



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년10월14일
(11) 등록번호 10-1662527
(24) 등록일자 2016년09월28일

(51) 국제특허분류(Int. Cl.)
G06F 17/21 (2006.01) G06F 17/27 (2006.01)
G06F 17/30 (2006.01)
(52) CPC특허분류
G06F 17/218 (2013.01)
G06F 17/277 (2013.01)
(21) 출원번호 10-2015-0090254
(22) 출원일자 2015년06월25일
심사청구일자 2015년06월25일
(56) 선행기술조사문헌
W02007052285 A1

(73) 특허권자
서울시립대학교 산학협력단
서울특별시 동대문구 서울시립대로 163 (전농동, 서울시립대학교)
(72) 발명자
구교진
서울특별시 강남구 선릉로69길 19, 102동 201호 (역삼동, 역삼래미안아파트)
박형진
서울특별시 중랑구 중랑천로 286, 106동 204호 (묵동, 묵동아이파크아파트)
(74) 대리인
특허법인 천지

전체 청구항 수 : 총 28 항

심사관 : 이복현

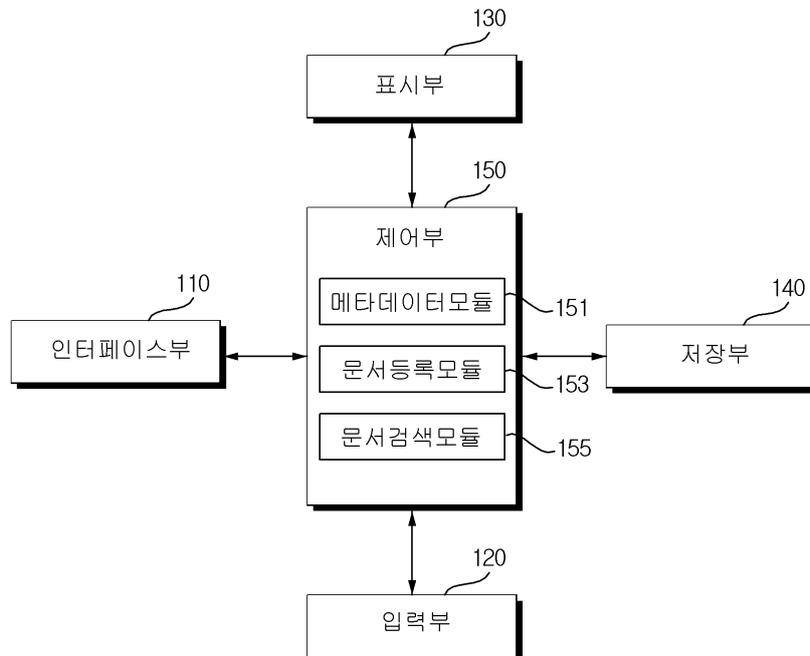
(54) 발명의 명칭 메타데이터 라이브러리를 이용한 문서 관리를 위한 장치, 이를 위한 방법 및 이 방법이 기록된 컴퓨터 판독 가능한 기록매체

(57) 요약

본 발명은 메타데이터 라이브러리를 이용한 문서 관리를 위한 장치, 이를 위한 방법 및 이 방법이 기록된 컴퓨터 판독 가능한 기록매체에 관한 것이다. 이러한 본 발명은 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체제로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으

(뒷면에 계속)

대표도 - 도2



로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하는 저장부와, 문서에서 복수의 키워드를 추출하고, 상기 기본어 테이블로부터 상기 키워드에 매칭되는 기본어를 추출하며, 상기 파생어 테이블로부터 상기 추출된 기본어의 파생어를 추출하고, 상기 유사어 테이블로부터 상기 추출된 기본어의 유사어를 추출하여, 상기 추출된 기본어, 파생어 및 유사어를 후보 태그로 선정하며, 선정된 후보 태그 중 적어도 하나를 태그로 선정하고, 선정된 태그를 상기 문서에 부여하여 상기 저장부에 저장하는 등록을 수행하는 문서등록모듈을 포함하는 것을 특징으로 하는 문서를 관리하기 위한 장치와 이를 위한 방법 및 이 방법이 기록된 컴퓨터 판독 가능한 기록매체를 제공한다.

(52) CPC특허분류

G06F 17/2795 (2013.01)

G06F 17/30967 (2013.01)

(72) 발명자

조동현

서울특별시 동대문구 전농로37길 41-9, 1층 (전농동)

박상현

서울특별시 서대문구 거북골로24길 17, 일원빌리지 401호 (북가좌동)

이 발명을 지원한 국가연구개발사업

과제고유번호 2010-0024159

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 이공분야기초연구사업→(구)일반연구자지원사업→(구)기본연구지원사업→(구)기본연구지원사업(유형Ⅱ)

연구과제명 건설지식기반 문서 재활용 그린인포 시스템 개발 (Construction Knowledge based Green Information System for Document Recycling)

기 여 율 1/1

주관기관 서울시립대학교

연구기간 2010.09.01 ~ 2015.08.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

문서를 관리하기 위한 장치에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하는 저장부; 및

문서에서 복수의 키워드를 추출하고, 상기 기본어 테이블로부터 상기 키워드에 매칭되는 기본어를 추출하며, 상기 파생어 테이블로부터 상기 추출된 기본어의 파생어를 추출하고, 상기 유사어 테이블로부터 상기 추출된 기본어의 유사어를 추출하여, 상기 추출된 기본어, 파생어 및 유사어를 후보 태그로 선정하며, 선정된 후보 태그 중 적어도 하나를 태그로 선정하고, 선정된 태그를 상기 문서에 부여하여 상기 저장부에 저장하는 등록을 수행하는 문서등록모듈;을 포함하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 2

제1항에 있어서,

상기 문서등록모듈은

상기 문서로부터 상기 기본어, 유사어 및 파생어가 아닌 키워드 중 기 설정된 출현 빈도를 가지는 키워드를 도출하고, 도출된 키워드 중 적어도 하나를 상기 문서에 태그로 부여하여 상기 저장부에 저장하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 3

제1항에 있어서,

상기 문서등록모듈은

기 설정된 시간 혹은 기 설정된 주기가 도래하면, 기 설정된 폴더에 저장된 복수의 문서 중 상기 등록이 이루어지지 않은 문서를 추출하고, 추출된 문서에 대해 상기 등록을 수행하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 4

제1항에 있어서,

건설업 분야에서 사용된 복수의 문서의 텍스트로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서 검색을 위해 사용된 검색어로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서를 등록할 때 문서에 부여된 핵심어로부터 복수의 키워드를 추출하며,

추출된 키워드를 상호 매칭하여 매칭된 키워드를 선정하고, 선정된 키워드로부터 기본어를 도출하여 도출된 기본어를 가지는 상기 기본어 테이블을 생성하는 메타데이터모듈;을 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 5

제4항에 있어서,

상기 메타데이터모듈은

상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 생성하는 것을 특징으로 하는 문서를 관리하기 위한 장

치.

청구항 6

제4항에 있어서,

상기 메타데이터모듈은

상기 복수의 키워드를 추출할 때,

상기 텍스트, 상기 검색어 및 상기 핵심어 중 2개 이상의 음절로 구성된 단어를 추출하고, 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하며, 분절된 어절들 중에서 중복되는 단어를 제거하고, 기 저장된 소거 용어 테이블에서 지정된 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외하여, 상기 복수의 키워드를 추출하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 7

문서를 관리하기 위한 장치에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하는 저장부;

사용자로부터 검색어를 입력 받는 입력부; 및

상기 검색어가 입력되면, 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 기본어이면, 상기 저장부에서 검색어인 기본어, 검색어인 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 문서검색모듈;을 포함하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 8

제7항에 있어서,

상기 문서검색모듈은

상기 메타데이터 라이브러리를 검색하여 상기 검색어가 파생어 혹은 유사어이면, 상기 파생어 혹은 유사어를 기본어로 변환하고, 상기 저장부로부터 변환된 기본어, 변환된 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 9

제7항 또는 제8항에 있어서,

상기 문서검색모듈은

상기 기본어 테이블로부터 상기 검색에 사용된 기본어의 상위 기본어를 도출하고, 상기 복합어 테이블로부터 도출된 상위 기본어를 포함하는 모든 복합어를 관련 검색어로 도출하여 제공하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 10

제7항에 있어서,

상기 문서검색모듈은

검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 11

제10항에 있어서,

상기 문서검색모듈은
수학식

$$S_T = F_T / \sum_{T=1}^n F_{Ti}$$

을 통해 상기 태그 일치도 점수를 산출하며,

상기 S_T 는 태그 일치도 점수이며,

상기 F_T 는 태그의 출현 빈도인 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 12

제7항에 있어서,

상기 문서검색모듈은

검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수, 상기 검색된 문서의 파일명 혹은 블록명이 검색어를 포함하는지 여부를 나타내는 파일명 일치도 점수 및 상기 검색된 문서에 대해 사용자에게 의해 부여된 점수인 문서 평점 각각에 가중치를 부여하고, 각각이 가중치가 부여된 태그 일치도 점수, 파일명 일치도 점수 및 문서 평점의 합산 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 13

제12항에 있어서,

상기 문서검색모듈은

수학식

$$S = W_a S_T + W_b S_F + W_c S_D$$

을 통해 상기 합산 점수를 산출하며,

상기 S 는 상기 합산 점수인 문서 랭킹 점수이며,

상기 S_T 는 정규화된 태그 일치도 점수이며,

상기 S_F 는 정규화된 파일명 일치도 점수이며,

상기 S_D 는 정규화된 문서 평점이며,

상기 W_a 는 태그 일치도에 대한 가중치이며,

상기 W_b 는 파일명 일치도에 대한 가중치이며,

상기 W_c 는 문서 평점에 대한 가중치인 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 14

문서를 관리하기 위한 장치에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하고, 미리 설정된 기준에 상기 태그가 속한 분류가 지정된 분류체계를 저장하는 저장부;

사용자로부터 상기 분류체계의 적어도 하나의 분류를 입력 받는 입력부; 및

상기 분류가 입력되면, 상기 분류에 속하는 태그를 추출하고, 추출된 태그가 부여된 문서를 검색하여 검색 결과를 제공하는 문서검색모듈;을 포함하는 것을 특징으로 하는 문서를 관리하기 위한 장치.

