



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년01월14일
(11) 등록번호 10-1008464
(24) 등록일자 2011년01월07일

(51) Int. Cl.
H04N 21/47 (2011.01)
(21) 출원번호 10-2004-0020018
(22) 출원일자 2004년03월24일
심사청구일자 2009년02월24일
(65) 공개번호 10-2004-0084774
(43) 공개일자 2004년10월06일
(30) 우선권주장
10/395,679 2003년03월24일 미국(US)
(56) 선행기술조사문헌
EP01094406 A2
GB2340633 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
마이크로소프트 코퍼레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
샌더스스코트디.
미국94587캘리포니아주유니온시자스민코트34495
(74) 대리인
주성민, 이중희, 백만기

전체 청구항 수 : 총 29 항

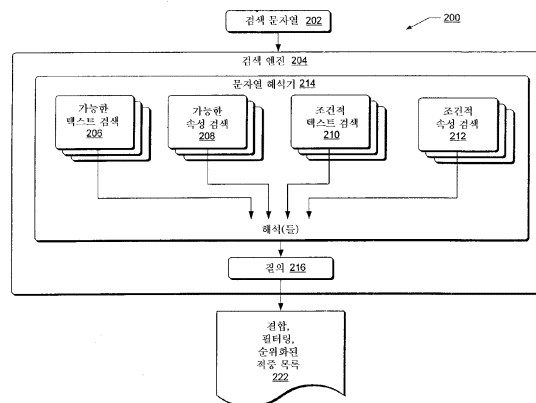
심사관 : 오석환

(54) 전자 프로그램 가이드(EPG) 데이터의 검색을 위한 방법, 검색 엔진, 시스템, 컴퓨터 판독가능한 기록 매체, 및 EPG 정보 검색 방법

(57) 요약

본 발명은 EPG 데이터를 위한 검색 엔진과 관련 방법들을 포함한다. 예시적인 방법에 있어서, 텍스트 검색 문자열은 검색가능한 용어들과 텍스트 검색 용어들 혹은 속성 검색 용어들로 해석되는 용어들로 표준화될 수 있다. 복잡성의 정도가 변화하는 검색 조건을 갖는 하나 이상의 질의가 검색 문자열의 용어 해석에 따라 생성된다. 텍스트 및 속성 용어들의 해석에 기초하여 EPG 데이터베이스 및/또는 웹 자원들 내에서 하나 이상의 검색이 수행되고 그 결과들에 이 해석에 따른 관련성 순위가 주어진다. 결합된 검색 결과들은 사용자에게 그룹화, 순위화, 및 디스플레이를 위해 필터링될 수 있다. 검색 문자열의 각 문자가 사용자에게 의해 입력됨에 따라 점진적으로 결과들이 디스플레이될 수도 있다.

대표도



특허청구의 범위

청구항 1

전자 프로그램 가이드(EPG: electronic program guide) 데이터를 검색하는 방법에 있어서,

검색 문자열 수신기(search string receiver)에 의해 텍스트 검색 문자열(text search string)을 수신하는 단계 - 상기 텍스트 검색 문자열은 하나 이상의 텍스트 용어들과 하나 이상의 속성 용어들을 포함함 - ;

검색가능한 용어들(searchable terms)을 만들기 위해, 상기 텍스트 검색 문자열에 대해 스펠링 체크를 적용하고, 단어를 잘못 스펠링된 것으로 식별하고, 상기 잘못 스펠링된 것으로 식별된 상기 단어의 정정된 형태를 상기 검색가능한 용어들에 첨부하여 상기 단어의 상기 정정된 형태에 기초한 조건 검색을 행함으로써, 상기 텍스트 검색 문자열을 표준화(normalizing)하는 단계;

문자열 해석기(string interpreter)에 의해 상기 검색가능한 용어들을 텍스트 용어들과 속성 용어들로 분리하는 단계 - 상기 문자열 해석기는 상기 텍스트 검색 문자열을 분석(parsing)하여, 상기 속성 용어들을 식별하기 위한 속성 목록(a list of attributes)과 상기 검색가능한 용어들을 비교함으로써 상기 검색가능한 용어들을 상기 텍스트 용어들과 상기 속성 용어들로 분리함 -;

각 텍스트 용어에 대해 상기 EPG 데이터에서 텍스트 검색을 수행하는 단계;

각 속성 용어에 대해 상기 EPG 데이터에서 속성 검색을 수행하는 단계; 및

상기 검색 문자열에 맞는(conforming) 검색 결과들을 디스플레이하기 위해 상기 텍스트 검색 및 상기 속성 검색으로부터의 결과들을 순위화(ranking)하는 단계 - 상기 순위화하는 단계는, 상기 잘못 스펠링된 것으로 식별된 단어에 기초한 검색 결과들이 얻어지는 경우, 상기 단어의 상기 정정된 형태에 대한 상기 조건 검색으로부터 얻어지는 적중(hits)의 관련성 순위를 낮추는 것을 포함함 -

를 포함하는 EPG 데이터 검색 방법.

청구항 2

제1항에 있어서,

상기 표준화하는 단계는 상기 검색가능한 용어들의 언어 규칙을 상기 EPG 데이터의 언어 규칙과 매치(match)시키기 위하여 변환하는 단계를 포함하는 EPG 데이터 검색 방법.

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 분리하는 단계는, 상기 속성 목록과 상기 검색가능한 용어들을 비교하는데 더하여 상기 텍스트 용어들을 상기 속성 용어들로부터 분리하기 위한 추가적 정보를 얻기 위해 상기 속성 목록과 상기 검색가능한 용어들을 비교함으로써 식별된 상기 속성 용어들 중 하나 이상을 이용한 대안적 텍스트 검색(alternate text search)으로부터 획득된 중간 결과들과 상기 검색가능한 용어들을 비교하는 단계를 포함하는 EPG 데이터 검색 방법.

청구항 5

제1항에 있어서,

상기 분리하는 단계는 다양한 조합들로 상기 검색가능한 용어들을 그룹화하여 검색가능한 어구들의 세트를 얻는(achieve) 단계를 더 포함하는 EPG 데이터 검색 방법.

청구항 6

제5항에 있어서,

상기 분리하는 단계는 상기 그룹화 전에 상기 검색가능한 용어들 중 적어도 두개의 순서를 바꾸는(transpose) 단계를 더 포함하는 EPG 데이터 검색 방법.

청구항 7

제6항에 있어서,

상기 검색가능한 용어들 중 상기 적어도 두개의 순서를 바꾸는 단계에 의해 생성된 상기 검색가능한 어구들을 상기 속성 목록과 비교하여 상기 검색가능한 어구들 중 어떤 것이 속성 용어들인지를 판정하는 단계를 더 포함하는 EPG 데이터 검색 방법.

청구항 8

제1항에 있어서,

상기 속성 목록은 각 색인된 속성에 대한 검색 값들과 목적 자원들(destination resources)에 색인된 속성들(indexed attributes)을 관련시키는 색인을 포함하는 EPG 데이터 검색 방법.

청구항 9

제1항에 있어서,

상기 속성 목록은, 상기 속성 목록 내의 속성을 이용한 검색을 지도(guide)하기 위한 상기 속성 목록 내의 상기 속성과 관련된 브랜칭 로직(branching logic)을 포함하는 EPG 데이터 검색 방법.

청구항 10

제7항에 있어서,

상기 속성 목록은, 상기 속성 목록 내의 속성을 이용한 검색을 지도하기 위한 상기 속성 목록 내의 상기 속성에 관한 메타데이터를 포함하는 EPG 데이터 검색 방법.

청구항 11

제1항에 있어서,

상기 검색 용어들을 분리하는 단계는 상기 텍스트 검색 문자열을 구성함에 있어서 사용자의 의도의 해석을 형성하는 단계를 포함하는 EPG 데이터 검색 방법.

청구항 12

제1항에 있어서,

상기 속성 목록과 상기 검색가능한 용어들을 비교하는데 더하여, 상기 검색가능한 용어들 중 하나 이상을 이용한 EPG 데이터의 대안적 속성 검색을 행하여, 상기 대안적 속성 검색의 결과들에 기초해서 상기 검색가능한 용어들을 상기 텍스트 용어들과 상기 속성 용어들로 더 분리하는 단계를 더 포함하는 EPG 데이터 검색 방법.

청구항 13

검색 엔진을 포함하는 컴퓨터 판독가능 기록 매체로서, 상기 검색 엔진은

검색 문자열에 대해 스펠링 체크를 적용하여 단어를 잘못 스펠링된 것으로 식별하는 노말라이저(normalizer) - 상기 노말라이저는 상기 잘못 스펠링된 것으로 식별된 상기 단어의 정정된 형태를 상기 검색 문자열의 검색 용어들에 첨부하여 상기 단어의 상기 정정된 형태에 기초한 조건 검색을 행함 - ;

속성 목록과 통신가능하게 결합된 문자열 해석기 - 상기 문자열 해석기는 상기 검색 문자열 내의 용어들의 의미를 가정(postulate)하고, 속성 용어들을 식별하기 위한 상기 속성 목록과 상기 검색 문자열 내의 후보 속성 용어들을 비교함으로써 상기 검색 문자열 내에서 텍스트 용어들로부터 상기 속성 용어들을 분리하는 속성 식별기를 더 포함함 -;

상기 문자열 해석기에 결합된 질의 생성기 - 상기 질의 생성기는 식별된 상기 속성 용어들에 기초하여 전자 프로그램 가이드(EPG) 데이터를 검색하기 위한 하나 이상의 속성 질의를 생성하고 상기 텍스트 용어들에 기초하여

상기 EPG 데이터를 검색하기 위한 하나 이상의 텍스트 질의들을 생성함 - ; 및

상기 하나 이상의 속성 질의 및 상기 하나 이상의 텍스트 질의 중 적어도 하나가 상기 잘못 스펠링된 것으로 식별된 단어를 포함하는 적중을 반환하는 경우 상기 단어의 상기 정정된 형태에 대한 상기 조건 검색으로부터 얻어지는 적중의 관련성 순위를 낮추는 결과 순위화기(results ranker)

를 포함하는, 컴퓨터 판독가능 기록 매체.

청구항 14

삭제

청구항 15

삭제

청구항 16

제13항에 있어서,

상기 검색 엔진은 EPG 데이터베이스들과 웹기반 EPG 자원들 내에서 검색을 수행하는, 컴퓨터 판독가능 기록 매체.

청구항 17

제13항에 있어서,

검색 비교기/필터 및 상기 결과 순위화기를 포함하는 결과 컴파일러를 더 포함하는, 컴퓨터 판독가능 기록 매체.

청구항 18

제17항에 있어서,

상기 검색 비교기/필터는 상기 문자열 해석기에 의해 가정된 용어들의 의미와 검색 결과들을 비교하는, 컴퓨터 판독가능 기록 매체.

청구항 19

제17항에 있어서,

상기 결과 순위화기는 상기 문자열 해석기에 의해 가정된 용어들의 의미에 기초하여 검색 결과들에 관련성 순위를 할당하는, 컴퓨터 판독가능 기록 매체.

청구항 20

전자 프로그램 가이드(EPG) 데이터에 대한 검색 엔진을 포함하는 컴퓨터 판독가능 기록 매체로서, 상기 검색 엔진은 컴퓨터 판독가능 기록 매체에 저장된 명령어들을 실행하는 프로세서에 의해 구현되고, 상기 검색 엔진은

텍스트 검색 문자열에 대해 스펠링 체크를 적용하여 단어를 잘못 스펠링된 것으로 식별하는 노말라이저(normalizer) - 상기 노말라이저는 상기 잘못 스펠링된 것으로 식별된 상기 단어의 정정된 형태를 상기 텍스트 검색 문자열의 검색 용어들에 첨부하여 상기 단어의 상기 정정된 형태에 기초한 조건 검색을 행함 - ;

상기 텍스트 검색 문자열의 문자 및 속성 인식 해석(text and attribute aware interpretation)을 달성하도록 상기 텍스트 검색 문자열을 분석하여 텍스트 용어들과 속성 용어들을 분리하는 속성 식별기;

상기 속성 용어들 및 상기 텍스트 용어들에 기초하여 EPG 데이터를 검색하기 위한 질의를 생성하는 질의 생성기; 및

상기 잘못 스펠링된 것으로 식별된 단어에 기초한 검색 결과들이 얻어지는 경우 상기 단어의 상기 정정된 형태에 대한 상기 조건 검색으로부터 얻어지는 적중의 관련성 순위를 낮추는 결과 순위화기

를 포함하는, 컴퓨터 판독가능 기록 매체.

청구항 21

전자 프로그램 가이드(EPG) 데이터 검색하는 방법에 있어서,

검색 문자열 수신기에 의해 텍스트 검색 문자열을 수신하는 단계 - 상기 텍스트 검색 문자열은 프로그램 제목들을 검색하기 위한 하나 이상의 텍스트 용어들과 프로그램 속성(program characteristic)들을 검색하기 위한 하나 이상의 속성 용어들을 포함함 - ;

검색가능한 용어들을 만들기 위해 상기 텍스트 검색 문자열에서 기호(symbols)들을 제거하거나 변환하여 상기 텍스트 검색 문자열을 표준화하는 단계 - 상기 표준화하는 단계는 또한 상기 텍스트 검색 문자열에 대해 스펠링 체크를 적용하는 단계, 단어를 잘못 스펠링된 것으로 식별하는 단계, 및 상기 잘못 스펠링된 것으로 식별된 상기 단어의 정정된 형태를 상기 검색가능한 용어들에 첨부하여 상기 단어의 상기 정정된 형태에 기초한 조건 검색을 행하는 단계를 포함함 - ;

프로세서에 의해, 공지 속성 용어들을 식별하기 위한 공지 속성 목록(a list of known attributes)과 상기 검색가능한 용어들을 비교하여 하나 이상의 상기 공지 속성 용어를 판정하는 단계 - 상기 공지 속성 목록은 상기 EPG 데이터에서 검색가능한 프로그램들의 제목이 아닌 속성들(non-title characteristics)에 관련됨 - ;

공지 속성 용어들로 인식되지 않은 남은 용어들을 텍스트 용어들로 식별하는 단계;

상기 텍스트 검색 문자열로부터 식별된 각 텍스트 용어에 대해 상기 EPG 데이터에서 프로그램 제목 데이터의 텍스트 검색을 수행하여 제1 중간(intermediate) 결과들을 획득하는 단계;

상기 텍스트 검색 문자열로부터 식별된 각각의 공지 속성 용어에 대해 상기 EPG 데이터에서 속성 검색을 수행하여 제2 중간 결과들을 획득하는 단계;

상기 텍스트 용어들 중 하나 이상을 이용하여 EPG 데이터에 대한 대안적 속성 검색(alternate attribute search)을 행하여 상기 하나 이상의 텍스트 용어들이 상기 공지 속성 목록에 리스팅되지 않은 속성 용어들인지 여부를 판정하기 위한 제3 중간 결과들을 획득하는 단계;

상기 속성 목록과 상기 검색가능한 용어들을 비교하여 식별된 상기 공지 속성 용어들 중 하나 이상을 이용해서 대안적 텍스트 검색을 행하여 상기 공지 속성 용어들이 텍스트 용어들이 아님을 검증하기 위한 제4 중간 결과들을 획득하는 단계; 및

상기 텍스트 검색 문자열에 맞는(conforming) 검색 결과들을 디스플레이하기 위해 상기 텍스트 검색 및 상기 속성 검색으로부터의 상기 제1, 제2, 제3 및 제4 중간 결과들을 결합, 순위화 및 필터링하는 것 중 하나를 수행하는 단계 - 상기 순위화는 상기 잘못 스펠링된 것으로 식별된 단어에 기초한 검색 결과들이 얻어지는 경우, 상기 단어의 상기 정정된 형태에 대한 상기 조건 검색으로부터 얻어지는 적중의 관련성 순위를 낮추는 것을 포함함 -

를 포함하는 EPG 데이터 검색 방법.

청구항 22

제21항에 있어서,

상기 속성 목록은 각 색인된 속성에 대한 검색 값들과 목적 자원들에 색인된 속성들을 관련시키는 색인을 포함하는 EPG 데이터 검색 방법.

청구항 23

제21항에 있어서,

상기 속성 목록은, 상기 속성 목록 내의 속성을 이용한 검색을 지도하기 위한 상기 속성 목록 내의 상기 속성에 관한 메타데이터를 포함하는 EPG 데이터 검색 방법.

