



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2010년06월07일  
(11) 등록번호 10-0960639  
(24) 등록일자 2010년05월24일

(51) Int. Cl.

G06T 1/00 (2006.01)

(21) 출원번호 10-2008-7004399

(22) 출원일자(국제출원일자) 2006년08월22일

심사청구일자 2008년02월22일

(85) 번역문제출일자 2008년02월22일

(65) 공개번호 10-2008-0034480

(43) 공개일자 2008년04월21일

(86) 국제출원번호 PCT/JP2006/316812

(87) 국제공개번호 WO 2007/023993

국제공개일자 2007년03월01일

(30) 우선권주장

11/461,147 2006년07월31일 미국(US)

(뒷면에 계속)

(56) 선행기술조사문헌

KR1020050075301 A\*

JP평성10228468 A

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

가부시키가이샤 리코

일본 도쿄도 오다꾸 나가마고메 1쵸메 3-6

(72) 발명자

힐 조나단 제이

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서  
치 센터 내

리 다 시양

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서  
치 센터 내

(뒷면에 계속)

(74) 대리인

김태홍, 신정건

전체 청구항 수 : 총 26 항

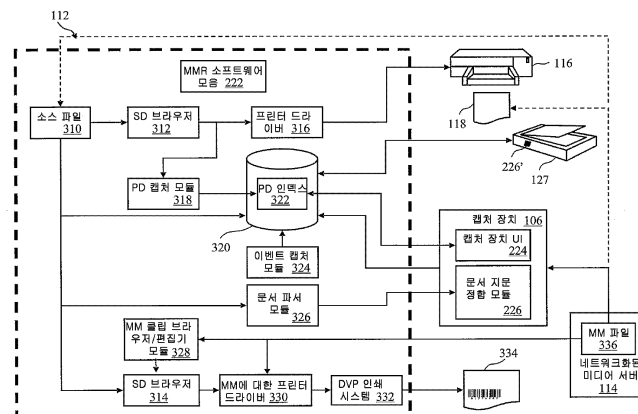
심사관 : 신재철

(54) 혼합 미디어 문서 시스템용의 데이터 구성 및 액세스

(57) 요약

MMR(Mixed Media Reality; 혼합 미디어 현실) 시스템 및 연관된 기법들이 개시되어 있다. 본 발명의 MMR 시스템은 적어도 2가지 유형(제1 미디어로서의 인쇄된 종이와, 제2 미디어로서의 디지털 콘텐츠 및/또는 웹 링크)의 미디어를 포함하는 혼합 미디어 문서(mixed media document)를 형성하는 메카니즘을 제공한다. 특정의 실시예에 따르면, MMR 시스템은 텍스트-기반 인덱스를 사용하여 탐색(look-up)을 가능하게 해주는 방식으로 인쇄된 문서로부터 추출된 객체들 사이의 2차원의 기하학적 관계를 표현하는 인덱스 테이블을 구비하도록 구성된 콘텐츠-기반 검색 데이터베이스를 포함한다. 순위가 매겨진 일련의 문서, 페이지 및 위치 가설은, 인덱스 테이블로부터 데이터가 제공될 경우에 계산될 수 있다. 이 기술은 이미지 패치에서 검출된 피처를, 피처 자체 및 이들 사이의 기하학적 관계 모두를 나타내는 문자 용어(또는 다른 검색 가능한 피처)로 효율적으로 변환시킨다. 저장 장비를 이용하여 각각의 문서 이미지 패치에 대한 추가의 특징을 저장할 수 있다.

대표도



(72) 발명자

**피에르술 커트**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

**하트 피터 이**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

**그라함 재메이**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

**에를 베르나**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

**오엘에스티 다니엘 지 벤**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

**루 시아오예**

미국 캘리포니아주 멘로 파크 스위트 115 샌드 힐  
로드 2882 리코이노베이션 인크 캘리포니아 리서치  
센터 내

(30) 우선권주장

11/461,164 2006년07월31일 미국(US)

60/710,767 2005년08월23일 미국(US)

60/792,912 2006년04월17일 미국(US)

60/807,654 2006년07월18일 미국(US)

## 특허청구의 범위

### 청구항 1

혼합 미디어 문서 시스템에서 정보를 구성하고 액세스하는 컴퓨터 구현 방법으로서,

종이 문서의 전자적 표시를 생성하는 단계와,

상기 종이 문서의 2차원 애스펙트(aspect)를 캡처하는, 상기 종이 문서상의 피쳐(feature)를 식별하는 단계와,

상기 피쳐의 위치를 식별하는 단계와,

각각의 위치에 의해 상기 피쳐를 인덱싱하여, 인덱스 테이블을 생성하는 단계

를 포함하고,

상기 종이 문서의 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피쳐(feature)를 식별하는 상기 단계는, 2개의 연속하는 시퀀스 사이의 수직 중첩(overlap)의 양을 체크함으로써 텍스트의 시퀀스를 논리적 라인(logical line)으로 그룹화하는 단계를 포함하는 것인 컴퓨터 구현 방법.

### 청구항 2

제1항에 있어서, 상기 종이 문서를 수신하는 예비 단계를 더 포함하는 컴퓨터 구현 방법.

### 청구항 3

제1항에 있어서, 상기 피쳐 중 적어도 하나와 연관된 하나 이상의 특징(characteristics)을 저장하는 단계를 더 포함하는 컴퓨터 구현 방법.

### 청구항 4

제3항에 있어서, 상기 하나 이상의 특징은, 문서 정보(textual information)의 검색, 그래픽 정보의 검색, 프로세스의 실행, 명령의 실행, 주문, 비디오의 검색, 사운드의 검색, 정보의 저장, 새로운 문서의 생성, 문서의 인쇄 및 문서의 디스플레이 중 적어도 하나를 포함한 하나 이상의 동작(action)을 포함하는 것인 컴퓨터 구현 방법.

### 청구항 5

제1항에 있어서, 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피쳐를 식별하는 상기 단계는, 수평으로 정렬되고 수직으로 정렬된 객체(objects)를 식별하는 단계를 포함하는 것인 컴퓨터 구현 방법.

### 청구항 6

제1항에 있어서, 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피쳐를 식별하는 상기 단계는, 수평 및 수직 단어 쌍을 식별하는 단계를 포함하는 것인 컴퓨터 구현 방법.

### 청구항 7

제1항에 있어서, 상기 종이 문서의 전자적 표시를 생성하는 상기 단계는 스캐닝 또는 인쇄 공정 중에 실행되는 것인 컴퓨터 구현 방법.

### 청구항 8

삭제

### 청구항 9

제1항에 있어서,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어(query term)를 수신하는 단계와,

상기 인덱스 테이블로부터의 데이터에 기초하여, 상기 질의어에 응답하는 적어도 하나의 혼합 미디어 문서 및

위치 가설(location hypotheses)을 계산하는 단계  
를 더 포함하는 컴퓨터 구현 방법.

#### 청구항 10

제9항에 있어서, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는, 타겟 문서를 수신하는 단계와, 타겟 문서의 적어도 패치의 이미지를 생성하는 단계와, 상기 이미지를 기초로 상기 하나 이상의 질의어를 발생시키는 단계 이후에, 실시되는 것인 컴퓨터 구현 방법.

#### 청구항 11

제10항에 있어서, 상기 이미지에 기초하여 상기 하나 이상의 질의어를 발생시키는 상기 단계는, 상기 이미지로부터 추출된 수평 및 수직 단어 쌍을 생성하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 12

제11항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,  
상기 타겟 문서의 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 단계와,  
상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 13

제12항에 있어서, 각각의 단어 쌍은 역 문서 빈도(inverse document frequency)와 연관되어 있으며, 상기 타겟 문서의 적어도 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 상기 단계는,  
각 단어 쌍에 대한 상기 역 문서 빈도를 단어 쌍이 보이는 문서 페이지에 의해 인덱싱된 누적기에 가산하는 단계와,  
그 누적기에서의 최대값이 임계값(threshold)을 초과하는 것에 응답하여, 상기 패치에 대한 정합(match)으로서 대응 문서 페이지를 출력하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 14

제13항에 있어서, 상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 상기 단계는,  
각 단어 쌍의 주변의 존(zone)에 있는 각 셀에 가중치를 추가하는 단계로서, 각 셀에 대한 상기 가중치는 그 단어 쌍의 상기 역 문서 빈도와 그 셀과 상기 존의 중심 사이의 정규화된 기하학적 거리의 곱에 의해 결정되는 것인 단계와,  
최대값을 갖는 상기 셀에 대하여 상기 누적기의 대응 누적 어레이를 검색하는 단계와,  
상기 최대값이 임계값을 초과하는 것에 응답하여, 상기 패치의 위치로서 상기 셀의 좌표를 보고하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 15

제9항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,  
각 질의어와 연관된 하나 이상의 위치를 검색하도록 상기 인덱스 테이블에서 상기 하나 이상의 질의어 각각을 탐색하는 단계와,  
각각의 식별된 위치에 대하여, 상기 위치를 포함하는 하나 이상의 후보 영역을 식별하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 16

제15항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,  
상기 하나 이상의 질의어 모두와 가장 부합하는 상기 하나 이상의 후보 영역 중 하나를 식별하는 단계와,

상기 하나 이상의 후보 영역 중 하나가 사전 정의한 정합 기준(matching criteria)을 만족시키는 것으로 결정되는 것에 응답하여, 그 영역을 상기 타겟 문서에 대하여 정합하는 것(match)으로서 확인하는 단계를 포함하는 것인 컴퓨터 구현 방법.

#### 청구항 17

하나 이상의 프로세서에 의해 실행될 때에, 상기 프로세서로 하여금 혼합 미디어 문서 시스템에서의 정보를 구성하고 액세스하는 프로세스를 실행하도록 하게 하는 명령어가 인코딩된 기계 판독 가능한 매체로서,

상기 프로세스는,

종이 문서의 전자적 표시를 생성하는 단계와,

상기 종이 문서의 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피처를 식별하는 단계와,

상기 피처의 위치를 식별하는 단계와,

각각의 위치에 의해 상기 피처를 인덱싱하여, 인덱스 테이블을 생성하는 단계

를 포함하고,

상기 종이 문서의 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피처를 식별하는 상기 단계는, 2개의 연속하는 시퀀스 사이의 수직 중첩(overlap)의 양을 체크함으로써 텍스트의 시퀀스를 논리적 라인(logical line)으로 그룹화하는 단계를 포함하는 것인 기계 판독 가능한 매체.

#### 청구항 18

제17항에 있어서, 상기 프로세스는, 상기 피처 중 적어도 하나와 연관된 하나 이상의 특징을 저장하는 단계를 더 포함하는 것인 기계 판독 가능한 매체.

#### 청구항 19

제17항에 있어서, 2차원 애스펙트를 캡처하는, 상기 종이 문서상의 피처를 식별하는 상기 단계는, 수평 및 수직 단어 쌍을 식별하는 단계를 포함하는 것인 기계 판독 가능한 매체.

#### 청구항 20

제17항에 있어서, 상기 프로세스는,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 단계와,

상기 인덱스 테이블로부터의 데이터를 기초로 하여, 상기 질의어에 응답하는 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계

를 더 포함하는 것인 기계 판독 가능한 매체.

#### 청구항 21

혼합 미디어 문서 시스템에서 정보에 액세스하는 컴퓨터 구현 방법으로서,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 단계와,

혼합 미디어 문서의 피처 위치와 문서 피처를 인덱싱하는 인덱스 테이블로부터의 데이터에 기초하여, 상기 질의어에 응답하는 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계

를 포함하고,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는,

상기 타겟 문서를 수신하는 단계와, 타겟 문서의 적어도 패치의 이미지를 생성하는 단계와, 상기 이미지에 기초하여 상기 하나 이상의 질의어를 발생시키는 단계 이후에 실시되며,

상기 이미지에 기초하여 상기 하나 이상의 질의어를 발생시키는 상기 단계는, 상기 이미지로부터 추출된 수평 및 수직 단어 쌍을 생성하는 단계를 포함하고,

적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,

상기 타겟 문서의 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 단계와,

상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 단계를 포함하며,

각각의 단어 쌍은 역 문서 빈도와 연관되어 있으며, 상기 타겟 문서의 적어도 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 상기 단계는,

각 단어 쌍에 대한 상기 역 문서 빈도를 그 단어 쌍이 보이는 문서 페이지에 의해 인덱싱된 누적기에 가산하는 단계와,

그 누적기에서의 최대값이 임계값을 초과하는 것에 응답하여, 상기 패치에 대한 정합(match)으로서 대응 문서 페이지를 출력하는 단계를 포함하는 것인 컴퓨터 구현 방법.

## 청구항 22

삭제

## 청구항 23

삭제

## 청구항 24

삭제

## 청구항 25

삭제

## 청구항 26

제21항에 있어서, 상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 상기 단계는,

각 단어 쌍의 주변의 존에 있는 각 셀에 가중치를 부가하는 단계로서, 각 셀에 대한 상기 가중치는 그 단어 쌍의 상기 역 문서 빈도와 그 셀과 상기 존의 중심 사이의 정규화된 기하학적 거리의 곱에 의해 결정되는 것인 단계와,

최대값을 갖는 상기 셀에 대하여 누적기의 대응 누적 어레이를 검색하는 단계와,

상기 최대값이 임계값을 초과하는 것에 응답하여, 상기 패치의 위치로서 상기 셀의 좌표를 보고하는 단계를 포함하는 것인 컴퓨터 구현 방법.

## 청구항 27

제21항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,

각 질의어와 연관된 하나 이상의 위치를 검색하도록 상기 인덱스 테이블에서 상기 하나 이상의 질의어 각각을 탐색하는 단계와,

각각의 식별된 위치에 대하여, 상기 위치를 포함하는 하나 이상의 후보 영역을 식별하는 단계를 포함하는 것인 컴퓨터 구현 방법.

## 청구항 28

제27항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,

상기 하나 이상의 질의어 모두와 가장 부합하는 상기 하나 이상의 후보 영역 중 하나를 식별하는 단계와,

상기 하나 이상의 후보 영역 중 하나가 사전 정의한 정합 기준(matching criteria)을 만족시키는 것으로 결정되는 것에 응답하여, 그 영역을 상기 타겟 문서에 대하여 정합하는 것(match)으로서 확인하는 단계를 포함하는 것인 컴퓨터 구현 방법.

## 청구항 29

하나 이상의 프로세서에 의해 실행될 때에, 상기 프로세서로 하여금 혼합 미디어 문서 시스템에서의 정보에 액세스하는 프로세스를 실행하도록 하게 하는 명령어가 인코딩된 기계 판독 가능한 매체로서,

상기 프로세서는,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 단계와,

혼합 미디어 문서의 피쳐 위치와 문서 피쳐를 인덱싱하는 인덱스 테이블로부터의 데이터에 기초하여, 상기 질의어에 응답하는 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계

를 포함하고,

타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는,

상기 타겟 문서를 수신하는 단계와, 타겟 문서의 적어도 패치의 이미지를 생성하는 단계와, 상기 이미지에 기초하여 상기 하나 이상의 질의어를 발생시키는 단계 이후에 실시되며,

상기 이미지를 기초로 상기 하나 이상의 질의어를 발생시키는 상기 단계는, 상기 이미지로부터 추출된 수평 및 수직 단어 쌍을 생성하는 단계를 포함하고,

적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,

상기 타겟 문서의 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 단계와,

상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 단계를 포함하며,

각각의 단어 쌍은 역 문서 빈도와 연관되어 있으며, 상기 타겟 문서의 적어도 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 상기 단계는,

각 단어 쌍에 대한 상기 역 문서 빈도를 그 단어 쌍이 보이는 문서 페이지에 의해 인덱싱된 누적기에 가산하는 단계와,

그 누적기에서의 최대값이 임계값을 초과하는 것에 응답하여, 상기 패치에 대한 정합으로서 대응 문서 페이지를 출력하는 단계를 포함하는 것인 기계 판독 가능한 매체.

## 청구항 30

삭제

## 청구항 31

삭제

## 청구항 32

삭제

## 청구항 33

삭제

## 청구항 34

제29항에 있어서, 상기 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 상기 단계는,

각 단어 쌍의 주변의 존에 있는 각 셀에 가중치를 부가하는 단계로서, 각 셀에 대한 상기 가중치는 그 단어 쌍의 상기 역 문서 빈도와 그 셀과 상기 존의 중심 사이의 정규화된 기하학적 거리의 곱에 의해 결정되는 것인 단계와,

최대값을 갖는 상기 셀에 대하여 상기 누적기의 대응 누적 어레이를 검색하는 단계와,

상기 최대값이 임계값을 초과하는 것에 응답하여, 상기 패치의 위치로서 상기 셀의 좌표를 보고하는 단계를 포

합하는 것인 기계 판독 가능한 매체.

### 청구항 35

제29항에 있어서, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는,

각 질의어와 연관된 하나 이상의 위치를 검색하도록 상기 인덱스 테이블에서 하나 이상의 질의어 각각을 탐색하는 단계와,

각각의 식별된 위치에 대하여, 상기 위치를 포함하는 하나 이상의 후보 영역을 식별하는 단계와,

상기 하나 이상의 질의어 모두와 가장 부합하는 상기 하나 이상의 후보 영역 중 하나를 식별하는 단계와,

상기 하나 이상의 후보 영역 중 하나가 사전 정의한 정합 기준을 만족시키는 것으로 결정되는 것에 응답하여, 그 영역을 상기 타겟 문서에 대하여 정합하는 것(match)으로서 확인하는 단계를 포함하는 것인 기계 판독 가능한 매체.

## 명세서

### 기술 분야

[0001] 본 발명은 적어도 2개의 미디어 유형으로부터 형성되는 혼합 미디어 문서를 생성하는 기법에 관한 것으로서, 보다 상세하게는 혼합 미디어 문서를 생성하기 위해 전자 미디어와 함께 인쇄된 미디어를 사용하는 MMR(Mixed Media Reality, 혼합 미디어 현실) 시스템에 관한 것이다.

### 배경 기술

[0002] 여러 해 동안 많은 정황에서 문서 인쇄 및 복사 기술이 사용되어 왔다. 예로서, 프린터 및 복사기가 개인적 및 상업적 사무실 환경에서, 퍼스널 컴퓨터를 갖춘 가정 환경에서, 그리고 문서 인쇄 및 출판 서비스 환경에서 사용되고 있다. 그렇지만, 인쇄 및 복사 기술이 이전에는 정적 인쇄된 미디어(즉, 종이 문서)와 디지털 통신, 네트워크, 정보 제공, 광고, 엔터테인먼트, 및 전자 상거래 등을 포함하는 상호작용성의 "가상 세계(virtual world)" 간의 차이를 줄이는 수단으로서 생각되지 않았다.

[0003] 인쇄된 미디어는 수세기 동안 뉴스 및 광고 정보 등의 정보를 전달하는 주요 소스이었다. 과거 수년에 걸쳐 퍼스널 컴퓨터 및 PDA(personal digital assistant) 및 셀룰러 전화(예를 들어, 셀룰러 카메라 전화) 등의 개인 전자 장치의 등장 및 점차적으로 증가하는 인기는 인쇄된 미디어를 전자적으로 판독가능하고 탐색가능한 형태로 이용가능하게 만듦으로써 또한 종래의 인쇄된 미디어와 비견되지 않는 상호작용적 멀티미디어 기능(interactive multimedia capabilities)을 도입함으로써 인쇄된 미디어의 개념을 확장시켰다.

[0004] 불행히도, 전자적으로 액세스가능한 가상의 멀티미디어-기반 세계와 인쇄 미디어의 물리적 세계 간에 차이가 존재한다. 예를 들어, 선진국의 거의 모든 사람이 인쇄된 미디어 및 전자적 정보에 매일같이 액세스하고 있지만, 인쇄된 미디어 및 개인 전자 장치의 사용자는 이 둘 간의 링크를 형성하는 데 필요한(즉, 혼합 미디어 문서를 용이하게 해주는) 도구 및 기술을 소유하고 있지 않다.

[0005] 게다가, 가상의 또는 디지털 미디어가 갖추고 있지 않는, 촉감, 전력을 필요로 하지 않는 것, 및 구성 및 저장의 연속성 등의 종래의 인쇄된 미디어가 제공하는 특성의 유익한 속성들이 있다. 이와 마찬가지로, 휴대성(예를 들어, 셀 전화 또는 랩톱의 저장 장치에 넣어 들고 다님) 및 전송의 편이성(예를 들어, 이메일) 등의 종래의 디지털 미디어가 제공하는 특성의 유익한 속성들이 있다.

[0006] 이들 이유로 인해, 인쇄된 미디어와 가상의 미디어 둘다에 연관된 이점들을 이용할 수 있는 기법들이 필요하다.

### 발명의 상세한 설명

[0007] 본 발명의 하나 이상의 실시예의 적어도 하나의 양태는, 혼합 미디어 문서 시스템에서 정보를 구성하고 액세스하는 컴퓨터 구현 방법을 제공한다. 이 컴퓨터 구현 방법은, 종이 문서의 전자적 표시를 생성하는 단계와, 종이 문서의 2차원 애스펙트(aspect)를 캡처하는, 종이 문서상의 피쳐(feature)를 식별하는 단계와, 피쳐의 위치를 식별하는 단계와, 각각의 위치에 의해 피쳐를 인덱싱하여, 인덱스 테이블을 생성하는 단계를 포함한다. 이 방법은 종이 문서를 수신하는 예비 단계를 포함할 수 있다. 이 방법은 피쳐 중 적어도 하나와 연관된 하나 이상의 특징(characteristics)을 저장하는 단계를 포함할 수 있다. 그러한 경우에, 하나 이상의 특징은, 문서 정



보(textual information)의 검색, 그래픽 정보의 검색, 프로세스의 실행, 명령의 실행, 주문, 비디오의 검색, 사운드의 검색, 정보의 저장, 새로운 문서의 생성, 문서의 인쇄 및 문서의 디스플레이 중 적어도 하나를 포함한 하나 이상의 동작(action)을 포함할 수 있다. 다른 특징의 경우에, 2차원 애스펙트를 캡처하는, 종이 문서상의 피처를 식별하는 상기 단계는, (인접하는 가의 여부에 관계없이) 수평으로 정렬되고 (인접하는 가의 여부에 관계없이) 수직으로 정렬된 객체(objects)를 식별하는 단계를 포함할 수 있다. 다른 특징의 경우에, 2차원 애스펙트를 캡처하는, 종이 문서상의 피처를 식별하는 상기 단계는, 수평 및 수직 단어 쌍(전체 또는 부분 단어 쌍)을 식별하는 단계를 포함할 수 있다. 다른 특징의 경우에, 종이 문서의 전자적 표시를 생성하는 상기 단계는 스캐닝 또는 인쇄 공정 중에 실행된다. 또 다른 특징의 경우에, 종이 문서의 2차원 애스펙트를 캡처하는, 종이 문서상의 피처를 식별하는 상기 단계는, 2개의 연속하는 시퀀스 사이의 수직 중첩(overlap)의 양을 체크함으로써 텍스트의 시퀀스를 논리적 라인(logical line)으로 그룹화하는 단계를 포함한다. 이 방법은 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어(query term)를 수신하는 단계와, 상기 인덱스 테이블로부터의 데이터를 기초로 하여, 상기 질의어에 잠재적으로 응답하여 적어도 하나의 혼합 미디어 문서 및 위치 가설(location hypotheses)을 계산하는 단계를 더 포함할 수 있다. 그러한 경우에, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는, 타겟 문서를 수신하는 단계와, 적어도 타겟 문서의 패치의 이미지를 생성하는 단계와, 이미지를 기초로 하나 이상의 질의어를 발생시키는 단계를 실행함으로써 실시된다. 그러한 경우에, 이미지를 기초로 하나 이상의 질의어를 발생시키는 상기 단계는, 이미지로부터 추출된 수평 및 수직 단어 쌍을 생성하는 단계를 포함한다. 다른 특징의 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 타겟 문서의 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 단계와, 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 단계를 포함한다. 그러한 경우에, 각각의 단어 쌍은 역 문서 빈도(inverse document frequency)와 연관되어 있으며, 타겟 문서의 적어도 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 상기 단계는, 각 단어 쌍에 대한 역 문서 빈도를 단어 쌍이 보이는 문서 페이지에 의해 인덱싱된 누적기에 가산하는 단계를 포함한다. 그 누적기에서의 최대값이 임계값(threshold)을 초과함에 따라, 상기 방법은 계속해서 패치에 대한 정합(match)으로서 대응 문서 페이지를 출력한다. 그러한 경우에, 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 상기 단계는, 각 단어 쌍의 주변의 존(zone)에 있는 각 셀에 가중치를 부가하는 단계[각 셀에 대한 가중치는 그 단어 쌍의 역 문서 빈도와 그 셀과 존의 중심 사이의 정규화된 기하학적 거리의 곱에 의해 결정될 수 있음]와, 최대값을 갖는 셀에 대하여 누적기의 대응 누적 어레이를 검색하는 단계를 포함한다. 최대값이 임계값을 초과함에 따라, 상기 방법은 패치의 위치로서 셀의 좌표를 보고하는 단계를 더 포함한다. 다른 특징의 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 각 질의어와 연관된 하나 이상의 위치를 검색하도록 인덱스 테이블에서 하나 이상의 질의어 각각을 탐색하는 단계와, 각각의 식별된 위치에 대하여, 위치를 포함하는 하나 이상의 후보 영역을 식별하는 단계를 포함한다. 그러한 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 하나 이상의 질의어 모두와 가장 부합하는 하나 이상의 후보 영역 중 하나를 식별하는 단계를 포함한다. 하나 이상의 후보 영역 중 하나가 사전 정의한 정합 기준(matching criteria)을 만족시키는 것으로 결정됨에 따라, 상기 방법은 계속해서 그 영역을 타겟 문서에 대하여 정합하는 것으로서 확인한다.

[0008] 본 발명의 하나 이상의 실시예의 적어도 하나의 다른 양태는, 하나 이상의 프로세서에 의해 실행될 때에, 프로세서로 하여금 혼합 미디어 문서 시스템에서의 정보를 구성하고 액세스하는 프로세스를 실행하도록 하게 하는 명령어가 인코딩된 기계 판독 가능한 매체(예컨대, 하나 이상의 콤팩트디스크, 디스켓, 서버, 메모리 스틱, 또는 하드 드라이브, ROMs, RAMs, 또는 전자 명령어를 저장하기에 적합한 기타 유형의 매체)를 제공한다. 이 프로세스는 예컨대 여기서 설명하는 방법 또는 그 방법의 변형예와 유사할 수 있다.

[0009] 본 발명의 하나 이상의 실시예의 적어도 하나의 다른 양태는, 혼합 미디어 문서 시스템에서 정보에 액세스하는 방법을 제공한다. 이 방법은, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 단계와, 혼합 미디어 문서의 피처 위치와 문서 피처를 인덱싱하는 인덱스 테이블로부터의 데이터를 기초로 하여, 상기 질의어에 잠재적으로 응답하여 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계를 포함한다. 특징의 경우에, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는, 타겟 문서를 수신하는 단계와, 적어도 타겟 문서의 패치의 이미지를 생성하는 단계와, 상기 이미지를 기초로 하나 이상의 질의어를 발생시키는 단계를 실행함으로써 진행된다. 그러한 경우에 있어서, 이미지를 기초로 하나 이상의 질의어를 발생시키는 상기 단계는, 이미지로부터 추출된 수평 및 수직 단어 쌍을 생성하는 단계를 포함한다. 다른 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 타겟 문서의 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 단계와, 패치의 중심일 가능성이 가

장 많은 그 페이지 내의 위치를 계산하는 단계를 포함한다. 다른 경우에, 각각의 단어 쌍은 역 문서 빈도와 연관되어 있으며, 타겟 문서의 적어도 패치와 정합할 가능성이 가장 높은 저장된 페이지를 찾아내는 상기 단계는, 각 단어 쌍에 대한 역 문서 빈도를 그 단어 쌍이 보이는 문서 페이지에 의해 인덱싱된 누적기에 가산하는 단계를 포함한다. 그 누적기에서의 최대값이 임계값을 초과함에 따라, 상기 방법은 계속해서 패치에 대한 정합으로서 대응 문서 페이지를 출력한다. 다른 경우에, 패치의 중심일 가능성이 가장 많은 그 페이지 내의 위치를 계산하는 상기 단계는, 각 단어 쌍의 주변의 존에 있는 각 셀에 가중치를 부가하는 단계(각 셀에 대한 가중치는 예컨대 그 단어 쌍의 역 문서 빈도와 그 셀과 존의 중심 사이의 정규화된 기하학적 거리의 곱에 의해 결정될 수 있음)와, 최대값을 갖는 셀에 대하여 누적기의 대응 누적 어레이를 검색하는 단계를 포함한다. 최대값이 임계값을 초과함에 따라, 상기 방법은 패치의 위치로서 셀의 좌표를 보고하는 단계를 포함한다. 다른 특정의 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 각 질의어와 연관된 하나 이상의 위치를 검색하도록 인덱스 테이블에서 하나 이상의 질의어 각각을 탐색하는 단계와, 각각의 식별된 위치에 대하여, 위치를 포함하는 하나 이상의 후보 영역을 식별하는 단계를 포함한다. 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 하나 이상의 질의어 모두와 가장 부합하는 하나 이상의 후보 영역 중 하나를 식별하는 단계와, 하나 이상의 후보 영역 중 하나가 사전 정의한 정합 기준을 만족시키는 것으로 결정됨에 따라, 그 영역을 타겟 문서에 대하여 정합하는 것으로서 확인하는 단계를 더 포함할 수 있다.

[0010] 본 발명의 다른 실시예는, 하나 이상의 프로세서에 의해 실행될 때, 프로세서로 하여금 혼합 미디어 문서 시스템에서의 정보에 액세스하는 프로세스를 실행하도록 하게 하는 명령어가 인코딩된 기계 판독 가능한 매체(예컨대, 하나 이상의 콤팩트디스크, 디스켓, 서버, 메모리 스틱, 또는 하드 드라이브, ROMs, RAMs, 또는 전자 명령어를 저장하기에 적합한 기타 유형의 매체)를 제공한다. 이 프로세스는 예컨대 여기서 설명하는 방법 또는 그 방법의 변형예와 유사할 수 있다.

[0011] 본 발명의 하나 이상의 실시예의 적어도 하나의 다른 양태는, 하나 이상의 프로세서에 의해 실행될 때, 프로세서로 하여금 혼합 미디어 문서 시스템에서 정보에 액세스하는 프로세스를 수행하게 하는 명령어로 인코딩된 기계-판독가능한 매체(예를 들어, 하나 이상의 콤팩트 디스크, 디스켓, 서버, 메모리 스틱, 또는 하드 드라이브, ROM, RAM, 또는 전자적 명령어를 저장하는 임의의 유형의 매체)를 제공한다. 이 프로세스는, 예를 들어, 여기에 기술된 방법과 유사하거나 그의 변형일 수 있다.

[0012] 본 명세서에 기술된 특징들 및 이점들은 모두 포함한 것이 아니며, 상세하게는 다수의 부가적인 특징들 및 이점들이 도면 및 설명을 참조하면 당업자에게 명백하게 될 것이다. 게다가, 유의할 점은 본 명세서에 사용된 표현이 주로 읽기 쉽고 도움이 되도록 선택되었으며 본 발명의 청구 대상의 범위를 제한하기 위한 것이 아니라는 것이다.

## 실시예

[0079] MMR(Mixed Media Reality) 시스템 및 연관된 방법이 기술되어 있다. MMR 시스템은 제1 미디어로서의 인쇄된 종이 및 제2 미디어로서의 디지털 사진, 디지털 영화, 디지털 오디오 파일, 디지털 텍스트 파일, 또는 웹 링크 등의 적어도 2가지 유형의 미디어를 포함하는 혼합 미디어 문서를 형성하는 메커니즘을 제공한다. MMR 시스템 및/또는 기법은 또한 혼합 미디어 문서를 제공하기 위해 휴대용 전자 장치(예를 들어, PDA 또는 셀룰러 카메라 전화) 및 종이 문서의 결합을 이용하는 다양한 영업 모델을 용이하게 해주는 데 더 사용될 수 있다.

[0080] 특정 일실시예에서, MMR 시스템은 텍스트-기반 인덱스를 사용하여 탐색(look-up)을 가능하게 해주는 방식으로 인쇄된 문서로부터 추출된 객체들 간의 2차원의 기하학적 관계를 나타내는 콘텐츠-기반 검색 데이터베이스를 포함한다. 증거 누적 기법(evidence accumulation technique)은 2차원 영역에서 피처의 발생 빈도수(frequency of occurrence)를 그의 위치의 확률과 결합시킨다. 한가지 이러한 실시예에서, MMR 데이터베이스 시스템은 MMR 피처 추출 알고리즘에 의해 계산된 기술(description)을 수신하는 인덱스 테이블(index table)을 포함한다. 인덱스 테이블은 각각의 피처가 발생하는 그 페이지들 내에서 문서, 페이지, 및 x-y 위치를 식별해준다. 증거 누적 알고리즘은 인덱스 테이블로부터의 데이터가 주어진 경우 순위가 매겨진 일련의 문서, 페이지 및 위치 가설들을 계산한다. 관계형 데이터베이스(또는 다른 적당한 저장 시설)은, 원하는 바대로, 각각의 문서, 페이지 및 위치에 관한 부가적인 특성을 저장하는 데 사용될 수 있다.

[0081] MMR 데이터베이스 시스템은 MMR 프로세서, 캡처 장치, 통신 메커니즘 및 MMR 소프트웨어를 포함하는 메모리 등의 다른 컴포넌트도 포함할 수 있다. MMR 프로세서는 또한 미디어 유형의 저장 장치 또는 소스, 입력 장치 및 출력 장치에 연결될 수 있다. 한가지 이러한 구성에서, MMR 소프트웨어는 부가적인 디지털 콘텐츠를 갖는 MMR 문서에 액세스하기 위해, MMR 문서를 생성 또는 수정하기 위해, 및 문서를 사용하여 이러한 비즈니스 거래, 데

이더 질의, 보고, 기타 등등의 다른 동작들을 수행하기 위해 MMR 프로세서에 의해 실행가능한 루틴을 포함한다.

[0082] MMR 시스템 개요

[0083] 이제, 도 1a를 참조하면, 본 발명의 일 실시예에 따른 MMR(Mixed Media Reality) 시스템(100a)이 도시되어 있다. MMR 시스템(100a)은 MMR 프로세서(102), 통신 메커니즘(104), 휴대용 입력 장치(168) 및 휴대용 출력 장치(170)를 갖는 캡처 장치(106), MMR 소프트웨어를 포함하는 메모리(108), 기본 미디어 저장 장치(base media storage)(160), MMR 미디어 저장 장치(162), 출력 장치(164), 및 입력 장치(166)를 포함한다. MMR 시스템(100a)은 기존의 인쇄된 문서로부터의 정보를, 오디오, 비디오, 텍스트, 갱신된 정보 및 서비스 등의 제2 미디어 유형(들)에 대한 인덱스로서 사용하는 방법을 제공함으로써 혼합 미디어 환경을 생성한다.

[0084] 캡처 장치(106)는 인쇄된 문서(예를 들어, 이미지, 드로잉, 또는 기타 이러한 표시)의 표시를 발생할 수 있고, 이 표시가 MMR 프로세서(102)에 전송된다. MMR 시스템(100a)은 이어서 이 표시를 MMR 문서 및 기타 제2 미디어 유형에 정합시킨다. MMR 시스템(100a)은 또한 표시의 입력 및 인식에 응답하여 동작을 취할 책임이 있다. MMR 시스템(100a)이 취하는 동작은, 예를 들어, 정보를 검색하는 것, 주문을 하는 것, 비디오 또는 사운드를 검색하는 것, 정보를 저장하는 것, 새 문서를 생성하는 것, 문서를 인쇄하는 것, 문서 또는 이미지를 디스플레이하는 것, 기타 등등을 비롯한 임의의 유형일 수 있다. 본 명세서에 기술된 콘텐츠-기반 검색 데이터베이스 기술을 사용함으로써, MMR 시스템(100a)은 인쇄된 텍스트를, 사용자에게 흥미가 있거나 귀중한 전자적 콘텐츠 또는 서비스에의 진입점을 제공하는 동적 미디어로 만드는 메커니즘을 제공한다.

[0085] MMR 프로세서(102)는 데이터 신호를 처리하고, CISC(complex instruction set computer) 아키텍처, RISC(reduced instruction set computer) 아키텍처, 또는 명령어 세트들의 조합을 구현하는 아키텍처를 비롯한 다양한 컴퓨팅 아키텍처를 포함할 수 있다. 한 특정의 실시예에서, MMR 프로세서(102)는 산술 논리 유닛(arithmetic logic unit), 마이크로프로세서, 범용 컴퓨터, 또는 본 발명의 동작들을 수행하도록 장비를 갖춘 어떤 다른 정보 가전 기기를 포함한다. 다른 실시예에서, MMR 프로세서(102)는, 예를 들어, WINDOWS 또는 UNIX 기반 운영 체제와 같은 운영 체제의 상부에서 실행되는 Java로 작성된 프로그램에 의해 발생될 수 있는 그래픽 사용자 인터페이스를 갖는 범용 컴퓨터를 포함한다. 도 1a에는 단 하나의 프로세서만이 도시되어 있지만, 다수의 프로세서가 포함될 수 있다. 이 프로세서는 MMR 메모리(108)에 연결되어 있으며 그 안에 저장된 명령어를 실행한다.

[0086] 통신 메커니즘(104)은 캡처 장치(106)를 MMR 프로세서(102)에 연결시키는 임의의 장치 또는 시스템이다. 예를 들어, 통신 메커니즘(104)은 네트워크(예를 들어, WAN 및/또는 LAN), 유선 링크(예를 들어, USB, RS232, 또는 이더넷), 무선 링크(예를 들어, 적외선, 블루투스, 또는 802.11), 모바일 장치 통신 링크(예를 들어, GPRS 또는 GSM), 공중 교환 전화망(PSTN) 링크, 또는 이들의 임의의 조합을 사용하여 구현될 수 있다. 수많은 통신 아키텍처 및 프로토콜이 여기에서 사용될 수 있다.

[0087] 캡처 장치(106)는 통신 메커니즘(104)과 인터페이스하는 송수신기 등의 수단을 포함하며, 입력 장치(168)를 통해 디지털적으로 이미지 또는 데이터를 캡처할 수 있는 임의의 장치이다. 캡처 장치(106)는 선택에 따라서는 출력 장치(170)를 포함할 수 있고, 선택에 따라서는 휴대용이다. 예를 들어, 캡처 장치(106)는 표준 셀룰러 카메라 전화, PDA 장치, 디지털 카메라, 바코드 판독기, RFID(radio frequency identification) 판독기, 표준 웹캠 등의 컴퓨터 주변 장치, 또는 PC의 비디오 카드 등의 내장된 장치이다. 캡처 장치(106a-d)의 몇가지 예에 대해 도 2a 내지 도 2d를 각각 참조하여 보다 상세히 기술한다. 그에 부가하여, 캡처 장치(106)는 콘텐츠-기반 검색을 가능하게 해주고 캡처 장치(106)를 MMR 시스템(100a, 100b)의 기반구조에 링크시키는 소프트웨어 애플리케이션을 포함할 수 있다. 도 2e를 참조하면 캡처 장치(106)의 더 많은 기능적 상세를 알 수 있다. 수많은 종래의 커스터마이징된 캡처 장치(106) 및 그들 각자의 기능 및 아키텍처는 본 개시 내용을 살펴보면 명백하게 될 것이다.

[0088] 메모리(108)는 프로세서(102)에 의해 실행될 수 있는 명령어 및/또는 데이터를 저장한다. 명령어 및/또는 데이터는 본 명세서에 기술된 기술들 중 임의의 것 및/또는 그 모두를 수행하는 코드를 포함할 수 있다. 메모리(108)는 DRAM(dynamic random access memory) 장치, SRAM(static random access memory) 장치, 또는 임의의 다른 적당한 메모리 장치일 수 있다. 메모리(108)에 대해서는 도 4를 참조하여 이하에서 보다 상세히 기술된다. 한 특정의 실시예에서, 메모리(108)는 MMR 소프트웨어 모음, 운영 체제 및 기타 애플리케이션 프로그램(예를 들어, 워드 프로세싱 애플리케이션, 전자 메일 애플리케이션, 금융 애플리케이션, 및 웹 브라우저 애플리케이션)을 포함한다.



- [0089] 기본 미디어 저장 장치(base media storage)(160)는 제2 미디어 유형을 그의 원래의 형태로 저장하기 위한 것이고, MMR 미디어 저장 장치(162)는 본 명세서에 상세히 기술된 바와 같이 MMR 환경을 생성하기 위해 MMR 문서, 데이터베이스 및 기타 정보를 저장하기 위한 것이다. 별개인 것으로 도시되어 있지만, 다른 실시예에서, 기본 미디어 저장 장치(160) 및 MMR 미디어 저장 장치(162)는 동일한 저장 장치의 일부분일 수 있거나 다른 방식으로 일체화되어 있을 수 있다. 데이터 저장 장치(160, 162)는 또한 MMR 프로세서(102)에 대한 데이터 및 명령어를 저장하고, 예를 들어, 하드 디스크 드라이브, 플로피 디스크 드라이브, CD ROM 장치, DVD ROM 장치, DVD RAM 장치, DVD RW 장치, 플래쉬 메모리 장치, 또는 임의의 다른 적당한 대용량 저장 장치를 비롯한 하나 이상의 장치를 포함한다.
- [0090] 출력 장치(164)는 MMR 프로세서(102)에 연결되어 동작하고 콘텐츠를 디스플레이하거나, 소리내거나 또는 다른 방식으로 제공하는 데이터 등의 데이터를 출력하도록 장비를 갖춘 임의의 장치를 나타낸다. 예를 들어, 출력 장치(164)는 프린터, 디스플레이 장치, 및/또는 스피커 등의 다양한 유형들 중 임의의 하나일 수 있다. 예시적인 디스플레이 출력 장치(164)는 음극선관(CRT), 액정 디스플레이(LCD), 또는 임의의 다른 유사하게 장비된 디스플레이 장치, 스크린 또는 모니터를 포함한다. 일 실시예에서, 출력 장치(164)는 터치 감응적인 투명한 패널이 출력 장치(164)의 스크린을 덮고 있는 터치 스크린을 갖추고 있다.
- [0091] 입력 장치(166)는 MMR 프로세서(102)에 연결되어 동작하며 키보드 및 커서 컨트롤러, 스캐너, 다기능 프린터, 정지 영상 또는 비디오 카메라, 키패드, 터치 스크린, 검출기, RFID 태그 판독기, 스위치, 또는 사용자가 시스템(100a)과 상호작용할 수 있게 해주는 임의의 메커니즘 등의 다양한 유형들 중 임의의 것이다. 일 실시예에서, 입력 장치(166)는 키보드 및 커서 컨트롤러이다. 커서 컨트롤러는, 예를 들어, 마우스, 트랙볼, 스타일러스, 펜, 터치 스크린 및/또는 패드, 커서 방향키, 또는 커서의 움직임을 야기하는 기타 메커니즘을 포함할 수 있다. 다른 실시예에서, 입력 장치(166)는 마이크, 범용 컴퓨터 시스템 내에서 사용하도록 설계된 오디오 애드인/확장 카드(audio add in/expansion card), 아날로그-디지털 변환기, 및 음성 인식 및/또는 오디오 처리를 용이하게 해주는 디지털 신호 처리기이다.
- [0092] 도 1b는 본 발명의 다른 실시예에 따라 구성된 MMR 시스템(100b)의 기능 블록도를 나타낸 것이다. 이 실시예에서, MMR 시스템(100b)은 MMR 컴퓨터(112)(사용자(110)에 의해 동작됨), 네트워크화된 미디어 서버(114), 및 인쇄된 문서(118)를 생성하는 프린터(116)를 포함한다. MMR 시스템(100b)은 또한 사무실 포털(120), 서비스 제공자 서버(122), 셋톱 박스(126)에 전기적으로 연결되어 있는 전자 디스플레이(124), 및 문서 스캐너(127)를 포함한다. MMR 컴퓨터(112), 네트워크화된 미디어 서버(114), 프린터(116), 사무실 포털(120), 서비스 제공자 서버(122), 셋톱 박스(126) 및 문서 스캐너(127) 간의 통신 링크는, LAN(예를 들어, 사무실 또는 홈 네트워크), WAN(예를 들어, 인터넷 또는 회사 네트워크), LAN/WAN 결합, 또는 다수의 컴퓨팅 장치가 통신할 수 있는 임의의 다른 데이터 경로일 수 있는 네트워크(128)를 통해 제공된다.
- [0093] MMR 시스템(100b)은 또한 셀룰러 기반구조(132), Wi-Fi(wireless fidelity) 기술(134), 블루투스 기술(136) 및/또는 적외선(IR) 기술(138)을 통해, 하나 이상의 컴퓨터(112), 네트워크화된 미디어 서버(114), 사용자 프린터(116), 사무실 포털(120), 서비스 제공자 서버(122), 전자 디스플레이(124), 셋톱 박스(126) 및 문서 스캐너(127)와 무선으로 통신할 수 있는 캡처 장치(106)를 포함한다. 다른 대안으로서 또는 그에 부가하여, 캡처 장치(106)는 유선 기술(140)을 통해 MMR 컴퓨터(112), 네트워크화된 미디어 서버(114), 사용자 프린터(116), 사무실 포털(120), 서비스 제공자 서버(122), 전자 디스플레이(124), 셋톱 박스(126), 및 문서 스캐너(127)와 유선 방식으로 통신을 할 수 있다. Wi-Fi 기술(134), 블루투스 기술(136), IR 기술(138) 및 유선 기술(140)이 도 1b에서 별도의 요소로서 도시되어 있지만, 이러한 기술도 역시 처리 환경(예를 들어, MMR 컴퓨터(112), 네트워크화된 미디어 서버(114), 캡처 장치(106), 기타 등등) 내에 통합될 수 있다. 그에 부가하여, MMR 시스템(110b)은 또한 서비스 제공자 서버(122) 또는 네트워크(128)와 무선 또는 유선 통신하고 있는 지리적 위치 확인 메커니즘(geo location mechanism)(142)도 더 포함한다. 이것도 역시 캡처 장치(106)에 통합될 수 있다.
- [0094] MMR 사용자(110)는 MMR 시스템(100b)을 사용하고 있는 임의의 개인이다. MMR 컴퓨터(112)는 임의의 데스크톱, 랩톱, 네트워크화된 컴퓨터, 또는 다른 이러한 처리 환경이다. 사용자 프린터(116)는, 하나 이상의 인쇄된 페이지로 이루어져 있는 종이 문서인 인쇄된 문서(118)를 생성할 수 있는 임의의 가정, 사무실, 또는 상업적 프린터이다.
- [0095] 네트워크화된 미디어 서버(114)는 네트워크(128)를 통해 MMR 시스템(110b)의 사용자에게 의해 액세스될 정보 및/또는 애플리케이션을 유지하고 있는 네트워크화된 컴퓨터이다. 한 특정의 실시예에서, 네트워크화된 미디어 서버(114)는, 텍스트 소스 파일, 웹 페이지, 오디오 및/또는 비디오 파일, 이미지 파일(예를 들어, 스틸 사진),

기타 등등의 다양한 미디어 파일이 저장되어 있는 중앙집중식 컴퓨터이다. 네트워크화된 미디어 서버(114)는, 예를 들어, Comcast Corporation의 Comcast Video-on-Demand 서버, Ricoh Innovations Inc.의 Ricoh Document Mall, 또는 Google Inc.의 Google Image 및/또는 Video 서버이다. 일반적으로 말하면, 네트워크화된 미디어 서버(114)는 캡처 장치(106)를 통해 인쇄된 문서(118)에 첨부되거나, 그와 통합되거나 또는 다른 방식으로 그와 연관될 수 있는 임의의 데이터에 액세스를 제공한다.

[0096] 사무실 포털(120)은 MMR 사용자(110)의 사무실에서 일어나는 이벤트 등의, MMR 사용자(110)의 환경에서 일어나는 이벤트를 캡처하기 위한 선택적인 메커니즘이다. 사무실 포털(120)은, 예를 들어, MMR 컴퓨터(112)와 별개인 컴퓨터이다. 이 경우에, 사무실 포털(120)은 MMR 컴퓨터(112)에 직접 연결되거나 네트워크(128)를 통해 MMR 컴퓨터(112)에 연결되어 있다. 다른 대안으로서, 사무실 포털(120)은 MMR 컴퓨터(112) 내에 내장되어 있다. 예를 들어, 사무실 포털(120)은 종래의 퍼스널 컴퓨터(PC)로 구성되고 이어서 임의의 연관된 캡처 장치(106)를 지원하는 적절한 하드웨어로 보강된다. 사무실 포털(120)은 비디오 카메라 및 오디오 녹음기 등의 캡처 장치를 포함할 수 있다. 그에 부가하여, 사무실 포털(120)은 MMR 컴퓨터(112)로부터의 데이터를 캡처 및 저장할 수 있다. 예를 들어, 사무실 포털(120)은 MMR 컴퓨터(112)에서 일어나는 이벤트 및 기능을 수신 및 모니터링할 수 있다. 그 결과, 사무실 포털(120)은 MMR 사용자(110)의 물리적 환경에서의 모든 오디오 및 비디오를 기록할 수 있고 MMR 컴퓨터(112)에서 일어나는 모든 이벤트를 기록할 수 있다. 한 특정의 실시예에서, 사무실 포털(120)은 MMR 컴퓨터(112)로부터의 이벤트, 예를 들어, 문서가 편집되고 있는 동안에 비디오 스크린 캡처를 캡처한다. 그렇게 함에 있어서, 사무실 포털(120)은 주어진 문서가 생성되고 있는 동안에 브라우징된 웹 사이트 및 참고된 기타 문서를 캡처한다. 그 정보는 나중에 MMR 사용자(110)가 그의 MMR 컴퓨터(112) 또는 캡처 장치(106)를 통해 이용가능하게 될 수 있다. 그에 부가하여, 사무실 포털(120)은 사용자가 그의 문서에 추가하는 클립들을 위한 멀티미디어 서버로서 사용될 수 있다. 게다가, 사무실 포털(120)은 종이 문서가 데스크 탑에 있는 동안에 일어나는 대화(예를 들어, 전화 또는 사무실 내), 전화를 통한 토의, 및 사무실 내에서의 소규모 회의 등의 기타 사무실 이벤트를 캡처할 수 있다. 사무실 포털(120) 상의 비디오 카메라(도시 생략)는, 캡처 장치(106)에 대해 개발된 동일한 콘텐츠-기반 검색 기술을 사용함으로써, MMR 사용자(110)의 물리적 책상 위의 종이 문서를 식별할 수 있다.

[0097] 서비스 제공자 서버(122)는 네트워크(128)를 통해 MMR 시스템(100b)의 MMR 사용자(110)에 의해 액세스될 수 있는 정보 또는 애플리케이션을 보유하는 임의의 상업적 서버이다. 상세하게는, 서비스 제공자 서버(122)는 MMR 시스템(110b)과 연관되어 있는 임의의 서비스 제공자를 나타낸다. 서비스 제공자 서버(122)는, 예를 들어, Comcast Corporation 등의 케이블 TV 제공자, Verizon Wireless 등의 셀 전화 서비스 제공자, Adelphia Communications 등의 인터넷 서비스 제공자, Sony Corporation 등의 온라인 음악 서비스 제공자, 기타 등등의 상업적 서버이지만, 이에 한정되는 것은 아니다.

[0098] 전자 디스플레이(124)는 표준 아날로그 또는 디지털 텔레비전(TV), 평면 스크린 TV, 평면 패널 디스플레이, 또는 프로젝션 시스템 등의 임의의 디스플레이 장치이지만, 이에 한정되는 것은 아니다. 셋톱 박스(126)는, 공지되어 있는 바와 같이, 위성 수신용 접시 안테나, 공중, 케이블, 네트워크 또는 전화선으로부터 들어오는 신호를 처리하는 수신기 장치이다. 셋톱 박스의 예시적인 제조업자는 Advanced Digital Broadcast이다. 셋톱 박스(126)는 전자 디스플레이(124)의 비디오 입력에 전기적으로 연결되어 있다.

[0099] 문서 스캐너(127)는 Panasonic Corporation의 KV-S2026C 완전 컬러 스캐너 등의 상업적으로 이용가능한 문서 스캐너 장치이다. 문서 스캐너(127)는 기존의 인쇄된 문서를 MMR-사용가능 문서(MMR-ready document)로 변환하는 데 사용된다.

[0100] 셀룰러 기반구조(132)는 복수의 셀 타워 및 기타 셀룰러 네트워크 상호 연결을 나타낸다. 상세하게는, 셀룰러 기반구조(132)를 사용함으로써, 양방향 음성 및 데이터 통신이 캡처 장치(106) 등의 장치 내에 포함된 무선 모뎀을 통해 핸드오프, 휴대용, 및 자동차-탑재 전화에 제공된다.

[0101] Wi-Fi 기술(134), 블루투스 기술(136) 및 IR 기술(138)은 전자 장치들 간의 무선 통신을 용이하게 해주는 기술을 나타낸다. Wi-Fi 기술(134)은, 공지되어 있는 바와 같이, 802.11 표준에 기초하고 있는 WLAN(wireless local area network) 제품과 연관되어 있는 기술이다. 블루투스 기술(136)은, 공지된 바와 같이, 셀룰러 전화, 컴퓨터 및 PDA가 단거리 무선 접속을 사용함으로써 어떻게 상호 연결되는지를 기술하는 전기 통신 산업 규격이다. IR 기술(138)은 전자 장치가 단거리 무선 신호를 통해 통신을 할 수 있게 해준다. 예를 들어, IR 기술(138)은 텔레비전 리모콘, 랩톱 컴퓨터, PDA, 및 기타 장치들에 의해 사용되는 LOS(line-of-sight) 무선 통신 매체이다. IR 기술(138)은 중간-마이크로파(mid-microwave)에서 가시광 아래까지의 스펙트럼에서 동작한다.

게다가, 하나 이상의 다른 실시예들에서, 무선 통신은 IEEE 802.15(UWB) 및/또는 802.16(WiMAX) 표준을 사용하여 지원될 수 있다.

[0102] 유선 기술(140)은 표준 이더넷 접속 또는 USB(universal serial bus) 접속 등의 임의의 유선 통신 메커니즘이다. 셀룰러 기반구조(132), Wi-Fi 기술(134), 블루투스 기술(136), IR 기술(138), 및/또는 유선 기술(140)을 사용함으로써, 캡처 장치(106)는 MMR 시스템(110b)의 임의의 또는 모든 전자 장치들과 양방향으로 통신을 할 수 있다.

[0103] 지리적 위치 확인 메커니즘(142)은 지리적 위치를 확인하기에 적합한 임의의 메커니즘이다. 공지된 바와 같이, 지리적 위치 확인 메커니즘(142)은, 예를 들어, 위치 데이터베이스를 지상 GPS 수신기 장치에 제공하는 GPS 위성이다. 도 1b에 도시되어 있는 이 예시적인 실시예에서, 위치 데이터베이스는 GPS 위성에 의해 GPS 수신기(도시 생략)와 함께 네트워크(128)에 연결되어 있는 서비스 제공자 서버(122)를 통해 MMR 시스템(100b)의 사용자들에게 제공된다. 다른 대안으로서, 지리적 위치 확인 메커니즘(142)은 삼각측량 메커니즘(triangulation mechanism)을 제공하는 일련의 셀 타워(예를 들어, 셀룰러 기반구조(132)의 서브셋), 셀 타워 ID(identification) 메커니즘, 및/또는 지리적 위치를 판정하는 수단으로서의 향상된 911 서비스이다. 다른 대안으로서, 지리적 위치 확인 메커니즘(142)은 WiFi 액세스 포인트 또는 블루투스 장치의 기지의 위치로부터의 신호 세기 측정에 의해 제공된다.

[0104] 동작을 설명하면, 캡처 장치(106)는 MMR 사용자(110)가 가지고 있는 클라이언트로서 기능한다. 콘텐츠-기반 검색 동작을 가능하게 해주고 셀룰러 기반구조(132), Wi-Fi 기술(134), 블루투스 기술(136), IR 기술(138), 및/또는 유선 기술(140)을 통해 캡처 장치(106)를 MMR 시스템(110b)의 기반구조에 링크시키는 소프트웨어 애플리케이션이 이 장치 상에 존재한다. 그에 부가하여, 인쇄 캡처 동작, 이벤트 캡처 동작(예를 들어, 문서의 편집 이력을 저장), 서버 동작(예를 들어, 나중에 다른 사람에게 서비스하기 위해 MMR 컴퓨터(112) 상에 저장된 데이터 및 이벤트) 또는 프린터 관리 동작(예를 들어, 프린터(116)가 문서 레이아웃 및 멀티미디어 클립들 등의 MMR에 필요한 데이터를 큐잉하도록 설정될 수 있음)(이에 한정되지 않음) 등의 몇가지 동작을 수행하는 소프트웨어 애플리케이션이 MMR 컴퓨터(112) 상에 존재한다. 네트워크화된 미디어 서버(114)는 MMR 사용자(110)에게 속한 MMR 컴퓨터(112)를 통해 인쇄되는 인쇄된 문서(118) 등의 인쇄된 문서에 첨부된 데이터에의 액세스를 제공한다. 그렇게 함에 있어서, 비디오 또는 오디오 등의 제2 미디어가 종이 문서 등의 제1 미디어와 연관된다. 제2 미디어의 제1 미디어에 대한 연관 관계를 형성하는 소프트웨어 애플리케이션 및/또는 메커니즘에 대한 더 많은 상세는 이하에서 도 2e, 도 3, 도 4 및 도 5를 참조하여 기술된다.

#### [0105] 캡처 장치

[0106] 도 2a, 도 2b, 도 2c 및 도 2d는 본 발명의 실시예들에 따른 예시적인 캡처 장치(106)를 나타낸 것이다. 보다 구체적으로는, 도 2a는 셀룰러 카메라 전화인 캡처 장치(106a)를 나타낸 것이다. 도 2b는 PDA 장치인 캡처 장치(106b)를 나타낸 것이다. 도 2c는 컴퓨터 주변 장치인 캡처 장치(106c)를 나타낸 것이다. 컴퓨터 주변 장치의 일례는 임의의 표준 웹캠이다. 도 2d는 (예를 들어, MMR 컴퓨터(112) 등의) 컴퓨팅 장치에 내장되어 있는 캡처 장치(106d)를 나타낸 것이다. 예를 들어, 캡처 장치(106d)는 컴퓨터 그래픽 카드이다. 캡처 장치(106)의 예시적인 상세는 도 2e를 참조하면 알 수 있다.

[0107] 캡처 장치(106a, 106b)의 경우에, 캡처 장치(106)는 MMR 사용자(110)가 가지고 있을 수 있으며, 그의 물리적 위치는 지리적 위치 확인 메커니즘(142)에 의해 또는 셀룰러 기반구조(132) 내의 각각의 셀 타워의 ID 번호에 의해 추적될 수 있다.

