



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년12월26일

(11) 등록번호 10-2746047

(24) 등록일자 2024년12월19일

(51) 국제특허분류(Int. Cl.)

G06T 7/50 (2017.01) G06F 18/00 (2023.01)  
G06F 18/00 (2023.01) G06T 7/11 (2017.01)  
G06T 7/143 (2017.01) G06V 10/24 (2022.01)  
H04N 13/00 (2018.01)

(52) CPC특허분류

G06T 7/50 (2017.01)  
G06F 18/2148 (2023.01)

(21) 출원번호 10-2017-7036816

(22) 출원일자(국제) 2016년05월06일

심사청구일자 2021년05월04일

(85) 번역문제출일자 2017년12월20일

(65) 공개번호 10-2018-0011215

(43) 공개일자 2018년01월31일

(86) 국제출원번호 PCT/EP2016/060221

(87) 국제공개번호 WO 2016/184700

국제공개일자 2016년11월24일

(30) 우선권주장

15168577.3 2015년05월21일  
유럽특허청(EPO)(EP)

(56) 선행기술조사문헌

US20140307056 A1\*

(뒷면에 계속)

전체 청구항 수 : 총 15 항

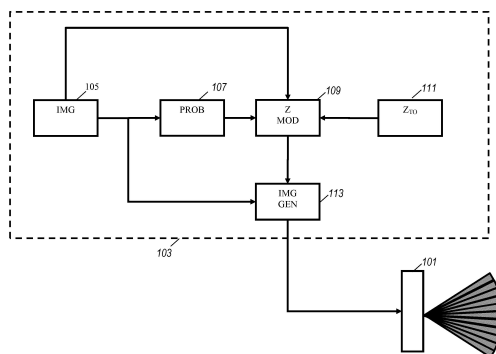
심사관 : 신지선

(54) 발명의 명칭 이미지를 위한 깊이 맵을 결정하기 위한 방법 및 장치

## (57) 요약

이미지를 위한 깊이 맵을 결정하기 위한 장치는 이미지의 적어도 일부 픽셀들에 대한 깊이 값들을 포함하는 연관된 깊이 맵을 이미지에 제공하는 이미지 유닛(105)을 포함한다. 확률 유닛(107)은 픽셀들이 텍스트 이미지 오브젝트에 속할 확률을 나타내는 확률 값들을 포함하는 이미지를 위한 확률 맵을 결정한다. 깊이 유닛(109)은 수정된 깊이 값들이 입력 값들과 텍스트의 바람직한 깊이에 대응하는 텍스트 이미지 오브젝트 깊이 값들의 가중된 조합들로서 결정되는 수정된 깊이 맵을 생성한다. 가중치는 픽셀들에 대한 확률 값에 의존한다. 이 접근법은 텍스트 오브젝트들에 대한 보다 부드러운 깊이 수정을 제공하여 아티팩트들 및 저하들을 감소시킨다(예를 들어, 깊이 맵들을 사용하여 뷰 이동을 수행할 때).

## 대표도



(52) CPC특허분류

**G06F 18/254** (2023.01)  
**G06T 7/11** (2017.01)  
**G06T 7/143** (2017.01)  
**G06V 20/635** (2022.01)  
**G06V 20/647** (2022.01)  
**H04N 13/128** (2021.08)  
**G06T 2207/10028** (2013.01)  
**G06T 2207/20182** (2013.01)

(56) 선행기술조사문헌

W02014072926 A1  
 Koppel, Martin, et al. (2013). "Temporally consistent adaptive depth map preprocessing for view synthesis." 2013 Visual Communications and Image Processing (VCIP). IEEE.\*  
 Socher, Richard, et al. (2012). "Convolutional-recursive deep learning for 3d object classification." Advances in neural information processing systems 25.\*  
 JP2007524950 A  
 US20110170801 A1  
 W02008062351 A1  
 US20130084007 A1  
 US20140055450 A1  
 \*는 심사관에 의하여 인용된 문헌

## 명세서

### 청구범위

#### 청구항 1

이미지에 대한 깊이 맵을 결정하기 위한 장치에 있어서:

제 1 이미지의 적어도 일부 픽셀들에 대한 제 1 깊이 값들을 포함하는 연관된 깊이 맵을 상기 제 1 이미지에 제공하기 위한 이미지 유닛(105);

상기 이미지의 적어도 일부 픽셀들에 대한 확률 값들을 포함하는 상기 제 1 이미지에 대한 확률 맵을 결정하기 위한 확률 유닛(107)으로서, 픽셀에 대한 상기 확률 값은 상기 픽셀이 텍스트 이미지 오브젝트에 속할 확률을 나타내는, 상기 확률 유닛(107); 및

상기 제 1 이미지에 대한 수정된 깊이 맵을 생성하기 위한 깊이 유닛(109)으로서, 상기 깊이 유닛은 제 1 픽셀에 대한 상기 연관된 깊이 맵의 제 1 깊이 값과 복수의 픽셀들에 대해 동일한 텍스트 이미지 오브젝트 깊이 값의 가중된 조합으로서 적어도 제 1 픽셀에 대한 수정된 깊이 값을 결정하도록 구성되고, 가중치는 상기 제 1 픽셀에 대한 상기 확률 값에 의존하는, 상기 깊이 유닛(109)을 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 2

제 1 항에 있어서,

상기 확률 유닛(107)은 상기 제 1 이미지에 대한 픽셀 그룹들을 결정하도록 구성되고, 각 픽셀 그룹은 복수의 픽셀들을 포함하고, 상기 확률 유닛은 픽셀 그룹들에 대한 확률 값들을 결정하도록 구성되는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 3

제 2 항에 있어서,

상기 깊이 유닛(109)은 상기 제 1 픽셀이 속하는 픽셀 그룹에 대한 제 1 확률 값 및 이웃하는 픽셀 그룹에 대한 제 2 확률 값에 응답하여 상기 가중된 조합에 대한 가중치들을 결정하도록 구성되고, 상기 제 1 확률 값 및 상기 제 2 확률 값의 가중치는 상기 제 1 픽셀이 속하는 상기 픽셀 그룹 내의 상기 제 1 픽셀의 위치에 의존하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 4

제 1 항에 있어서,

상기 확률 유닛(107)은 상기 이미지의 픽셀 그룹들에 대한 확률 값들을 결정하도록 구성되고, 각 픽셀 그룹은 적어도 하나의 픽셀을 포함하고, 상기 확률 유닛(107)은 상기 픽셀 그룹들이 텍스트 이미지 오브젝트에 속하는 것으로 지정되는지 또는 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정되는지의 여부를 나타내는 상기 픽셀 그룹들에 대한 값들을 포함하는 분류 맵을 먼저 결정하고, 상기 분류 맵의 필터링에 응답하여 상기 확률 맵을 생성하도록 구성되는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 5

제 4 항에 있어서,

상기 분류 맵은 상기 픽셀 그룹들에 대한 이진 값들을 포함하고, 각 이진 값은 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로 지정되거나 상기 픽셀그룹이 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정되는 것을 나타내는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 6

제 5 항에 있어서,

상기 필터링은 상기 분류 맵에 적용된 이진 팽창 필터링(binary dilation filtering)을 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 7

제 4 항에 있어서,

상기 제 1 이미지는 이미지들의 시간 시퀀스(temporal sequence)의 일부이고, 상기 필터링은 시간적 저역-통과 필터링을 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 8

제 7 항에 있어서,

상기 시간 필터링은, 상기 텍스트 이미지 오브젝트에 속할 픽셀 그룹의 확률을 증가시키기 위한 시간 상수가 상기 텍스트 이미지 오브젝트에 속할 픽셀 그룹의 확률을 감소시키기 위한 시간 상수와 상이하도록 비대칭인, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 9

제 4 항에 있어서,

상기 필터링은 공간 필터링을 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 10

제 9 항에 있어서,

상기 공간 필터링은 최대 출력값 한계를 갖는 필터인 소프트맥스 필터(softmax filter)를 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 11

제 9 항에 있어서,

상기 공간 필터링은 적어도 두 개의 순차 공간적 소프트맥스 필터들을 포함하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 12

제 1 항에 있어서,

상기 제 1 깊이 값들은 상기 제 1 이미지의 장면에 대한 상이한 뷰 방향들에 대응하는 복수의 이미지들 내의 이미지 오브젝트들에 대한 불일치들에 대응하는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 13

제 1 항에 있어서,

상기 텍스트 이미지 오브젝트 깊이의 가중치는 증가되고 상기 제 1 깊이 값의 가중치는 상기 확률 값의 증가하는 값에 대해 감소되는, 이미지에 대한 깊이 맵을 결정하기 위한 장치.

#### 청구항 14

이미지에 대한 깊이 맵을 결정하는 방법에 있어서,

제 1 이미지의 적어도 일부 픽셀들에 대한 제 1 깊이 값들을 포함하는 연관된 깊이 맵을 상기 제 1 이미지에 제공하는 단계;

상기 이미지의 적어도 일부 픽셀들에 대한 확률 값들을 포함하는 상기 제 1 이미지에 대한 확률 맵을 결정하는 단계로서, 픽셀에 대한 상기 확률 값은 상기 픽셀이 텍스트 이미지 오브젝트에 속할 확률을 나타내는, 상기 확

를 맵을 결정하는 단계; 및

상기 제 1 이미지에 대한 수정된 깊이 맵을 생성하는 단계로서, 상기 생성하는 단계는 제 1 픽셀에 대한 상기 연관된 깊이 맵의 제 1 깊이 값과 복수의 픽셀들에 대해 동일한 텍스트 이미지 오브젝트 깊이 값의 가중된 조합으로서 적어도 상기 제 1 픽셀에 대한 수정된 깊이 값을 결정하는 단계를 포함하고, 가중치는 상기 제 1 픽셀에 대한 상기 확률 값에 의존하는, 상기 수정된 깊이 맵을 생성하는 단계를 포함하는, 이미지에 대한 깊이 맵을 결정하는 방법.

## 청구항 15

프로그램이 컴퓨터 상에서 실행될 때, 제 14 항의 단계들 모두를 수행하도록 구성된 컴퓨터 프로그램 코드 수단들을 포함하는 컴퓨터 프로그램이 저장된 컴퓨터로 읽을 수 있는 저장매체.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 이미지를 위한 깊이 맵을 결정하기 위한, 특히 추정된 불일치 값들(disparity values)에 기초하여 깊이 맵을 결정하기 위한 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 3차원(3D) 디스플레이들은 시청되는 장면의 상이한 뷰들을 갖는 뷰어의 2개의 눈들을 제공함으로써 뷰잉 경험에 3차원을 추가한다. 이것은 사용자가 디스플레이 되는 2개의 뷰들을 분리하여 안경을 착용하게 함으로써 달성될 수 있다. 그러나, 이것은 사용자에게 불편한 것으로 간주될 수 있으므로, (렌티큘러 렌즈들, 또는 배리어들과 같은) 디스플레이에 수단을 사용하여 뷰들을 분리하고, 사용자의 눈에 개별적으로 도달할 수 있는 상이한 방향으로 상기 뷰들을 전송하는 오토스테레오스코픽(autostereoscopic) 디스플레이들을 사용하는 것이 많은 시나리오들에 있어서 바람직하다. 스테레오 디스플레이들에서, 2개의 뷰들이 요구되는 반면, 오토스테레오스코픽 디스플레이들에서는 전형적으로 더 많은 뷰들(예를 들어 9개의 뷰들과 같이)을 필요로 한다.

## 발명의 내용

### 해결하려는 과제

[0003] 그러나, 실제 디스플레이들은 이상적인 성능을 갖지 않는 경향이 있고, 전형적으로 완전한 3차원 이미지들을 나타낼 수 없다.

[0004] 예를 들어, 렌티큘러 기반 오토스테레오스코픽 3D 디스플레이들은 스크린 밖의 흐려짐(out-of-screen blur)을 겪는 경향이 있다. 이러한 효과는 카메라 시스템들에서의 피사체 심도 흐림(depth-of-field blur)으로 알려진 것과 유사하다.

[0005] 또한, 제시된 3차원 이미지의 품질은 수신된 이미지 데이터의 품질에 의존하고, 특히 3차원 지각은 수신된 깊이 정보의 품질에 의존한다.

[0006] 3차원 이미지 정보는 종종 장면에 대한 상이한 뷰 방향들에 대응하는 복수의 이미지들에 의해 제공된다. 특히, 영화들 또는 텔레비전 프로그램들과 같은 비디오 콘텐츠는 점점 일부 3D 정보를 포함하도록 생성된다. 이러한 정보는 약간 오프셋 된 카메라 위치들에서 2개의 동시 이미지들을 캡처하는 전용 3D 카메라들을 사용하여 캡처될 수 있다.