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

동작들을 수행하기 위한 컴퓨터에 의해 실행가능한 명령어들을 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체로서, 상기 동작들은

텍스트 검색 문자열을 수신하는 단계;

상기 텍스트 검색 문자열을 검색될 EPG 데이터와 호환가능한 검색 용어들로 포맷하는 단계;

속성 목록에 기초해서 속성 용어들을 식별하기 위해 상기 속성 목록과 상기 검색 용어들을 비교하여 상기 검색 용어들로부터 텍스트 용어들과 속성 용어들을 유도하는(deriving) 단계 - 속성은 검색가능한 프로그램 속성들임 - ;

상기 텍스트 검색 문자열에 대해 스펠링 체크를 적용하는 단계;

단어를 잘못 스펠링된 것으로 식별하는 단계;

상기 잘못 스펠링된 것으로 식별된 상기 단어의 정정된 형태를 상기 검색 용어들에 첨부하여 상기 단어의 상기 정정된 형태에 기초한 조건 검색을 행하는 단계;

각 텍스트 용어에 대해 상기 EPG 데이터에서의 프로그램 제목 데이터에 대해 하나 이상의 텍스트 검색을 수행하는 단계;

식별된 각 속성 용어에 대해 상기 EPG 데이터에서의 프로그램 속성 데이터에 대해 하나 이상의 속성 검색을 수행하는 단계;

상기 하나 이상의 텍스트 검색의 결과들을 상기 하나 이상의 속성 검색의 결과들과 결합하여 검색 결과 적중들을 획득하는 단계;

각 검색 결과 적중에 대해 관련성 순위를 할당하는 단계;

상기 잘못 스펠링된 것으로 식별된 단어에 기초한 검색 결과들이 얻어지는 경우, 상기 단어의 상기 정정된 형태에 대한 상기 조건 검색으로부터 얻어지는 적중의 관련성 순위를 낮추는 단계; 및

관련성에 의해 필터링된 적중들을 디스플레이하는 단계를 포함하는,

하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 28

제27항에 있어서,

상기 포맷하는 단계는, 자동 정정, 용어 분리, 레터케이스(letter case) 표준화, 구두점 제거, 기호 표준화, 관사 제거, 및 상기 검색 문자열의 복수화(pluralizing) 중 하나를 포함하는, 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 29

제27항에 있어서,

검색 용어의 단어 스템밍된 버전(word-stemmed version)을 사용하여 검색을 수행하기 위한 명령어들을 더 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 30

제27항에 있어서,

검색 용어의 단어 확장 버전(word-branched version)을 사용하여 검색을 수행하기 위한 명령어들을 더 포함하는

하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 31

삭제

청구항 32

제27항에 있어서,

검색 어구들의 세트를 얻기 위한 다양한 조합들로 상기 검색 용어들을 그룹화하기 위한 명령어들을 더 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 33

제32항에 있어서,

상기 검색 용어들을 그룹화하기 위한 상기 명령어들을 수행하기 전에 상기 검색 용어들 중 적어도 두개의 검색 용어의 순서를 바꾸기 위한 명령어들을 더 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 34

제32항에 있어서,

상기 속성 목록과 상기 검색 용어들을 비교하는데 더하여, 상기 텍스트 용어들 중 하나 이상을 이용한 EPG 데이터의 대안적 속성 검색을 행하여, 상기 대안적 속성 검색의 결과들에 기초해서 상기 검색 용어들을 상기 텍스트 용어들과 상기 속성 용어들로 더 분리하기 위한 명령어들을 더 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 35

제27항에 있어서,

상기 텍스트 용어들과 속성 용어들을 유도하는 단계는 상기 텍스트 검색 문자열을 구성함에 있어서 사용자의 의도의 해석을 형성하는 단계를 더 포함하는, 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 36

제35항에 있어서,

상기 해석에 따라 검색 결과들을 결합, 순위화, 및 필터링하기 위한 명령어들을 더 포함하는 하나 이상의 컴퓨터 판독가능한 기록 매체.

청구항 37

삭제

청구항 38

삭제

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

[0019] <컴퓨터 프로그램 목록>

[0020] 본 발명과 관련하여 사용된 XML 텍스트 파일 목록의 한 세트가 발명의 상세한 설명 뒤에 제공되며 발명의 상세한 설명에 참조로서 포함된다. 이 XML 텍스트 파일 목록은 예시적인 속성 목록(예컨대, 도 3의

블록(310))이다.

- [0021] 본 발명은 일반적으로 멀티미디어 데이터 통신에 관한 것이고 구체적으로는 전자 프로그램 가이드(EPG: electronic program guide)의 프리 텍스트(free text) 및 속성 검색에 관한 것이다.

발명이 이루고자 하는 기술적 과제

- [0022] 디지털 텔레비전 시스템에서, 사용자는 전형적으로 EPG의 내에서 정보를 보여주는 프로그램을 검색한다. EPG는 예를 들어, 한 주와 같은 시간 간격 동안 사용자의 로컬 헤드엔드(프로그램 제공자의 채널 구성)로부터 이용가능한 프로그램에 관련된 프로그램과 스케줄 메타데이터의 데이터 세트이다. 사용자는 키워드 용어를 지정함으로써 프로그램을 검색하여 프로그램 제목의 텍스트 검색을 수행한다. 검색 결과로, 전형적으로 "적중(hit)"이라고 알려진 검색 판단 기준에 부합하는 프로그램 제목이 나온다.

- [0023] 또한 프로그램 제목의 키워드만이 아니라, 프로그램 속성들을 검색함으로써 텔레비전 프로그램이 발견될 수도 있다. 프로그램 제목은 프로그램의 속성 중 하나로 볼 수 있지만, "속성(attribute)"이라는 용어는 멀티미디어 예술에서 보다 기술적으로 프로그램 제목이 아닌 많은 속성들을 기술하는데 사용된다. 이와 같이, 속성들은 프로그램과 그 표현에 대한 무제한적으로 무수히 많은 특성들을 포함한다. 프로그램과 그 속성들은 전형적으로 EPG 정보의 방대한 계층구조에 맞게 되어 있다. "카테고리"와 같은 일부 속성들은 탐색된 프로그램이 속하는 방대한 프로그램 세트를 기술하는 것으로, 즉, "카테고리" 속성은 탐색되는 프로그램을 포함하는 계층적 레벨을 기술하여, 사용자가 "카테고리" 속성을 검색하면 그 카테고리 내의 모든 프로그램이 적중으로 나오는 결과가 된다. 예를 들어, "서양(westerns)"을 검색하면, 서양 영화 및 텔레비전 장르에 속하는 많은 적중들이 나올 것이다.

- [0024] 속성들은 또한 프로그램 제목에 적합한 특성들을 기술할 수 있는데, 즉, 속성들은 동일 프로그램을 두가지 버전으로 변하게 하는 선택사항을 기술할 수 있다. 이러한 한정 형식의 속성은 특정 프로그램 "아래"의 EPG 계층의 엘리먼트들, 즉, 동일한 프로그램이 다른 느낌을 갖도록 하는 엘리먼트들을 기술한다. 예를 들어, 스펜서 트레이시사가 나오는 "산(The Mountains)"은 "흑백(black & white)" 그리고 "모노(monophonic)"라는 속성을 갖는 한 버전이 있지만, 다른 버전은 "컬러", "돌비 디지털(DOLBY® Digital)", "자막(close captioning)", "보호자 동반 관람가(rated PG)"의 속성들을 갖는 다른 버전도 있다. 제목의 키워드 엘리먼트에 한정된 텍스트 검색은 두 버전 모두가 검색될 것이지만, 제목에 "돌비(Dolby)"에 대한 속성 검색을 함께하는 키워드 엘리먼트들을 텍스트 검색하면 돌비 디지털(DOLBY® digital)이나 돌비 스테레오 사운드 시스템(DOLBY® stereo sound system) 버전들만이 검색될 것이다.

- [0025] 종래의 제목의 키워드 엘리먼트에 대한 텍스트 검색은 키워드 텍스트를 사용자 인터페이스(UI)의 검색 필드에 입력함으로써 수행되며, 종래의 속성 검색은 속성 판단 기준을 입력함으로써(예컨대, 체크박스) 텍스트 검색 키워드와는 별도로 UI의 별도의 부분에서, 혹은 별도의 UI에서 모두 함께 행해진다. 별도의 텍스트 검색에서 나오는 적중과 별도의 속성 검색은 양 검색의 마지막에서만 논리적으로 결합되어 적중 목록을 생성한다.

- [0026] 도 1은 종래의 텍스트 검색(100)과 종래의 속성 검색(102)을 도시하는데, 여기서 텍스트 키워드와 속성 판단 기준은 별도로, 하나 이상의 UI를 통해 디스플레이되는 속성 체크박스들과와 별도의 검색 필드에서와 같이 입력된다.

- [0027] 블록 104에서, 텍스트 키워드(들)이 입력된다. 블록(106)에서, 속성 판단 기준은 키워드(들)의 입력과는 별도의 입력이다. 전형적으로 프로그램 제목에 대한 텍스트 검색이, 입력된 텍스트 키워드(들)를 통해 수행되고(108), 입력된 속성 검색 판단 기준을 이용하여 속성 검색이 수행된다(110). 각각의 별도의 검색은 별도의 결과(112, 114)를 낳는다. 필터링 로직(116)은 이 결과들(112, 114)을 비교하여 텍스트 검색(100)과 속성 검색(102) 모두를 일반적으로 만족시키는 적중 목록(118)을 생성한다.

- [0028] 사용자는 일반적으로 UI의 별도의 부분에서 속성들을 검색할 것이 요구되므로, 많은 사용자들은 그들의 검색 입력을 프로그램 제목 글자만으로 된 키워드 형식의 질의로 제한하며, 어떤 사용자들은 디지털 텔레비전 시스템이 보통의 프로그램 제목 검색 용어들로서 입력된 속성들을 해석한다는 것을 인식하지 않고 프로그램 제목의 텍스트 검색에 대한 키워드로서 속성들을 잘못 입력한다.

발명의 구성 및 작용

- [0029] 본 발명은 EPG 데이터에 대한 검색 엔진과 관련 방법들을 포함한다. 예시적인 방법에서, 텍스트 검색 문자열

(string)은 검색가능한 용어들로 통일될 수 있고, 이 용어들은 텍스트 검색 용어나 속성 검색 용어들로 해석될 수 있다. 복잡성이 변화하는 검색 로직을 이용하는 질의들이 이 검색 문자열 내의 용어들의 해석에 따라 생성된다. EPG 데이터베이스 내 및/또는 EPG 웹 자원들 내의 하나 이상의 검색들이 각 텍스트 및 각 속성 용어에 대해 수행되고 그 결과들에는 검색 문자열의 해석에 따라 관련성 순위가 주어진다. 결합된 검색 결과들은 순위가 매겨지고, 필터링되고, 사용자에게 디스플레이하기 위해 그룹화된다. 결과들은 또한 검색 문자열의 각 텍스트가 사용자에게 의해 입력됨에 따라 점진적으로 디스플레이될 수도 있다.

[0030] 일 실시예에서, 텍스트 용어들을 속성 용어들로부터 구분하는데 속성 목록이 사용될 수 있다. 예시적인 속성 목록은 그 검색 문자열 내의 각 속성에 대해 수행된 검색 과정을 지시하는 각 속성에 대한 메타데이터나 색인을 추가로 포함할 수 있다.

[0031] 개요

[0032] 텔레비전 방송을 수신하고 기록할 수 있는 텔레비전 세트와 개인용 컴퓨터(PC)에 부가된 두 "셋톱 박스" 시스템을 포함하는 디지털 텔레비전 시스템의 말단 사용자들에게는 빠르고, 직관적이고, 정확한 EPG 검색이 중요하다. 예시적인 본 발명은 디지털 텔레비전 시스템에서 사용자가 텍스트와 속성 질의어를 입력하고 해석하기 위한 검색 엔진과 관련 방법들을 기술한다. 이러한 빠르고 정확한 검색 메카니즘은, 예를 들어, 사용자에게 의한 단일 문자열 입력을 해석하고 사용자의 의도를 추론하고, 사용자가 검색하고 있는 것에 대한 해석에 부합하는 적중 결과의 세트를 디스플레이한다. 이러한 향상된 검색 수행 능력은 PC DVR(PC digital video recorder)와 셋톱박스 시장이 급속도로 확장함에 따라 중요하다.

[0033] 도 2는 EPG 데이터베이스나 다른 자원들에서의 예시적인 검색(200)의 개요를 보여준다. 이 예시적인 검색(200)은 사용자에게 의해 입력되거나 사용자로부터 수신되거나 저장장치로부터 검색된 검색 문자열(202)과 함께 시작한다. 예시적인 검색 엔진(204)은 검색 문자열(202)을 수신하고, 그 문자열의 용어에 수차례 변형을 생성하여 다양한 수준의 복잡성을 갖는 가능한 질의 명령어들을 생성한다.

[0034] 단일한 검색 문자열(202)에서 하나 이상의 독립적인 검색들이, 일 실시예에서 검색 문자열(202)의 복잡성과 불명확성 등에 따라 나오고, 문자열 해석기(214)는 특정 검색 문자열(202)에 기초하여 가능한 적은 질의(가능하면 하나)와 가능한 제한적인 질의를 구성하려고 한다. 예를 들어, 일 실시예에서, 문자열 해석기(214)는 이 최소한의 수의 간결한 및/또는 제한적인 질의를 문자열(202)에서 어느 용어가 속성 검색 용어로서 검색될 공지의 속성인지를 판정함으로써 획득한다. 검색 문자열(202)에서 남은 용어들은 가능한 속성 용어들이거나 가능한 텍스트 용어들일 수 있지만, 공지의 속성들을 (예컨대, 공지의 속성 리스트나 데이터베이스를 사용하여) 판정하는 것은 흔히 질의를 상당히 제한시킨다. 따라서, 문자열 해석기(214)는 텍스트 검색(206), 속성 검색(208), 조건 텍스트 검색(210), 및 조건 속성 검색(212)의 시리즈를 이하에 보다 상세히 기술될도 3과 관련하여 다양한 방법들과 자원들을 사용하여 "고려(consider)"한다.

[0035] 일 실시예에서, 검색 문자열(202)의 가능성있는 의미를 강제로 결정하는 것은 EPG 데이터베이스 및/또는 온라인 디렉토리 및 순위 결과 내의 검색 문자열(202)의 각 통일된 용어에 대한 사전 검색을 수행함으로써 달성된다. 예를 들어, "Ryan"이 검색 용어인 경우, 그 데이터베이스나 디렉토리에서 더욱 통상적이면 "Saving Private Ryan"이 "Ryan O'Neal"에 우선하여 가능성 있는 검색 해석으로서 선택될 것이다. 다른 실시예에서, 문자열 해석기(214)는 검색될 EPG 자원 내에서 하나의 조합이 발견되거나 논리적으로 선호될 때까지 검색 용어들의 다른 조합을 시도한다. 그러므로, 검색 용어 "moon"과 "paper"가 무작위적으로 발생하지만 검색 문자열(202) 내의 많은 속성들 중 단 하나의 문자 용어로서, "Paper Moon"이라는 해석이 나올 가능성이 있고, 이는 질의를 통한 적중으로서 실제로 발견되면 높은 순위가 될 것이다.

[0036] 문자열 해석기(214)는 따라서 하나 이상의 해석에 도달하며, 질의(216)(또는 질의들)를 낳는다. 질의(216)는 사용자에게 디스플레이되는 결합, 필터링, 및/또는 순위화된 적중 목록(222)을 낳는 하나 이상의 검색을 수행하는데 사용된다. 이 예시적인 검색(200)은 검색에 민감한 하나 이상의 문자와 속성이 어떻게 단 하나의 검색 문자열(202)로부터 나올 수 있는지의 단 하나의 예이다. 검색 문자열(202)은 예를 들어, 문자열 해석기(214)에 의해 분해, 변형, 분석되어, 검색 문자열(202) 내의 용어들의 문자 및 속성 인식 해석을 얻는다.

[0037] 일 실시예에서, 조합, 필터링, 및/또는 순위화된 적중 목록(222)은 다양한 검색 경로들이 결과들을 반환할 때 갱신된다. 이와 달리, 검색 문자열(202)이 입력됨에 따라 조합, 필터링, 및/또는 순위화된 적중 목록에 결과들이 점진적으로 나타날 수도 있는데, 즉, 새로운 및/또는 변형된 결과가 각 문자가 검색 문자열(202)에 부가됨에 따라 개시된다.

