



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0111341  
(43) 공개일자 2014년09월18일

(51) 국제특허분류(Int. Cl.)  
G06K 9/00 (2006.01) G06K 9/22 (2006.01)  
(21) 출원번호 10-2014-7022199  
(22) 출원일자(국제) 2013년01월08일  
심사청구일자 없음  
(85) 번역문제출일자 2014년08월07일  
(86) 국제출원번호 PCT/US2013/020699  
(87) 국제공개번호 WO 2013/106346  
국제공개일자 2013년07월18일  
(30) 우선권주장  
13/736,527 2013년01월08일 미국(US)  
61/584,765 2012년01월09일 미국(US)

(71) 출원인  
켈컴 인코포레이티드  
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775  
(72) 발명자  
백 영기  
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
구 형일  
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
김 덕훈  
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
(74) 대리인  
특허법인코리어나

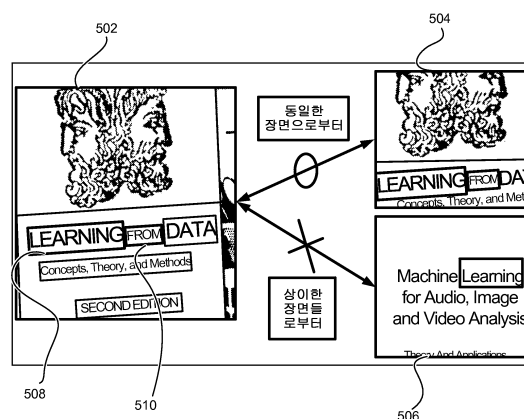
전체 청구항 수 : 총 38 항

(54) 발명의 명칭 OCR 캐시 업데이트

(57) 요약

본원에서 설명되는 기법들은, 컴퓨팅 디바이스를 사용하여 OCR을 수행하는 동안 OCR 캐시를 자동으로 및 지능적으로 생성하고 업데이트하는 방법을 제공한다. 컴퓨팅 디바이스에 연결된 카메라를 사용하여 캡처된 이미지는 OCR 캐시에 저장된 이전의 이미지에 대해 매칭될 수도 있다. 매치가 발견되면, OCR 캐시는 새로운 이미지를 이용하여 새로운 또는 양호한 정보로 업데이트될 수도 있다. 매칭된 이전의 이미지는 OCR 캐시에 유지될 수도 있거나, 또는 새로운 캡처된 이미지는 OCR 캐시에서의 매칭된 이전의 이미지를 대체할 수도 있다. 하나의 실시형태에서, OCR 캐시에 이미지를 저장하기 전에 글레이어를 제거하거나 또는 감소시키는 기법들이 설명된다. 일부 실시형태들에서, 글레이어는 OCR 수행의 부재 시에 제거되거나 또는 감소된다.

대표도 - 도5



## 특허청구의 범위

### 청구항 1

복수의 이미지들을 수신하는 단계;

상기 복수의 이미지들의 각각에서 데이터를 인식하는 단계로서, 상기 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 단계를 포함하는, 상기 복수의 이미지들의 각각에서 데이터를 인식하는 단계;

인식된 상기 데이터에 적어도 부분적으로 기초하여, 상기 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 단계; 및

상기 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 상기 인식된 데이터를 사용하여 상기 제 1 장면을 표현하는 최종 이미지를 저장하는 단계를 포함하는, 방법.

### 청구항 2

제 1 항에 있어서,

상기 수신하는 단계는, 카메라를 사용하여 상기 복수의 이미지들을 연속해서 캡처하는 단계를 포함하는, 방법.

### 청구항 3

제 1 항에 있어서,

상기 복수의 이미지들 중 상기 적어도 하나의 이미지 및 상기 적어도 하나의 이미지에 대한 상기 인식된 데이터의 저장은, 후속의 이미지가 캡처되고 상기 후속의 이미지에서 데이터가 인식되는 동안 유지되는, 방법.

### 청구항 4

제 1 항에 있어서,

상기 최종 이미지를 저장하는 단계 전에, 상기 방법은, 상기 적어도 2 개의 이미지들 및 상기 적어도 2 개의 이미지들과 연관된 상기 인식된 데이터를 사용하여 상기 최종 이미지를 합성하는 단계를 더 포함하며,

상기 저장하는 단계는 합성된 상기 이미지를 저장하는 단계를 포함하는, 방법.

### 청구항 5

제 4 항에 있어서,

상기 합성하는 단계는, 상기 최종 이미지를 생성하기 위해 상기 적어도 2 개의 이미지들 중 제 2 이미지로부터의 엘리먼트들을 상기 적어도 하나의 이미지에 통합하는 단계를 포함하는, 방법.

### 청구항 6

제 4 항에 있어서,

상기 최종 이미지는 상기 적어도 하나의 이미지를 포함하며,

상기 합성하는 단계는, 상기 적어도 2 개의 이미지들 중 제 2 이미지의 상기 인식된 데이터를 상기 적어도 하나의 이미지의 상기 인식된 데이터에 통합하는 단계를 포함하는, 방법.

### 청구항 7

제 1 항에 있어서,

상기 최종 이미지를 저장하는 단계 전에, 상기 방법은, 상기 최종 이미지로서 상기 제 1 장면을 표현하기 위해 상기 적어도 2 개의 이미지들로부터 상기 적어도 하나의 이미지를 선택하는 단계를 더 포함하며,

상기 저장하는 단계는 선택된 상기 이미지를 저장하는 단계를 포함하는, 방법.

## 청구항 8

제 7 항에 있어서,

상기 선택하는 단계는, 상기 적어도 2 개의 이미지들의 각각의 이미지의 에너지를 비교하는 단계를 포함하는, 방법.

## 청구항 9

제 8 항에 있어서,

각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초하여, 상기 적어도 2 개의 이미지들의 각각의 이미지의 상기 에너지를 계산하는 단계를 더 포함하는, 방법.

## 청구항 10

제 1 항에 있어서,

상기 결정하는 단계는, 상기 적어도 2 개의 이미지들의 각각의 이미지에서의 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색 및 전경색 중 하나 이상을 비교하는 단계를 포함하는, 방법.

## 청구항 11

제 1 항에 있어서,

상기 복수의 이미지들 중 하나 이상의 이미지들이 상기 제 1 장면 이외의 하나 이상의 장면들로부터의 정보를 포함한다고 결정하는 단계, 및 상기 최종 이미지 및 상기 하나 이상의 장면들을 각각 표현하는 하나 이상의 이미지들을 디스플레이하는 단계 또는 상기 적어도 하나의 이미지에 대한 및 상기 하나 이상의 장면들에 대한 상기 인식된 데이터를 디스플레이하는 단계를 더 포함하는, 방법.

## 청구항 12

제 11 항에 있어서,

사용자가 디스플레이된 상기 이미지들 중 하나의 이미지 또는 상기 하나의 이미지에 대한 디스플레이된 상기 인식된 데이터를 선택했다고 결정하는 단계, 및 선택된 상기 이미지 또는 상기 인식된 데이터에 관련된 정보를 프로그래밍 또는 애플리케이션에 입력하는 단계를 더 포함하는, 방법.

## 청구항 13

제 1 항에 있어서,

상기 최종 이미지를 저장하는 단계 전에, 상기 방법은,

상기 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하는 단계;

상기 제 1 장면으로부터의 상기 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하는 단계로서, 상기 제 1 지역 및 상기 제 2 지역은 상기 제 1 장면의 동일한 지역을 표현하는, 상기 제 2 지역을 식별하는 단계; 및

상기 제 1 이미지, 및 상기 제 2 이미지로부터의 글레이어가 없는 상기 제 2 지역을 사용하여 감소된 글레이어를 갖는 상기 최종 이미지를 합성하는 단계를 더 포함하는, 방법.

## 청구항 14

복수의 이미지들을 수신하도록 구성된 이미지 모듈;

광학적 문자 인식을 사용하여 상기 복수의 이미지들의 각각에서 데이터를 인식하도록 구성된 OCR 엔진;

인식된 상기 데이터에 적어도 부분적으로 기초하여, 상기 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1

장면으로부터의 정보를 포함한다고 결정하도록 구성된 매칭 모듈; 및

상기 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 상기 인식된 데이터를 사용하여 상기 제 1 장면을 표현하는 최종 이미지를 저장하도록 구성된 통합 모듈을 포함하는, 디바이스.

#### 청구항 15

제 14 항에 있어서,

상기 이미지 모듈은, 상기 디바이스에 연결된 카메라를 사용하여 상기 복수의 이미지들을 연속해서 캡처하도록 구성되는, 디바이스.

#### 청구항 16

제 14 항에 있어서,

상기 최종 이미지를 저장하기 전에, 상기 통합 모듈은, 상기 적어도 2 개의 이미지들 및 상기 적어도 2 개의 이미지들과 연관된 상기 인식된 데이터를 사용하여 상기 최종 이미지를 합성하도록 추가로 구성되며,

상기 저장은 합성된 상기 이미지를 저장하는 것을 포함하는, 디바이스.

#### 청구항 17

제 16 항에 있어서,

상기 합성은, 상기 최종 이미지를 생성하기 위해 상기 적어도 2 개의 이미지들 중 제 2 이미지로부터의 엘리먼트들을 상기 적어도 하나의 이미지에 통합하는 것을 포함하는, 디바이스.

#### 청구항 18

제 16 항에 있어서,

상기 최종 이미지는 상기 적어도 하나의 이미지를 포함하며,

상기 합성은, 상기 적어도 2 개의 이미지들 중 제 2 이미지의 상기 인식된 데이터를 상기 적어도 하나의 이미지의 상기 인식된 데이터에 통합하는 것을 포함하는, 디바이스.

#### 청구항 19

제 14 항에 있어서,

선택 모듈을 더 포함하며,

상기 통합 모듈이 상기 최종 이미지를 저장하기 전에, 상기 선택 모듈은 상기 최종 이미지로서 상기 제 1 장면을 표현하기 위해 상기 적어도 2 개의 이미지들로부터 상기 적어도 하나의 이미지를 선택하도록 구성되고,

상기 저장은 선택된 상기 이미지를 후속하여 저장하는 것을 포함하는, 디바이스.

#### 청구항 20

제 19 항에 있어서,

상기 선택은, 상기 적어도 2 개의 이미지들의 각각의 이미지의 에너지를 비교하는 것을 포함하는, 디바이스.

#### 청구항 21

제 20 항에 있어서,

상기 선택 모듈은, 각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초하여, 상기 적어도 2 개의 이미지들의 각각의 이미지의 상기 에너지를 계산하도록 추가로 구성되는, 디바이스.

#### 청구항 22

제 14 항에 있어서,

상기 결정은, 상기 적어도 2 개의 이미지들의 각각의 이미지에서의 인식된 텍스트 주위에서 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색, 전경색, 및 경계 박스 중 하나 이상을 비교하는 것을 포함하는, 디바이스.

#### 청구항 23

제 14 항에 있어서,

상기 매칭 모듈은, 상기 복수의 이미지들 중 하나 이상의 이미지들이 상기 제 1 장면 이외의 하나 이상의 장면들로부터의 정보를 포함한다고 결정하고,

상기 디바이스는, 상기 최종 이미지 및 상기 하나 이상의 장면들을 각각 표현하는 하나 이상의 이미지들 또는 상기 적어도 하나의 이미지에 대한 상기 인식된 데이터를 디스플레이하도록 또는 상기 적어도 하나의 이미지에 대한 및 상기 하나 이상의 장면들에 대한 상기 인식된 데이터를 디스플레이하도록 구성된 디바이스 디스플레이를 더 포함하는, 디바이스.

#### 청구항 24

제 23 항에 있어서,

디스플레이된 상기 이미지들 중 하나의 이미지 또는 상기 하나의 이미지에 대한 디스플레이된 상기 인식된 데이터의 사용자 선택을 수신하도록 구성된 입력 디바이스를 더 포함하며,

선택된 상기 이미지 또는 상기 인식된 데이터에 관련된 정보는, 프로그램 또는 애플리케이션에 입력되는, 디바이스.

#### 청구항 25

제 24 항에 있어서,

입력 정보를 사용하여 인터넷 쿼리가 제출되며, 상기 입력 정보에 기초하여 해석이 결정되거나, 또는 상기 입력 정보에 대응하는 텍스트가 텍스트 박스에 입력되는, 디바이스.

#### 청구항 26

제 14 항에 있어서,

상기 통합 모듈이 상기 최종 이미지를 저장하기 전에, 상기 매칭 모듈은,

상기 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하며; 그리고

상기 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하도록 추가로 구성되며,

상기 제 1 지역 및 상기 제 2 지역은 상기 제 1 장면의 동일한 지역을 표현하며,

상기 통합 모듈은, 상기 제 1 이미지, 및 상기 제 2 이미지로부터의 글레이어가 없는 상기 제 2 지역을 사용하여 감소된 글레이어를 갖는 상기 최종 이미지를 합성하도록 구성되는, 디바이스.

#### 청구항 27

프로세서에 의해 실행가능한 명령들을 포함하는 비일시적 컴퓨터 판독가능 저장 매체로서,

상기 명령들은,

복수의 이미지들을 수신하고;

상기 복수의 이미지들의 각각에서 데이터를 인식하는 것으로서, 상기 데이터를 인식하는 것은 상기 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 것을 포함하는, 상기 데이터를 인식하고,

인식된 상기 데이터에 적어도 부분적으로 기초하여, 상기 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1

장면으로부터의 정보를 포함한다고 결정하고; 그리고

상기 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 상기 인식된 데이터를 사용하여 상기 제 1 장면을 표현하는 최종 이미지를 저장하는 명령들을 포함하는, 비밀시적 컴퓨터 판독 가능 저장 매체.

#### 청구항 28

복수의 이미지들을 수신하는 수단;

상기 복수의 이미지들의 각각에서 데이터를 인식하는 수단으로서, 상기 복수의 이미지들의 각각에서 데이터를 인식하는 것이 상기 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 것을 포함하는, 상기 인식하는 수단;

인식된 상기 데이터에 적어도 부분적으로 기초하여, 상기 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 수단; 및

상기 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 상기 인식된 데이터를 사용하여 상기 제 1 장면을 표현하는 최종 이미지를 저장하는 수단을 포함하는, 장치.

#### 청구항 29

제 28 항에 있어서,

상기 수신하는 수단은, 카메라를 사용하여 상기 복수의 이미지들을 연속해서 캡처하는 수단을 포함하는, 장치.

#### 청구항 30

제 28 항에 있어서,

상기 복수의 이미지들 중 상기 적어도 하나의 이미지 및 상기 적어도 하나의 이미지에 대한 상기 인식된 데이터의 저장은, 후속의 이미지가 캡처되고 상기 후속의 이미지에서 데이터가 인식되는 동안 유지되는, 장치.

#### 청구항 31

제 28 항에 있어서,

상기 장치는, 상기 최종 이미지를 저장하기 전에, 상기 적어도 2 개의 이미지들 및 상기 적어도 2 개의 이미지들과 연관된 상기 인식된 데이터를 사용하여 상기 최종 이미지를 합성하는 수단을 더 포함하며,

상기 저장은 합성된 상기 이미지를 저장하는 것을 포함하는, 장치.

#### 청구항 32

제 28 항에 있어서,

상기 장치는, 상기 최종 이미지를 저장하기 전에, 상기 최종 이미지로서 상기 제 1 장면을 표현하기 위해 상기 적어도 2 개의 이미지들로부터의 상기 적어도 하나의 이미지를 선택하는 수단을 더 포함하며,

상기 저장은 선택된 상기 이미지를 저장하는 것을 포함하는, 장치.

#### 청구항 33

제 32 항에 있어서,

상기 선택하는 수단, 상기 적어도 2 개의 이미지들의 각각의 이미지의 에너지를 비교하는 수단을 포함하는, 장치.

#### 청구항 34

제 33 항에 있어서,

각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초하여, 상기 적어도 2 개의 이미지들의 각각의 이미지의 상기 에너지를 계산하는 수단을 더 포함하는, 장치.

### 청구항 35

제 28 항에 있어서,

상기 결정하는 수단은, 상기 적어도 2 개의 이미지들의 각각의 이미지에서의 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색 및 전경색 중 하나 이상을 비교하는 수단을 포함하는, 장치.

### 청구항 36

제 28 항에 있어서,

상기 복수의 이미지들 중 하나 이상의 이미지들이 상기 제 1 장면 이외의 하나 이상의 장면들로부터의 정보를 포함한다고 결정하는 수단, 및 상기 최종 이미지 및 상기 하나 이상의 장면들을 각각 표현하는 하나 이상의 이미지들을 디스플레이하는 수단 또는 상기 적어도 하나의 이미지에 대한 및 상기 하나 이상의 장면들에 대한 상기 인식된 데이터를 디스플레이하는 수단을 포함하는, 장치.

### 청구항 37

제 36 항에 있어서,

사용자가 디스플레이된 상기 이미지들 중 하나의 이미지 또는 상기 하나의 이미지에 대한 디스플레이된 데이터를 선택했다고 결정하는 수단, 및 선택된 상기 이미지 또는 상기 인식된 데이터에 관련된 정보를 프로그램 또는 애플리케이션에 입력하는 수단을 더 포함하는, 장치.

### 청구항 38

제 28 항에 있어서,

상기 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하는 수단;

상기 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하는 수단으로서, 상기 제 1 지역 및 상기 제 2 지역은 상기 제 1 장면의 동일한 지역을 표현하는, 상기 제 2 지역을 식별하는 수단; 및

상기 제 1 이미지, 및 상기 제 2 이미지로부터의 글레이어가 없는 상기 제 2 지역을 사용하여 감소된 글레이어를 갖는 상기 최종 이미지를 합성하는 수단을 더 포함하는, 장치.

## 명세서

### 배경 기술

[0001]

광학적 문자 인식 (Optical character recognition; OCR) 은 손으로 쓴, 타자로 친 또는 인쇄된 텍스트, 그래픽스 또는 심볼들의 스캔된 이미지들의 머신 인코딩된 텍스트로의 머신 또는 전자 번역이다. OCR은 컴퓨터 비전 및 증강 현실 애플리케이션들에서 성장하는 연구 분야이다. 컴퓨터 비전은 디바이스가 그것의 부근의 환경을 인지하는 것을 허용한다. 컴퓨터 비전은 실제 세계를 가상 세계로 연장하여 최종 사용자의 상호작용으로부터의 감각 입력 (sensory input) 을 허용하는 것에 의해 증강 현실 (augmented reality) 에서의 애플리케이션들을 가능하게 한다. 실생활의 예의 애플리케이션에서, 컴퓨터 비전 가능 (enabled) OCR은 쓰여진 텍스트 및 심볼들의 인지 시에 스마트 폰과 같은 개인 모바일 디바이스의 카메라의 시야에서 개체들을 시각적으로 디스플레이시키는 것을 도울 수도 있다.

### 발명의 내용

#### 과제의 해결 수단

[0002]

본원에서 설명되는 기법들은, 증강 현실의 응용들에서 모바일 디바이스를 사용하여 OCR을 수행하는 동안 OCR 캐시를 자동으로 및 능동적으로 생성하고 업데이트하는 방법을 제공한다. 모바일 디바이스에 연결된 카메라를 사용하여 캡처된 이미지가 OCR 캐시에 저장된 이전의 이미지에 대해 매칭된다. 매치가 발견되면, OCR 캐시는, 새로운 이미지를 OCR 캐시에 별도의 엔트리로서 저장하는 대신, 새로운 이미지를 이용하여 새로운 및/또는

양호한 정보로 업데이트된다. 대안으로, 새로운 이미지는 OCR 캐시에서의 이전의 이미지를 대체할 수도 있다. 덧붙여서, 다수의 이미지들을 사용하여 이미지로부터 글레어 (glare) 를 제거하는 기법들이 설명된다.

예를 들어, 이미지에 대한 지역으로부터의 글레어는, 동일한 장면으로부터, 키 이미지와 동일한 지역에 대한 글레어를 가지지 않는 다른 이미지로부터의 정보를 사용함으로써 제거될 수도 있다. 하나의 실시형태에서, 글레어는 심지어 OCR의 부재 시에도 이미지들로부터 감소되거나 또는 제거될 수도 있다. 하나의 실시형태에서, 글레어 감소와 OCR 캐시는 조합하여 수행될 수도 있다.

[0003] 지능적 및 자동 OCR 캐시 업데이트는 사용자에게 유익한데, 그것이 사용자가 동일한 장면과 연관된 정보의 모두에 액세스하는 것을 허용하고 및/또는 사용자가 복수의 장면들과 연관된 정보에 액세스하는 것을 허용할 수도 있어서이다. 더욱이, 시간이 지남에 따라 동일한 장면에 대한 다수의 이미지들 사이의 이미지 정보의 중복으로, 본원에서 설명되는 기법들은 장면과 연관된 최상의 또는 새로 획득된 정보를 저장하며, 그러므로 시간이 지남에 따라 결과들을 추가로 리파인할 수도 있다. 본원에서 설명되는 기법들과 연관된 다른 장점들은, OCR 결과들을 저장하는 감소된 메모리 공간과 OCR 캐시에 대한 감소된 히트 횟수를 포함하여, 전반적인 시스템에서 더 빠른 성능이 나타나게 한다.

[0004] 예로서, 사용자가 식당에 앉아 있으면서 스냅샷 모드 또는 미리보기 모드를 사용하여 이미지를 획득하는 전통적인 수단을 사용하여 정찬 메뉴에 대해 OCR을 수행하면, 그 사용자는 다수의 이미지들을 획득하는 것 및 이들 이미지들과 연관된 정보를 따로따로 프로세싱하는 것을 필요로 할 수도 있다. 사용자가 정찬 메뉴에서 너무 멀리 떨어져서 이미지를 획득하는 카메라를 보유하고 있으면, 그 이미지는 이미지에 대해 OCR을 적절하게 수행할 충분한 해상도를 가지지 못할 수도 있다. 한편 사용자가 이미지를 획득하는 카메라를 너무 가깝게 가져가면, 그 이미지는 전체 관심 영역을 포함하지 못할 수도 있다. 그러나, 자동 및 지능적 OCR 캐시 업데이트를 갖는 OCR 캐싱 모드를 사용하여, 본원에서 설명되는 기법들은, 적절한 초점으로 양호한 해상도에서 메뉴로부터 사용자에게 대한 관심 영역을 포함하는 단일 이미지를 생성하는 것을 사용자에게 허용한다. 어떤 경우에는, 사용자는 메뉴로부터 텍스트를 선택하고 그것을 다른 언어로 번역하거나 또는 온라인 검색 엔진을 사용하여 검색을 수행할 것을 선택할 수도 있다.

