법.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

- <19> 본 발명은 다시점 화상 및 3차원(이하, "3D") 오디오 송수신 장치 및 이를 이용한 송수신 방법에 관한 것이다. 보다 상세하게는 다시점 화상과 3D 오디오 데이터를 결합하되 시청자가 보는 시점화상에 따라 다른 오디오를 재생함으로써 보다 실감나는 미디어를 공급할 수 있고, 수신측에서 원하는 시점의 화상만을 선택적으로 전송할 수 있어 좁은 대역폭으로도 가능하며, 중간시점의 화상을 생성하여 보다 모니터 시차에 맞는 화상을 제공할 수 있는 다시점 화상 및 3D 오디오 송수신 장치 및 이를 이용한 송수신 방법에 관한 것이다.
- <20> 다시점(multi-view) 비디오는 평행 또는 아크 모양으로 배열된 동기화된 많은 개수의 카메라(예를 들면 8개)로 동일한 물체를 촬영해서 얻은 시점 화상(view image)의 집합이다. 이와 같은 다시점 비디오는 입체 디스플레이 장치뿐만 아니라 입체 방송, 실감 방송, 3D DMB 방송, FTV(Free-view TV) 등에서 사용자가 원하는 시점에서 시청하거나 콘텐츠를 3D 입체화상으로 보고자 할 때도 적용될 수 있는 응용의 폭이 넓은 기술이다.
- <21> 이러한 다시점 화상의 처리에 관한 요소 기술로는 획득 기술, 모델링/렌더링 기술, 부호화/복호화 기술 및 전송 기술 등이 있다.
- <22> 모델링/렌더링 기술은 다수개의 카메라를 통해 특정 객체(object)를 모델링하는 기법으로, 사용자는 임의의 시점에서 객체를 돌려볼 수 있다. 예를 들어, 일본 NHK 기술연구소의 "HD 카메라를 이용한 인체모델링 및 임의시점 영상 생성기술", 독일 MPI-infomatik의 "Free-viewpoint video기술" 등이 제안되었다. 모델링/렌더링 기술

은 송수신 개념과는 별도로 개발된 기술로, 순수 영상처리분야에서 사용되고 있다.

- <23> 한편, 다시점 동영상 기반의 부호화 및 전송기술로는 "Eye Vision"이 있다. Eye Vision은 예를 들어 스포츠 경기장에서 50대 정도의 카메라를 설치하고 임의의 시간동안 특정 객체에 대하여 사용자에게 360도 뷰를 생성하여 보여준다. 이 기술에서는 운용자가 임의의 시점을 선택하여 방송으로 송신하게 되므로, 시청자는 운용자가 전송해 주는 대로 시청할 수밖에 없다는 문제점이 있다.
- <24> 또한, 지금까지의 다시점 화상처리시스템은 다시점 화상 데이터와 오디오 데이터의 결합을 고려하지 않거나, 이를 고려하였다 하더라도 단일의 오디오 데이터와 다중화함으로써, 시청자가 어떤 시점의 화상을 시청하더라도 확실적인 사운드만 들려주어 현장감이 떨어진다는 문제점이 있다. 더욱이, 사용자가 시점을 선택할 수 있는 경우, 원하는 특정의 시점 화상에 맞추어 오디오 데이터가 실시간으로 적응하지 못한다는 문제점이 있다.
- <25> 또한, 기존의 다시점 화상처리시스템은 여러 대의 카메라로부터 획득한 다양한 시점 화상들 모두를 전송하고, 수신측에서 시청자가 원하는 시점 화상을 선택하는 방식이므로, 데이터의 전송에 있어 엄청난 대역폭을 차지할 수밖에 없는 문제점이 있다.
- <26> 또한, 기존의 다시점 화상처리시스템은 획득된 시점 화상 전부를 송신하였다 하더라도 수신측에서는 카메라 개수만큼의 시점 화상만 시청할 수 있다. 따라서, 수신측 3D 모니터의 시점 간격에 맞출 수 있는 충분한 시점 화상이 존재하지 않아 시점 화상간의 차이가 크고, 시점간의 부드러운 변화를 제공할 수 없다는 문제점이 있다.

발명이 이루고자 하는 기술적 과제

- <27> 본 발명은 상기와 같은 과제를 해결하기 위해 안출된 것으로, 특히 다시점 화상과 3D 오디오 데이터를 결합하되 시청자가 보는 시점화상에 따라 다른 오디오를 재생함으로써 보다 실감나는 미디어를 공급할 수 있고, 수신측에서 원하는 시점의 화상만을 선택적으로 전송할 수 있어 좁은 대역폭으로도 가능하며, 중간시점의 화상을 생성하여 보다 모니터 시차에 맞는 화상을 제공할 수 있는 다시점 화상 및 3D 오디오 송수신 장치 및 이를 이용한 송수신 방법을 제공하는 데 그 목적이 있다.

발명의 구성 및 작용

- <28> 상기 목적을 달성하기 위해 안출된 본 발명에 따른 다시점 화상 및 3D 오디오 송수신 장치는 다시점 화상 데이터와 3D 오디오 데이터를 송수신 처리하는 장치에 있어서, 부호화된 다시점 화상 데이터 일부를 다중화하는 다시점 화상 다중화부와, 부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화부를 포함하는 다시점 방송 서버부; 수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 다시점 적응 서버부; 시청자가 요구한 시점에 따라, 시점 화상의 기술자(descriptor)와 3D 오디오의 기술자를 생성하는 시점 변환부; 및 복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 중간시점화상 생성부를 포함하는 것을 특징으로 한다.
- <29> 또한, 상기 다시점 적응 서버부는 상기 시점 변환부로부터 전송된 시점 기술자(view descriptor)를 파싱(parsing)하는 시점 기술자 파싱부; 상기 시점 기술자 파싱부로부터 다시점 리소스(resource)를 적응하는 화상 리소스 적응부; 상기 화상 리소스 적응부로부터 다시점 화상 기술자를 적응하는 화상 기술자 적응부; 상기 시점 기술자 파싱부로부터 오디오 리소스를 적응하는 오디오 리소스 적응부; 및 상기 오디오 리소스 적응부로부터 오디오 기술자를 적응하는 오디오 기술자 적응부를 포함할 수 있다.
- <30> 또한, 상기 시점 변환부는 전송받은 복수개의 시점 화상으로부터 시청자가 원하는 적어도 하나의 시점을 선택하는 시점 결정부; 및 상기 시점 결정부로부터 선택된 시점의 기술자를 생성하여 상기 다시점 적응 서버부로 전송하는 시점 기술자 생성부를 포함할 수 있다.
- <31> 또한, 상기 중간시점화상 생성부는 카메라의 외부 및 내부 파라미터를 계산하는 카메라 파라메타 계산부; 이웃하는 시점 화상간의 매칭점을 찾아 초기 깊이정보를 찾아내는 깊이정보 계산부; 이웃하는 시점 화상간에 기준 화상을 정하고, 상기 기준 화상에 대한 색상 정보를 바탕으로 서로 다른 색상정보를 가지는 영역으로 분할하는 영역 분할부; 상기 영역 분할부를 통해 색상별로 분할된 영역에 상기 깊이정보 계산부를 통해 계산된 초기 깊이정보를 반영하여 깊이 계층을 생성하는 깊이계층 예측부; 상기 깊이계층 예측부로부터 생성된 깊이 계층에 반복적으로 깊이할당을 수행하는 깊이정보 개선부; 및 상기 깊이정보 개선부를 통해 최적으로 생성된 깊이 정보와, 상기 카메라 파라메타 계산부를 통해 계산된 카메라 파라메타를 이용하여 임의 시점의 화상을 생성하는 중간시점 제작부를 포함할 수 있다.