- [0038] 예시적인 검색 엔진
- [0039] 도 3은 도 2의 예시적인 검색 엔진을 상세히 도시한다. 이 예시적인 검색 엔진(204)은 하나 이상의 EPG 데이터베이스(들)(300)과 통신가능하게 결합된다. 예시적인 검색 엔진(204)에 의해 수행된 일부 특징과 프로세스들이 이제 기술될 것이나, 물론 예시적인 검색 엔진(204)은 예를 들어, 데이터베이스 콘텐츠 색인과 같이 기술된 것에 추가되는 특징들을 가지고 프로세스들을 수행할 것이다. 예시적인 검색 엔진(204)의 다양한 실시예들이 물론 다양한 다른 인간 언어들과 함께 동작할 것이다. 따라서, 예시적인 검색 엔진(204)의 일부 특징들은 어떤 언어에는 적용되지 않거나 필요없을 것이다.
- [0040] 보통, 예시적인 검색 엔진(204)은 EPG 데이터베이스 검색(들)을 먼저 API(application program interface)와 같은, 검색 표현들을 수신하고 프로그램으로 혹은 사용자에게 의해 생성된 요건 세트에 접근하는 검색 문자열 수신기(301)를 호출시킴으로써 실행한다. 부분적으로는 요건들에 기초하여 질의들을 생성한 뒤, 예시적인 검색 엔진(204)은 결과 세트를 반환한다.
- [0041] 검색 문자열 수신기(301)는 하나 이상의 UI를 통해 검색 문자열(202)을 수락할 수 있다. EPG 데이터 검색에 대한 UI는 "투풋(two-foot)" 혹은 "텐풋(ten foot)" 시스템에 대해 최적화될 수 있다. 투풋 시스템은 사용자가 PC 시스템에 전형적인 정보 입력 능력을 갖는 디지털 텔레비전 시스템이다. 예를 들어, 투풋 시스템은, 키보드, 마우스, 모니터를 완전히 사용할 것이다. 반면, 텐풋 시스템은, 투풋 시스템과 비교할 때 사용자가 제한된 입력 능력을 갖는 디지털 텔레비전 시스템이다. 텐풋 시스템은 셋톱박스과 동작에 상대적으로 적은 키를 갖는 리모트 콘트롤을 이용하여 입력된 검색 문자열(202)에 의해 생성될 수 있다.
- [0042] UI를 둘러싼 데이터 입력과 검색이 일반적으로 텐풋 시스템에서 더욱 어렵기 때문에, 검색의 정확성과 검색력을 최대화하면서, 텐풋 시스템으로 검색을 입력함에 있어서 사용자의 의도를 포착하는데 필요한 데이터의 양을 최대화하는 것이 유용하다.
- [0043] 검색 문자열 수신기(301)는, 물론, 임의의 수의 입력 필드, 체크박스, 스위치, 및/또는 장치로부터 사용자 입력을 수신할 수 있다. 그러나, 입력이 단지 단일한 검색 문자열(202)이면, 예시적인 검색 엔진(204)은 검색 문자열을 "복호화"하여 문자와 속성 검색 정보를 구별하고 적절한 질의들을 구성할 수 있다.
- [0044] 검색 문자열 수신기(301)에 통신가능하게 결합된 노말라이저(302)는 다양한 질의들에 꾸준히 적용될 수 있는 용어들로 검색 문자열(202) 내의 불규칙적인 것들을 변환한다(즉, "비표준화된(non-standized)" 용어들의 형식을 변환한다). 비표준 입력은 문자, 글자, 언어, 및/또는 검색될 EPG 데이터베이스(들)(300)의 형식과 맞지 않는 구두점 형식을 갖는 검색 문자열(202)의 용어나 일부이다. 비표준 입력은 부분적으로나 전체가 문자열 해석기(304)에 이해될 수 없을 위험을 가지며, 그 형식이 검색되는 EPG 데이터베이스(들)와 맞지 않을 때 적중을 정성화를 빼먹는 결과를 낳을 수 있다.
- [0045] 검색 요청의 일부로서, 검색 문자열 수신기(301)는 보통 사용자에게 의해 입력된 문자열, 예컨대, "The West Wing"를 지나친다. 노말라이저(302)나 예시적인 검색 엔진(204)의 다른 모듈은, 용어 분리를 채용하여 각 단어에 대한 텍스트 검색을 개별적으로 수행하면서 그것을 구성하는 단어들로 구문을 분리하려고 한다. 개별 단어들로 이렇게 분리하면 사용자가 널리 알려진 제목에서 단어들을 생략하거나 재구성할 때 매치되는 것을 찾을 수 있는 이점이 있으나, 사용자가 의도한 것보다 많은 적중을 생성할 수도 있다. 개별적인 단어 검색 용어들은 보통 공백(white space)이나 대시(dash) 경계로 분리된다. 노말라이저(302)는 또한 인용된 문구의 단어들 간에 추가의 공간을 제거할 수도 있다.
- [0046] 사용자들이 항상 정확한 레터케이싱(letter casing)을 사용하여 검색 문자열(202)을 입력하는 것은 아니다. 예를 들어, "The West Wing"이 "the west wing"으로 타이핑될 수 있을 것이다. 또한, EPG 데이터베이스 정보는 종종 "Sunday Ticket"이나 "sunday ticket" 대신 "SUNDAY TICKET"과 같이 문구 내에 비전형적인 레터케이싱을 포함한다. 노말라이저(302)는 각 검색 용어 및 검색될 EPG 데이터베이스 텍스트를 공통적인 것(대문자나 소문자)로 변환함으로써, 예컨대, 현재 규칙에 따라 레터케이스 표준화를 수행할 수 있다.
- [0047] 노말라이저(302)는 또한 어떤 기호나 마크를 표준화한다. 발음 기호는 액센트나 움라우트와 같은 많은 세계 언어의 알파벳에서 나타나는, 보통 모음을 동반하는 결합 문자이다. 본 발명의 일실시예에서, 검색 용어와 검색되는 데이터베이스 문자는 이 발음 기호가 존재하지 않는 것처럼 비교된다. 기호와 특수 문자는 또한 재기입 및/또는 노말라이저(302)에 의해 검색 용어에 더해져 정확한 검색을 용이하게 한다. 예를 들어, 독일어의 "ß"는 "ss"로 재기록되어 독일어 "ß"를 포함하는 검색이 독일어 "ß" 형태나 "ss" 형태가 검색되는 문자에 나타나면 적중을 생성할 것이다.

- [0048] 기호 표준화는 보다 정확한 매칭을 위해 검색 용어(들)와 검색되는 문자 모두로부터 구두점과 같은 기호를 제거하거나 변환하는 것을 포함한다. 예를 들어, "F.B.I"와 같은 약어 내의 점이나 "M*S*S*H" 내의 아스테리스크가 제거된다. "&"와 "@"와 같은 기호들은 그 단어 형태인 "and"과 "at"으로 변환될 것이다. 문자를 대체하고 있는 기호들이 변환될 것인데, 예컨대 "Vega\$"에서 "\$" 표시와 같은 것이다. 이 기호 표준화 결과 추가의 검색을 하게 된다. 예를 들어, 예시적인 검색 엔진(204)은 노말라이저(302)가 검색 문자열(202)을 진행한 뒤, 검색 용어와 종래의 형태에 따라 검색되는 텍스트를 모두 표준화하고 변환함에 의해 단 한가지 형태를 이용하는 대신, "Vegas"와 "Vega\$" 모두를 검색할 것이다.
- [0049] 어디에나 있지만 그다지 서술적이지 않은 단어들과 관사, "a", "the", "is", "are", "el", "la", "los", "les", 등은 노말라이저(302)에 의해 검색 용어들로부터 제거되어 고속 처리될 수 있다. 일부 실시예에서, 관사와 어디든 있는 단어들은 검색에서 존재하는 유일한 단어들이거나 첫번째 검색 후 아무런 결과가 없으면, 제거되지 않거나 조건적으로 검색된다.
- [0050] 예시적인 검색 엔진(204)은 어떤 방식으로 검색 의미를 변화시키는 표시에 우선하여 검색 문자열을 스캔할 것이다. 예를 들어, 마이너스 기호 "-"를 갖는 단어에 우선하는 것은 그 단어가 매칭에 나타나지 않아야 한다는 것을 나타낸다. 용어들을 인용부호로 그룹화하는 것은 그 그룹 내의 용어들이 주어진 순서로, 그리고 함께 나타날 때에만 고려되어야 한다는 것을 나타낼 것이다. "AND", "OR", "NOT"이 검색 연산에 영향을 주는 특별한 방식으로 사용될 수 있다. 이는 사용자가 검색을 보다 잘 조절할 수 있도록 한다.
- [0051] 일실시예에서, 노말라이저(302)나 검색 엔진(204)의 다른 컴포넌트들은 입력되면 검색 문자열(202)에 스펠링 체크를 적용할 것이다. 사용자가 검색 문자열(202)을 입력할 때 실수를 방지하는 것을 돕도록 몇가지 전략이 채용된다. 사용자는 검색이 스펠링 다이어로그 박스나 "wavy-line" INTELLISENSE[®] 하이라이트와 같은 다른 텍스트 표시기를 통해 실행되기 전에 검색 텍스트 내에 잘못 쓰여진 단어가 존재하는 것이 가능한 것으로 수정될 수 있다. 그와 달리, 스펠링 체크를 검색 실행 마지막으로 지연하거나 정확한 결과가 발견되지 않는 경우에만 수행할 수 있다. 마지막으로, 스펠링 체크는 투명하게 행해질 수 있다. 노말라이저(302)에 의해 잘못 쓰여진 것으로 해석된 단어들의 정정된 형태들은 자동적으로 검색 문자열(202)에 첨부되고 조건적 검색으로서 질의될 수 있다. 검색들이 실제로 잘못 쓰여진 검색 용어들을 포함하는 적중을 내면 이러한 조건적 검색 용어들의 순위를 낮추는데 관련성 순위화(이하 기술됨)가 그 다음에 이용될 수 있다.
- [0052] 노말라이저(302)는 또한 스펠링 체크의 단순화된 형태인 자동 정정을 이용할 수 있다. 각 검색 용어를 사전과 비교하는 대신, 단어들이 입력되면 통상적으로 잘못 스펠링되는 단어들의 단순한 록업 테이블을 이용하여 정정된다. 자동 정정이 잘못된 결정을 하는 경우에(사용자가 의도적으로 "잘못쓰는 것"을 의도한 단어를 정정하는 것), 사용자는 백스페이스를 한 번 쳐서 이 자동 정정을 취소할 수 있다. 사용자가 타이핑을 하면, 가능한 단어 완성의 팝업 목록이 사전 존재하는 완전한 텍스트 색인에 기초하여 노말라이저(302)에 의해 제시될 수도 있다. 이는 사용자가 검색 판단 기준을 보다 빨리 입력하도록 한다.
- [0053] 문자열 해석기(304)는 노말라이저(302)와 통신가능하게 결합되어 있다. 일실시예에서, 문자열 해석기(304)는 문자열 변환기(306) 및 속성들(310)의 목록과 통신가능하게 결합되어 있는 속성 식별기(308)를 포함한다. 문자열 해석기(304)는 또한 검색 요건표(312)에도 통신가능하게 결합되어 있다.
- [0054] 문자열 변환기(306)는 검색 문자열(202) 내의 검색 용어들을 단독으로 혹은 조합하여 테스트하여 검색 문자열(202)에 대한 해석을 달성하고 문자열이 재배열되어야 하는지를 결정한다. 문자열 변환기(306)는 재배열 및/또는 혼합 스킴을 사용하여 검색 문자열(202) 내의 용어들의 다양한 조합을 시도하고, 다양한 순서로 사용자가 검색하려고 하는 가능한 속성들을 식별한다. 그러므로, "dolby the closed mile green captioning"은 "dolby", "the Green Mile", 및 "close captioning"을 포함하는 검색 용어 후보들을 만들 것이다. 둘 이상의 용어가 함께 그룹화되는 것으로 보이면, 문자열 변환기는 그 용어들의 순서를 바꾸고, 혹은 그렇게 된 그룹을 검색 문자열(202) 내의 검색 용어들에 투명하게 더할 수 있다. 이를 위하여, 문자열 변환기(306)는 검색 용어들을 속성 목록(310) 및/또는 검색 요건표(312)에 대해 테스트할 수 있다.
- [0055] 문자열 변환기(306)는 검색의 범위를 넓히기 위해 검색 용어를 자를 수 있다. "스테밍(stemming)"은 접두어 및/또는 접미어를 검색 용어에서 제거하여 검색 용어의 어근을 원래의 검색 용어에 근접하게 하는 과정이다. 이는 원래 검색 용어의 변형체를 보다 많이 포착하는 보다 많은 적중을 생성한다. 예를 들어, "divine"이라는 단어는 "divin"으로 잘려 "divinity", "divination", "diviner", "divinity", 등과 매칭될 수 있다. 문자열 변환기(306)는 공지의 스테밍 알고리즘인 "Porter" 알고리즘 등을 사용할 것이다. 스테밍은 검색 용어가 정확히

질의되면 놓칠 수 있었던 적중들을 생성하는 장점이 있지만, 종종 관계없는 적중들을 생성하는 단점도 있다. 이 효과는 관련성 순위화기(relevance ranker)를 사용하여, 문자열 변환기(306)나 결과 컴파일러(이하 기술)와 같은 예시적인 검색 엔진(204)의 다른 부분에서 감소될 수 있다. 일부 실시예에서, 스테밍은 검색 속도를 줄일 수 있는 데이터 처리 단계를 요구한다. 그러므로, 일부 실시예에서, 스테밍은 특별한 단어들, 예를 들어, 발생 빈도 색인(frequency of occurrence index)에 의해 결정되는 드물게 나타나는 단어들에 제한된다.

[0056] 문자열 변환기(306)는 추가의 용어 및 문자열 변환을 수행할 수 있다. 단어 확장(word branching)은 단일 용어로부터 나오는 한 세트의 검색 용어들을 생성하는 것을 포함한다. 관련된 단어들의 목록이 사용될 수 있다. 예를 들어, 검색 용어 "NFL"은 용어 "football"을 검색 문자열(202)에 첨부할 수 있고 따라서 "NFL"은 "football"에 대한 추가의 검색을 낳을 것이다. 용어 "30"이 검색 문자열(202) 등에 나오면 용어 "thirty"가 검색 문자열(202)에 추가될 수 있다. 검색되는 데이터와 관련된 예외 목록에서, 또는 사용자 정의된 혹은 프로그램으로 결정된 요건표(312)에 저장된 예외 목록로부터 관련된 용어들이 나올 수 있다. 단어 확장은 또한 검색을 늦출 수 있고 관련되지 않은 결과들을 생성할 수 있지만, 아무런 것도 발견되지 않는 어떤 경우에도 적중을 생성할 수 있다. 관련성 순위화기는 덜 관련성 있는 결과들을 적중 목록(222)의 맨 아래에 두는데 사용될 수 있다.

[0057] 또한 복수화가 사용될 수 있다. 복수화는 단수 또는 복수 형태의 원래의 단어에 기초하여 추가의 검색 용어들을 생성하는 단어 스테밍의 간단한 형태이다. 예를 들어, 검색 용어가 "baby"이면, 추가 검색 용어로 "babies"가 포함될 수 있다.

[0058] 선택적으로, 문자열 해석기(304)나 검색 엔진(204)의 다른 컴포넌트는 다른 문자열 해석 특징과 프로세스에 자연어 분석(natural language parsing)을 추가할 수 있다. 이러한 형태의 분석은 "When is Star Wars shown in Dolby-Digital?"과 같은 검색 어구를 사용자의 검색 요청에 부응하는 작업가능한 검색 용어들로 변환시키는 것을 포함한다. 자연어 분석은 매우 강력한 것이지만, 어법이 일반적으로 사용자의 명사/동사 의미의 선택에 제한된다. 또한, 자연어 검색 문자열(202)에서 단어의 수가 증가되면(불리안 용어 검색에 대해), 자연어 검색은 검색 용어들이 원격 조정 키패드를 사용하여 입력될 필요가 있는 텐넷 시스템에 대한 최적의 선택이 되지 않을 수 있다. 마지막으로, 자연어 분석은 보다 많은 실행 처리 능력을 요구할 수 있다.

[0059] 문자열 변환기(306) 및 예시적인 검색 엔진(204)의 다른 모듈에 의해 수행된 특징과 절차들이 어떠한 순서로 기술되었으나, 이러한 특징들과 절차들이 제시된 순서는 프로세싱 파이프라인을 통해 제시된 순서를 나타내는 것을 의미하는 것은 아니다. 많은 경우, 이러한 특징들 및 프로세스들 각각의 순서 및/또는 성능은 생성된 결과에 중요하다.

[0060] 속성 식별기(308)는 텍스트 검색 용어들을 속성 검색 용어들로부터 다양한 방법, 예를 들어, 속성 목록(310) 내의 속성들에 대한 후보나 잠재적 검색 용어들 비교에 의해 분리시킨다. 속성 식별기(308)는 또한 텍스트 검색 결과로부터 피드백을 받고 검색 문자열(202) 내의 용어들에 제거 프로세스를 적용하여 속성들을 찾아낸다. 이 경우, 속성 식별기(308)는 검색 용어들 간에 다양한 불리안 관계를 가정할 것이다. 예를 들어, 검색 용어 "green mile dolby"는 "'green' OR 'mile' AND dolby" , 혹은 "'green mile' AND dolby"로 해석될 수 있지만, 검색 결과는 두번째 가정된 세트의 불리안 관계로 인한 단지 DOLBY®

digital 혹은 DOLBY®

Stereo sound 내의 "The Green Mile"의 경우만을 낳을 것이다.

[0061] 속성 목록(310)은 물론, 고속 문자열 해석을 위해 컴퓨터 메모리에 저장될 것이다. 약 200개의 텔레비전 속성들의 목록이 속성 목록(310)에 사용되면, 사용자에게 의해 입력될 텔레비전 프로그램에 관련된 모든 속성들이 가상적으로 속성 식별기(308) 및 분리된 텍스트 검색 용어들에 의해 식별될 수 있다.

