

건축 자재 유사도 검색 시스템 개발

Similarity Searching System Development for Construction Material

최 병 일* 이 은 지* 고 용 호** 한 승 우***
Choi, Byoung-Il Lee, Eun-Ji Ko, Yong-Ho Han, Seung-Woo

Abstract

The appropriate selection of a construction material is important considering the effect on the total project cost and quality. However, the methodologies of the existing systems and databases to suggest search results have been analyzed to lack reliability. Therefore, this study suggests a material searching methodology that results in quantitative results using similarity.

키 워 드 : 건축자재, 검색 시스템, 유사도
Keywords : Construction material, Searching System, Similarity

1. 서 론

1.1 연구의 목적

현재 세계적으로 전 산업에 걸쳐 정보의 공유 및 효율적 관리를 통한 생산성 향상에 대한 광범위한 연구가 진행되고 있다¹⁾. 하지만 건축 자재들은 같은 자재라 하더라도 사용하는 곳마다 다르게 불리기도하고, 자재 검색을 할 때 한 음절이라도 틀리면 검색이 안 되는 경우가 있다. 건설 분야에서 정보 공유가 미미한 이유로 다양한 형식의 자료와 서로 다른 시스템 환경, 정보 교환과 공유를 위한 표준 및 요소기술의 부재 때문이다²⁾. 그러므로 본 연구에서는 유사도를 이용한 건축 자재를 효율적으로 검색할 수 있는 시스템을 개발하는 것을 목적으로 한다.

2. 건축 자재 유사도 검색 시스템 format 제시

2.1 자재 검색 순서도

그림 1은 시스템의 순서를 순차적인 다이어그램으로 나타낸 것이다.

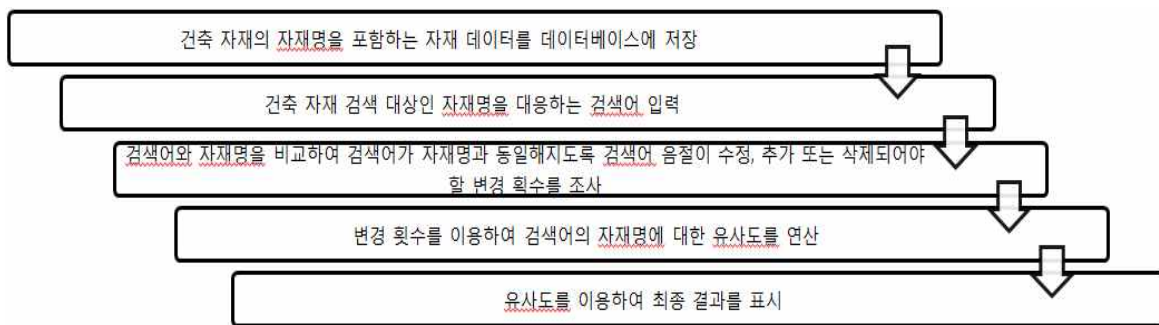


그림 1. 순서도

2.2 검색 및 유사도 설정

그림 2는 검색화면의 예시이다. 유사도는 $0 \leq \text{유사도} \leq 100$ 로 숫자 값으로 나타내고 검색어와 데이터 베이스에 저장 되어있는 자재명과 같아지기 위한 수정, 추가, 삭제에 대한 가중치를 기본 값으로 설정하거나 사용자가 직접 설정할 수 있다.

* 인하대학교 건축공학과 학사과정
** 인하대학교 건축공학과 석사과정
*** 인하대학교 건축공학과 부교수, 교신저자(shan@inha.ac.kr)

그림 2. 검색 화면

2.3 차이도, 가중 차이도

아래 식(1)과 (2)는 차이도와 가중차이도를 구하는 방법이다. 수정, 추가, 삭제 횟수는 사용자가 입력한 검색어가 데이터 베이스에 저장되어있는 자재명과 같아지기 위해 변화하는 음절의 수를 말한다. 식(2)에 X, Y, Z는 그림 2에서 사용자가 입력한 수정, 추가, 삭제에 따른 각각의 가중치를 말한다. $X, Y, Z \leq 1$ 의 조건을 가지고 있다.

$$\text{차이도} = \text{수정 횟수} + \text{추가 횟수} + \text{삭제 횟수} \text{ ----- (1)}$$

$$\text{가중차이도} = X \times \text{수정 횟수} + Y \times \text{추가 횟수} + Z \times \text{삭제 횟수} \text{ ----- (2)}$$

두 단어가 완벽히 일치하면 차이도는 0, 유사하지 않을수록 0에서 멀어진다. 가중치를 기본값으로 선택했으면 식(1), 직접 입력했다면 식(2)에 대한 값이 나오게 된다.

2.4 데이터 베이스 분류 및 유사도 계산

데이터 베이스를 총 6개로 나누어서 3개는 차이도를 사용할 때와 3개는 가중차이도를 사용할 때로 나누며 그 안에 각각의 3개의 데이터 베이스는 검색어의 음절수가 자재명의 음절수보다 많을 때, 같을 때, 적을 때로 나누어서 그림 3와 같은 계산식을 이용하여 유사도를 %로 나타낸다. 그렇게 되면 최종적으로 2.2에서 사용자가 입력한 유사도보다 높은 값의 유사도를 지닌 자재 데이터들이 나오게 된다.

$\begin{aligned} & \text{검색어 음절수} > \text{자재명 음절수} \\ \text{유사도} (\%) &= \frac{\text{검색어 음절수} - \text{차이도}}{\text{검색어 음절수}} \times 100 \\ \\ \text{유사도} (\%) &= \frac{\text{검색어 음절수} - \text{가중차이도}}{\text{검색어 음절수}} \times 100 \end{aligned}$	$\begin{aligned} & \text{검색어 음절수} < \text{자재명 음절수} \\ \text{유사도} (\%) &= \frac{\text{검색어 음절수} - \text{차이도}}{\text{자재명 음절수}} \times 100 \\ \\ \text{유사도} (\%) &= \frac{\text{검색어 음절수} - \text{가중차이도}}{\text{자재명 음절수}} \times 100 \end{aligned}$
--	--

그림 3. 계산식

3. 결 론

본 연구에서 제시한 건축자재 유사도 검색 시스템을 사용하여 사용자가 검색 하였을 때, 검색된 자재와 유사한 데이터를 제공함으로써 효율적인 데이터 검색이 가능할 것이다. 그러나 본 방법론을 적용하기 위해서는 방대한 양의 건축자재 데이터를 저장한 데이터베이스를 구축해야 하며 이를 관리할 인력이 필요할 것으로 판단된다. 따라서 향후 연구에서는 효율적인 데이터 수집 및 검색을 위한 데이터베이스 구축에 대한 연구가 필요할 것으로 사료된다.

Acknowledgement

본 논문은 2014년 한국연구재단 일반연구지원사업(과제번호: NRF-2012R1A2A2A01046193)의 일환으로 수행된 연구임을 밝히며 이에 감사를 드립니다.

참 고 문 헌

1. 송석기 외 3인, 건축자재정보 DB 및 검색시스템 구축 방안에 관한 연구, 대한건축학회 논문집, 계획계 제16권 제6호, pp.59~66, 2000.6
2. 정보처리학회지, 한국정보처리학회, 1997.1