[0007] 그러나, 많은 애플리케이션들에서, 제공된 이미지들은 원하는 방향들에 직접 대응할 수 없거나, 더 많은 이미지들이 필요로 될 수 있다. 예를 들어, 오토스테레오스코픽 디스플레이들의 경우, 2개 이상의 이미지들이 필요로 되고, 실제로 9 내지 26개의 뷰 이미지들이 사용된다.

[0008] 상이한 뷰 방향들에 대응하는 이미지들을 생성하기 위해, 뷰 포인트 이동 처리(view point shifting processing)가 채용될 수 있다. 이것은 전형적으로 연관된 깊이 정보와 함께 단일 뷰 방향에 대한 이미지를 사용하는 뷰 이동 알고리즘에 의해 수행된다. 그러나, 중요한 아티팩트들(artifact) 없이 새로운 뷰 이미지들을 생성하기 위해, 제공된 깊이 정보는 충분히 정확해야 한다.

[0009] 안타깝게도, 많은 애플리케이션과 사용 시나리오들에서, 깊이 정보는 원하는만큼 정확하지 않을 수 있다. 실제

로, 많은 시나리오들에서, 깊이 정보는 상이한 뷰 방향들에 대한 뷰 이미지들을 비교함으로써 깊이 값들을 추정하고 추출함으로써 생성된다.

[0010] 많은 애플리케이션들에서, 3차원 장면들은 약간 다른 위치들에서 두 대의 카메라들을 사용하여 스테레오 이미지들로 캡처된다. 특정 깊이 값들은 두 이미지들 내의 대응하는 이미지 오브젝트들 간의 불일치들을 추정함으로써 생성될 수 있다. 그러나, 이러한 깊이 추출 및 추정에는 문제가 있고 바람직하지 않은 깊이 값들을 초래하는 경향이 있다. 이는 다시 아티팩트들 및 저하된 3차원 이미지 품질을 초래할 수 있다.

[0011] 3차원 이미지 저하 및 아티팩트들은 예를 들어, 자막 블록들(subtitle blocks)과 같은 텍스트 이미지 오브젝트들에 대해 특히 중요한 경향이 있다. 장면의 부분이 아닌, 텍스트 이미지 오브젝트들은 장면에 통합되거나 포함된 것으로 인식되지 않는 고립된 오브젝트들인 경향이 있다. 또한, 텍스트 이미지 오브젝트들에 대한 깊이 변화들은 뷰어에게 더 잘 인식되는 경향이 있다. 또한, 전형적인 애플리케이션에서, (특히 자막들과 같은) 텍스트는 잘 정의된 에지들을 갖고 선명하고 초점이 맞춰질 것으로 예상된다. 따라서, 특히 높은 이미지 품질을 갖는 자막 블록들과 같은 텍스트 이미지 오브젝트들을 제공하는 것이 중요하다.

[0012] 그러므로, 텍스트 이미지 오브젝트들에 대한 적절한 깊이 정보를 결정하기 위한 개선된 접근법이 유리할 것이고, 특히 증가된 유연성, 용이한 구현, 감소된 복잡성, 개선된 3D 경험 및/또는 향상되고 인지된 이미지 품질이 유리할 것이다.

### 과제의 해결 수단

[0013] 따라서, 본 발명은 상기 언급된 단점들 중 하나 이상을 단독으로 또는 임의의 조합으로 바람직하게 경감, 완화 또는 제거하려고 시도한다.

[0014] 본 발명의 일 양상에 따르면, 이미지를 위한 깊이 맵을 결정하기 위한 장치가 제공되고, 상기 장치는: 제 1 이미지의 적어도 일부 픽셀들에 대한 제 1 깊이 값들을 포함하는 연관된 깊이 맵을 제 1 이미지에 제공하는 이미지 유닛; 이미지의 적어도 일부 픽셀들에 대한 확률 값들을 포함하는 제 1 이미지를 위한 확률 맵을 결정하기 위한 확률 유닛으로서, 픽셀에 대한 상기 확률 값은 픽셀이 텍스트 이미지 오브젝트에 속할 확률을 나타내는, 상기 확률 유닛; 제 1 이미지를 위한 수정된 깊이 맵을 생성하기 위한 깊이 유닛으로서, 깊이 유닛은 제 1 픽셀에 대한 연관된 깊이 맵의 제 1 깊이 값과 복수의 픽셀들에 대해 동일한 텍스트 이미지 오브젝트 깊이 값의 가중된 조합으로서 적어도 제 1 픽셀에 대한 수정된 깊이 값을 결정하도록 구성되고, 가중치는 제 1 픽셀에 대한 확률 값에 의존하는, 상기 깊이 유닛을 포함한다.

[0015] 상기 접근법은 많은 실시예들에서 이미지 처리에 사용될 때 개선된 품질을 제공할 수 있는 개선된 깊이 맵이 생성되게 허용할 수 있다. 예를 들어, 많은 실시예들 및 시나리오들에서, 수정된 깊이 맵을 사용하여 자막들과 같은 텍스트 이미지 오브젝트들의 개선된 표현이 달성될 수 있다. 특히, 수정된 깊이 맵이 이미지 뷰 이동에 사용될 때 및/또는 오토스테레오스코픽 디스플레이 상에 3차원 이미지들을 제공할 때 상기 개선이 중요할 수 있다. 상기 접근법은 또한 저하들, 불일치들 및/또는 아티팩트들을 완화 또는 감소시키면서 텍스트 이미지 오브젝트들에 대한 표현 깊이의 이동을 허용할 수 있다.

[0016] 맵들은 전체 또는 부분 맵들일 수 있다. 예를 들어, 확률 맵은 단지 픽셀들/픽셀 그룹들의 서브세트에 대한 확률 값들을 포함할 수 있다. 주어진 픽셀에 대한 확률 값은 픽셀이 텍스트 이미지 오브젝트에 속할(그 일부인) 확률 또는 가능성의 추정을 나타낼 수 있다. 예를 들어, 확률 값은 속성의 결정된 값이 텍스트 이미지 오브젝트에 대해 기대되는 값과 얼마나 근접하게 일치하는지의 표시일 수 있다. 예를 들어, 처리는 제 1 이미지에 적용되어 값(또는 값들의 세트)을 생성할 수 있다. 픽셀이 텍스트 이미지 오브젝트에 속할 확률은 그 값이 기대값과 얼마나 근접하게 일치하는지에 의해 표현될 수 있다. 예를 들어, 처리는 예를 들어,  $a=0$  및  $b=1$ 인  $[a;b]$ 의 간격으로 숫자를 생성하는 제 1 이미지에 적용될 수 있다. 텍스트 이미지 오브젝트에 속하는 픽셀의 경우, 결정된 특성은  $b$ 인 것으로 예상될 수 있지만, 텍스트 이미지 오브젝트에 속하지 않는 픽셀의 경우, 값은  $a$ 인 것으로 예상될 수 있다. 따라서 결정된 값은 직접적으로 대응하는 픽셀이 텍스트 이미지 오브젝트에 속할 또는 속하지 않을 확률을 나타내는 특성 값일 수 있다. 예를 들어,  $a=0$  및  $b=1$ 인 경우, 값이 높을수록 대응하는 픽셀이 텍스트 이미지 오브젝트에 속할 가능성이 높다. 값의 정확한 처리 및 결정은 개별적인 실시예의 선호도와 및 요구사항들에 따라 상이한 실시예들에서 다를 수 있다.

[0017] 텍스트 이미지 오브젝트는 텍스트를 포함하는 이미지의 영역(region)/구역(area)/서브셋 또는 세그먼트일 수 있다. 텍스트 이미지 오브젝트는 특히 자막 이미지 오브젝트일 수 있다. 텍스트 이미지 오브젝트는 하나 이상의

문자들에 대응할 수 있고, 예를 들어 텍스트의 경계 영역을 포함할 수 있다.

- [0018] 가중된 조합은 구체적으로 확률 값의 함수인 가중치를 갖는 제 1 깊이 값과 텍스트 이미지 오브젝트 깊이 값의 가중된 합일 수 있다. 일부 실시예들에서, 가중된 조합은 구체적으로 가중치가 확률 값의 함수인 제 1 깊이 값의 단조 함수(monotonic function)와 텍스트 이미지 오브젝트 깊이 값의 가중된 합일 수 있다.
- [0019] 텍스트 이미지 오브젝트 깊이 값은 미리 결정된 값일 수 있다. 텍스트 이미지 오브젝트 깊이 값은 구체적으로 시간 및/또는 공간 도메인에서 고정 및/또는 상수 값일 수 있다. 텍스트 이미지 오브젝트 깊이 값은 텍스트 이미지 오브젝트들에 대해 원하는 깊이일 수 있다. 텍스트 이미지 오브젝트 깊이 값은 텍스트 이미지 오브젝트들에 대한 바람직한 깊이를 나타낼 수 있고 고정된 값 및/또는 미리 결정된 값일 수 있다. 구체적으로, 텍스트 이미지 오브젝트 깊이 값은 텍스트 이미지 오브젝트들에 대한 바람직한 깊이를 나타낼 수 있고, 제 1 이미지에 의해 표현되는 장면의 깊이 특성들과 독립적일 수 있다. 실제로, 텍스트 이미지 오브젝트들은 전형적으로 제 1 이미지에 의해 표현되는 장면의 일부가 아닌 오버레이 그래픽들일 수 있고, 텍스트 이미지 오브젝트 깊이 값은 장면의 일부가 아닌 오버레이 그래픽들/텍스트 이미지 오브젝트들에 대한 바람직한 깊이를 나타낼 수 있다. 텍스트 이미지 오브젝트 깊이 값은 텍스트 이미지 오브젝트들과 연관되고, 복수의 픽셀들 또는 픽셀 그룹들에 대해 동일할 수 있다. 많은 실시예들에서, 텍스트 이미지 오브젝트 깊이 값은 제 1 이미지 및/또는 연관된 깊이 맵의 특성들과 무관하다.
- [0020] 제 1 이미지는 상이한 시야각들에 대응하는 복수의 이미지들의 일부인 이미지일 수 있거나, 예를 들어 하나의 고정되고 독립적인 (연관된 깊이 맵을 갖는) 이미지일 수 있다. 일부 실시예들에서, 제 1 이미지는 예를 들어 비디오 시퀀스의 프레임과 같은 이미지들의 시간 시퀀스의 이미지일 수 있다.
- [0021] 본 발명의 선택적 특징에 따라, 확률 유닛은 제 1 이미지에 대한 픽셀 그룹들을 결정하도록 구성되고, 각 픽셀 그룹은 복수의 픽셀들을 포함하고, 상기 확률 유닛은 픽셀 그룹들에 대한 확률 값들을 결정하도록 구성된다.
- [0022] 이는 개선된 성능 및/또는 용이한 동작 및/또는 감소된 복잡성 및/또는 자원 요구를 제공할 수 있다. 픽셀 그룹에 대해 결정된 확률 값은 픽셀 그룹에 속하는 모든 픽셀들에 할당될 수 있다.
- [0023] 본 발명의 선택적인 특징에 따라, 깊이 유닛은 제 1 픽셀이 속하는 픽셀 그룹에 대한 제 1 확률 값 및 이웃하는 픽셀 그룹에 대한 제 2 확률 값에 응답하여 가중된 조합에 대한 가중치들을 결정하도록 구성되고; 상기 제 1 확률 값 및 상기 제 2 확률 값의 가중치는 상기 제 1 픽셀이 속하는 상기 픽셀 그룹 내의 상기 제 1 픽셀의 위치에 의존한다.
- [0024] 이것은 많은 시나리오들에서 개선된 성능을 허용할 수 있다. 특히, 상기 접근법은 깊이 레벨들의 조합에 대해 유효 해상도가 증가되도록 허용하면서 이미지 픽셀 해상도보다 낮은 해상도에서 확률 값들의 결정을 허용할 수 있다.
- [0025] 일부 실시예들에서 가중된 조합의 가중치들은 보간된 확률 값에 응답하여 결정될 수 있고, 보간된 확률 값은 제 1 픽셀이 속하는 픽셀 그룹에 대한 제 1 확률 값과 이웃하는 픽셀 그룹에 대한 제 2 확률 값 사이의 공간 보간(spatial interpolation)에 응답하여 결정된다; 상기 보간은 제 1 픽셀이 속하는 픽셀 그룹 내의 제 1 픽셀의 위치에 의존한다.
- [0026] 확률 맵은 복수의 확률 값들을 포함할 수 있고, 구체적으로 복수의 픽셀들 또는 픽셀 그룹들의 각각에 대한 복수의 확률 값들을 포함할 수 있다. 유사하게, 조합은 상이한 확률 값들을 가지며 따라서 독립적인 조합들을 갖는 적어도 일부 픽셀들을 갖는 각 픽셀에 대해 수행될 수 있다. 따라서, 확률 값들과 조합은 이미지에 따라 다를 수 있다(상이한 픽셀들마다 다를 수 있다).
- [0027] 본 발명의 선택적 특징에 따라, 확률 유닛은 이미지의 픽셀 그룹들에 대한 확률 값들을 결정하도록 구성되고, 각 픽셀 그룹은 적어도 하나의 픽셀을 포함한다; 상기 확률 유닛은 픽셀 그룹들이 텍스트 이미지 오브젝트에 속하는 것으로 지정되는지 또는 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정되는지 여부를 나타내는 픽셀 그룹들에 대한 값들을 포함하는 분류 맵을 먼저 결정하고, 분류 맵의 필터링에 응답하여 확률 맵을 생성하도록 구성된다.
- [0028] 이것은 용이한 구현 및/또는 개선된 성능/결과들을 제공할 수 있다.
- [0029] 본 발명의 선택적인 특징에 따라서, 분류 맵은 픽셀 그룹들에 대한 이진 값을 포함하고, 각 이진 값은 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로 지정되거나 픽셀 그룹이 텍스트 이미지 오브젝트에 속하지 않는



것으로 지정된 것을 나타낸다.