[0062] 예시적인 속성 목록(310)의 표현으로 부록 A:"예시적인 속성 목록"에 있는 예시적인 XML(extensible markup language) 데이터 구조가 있다. 이 표현은 텔레비전 프로그래밍 속성들, 엘리먼트들, 및 값들을 단순히 선택하는 것만을 포함한다. 따라서, 전형적인 속성 목록(310)은 도시된 것보다 적거나 많은 엔트리 및/또는 다른 엘리먼트들, 속성들 및/또는 값들을 포함할 수 있다.

[0063] 부록 A의 예시적인 속성 목록(310)의 XML 버전에 의해 기술된 바와 같이, 속성 목록(310)은 속성 검색 용어 후보와 함께 제시될 때 참/거짓 또는 "존재/부존재"만을 제공하는 연속적인 속성 엘리먼트들의 배열에 제한될 필요는 없다. 속성 목록(310)은 적어도 일부의 속성들에 대한 메타데이터(311)를 포함할 것이고, 여기서 한 속성

의 메타데이터는 자원들을 검색하기 위한 속성 및 검색에 사용되는 값들에 관련된 것이다. 메타데이터(311)는 또한 요건표(312) 내의 논리적 검색 조건들의 세트에 관련될 것이다. 또한, 속성 목록(310)은 속성 목록(310) 내의 속성을 다른 관련된 속성들 및 합당하게 검색되어 그 검색 문자열(202)에 부합하는 적중을 생성하는 값들과 관련시키는 색인(들)(313)을 포함할 것이다.

[0064] 일부 실시예에서, 색인(들)(313) 및/또는 메타데이터(313)는 상대적으로 복잡한 브랜칭 로직 및/또는 검색 문자열(202)을 구축하는데 사용자의 의도에 부합하는 적중을 생성할 논리적 검색 조건들의 트리 구조를 포함할 것이다. 예를 들어, 속성 식별기(308)가 검색 용어 "stereo"에 대한 속성(310) 목록을 체크하면, 그 속성의 적절한 이름 지정 및/또는 프로그램의 스테레오 속성들을 포함하는 EPG 정보 내의 검색 필드 지정함에 있어서 매치가 발견될 것으로, 예를 들어, "stereo" 프로그램이 데이터베이스 기록의 "ScheduleEntryAttribute" 필드를 검색함으로써 발견될 수 있을 것이다. 또한, 후보 검색 용어가 속성 판단 기준에 맞는다는 것을 발견하고, EPG 데이터 내에 지정된 검색 필드가 있으면, 속성 목록(310)은 또한 하나 이상의 정확한 값들, 예를 들어, 스테레오 프로그램을 발견할(혹은 발견될) 검색 용어, 키, 또는 카테고리를 나타내는 "MSEPGA_STEREO"를 가져올 것이다.

[0065] 속성 목록(310)의 색인 능력은 다중 확장될 수 있다. 예를 들어, 속성 식별기(308)가 후보 검색 용어 "sports"에 대하여 속성 목록(310)을 체크하면, 속성 목록(310)은 "sports"를 속성으로서 식별하고, 그것을 또 "카테고리" 형의 속성으로 식별하고, 그 속성 식별자(308)가 상대적인 프로그램들에 대한 세가지 다른 카테고리, "news", "specials", 및 "MSEPGC_SPORTS"를 크로스 체크하도록 한다.

[0066] 문자열 변환기(306)에 대하여 이상 기술한 특징들과 프로세스들을 전개하고 속성들과 관련된 메타데이터를 속성 목록(310)을 통해 식별함에 의해, 문자열 해석기(304)는 검색 문자열(202)을 구축함에 있어서 사용자의 의도에 접근하려고 하는 검색 용어들의 세트를 얻는다. 다시 말해, 문자열 해석기(304)는 대부분의 사용자들이 검색 문자열(202)을 만들어냈으면 찾고 있는 것과 맞는 적중들을 반환할 확률이 매우 높은 텍스트 검색 용어들과 속성 검색 용어들의 세트를 얻는다. 예를 들어, 사용자가 검색 문자열 "scifi dolby letterbox"를 만들었다면, 예시적인 검색 엔진(204)은 letterbox에 DOLBY® Stereo sound라고 보이는 science fiction의 적중 목록(222)을 반환할 것이다. 사용자가 검색 문자열 "primetime network comedies"를 만든다면, 예시적인 검색 엔진(204)은 네트워크 스테이션 상에서 보여지는 primetime comedy의 적중 목록(222)을 반환할 것이다. 사용자가 검색 문자열 "drama rated pg"를 만든다면, 예시적인 검색 엔진(204)은 PG 등급(PG rating)을 갖는 드라마의 적중 목록(222)을 반환할 것이다. 사용자가 검색 문자열 "Discovery"를 만든다면, 예시적인 검색 엔진(204)은 제목에 "discovery"가 있는 방송과 채널 명에 "discovery"를 갖는, 예컨대 "DISCOVERY CHANNEL®"의 적중 목록(222)을 반환할 것이다.

[0067] 질의 생성기(314)는 문자열 해석기(304)와 통신가능하게 결합되어 있다. 질의 생성기(314)는 텍스트 질의 형성기(316) 및 속성 질의 형성기(318)를 포함할 것이다. 일 실시예에서, 텍스트 질의 형성기(316) 및 속성 질의 형성기(318)은 EPG 데이터베이스(들)(300) 및 웹기반 EPG 자원(334)에서 가능한 작고(흔히 하나) 가능한 제한적인 질의나 질의들을 구성한다. 다수의 질의가 구성되면, 결과들이 결합될 수 있다.

[0068] 문자열 해석기(304)로부터 수신된 각 검색 용어에 대해, 질의 생성기(314)는 매치가 EPG 데이터베이스(들)(300) 및 웹기반 EPG 자원(334)의 질의 동안 발견되는 방식으로 영향을 주는 연산자에 적용할 수 있다. 일 실시예에서, 각 연산자는 검색 용어가 검색되는 텍스트 내에서 발견되어야 하는(혹은 발견되지 않아야 하는) 위치나 구성을 특징한다. 몇몇 예시적인 연산자들이 이제 기술될 것이다. "정확한(exact)" 연산자는 검색 용어가 주어진 검색 텍스트의 위치에 정확히 맞아야 한다고, 즉, 질의 생성기(314)가 일정한 형태의 필드를 검색하면, 검색 용어가 그 필드에 정확히 맞아야 한다고 특징한다. "시작(절)(starts-with(phrase))" 연산자는 검색 용어가 검색될 텍스트 필드의 개시부에서 발견되어야 한다고 특징한다. "시작(단어)(starts-with(words))" 연산자는 검색 용어가 검색될 텍스트 필드 내의 단어의 시작이어야 한다고 특징한다. "포함(contains)" 연산자는 검색 용어가 검색될 텍스트 필드 어디에서든 발견될 수 있다고 특징한다.

[0069] 상기 연산자들 각각에 대해, "부정 연산자(counter-operator)"(NOT 형태)는 각 조건이 참이면 검색 엔진(204)에 의해 적중이 생성되어서는 안된다고 특징하는데 사용될 수 있다. 검색 연산자 "AND"와 "OR"은 또한 모든 검색 용어들이 검색될 텍스트 필드에서 발견되어야 하는지 혹은 검색 용어들의 일부만 매치되어도 되는지를 특징하는데도 사용된다.

[0070] 데이터 관리기(319)는 질의 생성기(314)에 의해 생성되는 다양한 검색 용어들을 이용하여 EPG 데이터베이스(들)(300) 내에서 검색을 수행한다. 본 발명의 일부 실시예에서, 데이터 관리기(319)는 SQL(structured query

language) 데이터베이스 엔진이나 JET 데이터베이스 엔진이 되거나 이를 채용할 것이며, 만일 그렇다면, 결과 컴파일러(320)의 일부 기능들을 결합할 것이다. 그러므로, 데이터 관리기(319)는 내부적으로 다수의 검색 경로를 구동하고 그들을 결합할 것이다. 따라서, EPG 데이터베이스(들)(300)은 다수의 다른 형식들(예를 들어, XML, 바이너리, JET, SQL, 등), 일부 저장장치에 보다 적합한 것들과 일부 네트워크를 통한 전송에 보다 적합한 것들을 가정한다. 전송 매체의 요건에 따라, 최적의 전송 형식은 최적의 저장장치 및/또는 검색 형식과 다를 수 있다. 몇가지의 이러한 형식 선택들과 그들의 상대적 강점과 약점이 이제 기술될 것이다.

[0071] 일부 가이드 서비스는 정해진 스키마에 맞도록 형식화된 XML을 사용하여 클라이언트에게 멀티데이(multi-day) 가이드를 송신한다. XML은 이해하기 쉽고, 인간이 판독가능하고, 다수의 기존의 기능을 가지기 때문에 작업하기 편리한 형식이다. 그러나, XML은 매우 간단한 형식이다. 이러한 이유로, 송신되기 전에 eHome 가이드는 CAB 파일로 압축되고 그 뒤 수신이 완료되면 압축해제된다.

[0072] 큰 XML 파일들은 쉽게 검색되지 않는다. 예를 들어, 10일분의 EPG 데이터는 6.6 메가바이트 파일을 채울 수 있다. 이러한 크기의 XML 파일을 DOM(document object model)로 로딩하는 것은 거의 90 메가바이트의 RAM을 요구한다. 이러한 큰 파일을 각 검색동안 스캐닝하는 것은 실용적이지 못하다. 이러한 이유로, XML 형식은 전송과 영구 저장에 주로 사용될 것이다. 많은 클라이언트들이 EPG 데이터 파일을 로드할 때, 텍스트가 아닌 데이터가 메모리에 저장되고 모든 텍스트 기반의 데이터는 검색을 위한 메모리 매핑된 파일에 기록된다. 일부 EPG 데이터 API들은 데이터 및/또는 검색 패턴 상의 변화를 수용하기 위해 상용 데이터베이스를 채용해야 할 것이다.

[0073] EPG 데이터베이스(들)(300)에 대한 다른 선택은 바이너리 형식이다. 바이너리 형식은 보통 매우 작고, 일반적으로 생성 및 브라우징에는 상용 도구들에 의존한다. XML 파일과 같이, 바이너리 파일들은 보통 검색에 최적이지 않고 이 목적을 위해 대안적인 중간 저장장치 형식을 요구할 수 있다.

[0074] 검색 엔진(204)은 또한 EPG 데이터베이스(들)(300)에 질의를 하는데에 데이터베이스 도구들이 잘 지원되는 제트 데이터베이스 엔진을 사용한다. MDAC라고 하는 재분배가능한 API는 SQL, ADO, 혹은 OLE DB를 사용하여 제트 데이터베이스를 생성하고 질의하는데 풍부한 API들을 제공한다. EPG 데이터베이스(들)(300)에 제트 데이터베이스를 사용하는 한가지 장점은 산업 표준 SQL 커맨드를 사용하여 쉽게 검색을 하는 것과 통상의 소프트웨어를 사용하여 쉽게 보기를 하는 것이다. 제트(Jet)를 사용하는 검색의 성능은 색인이 적절히 생성되고 저장되는 한 일반적으로 많은 다른 방법들보다 빠르다(한가지 예외는 SQL 서버). 제트 데이터베이스는 또한 전체 EPG 데이터베이스(들)(300)의 조작을 용이하게 하는 단일 파일에도 역시 저장된다.

[0075] 제트 데이터베이스 엔진과 같이, SQL은 도구들과 API들의 큰 라이브러리를 갖는 EPG 데이터베이스(들)(300)에 대해 잘 지원된 데이터베이스 플랫폼이다. 현재 공지된 버전의 SQL 서버는 단지 서버 기계에서만 구동될지라도, MSDE(혹은 SQL Desktop Edition)라고 하는 SQL 서버의 한 버전은 많은 현재의 동작 시스템 소프트웨어 버전에서 구동한다. 잘 색인된 SQL 데이터베이스를 검색하는 것은 아마도 다른 공지된 데이터 검색 기술보다 보다 빠르고 유연성있는 것이다. 그러나, SQL 데이터베이스는 네트워크를 통하여 쉽게 전송되지 않는다. EPG 데이터를 다른 형식으로부터 SQL 데이터베이스로 로딩한 후의 단계는 MSDE를 사용할 것이 요구된다. 또한, 서버 버전과 달리, (양자가 다른 방법들을 이용하여 에뮬레이트될 수 있을지라도) MSDE는 항상 자연어 검색이나 풀-텍스트 인덱싱(full-text indexing)을 지원하는 것은 아니다.

[0076] 많은 경우, EPG 데이터베이스(들)(300)의 원래의 형식은 예시적인 검색 엔진(204)에 의해 요구되는 최종 형식이 아니라는 것을 알아야 한다. 예를 들어, 데이터 색인이 생성될 필요가 있거나 데이터는 검색에 최적이나 네트워크를 통해 전송하기 어려운 다른 데이터베이스로 판독될 필요가 있을 수 있다. 일실시예에서, 예시적인 검색 엔진(204)은 원치 않는 데이터를 버리고 고속 검색을 위해 사용자가 특정한 데이터를 통합시키기도 한다.

[0077] 데이터 색인이 사용되면, 검색 엔진(204)은 풀-텍스트(full-text) 색인을 구현할 것이다. 이는 텍스트 기반 EPG 데이터의 각 엘리먼트로 역 단어 색인을 생성하는 것을 포함한다. 풀-텍스트 색인은 큰 EPG 데이터베이스(들)(300)과 색인 생성에 증가되는 시간을 희생하여 검색의 성능을 상당히 증가시킨다.

[0078] 검색 엔진(204)은 또한 서명 색인을 사용한다. 풀-텍스트 색인과 같이, 서명 색인은 가능한 검색 매치를 빠르게 식별하는 방법이다. EPG 데이터베이스(들)(300)에서, 각 데이터 로우(프로그램, 채널 등)는 어느 글자들이 어느 단어에서 나타나는지를 나타내는 해시와 유사한 바이너리 서명을 갖는다. 검색 용어와 검색될 텍스트의 단어에 대한 이러한 유사한 계산을 수행한 뒤, 그 결과들을 바이너리 산술법을 사용하여 매칭하여, 그 단어가 그 검색 용어에 맞는지 빠른 추측을 가능하게 한다. 제2의 보다 프로세서 위주의 검색이 그 다음 로우가 매칭

된 단어를 포함하는 로우가 적중을 생성해야 하는지를 결정하는데 필요할 것이다.