청구항 15

문서를 관리하기 위한 방법에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하는 단계;

문서에서 복수의 키워드를 추출하여, 상기 기본어 테이블로부터 상기 키워드에 매칭되는 기본어를 추출하고, 상기 파생어 테이블로부터 상기 추출된 기본어의 파생어를 추출하고, 상기 유사어 테이블로부터 상기 추출된 기본어의 유사어를 추출하는 단계;

상기 추출된 기본어, 파생어 및 유사어를 후보 태그로 선정하고, 선정된 후보 태그 중 적어도 하나를 태그로 선정하는 단계; 및

선정된 태그를 상기 문서에 부여하여 저장하는 등록을 수행하는 단계;를 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 16

제15항에 있어서,

상기 태그로 선정하는 단계는

상기 문서로부터 상기 기본어, 유사어 및 파생어가 아닌 키워드 중 기 설정된 출현 빈도를 가지는 키워드를 도출하여 상기 도출된 키워드 중 적어도 하나를 상기 태그로 선정하는 단계;를 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 17

제15항에 있어서,

상기 저장하는 단계 전,

건설업 분야에서 사용된 복수의 문서의 텍스트로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서 검색을 위해 사용된 검색어로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서를 등록할 때 문서에 부여된 핵심어로부터 복수의 키워드를 추출하는 단계; 및

상기 추출된 키워드를 상호 매칭하여 매칭된 키워드를 선정하고, 선정된 키워드로부터 기본어를 도출하여 도출된 기본어를 가지는 상기 기본어 테이블을 생성하는 단계;를 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 18

제17항에 있어서,

상기 기본어 테이블을 생성하는 단계 후, 상기 저장하는 단계 전,

상기 기본어 중 하나 이상의 조합으로 이루어진 상기 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 상기 유사어 테이블을 생성하는 단계;를 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 19

제17항에 있어서,

상기 복수의 키워드를 추출하는 단계는

상기 텍스트, 상기 검색어 및 상기 핵심어 중 2개 이상의 음절로 구성된 단어를 추출하는 단계;

상기 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하는 단계;

상기 분절된 어절들 중에서 중복되는 단어를 제거하는 단계; 및

기 저장된 소거 용어 테이블에서 지정된 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외하여 상기 복수의 키워드를 추출하는 단계를 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 20

문서를 관리하기 위한 방법에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하는 단계; 및

검색어가 입력되면, 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 기본어이면, 검색어인 기본어, 검색어인 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 단계;를 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 21

제20항에 있어서,

상기 메타데이터 라이브러리를 검색하여 상기 검색어가 파생어 혹은 유사어이면, 상기 파생어 혹은 유사어를 기본어로 변환하고, 변환된 기본어, 변환된 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 단계를 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 22

제21항에 있어서,

상기 기본어 테이블로부터 상기 검색에 사용된 기본어의 상위 기본어를 도출하고, 상기 복합어 테이블로부터 도출된 상위 기본어를 포함하는 모든 복합어를 관련 검색어로 도출하여 제공하는 단계;를 더 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 23

제20항에 있어서,

상기 검색 결과를 제공하는 단계는

상기 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 24

제23항에 있어서,

상기 검색 결과를 제공하는 단계는

수학식

$$S_T = F_T / \sum_{T=1}^n F_{Ti}$$

을 통해 상기 태그 일치도 점수를 산출하며,

상기 S_T 는 태그 일치도 점수이며,

상기 F_T 는 태그의 출현 빈도인 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 25

제20항에 있어서,

상기 검색 결과를 제공하는 단계는

검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수, 상기 검색된 문서의 파일명 혹은 블록명이 검색어를 포함하는지 여부를 나타내는 파일명 일치도 점수 및 상기 검색된 문서에 대해 사용자에게 의해 부여된 점수인 문서 평점 각각에 가중치를 부여하고, 각각이 가중치가 부여된 태그 일치도 점수, 파일명 일치도 점수 및 문서 평점의 합산 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 26

제25항에 있어서,

상기 검색 결과를 제공하는 단계는

수학식

$$S = W_a S_T + W_b S_F + W_c S_D$$

을 통해 상기 합산 점수를 산출하며,

상기 S 는 상기 합산 점수인 문서 랭킹 점수이며,

상기 S_T 는 정규화된 태그 일치도 점수이며,

상기 S_F 는 정규화된 파일명 일치도 점수이며,

상기 S_D 는 정규화된 문서 평점이며,

상기 W_a 는 태그 일치도에 대한 가중치이며,

상기 W_b 는 파일명 일치도에 대한 가중치이며,

상기 W_c 는 문서 평점에 대한 가중치인 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 27

문서를 관리하기 위한 방법에 있어서,

건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지

는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하고, 미리 설정된 기준에 상기 태그가 속한 분류가 지정된 분류체계를 저장하는 단계;

상기 분류가 입력되면, 상기 분류에 속하는 태그를 추출하는 단계; 및

상기 추출된 태그가 부여된 문서를 검색하여 검색 결과를 제공하는 단계;를 포함하는 것을 특징으로 하는 문서를 관리하기 위한 방법.

청구항 28

컴퓨터를 통해 제20항 내지 제27항 중 어느 한 항에 따른 문서를 관리하기 위한 방법을 실행시키기 위한 프로그램이 기록된 컴퓨터 판독 가능한 기록매체.

발명의 설명

기술 분야

[0001] 본 발명은 문서 관리 기술에 관한 것으로, 보다 상세하게는 메타데이터 라이브러리를 이용하여 문서를 관리할 수 있는 장치, 이를 위한 방법 및 이 방법이 기록된 컴퓨터 판독 가능한 기록매체에 관한 것이다.

배경 기술

[0002] 건설 프로세스에서 발생하는 문서들은 건설기술과 지식이 발현된 실체로서 실무적 가치가 높은 정보의 집성체이나, 건설 문서들의 물리적 내용적 특수성과 다양성으로 인해 건설 참여주체 내부 또는 주체간의 효율적인 공유 및 재활용이 이루어지지 않고 있다. 이러한 건설도메인 문서를 효율적으로 관리하는 기존의 방법은 유럽의 Uniclass, 북미의 Master Format, 국내 건설CALS와 같은 분류체계방식이다. 분류체계는 다양한 유형을 일관된 기준으로 분류하여 명확하게 구분될 수 있으나 분류가 애매모호하고 복수개 이상의 분류항목에 속하는 등의 문제점이 있다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) 한국공개특허 제2006-0130893호 2006년 12월 20일 등록 (명칭: CALS / EC 표준체계의 프로젝트성과 자동생성모듈)

발명의 내용

해결하려는 과제

[0004] 본 발명의 목적은 건설 메타데이터 라이브러리를 이용하여 건설 분야에서 사용되는 문서를 관리할 수 있는 장치, 이를 위한 방법 및 이 방법이 기록된 컴퓨터 판독 가능한 기록매체를 제공함에 있다.

과제의 해결 수단

[0005] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 장치는 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하는 저장부와, 문서에서 복수의 키워드를 추출하고, 상기 기본어 테이블로부터 상기 키워드에 매칭되는 기본어를 추출하며, 상기 파생어 테이블로부터 상기 추출된 기본어의 파생어를 추출하고, 상기 유사어 테이블로부터 상기 추출된 기본어의 유사어를 추출하여, 상기 추출된 기본어, 파생어 및 유사어를 후보 태그로 선정하며, 선정된 후보 태그 중 적어도 하나를 태그로 선정하고, 선정된 태그를 상기 문서에 부여하여 상기 저장부에 저장하는 등록을 수행하는 문서등록모듈을 포함한다.

[0006] 상기 문서등록모듈은 상기 문서로부터 상기 기본어, 유사어 및 파생어가 아닌 키워드 중 기 설정된 출현 빈도를 가지는 키워드를 도출하고, 도출된 키워드 중 적어도 하나를 상기 문서에 태그로 부여하여 상기 저장부에 저장하는 것을 특징으로 한다.

[0007] 상기 문서등록모듈은 기 설정된 시간 혹은 기 설정된 주기가 도래하면, 기 설정된 폴더에 저장된 복수의 문서 중 상기 등록이 이루어지지 않은 문서를 추출하고, 추출된 문서에 대해 상기 등록을 수행하는 것을 특징으로 한다.

[0008] 본 발명의 실시예에 따른 문서를 관리하기 위한 장치는 건설업 분야에서 사용된 복수의 문서의 텍스트로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서 검색을 위해 사용된 검색어로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서를 등록할 때 문서에 부여된 핵심어로부터 복수의 키워드를 추출하며, 추출된 키워드를 상호 매칭하여 매칭된 키워드를 선정하고, 선정된 키워드로부터 기본어를 도출하여 도출된 기본어를 가지는 상기 기본어 테이블을 생성하는 메타데이터모듈을 더 포함한다.

[0009] 상기 메타데이터모듈은 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 생성하는 것을 특징으로 한다.

[0010] 상기 메타데이터모듈은 상기 복수의 키워드를 추출할 때, 상기 텍스트, 상기 검색어 및 상기 핵심어 중 2개 이상의 음절로 구성된 단어를 추출하고, 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하며, 분절된 어절들 중에서 중복되는 단어를 제거하고, 기 저장된 소거 용어 테이블에서 지정된 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외하여, 상기 복수의 키워드를 추출하는 것을 특징으로 한다.

[0011] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 장치는 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하는 저장부와, 사용자로부터 검색어를 입력 받는 입력부와, 상기 검색어가 입력되면, 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 기본어이면, 상기 저장부에서 검색어인 기본어, 검색어인 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 문서검색모듈을 포함한다.

[0012] 상기 문서검색모듈은 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 파생어 혹은 유사어이면, 상기 파생어 혹은 유사어를 기본어로 변환하고, 상기 저장부로부터 변환된 기본어, 변환된 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 것을 특징으로 한다.

[0013] 상기 문서검색모듈은 상기 기본어 테이블로부터 상기 검색에 사용된 기본어의 상위 기본어를 도출하고, 상기 복합어 테이블로부터 도출된 상위 기본어를 포함하는 모든 복합어를 관련 검색어로 도출하여 제공하는 것을 특징으로 한다.

[0014] 상기 문서검색모듈은 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 한다.

[0015] 상기 문서검색모듈은 수학적

$$S_T = F_T / \sum_{T=1}^n F_{Ti}$$

을 통해 상기 태

그 일치도 점수를 산출하며, 상기 S_T 는 태그 일치도 점수이며, 상기 F_T 는 태그의 출현 빈도인 것을 특징으로 한다.

[0016] 상기 문서검색모듈은 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수, 상기 검색된 문서의 파일명 혹은 블록명이 검색어를 포함하는지 여부를 나타내는 파일명 일치도 점수 및 상기 검색된 문서에 대해 사용자에게 의해 부여된 점수인 문서 평점 각각에 가중치를 부여하고, 각각이 가중치가 부여된 태그 일치도 점수, 파일명 일치도 점수 및 문서 평점의 합산 점수가 높은 순으로 상기 검색 결

과를 정렬하여 제공하는 것을 특징으로 한다.

[0017] 상기 문서검색모듈은 수학적

$$S = W_a S_T + W_b S_F + W_c S_D$$

[0018] 을 통해 상기 합산 점수를 산출하며, 상기 S 는 상기 합산 점수인 문서 랭킹 점수이며, 상기 S_T 는 정규화된 태그 일치도 점수이며, 상기 S_F 는 정규화된 파일명 일치도 점수이며, 상기 S_D 는 정규화된 문서 평점이며, 상기 W_a 는 태그 일치도에 대한 가중치이며, 상기 W_b 는 파일명 일치도에 대한 가중치이며, 상기 W_c 는 문서 평점에 대한 가중치인 것을 특징으로 한다.

[0019] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 장치는 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하고, 미리 설정된 기준에 상기 태그가 속한 분류가 지정된 분류체계를 저장하는 저장부와, 사용자로부터 상기 분류체계의 적어도 하나의 분류를 입력 받는 입력부와, 상기 분류가 입력되면, 상기 분류에 속하는 태그를 추출하고, 추출된 태그가 부여된 문서를 검색하여 검색 결과를 제공하는 문서검색모듈을 포함한다.

[0020] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 방법은 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하는 단계와, 문서에서 복수의 키워드를 추출하여, 상기 기본어 테이블로부터 상기 키워드에 매칭되는 기본어를 추출하고, 상기 파생어 테이블로부터 상기 추출된 기본어의 파생어를 추출하고, 상기 유사어 테이블로부터 상기 추출된 기본어의 유사어를 추출하는 단계와, 상기 추출된 기본어, 파생어 및 유사어를 후보 태그로 선정하고, 선정된 후보 태그 중 적어도 하나를 태그로 선정하는 단계와, 선정된 태그를 상기 문서에 부여하여 저장하는 등록을 수행하는 단계를 포함한다.

[0021] 상기 태그로 선정하는 단계는 상기 문서로부터 상기 기본어, 유사어 및 파생어가 아닌 키워드 중 기 설정된 출현 빈도를 가지는 키워드를 도출하여 상기 도출된 키워드 중 적어도 하나를 상기 태그로 선정하는 단계를 더 포함한다.

[0022] 상기 저장하는 단계 전, 건설업 분야에서 사용된 복수의 문서의 텍스트로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서 검색을 위해 사용된 검색어로부터 복수의 키워드를 추출하고, 건설업 분야에서 문서를 등록할 때 문서에 부여된 핵심어로부터 복수의 키워드를 추출하는 단계와, 상기 추출된 키워드를 상호 매칭하여 매칭된 키워드를 선정하고, 선정된 키워드로부터 기본어를 도출하여 도출된 기본어를 가지는 상기 기본어 테이블을 생성하는 단계를 더 포함한다.

