

지능형 검색엔진을 위한 색상 질의 처리 방안*

홍정아

한양대학교 경영학부
(merrycats@hanyang.ac.kr)

구교정

한양대학교 파이낸스경영학과
(dell616@hanyang.ac.kr)

차지원

한양대학교 경영학부
(cha349@hanyang.ac.kr)

서아정

한양대학교 컴퓨터공학부
(rubyrang2@hanyang.ac.kr)

여운영

한양대학교 비즈니스인포매틱스학과
(ywy31730@hanyang.ac.kr)

김종우

한양대학교 경영대학 경영학부
(kjiw@hanyang.ac.kr)

.....

지능형 전자상거래 검색 엔진에 대한 관심이 커지면서, 검색 상품의 특징을 지능적으로 추출하고 활용하기 위한 연구들이 수행되고 있다. 특히 전자상거래 지능형 검색 엔진에서 상품을 검색 할 때, 제품의 색상은 상품을 묘사하는 중요한 특징 중에 하나이다. 따라서 사용자의 질의에 정확한 응답을 위해서는 사용자가 검색하려는 색상과 그 색상의 동의어 및 유의어에 대한 처리가 필요하다. 기존의 연구들은 색상 특징에 대한 동의어 처리를 주로 사전 방식으로 다루었다. 하지만 이러한 사전방식으로는 사전에 등록되지 않은 색상 용어가 질의에 포함된 경우 처리하지 못하는 한계점을 가지고 있다. 본 연구에서는 기존에 사용하던 방식의 한계점을 극복하기 위하여, 실시간으로 인터넷 검색 엔진을 통해 해당 색상의 RGB 값을 추출한 후 추출된 색상정보를 기반으로 유사한 색상명들을 출력하는 모델을 제안한다. 본 모델은 우선적으로 기본적인 색상 검색을 위해 671개의 색상명과 각 RGB값이 저장된 색상 사전을 구축하였다. 본 연구에서 제시한 모델은 특정 색상을 검색하는 것으로 시작하며, 검색된 색상이 색상 사전 내 존재하는 지 유무를 확인한다. 사전 내에 검색한 색상이 존재한다면, 해당 색상의 RGB 값이 기준 값으로 사용된다. 만일 색상사전 내에 존재하지 않는다면, Google 이미지 검색 결과를 크롤링하여 각 이미지의 특정 영역 내 RGB값들을 수집하여 구한 평균 RGB값을 검색한 색상의 기준 값으로 한다. 기준 RGB값을 앞서 구축한 색상 사전 내의 모든 색상의 RGB 값들과 비교하여 각 R, G, B 값에 있어서 ± 50 내의 색상 목록을 정렬하고, RGB값 간의 유클리디안 거리 유사도를 활용하여 최종적으로 유사한 색상명들을 출력한다. 제안 방안의 유용성을 평가하기 위해 실험을 진행하였다. 피설문자들이 생각하는 300 개의 색상 이름과 해당 색상 값을 얻어, 본 연구에서 제안한 방안을 포함한 총 네가지 방법을 통해 얻은 RGB 값들과 피설문자가 지정한 RGB값에 대한 비교를 진행했다. 인간의 눈을 반영하는 측정 기준인 CIELAB의 유클리드안 거리는 평균 13.85로 색상사전만을 활용한 방안의 30.88, 한글 동의어사전 사이트인 워드넷을 추가로 활용한 방안의 30.38에 비해 비교적 낮은 색상 간의 거리 값을 보였다. 연구에서 제시하는 방안에서 수집화 과정을 제외한 방안의 색 차는 13.88로 수집화 과정이 색 차를 줄여준다는 것을 확인할 수 있었다. 본 연구에서는 기존 동의어 처리 방식인 사전 방식이 지닌 한계에서 벗어나기 위해, 사전 방식에 새로운 색상명에 대한 실시간 동의어 처리 방식을 결합한 RGB값 기반의 새로운 색상 동의어 처리 방안을 제안한다. 본 연구의 결과를 활용하여 전자상거래 검색 시스템의 지능화에 크게 기여할 수 있을 것이다.

주제어 : 전자상거래, 색, 이미지 크롤링, RGB, 동의어 처리

.....