- [0079] 이상 기술한 바와 같이, 본 발명은 텍스트 검색 용어들과 속성 검색 용어들을 단일한 검색 문자열(202)로부터 유도하는데 검색 엔진(204)의 능력에 제한되는 것이 아니다. 예시적인 검색 엔진(204)은 텍스트 검색 용어들과 속성 검색 용어들을 단일 검색 문자열(202)로부터 유도하는 것에 더하여 검색 용어들을 UI 내의 다양한 엔트리 필드와 같은 다수의 소스로부터 유도할 것이다. 예를 들어, 어떤 제품에 있어서, "stereo", "closed captioning", "start time", 및 "year of release"와 같은 속성들이 흔히 검색 용어들로서 이용가능하지만 일부 UI에는 체크박스, 라디오 버튼 등을 통하여 입력된다. 이들은 또한 검색 용어들로서 검색 API에 통과됨으로써 예시적인 검색 엔진(204)에 의해 수용될 수도 있다. 이러한 수용은 검색 엔진(204)이 다른 속성 검색 용어들과 텍스트 검색 용어들을 단일한 검색 문자열(202)로부터 유도하는 것을 제한하지 않는다.
- [0080] 예시적인 검색 엔진(204)은 사용자가 검색하기를 원하는 필드를 선택하도록 할 수 있다. 그래서 모든 도메인(예를 들어, 제목, 설명, 채널, 캐스팅, 등)을 자동으로 검색하고 그들의 예견된 관련성에 의해 적중들을 순위화함으로써, 예시적인 검색 엔진(204)은 또한 사용자가 수동으로 필드 및/또는 검색 도메인을 좁히도록 할 수 있다.
- [0081] 예시적인 검색 엔진(204)은 또한 데이터 관리기(319)와 문자열 해석기(304)와 통신가능하게 결합된 결과 컴파일러(320)를 포함한다. 결과 컴파일러(320) 내의 적중 수신기(322)는 다수의 데이터베이스에서 검색하는 다양한 검색 양식이나 단일한 검색 양식으로부터 적중들을 수집한다. 다수의 검색이 동시에 일어날 수 있으므로, 적중 수신기는 임시적으로 계류중인 적중들을 저장하고 컴파일링하는 적중들을 저장할 수 있는 다수의 혹은 하나의 버퍼를 수신하는 것을 포함한다.
- [0082] 예시적인 검색 엔진(204)에 있어서, 검색 비교기/필터(324)는 적중 수신기(322) 및 문자열 해석기(304)와 통신가능하게 결합된다. 검색 비교기/필터(324)는 문자열 해석기(304)에 의해 셋업된 문자열 해석기(304)에 대해 적중들을 테스트한다. 이상 기술한 바와 같이, 문자열 해석기(304)는 문자열 변환기(306)와 속성 식별기(308)을 통해 검색 문자열(202)을 구성함에 있어서 사용자의 의도의 작용 가설, 혹은 해석을 얻는다. 결과 컴파일러(320)에서, 적중이 해석과 맞지 않으면, 적중은 필터링되고, 즉, 디스플레이된 적중 목록(222)로부터 제거되거나 디스플레이된 적중 목록(222) 내에서 순위화를 위해 낮은 순위로 가게 된다. 일실시예에서, 검색 비교기/필터(324)는 결과 세트를 하나씩 스캔하여, 각 결과가 정말로 적중인지를 판정한다. 이는 질의가 모든 적중들을 포함하면서도 모든 미스들을 배제하기에 충분한 미세한 레벨로 분해될 수 없으면 더욱 그러하다. 해석의 예로서, 원래의 검색 문자열(202)이 "duck dolby"이고 문자열 해석기(304)가 "Howard the Duck, by Thomas Dolby"라는 해석을 "Donald Duck in Dolby Digital"이라는 해석보다 선호하면, Donald Duck이라는 적중이 적중 목록(222)나 적어도 주어진 낮은 순위에서 완전히 제거될 것이다.
- [0083] 결과 순위화기(326)는 디스플레이에 대한 적중을 관련성이나 다른 판단 기준에 의해 디스플레이하기 위해, 모든 검색을 마친 때, 혹은 일부 실시예에서는, 결과들이 적중 수신기(322)에 의해 조금씩 수신되고 검색 비교기/필터(324)를 통해 통과해 내려갈 때, 정리할 수 있다. 검색 엔진(204)에 의해 생성된 적중은 종종 순위화되어 적중 결과가 가장 가능성 있는 결과가 제일 먼저 나오는 방식으로 디스플레이되도록, 즉, 검색 문자열(202)을 만듦에 있어서 사용자의 의도에 가장 가깝게 해석하는 적중들이 디스플레이되도록(문자열 해석기(304)에 의해 추측된 바대로) 정렬될 수 있다. 각 검색 용어의 정확한 매치는 예를 들어, 100%의 관련성 순위 및 주어진 최고 우선순위에 제공될 수 있다. 스테밍이나 스펠링 수정과 같은 어떠한 변형 뒤에 검색될 텍스트에만 매칭하는 적중들은 보통 낮은 관련성으로 주어진다.
- [0084] 포함 필터(inclusion filter)(328)는 결과 순위화기(326)에 통신가능하게 결합된다. 다른 변형예들에서는 포함 필터(328)가 사용자에 의해 조정될 수 있는 반면, 일부 변형예에서 포함 필터(328)는 일정한 방식으로 프로그램으로 설정될 수 있다. 예를 들어, 포함 필터(328)는 적어도 하나의 100% 관련 적중이 발견될 때 100% 미만의 관련성을 갖는 적중 결과를 완전히 배제하도록 조정되거나 설정될 수 있다. 어떤 경우에는, 포함 필터(328)가 긴 결과 목록을 잘라, 일부 결과들이 100% 관련성을 갖고 있을지라도 일정한 수의 최고 결과들만을 디스플레이하도록 설정되거나 조정될 수 있다. 다른 경우에서, 포함 필터(328)는 일정한 백분을 이상으로 할당된 순위를 갖는 적중 결과들만을 디스플레이하도록 할 수 있다. 그러므로, 만일 결과가 "Where the Red Fern Grows"이고 유일한 적중이 관련성 15%의 관련성 순위를 갖는 "The Red Skelton Hour"이면, 포함 필터(328)는 "결과 없음(no results found)"을 디스플레이할 것이다.
- [0085] 포함 필터(328)는 또한 결과의 개수를 정해진 값으로 제한할 것이며, 혹은 결과들을 시청 연령 제한(parental control) 및/또는 사용자 검색 수행에 따라 제한할 수 있다. 기록의 관점에서, 포함 필터(328)는 장래에 보이

지 않을 결과들을 제거할 것이다.

[0086] 디스플레이 갱신기(330)가 결과 컴파일러(320)에 포함되어 적중 목록(222)이 사용자에게 제공되는 디스플레이(332)를 제어한다. 검색 결과들의 적중 목록(222)을 제공하는 데에는 많은 방법들이 가능하다. 적중들은 알파벳순으로 혹은 시간순으로 디스플레이될 수 있다. 또한 결과들이 그룹화될 수 있고 관련성에 의해 순위화될 수 있다. 일부 실시예에서 디스플레이 갱신기(330)는 결과들의 집배소(clearinghouse)가 될 수 있는데, 즉, 디스플레이 갱신기(330)는 디스플레이될 것의 항목을 보유할 수 있다. 적중 결과들이 단편적으로 수신되면, 디스플레이 갱신기(330)는 오래된 결과들을 새로운 결과들과 결합시키고 어느 결과들이 디스플레이될 지를 결정하는 통합 기능을 수행할 것이다. 단편적인 검색 동안, 디스플레이 갱신기(330)는 디스플레이를 나누어 검색 결과들이 수신되는 순서를 보여주거나 또한 디스플레이를 나누어 어느 적중들이 어느 검색 용어들에 관련되는지를 보여줄 것이다. 점진적 검색(이하 보다 상세히 설명) 동안, 디스플레이 갱신기(330)는 디스플레이된 결과들의 모두 혹은 대부분을 사용자가 새로운 영숫자 문자를 검색 문자열(202)에 추가(예컨대, 타이핑 및/또는 입력)할 때마다 버릴 것이다. 따라서, 점진적 검색에서, 디스플레이 갱신기(330)는 입력되는 검색 문자열(202)의 의미가 전개됨에 따라 버린 것으로부터 검색 결과들을 재구성한다.

[0087] 물론, 일부 실시예들에서는, 결과 컴파일러(320)의 다양한 모듈들이 함께 결합되거나 달리 정렬될 수 있다. 변형예에서는 일부 모듈들이 제거되거나 다른 종래의 데이터베이스 모듈들이 추가될 수 있다.

[0088] 예시적인 검색 엔진(204)은 실시예에서 사용되는 가용한 하드웨어 및/또는 소프트웨어 환경의 속도와 처리력에 맞도록 구성될 수 있다. 검색은 사용자들이 수용할 수 있는 것보다 실행에 시간이 더 걸릴 수 있다. 검색 엔진(204)은 검색(들)이 여전히 수행되는 동안 발견되면 적중들을 비동기식으로 반환할 것이다. 달리, 예시적인 검색 엔진(204)은 결과가 완료될 때까지 결과 세트를 주기적으로 갱신하면서 지능적으로 이 검색을 부분들로 나눌 수 있다. 예를 들어, 스테밍 및 표준화 기능들이 적용되고 뒤이은 과정에서 반환되기 전에 정확한 단어 매치를 위한 텍스트 검색이 수행될(그리고 결과들이 반환될) 수 있다. 검색들은 또한 검색 프로그램들이 계획된 방송 시간에 따라 나누어질 수 있다.

[0089] 다른 방식으로, 추가의 데이터 저장기 예시적인 검색 엔진(204)에서 가능하다면, 검색 도메인이 확대될 수 있다. 검색 엔진(204)은 각 데이터 저장장치를 통해 검색을 수행하면서, 최종 결과 세트를 만들기 전에 결과들을 통합할 수 있다. 예를 들어, 이는 로컬 EPG 데이터베이스(들)(300) 검색을 웹기반 EPG 자원(334) 검색과 결합시키는데 사용할 수 있다. 또한, 로컬 매체에 사전 기록된 프로그램들이 또한 검색 도메인에 추가될 수 있다. 이러한 검색들은 다수의 API를 통해 수행될 수 있고 디스플레이되기 전에 통합된다.

[0090] 예시적인 검색 엔진(204)은 또한 이후 및/또는 배경 실행을 위한 검색 판단 기준을 지속할 수도 있다. 사용자들은 검색 문자열(202)을 재기록하지 않고 검색을 만드는데 관심이 있을 수 있다. 또한 사용자가 이미 검토하여, 일부 실시예에서는, 일부 결과들이 숨겨지고, 분리되거나, 보다 쉽게 읽기 위해 "비활성화(grayed out)"될 수 있도록 하는 적중들을 유지하는 것도 유용하다. EPG 데이터베이스(들)(300) 정보가 갱신되어 사용자가 관심 있는 장래의 이벤트들을 공지받을 수 있도록 할 때마다 지속적인 검색이 작동할 수 있다.

[0091] 검색 판단 기준은, 예를 들어, 매크로(macro)로 연결될(serialized) 수 있고, 자동 기록, 즉, 검색 히스토리 목록을 정의하기 위한 연결 기구로서 기능한다. 검색 용어들의 각 연결에 토큰(token)이 또한 할당되어, 사용자가 이 토큰을 선택함으로써 연결을 선택할 수 있다. 연결된 검색은 이메일, 웹 링크를 통해, 및/또는 TV 방송에 "click to record" 링크로 포함되어 다른 사용자들에게 교환될 수 있다. 예시적인 검색 엔진(204)은 또한 하나 이상의 연결 형식을 정의할 수 있다. 검색 및/또는 연결은 바이너리 형태로 소형 저장장치에 저장되거나 이메일을 통해 쉽게 전송하기 위해 ASCII 형태로 보내어질 수 있다. 이 연결(serialization)은 단독으로 혹은 자동 기록과 결합하여 사용될 수 있으며, 여기서 사용자는 검색을 정의하고 검색 판단 기준과 매칭하는 장래의 보여지는 모든 것(복제의 경우를 제외)이 기록될 것이라고 나타낸다.

[0092] 예시적인 검색

[0093] 도 4는 예시적인 검색 엔진(204)과 같은 검색 엔진에 의해 동반될 수 있는 본 발명에 따른 예시적인 검색(400)을 도시한다. "dolBy WEST-rated 'g'"와 같이 상대적으로 불명확한 검색 문자열(202)이 사용자나 영구 저장장치로부터 예시적인 검색 엔진(204)를 통해 수신된다. 이 검색 문자열(202)은 단일한 간결한 질의에 제한되지 않는다. 예시적인 검색 엔진(204)은, 그러나, 가능성 있는 적중들을 찾기 위해 다수의 검색들을 수행하는데 큰 융통성을 갖는다. 예시적인 노말라이저(302)는 문자열을 예시적인 표준 형태, 즉, 표준 검색 문자열(402): "DOLBY WEST RATED G"로 변환한다. 이 경우에, 노말라이저(302)는 "dolBy"의 일부 글자를 변경시켜서 EPG 데

이타베이스(들)(300) 등에서 검색될 텍스트의 것과 매칭될 텍스트 검색 용어를 얻는다. 노말라이저(302)는 단어들 간에 공백과 "WEST"와 "rates"간에 경계로서 대시를 사용하여 검색 용어들을 구분시킨다. 노말라이저(302)는 "g"를 둘러싼 따옴표를 제거하고 전체 검색 용어의 문자열을 한가지 경우, 즉, 대문자로 놓는다.

[0094] 표준화된 검색 문자열(402)은 검색 문자열(202)을 구성함에 있어서 명백히 결론적인 사용자의 의도의 해석에 도달할 예시적인 문자열 해석기(304)로 전달된다. 사용자에게 의해 입력된 속성 검색 용어들이 텍스트 검색 용어들로부터, 예를 들어, 속성 목록(310)를 통해 명백히 결론적인 해석이 명확히 기술될 수 있다. 문자열 해석기(304)는 텍스트와 속성 용어들 간에 결론적인 분리를 달성할 수 없을 것이나, 이 경우, 문자열 해석기(304)는 검색 용어들을 대체 및/또는 조건 검색을 위한 질의 생성기(314)에 검색 용어들을 전달할 것이다. 즉, 어떤 용어에 대해 모호성이 있으면 그 용어가 텍스트 검색 및 속성 검색에 사용되어, 예컨대, 얼마나 많은 적중이 각 검색 형태에서 나오는지 알아본다. 그러므로, 검색 용어 "DOLBY"는 "The Dolby Digital Experience"와 같이 프로그램 제목 내의 한 단어가 될 수도 있고, "Thomas Dolby"와 같이 프로그램 생성자 속성일 수도 있고, "DOLBY®

digital"과 같이 프로그램 정성적 속성이 될 수도 있다. 예시적인 질의 생성기(314)는 이러한 가능성들 각각을 두개의 속성 검색(406)(DOLBY® digital and DOLBY® Stereo sound)으로서, 대안적인 텍스트 검색(410)(프로그램 제목에서 "dolby"라는 단어)으로서, 그리고, 대안적인 속성 검색("Dolby"라는 이름을 갖는 캐스팅과 직원)으로서의 3가지 가능성들 각각을 테스트한다. 속성으로서의 "DOLBY"가 관련성 순위가 매우 높기 때문에 질의 생성기(314)는 주요 텍스트 검색(406)으로서 "DOLBY"를 검색하지 않을 것이라고 결정한다. 그러나, 만일 "DOLBY"가 텍스트 검색에서 발견될 수 있으면, 이는 다른 텍스트 검색(410)에 의해서 발견될 수 있을 것이다. 따라서, 다른 혹은 조건적 검색들이 검색에 의해 높은 우선 순위를 갖는 적중들이 없을 경우를 대비하여 사용될 수 있을 것이며, 또는 다른 검색들(410, 412)이 검색 문자열(202)에 대한 더 많은 정보를 얻기 위한, 즉, 검색 문자열(202)의 해석을 강화하거나 새로운 검색 결과들을 찾기 위한 첩자(feeler)나 변경자로서 사용될 수 있다.

[0095] 표준화된 검색 문자열(402)로부터의 검색 용어 "WEST"는 또한 다른 속성 검색(412)에서도 사용될 수 있는데, "WEST"는 프로그램 제목에서 나오는 단어이거나, 어떤 경우에는 프로그램과 관련된 사람 속성일 수 있기 때문이다. 그러나, "WEST"가 문자열 검색기(304)에 의해 속성 목록(310)를 스캔하는 중에 발견되지 않으면, 문자열 해석기(304)는 "WEST"가 프로그램 제목에 대한 검색을 위한 주요 텍스트 용어이어야 한다는 강력한 가정을 형성할 것이다.

[0096] 문자열 해석기(304)는 검색 용어 "G"에 앞서는 "RATED"를 감지하면, 두개의 용어는 "RATED G"로 해석될 수 있고 그에 따라 검색될 수 있다. 하나의 용어 "RATED"가 다른 텍스트 검색(410)(즉, 제목에 "rated"라는 단어가 있는 프로그램)에서 질의될 수 있을지라도 이러한 다른 검색은 "RATED"와 "G"가 공지의 속성의 적절한 순서에서 검색 문자열(202)에서 함께 발견될 때 낮은 관련성 순위로 주어질 가능성이 있다.

[0097] 질의 생성기(314) 및/또는 검색 엔진(204)은 다양한 텍스트 검색(406), 속성 검색(408), 대안적인 텍스트 검색(410), 및 대안적인 속성 검색(412)을 처리한다. 동시에 혹은 검색 문자열(202)의 길이와 의미에 의해 지정되는 논리적 순서로 수행될 수 있는 검색들의 수는 제한되지 않는다. 어떤 경우에는, 예시적인 검색 엔진(204)이 모든 검색들을 수행하고 모든 검색들을 통해 "첫번째 통과" 후에 관련성에 따라 결과들을 목록한다. 다른 경우에, 예시적인 검색 엔진(204)은 문자열 해석기(304)나 질의 생성기(314)에 의해 설정되는 논리적 조건들의 트리를 따른다. 예를 들어, 논리적 조건들은 "텍스트 검색(406)이 A이면, 속성 검색들(408)과 대안적인 텍스트 검색(410)을 검색 용어 X를 사용하여 처리하지만, 텍스트 검색(406)이 B이면, 세가지 다른 속성 검색들(412)을 처리한다. 이 로직은 또한 A가 적중들의 양이고 B가 특정 텍스트 매치를 포함하는 특정 적중라는 것을 더 가정할 수 있다.

[0098] 각 형태의 검색은 이 특정 검색의 결과들을 갖는 중간 적중 목록을 반환할 수 있다. 검색 용어 "WEST"에 대한 텍스트 검색 적중 목록(414)은 "How the West was Won", "West Side Story", "The Wild Wild West", "The West Wing" 등과 같은 프로그램 제목을 포함할 수 있다. "WEST"와 같이 상대적으로 모호한 단일한 검색 용어를 가지고, 질의 생성기(314)는 일반적인 "in-string" 연산자를 사용하여 "WEST"라는 단어를 포함하는 임의의 프로그램 제목을 발견할 가능성이 높다.

[0099] 속성 검색 적중 목록(415)은 "DOLBY® Stereo sound", "DOLBY® digital", 및 "rated G"라는 속성을 갖는 프로그램 제목의 큰 부분을 포함하거나 반환하려고 할 것이다. 이러한 속성들이 어디에나 있다는 것을 알게되면, 질의 생성기(314)는 이러한 속성 검색(408)을 일반적인 속성보다 우선하고 그대신 텍스트 검색(406) 동안 조건

및/또는 필터와 같은 속성들을 포함한다. 달리 말해서, 검색 용어 "WEST"를 포함하는 텍스트 검색(406) 적중은 Dolby 속성이나 "G" 등급 중 적어도 하나와 관련되지도 않으면 허가되지 않을 것이다.