- <32> 또한, 상기 다시점 화상 및 3D 오디오 송수신 장치는복수개의 카메라로부터 다시점 화상을 획득하는 다시점 화상 획득부; 상기 다시점 화상 획득부를 통해 얻은 다시점 화상 데이터를 부호화하는 다시점 화상 부호화부; 복수개의 마이크로폰으로부터 다채널 3D 오디오를 획득하는 3D 오디오 획득부; 상기 3D 오디오 획득부를 통해 얻은 3D 오디오 데이터를 부호화하는 3D 오디오 부호화부를 더 포함할 수 있다.
- <33> 또한, 상기 다시점 화상 및 3D 오디오 송수신 장치는 카메라로부터 각 시점 화상의 데이터를 가져 오는 획득 서버부; 수신측에서 요구된 시점을 선택적으로 전송하도록 제어 명령을 내리는 제어 서버부; 및 상기 제어 서버부로부터 명령받은 시점 화상의 데이터를 상기 획득 서버부로부터 가져와서 비압축 데이터 형태로 전송하는 IP(Internet Protocol)멀티캐스트 처리부를 구비하여 비압축 데이터 형태로 전송하는 다시점 데이터 전송부를 더 포함할 수 있다.
- <34> 또한, 상기 다시점 데이터 전송부는 상기 획득 서버부로부터 각 시점 화상 및 3D 오디오의 생성 시간 정보를 포함한 모든 시점 데이터를 받아 오는 시점 데이터 모음부; 및 상기 생성 시간 정보를 이용하여 상기 시점 데이터를 동기 다중화하는 다시점 다중화부를 구비하여 압축 데이터 형태로 더 전송할 수 있다.
- <35> 또한, 상기 다시점 화상 및 3D 오디오 송수신 장치는 수신된 n개의 다시점 부호화 데이터 중에서 j개의 시점을 복호화하는 j시점 복호화부; 상기 j시점 복호화부를 통해 복호화된 데이터를 하나로 모으는 시점 데이터 모음부; 및 수신된 다채널 오디오 데이터를 복호화하는 3D 오디오 복호화부를 구비하는 다시점 데이터 복호화부를 더 포함할 수 있다(단, $n \geq j$).
- <36> 또한, 상기 다시점 화상 및 3D 오디오 송수신 장치는 상기 시점 데이터 모음부를 통해 모인 데이터를 각각의 시점별로 프레임 단위의 시점화상을 생성하기 위해 저장하는 다시점 화상 저장부; 다시점 화상과 동기화하고 3D 오디오 합성을 위해 복호화된 3D 오디오 데이터를 저장하는 3D 오디오 저장부; 시청자에 의해 선택된 시점 화상에 일치하는 3D 오디오를 상기 3D 오디오 저장부로부터 가져와서 합성하는 3D 오디오 합성부; 및 상기 3D 오디오 합성부를 통해 합성된 3D 오디오 데이터를 다채널 스피커 또는 스테레오 스피커를 통해 재생하는 3D 오디오 재생부를 더 포함할 수 있다.
- <37> 또한, 상기 3D 오디오 합성부는 시청자의 시점 선택에 따라 서라운드 패닝(surround panning)을 통해 선택된 시점 화상에 일치하는 3D 오디오를 합성하는 서라운드 패닝부; 스테레오 재생시 3D 효과를 주기위해 다채널 오디오 데이터로부터 3D 파라미터를 추출하는 3D 파라미터 추출부; 및 상기 3D 파라미터 추출부를 통해 추출된 3D 파라미터를 이용하여 스테레오 오디오 데이터를 생성하는 다운 믹싱부를 포함할 수 있다.
- <38> 또한, 본 발명에 따른 다시점 화상 및 3D 오디오 송수신 방법은 다시점 화상 데이터와 3D 오디오 데이터를 송수신 처리하는 방법에 있어서, (a) 부호화된 다시점 화상 데이터 일부를 다중화하는 다시점 화상 다중화 단계; (b) 부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화 단계; (c) 시청자가 요구한 시점에 따라, 시점 화상의 기술자(descriptor)와 3D 오디오의 기술자를 생성하는 단계; (d) 수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 단계; 및 (e) 복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 단계를 포함하는 것을 특징으로 한다.
- <39> 또한, 상기 (d)단계는 상기 (c)단계를 통해 전송된 시점 기술자(view descriptor)를 파싱(parsing)하는 단계; 다시점 리소스(resource)를 적응하는 단계; 다시점 화상 기술자를 적응하는 단계; 오디오 리소스를 적응하는 단계; 및 오디오 기술자를 적응하는 단계를 포함할 수 있다.
- <40> 또한, 상기 (c)단계는 전송받은 복수개의 시점 화상으로부터 시청자가 원하는 적어도 하나의 시점을 선택하는 단계; 및 선택된 시점의 기술자를 생성하여 전송하는 단계를 포함할 수 있다.
- <41> 또한, 상기 (e)단계는 카메라의 외부 및 내부 파라미터를 계산하는 단계; 이웃하는 시점 화상 간의 매칭점을 찾아 초기 깊이 정보를 찾아내는 단계; 이웃하는 시점 화상 간에 기준 화상을 정하고, 상기 기준 화상에 대한 색상 정보를 바탕으로 서로 다른 색상 정보를 가지는 영역으로 분할하는 단계; 색상별로 분할된 영역에 상기 초기 깊이 정보를 반영하여 깊이 계층을 생성하는 단계; 상기 깊이 계층에 반복적으로 깊이 할당을 수행하는 단계; 및 상기 깊이 할당을 수행하는 단계를 통해 최적으로 생성된 깊이 정보와, 계산된 카메라 파라미터를 이용하여 임의 시점의 화상을 생성하는 단계를 포함할 수 있다.
- <42> 또한, 상기 다시점 화상 및 3D 오디오 송수신 방법은 수신측에서 요구된 시점 화상을 비압축 데이터 형태로 전송하는 비압축 데이터 전송단계; 및 모든 시점 화상과 이에 대응되는 3D 오디오를 하나의 데이터스트림(data stream)으로 다중화하여 압축 데이터 형태로 전송하는 압축 데이터 전송단계를 더 포함할 수 있다.