[0023] 상기 기본어 테이블을 생성하는 단계 후, 상기 저장하는 단계 전, 상기 기본어 중 하나 이상의 조합으로 이루어진 상기 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 상기 유사어 테이블을 생성하는 단계를 더 포함한다.

[0024] 상기 복수의 키워드를 추출하는 단계는 상기 텍스트, 상기 검색어 및 상기 핵심어 중 2개 이상의 음절로 구성된 단어를 추출하는 단계와, 상기 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하는 단계와, 상기 분절된 어절들 중에서 중복되는 단어를 제거하는 단계와, 기 저장된 소거 용어 테이블에서 지정된 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외하여 상기 복수의 키워드를 추출하는 단계를 포함한다.

[0025] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 방법은 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브

러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하는 단계와, 상기 검색어가 입력되면, 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 기본어이면, 검색어인 기본어, 검색어인 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 단계를 포함한다.

[0027] 또한, 본 발명의 실시예에 따른 문서를 관리하기 위한 방법은 상기 메타데이터 라이브러리를 검색하여 상기 검색어가 파생어 혹은 유사어이면, 상기 파생어 혹은 유사어를 기본어로 변환하고, 변환된 기본어, 변환된 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색하여, 검색 결과를 제공하는 단계를 더 포함한다.

[0028] 더욱이, 본 발명의 실시예에 따른 문서를 관리하기 위한 방법은 상기 기본어 테이블로부터 상기 검색에 사용된 기본어의 상위 기본어를 도출하고, 상기 복합어 테이블로부터 도출된 상위 기본어를 포함하는 모든 복합어를 관련 검색어로 도출하여 제공하는 단계를 더 포함한다.

[0029] 상기 검색 결과를 제공하는 단계는 상기 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 한다.

[0030] 상기 검색 결과를 제공하는 단계는 수학식

$$S_T = F_T / \sum_{T=1}^n F_{Ti}$$

을 통해 상기 태

그 일치도 점수를 산출하며, 상기 S_T 는 태그 일치도 점수이며, 상기 F_T 는 태그의 출현 빈도인 것을 특징으로 한다.

[0031] 상기 검색 결과를 제공하는 단계는 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도 점수, 상기 검색된 문서의 파일명 혹은 블록명이 검색어를 포함하는지 여부를 나타내는 파일명 일치도 점수 및 상기 검색된 문서에 대해 사용자에게 의해 부여된 점수인 문서 평점 각각에 가중치를 부여하고, 각각이 가중치가 부여된 태그 일치도 점수, 파일명 일치도 점수 및 문서 평점의 합산 점수가 높은 순으로 상기 검색 결과를 정렬하여 제공하는 것을 특징으로 한다.

[0032] 상기 검색 결과를 제공하는 단계는 수학식

$$S = W_a S_T + W_b S_F + W_c S_D$$

을 통해 상기 합산 점수를 산

출하며, 상기 S 는 상기 합산 점수인 문서 랭킹 점수이며, 상기 S_T 는 정규화된 태그 일치도 점수이며, 상기 S_F 는 정규화된 파일명 일치도 점수이며, 상기 S_D 는 정규화된 문서 평점이며, 상기 W_a 는 태그 일치도에 대한 가중치이며, 상기 W_b 는 파일명 일치도에 대한 가중치이며, 상기 W_c 는 문서 평점에 대한 가중치인 것을 특징으로 한다.

[0033] 상술한 바와 같은 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 문서를 관리하기 위한 방법은 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계로 이루어지는 기본어 테이블과, 상기 기본어 중 하나 이상의 조합으로 이루어진 복합어 테이블과, 상기 기본어의 파생어를 가지는 파생어 테이블 및 상기 기본어의 유사어를 가지는 유사어 테이블을 포함하여 구성되는 메타데이터 라이브러리를 저장하며, 상기 메타데이터 라이브러리의 기본어, 파생어, 유사어 중 적어도 하나가 태그로 부여된 문서를 저장하고, 미리 설정된 기준에 상기 태그가 속한 분류가 지정된 분류체계를 저장하는 단계와, 상기 분류가 입력되면, 상기 분류에 속하는 태그를 추출하는 단계와, 상기 추출된 태그가 부여된 문서를 검색하여 검색 결과를 제공하는 단계를 포함한다.

[0034] 본 발명의 다른 견지에 따르면 컴퓨터를 통해 전술한 본 발명이 실시예에 따른 문서를 관리하기 위한 방법을 실행시키기 위한 프로그램이 기록된 컴퓨터 판독 가능한 기록매체를 제공할 수 있다.

발명의 효과

[0035] 상술한 바와 같은 본 발명에 따르면, 건설 분야에 특화된 메타데이터 라이브러리를 이용하여 도면에 태그를 부여하여 등록할 수 있고, 이러한 태그를 통해 문서를 검색할 수 있어, 전문 분야, 즉, 건설 분야의 문서에 대해 정밀하고 정확한 검색을 제공할 수 있다.

도면의 간단한 설명

[0036] 도 1은 본 발명의 실시예에 따른 메타데이터 라이브러리를 이용한 문서 관리를 위한 시스템을 설명하기 위한 도면이다.

도 2는 본 발명의 실시예에 따른 문서 관리 장치의 구성을 설명하기 위한 블록도이다.

도 3은 본 발명의 실시예에 따른 메타데이터 라이브러리를 생성하는 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 메타데이터 라이브러리를 이용하여 문서를 등록하는 방법을 설명하기 위한 흐름도이다.

도 5는 본 발명의 일 실시예에 따른 문서의 일부를 등록 영역으로 지정하는 방법을 설명하기 위한 화면 예이다.

도 6은 본 발명의 다른 실시예에 따른 메타데이터 라이브러리를 이용하여 문서를 등록하는 방법을 설명하기 위한 흐름도이다.

도 7은 본 발명의 일 실시예에 따라 메타데이터 라이브러리를 이용하여 문서를 검색하는 방법을 설명하기 위한 흐름도이다.

도 8은 본 발명의 일 실시예에 따라 메타데이터 라이브러리를 이용하여 문서를 검색한 결과를 보이는 화면 예이다.

도 9는 본 발명의 다른 실시예에 따른 문서를 검색하는 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0037] 본 발명의 상세한 설명에 앞서, 이하에서 설명되는 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니 되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념으로 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 실시예에 불과할 뿐, 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형 예들이 있을 수 있음을 이해하여야 한다.

[0038] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예들을 상세히 설명한다. 이때, 첨부된 도면에서 동일한 구성 요소는 가능한 동일한 부호로 나타내고 있음을 유의해야 한다. 또한, 본 발명의 요지를 흐리게 할 수 있는 공지 기능 및 구성에 대한 상세한 설명은 생략할 것이다. 마찬가지로의 이유로 첨부 도면에 있어서 일부 구성요소는 과장되거나 생략되거나 또는 개략적으로 도시되었으며, 각 구성요소의 크기는 실제 크기를 전적으로 반영하는 것이 아니다.

[0039] 먼저, 본 발명의 실시예에 따른 메타데이터 라이브러리를 이용한 문서 관리를 위한 시스템에 대해서 설명하기로 한다. 도 1은 본 발명의 실시예에 따른 메타데이터 라이브러리를 이용한 문서 관리를 위한 시스템을 설명하기 위한 도면이다.

[0040] 본 발명의 실시예에 따른 문서 관리를 위한 시스템은 복수의 문서 관리 장치(100)로 이루어진다. 문서 관리 장치(100)는 컴퓨팅 연산 장치이다. 따라서 문서 관리 장치(100)는 기본적으로 컴퓨팅 연산을 위한 프로세서와, 메모리, 스토리지 등의 저장매체를 구비한다. 또한, 문서 관리 장치(100)는 입출력을 위한 I/O 장치, 화면 표시를 위한 디스플레이, 다른 문서 관리 장치(100) 혹은 네트워크와 통신하기 위한 통신 장치를 추가로 구비한다. 이러한 문서 관리 장치(100)는 대표적으로, 컴퓨터를 예시할 수 있다. 그 밖에 컴퓨팅 연산이 가능한 어떠한 장치라도 본 발명의 실시예에 따른 문서 관리 장치(100)가 될 수 있다. 문서 관리 장치(100) 중 일부는 서버로 동작하며, 다른 일부는 클라이언트로 동작할 수 있다. 도시된 바와 같이, 복수의 회사, 예컨대, 회사 A, B 및 C는 복수의 문서 관리 장치(100)로 이루어지며, 각 회사별로 어느 하나의 문서 관리 장치(100)는 서버로 동작하고, 나머지는 클라이언트로 동작할 수 있다. 또한, 복수의 회사의 모든 문서 관리 장치(100)를 클라이언트로 하고, 이에 대해 서버로 동작하는 문서 관리 장치(100)가 존재할 수 있다. 이러한 문서 관리 장치(100)는 상호 간에 인터넷 혹은 인트라넷을 통해 상호 연결될 수 있다. 이에 따라, 문서 관리 장치(100) 상호 간에 폴더를 상호 공

유하도록 설정될 수 있다.