- [0030] 이것은 용이한 구현 및/또는 개선된 성능/결과들을 제공할 수 있다. 특히, 그것은 많은 시나리오들에서 보다 견고하고 신뢰성 있는 픽셀 그룹들의 초기 분류를 허용할 수 있다. 필터링은 이진 분류를 픽셀 그룹의 시간적 및/또는 공간적 이웃의 특성들을 또한 반영하는 비-이진 확률 값들로 변환할 수 있다.
- [0031] 본 발명의 선택적 특징에 따라서, 필터링은 분류 맵에 적용되는 이진 팽창 필터링(binary dilation filtering)을 포함한다.
- [0032] 이것은 성능을 향상시킬 수 있고, 특히 텍스트 이미지 오브젝트들에 대응하는 검출된 영역들의 일관성을 향상시킬 수 있다. 많은 시나리오들에서 이러한 영역들의 홀 형성을 줄일 수 있다.
- [0033] 본 발명의 선택적인 특징에 따라서, 필터링은 시간 필터링(temporal filtering)을 포함한다.
- [0034] 이것은 예를 들어 향상된 안정성 및 일관성을 허용하고 향상된 사용자 경험을 제공할 수 있다(예를 들어, 수정된 깊이 맵에 기초하여 뷰 이동에 의해 생성된 이미지들을 볼 때).
- [0035] 본 발명의 선택적 특징에 따라, 시간 필터링은 비대칭이다.
- [0036] 이것은 많은 애플리케이션들 및 시나리오들에서 개선된 성능을 제공할 수 있다.
- [0037] 본 발명의 선택적 특징에 따라, 필터링은 공간 필터링(spatial filtering)을 포함한다.
- [0038] 이것은 많은 애플리케이션들 및 시나리오들에서 개선된 성능을 제공할 수 있다.
- [0039] 본 발명의 선택적인 특징에 따라서, 공간 필터링은 최대 출력값 한계를 갖는 필터인 소프트맥스 필터(softmax filter)를 포함한다.
- [0040] 이것은 많은 애플리케이션들 및 시나리오들에서 개선된 성능을 제공할 수 있다. 특히, 이는 텍스트 이미지 오브젝트들에 대응하는 영역들에 유리한 깊이 프로파일들이 생성되게 하는 것을 허용할 수 있다. 예를 들어, 많은 시나리오들에서, 소프트맥스 필터는 홀들의 양 또는 크기를 감소시키면서 및/또는 텍스트 이미지 오브젝트의 에지들에서 소프트 전이를 제공하면서 텍스트 이미지 오브젝트에 대응하는 일정한 깊이 영역의 생성을 허용할 수 있다.
- [0041] 소프트맥스 필터는 출력을 최댓값으로 제한하는 리미터와 필터의 캐스케이드 일 수 있다. 예를 들어, 소프트맥스 지역-통과 필터는 지역 통과 필터의 캐스케이드 및 지역 통과 필터의 출력 값을 최댓값으로 제한하는 리미터에 대응하는 동작을 수행할 수 있다. 그러므로, 소프트맥스 필터는 최대 출력값 한계를 갖는 필터에 대응할 수 있다.
- [0042] 본 발명의 선택적인 특징에 따르면, 공간 필터링은 적어도 두 개의 순차 공간적 소프트맥스 필터들(sequential spatial softmax filters)을 포함한다.
- [0043] 이것은 많은 애플리케이션들 및 시나리오들에서 개선된 성능을 제공할 수 있다. 특히, 텍스트 이미지 오브젝트들에 대응하는 영역들에 유리한 깊이 프로파일들이 생성되도록 허용할 수 있다. 예를 들어, 많은 시나리오들에서, 소프트맥스 필터들은 텍스트 이미지 오브젝트에 대응하는 일정한 깊이 영역의 생성을 허용할 수 있다. 많은 실시예들에서, 필터들은 텍스트 이미지 오브젝트의 에지들에서 소프트 전이를 또한 제공하면서 홀들의 양 또는 크기를 감소시킬 수 있다.
- [0044] 두 개의 순차 공간적 소프트맥스 필터들은 상이한 설계/동작 파라미터들을 갖도록 구성될 수 있고, 특히 커널 치수, 스케일 팩터 및/또는 최댓값은 두 개의 필터들에 대해 상이할 수 있다. 많은 실시예들에서, 제 1 필터에 대한 파라미터들은 홀 형성을 감소시키고 증가된 일관성을 제공하도록 최적화될 수 있는 반면, 제 2 필터의 파라미터들은 텍스트 이미지 오브젝트의 에지들에서 원하는 깊이 전이 프로파일을 제공하도록 최적화될 수 있다.
- [0045] 본 발명의 선택적인 특징에 따라서, 제 1 깊이 값들은 제 1 이미지의 장면에 대한 상이한 뷰 방향들에 대응하는 복수의 이미지들 내의 이미지 오브젝트들에 대한 불일치들에 대응한다.
- [0046] 본 발명은 많은 실시예들에서 불일치 추정에 의해 생성된 깊이 맵을 개선할 수 있고, 구체적으로 텍스트 이미지 오브젝트들과 연관된 저하들 및 아티팩트들을 완화시킬 수 있다.
- [0047] 일부 실시예들에서, 이미지 유닛은 복수의 이미지들 내의 이미지 오브젝트들에 대한 불일치들을 추정하도록 구성된다.



- [0048] 본 발명의 선택적인 특징에 따라서, 텍스트 이미지 오브젝트 깊이의 가중치는 증가되고, 제 1 깊이 값의 가중치는 확률값의 증가하는 값에 대해 감소된다.
- [0049] 이것은 많은 애플리케이션들 및 시나리오들에서 개선된 성능을 제공할 수 있다.
- [0050] 본 발명의 일 양상에 따르면, 이미지를 위한 깊이 맵을 결정하는 방법이 제공되고, 상기 방법은: 제 1 이미지의 적어도 일부 픽셀들에 대한 제 1 깊이 값들을 포함하는 연관된 깊이 맵을 제 1 이미지에 제공하는 단계; 이미지의 적어도 일부 픽셀들에 대한 확률 값들을 포함하는 제 1 이미지를 위한 확률 맵을 결정하는 단계로서, 픽셀에 대한 확률 값은 픽셀이 텍스트 이미지 오브젝트에 속할 확률을 나타내는, 상기 확률 맵을 결정하는 단계; 및 제 1 이미지를 위한 수정된 깊이 맵을 생성하는 단계로서, 상기 생성하는 단계는 제 1 픽셀에 대한 연관된 깊이 맵의 제 1 깊이 값과 복수의 픽셀들에 대해 동일한 텍스트 이미지 오브젝트 깊이 값의 가중된 조합으로서 적어도 제 1 픽셀에 대한 수정된 깊이 값을 결정하는 단계를 포함하고, 상기 가중치는 제 1 픽셀에 대한 확률 값에 의존하는, 상기 수정된 깊이 맵을 생성하는 단계를 포함한다.
- [0051] 본 발명의 이들 및 다른 양상들, 특징들, 및 이점들은 이하 설명된 실시예(들)로부터 명백해지고 이들을 참조하여 설명될 것이다.
- [0052] 본 발명의 실시예들은 오직 도면들을 참조하여 예시로서 기술될 것이다.

### 도면의 간단한 설명

- [0053] 도 1은 종래 기술의 일부 실시예들에 따른 장치를 포함하는 디스플레이 시스템의 예시도;
- 도 2는 오토스테레오스코픽 디스플레이로부터 투사되는 뷰 이미지들의 예를 도시한 도면;
- 도 3은 이미지 및 연관된 검출된 텍스트 영역들의 예를 도시한 도면;
- 도 4는 깊이 맵에 자막 블록을 위치시키는 예를 도시한 도면;
- 도 5는 본 발명의 일부 실시예들에 따른 장치에서 입력 이미지 및 그로 인해 생성된 확률 맵들의 예를 도시한 도면.

### 발명을 실시하기 위한 구체적인 내용

- [0054] 이하 설명은 장면의 상이한 뷰 방향들에 대한 이미지들을 생성하는 데 사용하기 위한 수정된 깊이 맵을 결정하기 위한 시스템에 적용가능한 본 발명의 실시예들(예를 들어, 오토스테레오스코픽 디스플레이 상에 입력 스테레오 이미지의 표현을 위한 추가적인 이미지들을 생성하기 위한 접근법과 같은)에 초점을 맞춘다. 그러나, 본 발명이 이러한 애플리케이션에 제한되지 않고 많은 다른 애플리케이션들 및 시스템들에 적용될 수 있다는 것이 이해될 것이다.
- [0055] 도 1은 윈도우의 일부 실시예들에 따른 시스템의 예를 도시한다. 특정 예에서, 오토스테레오스코픽 디스플레이(101)의 상이한 뷰들에 대응하는 이미지들은 입력 3차원 이미지로부터 생성된다. 입력 3차원 이미지는 예를 들어 연관된 깊이 맵을 갖는 단일 이미지로 표현될 수 있거나, 예를 들어 연관된 깊이 맵이 추출되는 스테레오 이미지들로 표현될 수 있다. 일부 실시예들에서, 이미지는 비디오 시퀀스/신호로부터의 프레임과 같은 이미지들의 시간 시퀀스로부터의 이미지일 수 있다.
- [0056] 일반적으로, 오토스테레오스코픽 디스플레이는 각각의 콘이 장면의 상이한 시야각들에 대응하는 다수의 뷰들을 포함하는 뷰들의 "콘들(cones)"을 생성한다. 인접한 (또는 더 바깥 일부 경우들에서) 뷰들 간의 시야각 차이는 사용자의 오른쪽 눈과 왼쪽 눈 사이의 시야각 차이에 대응하여 생성된다. 따라서, 왼쪽 및 오른쪽 눈이 두 개의 적절한 뷰들을 보는 뷰어는 3차원 효과를 인지할 것이다. 각각의 뷰잉 콘에서 9개의 상이한 뷰들을 생성하는 오토스테레오스코픽 디스플레이의 예시가 도 2에 도시된다.
- [0057] 오토스테레오스코픽 디스플레이는 렌티큘러 렌즈들 또는 패럴렉스 배리어들/배리어 마스크들과 같은 수단들을 사용하여 뷰들을 분리하고 뷰들을 상이한 방향으로 전송하여 그들이 개별적으로 사용자의 눈들에 도달하도록 하는 경향이 있다. 스테레오 디스플레이들의 경우, 두 개의 뷰들이 필요하지만 대부분의 오토스테레오스코픽 디스플레이는 전형적으로 더 많은 뷰들을 사용한다. 실제로, 일부 디스플레이들에서, 이미지의 상이한 부분들이 상이한 뷰잉 방향으로 투영될 수 있도록 이미지에 대해 뷰 방향들의 점진적인 전이가 수행된다. 그러므로, 일부 더욱 최근의 오토스테레오스코픽 디스플레이들에서, 고정된 수의 완전한 뷰들을 렌더링하는 오토스테레오스코픽 디스플레이보다 뷰 방향들에 대한 이미지 영역들의 더 점진적이고 연속적인 분포가 적용될 수 있다. 이러한 오

토스테레오스코픽 디스플레이는 전체 뷰들이 아닌 부분 뷰들을 제공하는 것으로 종종 지칭된다. 부분 뷰들에 대한 더 많은 정보는 예를 들어 WO 2006/117707에서 찾을 수 있다.