[0005] 위의 예에서, 장면으로부터의 이미지들 중 하나는 이미지의 지역과 연관된 글레어를 가질 수도 있다. 설명되는 기법들은, 제 1 이미지로부터 글레어를 갖는 제 1 지역을 식별하며, 제 2 이미지로부터 글레어가 없는 제 2 지역을 식별함으로써 글레어를 제거하는 것을 용이하게 할 수도 있는데, 제 1 지역 및 제 2 지역이 동일한 장면으로부터의 동일한 지역을 표현한다. 최종 이미지가 글레어를 제거하기 위해 제 1 이미지와, 제 2 이미지로부터의 글레어가 없는 제 2 지역을 사용하여 합성된다.

[0006] 본 발명의 실시형태들을 수행하는 예시적인 방법은, 복수의 이미지들을 수신하는 단계, 복수의 이미지들의 각각에서 데이터를 인식하는 단계로서, 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 단계를 포함하는 복수의 이미지들의 각각에서 데이터를 인식하는 단계, 인식된 데이터에 적어도 부분적으로 기초하여, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 단계, 및 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 적어도 하나의 이미지와 연관된 인식된 데이터를 사용하여 제 1 장면을 표현하는 최종 이미지를 저장하는 단계를 포함할 수도 있다. 하나의 실시형태에서, 수신하는 단계는, 카메라를 사용하여 복수의 이미지들을 연속해서 캡처하는 단계를 포함한다. 하나의 구현예에서, 복수의 이미지들 중 적어도 하나의 이미지 및 적어도 하나의 이미지에 대한 인식된 데이터의 저장은, 후속의 이미지가 캡처되고 상기 후속의 이미지에서 데이터가 인식되는 동안 유지된다. 하나의 양태에서, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 단계는, 2 개 이상의 이미지들의 각각의 이미지에서의 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색 및 전경색 (foreground color) 중 하나 이상을 비교하는 단계를 포함할 수도 있다.

[0007] 하나의 실시형태에서, 최종 이미지를 저장하는 단계 전에, 그 방법은, 적어도 2 개의 이미지들 및 적어도 2 개의 이미지들과 연관된 인식된 데이터를 사용하여 최종 이미지를 합성하는 단계, 및 후속하여 최종 이미지를 저장하는 단계를 더 포함한다. 하나의 구현예에서, 합성하는 단계는, 최종 이미지를 생성하기 위해 적어도 2 개의 이미지들 중 제 2 이미지로부터의 엘리먼트들을 적어도 하나의 이미지에 통합하는 단계를 포함한다. 다른 실시형태에서, 최종 이미지는 적어도 하나의 이미지를 포함하며, 합성하는 단계는, 적어도 2 개의 이미지들 중 제 2 이미지의 연관된 데이터를 적어도 하나의 이미지의 연관된 데이터에 통합하는 단계를 포함한다.

[0008] 다른 실시형태에서, 최종 이미지를 저장하는 단계 전에, 그 방법은, 최종 이미지로서 제 1 장면을 표현하기 위해 적어도 2 개의 이미지들로부터 적어도 하나의 이미지를 선택하는 단계 및 후속하여 최종 이미지를 저장하는



단계를 더 포함한다. 선택하는 단계는, 둘 이상의 이미지들의 각각의 이미지의 에너지를 비교하는 단계를 포함할 수도 있다. 둘 이상의 이미지들의 각각의 이미지의 에너지를 계산하는 단계는, 각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초할 수도 있다.

[0009] 그 방법은, 복수의 이미지들 중 하나 이상의 이미지들이 제 1 장면 이외의 하나 이상의 장면들로부터의 정보를 포함한다고 결정하는 단계, 및 최종 이미지 및 하나 이상의 장면들을 각각 표현하는 하나 이상의 이미지들 또는 적어도 하나의 이미지에 대한 및 하나 이상의 장면들에 대한 인식된 데이터를 디스플레이하는 단계를 더 포함한다. 더욱이, 본 발명의 실시형태들에 의해 수행되는 방법은, 사용자가 디스플레이된 이미지들 중 하나의 이미지 또는 하나의 이미지에 대한 디스플레이된 인식된 데이터를 선택했다고 결정하는 단계, 및 선택된 이미지 또는 인식된 데이터에 관련된 정보를 프로그램 또는 애플리케이션에 입력하는 단계를 또한 포함할 수도 있다.

[0010] 하나의 구현예에서, 그 방법은, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하며, 제 1 지역 및 제 2 지역이 제 1 장면의 동일한 지역을 표현하는, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하고, 제 1 이미지, 및 제 2 이미지로부터 글레이어가 없는 제 2 지역을 사용하여 감소된 글레이어를 갖는 최종 이미지를 합성함으로써 글레이어를 또한 제거하거나 또는 감소시킬 수도 있다.

[0011] 본 발명의 실시형태들을 수행하는 예시적인 디바이스는, 복수의 이미지들을 수신하도록 구성된 이미지 캡처 모듈, 광학적 문자 인식을 사용하여 복수의 이미지들의 각각에서 데이터를 인식하도록 구성된 OCR 엔진, 인식된 데이터에 적어도 부분적으로 기초하여, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하도록 구성된 매칭 모듈; 및 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 인식된 데이터를 사용하여 제 1 장면을 표현하는 최종 이미지를 저장하도록 구성된 통합 모듈을 구비할 수도 있다. 이미지들은 디바이스에 연결된 카메라를 사용하여 연속해서 캡처되고 획득될 수도 있다. 하나의 실시형태에서, 결정은 둘 이상의 이미지들의 각각의 이미지에서의 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색, 전경색 및 경계 박스 중 하나 이상을 비교하는 것을 포함한다.

[0012] 통합 모듈은, 적어도 2 개의 이미지들 및 적어도 2 개의 이미지들과 연관된 인식된 데이터를 사용하여 최종 이미지를 합성한 다음 후속하여 최종 이미지를 저장하도록 추가로 구성될 수도 있다. 하나의 실시형태에서, 합성하는 것은, 최종 이미지를 생성하기 위해 적어도 2 개의 이미지들 중 제 2 이미지로부터의 엘리먼트들을 적어도 하나의 이미지에 통합하는 것을 포함한다. 다른 실시형태에서, 최종 이미지는 적어도 하나의 이미지를 포함하며, 합성하는 것은, 적어도 2 개의 이미지들 중 제 2 이미지의 연관된 데이터를 적어도 하나의 이미지의 연관된 데이터에 통합하는 것을 포함한다.

[0013] 최종 이미지를 저장하기 전에, 선택 모듈은, 최종 이미지로서 제 1 장면을 표현하기 위해 적어도 2 개의 이미지들로부터 적어도 하나의 이미지를 선택하고 후속하여 최종 이미지를 저장하도록 구성될 수도 있다. 하나의 구현예에서, 선택하는 것은 상기 둘 이상의 이미지들의 각각의 이미지의 에너지를 비교하는 것을 포함한다. 둘 이상의 이미지들의 각각의 이미지의 에너지를 계산하는 것은, 각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초할 수도 있다.

[0014] 그 디바이스는, 적어도 2 개의 이미지들 또는 적어도 2 개의 이미지들에 대한 인식된 데이터를 디스플레이하도록 구성된 디바이스 디스플레이를 또한 포함할 수도 있다. 그 디바이스는, 사용자가 디스플레이된 이미지들 중 하나의 이미지 또는 하나의 이미지에 대한 디스플레이된 인식된 데이터를 터치했다고 결정하는 것, 및 터치된 이미지 또는 인식된 데이터에 관련된 정보를 프로그램 또는 애플리케이션에 입력하는 것을 더 포함할 수도 있다. 입력하는 것은, 인터넷 쿼리를 제출하는 것, 번역을 결정하는 것, 또는 텍스트를 입력하는 것을 포함할 수도 있다.

[0015] 디바이스의 하나의 예에서, 최종 이미지를 저장하기 전에, 그 디바이스는, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하며, 제 1 지역 및 제 2 지역이 제 1 장면의 동일한 지역을 표현하는, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하고, 제 1 이미지, 및 제 2 이미지로부터의 글레이어가 없는 제 2 지역을 사용하여 감소된 글레이어를 갖는 최종 이미지를 합성하도록 추가로 구성된다.

[0016] 예시적인 비밀시적 컴퓨터 판독가능 저장 매체인데, 비밀시적 컴퓨터 판독가능 저장 매체는 프로세서에 의해 실행

행되는 명령들을 포함하며, 그 명령들은, 복수의 이미지들을 수신하게 하는 명령, 복수의 이미지들에 대해 광학적 문자 인식을 수행하게 하는 것을 포함하여 복수의 이미지들의 각각에서 데이터를 인식하게 하는 명령, 인식된 데이터에 적어도 부분적으로 기초하여, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하게 하는 명령, 및 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 인식된 데이터를 사용하여 제 1 장면을 표현하는 최종 이미지를 저장하게 하는 명령을 포함한다.

[0017] 예시적인 장치는, 복수의 이미지들을 수신하는 수단, 복수의 이미지들의 각각에서 데이터를 인식하는 수단으로서, 복수의 이미지들의 각각에서 데이터를 인식하는 것이 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 것을 포함하는, 상기 인식하는 수단, 인식된 데이터에 적어도 부분적으로 기초하여, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 수단, 및 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 적어도 하나의 이미지와 연관된 인식된 데이터를 사용하여 제 1 장면을 표현하는 최종 이미지를 저장하는 수단을 포함할 수도 있다. 복수의 이미지들은 카메라를 사용하여 연속해서 캡처될 수도 있다. 복수의 이미지들 중 적어도 하나의 이미지 및 적어도 하나의 이미지에 대한 인식된 데이터의 저장은, 후속의 이미지가 캡처되고 상기 후속의 이미지에서 데이터가 인식되는 동안 유지될 수도 있다. 하나의 양태에서, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정하는 것은, 2 개 이상의 이미지들의 각각의 이미지에서의 그래픽 오브젝트 값들, 그래픽 오브젝트 컬러들, 배경색 및 전경색 중 하나 이상을 비교하는 것을 포함할 수도 있다.

[0018] 하나의 실시형태에서, 최종 이미지를 저장하기 전에, 그 장치는, 적어도 2 개의 이미지들 및 적어도 2 개의 이미지들과 연관된 인식된 데이터를 사용하여 최종 이미지를 합성한 다음 후속하여 최종 이미지를 저장하는 수단을 더 포함할 수도 있다. 다른 실시형태에서, 최종 이미지를 저장하기 전에, 그 장치는, 최종 이미지로서 제 1 장면을 표현하기 위해 적어도 2 개의 이미지들로부터 적어도 하나의 이미지를 선택한 다음 후속하여 최종 이미지를 저장하는 수단을 더 포함할 수도 있다. 선택하는 것은 둘 이상의 이미지들의 각각의 이미지의 에너지를 비교하는 수단을 포함할 수도 있다. 둘 이상의 이미지들의 각각의 이미지의 에너지를 계산하는 것은, 각각의 이미지에서 인식된 그래픽 오브젝트들의 개수 및 상기 개수의 그래픽 오브젝트들의 각각의 그래픽 오브젝트의 인식에 대한 신뢰도에 기초할 수도 있다.

[0019] 하나의 실시형태에서, 최종 이미지를 저장하기 전에, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 1 이미지로부터 글레이어를 갖는 제 1 지역을 식별하는 수단, 제 1 지역 및 제 2 지역이 제 1 장면의 동일한 지역을 표현하는, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이어가 없는 제 2 지역을 식별하는 수단, 제 1 이미지 및 제 2 이미지로부터의 글레이어가 없는 제 2 지역을 사용하여 감소된 글레이어를 갖는 최종 이미지를 합성하는 수단을 포함할 수도 있다.

[0020] 전술한 바는 다음의 상세한 설명이 양호하게 이해될 수 있도록 하기 위하여 예들의 특징들 및 기술적 장점들을 상당히 광범위하게 약술하고 있다. 부가적인 특징들 및 장점들은 이하에서 설명될 것이다. 개시된 개념 및 구체적인 예들은 본 개시물의 동일한 목적들을 수행하는 다른 구조들을 수정하거나 또는 설계하기 위한 기초로서 쉽사리 활용될 수 있다. 그런 동등한 구성들은 첨부 청구항들의 정신 및 범위로부터 벗어나지 않는다. 본원에서 개시된 개념들의 특징이 될 것이라 생각되는 특징부(feature) 들은, 그것들의 조직 및 방법 양쪽 모두에 관해, 연관된 이점들과 함께, 첨부 도면들에 관련하여 고려되는 경우에 다음의 설명으로부터 양호하게 이해될 것이다. 도면들의 각각은 예시 및 설명 목적으로만 제공되고 청구항들의 한계의 정의를 위해 제공되지는 않았다.

## 도면의 간단한 설명

[0021] 본 개시물의 양태들이 예로서 도시되어 있다. 다음의 설명이 도면들을 참조하여 설명되며, 그 도면들에서 유사한 참조 부호들은 전체에 걸쳐 유사한 엘리먼트들을 지칭하는데 사용된다. 하나 이상의 기법들의 다양한 세부사항들이 본원에서 설명되지만, 다른 기법들이 또한 가능하다. 일부 경우들에서, 잘 알려진 구조들 및 디바이스들은 다양한 기법들을 설명하는 것을 용이하게 하기 위하여 블록도 형태로 도시된다.

본 개시물에 의해 제공된 예들의 성질 및 장점들의 추가의 이해는 명세서의 나머지 부분들 및 도면들을 참조하여 실현될 수 있으며, 유사한 참조 번호들은 여러 도면들을 통해 유사한 컴포넌트들을 참조하는데 이용된다. 어떤 경우들에서는, 서브 라벨이 다수의 유사한 컴포넌트들 중 하나를 나타내는 참조 번호와 연관된다. 현존 서브 라벨에 대한 명세 없이 참조 번호가 참조되는 경우, 그 참조 번호는 모든 그러한 유사한 컴포넌트들

을 언급한다.

도 1은 본 발명의 실시형태들을 실시할 때에 채용되는 디바이스의 부분들을 통합하는 예시적인 컴퓨터 디바이스를 도시한다.

도 2는 시간이 지남에 따른 이미지들의 OCR 프로세싱을 위한 본 발명의 예시적인 실시형태를 도시하는 흐름도이다.

도 3a는 이미지들의 OCR 프로세싱을 위한 본 발명의 비제한적 예시적인 실시형태를 보여주는 흐름도를 도시한다.

도 3a는 이미지들의 OCR 프로세싱을 위한 본 발명의 다른 비제한적 예시적인 실시형태를 보여주는 흐름도를 도시한다.

도 4는 본 발명의 예시적인 실시형태에 따른 책 표지의 대표적인 이미지 및 연관된 장면 기술자 (scene descriptor) 를 보여주는 2 개의 블록 도면들을 묘사한다.

도 5는 복수의 이미지들로부터의 저장된 이미지와 동일한 장면을 표현하는 이미지의 선택을 보여주는 블록 도면들을 묘사한다.

도 6은 저장된 이미지와 동일한 장면으로부터 이미지를 검출하는 비제한적 예시적인 방법을 보여주는 흐름도를 도시한다.

도 7은 복수의 이미지들로부터 키 이미지로서의 이미지의 선택을 보여주는 블록 도면들을 묘사한다.

도 8은 복수의 이미지들로부터 키 이미지로서의 이미지의 선택에 대한 비제한적 예시적인 방법을 보여주는 흐름도를 도시한다.

도 9는 동일한 장면으로부터의 복수의 이미지들로부터의 최종 이미지로서의 이미지의 통합을 보여주는 블록 도면들을 묘사한다.

도 10은 동일한 장면으로부터의 복수의 이미지들로부터 최종 이미지를 합성하는 비제한적 예시적인 방법을 보여주는 흐름도를 도시한다.

도 11은 본 발명의 실시형태들에 의해 제공된 방법들을 수행하는 예시적인 컴포넌트들을 보여주는 블록도를 묘사한다.

도 12는 본 발명의 실시형태들에 의해 제공된 방법들을 수행하는 예시적인 컴포넌트들을 보여주는 다른 블록도를 묘사한다.

도 13a 및 도 13b는 텍스트를 포함하는 이미지들에 대한 예시적인 장면 기술자들을 묘사한다.

도 14는 이미지로부터 글레이어를 제거하거나 또는 감소시키는 본 발명의 실시형태들을 예시하는 도면을 묘사한다.

도 15는 본 개시물의 하나 이상의 예시적 양태들에 따라 이미지에서 반사 글레이어를 제거하거나 또는 감소시키는 본 발명의 실시형태를 예시하는 흐름도이다.

도 16a, 도 16b, 도 16c, 및 도 16d는 자동 및 지능적 OCR 캐시 업데이트들을 가지고서 OCR 캐싱 모드로 동작하는 몹시 단순화된 모바일 디바이스를 사용하여 사용자 인터페이스에 대한 예시적인 구현예를 묘사한다.

### 발명을 실시하기 위한 구체적인 내용

[0022]

모바일 디바이스에 연결된 카메라의 시야에서 텍스트를 포함하는 이미지들이 일부 구현예들에서 스냅샷 모드 또는 미리보기 모드를 사용하여 획득될 수도 있다. 스냅샷 모드에서, 사용자는 사용자가 시야에서 캡처하고자 하는 텍스트에 카메라 초점을 맞추고 사진을 찍는다. 결과적인 화상은 컴퓨터 디바이스에 의해 분석될 수도 있고 OCR 결과들이 생성된다. 한편, 미리보기 모드는 시야에서의 텍스트의 인식을 지속적으로 루핑하고 획득된 각각의 이미지 또는 프레임에 대해 OCR을 수행한다. 이 모드는, 어떤 경우들에서는 카메라 뷰파인더 바로 옆에서 디바이스가 인식하고 있는 것의 동적인 실시간 디스플레이를 보여준다. 사용자가 이미지의 OCR 결과들을 검토하는데 더 많은 시간을 소비하거나 또는 그 결과들을 이용하고 싶어하는 경우, 사용자는 이미지 캡처/OCR 인식을 중단할 수도 있고 그 후 정지시킨 (frozen) 캡처/인식에 의해 현재 디스플레이된 결과들로 진

행할 수도 있다.

[0023] 스냅샷 모드 및 미리보기 모드는, 관련 기술분야 응용들에서 구현된 바와 같이, 자동 이력 보관을 위한 메커니즘들을 가지지 않는다. 사용자가 큰 입력 타겟 또는 많은 타겟들과 마주하는 경우, 그 사용자는 스냅샷 모드 또는 미리보기 모드를 통해 다수의 이미지들을 취하는 것, 각각의 이미지에 대해 OCR 엔진을 정지시키는 것 및 각각의 이미지를 별도로 분석하는 것에 의존해야만 한다. 특히 미리보기 모드에서, 사용자에게는 그 사용자가 관심 영역 전체에 걸쳐 모바일 디바이스를 천천히 이동시키면서 반복적으로 취해진 동일한 장면의 많은 이미지들이 채도한다. 또한, 현재 OCR 시스템들은 OCR 성능을 개선하기 위해 상이한 OCR 결과들로부터의 정보를 종합하는, 또는 결과적인 캐싱된 이미지들에서 글레이어를 감소시키는 메커니즘을 가지지 않는다.

[0024] 본 발명의 실시형태들은 이들 및 다른 문제들을 해결한다.

[0025] 도 1은 본 발명의 실시형태들을 실시할 때 채용되는 디바이스의 부분들을 통합하는 예시적인 컴퓨터 디바이스를 도시한다. 도 1에 예시된 바와 같은 컴퓨터 디바이스가 본원에서 임의의 컴퓨터화된 시스템의 부분으로서 통합될 수도 있다. 예를 들어, 컴퓨터 디바이스 (100) 는 모바일 디바이스의 컴포넌트들의 일부를 나타낼 수 있다. 모바일 디바이스는 카메라 (150) 와 같은 하나 이상의 입력 감각 유닛 또는 입력 디바이스들 (115) 및 디스플레이 유닛 또는 터치 스크린과 같은 하나 이상의 입력/출력 디바이스들을 갖는 임의의 컴퓨팅 디바이스 (100) 일 수도 있다. 모바일 디바이스의 예들은 비디오 게임 콘솔들, 태블릿들, 스마트 폰들, 랩톱들, 넷북들, 또는 다른 휴대용 디바이스들을 포함하지만 그것들로 제한되지 않는다. 도 1은, 본원에서 설명되는 바와 같이, 다양한 다른 실시형태들에 의해 제공되는 방법들을 수행할 수 있고 및/또는 호스트 컴퓨터 디바이스, 원격 키오스크/단말, 판매시점관리 (point of sale) 디바이스, 모바일 디바이스, 셋톱 박스 및/또는 컴퓨터 디바이스로서 기능할 수 있는 컴퓨터 디바이스 (100) 의 하나의 실시형태의 개략적인 예시도를 제공한다. 도 1은 컴포넌트들 중 임의의 것 또는 모두가 적절하게 활용될 수도 있는 다양한 컴포넌트들의 일반화된 예시도를 제공하는 것만을 의미한다. 도 1은 그러므로 어떻게 개개의 시스템 엘리먼트들이 비교적 분리되거나 또는 비교적 더 통합되는 방식으로 구현될 수도 있는지를 광범위하게 예시한다.

[0026] 컴퓨터 디바이스 (100) 는 버스 (105) 를 통해 전기적으로 연결될 수 있는 (또는 그렇지 않으면 적절한 대로 통신될 수도 있는) 하드웨어 엘리먼트들을 포함하는 것으로 도시된다. 하드웨어 엘리먼트들은, 하나 이상의 범용 프로세서들 및/또는 하나 이상의 특수 목적 프로세서들 (이를테면 디지털 시그널 프로세싱 칩들, 그래픽 가속도 프로세서들 등) 을 제한 없이 포함하는 하나 이상의 프로세서들 (110); 카메라, 센서들 (관성 센서들을 포함함), 마우스, 키보드 등을 제한 없이 포함할 수 있는 하나 이상의 입력 디바이스들 (115); 및 도 11 및 도 12의 디바이스 디스플레이 (1118) 와 같은 디스플레이 유닛, 프린터 등을 제한 없이 포함할 수 있는 하나 이상의 출력 디바이스들 (120) 을 포함할 수도 있다.