논문접수일 : 2018년 11월 2일 논문수정일 : 2019년 2월 27일 게재확정일 : 2019년 2월 28일
원고유형 : 일반논문 교신저자 : 김종우

* 이 논문은 ‘비즈니스 랩 기반의 빅인텔리전스 경영교육 사업단(CK2)’의 지원을 받아서 수행된 연구임.

1. 서론

중국에서 2017년 11월 11일 진행된 온라인 쇼핑물 행사인 광군제에서 알리바바 기업은 하루 동안 약 1682억 위안, 한화로 약 28조원의 매출을 기록하면서 전자상거래 시장에 대한 세계인의 이목을 집중시켰다. 국내 전자상거래 시장 또한 2017년 78조 2,273억원을 기록해 전년대비 19.2%의 성장률을 보였다. 이와 같이 전자상거래 시장 규모는 점점 확대되고 있으며, 전자상거래를 사용한 상품 구매는 보편화 되어, 소비자의 변화에 맞춰 다양한 온라인 쇼핑물들이 끊임없이 개점하고 있다.

전자상거래에서 대부분의 고객과의 인터랙션은 상품 검색을 통해서 이루어진다. 하지만 등록되는 상품 수의 증가로 소비자들은 원하는 상품을 검색하는 데에 어려움을 겪고 있다. 구체적으로 소비자가 원하는 상품명을 세밀한 제품 별 특징 키워드와 함께 입력하는 경우에 원하는 정확한 검색 결과를 얻지 못하는 문제점이 있다. 이로 인해 각 전자상거래들은 지능형 검색엔진 구축을 위한 노력을 기울이고 있으며, 학술적으로는 소비자의 검색 의도를 파악하려는 연구들이 증가하고 있다. 전자상거래 상품 검색에 어려움을 겪는 주요한 원인은 상품 판매자가 상품을 올릴 때 등록한 몇 개의 키워드들과 소비자가 입력한 검색어 간의 텍스트 매칭으로 검색어 처리가 이루어진다는 것이다(Rose et al., 2004; Lei et al., 2006; Tran et al., 2007). 제품의 키워드와 일치하지 않은 검색 키워드는 실제로 검색 조건에 해당되는 상품이 존재함에도 불구하고 정확히 텍스트가 일치하지 않아 검색되지 않는다.

온라인 상품 검색에서의 ‘색상’ 특징은 동의어와 유의어가 많은 대표적인 특징으로 동의어

처리가 중요하다. 예를 들어, 검색자가 ‘붉은’ 제품을 검색하게 되면, 검색자가 찾는 제품은 적어도 ‘레드’, ‘빨강’ 과 같은 유사한 색상이 제목 텍스트에 포함된 제품일 것이다. 하지만, 현재 전자상거래의 검색 시스템은 위와 같은 색상 동의어 처리가 미흡해 유사어가 포함된 검색 결과를 제공하지 않는다. 이러한 문제점을 해결하기 위해 본 연구에서는 기존 전자상거래 검색 시스템에 색상 특징의 새로운 동의어 처리 방안을 제시한다.

현 전자상거래 검색 시스템의 색상 특징 동의어 처리는 대부분 온톨로지 서버를 통하여 검색 질의를 확장하는 방법을 사용한다. 이 방법은 정확도가 높을 수 있으나, 신조어나 모르는 단어가 나오면 찾지 못한다. 그와 함께, 새로운 표현 또는 색상명의 사용이 활발한 요즘, 동의어 사전을 자주 갱신 해야 한다는 한계점이 있다. 전자상거래 검색 시스템은 실시간 색인 기능을 통해 새롭게 추가, 변경, 삭제되는 상품 정보에 발빠르게 대처할 수 있어야 한다. 따라서 본 연구에서는 색상사전과 실시간 할당 방식을 통해 한국어에서의 색상 이름 동의어 간의 관계를 RGB 값을 통해 파악하고 유사 색상명들을 출력해내는 방안을 제시하고자 한다.

본 논문의 구성은 다음과 같다. 서론에 이어서 2장에서 본 연구에 활용된 관련 연구를 설명하며, 3장에서는 본 연구에서 제시하는 발전된 색상 특징 동의어 처리 모델을 제시한다. 이어서 4장에서 모델의 유용성을 확인하기 위한 실험 결과를 서술한다. 마지막 5장에서는 결론을 제시한다.