- [0058] 그러나, 대부분의 오토스테레오스코픽 디스플레이들에 대해 공통점은 그들이 상대적으로 많은 수의 상이한 뷰 방향들에 대해 생성된 이미지 정보를 필요로 한다는 것이다. 그러나, 전형적으로 3차원 이미지 데이터는 스테레오 이미지 또는 깊이 맵을 갖는 이미지로서 제공된다. 필요한 뷰 방향들을 생성하기 위해, 이미지 뷰 이동 알고리즘들은 전형적으로 렌더링을 위한 적합한 픽셀 값들을 생성하도록 적용된다. 그러나, 그러한 알고리즘들은 전형적으로 차선책이고, 아티팩트들 또는 왜곡들을 도입할 수 있다.
- [0059] 발명자들은 이러한 아티팩트들, 열화들, 및 품질 저하들이 특히 자막 이미지 오브젝트들과 같은 텍스트 이미지 오브젝트들과 관련하여 특히 보편적이고, 인지 가능하고, 및/또는 중요할 수 있다는 것을 인식했다. 그러므로, 텍스트를 포함하는 이미지들이 수신되어 표시될 때, 예를 들어 오토스테레오스코픽 디스플레이에 의해, 아티팩트들이 종종 자막 블록들 및 유사한 텍스트 오브젝트들 주의에서 인지될 수 있다.
- [0060] 이러한 현안들을 해결하기 위해, 알고리즘들은 이러한 이미지 오브젝트들을 식별하는 것을 시도하도록 적용될 수 있다. 따라서 고정된 깊이 레벨은 모든 텍스트를 고정된 깊이에 위치시키기 위해 식별된 이미지 오브젝트들에 적용될 수 있다. 그러나, 그러한 검출은 전형적으로 매우 어렵고, 텍스트 이미지 오브젝트들(및 비-텍스트 이미지 오브젝트들)로의 정확한 세분화는 전형적으로 실행 가능하지 않다. 결과적으로, 그러한 접근법들은 또한 종종 인지 가능한 아티팩트들을 초래한다.
- [0061] 예를 들어, 자막 박스가 식별될 수 있고, 이러한 박스의 깊이 레벨이 스크린 깊이로 설정될 수 있다. 도 3은 이를 어떻게 행하는지를 도시한다.
- [0062] 자막 박스는 스크린 깊이에 자막을 배치하는데 사용된다. 도 3은 이를 어떻게 행하는지를 도시한다. 이 예에서, 자막 박스는 자막 텍스트를 갖는 휘도 이미지(도 3a)에서 검출된다. 이 예에서, 각 8x8 블록 내에서 추출된 피쳐들을 사용하여 8x8 블록의 픽셀들 당 검출이 수행될 수 있다. 각각의 블록 블록은 먼저 자막의 일부(도 3b에서는 흰색으로 표시됨)로 분류되거나 분류되지 않을 수 있다(도 3b에서는 흑색으로 표시됨). 마지막으로, 검출 결과들을 먼저 수평적으로 적분하고 y-좌표를 따라 시작 및 정지 위치를 찾은 다음 검출 결과들을 수평적으로 적분하고 x-좌표의 시작 및 정지 위치를 찾음으로써 바운딩 자막 박스가 결정된다. 구체적으로, 자막 박스의 경계들은 자막 박스에 속하는 것으로 분류된 모든 블록들을 포함하는 가장 작은 직사각형으로 설정될 수 있다.
- [0063] 그러므로, 자막 박스 내의 모든 픽셀들에 대한 깊이는 특정 깊이 값으로 설정될 수 있고 구체적으로 스크린 또는 디스플레이 깊이로 설정될 수 있다. 이는 아티팩트들을 감소시킬 수 있고, 스크린 레벨에 있지 않은 오브젝트들에 오토스테레오스코픽 디스플레이로 인한 흐림 현상을 줄일 수 있다. 접근법은 스크린 깊이 뒤에 있는 섹션(401) 및 스크린 깊이 앞에 있는 섹션(403)을 갖는 이미지를 도시하는 도 4에 의해 도시될 수 있다. 기술된 접근법은 자막 블록(405)을 식별하고 이것을 스크린 깊이에 위치시킬 수 있다.
- [0064] 그러나, 이러한 솔루션이 날카로운 것(구체적으로, 스크린 외부 흐림 효과 없음) 및 기하학적으로 올바른 자막 오브젝트를 생성할 동안, 깊이 맵 내의 새롭게 도입된 가파른 에지들(텍스트 박스로 인한)은 종종 텍스트 박스(도 4에서 (407)로 표시되는)의 경계 부근에서 매우 선명한 텍스처 왜곡들(브레이크-업)을 유발할 수 있다. 게다가, 텍스트 박스 자체가 여러 개의 작은 텍스트 박스들로 분리될 때 렌더링 오류들이 표시될 것이다. 예를 들어, 한 문장의 일부 문자들은 날카로울 수 있고 다른 문자들은 흐려질 수 있다. 더욱이, 이러한 행동은 시간이 지남에 따라 급속히 변할 수 있고 실제로 이는 관찰자에게 잘 보여질 수 있다.
- [0065] 실제로 자막 영역을 식별하고 (특히) 이러한 영역의 깊이를 설정하는 것은 일반적으로 다양한 문제들을 야기한다는 것이 밝혀졌다. 먼저, 검출 알고리즘의 노이즈는 종종 검출된 박스가 두 개 이상의 개별 박스들로 떨어지는 것을 야기할 수 있다. 이것은 x 및/또는 y 좌표를 따라 가능하게 존재하는 다수의 시작 및 중지 위치들로 인해 발생할 수 있다. 게다가, 검출된 박스의 시작 및 중지 위치는 분류 에러들에 민감한 경향이 있다. 결과적으로, 결과적인 깊이 맵은 일시적으로 불안정해지고 오토스테레오스코픽 디스플레이에 의해 디스플레이된 이미지들에 가시적인 시간적 에러들을 유발할 수 있다.
- [0066] 도 1의 장치는 공지된 접근법들로부터 공지된 단점들의 일부를 완화하거나 감소시킬 수 있다. 장치는 오토스테레오스코픽 디스플레이(101)를 구동하는 디스플레이 드라이버(103)를 포함한다. 디스플레이 드라이버(103)는 오토스테레오스코픽 디스플레이(101)에 대한 복수의 뷰 이미지들을 생성하고 이들을 그에 공급하기 위한 기능을 포함한다. 뷰 이미지들은 입력이 입력 이미지 및 깊이 맵을 갖는 뷰 이동 알고리즘을 사용하여 상이한 뷰 방향들에 대해 생성된다.

- [0067] 그러나, 단지 자막 영역을 스크린 깊이로 식별하여 설정하는 것보다, 시스템은 입력 깊이 맵(외부 소스 또는 내부 소스로부터 수신되고, 구체적으로 스테레오 이미지와 같은 상이한 뷰 방향들에 대응하는 이미지들로부터 디스플레이 드라이버(103)에 의해 종종 생성되는)을 처리하여 자막 깊이의 보다 미세한 재구성이 사용되는 수정된 깊이 맵을 생성하도록 구성된다. 이것은 자막들이나 다른 텍스트 이미지 오브젝트들 주위에 가시적인 아티팩트들의 감소 또는 완화를 포함하는 충분히 개선된 시각 경험을 제공하는 경향이 있을 것이다.
- [0068] 디스플레이 드라이버(103)의 접근법은 먼저 대응하는 픽셀들이 텍스트 이미지 오브젝트에 속하는지 또는 속하지 않는지 여부를 나타내는 확률 값들을 포함하는 확률 맵( $\alpha$  맵이라고도 칭해짐)을 생성하는 단계를 포함한다. 따라서, 주어진 픽셀, 또는 픽셀 그룹이 텍스트 오브젝트인지 아닌지를 결정하는 단순한 이진 지정이기 보다는, 확률 맵은 비-이진수, 및 종종 픽셀, 또는 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는지 또는 속하지 않는지 추정된 확률을 나타내는 실질적으로 연속적인 값들을 제공한다.  $\alpha$  값들로 지칭되는 확률맵의 확률 값들은 일반적으로 이산 값들로 표현되며, 대부분의 실시예들에서 각 값은 적어도 4이지만 일반적으로 적어도 8, 16, 32, 64, 128, 256 또는 그 이상의 이산 레벨들을 갖는 값들로 표현될 수 있다. 많은 실시예들에서, 각 확률 값은 적어도 3, 4, 6, 8, 10 또는 16 비트들의 이진 값으로 표현될 수 있다.
- [0069] 이 접근법에서, 적어도 일부 픽셀들에 대한 깊이 레벨은 원래의 깊이 값을 선택하거나 이를 미리 결정된 깊이 레벨(스크린 레벨과 같은)로 설정함으로써 단순히 생성되지 않는다. 오히려, 수정된 깊이 맵의 적어도 일부 값이 값들은 원래의 깊이 값과 특히 텍스트 이미지 오브젝트들에 대한 바람직한 깊이 레벨(스크린 깊이와 같은)인 텍스트 이미지 오브젝트 깊이 값의 가중된 조합을 수행함으로써 생성된다. 주어진 픽셀에 대한 깊이 레벨들의 가중치는 상기 픽셀에 대한 확률 값에 기초하여 결정된다.
- [0070] 구체적으로,  $\alpha[i, j]$ 는 픽셀 위치  $[i, j]$ 에서 확률 맵의 값을 나타내고, 수정된 깊이 맵에 대한 수정된 깊이 레벨은 다음과 같이 결정될 수 있다:
- [0071] 
$$D_{\text{mod}}[i, j] = \alpha[i, j]D_{\text{text}} + (1 - \alpha[i, j])D[i, j].$$
- [0072] 여기서,  $D_{\text{text}}$ 는 텍스트 이미지 오브젝트 깊이를 나타내고,  $D[i, j]$ 는 원래의 깊이 레벨이다.
- [0073] 따라서, 이 접근법에서, 수정된 깊이 맵은 원래의 깊이 레벨과 텍스트 이미지 오브젝트들에 대해 요구되는 깊이 레벨 간의 차이를 더욱 연속적으로 확대하는 깊이 레벨들을 제공하도록 생성된다. 이런 방식으로, 텍스트 이미지 오브젝트는 효율적으로 원래의 깊이 표현과 원하는 깊이 사이에서 (깊이 방향으로) 공간적으로 "혼합"될 수 있다.
- [0074] 이러한 접근법은 실제로 구체적으로 자막 블록들과 같은 텍스트 이미지 오브젝트들 주위의 아티팩트들 및 저하들의 현저하게 감소된 인식으로 상당히 더 매력적이고 시각적 경험들을 제공하는 것으로 밝혀졌다.
- [0075] 이러한 접근법의 다른 이점은 텍스트 이미지 오브젝트 검출 결과들을 사용하여 확률 맵을 생성함으로써, 결과적인 렌더링 이미지들을 개선하기 위해 공간 및 시간 필터들을 적용할 수 있다는 것이다. 또한, 이러한 필터링은 원래 깊이 맵에 영향을 미치지 않고 수행될 수 있다.
- [0076] 디스플레이 드라이버(103)는 이미지의 적어도 일부 픽셀들에 대한 깊이 값들을 포함하는 연관된 깊이 맵을 이미지에 제공하도록 구성된 이미지 유닛(105)을 포함한다. 이미지는 입력 이미지로 지칭되고, 깊이 맵은 입력 깊이 맵(그리고 픽셀 및 깊이 값들은 입력 픽셀 및 깊이 값들로) 지칭될 것이다.
- [0077] 이미지 유닛(105)은 일부 실시예들에서 임의의 적합한 내부 또는 외부 소스로부터 입력 이미지 및 입력 깊이 맵을 단순히 수신하도록 구성될 수 있다. 예를 들어, 이미지로 표현되는 3차원 이미지들 및 연관된 깊이 맵을 포함하는 비디오 신호는 네트워크(인터넷과 같은), 방송 신호, 미디어 캐리어 등으로부터 수신될 수 있다.
- [0078] 일부 실시예들에서, 이미지 유닛(105)은 입력 이미지 및/또는 입력 깊이 맵을 생성하도록 구성될 수 있다. 구체적으로, 많은 실시예들에서, 이미지 유닛(105)은 구체적으로 스테레오 이미지와 같은 상이한 뷰 방향들에 대응하는 복수의 이미지들에 의해 형성된 3차원 이미지를 수신하도록 구성될 수 있고, 이는 입력 깊이 맵을 생성하도록 불일치 추정을 수행하도록 구성될 수 있다.
- [0079] 깊이 맵들에 포함된 깊이 값들은 구체적으로 깊이 좌표( $z$ ) 값 또는 상이한 뷰 방향들의 이미지들 사이의 이동들을 나타내는 불일치 값과 같은 임의의 적절한 깊이의 표시일 수 있다.
- [0080] 이 예에서, 입력 깊이 값들은 관찰되거나 추정된 깊이 값들이다. 특히, 이들은 외부 소스에서 또는 이미지 유닛

(105) 자체에 의해 수행된 불일치 추정에 의해 생성된 값들 일 수 있다. 따라서, 깊이 값들은 종종 상대적으로 부정확하고, 많은 실수들과 오류들을 포함할 수 있다.