[0027] 컴퓨터 디바이스 (100) 는, 로컬 및/또는 네트워크 액세스가능 스토리지를 제한 없이 포함할 수 있고 및/또는 디스크 드라이브, 드라이브 어레이, 광학적 저장 디바이스, 고체 상태 저장 디바이스 이를테면 프로그래밍가능, 플래시 업데이트 가능 등일 수 있는 랜덤 액세스 메모리 ("RAM") 및/또는 판독 전용 메모리 ("ROM") 를 제한 없이 포함할 수 있는 하나 이상의 비일시적 스토리지 디바이스들 (125) 을 추가로 구비할 수도 있다 (및/또는 그것들과 통신할 수도 있다). 이러한 저장 디바이스들은 다양한 파일 시스템들, 데이터베이스 구조들 등을 제한 없이 포함하는 임의의 적절한 데이터 스토리지를 구현하도록 구성될 수도 있다. 도 11 및 도 12로부터의 디바이스 스토리지 (1120) 는 위에서 논의된 바와 같은 하나 이상의 비일시적 스토리지 디바이스들을 사용할 수도 있다.

[0028] 컴퓨터 디바이스 (100) 는 모뎀, 네트워크 카드 (무선 또는 유선), 적외선 통신 디바이스, 무선 통신 디바이스 및/또는 칩셋 (이를테면 블루투스™ 디바이스, 802.11 디바이스, WiFi 디바이스, 와이맥스 디바이스, 셀룰러 통신 설비들 등) 등을 제한 없이 포함할 수 있는 통신 서브시스템 (130) 을 또한 구비할지도 모른다. 통신 서브시스템 (130) 은 네트워크 (이를테면 하나의 예를 명명하는 아래에서 설명되는 네트워크), 다른 컴퓨터 디바이스들, 및/또는 본원에서 설명되는 임의의 다른 디바이스들과 데이터가 상호교환되는 것을 허용할 수도 있다. 통신 서브시스템 (130) 은 일부 실시형태들에서, 예를 들어 다른 디바이스들과 통신하도록 구성되지 않은 디바이스들에서 생략될 수도 있다. 많은 실시형태들에서, 컴퓨터 디바이스 (100) 는 위에서 설명된 바와 같이 RAM 또는 ROM 디바이스를 포함할 수 있는 비일시적 작업 메모리 (135) 를 더 포함할 것이다. OCR 캐시 (1112) 는, 위에서 설명된 바와 같이, 비일시적 작업 메모리 (135) 를 사용하여 구현될 수도 있다.

[0029] 컴퓨터 디바이스 (100) 는 또한, 운영 체제 (140), 디바이스 드라이버들, 실행가능 라이브러리들, 및/또는 다른 코드, 이를테면 하나 이상의 애플리케이션 프로그램들 (145) 을 포함하는, 작업 메모리 (135) 내에 현재 위치되



어 있는 것으로서 도시되는 소프트웨어 엘리먼트들을 포함할 수 있으며, 그 소프트웨어 엘리먼트들은, 본원에서 설명되는 바와 같이, 다양한 실시형태들에 의해 제공되는 컴퓨터 프로그램들을 포함할 수 있고 및/또는 다른 실시형태들에 의해 제공되는 방법들을 구현하도록, 및/또는 시스템들을 구성하도록 설계될 수도 있다. 단순히 예로서, 위에서 논의된 방법(들)에 관해 설명된 하나 이상의 프로시저들은 컴퓨터 (및/또는 컴퓨터 내의 프로세서)에 의해 실행가능한 코드 및/또는 명령들로서 구현될지도 모르고; 일 양태에서, 그 다음에, 이러한 코드 및/또는 명령들은 설명된 방법들에 따라 하나 이상의 동작들을 수행하기 위해 범용 컴퓨터 (또는 다른 디바이스)를 구성 및/또는 적응하는데 이용될 수 있다.

[0030] 이들 명령들의 세트 및/또는 코드는 컴퓨터 판독가능 저장 매체, 이를테면 위에서 설명된 스토리지 디바이스(들) (125) 상에 저장될지도 모른다. 일부 경우들에서, 저장 매체는 컴퓨터 디바이스, 이를테면 컴퓨터 디바이스 (100) 내에 통합될지도 모른다. 다른 실시형태들에서, 저장 매체는 컴퓨터 디바이스로부터 분리되고 (예컨대, 콤팩트 디스크와 같은 착탈식 매체), 및/또는 설치 패키지 내에 제공될지도 몰라서, 저장 매체는 명령들/코드를 저장하고 있는 범용 컴퓨터를 프로그래밍하며, 구성하고 및/또는 적응시키는데 사용될 수 있다. 이들 명령들은, 컴퓨터 디바이스 (100)에 의해 실행가능한 실행가능 코드의 형태를 취할지도 모르고 (예컨대, 다양한 일반적으로 이용가능한 컴파일러들, 설치 프로그램들, 압축/압축해제 유틸리티들 등 중 임의의 것을 사용하여) 컴퓨터 디바이스 (100) 상의 컴파일 및/또는 설치 시, 실행가능 코드의 형태를 취하는 소스 및/또는 설치가능 코드의 형태를 취할지도 모른다.

[0031] 실질적인 변동들이 특정 요건들에 따라 이루어질 수도 있다. 예를 들어, 커스터마이징된 하드웨어가 또한 사용될 수도 있고, 및/또는 특정한 엘리먼트들은 하드웨어, 소프트웨어 (휴대용 소프트웨어, 이를테면 애플릿들 등을 포함함), 또는 양쪽 모두로 구현될지도 모른다. 게다가, 네트워크 입력/출력 디바이스들과 같은 다른 컴퓨팅 디바이스들에 대한 접속이 채용될 수도 있다.

[0032] 일부 실시형태들은 본 개시물에 따른 방법들을 수행하는 컴퓨터 디바이스 (이를테면 컴퓨터 디바이스 (100))를 채용할 수도 있다. 예를 들어, 설명된 방법들의 프로시저들의 일부 또는 전부는 작업 메모리 (135)에 포함된 (운영 체제 (140) 및/또는 다른 코드, 이를테면 애플리케이션 프로그램 (145) 속에 통합될 지도 모르는) 하나 이상의 명령들의 하나 이상의 시퀀스들을 실행하는 프로세서 (110)에 응답하여 컴퓨터 디바이스 (100)에 의해 수행될 수도 있다. 이러한 명령들은 다른 컴퓨터 판독가능 매체, 이를테면 스토리지 디바이스(들) (125) 중 하나 이상으로부터 작업 메모리 (135) 속으로 읽혀질 수도 있다. 단순히 예로서, 작업 메모리 (135)에 포함된 명령들의 시퀀스들의 실행은 프로세서(들) (110)로 하여금 위에서 설명된 방법들의 하나 이상의 프로시저들을 수행하게 할지도 모른다.

[0033] 용어들 "머신 판독가능 매체" 및 "컴퓨터 판독가능 매체"는, 본원에서 사용되는 바와 같이, 머신으로 하여금 특정 방식으로 동작하게 하는 데이터의 제공에 참여하는 임의의 매체를 말한다. 컴퓨터 디바이스 (100)를 사용하여 구현되는 실시형태에서, 다양한 컴퓨터 판독가능 매체들은 실행을 위해 프로세서(들) (110)로 명령들/코드의 제공에 관계될지도 모르고 및/또는 그런 명령들/코드를 (예컨대, 신호들로서) 저장 및/또는 반송하는데 사용될지도 모른다. 많은 구현예들에서, 컴퓨터 판독가능 매체는 물리적 및/또는 유형의 저장 매체이다. 이러한 매체는 비휘발성 매체들, 휘발성 매체들, 및 송신 매체들을 포함하지만 그것들로 한정되지 않는 많은 형태들을 취할 수도 있다. 비휘발성 매체들은, 예를 들어, 광 및/또는 자기 디스크들, 이를테면 스토리지 디바이스(들) (125)를 포함한다. 휘발성 매체들은 동적 메모리, 이를테면 작업 메모리 (135)를 제한 없이 포함한다. 송신 매체들은, 버스 (105)를 포함하는 와이어들뿐만 아니라 통신 서브시스템 (130)의 다양한 컴포넌트들 (및/또는 통신 서브시스템 (130)이 다른 디바이스들과의 통신을 하게 하는 매체들)을 포함하여, 동축 케이블들, 구리 선 및 광섬유들을 제한 없이 포함한다. 따라서, 송신 매체들은 또한 (라디오, 음향 및/또는 광 파들, 이를테면 라디오 파 및 적외선 데이터 통신들 동안에 생성된 것들을 제한 없이 포함한) 파들의 형태를 취할 수 있다.

[0034] 물리적 및/또는 유형의 컴퓨터 판독가능 매체들의 공통 형태들은, 예를 들어, 플로피 디스크, 플렉시블 디스크, 하드 디스크, 자기 테이프, 또는 임의의 다른 자기 매체, CD-ROM, 임의의 다른 광학적 매체, 천공카드들, 종이 테이프, 구멍들의 패턴들을 갖는 임의의 다른 물리적 매체, RAM, PROM, EPROM, FLASH-EPROM, 임의의 다른 메모리 칩 또는 카트리지, 이후로 설명되는 바와 같은 반송파, 또는 컴퓨터가 명령들 및/또는 코드를 읽을 수 있는 임의의 다른 매체를 포함한다.

[0035] 컴퓨터 판독가능 매체들의 다양한 형태들이 하나 이상의 명령들의 하나 이상의 시퀀스들을 실행을 위해 프로세서(들) (110)에 반송하는 것에 관련될 수도 있다. 단순히 예로서, 명령들은 초기에는 원격 컴퓨터의 자기

디스크 및/또는 광 디스크로 반송될 수도 있다. 원격 컴퓨터는 명령들을 그것의 동적 메모리에 로딩하고 그 명령들을 컴퓨터 디바이스 (100) 에 의해 수신되고 및/또는 실행되게 송신 매체를 통해 신호들로서 전송할지도 모른다. 전자기 신호들, 음향 신호들, 광 신호들 등의 형태일지도 모르는 이들 신호들은, 명령들이 본 발명의 다양한 실시형태들에 따라 인코딩될 수 있는 반송파들의 모든 예들이다.

[0036] 통신 서브시스템 (130) (및/또는 그것의 컴포넌트들) 은 일반적으로 신호들을 수신할 것이고, 버스 (105) 는 그 다음에 그 신호들 (및/또는 그 신호들에 의해 반송되는 데이터, 명령들 등) 을 작업 메모리 (135) 로 반송할지도 모르며, 그 작업 메모리로부터 프로세서(들) (110) 는 명령들을 추출하고 실행한다. 작업 메모리 (135) 에 의해 수신된 명령들은 옵션으로 프로세서(들) (110) 에 의한 실행 전 또는 후 중 어느 하나에 비일시적 스토리지 디바이스 (125) 상에 저장될 수도 있다.

[0037] 위에서 논의된 방법들, 시스템들, 및 디바이스들은 예들이다. 다양한 실시형태들이 다양한 프로시저들 또는 컴포넌트들을 적절한 대로 생략, 대체, 또는 추가할 수도 있다. 예를 들면, 대안적 구성들에서, 설명되는 방법들은 설명되는 것과 상이한 순서로 수행될 수도 있고, 및/또는 다양한 스테이지들이 추가되며, 생략되고, 및/또는 조합될 수도 있다. 또한, 특정 실시형태들에 관해 설명되는 특징들은 다양한 다른 실시형태들에 조합될 수도 있다. 실시형태들의 상이한 양태들 및 엘리먼트들이 유사한 방식으로 조합될 수도 있다. 또한, 기술이 진화하고, 따라서, 엘리먼트들의 대부분은 본 개시물의 범위를 그들 구체적인 예들로 제한하지 않는 예들이다.

[0038] 특정 세부사항들이 실시형태들의 철저한 이해를 제공하기 위해 상세한 설명에서 주어진다. 그러나, 실시형태들은 이들 특정 세부사항들 없이 실시될 수도 있다. 예를 들어, 잘 알려진 회로들, 프로세스들, 알고리즘들, 구조들, 및 기법들은 실시형태들을 불분명하게 하는 것을 피하기 위하여 불필요한 세부사항 없이 보여지고 있다. 이 설명은 예의 실시형태들만을 제공하고, 본 발명의 범위, 적용 가능성, 또는 구성을 제한하는 의도는 아니다. 오히려, 실시형태들의 앞서의 설명은 당업자들에게 본 발명의 실시형태들을 구현하기 위한 가능한 설명을 제공할 것이다. 다양한 변경들이 엘리먼트들의 기능 및 배열에서 본 발명의 사상 및 범위로부터 벗어남 없이 이루어질 수도 있다.

[0039] 또한, 일부 실시형태들은 흐름도들 또는 블록도들로서 묘사된 프로세스들로서 설명된다. 비록 각각이 순차적 프로세스로서 동작들을 설명할 수도 있지만, 그 동작들의 다수가 병행하여 또는 동시에 수행될 수 있다. 덧붙여서, 그 동작들의 순서는 재배열될 수도 있다. 프로세스가 도면에 포함되지 않은 부가적인 단계들을 가질 수도 있다. 더욱이, 그 방법의 실시형태들은 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 하드웨어 서술 언어들 또는 그것들의 임의의 조합에 의해 구현될 수도 있다. 소프트웨어, 펌웨어, 미들웨어, 또는 마이크로코드로 구현되는 경우, 연관된 태스크들을 수행하는 프로그램 코드 또는 코드 세그먼트들은 저장 매체와 같은 컴퓨터 판독가능 매체에 저장될 수도 있다. 프로세서들이 연관된 태스크들을 수행할 수도 있다.

[0040] 여러 실시형태들이 본원에서 설명되지만, 다양한 수정들, 대안적 구성들, 및 동등물이 본 개시물의 정신으로 벗어남 없이 사용될 수도 있다. 예를 들어, 위의 엘리먼트들은 단순히 큰 시스템의 컴포넌트일 수도 있으며, 다른 규칙들은 본 발명의 애플리케이션보다 우선시 되거나 또는 그렇지 않으면 본 발명의 애플리케이션을 수정할 수도 있다. 또한, 다수의 단계들은 엘리먼트들이 고려되기 전, 고려되는 동안, 또는 고려된 후에 착수될 수도 있다. 따라서, 본원에서의 설명은 본 개시물의 범위를 제한하지 않는다.

[0041] 본원에서 설명되는 바와 같이, 이미지로부터의 그래픽 오브젝트는 하나 이상의 심볼들, 영숫자 문자들, 단어들, 기호들, 숫자들, 텍스트 또는 제스처들을 포함할 수도 있지만 그것들로 제한되지 않는다. 예를 들어, 이미지에서의 그래픽 오브젝트는 임의의 언어에서의 단어 또는 문장일 수도 있다. 단어가 일부 언어들에서 심볼로서 표현될 수도 있다. 마찬가지로, 하나의 언어 또는 문화에서의 문장은 단순히 다른 언어 또는 문화에서의 심볼을 이용하여 표현될 수도 있다. 다른 예에서, 그래픽 오브젝트는 멈춤, 양보 등과 같은 도로 표지판일 수도 있다. 상이한 그래픽 오브젝트들이 상이한 지역들, 언어들 또는 문화들에서 동일한 의미를 표현하기 위해 사용될 수도 있다.

[0042] 도 2는 시간이 지남에 따라 이미지들의 프로세싱을 위한 본 발명의 예시적인 실시형태를 도시하는 흐름도이다. 위에서, 도 1을 참조하여 논의된 컴퓨터 디바이스 (100) 는, 도 2에서 설명된 본 발명의 실시형태들을 수행하는데 사용되는 모바일 디바이스의 컴포넌트들의 일부를 나타낼 수 있다. 도 11에서 논의되는 모듈들 및 컴포넌트들은 도 1에서 논의된 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 구현될 수도 있고, 도 2에서 논의되는 바와 같은 본 발명의 실시형태들을 수행하는데 사용될 수도 있다. 아래에서 논의되는 모듈들 및 컴포넌트

들은 소프트웨어, 하드웨어, 펌웨어 또는 그것들의 임의의 조합을 사용하여 수행될 수도 있다. 하나의 양태에서, OCR 캐싱 모드로서 본원에서 언급되는 모드에서, 컴퓨터 디바이스 (100) 에 연결된 카메라는 시간이 지남에 따라 이미지들을 지속적으로 획득한다.

[0043] 도 2에서, 복수의 이미지들은 I (202) 에 의해 표현되며, 여기서 각각의 I (202) 는 상이한 수신된 이미지를 나타낸다. 하나의 실시형태에서, 이미지 캡처 모듈 (1102) 은, 컴퓨팅 디바이스 (100) 에 연결된 하나 이상의 카메라들 (150) 을 사용하여 이미지들을 획득하거나 또는 수신할 수도 있다. 본 발명의 양태들에서, 이미지들이 수신되면, OCR 엔진 (206) (도 11의 1104) 은 그 이미지들을 프로세싱하여 각각의 이미지에 대한 OCR 결과들 (210) 을 생성하며, 복수의 이미지들의 각각에서 데이터를 인식한다. OCR 엔진 (206) 은 하드웨어, 소프트웨어 또는 펌웨어로 구현될 수도 있다. 소프트웨어/펌웨어로 구현되는 OCR에 대한 명령들은 작업 메모리 (135) 에 일시적으로 저장되고 컴퓨팅 디바이스 (100) 의 하나 이상의 프로세서들 (110) 상에서 실행될 수도 있다. 일부 실시형태들에서, OCR 엔진 (206) 에 의한 OCR 프로세싱의 일부 또는 전부는 네트워크를 통해 컴퓨팅 디바이스 (100) 에 접속된 원격 서버 (미도시) 에 의해 수행될 수도 있다. 일부 실시형태들에서, 저장된 결과의 이미지 품질은 시간이 지남에 따라 개선된다. 예를 들어, 하나의 실시형태에서, 글레이어는 OCR을 수행하는 것에 더하여 또는 그것과는 독립적으로 이미지들로부터 제거될 수도 있다. OCR 캐싱 모드에서, 시간이 지남에 따라 스냅샷 모드를 사용하여 획득된 이미지들은 이미지들을 OCR 캐시에 지능적으로 및 자동으로 저장하기 위해 본원에서 설명되는 기법들을 또한 사용할 수도 있다. OCR 캐시는 스토리지 디바이스 (125), 작업 메모리 (135), 하드웨어 레지스터들 (미도시) 또는 버퍼들을 사용하여 또는 그것들의 임의의 조합으로 구현될 수도 있다.

[0044] OCR 결과들은 이미지의 특성들을 설명하는 이미지에 관한 부가적인 데이터로 구성될 수도 있다. 하나의 구현예에서, 장면 기술자 생성기 모듈 (1106) 은 이미지에 대한 OCR 결과들을 이미지에 대한 장면 기술자로서 생성하고 그것을 OCR 캐시에 임시로 저장한다. 장면 기술자 모듈 (1106) 이 장면 기술자들을 생성한 후, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 매칭 모듈 (1108) 은, 매칭 프로세스를 사용함으로써 동일한 장면과 연관된 이미지들을 검출할 수도 있다. 하나의 구현예에서, 매칭 모듈은 매칭 프로세스 동안 이미지들과 연관된 장면 기술자들을 비교할 수도 있다. 도 2에서, OCR 결과들 (212 및 214) 로부터의 인식된 데이터는 장면 A와 연관되며, OCR 결과 (216) 는 장면 B와 연관되며, OCR 결과들 (218 및 222) 은 장면 C와 연관되고, OCR 결과 (220) 는 장면 D와 연관된다.

[0045] 212-222 OCR 결과들 및 연관된 이미지들의 모두를 OCR 캐시 (224) 에 저장하는 대신, OCR 캐싱 시스템은 각각의 장면에 대한 하나의 리파인된 이미지 및 OCR 결과들만을 지능적으로 및 자동으로 저장할 수도 있다. 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 선택 모듈 (1110) 은 장면으로부터의 복수의 이미지들로부터의 이미지를 OCR 캐시 (224) 에 저장하기 위해 선택할 수도 있다. 선택 모듈 (1110) 에 의한 이미지를 선택하는 프로세스는 선택 프로세스라고 지칭될 수도 있다. 예를 들면, 도 2에서, 장면 A를 표현하는 OCR 결과들 (212 및 214) 에 대응하는 이미지들은, OCR 캐시 (224) 에 A' (228) 으로서 저장될 수도 있다. 최종 합성된 또는 선택된 이미지 및 장면 기술자는 212 및 214에 대응하는 이미지들을 사용하여 OCR 캐시 (224) 에 저장된다. 하나의 양태에서, 선택 모듈은 선택 프로세스의 부분으로서 각각의 이미지와 연관된 에너지를 계산한다. 이미지와 연관된 에너지는 이미지와 연관된 선명도 (clarity) 및 해상도를 나타낼 수도 있다. 2 개의 이미지들 (212 및 214에 대응함) 로부터의 더 높은 에너지를 갖는 이미지 및 연관된 장면 기술자는 OCR 캐시 (224) 에 A' (228) 으로서 저장될 수도 있다. OCR 캐시 (224) 는, 도 1을 참조하여 위에서 설명된 바와 같이, 비일시적 작업 메모리 (135) 를 사용하여 구현될 수도 있다.

[0046] 다른 구현예에서, 통합 모듈 (1114) 은 최종 A' (228) 을 합성하는 동안 212 및 214 사이의 상이한 이미지 품질들 및 OCR 결과들을 사용하여 키 이미지로부터 누락된 이미지 데이터를 추가할 수도 있다. 최종 결과를 합성하는 프로세스는 통합 프로세스라고 지칭될 수도 있다. 하나의 예에서, 장면으로부터의 이미지들 중 하나의 이미지 (212에 대응한다고 함) 는 그 이미지의 지역과 연관된 글레이어를 가질 수도 있다. 통합 프로세스는, 212에 대응하는 이미지로부터 글레이어를 갖는 제 1 지역을 식별하며; 제 1 지역 및 제 2 지역이 장면 A의 동일한 지역을 표현하는, 214에 대응하는 이미지로부터 글레이어가 없는 제 2 지역을 식별하며; 및 글레이어를 제거하기 위해 212에 대응하는 이미지 및 214에 대응하는 이미지로부터의 글레이어가 없는 제 2 지역을 사용하여 최종 이미지를 합성함으로써 글레이어를 제거하는 것을 용이하게 할 수도 있다.