- <43> 또한, 상기 압축 데이터 전송단계의 다중화는 각 시점 화상 및 3D 오디오의 생성 시간 정보를 이용하여 동기화된 형태로 수행되는 것이 바람직하다.
- <44> 이하, 본 발명의 바람직한 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 우선 각 도면의 구성 요소들에 참조 부호를 부가함에 있어서, 동일한 구성 요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다. 또한, 이하에서 본 발명의 바람직한 실시예를 설명할 것이나, 본 발명의 기술적 사상은 이에 한정하거나 제한되지 않고 당업자에 의해 변형되어 다양하게 실시될 수 있음은 물론이다.
- <45> 먼저, 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 장치에 대해 설명한다.
- <46> 도 1은 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 장치의 블록도이다. 도 2는 도 1 중 다시점 화상 부호화부(200), 도 3은 도 1 중 다시점 방송 서버부(500), 도 4는 도 1 중 다시점 적응 서버부(600), 도 5는 도 1 중 다시점 데이터 전송부(700), 도 6은 도 1 중 시점 변환부(800), 도 7은 도 1 중 다시점 데이터 복호화부(900), 도 8은 도 1 중 중간시점 화상 생성부(1100), 그리고 도 9는 도 1 중 3D 오디오 합성부를 상세하게 도시한 블록도이다.
- <47> 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 장치는, 도 1을 참조하면, 다시점 화상 획득부(100), 다시점 화상 부호화부(200), 3D 오디오 획득부(300), 3D 오디오 부호화부(400), 다시점 방송 서버부(500), 다시점 적응 서버부(600), 다시점 데이터 전송부(700), 시점 변환부(800), 다시점 데이터 복호화부(900), 다시점 화상 저장부(1000), 중간시점 화상 생성부(1100), 다시점 화상 재생부(1200), 3D 오디오 저장부(1300), 3D 오디오 합성부(1400), 및 3D 오디오 재생부(1500)를 포함하여 형성된다.
- <48> 상기 다시점 화상 획득부(100)는 복수개의 카메라로부터 다시점 화상을 획득하는 부분이다. 예를 들어, 상기 다시점 화상 획득부(100)가 n개의 카메라를 구비하고 있다면, 동일한 촬영 대상을 각 카메라가 촬영하여 디지털 또는 아날로그 형태의 신호를 전송선을 통해 해당 버퍼에 저장한 후, 다시점 화상 부호화부(200)로 전송하게 된다.
- <49> 상기 다시점 화상 부호화부(200)는 다시점 화상 획득부(100)를 통해 얻은 다시점 화상 데이터를 부호화(encoding)하는 부분이다. 상기 다시점 화상 부호화부(200)는, 도 2를 참조하면, 다시점 화상 재배열부(210)와 재배열 화상 부호화부(220)를 포함하여 형성된다.
- <50> 상기 다시점 화상 재배열부(210)는 부호화기에 입력되는 복수개의 다시점 화상들(V1, V2, ..., Vn)을 GOP(Group Of Pictures) 단위로 분할하고, 이에 따라 순차적으로 재배열하는 부분이다. 상기 다시점 화상 재배열부(210)의 프레임 재배열 방법에 관해 간단히 설명하면 다음과 같다. 예를 들어, 8개의 카메라를 통해 8개의 시점 화상(S0~S7)이 입력되는 경우, 먼저 T0 시각에서 모든 시점(S0~S7)의 화면을 일렬로 나열한 다음, 소정의 시간(T1~T8) 동안 첫 시점의 화면(S0)에 대해서 일렬로 나열하여 출력한다. 이어서, 다음 시점의 화면(S1)에 대해서 소정 시간(T1~T8) 동안 나열하는 과정을 반복하여 마지막 시점(S7)의 화면에 대해서도 일렬로 나열하는 것이 완료되면, 다시 소정의 시간(T9~T16) 동안 첫 시점의 화면(S0)부터 나열하기 시작하여 마지막 시점(S7)의 화면까지 나열한다. 이와 같은 화면 재배열 방식은 현재 H.264/AVC 표준에서 다시점 화상에 대해 검토하고 있는 방식이다. 다만, 이와 다른 방식으로 화면을 재배열할 수도 있음은 물론이다.
- <51> 상기 재배열 화상 부호화부(220)는 재배열된 화상을 MPEG-2, H.264/AVC 등의 표준 부호화 방식에 따라 인코딩하는 부분이다. 각 시점 화상은 독립적으로 동시 부호화(simulcast encoding) 될 수도 있으나, 바람직하게는 시점 간 상관성을 이용하여 다시점 화상 부호화(multi-view encoding)된다. 다시점 화상에는 시간적으로 인접한 화상 간의 시간적 중복성뿐만 아니라, 공간적으로 인접한 카메라들에 의해 획득된 화상간의 공간적 중복성도 존재한다. 상기 동시 부호화 방식은 각 시점의 화상이 독립적으로 부호화되므로, 각 시점 화상 간에 존재하는 공간적 중복성을 제거할 수 없다는 단점이 있다. 이에 반해, 다시점 화상 부호화 방식은 시점 화상들을 재배열하여 단일 부호화기로 부호화함으로써, 공간적인 중복성 또한 제거할 수 있으므로, 보다 효율적으로 다시점 화상을 부호화할 수 있다.
- <52> 상기 3D 오디오 획득부(300)는 복수개의 마이크로폰(microphone)으로부터 다채널 3D 오디오 사운드 소스(sound source)를 획득하는 부분이다. 상기 3D 오디오 획득부(300)에서는 두 개의 마이크로폰을 이용하여 스테레오 사운드 소스를 획득할 수도 있으나, 바람직하게는 다수개의 마이크로폰을 이용하여 4채널 혹은 5.1채널 등의 다채널 사운드 소스를 획득하게 된다. 이러한 사운드 소스의 획득은 각 시점 화상별로 이루어지는 것이 바람직하다.

시청자는 시점 변환부(800)를 통해 자신이 선택한 시점 화상과 함께, 해당 시점 화상을 반영하는 3D 오디오를 시청할 수 있으므로, 보다 실감나는 입체 사운드를 체험할 수 있게 된다.