2. 관련 연구

2.1 지능형 검색 시스템

고객이 상품을 탐색하는 방법이 직접적인 질문을 통해서도 가능한 오프라인 상점과는 달리, 전자상거래 상에서는 텍스트와 키워드를 이용한 검색에 의존한다. 이러한 이유로 고객의 검색어에서 의도를 파악하는 것은 전자상거래 검색 시스템의 성능 향상에 있어서 매우 중요한 사안이다.

현재 국내에서 많이 사용되는 전자상거래들의 검색 시스템을 살펴보면 의미 기반의 검색 시스템이 일반적이다(Kim et al., 2005; Woo et al., 2005). 하지만 전자상거래 업체들은 텍스트 및 키워드를 이용한 일반적인 전자상거래 검색 시스템의 한계를 느끼고 이를 해결하기 위해 다양한 노력을 하고 있다. 현재 일반 검색 시스템 중에서 인공지능과 딥러닝을 활용하여 개체명 인식을 통해 검색한 텍스트의 카테고리 및 특징을 파악하는 기술이 가장 효과적인 검색 시스템 기술로 활용된다. 현재 개체명 인식 기술로 인해 질의 텍스트에서 검색어 별 카테고리를 추출하여 인식함으로써 검색의 의도 파악이 보다 용이해졌다. 애플사와 구글사의 검색 시스템이 개체명 인식을 통한 검색 시스템을 사용하는 것으로 검색 시스템의 시장에서 우위를 점한 것을 따라, 국내 검색 시스템과 전자상거래 또한 본 기술을 착안하여 검색 시스템의 향상을 꾀하고 있다 (Patent No.10-2017-0069606, 2017; Patent No.10-2017-7031186, 2017).

이와 함께 모바일로의 환경 변화에 맞춰 경쟁력 확대를 위해 텍스트 검색뿐만 아니라 다양한 검색 서비스 강화에 나서고 있다. 인공지능을 활용한 이미지 검색을 도입하는가 하면 음성에 중

점을 둔 ‘보이스 커머스’ 시대도 준비되고 있다 (Patent No.10-2015-7004968, 2015; Davis et al., 2016; Turban et al., 2018).

2.2 동의어 처리

현재 전자상거래 검색 시스템 동의어 처리의 경우, 일반적으로 온톨로지 기법을 활용하여 동의어가 처리된다(Gómez et al., 2006; Sudeepthi et al., 2012; Qu et al., 2011). 온톨로지는 공유된 개념과 그 개념들 사이의 관계로 표현되는 사전형식의 데이터 저장 모델이 일반적인 구조이다(Hwang et al., 2012). 온톨로지 기법을 이용하여, ‘빨간 의자’를 검색하면 ‘빨간’에 해당되는 동의어사전 내의 단어인 ‘빨강’, ‘레드’와 같이 기본적인 동의어 처리 후 ‘빨강 의자’, ‘레드 의자’로 질의가 확장되며 질의와 상품 이름 간의 텍스트 매칭으로 상품 검색이 되게끔 설정되어 있다. 이때 각 동의어는 한 대표 단어에 따르는 해당 키(key)에 동의어를 하나씩 온톨로지 서버에 추가하는 방식으로 이루어진다(Kim et al., 2012).

온톨로지 방식은 동의어 처리가 완벽하게 이루어질 수 없다는 단점이 존재한다. 텍스트 매칭으로 검색 결과가 도출되는 현 검색 시스템에서, 색상 동의어를 추가하는 방식으로는 기초적인 동의어 처리만 가능해 신조어 또는 모르는 색상 단어가 질의어에 포함되면 검색 결과가 존재하지 않는다. 또한 매번 새로운 색상명은 식별하여 주기적인 갱신을 해야한다는 한계점을 가진다.

2.3 RGB값의 활용

RGB 가산혼합은 빛의 삼원색인 빨강(Red), 초록(Green), 파란(Blue)을 이용하여 색을 표현하는 방식이다. R, G, B 각 값은 0부터 255 사이의 값

을 가지며, 1,680만 가지의 색상 표현이 가능하다. 이미지에서 색상을 추출해 내기 위한 방법으로 다양한 연구에서 이미지 크롤링이 사용된다(Lin et al., 2013; Cao et al., 2016; Cho et al., 2011). 그 한 예시로, 상품 이미지 상에서 RGB값을 추출해 색인하는 기존의 기술이 있다(Patent No.10-2007-0115690, 2007). 위 기술은 등록된 상품 이미지 별로 이미지 내의 지배적인 RGB값을 추출하여 대표 색상코드로 지정한다. 지정한 대표 색상 코드를 저장하고, 그에 해당되는 대표적인 색상 테이블 내의 색상명(키워드)에 대응시킨다.