[0081] 또한, 입력 신호에서, 자막들과 같은 텍스트 이미지 오브젝트는 오토스테레오스코픽 디스플레이 상에서 렌더링 될 때 바람직한 깊이에서 위치될 수 없다. 이는 예를 들어 상이한 디스플레이들에 대해 변하는 기술적으로 최적의 깊이(예를 들어 선명함에 관해)로 인해 또는 애플리케이션들 사이에서 변하는 주관적인 선호도로 인해 발생할 수 있다.

[0082] 예를 들어, 텔레비전 프로그램들 및 영화들에 경우, 자막들은 일반적으로 스크린의 앞쪽, 즉 자막들이 스크린의 앞에 보이도록 하는 깊이 레벨, 그리고 종종 가장 전방의 오브젝트로서 위치된다. 그러나, 많은 오토스테레오스코픽 디스플레이들에서, 이러한 깊이 레벨들에 대해 상당히 증가된 흐름이 도입되고, 그러므로, 스크린 깊이에 자막들을 배치하는 것이 바람직할 수 있다. 기술된 바와 같이, 디스플레이 드라이버(103)는 추정된 텍스트 이미지 오브젝트들에 대한 깊이 레벨들의 점진적인 재조정을 수행할 수 있다.

[0083] 구체적으로, 이미지 유닛(105)은 대응하는 픽셀이 텍스트 이미지 오브젝트에 속할 확률 또는 가능성을 나타내는 확률 값들을 포함하는 확률 맵을 생성하도록 구성되는 확률 유닛(107)에 연결된다.

[0084] 주어진 픽셀에 대한 확률 값은 텍스트 이미지 오브젝트에 속할 확률 또는 가능성의 추정치를 나타낼 수 있다. 확률은 예를 들어 픽셀을 포함하는 픽셀 그룹이 텍스트 이미지 오브젝트들에 대한 예상된 특성들의 세트와 얼마나 근접하게 일치하는지의 표시일 수 있다. 일치가 가까울수록 픽셀이 텍스트 이미지 객체에 속할 확률이 높아진다.

[0085] 텍스트 이미지 오브젝트에 속하는 픽셀의 확률은 텍스트 이미지 오브젝트가 텍스트 문자의 일부인 픽셀들만을 포함한다는 고려에 기초할 수 있다. 그러나, 많은 실시예들에서, 텍스트 이미지 오브젝트는 일반적으로 경계 박스를 포함하는 것으로 간주될 수 있다. 예를 들어, 텍스트 이미지 오브젝트는 주변 자막 박스를 포함하는 것으로 간주될 수 있다. 텍스트 이미지 오브젝트를 나타내는 것으로 고려되는 정확한 특징들 및 특성들은 개별적인 실시예의 특정 선호도들 및 요구사항들에 의존할 것이고, 특히 확률 값들의 결정은 바람직한 특징들 및 텍스트 박스로 여겨지는 것의 특성들을 반영하도록 적용될 수 있다는 것이 이해될 것이다.

[0086] 예를 들어, 일부 실시예들에서, 자막들은 예를 들어, 회색 또는 검은색 박스에 흰색 문자들로 제공될 수 있다. 그러한 예에서, 텍스트 이미지 오브젝트가 전체 자막 박스를 포함하는 것으로 고려하는 것이 매우 바람직할 수 있고, 즉 확률은 원하는 깊이로 자막 박스에 점진적으로 위치시키는 결과를 초래하는 픽셀들이 자막 박스에 속할 확률을 나타내도록 결정될 수 있다. 다른 애플리케이션들에서, 자막들은 기본 이미지 위에 단순히 흰색 문자들로 제공될 수 있다. 그러한 예에서, 경계 박스는 존재하지 않고, 확률은 픽셀이 텍스트 문자에 속할 확률을 반영할 수 있다.

[0087] 많은 실시예들에서, 텍스트 이미지 오브젝트는 특히 그래픽 오버레이 이미지 객체일 수 있고, 특히 자막 또는 제목 이미지 오브젝트일 수 있다.

[0088] 확률 맵을 결정하기 위한 다양한 접근법들이 상이한 실시예들에서 사용될 수 있고, 고성능을 갖는 고효율 동작을 제공하는 특정 예가 후술될 것이라는 것이 이해될 것이다.

[0089] 확률 유닛(107)은 수정된 깊이 맵을 생성하도록 구성된 깊이 유닛(109)에 연결된다. 깊이 유닛(109)은 이미지 유닛(105) 및 텍스트 이미지 오브젝트 깊이 값이 저장되는 메모리/저장소(111)에 더 연결된다. 깊이 유닛(109)은 따라서 입력 깊이 맵, 텍스트 이미지 오브젝트 깊이, 및 확률 맵을 수신하고, 가중치가 확률 값에 의존하는 입력 깊이 값들 및 텍스트 이미지 오브젝트 깊이 값들의 가중된 조합을 수행하도록 진행한다.

[0090] 특히, 이전에 기술된 바와 같이, 조합은 확률 값에 의존하는 가중치를 갖는 입력 깊이 값과 텍스트 이미지 오브젝트 깊이 값의 가중된 합과 같은 선형 조합일 수 있다. 그러나, 입력 깊이 레벨 및 텍스트 이미지 오브젝트 깊이 레벨로부터의 상대적 기여도가 확률 값에 의존하는 경우, 다른 가중치 조합들이 사용될 수 있음이 이해될 것이다. 예를 들어, 일부 실시예들에서, 비-선형 조합들이 사용될 수 있다.

[0091] 따라서, 조합(주어진 픽셀에 대한)은 확률 값에 의존하는 입력 깊이 값 및 텍스트 이미지 오브젝트 깊이 값 각각으로부터 출력 깊이 값에 기여하는 입력 깊이 값, 텍스트 이미지 오브젝트 깊이 값, 및 확률 값(픽셀에 대한)의 함수로서 출력 깊이 값을 제공한다.

[0092] 가중된 조합은 텍스트 이미지 오브젝트 깊이가 증가되고 제 1 깊이 값의 가중치가 확률 값의 증가하는 값에 대해 감소되는 것이다. 따라서, 주어진 픽셀이 텍스트 이미지 오브젝트에 속하는 것으로 추정될 가능성이 높을수



록, 출력 깊이 값은 텍스트 이미지 오브젝트 깊이 값에 더 가까워질 것이고, 주어진 픽셀이 텍스트 이미지 오브젝트에 속하는 것으로 추정될 가능성이 낮을수록, 출력 깊이 값은 입력 깊이 값에 더 가까워질 것이다. 정확한 관계는 개개의 실시예의 선호도들과 요구사항들에 따라 달라질 것이다.

[0093] 많은 실시예들에서, 출력 깊이 값은 주어진 픽셀에 대해 함수가 0의 확률 값에 대해 픽셀에 대한 입력 깊이 값을 출력하고 하나의 확률에 대해 텍스트 이미지 오브젝트 깊이 값을 출력하는 함수일 확률 값의 함수일 수 있다. 확률 값의 함수는 구체적으로 단조 함수일 수 있다. 확률 값의 함수는 주어진 픽셀 맵에 대해 픽셀에 대한 입력 깊이 값으로부터 텍스트 이미지 오브젝트 깊이 값인 깊이 값 범위에 대한 0부터 1까지의 확률 범위일 수 있다.

[0094] 대부분의 실시예들에서, 텍스트 이미지 오브젝트 깊이 값은 미리 결정된 값일 수 있고, 구체적으로 모든 이미지들에 대한 및/또는 전체 이미지에 대한 상수 값일 수 있다. 그러나, 일부 실시예들에서, 텍스트 이미지 오브젝트 깊이 값은 이미지의 상이한 영역들 사이에서 달라질 수 있다(예를 들어, 이미지의 하부에 대한 바람직한 텍스트 이미지 오브젝트 깊이 값은 이미지의 상부에 대한 바람직한 텍스트 이미지 오브젝트 깊이 값과 상이할 수 있다).

[0095] 깊이 유닛(109)은 예시에서, 오토스테레오스코픽 디스플레이(101)에 대한 뷰 이미지들을 생성하도록 구성된 이미지 생성기(113)에 연결된다. 이미지 생성기(113)는 깊이 유닛(109)으로부터 수정된 깊이 맵을 수신하고 입력 이미지를 수신하는 이미지 유닛(105)에 더 연결된다. 이미지 생성기(113)는 오토스테레오스코픽 디스플레이(101)에 의해 생성된 상이한 뷰들과 연관된 특정 뷰 방향들에 대한 뷰 이미지들을 생성하기 위해 뷰 이동을 수행함으로써 오토스테레오스코픽 디스플레이(101)에 대한 뷰 이미지들을 생성하도록 구성된다. 이미지 생성기(113)는 입력 이미지 및 수정된 깊이 맵에 기초한 뷰 이동 알고리즘에 의해 이들 이미지들을 생성하도록 구성된다. 따라서, 뷰 이미지들은 텍스트 이미지 오브젝트들에 속하는 것으로 또는 속하지 않는 것으로 간주되는지 여부를 점진적으로 반영하는 깊이들로 픽셀들을 제공하기 위해 생성될 것이다. 결과적으로, 시청자에 의해 인식된 투영된 3차원 이미지는 텍스트 이미지 오브젝트들과 다른 이미지 영역들 사이의 더 많은 점진적 전이를 갖는 텍스트 이미지 오브젝트의 더 많이 일관된 렌더링을 갖는 경향이 있을 것이다. 이는 전형적으로 3D 프리젠테이션의 인식된 불완전함들을 실질적으로 감소시킬 것이다.

[0096] 따라서, 이 예에서, 입력 이미지 및 수정된 깊이 맵은 이미지 생성기(113)에 의해 사용되어 입력 이미지의 장면의 이미지를 생성하지만 입력 이미지의 뷰 방향과 상이한 뷰 방향을 갖는다.

[0097] 당업자는 많은 다른 뷰 이동 알고리즘들을 인식하고 임의의 적절한 알고리즘이 본 발명을 손상시키지 않고 사용될 수 있음을 이해할 것이다.

[0098] 일부 실시예들에서, 이미지 유닛(105)은 입력 이미지와 함께 깊이 맵을 직접 수신할 수 있다. 일부 경우들에서, 깊이 맵들은 텍스트 이미지 오브젝트들이 예를 들어 캡처된 이미지에 포함되거나 추가될 때와 동일한 시간 및 장소에서 생성될 수 있다. 일부 경우들에서, 깊이 맵은 그러므로 특정한 바람직한 깊이에 있는 자막들과 같은 텍스트 이미지 오브젝트들에 대한 깊이 레벨들을 갖도록 생성될 수 있다. 그러므로, 깊이 맵은 텍스트 이미지 오브젝트에 대해 텍스트 이미지 오브젝트의 영역에 대응하는 완전하고 일관된 영역을 가질 수 있고, 이러한 영역의 깊이 값들은 모두 동일할 수 있다. 이것은 텍스트 이미지 오브젝트가 거의 에러가없고 적은 아티팩트들로 배치되는 것을 허용할 수 있다. 그러나, 이 경우에도 텍스트 이미지 오브젝트의 깊이 레벨은 특정 사용 시나리오의 바람직한 깊이 레벨이 아닐 수 있다. 예를 들어, 안경 기반 접근법을 사용하여 시청하기에 적합한 깊이 레벨은 오토스테레오스코픽 디스플레이를 사용하는 프리젠테이션에 이상적이지 않을 수 있다. 또한, 선명한 깊이 전이들은 여전히 뷰 이동을 수행할 때 아티팩트들을 유발할 수 있다.

[0099] 더욱이, 많은 애플리케이션들에서, 깊이 맵은 동시에 생성되지 않을 수 있고, 텍스트 이미지 오브젝트(자막들과 같은)의 포함으로 배치된다. 예를 들어, 많은 시나리오들에서, 3차원 이미지는 두 이미지들에 포함된 자막들을 갖고, 두 스테레오 이미지들 사이의 불일치에 의해 제어되는 자막들의 깊이를 갖는 스테레오 이미지들에 의해 표현될 수 있다. 이러한 3차원 스테레오 표현은 임의의 깊이 맵을 포함하지 않고, 따라서 필요하다면, 그러한 맵은 후-처리 동작으로 생성될 수 있다. 실제로, 많은 시나리오들에서, 깊이 맵은 최종 사용자 디바이스에서 생성될 수 있다.