[0108] 이제 도 2e를 참조하면, 본 발명에 따른 캡처 장치(106)의 일 실시예에 대한 기능 블록도가 도시되어 있다. 캡처 장치(106)는 프로세서(210), 디스플레이(212), 키패드(214), 저장 장치(216), 무선 통신 링크(218), 유선 통신 링크(220), MMR 소프트웨어 모음(222), 캡처 장치 사용자 인터페이스(UI)(224), 문서 지문 정합 모듈(226), 제3자 소프트웨어 모듈(228), 및 다양한 캡처 메커니즘들(230) 중 적어도 하나의 메커니즘을 포함한다. 예시적인 캡처 메커니즘(230)은 비디오 카메라(232), 정지 영상 카메라(234), 음성 녹음기(236), 전자적 하이라이터(electronic highlighter)(238), 레이저(240), GPS 장치(242) 및 RFID 판독기(244)를 포함하지만, 이에 한정되는 것은 아니다.

[0109] 프로세서(210)는 Intel Corporation에 의해 제조된 펜티엄 마이크로프로세서(이에 한정되지 않음) 등의 중앙 처리 장치(CPU)이다. 디스플레이(212)는 핸드헬드 전자 장치에서 사용되는 것 등의 임의의 표준 비디오 디스플레이 메커니즘이다. 보다 상세하게는, 디스플레이(212)는, 예를 들어, 액정 디스플레이(LCD) 또는 유기 발광 다이오드(OLED) 디스플레이 등의 임의의 디지털 디스플레이이다. 키패드(214)는 표준 컴퓨팅 장치 및 셀룰러 전

화 등의 핸드헬드 전자 장치에서 사용되는 키패드 등의 임의의 표준 영숫자 입력 메커니즘이다. 저장 장치(216)는, 공지된 바와 같이, 하드 디스크 드라이브 또는 랜덤 액세스 메모리(RAM) 장치 등의 임의의 휘발성 또는 비휘발성 메모리 장치이다.

- [0110] 무선 통신 링크(218)는, 공지된 바와 같이, 직접 포인트-투-포인트 통신(direct point-to-point communication) 또는 액세스 포인트(도시 생략) 및 LAN(예를 들어, IEEE 802.11 Wi-Fi 또는 블루투스 기술)을 통한 무선 통신을 제공하는 무선 데이터 통신 메커니즘이다. 무선 통신 링크(220)는, 예를 들어, 표준 이더넷 및/또는 USB 접속을 통해 직접 통신(direct communication)을 제공하는 유선 데이터 통신 메커니즘이다.
- [0111] MMR 소프트웨어 모듈(222)은 한 유형의 미디어와 제2 유형을 병합하는 것 등의 MMR 동작을 수행하는 전체적인 관리 소프트웨어이다. MMR 소프트웨어 모듈(222)에 대한 더 많은 상세는 도 4를 참조하면 알 수 있다.
- [0112] 캡처 장치 사용자 인터페이스(UI)(224)는 캡처 장치(106)를 동작시키는 사용자 인터페이스이다. 캡처 장치 UI(224)를 사용함으로써, 그 상에서 기능들을 선택하기 위한 다양한 메뉴가 MMR 사용자(110)에게 제공된다. 보다 구체적으로는, 캡처 장치 UI(224)의 메뉴는 MMR 사용자(110)가, 종이 문서와 상호작용하는 것, 기존의 문서로부터 데이터를 판독하는 것, 기존의 문서에 데이터를 기록하는 것, 그들 문서와 연관되어 있는 증강 현실(argumented reality)을 보면서 그와 상호작용하는 것, 사용자의 MMR 컴퓨터(112) 상에 디스플레이되는 문서와 연관되어 있는 증강 현실을 보면서 그와 상호작용하는 것(이에 한정되지 않음) 등의 작업들을 관리할 수 있게 해준다.
- [0113] 문서 지문 정합 모듈(226)은 캡처 장치(106)의 적어도 하나의 캡처 메커니즘(230)을 통해 캡처된 텍스트 이미지로부터 피쳐들을 추출하는 소프트웨어 모듈이다. 문서 지문 정합 모듈(226)은 또한 캡처된 이미지와 문서들의 데이터베이스 간의 패턴 정합을 수행할 수 있다. 가장 기본 레벨에서, 일 실시예에 따르면, 문서 지문 정합 모듈(226)은 더 큰 페이지 이미지 내에서 이미지 패치의 위치를 결정하고, 여기서 그 페이지 이미지는 문서들의 큰 컬렉션으로부터 선택된다. 문서 지문 정합 모듈(226)은 캡처된 데이터를 수신하고 캡처된 데이터로부터 이미지의 표시를 추출하고, 문서들 내에서 패치 인식 및 움직임 분석을 수행하며, 결정 조합들을 수행하고, 입력 이미지들이 위치하고 있는 페이지들 내에서의 x-y 장소들의 리스트를 출력하는 루틴 또는 프로그램을 포함한다. 예를 들어, 문서 지문 정합 모듈(226)은, 피쳐들이 추출된 문서 및 문서 내의 일부분을 식별하기 위해, 텍스트의 일부의 이미지로부터 추출되는 수평 및 수직 피쳐들을 결합시키는 알고리즘일 수 있다. 피쳐들이 추출되었으면, 기호적인 문서(symbolic document)를 식별하기 위해, 예를 들어, MMR 컴퓨터(112) 또는 네트워크화된 미디어 서버(114)에 존재하는 인쇄된 문서 인덱스(도시 생략)가 질의된다. 캡처 장치 UI(224)의 제어 하에서, 문서 지문 정합 모듈(226)은 인쇄된 문서 인덱스에 액세스한다. 인쇄된 문서 인덱스에 대해서는 도 3의 MMR 컴퓨터(112)를 참조하여 더 상세히 기술한다. 유의할 점은, 대안의 실시예에서, 문서 지문 정합 모듈(226)이 MMR 컴퓨터(112)의 일부일 수 있으며 캡처 장치(106) 내에 위치하지 않을 수 있다는 것이다. 이러한 실시예에서, 캡처 장치(106)는 미가공 캡처 데이터(raw captured data)를, 이미지 추출, 패턴 정합, 그리고 문서 및 위치 인식을 위해 MMR 컴퓨터(112)에 전송한다. 또다른 실시예에서, 문서 지문 정합 모듈(226)은 피쳐 추출만을 수행하고, 추출된 피쳐들이 패턴 정합 및 인식을 위해 MMR 컴퓨터(112)에 전송된다.
- [0114] 제3자 소프트웨어 모듈(228)은 캡처 장치(106)에서 일어날 수 있는 임의의 동작을 향상시키기 위한 임의의 제3자 소프트웨어 모듈을 나타낸다. 예시적인 제3자 소프트웨어는 보안 소프트웨어, 이미지 감지 소프트웨어, 이미지 처리 소프트웨어, 및 MMR 데이터베이스 소프트웨어를 포함한다.
- [0115] 상기한 바와 같이, 캡처 장치(106)는 임의의 수의 캡처 메커니즘(230)을 포함할 수 있으며, 이들의 예에 대해 이제부터 기술한다.
- [0116] 비디오 카메라(232)는 표준 디지털 카메라 또는 어떤 셀 전화에서 발견되는 것 등의 디지털 비디오 레코딩 장치이다.
- [0117] 정지 영상 카메라(234)는 디지털 이미지를 캡처할 수 있는 임의의 표준 디지털 카메라 장치이다.
- [0118] 음성 녹음기(236)는 오디오 신호를 캡처하고 이를 디지털 형태로 출력할 수 있는 임의의 표준 오디오 녹음 장치(마이크 및 연관된 하드웨어)이다.
- [0119] 전자적 하이라이터(238)는 인쇄된 텍스트, 바코드 및 작은 이미지를 스캔하여 저장하고 PC, 랩톱 컴퓨터, 또는 PDA 장치에 전송하는 기능을 제공하는 전자적 하이라이터이다. 전자적 하이라이터(238)는 정보가 펜 상에 저장되거나 직렬 포트, 적외선 통신 또는 USB 어댑터를 통해 컴퓨터 애플리케이션에 직접 전송될 수 있게 해주는,



예를 들어, Wizcom Technologies의 Quicklink Pen Handheld Scanner이다.

- [0120] 레이저(240)는, 공지된 바와 같이, 자극 방출(stimulated emission)을 통해, 코히런트하고 거의 단색인 광(coherent, near-monochromatic light)을 생성하는 광원이다. 레이저(240)는, 예를 들어, 순방향 바이어스될 때 코히런트 광을 방출하는 반도체 장치인 표준 레이저 다이오드이다. 레이저(240)가 향하고 있는 이미지에 의해 반사되는 광량을 측정하는 검출기가 레이저(240)와 연관되고 그 안에 포함되어 있다.
- [0121] GPS 장치(242)는 위치 데이터, 예를 들어, 디지털 위도 및 경도 데이터를 제공하는 임의의 휴대용 GPS 수신기 장치이다. 휴대용 GPS 장치(242)의 예는 Sony Corporation의 NV-U70 Portable Satellite Navigation System, 및 Thales North America, Inc.의 Magellan brand RoadMate Series GPS 장치, Meridian Series GPA 장치, 및 eXplorist Series GPS 장치이다. 공지된 바와 같이, GPS 장치(242)는 부분적으로 삼각 측량에 의해 실시간으로 캡처 장치(106)의 위치를 결정하는 방법을 복수의 지리적 위치 확인 메커니즘(142)에 제공한다.
- [0122] RFID 판독기(244)는 Texas Instruments에 의해 제조된 TI RFID 시스템 등의 상업적으로 이용가능한 RFID 태그 판독기 시스템이다. RFID 태그는 무선파를 사용하여 고유의 아이টে를 식별하는 무선 장치이다. 공지된 바와 같이, RFID 태그는 안테나에 부착되어 있는 마이크로칩(고유의 디지털 ID 번호가 저장되어 있음)으로 형성된다.
- [0123] 한 특정의 실시예에서, 캡처 장치(106)는 프로세서(210), 디스플레이(212), 키패드(214), 저장 장치(216), 무선 통신 링크(218), 유선 통신 링크(220), MMR 소프트웨어 모음(222), 캡처 장치 UI(224), 문서 지문 정합 모듈(226), 제3자 소프트웨어 모듈(228) 및 캡처 메커니즘들(230) 중 적어도 하나의 메커니즘을 포함한다. 그렇게 함에 있어서, 캡처 장치(106)는 완전-기능 장치(full-function device)이다. 다른 대안으로서, 캡처 장치(106)는 더 적은 기능을 가질 수 있으며, 따라서 제한된 일련의 기능 컴포넌트를 포함할 수 있다. 예를 들어, MMR 소프트웨어 모음(222) 및 문서 지문 정합 모듈(226)은 원격지에, 예를 들어, MMR 시스템(100b)의 MMR 컴퓨터(112) 또는 네트워크화된 미디어 서버(114)에 존재할 수 있으며 무선 통신 링크(218) 또는 유선 통신 링크(220)를 통해 캡처 장치(106)에 의해 액세스된다.
- [0124] MMR 컴퓨터
- [0125] 이제 도 3을 참조하면, 본 발명의 일 실시예에 따라 구성된 MMR 컴퓨터(112)가 도시되어 있다. 알 수 있는 바와 같이, MMR 컴퓨터(112)는 하나 이상의 멀티미디어(MM) 파일(336)을 포함하는 네트워크화된 미디어 서버(114), 인쇄된 문서(118)를 생성하는 사용자 프린터(116), 문서 스캐너(127), 및 캡처 장치 UI(224) 및 제1 문서 지문 정합 모듈 인스턴스(226)를 포함하는 캡처 장치(106)에 연결되어 있다. 이들 컴포넌트 간의 통신 링크는 직접 링크이거나 네트워크를 통할 것이다. 그에 부가하여, 문서 스캐너(127)는 제2 문서 지문 정합 모듈 인스턴스(226')를 포함한다.
- [0126] 이 예시적인 실시예의 MMR 컴퓨터(112)는 하나 이상의 소스 파일(310), 제1 소스 문서(SD) 브라우저(312), 제2 SD 브라우저(314), 프린터 드라이버(316), 인쇄된 문서(PD) 캡처 모듈(318), PD 인덱스(322)를 저장하는 문서 이벤트 데이터베이스(320), 이벤트 캡처 모듈(324), 문서 파서 모듈(326), 멀티미디어(MM) 클립 브라우저/편집기 모듈(328), MM에 대한 프린터 드라이버(330), DVP(document-to-video paper) 인쇄 시스템(332) 및 비디오 종이 문서(334)를 포함한다.
- [0127] 소스 파일(310)은 문서(또는 그의 일부분)의 전자적 표시인 임의의 소스 파일을 나타낸다. 예시적인 소스 파일(310)은 MMR 컴퓨터(112)의 하드 드라이브(또는 기타 적당한 저장 장치) 상에 저장되는, 하이퍼텍스트 마크업 언어(HTML) 파일, 마이크로소프트 워드 파일, 마이크로소프트 파워포인트 파일, 간단한 텍스트 파일, PDF(portable document format) 파일, 기타 등등을 포함한다.
- [0128] 제1 SD 브라우저(312) 및 제2 SD 브라우저(314)는 독립형 PC 애플리케이션 또는 소스 파일(310)과 연관되어 있는 데이터에의 액세스를 제공하는 플러그-인 중 어느 하나이다. 제1 및 제2 SD 브라우저(312, 314)는 MMR 컴퓨터(112) 상에 디스플레이하기 위한 원래의 HTML 파일 또는 MM 클립을 검색하는 데 사용될 수 있다.
- [0129] 공지된 바와 같이, 프린터 드라이버(316)는 애플리케이션들과 임의의 특정의 프린터에 의해 사용되는 페이지-기술 언어 또는 프린터 제어 언어 간의 통신 링크를 제어하는 프린터 드라이버 소프트웨어이다. 상세하게는, 인쇄된 문서(118) 등의 문서가 인쇄될 때마다, 프린터 드라이버(316)는 Ricoh Corporation에 의해 그의 인쇄 장치에 대해 제공된 것 등의 정확한 제어 명령을 갖는 데이터를 프린터(116)에 피드(feed)한다. 일 실시예에서, 프린터 드라이버(316)는 모든 인쇄된 페이지 상의 모든 문자의 x-y 좌표, 폰트, 및 포인트 크기의 표시를 자동적으로 캡처한다는 점에서 종래의 프린터 드라이버와 다르다. 환언하면, 이 드라이버는 인쇄된 모든 문서의 컨텐



츠에 관한 정보를 캡처하고 그 데이터를 PD 캡처 모듈(318)로 피드백한다.

[0130] PD 캡처 모듈(318)은 인쇄된 페이지 상의 문자 및 그래픽의 레이아웃이 검색될 수 있도록 문서들의 인쇄된 표시를 캡처하는 소프트웨어 애플리케이션이다. 그에 부가하여, PD 캡처 모듈(318)을 사용함으로써, 문서의 인쇄된 표시가 인쇄 시에 실시간으로 자동 캡처된다. 보다 구체적으로, PD 캡처 모듈(318)은 인쇄된 페이지 상의 2차원 텍스트 배열을 캡처하고 이 정보를 PD 인덱스(322)에 전송하는 소프트웨어 루틴이다. 일 실시예에서, PD 캡처 모듈(318)은 인쇄된 페이지 상의 모든 문자의 Windows 텍스트 레이아웃 명령(text layout command)을 트랩(trap)하는 것으로 동작한다. 텍스트 레이아웃 명령은 인쇄된 페이지 상의 모든 문자의 x-y 위치는 물론 폰트, 포인트 크기, 기타 등등을 운영 체제(OS)에 알려준다. 본질적으로, PD 캡처 모듈(318)은 프린터(116)에 전송되는 인쇄 데이터를 도청한다(eavesdrop). 도시된 예에서, PD 캡처 모듈(318)은 데이터를 캡처하기 위해 제1 SD 브라우저(312)의 출력부에 연결되어 있다. 다른 대안으로서, PD 캡처 모듈(318)의 기능은 프린터 드라이버(316) 내에 직접 구현될 수 있다. 이 개시 내용을 살펴보면 다양한 구성이 명백하게 될 것이다.

[0131] 문서 이벤트 데이터베이스(320)는 본 발명의 일 실시예에 따른 인쇄된 문서와 이벤트 간의 관계를 저장하도록 수정된 임의의 표준 데이터베이스이다. (문서 이벤트 데이터베이스(320)는 도 3a를 참조하여 MMR 데이터베이스로서 이하에 더 기술되어 있다.) 예를 들어, 문서 이벤트 데이터베이스(320)는 소스 파일(310)(예를 들어, Word, HTML, PDF 파일)에서 인쇄된 문서(118)와 연관되어 있는 이벤트로의 양방향 링크를 저장한다. 예시적인 이벤트는 Word 문서가 인쇄된 직후에 캡처 장치(106)에서 멀티미디어 클립을 캡처하는 것, 캡처 장치(106)의 클라이언트 애플리케이션으로 문서에 멀티미디어를 추가하는 것, 또는 멀티미디어 클립에 대한 주석 첨부를 포함한다. 그에 부가하여, 문서 이벤트 데이터베이스(320)에 저장될 수 있는 소스 파일(310)과 연관되어 있는 기타 이벤트는, 주어진 소스 파일(310)이 열릴 때, 닫힐 때, 또는 제거될 때를 로깅하는 것, 주어진 소스 파일(310)이 MMR 컴퓨터(112)의 바탕화면 상의 활성 애플리케이션에 있을 때, 문서 "복사" 및 "이동" 동작의 로깅 시간 및 목적지를 로깅하는 것, 및 주어진 소스 파일(310)의 편집 이력을 로깅하는 것을 포함한다. 이러한 이벤트는 이벤트 캡처 모듈(324)에 의해 캡처되고 문서 이벤트 데이터베이스(320)에 저장된다. 문서 이벤트 데이터베이스(320)는 소스 파일(310)을 수신하기 위해 이벤트 캡처 모듈(324), PD 캡처 모듈(318) 및 스캐너(127)의 출력에 연결되어 있고, 또한 질의 및 데이터를 수신하고 출력을 제공하기 위해 캡처 장치(106)에 연결되어 있다.

[0132] 문서 이벤트 데이터베이스(320)는 또한 PD 인덱스(322)도 저장한다. PD 인덱스(322)는 인쇄된 문서의 이미지로부터 추출되는 피쳐들을 그의 기호적인 형상으로(예를 들어, 스캔된 이미지를 Word에) 매핑하는 소프트웨어 애플리케이션이다. 일 실시예에서, PD 캡처 모듈(318)은 인쇄된 페이지 상의 모든 문자의 x-y 위치는 물론 폰트, 포인트 크기, 기타 등등을 PD 인덱스(322)에 제공한다. PD 인덱스(322)는 주어진 문서가 인쇄될 때 구성된다. 그렇지만, 나중에 조사될 수 있는 방식으로 모든 인쇄 데이터가 캡처되고 PD 인덱스(322)에 저장된다. 예를 들어, 인쇄된 문서(118)가 페이지 상에서 물리적으로 단어 "rose"보다 한 줄 위쪽에 위치한 단어 "garden"를 포함하는 경우, PD 인덱스(322)는 이러한 질의(즉, 단어 "rose" 위쪽의 단어 "garden")를 지원한다. PD 인덱스(322)는 어느 문서, 어느 페이지 및 그 페이지 내의 어느 위치에서 단어 "garden"이 단어 "rose"보다 위쪽에 나타나는지의 기록을 포함한다. 따라서, PD 인덱스(322)는 피쳐-기반 또는 텍스트-기반 질의를 지원하도록 구성되어 있다. 인쇄된 문서의 전자적 표시인 PD 인덱스(322)의 콘텐츠는 인쇄 동안에 PD 캡처 모듈(318)을 사용함으로써 및/또는 스캔 동작 동안에 문서 스캐너(127)의 문서 지문 정합 모듈(226')을 사용함으로써 발생된다. 데이터베이스(320) 및 PD 인덱스(322)의 부가적인 아키텍처 및 기능성에 대해 도 34a 내지 도 34c, 도 35 및 도 36을 참조하여 이하에서 기술한다.

[0133] 이벤트 캡처 모듈(324)은 주어진 인쇄된 문서(118) 및/또는 소스 파일(310)과 연관되어 있는 이벤트를 MMR 컴퓨터(112) 상에서 캡처하는 소프트웨어 애플리케이션이다. 이들 이벤트는 주어진 소스 파일(310)의 라이프사이클 동안에 캡처되고 문서 이벤트 데이터베이스(320)에 저장된다. 특정한 예에서, 이벤트 캡처 모듈(324)을 사용함으로써, MMR 컴퓨터(112)의 제1 SD 브라우저(312) 등의 브라우저에서 활성인 HTML 파일에 관련되어 있는 이벤트들이 캡처된다. 이들 이벤트는 HTML 파일이 MMR 컴퓨터(112) 상에 디스플레이된 시간 또는 HTML 파일이 디스플레이 또는 인쇄된 것과 동시에 열려 있는 기타 문서들의 파일 이름을 포함할 수 있다. 이러한 이벤트 정보는, 예를 들어, MMR 사용자(110)가 HTML 파일이 디스플레이 또는 인쇄된 것과 동시에 그가 어느 문서를 보고 있는지 또는 작업하고 있었는지를 (나중에) 알고자 하는 경우에 유용하다. 이벤트 캡처 모듈(324)에 의해 캡처되는 예시적인 이벤트는 문서 편집 이력, 주어진 소스 파일(310)이 바탕화면 상에 있었을 즈음에 일어났던 사무실 회의로부터의 비디오(예를 들어, 사무실 포털(120)에 의해 캡처됨), 및 주어진 소스 파일(310)이 열려 있을 때 행해졌던 전화 통화(예를 들어, 사무실 포털(120)에 의해 캡처됨)를 포함한다.

[0134] 이벤트 캡처 모듈(324)의 예시적인 기능은, 1) 추적 - 활성 파일 및 애플리케이션을 추적하는 것 -, 2) 키 스트

로크 캡처링 - 키 스트로크 캡처 및 활성 애플리케이션과 연관시키는 것 -, 3) 프레임 버퍼 캡처링 및 인덱싱 - 각각의 프레임 버퍼 이미지가 프레임 버퍼 데이터의 광학 문자 인식(OCR) 결과로 인덱싱되고, 따라서 인쇄된 문서의 일부분은 그것이 스크린 상에 디스플레이될 때까지 정합될 수 있다. 다른 대안으로서, PC 운영 체제에 의해 발행되는 PC 바탕화면에 대한 텍스트 드로잉 명령(text drawing command)을 트랩하는 그래픽 디스플레이 인터페이스(GDI) 새도우 d11로 텍스트가 캡처될 수 있다. MMR 사용자(110)는 캡처 장치(106)를 문서로 향하게 하고 그것이 MMR 컴퓨터(112)의 바탕화면 상에서 활성이었던 때를 결정할 수 있다 -, 및 4) 판독 이력 캡처 (reading history capture) - 얼마나 오래동안 또한 특정의 문서의 어느 부분이 MMR 사용자(110)에게 보였는지를 추적하기 위해, 프레임 버퍼 캡처 및 인덱싱의 데이터가 문서들이 그의 MMR 컴퓨터(112)의 바탕화면 상에서 활성이었던 때의 분석과 링크된다 - 를 포함한다. 그렇게 함에 있어서, MMR 사용자(110)가 문서를 읽고 있었는지를 추론하기 위해, 키스트로크 또는 마우스 움직임 등의 다른 이벤트들과 상관이 행해질 수 있다.

[0135] 문서 이벤트 데이터베이스(320), PD 인덱스(322) 및 이벤트 캡처 모듈(324)의 조합이 MMR 컴퓨터(112) 상에 로컬적으로 구현되거나, 다른 대안으로서, 공유 데이터베이스로서 구현된다. 로컬적으로 구현되는 경우, 공유 방식으로 구현되는 것과 비교하여 더 적은 보안이 요구된다.

[0136] 문서 파서 모듈(document parser module)(326)은 그 안에서 URL(uniform resource locator), 주소, 타이틀, 저작자, 시간, 또는 예를 들어, Hallidie Building 등의 위치를 표시하는 문구, 유용한 객체의 위치를 찾기 위해 각자의 인쇄된 문서(118)와 관계되어 있는 소스 파일(310)을 파싱하는 소프트웨어 애플리케이션이다. 그렇게 함에 있어서, 소스 파일(310)의 인쇄 버전에서의 그 객체들의 위치가 결정된다. 문서 파서 모듈(326)의 출력은 그 다음에 부가적인 정보로 문서(118)의 설명 체시를 보강하고 패턴 정합의 정확도를 향상시키기 위해 수신 장치에 의해 사용될 수 있다. 게다가, 수신 장치는 또한 URL의 경우에, URL과 연관된 웹 페이지를 검색하는 등, 위치들을 사용하여 동작을 취할 수 있다. 문서 파서 모듈(326)은 소스 파일(310)을 수신하도록 연결되어 있고 그의 출력을 문서 지문 정합 모듈(226)에 제공한다. 비록 캡처 장치의 문서 지문 정합 모듈(226)에만 연결되어 있는 것으로 도시되어 있지만, 문서 파서 모듈(326)의 출력은, 그들이 어디에 있든간에, 모든 또는 임의의 수의 문서 지문 정합 모듈(226)에 연결될 수 있다. 게다가, 문서 파서 모듈(326)의 출력은 또한 나중에 사용하기 위해 문서 이벤트 데이터베이스(320)에 저장될 수 있다.

[0137] MM 클립 브라우저/편집기 모듈(328)은 저작 기능(authoring function)을 제공하는 소프트웨어 애플리케이션이다. MM 클립 브라우저/편집기 모듈(328)은 독립형 소프트웨어 애플리케이션이거나, 다른 대안으로서, 문서 브라우저 상에서 실행되는 플러그-인(제2 SD 브라우저(314)로의 점선으로 표시됨)이다. MM 클립 브라우저/편집기 모듈(328)은 사용자에게 멀티미디어 파일을 디스플레이하며 멀티미디어 파일(336)을 수신하기 위해 네트워크화된 미디어 서버에 연결되어 있다. 그에 부가하여, MMR 사용자(110)가 문서를 작성하고 있을 때 (예를 들어, 멀티미디어 클립을 종이 문서에 첨부하고 있을 때), MM 클립 브라우저/편집기 모듈(328)은 이 기능을 위한 지원 도구이다. MM 클립 브라우저/편집기 모듈(328)은 멀티미디어가 캡처되었을 즈음에 인쇄되는 문서로부터 파싱된 정보 등의 메타데이터를 보여주는 애플리케이션이다.

[0138] MM에 대한 프린터 드라이버(330)는 MMR 문서를 작성하는 기능을 제공한다. 예를 들어, MMR 사용자(110)는 MM에 대한 프린터 드라이버(330)에 의해 발생된 UI에서 텍스트를 하이라이트하고, 멀티미디어 데이터를 검색하는 것 또는 네트워크(128) 또는 MMR 컴퓨터(112) 상에서 어떤 다른 프로세스를 실행하는 것을 포함하는 동작들을 텍스트에 추가할 수 있다. MM에 대한 프린터 드라이버(330) 및 DVP 인쇄 시스템(332)의 결합이 바코드를 사용하는 대안의 출력 형식을 제공한다. 이 형식은 콘텐츠-기반 검색 기술을 꼭 필요로 하지는 않는다. MM에 대한 프린터 드라이버(330)는 비디오 종이 기술, 즉 비디오 종이(334)를 지원하는 프린터 드라이버이다. MM에 대한 프린터 드라이버(330)는 멀티미디어에 액세스하는 방법으로서 바코드를 포함하는 종이 표시(paper representation)을 생성한다. 이와 반대로, 프린터 드라이버(316)는 멀티미디어에 액세스하는 방법으로서 MMR 기술을 포함하는 종이 표시를 생성한다. MM 클립 브라우저/편집기 모듈(328) 및 SD 브라우저(314)의 결합에서 구현되는 저작 기술은 SD 브라우저(312)와 동일한 출력 형식을 생성할 수 있으며, 따라서 콘텐츠-기반 검색에 대한 준비가 된 MMR 문서의 생성을 가능하게 해준다. DVP 인쇄 시스템(332)은 문서와 연관되어 있는 문서 이벤트 데이터베이스(320) 내의 임의의 데이터를, 명시적인 또는 암시적인 바코드를 갖는 그의 인쇄된 표시와 링크시키는 동작을 수행한다. 암시적인 바코드는 바코드처럼 사용되는 텍스트 피처의 패턴을 말한다.

[0139] 비디오 종이(334)는 종이 등의 인쇄가능한 매체상에 오디오-비주얼 정보를 제공하는 기술이다. 비디오 종이에서, 바코드는 컴퓨터에 저장되거나 액세스가능한 전자 콘텐츠에 대한 인덱스로서 사용된다. 사용자는 바코드를 스캔하고, 텍스트와 관계된 비디오 클립 또는 기타 멀티미디어 콘텐츠가 시스템에 의해 출력된다. 오디오 또는 비디오 종이를 인쇄하는 시스템이 있으며, 이들 시스템은 본질적으로 멀티미디어 정보에 대한 종이-기반 인터페

이스를 제공한다.

- [0140] 네트워크화된 미디어 서버(114)의 MM 파일(336)은 다양한 파일 유형 및 파일 형식 중 임의의 것의 컬렉션을 나타낸다. 예를 들어, MM 파일(336)은 텍스트 소스 파일, 웹 페이지, 오디오 파일, 비디오 파일, 오디오/비디오 파일, 및 이미지 파일(예를 들어, 정지 영상 사진)이다.
- [0141] 도 1b에 기술된 바와 같이, 문서 스캐너(127)는 기존의 인쇄된 문서를 MMR-준비 문서로 변환하는 데 사용된다. 그렇지만, 계속하여 도 3을 참조하면, 문서 스캐너(127)는 문서 지문 정합 모듈(226')의 피쳐 추출 동작을 스캔되는 문서의 모든 페이지에 적용함으로써 기존의 문서를 MMR-지원하도록 하는 데 사용된다. 그 다음에, PD 인덱스(322)는 스캐닝 및 피쳐 추출 동작의 결과로 채워지고, 따라서 스캔된 문서의 전자적 표시가 문서 이벤트 데이터베이스(320)에 저장된다. 이후, PD 인덱스(322) 내의 정보는 MMR 문서를 작성하는 데 사용될 수 있다.
- [0142] 계속하여 도 3을 참조하여, MMR 컴퓨터(112)의 소프트웨어 기능이 MMR 컴퓨터(112)에만 한정되지 않는다는 것에 유의한다. 다른 대안으로서, 도 3에 도시된 소프트웨어 기능은 임의의 사용자-정의된 구성으로 MMR 시스템(100b)의 MMR 컴퓨터(112), 네트워크화된 미디어 서버(114), 서비스 제공자 서버(122) 및 캡처 장치(106) 간에 분산될 수 있다. 예를 들어, 소스 파일(310), SD 브라우저(312), SD 브라우저(314), 프린터 드라이버(316), PD 캡처 모듈(318), 문서 이벤트 데이터베이스(320), PD 인덱스(322), 이벤트 캡처 모듈(324), 문서 파서 모듈(326), MM 클립 브라우저/편집기 모듈(328), MM에 대한 프린터 드라이버(330), 및 DVP 인쇄 시스템(332)는 전적으로 캡처 장치(106) 내에 존재할 수 있고, 그에 따라 캡처 장치(106)에 향상된 기능을 제공할 수 있다.
- [0143] MMR 소프트웨어 모음
- [0144] 도 4는 본 발명의 일 실시예에 따른 MMR 소프트웨어 모음(222)에 포함되어 있는 일련의 소프트웨어 컴포넌트를 나타낸 것이다. MMR 소프트웨어 모음(222)의 전부 또는 그 일부가 MMR 컴퓨터(112), 캡처 장치(106), 네트워크화된 미디어 서버(114) 및 기타 서버에 포함되어 있을 수 있다는 것을 잘 알 것이다. 그에 부가하여, MMR 소프트웨어 모음(222)의 다른 실시예들은 도시된 컴포넌트들 중의 하나로부터 그 전부까지 임의의 수의 컴포넌트를 가질 수 있다. 이 예의 MMR 소프트웨어 모음(222)은, 텍스트 콘텐츠-기반 검색 컴포넌트(412), 이미지 콘텐츠-기반 검색 컴포넌트(414), 및 스테가노그래피 수정 컴포넌트(steganographic modification component)(416)를 포함하는 멀티미디어 주석 첨부 소프트웨어(410), 종이 판독 이력 로그(paper reading history log)(418), 온라인 판독 이력 로그(420), 협동적 문서 검토 컴포넌트(collaborative document review component)(422), 실시간 통지 컴포넌트(424), 멀티미디어 검색 컴포넌트(426), 바탕화면 비디오 리마인더 컴포넌트(desktop video reminder component)(428), 웹 페이지 리마인더 컴포넌트(430), 물리적 이력 로그(432), 완성된 양식 검토자 컴포넌트(completed form reviewer component)(434), 시간 수송 컴포넌트(time transportation component)(436), 위치 인식 컴포넌트(location awareness component)(438), PC 저작 컴포넌트(440), 문서 저작 컴포넌트(442), 캡처 장치 저작 컴포넌트(444), 무의식적 업로드 컴포넌트(unconscious upload component)(446), 문서 버전 검색 컴포넌트(448), PC 문서 메타데이터 컴포넌트(450), 캡처 장치 UI 컴포넌트(452) 및 도메인-관련 컴포넌트(454)를 포함한다.
- [0145] 한 특정의 실시예에 따르면, 멀티미디어 주석 첨부 소프트웨어(410)는 문서 이벤트 데이터베이스(320)의 구성과 함께 MMR 시스템(100b)의 기본적인 기술을 형성한다. 보다 구체적으로는, 멀티미디어 주석 첨부 소프트웨어(410)는 종이 문서에 대한 멀티미디어 주석 첨부부를 관리하기 위한 것이다. 예를 들어, MMR 사용자(110)는 캡처 장치(106)를 종이 문서의 임의의 일부분으로 향하게 하고 이어서 그 일부분에 주석을 추가하기 위해 캡처 장치(106)의 적어도 하나의 캡처 메커니즘(230)을 사용한다. 특정의 예에서, 변호사는 계약의 일부분에 관한 메모를 받아 적게 한다(오디오 파일을 생성한다). 멀티미디어 데이터(오디오 파일)가 자동적으로 문서의 원래의 전자 버전에 첨부된다. 그 문서의 차후의 인쇄 출력물은 선택에 따라서는 그 주석이 존재한다는 표시를 포함한다.
- [0146] 텍스트 콘텐츠-기반 검색 컴포넌트(412)는 텍스트로부터 콘텐츠-기반 정보를 검색하는 소프트웨어 애플리케이션이다. 예를 들어, 텍스트 콘텐츠-기반 검색 컴포넌트(412)를 사용함으로써, 텍스트 패치로부터 콘텐츠가 검색되거나, 원래의 문서 및 문서 내의 일부분이 식별되거나, 그 패치에 링크되어 있는 다른 정보가 식별된다. 텍스트 콘텐츠-기반 검색 컴포넌트(412)는 OCR-기반 기법을 이용할 수 있다. 다른 대안으로서, 텍스트 동작으로부터 콘텐츠-기반 검색을 수행하는 비OCR-기반 기법은 텍스트 패치에 워드 길이를 2차원 배열하는 것을 포함한다. 텍스트 콘텐츠-기반 검색 컴포넌트(412)의 일례는, 피쳐들이 추출된 문서 및 문서 내의 일부분을 식별하기 위해, 텍스트의 일부의 이미지로부터 추출되는 수평 및 수직 피쳐를 결합하는 알고리즘이다. 수평 및 수직 피쳐는 직렬로, 병렬로, 또는 다른 방식으로 동시에 사용될 수 있다. 고속 구현 및 노이즈의 존재 시의 안정성



(robustness)을 제공하는 이러한 비OCR-기반 피쳐 세트가 사용된다.

- [0147] 이미지 콘텐츠-기반 검색 컴포넌트(414)는 이미지로부터 콘텐츠-기반 정보를 검색하는 소프트웨어 애플리케이션이다. 이미지 콘텐츠-기반 검색 컴포넌트(414)는 가능한 이미지 정합 및 연관된 신뢰 수준(level of confidence)의 리스트를 발생하기 위해 캡처된 데이터와 데이터베이스(320) 내의 이미지 간의 이미지 비교를 수행한다. 그에 추가하여, 각각의 이미지 정합은 연관된 데이터 또는 사용자 입력에 응답하여 수행되는 동작들을 가질 수 있다. 일례에서, 이미지 콘텐츠-기반 검색 컴포넌트(414)는 이미지를 동일한 피쳐 배열을 갖는 이미지에 대한 이미지 데이터베이스에 질의하는 데 사용될 수 있는 벡터 표시로 변환함으로써, 예를 들어, 래스터 이미지(예를 들어, 맵)에 기초하여 콘텐츠를 검색한다. 대안의 실시예는 데이터베이스에서 정합하는 이미지를 탐색하기 위해 이미지의 컬러 콘텐츠 또는 이미지 내의 객체들의 기하학적 배열을 사용한다.
- [0148] 스테가노그래피 수정 컴포넌트(416)는 인쇄에 앞서 스테가노그래피 수정을 수행하는 소프트웨어 애플리케이션이다. MMR 애플리케이션을 더 양호하게 인쇄가능시키기 위해, 텍스트 및 이미지가 인쇄되기 이전에, 디지털 정보가 텍스트 및 이미지에 추가된다. 대안의 실시예에서, 스테가노그래피 수정 컴포넌트(416)는 1) 텍스트, 오디오, 또는 비디오 정보 등의 원래의 기본 콘텐츠, 2) 텍스트, 오디오, 비디오, 애플릿, 하이퍼텍스트 링크, 기타 등등의 임의의 형태의 추가 콘텐츠를 포함하는 MMR 문서를 발생하고 저장한다. 스테가노그래피 수정은 컬러 또는 그레이스케일 이미지에 워터마크를 임베딩하는 것, 문서의 배경에 도트 패턴을 인쇄하는 것, 또는 디지털 정보를 인코딩하기 위해 인쇄된 문자의 윤곽선을 미묘하게 수정하는 것(subtle modification)을 포함할 수 있다.
- [0149] 종이 판독 이력 로그(418)는 종이 문서의 판독 이력 로그이다. 종이 판독 이력 로그(418)는, 예를 들어, 문서 이벤트 데이터베이스(320)에 존재한다. 종이 판독 이력 로그(418)는 Ricoh Innovations에 의해 개발된 비디오로부터의 문서 식별 기술(document identification-from-video technology)에 기초하며, 이는 MMR 사용자(110)에 의해 판독된 문서의 이력을 생성하는 데 사용된다. 종이 판독 이력 로그(418)는, 예를 들어, MMR 사용자(110)에게 판독된 문서 및/또는 임의의 연관된 이벤트를 잊지 않도록 알려주는(remind) 데 유용하다.
- [0150] 온라인 판독 이력 로그(420)는 온라인 문서의 판독 이력 로그이다. 온라인 판독 이력 로그(420)는 운영 체제 이벤트의 분석에 기초하며, 예를 들어, 문서 이벤트 데이터베이스(320)에 존재한다. 온라인 판독 이력 로그(420)는 MMR 사용자(110)에 의해 판독된 및 그의 일부분이 판독된 온라인 문서들의 기록이다. 온라인 판독 이력 로그(420)에서의 엔트리는, 각각의 페이지의 하단에 메모를 제공하는 것 또는 각각의 단락을 판독하는 데 소비된 시간량에 기초하는 서로 다른 컬러로 텍스트를 하이라이트하는 것 등의, 많은 방식으로 임의의 차후의 인쇄 출력물 상으로 인쇄될 수 있다. 그에 추가하여, 멀티미디어 주석 첨부 소프트웨어(410)는 이 데이터를 PD 인덱스(322)에 인덱싱할 수 있다. 선택에 따라서는, 온라인 판독 이력 로그(420)는, MMR 컴퓨터(112)를 모니터링하는 얼굴 검출 시스템 등의 장치들을 갖추고 있는 MMR 컴퓨터(112)의 도움을 받을 수 있다.
- [0151] 협동적 문서 검토 컴포넌트(collaborative document review component)(422)는, 동일한 종이 문서의 서로 다른 버전의 2명 이상의 판독자가 그의 캡처 장치(106)를 문서의 임의의 일부분으로 향하게 함으로써 다른 판독자들에게 의해 적용되는 코멘트를 검토할 수 있게 해주는 소프트웨어 애플리케이션이다. 예를 들어, 주석은 문서 섬네일의 상단에 오버레이(overlay)로서 캡처 장치(106) 상에 디스플레이될 수 있다. 협동적 문서 검토 컴포넌트(422)는 임의의 유형의 기존의 협동 소프트웨어로 구현되거나 다른 방식으로 그와 협력할 수 있다.
- [0152] 실시간 통지 컴포넌트(424)는 판독되고 있는 문서의 실시간 통지를 수행하는 소프트웨어 애플리케이션이다. 예를 들어, MMR 사용자(110)가 문서를 판독하고 있는 동안에, 그의 판독 자취(reading trace)가 블로그 또는 온라인 게시판 상에 게시된다. 그 결과, 동일한 주제에 관심이 있는 다른 사람들이 잠깐 들러서 그 문서에 관하여 이야기를 할 수 있다.
- [0153] 멀티미디어 검색 컴포넌트(426)는 임의적인 종이 문서로부터 멀티미디어를 검색하는 소프트웨어 애플리케이션이다. 예를 들어, MMR 사용자(110)는, 캡처 장치(106)를 문서로 향하게 함으로써, 임의적인 종이 문서가 MMR 사용자(110)의 책상에 존재한 동안에 일어난 모든 대화를 검색할 수 있다. 이것은 멀티미디어 데이터를 캡처하는 사무실 포털(120)(또는 다른 적당한 메커니즘)이 MMR 사용자(110)의 사무실에 존재하는 것으로 가정한다.
- [0154] 바탕화면 비디오 리마인더 컴포넌트(desktop video reminder component)(428)는 MMR 컴퓨터(112) 상에서 일어나는 이벤트를 MMR 사용자(110)에게 잊지 않도록 알려주는 소프트웨어 애플리케이션이다. 예를 들어, 캡처 장치(106)를 문서의 일부분으로 향하게 함으로써, MMR 사용자(110)는 그 일부분이 보여지는 동안에 일어난 MMR 컴퓨터(112)의 바탕화면에서의 변화를 보여주는 비디오 클립을 볼 수 있다. 그에 추가하여, 바탕화면 비디오 리마인더 컴포넌트(428)는, MMR 컴퓨터(112)의 근방에 존재하는 오디오 등의, MMR 컴퓨터(112)에 의해 녹화된 다

른 멀티미디어를 검색하는 데 사용될 수 있다.

- [0155] 웹 페이지 리마인더 컴포넌트(430)는 그의 MMR 컴퓨터(112) 상에 보여지는 웹 페이지들을 MMR 사용자(110)에게 잊지 않도록 알려주는 소프트웨어 애플리케이션이다. 예를 들어, 종이 문서 상에서 캡처 장치(106)를 패닝함으로써, MMR 사용자(110)는 문서의 대응하는 일부분이 MMR 컴퓨터(112)의 바탕화면 상에 나타내어진 동안에 보여지는 웹 페이지들의 자취를 볼 수 있다. 웹 페이지들은 SD 브라우저(312, 314) 등의 브라우저에 또는 캡처 장치(106)의 디스플레이(212) 상에 나타내어질 수 있다. 다른 대안으로서, 웹 페이지들은 캡처 장치(106)의 디스플레이(212) 상에 또는 MMR 컴퓨터(112) 상에 미가공 URL(raw URL)로서 제공된다.
- [0156] 물리적 이력 로그(432)는, 예를 들어, 문서 이벤트 데이터베이스(320)에 존재한다. 물리적 이력 로그(432)는 종이 문서들의 물리적 이력 로그이다. 예를 들어, MMR 사용자(110)는 그의 캡처 장치(106)를 종이 문서로 향하게 하고, 물리적 이력 로그(432)에 저장된 정보를 사용함으로써, 과거의 어떤 시간에 관심의 문서에 인접해 있던 다른 문서들이 결정된다. 이 동작은, 예를 들어, RFID와 같은 추적 시스템에 의해 용이하게 된다. 이 경우에, 캡처 장치(106)는 RFID 판독기(244)를 포함한다.
- [0157] 완성된 양식 검토자 컴포넌트(completed form reviewer component)(434)는 양식(form)을 완성하는 데 사용되는 이전에 획득된 정보를 검색하는 소프트웨어 애플리케이션이다. 예를 들어, MMR 사용자(110)는 그의 캡처 장치(106)를 비어있는 양식(예를 들어, 웹 사이트로부터 인쇄된 의료 보험금 청구서 양식(medical claim form))으로 향하게 하고 이전에 입력된 정보의 이력을 제공받는다. 그 후에, 이 양식은 이 완성된 양식 검토자 컴포넌트(434)에 의해 자동적으로 이 이전에 입력된 정보로 채워진다.
- [0158] 시간 수송 컴포넌트(time transportation component)(436)는 문서의 과거 및 미래의 버전에 대한 소스 파일을 검색하고 그 버전들과 연관되어 있는 이벤트의 리스트를 검색하여 디스플레이하는 소프트웨어 애플리케이션이다. 이 동작은 손에 있는 인쇄된 문서가 그와 연관된 가장 중요한 외부 이벤트(예를 들어, 토론 또는 회의)보다 수개월 이후에 생성된 문서의 버전으로부터 생성될 수 있다는 사실을 보상한다.
- [0159] 위치 인식 컴포넌트(location awareness component)(438)는 위치-인식 종이 문서(location-aware paper document)를 관리하는 소프트웨어 애플리케이션이다. 위치-인식 종이 문서의 관리, 예를 들어, RFID와 같은 추적 시스템에 의해 용이하게 된다. 예를 들어, 캡처 장치(106)는 하루 종일 MMR 사용자(110)의 지리적 위치의 자취를 캡처하고 문서 또는 문서를 포함하는 폴더에 첨부된 RFID 태그를 스캔한다. RFID 스캐닝 동작은 그의 범위 내에 있는 임의의 RFID 태그를 검출하기 위해 캡처 장치(106)의 RFID 판독기(244)에 의해 수행된다. MMR 사용자(110)의 지리적 위치는, 지리적 위치 확인 메커니즘(142)과 함께, 셀룰러 기반구조(132) 내의 각각의 셀 타워의 식별 번호에 의해 또는, 다른 대안으로서, 캡처 장치(106)의 GPS 장치(242)를 통해 추적될 수 있다. 다른 대안으로서, 문서 식별은 "항상 켜져 있는 비디오(always-on video)" 또는 캡처 장치(106)의 비디오 카메라(232)로 달성될 수 있다. 이 위치 데이터는 "지리적으로 참조되는(geo-referenced)" 문서를 제공하며, 이는 하루 종일 문서가 어디에 위치하고 있는지를 보여주는 맵-기반 인터페이스(map-based interface)를 가능하게 해준다. 응용은 먼 곳에 있는 고객을 방문할 시에 파일을 들고 가는 변호사이다. 대안의 실시예에서, 문서(118)는 문서가 이동될 때를 감지할 수 있고 또 어떤 기본적인 얼굴 검출 동작을 수행할 수 있는 그에 첨부된 감지 메커니즘을 포함한다. 이 감지 기능은 종이 문서에 첨부되어 있는 일련의 자이로스코프 또는 유사한 장치를 통한다. 위치 정보에 기초하여, MMR 시스템(100b)은 언제 소유자의 셀룰러 전화로 전화를 걸쳐 그에게 문서가 이동하고 있다고 말할지를 알려준다. 셀룰러 전화는 그 문서를 그의 가상의 서류 가방에 추가할 수 있다. 그에 부가하여, 이것은 "눈에 보이지 않는(invisible)" 바코드의 개념이며, 이는 캡처 장치(106)의 비디오 카메라(232) 또는 정지 영상 카메라(234)에 보이지만 사람에게는 보이지 않거나 아주 희미한 기계-판독가능 마킹이다. 캡처 장치(106)에서 디코딩될 수 있는 다양한 잉크 및 스테가노그래피 또는 인쇄된-이미지 워터마킹 기법이 위치를 결정하기 위해 고려될 수 있다.
- [0160] PC 저작 컴포넌트(440)는 MMR 컴퓨터(112) 등의 PC 상에서 저작 동작을 수행하는 소프트웨어 애플리케이션이다. PC 저작 컴포넌트(440)는 마이크로소프트 워드, 파워포인트, 및 웹 페이지 저작 패키지 등의 기존의 저작 애플리케이션에 대한 플러그-인으로서 제공된다. PC 저작 컴포넌트(440)는 MMR 사용자(110)가 그의 MMR 컴퓨터(112)로부터 이벤트트로의 또는 그의 환경에서의 이벤트트로의 링크를 갖는 종이 문서를 준비할 수 있게 해주거나, 인쇄된 문서(118)가 그것이 생성된 워드 파일로 자동적으로 링크되는 것과 같이, 링크를 갖는 종이 문서가 자동적으로 발생될 수 있게 해주거나, MMR 사용자(110)가 워드 파일을 검색하고 이것을 어떤 다른 사람에게 제공할 수 있게 해준다. 링크를 갖는 종이 문서는 지금까지 MMR 문서라고 하였다. MMR 문서에 대한 더 많은 상세에 대해 도 5를 참조하여 더 기술한다.

- [0161] 문서 저작 컴포넌트(442)는 기존의 문서에 대한 저작 동작을 수행하는 소프트웨어 애플리케이션이다. 문서 저작 컴포넌트(442)는, 예를 들어, 개인 에디션(personal edition)으로서 또는 엔터프라이즈 에디션(enterprise edition)중 하나로 구현될 수 있다. 개인 에디션에서, MMR 사용자(110)는 문서를 스캔하고 이들을 MMR 문서 데이터베이스(예를 들어, 문서 이벤트 데이터베이스(320))에 추가한다. 엔터프라이즈 에디션에서, 게시자(또는 제3자)는 원래의 전자 소스(예를 들어, 전자 교정쇄(electronic galley proof))로부터 MMR 문서를 생성한다. 이 기능은 고급 게시 패키지(예를 들어, 아도브 리더)에 내장될 수 있으며 다른 개체에 의해 제공되는 백엔드 서비스(backend service)와 링크될 수 있다.
- [0162] 캡처 장치 저작 컴포넌트(444)는 캡처 장치(106) 상에서 직접 저작 동작을 수행하는 소프트웨어 애플리케이션이다. 캡처 장치 저작 컴포넌트(444)를 사용하여, MMR 사용자(110)는 그의 손에 있는 종이 문서로부터 주요 구문을 추출하고 임시적인 MMR 문서를 생성하기 위해 동작 중에 캡처된 부가적인 콘텐츠와 함께 주요 구문을 저장한다. 그에 부가하여, 캡처 장치 저작 컴포넌트(444)를 사용함으로써, MMR 사용자(110)는 그의 MMR 컴퓨터(112)로 복귀할 수 있고 그가 생성한 임시적인 MMR 문서를 파워포인트 등의 기존의 문서 애플리케이션으로 다운로드할 수 있으며, 이어서 이를 편집하여 MMR 문서의 최종 버전으로 또는 다른 애플리케이션에 대한 다른 표준 유형의 문서로 만든다. 그렇게 함에 있어서, 이미지 및 텍스트가, 파워포인트 문서의 페이지들과 같은, 기존의 문서의 페이지들에 자동적으로 삽입된다.
- [0163] 무의식적 업로드 컴포넌트(unconscious upload component)(446)는 인쇄된 문서들을 캡처 장치(106)로 무의식적으로(자동적으로, 사용자 간섭 없이) 업로드하는 소프트웨어 애플리케이션이다. MMR 사용자(110)가 그의 MMR 컴퓨터(112)에 있을 때를 비롯하여 대부분의 시간에, 캡처 장치(106)가 MMR 사용자(110)의 소유이기 때문에, 프린터 드라이버(316)는, 문서를 프린터(116)에 전송하는 것에 부가하여, Wi-Fi 기술(134) 또는 블루투스 기술(136)을 조합하여 그 동일한 문서를 캡처 장치(106)의 무선 통신 링크(218)를 통해, 또는 캡처 장치(106)가 MMR 컴퓨터(112)에 연결되거나 그와 도킹되어 있는 경우 유선 접속에 의해 캡처 장치(106)의 저장 장치(216)에 푸싱할 수도 있다. 이와 같이, MMR 사용자(110)는 문서가 인쇄된 후에 잊지 않고 그 문서를 집어드는데, 그 이유는 그것이 자동적으로 캡처 장치(106)로 업로드되기 때문이다.
- [0164] 문서 버전 검색 컴포넌트(448)는 주어진 소스 파일(310)의 과거 및 장래의 버전을 검색하는 소프트웨어 애플리케이션이다. 예를 들어, MMR 사용자(110)는 캡처 장치(106)를 인쇄된 문서로 향하게 하고, 이어서 문서 버전 검색 컴포넌트(448)는 현재의 소스 파일(310)(예를 들어, 워드 파일) 및 소스 파일(310)의 다른 과거 및 장래의 버전의 위치를 찾는다. 한 특정의 실시예에서, 이 동작은 소스 파일(310)이 복사 및 이동되는 위치들을 추적하는 Windows 파일 추적 소프트웨어를 사용한다. 다른 이러한 파일 추적 소프트웨어도 역시 여기에서 사용될 수 있다. 예를 들어, 구글 데스크톱 검색(Google Desktop Search) 또는 마이크로소프트 윈도우즈 검색 도우미(Microsoft Windows Search Companion)는 소스 파일(310)로부터 선택된 단어들로 작성된 질의로 파일의 현재 버전을 찾아낼 수 있다.
- [0165] PC 문서 메타데이터 컴포넌트(450)는 문서의 메타데이터를 검색하는 소프트웨어 애플리케이션이다. 예를 들어, MMR 사용자(110)는 캡처 장치(106)를 인쇄된 문서로 향하게 하고, PC 문서 메타데이터 컴포넌트(450)는 누가 문서를 인쇄했는지, 언제 문서가 인쇄되었는지, 어디서 문서가 인쇄되었는지, 및 인쇄 시에 주어진 소스 파일(310)에 대한 파일 경로를 결정한다.
- [0166] 캡처 장치 UI 컴포넌트(452)는 캡처 장치(106)의 UI의 동작을 관리하는 소프트웨어 애플리케이션으로서, MMR 사용자(110)가 종이 문서와 상호작용할 수 있게 해준다. 캡처 장치 UI 컴포넌트(452) 및 캡처 장치 UI(224)의 결합은 MMR 사용자(110)가 기존의 문서들로부터 데이터를 판독하고 데이터를 기존의 문서들에 기록하며, 그 문서들과 연관되어 있는 증강 현실을 보고 그와 상호작용하며(즉, 캡처 장치(106)를 통해, MMR 사용자(110)는 문서가 생성될 때 또는 그것이 편집되는 동안에 무슨 일이 일어났는지를 볼 수 있으며), 그의 캡처 장치(106) 상에 디스플레이되는 문서들과 연관되어 있는 증강 현실을 보고 그와 상호작용할 수 있게 해준다.
- [0167] 도메인-관련 컴포넌트(454)는 도메인-관련 기능을 관리하는 소프트웨어 애플리케이션이다. 예를 들어, 음악 애플리케이션에서, 도메인-관련 컴포넌트(454)는, 예를 들어, 캡처 장치(106)의 음성 녹음기(236)를 통해 검출되는 음악을 타이틀, 음악가, 또는 작곡가와 정합시키는 소프트웨어 애플리케이션이다. 이와 같이, 검출된 음악에 관계되어 있는 시트 음악(sheet music) 또는 음악 CD 등의 관심의 항목이 MMR 사용자(110)에게 제공될 수 있다. 이와 유사하게, 도메인-관련 컴포넌트(454)는 비디오 콘텐츠, 비디오 게임, 및 임의의 엔터테인먼트 정보에 대해 유사한 방식으로 동작하도록 구성되어 있다. 장치 관련 컴포넌트(454)는 또한 임의의 매스 미디어 콘텐츠의 전자 버전에 맞게 구성될 수 있다.