하지만 색상코드 추출 방식, 색상코드와 색상명 간의 대응에는 한계가 있다. 이미지의 색상코드에 따라 색상명을 대응시키는 과정에서 대응시킬 수 있는 색상명 키워드는 한정되어 있기 때문에, 이미지에 적합한 색상 키워드가 대응되어 적용된다고 보기 어렵다. 또한 검색을 했을 때 색상코드에 따라 지정된 색상명 키워드에 해당되지 않는 색상명을 검색하는 경우에 있어서는 검색 결과가 여전히 부족하다.

2.4 DBSCAN 군집화

DBSCAN 군집화는 데이터 클러스터링 알고리즘의 일종으로 K-평균 알고리즘이 거리 기반 군집 모델이라면 DBSCAN 밀도 기반 군집화 방법이다(Chakraborty et al., 2014). 어느 포인트를 기준으로 지정 반경 내에 데이터 포인트가 n개 이상 있으면 하나의 클러스터로 인식하는 방식이다. 만약 임의로 선정한 하나의 데이터 포인트에서부터 지정 거리 이내에 포인트 수가 최소 샘플 수보다 적다면 그 포인트는 어떤 클래스에도 속하지 않는 잡음(Noise)으로 레이블 된다. 최소 샘플 수보다 많으면 그 포인트는 핵심 샘플

플로 레이블 되고 새로운 클러스터 레이블에 할당된다. 그 다음 그 포인트의 지정 거리 안의 모든 이웃을 살펴 아직 어떤 클러스터에도 할당되지 않았다면 바로 전에 만든 클러스터 레이블에 할당된다. 이미 할당이 완료된 핵심 샘플이면 그것의 이웃 포인트로 넘어가서 일련의 과정을 반복한다.

DBSCAN의 경우 K-평균 군집화와 달리 클러스터의 수를 지정하지 않아도 된다는 장점이 있다. 그리고 K-평균 군집화에서는 적절히 클러스터링 될 수 없는 비선형의 클러스터를 찾을 수 있어 본 연구에서는 DBSCAN 군집화를 활용했다.

2.5 유클리디안 거리 유사도

유클리디안 거리 유사도는 2-노름 거리(L2 Distance)라고도 불리며, 일반적으로 두 점 사이의 거리를 자로 재었을 때의 "직관적인" 거리 값을 나타낸다. 유클리디안 거리 유사도 값은 거리 값을 보여주므로 최대 값은 존재하지 않지만 최소 값은 같은 위치에 있는 두 점인 경우로 0이다. 그러므로 유사도 값이 0에 가까울수록 두 점이 가깝다는 것을 나타낸다. n차원의 튜플에 대해서 유클리디언 거리는 아래와 같은 수식을 통해 계산된다.

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (1)$$

$$\text{where } x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n)$$

본 연구에서 유클리디안 거리 유사도는 RGB 값의 DBSCAN 군집화 과정과 최종 출력을 위한 색상명의 정렬에 활용되었다.

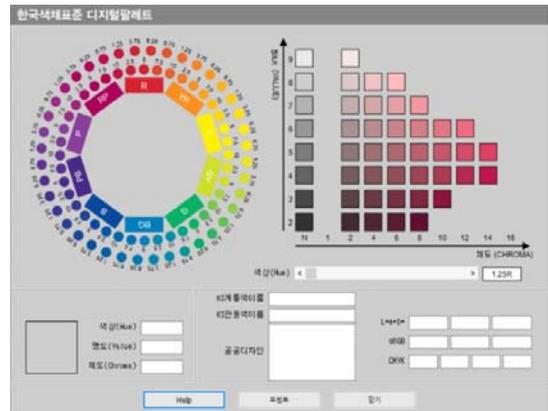
3. 제안 모델

본 절에서는 이미 출현한 색상 동의어 처리 기술을 기반으로 하여, 실시간 색상 동의어처리 기술을 융합한 방안을 제시한다. 본 연구에서 제안하는 모델은 편의상 CQP-RGB (Color-related Query Process based on RGB Values)로 명명하였다. CQP-RGB는 다음의 <Figure 1>에서 제시된 것처럼 구현된다.