[0100] "WEST"에 대한 텍스트 검색(406)이 많은 다양한 적중을 나오게 하므로, 검색 엔진(204)은 "DOLBY", "RATED", 및 "G"에 대한 대안적인 텍스트 검색(410)을 수행할 가능성이 있고 아마도 각각이 더 많은 정보를 수집하는 적중들의 카운트를 수행할 것이다. "Dolby West"라는 제목의 프로그램이 있으면, 검색 엔진(204)은 정확한 제목 매치라는 관점에서 초기의 해석-검색 엔진(204)이 초기에 가정했던 인접 검색 용어들은 텍스트와 속성 검색 용어들의 혼합이라는 해석을 변경할 것이다.

[0101] 대안적인 속성 검색(412)은 예컨대, 캐스팅과 직원 이름을 속성으로 사용하는 스킵에서 추가의 결과들을 발생시킬 것이다. 그러므로, "DOLBY"와 "WEST"의 대안적인 속성 검색 용어들은 4개의 카테고리로나타난다: Thomas Dolby movies, Steve Dolby movies, Dominic West movies, Shane West movies. 대안적인 속성 검색 용어 "G"도 프로그램 제목에서 큰 부분이 "G" 등급과 관련된 것으로 나타나나 이는 그 크기나 이 검색이 주 속성 검색(408)에서 이미 커버되었으므로 필터링되어 나갔을 것이다.

[0102] 마지막으로, 다양한 검색들(406, 408, 410, 412)로부터 나온 모든 적중 목록(414, 416, 418, 420)이 결과 컴파일러(320)에 의해 처리된다. 검색 비교기/필터(324)는 수신한 적중들을 문자열 해석기(304)에 의해 설정된 검색 문자열(202)을 해석한 것과 비교할 것이다. 결과 순위화기(326)는 각각의 적중에 관련성 순위를 줄 것이다. 포함 필터(328)는 누적적인, 저순위의, 및/또는 범람하는(너무 많은) 적중들을 제거할 뿐만 아니라, 검색 문자열(202)의 해석에 맞지 않는 이러한 적중들을 제거할 것이다. 이 기술된 필터링된 적중 목록(422)에서, 문자열 해석기(304)가 이전에 검색 용어들 간의 불리안 관계를 판정했으면, 즉, 프로그램 제목이 "WEST"를 포함 AND 프로그램이 DOLBY[®]

digital이고 AND "G" 등급이어야 한다고 검색 문자열(202)의 해석을 했으면, "How the West was Won"과 상대적으로 적은 다른 프로그램 제목들만이 디스플레이될 것이다. 물론, 예시적인 검색 엔진(204)에 의해 얻어지는 해석이 사용자의 의도에 가깝지 않으면(사용자에 의해 결정되는 바와 같은 유용한 결과들이 없음으로 판명되는 것과 같이), 사용자는 항상, 특히 많은 실시예에서 전체 프로세스가 거의 즉시에 일어나는 경우에 그 검색을 수정한다.

[0103] 예시적인 점진적(progressive) 검색

[0104] 도 5는 본 발명에 따른 예시적인 점진적 검색을 도시한다. 예시적인 검색 엔진(204)와 같은 검색 엔진의 모듈을 구축하는데 이용가능한 하드웨어 및/또는 소프트웨어는 뛰어난 검색 속도를 제공할만큼 충분히 강력하기 때문에, 예시적인 검색 엔진(204)의 일실시예에서 점진적인 검색을 수행하는 것이 가능하다.

[0105] 점진적인 검색에서, 검색 결과들은 사용자가 타이핑을 하거나 문자들을 UI에 입력함에 따라 생성되는대로 빨리 동적으로 갱신된다. 각 문자가 입력(혹은 제거)된 뒤 검색의 새로운 일괄처리가 시작된다. 비점진적 검색에서 사용자는 검색을 "Enter"키를 침 및/또는 다른 어떤 방식으로 검색 문자열(202)이 완료되었다는 것을 나타냄으로써 수동적으로 검색을 시작할 것이 요구된다.

[0106] 도 5는 몇주분의 EPG 정보를 갖는 EPG 데이터베이스(들)(300) 내에서 검색을 위한 입력 동안 검색 문자열(202) "The West Wing Dolby Election"의 5개의 스냅 사진과 디지털 비디오 기록 시스템의 로컬 하드 드라이브 내에 로컬로 저장된 프로그램의 집합을 보여준다. 블럭(502)에서, "The"라는 단어가 입력되고 예시적인 검색 엔진(204)는 아무런 결과도 없다고 표시한다(504). 이는 노말라이저(302)가 "The"를 관사로서 인식하고 이 관사에만 기초한 검색을 제한했기 때문일 것이다. 그러나, "The"의 "T"가 입력될 때, 매우 빠른 시스템이라면 프로그램 제목에서, "Mr.T"와 같이, "T"만을 독립적으로 가지고 있는 프로그램을 반환할 것이다.

[0107] 블럭(506)에서, 이제 "The West"가 입력되고 예시적인 검색 엔진(204)은 "The West Wing", "How the West was Won" 등과 같은 제목을 포함하는 50개의 결과들(508)을 반환할 것이다.

[0108] 블럭(510)에서, 이제 "The West Wing"이 입력되면 적중의 수를 50에서 10개의 결과(512)로 줄인다. 이 때, 프로그램 제목이 아마도 어떤, 10개의 결과들은 아마도 각각이 자신의 프로그램 ID 번호를 갖고 있는 "The West Wing"의 TV와 영화 에피소드들을 포함할 것이다.

[0109] 블럭(514)에서, 이제 "The West Wing Dolby"가 입력되고, 자격있는 적중들이 5개의 결과로(516) 더 준다. DOLBY[®]

digital이 아니고 다른 Dolby 속성을 가진 이전의 10개의 에피소드들이 제거된다.

- [0110] 블럭(518)에서, 이제 "The West Wing Dolby Election"이 입력되었다. 이는 단 두개의 결과(520)만을 만들었다. 이 두개의 적중들은 "Dolby" 속성을 갖고 "Election" 속성을 갖는 "The West Wing" 에피소드를 이 에피소드의 설명 필드의 단어로서 포함된다.
- [0111] 다양한 적중 결과들이 검색 문자열(202)로서 입력되기 때문에, 원하는 에피소드인 "The West Wing"이 이 과정에서 이미 디스플레이(332)에 나타날 것인데, 예를 들어 "The West"라는 단어만 블럭(506)에서 입력되면, 사용자는 즉시 원하는 에피소드를 강조하여 선택할 수 있을 것이다.
- [0112] 예시적인 방법들
- [0113] 도 6은 본 발명에 따른 프리 텍스트 및 속성 검색을 수행하는 예시적인 방법(600)을 도시한다. 이 방법(600)은 도 2, 3, 5에 도시된 예시적인 검색 엔진(204)와 같은 모듈에 의해 수행될 수 있다. 이 흐름도에서, 동작들은 개별 블럭들에서 요약된다. 동작들은 하드웨어 및/또는 프로세서에 의해 실행될 수 있는 기계 판독가능한 명령들(소프트웨어나 펌웨어)에서 수행될 수 있다.
- [0114] 블럭(602)에서, 검색 문자열이 수신된다. 문자열은 하나 이상의 UI를 통하여 사용자로부터 수신되거나, 저장된 매크로나 연결된 검색으로써 저장장치로부터 수신될 수 있다. 검색 문자열은 또한 한번에 한 글자씩 수신될 수 있고, 새로운 글자가 수신되고 늘어나는 검색 문자열에 더해질 때마다 점진적으로 수행되는 새로운 세트의 검색들이 수행될 수 있다. 수신된 검색 문자열은 텍스트 형태로 속성들이 쓰여진 텍스트 문자열일 것으로 기대된다. 그러나, 수신된 검색 문자열이 표준화되어 텍스트가 아닌 엘리먼트들이 표준화된 텍스트나 표준화된 검색 기호로 변환될 것이기 때문에, 수신된 검색 문자열은 영숫자 문자, ASCII 세트, 토큰, 및/또는 국제 알파벳 문자의 임의의 조합이 될 수 있다. 여러 변형예에서, 검색 문자열은 또한 체크박스 입력, 소프트 키 입력, 아이콘 클릭, 등을 포함할 수 있다.
- [0115] 블럭(604)에서, 검색 문자열이 표준화된다. 표준화는 수신된 문자열을 특정 EPG 정보원을 검색하기 위하여 이해할 수 있는 검색가능한 용어들, 즉, 검색될 EPG 정보 내의 데이터 형식과 맞는 형식을 갖는 용어들로 변환하는 것을 포함한다. 텍스트, 케이스, 및 문자 형식의 단순한 차이로 형식상의 차이에도 불구하고 적중을 생성하는 검색 용어들과 EPG 정보 간에 매칭이 되지 않게 되는 것을 제거하는 것이 그 목적이다. 따라서, 표준화는 스펠링 체크, 자동 정정, 용어 분리, 레터케이싱 표준화, 기호 표준화, 관사 제거, 구두점 제거 등 검색 용어와 가능하면 목적 데이터베이스 정보 모두의 전체 도구이다. 예를 들어, "mash"와 "M*A*S*H"는 검색 용어, 검색될 정보의 에서 어떠한 형태로 발견되든지에 상관없이,
- [0116] 블럭(606)에서, 검색 문자열이 텍스트 용어들과 속성 용어들로 분리된다. 이 분리는 검색 문자열의 해석에 영향을 주고 이를 동반한다. 텍스트 용어들과 속성 용어들로 분리함으로써 표준화에 의해 제공된 검색가능한 용어들의 변환과 전송뿐만 아니라 어떤 해석을 얻게 된다. 평균 검색 문자열이 전형적으로 5개나 6개 혹은 그 이하의 검색 용어들이므로, 스킴에 다양한 크기로 그룹화하고 및 가능한 모든 순서로 검색 용어들의 가능한 모든 조합을 시도하는 것이 뒤따를 수 있다. 그러므로, 검색 문자열은 다양한 그룹화와 검색 용어들의 조합을 테스트하기 위하여 분리될 수 있다.
- [0117] 다양한 후보 검색 용어와 검색 구문들(그룹화)이 생성되므로, 이들은 사전에 존재하는 속성 목록에 대하여 테스트되거나 해석되어 어느 검색 용어와 구문들이 공지의 속성들인지를 결정할 수 있다. 후보 검색 용어들은 또한 이미 검색 문자열에서 수행된 임의의 검색 결과들에 대해 테스트될 수 있는데, 종종 검색 문자열에 수행된 다양한 검색이 동시에 그러나 순차적으로 혹은 상기 검색 문자열 해석 동안 검색 요건으로서 설정되는 조건적 검색 로직에 따라 반복적으로 수행되기 때문이다.
- [0118] 검색 구문에서 단어들을 빠르게 스캔하는 것은 또한 사용자가 의도하는, 또다른 해석 엘리먼트를 구성하는, 다른 상세한 사항을 노출시킬 수 있다. "in", "and", "is", "the"와 같은 연결 어구와 관사들이 많이 있으면, 검색 문자열은 자연어 엔트리로 해석될 것이다. 그러나, 단지 몇개의 간결한 용어들만을 갖는 짧은 검색 문자열은 불리안 스타일 검색으로서 해석될 가능성이 있다.
- [0119] 블럭(608)에서, 각 텍스트 용어와 각 속성 용어에 대해 하나 이상의 검색들이 수행된다. 상대적으로 복잡한 검색 질의들이 존재하는 텍스트와 속성 용어들에 따라 그리고 그것들이 어떻게 서로 관련되어 있는지에 따라 설정될 수 있다. 있음직하지 않지만 예시적인 예로서, 검색 문자열 "western after 1980 sci-fi actors"가 "search for westerns created after 1980 but only those with actors who have also starred in science

fiction movies"로 해석될 수도 있다. 이러한 검색은 "westerns" 카테고리를 검색하는 것과 이러한 1980년 이전에 제작된 서양물(those westerns created before 1980)과 공상과학 영화에도 출연한 배우가 없는 서양물(westerns having no actors who also starred in science fiction movies) 모두를 제거하는 것을 수반한다. 많은 텍스트 용어와 속성 용어 검색들이 한번에, 또한 수신된 결과에 따라 순차적으로 혹은 반복적으로 수행될 수 있다. 따라서, 예시적인 검색 엔진(204)과 같은 예시적인 방법(600)을 수행하는 검색 엔진은 "한 번에 모든 방향으로(in all directions at once)"로 발산하는 검색들을 개시하는 검색 발산기(search manifold)가 될 수 있다.

[0120] 블럭(610)에서, 검색 결과들은 결합되고, 순위화되고, 및/또는 디스플레이를 위해 필터링된다. 결합은 흔히 텍스트 용어들로부터 속성 용어들을 분리하는 동안 가정되는 불리안 관계에 따라 수행된다. 즉, 검색 결과가 어떻게 결합되는지는 텍스트 및 속성 용어 분리와 질의 형성 동안 얻어진 검색 문자열을 해석함에 의해서 부분적으로 예정된다. 주어진 관련성 순위화에서 결과들이 어떠한지는 또한 이 프로세스에서 먼저 얻어진 해석에 달려있다. 필터링은 사전 설정된 사용자의 디스플레이 선호도에 달려 있거나, 임계값 이하의 관련성 순위를 갖는 결과들을 시야로부터 제거시키는 스킴에 달려있다.

[0121] 블럭(612)에서, 검색 결과들이 사용자에게 디스플레이된다. 결과들은 모든 검색이 완료된 후 디스플레이될 수 있지만, 결과들이 수신되고 컴파일되면서 디스플레이될 수도 있다. 점진적 검색에서, 결과들은 검색 문자열이 입력되고 각 입력된 문자들이 새로운 전체의 새로운 검색 세트를 시작하는대로 빨리 디스플레이된다.

[0122] 이상의 예시적인 방법(600)은 본 발명의 단지 한 예일 뿐이다. 이는 발명을 제한하려고 하는 것이 아니라, 사용자에게 가능한 많은 변형예들을 제공할 예시적인 프로세스를 기술하려고 하는 것이다.

[0123] 예시적인 컴퓨팅 장치

[0124] 도 7은 본 발명의 측면들을 실시하기 위한 환경으로서 적절한 예시적인 컴퓨터(700)를 도시한다. 예시적인 컴퓨터(700)의 컴포넌트들은 처리부(720), 시스템 메모리(730), 및 시스템 메모리(730)를 포함하는 다양한 시스템 컴포넌트들을 처리부(720)에 결합시키는 시스템 버스(721)를 포함하나, 이에 제한되지는 않는다. 시스템 버스(721)는 메모리 버스나 메모리 제어기, 주변 버스, 및 임의의 다양한 버스 구조를 사용하는 로컬 버스를 포함하는 버스 구조의 임의의 몇가지 형태가 될 수 있다. 예로서, 이에 제한되지는 않지만, 이러한 아키텍처는 ISA(Industry Standard Architecture) 버스, MCA(Micro Channel Architecture) 버스, EISA(Enhanced ISA) 버스, VESA(Video Electronics Standards Association) 로컬 버스, 및 메자닌(Mezzanine) 버스라고도 알려진 PCI(Peripheral Component Interconnect) 버스를 포함한다.

[0125] 예시적인 컴퓨터(700)는 전형적으로 다양한 컴퓨터 관독가능한 매체를 포함한다. 컴퓨터 관독가능한 매체는 예시적인 컴퓨터(700)에 의해 액세스가능하며 휘발성 및 불휘발성, 분리형 및 비분리형 매체를 포함하는 임의의 가능한 매체가 될 수 있다. 예를 들어, 이에 제한되지는 않지만, 컴퓨터 관독가능한 매체는 컴퓨터 저장 매체와 통신 매체를 포함할 것이다. 컴퓨터 저장 매체는 컴퓨터 관독가능한 명령, 데이터 구조, 프로그램 모듈, 혹은 다른 데이터와 같은 정보를 저장하기 위한 임의의 방법이나 기술로 구현되는 휘발성 및 불휘발성 매체, 분리형 및 비분리형 매체를 포함한다. 컴퓨터 저장 매체는 RAM, ROM, EEPROM, 플래시 메모리나 다른 메모리 기술, CD-ROM, DVD, 혹은 다른 광 디스크 저장장치, 자기 카세트, 자기 테이프, 자기 디스크 저장장치 또는 기타의 자기 저장장치, 혹은 원하는 정보를 저장하는데 사용될 수 있고 예시적인 컴퓨터(700)에 의해 액세스될 수 있는 임의의 다른 매체를 포함하나, 이에 제한되지 않는다. 통신 매체는 전형적으로 컴퓨터 관독가능한 명령, 데이터 구조, 프로그램 모듈이나 반송파나 다른 전송 메카니즘과 같은 변조된 데이터 신호 내의 다른 데이터를 포함하고 임의의 정보 전달 매체를 포함한다. "변조된 데이터 신호"란 하나 이상의 그 특성 세트를 갖거나 신호 내의 정보를 인코딩하는 방식으로 변경된 신호를 의미한다. 예를 들어, 통신 매체는 유선 네트워크나 직접 유선 접속 및 음파, RF, 적외선 및 다른 무선 매체와 같은 무선 매체를 포함한다. 상기의 것들의 어떤 조합도 컴퓨터 관독가능한 매체의 범위 내에 포함된다.