- <53> 상기 3D 오디오 부호화부(400)는 3D 오디오 획득부를 통해 얻은 3D 오디오 데이터를 부호화하는 부분이다. 이때, 상기 3D 오디오 부호화부(400)에서는 AAC(Advanced Audio Coding) 또는 AC-3(Audio Coding-3)와 같은 다 채널 오디오 압축기법을 이용하여 사운드 소스를 인코딩하게 된다.
- <54> 상기 다시점 방송 서버부(500)는 다시점 시스템을 관리하고, 다시점 화상 및 3D 오디오 데이터를 처리하는 부분이다. 상기 다시점 방송 서버부(500)는, 도 3을 참조하면, 다시점 화상 데이터와 3D 오디오 데이터 일부를 다중화하는 다시점 화상 다중화부(510), 부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화부(520)를 포함하여 형성된다. 예를 들어, 다시점 화상 부호화 데이터는 V1, V2, ..., Vn으로 구성되어 있고, 오디오 부호화 데이터는 A 하나로 구성된다면, 상기 다시점 화상 다중화부(510)는 V1, V2, ..., Vn-1의 화상 데이터를 다중화하고 상기 화상 오디오 다중화부(520)는 Vn의 시점 화상과 A의 오디오를 다중화할 수 있다. 한편, 각 시점 화상별로 오디오 부호화 데이터가 존재하는 경우에는 여러 개의 화상 오디오 다중화부(520)가 구비되거나, 다수의 데이터를 적절히 그룹지어 하나의 화상 오디오 다중화부(520)를 통해 다중화될 수 있다.
- <55> 상기 다시점 적응 서버부(600)는 수신측(클라이언트)으로 다시점 화상을 전송하고, 메시지 교환을 하며, 수신측이 요구하는 시점에 따른 시점 화상 및 3D 오디오 데이터를 생성하는 부분이다. 즉, 상기 다시점 적응 서버부(600)는 다시점 화상과 3D 오디오를 소비하는 사용 환경, 즉 수신측 단말 특성 및/또는 시청자의 재생 취향 특성에 대한 정보를 이용하여 하나의 콘텐츠를 각기 다른 사용 환경에 부합할 수 있도록 적응 변환처리하는 단일 소스 복수사용(Single-Source Multi-Use) 환경을 제공한다. 상기 다시점 적응 서버부(600)는, 도 4를 참조하면, 시점 기술자 파싱부(610), 화상 리소스 적응부(620), 화상 기술자 적응부(630), 오디오 리소스 적응부(640), 및 오디오 기술자 적응부(650)를 포함하여 형성된다.
- <56> 상기 시점 기술자 파싱부(610)는 시점 변환부(800)로부터 전송된 시점 기술자(view descriptor)를 파싱(parsing)하는 부분이다. 여기서 기술자(descriptor)란, 표준화된 표현, 식별 및 메타 데이터를 갖는 구조화된 디지털 객체인 디지털 아이템(Digital Item;DI) 내의 항목 또는 컴포넌트(component)에 관련된 정보를 의미한다. 즉, 상기 시점 기술자 파싱부(610)는 시점 변환부(800)의 시점 결정부(810)에서 시청자가 선택한 시점에 따라 시점 기술자 생성부(820)에서 생성된 시점 기술자를 입력받아 관련 정보를 파싱한다.
- <57> 상기 화상 리소스 적응부(620)는 시점 기술자 파싱부(610)를 통해 파싱된 다시점 리소스를 적응 변환하는 부분이다. 여기서 리소스(resource)란, 비디오, 오디오, 이미지 또는 텍스트 항목과 같이 개별적으로 식별가능한 항목을 의미한다. 즉, 상기 화상 리소스 적응부(620)는 다시점 리소스에 시청자의 재생 취향을 반영하여 화상 콘텐츠를 적응 변환시킨다. 이를 통해 적응 리소스 Rv가 생성되고, 소정의 변환과정을 거친 후 클라이언트로 전송된다.
- <58> 상기 화상 기술자 적응부(630)는 다시점 화상 기술자를 적응 변환하는 부분이다. 일반적으로 화상 기술자는 수신측 단말기의 화상 포맷, 초당 처리되는 최대 교점수(maximum vertices), 초당 최대 픽셀수(maximum pixels), 최대 전송율(maximum rate)과 같은 특성을 기술하거나, 특정 시점 화상을 선호하는 시청자의 재생 취향 특성을 기술한다. 상기 화상 기술자 적응부(630)는 시점 기술자 생성부(820)로부터 전송받은 다시점 화상 기술자를 반영하여 화상 콘텐츠를 적응 변환하게 된다. 이를 통해 적응 기술자 Dv가 생성되고, 소정의 변환과정을 거친 후 클라이언트로 전송된다.
- <59> 상기 오디오 리소스 적응부(640)는 시점 기술자 파싱부(610)를 통해 파싱된 오디오 리소스를 적응 변환하는 부분이다. 이를 통해 적응 리소스 Ra가 생성되고, 소정의 변환과정을 거친 후 클라이언트로 전송된다.
- <60> 상기 오디오 기술자 적응부(650)는 시점 기술자 생성부(820)로부터 전송받은 오디오 기술자를 반영하여 오디오 콘텐츠를 적응 변환하게 된다. 이를 통해 적응 기술자 Da가 생성되고, 소정의 변환과정을 거쳐 클라이언트로 전송된다.
- <61> 상기 다시점 데이터 전송부(700)는 클라이언트에게 효율적으로 데이터를 전송하기 위한 전송 프로토콜 및 메커니즘을 제공한다. 이때, 상기 다시점 데이터 전송부(700)는 비압축 데이터 전송과 압축 데이터 전송을 위한 두 가지 형태의 데이터 전송을 위한 구조를 구비한다. 비압축 데이터 전송은 클라이언트가 요구한 시점 화상이 지연되거나 화질이 손상되는 것을 방지하기 위해 비압축 데이터 형태로 전송한다. 반면, 압축 데이터 전송은 클라이언트가 요구한 시점 화상을 포함한 모든 시점 화상을 압축하고, 이에 3D 오디오 데이터를 다중화하여 전송한다.

다.