[0100] 예를 들어, 이미지 유닛(105)은 동일한 장면에 대한 상이한 뷰 방향들에 대응하는 복수의 이미지들을 수신하도록 구성될 수 있다. 구체적으로, 좌측 눈 이미지와 우측 눈 이미지를 포함하는 스테레오 이미지가 수신될 수 있다. 스테레오 이미지는 텍스트 이미지 오브젝트, 예를 들어 자막을 포함할 수 있고, 이것의 깊이는 두 이미지들

사이의 불일치에 의해 반영된다.

- [0101] 이후, 이미지 유닛(105)은 뷰 방향 이미지들 사이의 불일치 검출에 응답하여 깊이 맵을 생성하도록 구성될 수 있다. 따라서, 이미지 유닛(105)은 이미지들 내의 대응하는 이미지 오브젝트들을 발견하고, 그들 사이의 상대적인 이동/불일치를 결정하고, 대응하는 깊이 레벨을 이미지 오브젝트들에 할당할 수 있다. 불일치 추정에 기초하여 깊이를 결정하는 임의의 적합한 알고리즘이 사용될 수 있다는 것이 명백할 것이다.
- [0102] 이러한 불일치 추정은 비교적 정확한 깊이 맵들을 유도할 수 있다. 그러나, 깊이 맵들은 여전히 일반적으로 비교적 많은 수의 오류들을 포함할 것이고 일반적으로 완전히 일관되지 않을 것이다. 특히, 아티팩트들 및 불일치들은 특히 텍스트 이미지 오브젝트들에 대해 종종 발생할 수 있는 크고 날카로운 깊이 전이들 주변에서 발생할 수 있다.
- [0103] 그러므로, 상이한 방향들에 대한 이미지들에 대한 불일치 추정으로부터 생성된 깊이 맵을 직접 사용하는 것은 예를 들어, 뷰 이동을 수행할 때 지각된 품질 저하 및 아티팩트들의 도입을 초래하는 경향이 있을 것이다.
- [0104] 그러나, 확률 검출들 및 텍스트 이미지 오브젝트에 대한 불일치 결정된 깊이 값 및 바람직한 깊이 값의 병합에 기초하여 수정된 깊이 맵의 생성은 실질적으로 개선된 이미지 및 구체적으로 증가된 일관성 및 감소된 아티팩트들을 갖는 텍스트 이미지 오브젝트들 주위의 실질적으로 개선된 지각된 품질을 초래하는 경향이 있다. 특히, 불일치 추정을 수행할 때 텍스트 이미지 오브젝트들 주위에 일반적으로 발생할 수 있는 에러들, 불일치들 또는 아티팩트들의 존재는 설명된 접근법에 의해 실질적으로 완화될 수 있다.
- [0105] 이하에서, 확률 맵을 결정하기 위한 특정 접근법이 설명될 것이다.
- [0106] 접근법에서, 확률 값들의 결정은 복수의 픽셀들을 포함할 수 있는 픽셀 그룹들에 대해 수행될 수 있다. 따라서, 이하에서 기술된 원리들이 개별적인 픽셀들에 적용될 수 있지만, 이들은 특정 예에서 픽셀 그룹 단위로, 구체적으로 직사각형 블록 단위로 수행된다. 특정 예에서, 각 픽셀 그룹은 8x8 픽셀들의 블록이다.
- [0107] 따라서, 예에서, 확률 유닛(107)은 각 픽셀 그룹이 복수의 픽셀들을 포함하는 입력 이미지에 대한 픽셀 그룹들을 결정하기 위한 기능을 포함한다. 따라서 확률 값들의 결정은 이러한 픽셀 그룹들에 기초한다.
- [0108] 상기 접근법에서, 확률 유닛(107)은 각 값이 대응하는 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로 또는 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정되는지를 표시하는 픽셀 그룹들에 대한 값들을 포함하는 분류 맵을 먼저 결정하도록 구성된다.
- [0109] 특정 기술된 실시예에서, 분류 값들은 이진 값들이고, 따라서 각 값은 대응하는 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로서 지정되는 것 또는 픽셀 그룹이 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정되는 것을 나타낸다. 이러한 접근법은 많은 실시예들에서 분류 프로세스를 용이하게 할 수 있고, 강력한 결정들을 생성할 수 있다. 더욱이, 이것은 처리를 용이하게 할 수 있고, 상이한 깊이 레벨들의 다음 블렌딩 또는 조합에 매우 적합한 확률 맵들을 야기하는 것으로 발견되었다. 실제로, 고품질인 것으로 인식되는 3차원 이미지들의 생성을 초래하는 것이 발견되었다.
- [0110] 그러나, 다른 실시예들에서 분류 맵의 값들이 비-이진 값들일 수 있다는 것이 이해될 것이다. 예를 들어, 오브젝트 분류 프로세스는 픽셀 그룹들이 텍스트 이미지 오브젝트에 속하는지 아닌지의 여부를 나타내는 소프트 판정 값들을 생성할 수 있다. 예를 들어, 값들은 [0;1]의 간격에서 이산 값들의 세트일 수 있고, 1은 텍스트 이미지 오브젝트에 속하는 것으로서의 지정을 나타내고, 0은 텍스트 오브젝트에 속하지 않는 지정을 나타내고, 값들은 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로서 지정되기 위한 요구사항들과 얼마나 가깝게 일치하는 것으로 고려되는 것을 반영하는 것 사이의 값들이다. 실제로, 일부 실시예들에서, 값들은 초기 확률 값들로서 고려될 수 있고, 분류 맵은 초기 확률 맵으로서 고려될 수 있다.
- [0111] 확률 유닛(107)은 각 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는지 아닌지 여부를 나타내는 이진 결정을 제공하는 입력 이미지에 텍스트 이미지 오브젝트 분류 프로세스를 적용할 수 있다.
- [0112] 상이한 알고리즘들이 픽셀 그룹들을 텍스트 이미지 오브젝트에 속하는 것으로 또는 속하지 않는 것으로 지정하기 위해 사용될 수 있다는 것이 이해될 것이다. 간단한 예로서, 픽셀 그룹 내의 컬러 분포가 평가될 수 있고, 주로 자막 박스들(예를 들어, 흰색 및 검은색)에 사용된 것에 대응하는 픽셀 컬러들을 포함하는 경우, 픽셀 그룹은 텍스트 이미지 오브젝트로 지정될 수 있고, 그렇지 않으면 텍스트 이미지 오브젝트가 아닌 것으로 지정될 수 있다.



- [0113] 대부분의 실시예들에서, 실질적으로 더 복잡한 분류 또는 지정 알고리즘이 사용될 수 있음이 이해될 것이다. 예를 들어, 분류는 각 블록에 대해 계산되는 다수의 특징들에 기초할 수 있다. 특징들은 예를 들어 평균 수평 픽셀 경사 및 픽셀 강도 히스토그램의 특정 bin(certain bin) 내의 픽셀들의 수일 수 있다. 이후, AdaBoost(<http://cseweb.ucsd.edu/~yfreund/papers/IntroToBoosting.pdf>; 날짜: 20-05-2015)와 같은 기계 학습 접근법이 예를 들어 50 '약한' 분류 규칙들을 선형적으로 결합함으로써 소위 '강력한 분류자(strong classifier)'를 (자동으로) 트레인하기 위해 사용될 수 있다. 일반적으로 후보 특징들의 훨씬 더 큰 세트(예: > 300)가 선택하는 데 사용되므로 좋은 선택이 이루어질 수 있다. 이러한 큰 세트의 후보 피쳐들을 얻기 위해, 평균 수평 픽셀 경사와 같은 기본 피쳐들이 다양한 커널 크기를 갖는 다중 공간 컨볼루션(필터) 연산들에 대한 입력으로서 각각 사용되고, 따라서 예를 들어 10 개의 새로운 피쳐 후보들을 생성할 수 있다. 분류자의 트레이닝은 트레이닝 이미지들의 주어진 세트를 사용하는 8x8 블록 레벨이다. 그 후, 결과 알고리즘은 각 블록을 텍스트 이미지 오브젝트에 속하는 것으로 또는 속하지 않는 것으로 분류하기 위해 사용될 수 있다.
- [0114] 언급된 바와 같이, 생성된 분류 값들은 픽셀 그룹이 텍스트 이미지 오브젝트에 속하는 것으로 지정되었거나 텍스트 이미지 오브젝트에 속하지 않는 것으로 지정된 것을 나타내는 이진 값들로 기술된 예이다. 그러나, 일부 실시예들에서, 분류 프로세스는 소프트-결정 값들을 생성할 수 있고, 이들은 이진값들 대신 사용될 수 있다. 예를 들어, AdaBoost 분류자는 내부적으로 픽셀 그룹을 지정하기 위해 임계값과 비교되는 소프트 판정 표시들을 사용하고 생성할 수 있다. 그러나, 일부 실시예들에서, 이들 소프트-판정 값들은 이진 값들을 생성하기 위한 임계값에 대한 임의의 비교 없이 분류 맵에 대신 사용될 수 있다.
- [0115] 확률 맵을 생성하는 제 2 단계는 생성된 분류 맵에 필터링을 적용하는 단계로 진행되는 것이다. 필터링은 시간적 및 공간적 제한들을 가하고, 분류 맵에 평활화를 적용할 수 있다.
- [0116] 특히, 분류 프로세스가 이진 경 판정 분류 결과들(binary hard decision classification results)을 반환할 때, 이들 초기 이진 값들은 개개의 픽셀 그룹들(또는 픽셀들)이 텍스트 이미지 오브젝트에 속하는 또는 속하지 않는 확률을 나타내는 비-이진 값들로 변환될 수 있다. 따라서, 초기 이진 분류/지정은 분류 맵에 적용되는 필터링에 의한 점진적인 비-이진 확률 값들로 변환될 수 있다. 이러한 변환은 비교적 낮은 복잡도 및 자원 요구사항들을 갖는 효율적인 처리를 허용할뿐만 아니라, 그룹 자체의 분류뿐만 아니라 이것이 픽셀 그룹의 시간적 및/또는 공간적 인접면에서 다른 분류들과 어떻게 관련되는지를 반영하는 확률 값들로 매우 양호한 결과들을 제공하는 것으로 밝혀졌다. 실제로, 상기 접근법은 일관성이 높고 텍스트 이미지 오브젝트들의 깊이들을 조작하는 데 매우 적합한 확률 맵의 생성을 초래하는 경향이 있다.
- [0117] 필터링은 복수의 필터링 동작들을 포함할 수 있다. 많은 실시예들에서, 필터링은 복수의 순차 필터링 동작들을 포함할 수 있고, 전형적으로 후속 필터링 동작은 이전 필터링 동작의 결과에 대해 수행된다.
- [0118] 이하에서, 수행되는 복수의 순차 필터링 동작들을 갖는 특정 접근법의 예가 설명될 것이다. 그러나, 이것은 단지 예일 뿐이고, 상이한 실시예들에서 설명된 필터 동작들의 서브세트만이 사용될 수 있고, 필터 동작들의 순서는 다른 실시예들에서 상이할 수 있다는 것이 이해될 것이다. 예를 들어, 일부 실시예들에서, 시간 필터링은 공간 필터링 이전에 적용될 수 있는 반면, 다른 실시예들에서, 공간 필터링이 시간 필터링 이전에 적용될 수 있다.
- [0119] 상기 예에서, 필터 동작들은 각 픽셀 그룹이 복수의 픽셀들을 포함하는 픽셀 그룹 해상도에서 수행된다. 특히, 모든 필터들은 8x8 블록 해상도로 작동한다. 상기 설명에서, 인덱스들 $[m,n]$ 은 따라서 블록 인덱스들을 참조할 것이다. 결과적인 확률 맵은 또한 블록 해상도로 될 것이다. 수정된 깊이 맵의 생성 동안, 확률 값들은 높은 해상도를 제공하고, 특히 픽셀 레벨 해상도를 제공하기 위해 보간될 수 있다. 특정 예로서,  $a[m,n]$ 은 픽셀 레벨에 있는  $a[i,j]$ 를 계산하기 위해 (예를 들어, 선형적으로) 보간될 수 있다.
- [0120] 일부 실시예들에서, 필터링은 분류 맵에 적용된 이진 팽창 필터링을 포함할 수 있다.
- [0121] 예를 들어,  $a_{init}[m,n]$ 은 블록 인덱스들 $[m,n]$ , 즉 분류 맵의 이진 값들에서의 미가공 검출 결과를 나타낼 수 있다. 미가공 검출 결과는 1(텍스트 이미지 오브젝트에 속함) 또는 0(텍스트 이미지 오브젝트에 속하지 않음)의 값을 취한다. 결과의 밀도를 증가시키기 위해(즉, 검출된 영역들의 크기를 증가시키기 위해),  $a_{init}[m,n]$ 은 구체적으로 '십자형(cross-shape)'을 갖는 형태학적 팽창 필터와 같은 (일반적으로 이진) 팽창 필터를 사용하여 먼저 필터링될 수 있다:

$$\alpha_{dil}[m,n] = \max(\alpha_{init}[m,n], \alpha_{init}[m+1,n], \alpha_{init}[m,n+1], \alpha_{init}[m-1,n], \alpha_{init}[m,n-1])$$

[0122]

[0123]

이런 방식으로, 텍스트 이미지 오브젝트들에 속하는 것으로 검출된 영역들이 증가되어, 그에 의해 텍스트 이미지 오브젝트 픽셀 그룹들이 포함될 가능성을 증가시키고 이들 사이의 증가된 중첩 등을 제공한다.