- [0168] 계속하여 도 3 및 도 4를 참조하여, MMR 소프트웨어 모음(222)의 소프트웨어 컴포넌트가 전체적으로 또는 부분적으로 MMR 시스템(100b)의 하나 이상의 MMR 컴퓨터(112), 네트워크화된 서버(114), 서비스 제공자 서버(122), 및 캡처 장치(106) 상에 존재할 수 있다. 환언하면, MMR 소프트웨어 모음(222)에 의해 수행되는 것 등의 MMR 시스템(100b)의 동작은 임의의 사용자 정의 구성에서 MMR 컴퓨터(112), 네트워크화된 서버(114), 서비스 제공자 서버(122), 와 캡처 장치(106)(또는 시스템(100b)에 포함되어 있는 다른 이러한 처리 환경들) 사이에 분산될 수 있다.
- [0169] 이 개시 내용을 살펴보면 MMR 시스템(100a/100b)의 기본적인 기능이 MMR 소프트웨어 모음(222)의 소프트웨어 컴포넌트들의 어떤 조합으로 수행될 수 있다는 것이 명백할 것이다. 예를 들어, MMR 시스템(100a/100b)의 일 실시예의 기본적인 기능은 이하의 것들을 포함한다.
- [0170] • 제1 미디어 부분 및 제2 미디어 부분을 포함하는 MMR 문서를 생성하거나 그에 추가하는 것
  - [0171] • 제2 미디어 부분에 있는 정보에 액세스하기 위해 MMR 문서의 제1 미디어 부분(예를 들어, 종이 문서)을 사용하는 것
  - [0172] • 전자적 도메인에서 프로세스를 트리거 또는 시작하기 위해 MMR 문서의 제1 미디어 부분(예를 들어, 종이 문서)을 사용하는 것
  - [0173] • 제2 미디어 부분을 생성 또는 그에 추가하기 위해 MMR 문서의 제1 미디어 부분(예를 들어, 종이 문서)을 사용하는 것
  - [0174] • 제1 미디어 부분을 생성 또는 그에 추가하기 위해 MMR 문서의 제2 미디어 부분을 사용하는 것
  - [0175] • 전자적 도메인 내의 또는 제1 미디어 부분과 관계되어 있는 프로세스를 트리거 또는 시작하기 위해 MMR 문서의 제2 미디어 부분을 사용하는 것.
- [0176] MMR 문서
- [0177] 도 5는 본 발명의 일 실시예에 따른 MMR 문서(500)를 나타낸 도면이다. 보다 구체적으로는, 도 5는 인쇄된 문서(118)의 일부분의 표시(502), 동작 또는 제2 미디어(504), 인덱스 또는 핫스팟(506), 및 전체 문서(118)의 전자적 표시(508)를 포함하는 MMR 문서(500)를 나타낸 것이다. MMR 문서(500)가 통상적으로 문서 이벤트 데이터 베이스(320)에 저장되지만, 네트워크(128)에 연결된 캡처 장치 또는 임의의 다른 장치에 저장될 수도 있다. 일 실시예에서, 다수의 MMR 문서가 인쇄된 문서에 대응할 수 있다. 다른 실시예에서, 단일의 인쇄된 문서에 다수의 핫스팟(506)을 생성하기 위해 도 5에 나타난 구조가 복제된다. 한 특정의 실시예에서, MMR 문서(500)는 페이지 및 페이지 내에서의 위치를 갖는 표시(502) 및 핫스팟(506)을 포함하고, 제2 미디어(504) 및 전자적 표시(508)는 선택적이고 그 자체로서 점선으로 표시되어 있다. 유의할 점은 제2 미디어(504) 및 전자적 표시(508)가, 원하는 경우, 나중에 MMR 문서가 생성된 후에 추가될 수 있다는 것이다. 이 기본적인 실시예는 문서 또는 그 표시에 대응하는 문서에서의 특정의 위치를 찾아내는 데 사용될 수 있다.
- [0178] 인쇄된 문서(118)의 일부분의 표시(502)는 패턴 정합에 사용가능하고 문서에서의 적어도 하나의 위치를 식별해주는 임의의 형태(이미지, 벡터, 픽셀, 텍스트, 코드, 기타)일 수 있다. 표시(502)이 인쇄된 문서에서의 위치를 일의적으로 식별해주는 것이 더 바람직하다. 일 실시예에서, 표시(502)는 도 5에 나타난 텍스트 지문이다. 텍스트 지문(502)은 PD 캡처 모듈(318)을 통해 자동적으로 캡처되고 인쇄 동안에 PD 인덱스(322)에 저장된다. 다른 대안으로서, 텍스트 지문(502)은 문서 스캐너(127)의 문서 지문 정합 모듈(226')을 통해 자동적으로 캡처되고 스캔 동작 동안에 PD 인덱스(322)에 저장된다. 표시(502)은 다른 대안으로서 전체 문서, 텍스트 패치, 문서에서의 고유한 인스턴스인 경우 단일 단어, 이미지의 일부분, 고유의 속성 또는 문서의 정합가능한 부분의 임의의 다른 표시일 수 있다.
- [0179] 동작 또는 제2 미디어(504)는 양호하게는 임의의 유형의 디지털 파일 또는 데이터 구조이다. 가장 기본적인 실시예에서의 제2 미디어(504)는 제공될 텍스트 또는 실행될 하나 이상의 명령일 수 있다. 제2 미디어 유형(504)은 보다 통상적으로는 표시(502)에 의해 식별되는 문서의 일부분과 관계되어 있는 텍스트 파일, 오디오 파일, 또는 비디오 파일이다. 제2 미디어 유형(504)은 데이터 구조 또는 다수의 서로 다른 미디어 유형을 참조하거나 포함하는 파일, 및 동일한 유형의 다수의 파일일 수 있다. 예를 들어, 제2 미디어(504)는 텍스트, 명령, 이미지, PDF 파일, 비디오 파일, 오디오 파일, 애플리케이션 파일(예를 들어, 스프레드시트 또는 워드 프로세싱 문



서), 기타 등등일 수 있다.

- [0180] 인덱스 또는 핫스팟(506)은 표시(502)와 동작 또는 제2 미디어(504) 사이의 링크이다. 핫스팟(506)은 표시(502)와 제2 미디어(504)를 연관시킨다. 일 실시예에서, 인덱스 또는 핫스팟(506)은 문서 내에서의 x 및 y 좌표 등의 위치 정보를 포함한다. 핫스팟(506)은 점, 영역 또는 심지어 전체 문서일 수 있다. 일 실시예에서, 핫스팟은 표시(502)에 대한 포인터, 제2 미디어(504)에 대한 포인터, 및 문서 내에서의 위치를 갖는 데이터 구조이다. MMR 문서(500)가 다수의 핫스팟(506)을 가질 수 있고 이러한 경우에 데이터 구조가 다수의 표시, 다수의 제2 미디어 파일, 및 인쇄된 문서(118) 내의 다수의 위치 간의 링크를 생성한다는 것을 잘 알 것이다.
- [0181] 대안의 실시예에서, MMR 문서(500)는 전체 문서(118)의 전자적 표시(508)을 포함한다. 이 전자적 표시는 핫스팟(506)의 위치를 결정하는 데 또한 캡처 장치(106) 또는 MMR 컴퓨터(112) 상에 문서를 디스플레이하기 위해 사용자 인터페이스에 의해 사용될 수 있다.
- [0182] MMR 문서(500)의 예시적인 용도는 다음과 같다. 텍스트 지문 또는 표시(502)를 분석함으로써, 캡처된 텍스트 프래그먼트(captured text fragment)가 캡처 장치(106)의 문서 지문 정합 모듈(226)을 통해 식별된다. 예를 들어, MMR 사용자(110)는 그의 캡처 장치(106)의 비디오 카메라(232) 또는 정지 영상 카메라(234)를 인쇄된 문서(118)로 향하게 하고 이미지를 캡처한다.
- [0183] 그 후에, 문서 지문 정합 모듈(226)은 캡처된 이미지에 대해 그의 분석을 수행하여, 연관된 엔트리가 PD 인덱스(322) 내에 존재하는지 여부를 판정한다. 정합이 발견되는 경우, 핫스팟(506)의 존재가 그의 캡처 장치(106)의 디스플레이(212) 상에서 MMR 사용자(110)에게 하이라이트된다. 예를 들어, 도 5에 나타난 바와 같이, 단어 또는 어구가 하이라이트된다. 인쇄된 문서(118) 내의 각각의 핫스팟(506)은, 네트워크화된 미디어 서버(114) 상에 존재하는 MM 파일들(336) 중의 하나 등의, 다른 사용자-정의된 또는 미리 정해진 데이터로의 링크로서 역할한다. PD 인덱스(322)에 저장되어 있는 텍스트 지문 또는 표시(502)에의 액세스는 전자적 데이터가 임의의 MMR 문서(500) 또는 문서 내의 임의의 핫스팟(506)에 추가될 수 있게 해준다. 도 4를 참조하여 기술한 바와 같이, 적어도 하나의 핫스팟(506)(예를 들어, 링크)을 포함하는 종이 문서는 MMR 문서(500)라고 언급된다.
- [0184] 계속하여 도 1b, 도 2a 내지 도 2d, 도 3, 도 4 및 도 5를 참조하면, MMR 시스템(100b)의 예시적인 동작은 다음과 같다. MMR 사용자(110) 또는 출판 회사 등의 임의의 다른 개체는 주어진 소스 파일(310)을 열고, 인쇄된 문서(118) 등의 종이 문서를 생성하기 위해 인쇄 동작을 시작한다. 인쇄 동작 동안에, (1) 인쇄 시에 PD 캡처 모듈(318)을 통해 인쇄된 형식을 자동적으로 캡처하고 이를 캡처 장치(106)에 전송하는 것 - 문서의 전자적 표시(508)는, 예를 들어, SD 브라우저(312)의 출력에서 PD 캡처 모듈(318)을 사용함으로써 인쇄 시에 자동적으로 캡처된다. 예를 들어, MMR 사용자(110)는 SD 브라우저(312)로부터의 콘텐츠를 인쇄하고 이 콘텐츠는 PD 캡처 모듈(318)을 통해 필터링된다. 앞서 논의한 바와 같이, 페이지 상의 텍스트의 2차원 배열은 인쇄를 위해 문서가 레이아웃될 때 결정될 수 있다. -, (2) PD 캡처 모듈(318)을 통해 인쇄 시에 주어진 소스 파일(310)을 자동적으로 캡처하는 것, 및 (3) 캡처 장치(106) 상의 멀티미디어 주식 첨부 인터페이스를 채울 수 있는 "명명된 개체(named entity)" 또는 다른 흥미있는 정보를 찾아내기 위해, 문서 파서 모듈(326)을 통해, 인쇄된 형식 및/또는 소스 파일(310)을 파싱하는 것 - 명명된 개체는, 예를 들어, 나중에 멀티미디어를 추가하기 위한 "앵커(anchor)", 즉 자동적으로 발생된 핫스팟(506)이다. 문서 파서 모듈(326)은 입력으로서 주어진 인쇄된 문서(118)와 관계되어 있는 소스 파일(310)을 수신한다. 문서 파서 모듈(326)은 인쇄된 문서(118)에서의 타이틀, 저작자, 시간, 또는 장소 등의 핫스팟(506)을 이용하기 위한 표시(502)를 식별해주고 따라서 캡처 장치(106) 상에서 수신될 정보를 프롬프트하는 애플리케이션이다. -, (4) 콘텐츠-기반 검색을 위해 인쇄된 형식 및/또는 소스 파일(310)을 자동적으로 인덱싱하는 것, 즉 PD 인덱스(322)를 작성하는 것, (5) 소스 파일(310)과 연관된 문서 및 이벤트에 대해 문서 이벤트 데이터베이스(320) 내에 엔트리, 예를 들어, 편집 이력 및 현재의 위치를 만드는 것, 및 (6) MMR 사용자(110)가, 인쇄되기 이전에, 따라서 MMR 문서(500)이 형성되기 이전에, 핫스팟(506)을 문서에 추가할 수 있게 해주는, 프린터 드라이버(316) 내의 상호작용적 대화상자를 수행하는 것 등의 어떤 동작들이 자동적으로 수행된다. 연관된 데이터는 MMR 컴퓨터(112) 상에 저장되거나 네트워크화된 미디어 서버(114)로 업로드된다.
- [0185] 예시적인 대안의 실시예
- [0186] MMR 시스템(100)(100a 또는 100b)은 도 1a 및 도 1b, 도 2a 내지 도 2d, 및 도 3 내지 도 5에 도시된 구성으로 한정되지 않는다. MMR 소프트웨어는 전체적으로 또는 부분적으로 캡처 장치(106)와 MMR 컴퓨터(112) 간에 분산되어 있을 수 있고, 도 3 및 도 4를 참조하여 상기한 모듈들 전부보다 상당히 더 적은 모듈이 필요하다. 이하의 것들을 비롯하여 다수의 구성들이 가능하다.



- [0187] MMR 시스템(100)의 제1 대안의 실시예는 캡처 장치(106) 및 캡처 장치 소프트웨어를 포함한다. 캡처 장치 소프트웨어는 캡처 장치 UI(224) 및 문서 지문 정합 모듈(226)(예를 들어, 도 3에 도시됨)이다. 캡처 장치 소프트웨어는 캡처 장치(106) 상에서, 또는 다른 대안으로서, 캡처 장치(106)가 액세스가능한 네트워크화된 미디어 서버(114) 또는 서비스 제공자 서버(122) 등의 외부 서버 상에서 실행된다. 이 실시예에서, 게시(publication)에 링크되어 있는 데이터를 공급하는 네트워크화된 서비스가 이용가능하다. 계층적 인식 방식이 사용될 수 있으며, 이 경우 게시가 먼저 식별되고 이어서 그 게시 내의 페이지 및 섹션이 식별된다.
- [0188] MMR 시스템(100)의 제2 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어 및 문서 사용 소프트웨어를 포함한다. 제2 대안의 실시예는 인쇄된 문서를 캡처 및 인덱싱하고 문서의 편집 이력 등의 기본적인 문서 이벤트를 링크하는, 도 4를 참조하여 도시되고 기술된 것과 같은 소프트웨어를 포함한다. 이것은 MMR 사용자(110)가 그의 캡처 장치(106)를 임의의 인쇄된 문서로 향하게 할 수 있게 해주고, 또 문서를 발생한 소스 파일(310)의 이름 및 위치를 결정하는 것은 물론 인쇄 시간 및 장소를 결정할 수 있게 해준다.
- [0189] MMR 시스템(100)의 제3 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어, 문서 사용 소프트웨어, 및 이벤트 캡처 모듈(324)을 포함한다. 문서가 MMR 컴퓨터(112)의 바탕화면 상에 보인 때(GDI 문자 발생기를 모니터링하는 것에 의해 결정됨), 문서가 열려 있는 동안에 액세스된 URL, 또는 문서가 열려 있는 동안에 키보드 상에서 타이핑된 문자 등의 문서와 연관되어 있는 이벤트를 캡처하는 이벤트 캡처 모듈(324)은 MMR 컴퓨터(112)에 추가된다.
- [0190] MMR 시스템(100)의 제4 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어, 및 프린터(116)를 포함한다. 이 제4 대안의 실시예에서, 프린터(116)는 근접하여 있는 임의의 MMR 사용자(110)의 캡처 장치(106)와 통신하는 블루투스 송수신기 또는 유사한 통신 링크를 갖추고 있다. MMR 사용자(110)가 프린터(116)로부터 문서를 집을 때마다, 프린터(116)는 MMR 데이터(문서 레이아웃 및 멀티미디어 클립)를 그 사용자의 캡처 장치(106)로 푸시(push)한다. 사용자 프린터(116)는 키패드를 포함하며, 이를 사용하여 사용자는 특정의 문서와 연관되어 있는 멀티미디어 데이터를 획득하기 위해 로그인하고 코드를 입력한다. 문서는 프린터 드라이버(316)에 의해 삽입될 수 있는 그의 꼬리말/footer)에 코드의 인쇄된 표시를 포함할 수 있다.
- [0191] MMR 시스템(100)의 제5 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어, 및 사무실 포털(120)을 포함한다. 사무실 포털 장치는 양호하게는 사무실 포털(120)의 개인화된 버전이다. 사무실 포털(120)은, 대화, 회의/전화 통화 및 회합 등의, 사무실에서의 이벤트를 캡처한다. 사무실 포털(120)은 물리적 책상 위의 특정의 종이 문서를 식별하고 추적한다. 사무실 포털(120)은 그에 부가하여 문서 식별 소프트웨어를 실행한다(즉, 문서 지문 정합 모듈(226) 및 문서 이벤트 데이터베이스(320)를 호스팅한다). 이 제5 대안의 실시예는 MMR 컴퓨터(112)로부터 계산 작업 부하를 덜어주는 역할을 하며, MMR 시스템(100b)을 소비자 장치로서 패키징하는 편리한 방법을 제공한다(예를 들어, MMR 시스템(100b)은 Apple Computer, Inc.의 Mac Mini 컴퓨터 상에서 실행되고 있는 하드웨어 및 소프트웨어 제품으로서 판매된다).
- [0192] MMR 시스템(100)의 제6 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어, 및 네트워크화된 미디어 서버(114)를 포함한다. 이 실시예에서, 멀티미디어 데이터는 Comcast Video-on-Demand 서버 등의 네트워크화된 미디어 서버(114) 상에 존재한다. MMR 사용자(110)가 그의 캡처 장치(106)를 사용하여 문서 텍스트 패치를 스캔할 때, 그 결과 얻어지는 탐색 명령(lookup command)이 MMR 사용자(110)의 케이블 TV와 연관되어 있는 셋톱 박스(126)로 (무선으로, 인터넷을 통해 또는 전화로 셋톱 박스(126)를 호출함으로써) 또는 Comcast 서버에 전송된다. 양 경우에, 멀티미디어는 Comcast 서버로부터 셋톱 박스(126)로 스트리밍된다. 시스템(100)은 데이터를 어디에 전송할지를 알고 있는데, 그 이유는 MMR 사용자(110)가 이미 그의 전화를 등록했기 때문이다. 따라서, 캡처 장치(106)는 셋톱 박스(126)의 액세스 및 제어를 위해 사용될 수 있다.
- [0193] MMR 시스템(100)의 제7 대안의 실시예는 캡처 장치(106), 캡처 장치 소프트웨어, 네트워크화된 미디어 서버(114) 및 위치 서비스를 포함한다. 이 실시예에서, 위치-인식 서비스는 Comcast 시스템(또는 다른 적당한 통신 시스템)으로부터의 출력에 대한 다수의 목적지를 구별한다. 이 기능은 셀룰러 전화 타워 ID들을 자동적으로 구별함으로써 또는 MMR 사용자(110)로 하여금 데이터가 디스플레이될 위치를 선택하게 하는 키패드 인터페이스에 의해 수행된다. 따라서, 사용자는, 다른 장소를 방문하고 있는 동안에 그 다른 장소가 케이블에 액세스하고 있는 한, 그의 케이블 통신 사업자에 의해 제공되는 프로그래밍 및 다른 케이블 TV 피쳐들에 액세스할 수 있다.
- [0194] 문서 지문 정합("이미지-기반 패치 인식")
- [0195] 앞서 기술한 바와 같이, 문서 지문 정합은 MMR 문서의 일부분 또는 "패치"를 일의적으로 식별하는 것을 필요로

한다. 도 6을 참조하면, 문서 지문 정합 모듈/시스템(610)은 캡처된 이미지(612)를 수신한다. 문서 지문 정합 시스템(610)은 이어서 문서 데이터베이스(3400) 내의 페이지 컬렉션에 질의를 하고(예를 들어, 도 34a를 참조하여 이하에서 더 기술함), 캡처된 이미지(612)가 포함되어 있는 것들을 포함하는 페이지 및 문서의 리스트를 반환한다. 각각의 결과는 캡처된 입력 이미지(612)가 있는 x-y 위치이다. 당업자라면 데이터베이스(3400)가 문서 지문 정합 모듈(610) 외부에 있을 수 있지만(예를 들어, 도 6에 도시됨), 또한 문서 지문 정합 모듈(610) 내부에 있을 수도 있다(예를 들어, 문서 지문 정합 모듈(610)은 데이터베이스(3400)를 포함하는 도 7, 도 11, 도 12, 도 14, 도 20, 도 24, 도 26, 도 28 및 도 30 내지 도 32에 도시되어 있음)는 것을 잘 알 것이다.

[0196] 도 7은 본 발명의 일 실시예에 따른 문서 지문 정합 시스템(610)의 블록도를 나타낸 것이다. 캡처 장치(106)는 이미지를 캡처한다. 캡처된 이미지는 품질 평가 모듈(712)에 전송되고, 이 품질 평가 모듈(712)은 다운스트림 처리의 필요성 및 기능에 기초하여 캡처된 이미지의 콘텐츠에 관한 예비적인 판단을 효과적으로 행한다. 예를 들어, 캡처된 이미지가 다운스트림에 있는 문서 지문 정합 시스템(610)에서 처리될 수 없을 정도의 품질을 가지고 있는 경우, 품질 평가 모듈(712)은 캡처 장치(106)로 하여금 더 높은 해상도로 그 이미지를 재캡처하게 한다. 게다가, 품질 평가 모듈(712)은, 예를 들어, 캡처된 이미지가 "초점이 맞는지" 여부의 표시인 캡처된 이미지에 포함되어 있는 텍스트의 선명함(sharpness) 등의 캡처된 이미지의 많은 다른 관련 특성들을 검출할 수 있다. 게다가, 품질 평가 모듈(712)은 캡처된 이미지가 문서의 일부일 수 있는 무언가를 포함하고 있는지 여부를 결정할 수 있다. 예를 들어, 비문서 이미지(예를 들어, 책상, 실외 장면)를 포함하는 이미지 패치는 사용자가 캡처 장치(106)의 시야를 새 문서로 전환시키고 있음을 나타낸다.

[0197] 게다가, 하나 이상의 실시예들에서, 품질 평가 모듈(712)은 인식가능한 텍스트를 포함할 가능성이 있는 이미지만을 통과시키기 위해 텍스트/비텍스트 구별을 수행할 수 있다. 도 8은 하나 이상의 실시예들에 따른 텍스트/비텍스트 구별을 위한 흐름 프로세스를 나타낸 것이다. 단계(810)에서, 다수의 컬럼의 픽셀이 입력 이미지 패치로부터 추출된다. 통상적으로, 입력 이미지는 그레이-스케일이고, 컬럼에서의 각각의 값은 0 내지 255의 정수이다(8 비트 픽셀의 경우). 단계(812)에서, 각각의 컬럼에서의 로컬 피크(local peak)가 검출된다. 이것은 통상 알고 있는 "슬라이딩 윈도우" 방법(고정된 길이(예를 들어, N개의 픽셀)의 윈도우가 컬럼 상을 한번에 M개의 픽셀씩 슬라이딩함, 단  $M < N$ 임)으로 행해질 수 있다. 각각의 단계에서, 그레이 레벨 값의 상당한 차이(예를 들어, 40보다 큼)를 찾아냄으로써 피크의 존재가 결정된다. 피크가 윈도우의 한 위치에 있는 경우, 슬라이딩 윈도우가 이 위치를 덮고 있을 때마다 다른 피크들의 검출이 억제된다. 연속적인 피크들 간의 갭도 역시 단계(812)에서 검출될 수 있다. 단계(812)는 이미지 패치에서의 C개의 컬럼들에 적용되고, 단계(814)에서 간극 값이 히스토그램 형태로 누적된다.

[0198] 이 갭 히스토그램이 데이터베이스(818)에 저장된 공지의 분류(단계(816))를 갖는 훈련 데이터로부터 도출된 다른 히스토그램들과 비교되고, 패치의 카테고리(텍스트 또는 비텍스트)에 관한 결정이 그 결정에서의 신뢰 척도와 함께 출력된다. 단계(816)에서의 히스토그램 분류는 텍스트의 이미지로부터 도출된 히스토그램의 통상적인 모습과 2개의 대등한 피크(하나는 라인들 간의 거리의 중앙에 있고, 아마도 하나 또는 2개의 다른 훨씬 더 작은 피크가 히스토그램에서 다른 피크들로부터 멀리 떨어져 정수배 더 높음)를 포함한다는 것을 고려한다. 이 분류는 통계적 분산의 척도로 히스토그램의 형상을 결정할 수 있거나, 이는 예를 들어, 해밍 거리 또는 유클리드 거리 등의 거리 척도를 사용하여 히스토그램을 저장된 프로토타입과 하나씩 비교할 수 있다.

[0199] 이제 도 9도 참조하면, 동 도면은 텍스트/비텍스트 구별의 일례를 나타낸 것이다. 입력 이미지(910)가 처리되어 다수의 컬럼을 샘플링하고, 그의 서브셋이 점선으로 나타내어져 있다. 통상적인 컬럼(912)에 대한 그레이 레벨 히스토그램이 914에 나타내어져 있다. Y 값은 910에서 그레이 레벨이고 X 값은 910에서 행이다. 히스토그램에서 피크들 간의 검출된 갭은 916에 나타내어져 있다. 샘플링된 컬럼들 전부로부터의 갭 값의 히스토그램은 918에 나타내어져 있다. 이 예는 텍스트를 포함하는 패치로부터 도출된 히스토그램의 형상을 나타낸 것이다.

[0200] 이미지 패치에서의 텍스트의 포인트 크기를 추정하는 흐름 프로세스가 도 10에 도시되어 있다. 이 흐름 프로세스는 이미지에서의 흐려짐(blur)이 캡처 장치의 페이지로부터의 거리에 역비례한다는 사실을 이용한다. 흐려짐의 정도를 추정함으로써, 거리가 추정될 수 있으며, 그 거리는 이미지에서의 객체의 크기를 기지의 "정규화된" 높이로 스케일링하는 데 사용될 수 있다. 이 동작은 새 이미지에서의 텍스트의 포인트 크기를 추정하는 데 사용될 수 있다.

[0201] 훈련 과정(1010)에서, 단계(1012)에서, 기지의 폰트 및 포인트 크기의 텍스트 패치의 이미지("캘리브레이션 이미지(calibration image)"라고 함)가 기지의 거리에 있는 이미지 캡처 장치로 획득된다. 단계(1014)에서, 픽셀

의 수로 표시된 그 이미지에서의 텍스트 문자의 높이가 측정된다. 이것은, 예를 들어, 마이크로소프트 Photo Editor 등의 이미지 주석 첨부 도구를 사용하여 수작업으로 행해질 수 있다. 단계(1016)에서, 캘리브레이션 이미지에서의 흐려짐이 추정된다. 이것은, 예를 들어, 2차원 고속 푸리에 변환의 차단 스펙트럼(spectral cutoff)의 기지의 측정으로 행해진다. 이것은 또한 픽셀 수(1020)의 단위로 표시현될 수 있다.

[0202] 단계(1024)에서 "새" 이미지를 제공할 때, MMR 인식 시스템에서 런타임 시와 같이, 단계(1026)에서 이 이미지가 처리되어 각각의 문자 주변에 경계 박스를 생성하는 라인 세그먼트화 및 문자 세그먼트화의 통상적으로 알고 있는 방법으로 텍스트를 찾아낸다. 이들 박스의 높이는 픽셀 단위로 표시될 수 있다. 단계(1028)에서, 새 이미지의 흐려짐이 단계(1016)와 유사한 방식으로 추정된다. 단계(1030)에서, 이들 척도가 결합되어 각각의 문자(또는 등가적으로, 각각의 라인)의 포인트 크기의 제1 추정치(1032)를 발생한다. 이것은 이하의 방정식을 계산함으로써 행해질 수 있다:  $(\text{캘리브레이션 이미지 흐려짐 크기} / \text{새 이미지 흐려짐 크기}) * (\text{새 이미지 텍스트 높이} / \text{캘리브레이션 이미지 텍스트 높이}) * (\text{포인트 단위의 캘리브레이션 이미지 폰트 크기})$ . 이것은 캘리브레이션 이미지의 텍스트의 포인트 크기를 스케일링하여 입력 이미지 패치에서의 텍스트의 추정된 포인트 크기를 생성한다. 동일한 스케일링 함수가 모든 문자의 경계 박스의 높이에 적용될 수 있다. 이것은 패치에서의 모든 문자에 대한 결정을 생성한다. 예를 들어, 패치가 50개 문자를 포함하는 경우, 이 절차는 패치에서의 폰트의 포인트 크기에 대한 50 투표(vote)를 생성한다. 이어서, 투표들의 메디안(median)으로 포인트 크기에 대한 단일의 추정치가 도출될 수 있다.

[0203] 게다가, 보다 구체적으로 도 7을 다시 참조하면, 하나 이상의 실시예들에서, 품질 평가 모듈(712)의 캡처 장치(106)로의 피드백이 캡처 장치(106)의 사용자 인터페이스(UI)로 보내질 수 있다. 예를 들어, 이 피드백은 캡처된 이미지가 텍스트처럼 보이지만 흐려져 있는 무언가를 포함하고 있고 또 사용자가 캡처 장치(106)를 안정화시켜야만 한다는 것을 나타내는 사운드 또는 진동의 형태의 표시를 포함할 수 있다. 이 피드백은 또한 캡처된 이미지의 품질을 향상시키기 위해 캡처 장치(106)의 광학계의 파라미터를 변경하는 명령을 포함할 수 있다. 예를 들어, 캡처된 이미지의 품질을 향상시키기 위해 초점, F-스톱(F-stop), 및/또는 노출 시간이 조정될 수 있다.

[0204] 게다가, 품질 평가 모듈(712)의 캡처 장치(106)로의 피드백은 사용되고 있는 특정의 피쳐 추출 알고리즘의 요구에 의해 전문화될 수 있다. 이하에서 더 기술하는 바와 같이, 피쳐 추출은 이미지를 기호적인 표시로 변환한다. 단어의 길이를 계산하는 인식 시스템에서, 캡처 장치(106)의 광학계가 캡처된 이미지를 흐려지게 하는 것이 바람직할 수 있다. 당업자라면 이러한 조정이, 비록 아마도 사람 또는 광학 문자 인식(OCR) 프로세스에 의해 인식가능하지 않지만, 피쳐 추출 기법에 아주 적합한 이미지를 생성할 수 있다는 것을 잘 알 것이다. 품질 평가 모듈(712)은 캡처 장치(106)로 명령어를 피드백하여 캡처 장치(106)로 하여금 렌즈의 초점을 흐리게 하고 그에 의해 흐려진 이미지를 생성함으로써 이것을 구현할 수 있다.

[0205] 피드백 프로세스는 제어 구조(control structure)(714)에 의해 수정된다. 일반적으로, 제어 구조(714)는 문서 지문 정합 시스템(610) 내의 다른 컴포넌트들로부터 데이터 및 기호적인 정보를 수신한다. 제어 구조(714)는 문서 지문 정합 시스템(610)에서의 다양한 단계들의 실행 순서를 결정하고 계산 부하를 최적화할 수 있다. 제어 구조(714)는 수신된 이미지 패치의 x-y 위치를 식별한다. 보다 상세하게는, 제어 구조(714)는 피쳐 추출 프로세스의 요구, 품질 평가 모듈(712)의 결과, 및 캡처 장치(106) 파라미터에 관한 정보를 수신하고, 이들을 적절히 변경할 수 있다. 이것은 프레임별로 동적으로 행해질 수 있다. 다수의 피쳐 추출 방법을 사용하는 시스템 구성에서, 한 방법은 큰 텍스트 패치들의 흐린 이미지를 필요로 하고 다른 방법은 고해상도 선명한 초점의 종이결(paper grain) 이미지를 필요로 할 수 있다. 이러한 경우에, 제어 구조(714)는 품질 평가 모듈(712)이 텍스트를 보고 있을 때 그에 적절한 이미지 품질을 생성하도록 지시하는 명령을 품질 평가 모듈(712)에 전송할 수 있다. 품질 평가 모듈(712)은 캡처 장치(106)와 상호작용하여 정확한 이미지를 생성하게 된다(예를 들어, 큰 패치의 N개의 흐린 이미지 및 그에 뒤이은 선명한 초점의 종이결의 M개의 이미지(고해상도)). 제어 구조(714)는 대응하는 피쳐 추출 및 분류가 적용되도록 하기 위해 처리 파이프라인을 통해 그 이미지들의 진행을 추적한다.

[0206] 이미지 처리 모듈(716)은 인식 시스템의 요구에 기초하여 입력 이미지의 품질을 수정한다. 이미지 수정의 유형들의 예는 샤프닝(sharpening), 디스큐잉(deskewing) 및 이진화(binartization)를 포함한다. 이러한 알고리즘은 마스크 크기, 예상된 회전(expected rotation) 및 임계값 등의 많은 조정가능한 파라미터를 포함한다.

[0207] 도 7에 나타난 바와 같이, 문서 지문 정합 시스템(610)은 이미지 처리 모듈(716)의 파라미터를 동적으로 수정하기 위해 피쳐 추출 및 분류 모듈(718, 720)(이하에서 기술함)로부터의 피드백을 사용한다. 이것이 동작하는 이유는 사용자가 통상적으로 그의 캡처 장치(106)를 수초 동안 계속하여 문서 내의 동일한 위치로 향하게 하기 때



문이다. 예를 들어, 캡처 장치(106)가 초당 30 프레임을 처리하고, 임의의 시퀀스에서의 처음 수 프레임을 처리한 결과가 나중에 캡처된 프레임이 어떻게 처리되는지에 영향을 줄 수 있다.

[0208] 피쳐 추출 모듈(718)은 캡처된 이미지를 기호적인 표시로 변환한다. 한 예에서, 피쳐 추출 모듈(718)은 단어를 찾아내고 그의 경계 박스를 계산한다. 다른 예에서, 피쳐 추출 모듈(718)은 연결된 컴포넌트를 찾아내고 그 형상에 대한 기술자를 계산한다. 게다가, 하나 이상의 실시예에서, 문서 지문 정합 시스템(610)은 피쳐 추출의 결과에 관한 메타데이터를 제어 구조(714)와 공유하고 그 메타데이터를 사용하여 다른 시스템 컴포넌트의 파라미터를 조정한다. 당업자라면 이것이 계산 요구사항을 상당히 감소시킬 수 있고 품질이 좋지 않은 데이터의 인식을 금지시킴으로써 정확도를 향상시킬 수 있다는 것을 잘 알 것이다. 예를 들어, 단어 경계 박스를 식별해주는 피쳐 추출 모듈(718)은 그가 발견한 라인 및 "단어"의 수를 제어 구조(714)에 알려준다. 단어의 수가 너무 높은 경우(예를 들어, 입력 이미지가 프래그먼트화되어 있음을 나타냄), 제어 구조(714)는 품질 평가 모듈(712)에 더 흐린 이미지를 생성하도록 지시할 수 있다. 품질 평가 모듈(712)은 이어서 캡처 장치(106)에 적절한 신호를 전송한다. 다른 대안으로서, 제어 구조(714)는 이미지 처리 모듈(716)에 평활화 필터(smoothing filter)를 적용하도록 지시할 수 있다.

[0209] 분류 모듈(720)은 피쳐 추출 모듈(718)로부터의 피쳐 기술(feature description)을 문서 내의 하나 이상의 페이지 및 입력 이미지 패치(input image patch)가 일어나는 그 페이지 내의 x,y 위치의 ID로 변환한다. 차례로 설명하는 바와 같이 데이터베이스(3400)로부터의 피드백에 의존하여 식별이 행해진다. 게다가, 하나 이상의 실시예들에서, 신뢰도없이 각각의 결정과 연관될 수 있다. 문서 지문 정합 시스템(610)은 이러한 결정을 사용하여 시스템 내의 다른 컴포넌트들의 파라미터를 결정할 수 있다. 예를 들어, 제어 구조(714)는 상위 2개의 결정의 신뢰도가 서로 비슷한 경우, 이미지 처리 알고리즘의 파라미터가 변경되어야 하는 것으로 결정할 수 있다. 이 결과 메디안 필터에 대한 크기 범위 및 그 결과의 나머지 컴포넌트들로의 전달(carry-through)이 증가할 수 있다.

[0210] 게다가, 도 7에 도시한 바와 같이, 분류 모듈(720)과 데이터베이스(3400) 사이에 피드백이 있을 수 있다. 게다가, 당업자라면 데이터베이스(3400)가 도 6에 도시한 바와 같이 모듈(610) 외부에 있을 수 있다는 것을 잘 알 것이다. 패치의 ID에 관한 결정이 유사한 모습을 갖는 다른 패치들에 대한 데이터베이스(3400)에 질의하는 데 사용될 수 있다. 이것은, 입력 이미지 패치를 데이터베이스(3400)와 비교하기 보다는, 데이터베이스(3400)에 저장된 패치의 완전한 이미지 데이터를 데는 데이터베이스(3400) 내의 다른 이미지와 비교한다. 이것은 분류 모듈(720)의 결정에 대한 추가적인 확인 레벨을 제공할 수 있으며 정합하는 데이터의 어떤 전처리를 가능하게 해줄 수 있다.

[0211] 데이터베이스 비교는 또한 이미지 데이터만이라기 보다는 패치의 기호적인 표시에 대해 행해질 수 있다. 예를 들어, 최상의 결정은 이미지 패치가 12-포인트 Arial 폰트 한줄 띄기(double-spaced)를 포함한다는 것을 나타낼 수 있다. 데이터베이스 비교는 이미지 비교보다는 단지 텍스트 메타데이터만을 사용하여 유사한 폰트, 간격, 및 워드 레이아웃(word layout)을 갖는 다른 문서들에 있는 패치를 찾아낼 수 있다.

[0212] 데이터베이스(3400)는 몇가지 유형의 콘텐츠-기반 질의를 지원할 수 있다. 분류 모듈(720)은 데이터베이스(3400)에 피쳐 배열을 전달할 수 있고 그 배열이 행해지는 문서 및 x-y 위치의 리스트를 수신할 수 있다. 예를 들어, 피쳐들은 수평으로 또는 수직으로 단어 길이들의 트라이그램(trigram)(이하에서 기술함)일 수 있다. 데이터베이스(3400)는 어느 한 유형의 질의에 응답하여 결과들의 리스트를 반환하도록 구성될 수 있다. 분류 모듈(720) 또는 제어 구조(714)는 그 순위들을 결합하여 하나의 정렬된 결정 리스트를 생성할 수 있다.

[0213] 게다가, 데이터베이스(3400), 분류 모듈(720) 및 제어 구조(714) 간에 피드백이 있을 수 있다. 피쳐 벡터로부터 위치를 확인하기에 충분한 정보를 저장하는 것에 부가하여, 데이터베이스(3400)는 문서의 원시 이미지는 물론 그것의 그래픽 컴포넌트에 대한 기호적인 표시를 비롯한 관련 정보를 저장할 수 있다. 이것은 제어 구조(714)가 동작 중에 다른 시스템 컴포넌트의 거동을 수정할 수 있게 해준다. 예를 들어, 주어진 이미지 패치에 대해 2개의 그럴듯한 결정이 있는 경우, 데이터베이스(3400)는 이미지의 존재를 위한 오른쪽에 있는 영역을 좁아주어 조사함으로써 그 결정들이 명확하게 될 수 있음을 나타낼 수 있다. 제어 구조(714)는 캡처 장치(106)에 좁아주라라고 지시하는 적절한 메시지를 캡처 장치(106)에 전송할 수 있다. 피쳐 추출 모듈(718) 및 분류 모듈(720)은 문서 상에 인쇄된 이미지에 대한 이미지의 우측면을 검사할 수 있다.

[0214] 게다가, 유의할 점은, 패치가 정확하게 문서 내에 위치하고 있다고 하면, 데이터베이스(3400)가 이미지 패치를 둘러싸고 있는 데이터에 관한 상세한 정보를 저장한다는 것이다. 이것은 종래의 기술에서 예견되지 않은 추가적인 하드웨어 및 소프트웨어 이미지 분석 단계들을 트리거하는 데 사용될 수 있다. 한 경우에 그 상세한 정보

는 문서의 상세한 기호적인 설명을 저장하는 인쇄 캡처 시스템에 의해 제공된다. 하나 이상의 다른 실시예에서, 유사한 정보는 문서를 스캐닝함으로써 획득될 수 있다.

[0215] 여전히 도 7을 참조하면, 위치 추적 모듈(724)은 제어 구조(714)로부터 이미지 패치의 ID에 관한 정보를 수신한다. 위치 추적 모듈(724)은 그것을 사용하여 데이터베이스(3400)로부터 전체 문서 페이지의 복사본 또는 문서를 기술하는 데이터 구조를 검색할 수 있다. 이 초기 위치는 위치 추적 프로세스의 시작에 대한 앵커이다. 품질 평가 모듈(712)이 캡처된 이미지가 추적에 적합하다고 결정할 때, 위치 추적 모듈(724)은 캡처 장치(106)로부터 이미지 데이터를 수신한다. 위치 추적 모듈(724)은 또한 마지막 프레임이 성공적으로 인식된 후에 경과한 시간에 관한 정보도 갖는다. 위치 추적 모듈(724)은 이 모듈이 캡처 장치(106)가 연속적인 프레임들 간에 이동된 문서 상에서의 거리를 추정할 수 있게 해주는 광 흐름 기법을 적용한다. 캡처 장치(106)의 샘플링 레이트가 주어진 경우, 캡처 장치(106)가 보는 데이터가 인식가능하지 않을지도 모르지만, 그의 대상(target)이 추정될 수 있다. 캡처 장치(106)의 추정된 위치가 그의 이미지 데이터를 데이터베이스 문서로부터 도출된 대응하는 이미지 데이터와 비교함으로써 확인될 수 있다. 간단한 예는 캡처된 이미지의 데이터베이스(3400) 내의 예상된 이미지와의 교차 상관을 계산한다.

[0216] 따라서, 위치 추적 모듈(724)은 위치 추적 알고리즘의 진행을 안내하기 위해 데이터베이스 이미지의 상호작용적 사용에 대비하고 있다. 이것은 그래픽 및 이미지 등의 비텍스트 객체에 대한 전자적 상호작용의 첨부를 고려하고 있다. 게다가, 하나 이상의 다른 실시예에서, 이러한 첨부는 상기한 이미지 비교/확인 단계 없이 구현될 수 있다. 환언하면, 페이지 상에서의 캡처 장치(106)의 순간 움직임을 추정함으로써, 캡처된 이미지와 독립적으로 보여야만 하는 전자 링크가 추정될 수 있다.

[0217] 도 11은 본 발명의 일 실시예에 따른 문서 지문 정합 기법을 나타낸 것이다. 도 11에 나타낸 "피드-포워드" 기법은 각각의 패치를 독립적으로 처리한다. 이 기법은 하나 이상의 페이지 및 패치가 있는 그 페이지 상에서의 x-y 위치를 찾아내는 데 사용되는 피쳐들을 이미지 패치로부터 추출한다. 예를 들어, 하나 이상의 실시예에서, 문서 지문 정합을 위한 피쳐 추출은 캡처된 이미지의 피쳐들의 수평 및 수직 그룹(예를 들어, 단어, 문자, 블록)에 의존할 수 있다. 추출된 피쳐들의 이들 그룹은 이어서 추출된 피쳐들을 포함하는 문서들(및 그 문서들 내의 패치)을 탐색하는 데 사용될 수 있다. 캡처된 이미지에서 수평 단어 쌍을 식별하기 위해 OCR 기능이 사용될 수 있다. 각각의 식별된 수평 단어쌍은 이어서 식별된 수평 단어쌍을 포함하는 모든 문서 및 그 문서들 내의 단어쌍의 x-y 위치를 결정하기 위해 데이터베이스(3400)에 대한 검색 질의를 형성하는 데 사용된다. 예를 들어, 수평 단어쌍 "the, cat"에 대해, 데이터베이스(3400)는 (15,x,y), (20,x,y)를 반환하여, 문서(15, 20)에서 표시된 x-y 위치에 수평 단어쌍 "the, cat"이 있음을 나타낼 수 있다. 이와 유사하게, 각각의 수직으로 인접한 단어쌍에 대해, 단어쌍 및 그 문서 내에서의 단어쌍의 x-y 위치의 예를 포함하는 모든 문서에 대해 데이터베이스(3400)에 질의된다. 예를 들어, 수직으로 인접한 단어쌍 "in, hat"에 대해, 데이터베이스(3400)는 (15,x,y), (7,x,y)를 반환하여, 수직으로 인접한 단어쌍 "in, hat"가 문서(15, 7)에서 표시된 x-y 위치에 있다는 것을 나타낼 수 있다. 그 다음에, 데이터베이스(3400)에 의해 반환된 문서 및 위치 정보를 사용하여, 캡처된 이미지로부터 추출된 다양한 수평 단어쌍 및 수직으로 인접한 단어쌍 간에 어느 문서가 가장 많은 위치 중첩이 있는지에 관한 결정이 행해질 수 있다. 이 결과, 핫스팟 및 링크된 미디어의 존재가 결정될 수 있는지에 응답하여 캡처된 이미지를 포함하는 문서를 식별할 수 있다.

[0218] 도 12는 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 12에 나타낸 "상호작용적 이미지 분석(interactive image analysis)" 기법은 이미지 패치가 인식되기 이전에 일어날 수 있는 이미지 처리와 피쳐 추출 간의 상호작용을 필요로 한다. 예를 들어, 이미지 처리 모듈(716)은 먼저 입력 이미지에서 흐려짐(blur)을 추정할 수 있다. 이어서, 피쳐 추출 모듈(718)은 그 페이지로부터의 거리 및 이미지 텍스트의 포인트 크기를 계산한다. 이어서, 이미지 처리 모듈(716)은 그 포인트 크기의 폰트의 특성을 사용하여 이미지에 대한 템플릿 정합 단계를 수행할 수 있다. 그 다음에, 피쳐 추출 모듈(718)은 이어서 그 결과로부터 문자 또는 단어 피쳐를 추출할 수 있다. 게다가, 당업자라면 폰트, 포인트 크기 및 피쳐가 데이터베이스(3400) 문서들에서의 폰트에 의해 제약될 수 있다는 것을 잘 알 것이다.

[0219] 도 12를 참조하여 상기한 상호작용적 이미지 분석의 일례가 도 13에 도시되어 있다. 단계(1310)에서 이미지 패치 내의 텍스트의 폰트 및 포인트 크기는 물론 이미지 패치의 카메라로부터의 거리를 추정하기 위해 입력 이미지 패치가 처리된다. 당업자라면 폰트 추정(즉, 패치 내의 텍스트의 폰트에 대한 후보들의 식별)이 공지된 기법으로 행해질 수 있다는 것을 잘 알 것이다. 포인트 크기 및 거리 추정은, 예를 들어, 도 10을 참조하여 기술한 흐름 프로세스를 사용하여 수행될 수 있다. 게다가, 캡처 장치에 맞게 손쉽게 조정될 수 있는 초점으로부터

의 거리의 공지의 방법 등의 다른 기법들이 사용될 수 있다.

- [0220] 여전히 도 13을 참조하면, 단계(1312)에서 패치 내의 텍스트의 라인들 주위에 경계 박스를 구성하는 라인 세그먼트화 알고리즘이 적용된다. 단계(1314)에서 비례 스케일링(proportional scaling) 등의 공지의 기법들을 사용하여 각각의 라인 이미지의 높이는 고정된 크기로 정규화된다. 이미지에서 검출된 폰트에 대한 ID는 물론 그것의 포인트 크기가 폰트 프로토타입의 컬렉션(1322)에 전달(1324)되며, 여기에서 이들은 각각의 지명된 폰트에서의 문자들에 대한 이미지 프로토타입을 검색하는 데 사용된다.
- [0221] 폰트 데이터베이스(1322)는 문서를 인쇄하기 위해 운영 체제 및 기타 소프트웨어 애플리케이션에 의해 사용되는 사용자의 시스템 상의 폰트 컬렉션(예를 들어, 트루타입(TrueType), 오픈타입(OpenType), Microsoft Windows에서의 래스터 폰트(raster font))으로부터 구성될 수 있다. 하나 이상의 다른 실시예들에서, 폰트 컬렉션은 데이터베이스(3400) 내의 문서들의 원시 이미지로부터 생성될 수 있다. 데이터베이스(3400) xml 파일들은 원시 이미지로부터 문자들의 프로토타입 이미지를 추출하는 데 사용될 수 있는 x-y 경계 박스 좌표를 제공한다. xml 파일은 폰트의 이름 및 문자의 포인트 크기를 정확하게 식별한다.
- [0222] 단계(1320)에서, 선택된 폰트에서의 문자 프로토타입은 단계(1314)에서 사용된 파라미터의 함수에 기초하여 크기 정규화된다. 단계(1316)에서의 이미지 분류는 단계(1320)에서 출력된 크기 정규화된 문자를 단계(1314)에서의 출력과 비교하여, 이미지 패치에서의 각각의 x-y 위치에서 결정을 생성할 수 있다. 공지의 이미지 템플릿 정합 방법은  $(c_i, x_i, y_i, w_i, h_i)$  등의 출력을 생성하는 데 사용될 수 있으며, 여기서  $c_i$ 는 이미지 패치에서 검출된 모든 문자  $i$  ( $i = 1 \dots n$ )에 대한 문자의 ID이고,  $(x_i, y_i)$ 는 그의 경계 박스의 좌측 상부 모서리이며,  $h_i, w_i$ 는 그의 폭과 높이이다.
- [0223] 단계(1318)에서, 기하학적 관계-제약이 있는(geometric relation-constrained) 데이터베이스 탐색이 상기한 바와 같이 수행될 수 있지만, 단어쌍 대신에 문자쌍에 대한 경우에 전문화될 수 있다. 이러한 경우에, "a-b"는 문자 a 및 b가 수평으로 인접해 있다는 것을 나타낼 수 있고, "a+b"는 이들이 수직으로 인접해 있다는 것을 나타낼 수 있으며, "a/b"는 a가 b의 남서쪽에 있다는 것을 나타낼 수 있고, "a\b"는 a가 b의 남동쪽에 있다는 것을 나타낼 수 있다. 기하학적 관계는 각각의 문자쌍의  $x_i, y_i$  값으로부터 도출될 수 있다. MMR 데이터베이스(3400)는 단어쌍 대신에 문자쌍을 포함하는 문서 페이지들의 리스트를 반환하도록 구성될 수 있다. 단계(1326)에서의 출력은 점수로 순위가 매겨지는 x-튜플( $document_i, page_i, x_i, y_i, action_i, score_i$ )로서 표시되는 입력 이미지와 정합하는 후보들의 리스트이다.
- [0224] 도 14는 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 14에 나타낸 "생성 및 테스트(generate and test)" 기법은 각각의 패치를 독립적으로 처리한다. 이 기법은 주어진 이미지 패치를 포함할 수 있는 다수의 페이지 이미지들을 찾아내는 데 사용되는 이미지 패치로부터 피쳐들을 추출한다. 게다가, 하나 이상의 실시예들에서, 페이지가 이미지 패치를 포함할 가능성에 의해 페이지들에 순위를 매기기 위해 부가적인 추출-분류 단계가 수행될 수 있다.
- [0225] 여전히 도 14를 참조하여 상기한 "생성 및 테스트" 기법을 참조하면, 캡처된 이미지의 피쳐들이 추출될 수 있고 이들 추출된 피쳐들을 가장 많이 포함하고 있는 데이터베이스(3400) 내의 문서 패치가 식별될 수 있다. 가장 정합하는 피쳐를 갖고 있는 처음 X개의 문서 패치들("후보들")이 이어서 추가적으로 처리된다. 이 처리에서, 정합하는 문서 패치에서의 피쳐들의 상대적 위치는 질의 이미지에서의 피쳐들의 상대 위치와 비교된다. 이 비교에 기초하여 점수가 계산된다. 이어서, 최상의 정합하는 문서 패치 P에 대응하는 최고 점수가 식별된다. 최고 점수가 적응적 임계값보다 큰 경우, 문서 패치 P는 질의 이미지와 정합하는 것으로 밝혀진다. 이 임계값은, 예를 들어, 추출된 피쳐들의 수를 비롯한, 많은 파라미터에 대해 적응적이다. 데이터베이스(3400)에서, 문서 패치 P가 어디에서 오는 것인지를 알게 되고, 따라서 질의 이미지가 동일한 장소로부터 온 것으로 결정된다.
- [0226] 도 15는 단어 경계 박스 검출 알고리즘의 일례를 나타낸 것이다. 입력 이미지 패치(1510)는 회전을 정정하는 이미지 처리 이후에 도시되어 있다. 스큐 정정 알고리즘이라고 통상적으로 알려져 있는 이러한 부류의 기법은 텍스트 이미지를 회전시켜 수평축과 일렬로 정렬되게 한다. 경계 박스 검출 알고리즘에서의 그 다음 단계는 수평 투사 프로파일(1512)의 계산이다. "임계값을 넘는" 영역들이 텍스트의 라인들에 대응하도록, 공지의 적응적 임계값 지정 또는 슬라이딩 윈도우 알고리즘(adaptive thresholding or sliding window algorithm)에 의해 라인 검출을 위한 임계값이 선택된다(1516). 라인 내의 단어들을 나타내는 임계값을 넘는 영역들을 찾아내기 위해, 각각의 라인 내의 영역들이 유사한 방식으로 추출되고 처리된다(1514, 1518). 한 텍스트 라인에서 검출된 경계 박스의 일례가 1520에 도시되어 있다.

- [0227] 문서 패치 후보들과 비교하기 위해 다양한 피쳐들이 추출될 수 있다. 예를 들어, SIFT(Scale Invariant Feature Transform) 피쳐들, 모서리 피쳐들, 돌출 포인트, 어센더(ascender), 디센더(descender), 단어 경계 및 공백이 정합을 위해 추출될 수 있다. 문서 이미지로부터 신뢰성있게 추출될 수 있는 피쳐들 중 하나가 단어 경계이다. 일단 단어 경계가 추출되면, 이들은 도 16에 도시된 바와 같이 그룹들로 형성될 수 있다. 도 16에서, 예를 들어, 수직 그룹은 단어 경계가 위아래 양쪽으로 중첩하는 단어 경계를 가지며 중첩하는 단어 경계의 총 수가 적어도 3개이도록 형성된다(중첩하는 단어 경계의 최소 수가 하나 이상의 다른 실시예에서 다를 수 있다는 것에 유의한다). 예를 들어, 제1 피쳐 포인트(제2 라인에 있는 제2 단어 박스, 길이가 6임)는 위쪽으로 2개의 단어 경계(길이가 5 및 7임)를 가지고 아래쪽으로 하나의 단어 경계(길이가 5임)를 갖는다. 제2 피쳐 포인트(제3 라인에 있는 제4 단어 박스, 길이가 5임)는 위쪽으로 2개의 단어 경계(길이가 4 및 5임)를 가지고 아래쪽으로 2개의 단어 경계(길이가 8 및 7임)를 갖는다. 따라서, 도 16에 도시한 바와 같이, 표시된 피쳐들은 중간 단어 경계의 길이와 그에 뒤이은 위쪽의 단어 경계의 길이, 또 아래쪽 단어 경계의 길이로 표시될 수 있다. 게다가, 단어 박스의 길이가 임의의 메트릭에 기초할 수 있다는 것에 유의한다. 따라서, 어떤 단어 박스에 대해 대안의 길이를 가질 수 있다. 이러한 경우에, 이들의 대안들 전부 또는 그 일부를 포함하는 피쳐들이 추출될 수 있다.
- [0228] 게다가, 하나 이상의 실시예들에서, 공백이 0으로 표시되고 단어 구역이 1로 표시되도록 피쳐들이 추출될 수 있다. 일례가 도 17에 도시되어 있다. 오른쪽에 있는 블록 표시는 왼쪽에 있는 문서 패치의 단어/공백 구역에 대응한다.
- [0229] 추출된 피쳐들은, 예를 들어, 노름(norm) 및 해밍 거리(Hamming distance)를 비롯한 다양한 거리 척도와 비교될 수 있다. 다른 대안으로서, 하나 이상의 실시예들에서, 질의 이미지와 동일한 피쳐를 갖는 문서 패치를 식별하기 위해 해쉬 테이블이 사용될 수 있다. 이러한 패치가 식별되면, 각각의 피쳐 포인트로부터 다른 피쳐 포인트로의 각도가 도 18에 나타난 바와 같이 계산될 수 있다. 다른 대안으로서, 피쳐 포인트들의 그룹들 간의 각도가 계산될 수 있다. 1802는 3쌍의 피쳐 포인트들로부터 계산된 각도(1803, 1804, 1805)를 나타낸 것이다. 계산된 각도들은 이어서 질의 이미지에서의 각각의 피쳐 포인트로부터 다른 피쳐 포인트들로의 각도들과 비교될 수 있다. 정합하는 점들에 대한 임의의 각도가 비슷한 경우, 유사성 점수(similarity score)가 증가될 수 있다. 다른 대안으로서, 각도의 그룹들이 사용되고 또 2개의 이미지에서의 유사한 피쳐 포인트 그룹들 간의 각도 그룹이 수치적으로 유사한 경우, 유사성 점수가 증가된다. 질의 이미지와 각각의 검색된 문서 패치 간의 점수가 계산될 때, 최고 점수를 얻는 문서 패치가 선택되고 그 정합이 어떤 미리 정해진 기준을 만족시키는지 결정하기 위해 적응적 임계값과 비교된다. 이 기준이 만족되는 경우, 정합하는 문서 경로가 발견된 것으로 표시된다.
- [0230] 게다가, 하나 이상의 실시예들에서, 추출된 피쳐들은 단어의 길이에 기초할 수 있다. 각각의 단어는 단어 높이 및 폭에 기초하여 추정된 문자들로 나누어진다. 주어진 단어 위아래에 있는 단어 라인이 스캔될 때, 위아래의 라인들에서의 공백 정보에 따라 추정된 문자들 각각에 이진값이 할당된다. 이어서 이진 코드가 정수로 표시된다. 예를 들어, 도 19를 참조하면, 동 도면은 단어 박스의 배열을 나타낸 것으로서, 각각의 단어 박스는 캡처된 이미지에서 검출된 단어를 나타낸다. 단어(1910)는 추정된 문자들로 나누어져 있다. 이 피쳐는 (i) 단어(1910)의 길이, (ii) 단어(1910) 위쪽의 라인의 텍스트 배열, 및 (iii) 단어(1910) 아래쪽의 라인의 텍스트 배열로 기술된다. 단어(1910)의 길이는 추정된 문자의 수로 측정된다. 텍스트 배열 정보는 현재의 추정된 문자의 위쪽 또는 아래쪽의 공백 정보의 이진 코딩으로부터 추출된다. 단어(1910)에서, 마지막 추정된 문자만이 공백 위에 있고, 두번째 및 세번째 추정된 문자는 공백 아래에 있다. 그에 따라, 단어(1910)의 피쳐는 (6, 100111, 111110)으로 코딩되고, 여기서 0은 공백을 의미하고 1은 공백이 없음을 의미한다. 정수 형태로 다시 쓰면, 단어(1910)는 (6, 39, 62)로 코딩된다.
- [0231] 도 20은 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 20에 도시한 "다중 분류자(multiple classifier)"는 서로 다른 피쳐 기술들을 독립적으로 분류하고 그 결과를 결합함으로써 서로 다른 피쳐 기술들의 상보적 정보를 이용한다. 텍스트 패치 정합에 적용되는 이 패러다임의 일례는 수평으로 및 수직으로 인접해 있는 단어쌍의 길이를 추출하고 데이터베이스 내의 패치들의 순위를 개별적으로 계산하는 것이다. 보다 상세하게는, 예를 들어, 하나 이상의 실시예들에서, 피쳐들의 위치가 분류 모델(720)에 부속하는 "분류자"에 의해 결정된다. 캡처된 이미지는 캡처된 이미지의 수평 및 수직 피쳐들을 결정하기 위해 분류자들의 조합을 사용하여 핑거프린팅(fingerprint)된다. 이것은 텍스트의 이미지가 그의 ID에 관한 2개의 독립적인 정보 소스를 포함한다는 관찰을 바탕으로 수행된다. 수평의 단어 시퀀스에 부가하여, 단어들의 수직 레이아웃도 또한 이미지가 추출된 문서를 식별하기 위해 사용될 수 있다. 예를 들어, 도 21에 도시된 바와 같이, 캡처된 이



미지(2110)는 수평 분류자(2112) 및 수직 분류자(2114)에 의해 분류된다. 분류자(2112, 2114) 각각은, 캡처된 이미지를 입력하는 것에 부가하여, 데이터베이스(3400)로부터 정보를 받고 차례로 각자의 분류가 적용될 수 있는 이들 문서 페이지의 순위를 출력한다. 환언하면, 도 21에 도시된 다중-분류자 기법은 수평 및 수직 피처를 사용하여 캡처된 이미지를 독립적으로 분류한다. 순위가 매겨진 문서 페이지들의 리스트가 이어서 결합 알고리즘(2118)에 따라 결합되고, 차례로 순위가 매겨진 문서 페이지들의 리스트를 출력하며, 이 리스트는 캡처된 이미지(2110)의 수평 및 수직 피처들 둘다에 기초한다. 상세하게는, 하나 이상의 실시예들에서, 수평 분류자(2112) 및 수직 분류자(2114)로부터의 개별적인 순위가 검출된 피처들이 데이터베이스(3400)에서 어떻게 공존하는지에 관한 정보를 사용하여 결합된다.