- <62> 첫째, 비압축 데이터 전송을 위해 상기 다시점 데이터 전송부(700)는, 도 5를 참조하면, 획득 서버부(710), IP 멀티캐스트 처리부(720), 제어 서버부(740)를 구비한다. 상기 획득 서버부(710)는 다수개 구비되며, 바람직하게는 다시점 카메라의 개수만큼 구비되어 각 다시점 카메라로부터 시점 데이터를 얻어 온다. 상기 제어 서버부(740)는 클라이언트가 요구한 시점을 선택적으로 전송하기 위해 제어 명령을 내리는 부분으로, 클라이언트가 원하는 시점에 해당하는 획득 서버부(710)로부터 데이터를 가져와 비압축 데이터 형태로 각 IP 멀티캐스트 처리부(720)를 이용하여 전송한다. 즉, 클라이언트가 원하는 시점만이 제어 서버부(740)의 제어 데이터를 통하여 선택적으로 전송된다.
- <63> 둘째, 압축 데이터 전송을 위해 상기 다시점 데이터 전송부(700)는 획득 서버부(710), 시점 데이터 모음부(730), 동기 다중화부(750), 및 IP 멀티캐스트 처리부(760)를 구비한다. 상기 시점 데이터 모음부(730)는 획득 서버부(710)들로부터 각 시점 화상 및 3D 오디오의 생성 시간 정보를 포함하여, 모든 시점의 데이터를 받아 온다. 수집된 데이터들은 동기 다중화부(750)를 통해 하나의 데이터 스트림으로 다중화된다. 이때, 상기 동기 다중화부(750)는 각 시점 화상 및 3D 오디오의 생성 시간 정보를 이용하여 동기화된 다중화를 수행한다. 이렇게 다중화된 하나의 스트림은 IP 멀티캐스트 처리부(760)를 통해 클라이언트에게 전송된다.
- <64> 상기 시점 변환부(800)는 시청자가 요구한 시점에 따라 시점 화상의 기술자와 3D 오디오의 기술자를 생성하는 부분이다. 상기 시점 변환부(800)는, 도 6을 참조하면, 시점 결정부(810)와 시점 기술자 생성부(820)를 포함한다.
- <65> 상기 시점 결정부(810)는 전송받은 복수개의 시점 화상으로부터 시청자가 원하는 시점을 선택하는 부분이다. 이때, 시청자가 원하는 시점은 하나일 수도 있고 둘 이상일 수도 있다. 즉, 상기 시점 결정부(810)는 다시점 카메라의 시점 화상 V_1, V_2, \dots, V_n 으로부터 사용자가 원하는 시점 V_k ($k=1\sim n$)를 선택한다.
- <66> 상기 시점 기술자 생성부(820)는 시점 결정부(810)를 통해 시점이 결정되면 그에 따른 시점 기술자(view descriptor)를 생성한다. 시점 기술자는 XML(eXtensible Markup Language) 형식의 기계판독 가능한 언어로 표현될 수 있으며, 다시점 화상 기술자와 오디오 기술자가 있다. 상기 시점 기술자 생성부(820)를 통해 생성된 시점 기술자는 다시점 적응 서버부(600)로 전송되어 시점 화상과 3D 오디오를 적응 변환시키도록 한다.
- <67> 상기 다시점 데이터 복호화부(900)는 수신된 다시점 화상 및 3D 오디오 데이터를 복호화하여, 다시점 화상 저장부(1000)와 3D 오디오 저장부(1400)로 전송한다. 상기 다시점 데이터 복호화부(900)는, 도 7을 참조하면, j 시점 복호화부(910), 시점 데이터 모음부(920), 및 3D 오디오 복호화부(930)를 구비한다.
- <68> 상기 j 시점 복호화부(910)는 수신된 n 개의 다시점 부호화 데이터 중에서 j 개의 시점을 복호화하는 부분이다. 이때, j 값은 n 값보다 작거나 같은 것이 바람직하다. 즉, 상기 j 시점 복호화부(910)는 n 개의 시점 중에서 일부인 j 개의 시점을 복호화하거나, 모든 시점에 대해 복호화하게 된다. 모든 시점에 대해 복호화하는 경우($j=n$), 도시되지 않았으나 1개의 복호기만을 이용하여 n 개의 다시점 부호화 데이터를 복호화할 수도 있다. 이는 클라이언트가 원하는 품질 정도에 따라 부호화기에 비해 복잡도가 낮은 복호화기를 사용할 수 있기 때문이다. 이와 같이 1개의 복호기만을 이용하여 복호화하는 경우에는 일정 이상의 품질을 확보하면서도 효율적인 통제가 가능하며, 시점 개수가 증감하더라도 유연하게 대처할 수 있다는 장점이 있다. 도 7에 도시된 바와 같이, 필요에 따라 복호화기의 개수를 증가시켜 보다 향상된 성능을 확보할 수도 있음은 물론이다. j 값은 복호화기의 성능을 비롯하여 사용 환경 등 몇 가지 변수에 의해 결정된다. 상기 j 시점 복호화부(910)는 MPEG-2, H.264/AVC와 같은 표준 시스템의 복호 규칙을 일부 수용하여 형성된다.
- <69> 상기 시점 데이터 모음부(920)는 j 시점 복호화부(910)를 통해 복호화된 데이터를 하나로 모으는 부분이다. 상기 시점 데이터 모음부(920)는 수집한 데이터를 다시점 화상 저장부(1000)로 전송하게 된다. 상기 다시점 화상 저장부(1000)에서는 전송받은 데이터를 각 시점별로 프레임 단위의 시점 화상을 생성하기 위해 일시적으로 저장한다.
- <70> 상기 3D 오디오 복호화부(930)는 수신된 오디오 데이터를 복호화하는 부분이다. 이때, 오디오 데이터는 다채널 또는 스테레오 채널일 수 있으며, 3D 오디오 저장부(1300)로 전송된다. 상기 3D 오디오 저장부(1300)는 다시점 화상과 동기화하여 3D 오디오를 합성하기 위해 복호화된 3D 오디오 데이터를 일시적으로 저장한다.
- <71> 상기 중간시점 화상 생성부(1100)는 다시점 화상의 중간시점 화상을 생성하여 다시점 화상 재생부(1200)로 전송한다. 즉, 상기 중간시점 화상 생성부(1100)는 복호화된 시점 화상 간의 중간시점 화상을 생성하여, 3D 모니터의 시점 간격에 맞추어 부드러운 시점 화상을 제공하게 된다. 상기 중간시점 화상 생성부(1100)는, 도 8을 참조

하면, 카메라 파라미터 계산부(1110), 깊이정보 계산부(1120), 영역 분할부(1130), 깊이계층 예측부(1140), 깊이정보 개선부(1150), 및 중간시점 화상 제작부(1160)를 구비한다.