[0047] 마찬가지로, 218 및 222에 대응하는 이미지들은 동일한 장면 C와 연관된 것으로서 인식되고 최종 이미지 및 장면 기술자 C' (232) 을 형성하기 위해 합성된다. 이미지를 합성하는 것은, 본원에서 설명된 바와 같이, 이미지 또는 그 이미지에 대한 장면 기술자 또는 그 이미지의 표현을 포함하는 임의의 연관된 데이터를 OCR 캐시

에서 업데이트하거나 또는 대체하는 것을 포함하지만 그것으로 제한되지 않을 수도 있다. 따라서, 일부 실시형태들에서, 이미지를 합성하는 것 또는 복수의 이미지들을 함께 합성하는 것은, 예를 들어, 제 1 이미지의 외관 또는 시각적 표현을 변경하거나 또는 수정하는 일 없이, 제 2 이미지로부터의 데이터로써 (또는 복수의 이미지들로써), 제 1 이미지와 연관된 데이터를 업데이트하는 것 또는 수정하는 것을 포함할 수도 있다. 216에 대응하는 이미지는 장면 B를 표현하는 유일한 이미지이다. 일단 216에 대응하는 이미지가 장면 B를 표현하는 유일한 이미지라고 매칭 모듈 (1108) 이 결정하면, 삽입 모듈 (1116) 은 216에 대응하는 이미지를 OCR 캐시 (224) 에서의 캐시 로케이션 (230) 에 저장할 수도 있다. 마찬가지로 220에 대응하는 이미지는 장면 D를 표현하는 유일한 이미지이고 OCR 캐시 (224) 에서의 캐시 로케이션 (234) 에 저장될 수도 있다.

[0048] 수신된 이미지는 임시 버퍼, OCR 캐시 (224) 또는 다른 적합한 저장 매체에 저장된 동일한 장면과 연관된 다른 이미지들을 가지는 장면과 연관될 수도 있다. 최근에 수신된 이미지가 동일한 장면에 대해 OCR 캐시 (224) 에 저장된 이미지 및 OCR 결과와 매치되면, OCR 캐시 (224) 에서의 정보는 가장 최근에 수신된 이미지의 결과들을 통합한 후 그 장면에 대해 통합 모듈 (1114) 에 의해 업데이트될 수도 있다.

[0049] 하나의 양태에서, 이미지들은 그것들이 수신될 때 즉시 프로세싱된다. 다른 실시형태에서, 획득된 이미지들은 임시 저장되고 함께 일괄 프로세싱되어, 더 적은 OCR 캐시 업데이트들을 잠정적으로 허용한다. 또 다른 실시형태에서, 슬라이딩 윈도우가 함께 시간의 특정 윈도우 내에 속하는 이미지들을 프로세싱하는데 사용될 수도 있다. 프로세싱은 모바일 디바이스, 예를 들어 컴퓨팅 디바이스 (100) 에서 수행될 수도 있거나, 또는 정보는 프로세싱을 위해 원격 서버 또는 프로세서로 송신될 수도 있다. 예를 들어, 하나의 구현예에서, 컴퓨팅 디바이스 (100) 는 복수의 이미지들을 수신하고 그 이미지들 또는 그 이미지들에 관련된 정보를 OCR, 매칭 프로세스, 선택 프로세스, 및/또는 통합 프로세스를 수행하는 원격 서버로 송신할 수도 있다.

[0050] 이 지능적 및 자동 OCR 캐시 업데이트는 사용자에게 유익한데, 이것이 사용자에게 동일한 장면에 또는 복수의 장면들과 연관된 정보의 모두에 액세스하는 것을 허용해서이다. 더욱이, 시간이 지남에 따라 동일한 장면에 대한 다수의 이미지들 사이의 이미지 정보의 중복으로, 본원에서 설명되는 기법들은 장면과 연관된 최상의 정보 또는 새로 획득된 정보를 저장하며, 그러므로 시간이 지남에 따라 결과들을 추가로 리파인할 수도 있다. 본원에서 설명되는 기법들과 연관된 다른 장점들은, OCR 결과들을 저장하는데 필요한 감소된 메모리 공간과 OCR 캐시에 대한 감소된 히트 횟수를 포함하여 전반적인 시스템에서 더 빠른 성능이 나타나게 한다.

[0051] 예로서, 사용자가 식당에 앉아 있으면서 스냅샷 모드 또는 미리보기 모드를 사용하여 이미지를 획득하는 전통적인 수단을 사용하여 정찬 메뉴에 대해 OCR을 수행하면, 그 사용자는 다수의 이미지들을 따로따로 획득하는 것 및 그들 이미지들의 각각과 연관된 정보를 순차적으로 프로세싱하는 것을 필요로 할 수도 있다. 사용자가 정찬 메뉴에서 너무 멀리 떨어져서 이미지를 획득하는 카메라를 보유하고 있으면, 그 이미지는 이미지에 대해 OCR을 적절하게 수행할 충분한 해상도를 가지지 못할 수도 있다. 한편 사용자가 이미지를 획득하는 카메라를 너무 가깝게 가져가면, 그 이미지는 전체 관심 영역을 포함하지 못할 수도 있다. 그러나, 자동 및 지능적 OCR 캐시 업데이트를 갖는 OCR 캐싱 모드에서 본 발명의 양태들을 사용하여, 본원에서 설명되는 기법들은, 적절한 초점으로 양호한 해상도에서 메뉴로부터 사용자에게 대한 관심 영역을 포함하는 단일 이미지를 생성하는 것을 사용자에게 허용한다. 어떤 경우에는, 사용자는 메뉴로부터 텍스트를 선택하고 그것을 다른 언어로 번역하거나 사설 또는 온라인 검색 엔진을 사용하여 검색을 수행할 것을 선택할 수도 있다. 일부 실시형태들에서, 사용자는 애플리케이션 또는 프로그램, 예를 들어 텍스트 메시징 또는 이메일 프로그램 또는 워드 프로세서에 입력할 텍스트를 메뉴로부터 선택할 수도 있다.

[0052] 도 3a는 본 개시물의 하나 이상의 예시적 양태들에 따른 이미지들의 프로세싱을 위한 본 발명의 비제한적 예시적인 실시형태를 보여주는 흐름도를 도시한다. 하나 이상의 양태들에 따르면, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 예를 들면, 컴퓨팅 디바이스 (100) 및/또는, 도 1에서 더 상세히 설명된 디바이스와 같은 컴퓨팅 디바이스 (100) 에 의해 및/또는 그것에서 구현될 수도 있다. 하나의 실시형태에서, 도 3a를 참조하여 아래에서 설명되는 방법 단계들 중 하나 이상은 프로세서 (110) 또는 다른 프로세서와 같은 디바이스 (100) 의 프로세서에 의해 구현된다. 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들은 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 또한 구현될 수도 있고, 도 3a에서 논의되는 본 발명의 실시형태들을 수행함에 있어서 사용될 수도 있다. 부가적으로 또는 대안으로, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 컴퓨터 판독가능 명령들, 이를테면 메모리 (135), 스토리지 (125) 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체 상에 저장된 컴퓨터 판독가능 명령들로 구현될 수도 있다.



- [0053] 도 3a에서, 도 1을 참조하여 위에서 논의된 컴퓨팅 디바이스 (100) 와 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들은 본 발명의 실시형태들을 수행하는데 사용되는 디바이스의 모듈들 및 컴포넌트들의 일부를 나타낼 수 있다. 단계 302에서, 컴퓨팅 디바이스 (100) 에 연결된 카메라 (150) 는 복수의 이미지들을 수신하거나 또는 획득한다. 카메라 (150) 는 도 1에서 설명된 입력 디바이스들 (115) 중 하나를 나타낼 수도 있다. 하나의 구현예에서, 도 11로부터의 이미지 캡처 모듈 (1102) 은 이미지를 획득하는데 사용될 수도 있다. 복수의 이미지들은 일부 실시형태들에서 연속해서 수신될 수도 있다. 다른 실시형태들에서, 예를 들면 다수의 카메라들 (150) 또는 스테레오 카메라가 컴퓨팅 디바이스 (100) 에 연결되는 경우, 이미지들은 동시에 수신될 수도 있다. 단계 304에서, 수신된 이미지들 및 그 이미지들과 연관된 데이터는 하나를 초과하는 이미지가 동일한 장면과 연관되는지를 결정하기 위해 서로에 대해 그리고 또 OCR 캐시 (224) 에 저장된 OCR 결과들 및 이미지들에 대해 비교된다. 하나의 구현예에서, OCR 캐시는 도 1에서 설명된 작업 메모리 (135) 에서 동작한다.
- [0054] 블록 306에서, 동일한 장면과 연관된 하나를 초과하는 이미지들이 검출되면, 그 이미지들은 통합 프로세스를 통해 합성된 이미지를 형성하기 위해 함께 통합되거나 또는 합성될 수도 있다. 하나의 양태에서, 선택 프로세스는 통합 프로세스에 선행한다. 선택 프로세스 동안, 장면을 표현하는 키 이미지가 그 장면을 표현하는 복수의 이미지들로부터 선택될 수도 있으며, 그 중에서 복수의 이미지들의 하나 이상이 컴퓨팅 디바이스 (100) 상에 저장될 수도 있다. 키 이미지는 각각의 이미지와 연관된 에너지를 계산함으로써 결정될 수도 있다. 하나의 구현예에서, 각각의 이미지와 연관된 에너지는 OCR 엔진을 사용하여 인식된 단어들의 수 및 각각의 이미지에 대한 OCR 결과들과 연관된 신뢰도에 정비례한다. 하나의 실시형태에서, 키 이미지는 OCR 캐시에 저장된 이미지일 수도 있다. 다른 구현예들에서, 키 이미지는 이미지를 OCR 캐시에 저장하기 전에 동일한 장면으로부터의 다양한 이미지들로부터의 정보를 사용하여 최종 이미지를 합성함으로써 향상되거나 또는 증강될 수도 있다. 예를 들면, 키 이미지에서 디스플레이된 단어가 암호해독될 수 없거나 또는 검출될 수 없다면, 동일한 장면으로부터의 다른 이미지들로부터의 정보가 이미지 및 이미지와 연관된 데이터를 향상시키는데 사용될 수도 있다. 다른 예에서, 키 이미지의 지역으로부터의 글레어는, 동일한 장면으로부터, 키 이미지와 동일한 지역에 대한 글레어를 가지지 않는 다른 이미지로부터의 정보를 사용함으로써 제거될 수도 있다. 위에서 설명된 바와 같이, 글레어는 심지어 OCR의 부재 시에도 이미지들로부터 감소되거나 또는 제거될 수도 있다. 다양한 이미지들로부터의 정보는, OCR 캐시 또는 다른 캐시, 예를 들어 OCR 정보를 포함할 수도 있거나 또는 포함하지 않을 수도 있는 이미지 캐시에 저장될 수도 있는 최종 이미지를 합성할 시에 사용될 수도 있다. 예를 들어, 이미지 캐시가 감소된 글레어를 갖는 최종 이미지를 저장하는데 사용될 수도 있다. 이러한 이미지 캐시는 일부 실시형태들에서 OCR 정보를 생략할 수도 있거나 또는 OCR 정보를 포함할 수도 있다.
- [0055] 블록 308에서, 일단 최종 이미지가 컴퓨팅 디바이스 (100) 에 의해 합성되면, 그 이미지는 OCR 캐시에 저장될 수도 있다. 어떤 경우들에서는, OCR 캐시에서의 미리 존재하는 이미지 및 OCR 결과가 업데이트되고 대체된다. 예를 들어, 수신된 또는 획득된 이미지가 OCR 캐시에 저장된 것과 동일한 장면과 연관되지만 OCR 캐시에 저장된 이미지보다 양호한 해상도 및 선명도를 가진다면, OCR 캐시에서의 이미지 및 OCR 결과들은 새로이 수신된 이미지와 연관된 정보를 반영하기 위해 업데이트되거나 또는 대체될 수도 있다. 이미지를 합성하는 것은, 본원에서 설명된 바와 같이, 이미지 또는 그 이미지에 대한 장면 기술자의 부분들 또는 그 이미지의 표현을 포함하는 임의의 연관된 데이터를 OCR 캐시에서 업데이트하거나 또는 대체하는 것을 포함하지만 그것으로 제한되지 않을 수도 있다.
- [0056] 도 3a에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약술된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 3a에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.
- [0057] 도 3b는 본 개시물의 하나 이상의 예시적 양태들에 따른 이미지들의 프로세싱을 위한 본 발명의 다른 비제한적 예시적인 실시형태를 보여주는 흐름도를 도시한다. 하나 이상의 양태들에 따르면, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 예를 들면, 컴퓨팅 디바이스 (100) 및/또는, 도 1에서 더 상세히 설명된 디바이스와 같은 컴퓨팅 디바이스 (100) 에 의해 및/또는 그것에서 구현될 수도 있다. 하나

의 실시형태에서, 도 3b를 참조하여 아래에서 설명되는 방법 단계들 중 하나 이상은 프로세서 (110) 또는 다른 프로세서와 같은 디바이스 (100) 의 프로세서에 의해 구현된다. 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들은 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 또한 구현될 수도 있고, 도 3b에서 논의된 바와 같은 본 발명의 실시형태들을 수행함에 있어서 사용될 수도 있다. 부가적으로 또는 대안으로, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 컴퓨터 판독가능 명령들, 이를테면 메모리 (135), 스토리지 (125) 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체 상에 저장된 컴퓨터 판독가능 명령들로 구현될 수도 있다.

[0058] 도 3b에서, 도 1을 참조하여 위에서 논의된 컴퓨팅 디바이스 (100) 와 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들은 본 발명의 실시형태들을 수행하는데 사용되는 디바이스의 모듈들 및 컴포넌트들의 일부를 나타낼 수 있다. 단계 310에서, 컴퓨팅 디바이스 (100) 에 연결된 카메라 (150) 는 복수의 이미지들을 수신하거나 또는 획득한다. 카메라 (150) 는 도 1에서 설명된 입력 디바이스들 (115) 중 하나를 나타낼 수도 있다.

하나의 구현예에서, 도 11로부터의 이미지 캡처 모듈 (1102) 은 이미지를 획득하는데 사용될 수도 있다. 복수의 이미지들은 일부 실시형태들에서 연속해서 수신될 수도 있다. 다른 실시형태들에서, 예를 들면 다수의 카메라들 (150) 또는 스테레오 카메라가 컴퓨팅 디바이스 (100) 에 연결되는 경우, 이미지들은 동시에 수신될 수도 있다.

[0059] 단계 312에서, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 OCR 엔진 (1104) 또는/및 장면 기술자 모듈 (1106) 은, 복수의 이미지들의 각각에서 데이터를 인식할 수도 있으며, 복수의 이미지들의 각각에서 데이터를 인식하는 것은 복수의 이미지들에 대해 광학적 문자 인식을 수행하는 것을 포함한다. 데이터를 인식함에 있어서 OCR 엔진 (1104) 및 장면 기술자 (1106) 에 의해 수행된 단계들은 도 4 및 도 11에서 더 상세히 설명된다.

[0060] 단계 314에서, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 매칭 모듈 (1108) 은, 인식된 데이터에 적어도 부분적으로 기초하여, 복수의 이미지들 중 적어도 2 개의 이미지들이 제 1 장면으로부터의 정보를 포함한다고 결정할 수도 있다. 매칭 모듈 (1108) 에 의해 수행되는 단계들은 도 5, 도 6, 도 7, 도 8 및 도 11에서 더 상세히 설명된다.

[0061] 단계 316에서, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 통합 모듈 (1114) 은, 적어도 2 개의 이미지들 중 적어도 하나의 이미지 및 상기 적어도 하나의 이미지와 연관된 인식된 데이터를 사용하여 제 1 장면을 표현하는 최종 이미지를 저장할 수도 있다. 하나의 실시형태에서, 동일한 장면과 연관된 하나를 초과하는 이미지들이 검출되면, 그 이미지들은 통합 프로세스를 통해 합성된 이미지를 형성하기 위해 함께 통합되거나 또는 합성될 수도 있다. 통합 모듈 (1114) 에 의해 수행되는 단계들은 도 8, 도 9, 도 10 및 도 11에서 더 상세히 설명된다.

[0062] 도 3b에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약술된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 3b에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.

[0063] 컴퓨팅 디바이스 (100) 에 의해 수신된 이미지는 장면 기술자를 생성하기 위해 프로세싱될 수도 있다. 장면 기술자를 생성하는 구조 및 프로세스는 도 4 및 도 13에서 더 상세히 설명된다. 일단 장면 기술자가 수신된 이미지에 대해 생성되면, 수신된 이미지는, 그것이 컴퓨팅 디바이스 (100) 상에 저장된 이미지들에 대한 장면들 중 적어도 하나에 속하는지를 결정하기 위해 다른 이미지들에 대해 매칭되거나 또는 비교된다. 도 5 및 도 6은 매칭 프로세스를 더 상세히 설명한다. 매치가 발견되면, 키 이미지는 동일한 장면에 속한 다수의 이미지들로부터 각각의 이미지와 연관된 에너지를 사용하여 선택된다. 도 7 및 도 8은 키 이미지에 대한 선택 프로세스를 더 상세히 설명한다. 일단 키 이미지가 선택되면, 이미지들은 새로운 합성된 이미지를 형성하기 위해 업데이트된 장면 기술자와 함께 통합될 수도 있고 OCR 캐시는 업데이트될 수도 있다. 통합 프로세스는 도 9 및 도 10에서 더 상세히 설명된다. 도 11은 본 발명의 실시형태의 상이한 모듈들 및 컴포넌트들과 그들 컴포넌트들 사이의 연관된 흐름을 설명한다. 도 12는 부가적인 추적 컴포넌트와 함께 도 11을 추가로 설명한다. 도 14 및 도 15는 이미지를 캐시에 저장하는 경우 이미지들로부터의 글레이어를 감소시키거나 또는

제거하는 본 발명의 예시적인 실시형태들을 설명한다. 도 16a, 도 16b, 도 16c 및 도 16d는 예시적인 컴퓨팅 디바이스 (100) 에 대한 본원에서 설명되는 기법들을 사용하여 예시적인 OCR 캐시 업데이트 시나리오를 설명한다.

[0064] 도 4는 본 발명의 예시적인 실시형태에 따른 책 표지의 대표적인 이미지 및 연관된 장면 기술자를 보여주는 도면을 묘사한다. 본 발명의 양태들에서, 이미지 캡처 모듈 (1102) 은 하나 이상의 이미지들을 수신하고 OCR 엔진 (1104) 은 하나 이상의 이미지들에 대해 OCR을 수행한다. 장면 기술자 모듈 (1106) 은 이미지를 프로세싱할 수도 있고 장면 기술자는 이미지와 연관된 특성들을 표현하는 이미지에 대해 생성될 수도 있다. 장면 기술자는, 수신된 또는 저장된 이미지들 중 임의의 것이 동일한 장면에 속하는지를 결정함에 있어서 매칭 프로세스 동안 다양한 이미지들을 비교할 시에 매칭 모듈에 의해 사용될 수도 있다.

[0065] 장면 기술자는 그래픽 오브젝트들을 나타내는 이미지의 부분들의 측면에서 정의될 수도 있다. 그래픽 오브젝트들은 심볼들, 문자들, 단어들, 텍스트, 기호들, 및 숫자들을 포함할 수도 있지만 그것들로 제한되지 않는다. 도 4에 도시된 바와 같이, 하나의 구현예에서, 장면 기술자는 텍스트에 관한 데이터, 이를테면 하나 이상의 단어들 주위의 텍스트 컬러, 배경색 및 경계 박스를 카탈로그화한다. 컴퓨팅 디바이스 (100) 에 의해 검출된 그래픽 오브젝트에 대한 특성들은 단항 (unary) 특성들 및 이항 (binary) 특성들을 포함할 수도 있다. 단항 특성들은 그래픽 오브젝트 콘텐츠, 그래픽 오브젝트 컬러, 전경색, 배경색 및 그래픽 오브젝트를 포함하는 이미지의 부분들의 치수들을 포함할 수도 있지만 그것들로 제한되지 않는다. 그래픽 오브젝트들의 이항 특성들은 그래픽 오브젝트들의 서로 및 이미지에 관한 기하학적 관계를 포함할 수도 있다. 본 발명의 일부 양태들에서, 이항 특성들은 이미지에서의 그래픽 오브젝트들의 서로의 상이한 기하학적 관계들을 설명하는 친화도 매트릭스를 사용하여 표현될 수도 있다.

[0066] 다시 도 4의 블록 402를 참조하면, 장면 기술자 모듈 (1106) 은 단어들 "LEARNING from DATA", "Concepts, Theory, and Methods", 및 "SECOND EDITION"을 포함하는 이미지의 부분들을 식별할 수도 있다. 각각의 단어는 장면 기술자 내의 별도의 기술 엔티티, 이를테면 도 4에서의 "DATA" (406) 로서 취급될 수도 있다. 또한, 단어들은 집합적 기술 엔티티, 이를테면 "SECOND EDITION" (412) 을 형성하기 위해 함께 그룹화될 수도 있다.

[0067] 도 4에서, 텍스트 "DATA"를 포함하는 블록 406에 의해 표현된 부분은, 장면 기술자 (404) 에서 "Word 3"에 의해 표현된 별도의 기술 엔티티 (408) 로서 설명된다. 단어 기술 엔티티 (408) 에 대한 단항 특성들은 텍스트/콘텐츠 (즉, DATA), 텍스트 컬러 (즉, 백색), 배경색 (즉, 흑색) 및 경계 박스 또는/및 치수들을 표현하는 필드를 갖는 데이터 구조를 통해 표현될 수도 있다. 장면 기술자 모듈 (1106) 은 각각의 단어에 대한 정보를 편성하고 그것을 데이터 구조, 클래스 오브젝트 또는 임의의 다른 적합한 수단 내에 저장할 수도 있다. 마찬가지로, 이미지에 대한 장면 기술자와 연관된 단어들을 표현하는 구조들은, 스토리지 디바이스 (125) 또는 메모리 (135) 에 저장된 정보를 추출, 인덱싱 및/또는 업데이트하는 링크형 리스트들, 큐 (queue) 들 또는 임의의 다른 적합한 수단을 사용하여 편성될 수도 있다.