[0232] 이제부터, 도 22를 참조하여, 피처 추출을 위해 수직 레이아웃이 어떻게 수평 레이아웃과 통합되는지의 일례를 나타낸 것이다. (a)에서, 단어 경계선을 갖는 캡처된 이미지(2200)가 도시되어 있다. 캡처된 이미지(2200)로부터, 수평 및 수직 "n-그램(n-gram)"이 결정된다. "n-그램"은 n개의 숫자의 시퀀스이며, 각각의 숫자는 어떤 특성의 양을 나타낸다. 예를 들어, 수평 트라이그램은 3개의 단어로 된 수평 시퀀스의 각각의 단어에서의 문자들의 수를 규정한다. 예를 들어, 캡처된 이미지(2200)의 경우, (b)는 수평 트라이그램들, 즉 5-8-7(캡처된 이미지(2200)의 첫번째 라인에서의 수평으로 있는 일련의 단어 "upper", "division" 및 "courses" 각각에서의 문자의 수), 7-3-5(캡처된 이미지(2200)에서의 두번째 라인에서의 수평으로 있는 일련의 단어 "Project", "has" 및 "begun" 각각에서의 문자의 수), 3-5-3(캡처된 이미지(2200)의 두번째 라인에서의 수평으로 있는 일련의 단어 "has", "begun" 및 "The" 각각에서의 문자의 수), 3-3-6(캡처된 이미지(2200)의 세번째 라인에서의 수평으로 있는 일련의 단어 "461", "and" 및 "permit" 각각에서의 문자의 수) 및 3-6-8(캡처된 이미지(2200)의 세번째 라인에서의 수평으로 있는 일련의 단어 "and", "permit" 및 "projects" 각각에서의 문자의 수)을 나타낸 것이다.

[0233] 수직 트라이그램들은 주어진 단어의 위아래쪽에 있는 수직 단어 시퀀스의 각각의 단어에서의 문자의 수를 규정한다. 예를 들어, 캡처된 이미지(2200)의 경우, (c)는 수직 트라이그램들, 5-7-3(수직으로 있는 일련의 단어 "upper", "Project" 및 "461" 각각에서의 문자의 수), 8-7-3(수직으로 있는 일련의 단어 "division", "Project" 및 "461" 각각에서의 문자의 수), 8-3-3(수직으로 있는 일련의 단어 "division", "has" 및 "and" 각각에서의 문자의 수), 8-3-6(수직으로 있는 일련의 단어 "division", "has" 및 "permit" 각각에서의 문자의 수), 8-5-6(수직으로 있는 일련의 단어 "division", "begun" 및 "permit" 각각에서의 문자의 수), 8-5-8(수직으로 있는 일련의 단어 "division", "begun" 및 "projects" 각각에서의 문자의 수), 7-5-6(수직으로 있는 일련의 단어 "courses", "begun" 및 "permit" 각각에서의 문자의 수), 7-5-8(수직으로 있는 일련의 단어 "courses", "begun" 및 "projects" 각각에서의 문자의 수), 7-3-8(수직으로 있는 일련의 단어 "courses", "The" 및 "projects" 각각에서의 문자의 수), 7-3-7(수직으로 있는 일련의 단어 "Project", "461" 및 "student" 각각에서의 문자의 수), 및 3-3-7(수직으로 있는 일련의 단어 "has", "and" 및 "student" 각각에서의 문자의 수)을 나타낸 것이다.

[0234] 도 22에 도시한 캡처된 이미지(2200)로부터의 결정된 수평 및 수직 트라이그램에 기초하여, 수평 및 수직 트라이그램들 각각을 포함하는 문서들을 나타내는 문서들의 리스트 (d) 및 (e)가 생성된다. 예를 들어, (d)에서, 수평 트라이그램 7-3-5는 문서(15, 22, 134)에서 발생한다. 게다가, 예를 들어, (e)에서, 수직 트라이그램 7-5-6은 문서(15, 17)에서 발생한다. (d) 및 (e)의 문서 리스트를 사용하여, 모든 참조된 문서들의 순위가 매겨진 리스트가 각각 (f) 및 (g)에 도시되어 있다. 예를 들어, (f)에서, 문서(15)는 (d)에 있는 5개의 수평 트라이그램에 의해 참조되는 반면, 문서(9)는 (d)에 있는 하나의 수평 트라이그램에 의해서만 참조된다. 게다가, 예를 들어, (g)에서, 문서(15)는 (e)에서 11개 수직 트라이그램에 의해 참조되는 반면, 문서(18)는 (e)에서 하나의 수직 트라이그램에 의해서만 참조된다.

[0235] 이제부터 또한 도 23을 참조하면, 동 도면은 도 22를 참조하여 기술한 수평 및 수직 트라이그램 정보를 결합하는 기법을 나타낸 것이다. 이 기법은 원래의 인쇄된 페이지 상의 트라이그램들의 기지의 물리적 위치에 관한 정보를 사용하여 수평 및 수직 피처 추출로부터의 투표(vote)의 리스트를 결합한다. 수평 및 수직 분류자들 각각에 의해 출력되는 상위 M개의 선택요소들 중에 공통인 모든 문서에 대해, 그 문서에 투표한 모든 수평 트라이그램의 위치가 그 문서에 투표한 모든 수직 트라이그램의 위치와 비교된다. 문서는 임의의 수직 트라이그램과 중첩하는 수평 트라이그램의 수와 같은 투표수를 받으며, 여기서 "중첩"은 2개의 트라이그램의 경계 박스가 중첩할 때 일어난다. 그에 부가하여, 중첩의 중심의 x-y 위치가 도 34a의 3406을 참조하여 이하에 기술되는 증거 누적 알고리즘의 적절히 수정된 버전에 의해 카운트된다. 예를 들어, 도 23에 나타낸 바와 같이, 수평 및 수직 트라이그램 둘다에 의해 참조되는 페이지들의 리스트(c)를 결정하기 위해 (a) 및 (b)(각각 도 22의 (f) 및



(g))에 있는 리스트의 교집합이 구해진다. 교집합 리스트 (c), 리스트 (d) 및 (e)(식별된 트라이그램에 의해 참조된 교집합 문서들만을 나타냄), 그리고 인쇄된 문서 데이터베이스(3400)를 사용하여, 문서들의 중첩이 결정된다. 예를 들어, 문서(6)은 수평 트라이그램 3-5-3 및 수직 트라이그램 8-3-6에 의해 참조되고, 이들 2개의 트라이그램 자체는 캡처된 이미지(2200)에서의 단어 "has"에서 중첩하고, 따라서 문서(6)은 하나의 중첩에 대해 하나의 투표를 받는다. (f)에 나타난 바와 같이, 특정의 캡처된 이미지(2200)의 경우, 문서(15)는 가장 많은 투표를 받으며, 따라서 캡처된 이미지(2200)를 포함하는 문서인 것으로 식별된다. (x1,y1)은 문서(15) 내의 입력 이미지의 위치로서 식별된다. 이와 같이, 도 22 및 도 23을 참조하여 상기한 문서 지문 정합 기법의 요약에서, 수평 분류자는 텍스트의 단어들의 수평 배열로부터 도출된 피쳐들을 사용하고, 수직 분류자는 그 단어들의 수직 배열로부터 도출된 피쳐들을 사용하며, 이들 결과가 원래의 문서에서의 그 피쳐들의 중첩에 기초하여 결합된다. 이 피쳐 추출의 수평 측면이 적절한 문법 및 언어의 제약조건을 받는 반면, 수직 측면이 이러한 제약조건을 받지 않는다는 점에서, 이러한 피쳐 추출은 문서들을 일의적으로 식별하는 메커니즘을 제공한다.

[0236] 게다가, 도 22 및 도 23과 관련한 설명이 트라이그램의 사용과 관련되어 있지만, 수평 및 수직 피쳐 추출/분류 중 하나 또는 그 둘다에 대해 임의의 n-그램이 사용될 수 있다. 예를 들어, 하나 이상의 실시예에서, 수직 및 수평 n-그램(단, n=4)은 다중-분류자 피쳐 추출에 사용될 수 있다. 하나 이상의 다른 실시예들에서, 수평 분류자는 n-그램(단, n=3)에 기초하여 피쳐를 추출할 수 있는 반면, 수직 분류자는 n-그램(단, n=5)에 기초하여 피쳐를 추출할 수 있다.

[0237] 게다가, 하나 이상의 실시예들에서, 분류가 엄격히 말해 수직 또는 수평이 아닌 인접 관계(adjacency relationship)에 기초할 수 있다. 예를 들어, NW, SW, NW 및 SE 인접 관계가 추출/분류에 사용될 수 있다.

[0238] 도 24는 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 24에 도시된 "데이터베이스-위주 피드백(database-driven feedback)" 기법은 입력과 정합할 수 있는 문서들의 이미지를 이용하여 원시 이미지로부터의 서브-이미지가 입력 이미지와 정합되는 이미지 분석의 후속 단계를 결정함으로써 문서 이미지 정합 시스템의 정확도가 향상될 수 있다는 것을 고려한다. 이 기법은 입력 이미지에 존재하는 노이즈를 복제하는 변환을 포함한다. 이것은 템플릿 정합 분석이 뒤따를 수 있다.

[0239] 도 25는 본 발명의 일 실시예에 따른 데이터베이스-위주 피드백에 대한 흐름 프로세스를 나타낸 것이다. 입력 이미지 패치는 먼저 상기한 바와 같이 (예를 들어, 단어 OCR 및 단어쌍 탐색, 문자 OCR 및 문자쌍 탐색, 단어 경계 박스 구성을 사용하여) 단계(2510, 2512)에서 전처리되고 인식되어 이미지 패치(2522)의 식별을 위한 다수의 후보를 생성한다. 이 리스트 내의 각각의 후보는 이하의 항목 (doci, pagei, xi, yi)를 포함할 수 있으며, 여기서 doci는 문서의 식별자이고, pagei는 문서 내의 페이지이며, (xi,yi)는 그 페이지 내의 이미지 패치의 중심의 x-y 좌표이다.

[0240] 원시 패치 검색 알고리즘은 단계(2514)에서 기지의 공간 해상도, 예를 들어, 100dpi로 변환되도록 하기 위해 선택에 따라서는 페이지로부터의 거리에 관한 정보를 사용하여 입력 이미지 패치 전체의 크기를 고정된 크기로 정규화한다. 상기한 폰트 크기 추정 알고리즘이 이 작업에 적합할 수 있다. 이와 유사하게, 공지의 초점으로부터의 거리 또는 초점으로부터의 깊이 기법이 사용될 수 있다. 또한, 크기 정규화는 그의 단어 경계 박스의 높이에 기초하여 이미지 패치를 비례 스케일링할 수 있다.

[0241] 원시 패치 검색 알고리즘은 MMR 데이터베이스가 발생하는 패치에 대한 경계 박스의 중심과 함께 수신하는 각각의 문서 및 페이지에 대한 식별자를 갖는 MMR 데이터베이스(3400)에 질의를 한다. 발생한 패치의 범위는 정규화된 입력 패치의 크기에 기초한다. 이러한 방식으로, 동일한 공간 해상도 및 차원의 패치가 획득될 수 있다. 예를 들어, 100dpi로 정규화될 때, 입력 패치는 그의 중심의 양측에서 50 픽셀 확장될 수 있다. 이 경우에, MMR 데이터베이스는 지정된 x-y 값에 중심이 있는 100 픽셀 높이 및 폭인 100dpi 원시 패치를 발생하도록 지시 받는다.

[0242] MMR 데이터베이스(2524)로부터 반환된 각각의 원시 이미지 패치가 이하의 항목들(doci, pagei, xi, yi, widthi, heighti, actioni)과 연관될 수 있으며, 여기서 (doci, pagei, xi, yi)는 상기한 바와 같고, widthi 및 heighti는 원시 패치의 폭 및 높이(단위: 픽셀)이고, actioni는 데이터베이스 내의 doci 엔트리 내의 대응하는 영역과 연관될 수 있는 선택적인 동작이다. 원시 패치 검색 알고리즘은 그가 구성한 크기 정규화된 입력 패치와 함께 이미지 패치 및 데이터의 리스트를 출력(2518)한다.

[0243] 게다가, 하나 이상의 실시예들에서, 패치 정합 알고리즘(2516)은 크기 정규화된 입력 패치를 각각의 원시 패치와 비교하고, 이들이 서로 얼마나 잘 정합하는지를 나타내는 점수를 할당(2520)한다. 당업자라면 패치들의 크

기가 비슷하도록 하기 위해 사용되는 메커니즘으로 인해 해밍 거리에 대한 단순 교차 상관이 많은 경우에 만족된다는 것을 잘 알 것이다. 게다가, 이 프로세스는 입력에서 검출된 이미지 노이즈와 흡사한 노이즈를 원시 패치에 유입시키는 것을 포함할 수 있다. 이 비교는 또한 아주 복잡할 수 있으며 2개의 패치의 OCR 결과 및 문자, 문자쌍 또는 단어쌍의 숫자에 기초한 순위를 비롯한 임의의 피쳐 세트의 비교를 포함할 수 있으며, 이들 쌍은 이전과 같이 기하학적 관계에 의해 제약될 수 있다. 그렇지만, 이 경우에, 입력 패치와 원시 패치 간에 공통인 기하학적 쌍의 수가 추정되고 순위 메트릭으로서 사용될 수 있다.

[0244] 게다가, 출력(2520)은 n-튜플(doc<sub>i</sub>, page<sub>i</sub>, x<sub>i</sub>, y<sub>i</sub>, action<sub>i</sub>, score<sub>i</sub>)의 형태일 수 있으며, 여기서 점수는 패치 정합 알고리즘에 의해 제공되고 입력 패치가 doc<sub>i</sub>, page<sub>i</sub>의 대응하는 구역과 얼마나 잘 정합하는지를 나타낸다.

[0245] 도 26은 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 26에 나타낸 바와 같이 "데이터베이스-위주 분류자" 기법은 입력 이미지를 포함할 수 있는 일련의 가설을 발생하기 위해 초기 분류(initial classification)를 사용한다. 이들 가설은 데이터베이스(3400)에서 탐색되며, 피쳐 추출과 분류 전략이 이들 가설에 대해 자동적으로 설계된다. 일례는 입력 패치를 Times 또는 Arial 폰트를 포함하는 것으로 식별하는 것이다. 이 경우에, 제어 구조(714)는 serif/san serif 구별을 위한 특수한 피쳐 추출기 및 분류자를 호출한다(invokes).

[0246] 도 27은 본 발명의 일 실시예에 따른 데이터-위주 분류를 위한 흐름 프로세스를 나타낸 것이다. 제1 피쳐 추출(2710) 이후에, 문서, 페이지, 및 그 페이지들 내에서의 x-y 위치의 순위를 생성하기 위해 상기한 인식 방법들 중 임의의 하나 이상에 의해 입력 이미지 패치가 분류(2712)된다. 이 리스트에서의 각각의 후보는, 예를 들어, 이하의 항목(doc<sub>i</sub>, page<sub>i</sub>, x<sub>i</sub>, y<sub>i</sub>)을 포함할 수 있으며, 여기서 doc<sub>i</sub>는 문서에 대한 식별자이고, page<sub>i</sub>는 문서 내의 페이지이며, (x<sub>i</sub>, y<sub>i</sub>)는 그 페이지 내에서의 이미지 패치의 중심의 x-y 좌표이다. 도 25를 참조하여 기술되는 원시 패치 검색 알고리즘(2714)은 각각의 후보에 대한 패치 이미지를 발생하는 데 사용될 수 있다.

[0247] 여전히 도 27을 참조하면, 제2 피쳐 추출이 원시 패치(2716)에 적용된다. 이것은 제1 피쳐 추출과 다를 수 있으며, 예를 들어, 폰트 검색 알고리즘, 문자 인식 기법, 경계 박스, 및 SIFT 피쳐 중 하나 이상을 포함할 수 있다. 각각의 원시 패치에서 검출되는 피쳐들은, 예를 들어, 미지의 샘플들을 원시 패치들 중 하나로 분류하도록 설계되어 있는 신경망, 지원 벡터 머신, 및/또는 최근접 이웃 분류자를 포함하는 자동 분류자 설계 방법(2720)에 입력된다. 동일한 제2 피쳐 추출이 입력 이미지 패치에 적용(2718)될 수 있으며, 그가 검출하는 피쳐들이 원시 패치에 대해 전문화되어 있을 수 있는 이 새로 설계된 분류자에 입력된다.

[0248] 출력(2724)은 n-튜플(doc<sub>i</sub>, page<sub>i</sub>, x<sub>i</sub>, y<sub>i</sub>, action<sub>i</sub>, score<sub>i</sub>)의 형태일 수 있으며, 여기서 점수는 2720에 의해 자동적으로 설계된 분류 기법(2722)에 의해 제공된다. 당업자라면 점수가 입력 패치가 doc<sub>i</sub>, page<sub>i</sub>의 대응하는 구역과 얼마나 잘 정합하는지를 나타낸다는 것을 잘 알 것이다.

[0249] 도 28은 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 28에 도시한 "데이터베이스-위주 다중 분류자(database-driven multiple classifier)" 기법은 결정 프로세스 전체에 걸쳐 다수의 후보를 전달함으로써 인식 프로세스에서 조기에 복원가능하지 않은 예러의 가능성을 감소시킨다. 몇가지 초기 분류가 수행된다. 각각은 서로 다른 피쳐 추출 및 분류에 의해 구별될 수 있는 입력 패치의 서로 다른 순위를 발생한다. 예를 들어, 그 세트들 중 하나가 수평 n-그램에 의해 발생될 수 있고 serif를 san-serif와 구별함으로써 일의적으로 인식될 수 있다. 다른 예는 수직 n-그램에 의해 발생될 수 있고 라인 간격의 정확한 계산으로 일의적으로 인식될 수 있다.

[0250] 도 29는 본 발명의 일 실시예에 따른 데이터베이스-위주 다중 분류를 위한 흐름 프로세스를 나타낸 것이다. 이 흐름 프로세스는 도 27에 나타낸 것과 유사하지만, 다수의 서로 다른 피쳐 추출 알고리즘(2910, 2912)을 사용하여 분류자(2914, 2916)로 입력 이미지 패치의 독립적인 순위를 생성한다. 피쳐들 및 분류 기법들의 예는 상기한 수평 및 수직 단어-길이 n-그램을 포함한다. 각각의 분류자는 각각의 후보에 대해 적어도 이하의 항목들(doc<sub>i</sub>, page<sub>i</sub>, x<sub>i</sub>, y<sub>i</sub>, score<sub>i</sub>)을 포함하는 순위가 매겨진 패치 ID 리스트를 생성할 수 있으며, doc<sub>i</sub>는 문서에 대한 식별자이고, page<sub>i</sub>는 문서 내의 페이지이며, (x<sub>i</sub>, y<sub>i</sub>)는 그 페이지 내에서의 이미지 패치의 중심의 x-y 좌표이고, score<sub>i</sub>는 입력 패치가 데이터베이스 문서 내의 대응하는 위치와 얼마나 잘 정합하는지를 나타낸다.

[0251] 도 25를 참조하여 상기한 원시 패치 검색 알고리즘은 2914 및 2916의 출력에서 패치 ID 리스트에서의 엔트리에 대응하는 일련의 원시 이미지 패치를 생성하는 데 사용될 수 있다. 제3 및 제4 피쳐 추출(2918, 2920)은 도 27에서 상기한 바와 같이 자동적으로 설계되고 적용되는 원시 패치 및 분류자에 이전과 같이 적용될 수 있다.

- [0252] 여전히 도 29를 참조하면, 이들 분류자에 의해 생성된 순위들이 결합되어 엔트리(doc<sub>i</sub>, page<sub>i</sub>, xi, yi,)(단, i=1...후보의 수)를 갖는 단일의 순위(2924)를 생성하며, 여기서 각각의 엔트리에서의 값들은 상기한 바와 같다. 예를 들어, 2개의 순위에서의 그의 공통의 위치에 기초하여 항목에 점수를 할당하는 공지의 Borda 카운트 척도에 의해 순위 결합(ranking combination)(2922)이 수행될 수 있다. 이것은 개개의 분류자에 의해 할당된 점수와 결합되어 합성 점수를 생성할 수 있다. 게다가, 당업자라면 다른 순위 결합 방법이 사용될 수 있다는 것을 잘 알 것이다.
- [0253] 도 30은 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 30에 도시된 "비디오 시퀀스 이미지 누적(video sequence image accumulation)" 기법은 근방의 또는 인접한 프레임들로부터의 데이터를 통합함으로써 이미지를 구성한다. 한 예는 "수퍼-해상도(super-resolution)"를 포함한다. 이는 N개의 시간적으로 인접한 프레임들을 등록하고 렌즈의 점 퍼짐 함수(point spread function)에 대해 아는 것을 사용하여 본질적으로 서브-픽셀 엣지 향상(sub-pixel edge enhancement)이 되는 것을 수행한다. 그 효과는 이미지의 공간 해상도를 향상시키는 것이다. 게다가, 하나 이상의 실시예들에서, 수퍼-해상도 방법은 구멍, 모서리 또는 도트 등의 텍스트-관련 피쳐들을 강조하도록 전문화될 수 있다. 추가적인 확장은 수퍼-해상도 통합 기능을 전문화하기 위해 데이터베이스(3400)로부터 결정되는 후보 이미지 패치의 특성들을 사용한다.
- [0254] 도 31은 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 31에 도시된 "비디오 시퀀스 피쳐 누적(video sequence feature accumulation)" 기법은 결정을 하기 이전에 다수의 시간적으로 인접한 프레임들에 걸쳐 피쳐들을 누적한다. 이것은 캡처 장치의 높은 샘플링 레이트(예를 들어, 초당 30 프레임) 및, 캡처 장치를 적어도 몇 초 동안 문서 상의 동일한 점으로 향한 상태로 있게 하는 사용자의 의도를 이용한다. 각각의 프레임 상에서 독립적으로 피쳐 추출이 수행되고, 그 결과들이 결합되어 하나의 통합된 피쳐 맵을 생성한다. 이 결합 프로세스는 묵시적인 등록 단계를 포함한다. 텍스트 패치의 비디오 클립의 조사 시에 이 기법이 필요하다는 것이 곧 명백해진다. 통상적인 캡처 장치에서의 자동-초점 및 콘트라스트 조정이 인접한 비디오 프레임에서 상당히 서로 다른 결과를 야기할 수 있다.
- [0255] 도 32는 본 발명의 일 실시예에 따른 다른 문서 지문 정합 기법을 나타낸 것이다. 도 32에 도시된 "비디오 시퀀스 결정 결합(video sequence decision combination)" 기법은 다수의 시간적으로 인접한 프레임들로부터의 결정들을 결합한다. 이것은 통상적인 캡처 장치의 높은 샘플링 레이트 및, 적어도 몇 초 동안 캡처 장치를 문서 상의 동일한 점으로 향하게 유지하는 사용자의 의도를 이용한다. 각각의 프레임이 독립적으로 처리되고 그 자신의 순위가 매겨진 결정 리스트를 발생한다. 이들 결정이 결합되어 입력 이미지 세트의 하나의 통합된 순위를 발생한다. 이 기법은 결정 결합 프로세스를 제어하는 묵시적인 등록 방법을 포함한다.
- [0256] 하나 이상의 실시예에서, 도 6 내지 도 32를 참조하여 상기한 다양한 문서 지문 정합 기술들 중 하나 이상과 관련하여 사용될 수 있으며, 이러한 결합을 본 명세서에서 "다중-계층(multi-tier)(또는 다중-인자(multi-factor)) 인식"이라고 한다. 일반적으로, 다중-계층 인식(multi-tier recognition)에서, 제1 정합 기술은 문서 데이터베이스에서 특정의 기준을 갖는 일련의 페이지를 찾아내는 데 사용되며, 이어서 제2 정합 기술은 그 일련의 페이지 내의 페이지들 중으로부터 패치를 일의적으로 식별하는 데 사용된다.
- [0257] 도 33은 본 발명의 일 실시예에 따른 다중-계층 인식의 흐름 프로세스의 일례를 나타낸 것이다. 처음에, 단계(3310)에서, 캡처 장치(106)는 관심의 문서 상의 "컬링(culling)" 피쳐를 캡처/스캔하는 데 사용된다. 컬링 피쳐는 임의의 피쳐일 수 있으며, 그의 캡처의 결과 사실상 문서 데이터베이스 내의 일련의 문서의 선택이 있게 된다. 예를 들어, 컬링 피쳐는 숫자 전용 바코드(numeric-only bar code)(예를 들어, UPC(universal product code)), 영숫자 바코드(alphanumeric bar code)(예를 들어, 코드 39, 코드 93, 코드 128), 또는 2-차원 바코드(예를 들어, QR 코드, PDF417, DataMatrix, Maxicode)일 수 있다. 게다가, 컬링 피쳐는, 예를 들어, 그래픽, 이미지, 상표, 로고, 특정의 컬러 또는 컬러들의 조합, 키워드, 또는 구문일 수 있다. 게다가, 하나 이상의 실시예에서, 컬링 피쳐는 캡처 장치(106)에 의해 인식하기에 적당한 피쳐들로 제한될 수 있다.
- [0258] 단계(3310)에서 컬링 피쳐가 캡처되면, 단계(3312)에서 문서 데이터베이스 내의 일련의 문서 및/또는 문서들의 페이지가 캡처된 컬링 피쳐와의 연관성에 기초하여 선택된다. 예를 들어, 캡처된 컬링 피쳐가 회사의 로고인 경우, 그 로고를 포함하는 것으로 인덱싱된 데이터베이스 내의 모든 문서가 선택된다. 다른 예에서, 데이터베이스는 캡처된 컬링 이미지와 비교되는 상표의 라이브러리를 포함할 수 있다. 라이브러리에 "히트(hit)"가 있는 경우, 그 히트 상표(hit trademark)와 연관된 모든 문서들이 이하에 기술되는 바와 같이 차후의 정합을 위해 선택된다. 게다가, 하나 이상의 실시예에서, 단계(3312)에서의 문서/페이지의 선택은 캡처된 컬링 피쳐 및 스캔된 문서 상에서의 그 컬링 피쳐의 위치에 의존할 수 있다. 예를 들어, 캡처된 컬링 피쳐와 연관된 정보는 그



컬링 이미지가 문서의 좌측 하부 모서리와 달리 문서의 우측 상부 모서리에 위치하는지 여부를 규정할 수 있다.

- [0259] 게다가, 당업자라면 특정의 캡처된 이미지가 컬링 피처의 이미지를 포함한다는 결정이 캡처 장치(106) 또는 캡처 장치(106)로부터 미가공 이미지 데이터를 수신하는 어떤 다른 컴포넌트에 의해 행해질 수 있다는 것을 잘 알 것이다. 예를 들어, 데이터베이스 자체는 캡처 장치(106)로부터 전송된 특정의 캡처된 이미지가 컬링 피처를 포함하는 것으로 결정할 수 있으며, 그에 응답하여 데이터베이스는 캡처된 컬링 피처와 연관된 일련의 문서를 선택한다.
- [0260] 단계(3312)에서 특정의 일련의 문서가 선택된 후에, 단계(3314)에서 캡처 장치(106)는 계속하여 관심의 문서의 이미지를 스캔하고 그에 따라 캡처한다. 문서의 캡처된 이미지는 이어서 도 6 내지 도 32를 참조하여 기술된 다양한 문서 지문 정합 기법들 중 하나 이상을 사용하여, 단계(3312)에서 선택된 문서에 대해 정합된다. 예를 들어, 신발 그래픽의 컬링 피처를 포함하는 것으로 인덱싱된 일련의 문서들이 단계(3310)에서의 관심의 문서의 신발 그래픽 이미지의 캡처에 기초하여 단계(3312)에서 선택된 후에, 관심의 문서의 차후의 캡처된 이미지가 앞서 기술한 바와 같이 다중 분류자 기법을 사용하여 일련의 선택된 문서들에 대해 정합될 수 있다.
- [0261] 따라서, 도 33을 참조하여 상기한 다중-계층 인식 흐름 프로세스의 구현을 사용하여, 차후의 캡처된 이미지가 정합되는 페이지/문서의 양을 처음에 감소시킴으로써 패치 인식 횟수가 감소될 수 있다. 게다가, 사용자는 먼저 이미지, 바코드, 그래픽 또는 다른 유형의 컬링 피처가 있는 위치 상에서 문서를 스캔함으로써 이러한 향상된 인식 횟수를 이용할 수 있다. 이러한 조치를 취함으로써, 사용자는 차후의 캡처된 이미지가 정합되는 문서의 양을 신속하게 감소시킬 수 있다.
- [0262] MMR 데이터베이스 시스템
- [0263] 도 34a는 본 발명의 일 실시예에 따라 구성된 MMR 데이터베이스 시스템(3400)의 기능 블록도를 나타낸 것이다. 시스템(3400)은 콘텐츠-기반 검색을 하도록 구성되어 있으며, 여기서 객체들 간의 2차원 기하학적 관계가 텍스트-기반 인덱스(또는 임의의 다른 검색가능한 인덱스)로 탐색할 수 있는 방식으로 표시되어 있다. 시스템(3400)은 증거 누적을 이용하여, 예를 들어 피처의 발생 빈도를 2차원 구역에서의 그것의 위치의 확률과 결합시킴으로써 탐색 효율을 향상시킨다. 한 특정의 실시예에서, 데이터베이스 시스템(3400)은 (PD 인덱스(322)를 포함하는) 문서 이벤트 데이터베이스(320)의 상세한 구현이며, 이 데이터베이스의 내용은 도 3을 참조하여 상기한 캡처 모듈(318) 및/또는 문서 지문 정합 모듈(226)에 의해 발생된 인쇄된 문서의 전자적 표시를 포함한다. 시스템(3400)의 다른 응용 및 구성은 이 개시 내용으로부터 명백하게 될 것이다.
- [0264] 알 수 있는 바와 같이, 데이터베이스 시스템(3400)은 MMR 피처 추출 모듈(3402)에 의해 계산되는 기술을 수신하는 MMR 인덱스 테이블 모듈(3404), 증거 누적 모듈(3406) 및 관계형 데이터베이스(3408)(또는 임의의 다른 적당한 저장 설비)을 포함한다. 인덱스 테이블 모듈(3404)은 문서, 페이지, 및 각각의 피처가 있는 그 페이지 내에서의 x-y 위치를 식별하는 인덱스 테이블을 조사한다(interrogate). 인덱스 테이블은, 예를 들어, MMR 인덱스 테이블 모듈(3404) 또는 어떤 다른 전용 모듈에 의해 발생될 수 있다. 인덱스 테이블 모듈(3404)로부터 데이터가 주어지면, 증거 누적 모듈(3406)은 순위가 매겨진 일련의 문서, 페이지 및 위치 가설(3410)을 계산하도록 프로그래밍되거나 다른 방식으로 구성된다. 관계형 데이터베이스(3408)는 각각의 패치에 대한 부가적인 패치(3412)를 저장하는 데 사용될 수 있다. 이들은 도 5의 504 및 508을 포함하지만, 이에 한정되지 않는다. 패치에 대한 서명 또는 지문(즉, 고유의 검색어)을 도출하는 데 패치 내에서 텍스트의 2차원 배열을 사용함으로써, 심지어 텍스트의 작은 프래그먼트의 고유성이 상당히 향상된다. 다른 실시예들도 이와 유사하게 패치에 대한 서명 또는 지문을 도출하는 데 패치 내의 객체/특징의 임의의 2차원 배열을 이용할 수 있으며, 본 발명의 실시예들이 패치를 일의적으로 식별하기 위한 텍스트의 2차원 배열로 한정되는 것으로 보아서는 안된다. 도 34a에 도시된 데이터베이스 시스템(3400)의 다른 컴포넌트 및 기능은 피드백-관련 피처 검색 모듈(feedback-directed features search module)(3418), 문서 렌더링 적용 모듈(document rendering application module)(3414) 및 서버-이미지 추출 모듈(3416)을 포함한다. 이들 컴포넌트는 피드백-관련 피처 검색은 물론 동적 원시 이미지 발생을 제공하기 위해 다른 시스템(3400) 컴포넌트들과 상호작용한다. 그에 부가하여, 시스템(3400)은 동작들을 수신하는 동작 처리기(3413)를 포함한다. 이들 동작은 데이터베이스 시스템(3400)에 의해 수행되는 동작 및 이 시스템(3400)이 제공하는 출력을 결정한다. 이들 다른 컴포넌트 각각에 대해 차례로 설명한다.
- [0265] 패치 내에서 이러한 2차원 텍스트 배열을 이용하는 MMR 피처 추출 모듈(3402)의 일례가 도 34b에 도시되어 있다. 한가지 이러한 실시예에서, MMR 피처 추출 모듈(3402)은 OCR-기반 기법을 이용하여 이미지 패치로부터 피처들(텍스트 또는 다른 타겟 피처들)을 추출하도록 프로그래밍되거나 다른 방식으로 구성되어 있다. 이 특정의 실시예에서, 피처 추출 모듈(3402)은 패치의 이미지 내의 단어들의 x-y 위치를 추출하고, 그 위치들을 그 위치

가 포함하는 일련의 수평 및 수직으로 인접한 단어쌍으로 표시한다. 이미지 패치는 이들이 수평으로 인접해 있는 경우 "-"(예를 들어, the-cat, in-the, the-hat, 및 is-back)로, 또한 이들이 수직으로 중첩하고 있는 경우 "+"(예를 들어, the+in, cat+the, in+is, 및 the+back)로 연결되어 있는 단어쌍으로 효과적으로 변환된다. x-y 위치는, 예를 들어, 문서 이미지 내의 어떤 고정된 점으로부터(문서의 좌측 최상부 모서리 또는 중심으로부터) x 및 y 평면 방향으로의 픽셀 카운트에 기초하고 있다. 유의할 점은 이 예에서 수평으로 인접한 쌍이 많은 다른 텍스트 구절에서 빈번히 있을 수 있지만, 수직으로 중첩하는 쌍이 다른 텍스트 구절에서는 드물게 있을 수 있다는 것이다. 단어들 사이에 "/"를 갖는 SW-NE 인접성, "\"를 갖는 NW-SE 인접성 기타 등등의, 이미지 피쳐들 간의 다른 기하학적 관계가 이와 유사하게 인코딩될 수 있다. 또한, "피쳐들"은 임의적이지만 일관성있는 문자열로 인코딩될 수 있는 단어 경계 박스(또는 다른 피쳐 경계 박스)로 일반화될 수 있다. 예를 들어, 울퉁불퉁한 상부 윤곽 및 매끈한 하부 윤곽을 갖는 높이보다 4배 긴 경계 박스가 문자열 "4rusl"로 표시될 수 있다. 그에 부가하여, 기하학적 관계는 피쳐들 간의 임의의 각도 및 거리로 일반화될 수 있다. 예를 들어, NW-SE 인접해 있지만 2 단어-높이만큼 떨어져 있는 "4rusl" 기술을 갖는 2개의 단어가 "4rusl\ \4rusl"로 표시될 수 있다. 이 개시 내용을 살펴보면 수많은 인코딩 방식들이 명백하게 될 것이다. 게다가, 유의할 점은 숫자, 부울값, 기하학적 형상, 및 다른 이러한 문서 피쳐들이 패치를 식별하는 데 단어쌍 대신에 사용될 수 있다는 것이다.

[0266] 도 34c는 본 발명의 일 실시예에 따른 예시적인 인덱스 테이블 구성을 나타낸 것이다. 알 수 있는 바와 같이, MMR 인덱스 테이블은 반전된 용어 인덱스 테이블(inverted term index table)(3422) 및 문서 인덱스 테이블(document index table)(3424)을 포함한다. 각각의 고유의 용어 또는 피쳐(예를 들어, 키(3421))는 레코드(3423)의 리스트(예를 들어, Rec#1, Rec#2, 기타 등등)를 가리키는 피쳐의 함수값(예를 들어, 키 x)을 보유하는 용어 인덱스 테이블(3422) 내의 위치를 가리키고, 각각의 레코드는 문서 내의 페이지 상의 후보 구역을 식별해 주며, 이에 대해 차례로 기술한다. 일례에서, 키 및 키(키 x)의 함수값이 동일하다. 다른 예에서, 해쉬 함수가 키에 적용되고, 함수의 출력이 키 x이다.

[0267] 질의어(query term)의 리스트가 주어지면, 키에 의해 인덱싱된 모든 레코드가 검사되고, 모든 질의어와 가장 부합하는 구역이 식별된다. 구역이 (예를 들어, 사전 정의된 정합 임계값에 기초하여) 충분히 높은 정합 점수를 포함하는 경우, 가설이 확인된다. 그렇지 않은 경우, 정합이 실패한 것으로 선언되고 어떤 구역도 반환되지 않는다. 이 예시적인 실시예에서, 키는 앞서 기술한 바와 같이 "-" 또는 "+"에 의해 분리되어 있는 단어쌍(예를 들어, "the-cat" 또는 "cat+the")이다. 키 자체에 기하학적 관계를 포함시키는 이 기법은 2차원 기하학적 질의에 종래의 텍스트 검색 기술을 사용하는 것을 가능하게 해준다.

[0268] 따라서, 인덱스 테이블 구성은 이미지 패치에서 검출된 피쳐들을 피쳐들 자체와 이들 간의 기하학적 관계를 나타내는 텍스트 용어(textual term)로 변환한다. 이것은 종래의 텍스트 인덱싱 및 검색 방법의 이용을 가능하게 해준다. 예를 들어, 수직으로 인접한 용어 "cat" 및 "the"가, 본 발명의 개시 내용으로부터 명백하게 되는 바와 같이, "질의어"라고 할 수 있는 심볼 "cat+the"로 표시된다. 종래의 텍스트 검색 데이터 구조 및 방법의 이용은 인터넷 텍스트 검색 시스템(예를 들어, 구글, 야후, 마이크로소프트, 기타 등등)의 상단에서 본 명세서에 기술된 MMR 기법의 이식(grafting)을 용이하게 해준다.

[0269] 이 예시적인 실시예의 반전된 용어 인덱스 테이블(3422)에서, 각각의 레코드는 6개의 파라미터, 문서 ID(DocID), 페이지 번호(PG), x/y 오프셋(각각 X 및 Y), 그리고 직사각형 영역의 폭 및 높이(각각 W 및 H)를 사용하여 문서 내의 페이지 상의 후보 구역을 식별해준다. DocID는 문서가 인쇄될 때 타임스탬프(또는 다른 메타 데이터)에 기초하여 발생하는 고유의 문자열이다. 그렇지만, 이는 임의의 문자열 결합 장치 ID 및 사람 ID일 수 있다. 여하튼, 문서는 고유의 DocID에 의해 식별되고 문서 인덱스 테이블에 저장되어 있는 레코드를 갖는다. 페이지 번호는 용지 출력에 대응하는 페이지 번호 매김(pagination)이며, 1에서 시작한다. 직사각형 영역은 좌측 상부 모서리의 X-Y 좌표는 물론 정규화된 좌표 시스템에서의 경계 박스의 폭 및 높이에 의해 파라미터화된다. 수많은 내부 문서 위치/좌표 방식이 이 개시 내용으로부터 명백하게 되며, 본 발명이 임의의 특성의 것으로 제한되는 것으로 보아서는 안된다.

[0270] 본 발명의 일 실시예에 따라 구성된 예시적인 레코드 구조는 24-비트 DocID 및 8-비트 페이지 번호를 사용하여, 1600만개 문서 및 40억 페이지까지 가능하게 해준다. 경계 박스의 각각의 X 및 Y 오프셋에 대한 부호없는 1 바이트(one unsigned byte)는 (8.5" x 11" 페이지를 가정하는 경우) 30dpi 수평 및 23dpi 수직의 공간 해상도를 제공한다(그렇지만, 다른 페이지 크기 및/또는 공간 해상도가 사용될 수 있다). 경계 박스의 폭 및 높이에 대한 유사한 취급(예를 들어, W 및 H 각각에 대해 부호없는 1 바이트)이 마침표 또는 "i" 상의 점만큼 작은 또는 전체 페이지(예를 들어, 8.5" x 11" 또는 기타)만큼 큰 구역의 표시를 가능하게 해준다. 따라서, 레코드당 8

바이트(DocID에 3 바이트, PG에 1 바이트, X에 1 바이트, Y에 1 바이트, W에 1 바이트, 및 H에 1 바이트는 총 8 바이트임)는 많은 수의 구역을 수용할 수 있다.

[0271] 문서 인덱스 테이블(3424)은 각각의 문서에 관한 관련 정보를 포함한다. 한 특징의 실시예에서, 이 정보는 인쇄 해상도, 인쇄 일자, 용지 크기, 쉼도우 파일 이름, 페이지 이미지 위치, 기타 등등을 비롯한 문서-관련 필드를 XML 파일에 포함한다. 문서를 인덱싱할 때 인쇄 좌표가 정규화된 좌표계로 변환되기 때문에, 검색 가설을 계산하는 것은 이 테이블을 필요로 하지 않는다. 따라서, 문서 인덱스 테이블(3424)은 정합된 후보 구역에 대해서만 조회된다. 그렇지만, 이 결정은 인덱스에서의 얼마간의 정보 손실을 암시하는데, 그 이유는 정규화된 좌표가 보통 인쇄 해상도보다 낮은 해상도에 있기 때문이다. 대안의 실시예가 검색 가설을 계산할 때, 원하는 경우, 문서 인덱스 테이블(3424)(또는 정규화된 좌표보다 더 높은 해상도)을 사용할 수 있다.

[0272] 따라서, 인덱스 테이블 모듈(3404)은 주어진 이미지 질의가 행해지는 객체(예를 들어, 문서 페이지) 및 그 객체 내에서의 x-y 위치의 콘텐츠-기반 검색을 가능하게 해주는 이미지 인덱스를 효과적으로 제공하는 동작을 한다. 이러한 이미지 인덱스와 관계형 데이터베이스(3408)의 결합은 이미지 패치 및 패치의 특성(예를 들어, 패치에 부착된 "동작" 또는 패치와 관련된 다른 콘텐츠를 검색하게 하기 위해 스캔될 수 있는 바코드 등)을 정합시키는 객체의 위치를 고려한다. 관계형 데이터베이스(3408)는 또한 문서 내의 다른 패치들에 패치로부터 인덱스 테이블 내의 피쳐들의 "역방향 링크(reverse link)"를 위한 수단을 제공한다. 역방향 링크는 인식 알고리즘이 문서 이미지의 한 부분에서 다른 부분으로 이동할 때 보게 될 것으로 예상하는 피쳐들을 찾아내는 방법을 제공하며, 이는 본 명세서에서 논의되는 MMR 시스템에서의 프론트-엔드 이미지 분석 알고리즘의 수행을 상당히 향상시킬 수 있다.

[0273] 피드백-관련 피쳐 검색

[0274] 이미지 패치의 x-y 좌표(예를 들어, 이미지 패치의 중심에 대한 x-y 좌표)는 물론 문서 및 페이지의 ID는 또한 피드백-관련 피쳐 검색 모듈(feedback-directed feature search module)(3418)에 입력될 수 있다. 피드백-관련 피쳐 검색 모듈(3418)은 이미지 패치의 중심으로부터 주어진 거리 내에 있는 레코드(3423)를 용어 인덱스 테이블(3422)에서 검색한다. 이 검색은, 예를 들어, X 또는 Y 값의 순서로 정렬된 인접한 메모리 블록들에서 각각의 DocID-PG 조합에 대한 레코드(3423)를 저장함으로써 용이하게 될 수 있다. 탐색은 주어진 값(X 또는 Y는 데이터가 저장될 때 어떻게 정렬되는가에 의존함)에 대한 이진 검색(binary search) 및 주어진 X 및 Y 값을 갖는 모든 레코드에 대해 그 위치로부터 순차적으로 검색하는 것에 의해 수행된다. 통상적으로, 이것은 주어진 문서 및 페이지에서의 W 인치 폭 및 H 인치 높이를 나타내는 패치 바깥쪽에 있는 M-인치 링에서의 x-y 좌표를 포함한다. 이 링에 있는 레코드를 찾아내고 포인터를 역추적함으로써 그의 키 또는 피쳐(3421)를 찾아낸다. 링에 있는 피쳐 및 그의 x-y 좌표의 리스트가 도 34a의 3417에 나타난 바와 같이 보고된다. 3415에 나타난 W, H 및 M의 값은 피쳐(3417)가 입력 이미지 패치 밖에 있도록 입력 이미지의 크기에 기초하여 인식 시스템에 의해 동적으로 설정될 수 있다.

[0275] 이미지 데이터베이스 시스템(3400)의 이러한 특징들은, 예를 들어, 다수의 가설들을 명확하게 하는 데 유용하다. 데이터베이스 시스템(3400)이 2개 이상의 문서가 입력 이미지 패치와 정합할 수 있는 것으로 보고하는 경우, 패치 주변에 있는 링에 있는 피쳐들에 의해 인식 시스템(예를 들어, 지문 정합 모듈(226) 또는 다른 적당한 인식 시스템)이 사용자에게 이미지 캡처 장치를 결정을 명확하게 해주는 방향으로 약간 이동시키도록 지시함으로써 어느 문서가 사용자가 보유하고 있는 문서와 가장 잘 정합하는지를 결정할 수 있게 된다. 예를 들어(OCR-기반 피쳐가 사용되는 것으로 가정하면, 개념이 임의의 기하학적으로 인덱싱된 피쳐 세트로 확장되지 만), 문서 A 내의 이미지 패치는 단어쌍 "blue-xylophone" 바로 아래에 있을 수 있다. 문서 B 내의 이미지 패치는 단어쌍 "blue-thunderbird" 바로 아래에 있을 수 있다. 데이터베이스 시스템(3400)은 이들 피쳐의 예상된 위치를 보고할 수 있으며, 인식 시스템은 사용자에게 패치의 상부에서의 피쳐들의 y 좌표의 차이로 표시된 양만큼 (예를 들어, 사용자 인터페이스를 통해) 카메라를 위쪽으로 이동시키도록 지시할 수 있다. 인식 시스템은 그 차이 영역에서의 피쳐들을 계산할 수 있고 어느 것이 가장 잘 정합하는지를 결정하기 위해 문서 A 및 B로부터의 피쳐들을 사용할 수 있다. 예를 들어, 인식 시스템은 (xylophone, thunderbird)로 이루어지는 피쳐들의 "사전(dictionary)"을 사용하여 차이 영역으로부터의 OCR 결과를 후처리(post-process)할 수 있다. OCR 결과와 가장 잘 정합하는 단어가 입력 이미지와 가장 잘 정합하는 문서에 대응한다. 후처리 알고리즘의 일례는 통상적으로 알고 있는 스펠링 정정 기법(워드 프로세서 및 이메일 애플리케이션에 의해 사용되는 것 등)을 포함한다.

[0276] 이 예가 나타내는 바와 같이, 데이터베이스 시스템(3400) 설계에 의해 인식 시스템이 추가적인 데이터베이스 액



세스를 할 필요가 없도록 피쳐 기술들을 정합시킴으로써 효율적인 방식으로 다수의 좌표를 명확하게 할 수 있게 된다. 대안의 해결책은 각각의 이미지를 독립적으로 처리하는 것이다.

[0277] 동적 원시 이미지 발생

[0278] 이미지 패치의 위치에 대한  $x-y$  좌표(예를 들어, 이미지 패치의 중심에 대한  $x-y$  좌표)는 물론 문서 및 페이지의 ID도 역시 관계형 데이터베이스(3408)에 입력될 수 있으며, 여기서 이들은 그 문서 및 페이지에 대한 저장된 전자 원본을 검색하는 데 사용될 수 있다. 그 문서는 이어서 문서 렌더링 애플리케이션 모듈(3414)에 의해 비트맵 이미지로 렌더링될 수 있다. 또한, 모듈(3414)에 의해 제공되는 부가적인 "박스 크기" 값은 중심 부근의 비트맵의 일부분을 추출하기 위해 서브-이미지 추출 모듈(3416)에 의해 사용된다. 이 비트맵은 이미지 패치의 예상된 모습에 대한 "원시" 표시(pristine representation)이고, 이 비트맵은 입력 이미지에 존재해야만 하는 모든 피쳐들에 대한 정확한 표시를 포함한다. 원시 패치는 이어서 패치 특성(3412)으로서 반환될 수 있다. 이 해결책은 요구 시에 비트맵 데이터로 차후에 변환될 수 있는 컴팩트한 비이미지 표시(compact non-image representation)를 저장함으로써 이미지 비트맵을 저장하는 종래 기술의 요구된 과도한 저장 장치를 극복한다.

[0279] 이러한 저장 방식은 유익한데, 그 이유는 이 방식에 의해 이미지로부터 추출된 피쳐 표시가 상세한 피쳐 분석에 의해 명확하게 되는 일련의 후보를 검색하는 데 사용되는 가설 및 테스트 인식 전략(hypothesize-and-test recognition strategy)의 사용이 가능하게 되기 때문이다. 종종, 임의적인 일련의 후보들을 최적으로 명확하게 해주는 피쳐들을 예측하는 것이 가능하지 않으며, 이것이 그 후보들의 원래의 이미지로부터 결정되는 것이 바람직하다. 예를 들어, 단어쌍 "the cat"의 이미지는 2개의 데이터베이스 문서에 위치할 수 있으며, 그 중 하나는 처음에 Times Roman 폰트로 인쇄되고 다른 하나는 Helvetica 폰트로 인쇄되었다. 입력 이미지가 이들 폰트 중 하나를 포함하는지를 단지 결정하는 것으로 정확하게 정합하는 데이터베이스 문서를 식별하게 된다. 유클리드 거리와 같은 템플릿 정합 비교 메트릭을 사용하여 이들 문서에 대한 원시 패치를 입력 이미지 패치와 비교하는 것은 정확한 후보를 식별해준다.

[0280] 일례는 마이크로소프트 워드 ".doc" 파일을 저장하는 관계형 데이터베이스(3408)를 포함한다(유사한 방법이 포스트스크립트, PCL, pdf, 또는 마이크로소프트의 XML 용지 규격 XPS, 또는 고스트스크립트(ghostscript) 또는, XPS의 경우에, WinFX 컴포넌트가 설치되어 있는 마이크로소프트의 인터넷 익스플로러 등의 렌더링 애플리케이션에 의해 비트맵으로 변환될 수 있는 다른 이러한 형식 등의 다른 문서 형식에 대해 동작한다). 문서, 페이지,  $x-y$  위치, 박스 치수, 및 원하는 해상도가 600dpi(dots per inch)라는 것을 나타내는 시스템 파라미터에 대한 ID가 주어지면, 비트맵 이미지를 발생하기 위해 워드 애플리케이션이 호출될 수 있다. 이것은 6600 행 및 5100 열을 갖는 비트맵을 제공한다. 부가적인 파라미터  $x=3"$ ,  $y=3"$ ,  $height=1"$  및  $width=1"$ 는 데이터베이스가 페이지의 좌측 상부 모서리로부터  $x$  및  $y$ 에서 1800 픽셀 떨어져 있는 점에 중심이 있는 600 픽셀 높이 및 폭의 패치를 반환해야 한다는 것을 나타낸다.

[0281] 다중 데이터베이스

[0282] 각각이 서로 다른 문서 컬렉션을 포함할 수 있는 다중 데이터베이스 시스템(3400)이 사용될 때, 2개의 데이터베이스가 동일한 문서를 반환하는지 또는 어느 데이터베이스가 입력에 더 잘 정합하는 후보를 반환했는지를 결정하기 위해 원시 패치가 사용될 수 있다.

[0283] 2개의 데이터베이스가, 아마도 서로 다른 식별자(3410)(즉, 원래의 문서가 동일하지가 명백하지 않는데, 그 이유는 이들이 서로 다른 데이터베이스에 따라 입력되어 있기 때문임) 및 특성(3412)를 갖는, 동일한 문서를 반환할 때, 원시 패치는 거의 정확히 동일하다. 이것은, 예를 들어, 서로 다른 픽셀의 수를 카운트하는 해밍 거리를 사용하여 원시 패치들을 서로 비교함으로써 결정될 수 있다. 원래의 문서들이 픽셀끼리 정확하게 동일한 경우 해밍 거리는 0이 된다. 패치가 약간 다른 경우(사소한 폰트 차이로 야기될 수 있음) 해밍 거리는 0보다 약간 더 크다. 이것은 해밍 연산자에서 이미지 차이가 계산될 때 문자의 엣지 근방에 "후광" 효과(halo effect)를 야기할 수 있다. 이와 같은 폰트 차이는 원래의 렌더링 애플리케이션의 서로 다른 버전, 데이터베이스를 실행시키는 서버 상의 운영 체제의 서로 다른 버전, 서로 다른 프린터 드라이버, 또는 서로 다른 폰트 컬렉션에 의해 야기될 수 있다.

[0284] 2개의 문서에 있는 2개 이상의  $x-y$  위치로부터의 패치에 대해 원시 패치 비교 알고리즘이 수행될 수 있다. 이들 모두가 동일해야 하지만, 이것과 같은 샘플링 절차는 데이터베이스 시스템들 간의 렌더링 차이를 극복할 수 있는 중복성을 고려한다. 예를 들어, 2개의 시스템 상에서 렌더링될 때 하나의 폰트가 근본적으로 다르게 보일 수 있지만, 다른 폰트가 정확하게 동일할 수 있다.

- [0285] 2개 이상의 데이터베이스가 입력 이미지에 대한 그의 최상의 정합으로서 서로 다른 문서를 반환하는 경우, 어느 것이 올바른 것인지를 결정하기 위해 해밍 거리 등의 픽셀 기반 비교 메트릭에 의해 원시 패치가 입력 이미지와 비교될 수 있다.
- [0286] 2개 이상의 데이터베이스로부터의 결과를 비교하는 대안의 전략은 각각의 데이터베이스에 의해 보고된 문서들에서의 피쳐들의 기하학적 분포를 측정하는 누적기 어레이(accumulator array)의 콘텐츠를 비교하는 것이다. 원래의 피쳐 세트의 개별적인 탐색을 수행할 필요가 없도록 하기 위해 이 누적기가 데이터베이스에 의해 직접 제공되는 것이 바람직하다. 또한, 이 누적기는 데이터베이스 시스템(3400)의 콘텐츠와 독립적이어야만 한다. 도 34a에 도시된 실시예에서, 활동 어레이(activity array)(3420)가 내보내기(export)된다. 2개의 활동 어레이가 그의 값의 내부 분포(internal distribution)를 측정함으로써 비교될 수 있다.
- [0287] 보다 상세하게는, 2개 이상의 데이터베이스가 아마도 서로 다른 식별자(3410)(즉, 원래의 문서가 동일한지가 명백하지 않은데, 그 이유는 이들이 서로 다른 데이터베이스에 개별적으로 입력되었기 때문임) 및 특성(3412)을 갖는 동일한 문서를 반환하는 경우, 각각의 데이터베이스로부터의 활동 어레이(3420)가 거의 정확히 동일하다. 이것은, 예를 들어, 서로 다른 픽셀의 수를 카운트하는 해밍 거리를 사용하여 어레이들을 서로 비교함으로써 결정될 수 있다. 해밍 거리는 원래의 문서가 정확하게 동일한 경우 0이 될 것이다.
- [0288] 2개 이상의 데이터베이스가 입력 피쳐들에 대한 최상의 정합으로서 서로 다른 문서를 반환하는 경우, 이들의 활동 어레이(3420)는 어느 문서가 입력 이미지에 "가장 잘" 정합하는지를 결정하기 위해 비교될 수 있다. 이미지 패치에 정확하게 정합하는 활동 어레이는 패치가 있는 장소에 거의 중심이 있는 높은 값들의 클러스터(cluster)를 포함한다. 이미지 패치와 부정확하게 정합하는 활동 어레이는 랜덤하게 분포된 값들을 포함하게 된다. 엔트로피 등의 이미지의 분산(dispersion) 또는 랜덤성(randomness)을 측정하는 많은 공지된 전략들이 있다. 이러한 알고리즘은 클러스터의 존재를 나타내는 척도를 획득하기 위해 활동 어레이(3420)에 적용될 수 있다. 예를 들어, 이미지 패치에 대응하는 클러스터를 포함하는 활동 어레이(3420)의 엔트로피는 값들이 랜덤하게 분포되어 있는 활동 어레이(3420)의 엔트로피와 상당히 다르게 된다.
- [0289] 게다가, 유의할 점은 개개의 클라이언트(106)가 언제라도 콘텐츠가 반드시 서로 충돌하는 것은 아닌 다수의 데이터베이스(3400)에 액세스할 수 있다는 것이다. 예를 들어, 회사는 각각이 단일의 문서를 참조하는, 공개적으로 액세스가능한 패치 및 회사에 비밀인 패치 둘다를 가질 수 있다. 이러한 경우에, 클라이언트 장치(106)는 순서대로 조회되는 데이터베이스(D1, D2, D3, ...)의 리스트를 유지하고 사용자에게 통합하여 디스플레이하기 위해 결합된 활동 어레이(3420) 및 식별자(3410)를 생성한다. 주어진 클라이언트 장치(106)는 모든 데이터베이스로부터 이용가능한 패치를 디스플레이할 수 있거나 사용자가 데이터베이스의 서브셋(예를 들어, D1, D3 및 D7만)을 선택하고 그 데이터베이스로부터의 패치만을 보여줄 수 있게 해준다. 데이터베이스는 서비스에 가입함으로써 리스트에 추가될 수 있거나 클라이언트 장치(106)가 어떤 위치에 있을 때, 또는 데이터베이스가 클라이언트 장치(106)에 로드된 몇개 중의 하나이기 때문에, 또는 어떤 사용자가 현재 장치를 사용하고 있도록 인증되어 있기 때문에, 또는 심지어 장치가 어떤 모드에서 동작하고 있기 때문에, 무선으로 이용가능하게 될 수 있다. 예를 들어, 어떤 데이터베이스가 이용가능할 수 있는데, 그 이유는 특정의 클라이언트 장치가 그의 오디오 스피커를 턴온 또는 오프하기 때문이거나 비디오 프로젝터와 같은 주변 장치가 현재 클라이언트에 부착되어 있기 때문이다.
- [0290] 동작
- [0291] 도 34a를 더 참조하면, MMR 데이터베이스(3400)는 MMR 피쳐 추출 모듈(3402)로부터 일련의 피쳐들과 함께 동작을 수신한다. 동작은 명령 및 파라미터를 지정한다. 이러한 실시예에서, 명령 및 그것의 파라미터는 반환되는 패치 특성(3412)을 결정한다. 동작들은, 예를 들어, 용이하게 텍스트로 변환될 수 있는 http를 비롯한 형식으로 수신된다.
- [0292] 동작 처리기(3413)는 문서의 ID, 페이지 및 증거 누적 모듈(3406)에 의해 결정된 페이지 내에서의 x-y 위치를 수신한다. 이는 또한 명령 및 그의 파라미터를 수신한다. 동작 처리기(3413)는 명령을 관계형 데이터베이스(3408)를 사용하여 주어진 문서, 페이지 및 x-y 위치에 대응하는 위치에서 데이터를 검색하거나 저장하는 명령어로 변환하도록 프로그램되거나 다른 방식으로 구성된다.
- [0293] 한가지 이러한 실시예에서, 명령은 RETRIEVE, INSERT\_TO <DATA>, RETRIEVE\_TEXT <RADIUS>, TRANSFER <AMOUNT>, PURCHASE, PRISTINE\_PATCH <RADIUS [DOCID PAGEID X Y DPI]> 및 ACCESS\_DATABASE <DBID>를 포함한다. 각각에 대해 이제부터 차례로 설명할 것이다.