- <72> 상기 카메라 파라미터 계산부(1110)는 카메라의 외부 및 내부 파라미터를 계산하는 부분이다. 카메라의 외부 파라미터는 카메라의 세팅과 관련된 변수들, 즉 카메라의 위치, 높이, 각도 등에 관한 파라미터이며, 내부 파라미터는 카메라 자체의 특성과 관련된 변수들, 즉 광심, 초점 거리, 해상도 등에 관한 파라미터이다.
- <73> 상기 깊이정보 계산부(1120)는 이웃하는 시점 화상간의 매칭점을 찾아 초기 깊이정보를 찾아내는 부분이다. 즉, 상기 깊이정보 계산부(1120)는 주어진 각 시점 화상 사이에 블록 단위로 유사성을 측정하고, 그 결과 가장 유사한 매칭점을 찾아 초기 깊이정보를 계산한다.
- <74> 상기 영역 분할부(1130)는 이웃하는 시점 화상간에 기준 화상을 정하고, 이러한 기준 화상에 대한 색상 정보를 바탕으로 서로 다른 색상정보를 가지는 영역으로 분할하는 부분이다.
- <75> 상기 깊이계층 예측부(1140)는 영역 분할부(1130)를 통해 색상별로 분할된 영역에 깊이정보 계산부(1120)를 통해 계산된 초기 깊이정보를 반영하여 깊이계층을 생성하는 부분이다. 이때, 상기 깊이계층 예측부(1140)는 같은 영역에 대해서는 대표 깊이정보를 할당한다. 이러한 작업을 반복하여 모든 영역에 대하여 유사한 대표 깊이정보끼리 그룹으로 묶은 후, 그룹 내에서 대표깊이값 하나를 추출하여 깊이계층을 생성한다.
- <76> 상기 깊이정보 개선부(1150)는 깊이계층 예측부(1140)로부터 생성된 깊이 계층에 반복적으로 깊이할당을 수행하는 부분이다. 보다 상세하게는, 상기 깊이정보 개선부(1150)는 각 영역별로 대표 깊이 정보를 이용하여 깊이계층 예측부(1140)에서 생성된 깊이계층에 할당하고, 대표깊이 정보와 위치상 인접한 색상 영역 및 색상의 유사성 등을 반영하는 비용함수를 정의한다. 이때, 깊이할당은 이러한 비용함수값이 최소가 되는 지점을 찾기 위해 반복적으로 수행되는 것이 바람직하다.
- <77> 상기 중간시점 제작부(1160)는 깊이정보 개선부(1150)를 통해 최적으로 생성된 깊이정보와, 카메라 파라미터 계산부(1110)를 통해 계산된 카메라 파라미터를 이용하여 임의 시점의 화상을 생성하는 부분이다. 이와 같이 하여 생성된 임의의 시점 화상에는 가려짐 등에 기인한 빈 영역들이 발생할 수 있다. 이때, 이러한 빈 영역은 이웃하는 화소를 바탕으로 복원하고, 보간법(interpolation)을 이용하여 자연스러운 중간시점 화상을 생성한다.
- <78> 상기 3D 오디오 합성부(1400)는, 도 9를 참조하면, 서라운드 패닝(surround panning)부(1410), 3D 파라미터 추출부(1420) 및 다운 믹싱부(1430)를 구비한다. 상기 서라운드 패닝부(1410)는 사용자가 선택한 시점에 따라 서라운드 패닝을 통해 선택된 시점 화상에 일치하는 3D 오디오를 합성한다. 합성된 3D 오디오 데이터는 다채널 스피커를 통해 재생된다. 상기 3D 파라미터 추출부(1420)는 스테레오 재생시 3D 효과를 주기 위해 다채널 오디오 데이터로부터 3D 파라미터를 추출한다. 상기 다운 믹싱부(1430)는 추출된 3D 파라미터와 선택된 시점을 이용하여 스테레오 오디오 데이터를 생성한다. 생성된 스테레오 오디오 데이터는 스테레오 스피커를 통해 재생된다.
- <79> 상기 3D 오디오 재생부(1500)는 합성된 3D 오디오 데이터를 다채널 스피커 또는 스테레오 스피커를 통해 재생한다. 이때, 상기 서라운드 패닝부(1410)를 통해 전송된 데이터는 다채널 스피커로 재생되고, 다운 믹싱부(1430)를 통해 전송된 데이터는 스테레오 스피커로 재생됨은 상기에서 언급한 바와 같다.
- <80> 다음으로, 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 방법에 대해 설명한다.
- <81> 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 방법은 다시점 화상 획득단계, 다시점 화상 부호화단계, 3D 오디오 획득단계, 3D 오디오 부호화단계, (a)부호화된 다시점 화상 데이터 일부를 다중화하는 다시점 화상 다중화 단계, (b)부호화된 나머지 다시점 화상 데이터와 부호화된 3D 오디오 데이터를 다중화하는 화상 오디오 다중화단계, (c)시청자가 요구한 시점에 따라 시점 화상의 기술자와 3D 오디오의 기술자를 생성하는 단계, 수신측에서 요구된 시점 화상을 비압축 데이터 형태로 전송하는 비압축 데이터 전송단계, 모든 시점 화상과 이에 대응되는 3D 오디오를 하나의 데이터스트림으로 다중화하여 압축 데이터 형태로 전송하는 압축 데이터 전송단계, (d)수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 단계, 다시점 데이터 복호화단계, 다시점 화상 저장단계, (e)복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 단계, 다시점 화상 재생 단계, 3D 오디오 저장단계, 3D 오디오 합성 단계, 및 3D 오디오 재생단계를 포함하여 이루어진다.
- <82> 상기 (c)시청자가 요구한 시점에 따라 시점 화상의 기술자와 3D 오디오의 기술자를 생성하는 단계는 전송받은 복수개의 시점 화상으로부터 시청자가 원하는 적어도 하나의 시점을 선택하는 단계 및 선택된 시점의 기술자를 생성하여 전송하는 단계를 포함하여 이루어진다.

상기 (d)수신측에서 요구하는 시점 화상과 그에 대응되는 3D 오디오를 생성하는 단계는 전송된 시점 기술자를 파악하는 단계, 다시점 리소스를 적용하는 단계, 다시점 화상 기술자를 적용하는 단계, 오디오 리소스를 적용하는 단계, 및 오디오 기술자를 적용하는 단계를 포함하여 이루어진다.

<83> 삭제

<84> 또한, 상기 (e)복호화된 다시점 화상 데이터를 통해 중간 시점의 화상을 생성하는 단계는 카메라의 외부 및 내부 파라미터를 계산하는 단계, 이웃하는 시점 화상 간의 매칭점을 찾아 초기 깊이정보를 찾아내는 단계, 이웃하는 시점 화상간에 기준 화상을 정하고, 기준 화상에 대한 색상정보를 바탕으로 서로 다른 색상 정보를 가지는 영역으로 분할하는 단계, 색상별로 분할된 영역에 초기 깊이 정보를 반영하여 깊이 계층을 생성하는 단계, 깊이 계층에 반복적으로 깊이 할당을 수행하는 단계 및 깊이 할당을 수행하는 단계를 통해 최적으로 생성된 깊이 정보와, 계산된 카메라 파라미터를 이용하여 임의 시점의 화상을 생성하는 단계를 포함하여 이루어진다. 상기 각 단계에 대해서는 상기 다시점 화상 및 3D 오디오 송수신 장치 구조에서 상술하였으므로, 여기서는 생략하기로 한다.

<85> 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 방법이 수행되는 대략적인 순서는 다음과 같다.

<86> 먼저, 다시점 화상 획득부(100)로 입력된 다시점 화상은 다시점 화상 부호화부(200)를 통해 인코딩된 후, 다시점 방송 서버부(500)로 전송된다. 마찬가지로, 3D 오디오 획득부(300)로 입력된 사운드 소스는 3D 오디오 부호화부(400)를 통해 인코딩된 후, 다시점 방송 서버부(500)로 전송된다.

<87> 다시점 방송 서버부(500)에서는 부호화된 다시점 화상 데이터 일부를 다중화하고(a단계), 나머지 다시점 화상 데이터와 3D 오디오 데이터를 다중화한다(b단계). 시점 변환부(800)에서는 시청자가 요구한 시점에 따라, 시점 화상 기술자와 3D 오디오 기술자를 생성(c단계)한 후, 다시점 적용 서버부(600)로 전송한다. 다시점 적용 서버부(600)에서는 수신측 시점 변환부(800)에서 요구하는 시점에 적용하여 시점 화상과 그에 대응하는 3D 오디오를 생성(d단계)하여 다시점 데이터 전송부(700)로 전송한다. 다시점 데이터 전송부(700)에서는 비압축 데이터 전송과 압축 데이터 전송을 수행하여 전송하고, 다시점 데이터 복호화부(900)에서 이를 디코딩하여 다시점 화상 저장부(1000)와 3D 오디오 저장부(1300)에 저장한다.

<88> 중간시점 화상 생성부(1100)에서는 다시점 화상 저장부(1000)로부터 복호화된 다시점 화상 데이터를 가져와서 중간시점의 화상을 생성한다(e단계). 이러한 시점 화상들은 다시점 화상 재생부(1200)를 통하여 재생된다. 마찬가지로, 3D 오디오 합성부(1400)에서는 3D 오디오 저장부(1300)로부터 복호화된 3D 오디오 데이터를 가져와서 다채널 혹은 스테레오 채널의 3D 오디오를 합성하고, 3D 오디오 재생부(1500)를 통해 재생된다.