[0068] 단항 특성들에 더하여, 각각의 단어에 대한 이항 특성들이 이미지 (미도시) 에서의 각각의 단어에 대한 장면 기술 모듈 (1106) 에 의해 또한 저장될 수도 있다. 이항 특성들은 단어들의 서로에 관한 기하학적 관계를 포함하지만 그것으로 제한되지 않는다. 도 4에서, 단어 "LEARNING"에 대한 이항 특성들은 다른 단어들 및 이미지 자체에 관한 단어의 공간적 관계를 포함할 수도 있다. 예를 들면, "LEARNING"은 공간적으로는 좌측 및 상단으로부터의 첫 번째 단어이고, 그것은 우측의 단어 "from"과 그것 아래의 "Concepts"을 가진다. 상이한 장면들에 대한 이항 특성들을 비교하는 것은, 유사한 글자들을 갖는 유사한 단어들이 상이한 문장들을 생성하기 위해 상이한 장면들에서 사용되는 경우에 유익할 수도 있다. 하나의 예에서, 단어의 이항 특성들을 비교하는 것은, 컴퓨팅 디바이스로 하여금 단어들의 서로에 관한 상대 위치를 사용하여 매치들을 더 효율적으로 발견하는 것을 허용할 수도 있다.

[0069] 도 5는 컴퓨팅 디바이스 (100) 에 의해 수신된 복수의 이미지들과 다양한 장면들 사이의 매칭 프로세스를 예시하는 도면을 묘사한다. 도 1을 참조하여 위에서 논의된 컴퓨팅 디바이스 (100) 는, 도 5에서 설명된 본 발명의 실시형태들을 수행하는데 사용되는 디바이스의 컴포넌트들의 일부를 나타낼 수 있다. 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들, 예를 들어 매칭 모듈 (1108) 은 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 또한 구현될 수도 있고, 도 5에서 논의된 바와 같은 본 발명의 실시형태들을 수행할 시에 사용될 수도 있다. 이미지 (502) 는 컴퓨팅 디바이스 (100) 에 연결된 카메라 (150) 를 사용하여 수신된 이미지를 나타낸다. 하나의 실시형태에서, 이미지 캡처 모듈 (1102) 은 추가의 프로세싱을 위해 이미지를 캡처하고 그 이미지 데

이터를 OCR 엔진 (1104) 으로 전달한다. OCR 엔진 (1104) 은 이미지 (502) 를 프로세싱하여 OCR 결과들, 이를테면 이미지와 연관된 데이터를 생성한다. 도 4에서 논의된 바와 같이, 장면 기술자가 생성되고 이미지 (502) 와 연관된다. 이미지 (504) 및 이미지 (506) 는 이미지 (502) 보다 먼저 획득된 이미지들이고 연관된 장면 기술자들을 가진다. 컴퓨팅 디바이스 (100) 의 컴포넌트들은, 임시 버퍼 또는 OCR 캐시 중 어느 하나에 이미지들 (504 및 506) 을 저장할 수도 있다. 하나의 구현예에서, 임시 버퍼 또는 OCR 캐시 (1112) 는 도 1에서 설명된 작업 메모리 (135) 의 섹션들을 사용할 수도 있다.

[0070] 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 매칭 모듈 (1108) 은, 이미지 (502) 에 대한 장면 기술자에 대한 단항 및/또는 이항 특성들을, 이미지들 (504 및 506) 에 대한 장면 기술자들에 대한 단항 및/또는 이항 특성들에 대해 비교할 수도 있다. 예를 들면, 이미지 (502) 는 "LEARNING", "from" 및 "DATA"에 대한 단어 기술 엔티티를 가진다. 그러나, 이미지 (506) 는 "Learning"에 대한 단어 기술 엔티티만을 가진다. 심지어 이미지 (502) 로부터의 "LEARNING" 및 이미지 (506) 로부터의 "Learning" 사이에 긍정적 단항 매치가 있었다 하더라도, 2 개의 단어들에 대한 이항 특성들은 매치되지 않을 것이다. 이미지 (506) 에서, "Learning"은 좌측으로부터의 첫 번째 단어가 아니고 이미지 (502) 에서의 경우와 마찬가지로 그것의 우측에는 from을 가지지 않는다. 그러므로, 본 발명의 실시형태들은 이미지 (502) 및 이미지 (506) 가 동일한 장면에 속하지 않는다고 결론을 내릴 수도 있다.

[0071] 그 반면, 이미지 (502) 및 이미지 (504) 에 대한 장면 기술자와 연관된 단항 및 이항 특성들은 긍정적 매치를 형성할 수도 있다. 양쪽 모두의 이미지들은 상이한 해상도들을 설명하는 팩터 (factor) 에 의해 스케일링된 이미지에서의 각각의 단어에 대한 동일한 단항 특성들을 가진다. 또한, 단어들 사이의 기하학적 관계는 양쪽 모두의 이미지들에서 동일하다. 이미지 (502) 및 이미지 (504) 사이의 유일한 중요한 차이는, 이미지 (504) 가 이미지의 부분으로서 "SECOND EDITION"을 가지지 않는다는 것이다. 본 발명의 양태들은 동일한 장면으로부터의 이미지들 사이의 이러한 차이들을 무시하는 메커니즘들을 구현할 수도 있다.

[0072] 하나의 구현예에서, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 매칭 모듈 (1108) 은 장면 기술자에 대한 임의의 2 개의 주어진 특성들 사이의 각각의 유사도 및 차이에 가중치를 할당할 수도 있다. 할당된 가중치들은 유사도들에 대해서는 가산되고 차이들에 대해서는 감산될 수도 있다. 임계치가 생성되거나 또는 미리 결정될 수도 있다. 2 개의 이미지들에 대한 장면 기술자들 사이의 유사도들 및 차이들의 가중된 합이 임계치를 초과하면, 매칭 모듈 (1108) 은 2 개의 주어진 이미지들을 동일한 장면과 연관시킬 수도 있다. 2 개의 이미지들에 대한 장면 기술자들 사이의 유사도들 및 차이들의 가중된 합이 임계치 미만이면, 그 이미지들은 상이한 장면들을 표현하는 것으로서 간주될 수도 있다.

[0073] 도 6은 획득된 이미지와 저장된 이미지에 대한 장면을 연관시키는 비제한적 예시적인 매칭 프로세스를 보여주는 흐름도를 도시한다. 단계 602에서, 컴퓨팅 디바이스 (100) 의 컴포넌트들, 이를테면 매칭 모듈 (1108) 은, 수신된 이미지에 대해 OCR을 수행한다. 도 1을 참조하여 위에서 논의된 컴퓨터 디바이스 (100) 는, 도 6에서 설명된 본 발명의 실시형태들을 수행하는데 사용되는 모바일 디바이스의 컴포넌트들의 일부를 나타낼 수 있다. 이미지는 컴퓨팅 디바이스 (100) 에 연결된 카메라 (150) 를 사용하여 획득된다. 단계 604에서, 장면 기술자는, 예를 들어 장면 기술자 생성기 (1106) 를 사용하여 수신된 이미지에 대해 생성될 수도 있다. 도 4를 참조하여 설명된 바와 같이, 그래픽 오브젝트들, 이를테면 텍스트를 포함하는 이미지의 부분들은, 인식되고 장면 기술자에서 추가로 특징화될 수도 있다. 각각의 그래픽 오브젝트에 대해, 단항 및 이항 특성들 양쪽 모두가 정의될 수도 있다. 단항 특성들은 그래픽 오브젝트 콘텐츠, 그래픽 오브젝트 컬러, 전경색, 배경색 및 그래픽 오브젝트를 포함하는 이미지의 부분들의 치수들을 포함할 수도 있지만 그것들로 제한되지 않는다. 그래픽 오브젝트들의 이항 특성들은 그래픽 오브젝트들의 서로 및 이미지에 관한 기하학적 관계를 포함할 수도 있다. 각각의 그래픽 오브젝트에 대한 정의된 특성들은 데이터 구조, 클래스 오브젝트에 또는 임의의 다른 적합한 수단을 사용하여 저장될 수도 있다.

[0074] 단계 610에서, 수신된 이미지의 장면 기술자는 컴퓨팅 디바이스 (100) 상에서 이전에 획득되고 저장되거나 또는 버퍼링된 이미지의 장면 기술자와, 예를 들어 매칭 모듈 (1108) 에 의해 비교될 수도 있다. 이전에 획득된 이미지 및 그것의 연관된 장면 기술자는 스토리지 (125), 메모리 (135) 또는 하드웨어 레지스터들을 사용하여 임시 버퍼, OCR 캐시 또는 임의의 다른 적합한 저장 매체에 저장될 수도 있다. 장면 기술자들과 비교하는 동안 (단계 610), 단계 606에서, 수신된 이미지에 대한 장면 기술자의 단항 특성들은 저장된 이미지와 연관된 장면 기술자의 단항 특성들에 대해 비교될 수도 있다. 단계 608에서, 수신된 이미지에 대한 장면 기술자의 이항 특성들은 저장된 이미지와 연관된 장면 기술자의 이항 특성들에 대해 비교될 수도 있다.



- [0075] 하나의 구현예에서, 장면 기술자의 임의의 2 개의 주어진 특성들 사이의 각각의 유사도는 그것과 연관된 할당된 가중치를 가질 수도 있다. 할당된 가중치들은 유사도들에 대해서는 반영되고 차이들에 대해서는 감산될 수도 있다. 임계치가 생성되거나 또는 미리 결정될 수도 있다. 단계 612에서, 2 개의 이미지들에 대한 장면 기술자들에서의 유사도들의 가중된 합이 임계치를 초과하면, 컴퓨팅 디바이스 (100)의 컴포넌트들은 그 이미지들과 동일한 장면을, 예를 들어, 선택 모듈 (1110) 및/또는 통합 모듈 (1114)을 사용하여 연관시킬 수도 있다 (블록 616). 한편, 2 개의 이미지들에 대한 장면 기술자들에서의 유사도들의 가중된 합이 임계치 미만이면, 그 이미지들은 상이한 장면들을 표현하는 것으로서 간주될 수도 있다 (블록 614). 방법 (600)에서의 부분적 흐름도들은 모든 이전에 저장된 이미지들의 장면 기술자와 임시 메모리, OCR 캐시 (1112) 또는 임의의 다른 적합한 저장 매체로부터의 그것들의 연관된 장면 기술자들을 비교하기 위해 되풀이하여 또는 순환적으로 (recursively) 반복될 수도 있다.
- [0076] 도 6에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약속된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 6에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.
- [0077] 도 7은 동일한 장면내에 속한 복수의 이미지들로부터의 키 이미지로서의 이미지의 선택 프로세스를 보여주는 블록도를 묘사한다. 도 4, 도 5 및 도 6에 설명된 바와 같이, 이미지가 컴퓨팅 디바이스 (100)에 의해 수신된 후, 컴퓨팅 디바이스 (100)는 이미지에 대한 장면 기술자를 생성하고 그 이미지를 이전에 획득되고 저장된 다른 이미지들에 대해 매치시킬 수도 있다. 하나의 구현예에서, 컴퓨팅 디바이스는 동일한 장면내에 속한 복수의 이미지들로부터 키 이미지를 선택한다. 복수의 이미지들로부터 키 이미지를 선택하는 프로세스는 선택 프로세스라고 지칭될 수도 있고, 컴퓨팅 디바이스 (100)의 컴포넌트들, 이를테면 도 11 및 도 12로부터의 선택 모듈 (1110)에 의해 수행될 수도 있다. 도 7을 참조하면, 그룹 I은 제 1 장면과 연관된 3 개의 이미지들을 가지고 그룹 II는 제 2 장면과 연관된 3 개의 이미지들을 가진다. 최종 키 이미지는 그룹 I 및 그룹 II에 대한 맨 오른쪽 열에서 선택되고 디스플레이된다. 키-이미지는 장면을 표현하는, OCR 캐시 (1112)에 저장된 후보 이미지일 수도 있다. 어떤 경우들에서는, 키-이미지 또는 키-이미지에 관련된 정보는 키-이미지를 저장하기 전에 다른 이미지들로부터의 정보를 사용하여 향상될 수도 있다. 예를 들어, 키 이미지가 아닌 하나 이상의 이미지들에서 인식되는 단어들은 키 이미지와 연관된 데이터에 추가될 수도 있다. 아래에서 설명되는 메커니즘들은 이미지들의 각각의 그룹으로부터 키 이미지를 찾을 시에 추구될 수도 있다.
- [0078] 키 이미지는 동일한 장면과 연관된 다른 이미지들보다 양호한 해상도 및 선명도를 갖는 이미지를 나타낼 수도 있다. 하나의 양태에서, 컴퓨팅 디바이스 (100)의 컴포넌트들은 각각의 이미지와 연관된 에너지를 계산할 수도 있다. 에너지 값은 이미지에 대한 장면 기술자의 부분으로서 저장될 수도 있다. 최고 에너지를 갖는 이미지는 동일한 장면과 연관된 복수의 이미지들로부터 키 이미지로서 선택될 수도 있다. 하나의 구현예에서, 컴퓨팅 디바이스 (100)는 (예를 들어, 선택 모듈 (1110)을 사용하여) 이미지로부터의 인식된 단어들의 수 및 OCR로부터의 단어들의 신뢰도 평균을 사용하여 에너지 값을 결정할 수도 있다.
- [0079] 도 8은 동일한 장면으로부터의 복수의 이미지들로부터의 키 이미지로서의 이미지의 선택 프로세스에 대한 비제한적 예시적인 방법을 보여주는 흐름도를 도시한다. 하나 이상의 양태들에 따르면, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는, 예를 들면, 컴퓨팅 디바이스 (100) 및/또는, 도 1에서 더 상세히 설명된 디바이스와 같은 컴퓨팅 디바이스 (100)에 의해 및/또는 그것에서 구현될 수도 있다. 하나의 실시형태에서, 도 8을 참조하여 아래에서 설명되는 방법 단계들 중 하나 이상은 프로세서 (110) 또는 다른 프로세서와 같은 디바이스 (100)의 프로세서에 의해 구현된다. 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들, 예를 들어 선택 모듈 (1110)은 컴퓨팅 디바이스 (100)의 컴포넌트들로서 또한 구현될 수도 있고, 도 8에서 논의되는 바와 같은 본 발명의 실시형태들을 수행함에 있어서 사용될 수도 있다. 부가적으로 또는 대안으로, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 컴퓨터 판독가능 명령들, 이를테면 메모리 (135), 스토리지 (125) 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체에 저장된 컴퓨터 판독가능 명령들로 구현될 수도 있다.
- [0080] 단계 802에서, 하나 이상 (M)의 이미지들이 도 5 및 도 6에서 논의된 바와 동일한 장면내에 속하는 것으로서 검

출된다. 일부 구현예들에서, OCR 캐시 (1112) 는 임의의 주어진 장면과 연관된 하나의 이미지만을 가진다.

OCR 캐시 (1112) 에 저장된 것과 동일한 장면과 연관된 임의의 획득된 이미지는, 즉시 프로세싱되고 OCR 캐시 (1112) 속으로 통합될 수도 있다. 새로 획득된 이미지가 키 이미지로서 선택되거나 또는 부가적인 정보를 가진다면 OCR 이미지는 업데이트될 수도 있다. 이러한 구현예에서, 선택 프로세스는 2 개의 이미지들, 즉, OCR 캐시에 저장된 이미지 및 카메라에 의해 최근에 획득된 이미지 사이에만 있을 수도 있다.

[0081] 그러나, 본 발명의 다른 구현예들에서, 다수의 이미지들이 키 이미지를 찾기 위해 동시에 비교될 수도 있다. 예를 들면, 하나의 구현예에서, 하나를 초과하는 이미지가 기간에 걸쳐 획득되고 동시에 프로세싱될 수도 있는 윈도우 프로토콜이 구현될 수도 있다. 획득되고 임시로 저장된 다수의 이미지들로부터, 일반적으로 하나를 초과하는 이미지들이 동일한 장면에 속하는데, 사용자가 동일한 장면에 대해 짧은 시간 기간에 걸쳐 다수 회 모바일 디바이스를 글라이딩 (gliding) 할 수도 있기 때문이다. 그러므로, 동일한 장면과 연관된 시간의 임의의 주어진 윈도우로부터의 하나 이상의 이미지들은, 서로에 대해 또는 상기 동일한 장면과 연관된 OCR 캐시 (1112) 에 저장된 이미지에 대해 비교될 수도 있다.

[0082] M은 임의의 주어진 시각점 (point in time) 에서 특정 장면과 연관된 이미지들의 수를 나타낸다. 각각의 이미지와 연관된 에너지는 반복적으로 또는 순환적으로 계산될 수도 있다. 반복 횟수는 X에 의해 표현된다. 하나의 예시적인 구현예에서, X는 1로 초기화될 수도 있다. X는 또한 현재 프로세싱되는 이미지를 나타낸다. 하나의 양태에서, 이미지의 에너지는 이미지의 장면 기술자에 저장된다.

[0083] 단계 804에서, 이미지 X에서의 단어들의 수 (N) 가 OCR을 사용하여 인식된다. 단계 806에서, OCR로부터의 단어들의 신뢰도 평균 (C) 이 도출된다. 단계 808에서, 인식된 단어와 연관된 신뢰도가 스케일 팩터 ( $\lambda$ ) 에 의해 스케일링된다. 하나의 실시형태에서, 스케일 팩터 ( $\lambda$ ) 은 경험적으로 결정될 수도 있다. 단계 810에서, 이미지 X에 대한 에너지가 다음의 수식에 따라 계산된다;

[0084]  $EX = N + \lambda C.$

[0085] 단계 812에서, 이미지에 대해 도출된 에너지는 선택적으로 이미지의 장면 기술자에 저장될 수도 있다. 단계 814에서,  $X=M$ 이면, 최대 에너지 값을 갖는 이미지는 동일한 장면과 연관된 복수의 이미지들로부터의 키 이미지로서 선택된다 (단계 816). X가 M과 동일하지 않으면, X는 증가되고 (단계 818) 후속의 이미지 X가 폐치되고 단계 804으로 되돌아가서 시작하는 방법 (800) 을 통해 프로세싱될 수도 있다.

[0086] 다른 실시형태에서, 이미지들에 대한 에너지의 반복적/순환적 계산이 사용되지 않는다. 이미지가 획득된 후 또는 그 이미지가 이미지와 연관된 에너지에 영향을 미칠 수도 있는 부가적인 정보로 업데이트되면, 그 이미지에 대해 그 이미지의 에너지가 계산될 수도 있다. 하나의 구현예에서, 이미지의 에너지는 이미지에 대한 장면 기술자에 저장된다.

[0087] 도 8에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약술된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 8에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.

[0088] 도 9는 동일한 장면으로부터의 복수의 이미지들로부터의 이미지의 OCR 캐시 (1112) 에 저장된 최종 이미지로서의 합성을 보여주는 도면을 묘사한다. 하나의 실시형태에서, 이미지들 (902 및 904) 은 OCR 엔진 (1104) 을 통해 라우팅된 다음 장면 기술자 생성기 (1106) 는 이미지와 연관된 그래픽 오브젝트들을 식별하고 장면 기술자들을 각각 생성한다. 매칭 모듈 (1108) 은, 이미지들 (902 및 904) 과 동일한 장면을 연관시킨다. 하나의 실시형태에서, 매칭 프로세스 후에 이미지들 및 그것들의 연관된 장면 기술자들은 키 이미지를 결정하기 위해 선택 모듈 (1110) 에 의해 선택 프로세스를 통해 라우팅된다. 예시 목적을 위해, 이미지 (902) 는 선택 프로세스로부터의 키 이미지로서 가상적으로 선택된다.

[0089] 선택 프로세스 후, 이미지들은 통합 프로세스를 위해 통합 모듈 (1114) 을 통해 라우팅될 수도 있으며, 그 통합 모듈에서 2 개의 이미지들은 단일 합성 이미지 (906) 로 통합될 수도 있다. 도 9에서, 이미지 (902) 는 OCR 캐시에 저장될 수도 있는 이미지로서 선택된다. 그러나, 이미지 (904) 는 이미지 (902) 에서 묘사되지 않은

부가적인 고유 정보를 가질 수도 있다. 이미지 (904) 에 고유한 부가적인 정보는 이미지 (902) 와 합성되어 최종 이미지 (906) 를 형성할 수도 있다. 하나의 실시형태에서, 이미지 (906) 는 모든 OCR 결과들이 이미지 (906) 에서 보여질 수 있도록 이미지 (902) 및 이미지 (904) 로부터 새로 생성될 수도 있다.

[0090] 장면 기술자는 또한 업데이트될 수도 있다. 예를 들면, 이미지 (906) 에 대한 장면 기술자는 902에 대한 장면 기술자를 위해 정의된 특성들을 가질 것이다. 그러나, 부가적인 고유한 단어가 이미지 (904) 에 추가되므로, 이미지 (906) 에 대한 장면 기술자는 이미지 (904) 로부터의 "Methods"에 대해 정의된 특성들을 상속할 수도 있다.

[0091] 도 10은 동일한 장면으로부터의 복수의 이미지들로부터 이미지를 합성하는 비제한적 예시적인 방법을 보여주는 흐름도를 도시한다. 하나 이상의 양태들에 따르면, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 예를 들면, 컴퓨팅 디바이스 (100) 및/또는, 도 1에서 더 상세히 설명된 디바이스와 같은 컴퓨팅 디바이스 (100) 에 의해 및/또는 그것에서 구현될 수도 있다. 하나의 실시형태에서, 도 10을 참조하여 아래에서 설명되는 방법 단계들 중 하나 이상은 프로세서 (110) 또는 다른 프로세서와 같은 디바이스 (100) 의 프로세서에 의해 구현된다. 도 11 및 도 12에서 논의되는 모듈들 및 컴포넌트들, 예를 들어 통합 모듈 (1114) 은 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 또한 구현될 수도 있고, 도 8에서 논의되는 바와 같은 본 발명의 실시형태들을 수행할 시에 사용될 수도 있다. 부가적으로 또는 대안으로, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 컴퓨터 판독가능 명령들, 이를테면 메모리 (135), 스토리지 (125) 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체 상에 저장된 컴퓨터 판독가능 명령들로 구현될 수도 있다.