- [0294] RETRIEVE - 주어진 문서 페이지에서 x-y 위치에 링크된 데이터를 검색한다. 동작 처리기(3413)는 RETRIEVE 명령을 이 x-y 위치 근방에 저장될 수 있는 데이터를 검색하는 관계형 데이터베이스 질의로 변환한다. 이렇게 하려면 x-y 위치를 둘러싸고 있는 영역을 검색하기 위해 2개 이상의 데이터베이스 질의를 발행할 것을 필요로 할 수 있다. 검색된 데이터는 패치 특성(3412)으로서 출력된다. RETRIEVE 명령의 예시적인 애플리케이션은 비디오 클립 또는 동적 정보 객체(예를 들어, 현재의 정보가 검색될 수 있는 전자적 주소)를 검색하는 멀티미디어 브라우징 애플리케이션이다. 검색된 데이터는 MMR 장치 상에서 수행될 후속 단계들을 지정하는 메뉴를 포함할 수 있다. 이는 또한 JPEG 이미지 또는 비디오 클립 등의 전화 (또는 다른 디스플레이 장치) 상에 디스플레이될 수 있는 정적 데이터일 수 있다. 패치 특성을 찾기 위해 검색되는 영역을 결정하는 파라미터가 RETRIEVE 명령에 제공될 수 있다.
- [0295] INSERT\_TO <DATA> - <DATA>를 이미지 패치에 의해 지정되는 x-y 위치에 삽입한다. 동작 처리기(3413)는 INSERT\_TO 명령을, 지정된 x-y 위치에 데이터를 추가하는 관계형 데이터베이스에 대한 명령어로 변환한다. INSERT\_TO 명령의 성공적인 완료의 확인 응답이 패치 특성(3412)으로서 반환된다. INSERT\_TO 명령의 예시적인 애플리케이션은 사용자가 텍스트의 구절에서의 임의적인 x-y 위치에 데이터를 첨부할 수 있게 해주는 MMR 장치 상에서의 소프트웨어 애플리케이션이다. 데이터는 JPEG 이미지, 비디오 클립 또는 오디오 파일 등의 정적 멀티미디어일 수 있지만, 주어진 위치와 연관된 동작들을 지정하는 메뉴 등의 임의적인 전자 데이터일 수도 있다.
- [0296] RETRIEVE\_TEXT <RADIUS> - 이미지 패치에 의해 결정되는 x-y 위치의 <RADIUS> 내에서 텍스트를 검색한다. <RADIUS>는, 예를 들어, 이미지 공간에서의 픽셀의 수로서 지정될 수 있거나, 증거 누적 모듈(3406)에 의해 결정되는 x-y 위치 주변의 단어들의 문자의 수로서 지정될 수 있다. <RADIUS>는 또한 파싱된 텍스트 객체를 말할 수 있다. 이 특성의 실시예에서, 동작 처리기(3413)는 RETRIEVE\_TEXT 명령을 적절한 텍스트를 검색하는 관계형 데이터베이스 질의로 변환한다. <RADIUS>가 파싱된 텍스트 객체를 지정하는 경우, 동작 처리기는 파싱된 텍스트 객체를 반환할 뿐이다. 파싱된 텍스트 객체가 지정된 x-y 위치 근방에 위치하지 않는 경우, 동작 처리기는 널 표시(null indication)를 반환한다. 대안의 실시예에서, 동작 처리기는 주어진 x-y 위치의 반경 내에 있는 텍스트를 검색하기 위해 피드백-관련 피쳐 검색 모듈을 호출한다. 텍스트 문자열이 패치 특성(3412)으로서 반환된다. 텍스트 문자열 내의 각각의 단어와 연관되어 있는 선택적인 데이터는 원래의 문서 내의 그의 x-y 경계 박스를 포함한다. RETRIEVE\_TEXT 명령의 예시적인 응용은 다른 문서 내에 포함시키기 위해 인쇄된 문서로부터 텍스트 구문을 선택하는 것이다. 이것은, 예를 들어, MMR 시스템 상에서 프리젠테이션 파일을 (예를 들어, 파워포인트 형식으로) 작성하는 데 사용될 수 있다.
- [0297] TRANSFER <AMOUNT> - 문서 전체 및 다른 데이터베이스로 로드될 수 있는 형태로 그에 링크되어 있는 데이터 중 일부를 검색한다. <AMOUNT>는 검색되는 데이터의 수 및 유형을 지정한다. <AMOUNT>가 ALL인 경우, 동작 처리기(3413)는 문서와 연관되어 있는 데이터 전부를 검색하는 명령을 데이터베이스(3408)에 발행한다. 이러한 명령의 예는 DUMP 또는 Unix TAR을 포함한다. <AMOUNT>가 SOURCE인 경우, 문서에 대한 원래의 소스 파일이 검색된다. 예를 들어, 이것은 인쇄된 문서에 대한 워드(Word) 파일을 검색할 수 있다. <AMOUNT>가 BITMAP인 경우, 인쇄된 문서에 대한 비트맵의 JPEG-압축된 버전(또는 다른 통상적으로 사용되는 형식)이 검색된다. <AMOUNT>가 PDF인 경우, 문서의 PDF 표시가 검색된다. 검색된 데이터는 명령 이름(command name)의 덕분에 호출 애플리케이션이 알고 있는 형식으로 패치 특성(3412)으로서 출력된다. TRANSFER 명령의 예시적인 애플리케이션은 사용자가 작은 텍스트 영역을 이미지화(image)함으로써 문서에 대한 PDF 표시를 MMR 장치에 전송할 수 있게 해주는 "문서 그래버(document grabber)"이다.
- [0298] PURCHASE - 문서에서의 x-y 위치에 링크된 제품 상세(product specification)를 검색한다. 동작 처리기(3413)는 먼저 주어진 x-y 위치 근방에 있는 제품 상세를 획득하기 위해 일련의 하나 이상의 RETRIEVE 명령을 수행한다. 제품 상세는, 예를 들어, 판매자 이름, 제품 ID(예를 들어, 재고 번호) 및 판매자의 전자 주소(electronic address)를 포함한다. 제품 상세는 근방에 위치할 수 있는 다른 데이터 유형에 우선하여 검색된다. 예를 들어, jpeg가 이미지 패치에 의해 결정된 x-y 위치에 저장되어 있는 경우, 그 다음으로 가장 가까운 제품 상세가 그 대신에 검색된다. 검색된 제품 상세는 패치 특성(3412)으로서 출력된다. PURCHASE 명령의 예시적인 애플리케이션은 인쇄된 문서에 광고하는 것과 연관되어 있다. MMR 장치 상의 소프트웨어 애플리케이션은 광고와 연관된 제품 상세를 수신하고, 이를 지정된 전자 주소에 있는 지정된 판매자에 전송하기 이전에 사용자의 개인 식별 정보(예를 들어, 이름, 발송 주소, 신용 카드 번호, 기타 등등)를 추가한다.
- [0299] PRISTINE\_PATCH <RADIUS [DOCID PAGEID X Y DPI]> - 지정된 문서에 대한 전자적 표시를 검색하고 반경 RADIUS을 갖는 x-y에 중심이 있는 이미지 패치를 추출한다. RADIUS는 원의 반경을 지정할 수 있지만, 직사각형 패치(예를 들어, 2 인치 높이 x 3인치 폭)도 지정할 수 있다. 이는 또한 문서 페이지 전체를 지정할 수도 있다.

(DocID, PG, x, y) 정보가 동작의 일부로서 명시적으로 제공될 수 있거나 텍스트 패치의 이미지로부터 도출될 수 있다. 동작 처리기(3413)는 관계형 데이터베이스(3408)로부터 문서에 대한 원래의 표시를 검색한다. 그 표시가 비트맵일 수 있지만 렌더링가능한 전자 문서일 수도 있다. 원래의 표시가 문서 렌더링 애플리케이션(3414)으로 전달되며, 이곳에서 그 표시가 (파라미터 DPI(dots per inch)로 제공되는 해상도를 갖는) 비트맵으로 변환된 다음에 서브-이미지 추출(3416)에 제공되고, 여기서 원하는 패치가 추출된다. 이 패치 이미지는 패치 특성(3412)으로서 반환된다.

[0300] ACCESS\_DATABASE <DBID> - 데이터베이스(3400)를 클라이언트(106)의 데이터베이스 리스트에 추가한다. 클라이언트는 이제, 현재 리스트에 있는 임의의 기존의 데이터베이스에 부가하여, 이 데이터베이스(300)에 조회할 수 있다. DBID는 파일 또는 지정된 데이터베이스에 대한 원격 네트워크 참조를 지정한다.

[0301] 인덱스 테이블 발생 방법

[0302] 도 35는 본 발명의 일 실시예에 따른 MMR 인덱스 테이블을 발생하는 방법(3500)을 나타낸 것이다. 이 방법은, 예를 들어, 도 34a의 데이터베이스 시스템(3400)에 의해 수행될 수 있다. 한가지 이러한 실시예에서, MMR 인덱스 테이블은, 예를 들어, 스캔된 또는 인쇄된 문서로부터 MMR 인덱스 테이블 모듈(3404)(또는 어떤 다른 전용 모듈)에 의해 발생된다. 이 발생 모듈은, 본 명세서에 기술된 다른 모듈들처럼, 소프트웨어로, 하드웨어(예를 들어, 게이트-레벨 로직)로, 펌웨어(예를 들어, 이 방법을 수행하는 임베딩된 루틴들로 구성된 마이크로컨트롤러)로, 또는 이들의 어떤 조합으로 구현될 수 있다.

[0303] 이 방법은 종이 문서를 수신하는 단계(3510)를 포함한다. 이 종이 문서는 임의의 수의 페이지를 갖는 메모(예를 들어, 업무-관련, 사적인 편지), 제품 라벨(예를 들어, 통조림 제품, 의약, 박스 포장된 전자 장치), 제품 상세(예를 들어, 제설기(snow blower), 컴퓨터 시스템, 제조 시스템), 제품 브로셔 또는 광고 자료(예를 들어, 자동차, 투표, 휴양지), 서비스 설명 자료(예를 들어 인터넷 서비스 제공자, 청소 서비스), 책, 잡지 또는 다른 이러한 간행물로부터의 하나 이상의 페이지, 웹사이트로부터의 인쇄된 페이지, 필기 노트, 화이트보드로부터의 캡처된 및 인쇄된 노트, 또는 임의의 처리 시스템(예를 들어, 데스크톱 또는 휴대용 컴퓨터, 카메라, 스마트폰, 원격 단말기)으로부터의 인쇄된 페이지 등의 임의의 문서일 수 있다.

[0304] 이 방법은 종이 문서의 전자 표시를 발생하는 단계(3512)로 계속되며, 이 표시는 문서에 나타난 피쳐들의 x-y 위치를 포함한다. 타겟 피쳐들은, 예를 들어, 문서 내의 개개의 단어, 문자 및/또는 글자일 수 있다. 예를 들어, 원래의 문서가 스캔되는 경우, 이는 먼저 OCR되고, 단어(또는 다른 타겟 피쳐) 및 그의 x-y 위치가 (예를 들어, 스캐너(127)의 문서 지문 정합 모듈(226'))의 동작에 의해) 추출된다. 원래의 문서가 인쇄되는 경우, 인덱싱 프로세스는 모든 글자(또는 다른 타겟 피쳐)의 폰트, 포인트 크기 및 x-y 경계 박스의 XML 형식으로 (예를 들어, 프린터(116)의 프린트 드라이버(316)의 동작에 의해) 정확한 표시를 수신한다. 이 경우에, 인덱스 테이블 발생은 단계(3514)에서 시작하는데 그 이유는 전자 문서가 정확하게 식별된 x-y 피쳐 위치와 함께 (예를 들어, 프린트 드라이버(316)로부터) 수신되기 때문이다. XML 이외의 형식이 본 개시 내용으로부터 명백하게 될 것이다. 마이크로소프트 워드, 아도브 아크로벳 및 포스트스크립트 등의 전자 문서가 이들을 프린트 드라이버로 "인쇄(printing)"함으로써 데이터베이스에 입력될 수 있으며, 프린트 드라이버의 출력은 파일로 보내지며 따라서 종이가 반드시 발생될 필요는 없다. 이것은 이하에 나타내는 XML 파일 구조의 생성을 트리거한다. 어떤 경우에도, XML은 물론 원래의 문서 형식(워드, 아크로벳, 포스트스크립트, 기타)이 식별자(데이터베이스에 추가된 i번째 문서에 대한 doci)를 할당받고, 그 식별자로 또한 문서가 캡처된 시간, 인쇄된 날짜, 인쇄를 트리거한 애플리케이션, 출력 파일의 이름, 기타를 비롯한 문서의 다른 "메타데이터" 특성에 기초하여 나중에 검색가능하도록 관계형 데이터베이스(3408)에 저장된다.

[0305] XML 파일 구조의 일례가 아래에 나타내어져 있다.

[0306] \$docID.xml :

[0307] <?xml version="1.0" ?>

[0308] <doclayout ID="00001234">

[0309] <setup>

[0310] <url>file url/path or null if not known</url>

[0311] <date>file printed date</date>

- [0312] <app>application that triggered print</app>
- [0313] <text>\$docID.txt</text>
- [0314] <prfile>name of output file</prfile>
- [0315] <dpi>dpi of page for x, y coordinates, eg.600</dpi>
- [0316] <width>in inch, like 8.5</width>
- [0317] <height>in inch, eg. 11.0</height>
- [0318] <imagescale>0.1 is 1/10th scale of dpi</imagescale>
- [0319] </setup>
- [0320] <page no="1">
- [0321] \$docID\_1.jpeg</image>
- [0322] <sequence box="x y w h">
- [0323] <text>this string of text</text>
- [0324] <font>any font info</font>
- [0325] <word box="x y w h">
- [0326] <text>word text</text>
- [0327] <char box="x y w h">a</char>
- [0328] <char box="x y w h">b</char>
- [0329] <char>l entry per char, in sequence</char>
- [0330] </word>
- [0331] </sequence>
- [0332] </page>
- [0333] </doclayout>
- [0334] 한 특정의 실시예에서, 단어는 a-z, A-Z, 0-9 및 @%\$#;로부터의 임의의 문자를 포함할 수 있고, 그 밖의 모든 것은 구분 문자(delimiter)이다. .xml 파일의 원래의 설명은 (예를 들어, 데이터베이스(320) 서버 등의 서버 상에서 실행되는) 인덱싱 프로세스에 의해 사용되는 프린트 캡처 소프트웨어(print capture software)에 의해 생성될 수 있다. 실제의 형식은 항상 진화하고 있으며 더 많은 요소를 포함하는데, 그 이유는 시스템에 의해 새 문서가 획득되기 때문이다.
- [0335] 프린트 드라이버(예를 들어, 프린트 드라이버(316))에 의해 수신되는 원래의 텍스트 시퀀스가 보존되고 논리적 단어 구조가 "\_@%\$#"을 제외한 구두점(punctuation mark)에 기초하여 강요된다. XML 파일을 입력으로 사용하여, 인덱스 테이블 모듈(3404)은 페이지 경계를 존중하고, 먼저 2개의 연속적인 시퀀스 간의 수직 중첩의 정도를 검사함으로써 시퀀스들을 논리적 라인으로 그룹화하려고 시도한다. 한 특정의 실시예에서, 2개의 시퀀스가 그의 평균 높이의 절반보다 적게 중첩하는 경우 라인 분할이 일어났다는 발견적 방법(heuristic)이 사용된다. 이러한 발견적 방법은 통상적인 텍스트 문서(예를 들어, 마이크로소프트 워드 문서)에 잘 동작한다. 복잡한 레이아웃을 갖는 html 페이지의 경우, 부가적인 기하학적 분석이 필요할 수 있다. 그렇지만, 일관된 인덱싱 용어가 질의 프로세스에 의해 발생될 수 있는 한, 완벽한 의미론적 문서 구조를 추출하는 것이 필요하지 않다.
- [0336] 종이 문서의 전자 표시의 구조에 기초하여, 이 방법은 종이 문서의 모든 페이지 상의 모든 타겟 피처의 위치를 인덱싱하는 단계(3514)로 계속된다. 한 특정의 실시예에서, 이 단계는 종이 문서의 모든 페이지 상의 수평으로 및 수직으로 인접한 단어들의 모든 쌍의 위치를 인덱싱하는 단계를 포함한다. 앞서 설명한 바와 같이, 수평으로 인접한 단어는 라인 내에서의 이웃하는 단어들의 쌍이다. 수직으로 인접한 단어들은 수직으로 일렬로 정렬되어 있는 이웃하는 라인에 있는 단어들이다. 페이지의 다른 다차원 애스펙트(aspect)도 이와 유사하게 이

용될 수 있다.

- [0337] 이 방법은 또한 각각의 타겟 피처와 연관되어 있는 패치 특성을 저장하는 단계(3516)를 포함한다. 한 특성의 실시예에서, 패치 특성은 패치에 첨부된 동작들을 포함하며, 관계형 데이터베이스에 저장된다. 앞서 설명한 바와 같이, 이러한 이미지 인덱스 및 저장 설비의 결합은 이미지 패치와 정합하는 객체의 위치 및 그 패치의 특성을 고려한다. 이 특성은 메타데이터 등의 패치와 관련된 임의의 데이터일 수 있다. 이 특성은, 예를 들어, 특정의 기능을 수행하는 동작, 패치와 관련된 다른 콘텐츠에의 액세스를 제공하기 위해 선택될 수 있는 링크, 및/또는 패치와 관련된 다른 콘텐츠를 검색하게 하기 위해 스캔되거나 다른 방식으로 검색될 수 있는 바코드를 포함할 수 있다.
- [0338] 검색어 발생을 위해 보다 정확한 정의가 제공되며, 라인 구조의 프래그먼트만이 관찰된다. 수평으로 인접한 쌍의 경우, 질의어는 단어들을 "-" 분리자(separator)로 연결시킴으로써 형성된다. 수직쌍은 "+"를 사용하여 연결된다. 이들 단어는 원하는 경우 대문자로 시작하는 것(capitalization)을 보존하기 위해 그의 원래의 형태로 사용될 수 있다(이것은 더 독특한 용어를 생성하지만 대소문자 구별(case sensitivity) 등을 고려하기 위해 부가적인 질의 문제를 갖는 더 큰 인덱스를 생성한다). 이 인덱싱 방식은 동일한 검색 전략이 수평 또는 수직 단어쌍 또는 이 둘의 조합에 적용될 수 있게 해준다. 이들 경우들 중 임의의 것에 대해 역 문서 빈도(inverse document frequency)에 의해 용어들의 구별 능력이 고려된다.
- [0339] 증거 누적 방법(Evidence Accumulation Methodology)
- [0340] 도 36은, 본 발명의 일 실시예에 따른, 순위가 매겨진 일련의 문서, 페이지 및 위치 가설을 계산하는 방법(3600)을 나타낸 것이다. 이 방법은, 예를 들어, 도 34a의 데이터베이스 시스템(3400)에 의해 수행될 수 있다. 한가지 이러한 실시예에서, 증거 누적 모듈(3406)은 앞서 기술한 바와 같이 인덱스 테이블 모듈(3404)로부터의 데이터를 사용하여 가설을 계산한다.
- [0341] 이 방법은, 더 큰 문서 이미지의 이미지 패치 또는 문서 이미지 전체 등의, 타겟 문서 이미지를 수신하는 단계(3610)로 시작한다. 이 방법은 타겟 문서 이미지 내의 객체들 간의 2차원 관계를 캡처하는 하나 이상의 질의어를 발생하는 단계(3612)로 계속된다. 한 특성의 실시예에서, 질의어는, 도 34b를 참조하여 기술한 바와 같이, 수평 및 수직 단어쌍을 생성하는 피처 추출 프로세스에 의해 발생된다. 그렇지만, 본 명세서에 기술된 임의의 수의 피처 추출 프로세스는, 본 개시 내용으로부터 명백하게 될 것인 바와 같이, 타겟 이미지 내의 객체들 간의 2차원 관계를 캡처하는 질의어를 생성하는 데 사용될 수 있다. 예를 들어, 방법(3500)의 인덱스를 작성하는 데 사용되는 동일한 피처 추출 기법이, 단계(3512)(중이 문서의 전자 표시를 발생함)를 참조하여 논의된 것 등의, 질의어를 발생하는 데 사용될 수 있다. 게다가, 유의할 점은 질의어의 2차원 애스펙트가 각각의 질의어에 개별적으로(예를 들어, 타겟 문서 내의 수평 및 수직 객체 둘다를 나타내는 단일의 질의어) 또는 일련의 검색어(수평 단어쌍인 제1 질의어 및 수직 단어쌍인 제2 질의어)에 적용될 수 있다는 것이다.
- [0342] 이 방법은 각각의 질의어와 연관된 위치들의 리스트를 검색하기 위해 용어 인덱스 테이블(3422) 내의 각각의 질의어를 탐색하는 단계(3614)로 계속된다. 각각의 위치에 대해, 이 방법은 위치를 포함하는 다수의 구역을 발생하는 단계(3616)로 계속된다. 모든 질의어가 처리된 후에, 이 방법은 모든 질의어와 가장 부합하는 구역을 식별하는 단계(3618)를 더 포함한다. 한가지 이러한 실시예에서, 모든 후보 구역에 대한 점수는 가중치만큼(예를 들어, 각각의 구역이 모든 질의어와 얼마나 부합하는지에 기초하여) 증가된다. 이 방법은(예를 들어, 사전 정의된 정합 임계값에 기초하여) 식별된 구역이 사전 정의된 정합 기준을 만족시키는지를 결정하는 단계(3620)로 계속된다. 만족시키는 경우, 이 방법은 그 구역을 타겟 문서 이미지에 정합하는 것(match)으로서 확인하는 단계(3622)로 계속된다(예를 들어, 그 구역을 포함할 가능성이 가장 많은 페이지가 액세스되어 다른 방식으로 사용될 수 있다). 그렇지 않은 경우, 이 방법은 그 구역을 거부하는 단계(3624)로 계속된다.
- [0343] 단어쌍이 "정규화된" 좌표 공간에서의 위치와 함께 용어 인덱스 테이블(3422)에 저장된다. 이것은 서로 다른 프린터 및 스캐너 해상도 간에 균일성을 제공한다. 한 특성의 실시예에서, 8.5" x 11" 페이지에 대해 85x110 좌표 공간이 사용된다. 이러한 경우에, 모든 단어쌍이 이 85x110 공간에서의 그것의 위치에 의해 식별된다.
- [0344] 검색의 효율성을 향상시키기 위해, 2-단계 프로세스가 수행될 수 있다. 첫번째 단계는 입력 이미지 패치를 포함할 가능성이 가장 많은 페이지를 찾아내는 단계를 포함한다. 두번째 단계는 패치의 중심일 가능성이 가장 많은 그 페이지 내에서의 x-y 위치를 계산하는 단계를 포함한다. 이러한 접근 방법은 첫번째 단계에서 진짜 최상의 정합이 누락될 수 있을 가능성을 도입한다. 그렇지만, 희박한 인덱싱 공간의 경우, 이러한 가능성이 좀처럼 없다. 따라서, 인덱스의 크기 및 원하는 성능에 따라, 이러한 효율성 향상 기법이 이용될 수 있다.



[0345] 한가지 이러한 실시예에서, 이하의 알고리즘이 입력 이미지 패치에서 검출된 단어쌍을 포함할 가능성이 가장 많은 페이지를 찾아내는 데 사용된다.

[0346] 각각의 주어진 단어쌍 wp에 대해,

[0347]     idf = 1/log(2 + num\_docs(wp))

[0348] wp가 있는 각각의 (doc, page)에 대해,

[0349]     Accum[doc, page] += idf;

[0350] end /\* For each (doc, page) \*/

[0351] end /\* For each wp \*/

[0352]     (maxdoc, maxpage) = max ( Accum[doc, page] );

[0353]     if (Accum[ maxdoc, maxpage ] > thresh\_page)

[0354]     return ( maxdoc, maxpage);

[0355] 이 기법은 각각의 단어쌍에 대한 역 문서 빈도(idf)를 그 단어쌍이 나타나는 문서 및 페이지에 의해 인덱싱된 누적기에 가산한다. num\_docs(wp)는 단어쌍 wp를 포함하는 문서의 수를 반환한다. 누적기는 증거 누적 모듈(3406)에 의해 구현된다. 그 누적기에서의 최대값이 임계값을 초과하는 경우, 그것이 패치에 대한 최상의 정합인 페이지로서 출력된다. 따라서, 이 알고리즘은 질의에서의 단어쌍에 가장 잘 정합하는 페이지를 식별하는 동작을 한다. 다른 대안으로서, 누적 어레이가 정렬될 수 있고 상위 N개의 페이지가 입력 문서와 정합하는 "N개의 최상의" 페이지로서 보고될 수 있다.

[0356] 이하의 증거 누적 알고리즘은, 본 발명의 일 실시예에 따른, 단일의 페이지 내의 입력 이미지 패치의 위치에 대한 증거를 누적한다.

[0357] 각각의 주어진 단어쌍 wp에 대해,

[0358]     idf = 1/log(2 + num\_docs(wp))

[0359] wp가 있는 각각의 (x,y)에 대해,

[0360]     (minx, maxx, miny, maxy) = extent(x,y);

[0361]     maxdist = maxdist(minx, maxx, miny, maxy);

[0362]     For i=miny to maxy do

[0363]     For j = minx to maxx do

[0364]     norm\_dist = Norm\_geometric\_dist (i, j, x, y, maxdist)

[0365]     Activity [i, j ] += norm\_dist;

[0366]     weight = idf \* norm\_dist;

[0367]     Accum2[i,j] += weight;

[0368]     end /\* for j \*/

[0369]     end /\* for I \*/

[0370]     end /\* For each (y,y) \*/

[0371]     end /\* For each \*/

[0372] 이 알고리즘은 입력 이미지 패치의 중심일 가능성이 가장 많은 85x110 공간에서 셀을 찾아내는 동작을 한다. 여기에 나타난 실시예에서, 이 알고리즘은 각각의 단어쌍(존(zone)이라고 함) 주변의 고정된 영역에 있는 셀에

가중치를 부가함으로써 이것을 행한다. 범위 함수(extent function)가 x,y 쌍을 제공받고, 이 함수는 주변의 고정된 크기 구역(1.5" 높이 및 2" 폭이 통상적임)에 대한 최소 및 최대값을 반환한다. 범위 함수는 경계 조건에 주의하며 그가 반환하는 값이 누적기의 외부에 있지 않도록(즉, 0보다 작거나 또는 x가 85보다 크거나 y가 110보다 크지 않도록) 한다. maxdist 함수는 경계 박스 좌표 (minx, maxx, miny, maxy)로 기술되는 경계 박스 내의 2개의 점 사이의 최대 유클리드 거리를 찾아낸다. 단어쌍의 역 문서 빈도와 셀과 존의 중심 간의 정규화된 기하학적 거리의 곱에 의해 결정되는 가중치가 존 내의 각각의 셀에 대해 계산된다. 이것은 중심에 가까운 셀을 더 멀리 떨어져 있는 셀보다 더 높게 가중한다. 모든 단어쌍이 이 알고리즘에 의해 처리된 후에, 최대값을 갖는 셀을 찾기 위해 Accum2 어레이가 검색된다. 그것이 임계값을 초과하는 경우, 그의 좌표가 이미지 패치의 위치로서 보고된다. Activity 어레이는 누적된 norm\_dist 값을 저장한다. 이들이 idf에 의해 스케일링되지 않기 때문에, 이들은 특정한 단어쌍을 포함하는 데이터베이스 내의 문서의 수를 고려하지 않는다. 그렇지만, 이들은 주어진 일련의 워드쌍에 가장 잘 정합하는 x-y 위치에 대한 2차원 이미지 표시를 제공한다. 게다가, Activity 어레이 내의 엔트리들은 데이터베이스에 저장된 문서들에 독립적이다. 통상적으로 내부적으로 사용되는 이 데이터 구조가 내부내기(3420)될 수 있다.

[0373] 정규화된 기하학적 거리가 본 발명의 일 실시예에 따라 여기에 나타난 바와 같이 계산된다.

[0374] Norm\_geometric\_dist(i, j, x, y, maxdist)

[0375] begin

[0376]  $d = \sqrt{(i-x)^2 + (j-y)^2};$

[0377] return ( maxdist - d );

[0378] end

[0379] 단어쌍의 위치와 존의 중심 간의 유클리드 거리가 계산되고 이것과 계산될 수 있는 최대 거리 간의 차이가 반환된다.

[0380] 모든 단어쌍이 증거 누적 알고리즘에 의해 처리된 후에, 최대값을 갖는 셀을 찾기 위해 Accum2 어레이가 검색된다. 그 값이 사전-정의된 임계값을 초과하는 경우, 그의 좌표가 이미지 패치의 중심의 위치로서 보고된다.

[0381] MMR 인쇄 아키텍처

[0382] 도 37a는 본 발명의 일 실시예에 따른 MMR 컴포넌트의 기능 블록도를 나타낸 것이다. 주된 MMR 컴포넌트는 연관된 프린터(116)를 갖는 컴퓨터(3705) 및/또는 공유 문서 주석(shared document annotation, SDA) 서버(3755)를 포함한다.

[0383] 컴퓨터(3705)는 기술 분야에 공지되어 있는 임의의 표준 데스크톱, 랩톱, 또는 네트워크화된 컴퓨터이다. 일 실시예에서, 컴퓨터는 도 1b를 참조하여 기술된 MMR 컴퓨터이다. 사용자 프린터(116)는, 본 명세서에 기술된 바와 같이, 임의의 표준 가정용, 사무실용 또는 상업용 프린터이다. 사용자 프린터(116)는 하나 이상의 인쇄된 페이지로 이루어진 종이 문서인 인쇄된 문서(118)를 생성한다.

[0384] SDA 서버(3755)는 공유 주석(shared annotation)의 방법과 연관되어 있는 정보, 애플리케이션, 및/또는 각종의 파일을 보유하는 표준 네트워크화된 또는 중앙집중식 컴퓨터이다. 예를 들어, 웹 페이지 또는 다른 문서들과 연관되어 있는 공유 주석이 SDA 서버(3755)에 저장된다. 이 예에서, 주석은 본 명세서에 기술되어 있는 바와 같이 MMR에서 사용되는 데이터 또는 상호작용이다. SDA 서버(3755)는 일 실시예에 따라 네트워크 접속을 통해 액세스가능하다. 일 실시예에서, SDA 서버(3755)는 도 1b를 참조하여 기술된 네트워크화된 미디어 서버(114)이다.

[0385] 컴퓨터(3705)는 각종의 컴포넌트를 더 포함하며, 이들 중 일부 또는 그 전부는 다양한 실시예에 따라 선택적이다. 일 실시예에서, 컴퓨터(3705)는 소스 파일(3710), 브라우저(3715), 플러그-인(3720), 기호적인 핫스팟 기술(symbolic hotspot description)(3725), 수정된 파일(3730), 캡처 모듈(3735), page\_desc.xml(3740), hotspot.xml(3745), 데이터 저장소(3750), SDA 서버(3755), 및 MMR 프린터 소프트웨어(3760)를 포함한다.

[0386] 소스 파일(3710)은 문서의 전자적 표시는인 임의의 소스 파일을 나타낸다. 예시적인 소스 파일(3710)은 하이퍼 텍스트 마크업 언어(HTML) 파일, 마이크로소프트® 워드® 파일, 마이크로소프트® 파워포인트® 파일, 간단한 텍스트 파일, PDF(portable document format) 파일, 기타 등등을 포함한다. 본 명세서에 기술된 바와

같이, 브라우저(3715)에 수신된 문서는 많은 경우에 소스 파일(3710)로부터 온 것이다. 일 실시예에서, 소스 파일(3710)은 도 3을 참조하여 기술된 소스 파일(310)과 동등하다.

[0387] 브라우저(3715)는 소스 파일(3710)과 연관되어 있는 데이터에의 액세스를 제공하는 애플리케이션이다. 예를 들어, 브라우저(3715)는 소스 파일(3710)로부터 웹 페이지 및/또는 문서를 검색하는 데 사용될 수 있다. 일 실시예에서, 브라우저(3715)는 도 3을 참조하여 기술된 SD 브라우저(312, 314)이다. 일 실시예에서, 브라우저(3715)는 인터넷 익스플로러 등의 인터넷 브라우저이다.

[0388] 플러그-인(3720)은 저작 기능(authoring function)을 제공하는 소프트웨어 애플리케이션이다. 플러그-인(3720)은 독립형 소프트웨어 애플리케이션 또는 다른 대안으로서 브라우저(3715) 상에서 실행되는 플러그-인이다. 일 실시예에서, 플러그-인(3720)은 본 명세서에 기술된 특정의 기능을 제공하기 위해 브라우저(3715) 등의 애플리케이션과 상호작용하는 컴퓨터 프로그램이다. 플러그-인(3720)은 여러가지 실시예들에 따라 브라우저(3715)에 디스플레이된 문서 또는 웹 페이지에 다양한 변환 및 다른 수정을 수행한다. 예를 들어, 플러그-인(3720)은 핫스팟을 생성하기 위해 개별적으로 구별가능한 기점 마크(fiducial mark)로 핫스팟 지정(hotspot designation)을 둘러싸고 HTML 파일의 "마크업" 버전을 브라우저(3715)에 반환하며, 브라우저(3715)에 디스플레이된 문서의 일부분에 변환 규칙을 적용하고, 브라우저(3715)에 디스플레이된 문서에 대한 공유 주석을 검색 및/또는 수신한다. 그에 추가하여, 플러그-인(3720)은, 본 명세서에 기술된 바와 같이, 수정된 문서를 생성하는 것 및 기호적인 핫스팟 기술(3725)을 생성하는 것 등의 다른 기능들을 수행할 수 있다. 플러그-인(3720)은, 캡처 모듈(3735)을 참조하여, 도 38, 도 44, 도 45, 도 48 그리고 도 50a 및 도 50b를 참조하여 기술된 방법들을 용이하게 해준다.

[0389] 기호적인 핫스팟 기술(3725)은 문서 내의 핫스팟을 식별해주는 파일이다. 기호적인 핫스팟 기술(3725)은 핫스팟 번호 및 콘텐츠를 식별해준다. 이 예에서, 기호적인 핫스팟 기술(3725)은 데이터 저장소(3750)에 저장된다. 기호적인 핫스팟 기술의 일례가 도 41에 더 상세히 나타내어져 있다.

[0390] 수정된 파일(3730)은 플러그-인(3720)에 의해 소스 파일(3710)의 수정 및 변환의 결과로서 생성되는 문서 및 웹 페이지이다. 예를 들어, 상기한 바와 같은 마크업 HTML 파일은 수정된 파일(3730)의 일례이다. 수정된 파일(3730)은, 본 개시 내용으로부터 명백하게 될 것인 바와 같이, 어떤 경우에 사용자에게 디스플레이하기 위해 브라우저(3715)에 반환된다.

[0391] 캡처 모듈(3735)은, 인쇄된 페이지 상의 문자 및 그래픽의 레이아웃이 검색될 수 있도록, 문서들의 인쇄된 표시에 대해 피쳐 추출 및/또는 좌표 캡처를 수행하는 소프트웨어 애플리케이션이다. 레이아웃, 즉 인쇄된 페이지 상에서의 텍스트의 2차원 배열이 인쇄 시에 자동적으로 캡처될 수 있다. 예를 들어, 캡처 모듈(3735)은 모든 텍스트 및 도면 인쇄 명령을 실행하고, 그에 추가하여, 인쇄된 표시에서의 모든 문자 및/또는 이미지의 x-y 좌표 및 다른 특성을 가로채기(intercept)하여 기록(record)한다. 일 실시예에 따르면, 캡처 모듈(3735)은, 기존의 DLL의 기능의 추가 또는 수정을 가능하게 해주는 포워딩 DLL(forwarding Dynamically Linked Library)인, 본 명세서에 기술되는 Printcapture DLL이다. 캡처 모듈(3735)의 기능에 대한 보다 상세한 설명이 도 44를 참조하여 기술된다.

[0392] 당업자라면 캡처 모듈(3735)이 데이터의 캡처를 위해 브라우저(3715)의 출력에 연결되어 있다는 것을 잘 알 것이다. 다른 대안으로서, 캡처 모듈(3735)의 기능들은 프린터 드라이버 내에 직접 구현될 수 있다. 일 실시예에서, 캡처 모듈(3735)은, 도 3을 참조하여 기술된 바와 같이, PD 캡처 모듈(318)과 동등하다.

[0393] page\_desc.xml(3740)은 텍스트 관련되어 있는 캡처 모듈(3735)에 의해 처리되는 함수 호출을 위해 텍스트-관련 출력이 기록되는 확장가능 마크업 언어("XML") 파일이다. page\_desc.xml(3740)은 단어별로 또 문자별로 모든 인쇄된 텍스트의 문서에 대한 좌표 정보는 물론 핫스팟 정보, 프린터 포트 이름, 브라우저 이름, 인쇄 날짜 및 시간, dpi(dots per inch) 및 해상도(res) 정보를 포함한다. page\_desc.xml(3740)은 예를 들어 데이터 저장소(3750)에 저장된다. 데이터 저장소(3750)는 도 34a를 참조하여 기술된 MMR 데이터베이스(3400)와 동등하다. 도 42a 및 도 42b는 HTML 파일에 대한 page\_desc.xml(3740)의 일례를 더 상세히 나타낸 것이다.

[0394] hotspot.xml(3745)은 문서가 인쇄될 때 (예를 들어, 앞서 기술한 바와 같이, 프린트 드라이버(316)의 동작에 의해) 생성되는 XML 파일이다. hotspot.xml은 기호적인 핫스팟 기술(3725)과 page\_desc.xml(3740)을 병합한 결과이다. hotspot.xml은 핫스팟 번호, 좌표 정보, 크기 정보 및 핫스팟의 콘텐츠 등의 핫스팟 식별자 정보를 포함한다. hotspot.xml 파일의 일례가 도 43에 나타내어져 있다.

[0395] 데이터 저장소(3750)는 파일을 저장하는 기술 분야에 공지된 임의의 데이터베이스이며, 본 명세서에 기술된 방

법들에서 사용하기 위해 수정된다. 예를 들어, 일 실시예에 따르면, 데이터 저장소(3750)는 소스 파일(3710), 기호적인 핫스팟 기술(3725), page\_desc.xml(3740), 렌더링된 페이지 레이아웃, 공유 주식, 이미지화된 문서, 핫스팟 정의, 및 피쳐 표시를 저장한다. 일 실시예에서, 데이터 저장소(3750)는 도 3을 참조하여 기술된 문서 이벤트 데이터베이스(320) 및 도 34a를 참조하여 기술된 데이터베이스 시스템(3400)과 동등하다.

- [0396] MMR 인쇄 소프트웨어(3760)는 본 명세서에 기술된, 예를 들어, 앞서 기술한 바와 같이 컴퓨터(3705)의 컴포넌트에 의해 수행되는 MMR 인쇄 동작을 용이하게 해주는 소프트웨어이다. MMR 인쇄 소프트웨어(3760)에 대해서는 도 37b를 참조하여 이하에서 더 상세히 기술된다.
- [0397] 도 37b는 본 발명의 일 실시예에 따른 MMR 인쇄 소프트웨어(3760)를 비롯한 일련의 소프트웨어 컴포넌트를 나타낸 것이다. MMR 인쇄 소프트웨어(3760)의 전부 또는 그 일부가 컴퓨터(112, 905), 캡처 장치(106), 네트워크화된 미디어 서버(114) 및 본 명세서에 기술된 다른 서버들에 포함될 수 있다는 것을 잘 알 것이다. 이제부터, MMR 인쇄 소프트웨어(3760)가 이들 다른 컴포넌트를 포함하는 것으로 기술될 것이지만, 당업자라면 MMR 인쇄 소프트웨어(3760)가 하나부터 그 전부까지 임의의 수의 이들 컴포넌트를 가질 수 있다는 것을 잘 알 것이다. MMR 인쇄 소프트웨어(3760)는 컨버전 모듈(conversion module)(3765), 임베딩 모듈(embed module)(3768), 파싱 모듈(parse module)(3770), 변환 모듈(transform module)(3775), 피쳐 추출 모듈(feature extraction module)(3778), 주석 모듈(annotation module)(3780), 핫스팟 모듈(3785), 렌더링/디스플레이 모듈(3790), 및 저장 모듈(3795)을 포함한다.
- [0398] 컨버전 모듈(3765)은 소스 문서를 이미지화된 문서(이 이미지화된 문서로부터 피쳐 표시가 추출될 수 있음)로 컨버전하는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다.
- [0399] 임베딩 모듈(3768)은 전자 문서에서의 핫스팟에 대한 지점에 대응하는 마크를 임베딩하는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다. 한 특정의 실시예에서, 임베딩된 마크는 핫스팟의 시작점 및 핫스팟의 종료점을 나타낸다. 다른 대안으로서, 실시예 마크(embodiment mark) 주변의 사전-정의된 영역이 전자 문서에서의 핫스팟을 식별하는 데 사용될 수 있다. 다양한 이러한 마킹 방식이 사용될 수 있다.
- [0400] 파싱 모듈(3770)은 핫스팟에 대한 시작점을 나타내는 마크를 찾아 전자 문서(프린터에 전송됨)를 파싱하는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다.
- [0401] 변환 모듈(3775)은 변환 규칙을 전자 문서의 일부분에 적용하는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다. 한 특정의 실시예에서, 그 일부분은 핫스팟의 시작점을 나타내는 마크와 핫스팟의 종료점을 나타내는 마크 사이의 문자들의 스트림이다.
- [0402] 피쳐 추출 모듈(3778)은 문서 및 핫스팟의 인쇄된 표시에 대응하는 피쳐의 추출 및 좌표의 캡처를 가능하게 해주며 그렇게 하는 하나의 수단이다. 좌표 캡처는 포워딩 동적 링크된 라이브러리를 사용하여 인쇄 명령을 태핑하는 것(tapping print command) 및 핫스팟 또는 변환된 문자에 대응하는 좌표의 서브셋에 대한 인쇄된 표시를 파싱하는 것을 포함한다. 피쳐 추출 모듈(3778)은 일 실시예에 따라 캡처 모듈(3735)의 기능을 가능하게 해준다.
- [0403] 주석 모듈(3780)은 공유 주식 및 그에 수반하는 공유 주식과 연관된 문서의 일부분의 표시를 수신하는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다. 공유 주석을 수신하는 단계는 최종 사용자로부터 또 SDA 서버로부터 주석을 수신하는 것을 포함한다.
- [0404] 핫스팟 모듈(3785)은 하나 이상의 클립을 하나 이상의 핫스팟과 연관시키는 것을 가능하게 해주며, 그렇게 하는 하나의 수단이다. 핫스팟 모듈(3785)은 또한 먼저 문서 내의 핫스팟에 대한 위치를 지정하고 핫스팟과 연관된 클립을 정의함으로써 핫스팟 정의를 작성(formulation)하는 것을 가능하게 해준다.
- [0405] 렌더링/디스플레이 모듈(3790)은 문서 또는 문서의 인쇄된 표시가 렌더링 또는 디스플레이될 수 있게 해주고, 그렇게 하는 하나의 수단이다.
- [0406] 저장 모듈(3795)은 페이지 레이아웃, 이미지화된 문서, 핫스팟 정의, 및 피쳐 표시를 비롯한 다양한 파일의 저장을 가능하게 해주고, 그렇게 하는 하나의 수단이다.
- [0407] 소프트웨어 일부분(3765-3795)은 개별의 소프트웨어 모듈일 필요가 없다. 도시된 소프트웨어 구성은 단지 예에 불과하며, 본 개시 내용으로부터 명백하게 될 것인 바와 같이, 본 발명의 범위 내에서 다른 구성이 생각된다.



- [0408] 핫스팟을 문서에 임베딩
- [0409] 도 38은 본 발명의 일 실시예에 따라 문서에 핫스팟을 임베딩하는 방법의 플로우차트를 나타낸 것이다.
- [0410] 이 방법에 따르면, 문서 내의 핫스팟에 대한 지정에 대응하는 마크가 문서에 임베딩(3810)된다. 일 실시예에서, 핫스팟 지정 위치를 포함하는 문서가 브라우저에 디스플레이하기 위해 수신되며, 예를 들어, 소스 파일(3710)로부터 문서가 브라우저(3715)에서 수신된다. 핫스팟은 얼마간의 텍스트 또는 그래픽이나 사진은 물론 전자 데이터 등의 다른 문서 객체를 포함한다. 전자 데이터는 오디오 또는 비디오 등의 멀티미디어를 포함할 수 있거나 핫스팟에 액세스될 때 캡처 장치 상에서 수행되는 일련의 단계일 수 있다. 예를 들어, 문서가 하이퍼텍스트 마크업 언어(HTML) 파일인 경우, 브라우저(3715)는 인터넷 익스플로러일 수 있으며, 그 지정은 HTML 파일 내의 URL(Uniform Resource Locator)일 수 있다. 도 39a는 URL(3920)을 갖는 이러한 HTML 파일(3910)의 일례를 나타낸 것이다. 도 40a는 브라우저(4010), 예를 들어, 인터넷 익스플로러에 디스플레이된 바와 같이 도 39a의 HTML 파일(3910)의 텍스트를 나타낸 것이다.
- [0411] 마크를 임베딩(3810)하기 위해, 브라우저(3715)에 대한 플러그-인(3720)은 핫스팟을 생성하기 위해 개별적으로 구별가능한 기점 마크로 각각의 핫스팟 지정 위치를 둘러싼다. 일 실시예에서, 플러그-인(3720)은 브라우저(3715)에 디스플레이된 문서, 예를 들어, 상기한 예에 계속하여 인터넷 익스플로러에 디스플레이된 HTML을 수정하고, 핫스팟 지정 위치(예를 들어, URL)를 괄호 안에 넣은 마크, 또는 태그를 삽입한다. 이 마크는 브라우저(3715) 또는 문서의 인쇄 버전에서 문서를 보는 최종 사용자에게 보이지 않지만, 인쇄 명령에서 검출될 수 있다. 이 예에서, 본 명세서에서 MMR Courier New라고 하는 새로운 폰트가 시작 및 종료 기점 마크를 추가하는 데 사용된다. MMR Courier New 폰트에서, 문자 "b", "e" 및 숫자에 대한 통상적인 그림 문자(glyph) 또는 도트 패턴이 빈 공백으로 표시된다.
- [0412] 다시 도 39a 및 도 40a에 도시된 예시적인 HTML 페이지를 참조하면, 플러그-인(3720)은 식별자 "0."를 갖는 핫스팟을 나타내기 위해 기점 마크 "b0"를 URL의 시작("here")에 또 기점 마크 "e0"를 URL의 끝에 임베딩(3810)한다. b, e 및 숫자 문자가 공백으로 나타내어지기 때문에, 사용자는 문서의 외관에서 변화를 거의 또는 전혀 알지 못한다. 게다가, 플러그-인(3720)은, 도 41에 나타낸 바와 같이, 이들 마크를 나타내는 기호적인 핫스팟 기술(3725)을 생성한다. 기호적인 핫스팟 기술(3725)은 핫스팟 번호를 0으로 식별하며(4120), 이는 "b0" 및 "e0" 기점 마커에서의 0에 대응한다. 이 예에서, 기호적인 핫스팟 기술(3725)은, 예를 들어, 데이터 저장소(3750)에 저장된다.
- [0413] 플러그-인(3720)은, 도 39b에 나타낸 바와 같이, HTML(3950)의 "마크업" 버전을 브라우저(3715)로 반환한다. 마크업 HTML(3950)은 폰트를 1-포인트 MMR Courier New로 변경하는 스패ن 태그(span tag)(3960)로 기점 마크를 둘러싼다. b, e 및 숫자 문자가 공백으로 나타내어지기 때문에, 사용자는 문서의 외관에서의 변화를 거의 또는 전혀 보지 못한다. 마크업 HTML(3950)은 수정된 파일(3730)의 일례이다. 이 예는 간단함을 위해 단일 페이지 모델을 사용하지만, 다중 페이지 모델이 동일한 파라미터를 사용한다. 예를 들어, 핫스팟이 페이지 경계에 걸쳐 있는 경우, 이는 각각의 페이지 위치에 대응하는 기점 마크를 가지며, 각각에 대한 핫스팟 식별자는 동일하다.
- [0414] 그 다음에, 인쇄 명령에 응답하여, 인쇄된 표시는 및 핫스팟에 대응하는 좌표가 캡처된다(3820). 일 실시예에서, 캡처 모듈(3735)은 인쇄 명령 내의 텍스트 및 드로잉 명령을 "태핑"한다. 캡처 모듈(3735)은 모든 텍스트 및 드로잉 명령을 실행하고, 그에 부가하여, 인쇄된 표시는 내의 모든 문자 및/또는 이미지의 x-y 좌표 및 다른 특성을 가로채기하여 기록한다. 이 예에서, 캡처 모듈(3735)은, 출력 형식(즉, 프린터, 윈도우, 파일 형식, 메모리 버퍼, 기타)에 따라 출력될 텍스트 및/또는 이미지의 속성을 정의하는 인쇄된 표시는의 구조에 대한 핸들인, 인쇄된 표시는에 대한 장치 컨텍스트(Device Context, DC)를 참조한다. 인쇄된 표시에 대한 좌표를 캡처(3820)하는 프로세스에서, 핫스팟은 HTML 내의 임베딩된 기점 마크를 사용하여 용이하게 식별된다. 예를 들어, 시작 마크를 만날 때, 종료 마크가 발견될 때까지 모든 문자의 x-y 위치가 기록된다.
- [0415] 일 실시예에 따르면, 캡처 모듈(3735)은, 기존의 DLL의 기능을 추가 또는 수정할 수 있는, 본 명세서에서 "Printcapture DLL"이라고 하는 포워딩 DLL이다. 포워딩 DLL은 클라이언트에게 정확하게 원래의 DLL인 것처럼 보이지만, 호출이 타겟(원래의) DLL로 포워딩되기 이전에 기능들 중 일부 또는 그 전부에 부가적인 코드("탭(tap)")가 추가된다. 이 예에서, Printcapture DLL은 Windows GDI(Windows Graphics Device Interface) DLL gdi32.dll에 대한 포워딩 DLL이다. gdi32.dll은 600개 이상의 내보내기된 기능들을 가지며, 이들 모두가 포워딩될 필요가 있다. 본 명세서에서 gdi32.dll이라고 하는 Printcapture DLL은 클라이언트가 드로잉을 위해 DLL gdi32.dll을 사용하는 임의의 Windows 애플리케이션으로부터의 인쇄 출력(printout)을 캡처할 수 있게 해주며,

이는 원격 서버로 인쇄하는 경우에도 로컬 컴퓨터 상에서 실행되기만 하면 된다.

- [0416] 일 실시예에 따르면, gdi32\_mmr.dll은 gdi32.dll으로 이름 변경되고, C:\Windows\system32로 복사되어, 그로 하여금 거의 모든 Windows 애플리케이션으로부터의 인쇄를 모니터링하게 한다. 다른 실시예에 따르면, gdi32\_mmr.dll은 gdi32.dll으로 명명되고 인쇄가 모니터링되는 애플리케이션의 홈 디렉토리로 복사된다. 예를 들어, C:\Program Files\Internet Explorer는 Windows XP 상에서 인터넷 익스플로러를 모니터링한다. 이 예에서, 이 애플리케이션만이 자동적으로 Printcapture DLL 내의 기능들을 호출한다.
- [0417] 도 44는 본 발명의 일 실시예에 따른 포워딩 DLL에 의해 사용되는 프로세스의 플로우차트를 나타낸 것이다. Printcapture DLL gdi32\_mmr.dll은 먼저 gdi32.dll에 관련된 함수 호출을 수신한다(4405). 일 실시예에서, gdi32\_mmr.dll은 gdi32.dll에 관련된 모든 함수 호출을 수신한다. gdi32.dll은, 어떤 방식으로 인쇄된 페이지의 외관에 영향을 미치는 함수들에 대한 것인, 총 약 600개의 함수 호출 중 대략 200개를 모니터링한다. 따라서, Printcapture DLL은 이어서 수신된 호출이 모니터링된 함수 호출인지를 결정한다(4410). 수신된 호출이 모니터링된 함수 호출이 아닌 경우, 이 호출은 단계(4415 내지 4435)를 우회하고 gdi32.dll로 포워딩된다(4440).
- [0418] 수신된 호출이 모니터링된 함수 호출인 경우, 이 방법은 그 다음에 함수 호출이 "새로운" 프린터 장치 컨텍스트(DC)를 지정하는지, 즉 이전에 수신되지 않은 프린터 DC인지 여부를 결정한다(4415). 이것은 내부 DC 테이블과 대조하여 프린터 DC를 검사함으로써 결정된다. DC는 앞서 언급한 드로잉을 위한 타겟(프린터, 메모리 버퍼, 기타 등등일 수 있음)은 물론 폰트, 컬러, 기타 등등의 드로잉 설정을 캡슐화한다. 모든 드로잉 동작(예를 들어, LLineTo(), DrawText(), 기타 등등)이 DC에 대해 수행된다. 프린터 DC가 새로운 것이 아닌 경우, 프린터 DC에 대응하는 메모리 버퍼가 이미 존재하고, 단계(4420)를 건너뛴다. 프린터 DC가 새로운 것인 경우, 새로운 프린터 DC에 대응하는 메모리 버퍼 DC가 생성된다(4420). 이 메모리 버퍼 DC는 인쇄된 페이지의 외관을 미러링하고, 이 예에서, 상기한 인쇄된 표시와 동등하다. 따라서, 프린터 DC가 내부 DC 테이블에 추가될 때, 동일한 크기의 메모리 버퍼 DC(및 메모리 버퍼)가 생성되고 내부 DC 테이블 내의 프린터 DC와 연관된다.
- [0419] gdi32\_mmr.dll은 그 다음에 호출이 텍스트-관련 함수 호출인지 여부를 결정한다(4425). 200개의 모니터링된 gdi32.dll 호출 중 대략 12개가 텍스트-관련되어 있다. 텍스트-관련되어 있지 않은 경우, 단계(4430)를 건너뛴다. 함수 호출이 텍스트-관련되어 있는 경우, 텍스트-관련 출력이 본 명세서에서 page\_desc.xml(3740)(도 37a에 나타냄)이라고 하는 xml 파일에 기록된다(4430). page\_desc.xml(3740)은, 예를 들어, 데이터 저장소(3750)에 저장된다.
- [0420] 도 42a 및 도 42b는 도 39a 및 도 40a를 참조하여 기술된 HTML 파일(3910) 예에 대한 예시적인 page\_desc.xml(3740)을 나타낸 것이다. page\_desc.xml(3740)은 단어(4210)별도(예를 들어, Get), x, y, 폭 및 높이별도, 또한 문자(4220)별도(예를 들어, G), 모든 인쇄된 텍스트에 대한 좌표 정보를 포함한다. 모든 좌표는, 달리 언급하지 않는 한, 페이지의 좌측 상부 모서리에 대해 도트(dot)(픽셀에 대응하는 프린터에서의 용어)로 되어 있다. page\_desc.xml(3740)은 또한 시작 마크(4230) 및 종료 마크(4240) 등의 핫스팟 정보를 "시퀀스"의 형태로 포함한다. (예를 들어, 페이지 N과 페이지 N+1의) 페이지 경계에 걸쳐 있는 핫스팟의 경우, 이는 양쪽 페이지(N 및 N+1) 둘다에 나타나며, 양 경우에서의 핫스팟 식별자는 동일하다. 그에 추가하여, 생성된 .xml 및 .jpg 파일에 상당한 영향을 미칠 수 있는 프린터 포트 이름(4250), 브라우저(3715)(또는 애플리케이션) 이름(4260), 및 인쇄 날짜 및 시간(4270)은 물론 페이지(4280) 및 인쇄가능 구역(4290)에 대한 dpi(dots per inch) 및 해상도(res) 등의 다른 중요한 정보가 page\_desc.xml(3740)에 포함된다.
- [0421] 다시 도 44를 참조하면, 호출이 텍스트 관련되어 있지 않다는 결정 이후에, 또는 텍스트-관련 출력을 page\_desc.xml(3740)에 기록(4430)한 이후에, gdi32\_mmr.dll은 DC에 대한 메모리 버퍼에 대해 함수 호출을 실행한다(4435). 이 단계(4435)는 역시 로컬 컴퓨터 상의 메모리 버퍼로의 출력도 얻기 위해 프린터에 출력을 제공한다. 이어서, 페이지가 증가될 때, 메모리 버퍼의 내용이 압축되고 JPEG 및 PNG 형식으로 기록된다. 함수 호출은 이어서 gdi32.dll로 포워딩되어(4440), 정상적인 것처럼 그를 실행한다.
- [0422] 다시 도 38을 참조하면, 핫스팟을 포함하는 인쇄된 표시를 포함하는 페이지 레이아웃이 렌더링된다(3830). 일 실시예에서, 렌더링(3830)은 문서를 인쇄하는 것을 포함한다. 도 40b는 도 39a 및 도 40a의 HTML 파일(3910)의 인쇄 버전(4011)의 일례를 나타낸 것이다. 유의할 점은 기점 마크가 최종 사용자에게 시각적으로 지각되지 않는다는 것이다. 렌더링된 레이아웃은, 예를 들어, 데이터 저장소(3750)에 저장된다.
- [0423] 일 실시예에 따르면, Printcapture DLL은 기호적인 핫스팟 기술(3725) 내의 데이터와 page\_desc.xml(3740)(예를 들어, 도 42a 및 도 42b에 나타냄)을 hotspot.xml(3745)(도 43에 나타냄)로 병합한다. 이 예에서,

hotspot.xml(3745)은 문서가 인쇄될 때 생성된다. 도 43에서의 예는 핫스팟 0이 x=1303, y=350에 있고, 190 픽셀 폭 및 71 픽셀 높이임을 나타낸다. 핫스팟의 콘텐츠도 나타내어져 있으며, 즉 http://www.ricoh.com이다.

[0424] 캡처 모듈(3820)의 대안의 실시예에 따르면, 마이크로소프트 XPS(XML print specification) 프린트 드라이버에서의 필터(통상 "XPSDrv 필터"라고 함)는 텍스트 드로잉 명령을 수신하고 상기한 바와 같이 page\_desc.xml 파일을 생성한다.

[0425] 시각적으로 지각가능한 핫스팟

[0426] 도 45는 본 발명의 일 실시예에 따른 문서에서의 핫스팟에 대응하는 문자를 변환하는 방법의 플로우차트를 나타낸 것이다. 이 방법은 최종 사용자 및 MMR 인식 소프트웨어 둘다에 핫스팟이 존재한다는 것을 알려주도록 인쇄된 문서를 수정한다.

[0427] 처음에, 인쇄된 전자 문서가 문자 스트림으로서 수신된다(4510). 예를 들어, 문서는 프린터 드라이버에서 또는 문자 스트림을 필터링할 수 있는 소프트웨어 모듈에서 수신될 수 있다(4510). 일 실시예에서, 문서는 소스 파일(3710)로부터 브라우저(3715)에서 수신된다(4510). 도 46은 본 발명의 일 실시예에 따라 문서(4610)의 전자 버전의 일례를 나타낸 것이다. 이 예에서의 문서(4610)는 2개의 핫스팟을 가지며, 하나는 "이하에 열거된 것"과 관련되어 있고 하나는 "가능한 종래 기술"과 관련되어 있다. 핫스팟은 일 실시예에 따르면 최종 사용자에게 의해 시각적으로 지각가능하지 않다. 핫스팟은 도 38을 참조하여 기술된 좌표 캡처 방법을 통해 또는 본 명세서에 기술된 다른 방법들 중 임의의 것에 따라 설정될 수 있다.