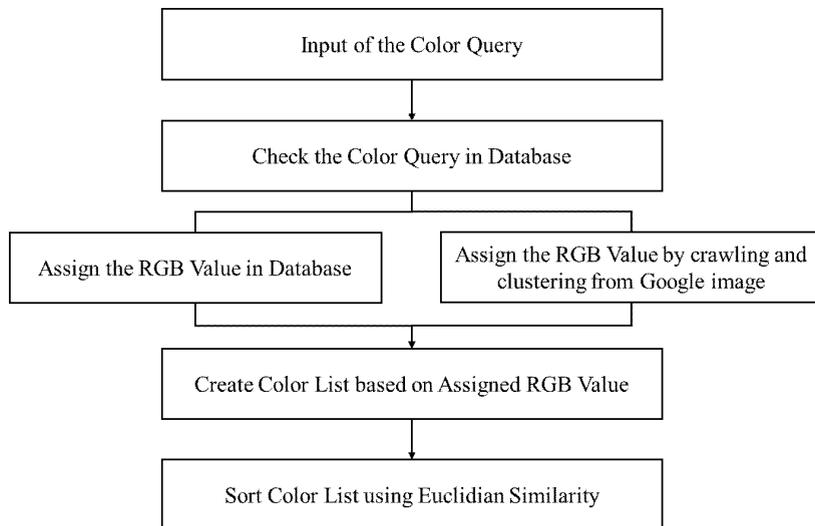
3.1 데이터베이스 구축

본 연구에서 제안하는 CQP-RGB의 기본적인 색상 검색, 동의어 및 유의어 출력을 위해 기본 색상 데이터베이스가 필요했다. 이를 위해서 <Figure 2>와 같은 공인된 한국 색채 표준 디지털 팔레트 프로그램의 데이터베이스, 그리고 <Figure 3>의 위키피디아(Wikipedia) 색 목록에 정리된 색상표의 색상명과 RGB 값을 활용하여 색상 데이터베이스를 구축하였다.

‘민트색’과 ‘화이트’와 같이 흔히 찾을 수 있는 외래어 색상명은 현 전자상거래 제품 색상 검색에서 굉장히 중요한 부분이다. 따라서 영어 색상명을 한글 외래어 표기로 변환한 138개의 색상명과 해당 RGB값 리스트를 추가하여 총 671개의 색상명과 RGB값을 갖게 함으로써 데이터베이스를 보강하였다.



<Figure 2> Korean Standard Color Digital Palette



<Figure 1> Procedure of the Proposed Model

ㄱ [편집]

이름	샘플	RGB (빨강, 초록, 파랑)		
갈색		150	75	0
개나리색		247	230	0
검정		0	0	0
귤색		248	155	0
금색		255	215	0
군청색		70	73	100

ㄴ [편집]

이름	샘플	RGB (빨강, 초록, 파랑)		
남색		75	0	130
노랑		255	212	0
녹색		0	153	0

ㄷ [편집]

이름	샘플	RGB (빨강, 초록, 파랑)		
다홍색		255	36	0
담청색		62	145	181
대남		21	96	189
등색		255	183	76