[0124]

확률 유닛(107)은 또한 저역 통과 시간 필터링을 적용할 수 있다. 팽창 필터링이 수행되면, 시간 필터링이 상기 필터링의 결과에 적용될 수 있다. 그렇지 않으면, 이것은 예를 들어, 분류 맵에 직접 적용될 수 있다.

[0125]

많은 실시예들에서, 시간 필터링은 비대칭일 수 있다(즉, 저역 통과 효과는 다른 것보다 한 방향으로 더 중요할 수 있다). 구체적으로, 픽셀 그룹들이 텍스트 이미지 오브젝트들에 속할 가능성을 증가시키기 위해 변경하는 시정수는 픽셀 그룹들이 텍스트 이미지 오브젝트들에 속할 가능성을 감소시키기 위해 변경하는 시정수보다 낮다.

[0126]

특정 예로서, 시간적, 비대칭 재귀 필터는 예를 들어 팽창 필터의 출력(또는 예를 들어, 분류 맵에 직접적으로)에 적용될 수 있다. 많은 실시예들에 대해 적합한 필터의 특정 예는 다음일 수 있다:

$$\alpha_{tmp}^{(t)} = \begin{cases} \frac{1}{4}\alpha_{tmp}^{(t-1)} + \frac{3}{4}\alpha_{dil}^{(t)} & \text{if } \alpha_{dil}^{(t)} > \alpha_{tmp}^{(t-1)} \\ \frac{7}{8}\alpha_{tmp}^{(t-1)} + \frac{1}{8}\alpha_{dil}^{(t)} & \text{otherwise} \end{cases}$$

[0127]

[0128]

여기서, 위첨자 t는 프레임 번호를 나타낸다.

[0129]

이러한 필터는 텍스트가 먼저 검출될 때, 깊이가 빠르게 수정될 것이라는 것을 보장할 수 있다. 그러나, 텍스트가 깊이 맵에 미치는 영향은 오직 텍스트가 제거될 때 점진적으로 감소할 것이다. 순수 효과(net effect)는 더 나은 시간적 안정성이다.  $\alpha_{dil}^{(t)}$ 은 전형적으로 t=0에서 모든 블록들에 대해 0으로 설정된다.

[0130]

일부 실시예들에서, 필터링은 공간 필터링을 포함한다. 필터링은 분류 맵 상에서 직접 수행될 수 있거나(예를 들어, 그와 같이 포함된다면 팽창 필터링 이후에) 예를 들어 시간 필터링 이후 수행될 수 있다.

[0131]

일부 실시예들에서, 공간 필터링은 선형 저역 통과 필터일 수 있다. 그러나, 많은 실시예들에서, 공간 필터링은 구체적으로 적어도 하나의 소프트맥스 필터를 포함할 수 있다.

[0132]

소프트맥스 필터는 필터의 케이스케이드와 출력을 최댓값으로 제한하는 리미터일 수 있다. 예를 들어, 소프트맥스 저역-통과 필터는 저역-통과 필터의 케이스케이드 및 저역-통과 필터의 출력값을 최댓값으로 제한하는 리미터에 대응하는 동작을 수행할 수 있다. 그러므로, 소프트맥스 필터는 최대 출력 값 한계를 갖는 필터에 대응할 수 있다. 최댓값은 1의 값일 수 있지만, 이것이 상이한 실시예들 간에 다를 수 있는 것이 이해될 것이다.

[0133]

구체적으로, 소프트맥스 필터는 필터 출력이 더 높은 값들로 바이어싱 되도록, 저역-통과 필터와 같은 표준 필터 동작을 수행하지만, 필터 출력에 1보다 큰 값을 곱하는 표준 필터 동작을 수행하는 필터일 수 있고, 그 후 얻어진 출력과 미리 정의된 최대값의 최소치가 취해진다. 이런 방식으로 정의된 소프트맥스 필터는 따라서 세 가지 구성요소들로 구성된다: (표준) 필터 동작, 곱셈 계수, 및 미리 설정된 최댓값. 따라서, 소프트맥스 필터링은 필터 출력을 최댓값으로 제한하면서 공간 (저역 통과) 필터링을 적용하는 것을 포함할 수 있다. 전형적으로, 저역 통과 필터링은 일반적으로 미리 결정된 스케일 인자에 의해 필터 출력 값들을 스케일링 한 후 (전형적으로 정규화된) 저역 통과 필터를 사용하는 제 1 필터링의 2단계 프로세스일 수 있다. 이후 결과적인 스케일된 출력은 최댓값으로 제한된다. 그러나, 스케일링은 전반적인 필터링의 일부로서 고려될 수 있고, 예를 들어 예로서 공간 필터의 스케일링 (커널) 계수들에 의해 구현될 수 있다. 최댓값은 전형적으로 미리 결정된다.

[0134]

소프트맥스 필터의 사용은 개선된 성능을 제공할 수 있다. 특히, 영역의 경계들을 향하여 부드럽고 점진적인 전이를 제공하는 동안 내부 픽셀들이 텍스트 이미지 오브젝트들로 표시되는(거의 확실하게) 영역들을 생성하려는 경향이 있을 수 있다. 이러한 특성은 특히 예를 들어, 확률 값들을 사용하는 자막 박스들을 식별하는데 적합하다.

[0135]

일부 실시예들에서, 공간 필터링은 적어도 두 개의 순차적 공간 소프트-맥스 필터들을 포함할 수 있다. 상이한 소프트맥스 필터들은 상이한 특성들을 가질 수 있다. 구체적으로, 공간 필터링은 두 개의 기능들을 제공할 수

있다. 하나는 텍스트 이미지 오브젝트들에 대응하는 영역들을 채우기 위한 것, 즉 검출된 영역들 내에 어떠한 갭들도 바람직하게 존재하지 않도록, 검출된 영역들을 중첩시키기 위한 것이다. 다른 하나는 검출된 영역들 주위의 전이들이 부드럽고 점진적이게 되도록 공간 평활화를 제공하는 것이다. 두 개의 기능들은 개선된 시각적 경험을 제공한다.

[0136] 따라서, 특정 예에서, 제 1 소프트맥스 필터는 텍스트 이미지 오브젝트들에 대응하는 검출된 영역을 "채우기" 위해 적용될 수 있다. 이후, 제 2 소프트맥스 필터는 부드러운 전이들을 생성하기 위해 제 1 소프트맥스 필터의 결과에 적용될 수 있다.

[0137] 특정 예에서, 시간 필터링 후에, 일부 텍스트 오브젝트들은 텍스트 이미지 오브젝트들에 속하는 영역들의 표시들의 매우 드물게 채워진 맵을 초래할 수 있다. 그러므로, 예를 들어 자막 박스 안의 홀들을 채우기 위해, '소프트-맥스' 필터가 적용될 수 있다. 예를 들어, 21 블록들의 커널 높이 및 35 블록들의 커널 너비를 갖는 소프트맥스 필터가 사용될 수 있다:

$$\alpha_{\text{softmax}}[m,n] = \min \left( \frac{\sum_{p=m-10}^{m+10} \sum_{q=n-17}^{n+17} \alpha_{\text{tmp}}[p,q]}{K_1}, 1 \right),$$

[0138]

[0139]

$$K_1 < \sum_{p=m-10}^{m+10} \sum_{q=n-17}^{n+17} \alpha_{\text{tmp}}[p,q].$$

[0140]

[0141] 특정 예로서, 1의 값이 255로 표현되도록  $a$ 는 8-비트 수들로 표현될 수 있다. 이러한 예에서,  $K_1=32$ 의 값이 종종 사용될 수 있다. 이러한 소프트맥스 필터는 많은 실시예들에서 문자들과 단어들 사이에 홀들을 채울 수 있다.

[0142] 이러한 제 1 소프트맥스 필터는 텍스트 이미지 오브젝트에 대응하는 깊이 맵의 영역과 깊이 맵의 주변 영역들 사이의 하드 공간 경계를 제거하고, 따라서 이것이 표현될 때 텍스트 및 3차원 이미지의 나머지 사이의 부드러운 전이를 제공하려고 시도하는 제 2 소프트맥스 필터에 의해 뒤따라질 수 있다.

[0143] 제 2 소프트맥스 필터는 제 1 소프트맥스 필터에 대응할 수 있지만, 상이한 파라미터들을 갖고, 구체적으로, 필터가 더 많은 저역-통과 필터와 같이 그리고 더 적은 맥스-필터와 같이 동작하도록, 더 큰 분모를 갖는다. 특정 예에서, 높이 11 및 너비 25 픽셀들의 커널 사이즈가 사용될 수 있다:

$$\alpha[m,n] = \min \left( \frac{\sum_{p=m-5}^{m+5} \sum_{q=n-12}^{n+12} \alpha_{\text{softmax}}[p,q]}{K_2}, 1 \right),$$

[0144]

[0145] 다시,  $a$ 는 8비트 수들로 표현될 수 있고, 분모에 적합한 값은  $K_2=256$ 일 수 있다.

[0146] 두 개의 순차적 소프트 필터들의 사용은 개선된 성능을 초래할 뿐만 아니라, 용이한 구현을 허용할 수 있다. 구체적으로, 동일한 알고리즘 또는 함수가 변경되는 파라미터 값들로 단지 사용될 수 있다. 실제로, 많은 실시예들에서, 예를 들어 적분 이미지 접근법을 사용하여, 심지어 이미지 데이터가 더 효율적으로 (재)사용될 수 있다.

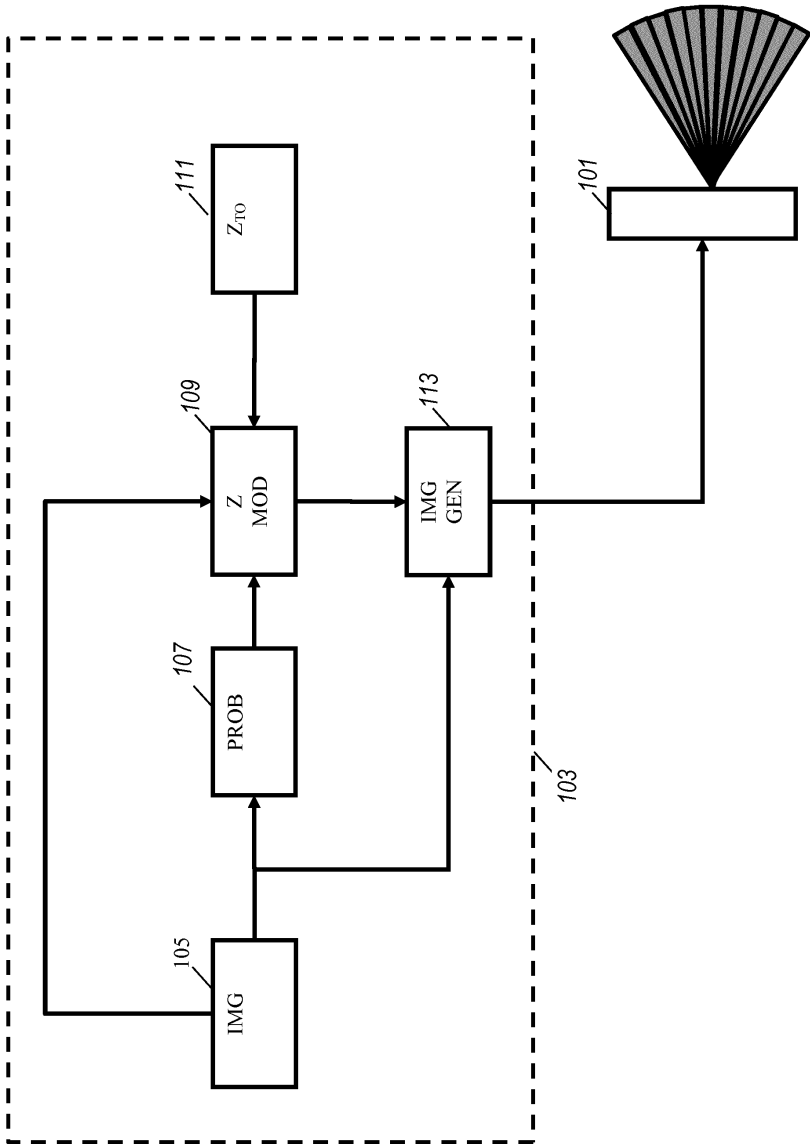
[0147] 도 5는 확률 맵을 결정하기 위해 기술된 예시적인 접근법의 결과들의 예들을 도시한다. 도 5는 이진 팽창 필터링 및 시간 필터링 이후의 입력 이미지(501) 및 분류 맵(503)을 도시한다. 알 수 있는 바와 같이, 이 단계의 확률 맵은 매우 희박하게 채워질 수 있다(흰색은 텍스트 이미지 오브젝트에 속할 높은 확률을 갖는 영역들을 나타내고, 검은색은 텍스트 이미지 오브젝트에 속할 낮은 확률을 갖는 낮은 영역들을 나타냄). 제 1 소프트맥스 필터의 출력(505)이 도시된다. 알 수 있는 바와 같이, 이러한 필터링은 텍스트 이미지 오브젝트에 속하는 것으로 간주되는 상당히 증가된 영역을 초래한다. 그러나, 전이는 아주 갑작스럽다. 제 2 소프트맥스 필터의 출력(507)이 또한 도시된다. 알 수 있는 바와 같이, 이것은 제 1 소프트맥스 필터의 출력에 상당히 가깝게 대응하지만,

오토스테레오스코픽 디스플레이(101)에 의해 렌더링된 이미지에서 훨씬 더 많은 점진적인 깊이 전이를 초래하는 실질적으로 평활한 전이들을 갖는다.