- [0041] 이하에서 설명되는 문서 등록 및 문서 검색은 문서 관리 장치(100)가 스탠드얼론(stand-alone) 기기이고, 하나의 문서 관리 장치(100)에서 이루어지는 것과 같이 설명될 것이다. 하지만, 본 발명은 이에 한정되는 것은 아니다. 가령, 클라이언트/서버 방식 혹은 클라우드 컴퓨팅 방식으로, 문서 등록을 위해 필요한 데이터는 클라이언트로부터 제공되며, 그 데이터를 이용하여 문서를 등록하는 컴퓨팅 연산은 서버에서 이루어질 수 있다. 그리고 문서 검색을 위한 검색어 입력은 클라이언트에서 이루어지고, 검색을 위한 컴퓨팅 연산은 서버에서 이루어지며, 검색 결과는 다시 클라이언트에 표시될 수 있다.
- [0042] 그러면, 보다 상세히 본 발명의 실시예에 따른 문서 관리 장치(100)에 대해서 설명하기로 한다. 도 2는 본 발명의 실시예에 따른 문서 관리 장치의 구성을 설명하기 위한 블록도이다.
- [0043] 도 2를 참조하면, 본 발명의 실시예에 따른 문서 관리 장치(100)는 인터페이스부(110), 입력부(120), 표시부(130), 저장부(140) 및 제어부(150)을 포함한다.
- [0044] 인터페이스부(110)는 네트워크 연결 혹은 주변 장치와의 연결 등을 통해 외부와 연결되는 인터페이스 수단이다. 즉, 인터페이스부(110)는 문서 관리 장치(100)와 다른 문서 관리 장치(100)와의 인터페이스이자, 통신 수단이다. 이러한 인터페이스부(110)는 모뎀, 인터페이스 카드, 유무선 랜카드, USB 포트, 시리얼 포트, 패러럴 포트, 데이터 버스 등으로 구성될 수 있다. 특히, 인터페이스부(110)는 다양한 데이터 송수신 수단으로 문서 관리 장치(100) 외부와 데이터를 송수신할 수 있는 어떠한 것이라도 될 수 있다.
- [0045] 입력부(120)는 문서 관리 장치(100)의 각 종 기능, 동작 등을 제어하기 위한 사용자의 키 조작을 입력받고 입력 신호를 생성하여 제어부(150)에 전달한다. 입력부(120)는 키보드, 마우스 등을 예시할 수 있다. 입력부(120)는 전원 on/off를 위한 전원 키, 문자 키, 숫자 키, 방향키 중 적어도 하나를 포함할 수 있다. 입력부(120)의 기능은 표시부(130)가 터치스크린으로 구현된 경우, 표시부(130)에서 이루어질 수 있으며, 표시부(130)만으로 모든 기능을 수행할 수 있는 경우, 입력부(120)는 생략될 수도 있다.
- [0046] 표시부(130)는 제어부(150)로부터 화면 표시를 위한 데이터를 수신하여 수신된 데이터를 화면으로 표시할 수 있다. 또한, 표시부(130)는 문서 관리 장치(100)의 메뉴, 데이터, 기능 설정 정보 및 기타 다양한 정보를 사용자에게 시각적으로 제공할 수 있다. 표시부(130)가 터치스크린으로 형성되는 경우, 입력부(120)의 기능의 일부 또는 전부를 대신 수행할 수 있다. 표시부(130)는 액정표시장치(LCD, Liquid Crystal Display), 유기 발광 다이오드(OLED, Organic Light Emitting Diodes), 능동형 유기 발광 다이오드(AMOLED, Active Matrix Organic Light Emitting Diodes) 등으로 형성될 수 있다.
- [0047] 저장부(140)는 문서 관리 장치(100)의 동작에 필요한 각 종 데이터, 어플리케이션, 문서 관리 장치(100)의 동작에 따라 발생된 각 종 데이터를 저장하는 역할을 수행한다. 이러한 저장부(140)는 크게 프로그램 영역과 데이터 영역을 포함할 수 있다. 프로그램 영역은 문서 관리 장치(100)의 부팅(booting) 및 운영(operation)을 위한 운영체제(OS, Operating System), 본 발명의 실시예에 따른 도면 관리를 위한 방법을 실행하는 어플리케이션 등을 저장할 수 있다. 데이터 영역은 문서 관리를 위한 각 종 데이터를 저장할 수 있다. 저장부(140)에 저장되는 각 종 데이터는 사용자의 조작에 따라, 삭제, 변경, 추가될 수 있다.
- [0048] 제어부(150)는 문서 관리 장치(100)의 전반적인 동작 및 문서 관리 장치(100)의 내부 블록들 간 신호 흐름을 제어하고, 데이터를 처리하는 데이터 처리 기능을 수행할 수 있다. 이러한 제어부(150)는 중앙 처리 장치(Central Processing Unit : CPU), 어플리케이션 프로세서(Application Processor), GPU(Graphic Processing Unit) 등이 될 수 있다.
- [0049] 제어부(150)는 메타데이터모듈(451), 문서등록모듈(453) 및 문서검색모듈(455)를 포함한다. 본 발명의 실시예에서 메타데이터모듈(451), 문서등록모듈(453) 및 문서검색모듈(455)는 하드웨어로 구현된 것과 같이 설명될 것이나, 이에 한정되는 것은 아니며, 메타데이터모듈(451), 문서등록모듈(453) 및 문서검색모듈(455)는 저장부(140)에 저장되었다가, 제어부(150)에서 실행되는 어플리케이션으로 구현될 수도 있다. 메타데이터모듈(451)은 본 발명의 실시예에 따른 메타데이터 라이브러리를 생성한다. 문서등록모듈(453)은 메타데이터 라이브러리를 이용하여 문서에 태그를 부여하고, 이를 저장부(140)에 저장하는 등록을 수행한다. 문서검색모듈(455)는 사용자가 입력한 검색어에 따라 문서를 검색하여 제공한다. 이러한 메타데이터모듈(451), 문서등록모듈(453) 및 문서검색모듈(455)를 포함하는 제어부(150)의 동작은 아래에서 보다 상세하게 설명될 것이다.
- [0050] 본 발명은 메타데이터 라이브러리를 이용하여 도면을 관리한다. 이러한 메타데이터 라이브러리를 생성하는 방법에 대해서 설명하기로 한다. 도 3은 본 발명의 실시예에 따른 메타데이터 라이브러리를 생성하는 방법을 설명하

기 위한 흐름도이다.

[0051] 제어부(150)의 메타데이터모듈(151)은 S110 단계에서 건설업 분야에서 사용된 복수의 문서의 텍스트로부터 복수의 키워드를 도출한다. 여기서, 건설업 분야에서 사용된 복수의 문서는 인터페이스부(110)를 통해 입력되거나, 저장부(140)에 저장된 문서가 될 수 있다. 이러한 S110 단계에서 메타데이터모듈(151)은 텍스트 중 ① 2개 이상의 음절로 구성된 단어를 추출하고, ② 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하며, ③ 분절된 어절들 중에서 중복되는 단어를 제거하고, ④ 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외한다.

[0052] 여기서, 메타데이터모듈(151)은 본 발명의 실시예에 따른 등록 용어 테이블에 미리 등록된 건설업 분야의 용어를 선택할 수 있다. 특히, 메타데이터모듈(151)은 본 발명의 실시예에 따른 소거 용어 테이블에 미리 등록된 건설업 분야의 용어가 아닌 용어와 일반 용어를 제외한다. 다음의 표 1은 소거 용어 테이블의 일례를 보인다.

표 1

구분		소거 용어
건설업분야가 아닌 키워드		교환업무, 국가경쟁, 기계화, 기능인력
건설업/타분야에서 함께 활용되는 일반적인 키워드		
사물명사(대명사) 또는 고유명사	사물명사	단풍나무, 환풍기, 덤프트럭, 음향계 등
	고유명사	국가명, 도시명, 하천교량명 등
분야 관계없이 사용되는 키워드		교육훈련
기타		관리항목, 문서유형, 자재, 시설부위, 공중에 해당하지 않는 키워드

[0054] 다음으로, 메타데이터모듈(151)은 S120 단계에서 건설업 분야에서 문서 검색을 위해 사용된 검색어로부터 복수의 키워드를 도출한다. 이러한 검색어는 복수의 문서 관리 장치(100)를 통해 검색이 이루어질 때 사용된 검색어가 누적 저장된 것이 될 수 있으며, 이러한 검색어들은 복수의 문서 검색 장치(100)가 공유하거나, 관리 서버(200)에서 취합하여 제공될 수 있다. 이러한 S120 단계에서 메타데이터모듈(151)은 검색어 중 ① 2개 이상의 음절로 구성된 단어를 추출하고, ② 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하며, ③ 분절된 어절들 중에서 중복되는 단어를 제거하고, ④ 건설업 분야의 용어가 아닌 용어와 및 일반 용어를 제외한다.

[0055] 다음으로, 메타데이터모듈(151)은 S130 단계에서 건설업 분야에서 문서를 등록할 때 문서에 부여된 핵심어로부터 복수의 키워드를 도출한다. 이러한 핵심어는 복수의 문서 관리 장치(100)를 통해 문서가 등록될 때 사용자가 입력한 용어이며, 문서 등록 시, 누적 저장된다. 그리고 이러한 핵심어는 복수의 문서 검색 장치(100)가 공유하거나, 관리 서버(200)가 취합하여 제공될 수 있다. 이러한 S130 단계에서 메타데이터모듈(151)은 핵심어 중 ① 2개 이상의 음절로 구성된 단어를 추출하고, ② 추출된 단어들을 의미 파악이 가능한 최소 어절 단위로 분절하며, ③ 분절된 어절들 중에서 중복되는 단어를 제거하고, ④ 건설업 분야의 용어가 아닌 용어와 및 일반 용어를 제외한다.

[0056] 다음으로, 메타데이터모듈(151)은 S140 단계에서 앞서 S110, S120 및 S130 단계에서 도출된 키워드를 상호 매칭하고, S150 단계에서 상호 매칭되는 키워드를 선정한다.

[0057] 이어서, 메타데이터모듈(151)은 S160 단계에서 앞서 선정된 키워드를 기 설정된 분류에 따라 구분한다. 예컨대, 메타데이터모듈(151)은 다음의 표 2와 같이, 선정된 키워드를 관리항목, 문서유형, 자재, 시설부위, 공중 및 기타를 포함하는 6개의 분류로 구분할 수 있다.

표 2

관리항목	문서유형	자재	시설부위	공중	기타
공정	검토서	강관과일	경사로	가설	검토항목
공정관리	계약서	강판	계단	가설공사	결론
사업비	계획서	고강도콘크리트	공장	가스	공사
사업비관리	관리대장	골재	교량	금속	대상지역설정
설계	기준	기동콘크리트	놀이시설	기계	대안평가
설계관리	내역서	기초콘크리트	도로	기계설비공사	부도
설계단계	법	기포콘크리트	발코니	방수	사용계획
설계변경	법규	노출콘크리트	샤워부스	전기	수행기간

...
-----	-----	-----	-----	-----	-----

[0059] 다음으로, 메타데이터모듈(151)은 S170 단계에서 앞서 선정된 키워드 중 기본어를 도출하고, 기본어를 상호 위계를 가지는 복수의 레벨의 분류 체계로 구성하는 기본어 테이블을 생성한다. 예컨대, 메타데이터모듈(151)은 다음의 표 3과 같이, 선정된 키워드 중 기본어를 도출하고, 도출된 기본어를 위계를 가지는 3개의 레벨의 분류 체계로 구성하는 기본어 테이블을 생성한다.

표 3

[0060]

레벨1	CODE	레벨2	CODE	레벨3	CODE		
공중	W	가설	W01	비계	W0101		
				복공	W0102		
		철근콘크리트공사	W02	거푸집	W0201		
				레미콘	W0202		
				배근	W0203		
				동바리	W0204		
		토(토목)공사	W03	굴착	W0301		
				말뚝	W0302		
				발파	W0303		
				터파기	W0304		
		...					
		원가	C	간접비	C01		
직접비	C02						
견적	C03			내역	C0301		
...							

[0061] 이어서, 메타데이터모듈(151)은 S180 단계에서 앞서 도출된 기본어로부터 복합어, 파생어 및 유사어 도출하고, 복합어, 파생어 및 유사어에 따른 기본어간 연관 관계를 설정하는 복합어, 파생어 및 유사어 테이블을 생성한다. 이러한 연관 관계는 코드를 통해 설정할 수 있다. 다음의 표 4는 복합어 테이블의 일례이며, 표 5는 파생어 테이블의 일례이고, 표 6은 유사어 테이블의 일례를 보인다.

표 4

[0062]

복합어	코드	속성 1	속성 2	속성 3
가설복공	W010015	W01	W0102	
가설복공계획	W010016	W01	W0102	N07
가설복공구조	W010017	W01	W0102	N01
가설복공내역	W010018	W01	W0102	C0301
가설비계	W010018	W01	W0101	

표 5

[0063]

코드	기본어	파생어 1	파생어 2	...
EW0001	내역	내역서		...
EW0002	계획	계획서		...
EW0003	단가	단가표	단가집	...
EW0004	거푸집	거푸집틀		...

표 6

[0064]

코드	기본어	유사어 1	유사어 2	...
SW0001	해체	제거		...
SW0002	검사	검측	시험	...
SW0003	방음재	방음제	흡음재	...
SW0004	스테인레스	스텐레스	stainless	...