[0092] 이미지를 합성하는 것은, 본원에서 설명된 바와 같이, 이미지 또는 그 이미지에 대한 장면 기술자 또는 그 이미지의 표현을 포함하는 임의의 연관된 데이터를 OCR 캐시에서 업데이트하거나 또는 대체하는 것을 포함하지만 그것으로 제한되지 않을 수도 있다. 예를 들면, OCR 캐시에 저장된 식당으로부터의 정찬 메뉴의 이미지는, 카메라가 미리보기 모드 또는 스냅샷 모드에서 정찬 메뉴 위를 선회하고 더 많은 이미지들을 수신하므로, 수신된 이미지들로부터의 부가적인 정보로 증강될 수도 있다. 단계 1002에서, 키 이미지가, 도 7 및 도 8에서 논의된 바와 같이, 선택 모듈 (1110) 을 사용하여 동일한 장면에 속한 복수의 이미지들로부터 선택된다. 단계 1004에서, 이미 키 이미지의 부분이 아닌, 본 발명의 컴포넌트들, 예를 들어 장면 기술자 생성기 (1106) 에 의해 식별된 임의의 고유 그래픽 오브젝트들이 키 이미지에 추가될 수도 있다. 일부 경우들에서 수신된 이미지는 부가적인 정보, 이를테면 OCR 캐시에 이미 저장되지 않은 장면의 지역을 가질 수도 있다. 반대로, 수신된 이미지는 OCR 캐시 (1112) 에서의 저장된 이미지에 포함된 모든 정보를 가지지 않을 수도 있다. 어떤 경우에는, 일단 키 이미지가 선택되면, 이미 키 이미지의 부분이 아닌 다른 이미지들로부터의 임의의 그래픽 오브젝트들이 그 이미지에 추가된다.

[0093] 단계 1006에서, 선택적으로, 합성된 이미지에 대한 에너지는 그 이미지에 추가된 임의의 새로운 정보를 설명하기 위해 또한 업데이트될 수도 있다. 하나의 실시형태에서, 컴퓨팅 디바이스 (100) 는 도 8의 810에서 논의된 바와 같은 합성된 이미지에 대한 에너지를 계산한다. 단계 1008에서, 장면 기술자는 부가적인 정보로 또한 업데이트될 수도 있다. 예를 들면, 그래픽 오브젝트들과 연관된 부가적인 이미지 데이터 및 장면 기술자 데이터는, OCR 캐시 (1112) 에 합성된 이미지를 저장하기 전에 그 합성된 이미지를 형성하는 키 이미지에 대한 데이터와 통합될 수도 있다. 위에서 논의된 동일한 프로세스가 임의의 둘 이상의 이미지들, 이를테면 다수의 수신된 이미지들 및 OCR 캐시 (1112) 에 저장된 이미지들에 대해 동시에 일어날 수도 있다.

[0094] 도 10에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약술된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 10에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.

[0095] 도 11은 본 발명의 실시형태들에 의해 제공된 방법들을 수행하는 예시적인 컴포넌트들 및 모듈들을 보여주는 블록도를 묘사한다. 도 1을 참조하여 위에서 논의된 컴퓨터 디바이스 (100) 는, 도 11에서 설명된 본 발명의 실시형태들을 수행하는데 사용되는 모바일 디바이스의 컴포넌트들의 일부를 나타낼 수 있다. 도 11에서 논

의된 컴포넌트들 및 모듈들은 하드웨어, 소프트웨어, 펌웨어, 또는 그것들의 임의의 조합으로 구현될 수도 있다.

[0096] 하나의 실시형태에서, 컴퓨팅 디바이스 (100) 는 자동 및 지능적 OCR 캐시들을 사용하는 OCR 캐싱 모드로 스위칭될 수도 있다. OCR 캐시 (1112) 는 도 1에 설명된 작업 메모리 (135) 의 부분들을 이용할 수도 있다.

하나의 실시형태에서, 자동 및 지능적 OCR 캐시 업데이트를 갖는 미리보기 모드는 프로그램 또는 애플리케이션, 이를테면 카메라 키보드가 온 되는 경우 트리거되고 카메라 키보드가 턴 오프되는 경우 턴 오프된다.

[0097] 이미지 캡처 모듈 (1102) 은 하나 이상의 이미지들을 획득, 캡처 또는 수신할 수도 있다. 하나의 실시형태에서, 이미지 캡처 모듈 (1102) 은 하나 이상의 이미지들을 획득하기 위해 컴퓨팅 디바이스에 연결된 하나 이상의 카메라들 (150) 을 사용한다. OCR 엔진 (1104) 은 이미지들, (예를 들면, 프로세서 (110) 를 사용하여) 프로세싱하고, 이미지에 내장된 그래픽 오브젝트들을 포함하는 이미지의 부분들을 인식한다. 그래픽 오브젝트들은 심볼들, 문자들, 단어들, 기호들, 및 숫자들을 포함할 수도 있지만 그것들로 제한되지 않는다. 덧붙여, OCR 엔진은 이미지를 프로세싱하고 각각의 이미지에 대한 장면 기술자를 채우기 시작할 수도 있다.

[0098] 장면 기술자 생성기 (1106) 는 OCR 캐시 시스템 (1124) 의 부분으로서 구현될 수도 있고, 장면 기술자를 생성하고 채울 수도 있다. 장면 기술자는 이미지와 연관된 그래픽 오브젝트들의 특성들을 설명한다. 장면 기술자는 다수의 그래픽 오브젝트들을 나타내는 이미지의 부분들의 측면에서 정의될 수도 있다. 그래픽 오브젝트에 대한 특성들은 단항 특성들 및 이항 특성들을 포함할 수도 있다. 단항 특성들은 그래픽 오브젝트 콘텐츠, 그래픽 오브젝트 컬러, 전경색, 배경색, OCR 결과와 연관된 신뢰도, 및 그래픽 오브젝트를 포함하는 이미지의 부분들의 치수들 및 경계를 포함할 수도 있지만 그것들로 제한되지 않는다. 그래픽 오브젝트들의 이항 특성들은 그래픽 오브젝트들의 서로 및 이미지에 관한 기하학적 관계를 포함할 수도 있다. 본 발명의 일부 양태들에서, 이항 특성들은 이미지에서의 그래픽 오브젝트들의 서로의 상이한 기하학적 관계들을 설명하는 친화도 매트릭스를 사용하여 표현될 수도 있다. 도 4 및 도 6은 장면 기술자 생성기 (1106) 에 의해 수행되는 단계들의 일부를 더 상세히 논의한다.

[0099] 매칭 모듈 (1108) 은 하나 이상의 수신된 이미지들 및 OCR 캐시 (1112) 로부터의 저장된 이미지들의 장면 기술자들을 매칭시킬 수도 있다. 임의의 2 개의 이미지들의 장면 기술자들 사이의 매치는, 2 개의 이미지들이 동일한 장면과 연관된다는 것을 나타낸다. 이미지들 사이의 장면 기술자들에서의 유사도들 및 차이들은, 동일한 장면에 합리적으로 속하기 위해 서로 밀접하게 충분히 정렬되는 이미지들을 찾기 위해 활용될 수도 있다. 도 5 및 도 6은 매치를 찾기 위한 예시적인 실시형태들을 더 상세히 설명한다.

[0100] 매칭 모듈 (1108) 이 이미지 캡처 모듈 (1102) 로부터의 수신된 이미지 및 OCR 캐시 (1112) 에 저장된 이미지들 사이의 매치를 발견하지 못하면, 그 이미지는 OCR 캐시 (1112) 에 저장되기 위해 삽입 모듈 (1116) 로 전해진다. OCR 캐시 (1112) 에 저장된 새로운 이미지는 컴퓨팅 디바이스 (100) 에 의해 획득된 새로운 장면을 나타낸다.

[0101] 매칭 모듈 (1108) 이 수신된 이미지 및 OCR 캐시 (1112) 로부터의 이미지 사이에 매치를 발견하면, 매칭된 이미지들 (또는 그 이미지들에 대한 포인터들) 은 키 이미지를 선택하기 위해 선택 모듈 (1110) 로 포워딩된다. 블록 1114에서, 수신된 이미지 및 OCR 캐시 (1112) 로부터의 매칭된 이미지와 연관된 에너지를 사용하는 선택 프로세스를 사용하여 2 개의 이미지들로부터의 키 이미지가 결정된다. 하나의 실시형태에서, 에너지는 이미지에 대한 장면 기술자에 저장된다. 다른 실시형태에서, 각각의 이미지와 연관된 에너지는 선택 모듈 (1110) 에서 도출된다. 하나의 구현예에서, 이미지의 에너지는 이미지에 대해 인식된 단어들의 수 및 그것과 연관된 신뢰도에 정비례하게 정의된다. 다음의 수식은 에너지를 도출하는 구현예를 나타낼 수도 있다,

[0102]  $E = N + \lambda C$ , 여기서

[0103] N은 이미지에서 인식되는 단어들의 수이며, C는 OCR로부터의 단어들의 신뢰도이고  $\lambda$  는 스케일 팩터이다. 하나의 양태에서,  $\lambda$  는 경험적 데이터를 사용하여 도출된다. 이미지가 수신된 후 또는 그 이미지가 이미지와 연관된 에너지에 영향을 미칠 수도 있는 부가적인 정보로 업데이트되면, 그 이미지에 대해 그 이미지의 에너지가 계산될 수도 있다.

[0104] 통합 모듈 (1114) 은 합성된 이미지를 형성하기 위해 키 아닌 이미지들로부터의 장면 기술자 및 다른 이미지 데이터를 키 이미지와 통합할 수도 있다. 이미지를 합성하는 것은, 본원에서 설명된 바와 같이, 이미지 또는 그 이미지에 대한 장면 기술자 또는 그 이미지의 표현을 포함하는 임의의 연관된 데이터를 OCR 캐시에서 업데이트



트하거나 또는 대체하는 것을 포함하지만 그것으로 제한되지 않을 수도 있다. 어떤 경우에는, OCR 캐시 (1112) 에서의 저장된 이미지와 수신된 이미지는 동일한 그래픽 오브젝트들에 대해 2 개의 이미지들 간에 동일한 장면 기술자 정보를 갖는 중복 지역들을 가질 수도 있다. 그러나, 어떤 경우들에서 수신된 이미지는 부가적인 정보, 이를테면 OCR 캐시에 이미 저장되지 않은 장면의 지역을 가질 수도 있다. 반대로, 수신된 이미지는 OCR 캐시 (1112) 에서의 저장된 이미지에 포함된 모든 정보를 가지지 않을 수도 있다. 어떤 경우에는, 일단 키 이미지가 선택되면, 이미 키 이미지의 부분이 아닌 다른 이미지들로부터의 임의의 그래픽 오브젝트들이 그 이미지에 추가된다. 도 10은 통합 모듈 (1114) 에 의해 수행되는 단계들의 일부를 더 상세히 논의한다.

[0105] 하나의 예에서, OCR 캐시에 저장된 식당으로부터의 정찬 메뉴의 이미지는, 카메라가 미리보기 모드 또는 스냅샷 모드에서 정찬 메뉴 위를 선회하고 더 많은 이미지들을 수신하므로, 수신된 이미지들로부터의 부가적인 정보로 증강될 수도 있다. 다른 예에서, 통합 모듈은 OCR 캐시 (1112) 와 같은 캐시에 합성된 이미지를 저장하기 전에 이미지로부터 글레이어를 제거하거나 또는 감소시킬 수도 있다. 통합 프로세스는, 키 이미지로부터 글레이어를 갖는 제 1 지역을 식별하며; 제 1 지역 및 제 2 지역이 장면의 동일한 지역을 나타내는, 키 아닌 이미지로부터 글레이어가 없는 제 2 지역을 식별하며; 및 글레이어를 제거하기 위해 키 이미지 및 키 아닌 이미지로부터의 글레이어가 없는 제 2 지역을 사용하여 최종 이미지를 합성함으로써 글레이어를 제거하는 것을 용이하게 할 수도 있다.

[0106] 하나의 실시형태에서, 자동 및 지능적 OCR 캐시 업데이트를 갖는 OCR 캐싱 모드는, 프로그램 또는 애플리케이션, 이를테면 카메라 키보드가 턴 온 되는 경우에 동작가능하고, 이 모드는, 카메라 키보드가 턴 오프되거나 또는 그렇지 않고 일시 중단 또는 보류되는 경우 비활성화된다. 일부 실시형태들에서, 카메라 키보드 카메라에 의해 캡처된 정보, 예를 들어 이미지에서 인식된 또는 장면의 키 이미지와 연관된 텍스트가, 애플리케이션에 입력되는 것을 허용할 수도 있다. 더욱이, 컴퓨팅 디바이스 (100) 에 연결된 디바이스 디스플레이는, OCR 캐싱 모드 (일부 실시형태들에서 자동 및 지능적 OCR 캐시 업데이트를 가짐) 가 동작가능한 (즉, 카메라 키보드가 턴 온되거나 또는 프로그램 또는 애플리케이션이 재개되거나 또는 그렇지 않으면 동작가능한) 경우, OCR 캐시 (1112) 의 콘텐츠들을 또한 지속적으로 디스플레이할 수도 있다. 또한, 카메라 키보드가 중지되는 경우, OCR 캐시 (1112) 에 대한 캐시 데이터는 메모리로부터 디바이스 스토리지 (1120) 에 저장될 수도 있다. 디바이스 스토리지 (1120) 는 도 1에서 설명된 스토리지 디바이스들 (125) 중 하나일 수도 있다. 그리고 카메라 키보드가 시작되는 경우, OCR 캐시 (1112) 에 대한 캐시 데이터는 디바이스 스토리지 (1120) 로부터 메모리에 로딩될 수도 있다. 하나의 구현예에서, OCR 캐시 (1112) 는 이미지들의 큐를 사용하여 구현된다. 그러나, 링크형 리스트, 클래스 오브젝트, 또는 이미지들의 수 또는 데이터의 양을 포함하는 임의의 다른 적합한 수단이 OCR 캐시 (1112) 를 구현하는데 사용될 수도 있다.

[0107] 도 12는 본 발명의 실시형태들에 의해 제공된 방법들을 수행하는 예시적인 컴포넌트들을 보여주는 다른 블록도를 묘사한다. 도 11에서 설명된 컴포넌트들에 더하여, 도 12는 OCR 캐싱 모드에서 이미지 캡처 모듈 (1102) 로부터 장면을 지속적으로 추적하고 입력을 블록 1108에서의 매칭 로직에 피드하는 추적 모듈 (1222) 을 가진다. 장면을 추적하는 것은, 장면과 연관된 양호한 또는 더 큰 이미지를 생성하기 위해 사용자가 동일한 장면 위에서 모바일 디바이스를 지속적으로 글라이딩하는 경우, 매칭 로직의 지원 시에 특히 도움이 된다. 이 시나리오에서, 매칭 로직에 대한 복잡도와 프로세싱 시간 및 전력은 획득되는 모든 새로운 이미지에 대해 상당히 감소된다. 하나의 실시형태에서, 그래픽 오브젝트들 (및 그래픽 오브젝트들을 만들기 위해 그래픽 오브젝트들 주위에서 사용되는 연관된 박스들) 의 움직임은 매칭 문제의 해결 시에 사용될 수 있다. 도 11로부터 정찬 메뉴 예를 다시 참조하면, 추적 모듈은, 컴퓨팅 디바이스 (100) 가 정찬 메뉴 위에서 선회할 때에 이미지를 끊임없이 추적할 수 있고, 컴퓨팅 디바이스 (100) 가 정찬 메뉴의 단일의 통합된 이미지를 구축하는 것을 도울 수 있다. 일부 실시형태들에서, 모듈들 (1104, 1106, 1108, 1110, 1112, 1114, 1116, 1124, 및 1222) 중 하나 이상은 디바이스 (100) 의 하나 이상의 프로세서들 또는 프로세싱 코어들에서, 예를 들어 프로세서 (110) 에서 구현된다.

[0108] 도 13a 및 도 13b는 텍스트를 포함하는 이미지에 대한 예시적인 장면 기술자들을 묘사한다. 도 13a는 장면 기술자 생성기 (1106) 에 의해 생성된 장면 기술자이다. 장면 기술자는 이미지 해상도, 컬러 깊이 및 다른 이러한 이미지 특성들과 같은 이미지 데이터를 포함한다. 덧붙여, 장면 기술자는, 텍스트 콘텐츠, 관심 지역 (ROI), OCR을 사용한 단어의 적절한 인식과 연관된 신뢰도, 추정된 텍스트 컬러 및 배경색과 같은 각각의 텍스트 단어와 연관된 특성들을 포함할 수도 있다.

[0109] 도 13b는 업데이트된 장면 기술자를 예시한다. 장면 기술자에 대한 에너지 값은, 일단 각각의 이미지에 대

한 에너지가 계산되거나 또는 업데이트되면, 선택 모듈 (1110) 및 통합 모듈 (1114) 에 의해 업데이트될 수도 있다. 마찬가지로, 새로운 텍스트가 통합 모듈 (1114) 에서 OCR 캐시 (1112) 에 대한 키 이미지에 추가되면, 새로운 텍스트에 대한 부가적인 엔트리들이 추가될 수도 있다 (예, 도 13b의 TextN+1).

[0110] 도 14는 이미지로부터 글레이어를 제거하거나 또는 감소시키는 본 발명의 실시형태들을 예시하는 도면을 묘사한다. 카메라 (150) 를 사용하여, 실내 설정들에서 획득된 이미지들은 때때로 반사된 글레이어에 의해 영향을 받는다. 예를 들면, 정찬 메뉴는 정찬 조명으로부터 메뉴에서 반사되는 글레이어를 가질 수도 있다. 반사된 글레이어는 이미지들의 캡처링, 검출 및 인식을 위한 카메라 기반 애플리케이션들, 이를테면 광학적 문자 인식 (OCR) 의 성능을 떨어뜨릴 수도 있다. 반사된 글레이어는 관심 오브젝트 및 글레이어 소스 사이의 루미넌스의 상당한 비율에 의해 유발될 수도 있다. 관심 오브젝트, 글레이어 소스 및 카메라 사이의 각도와 같은 팩터들은 글레이어의 경험에 상당한 영향을 미칠 수도 있다.

[0111] 본 발명의 실시형태들은 둘 이상의 포지션들로부터의 동일한 관심 피사체의 이미지들을 사용하여 반사된 글레이어를 검출하고 감소시키는 기법들을 제공한다. 이미지는 둘 이상의 포지션들 사이에서 카메라 (150) 를 이동시킴으로써 또는 다수의 카메라들 (150), 이를테면 스테레오 카메라를 사용함으로써 획득될 수도 있다. 둘 이상의 포지션들로부터 이미지를 획득하는 것은, 글레이어 소스 및 관심 오브젝트에 관해 카메라와의 각도를 변경하는 것을 허용한다. 약간 상이한 관점 (vantage point) 들로부터의 2 개의 획득된 이미지들은, 동일한 관심 오브젝트에 대해 이미지의 상이한 부분들에서 글레이어를 나타낼 수도 있어, 감소된 글레이어를 갖는 이미지의 합성을 허용한다.

[0112] 도 14는 본 발명의 예시적인 실시형태를 나타내는 블록도이다. 도 14에서, 디바이스 (디바이스 (1402)) 에 연결된 스테레오 카메라 (150) 를 갖는 스테레오 카메라 컴퓨팅 디바이스가, 동일한 관심 오브젝트의 둘 이상의 이미지들을 획득하기 위해 사용된다. 하나의 구현예에서, 이미지들은 동시에 획득된다. 다른 실시형태에서, 이미지들은 상이한 시간들에 상이한 관점들로부터의 단일 카메라를 사용하여 차례차례 획득될 수도 있다. 도 14에서 알 수 있듯이, 왼쪽 뷰를 사용하여 획득된 이미지 (이미지 (1404)) 및 오른쪽 뷰를 사용하여 획득된 이미지 (이미지 (1406)) 양쪽 모두는 반사된 글레이어 (1410 및 1412) 를 가진다. 그러나, 2 개의 카메라들이 글레이어 소스 및 관심 오브젝트에 관해 이미지를 캡처하는 각도들이 상이하므로, 글레이어는 이미지의 상이한 부분들에서 존재한다. 도 14에 도시된 바와 같이, 본 발명의 실시형태들은 왼쪽 뷰 이미지 및 오른쪽 뷰 이미지를 사용하여 향상된 이미지를 합성한다. 하나의 실시형태에서, 2 개의 이미지들의 하나의 향상된 이미지의 합성은 통합 모듈 (1114) 에 의해 수행될 수도 있다. 일단 컴퓨팅 디바이스 (100) 가 매칭 모듈 (1110) 에서 2 개의 매칭 이미지들을 식별하면, 통합 모듈은, 아래의 도 15에서 더 논의되는 바와 같이, 최종 이미지를 OCR 캐시 (1112) 에 저장하기 전에 최종 이미지 (1408) 를 생성하기 위해 글레이어를 제거하거나 또는 감소시킴으로써 이미지를 향상시킬 수 있다.

[0113] 도 15는 본 개시물의 하나 이상의 예시적 양태들에 따라 이미지에서 반사 글레이어를 제거하거나 또는 감소시키는 본 발명의 실시형태를 예시하는 흐름도이다. 하나 이상의 양태들에 따르면, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 예를 들면, 컴퓨팅 디바이스 (100) 및/또는, 도 1에서 더 상세히 설명된 디바이스와 같은 컴퓨팅 디바이스 (100) 에 의해 및/또는 그것에서 구현될 수도 있다. 하나의 실시형태에서, 도 15를 참조하여 아래에서 설명되는 방법 단계들 중 하나 이상은 프로세서 (110) 또는 다른 프로세서와 같은 디바이스 (100) 의 프로세서에 의해 구현된다. 도 11 및 도 12에서 논의된 모듈들 및 컴포넌트들, 이를테면 장면 기술자 생성기 (1106), 매칭 모듈 (1108), 및/또는 통합 모듈 (1114) 은 컴퓨팅 디바이스 (100) 의 컴포넌트들로서 또한 구현될 수도 있고, 도 16에서 논의된 바와 같이 본 발명의 실시형태들의 구현 시에 사용될 수도 있다. 부가적으로 또는 대안으로, 본원에서 설명되는 방법들 및/또는 방법 단계들의 임의의 것 및/또는 모두는 컴퓨터 판독가능 명령들, 이를테면 메모리 (135), 스토리지 (125) 또는 다른 컴퓨터 판독가능 매체와 같은 컴퓨터 판독가능 매체 상에 저장된 컴퓨터 판독가능 명령들로 구현될 수도 있다.