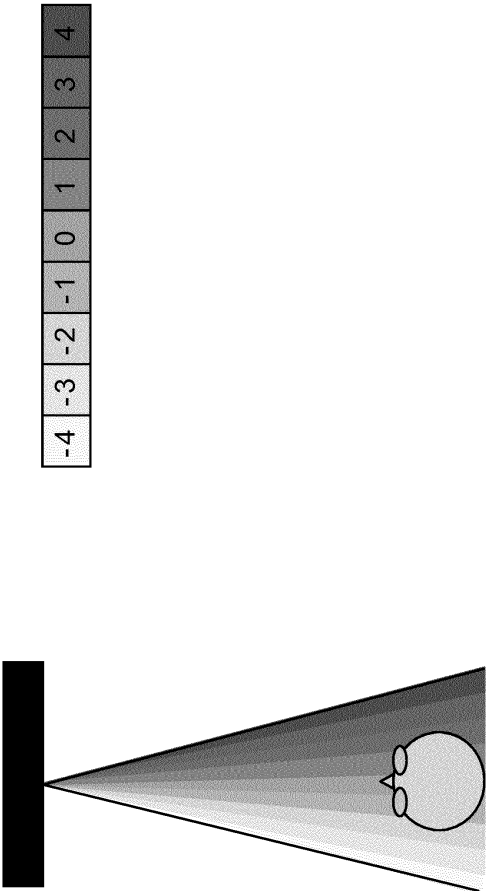
- [0148] 설명된 예시에서, 분류 맵 및 필터링/처리하는 픽셀 그룹들, 및 구체적으로 복수의 픽셀들을 포함하는 픽셀 블록들을 사용하여 수행된다. 실제로, 이 예에서, 8x8 픽셀들의 블록들이 사용된다. 이것은 64의 인자에 의한 해상도를 감소시키고, 따라서 실질적으로 더 효율적이고 덜 까다로운 작업을 허용할 수 있다. 더욱이, 발명자는 바람직한 성능 및 인지된 이미지 품질을 달성하면서 이러한 해상도 감소가 적용될 수 있는 것을 깨달았다.
- [0149] 그러나, 인지된 이미지 품질을 개선하기 위해, 가중된 조합에 대한 가중치들은 블록 레벨 해상도보다 높은 해상도로 생성될 수 있고, 특히 픽셀 해상도로 생성될 수 있다.
- [0150] 이는 블록 기반 깊이 값들 사이의 보간을 사용함으로써 달성될 수 있다. 예를 들어, 인접 블록들에 대한 확률 값들 사이의 보간은 블록들의 개별적인 픽셀들에 대한 깊이 값들을 생성하기 위해 사용될 수 있다. 보간에 대한 가중치들은 픽셀 그룹 내의 픽셀의 위치에 기초하여 결정될 수 있다.
- [0151] 예를 들어, 제 1 및 제 2 깊이 값 사이의 보간이 수행될 수 있다. 제 1 및 제 2 깊이 값에 대한 픽셀 그룹들 사이의 경계상의 실질적으로 픽셀의 경우, 깊이 값은 제 1 및 제 2 깊이 값들을 실질적으로 동등하게 가중함으로써 결정될 수 있다. 예를 들어, 간단한 평균이 수행될 수 있다. 그러나, 하나의 픽셀 그룹의 중심에 있는 픽셀의 경우, 깊이 값은 단순히 상기 픽셀 그룹에 대해 결정된 깊이 값으로 결정될 수 있다.
- [0152] 명료화를 위한 상기 설명이 상이한 기능 회로들, 유닛들, 및 프로세서들을 참조하여 본 발명의 실시예들을 설명하였음이 이해될 것이다. 그러나, 본 발명을 손상시키지 않으면서 다른 기능 회로들, 유닛들 또는 프로세서들 간의 기능의 임의의 적절한 분배가 사용될 수 있음이 명백할 것이다. 예를 들어, 별도의 프로세서들 또는 제어기들에 의해 수행되도록 도시된 기능은 동일한 프로세서 또는 제어기들에 의해 수행될 수 있다. 그러므로, 특정 기능 유닛들 또는 회로들에 대한 참조들은 엄격한 논리적 또는 물리적 구조 또는 조직을 나타내기보다는 기술된 기능을 제공하기 위한 적절한 수단에 대한 참조들로서만 이해되어야 한다.
- [0153] 본 발명은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 임의의 조합을 포함하는 임의의 적합한 형태로 구현될 수 있다. 본 발명은 적어도 부분적으로 하나 이상의 데이터 프로세서들 및/또는 디지털 신호 프로세서들 상에서 실행되는 컴퓨터 소프트웨어로서 선택적으로 구현될 수 있다. 본 발명의 일 실시예의 요소들 및 구성 요소들은 물리적으로, 기능적으로 및 논리적으로 임의의 적절한 방식으로 구현될 수 있다. 실제로, 기능은 단일 유닛, 복수의 유닛들 또는 다른 기능 유닛들의 일부로서 구현될 수 있다. 이와 같이, 본 발명은 단일 유닛으로 구현될 수 있거나 또는 상이한 유닛들, 회로들 및 프로세서들 간에 물리적 및 기능적으로 분배될 수 있다.
- [0154] 본 발명이 몇몇 실시예들과 관련하여 설명되었지만, 본원에 설명된 특정 형태에 한정되도록 의도되지는 않는다. 오히려, 본 발명의 범위는 첨부된 청구항들에 의해서만 한정된다. 또한, 특징이 특정 실시예들과 관련하여 기술되는 것처럼 보일 수 있지만, 당업자는 기술된 실시예들의 다양한 특징들이 본 발명에 따라 결합될 수 있음을 인식할 것이다. 청구항들에서, 용어(포함하는)는 다른 요소들 또는 단계들의 존재를 배제하지 않는다.
- [0155] 또한, 개별적으로 열거되지만, 복수의 수단, 요소들, 회로들 또는 방법 단계들은 예를 들어, 단일 회로, 유닛 또는 프로세서에 의해 구현될 수 있다. 추가로, 개개의 피쳐들이 상이한 청구항들에 포함될 수 있지만, 이들은 아마도 유리하게 결합될 수 있고, 상이한 청구항들에 포함된다는 것은 피쳐들의 조합이 실현 가능 및/또는 유리하지 않다는 것을 의미하지 않는다. 또한, 청구항들의 하나의 카테고리에 기능을 포함하는 것은 이 카테고리에 대한 제한을 의미하지 않고, 해당 기능이 다른 청구항 카테고리들에도 동등하게 적용가능함을 나타낸다. 또한, 청구항들의 특징들의 순서는 특징들이 작동되어야 하는 임의의 특정 순서를 의미하지 않고, 특히 방법 청구항의 개별적인 단계들의 순서는 단계들이 이러한 순서로 수행되어야 함을 의미하지 않는다. 오히려, 단계들은 임의의 적절한 순서로 수행될 수 있다. 게다가, 단수의 참조들은 복수를 배제하지 않는다. 따라서, 참조들("a", "an", "제 1", "제 2" 등)은 복수를 배제하지 않는다. 청구항들에서 기호 부호들은 단지 명확한 예가 임의의 방식으로 청구항들의 범위를 제한하는 것으로 해석되지 않아야 한다.

도면

도면1

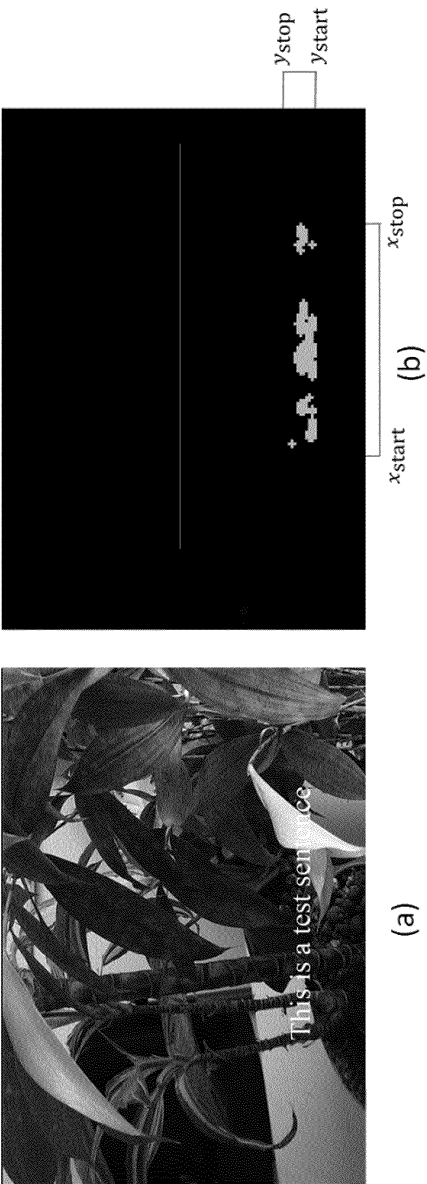


도면2

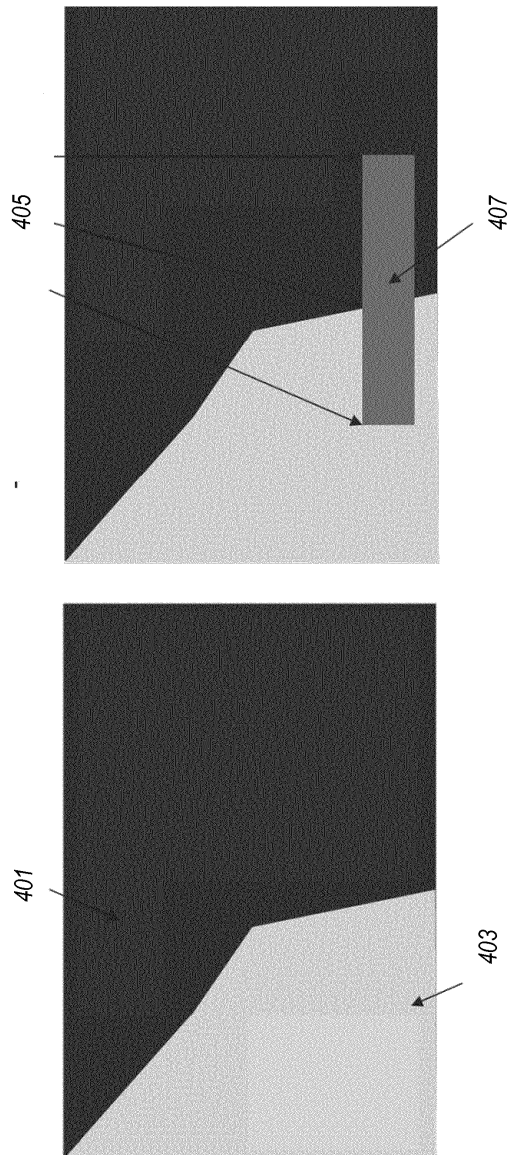




도면3



도면4



도면5