[0126] 시스템 메모리(730)는 ROM(731)과 RAM(732)와 같은 휘발성 및/또는 불휘발성 메모리의 형태의 컴퓨터 저장 매체를 포함한다. 예시적인 컴퓨터(700) 내의 엘리먼트들 간에, 시동시 동안 등에 정보를 전달하는 것을 돕는 기본적인 루틴을 포함하는 BIOS(733)가 전형적으로 ROM(731)에 포함된다. RAM(732)은 전형적으로 처리부(720)에 의해 즉시 접속가능 및/또는 현재 동작되는 데이터 및/또는 프로그램 모듈들을 포함한다. 예를 들어, 도 7은 운영 체제(734), 애플리케이션 프로그램(735), 예시적인 검색 엔진(204), 다른 프로그램 모듈(736), 및 프로그램 데이터(737)를 도시한다. 예시적인 검색 엔진(204)이 RAM(732)에서 소프트웨어로서 도시되어 있지만, 예시적인 검색 엔진(204)의 다른 실시예들은 하드웨어나 소프트웨어와 하드웨어의 조합이 될 수 있다.

- [0127] 예시적인 컴퓨터(700)는 또한 다른 분리형/비분리형, 휘발성/불휘발성 컴퓨터 저장 매체를 포함한다. 단지 예로서, 도 7은 비분리형, 불휘발성 자기 매체에 관독이나 기록을 하는 하드 디스크 드라이브(741), 제거형, 불휘발성 자기 디스크(752)에 관독이나 기록을 하는 자기 디스크 드라이브(751), 및 CD ROM이나 다른 광 매체와 같은 분리형, 불휘발성 광학 디스크(756)에 관독이나 기록을 하는 광 디스크 드라이브(755)를 도시한다. 예시적인 동작 환경에서 사용될 수 있는 다른 제거가능/제거불가능, 휘발성/불휘발성 컴퓨터 저장 매체는 자기 테이프 카세트, 플래시 메모리 카드, DVD, 디지털 비디오 테이프, 고상 RAM, 고상 ROM 등을 포함하나 이에 제한되지 않는다. 하드 디스크 드라이브(741)는 전형적으로 인터페이스(740)와 같은 비분리형 메모리 인터페이스를 통해 시스템 버스(721)에 접속되고 자기 디스크 드라이브(751)와 광 디스크 드라이브(755)는 인터페이스(750)와 같은 제거가능한 메모리에 의해 시스템 버스(721)에 접속된다.
- [0128] 이상 기술되고 도 7에 도시된 드라이브와 그들의 관련 컴퓨터 저장 매체는 컴퓨터 관독가능한 명령, 데이터 구조, 프로그램 모듈, 및 예시적인 컴퓨터(700)에 대한 다른 데이터에 대한 저장장치를 제공한다. 예를 들어, 도 7에서, 하드 디스크 드라이브(741)는 운영 체제(744), 애플리케이션 프로그램(745), 다른 프로그램 모듈(746), 및 프로그램 데이터(747)를 저장하는 것으로 도시된다. 이러한 컴포넌트들은 운영 체제(734), 애플리케이션 프로그램(735), 다른 프로그램 모듈(736), 및 프로그램 데이터(737)와 동일하거나 다를 수 있다. 운영 체제(744), 애플리케이션 프로그램(745), 다른 프로그램 모듈(746), 및 프로그램 데이터(747)는 최소한 다른 것들이라는 것을 설명하기 위해 여기에 다른 번호들로 주어진다. 사용자는 커맨드와 정보를 예시적인 컴퓨터(700)로 키보드(762) 및 통상 마우스, 트랙볼, 또는 터치 패드라고 하는 지시 장치(761)를 통해 입력할 수 있다. 다른 입력 장치들(도시되지 않음)로 마이크로폰, 조이스틱, 게임 패드, 위성 접시, 스캐너 등이 포함된다. 이러한 것들과 기타의 입력 장치들이 보통 시스템 버스에 결합된 사용자 입력 인터페이스(760)를 통해 처리부(720)에 접속되나, 병렬 포트, 게임 포트, USB와 같은 다른 인터페이스와 버스 구조들에는 접속되지 않는다. 모니터(791)와 다른 형태의 디스플레이 장치도 비디오 인터페이스(790)와 같은 인터페이스를 통해 시스템 버스(721)에 접속된다. 모니터 외에, 컴퓨터는 출력 병렬 인터페이스(795)를 통해 접속될 수 있는 스피커(797)와 프린터(796)와 같은 다른 주변 출력 장치들을 포함할 것이다.
- [0129] 컴퓨터는 원격 컴퓨터(780)와 같은 하나 이상의 원격 컴퓨터에 논리적 접속을 이용하여 네트워킹된 환경에서 동작할 수 있다. 원격 컴퓨터(780)는 개인용 컴퓨터, 서버, 라우터, 네트워크 PC, 피어 장치 또는 다른 공통 네트워크 노드가 될 수 있으며, 도 7에는 하나의 메모리 저장 장치(781)만이 도시되었지만 예시적인 컴퓨터(700)와 관련된 이상 기술된 많은 혹은 모든 엘리먼트들을 보통 포함한다. 도 7에 도시된 논리적 접속은 LAN(771)과 WAN(773)을 포함하나, 다른 네트워크도 역시 포함할 수 있다. 이러한 네트워킹 환경들은 사무실, 기업형 컴퓨터 네트워크, 인트라넷, 인터넷에서 흔하다.
- [0130] LAN 네트워킹 환경에서 사용될 때, 예시적인 컴퓨터(700)는 네트워크 인터페이스나 어댑터(770)를 통해 LAN(771)에 접속된다. WAN 네트워킹 환경에서 사용될 때, 예시적인 컴퓨터(700)는 전형적으로 모뎀(772)이나 WAN(773)을 통해 통신을 확립하는 인터넷과 같은 다른 수단들을 포함한다. 모뎀(772)은 내장형이거나 외장형인데, 사용자 입력 인터페이스(760)나 기타의 적절한 기구를 통해 시스템 버스(721)에 접속될 것이다. 네트워킹된 환경에서, 예시적인 컴퓨터(700)와 관련하여 도시된 프로그램 모듈들, 혹은 그들의 부분들은 원격 메모리 저장 장치에 저장될 수 있다. 예를 들어, 도 7은 원격 애플리케이션 프로그램(785)이 메모리 장치(781)에 상주하는 것으로 도시하고 있다. 도시된 네트워크 접속들은 예시적인 것이며 컴퓨터들 간에 통신 링크를 확립하는 다른 수단들이 사용될 수 있다는 것을 이해해야 할 것이다.
- [0131] 물론, 본 발명은 적절한 하드웨어, 게임 콘솔, 및 많은 다른 형태의 처리 장치들을 갖는 셋톱 박스와 같은 많은 다른 환경들에서도 사용될 수 있다. 이러한 환경들 각각은 이 예시적인 컴퓨터(700)에 대하여 이상 기술한 컴포넌트들의 전부는 아니지만 일부를 포함할 수 있다.
- [0132] 결론
- [0133] 이상 기술된 본 발명은 하드웨어, 소프트웨어, 혹은 하드웨어 및 소프트웨어 모두로 구현될 수 있다. 일부 실시예에서, 예시적인 시스템과 관련 방법들이 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터 실행가능한 명령들의 일반적인 문맥으로 기술될 수 있다. 일반적으로, 프로그램 모듈은 루틴, 프로그램, 객체, 컴포넌트, 데이터 구조 등 특정 작업을 수행하거나 특정 추상 데이터 타입을 실현하는 것들을 포함한다. 본 발명은 또한 작업들이 통신 네트워크를 통해 링크된 원격 프로세싱 장치에 의한 무선 통신 상에서 수행되는 분산 통신 환경에서 실행될 수 있다. 무선 네트워크에서, 프로그램 모듈은 메모리 저장 장치를 포함하는 로컬 및 원격 통신 장치 저장 매체에 위치할 수 있다.

- [0134] 이상의 논의는 EPG 데이터의 프리 텍스트 및 속성 검색을 위한 예시적인 시스템 및 방법을 기술한다. 본 발명이 특정 구조적 특징 및/또는 방법론적 액트들에 구체적인 언어로 기술되었지만, 첨부된 청구항들에 정의된 본 발명이 반드시 기술된 특정 특징이나 액트들에 제한되는 것은 아니다. 오히려, 특정 특징들과 액트들은 청구항의 발명을 구현하는 예시적인 형태로서 개시된 것이다.

발명의 효과

- [0135] 텔레비전 방송을 수신하고 기록할 수 있는 텔레비전 세트와 개인용 컴퓨터(PC)에 부가된 두 "셋톱 박스" 시스템을 포함하는 디지털 텔레비전 시스템의 말단 사용자들에게 빠르고, 직관적이고, 정확한 EPG 검색을 제공한다. 본 발명은 디지털 텔레비전 시스템의 사용자의 문자와 속성 질의어를 입력하고 해석하기 위한 검색 엔진과 관련 방법들을 기술한다. 이러한 빠르고 정확한 검색 메카니즘은 예를 들어, 사용자에게 의한 단일 문자열 입력을 해석하고 사용자의 의도를 추론하고, 사용자가 검색하고 있는 해석에 부합하는 적중 결과의 세트를 디스플레이한다.

- [0136] <첨부 A: 예시적인 속성 리스트>

```
<?xml version="1.0" encoding="utf-8" standalone="yes"?>
<searchInfo>
  <categoryAndAttributeMappings separatorChar="|">
    <mapping keywords="closed captioned|closed captioning|closed
captions" field="ScheduleEntryAttribute" value="MSEPGA_CC" />
    <mapping keywords="stereo" field="ScheduleEntryAttribute"
value="MSEPGA_STEREO" />
    <mapping keywords="repeat|rerun|re run"
field="ScheduleEntryAttribute" value="MSEPGA_REPEAT" />
    <mapping keywords="live" field="ScheduleEntryAttribute"
value="MSEPGA_LIVE" />
    <mapping keywords="subtitled|subtitles"
field="ScheduleEntryAttribute" value="MSEPGA_SUBTITLED" />
    <mapping keywords="season premiere|season premier"
field="ScheduleEntryAttribute" value="MSEPGA_SEASONPREMIERE" />
    <mapping keywords="season finale" field="ScheduleEntryAttribute"
value="MSEPGA_SEASONFINALE" />
    <mapping keywords="series premiere|season premier"
field="ScheduleEntryAttribute" value="MSEPGA_SERIESPREMIERE" />
    <mapping keywords="series finale" field="ScheduleEntryAttribute"
value="MSEPGA_SERIESFINALE" />
    <mapping keywords="premiere" field="ScheduleEntryAttribute"
value="MSEPGA_PREMIERE" />
    <mapping keywords="finale" field="ScheduleEntryAttribute"
value="MSEPGA_FINAL" />
    <mapping keywords="letterbox" field="ScheduleEntryAttribute"
value="MSEPGA_LETTERBOX" />
    <mapping keywords="hdtv" field="ScheduleEntryAttribute"
value="MSEPGA_HDTV" />
    <mapping keywords="sap" field="ScheduleEntryAttribute"
value="MSEPGA_SAP" />
    <mapping keywords="dolby" field="ScheduleEntryAttribute"
value="MSEPGA_DOLBY" />
    <mapping keywords="dvs" field="ScheduleEntryAttribute"
value="MSEPGA_DVS" />
    <mapping keywords="rated pg 13|rated pg13"
field="ProgramRatingAttribute" value="MSEPRG_PG13" />
    <mapping keywords="rated pg" field="ProgramRatingAttribute"
value="MSEPRG_PG" />
    <mapping keywords="rated nc 17|rated nc17"
field="ProgramRatingAttribute" value="MSEPRG_NC17" />
    <mapping keywords="rated g" field="ProgramRatingAttribute"
value="MSEPRG_G" />
    <mapping keywords="rated r" field="ProgramRatingAttribute"
value="MSEPRG_R" />
    <mapping keywords="rated x" field="ProgramRatingAttribute"
value="MSEPRG_X" />
    <mapping keywords="rated ao" field="ProgramRatingAttribute"
value="MSEPRG_AO" />
    <mapping keywords="paid programming" field="ProgramAttribute"
value="MSEPGA_PAIDPROGRAMMING" />
    <mapping keywords="miniseries" field="ProgramAttribute"
value="MSEPGA_MINISERIES" />
```

- [0137]

```

    <mapping keywords="action adventure series" field="Category"
value="Series-701" />
    <mapping keywords="action adventure movie|action adventure movies"
field="Category" value="CMOVIE_ACTIONADV" />
    <mapping keywords="action adventure" field="Category"
value="CMOVIE_ACTIONADV|Series-701" />
    <mapping keywords="comedy|comedies" field="Category" value="Movies-
504|Series-703|Kids-303" />
    <mapping keywords="comedy series" field="Category" value="Series-
703" />
    <mapping keywords="comedy movie" field="Category" value="Movies-504"
/>
    <mapping keywords="documentary|documentaries" field="Category"
value="Educational-103" />
    <mapping keywords="drama" field="Category"
value="CMOVIE_DRAMA|Series-705" />
    <mapping keywords="educational series" field="Category"
value="Series-706" />
    <mapping keywords="educational" field="Category" value="Series-
706|Kids-304|Educational-1" />
    <mapping keywords="movies" field="Category" value="MSEPGC_MOVIES" />
    <mapping keywords="películas" field="Category" value="MSEPGC_MOVIES"
/>
    <mapping keywords="music" field="Category" value="Movies-508|Series-
709|Special-803" />
    <mapping keywords="news" field="Category" value="MSEPGC_NEWS" />
    <mapping keywords="science fiction movie|science fiction movies"
field="Category" value="Movies-511" />
    <mapping keywords="science fiction" field="Category" value="Movies-
511" />
    <mapping keywords="fantasy movie|fantasy movies" field="Category"
value="Movies-506" />
    <mapping keywords="fantasy" field="Category" value="Movies-506" />
    <mapping keywords="soap operas|soaps" field="Category"
value="Series-711" />
    <mapping keywords="sports" field="Category" value="News-605|Special-
805|MSEPGC_SPORTS" />
    <mapping keywords="adults only|xxx" field="Category" value="Movies-
502|Lifestyle-401" />
    <mapping keywords="kids movies" field="Category" value="Movies-503"
/>
    <mapping keywords="kids series" field="Category" value="Series-702"
/>
    <mapping keywords="kids" field="Category" value="Series-702|Movies-
503|Kids-3" />
    <mapping keywords="musicals" field="Category" value="Movies-508" />
    <mapping keywords="mystery movies|mysteries|mystery"
field="Category" value="Movies-509" />
    <mapping keywords="love stories|romance" field="Category"
value="Movies-510" />
    <mapping keywords="suspense" field="Category" value="Movies-512" />
    <mapping keywords="westerns|western" field="Category" value="Movies-
513" />
    <mapping keywords="animated|cartoon" field="Category" value="Kids-
302" />
    <mapping keywords="baseball|mlb" field="Category" value="CBASEBALL"
/>

```

[0138]

```

    <mapping keywords="basketball|nba" field="Category"
value="CBASKETBALL" />
    <mapping keywords="boxing" field="Category" value="CBOXING" />
    <mapping keywords="football|nfl" field="Category" value="CFOOTBALL"
/>
    <mapping keywords="golf|pga|lpga|usga" field="Category"
value="CGOLF" />
    <mapping keywords="hockey|nhl" field="Category" value="CHOCKEY" />
    <mapping keywords="outdoor sports" field="Category" value="COUTDOOR"
/>
    <mapping keywords="racing" field="Category" value="CRACING" />
    <mapping keywords="soccer" field="Category" value="CSOCCER" />
    <mapping keywords="tennis" field="Category" value="CTENNIS" />
    <mapping keywords="sports news" field="Category" value="News-605" />
    <mapping keywords="sports" field="Category"
value="MSEPGC_SPORTS|News-605" />
    <mapping keywords="business news|financial news" field="Category"
value="News-601" />
    <mapping keywords="interviews|interview" field="Category"
value="News-602" />
    <mapping keywords="public affairs" field="Category" value="News-604"
/>
    <mapping keywords="weather" field="Category" value="News-606" />
    <mapping keywords="cooking" field="Category" value="Series-704" />
    <mapping keywords="game shows|game show" field="Category"
value="Series-707" />
    <mapping keywords="how to|howto" field="Category" value="Series-708"
/>
    <mapping keywords="talk shows|talk show" field="Category"
value="Series-712" />
    <mapping keywords="travel" field="Category" value="Series-713" />
    <mapping keywords="religious" field="Category" value="Special-
804|Lifestyle-409" />
    <mapping keywords="holidays|holiday" field="Category"
value="Special-802" />
    <mapping keywords="biography" field="Category" value="Educational-
102" />
    <mapping keywords="science" field="Category" value="Educational-105"
/>
    <mapping keywords="collectibles" field="Category" value="Lifestyle-
402" />
    <mapping keywords="exercise" field="Category" value="Lifestyle-404"
/>
    <mapping keywords="health" field="Category" value="Lifestyle-405" />
    <mapping keywords="home garden" field="Category" value="Lifestyle-
406" />
    <mapping keywords="international" field="Category" value="Lifestyle-
407" />
    <mapping keywords="outdoors" field="Category" value="Lifestyle-408" />
    <mapping keywords="primetime|prime time" field="TimeOfDay"
value="1140|1380" />
    <mapping keywords="morning" field="TimeOfDay" value="300|720" />
    <mapping keywords="afternoon" field="TimeOfDay" value="720|1080" />
    <mapping keywords="evening" field="TimeOfDay" value="1020|1200" />
    <mapping keywords="late night" field="TimeOfDay" value="1380|300" />
</categoryAndAttributeMappings>
</searchInfo>

```

[0139]

도면의 간단한 설명

[0001]

도 1은 종래 기술의 텍스트 및 속성 검색의 도식적 표현.