[0114] 하나의 실시형태에서, 스테레오 카메라 (150) 가 2 개의 상이한 포지션들로부터 관심 오브젝트의 2 개의 이미지들을 획득하는데 사용될 수도 있다. 다른 실시형태에서, 2 개의 이미지들이 2 개의 상이한 포지션들로부터 차례차례 획득된다. 단계 1502에서, 왼쪽 뷰로부터의 제 1 이미지가 획득되고 블록 1504에서, 오른쪽 뷰로부터의 제 2 이미지가 획득된다. 블록들 (1502 및 1504) 에서 수행된 단계들은 동시에 또는 하나씩 하나씩 발생할 수도 있다. 다른 실시형태에서, 카메라 (150) 로부터의 수신된 이미지 및 OCR 캐시 (1112) 로부터의 이미지는 OCR 캐시 (1112) 에 저장된 이미지로부터 글레이어를 제거하기 위해 사용될 수도 있거나, 또는 OCR 캐시 (1112) 로부터의 2 개의 이미지들이 사용될 수도 있다. 단계 1506에서, 양쪽 모두의 이미지들로부터의 특징들은 추출될 수도 있고 2 개의 이미지들 사이의 기하학적 관계들은 도출될 수도 있다. 하나의

실시형태에서, OCR 엔진 (1104) 및 장면 기술자 생성기 (1106) 는 도 4에서 이전에 논의했듯이 이미지에서의 다양한 그래픽 오브젝트들 및 그것들의 서로에 대한 관계를 식별할 수도 있다. 다른 실시형태에서, 특징들은 해리스 코너 (Harris corner), 패스트 코너 등과 같은 기존의 특징 추출 알고리즘들을 사용하여 추출된다. 글레이 지역들에서의 특징들은 추정되고 제거될 수도 있다. 매칭 모듈 (1110) 은 매칭 프로세스를 사용하여 이미지들 및 그래픽 오브젝트들과 같은 이미지들의 특징들을 매치시킬 수도 있다. 다른 실시형태에서, 특징들은, 기존의 기술자들 (SIFT, SURF 등) 및 매칭 체제들, 이를테면 RANSAC (RANDOM SAmple Consensus) 를 사용하여 매치될 수도 있다. 기본 제약 시스템이 특징들 사이의 대응을 발견하기 위해 검색 공간을 제약할 시에 이용될 수도 있다. 블록 1508에서, 이미지들은 그것들의 기하학적 관계에 기초하여 정렬될 수도 있다. 더욱이, 이미지들의 글레이 지역들이 검출된다. 하나의 실시형태에서, 글레이 지역들은 2 개의 이미지들의 중첩 지역들 간에 차이를 검출함으로써 결정된다. 하나의 구현예에서, 컴퓨팅 디바이스 (100) 는 동일한 장면으로부터의 제 1 이미지 및 제 2 이미지의 복수의 특징들에서의 차이를 검출함으로써 이미지들 중 하나의 이미지의 하나의 지역에서 글레이를 식별한다. 예를 들어, 지역에서의 감소된 수의 특징들이 일부 실시형태들에서 글레이를 식별하는데 사용될 수도 있다.

[0115] 마지막으로, 블록 1510에서, 통합 프로세스를 사용하여, 예를 들어 통합 모듈 (1114) 로, 글레이 영향을 감소시키기 위해 적은 글레이를 갖는 2 개의 이미지들로부터의 중첩 지역들을 사용하여 2 개의 이미지들을 합성함으로써 향상된 이미지가 생성될 수도 있다. 하나의 실시형태에서, 본 발명의 실시형태들은, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 이미지들 중 하나의 이미지로부터 글레이를 갖는 제 1 지역을 식별하며, 제 1 장면으로부터의 둘 이상의 이미지들로부터의 제 2 이미지로부터 글레이가 없는 제 2 지역을 식별하는데, 제 1 지역 및 제 2 지역은 제 1 장면의 동일한 지역을 표현한다. 제 1 이미지 및 제 2 이미지를 사용하여, 최종 이미지가 감소된 글레이로 또는 글레이 없이 합성될 수도 있다. 상이한 기법들, 이를테면 2 개의 이미지들 사이에서 정보를 주는 패치들을 선택하는 것, 양호한 인식을 위해 전경 (foreground) 및 배경을 세그먼트화하는 것이, 이미지를 합성할 시에 사용될 수도 있다. 감소된 또는 제거된 글레이를 갖는 합성된 이미지는 OCR 캐시 (1112) 와 같은 캐시에 저장될 수도 있다.

[0116] 도 15에 예시된 구체적인 단계들은 본 발명의 일 실시형태에 따라 동작의 모드들 간을 스위칭하는 특정 방법을 제공한다는 것이 이해되어야 한다. 단계들의 다른 시퀀스들이 또한 그에 따라 대체 실시형태들에서 수행될 수도 있다. 예를 들어, 본 발명의 대안적 실시형태들은 위에서 약속된 단계들을 상이한 순서로 수행할 수도 있다. 예시를 위해, 사용자가 동작의 제 3 모드로부터 동작의 제 1 모드로, 제 4 모드에서 제 2 모드, 또는 그것들 사이의 임의의 조합으로 변경할 것을 선택할 수도 있다. 더구나, 도 15에 예시된 개개의 단계들은 개개의 단계에 적절한 대로 다양한 시퀀스들에서 수행될 수도 있는 다수의 서브단계들을 포함할 수도 있다. 더욱이, 부가적인 단계들은 특정 애플리케이션들에 의존하여 추가되거나 또는 제거될 수도 있다. 당업자는 프로세스의 많은 변형예들, 수정예들, 및 대체예들을 인식하고 이해할 것이다.

[0117] 도 16은 컴퓨팅 디바이스 (100), 이를테면 자동 및 지능적 OCR 캐시 업데이트들을 가지고서 OCR 캐싱 모드에서 동작하는 모바일 디바이스의 단순화된 표현을 사용하여 사용자 인터페이스에 대한 예시적인 구현예를 묘사한다. 예를 들어, 사용자 인터페이스는 출력 디바이스 (120) 에 의해 디스플레이될 수도 있다. 도 16a, 도 16b, 도 16c, 및 도 16d는 연속적인 시각점들에서 컴퓨팅 디바이스 (100) 를 나타낸다. 컴퓨팅 디바이스 (100) 의 상단 부분에 위치한 영역 (1604) 은, 컴퓨팅 디바이스 (100) 에 연결된 카메라의 렌즈를 통해 사용자가 볼 수 있는 시야를 나타낸다. 도면들의 하단 부분 쪽에서 묘사된 좁은 스트립은 OCR 캐시 (1112) 를 나타낸다. 하나의 양태에서, OCR 캐시 (1112) 에서의 이미지들의 표현들은 모바일 디바이스의 하단 부분 쪽에서 묘사된 좁은 스트립에서의 섬네일들 (1606, 1608 및 1610) 로서 사용자가 볼 수 있다. 하나의 실시형태에서, OCR 결과들은 이미지들 대신에 스트립으로 묘사된다. 다른 실시형태에서, 이미지들 및 OCR 결과들 양쪽 모두가 묘사된다. 사용자는 OCR 캐시 이미지를 선택하고, 텍스트를 선택하고 공통 텍스트 편집 기능들, 이를테면 편집, 복사 및 붙여넣기를 수행함으로써 그 이미지에서의 텍스트를 다룰 수도 있다. 예시적인 사용 사례에서, 사용자는, 예를 들어 위에서 설명된 사용자 인터페이스를 디스플레이하는 터치스크린 상의 이미지 또는 OCR 결과를 터치함으로써 OCR 캐시로부터 텍스트를 선택할 수도 있다. 좁은 스트립에서의 이미지를 선택하는 것은, 예를 들어 영역 (1604) 에 디스플레이될 이미지로 하여금 사이즈가 증대되게 할 수도 있어서, OCR 결과들은 더 큰 정밀도로 선택될 수도 있다. 일부 실시형태들에서, 선택은 좁은 스트립으로부터 직접 될 수도 있다. 예를 들어 사용자의 터치 또는 터치 없는 제스처 - 사용자의 손, 손가락, 눈, 머리, 또는 사용자에 의해 제어되는 다른 아이템으로 수행될 수도 있음- 에 기초하여 사용자에게 의해 식별되거나 또는 선택된 정보는, 번역되거나, 검색을 수행하기 위해 제출되거나, 또는 그렇지 않으면, 예를 들어 텍스트 박스에 입력하는 것에 의해 프로그램 또는 애플리케이션에 입력될 수도 있다. 일부 실시형태들에서, 이미지들 또는 OCR 결과들

의 수는 컴퓨팅 디바이스 (100) 의 스크린 상에 디스플레이될 수 있는 것보다 클 수도 있다. 하나의 이러한 실시형태들에서, 사용자는 이미지들 또는 결과들 사이에서 스크롤할 수도 있거나 또는 그렇지 않으면, 예를 들어, 스트립을 전진시키거나 또는 후퇴시키기 위하여 터치 스크린 상에서 손가락을 스윙핑함으로써 가시적 이미지들 또는 결과들을 토글할 수도 있다.

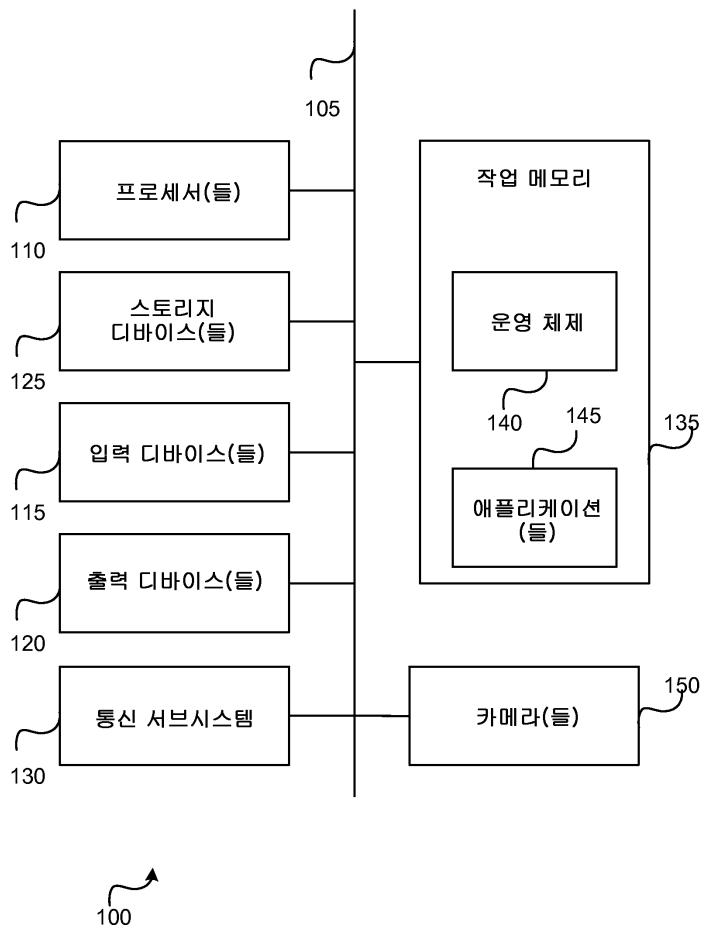
[0118] 도 16a를 참조하면, 장면 1이 카메라 (150) 의 시야를 나타낼 수도 있다. 카메라 (150) 는 본 발명의 실시 형태들에 따라 장면 1을 표현하는 이미지를 캡처하고 그 이미지를 장면 기술자와 함께 OCR 캐시에 저장할 수도 있는데, OCR 캐시 (1112) 가 도 16a에 연관된 시간에 비어있기 때문이다. 도 16b에서, 장면 2가 카메라의 시야에 있다. 카메라 (150) 는 장면 2를 표현하는 이미지를 캡처할 수도 있다. 본원에서 논의되는 기법들을 사용하여, 디바이스 (100) 의 컴포넌트들은 매칭 프로세스를 수행하고 OCR 캐시로부터의 장면 1에 대해 캡처된 이미지와 장면 2에 대해 캡처된 이미지를 비교한다. 그 이미지들이 상이한 장면들에 속하므로, 매치는 발견되지 않고 OCR 캐시는 장면 2에 대해 캡처된 이미지로 업데이트된다.

[0119] 도 16c에서, 장면 1이 다시 카메라 (150) 의 시야에 있다. 장면 1을 표현하는 이미지는 카메라 (150) 에 의해 획득된다. 이제 디바이스의 컴포넌트들이 수신된 이미지에 대한 장면 기술자 및 OCR 캐시에 저장된 이미지들의 장면 기술자 사이에서 비교를 수행하는 경우, 매치가 발견된다. 장면 기술자들로부터의 양쪽 모두의 이미지들의 계산된 에너지들을 사용하여, 키 이미지가 2 개의 이미지들로부터 선택될 수도 있다. 통합 프로세스의 부분으로서, 키 아닌 이미지에서 발견된 임의의 고유 단어들이 이미지에 추가될 수도 있고 장면 기술자는 업데이트될 수도 있다. 게다가, 장면 기술자에서의 임의의 단어들 또는 결과들은 높은 신뢰도를 갖는 다른 단어들 또는 결과들에 의해 업데이트되거나 또는 대체될 수도 있다. 통합 프로세스는 도 9 및 도 10에서 더 상세히 설명된다. 하나의 구현예에서, 부가적으로, 글레이는 또한, 도 14 및 도 15에서 논의된 바와 같이, 키 이미지로부터 감소되거나 또는 제거될 수도 있다. 도 16c에 도시된 바와 같이, 장면 1에 대한 OCR 캐시 (1112) 는 새로운 이미지 및 장면 기술자로 업데이트된다. 도 16b와 유사하게, 도 16d는 시야에서 새로운 장면을 가진다. 장면 3을 표현하는 이미지가 캡처되고 OCR 캐시 (1112) 는 새로운 이미지로 업데이트된다. 일부 실시형태들에서, 키 이미지들은, 예를 들어 도 16a 내지 도 16d에서 예시된 바와 같이, 장면들이 식별되게 하기 위하여 좁은 스트립으로 디스플레이된다. 일부 실시형태들에서, 키 이미지들의 순서는 어떤 장면이 가장 최근에 검출되었는지에 기초하여 가변한다. 예를 들어, 가장 최근에 검출된 장면을 표현하는 키 이미지는 항상 좁은 스트립의 가장 왼쪽에서 보여질 수도 있다. 키 이미지(들)를 디스플레이하는 다른 순서들 또는 기법들이 부가적으로 또는 대신 사용될 수도 있다.

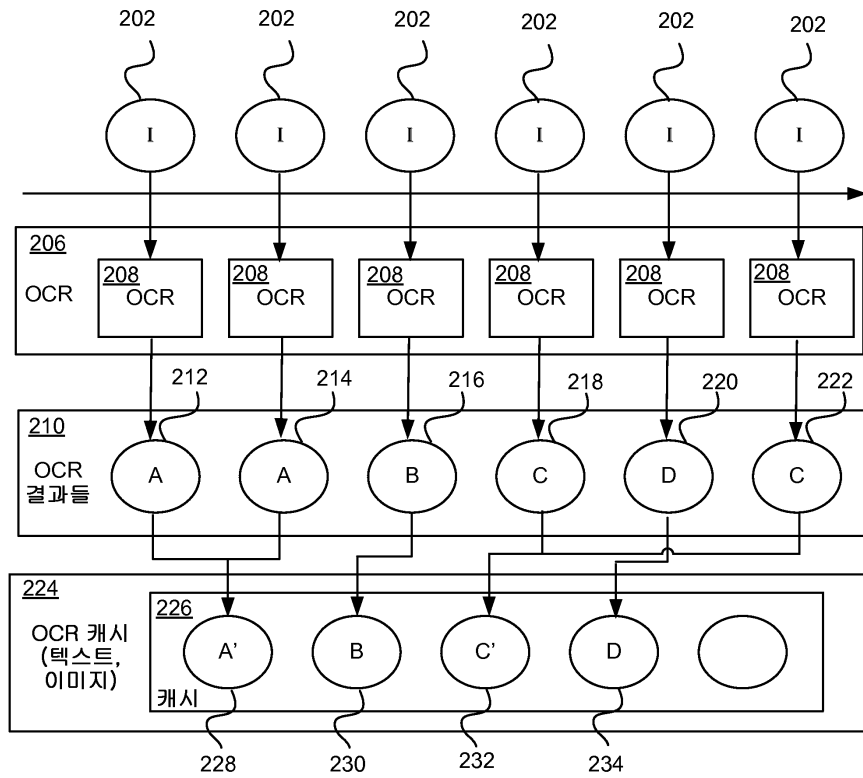
[0120] 본원에서 설명된 바와 같은 캐시 업데이트는 사용자에게 유익할 수도 있는데, 사용자가 동일한 장면에 또는 복수의 장면들과 연관된 정보의 모두에 액세스하는 것을 허용하기 때문이다. 더욱이, 시간이 지남에 따라 동일한 장면에 대한 다수의 이미지들 사이의 이미지 정보의 중복으로, 본원에서 설명되는 기법들은 장면과 연관된 최상의 정보 또는 새로 획득된 정보를 저장하며, 그러므로 시간이 지남에 따라 결과들을 추가로 리파인할 수도 있다. 본원에서 설명되는 기법들과 연관된 다른 장점들은, OCR 결과들을 저장하는데 필요한 감소된 메모리 공간과 OCR 캐시에 대한 감소된 히트 횟수를 포함하여 전반적인 시스템에서 더 빠른 성능이 나타나게 한다.