[0428] 핫스팟의 시작을 나타내는 시작 마크(begin mark)를 찾기 위해 문서가 파싱된다(4520). 시작 마크는 앞서 기술한 기점 마크 또는 핫스팟을 식별해주는 임의의 다른 개별적으로 구별가능한 마크일 수 있다. 일단 시작 마크가 발견되면, 종료 마크가 발견될 때까지, 문서의 일부분에, 즉 시작 마크 이후의 문자에 변환 규칙이 적용된다(4530). 이 변환 규칙은, 예를 들어, 문자 폰트 또는 컬러를 수정함으로써 일 실시예에 따라 핫스팟에 대응하는 문서의 일부분의 시각적 수정을 야기한다. 이 예에서, 원래의 폰트, 예를 들어, Times New Roman은 다른 기점의 폰트, 예를 들어, OCR-A로 컨버전될 수 있다. 다른 예에서, 텍스트가 다른 폰트 컬러, 예를 들어, 청색 #F86A로 렌더링된다. 폰트를 변환하는 프로세스는 일 실시예에 따라 상기한 프로세스와 유사하다. 예를 들어, 문서(4610)가 HTML 파일인 경우, 문서(4510)에서 기점 마크를 만나는 경우, HTML 파일에서 폰트가 대체된다.

[0429] 일 실시예에 따르면, 변환 단계는 브라우저(3715)에 대한 플러그-인(3720)에 의해 달성되어, 수정된 문서(3730)를 생성한다. 도 47은 본 발명의 일 실시예에 따른 인쇄된 수정된 문서(4710)의 일례를 나타낸 것이다. 예시한 바와 같이, 핫스팟(4720, 4730)은 나머지 텍스트와 시각적으로 구별가능하다. 상세하게는, 핫스팟(4720)은 그의 다른 폰트에 기초하여 시각적으로 구별가능하고, 핫스팟(4730)은 그의 다른 컬러 및 밑줄에 기초하여 시각적으로 구별가능하다.

[0430] 그 다음에, 변환된 부분을 갖는 문서가 전자 문서 및 전자 문서 내에서의 핫스팟의 위치를 포함하는 페이지 레이아웃 내에 렌더링된다(4540). 일 실시예에서, 문서를 렌더링하는 것은 문서를 인쇄하는 것이다. 일 실시예에서, 렌더링하는 것은, 본 명세서에 기술된 그렇게 하는 방법들 중 임의의 방법에 따라, 변환된 부분을 갖는 문서에 대해 피쳐 추출을 수행하는 것을 포함한다. 일 실시예에서, 피쳐 추출은, 인쇄 명령에 응답하여, 일 실시예에 따라 전자 문서에 대응하는 페이지 좌표를 캡처하는 것을 포함한다. 변환된 문자에 대응하는 좌표의 서브셋을 찾기 위해 전자 문서가 파싱된다. 일 실시예에 따르면, 도 37a의 캡처 모듈(3735)은 피쳐 추출 및/또는 좌표 캡처를 수행한다.

[0431] MMR 인식 소프트웨어는 동일한 변환 규칙을 사용하여 모든 이미지를 전처리한다. 먼저, 이는 이 규칙을 따르는, 예를 들어, OCR-A 또는 청색 #F86A인 텍스트를 찾고, 이어서 그것의 정상적인 인식 알고리즘을 적용한다.

[0432] 본 발명의 이 양태가 유익한 이유는 MMR 인식 소프트웨어의 계산 부하를 상당히 감소시키기 때문인데, 왜냐하면 많은 양의 계산 오버헤드를 제거하는 아주 간단한 이미지 전처리 루틴을 사용하기 때문이다. 그에 부가하여, 이는, 예를 들어, 도 51a 내지 도 51d를 참조하여 기술된 바와 같이, 예를 들어, 문서의 일부분에 걸쳐 경계 박스가 있는 경우, 선택으로부터 적용될 수 있는 많은 수의 대안의 해결책을 제거함으로써 피쳐 추출의 정확도를 향상시킨다. 그에 부가하여, 텍스트의 가시적인 수정은 어느 텍스트(또는 다른 문서 객체)가 핫스팟의 일부인지를 최종 사용자에게 알려준다.

[0433] 공유 문서 주석

- [0434] 도 48은 본 발명의 일 실시예에 따른 공유 문서 주석의 방법의 플로우차트를 나타낸 것이다. 이 방법은 사용자가 공유 환경에서 문서에 주석을 첨부할 수 있게 해준다. 이하에 기술되는 실시예에서, 공유 환경은 다양한 사용자가 보는 웹 페이지이지만, 공유 환경은 다른 실시예에 따른, 작업 그룹 등의, 리소스가 공유되는 임의의 환경일 수 있다.
- [0435] 이 방법에 따르면, 소스 문서가 브라우저, 예를 들어, 브라우저(3715)에 디스플레이된다(4810). 일 실시예에서, 소스 문서는 소스 파일(3710)로부터 수신되고, 다른 실시예에서, 소스 문서는 네트워크, 예를 들어, 인터넷 접속을 통해 수신되는 웹 페이지이다. 웹 페이지 예를 사용하여, 도 49a는 본 발명의 일 실시예에 따른 브라우저 내의 샘플 소스 웹 페이지(4910)를 나타낸 것이다. 이 예에서, 웹 페이지(4910)는 인기있는 어린이 책 캐릭터와 관련된 게임, 즉 Jerry Butter Game에 대한 HTML 파일이다.
- [0436] 소스 문서의 디스플레이(4810) 시에, 공유 주석 및 소스 문서와 연관된 공유 주석과 연관된 소스 문서의 일부분의 지정이 수신된다(4820). 이 예에서 설명의 명확함을 위해 단일의 주석이 사용되지만, 다수의 주석이 가능하다. 이 예에서, 주석은 본 명세서에 기술된 바와 같이 MMR에서 사용되는 데이터 또는 상호작용이다. 일 실시예에 따르면, 주석은 공유 문서 주석 서버(SDA 서버)(예를 들어, 도 37a에 도시된 3755)에 저장되고 그로부터의 검색에 의해 수신된다. 일 실시예에서, SDA 서버(3755)는 네트워크 접속을 통해 액세스가능하다. 이 예에서, 공유 주석의 검색을 위한 플러그-인(예를 들어, 도 37a에 도시된 플러그-인(3720))이 이 기능을 용이하게 해준다. 다른 실시예에 따르면, 주석 및 지정이 사용자로부터 수신된다. 사용자는 어떤 주석도 갖지 않는 문서에 대한 공유 주석을 생성할 수 있거나, 기존의 공유 주석을 문서에 추가하거나 수정할 수 있다. 예를 들어, 사용자는 소스 문서의 일부분을 하이라이트하여, 역시 본 명세서에 기술된 다양한 방법들을 통해 사용자에게 의해 제공되는 공유 문서와 연관시키도록 그를 지정할 수 있다.
- [0437] 그 다음에, 수정된 문서가 브라우저에 디스플레이된다(4830). 수정된 문서는 단계(4820)에서 지정된 소스 문서의 일부분에 대응하는 핫스팟을 포함한다. 핫스팟은 공유 주석에 대한 위치를 지정한다. 수정된 문서는 일 실시예에 따라 플러그-인(3720)에 의해 생성되고 브라우저(3715)로 반환된 수정된 파일(3730)의 일부이다. 도 49b는 본 발명의 일 실시예에 따라 브라우저에 있는 수정된 웹 페이지(4920) 샘플을 나타낸 것이다. 웹 페이지(4920)는 핫스팟(4930)에 대한 지정 및 연관된 주석(4940)(이 예에서 비디오 클립임)을 나타내고 있다. 이 지정(4930)은, 예를 들어, 하이라이트에 의해 나머지 웹 페이지(4920) 텍스트와 시각적으로 구별될 수 있다. 일 실시예에 따르면, 지정(4930)이 클릭되거나 그 위에 마우스가 있을 때, 주석(4940)이 디스플레이된다.
- [0438] 인쇄 명령에 응답하여, 수정된 문서의 인쇄된 표시 및 핫스팟에 대응하는 텍스트 좌표가 캡처된다(4840). 좌표 캡처의 상세는 본 명세서에 기술된 그 목적을 위한 방법들 중 임의의 방법에 따른다.
- [0439] 이어서, 핫스팟을 포함하는 인쇄된 표시의 페이지 레이아웃이 렌더링된다(4850). 일 실시예에 따르면, 렌더링(4850)은 문서를 인쇄하는 것이다. 도 49c는 본 발명의 일 실시예에 따른 인쇄된 웹 페이지(4950) 샘플을 나타낸 것이다. 인쇄된 웹 페이지 레이아웃(4950)은 지정된 핫스팟(4930)을 포함하지만, 인쇄 레이아웃(4950)에서의 라인 단절이 웹 페이지(4920)와 다르다. 이 예에서 핫스팟(4930) 경계가 인쇄된 레이아웃(4950) 상에서 보이지 않는다.
- [0440] 선택적인 최종 단계에서, 공유 주석은 로컬적으로, 예를 들어, 데이터 저장소(3750)에 저장되고, 인쇄된 문서(4950)에서의 핫스팟(4930)과의 연관관계를 사용하여 인덱싱된다. 인쇄된 표시도 역시 로컬적으로 저장될 수 있다. 일 실시예에서, 인쇄하는 동작은 로컬 복사본의 다운로드 및 생성을 트리거한다.
- [0441] 이미지화된 문서에 대한 핫스팟
- [0442] 도 50a는 본 발명의 일 실시예에 따른 이미지화된 문서에 핫스팟을 추가하는 방법의 플로우차트를 나타낸 것이다. 이 방법은 종이 문서가 스캔된 후에 핫스팟이 종이 문서에 또는 인쇄를 위해 렌더링된 후에 기호적인 전자 문서에 추가될 수 있게 해준다.
- [0443] 먼저, 소스 문서가 이미지화된 문서로 컨버전된다(5010). 일 실시예에 따르면, 소스 문서가 소스 파일(3710)로부터 브라우저(3715)에서 수신된다. 이 컨버전(5010)은 피쳐 표시를 생성하기 위해 피쳐 추출이 수행될 수 있는 문서를 생성하는 임의의 방법에 의한다. 일 실시예에 따르면, 종이 문서가 이미지화된 문서가 되기 위해 스캔된다. 다른 실시예에 따르면, 전자 문서에 대한 렌더링가능한 페이지 교정쇄(page proof)가 적절한 애플리케이션을 사용하여 렌더링된다. 예를 들어, 렌더링가능한 페이지 교정쇄가 PostScript 형식으로 되어 있는 경우, Ghostscript가 사용된다. 도 51a는 일 실시예에 따라 스캔된 신문 페이지(5110)의 일부분을 보여주는 사용자 인터페이스(5105)의 일례를 나타낸 것이다. 주 윈도우(5115)는 신문 페이지(5110)의 확대된 일부분을



보여주며, 섬네일(5120)은 페이지의 어느 일부분이 디스플레이되고 있는지를 보여준다.

[0444] 그 다음에, 피쳐 표시를 생성하기 위해 이미지화된 문서에 피쳐 추출이 적용된다(5020). 본 명세서에 기술된 다양한 피쳐 추출 방법들 중 임의의 방법이 이를 위해 사용될 수 있다. 피쳐 추출은 일 실시예에 따라 도 37a를 참조하여 기술된 캡처 모듈(3735)에 의해 수행된다. 이어서, 하나 이상의 핫스팟(5125)이 이미지화된 문서에 추가된다(5030). 핫스팟은 사전-정의될 수 있거나 여러가지 실시예들에 따라 정의될 필요가 있을 수 있다. 핫스팟이 이미 정의되어 있는 경우, 이 정의는 페이지 번호, 페이지 상의 핫스팟에 대한 경계 바가스의 좌표 위치, 및 핫스팟에 첨부된 전자 데이터 또는 상호작용을 포함한다. 일 실시예에서, 핫스팟 정의는, 도 43에 나타난 바와 같이, hotspot.xml의 형태를 갖는다.

[0445] 핫스팟이 정의되지 않은 경우, 최종 사용자는 핫스팟을 정의할 수 있다. 도 50b는 본 발명의 일 실시예에 따라 이미지화된 문서에 추가하기 위한 핫스팟을 정의하는 방법의 플로우차트를 나타낸 것이다. 먼저, 후보 핫스팟이 선택된다(5032). 예를 들어, 도 51a에서, 최종 사용자는 경계 박스(5125)를 사용하여 문서의 일부분을 핫스팟으로서 선택한다. 그 다음에, 주어진 데이터베이스에 대해, 선택적인 단계(5034)에서, 핫스팟이 고유한 것인지가 결정된다. 예를 들어, 핫스팟을 일의적으로 식별하기 위해 주변의  $n \times n$  패치에 충분한 텍스트가 있어야만 한다.  $n$ 의 통상적인 값의 예는 2이다. 핫스팟이 데이터베이스에 대해 충분히 고유하지 않은 경우, 일 실시예에서 최종 사용자는 불명료를 어떻게 처리할지에 관한 옵션들을 제공받는다. 예를 들어, 사용자 인터페이스는 더 큰 영역을 선택하는 것 또는 불명료를 인정하지만 그에 대한 기술을 데이터베이스에 추가하는 것 등의 대안들을 제공할 수 있다. 다른 실시예들은 핫스팟을 정의하는 다른 방법들을 사용할 수 있다.

[0446] 일단 핫스팟 위치가 선택되면(5032), 데이터 또는 상호작용이 정의되고(5036) 핫스팟에 첨부된다. 도 51b는 선택된 핫스팟과 연관시킬 데이터 또는 상호작용을 정의하기 위한 사용자 인터페이스를 나타낸 것이다. 예를 들어, 일단 사용자가 경계 박스(5125)를 선택하였으면, 편집 박스(5130)가 디스플레이된다. 연관된 버튼들을 사용하여, 사용자는 동작을 취소할 수 있거나(5135), 간단히 경계 박스(5125)를 저장할 수 있거나(5140), 데이터 또는 상호작용을 핫스팟에 할당할 수 있다(5145). 사용자가 데이터 또는 상호작용을 핫스팟에 할당하기로 선택하는 경우, 도 51c에 나타난 바와 같이, 할당 박스(5150)가 디스플레이된다. 할당 박스(5150)는 최종 사용자가 이미지(5155), 다양한 다른 미디어(5160), 및 웹 링크(5165)를, ID 번호(5170)에 의해 식별되는 핫스팟에 할당할 수 있게 해준다. 이어서, 사용자는 핫스팟 정의를 저장(5175)하기로 선택할 수 있다. 간단함을 위해 하나의 핫스팟이 기술되어 있지만, 다수의 핫스팟이 가능하다. 도 51d는 문서 내에 핫스팟을 디스플레이하기 위한 사용자 인터페이스를 나타낸 것이다. 일 실시예에서, 서로 다른 컬러 경계 박스는 서로 다른 데이터 및 상호작용 유형에 대응한다.

[0447] 선택적인 단계에서, 이미지화된 문서, 핫스팟 정의, 및 피쳐 표시가 함께, 예를 들어, 데이터 저장소(3750)에 저장된다(5040).

[0448] 도 52는 본 발명의 일 실시예에 따른 MMR 문서(500) 및 MMR 시스템(100b)을 사용하는 방법(5200)을 나타낸 것이다.

[0449] 방법(5200)은 제1 문서 또는 제1 문서의 표시를 획득(5210)하는 것으로 시작한다. 제1 문서를 획득하는 예시적인 방법은 이하의 것들을 포함한다. (1) 제1 문서가 MMR 컴퓨터(112)의 운영 체제 내에서 인쇄된 문서의 텍스트 레이아웃을 PD 캡처 모듈(318)을 통해 자동적으로 캡처함으로써 획득된다. (2) 제1 문서가 MMR 컴퓨터(112)의 프린터 드라이버(316) 내에서 인쇄된 문서의 텍스트 레이아웃을 자동적으로 캡처함으로써 획득된다. (3) 제1 문서가, 예를 들어, MMR 컴퓨터(112)에 연결되어 있는 표준 문서 스캐너 장치(127)를 통해 종이 문서를 스캔함으로써 획득된다. (4) 제1 문서가 인쇄된 문서의 표시인 파일을 MMR 컴퓨터(112)로 자동적으로 또는 수작업으로 전송, 업로드 또는 다운로드함으로써 획득된다. 획득하는 단계가 인쇄된 문서의 대부분 또는 그 전부를 획득하는 것으로 기술되어 있지만, 획득하는 단계(5210)가 인쇄된 문서의 가장 작은 일부분에 대해서만 수행될 수 있다는 것을 잘 알 것이다. 게다가, 이 방법이 단일 문서를 획득하는 것과 관련하여 기술되어 있지만, 이 단계는 다수의 문서를 획득하고 제1 문서의 라이브러리를 생성하기 위해 수행될 수 있다.

[0450] 일단 획득하는 단계(5210)가 수행되면, 이 방법(5200)은 제1 문서에 대해 인텍싱 동작을 수행한다(5212). 인텍싱 동작은 획득된 제1 문서 또는 그의 일부분과 정합하는 입력을 위한 문서 및 연관된 제2 미디어 유형의 대응하는 전자적 표시의 식별을 가능하게 해준다. 이 단계의 일 실시예에서, 문서 인텍싱 동작은 PD 인텍싱(322)를 발생하는 PD 캡처 모듈(318)에 의해 수행된다. 예시적인 인텍싱 동작은 이하의 것들을 포함한다. (1) 인쇄된 문서의 문자들의 x-y 위치가 인텍싱된다. (2) 인쇄된 문서의 단어의 x-y 위치가 인텍싱된다. (3) 인쇄된 문서 내의 이미지 또는 이미지의 일부분의 x-y 위치가 인텍싱된다. (4) OCR 이미징 동작이 수행되고, 문자 및/또는

단어의 x-y 위치가 인덱싱된다. (5) 렌더링된 페이지의 이미지로부터의 피쳐 추출이 수행되고, 피쳐의 x-y 위치가 인덱싱된다. (6) 페이지의 기호적인 버전에 대한 피쳐 추출이 시물레이션되고, 피쳐의 x-y 위치가 인덱싱된다. 인덱싱 동작(5212)은 본 발명의 응용에 따라 상기한 인덱싱 동작들 또는 그의 그룹들 중 임의의 것을 포함할 수 있다.

[0451] 방법(5200)은 또한 제2 문서를 획득한다(5214). 이 단계(5214)에서, 획득된 제2 문서는 문서 전체 또는 제2 문서의 단지 일부분(패치)일 수 있다. 제2 문서를 획득하는 예시적인 방법들은 이하의 것들을 포함한다. (1) 캡처 장치(106)의 하나 이상의 캡처 메커니즘(230)에 의해 텍스트의 패치를 스캔하는 것. (2) 캡처 장치(106)의 하나 이상의 캡처 메커니즘(230)에 의해 텍스트의 패치를 스캔하고 이어서 의도된 피쳐 설명이 정확하게 추출될 가능성을 결정하기 위해 이미지를 전처리하는 것. 예를 들어, 인덱스가 OCR에 기초하는 경우, 이 시스템은 이미지가 텍스트 라인들을 포함하는지 및 이미지 선명도가 성공적인 OCR 동작에 충분한지를 결정할 수 있다. 이 결정이 실패하는 경우, 텍스트의 다른 패치가 스캔된다. (3) 스캔된 문서를 식별해주는 기계 판독가능 식별자(예를 들어, ISBN(international standard book number) 또는 UPC(universal produce code) 코드)를 스캔하는 것. (4) 요청되는 문서 또는 일련의 문서들(예를 들어, Sports Illustrated 잡지의 2003년판)을 식별해주는 데이터를 입력하고 이어서 텍스트 패치가 이 방법 단계의 항목 (1) 또는 (2)를 사용하여 스캔되는 것. (5) 제2 문서가 첨부되어 있는 이메일을 수신하는 것. (6) 파일 전송에 의해 제2 문서를 수신하는 것. (7) 캡처 장치(106)의 하나 이상의 캡처 메커니즘(230)으로 이미지의 일부분을 스캔하는 것. (8) 입력 장치(166)로 제2 문서를 입력하는 것.

[0452] 일단 단계(5210, 5214)가 수행되면, 이 방법은 제1 문서와 제2 문서 사이의 문서 또는 패턴 정합을 수행한다(5216). 일 실시예에서, 이것은 제2 문서의 제1 문서에 대한 문서 지문 정합을 수행함으로써 행해진다. 문서 지문 정합 동작은 PD 인덱스(322)에 질의를 함으로써 제2 미디어 문서에 대해 수행된다. 문서 지문 정합의 일례는 단계(5214)에서 캡처된 이미지로부터 피쳐들을 추출하고, 그 피쳐들로부터 기술자(descriptor)를 작성하며, 그 기술자들을 일정 비율 포함하는 문서 및 패치를 탐색하는 것이다. 각각의 문서에 대해 한번씩 이 패턴 정합 단계가 여러번 수행될 수 있고 라이브러리 또는 데이터베이스 내의 임의의 문서가 제2 문서와 정합하는지를 결정하기 위해 데이터베이스가 다수의 문서를 저장하고 있다는 것을 잘 알 것이다. 다른 대안으로서, 인덱싱 단계(5212)가 문서의 컬렉션을 나타내는 인덱스에 문서(5210)를 추가하고, 패턴 정합 단계가 한번 수행된다.

[0453] 마지막으로, 방법(5200)은 단계(5216)의 결과에 기초하여 또 선택에 따라서는 사용자 입력에 기초하여 동작을 실행한다(5218). 일 실시예에서, 방법(5200)은, 예를 들어, 단계(5216)에서 정합하는 것으로 밝혀진 핫스팟(506)과 연관된 제2 미디어(504)에 저장되어 있는 주어진 문서 패치와 연관되어 있는 미리 정해진 동작을 탐색한다. 미리 정해진 동작의 예는, (1) 문서 이벤트 데이터베이스(320), 인터넷, 또는 다른 곳에서 정보를 검색하는 것, (2) 시스템의 출력을 수신할 준비가 되어 있는 MMR 시스템(100b)에 의해 검증되는 위치에 정보를 기록하는 것. (3) 정보를 탐색하는 것. (4) 캡처 장치(106) 등의 클라이언트 장치 상에 정보를 디스플레이하고 사용자와 상호작용적 대화를 수행하는 것. (5) 나중에 실행하기 위해(사용자의 참여는 선택적일 수 있음) 방법 단계(5216)에서 결정되는 동작 및 데이터를 큐잉하는 것. (6) 방법 단계(5216)에서 결정되는 동작 및 데이터를 즉각 실행하는 것을 포함한다. 이 방법 단계의 예시적인 결과는 정보의 검색, 수정된 문서, 어떤 다른 동작(예를 들어, 주식 또는 제품의 구입)의 실행, 또는 비디오를 다시 케이블 TV 박스로 스트리밍하는 케이블 TV 서버(예를 들어, 서비스 제공자 서버(122))에 연결되어 있는 셋톱 박스(126) 등의 케이블 TV 박스에 전송되는 명령의 입력을 포함한다. 단계(5218)가 행해지면, 방법(5200)은 완료되어 종료된다.

[0454] 도 53은 본 발명의 일 실시예에 따른, MMR 시스템(100b)과 연관되어 있는 예시적인 일련의 비즈니스 개체(5300)의 블록도를 나타낸 것이다. 일련의 비즈니스 개체(5300)는 MMR 서비스 제공자(5310), MMR 소비자(5312), 멀티미디어 회사(5314), 프린터 사용자(5316), 셀 전화 서비스 제공자(5318), 하드웨어 제조업자(5320), 하드웨어 소매업자(5322), 금융 기관(5324), 신용 카드 처리기(5326), 문서 출판업자(5328), 문서 프린터(5330), 이행 하우스(fulfillment house)(5332), 케이블 TV 제공자(5334), 서비스 제공자(5336), 소프트웨어 제공자(5338), 광고 회사(5340), 및 비즈니스 네트워크(5370)를 포함한다.

[0455] MMR 서비스 제공자(5310)는 도 1a 내지 도 5 및 도 52를 참조하여 기술한 MMR 시스템(100)의 소유자 및/또는 관리자이다. MMR 소비자(5312)는, 도 1b를 참조하여 상기한 바와 같이, 임의의 MMR 사용자(110)를 나타낸다.

[0456] 멀티미디어 회사(5314)는 디지털 영화 및 비디오 게임을 제공하는 Blockbuster Inc.(Dallas, TX) 및 디지털 음악, 영화 및 TV 쇼를 제공하는 Sony Corporation of America(New York, NY) 등의 임의의 디지털 멀티미디어 제

품 제공자이다.

- [0457] 프린터 사용자(5316)는 인쇄된 종이 문서를 생성하기 위해 임의의 종류의 임의의 프린터를 이용하는 임의의 개인 또는 개체이다. 예를 들어, MMR 소비자(5312)는 프린터 사용자(5316) 또는 문서 프린터(5330)일 수 있다.
- [0458] 셀 전화 서비스 제공자(5318)는 Verizon Wireless(Bedminster, NJ), Cingular Wireless(Atlanta, GA), T-Mobile USA(Bellevue, WA), 및 Sprint Nextel(Reston, VA) 등의 임의의 셀 전화 서비스 제공자이다.
- [0459] 하드웨어 제조업자(5320)는 프린터, 셀룰러 전화, 또는 PDA의 제조업자 등의 임의의 하드웨어 장치의 제조업자이다. 예시적인 하드웨어 제조업자는 Hewlett-Packard(Houston, TX), Motorola, Inc.(Schaumburg, IL) 및 Sony Corporation of America(New York, NY)를 포함한다. 하드웨어 소매업자(5322)는 프린터, 셀룰러 전화, 또는 PDA의 소매업자 등의 임의의 하드웨어 장치 소매업자이다. 예시적인 하드웨어 소매업자는 RadioShack Corporation(Fort Worth, TX), Circuit City Stores, Inc.(Richmond, VA), Wal-Mart(Bentonville, AR), 및 Best Buy Co.(Richfield, MN)를 포함하지만, 이에 한정되는 것은 아니다.
- [0460] 금융 기관(5324)은 은행 계좌 및 다른 은행 또는 금융 기관으로의 자금 이체를 처리하는 임의의 은행 또는 신용 조합 등의 임의의 금융 기관이다. 신용 카드 처리기(5326)는 구입 거래를 위한 신용 카드 인증 및 승인 프로세스를 관리하는 임의의 신용 카드 기관이다. 예시적인 신용 카드 처리기는 Click Sales Inc, (Boise ID), ShareIt ! Inc.(Eden Prairie, MN), 및 CCNow Inc.(Eden Prairie, MN)의 서비스인 ClickBank를 포함하지만, 이에 한정되는 것은 아니다.
- [0461] 문서 출판업자(5328)는 The Gregath Publishing Company(Wyandotte, OK), Prentice Hall(Upper Saddle River, NJ), 및 Pelican Publishing Company(Gretna, LA) 등의 임의의 문서 출판 회사이지만, 이에 한정되는 것은 아니다. 문서 프린터(5330)는 PSPrint LLC(Oakland CA), PrintLizard, Inc.(Buffalo, NY), 및 Mimeo, Inc.(New York, NY) 등의 임의의 문서 인쇄 회사이지만, 이에 한정되는 것은 아니다. 다른 예에서, 문서 출판업자(5328) 및/또는 문서 프린터(5330)는 신문 또는 잡지를 제작 및 배포하는 임의의 개체이다.
- [0462] 이행 하우스(fulfillment house)(5332)는 공지된 바와 같이 주문의 이행을 전문으로 하는 임의의 제3자 물류 창고이다. 예시적인 이행 하우스는 Corporate Disk Company(McHenry, IL), OrderMotion, Inc.(New York, NY), 및 Shipwire.com(Los Angeles, CA)을 포함하지만, 이에 한정되는 것은 아니다.
- [0463] 케이블 TV 제공자(5334)는 Comcast Corporation(Philadelphia, PA) 및 Adelphia Communications(Greenwood Village, CO) 등의 임의의 케이블 TV 서비스 제공자이지만, 이에 한정되는 것은 아니다. 서비스 제공자(5336)는 임의의 종류의 서비스를 제공하는 임의의 개체를 나타낸다.
- [0464] 소프트웨어 제공자(5338)는 Art & Logic, Inc.(Pasadena, CA), Jigsaw Data Corp.(San Mateo, CA), DataMirror Corporation(New York, NY), 및 DataBank IMX, LCC(Beltsville, MD) 등의 임의의 소프트웨어 개발 회사이지만, 이에 한정되는 것은 아니다.
- [0465] 광고 회사(5340)는 D and B Marketing(Elhurst, IL), BlackSheep Marketing(Boston, MA), 및 Gotham Direct, Inc.(New York, NY) 등의 임의의 광고 회사 또는 대행사이지만, 이에 한정되지 않는다.
- [0466] 비즈니스 네트워크(5370)는 비즈니스 관계가 확립 및/또는 촉진되는 임의의 메커니즘을 나타낸다.
- [0467] 도 54는 본 발명의 일 실시예에 따른, MMR 시스템(100b)을 사용하여 용이하게 되는 일반화된 영업 방법인 방법(5400)을 나타낸 것이다. 방법(5400)은 적어도 2개의 개체 간의 관계를 확립하는 단계, 가능한 비즈니스 거래를 결정하는 단계, 적어도 하나의 비즈니스 거래를 실행하는 단계, 및 거래를 위해 제품 또는 서비스를 전달하는 단계를 포함한다.
- [0468] 먼저, 적어도 2개의 비즈니스 개체(5300) 사이에 관계가 확립된다(5410). 비즈니스 개체(5300)는, 예를 들어, 4개의 광역 카테고리, 즉 (1) MMR 생성자, (2) MMR 분배자, (3) MMR 사용자, 및 (4) 기타 내에 정렬될 수 있으며, 이들 내에서 어떤 비즈니스 개체는 2개 이상의 카테고리에 속한다. 이 예에 따르면, 비즈니스 개체(5300)는 다음과 같이 분류된다.
- [0469]
  - MMR 생성자 - MMR 서비스 제공자(5310), 멀티미디어 회사(5314), 문서 출판업자(5328), 문서 프린터(5330), 소프트웨어 제공자(5338) 및 광고 회사(5340)
- [0470]
  - MMR 분배자 - MMR 서비스 제공자(5310), 멀티미디어 회사(5314), 셀 전화 서비스 제공자(5318), 하드웨어

제조업자(5320), 하드웨어 소매업자(5322), 문서 출판업자(5328), 문서 프린터(5330), 이행 하우스(5332), 케이블 TV 제공자(5334), 서비스 제공자(5336) 및 광고 회사(5340)

[0471] • MMR 사용자 - MMR 소비자(5312), 프린터 사용자(5316) 및 문서 프린터(5330)

[0472] • 기타 - 금융 기관(5324) 및 신용 카드 처리기(5326)

[0473] 예를 들어, 이 방법 단계에서, MMR 생성자인 MMR 서비스 제공자(5310)와 MMR 사용자인 MMR 소비자(5312)와 MMR 분배자인 셀 전화 서비스 제공자(5318) 및 하드웨어 소매업자(5322) 사이에 비즈니스 관계가 확립된다. 게다가, 하드웨어 제조업자(5320)는 하드웨어 소매업자(5322)와 비즈니스 관계를 가지며, 이 둘다는 MMR 분배자이다.

[0474] 그 다음에, 방법(5400)은 단계(5410)에서 확립된 관계를 갖는 당사자들 간에 가능한 비즈니스 거래를 결정한다(5412). 상세하게는, 임의의 2개 이상의 비즈니스 개체(5300) 간에 다양한 거래가 있을 수 있다. 예시적인 거래는 정보를 구입하는 것, 물리적 상품을 구입하는 것, 서비스를 구입하는 것, 대역폭을 구입하는 것, 전자 저장소를 구입하는 것, 광고를 구입하는 것, 광고 통계를 구입하는 것, 상품을 발송하는 것, 정보를 판매하는 것, 물리적 상품을 판매하는 것, 서비스를 판매하는 것, 대역폭을 판매하는 것, 전자 저장소를 판매하는 것, 광고를 판매하는 것, 광고 통계를 판매하는 것, 임대/리스, 및 의견/평가/투표를 수집하는 것을 포함한다.

[0475] 방법(5400)이 당사자들 간의 가능한 비즈니스 거래를 결정하였으면, MMR 시스템(100)은 적어도 하나의 비즈니스 거래에 관한 합의에 도달(5414)하는 데 사용된다. 상세하게는, 임의의 2개 이상의 비즈니스 개체(5300) 사이에 거래의 결과인 다양한 동작이 있을 수 있다. 예시적인 동작은 정보를 구입하는 것, 주문을 받는 것, 더 많은 정보를 얻기 위해 클릭하는 것, 로컬/원격 액세스를 제공하는 것, 호스팅, 발송, 비즈니스 관계를 생성하는 것, 개인 정보를 저장하는 것, 다른 사람들에게 정보를 전달(pass-through)하는 것, 콘텐츠를 추가하는 것, 및 포드캐스트(podcasting)을 포함한다.

[0476] 방법(5400)이 비즈니스 거래에 관한 합의에 도달하였으면, MMR 시스템(100)은 거래를 위한 제품 또는 서비스를, 예를 들어, MMR 소비자(5312)로 전달(5416)하는 데 사용된다. 상세하게는, 방법 단계(5414)에서 합의된 비즈니스 거래의 결과로서, 다양한 콘텐츠가 임의의 2개 이상의 비즈니스 개체(5300) 사이에서 교환될 수 있다. 예시적인 콘텐츠는 텍스트, 웹 링크, 소프트웨어, 정지 영상 사진, 비디오, 오디오, 및 상기한 것의 임의의 조합을 포함한다. 그에 부가하여, 거래를 용이하게 하기 위해 임의의 2개 이상의 비즈니스 개체(5300) 사이에서 다양한 전달 메커니즘이 이용될 수 있다. 예시적인 전달 메커니즘은 종이, 퍼스널 컴퓨터, 네트워크화된 컴퓨터, 캡처 장치(106), 퍼스널 비디오 장치, 퍼스널 오디오 장치, 및 상기한 것의 임의의 조합을 포함한다.

[0477] 진술한 실시예에서 언급하고 청구범위에 청구하고 있는 바와 같은 발명 이외에도, 본 발명의 하나 이상의 실시예의 적어도 하나의 양태는, 혼합 미디어 문서를 제공하는 컴퓨터 구현 방법을 제공한다. 이 방법은 종이 문서로부터 추출된 피처의 전자적 설명을 (인덱스 테이블에서) 수신하는 단계를 포함한다. 인덱스 테이블은 종이 문서와 그 문서 내에서의 피처의 위치를 인쇄 미디어와 디지털 미디어를 조합한 혼합 미디어 문서와 연관시키기 위한 것이다. 이 방법은 계속해서 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 단계와, 상기 인덱스 테이블로부터의 데이터를 기초로 하여, 상기 질의어에 잠재적으로 응답하여 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계를 포함한다. 특정의 경우에, 상기 방법은 타겟 문서와 연관된 추가의 특징을 저장하는 단계를 더 포함한다. 그러한 경우에, 추가의 특징은 문서 정보(textual information)의 검색, 그래픽 정보의 검색, 프로세스의 실행, 명령의 실행, 주문, 비디오의 검색, 사운드의 검색, 정보의 저장, 새로운 문서의 생성, 문서의 인쇄 및/또는 문서의 디스플레이를 포함한 하나 이상의 동작을 포함한다. 다른 특정의 경우에, 종이 문서로부터 추출된 피처의 전자적 설명을 (인덱스 테이블에서) 수신하는 단계는 복수의 종이 문서로부터 추출된 피처의 전자적 설명을 수신하는 단계를 포함한다. 다른 특정의 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계는 인덱스 테이블로부터의 데이터를 기초로 순위가 매겨진 일련의 혼합 미디어 문서, 페이지 및 위치 가설을 계산하는 단계를 포함한다. 다른 특정의 경우에, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하는 상기 단계는, 타겟 문서로부터 추출된 한 세트의 수평 및 수직으로 인접한 단어 쌍을 수신하는 단계를 포함한다. 다른 특정의 경우에, 인덱스 테이블은 반전된 용어 인덱스 테이블(inverted term index table)을 포함하며, 여기서 반전된 용어 인덱스 테이블 내의 각각의 고유 용어는 레코드의 리스트를 가리키며, 각 레코드는 혼합 미디어 문서 내의 페이지 상의 후보 영역을 식별한다. 그러한 경우에, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 상기 단계는, 질의어에 대응하는 키에 의해 인덱싱된 모든 레코드를 검사하는 단계와, 모든 질의어와 가장 부합하는



영역을 식별하는 단계를 포함한다. 식별된 영역이 정합 기준을 만족시키는 정합 점수를 갖는 경우에, 상기 방법은 대응하는 혼합 미디어 문서 및 위치 가설을 확인하는 단계를 더 포함할 수 있다. 특정의 경우에 있어서, 타겟 문서는 종이 문서의 이미지 또는 종이 문서의 패치이다.

[0478] 본 발명의 하나 이상의 실시예의 적어도 하나의 다른 양태는, 하나 이상의 프로세서에 의해 실행될 때에, 프로세서로 하여금 혼합 미디어 문서 시스템을 제공하는 프로세스를 실행하도록 하게 하는 명령어가 인코딩된 기계 판독 가능한 매체(예컨대, 하나 이상의 콤팩트디스크, 디스켓, 서버, 메모리 스틱, 또는 하드 드라이브, ROMs, RAMs, 또는 전자 명령어를 저장하기에 적합한 기타 유형의 매체)를 제공한다. 이 프로세스는 예컨대 여기서 설명하는 방법 또는 그 방법의 변형예와 유사할 수 있다.

[0479] 본 발명의 하나 이상의 실시예의 적어도 하나의 다른 양태는, 혼합 미디어 문서를 제공하는 데이터베이스 시스템을 제공한다. 이 데이터베이스 시스템은, 종이 문서로부터 추출된 피처의 전자적 설명을 수신하고, 종이 문서와 그 문서 내에서의 피처의 위치를 인쇄 미디어와 디지털 미디어를 조합한 혼합 미디어 문서와 연관시키는 인덱스 테이블을 포함한다. 이 시스템은, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 하나 이상의 질의어를 수신하고, 상기 인덱스 테이블로부터의 데이터를 기초로 하여, 질의어에 잠재적으로 응답하여 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 누적기 모듈을 더 포함한다. 특정의 경우에, 상기 시스템은 타겟 문서와 연관된 추가의 특징을 저장하는 저장 설비(예컨대, 관련 데이터베이스)를 포함한다. 그러한 경우에, 추가의 특징은 문서 정보의 검색, 그래픽 정보의 검색, 프로세스의 실행, 명령의 실행, 주문, 비디오의 검색, 사운드의 검색, 정보의 저장, 새로운 문서의 생성, 문서의 인쇄 및/또는 문서의 디스플레이를 포함한 하나 이상의 동작을 포함한다. 다른 특정의 경우에, 인덱스 테이블은, 복수의 종이 문서로부터 추출된 피처의 전자적 설명을 수신할 수 있다. 다른 특정의 경우에, 종이 문서는 다수의 페이지를 포함하고, 인덱스 테이블은 이들 페이지 내의 혼합 미디어, 페이지 및 x-y 피처 위치를 식별하도록 추가로 구성된다. 다른 특정의 경우에, 누적기 모듈에 의해 실시되는, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계는, 인덱스 테이블로부터의 데이터를 기초로 순위가 매겨진 일련의 혼합 미디어 문서, 페이지 및 위치 가설을 계산하는 단계를 포함한다. 다른 특정의 경우에, 누적기 모듈에 의해 실시되는, 타겟 문서 내의 객체들 사이의 2차원 관계를 캡처하는 질의어를 수신하는 단계는, 타겟 문서로부터 추출된 한 세트의 수평 및 수직으로 인접한 단어 쌍을 수신하는 단계를 포함한다. 다른 특정의 경우에, 인덱스 테이블은 반전된 용어 인덱스 테이블을 포함하며, 여기서 반전된 용어 인덱스 테이블 내의 각각의 고유 용어는 레코드의 리스트를 가리키며, 각 레코드는 혼합 미디어 문서 내의 페이지 상의 후보 영역을 식별한다. 그러한 경우에, 누적기 모듈에 의해 실시되는, 적어도 하나의 혼합 미디어 문서 및 위치 가설을 계산하는 단계는, 질의어에 대응하는 키에 의해 인덱싱된 모든 레코드를 검사하는 단계와, 모든 질의어와 가장 부합하는 영역을 식별하는 단계를 포함한다. 식별된 영역이 정합 기준을 만족시키는 정합 점수를 갖는 경우에, 상기 누적기 모듈은, 대응하는 혼합 미디어 문서 및 위치 가설을 확인할 수 있다. 다른 경우에 있어서, 인덱스 테이블은 각각의 혼합 미디어 문서에 대한 관련 정보를 포함하는 문서 인덱스 테이블을 더 포함하며, 상기 정보는 인쇄 해상도, 인쇄 일자, 용지 크기, 쉼도우 파일 이름, 페이지 이미지 위치 중 적어도 하나를 포함한다. 다른 특정의 경우에, 인덱스 테이블이 수신한 설명은, 추출된 피처와 이들 피처의 내부 문서(inner-document) 위치 데이터를 연관시키는 피처 추출 모듈에 의해 계산된다. 특정의 경우에 있어서, 타겟 문서는 종이 문서의 이미지 또는 그 종이 문서의 패치이다. 시스템의 기능은, 소프트웨어(예컨대, 하나 이상의 컴퓨터 판독 가능한 매체에 인코딩된 실행 명령어), 하드웨어(예컨대, 게이트 레벨 로직 또는 하나 이상의 ASICs), 펌웨어(예컨대, 본 명세서에서 설명하는 기능을 실시하는 임베딩된 루틴과 I/O 능력을 갖는 하나 이상의 마이크로컨트롤러) 또는 이들의 조합을 비롯한 많은 수단에 의해 구현될 수 있다. 데이터베이스 시스템은 본 명세서에서 설명하고 있는 바와 같이 MMR(Mixed Media Reality) 시스템에서 구현될 수 있으며, 그 기능은 예컨대 하나 이상의 서버, 컴퓨터 시스템, 휴대용 장치 또는 이들을 일부 조합한 설비에서 달성된다.

[0480] 특정의 일 실시예에 따르면, MMR 시스템은 텍스트-기반 인덱스를 사용하여 탐색(look-up)을 가능하게 해주는 방식으로 인쇄된 문서로부터 추출된 객체들 사이의 2차원의 기하학적 관계를 표현하는 인덱스 테이블을 구비하도록 구성된 콘텐츠-기반 검색 데이터베이스를 포함한다. 순위가 매겨진 일련의 문서, 페이지 및 위치 가설은, 인덱스 테이블로부터 데이터가 제공될 경우에 계산될 수 있다. 이 기술은 이미지 패치에서 검출된 피처를, 피처 자체 및 이들 사이의 기하학적 관계 모두를 나타내는 문자 용어(또는 다른 검색 가능한 피처)로 효율적으로 변환시킨다. 저장 장비를 이용하여 각각의 문서 이미지 패치에 대한 추가의 특징을 저장할 수 있다.

[0481] 본 명세서에 제공된 알고리즘들은 본질적으로 임의의 특정의 컴퓨터 또는 다른 장치에 관련되어 있지 않다. 다양한 범용 및/또는 특수 목적의 시스템이 본 발명의 실시예들에 따라 프로그램되거나 다른 방식으로 구성될 수 있다. 이 개시 내용으로부터 명백하게 될 것인 바와 같이, 수많은 프로그래밍 언어 및/또는 구조가 다양한 이

러한 시스템을 구현하는 데 사용될 수 있다. 게다가, 본 발명의 실시예들은 정보 시스템 또는 네트워크 상에서 동작하거나 그와 관련하여 작동할 수 있다. 예를 들어, 본 발명은 구성에 따라 변하는 기능을 갖는 독립형 다 기능 프린터 또는 네트워크화된 프린터 상에서 동작할 수 있다. 본 발명은 최소 기능을 갖는 것부터 본 명세서에 개시된 기능 전부를 제공하는 것까지 임의의 정보 시스템에서 동작할 수 있다.

[0482] 본 발명의 실시예들에 대한 이상의 설명은 예시 및 설명을 위해 제공된 것이다. 이는 전수적인 것이 아니며 본 발명을 개시된 정확한 형태로 제한하려는 것이 아니다. 이상의 개시 내용을 바탕으로 많은 수정 및 변형이 가능하다. 본 발명의 범위는 이 상세한 설명이 아니라 본 출원의 청구항들에 의해 제한되는 것으로 보아야 한다. 당업자라면 잘 알 것인 바와 같이, 본 발명이 본 발명의 정신 또는 필수적인 특징을 벗어나지 않고 다른 구체적인 형태로 구현될 수 있다. 이와 마찬가지로, 모듈, 루틴, 특징, 애플리케이션, 방법 및 기타 측면들의 특징의 명명 및 구분이 필수적이거나 중요한 것이 아니며, 본 발명 또는 그의 특징들을 구현하는 메커니즘들이 서로 다른 이름, 구분 및/또는 형식을 가질 수 있다. 게다가, 당업자에게는 명백하게 될 것인 바와 같이, 본 발명의 모듈, 루틴, 피쳐, 애플리케이션, 방법 및 다른 특징은 소프트웨어, 하드웨어, 펌웨어 또는 이들의 임의의 조합으로서 구현될 수 있다. 또한, 본 발명의 컴포넌트(이의 일례가 모듈임)가 소프트웨어로서 구현되는 경우, 이 컴포넌트는 독립형 프로그램으로서, 더 큰 프로그램의 일부로서, 복수의 개별적인 프로그램으로서, 정적으로 또는 동적으로 링크된 라이브러리로서, 커널 로드가능한 모듈로서, 장치 드라이버로서, 및/또는 현재 또는 장래에 컴퓨터 프로그래밍 분야의 당업자라면 알고 있는 모든 다른 방법으로 구현될 수 있다. 그에 부가하여, 본 발명은 임의의 특징의 프로그래밍 언어로 된 또는 임의의 특징의 운영 체제 또는 환경에 대한 구현으로 결코 제한되지 않는다. 그에 따라, 본 발명의 개시 내용은 예시적인 것이며 이하의 청구항들에 기술되어 있는 본 발명의 범위를 제한하는 것이 아니다.

[0483] 본 발명은 미국 우선권 출원들, 2005년 8월 23일자로 출원된 제60/710,767호, 2006년 4월 17일자로 출원된 제60/792,912호, 2006년 7월 18일자로 출원된 제60/807,654호, 2006년 7월 31일에 출원된 제11/461,147호, 2006년 7월 31일자로 출원된 제11/461,164호에 기초하며, 그 전체 내용은 본원 명세서에 참고로 인용된다.

### 도면의 간단한 설명

- [0013] 도 1a는 본 발명의 일 실시예에 따라 구성된 MMR(Mixed Media Reality) 시스템의 기능 블록도.
- [0014] 도 1b는 본 발명의 다른 일 실시예에 따라 구성된 MMR 시스템의 기능 블록도.
- [0015] 도 2a, 도 2b 도 2c 및 도 2d는 본 발명의 일 실시예에 따른 캡처 장치를 나타낸 도면.
- [0016] 도 2e는 본 발명의 일 실시예에 따라 구성된 캡처 장치의 기능 블록도.
- [0017] 도 3은 본 발명의 일 실시예에 따라 구성된 MMR 컴퓨터의 기능 블록도.
- [0018] 도 4는 본 발명의 일 실시예에 따라 구성된 MMR 소프트웨어 모음(software suite)에 포함된 일련의 소프트웨어 컴포넌트를 나타낸 도면.
- [0019] 도 5는 본 발명의 일 실시예에 따라 구성된 MMR 문서의 일 실시예를 나타낸 도면.
- [0020] 도 6은 본 발명의 일 실시예에 따른 문서 지문 정합 방법(document fingerprint matching methodology)을 나타낸 도면.
- [0021] 도 7은 본 발명의 일 실시예에 따라 구성된 문서 지문 정합 시스템을 나타낸 도면.
- [0022] 도 8은 본 발명의 일 실시예에 따른 텍스트/비텍스트 구분(text/non-text discrimination)을 위한 흐름 프로세스를 나타낸 도면.
- [0023] 도 9는 본 발명의 일 실시예에 따른 텍스트/비텍스트 구분의 일례를 나타낸 도면.
- [0024] 도 10은 본 발명의 일 실시예에 따른 이미지 패치(image patch) 내의 텍스트의 포인트 크기(point size)를 추정하는 흐름 프로세스를 나타낸 도면.
- [0025] 도 11은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0026] 도 12는 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0027] 도 13은 본 발명의 일 실시예에 따른 상호작용적 이미지 분석의 일례를 나타낸 도면.

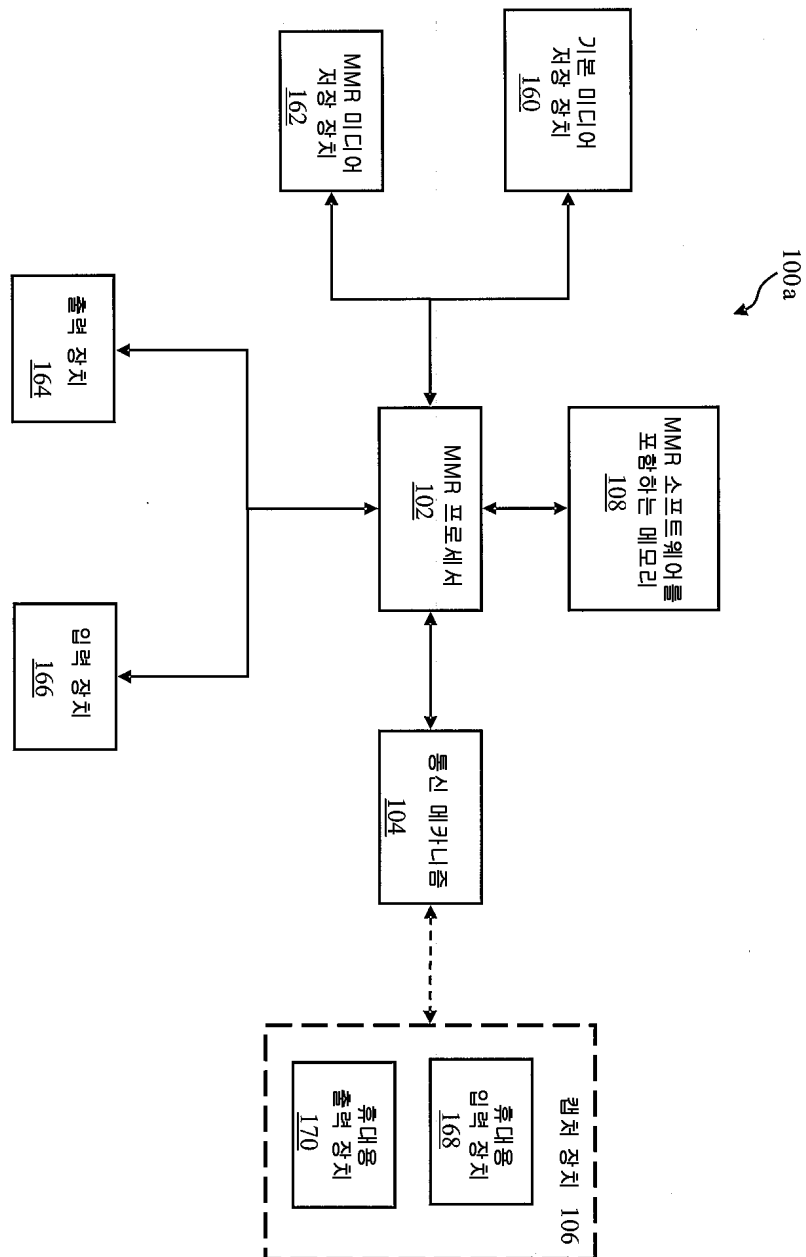
- [0028] 도 14는 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0029] 도 15는 본 발명의 일 실시예에 따른 단어 경계 박스 검출(word bounding box detection)의 일례를 나타낸 도면.
- [0030] 도 16은 본 발명의 일 실시예에 따른 피쳐 추출 기법을 나타낸 도면.
- [0031] 도 17은 본 발명의 다른 실시예에 따른 피쳐 추출 기법을 나타낸 도면.
- [0032] 도 18은 본 발명의 다른 실시예에 따른 피쳐 추출 기법을 나타낸 도면.
- [0033] 도 19는 본 발명의 다른 실시예에 따른 피쳐 추출 기법을 나타낸 도면.
- [0034] 도 20은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0035] 도 21은 본 발명의 일 실시예에 따른 문서 지문 정합을 위한 다중-분류자 피쳐 추출(multi-classifier feature extraction)을 나타낸 도면.
- [0036] 도 22 및 도 23은 본 발명의 일 실시예에 따른 문서 지문 정합 기법의 일례를 나타낸 도면.
- [0037] 도 24는 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0038] 도 25는 본 발명의 일 실시예에 따른 데이터베이스-위주 피드백(database-driven feedback)을 위한 흐름 프로세스를 나타낸 도면.
- [0039] 도 26은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0040] 도 27은 본 발명의 일 실시예에 따른 데이터베이스-위주 분류(database-driven classification)를 위한 흐름 프로세스를 나타낸 도면.
- [0041] 도 28은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0042] 도 29는 본 발명의 일 실시예에 따른 데이터베이스-위주 다중 분류(database-driven multiple classification)를 위한 흐름 프로세스를 나타낸 도면.
- [0043] 도 30은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0044] 도 31은 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0045] 도 32는 본 발명의 다른 실시예에 따른 문서 지문 정합 기법을 나타낸 도면.
- [0046] 도 33은 본 발명의 일 실시예에 따른 다중-계층 인식(multi-tier recognition)을 위한 흐름 프로세스를 나타낸 도면.
- [0047] 도 34a는 본 발명의 일 실시예에 따라 구성된 MMR 데이터베이스 시스템의 기능 블록도.
- [0048] 도 34b는 본 발명의 일 실시예에 따른 OCR-기반 기법에 대한 MMR 피쳐 추출의 일례를 나타낸 도면.
- [0049] 도 34c는 본 발명의 일 실시예에 따른 예시적인 인덱스 테이블 구성(index table organization)을 나타낸 도면.
- [0050] 도 35는 본 발명의 일 실시예에 따른 MMR 인덱스 테이블을 발생하는 방법을 나타낸 도면.
- [0051] 도 36은 본 발명의 일 실시예에 따른, 대상 문서(target document)에 대한 순위가 매겨진 일련의 문서, 페이지, 및 위치 가설들(document, page and location hypotheses)을 계산하는 방법을 나타낸 도면.
- [0052] 도 37a는 본 발명의 다른 실시예에 따라 구성된 MMR 컴포넌트의 기능 블록도.
- [0053] 도 37b는 본 발명의 일 실시예에 따른 MMR 인쇄 소프트웨어에 포함되어 있는 일련의 소프트웨어 컴포넌트를 나타낸 도면.
- [0054] 도 38은 본 발명의 일 실시예에 따른, 문서에 핫스팟(hot spot)을 임베딩(embed)하는 방법의 플로우차트.
- [0055] 도 39a는 본 발명의 일 실시예에 따른 HTML 파일의 일례를 나타낸 도면.
- [0056] 도 39b는 도 39a의 HTML 파일의 마크-업 버전의 일례를 나타낸 도면.
- [0057] 도 40a는 본 발명의 일 실시예에 따른 브라우저에 디스플레이되는 도 39a의 HTML 파일의 일례를 나타낸 도면.

- [0058] 도 40b는 본 발명의 일 실시예에 따른, 도 40a의 HTML 파일의 인쇄 버전의 일례를 나타낸 도면.
- [0059] 도 41은 본 발명의 일 실시예에 따른 기호적인 핫스팟 서술(symbolic hotspot description)을 나타낸 도면.
- [0060] 도 42a 및 도 42b는 본 발명의 일 실시예에 따른, 도 39a의 HTML 파일에 대한 예시적인 page\_desc.xml 파일을 나타낸 도면.
- [0061] 도 43은 본 발명의 일 실시예에 따른, 도 41, 도 42a 및 도 42b에 대응하는 hotspot.xml 파일을 나타낸 도면.
- [0062] 도 44는 본 발명의 일 실시예에 따른 포워딩 DLL(forwarding DLL)에 의해 사용되는 프로세스의 플로우차트.
- [0063] 도 45는 본 발명의 일 실시예에 따른 문서에서의 핫스팟에 대응하는 문자들을 변환하는 방법의 플로우차트.
- [0064] 도 46은 본 발명의 일 실시예에 따른 문서의 전자 버전의 일례를 나타낸 도면.
- [0065] 도 47은 본 발명의 일 실시예에 따른 인쇄된 수정된 문서의 일례를 나타낸 도면.
- [0066] 도 48은 본 발명의 일 실시예에 따른 공유 문서 주석 첨부(shared document annotation) 방법의 플로우차트.
- [0067] 도 49a는 본 발명의 일 실시예에 따른 브라우저에서의 소스 웹 페이지의 샘플을 나타낸 도면.
- [0068] 도 49b는 본 발명의 일 실시예에 따른 브라우저에서의 수정된 웹 페이지의 샘플을 나타낸 도면.
- [0069] 도 49c는 본 발명의 일 실시예에 따른 인쇄된 웹 페이지의 샘플을 나타낸 도면.
- [0070] 도 50a는 본 발명의 일 실시예에 따른 이미지화된 문서에 핫스팟을 추가하는 방법의 플로우차트.
- [0071] 도 50b는 본 발명의 일 실시예에 따른 이미지화된 문서에 추가하기 위한 핫스팟을 정의하는 방법의 플로우차트.
- [0072] 도 51a는 일 실시예에 따라 스캔된 신문 페이지의 일부분을 보여주는 사용자 인터페이스의 일례를 나타낸 도면.
- [0073] 도 51b는 선택된 핫스팟과 연관시킬 데이터 또는 상호작용을 정의하는 사용자 인터페이스를 나타낸 도면.
- [0074] 도 51c는 본 발명의 일 실시예에 따른 할당 박스(assign box)를 포함하는 도 51b의 사용자 인터페이스를 나타낸 도면.
- [0075] 도 51d는 본 발명의 일 실시예에 따른 문서 내에 핫스팟을 디스플레이하는 사용자 인터페이스를 나타낸 도면.
- [0076] 도 52는 본 발명의 일 실시예에 따른 MMR 문서 및 MMR 시스템을 사용하는 방법의 플로우차트.
- [0077] 도 53은 본 발명의 일 실시예에 따른, MMR 시스템과 연관된 예시적인 일련의 비즈니스 개체의 블록도.
- [0078] 도 54는 본 발명의 일 실시예에 따른, MMR 시스템의 사용에 의해 용이하게 되는 일반화된 영업 방법인 방법의 플로우차트.