- [0065] 보인 바와 같이, 표 4의 복합어 테이블은 복수의 복합어를 포함하며, 복수의 복합어는 복수의 기본어로 이루어진다. 이에 따라, 각 복합어는 속성(속성 1, 2, 3)으로 기본어의 코드를 가진다. 또한, 표 5의 파생어 테이블은 기본어와 그 기본어에 대한 하나 이상의 파생어를 가진다. 그리고 표 6의 유사어 테이블은 기본어와 그 기본어에 대한 하나 이상의 유사어를 가진다.
- [0066] 전술한 바와 같은 방법에 따라 본 발명의 실시예에 따른 메타데이터 라이브러리를 구축할 수 있으며, 이러한 메타데이터 라이브러리는 건설업 분야에서 사용되는 키워드로 이루어진 복수의 기본어가 복수 레벨의 위계를 가지는 분류체계에 이루어지는 기본어 테이블과, 복수의 기본어로부터 도출되는 복합어(복합단어) 테이블과, 복수의 기본어로부터 도출되는 파생어 테이블 및 복수의 기본어로부터 도출되는 유사어 테이블을 포함하여 구성된다. 특히, 본 발명의 실시예에 따른 메타데이터 라이브러리는 실무에서 축적되는 데이터(키워드)를 지속적으로 반영하여 전술한 기본어 테이블, 복합어 테이블, 파생어 테이블 및 유사어 테이블을 갱신할 수 있다.
- [0067] 다음으로, 본 발명의 일 실시예에 따른 메타데이터 라이브러리를 이용하여 문서에 대한 태그를 도출하고, 이를 문서에 부여하는 등록 절차에 대해서 설명하기로 한다. 도 4는 본 발명의 일 실시예에 따른 메타데이터 라이브러리를 이용하여 문서를 등록하는 방법을 설명하기 위한 흐름도이다. 도 5는 본 발명의 일 실시예에 따른 문서의 일부를 등록 영역으로 지정하는 방법을 설명하기 위한 화면 예이다.
- [0068] 도 4를 참조하면, 제어부(150)의 문서등록모듈(153)은 S210 단계에서 입력부(120)를 통한 사용자의 입력에 따라 등록하고자 하는 문서를 로드하여 표시부(130)를 통해 표시한다.
- [0069] 이러한 문서는 다양한 파일 형식을 가질 수 있다. 예컨대, 파일 형식은 HWP, DOC, PDF, XLS, PPT, DWG 등이 될 수 있다. 사용자는 이러한 문서의 일부 혹은 전부를 선택하여 등록할 수 있다. 이에 따라, 사용자는 문서 중 일부 또는 전부를 지정하는 입력을 할 수 있다. 본 발명의 실시예에 따르면, 사용자가 문서의 일부를 선택하는 경우, 파일 형식에 상응하는 방식에 따라 문서의 일부를 등록 영역으로 선택하게 할 수 있다. 문서등록모듈(153)은 해당 파일 형식을 사용하는 어플리케이션으로부터 등록 영역을 지정하는 기능을 도출하고, 이를 사용자가 사용할 수 있도록 제공한다. 예컨대, 도 5에 보인 바와 같이, HWP 파일을 운영하는 어플리케이션은 '쉬프트 키와 방향키', 마우스 드래그, 'ctrl 및 F3 키' 등을 이용하여 블록을 지정할 수 있다. 이에 따라, 문서등록모듈(153)은 해당 어플리케이션으로부터 블록을 지정하는 기능을 도출하여 사용자가 등록 영역을 지정하는 데에 사용할 수 있도록 제공한다. DOC 파일 또한 HWP 파일과 유사한 기능을 제공하며, 이에 따라, 문서등록모듈(153)은 해당 어플리케이션으로부터 블록을 지정하는 기능을 도출하여 사용자가 등록 영역을 지정하는 데에 사용할 수 있도록 제공할 수 있다. 또한, XLS 파일에서 제공하는 셀 선택 기능, PPT 파일에서 페이지 번호를 선택하는 기능 등을 사용자가 등록 영역을 지정하는 데에 사용할 수 있도록 제공할 수 있다. 그리고 기타 다양한 형식으로 사용자가 등록하고자 하는 영역을 지정하도록 할 수 있다.
- [0070] 이와 같이, 사용자가 파일 형식에 따라 등록하고자 하는 영역을 지정하는 입력을 하면, 문서등록모듈(153)은 S220 단계에서 입력부(120)를 통해 사용자의 등록 영역을 지정하는 입력을 감지하여 입력에 따라 등록 영역을 지정한다. 등록 영역이 지정되면, 문서등록모듈(153)은 S230 단계에서 문서의 지정 영역 내에서 모든 키워드를 추출한다.
- [0071] 이어서, 문서등록모듈(153)은 S240 단계에서 기본어 테이블에서 앞서(S230) 추출된 키워드에 매칭되는 기본어를 추출한다. 이어서, 문서등록모듈(153)은 S250 단계에서 복합어 테이블 및 파생어 테이블로부터 앞서(S240) 추출된 기본어와 연관된 복합어 및 파생어를 추출한다. 일례로, S240 단계에서 기본어 테이블로부터 추출된 키워드에 매칭되는 기본어 '입찰'이 추출되고, S250 단계에서 추출된 기본어와 연관된 복합어 및 파생어인 '입찰보증', '입찰자', '입찰제도' 등이 추출될 수 있다.
- [0072] 문서등록모듈(153)은 S260 단계에서 앞서 추출된 기본어, 복합어 및 파생어를 후보 태그로 표시부(130)를 통해 표시한다. 이때, 메타데이터 라이브러리로부터 얻어진 후보 태그는 문서에서 출현 빈도에 따라 정렬하여 출현 빈도와 함께 표시하는 것이 바람직하다.
- [0073] 다음으로, 문서등록모듈(153)은 S270 단계에서 메타데이터 라이브러리와 매칭되지 않지만, 문서 내 기 설정된 횟수 이상 출현한 다출현 키워드를 후보 태그로 표시부(130)를 통해 표시한다. 이때, 다출현 키워드에 따른 후보 태그 또한 문서에서 출현 빈도에 따라 정렬하여 출현 빈도와 함께 표시하는 것이 바람직하다.
- [0074] 한편, 사용자는 표시된 후보 태그 중 원하는 태그가 없거나, 사용자가 추가의 태그가 필요하다고 판단하는

경우, 사용자는 태그로 사용할 키워드를 입력할 수 있다. 그러면, 문서등록모듈(153)은 S280 단계에서 입력부(120)를 통해 사용자가 입력한 키워드를 입력 받고, 입력된 키워드를 후보 태그로 표시부(130)를 통해 표시할 수 있다.

- [0075] 사용자는 표시부(130)에 표시된 복수의 후보 태그 중 적어도 하나를 선택할 수 있다. 이에 따라, 사용자가 입력부(120)를 통해 복수의 후보 태그 중 적어도 하나를 선택하는 입력을 하면, 문서등록모듈(153)은 S290 단계에서 이를 감지하여 선택된 키워드를 해당 문서의 태그로 결정한다. 이와 같이, 태그가 결정되면, 문서등록모듈(153)은 S300 단계에서 태그와 문서를 매핑하여 저장부(140)에 저장하는 등록을 수행한다.
- [0076] 한편, 전술한 바와 같이, 문서 전체에 대해 등록하지 않고, 일부분을 등록하는 경우, 원본 문서와 따로 지정된 등록 영역만 파일로 저장될 수 있다. 이때, 따로 저장된 저장되는 등록 영역의 파일에 대한 파일이름은 블록명이라고 한다. 이러한 블록명은 사용자가 입력한 명칭으로 지정되거나, 원본 문서의 파일 이름에 추가 일련번호가 붙는 형식으로 지정되거나, 해당 등록 영역의 태그 중 어느 하나로 지정될 수 있다.
- [0077] 다음으로, 본 발명의 다른 실시예에 따른 메타데이터 라이브러리를 이용하여 문서를 태깅하는 방법에 대해서 설명하기로 한다. 도 6은 본 발명의 다른 실시예에 따른 메타데이터 라이브러리를 이용하여 문서를 등록하는 방법을 설명하기 위한 흐름도이다.
- [0078] 도 6을 참조하면, 본 발명의 다른 실시예는 제어부(150)의 문서등록모듈(153)이 기 설정된 주기 혹은 기 설정된 시간이 도래한 것을 감지한 때로부터 시작된다. 문서등록모듈(153)은 S310 단계에서 기 설정된 주기 혹은 기 설정된 시간이 도래한 것을 감지하면, S320 단계에서 지정된 폴더를 스캔하여 그 폴더 내에 저장된 문서 중 아직 등록되지 않은 문서(미등록 문서)를 추출한다. 여기서, 폴더는 문서 관리 장치(100) 자신의 폴더 혹은 네트워크로 연결된 다른 문서 관리 장치(100)의 폴더가 될 수 있다. 그런 다음, 아래의 S330 단계 내지 S390 단계는 지정된 폴더 내의 모든 미등록 문서에 대해서 수행된다.
- [0079] 즉, 문서등록모듈(153)은 S330 단계에서 문서에서 키워드를 추출한다. 이어서, 문서등록모듈(153)은 S340 단계에서 기본어 테이블에서 앞서(S330) 추출된 키워드에 매칭되는 기본어를 추출한다. 이어서, 문서등록모듈(153)은 S350 단계에서 복합어 테이블 및 파생어 테이블로부터 앞서(S340) 추출된 기본어와 연관된 복합어 및 파생어를 추출하고, S360 단계에서 앞서 추출된 기본어, 복합어 및 파생어를 후보 태그로 선정한다.
- [0080] 다음으로, 문서등록모듈(153)은 S370 단계에서 메타데이터 라이브러리와 매칭되지 않지만, 문서 내 기 설정된 횟수 이상 출현한 다출현 키워드를 후보 태그로 선정한다.
- [0081] 이어서, 문서등록모듈(153)은 S380 단계에서 복수의 후보 태그 중 적어도 하나를 기 설정된 규칙에 따라 해당 문서에서 사용할 태그로 결정한다. 일례로, 문서등록모듈(153)은 복수의 후보 태그 중 기 설정된 횟수 이상 출현한 것을 해당 문서에서 사용할 태그로 결정할 수 있다. 다른 예로, 문서등록모듈(153)은 복수의 후보 태그 중 출현 횟수가 많은 순으로 소정 수를 해당 문서에서 사용할 태그로 결정할 수 있다.
- [0082] 이와 같이, 태그가 결정되면, 문서등록모듈(153)은 S390 단계에서 태그와 문서를 매핑하여 저장부(140)에 저장하는 등록을 수행한다.
- [0083] 전술한 바와 같이, 메타데이터 라이브러리를 이용하여 태그를 도출하고 그 태그를 문서에 부여하여 등록할 수 있다. 그러면, 메타데이터 라이브러리와 태그를 이용하여 문서를 검색하는 방법에 대해서 설명하기로 한다. 도 7은 본 발명의 일 실시예에 따라 메타데이터 라이브러리를 이용하여 문서를 검색하는 방법을 설명하기 위한 흐름도이다. 도 8은 본 발명의 일 실시예에 따라 메타데이터 라이브러리를 이용하여 문서를 검색한 결과를 보이는 화면 예이다.
- [0084] 도 7을 참조하면, 사용자의 선택에 따라 검색 프로세스가 시작되면, 제어부(150)의 문서검색모듈(155)은 S410 단계에서 검색 화면을 표시한다. 검색 화면은 검색어를 입력할 수 있는 UI인 대화 상자(Dialogue Box)를 포함한다. 이에 따라, 사용자는 입력부(120)를 통해 검색어를 입력할 수 있다. 그러면, 문서검색모듈(155)은 S420 단계에서 입력부(120)를 통해 검색어를 입력 받고, S430 단계에서 메타데이터 라이브러리의 각 테이블과 검색어를 매칭한다.
- [0085] 이때, S440 단계에서 검색어와 매칭되는 기본어가 존재하는지 여부를 판단한다. 즉, 표 4와 같은 기본어 테이블을 기초로 검색어와 매칭되는 기본어가 존재하면, 문서검색모듈(155)은 S470 단계로 진행하여, 매칭된 기본어를 이용하여 파일명, 블록명 및 태그 검색을 수행한다. 즉, 이 단계에서 문서검색모듈(155)은 메타데이터 라이브러리를 참조하여 저장부(140)로부터 매칭된 기본어, 그 기본어의 파생어 또는 유사어를 파일명으로 가지거나, 매

칭된 기본어, 그 기본어의 파생어 또는 유사어를 블록명으로 가지거나, 매칭된 기본어, 그 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색한다. 반면, 검색어와 매칭되는 기본어가 존재하지 않으면, 문서검색모듈(155)은 S450 단계로 진행한다.