도면

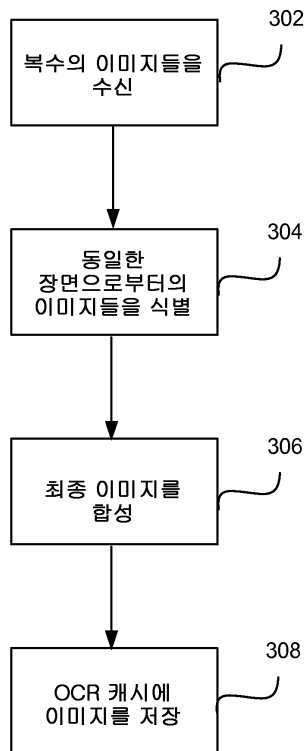
도면1



도면2

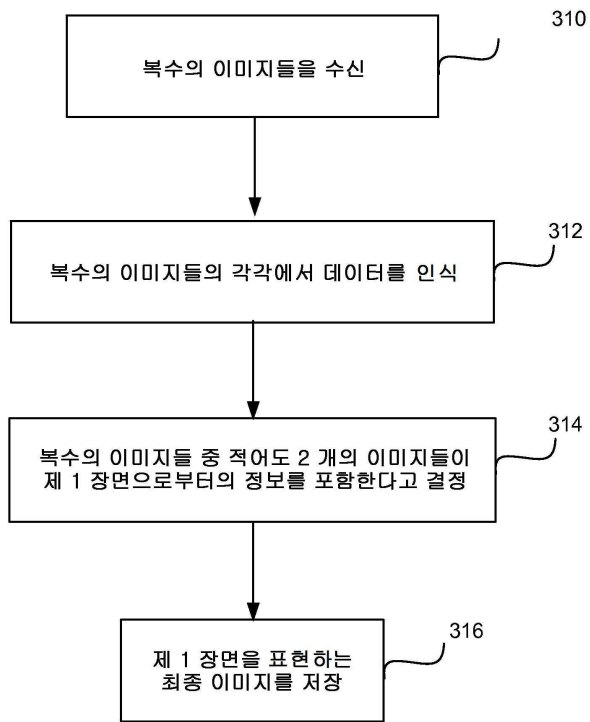


도면3a

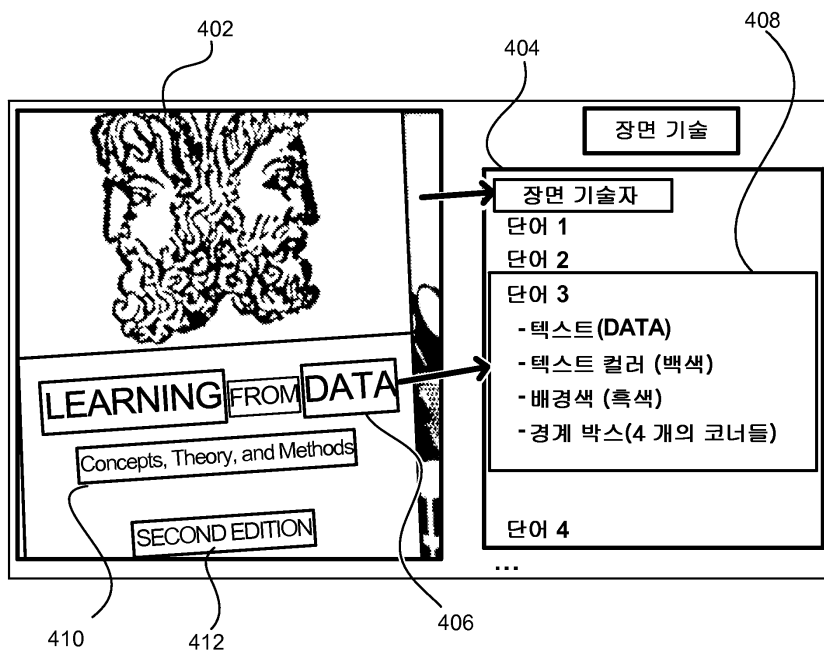




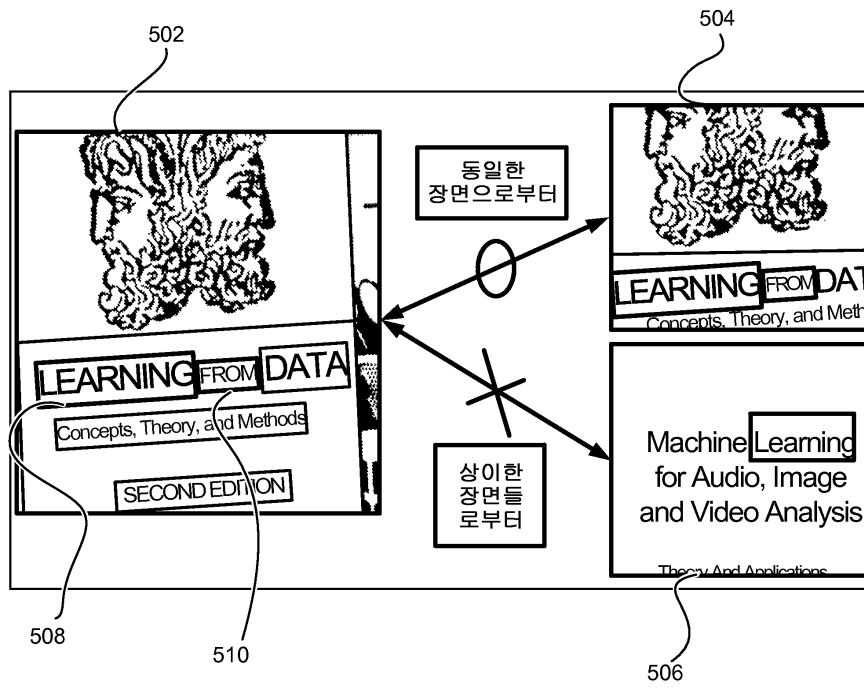
도면3b



도면4

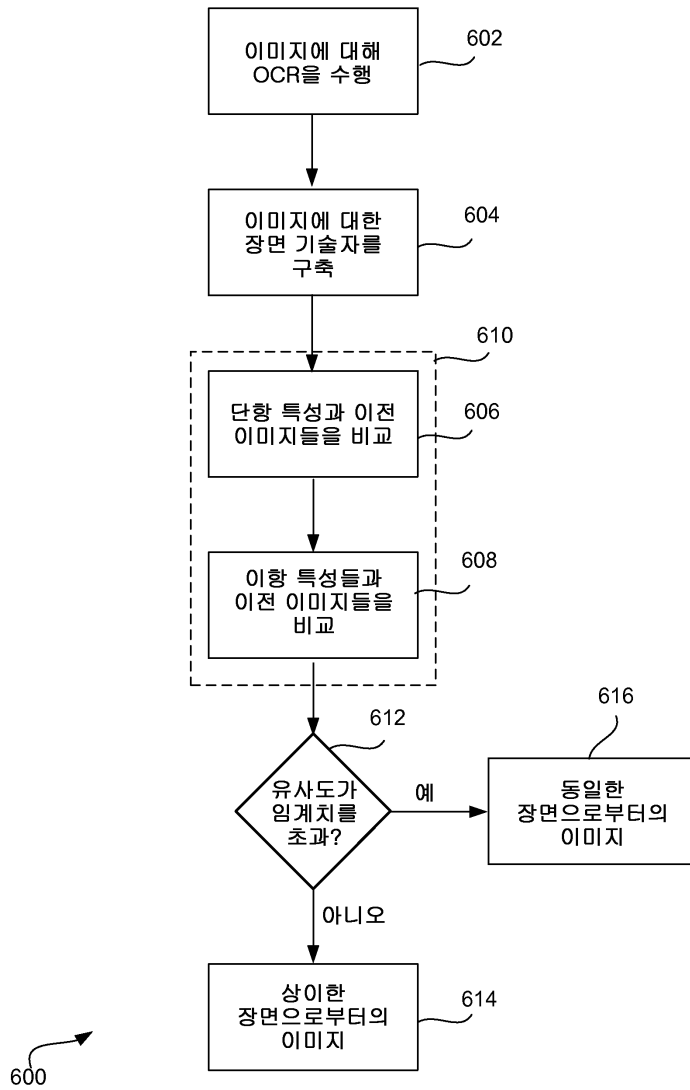


도면5

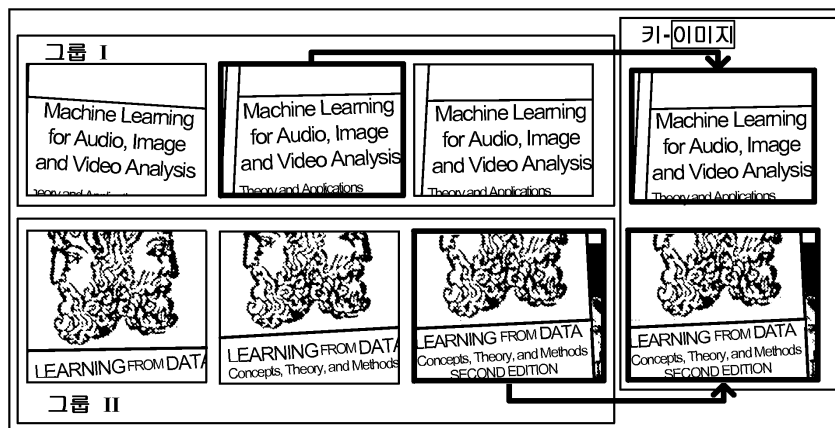




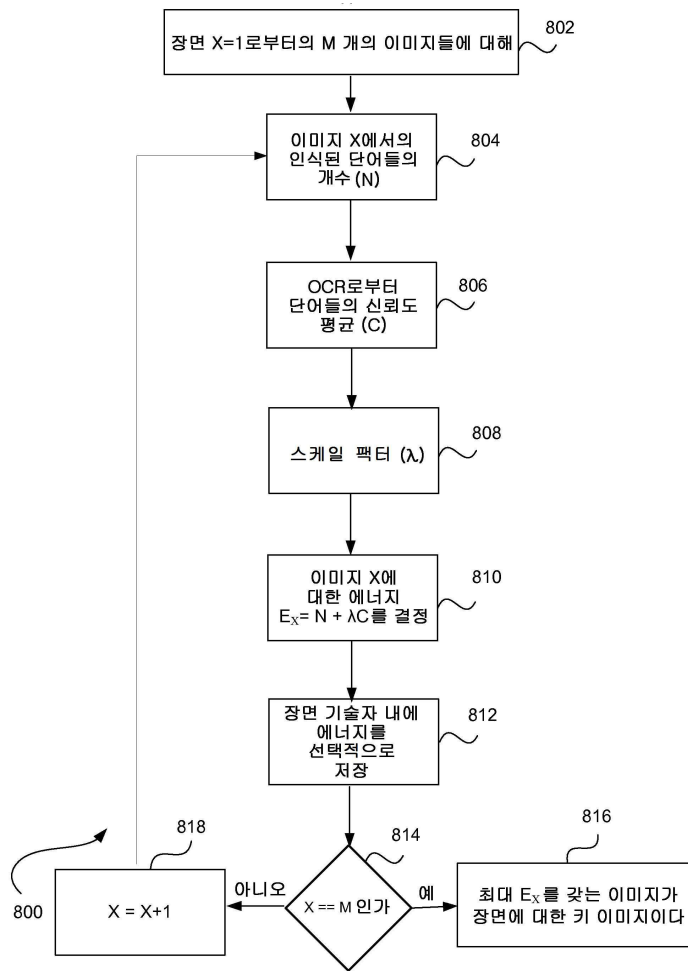
도면6



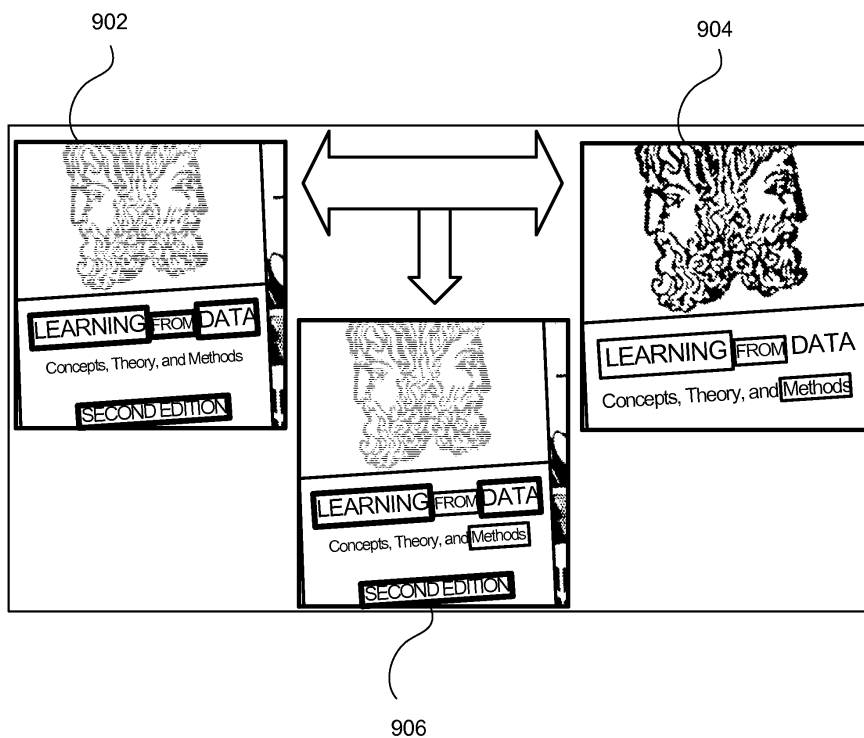
도면7



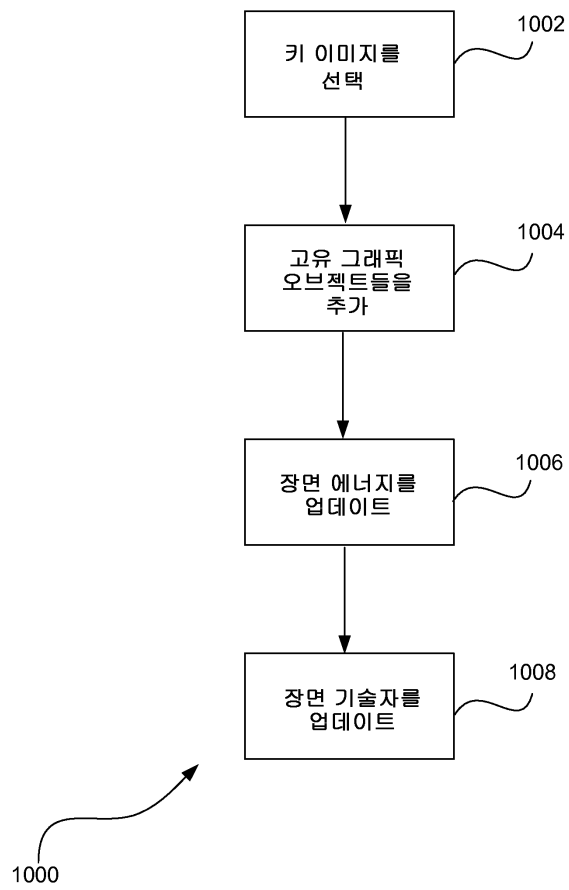
도면8



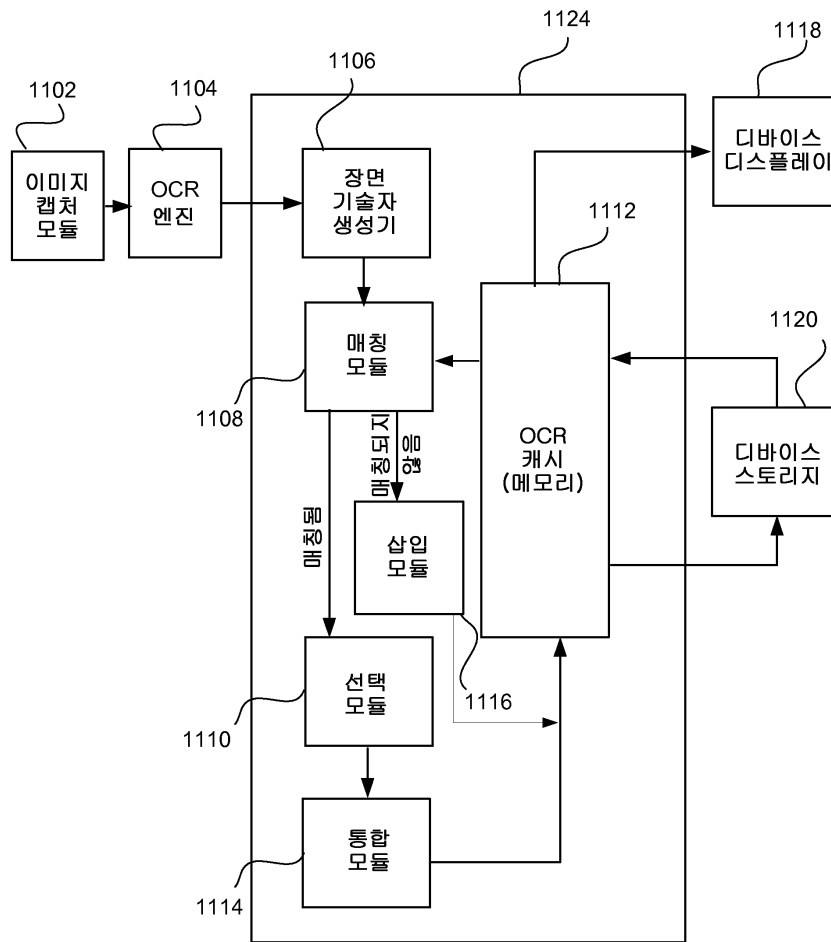
도면9



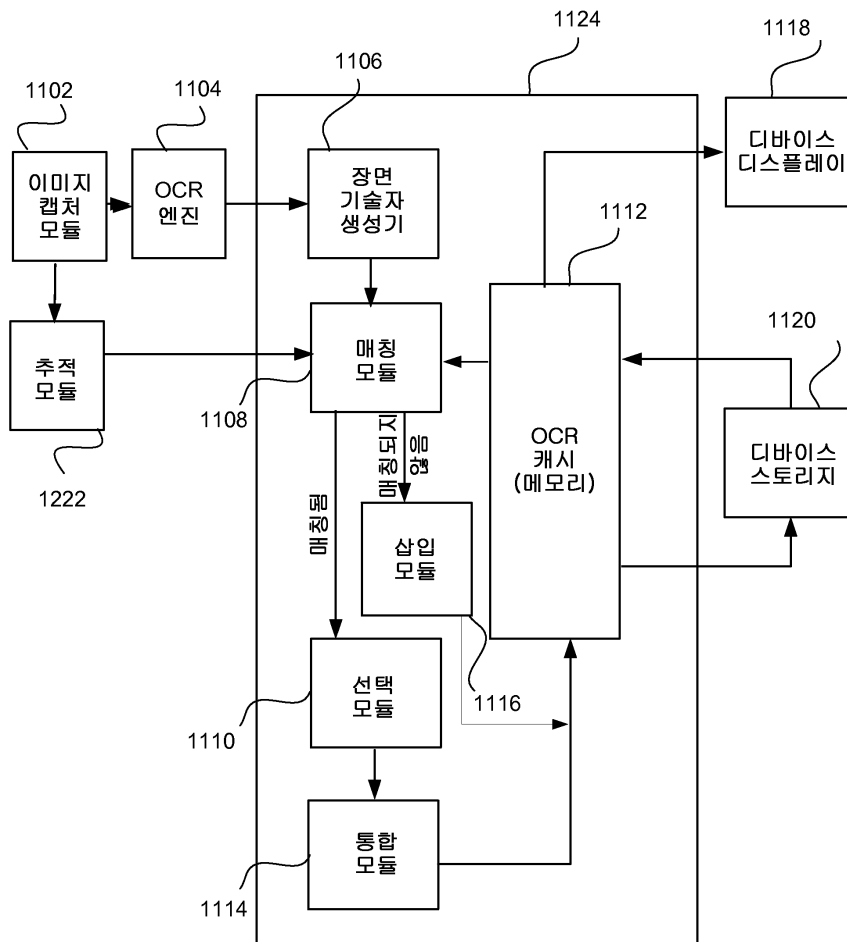
도면10



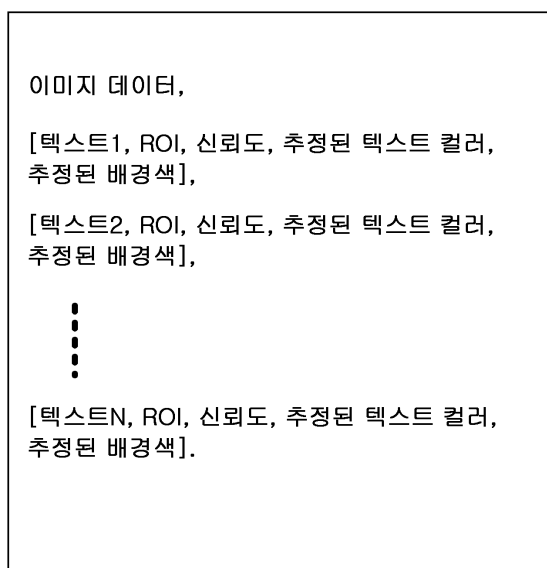
도면11



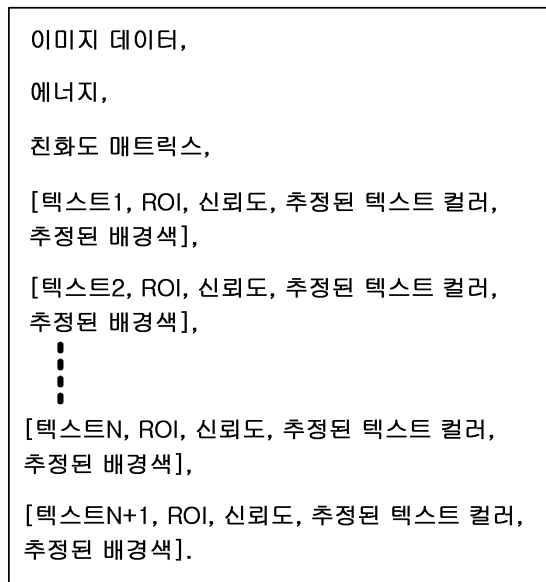
도면12



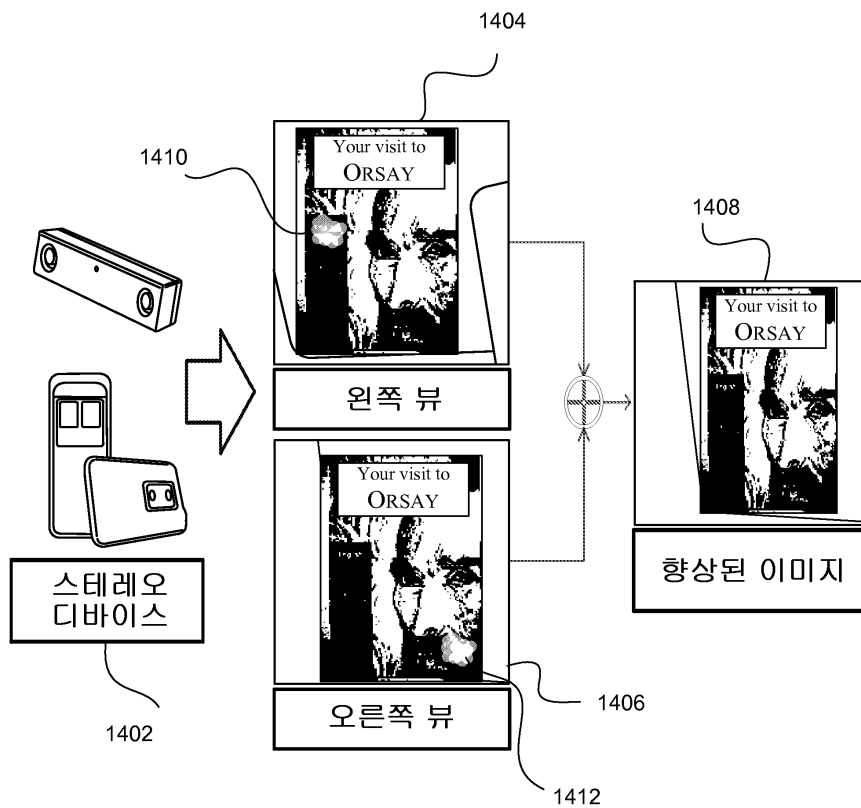
도면13a



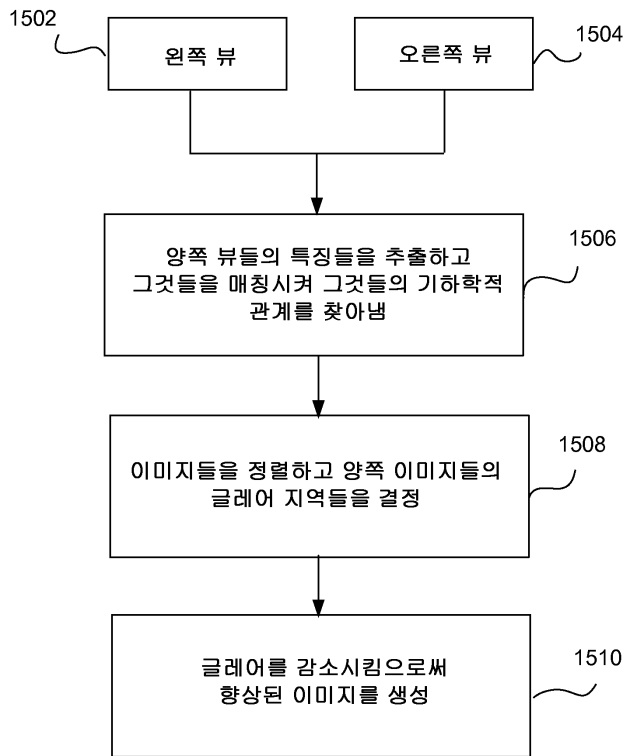
도면13b



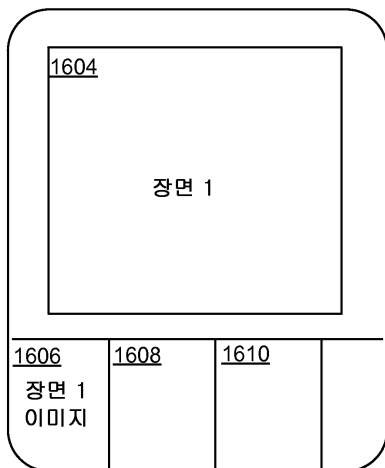
도면14



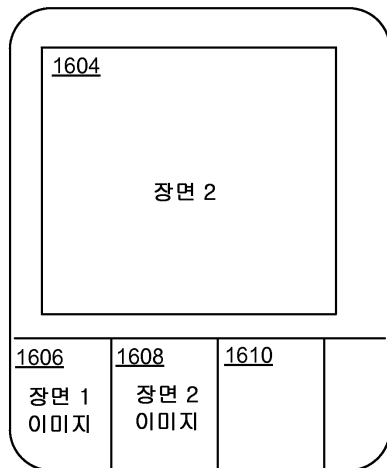
도면15



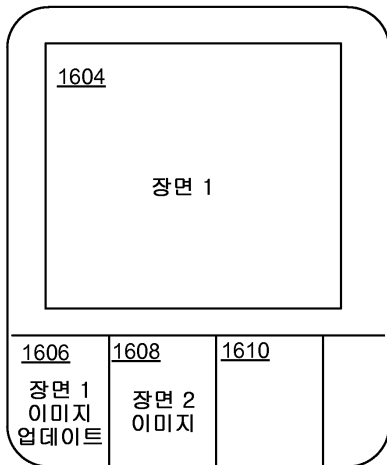
도면16a



도면16b



도면16c



도면16d