[0002]

도 2는 본 발명에 따른 예시적인 검색의 도식적 표현.

[0003]

도 3은 예시적인 검색 엔진의 블록도.

[0004]

도 4는 도 3의 예시적인 검색 엔진에 의해 수행된 예시적인 검색의 도식적 표현.

[0005]

도 5는 도 3의 예시적인 검색 엔진에 의해 수행된 예시적인 점진적 검색의 도식적 표현.

[0006]

도 6은 본 발명에 따른 예시적인 방법의 흐름도.

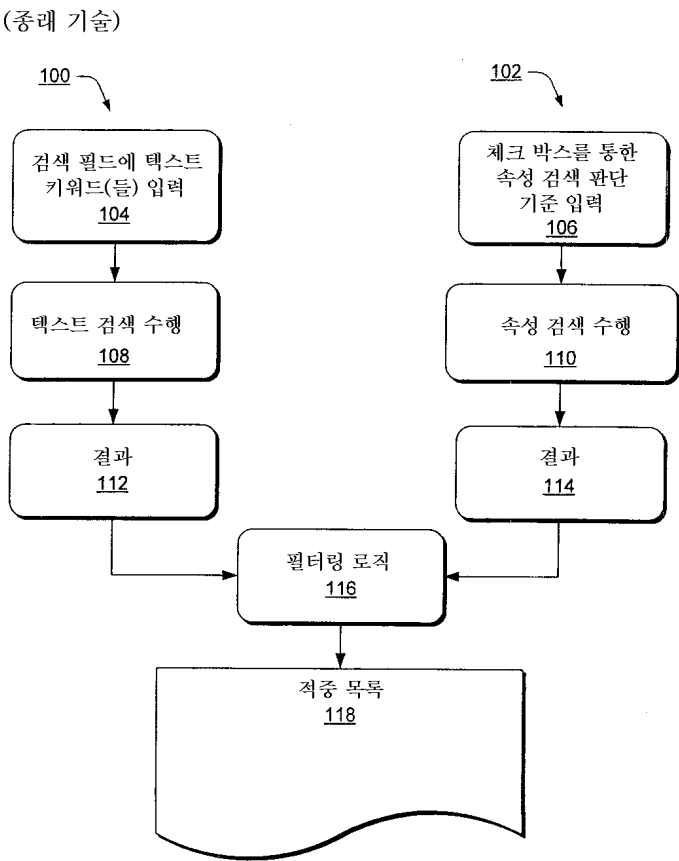
[0007]

도 7은 본 발명을 실행하는 예시적인 컴퓨팅 장치 환경의 블록도.

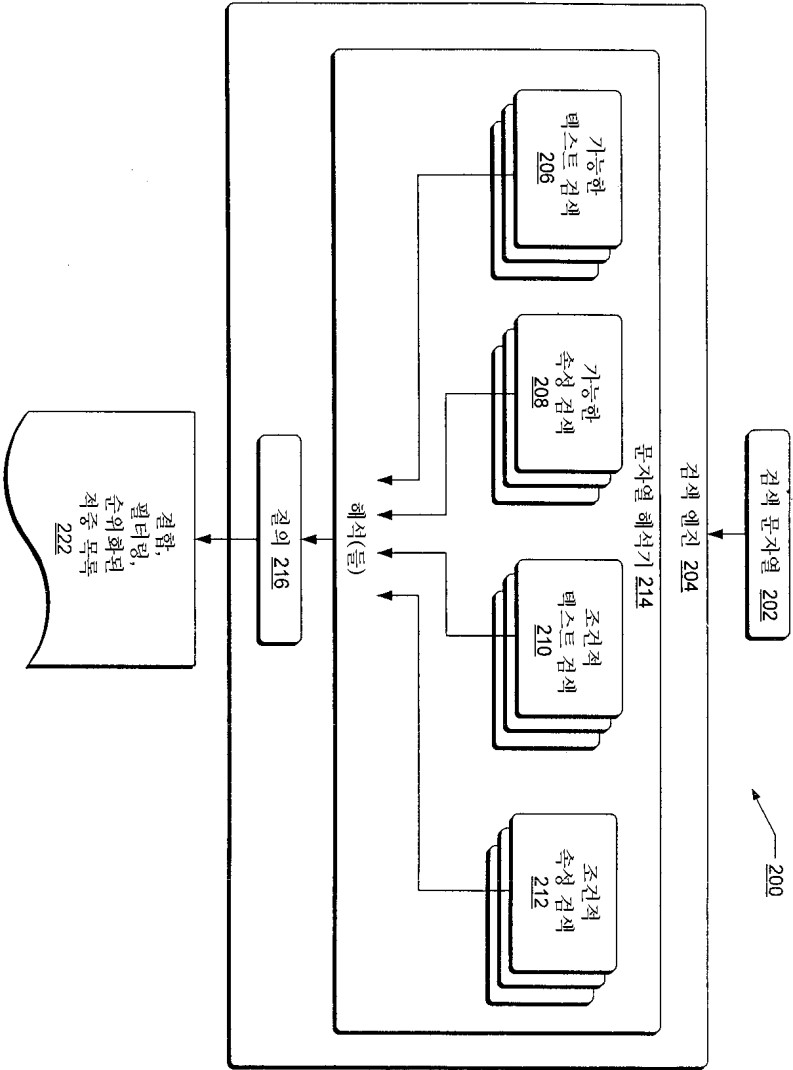
- [0008] (도면의 주요부분에 대한 부호의 설명)
- [0009] 204: 검색 엔진
- [0010] 301: 속성 목록
- [0011] 312: 검색 요건표
- [0012] 334: 웹기반 EPG 자원
- [0013] 300: EPG 데이터베이스
- [0014] 302: 노말라이저
- [0015] 214: 문자열 해석기
- [0016] 314: 질의 생성기
- [0017] 320: 결과 컴파일러
- [0018] 332: 디스플레이

도면

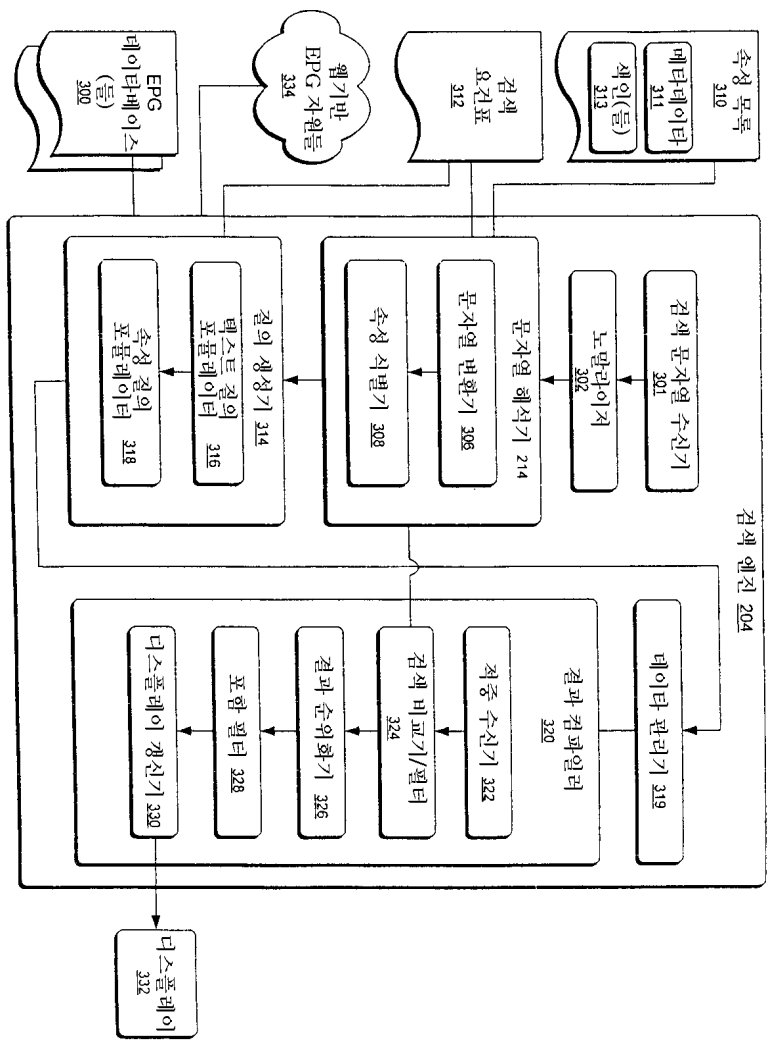
도면1



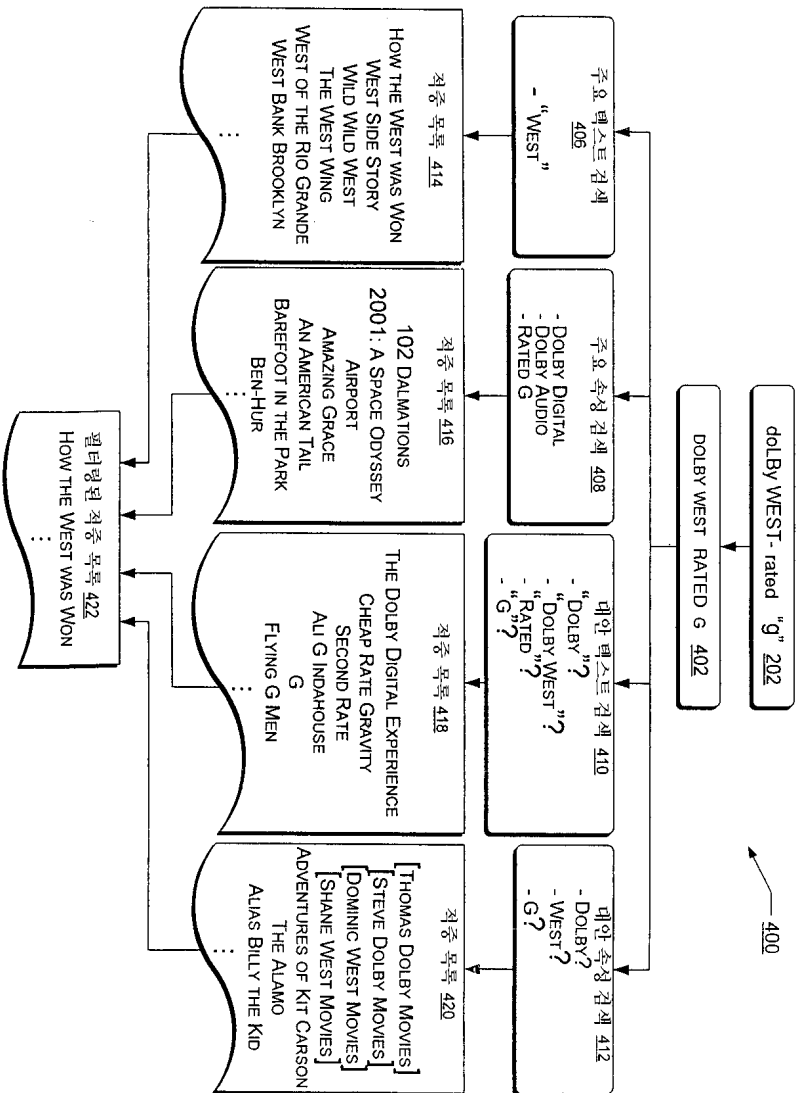
도면2



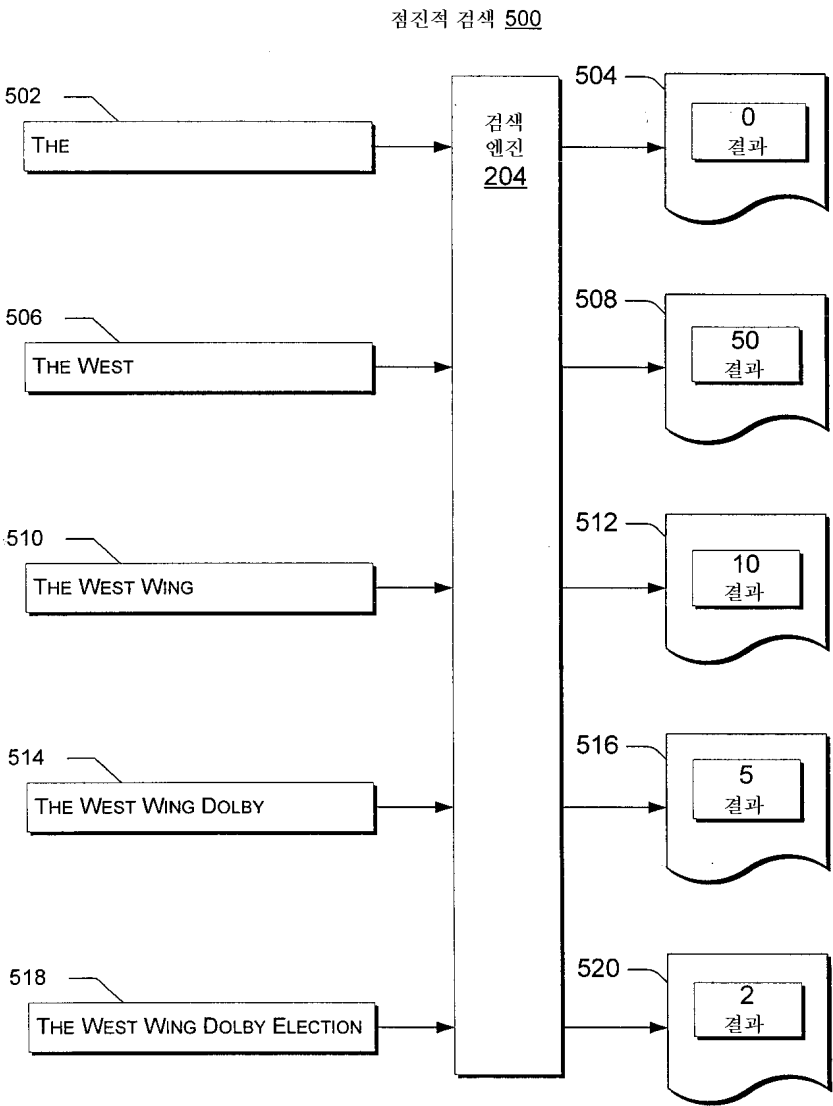
도면3



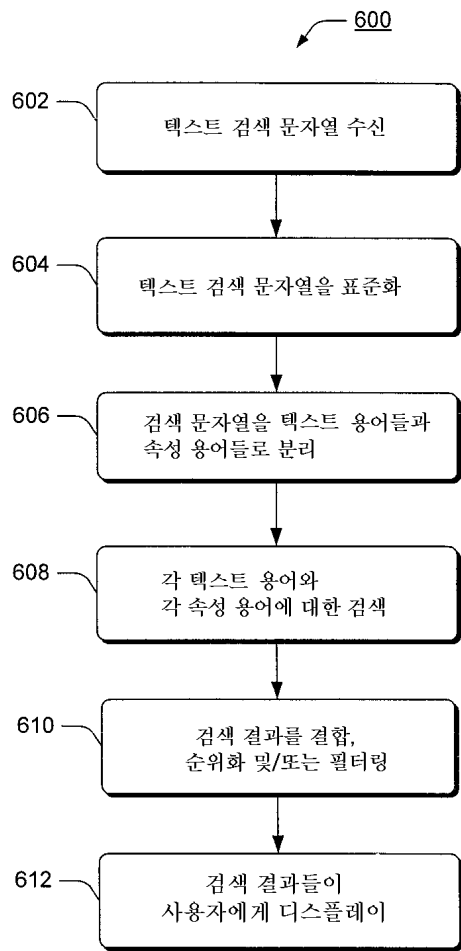
도면4



도면5



도면6



도면7