[0086] 한편, S450 단계에서 문서검색모듈(155)은 검색어와 매칭되는 유사어 혹은 파생어가 존재하는지 여부를 판단한다. 즉, 표 5와 같은 유사어 테이블 및 표 6과 같은 파생어 테이블을 기초로 검색어와 매칭되는 파생어 혹은 유사어가 존재하는지 판단한다. 이에 따라, 파생어 혹은 유사어가 존재하면, 문서검색모듈(155)은 S460 단계로 진행하고, 그렇지 않으면, S480 단계로 진행한다. S460 단계에서 문서검색모듈(155)은 파생어 혹은 유사어인 검색어를 기본어로 변환한다. 그런 다음, 문서검색모듈(155)은 S470 단계에서 앞서 변환된 기본어를 이용하여 파일명, 블록명 및 태그 검색을 수행한다. 즉, 문서검색모듈(155)은 메타데이터 라이브러리를 참조하여 저장부(140)에서 변환된 기본어, 그 기본어의 파생어 또는 유사어를 파일명으로 가지거나, 변환된 기본어, 그 기본어의 파생어 또는 유사어를 블록명으로 가지거나, 변환된 기본어, 그 기본어의 파생어 또는 유사어가 태그로 부여된 문서를 검색한다. 예컨대, 표 6을 참조하면, 문서검색모듈(155)은 검색어로 '흡음재'가 입력되면, 유사어의 기본어인 방음재, 방음재의 유사어, 방음재 혹은 흡음재를 파일명으로 가지는 문서, 방음재, 방음재 또는 흡음재를 블록명으로 가지는 문서, 방음재, 방음재 또는 흡음재가 태그로 부여된 문서를 검색한다.

[0087] 한편, 검색어가 기본어, 파생어, 유사어 어디에도 매칭되지 않을 때, 문서검색모듈(155)은 S480 단계에서 입력된 검색어 그대로 파일명, 블록명, 태그 검색을 수행한다. 즉, 문서검색모듈(155)은 저장부(140)에서 검색어를 파일명으로 가지거나, 검색어를 블록명으로 가지거나, 검색어가 태그로 부여된 문서를 검색한다.

[0088] 전술한 바와 같이, 검색이 완료된 후, 문서검색모듈(155)은 S490 단계에서 만약, 존재하는 경우, 기본어 혹은 검색어에 대한 복합어를 관련 검색어로 도출한다. 이때, 문서검색모듈(155)은 먼저, 기본어 테이블로부터 검색에 사용된 기본어의 상위 기본어를 도출하고, 복합어 테이블로부터 도출된 상위 기본어를 속성으로 가지는 모든 복합어를 관련 검색어로 도출할 수 있다. 즉, 문서검색모듈(155)은 복합어 테이블로부터 도출된 상위 기본어를 포함하는 모든 복합어를 관련 검색어로 도출하여 제공할 수 있다.

[0089] 보다 자세히, 관련 검색어를 도출하는 방법에 대해 표 3 및 표 4를 참조하여 설명한다. 표 3 및 표 4를 참조하면, 각각의 복합어는 속성으로 기본어에 매핑되는 기본어의 코드들을 가지며, 이를 통해 문서검색모듈(155)은 속성에 포함된 기본어 코드를 가지는 다른 복합어를 관련 검색어로 추출할 수 있다. 예컨대, 사용자가 검색을 할 때, 검색어로 '비계(코드: W0101)'를 입력했다고 가정하면, 기본어 '비계'의 상위 기본어인 '가설(코드: W01)'을 속성(속성 1)으로 가진 복합어 '가설복공', '가설복공계획', '가설복공구조', '가설복공내역' 등을 관련 검색어로 추출할 수 있다. 이와 같이, 본 발명은 위계를 가진 기본어들을 속성으로 활용하기 때문에, 기존 분류 체계에서는 불가능한 레벨1 분야가 다른 '원가' 분야의 '내역(C0301)'을 속성으로 가진 '가설복공내역'을 관련 검색어로 제시할 수 있다.

[0090] 마지막으로, 문서검색모듈(155)은 S500 단계에서 검색 결과를 표시부(130)를 통해 표시한다. 검색 결과는 기본적으로, S470 단계 및 S480 단계의 검색 결과를 포함한다. 즉, 기본어 혹은 검색어를 파일명으로 가지거나, 기본어 혹은 검색어를 블록명으로 가지거나, 기본어 혹은 검색어가 태그로 부여된 문서를 표시한다. 이러한 화면 예를 도 8에 도시하였다. 도면 부호 81이 지시하는 바와 같이, 검색어 '감리'에 대해 검색된 문서(문서 파일)이 정렬되어 표시된다. 또한, 검색 결과는 S490 단계에서 도출된 관련 검색어를 더 포함할 수 있다. 도면 부호 83은 도출된 관련 검색어를 보인다.

[0091] 한편, 본 발명의 실시예에서 검색된 문서는 다음과 같은 정렬 방식을 이용하여 정렬될 수 있다. 문서를 등록할 때 빈도가 높은 태그는 그 문서를 가장 잘 설명하고 있다고 할 수 있다. 이에 따라, 본 발명의 일 실시예에 따르면, 태그일치도 점수를 이용한다. 태그일치도는 문서에서 다른 태그의 출현 빈도 대비 해당 태그(검색어)의 출현 빈도를 나타내며, 이를 점수로 나타낸 것이 태그일치도 점수이다. 일 실시예에 따르면, 문서검색모듈(155)은 검색된 모든 문서에 대해 다음의 수학적 식 1에 따라 태그 일치도 점수를 구하고, 태그 일치도 점수가 높은 순으로 문서를 정렬하여 표시부(130)를 통해 표시한다.

수학식 1

$$S_T = F_T / \sum_{T=1}^n F_{Ti}$$

[0092]

[0093] S_T : 태그 일치도 점수

[0094] F_T : 태그의 출현 빈도

[0095] 예를 들어 '감리'라는 검색어로 검색하였을 때 '감리'라는 태그를 포함하는 문서로 문서 A와 문서 B가 검색되었다고 가정한다. 문서 A의 태그는 '감리(빈도: 8)', '제도(빈도: 3)', '시공(빈도: 7)'이며, 문서 B의 태그는 '감리(빈도: 6)', '건설(빈도: 2)', '보고서(빈도: 3)'이라고 가정한다. 이때, '감리'라는 검색어에 대해 문서 A의 태그일치도 점수는 수학식 1에 따라 '감리(빈도: 8)' / ['감리(빈도: 8)' + '제도(빈도: 3)' + '시공(빈도: 7)'] = 8/18 = 0.44점이다. 또한, 문서B의 태그일치도 점수는 '감리(빈도:6)' / ['감리(빈도:6)' + '건설(빈도:2)' + '보고서(빈도:3)'] = 6/11 = 0.54점이다. 따라서 문서검색모듈(155)은 '감리'라는 검색어의 검색결과, 문서 B를 문서 A 보다 검색 결과에서 상위에 위치시킨다.

[0096] 한편, 본 발명의 다른 실시예에 따르면, 정렬을 위한 파라미터로 태그일치도 뿐만 아니라, 파일명 일치도 및 문서 평점을 더 사용할 수 있다. 여기서, 파일명 일치도는 검색어와 파일명(블록명)의 일치하는 정도이다. 파일명 일치도 점수는 검색어가 파일명 혹은 블록명에 있는 단어에 포함이 되는 경우를 1점, 아니면 0점을 부여한다. 문서 평점은 사용자들이 해당 문서를 열람한 후 해당 문서에 대한 효용성을 점수로 부여한 것이다.

[0097] 예를 들면, 검색어는 '감리'이며, 검색 결과, 문서 1 및 문서 2가 검색되었다고 가정한다. 여기서, 문서 1의 파일명은 '실시설계 시공성 검토'이며, 블록명은 파일명과 동일한 '실시설계 시공성 검토'라고 가정한다. 또한, 문서 2의 파일명은 '유지관리 지침계획'이며, 블록명은 '안전 및 유지관리 계획'이라고 가정한다.

[0098] 또한, 문서 1 등록 시, 부여된 태그 및 그 태그의 빈도는 다음의 표 7과 같고, 문서 2 등록 시 부여된 태그 및 그 태그의 빈도는 다음의 표 8과 같다고 가정한다.

표 7

[0099]

태그	빈도
시설	12
건설	5
보고	3
감리	3
토지	1
건축	1
계	25

표 8

[0100]

태그	빈도
시설	5
건축	4
보고	4
건축물	2
감리	1
계	16

[0101] 본 발명의 다른 실시예에 따르면, 문서검색모듈(155)은 다음의 수학식 2와 같이, 파일명 일치도, 태그명 일치도

및 문서 평점을 포함하는 3개의 파라미터에 대한 점수에 가중치를 부여하고, 가중치가 부여된 점수를 합산하여 문서 랭킹 점수를 도출하고, 문서 랭킹 점수가 높은 순으로, 문서의 검색 결과를 정렬하여 표시한다.

수학식 2

$$S = W_a S_T + W_b S_F + W_c S_D$$

S : 문서 랭킹 점수

S_T : 정규화된 태그 일치도 점수

S_F : 정규화된 파일명 일치도 점수

S_D : 정규화된 문서 평점

W_a : 태그 일치도에 대한 가중치

W_b : 파일명 일치도에 대한 가중치

W_c : 문서 평점에 대한 가중치

수학식 2를 참조하면 문서 1의 문서 랭킹 점수는 다음의 표 9와 같이 도출될 수 있다.

표 9

문서 1	기본점수A	정규화(max=1) B	가중치C	문서 랭킹 점수D=B×C
파일명(블록명) 일치도	0	0	0.5	0
태그 일치도	0.12	0.12	0.4	0.048
문서 평점	10	1	0.1	0.1
				0.148

먼저, 문서검색모듈(155)은 파일명(블록명) 일치도, 태그 일치도, 문서 평점의 점수를 구한다. 문서 1의 파일명 일치도의 기본 점수는 문서 1의 파일명인 '실시설계 시공성 검토'와, 블록명인 '실시설계 시공성 검토' 공히 검색어 '감리'를 포함하지 않기 때문에 0이다. 태그 일치도 점수는 표 7에 따라 $3/25 = 0.12$ 이다. 그리고 복수의 사용자가 부여한 문서 평점(평균 점수)은 10점이라고 가정한다. 그런 다음, 문서검색모듈(155)은 파일명(블록명) 일치도, 태그 일치도, 문서 평점의 점수를 정규화한다. 여기서, 0 내지 1의 점수로 정규화할 수 있다. 파일명(블록명) 일치도, 태그 일치도, 문서 평점의 정규화된 점수는 각각 0, 0.12 및 1이 된다. 이어서, 문서검색모듈(155)은 파일명(블록명) 일치도, 태그 일치도, 문서 평점 각각에 대한 가중치를 부여한다. 본 발명의 실시예에 따르면, 가중치를 조절하여 문서 검색을 사용하는 분야에 따라 파일명 검색이 중요한지, 문서 내용을 기초로 하는 태그명 검색이 중요한지, 또는 문서를 열람한 사용자가 부여한 문서 평점이 중요한지 선택할 수 있다. 여기서, 파일명(블록명) 일치도, 태그 일치도, 문서 평점 각각에 대한 가중치로 0.5, 0.4 및 0.1이 설정되었다고 가정한다. 그리고 문서검색모듈(155)은 부여된 가중치에 따라 수학식 2와 같이 문서 랭킹 점수를 구한다. 문서 1에 대해 구해진 문서 랭킹 점수는 0.148이 된다.

또한, 수학식 2를 참조하면 문서 2의 문서 랭킹 점수는 다음의 표 10과 같이 구해질 수 있다.