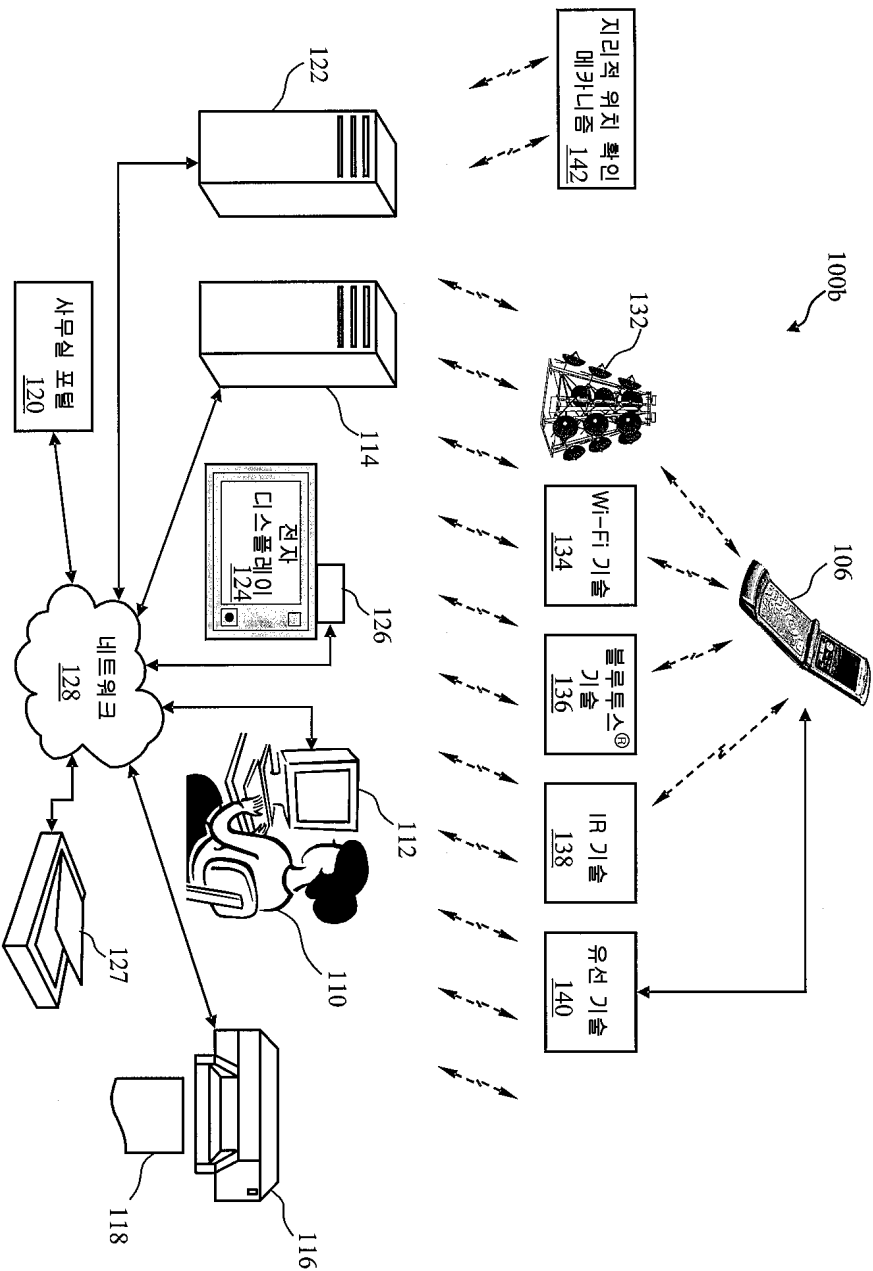


도면

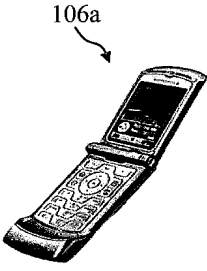
도면1a



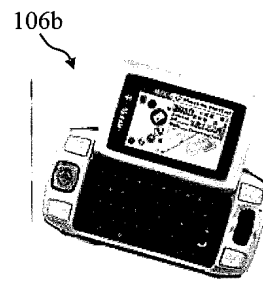
도면1b



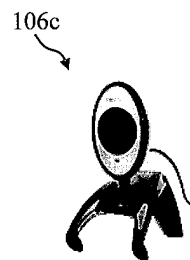
도면2a



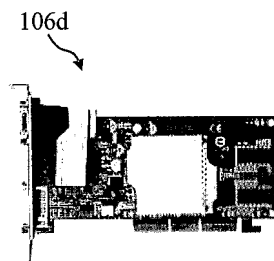
도면2b



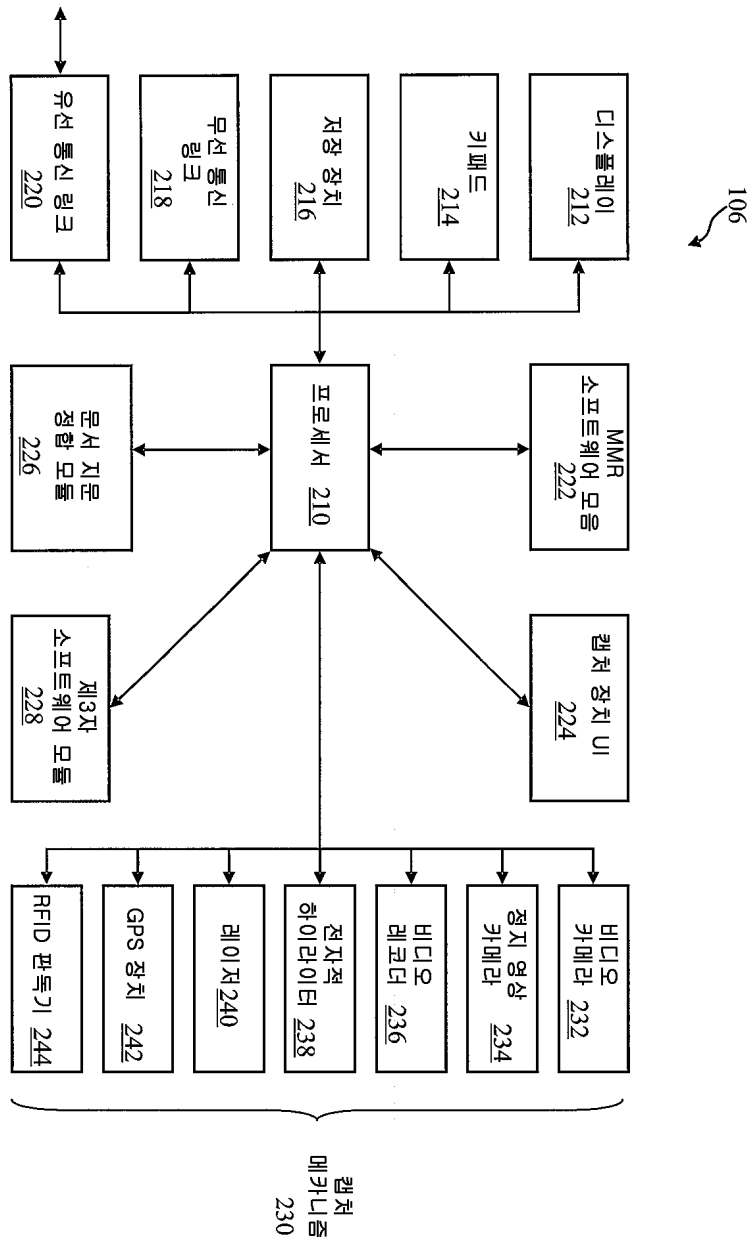
도면2c



도면2d

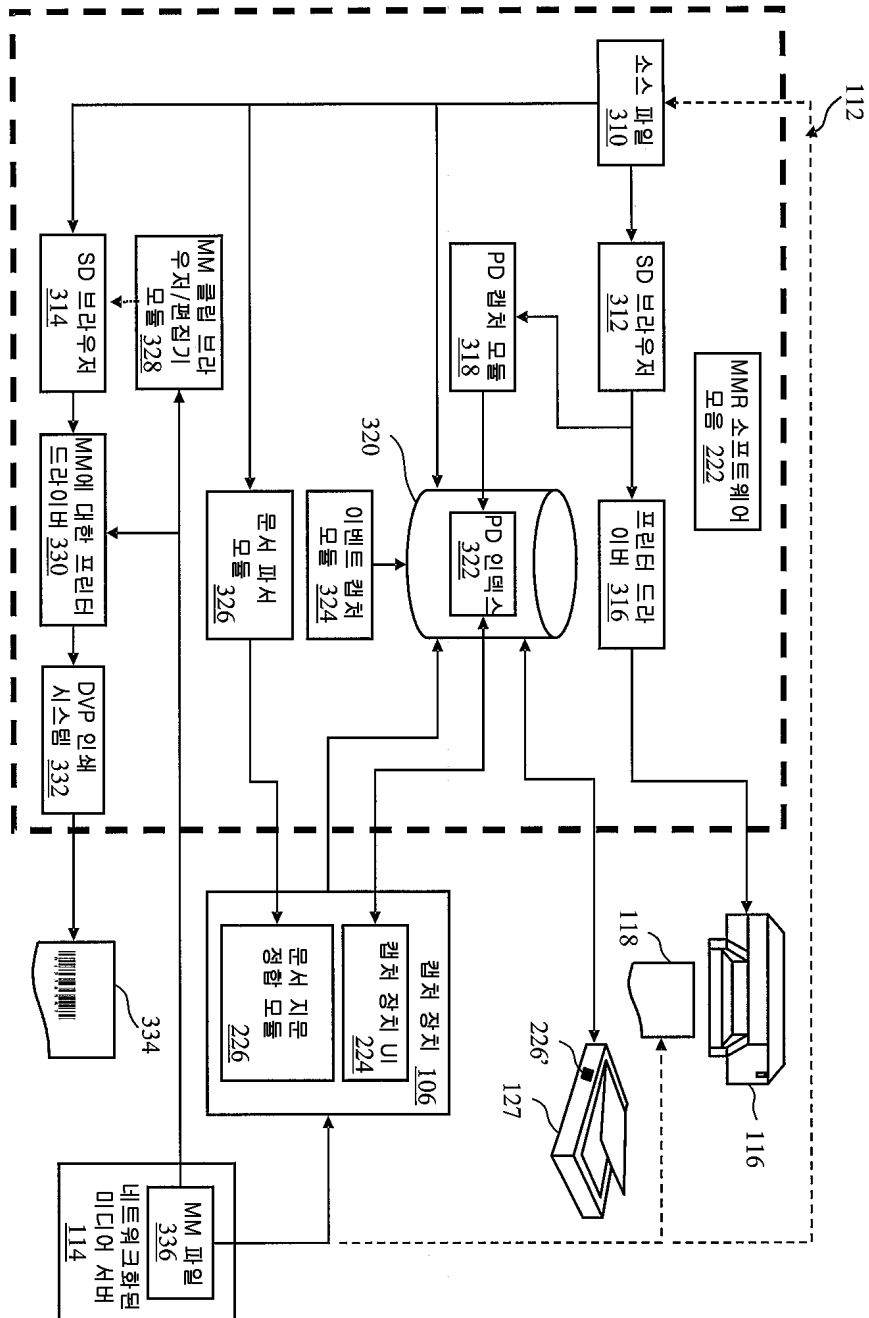


도면2e





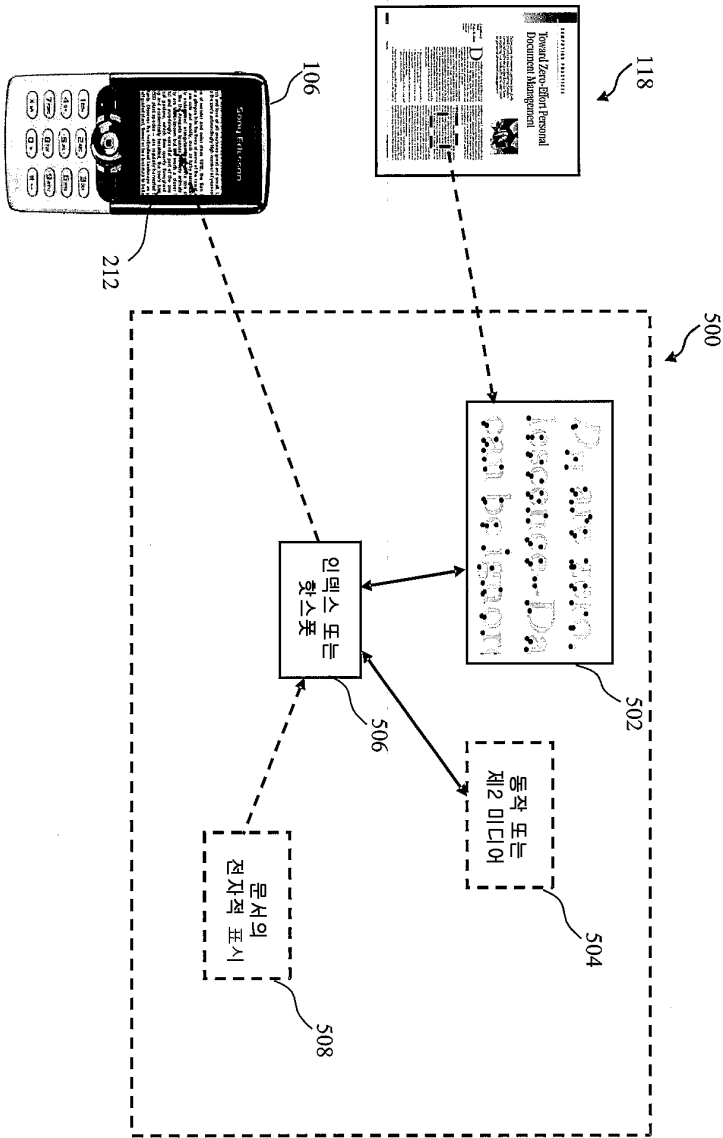
도면3



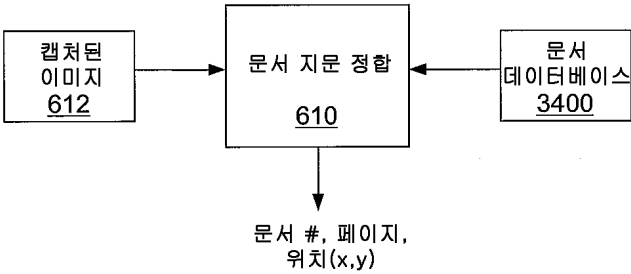
멀티미디어 주석 첨부 소프트웨어 410			
텍스트 쿼리존-기반 검색 컴포넌트 412	실시간 통지 컴포넌트 424	시간 수송 컴포넌트 436	문서 버전 검색 컴포넌트 448
이미지 쿼리존-기반 검색 컴포넌트 414	멀티미디어 검색 컴포넌트 426	위치 인식 컴포넌트 438	PC 문서 메타데이터 컴포넌트 450
스태가노그래픽 수정 컴포넌트 416	바탕화면 비디오 리마인더 컴포넌트 428	PC 저작 컴포넌트 440	웹처 장치 UI 컴포넌트 452
종이 판독 이력 로그 418	웹 페이지 리마인더 컴포넌트 430	문서 저작 컴포넌트 442	도메인-관련 컴포넌트 454
온라인 판독 이력 로그 420	물리적 이력 로그 432	웹처 장치 저작 컴포넌트 444	
협동적 문서 검토 컴포넌트 422	완성된 양식 검토자 컴포넌트 434	무의식적 업로드 컴포넌트 446	

222 ↗

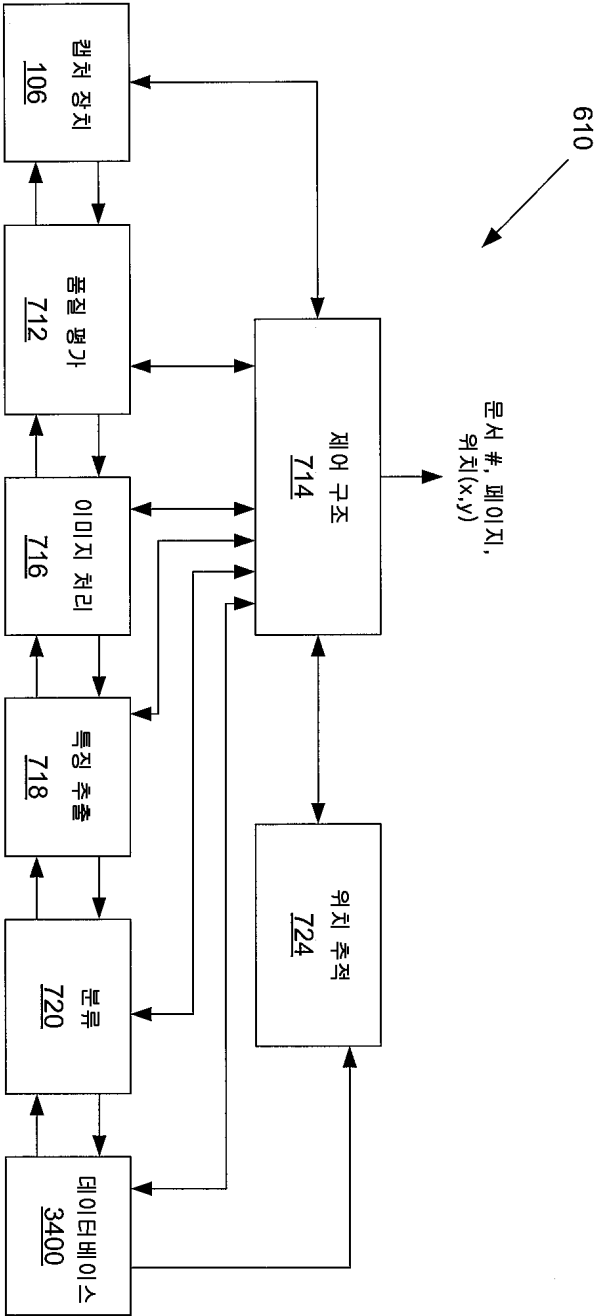
도면5



도면6

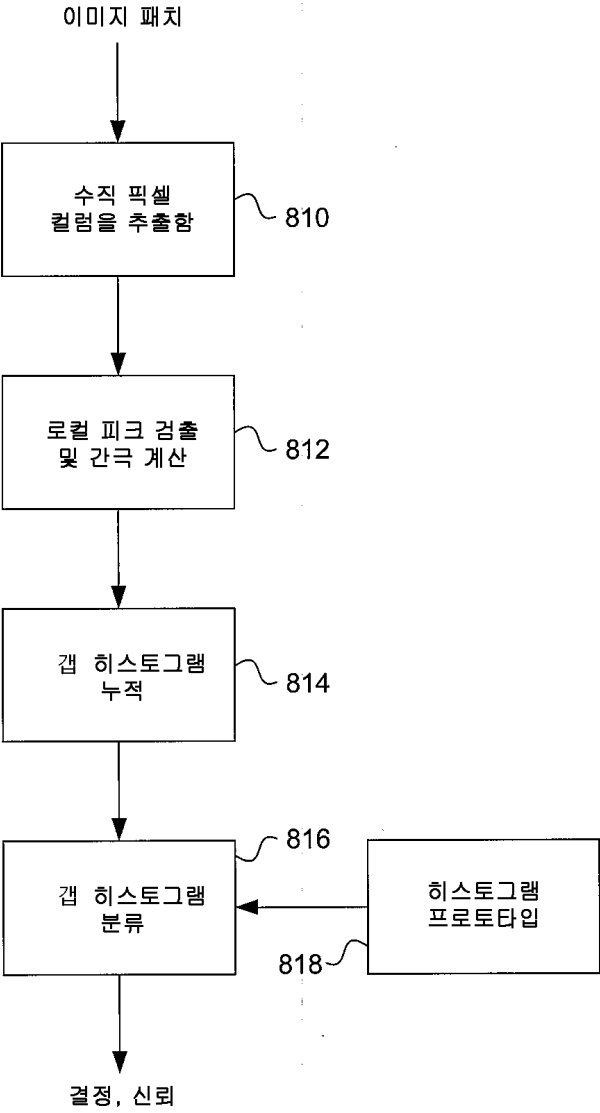


도면7

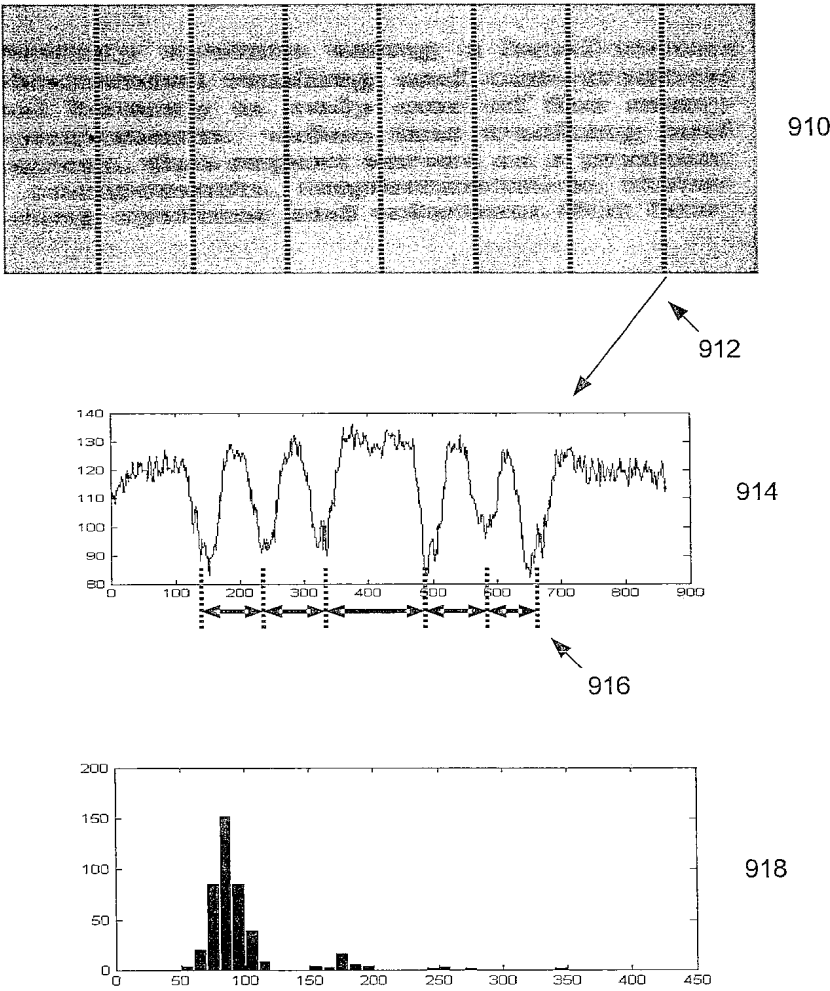




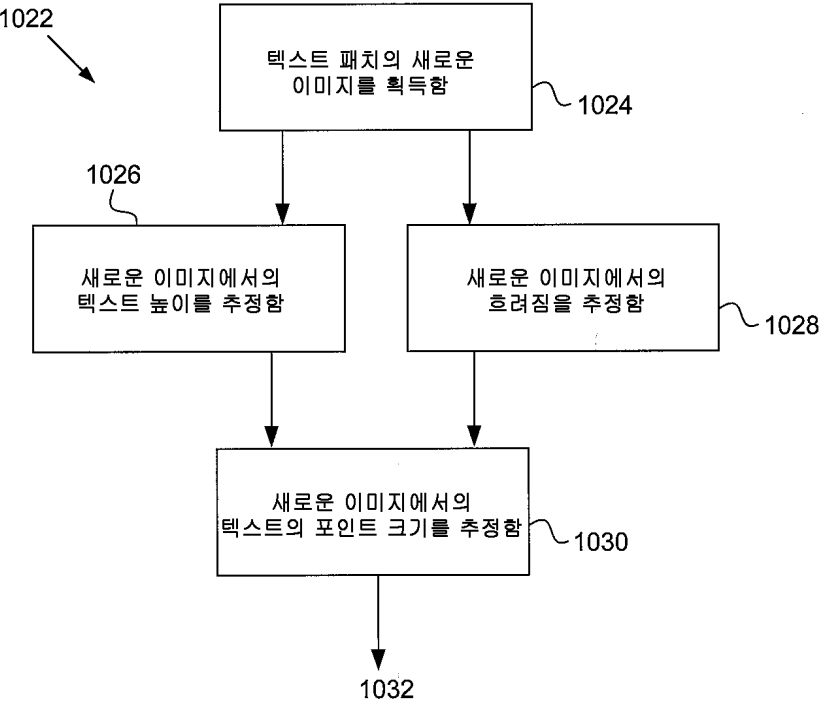
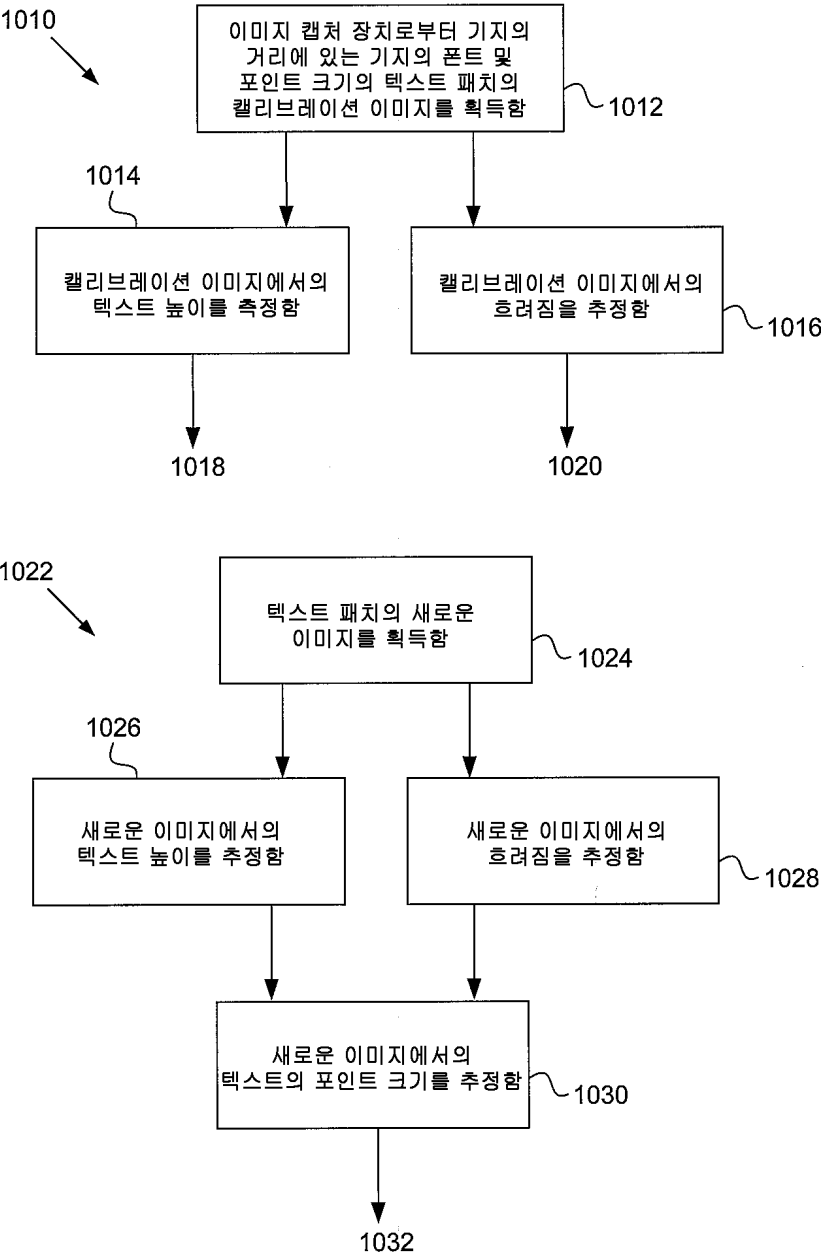
도면8



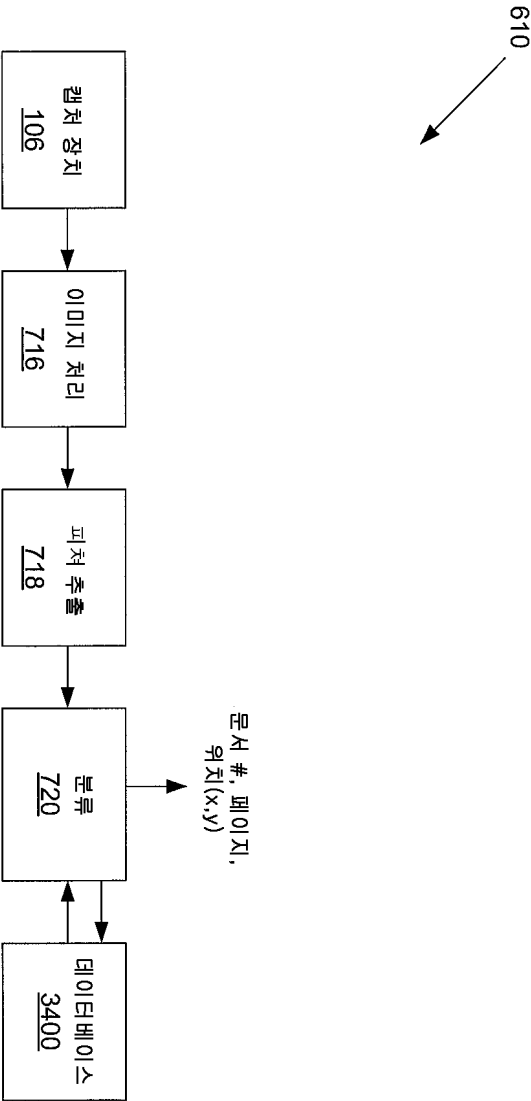
도면9



도면10

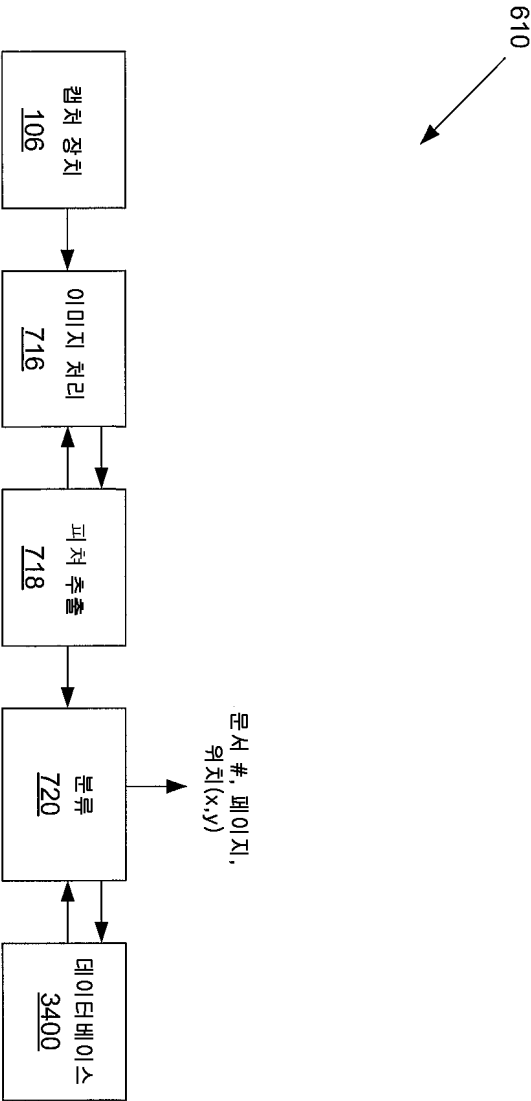


도면11

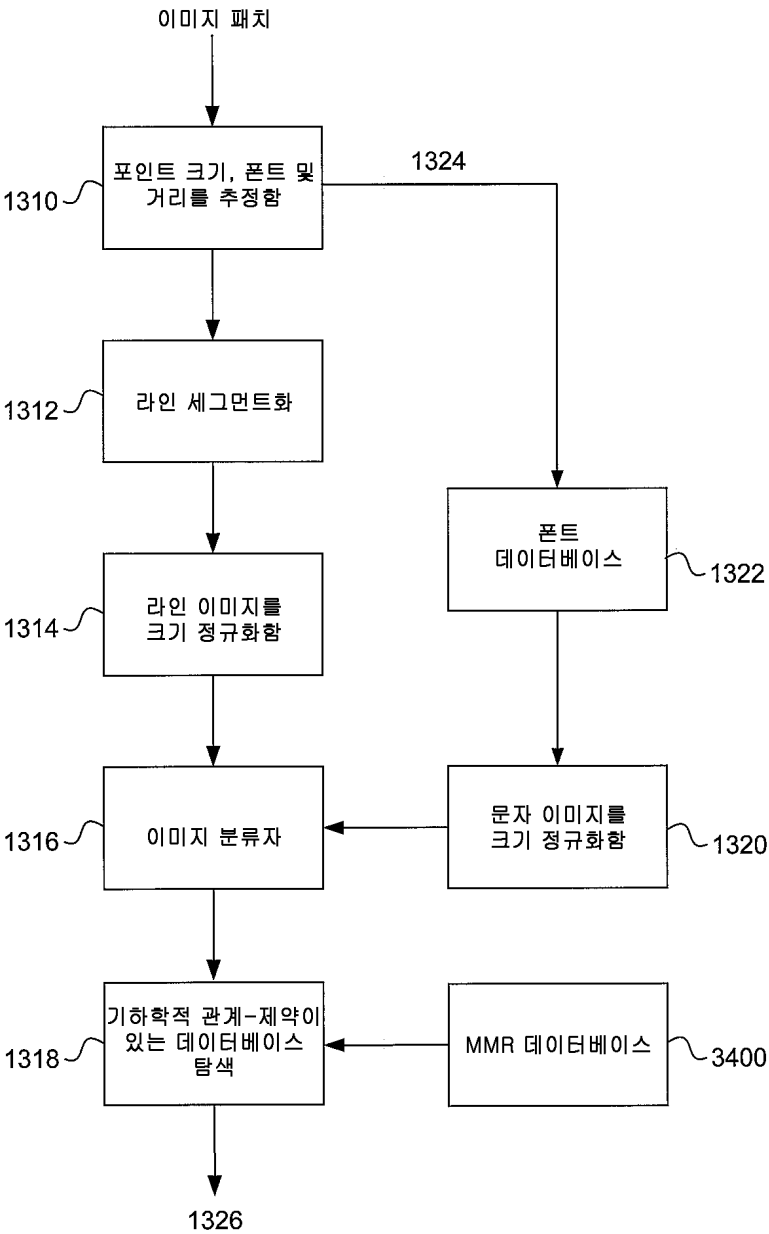




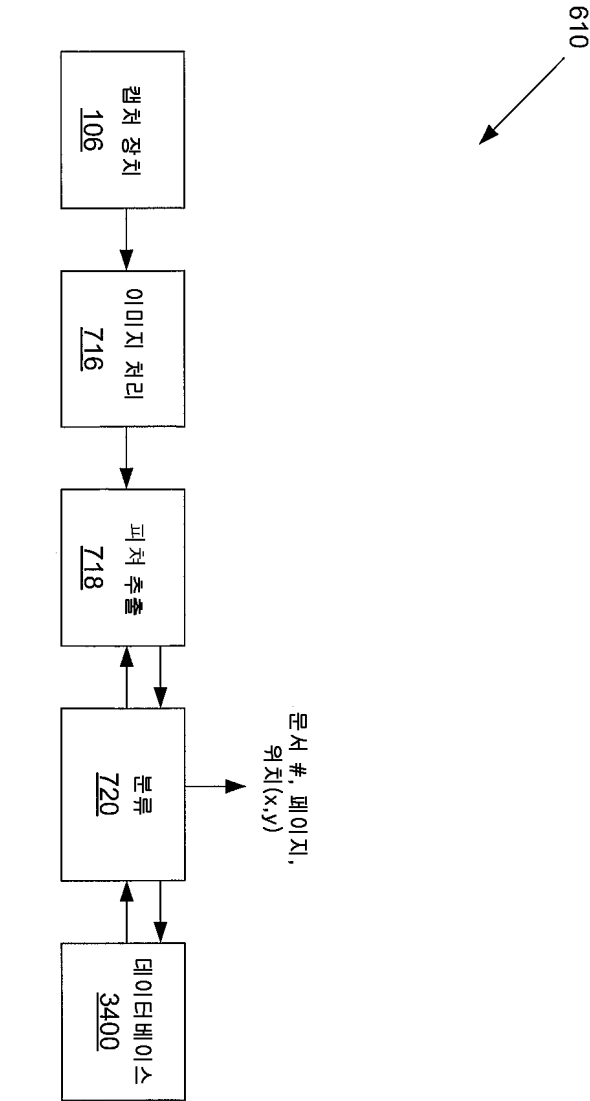
도면12



도면13

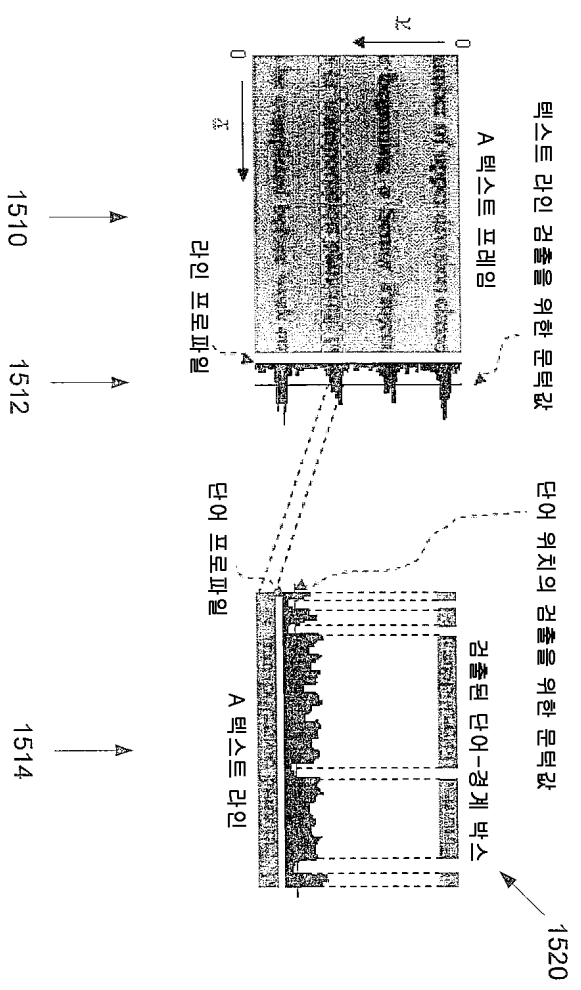


도면14



1516

1518

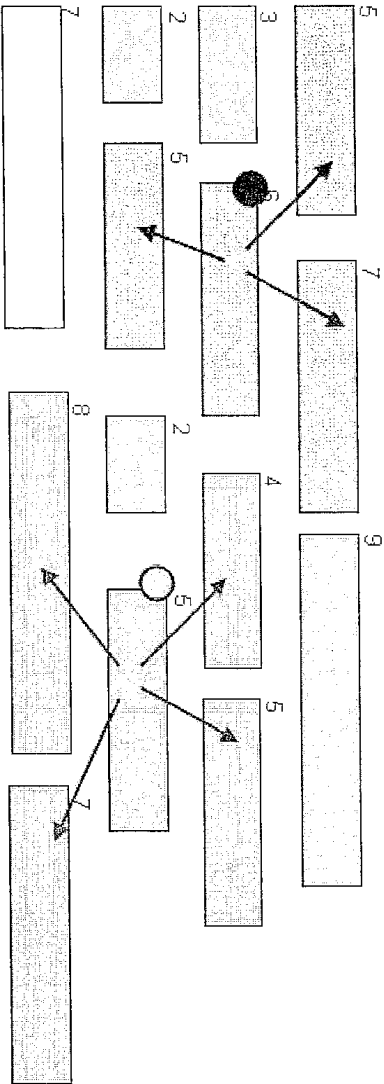


도면15

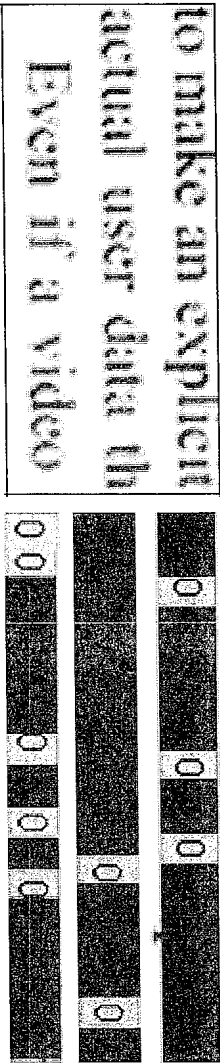


도면16

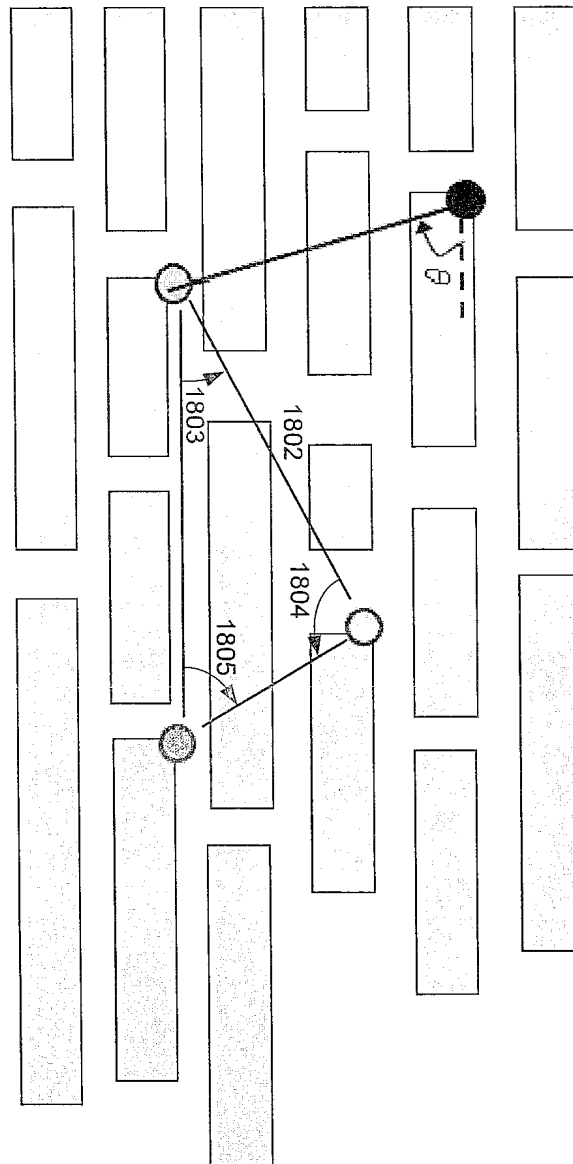
- 피쳐 포인트 : 6 57 5
- 피쳐 포인트 : 5 45 87



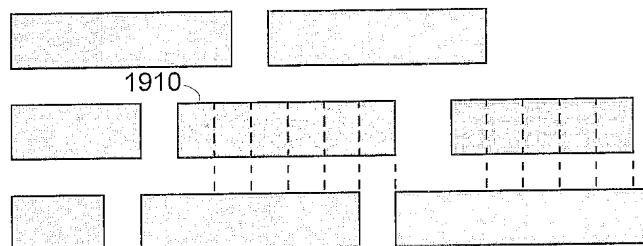
도면17



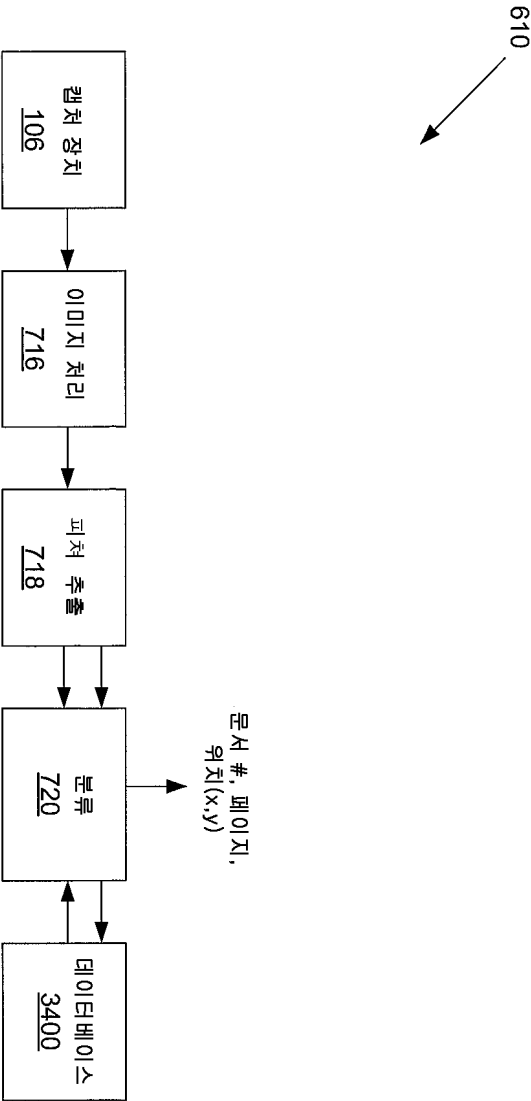
도면18



도면19

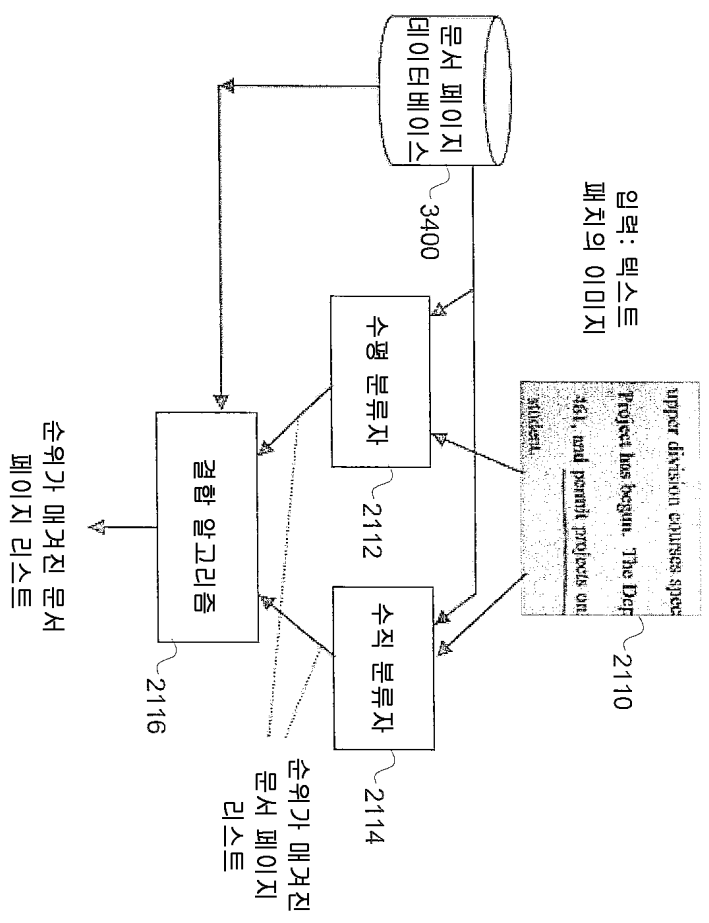


도면20

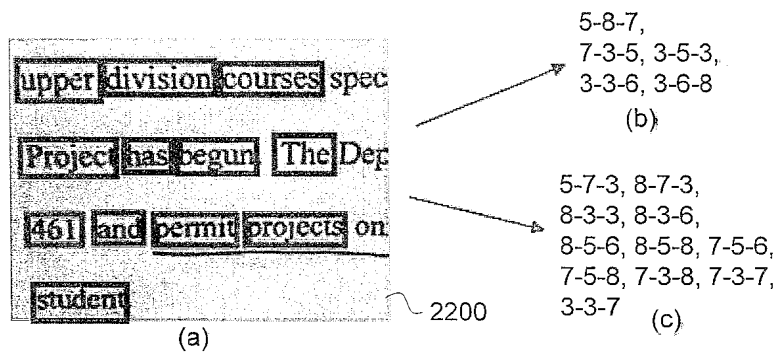




도면21



도면22



5-7-8 => 12, 15, 33  
7-3-5 => 15, 22, 134  
3-5-3 => 4, 6, 15, 56  
3-3-6 => 6, 9, 15  
3-6-8 => 12, 15, 134

(d)

5-7-3 => 15, 18  
8-7-3 => 2, 15  
8-3-3 => 2, 4, 15  
8-3-6 => 4, 6, 15  
8-5-6 => 12, 15  
8-5-8 => 15, 17  
7-5-6 => 15, 17  
7-5-8 => 4, 15  
7-3-8 => 2, 15  
7-3-7 => 7, 10, 15  
3-3-7 => 2, 15

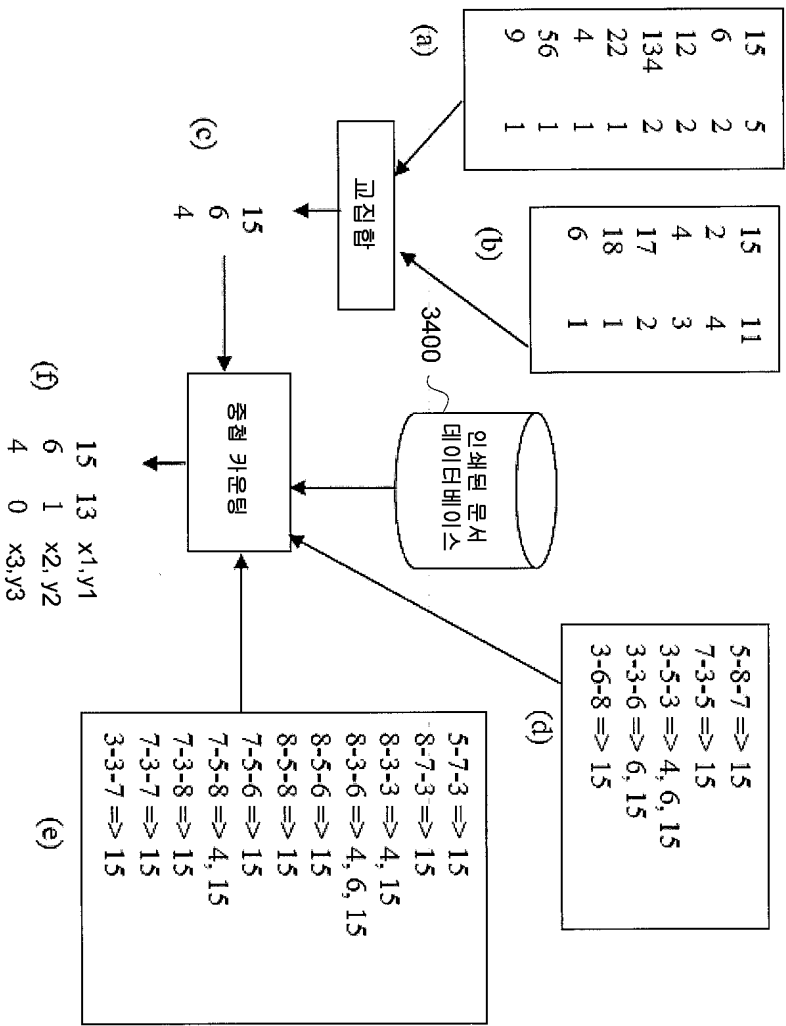
(e)

15	5
6	2
12	2
134	2
22	1
4	1
56	1
9	1

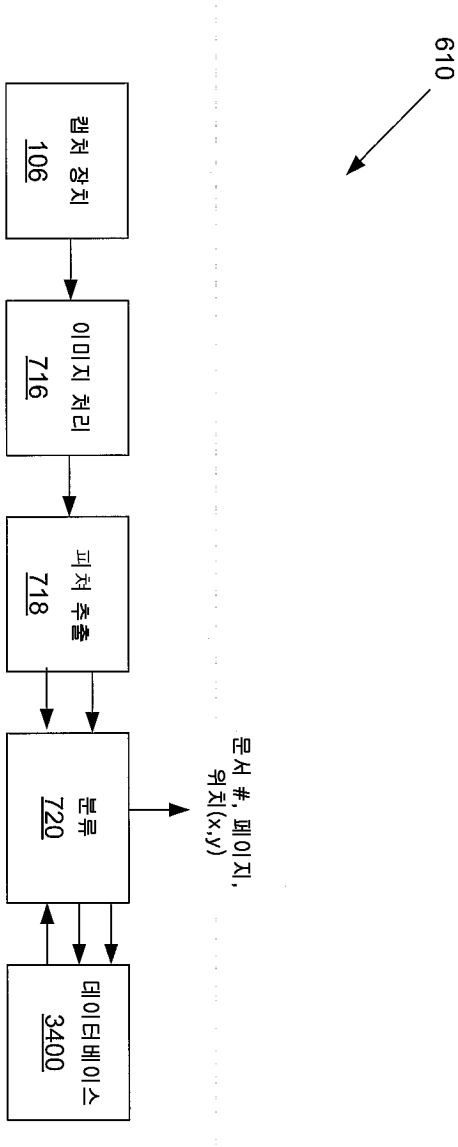
(f)

15	11
2	4
4	3
17	2
18	1
6	1

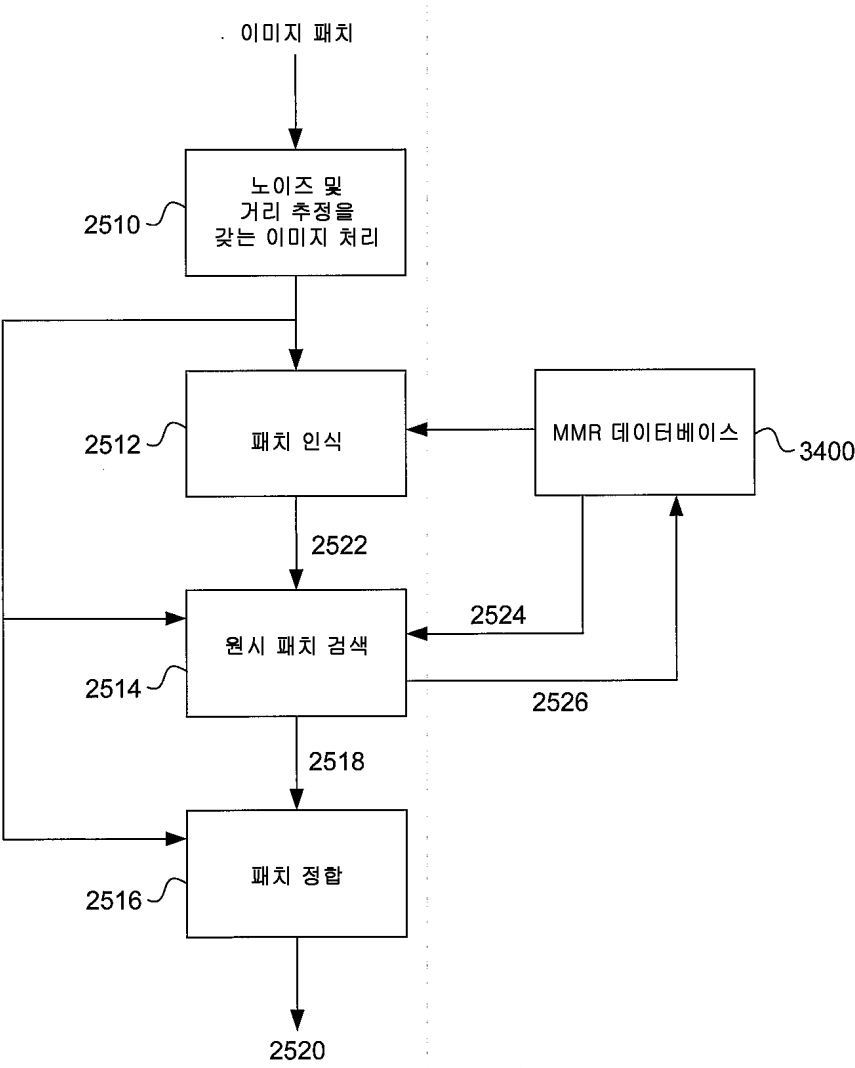
(g)