<89> 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

발명의 효과

<90> 본 발명에 의하면 다시점 화상과 3D 오디오 데이터를 결합하되 시청자가 보는 시점화상에 따라 다른 오디오를 재생함으로써 보다 실감나는 미디어를 공급할 수 있고, 수신측에서 원하는 시점의 화상만을 선택적으로 전송할 수 있어 좁은 대역폭으로도 가능하며, 중간시점의 화상을 생성하여 모니터의 시점 간격에 맞추어 부드럽고 자연스러운 시점 화상을 제공할 수 있는 효과가 있다.

도면의 간단한 설명

- <1> 도 1은 본 발명의 바람직한 실시예에 따른 다시점 화상 및 3D 오디오 송수신 장치의 블록도,
- <2> 도 2는 도 1 중 다시점 화상 부호화부의 상세한 구성을 도시한 블록도,
- <3> 도 3은 도 1 중 다시점 방송 서버부의 상세한 구성을 도시한 블록도,

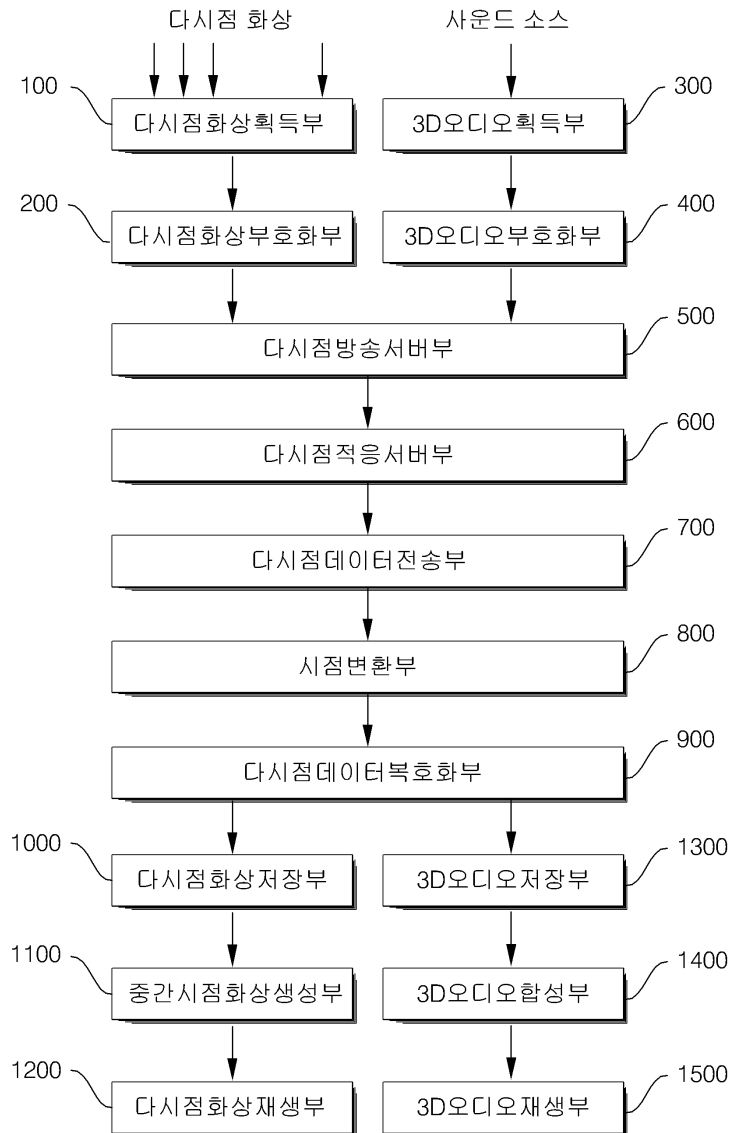
- <4> 도 4는 도 1 중 다시점 적응 서버부의 상세한 구성을 도시한 블록도,
- <5> 도 5는 도 1 중 다시점 데이터 전송부의 상세한 구성을 도시한 블록도,
- <6> 도 6은 도 1 중 시점 변환부의 상세한 구성을 도시한 블록도,
- <7> 도 7은 도 1 중 다시점 데이터 복호화부의 상세한 구성을 도시한 블록도,
- <8> 도 8은 도 1 중 중간시점 화상 생성부의 상세한 구성을 도시한 블록도,
- <9> 도 9는 도 1 중 3D 오디오 합성부의 상세한 구성을 도시한 블록도이다.

<10> <도면의 주요 부분에 대한 부호의 설명>

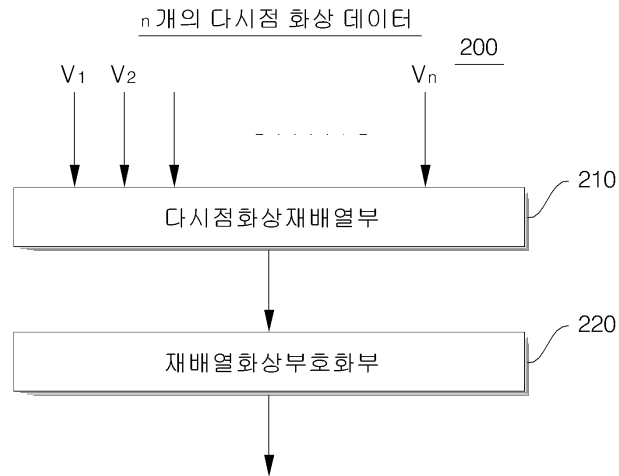
- | | |
|-------------------------|-------------------|
| <11> 100 - 다시점 화상 획득부 | 200 - 다시점 화상 부호화부 |
| <12> 300 - 3D 오디오 획득부 | 400 - 3D 오디오 부호화부 |
| <13> 500 - 다시점 방송 서버부 | 600 - 다시점 적응 서버부 |
| <14> 700 - 다시점 데이터 전송부 | 800 - 시점 변환부 |
| <15> 900 - 다시점 데이터 복호화부 | 1000 - 다시점 화상 저장부 |
| <16> 1100 - 중간시점 화상 생성부 | 1200 - 다시점 화상 재생부 |
| <17> 1300 - 3D 오디오 저장부 | 1400 - 3D 오디오 합성부 |
| <18> 1500 - 3D 오디오 재생부 | |