표 10

문서 2	기본점수A	정규화(max=1) B	가중치C	문서 랭킹 점수D=B×C
파일명(블록명) 일치도	0	0	0.5	0
태그 일치도	0.063	0.063	0.4	0.025
문서 평점	10	1	0.1	0.1
				0.125

- [0115] 문서 1과 동일한 방법으로 문서 2의 구해진 문서 랭킹 점수는 0.125이다. 이와 같이, 문서 랭킹 점수가 구해지면, 문서검색모듈(155)은 검색된 모든 문서에 대해 문서 랭킹 점수가 높은 순으로 문서를 정렬하여 표시부(130)를 통해 표시한다.
- [0116] 전술한 바와 같이, 본 발명의 다른 실시예에 따르면, 문서검색모듈(155)은 검색 결과를 정렬하여 제공할 때, 검색된 문서에서 다른 태그의 출현 빈도 대비 검색어인 태그의 출현 빈도를 나타내는 태그 일치도, 검색된 문서의 파일명 혹은 블록명이 검색어를 포함하는지 여부를 나타내는 파일명 일치도 및 검색된 문서에 대해 사용자에게 부여된 점수인 문서 평점을 고려한다. 특히, 문서검색모듈(155)은 태그 일치도 점수, 파일명 일치도 점수 및 문서 평점 각각에 가중치를 부여하고, 합산 점수, 즉, 문서 랭킹 점수가 높은 순으로 검색 결과를 정렬하여 제공할 수 있다.
- [0117] 다음으로, 본 발명의 다른 실시예에 따른 문서를 검색하는 방법에 대해서 설명하기로 한다. 도 9는 본 발명의 다른 실시예에 따른 문서를 검색하는 방법을 설명하기 위한 흐름도이다. 본 발명의 다른 실시예에 따르면, 분류체계(프로젝트 분류체계, 연도별 분류체계, 공종분류체계 등등)를 가지며, 이러한 분류체계는 임의의 태그는 어느 분류체계 중 어느 하나에 속한다. 즉, 분류체계는 미리 설정된 기준에 상기 태그가 속한 분류를 지정한다. 그리고 이러한 분류체계는 저장부(140)에 미리 저장된다.
- [0118] 도 9를 참조하면, 사용자의 선택에 따라 검색 프로세스가 시작되면, 제어부(150)의 문서검색모듈(155)은 S510 단계에서 검색 화면을 표시한다. 검색 화면은 분류체계를 선택할 수 있는 UI를 제공할 수 있다. 이에 따라, 사용자는 입력부(120)를 통해 분류체계 중 어느 하나의 분류를 선택하는 입력을 할 수 있다. 그러면, 문서검색모듈(155)은 S520 단계에서 입력부(120)를 통해 사용자가 선택한 분류를 입력 받는다.
- [0119] 그러면, 문서검색모듈(155)은 S530 단계에서 분류체계를 참조로 선택된 분류에 매핑되는 모든 태그를 추출한다. 이어서, 문서검색모듈(155)은 S540 단계에서 앞서 추출된 태그가 부여된 문서를 검색한다. 그리고 문서검색모듈(155)은 S550 단계에서 앞서 검색된 문서를 표시부(130)를 통해 표시한다.
- [0120] 한편, 전술한 본 발명의 실시예에 따른 문서 등록 방법 및 문서 검색 방법을 포함하는 메타데이터 라이브러리를 이용하여 문서를 관리를 위한 방법은 다양한 컴퓨터수단을 통하여 관독 가능한 프로그램 형태로 구현되어 컴퓨터로 관독 가능한 기록매체에 기록될 수 있다. 여기서, 기록매체는 프로그램 명령, 데이터 파일, 데이터구조 등을 단독으로 또는 조합하여 포함할 수 있다. 기록매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 예컨대 기록매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함한다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 와이어뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 와이어를 포함할 수 있다. 이러한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0121] 이상 본 발명을 몇 가지 바람직한 실시예를 사용하여 설명하였으나, 이들 실시예는 예시적인 것이며 한정적인 것이 아니다. 이와 같이, 본 발명이 속하는 기술분야에서 통상의 지식을 지닌 자라면 본 발명의 사상과 첨부된 특허청구범위에 제시된 권리범위에서 벗어나지 않으면서 균등론에 따라 다양한 변화와 수정을 가할 수 있음을 이해할 것이다.

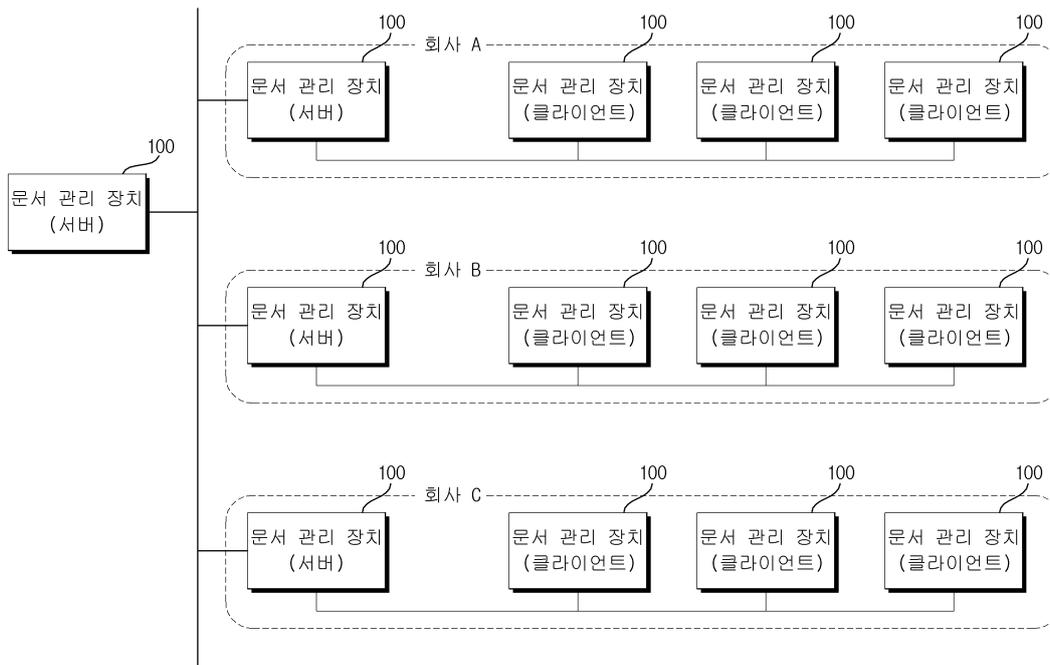
부호의 설명

- [0122] 100: 문서 관리 장치
- 110: 인터페이스부
- 120: 입력부
- 130: 표시부
- 140: 저장부

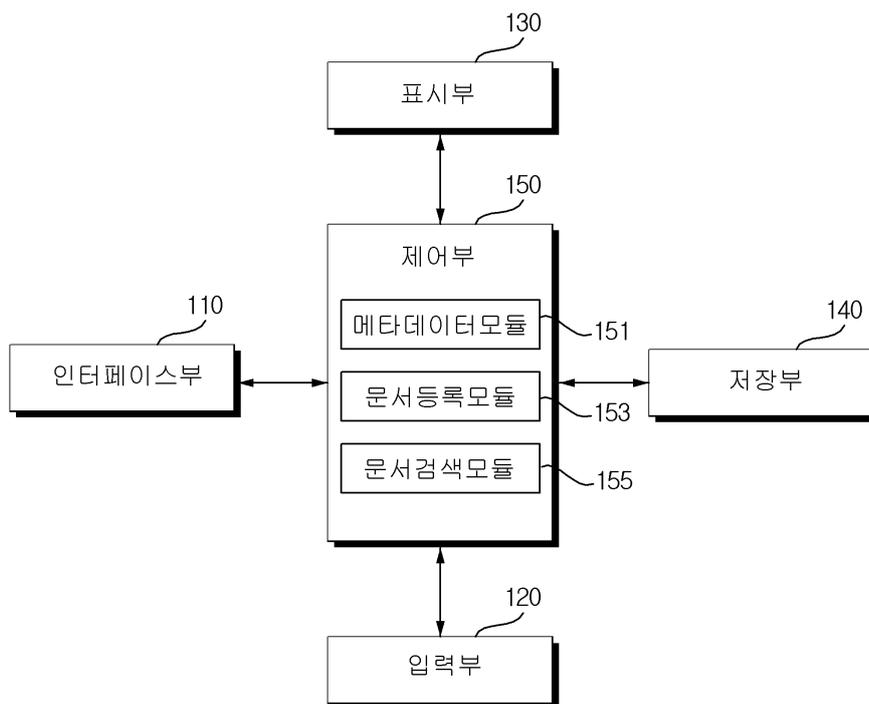
- 150: 제어부
- 151: 메타데이터모듈
- 153: 문서등록모듈
- 155: 문서검색모듈