도면24

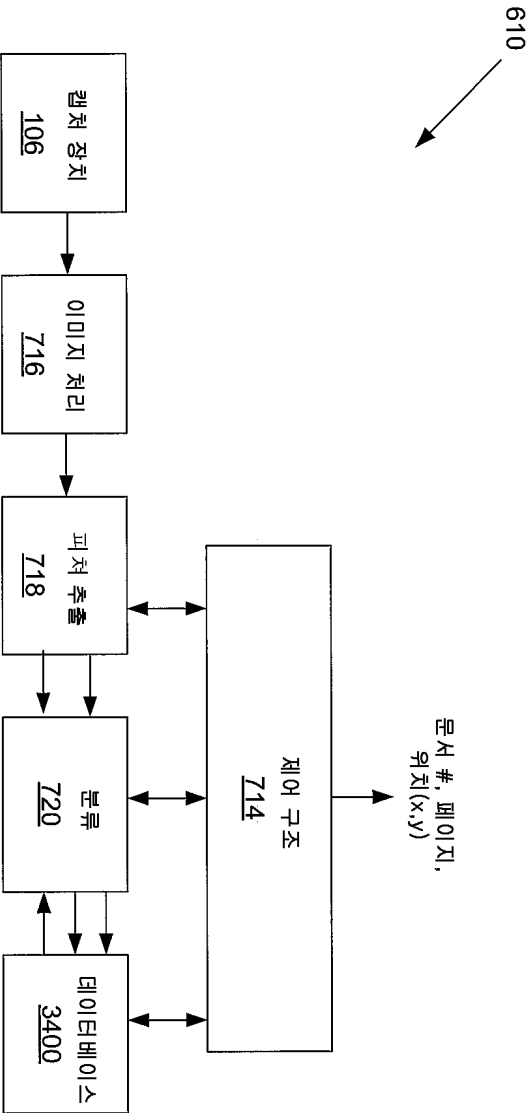


도면25

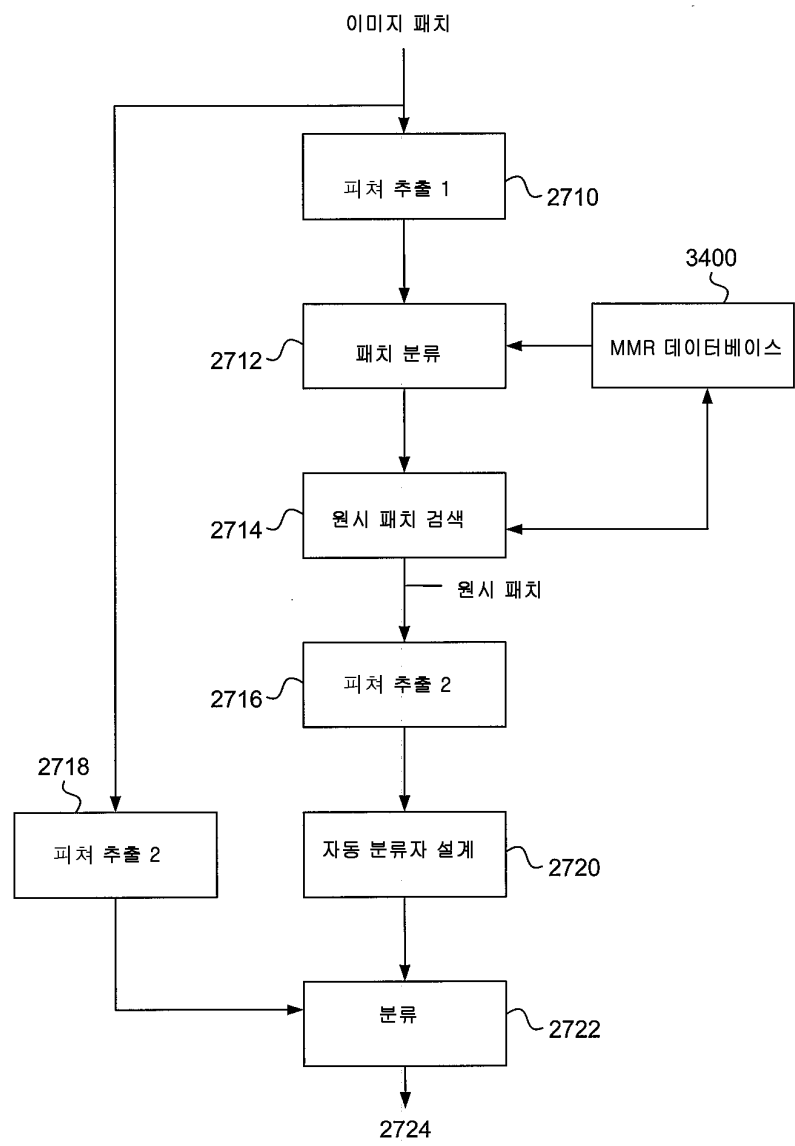




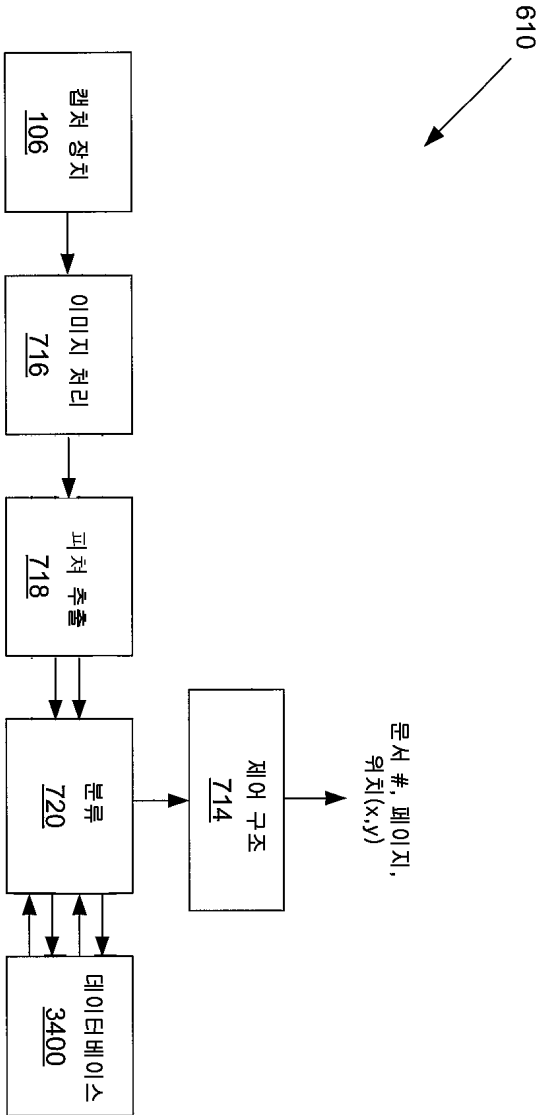
도면26



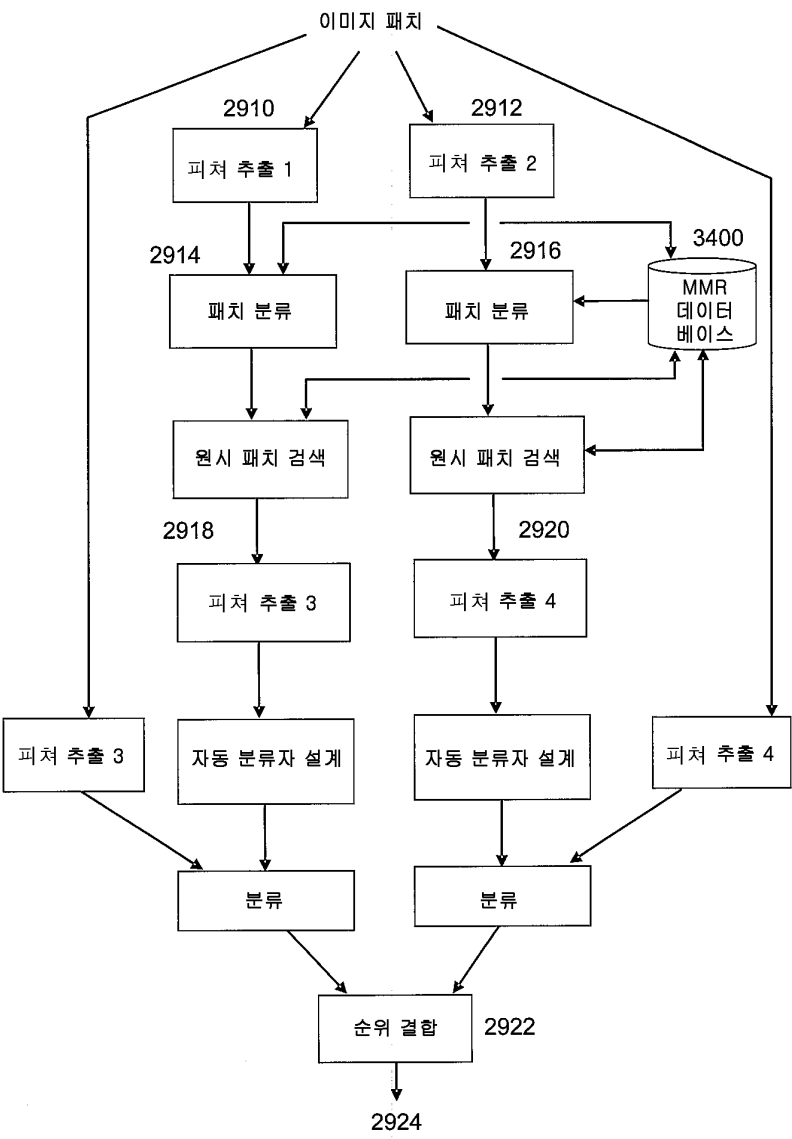
도면27



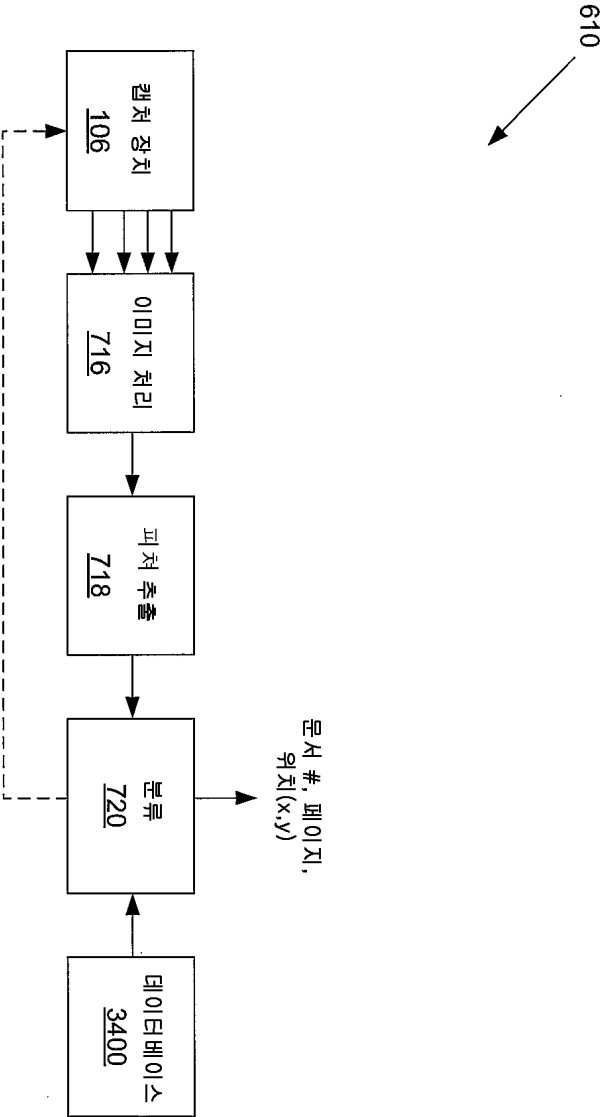
도면28



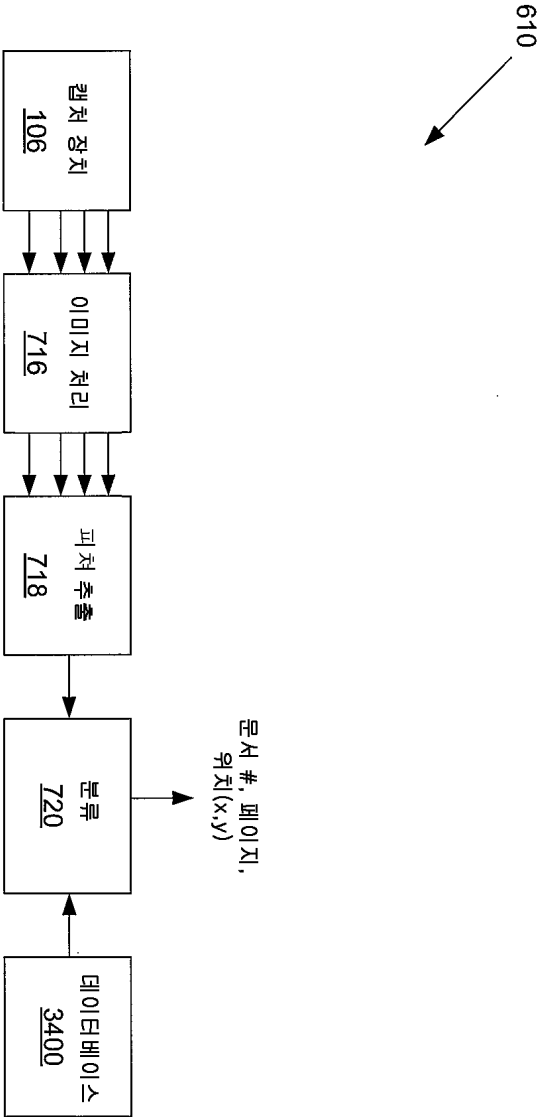
도면29



도면30

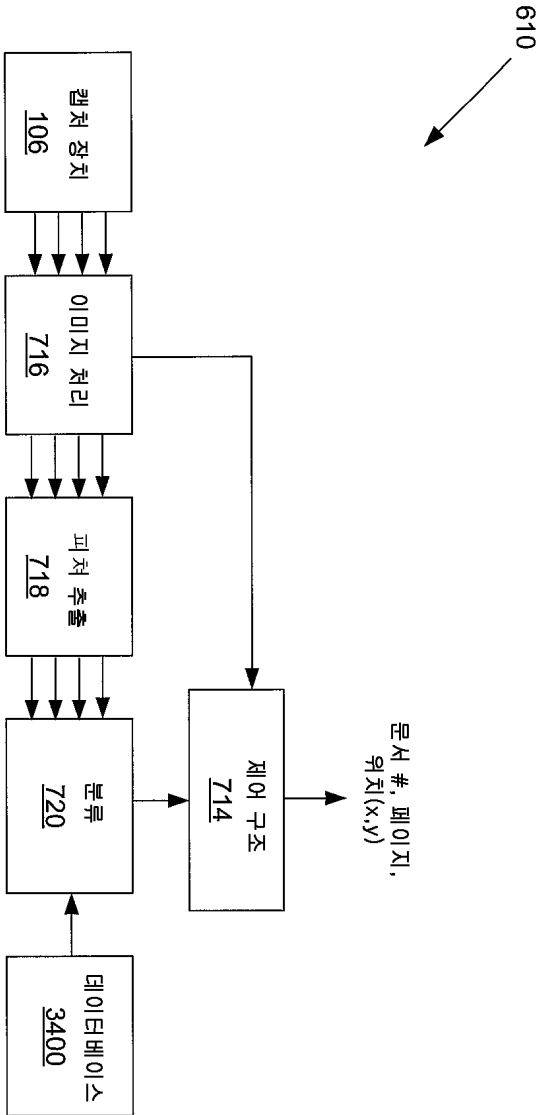


도면31

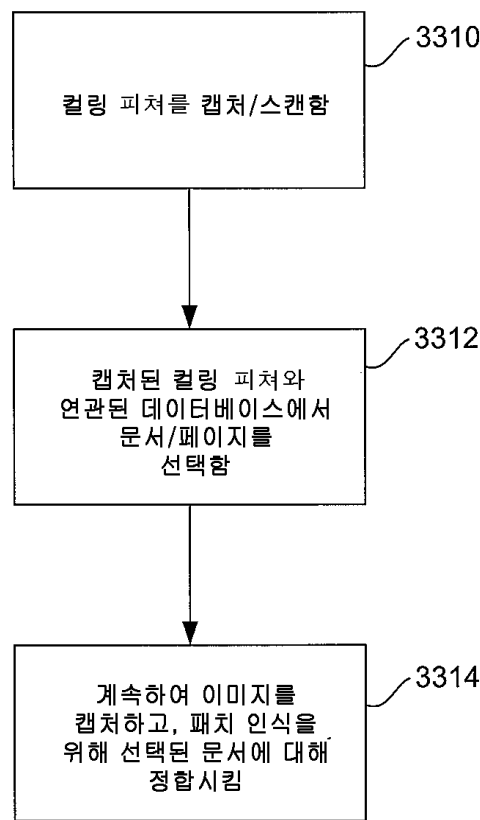




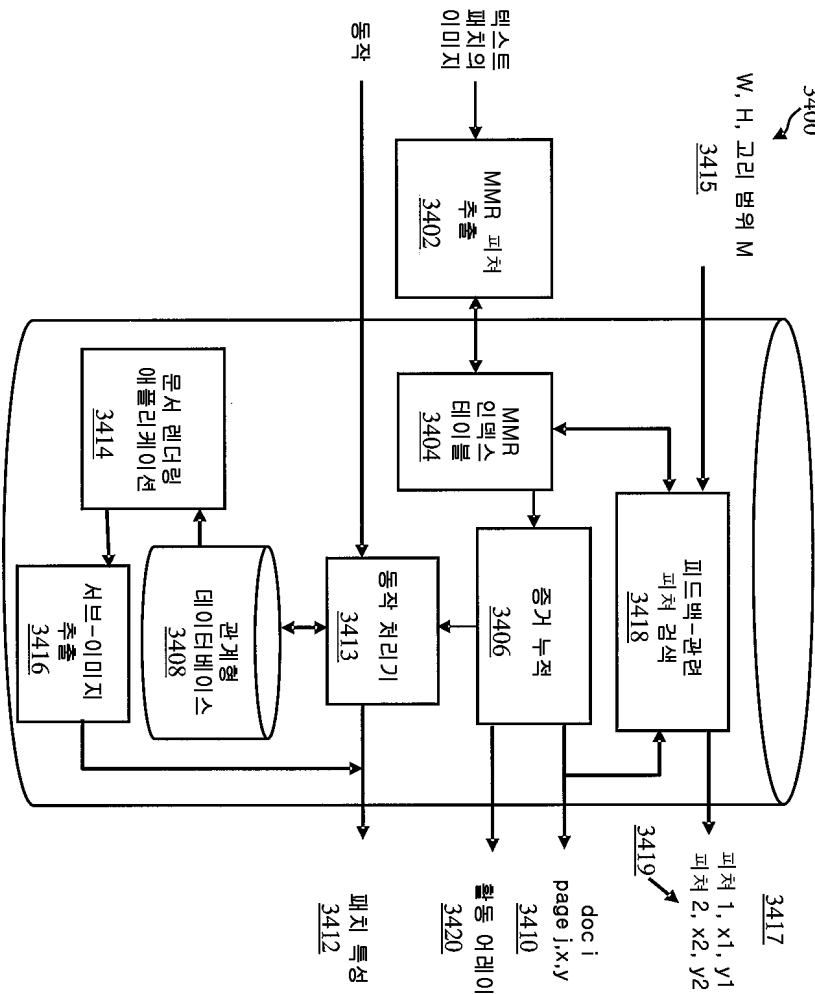
도면32



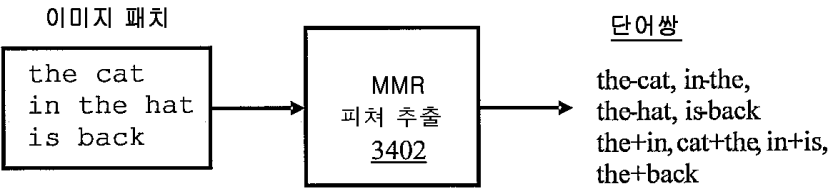
도면33



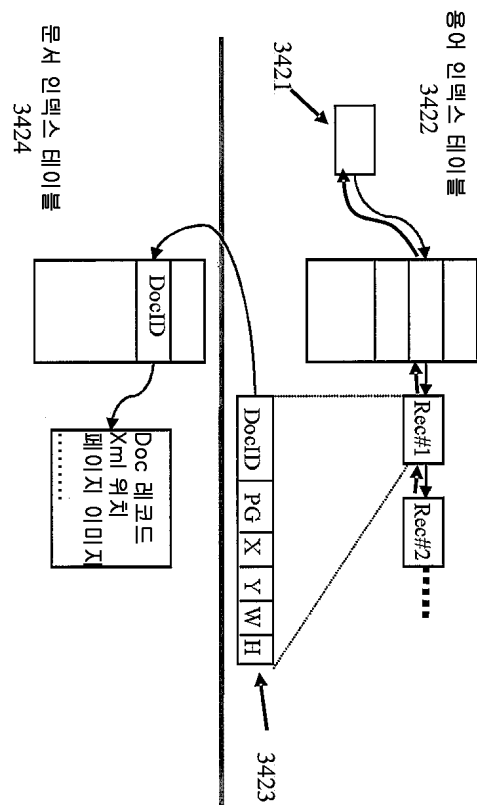
도면34a



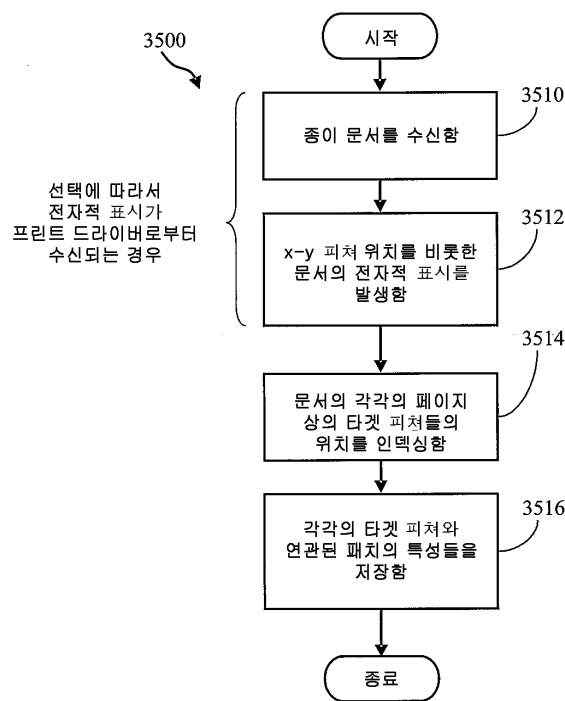
도면34b



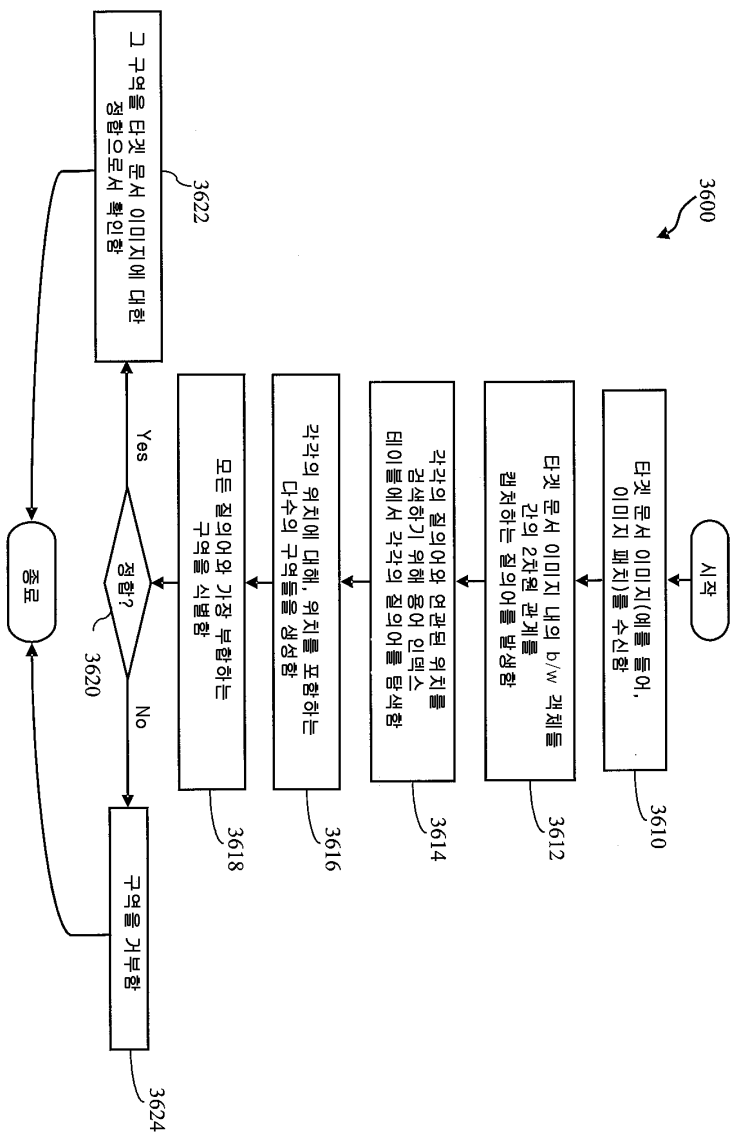
도면34c



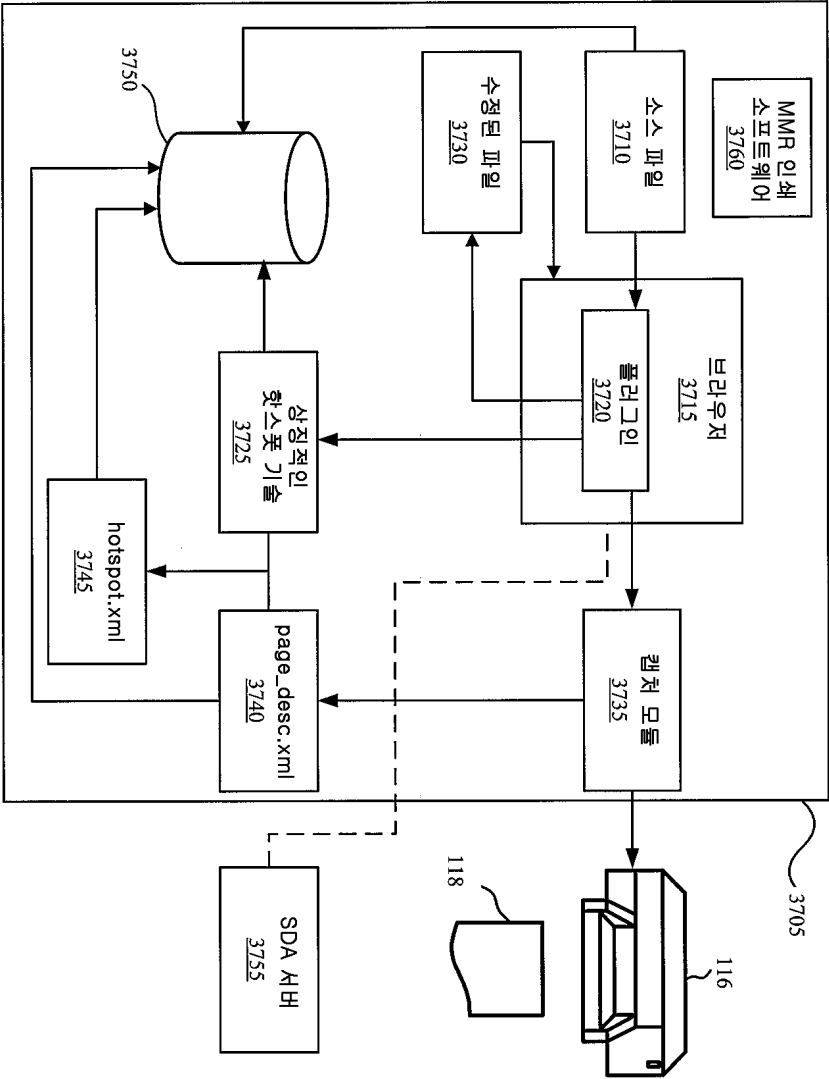
도면35



도면36

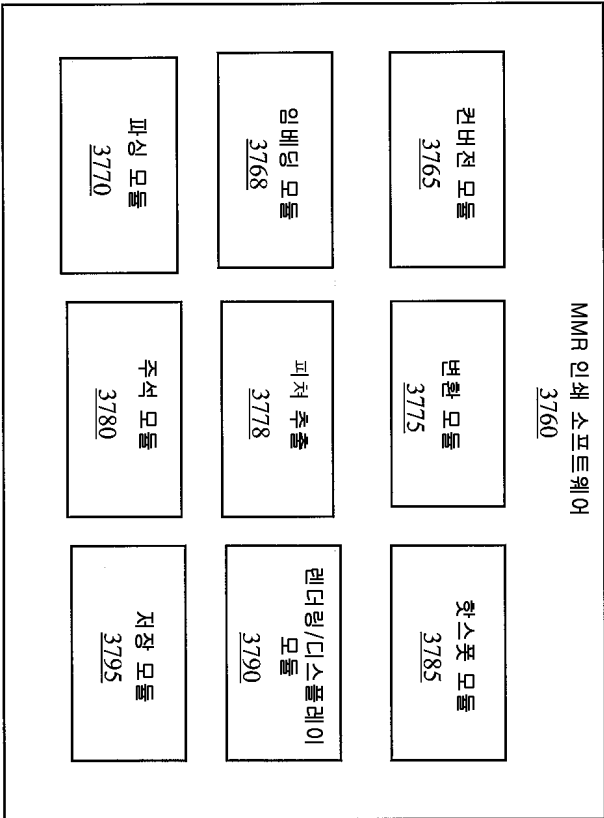


도면37a

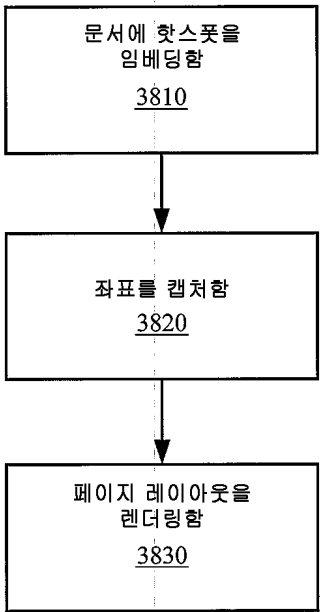




도면37b



도면38



도면39a

```
<html>
Get the greatest copiers
<a
href=http://www.ricoh.com>here.
</a>
</html>
```

3910

3920

도면39b

```
<BODY style="TEXT-DECORATION: none">
Get the greatest copiers

<SPAN>

<SPAN style="FONT-SIZE: 1px;
COLOR: blue;
FONT-FAMILY: 'MMR Courier New'">b0</SPAN>

<A style="TEXT-DECORATION: none"
href="http://www.ricoh.com/">here.</A>

<SPAN style="FONT-SIZE: 1px;
COLOR: blue;
FONT-FAMILY: 'MMR Courier New'">e0</SPAN>

</SPAN> </BODY>
```

3950

3960

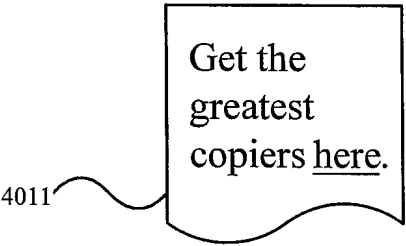
3960

도면40a

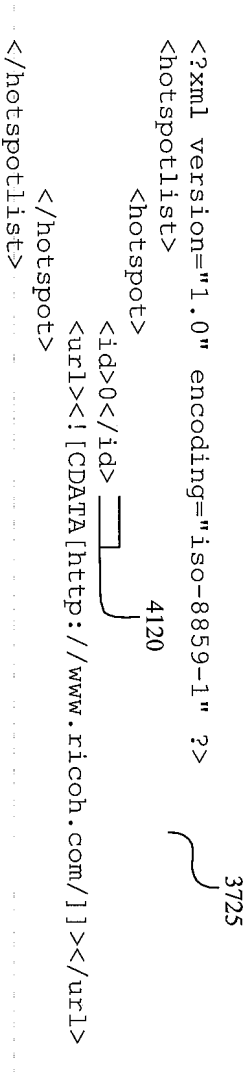
Get the greatest  
copiers here.

4010

도면40b



도면41



도면42a

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<doclayout ID="01150225563">
  <setup>
    <printername>ricoh-lpr</printername>
    <documentname>file://C:\export_to_paper.htm</documentname>
    <prfile>01150225563.spool</prfile>
    <textfile>01150225563.txt</textfile>
    <app>EXPLORE.EXE</app>
    <date>Tue Jun 13 12:06:08 2006</date>
    <dpix>600</dpix>
    <dpiy>600</dpiy>
    <resx>5100</resx>
    <resy>6600</resy>
    <printable_resx>4860</printable_resx>
    <printable_resy>6360</printable_resy>
    <printable_offx>120</printable_offx>
    <printable_offy>120</printable_offy>
    <width>8.5</width>
    <height>11.0</height>
    <imagescale>0.50</imagescale>
  </setup>
  <page no="1">
    <image>01150225563_1.jpg</image>
    <sequence box="333 350 935 91">
      <text>Get the greatest copiers</text>
      <font>
        <face>Times New Roman</face>
        <height>114</height>
        <weight>400</weight>
        <italic>0</italic>
        <underlined>0</underlined>
        <type>0xf</type>
        <truetype>1</truetype>
        <points>12</points>
      </font>
      <word box="333 352 141 69">
        <text>Get</text>
        <char box="333 352 68 69">G</char>
        <char box="405 374 39 47">e</char>
        <char box="447 360 27 61">t</char>
      </word>
      <word box="500 350 119 71">
        <text>the</text>
        <char box="500 360 27 61">t</char>
        <char box="527 350 50 70">h</char>
        <char box="580 374 39 47">e</char>
      </word>
      <word box="648 360 309 81">
        <text>greatest</text>
        <char box="648 374 46 67">g</char>
        <char box="696 374 34 46">r</char>
        <char box="732 374 39 47">e</char>
        <char box="777 374 41 47">a</char>
        <char box="819 360 27 61">t</char>
        <char box="849 374 39 47">e</char>
        <char box="894 374 32 47">s</char>
        <char box="930 360 27 61">t</char>
      </word>
      <word box="985 350 283 91">
        <text>copiers</text>
        <char box="985 374 38 47">c</char>
        <char box="1030 374 44 47">o</char>
        <char box="1076 374 48 67">p</char>
        <char box="1130 350 23 70">i</char>
        <char box="1157 374 39 47">e</char>
        <char box="1199 374 34 46">r</char>
      </word>
    </sequence>
  </page>
</doclayout>

```

도면42b

```

    <char box="1236 374 32 47">s</char>
  </word>
</sequence>
<sequence box="1296 418 4 1">
  <text>b0</text>
  <font>
    <face>MMR Courier New</face>
    <height>6</height>
    <weight>400</weight>
    <italic>0</italic>
    <underlined>0</underlined>
    <type>0xf</type>
    <truetype>1</truetype>
    <points>0</points>
  </font>
  <word box="1296 418 4 1">
    <text>b0</text>
    <char box="1296 418 1 1">b</char>
    <char box="1299 418 1 1">0</char>
  </word>
</sequence>
<sequence box="1303 350 190 71">
  <text>here.</text>
  <font>
    <face>Times New Roman</face>
    <height>114</height>
    <weight>400</weight>
    <italic>0</italic>
    <underlined>0</underlined>
    <type>0xf</type>
    <truetype>1</truetype>
    <points>12</points>
  </font>
  <word box="1303 350 169 71">
    <text>here</text>
    <char box="1303 350 50 70">h</char>
    <char box="1356 374 39 47">e</char>
    <char box="1397 374 34 46">r</char>
    <char box="1433 374 39 47">e</char>
  </word>
  <punctuation box="1482 410 11 11">
    <text>.</text>
    <char box="1482 410 11 11">.</char>
  </punctuation>
</sequence>
<sequence box="1500 418 4 1">
  <text>e0</text>
  <font>
    <face>MMR Courier New</face>
    <height>6</height>
    <weight>400</weight>
    <italic>0</italic>
    <underlined>0</underlined>
    <type>0xf</type>
    <truetype>1</truetype>
    <points>0</points>
  </font>
  <word box="1500 418 4 1">
    <text>e0</text>
    <char box="1500 418 1 1">e</char>
    <char box="1503 418 1 1">0</char>
  </word>
</sequence>

```

도면43

3745

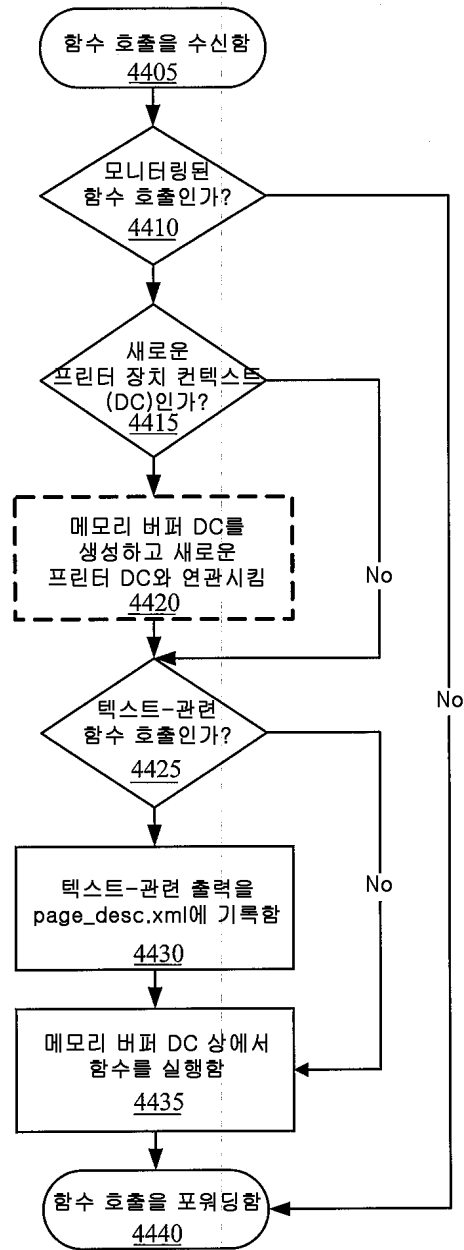
```

<?xml version="1.0" encoding="iso-8859-1" ?>
<hotspotlayout>
  <setup>
    <printcapfile>01150225563.xml</printcapfile>
    <printername>ricoh-lpr</printername>
    <documentname>
      <![CDATA[
        file://C:\export_to_paper.htm
      ]]>
    </documentname>
    <prfile>01150225563.spool</prfile>
    <textfile>01150225563.txt</textfile>
    <app>IEXPLORE.EXE</app>
    <date>Tue Jun 13 12:06:08 2006</date>
    <dpix>600</dpix>
    <dpiy>600</dpiy>
    <resx>5100</resx>
    <resy>6600</resy>
    <width>8.5</width>
    <height>11</height>
    <imagescale>0.5</imagescale>
  </setup>
  <page no="1">
    <hotspot box="1303 350 190 71">
      <id>0</id>
      <url>
        <![CDATA[ http://www.ricoh.com/ ]]>
      </url>
    </hotspot>
  </page>
</hotspotlayout>

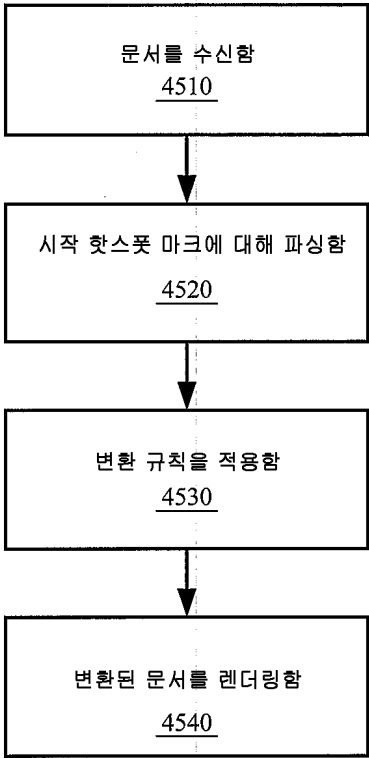
```



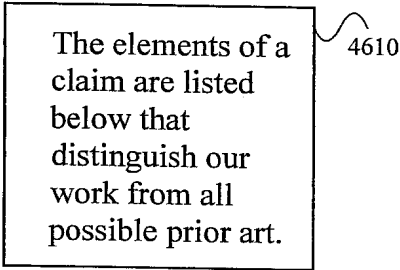
도면44



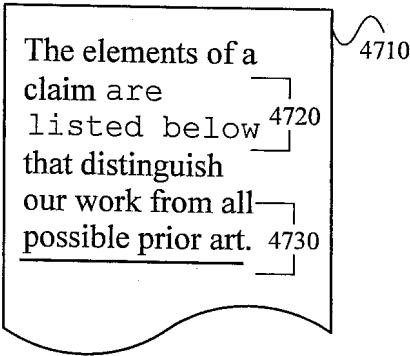
도면45



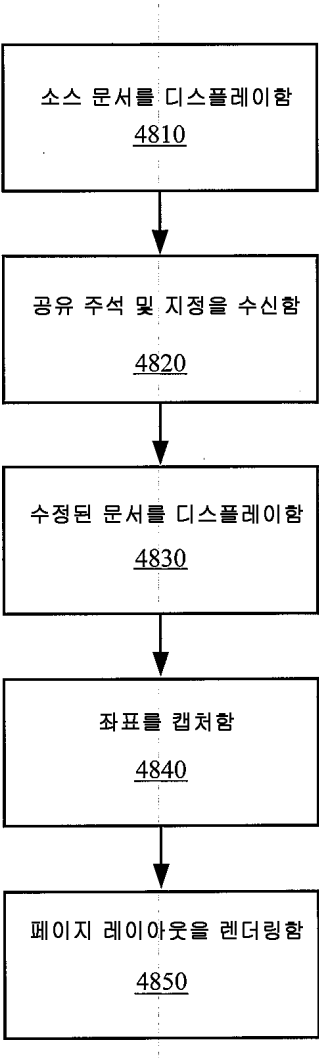
도면46



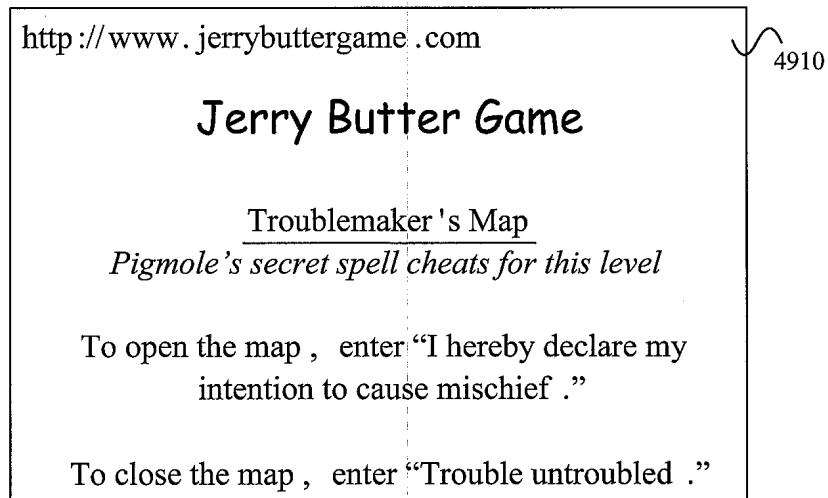
도면47



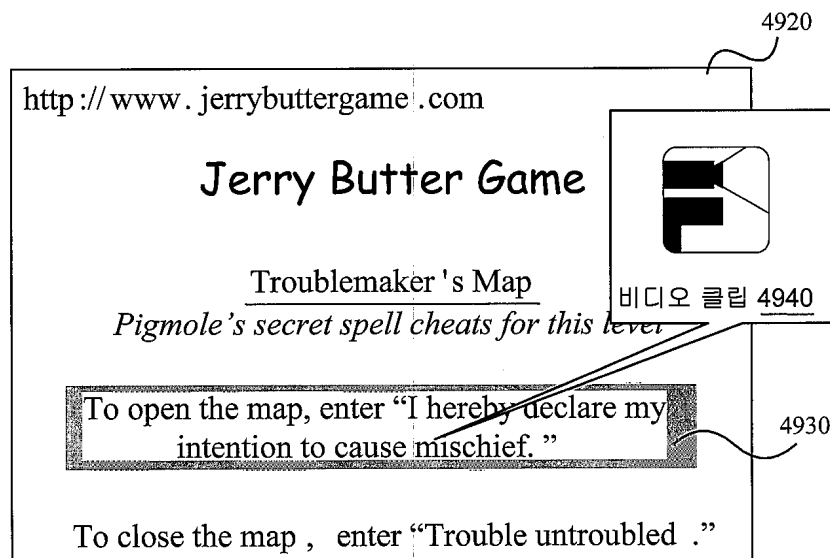
도면48



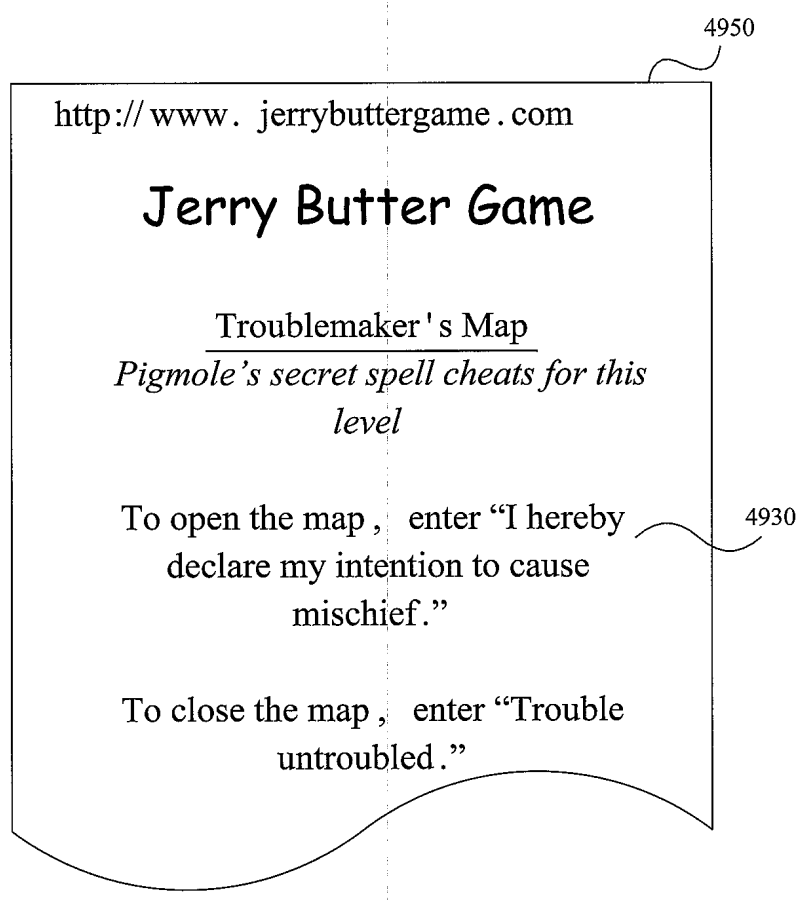
도면49a



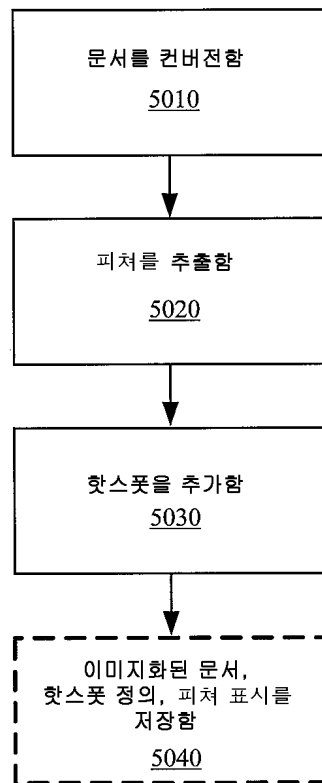
도면49b



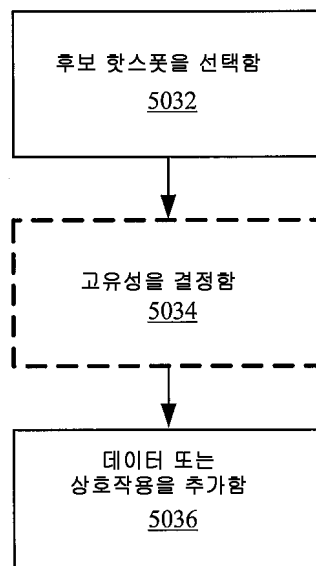
도면49c



도면50a



도면50b



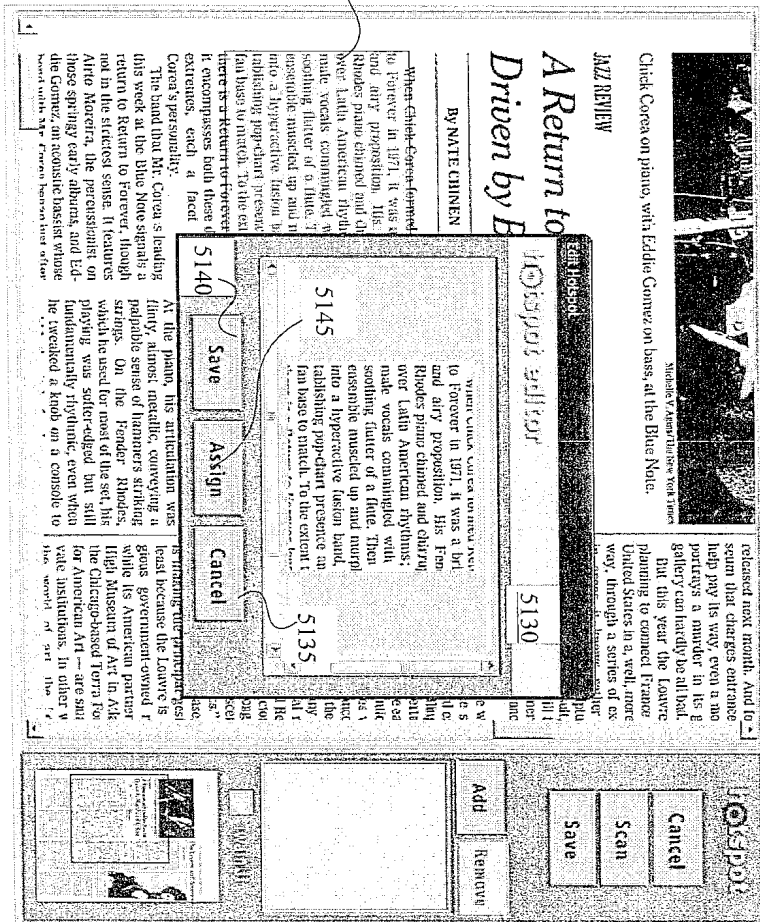
도면51a

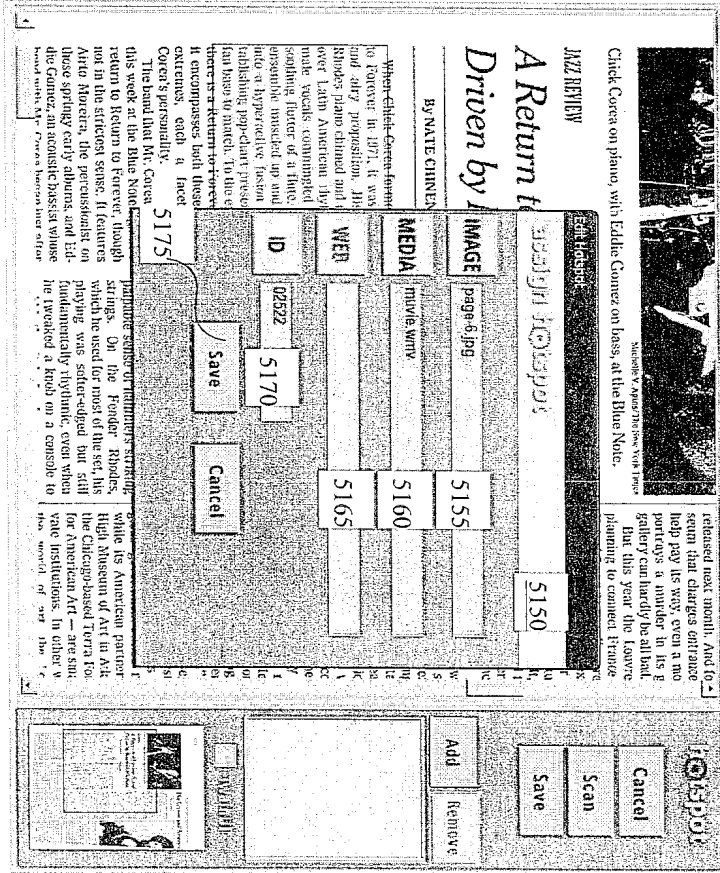




도면51b

5125





도면51d

# Return to a Familiar Sound Written by Brazilian Rhythms

By NATE CHENEN

When Chick Corea formed Return

Forever in 1971, it was a bright  
any proposition. His Federal  
uses piano chords and chirped  
Latin American rhythms, lo-  
le vocals commingled with the  
thing factor of a tale. Then the  
ruble muscled up and morphed  
a hyperactive fusion band, es-  
tablishing pop-chart presence and a  
basso to match. To the extent that  
it is a Return to Forever legacy,  
it encompasses both these dynamic  
genres, each a facet of Mr.  
Corea's personality.

the band that Mr. Corea is leading  
work at the Blue Note signals a  
turn to Return to Forever, though  
in the strictest sense. It features  
to Moreira, the percussionist on  
so springy early albums, and Ed-  
Gomez, an acoustic bassist whose  
droll Mr. Corea began just after  
World War II. Mr. Gomez fills  
role that originally belonged to  
Ney Clarke, the only musician be-  
as Mr. Corea to last through every  
Return to Forever permutation. That  
core involvement explains one  
sequence of Mr. Corea's nonpal-  
pation here: the Blue Note is dif-  
ferent, too. Long, yes, however, not

Chick Corea  
Blue Note

both Mr. Gomez and Mr. Moreira  
Less especially, the set included as  
many standards as original composi-  
tions, one by Antonio Carlos Jobim  
and a con. of course, the John Tume  
"Desafinado," was a bossa nova; one  
of the Rodgers songs, "With a Song in  
My Heart," was least as it sounds like  
Mr. Corea has always been a per-  
cussionist at heart, and he indulged  
that affinity with a cobble, some  
shakers and a Brazilian hand drum.  
At the piano, his articulation was  
flurry, almost metallic, conveying a  
palpable sense of hurried striding,  
strings. On the Federal Rhodes,  
which he used for most of the set, his  
playing was seldom edged but still  
fundamentally rhythmic, even when  
he tweaked a knob on a console to  
wobble the pitch to a phrase.

There was suddenly just as much  
rhythm in the totally virtuosic of  
Mr. Gomez, whose strongest solo  
was on "But Beautiful," the second  
Rodgers tune, and in the drumming  
of Mr. Moreira, which embedded the  
sort of earthy impression that rare-  
ly has occasion to flourish in Mr.  
Corea's working hands.

paintings, drawings, sculptures and  
decorative arts. As a result, part of  
the Louvre's collection will travel to  
the United States and American art  
will make a rare appearance in the  
Louvre.

If relatively few people will actu-  
ally see these works, the symbolic  
importance of the cultural exchange  
is that it indicates a willingness to  
collaborate on matters pertaining to  
the spirit, even when more earthly is-  
sues are subjects of contention.

Indeed, these exchanges were be-  
ing worked out when Franco-Ameri-  
can relations were at their most  
strained over Iraq. "Many felt this  
need to reinforce cultural relations  
to help get over this," said Henri Loy-  
rette, the president-director of the  
Louvre. "This was a strong feeling  
that culture could transcend what  
was taking place in politics."

At it happens, in this case, France  
is making the principal gesture, not  
least because the Louvre is a pres-  
tigious government-owned museum,  
while its American partners — the  
High Museum of Art in Atlanta and  
the Chicago-based Terra Foundation  
for American Art — are smaller pri-  
vate institutions. In other words, in  
the world of art, the Louvre is  
France.

The most ambitious itinerary in-  
cludes showing art from the Louvre  
at the High Museum over a three-  
year period, starting in October with  
the temporary exhibitions built  
around three themes: the royal col-  
lections from France's 16th-19th cen-  
turies; the treasures of ancient civil-  
izations; and "The Louvre Today and

Cancel

Save

Add

Remove

X=50, Y=57.4, W=100, H=100


X=48.4, Y=61.0, W=100, H=100

X=38.6, Y=75.5, W=100, H=100

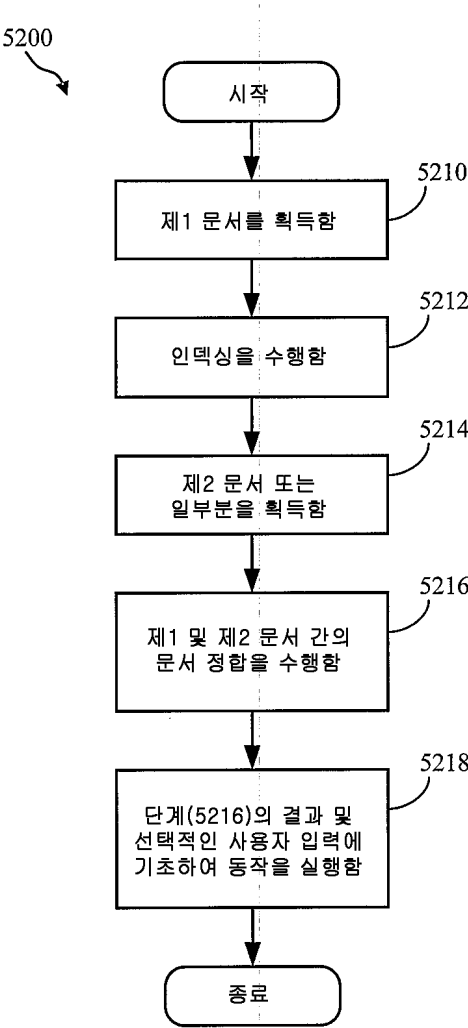
X=49.0, Y=89.2, W=100, H=100

X=52, Y=87.3, W=100, H=100

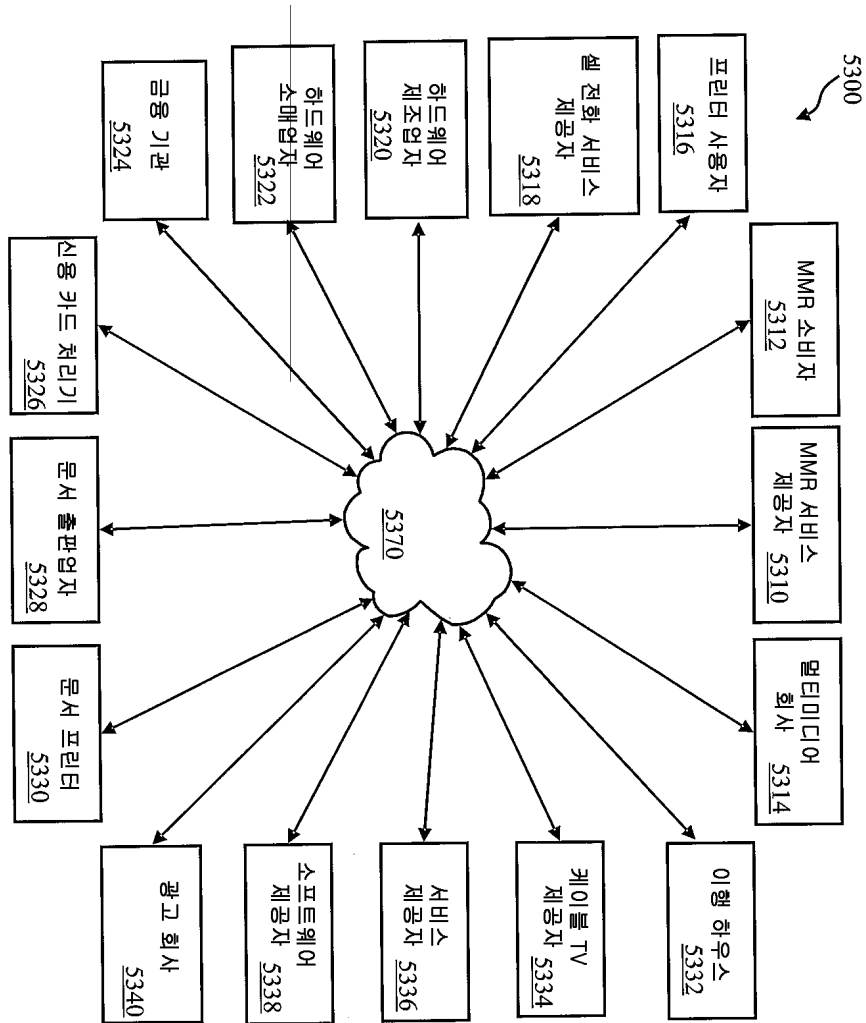
X=49.2, Y=23.5, W=100, H=100



도면52



도면53



도면54