도면

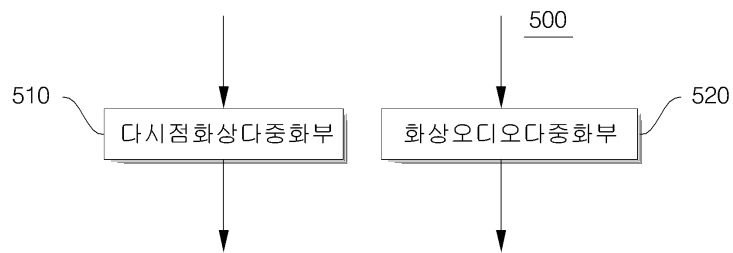
도면1



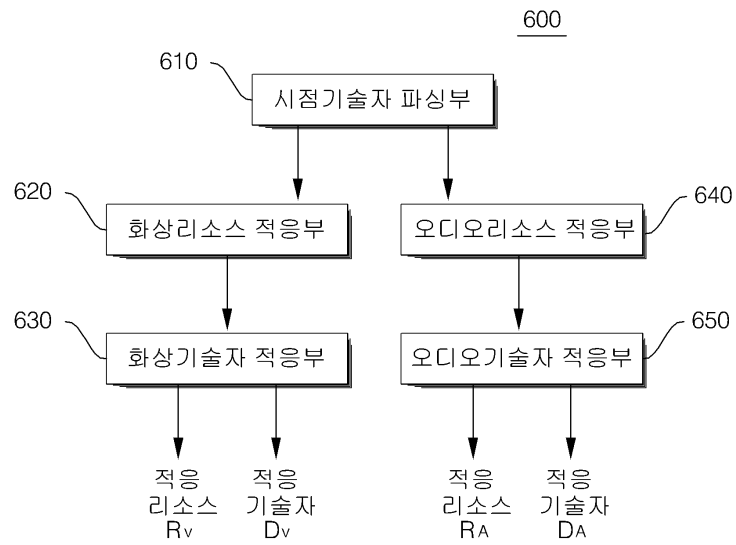
도면2



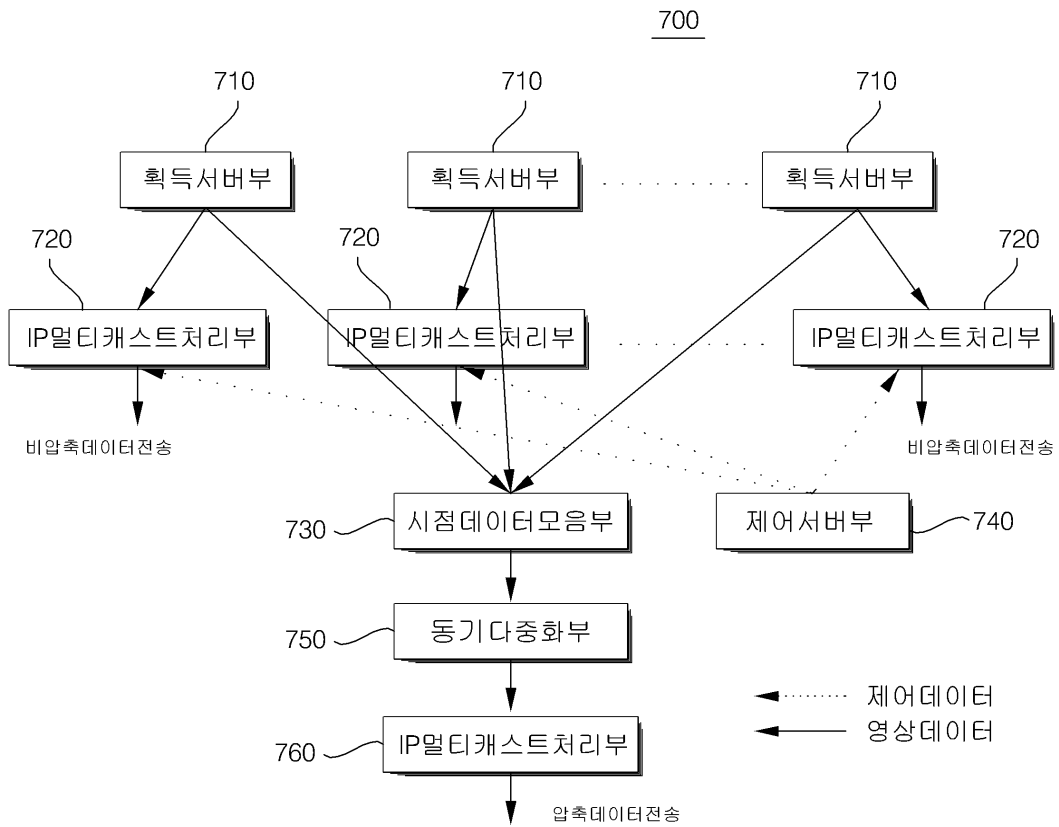
도면3



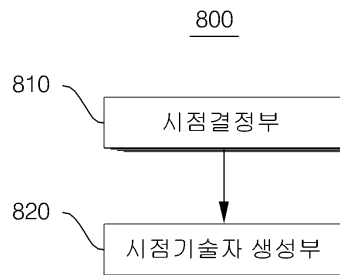
도면4



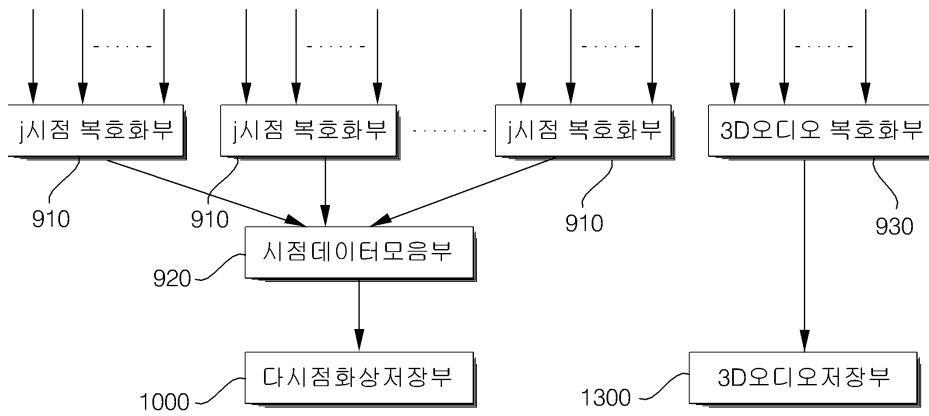
도면5



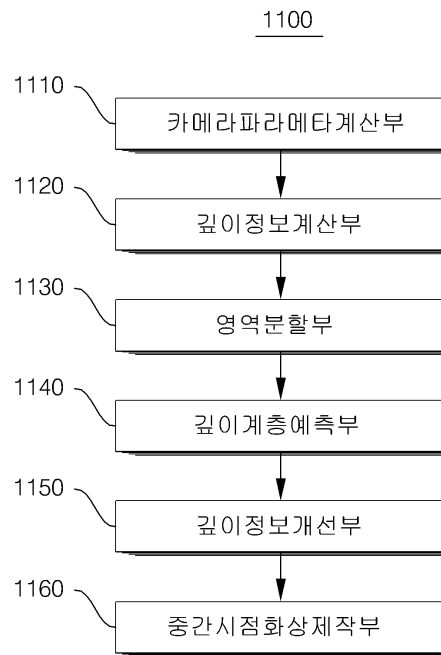
도면6



도면7



도면8



도면9