도면

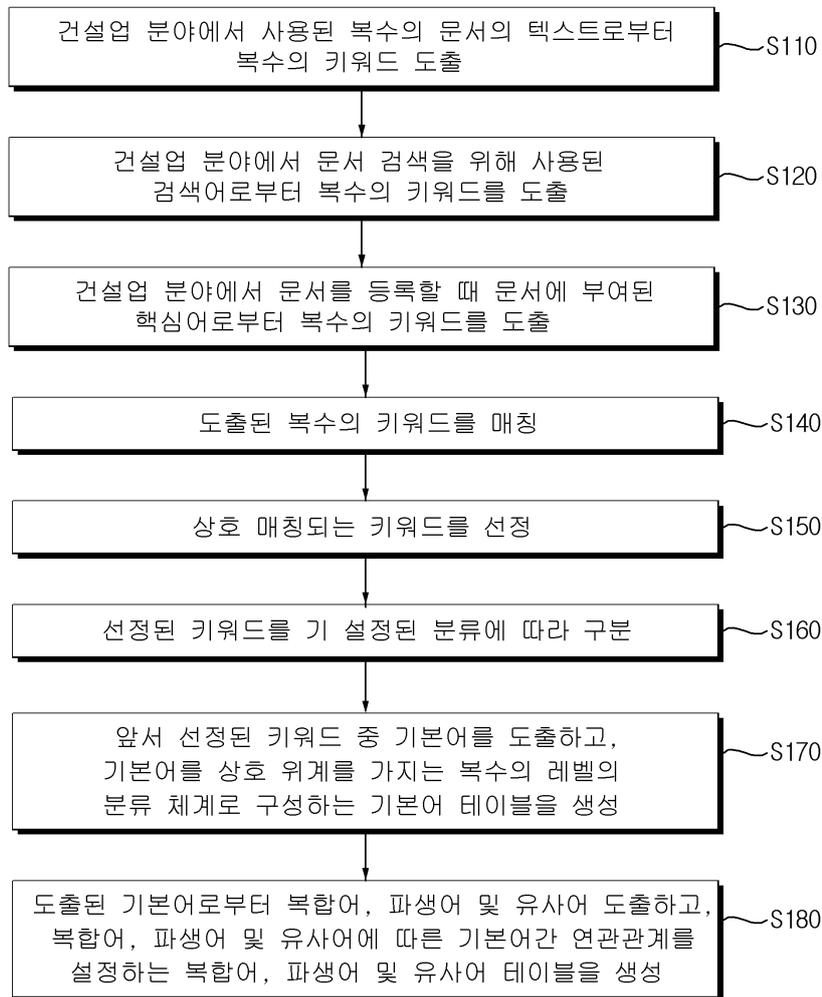
도면1



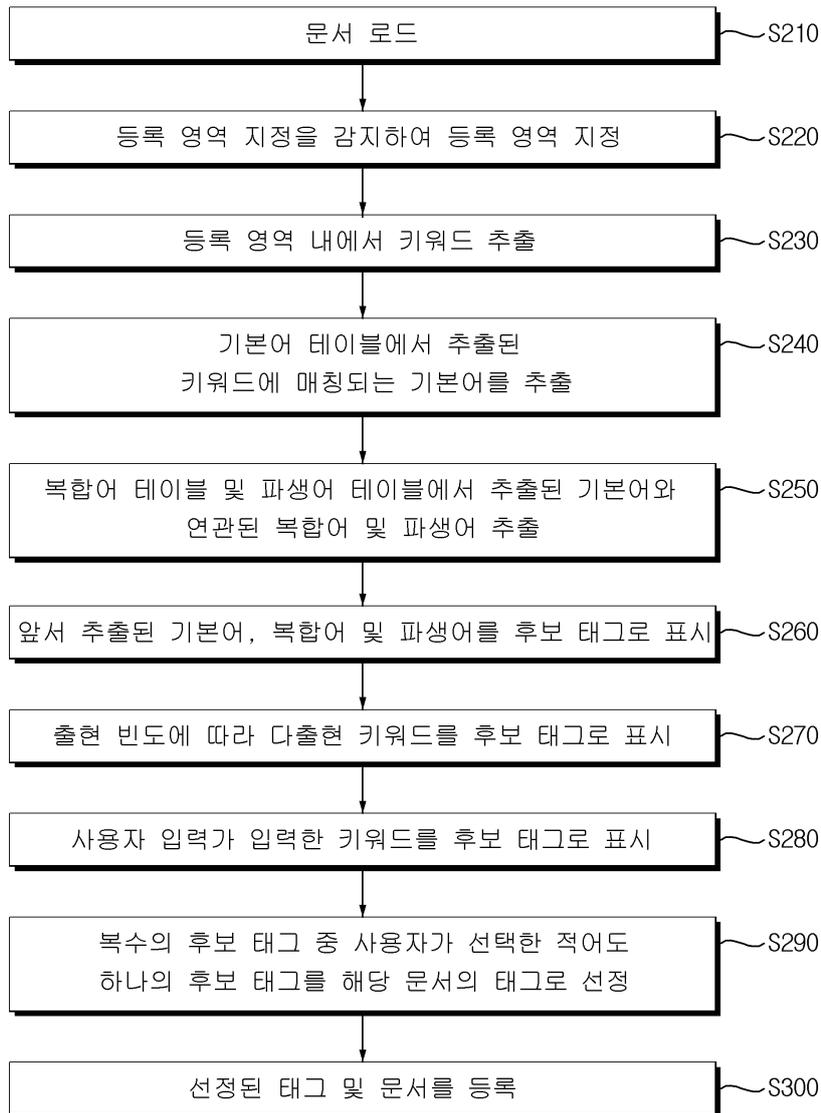
도면2



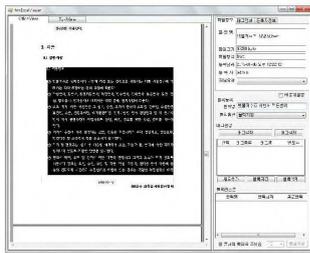
도면3



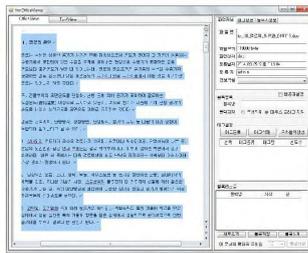
도면4



도면5



HWP 파일



DOC 파일



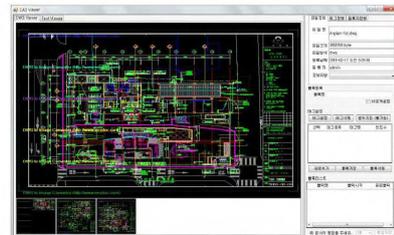
XLS 파일



PPT 파일

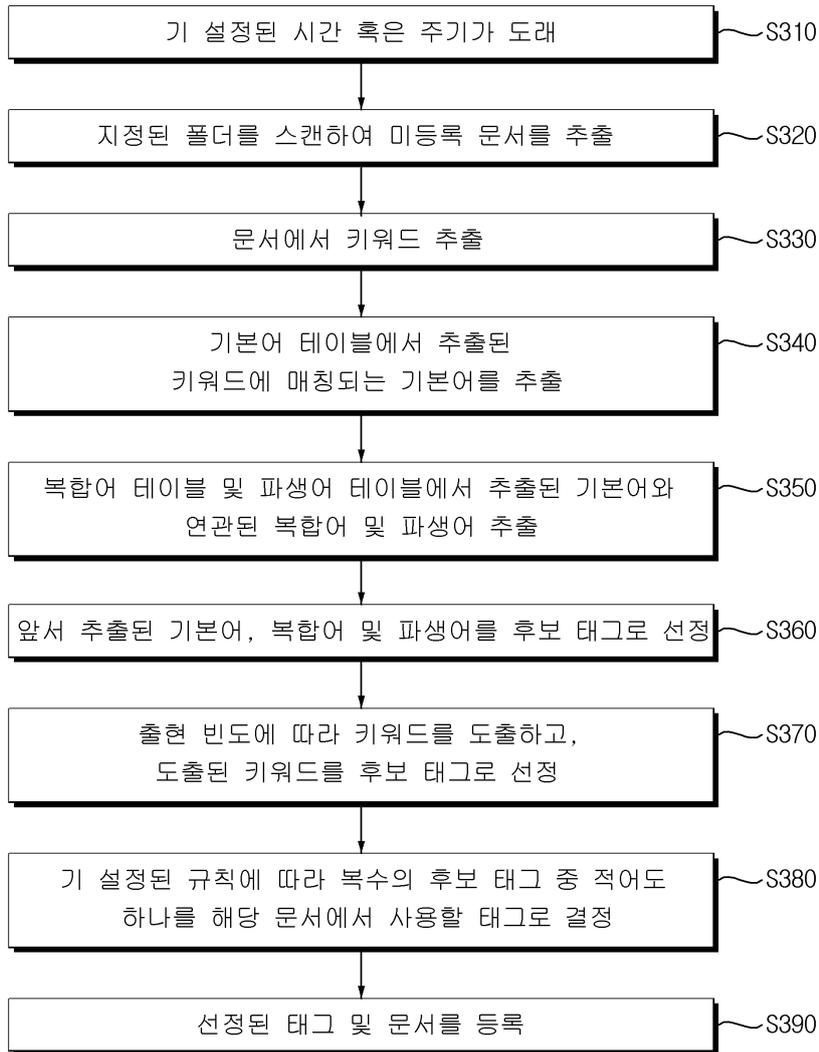


PDF 파일

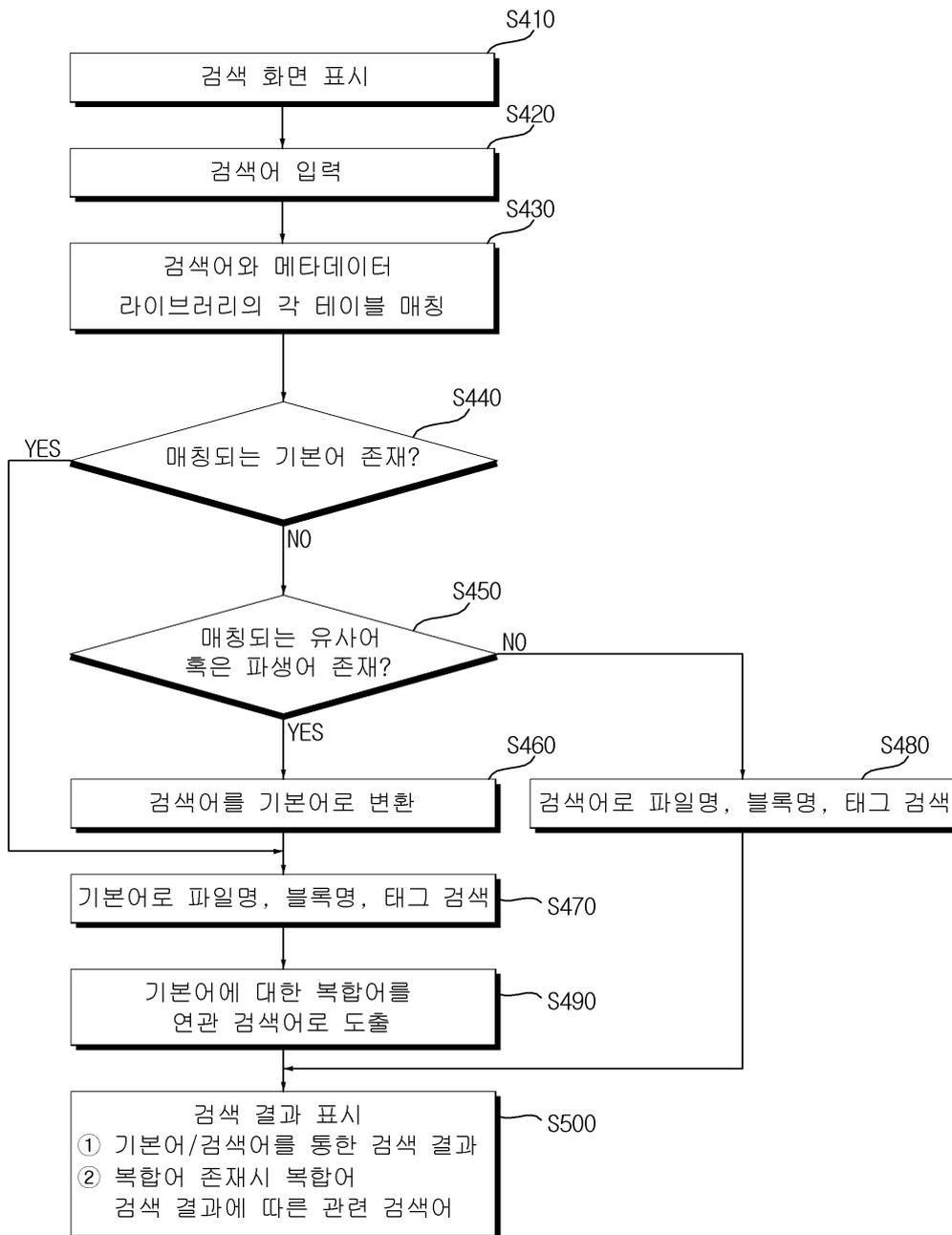


DWG 파일

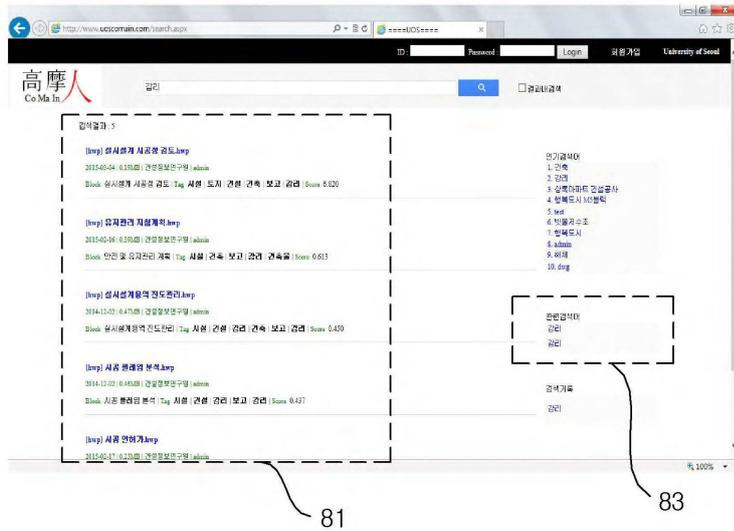
도면6



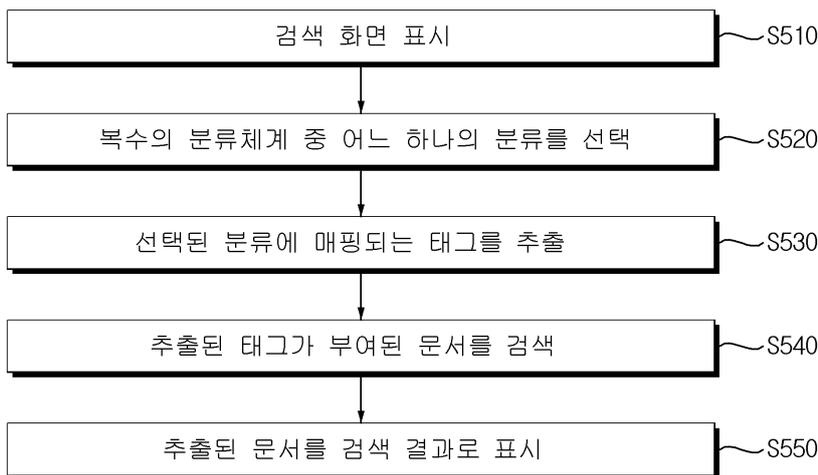
도면7



도면8



도면9



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 20항

【변경전】

상기 검색어가 입력되면

【변경후】

검색어가 입력되면