<Figure 3> Wikipedia 'Color List' search result

<Table 1> Part of final color database

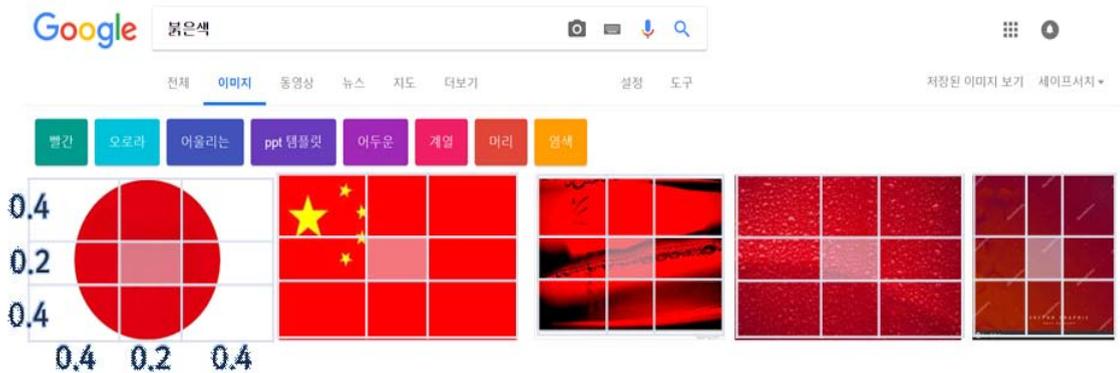
R	G	B	Color
0	0	255	Blue
72	61	139	Dark slate blue
95	158	160	Cardet blue
135	206	235	Sky blue
65	105	225	Royal blue
176	224	230	Powder blue
0	0	128	Navy
0	0	139	Dark blue
138	43	226	Blue violet
139	0	139	Dark magneta
153	50	204	Dark ochid
148	0	211	Dark violet
255	0	255	Magneta
255	0	255	Pattusa
199	21	133	Medium violet red
186	85	211	Medium ochid
147	112	219	Medium purple
220	20	60	Crimson
255	20	147	Deep pink

총 세 개의 색상 데이터 집합을 합침으로 인해 데이터베이스 내에 동일한 이름의 색상이 다른 R, G, B 값을 갖는 경우가 발생했다. 그러한 경우에는, 각 R, G, B 값의 평균 값으로 비교하여 출력할 수 있게 설정하였다. 그 결과 <Table 1>과 같이 총 671개의 색상명과 해당 RGB값을 가진 색상 기본 데이터베이스를 구축하게 되었다.

3.2 RGB값 할당

본 연구에서 제시하는 모형은 특정 색상명을 검색을 하는 것으로 시작된다. 먼저 질의어에 포함된 색상명이 앞서 구축한 색상 데이터베이스에 존재하는지 여부를 확인한다. 데이터베이스 내에 검색한 색상이 존재한다면, 그 색상의 RGB 값을 검색한 색상명의 대표 값으로 한다. 반면, 검색한 색상이 사전에 존재하지 않는다면, 검색한 색상의 Google 이미지 검색 결과 중 상위 5개를 크롤링(crawling)하는 방식을 활용한다. 크롤링한 각 이미지의 가운데 특정 영역인 가로 2/5부터 3/5, 세로 2/5부터 3/5의 직사각형 범위 내의 각 R, G, B 값들의 평균 R, G, B 값을 검색한 색상명의 대표 값으로 한다.

Google 이미지 검색은 일반적으로 그 검색어가 제목 또는 작성 글 내에 있는 이미지를 출력해준다는 것을 확인 할 수 있었다. 그 색상명을 나타내는 이미지와 그 색상을 띄는 상품 이미지가 많은데 상품 이미지는 이미지 전체의 가운데 영역에 배치된 이미지들이 대다수라고 판단하여 가운데 영역의 RGB값 추출을 설정했다. 아래 <Figure 4>은 '붉은색'을 검색했을 때의 예시로 대표 RGB 값을 추출하는 방식과 영역을 보여준다.



<Figure 4> Area where RGB values are extracted

단순히 특정 가운데 영역의 RGB값의 평균을 가져오는 것에는 한계가 존재한다고 판단하여 다른 다양한 방법으로도 시도했다. 이미지 내의 RGB값을 클러스터링 해보고 가장 밀도가 높은 군집의 평균 값을 추출하는 방식을 시도했다. 앞서 기존의 방식으로 추출된 이미지 내의 모든 RGB 값들을 DBSCAN 클러스터링을 해보면, 가장 점들이 많이 몰려 있는 군집 내의 RGB값들이 이미지 내의 지배적이고 색상명에 적합한 색상일 것이라고 판단했다(Ester et al., 1996). K-평균 군집화는 클러스터링을 하기 위해서는 군집 수를 정해야 하는 단점이 있어 RGB값들을 클러스터링 할 때 적합한 밀도 기반 군집화 방법인 DBSCAN 클러스터링을 활용하였다. DBSCAN 클러스터링 매개변수 중 지정거리(epsilon) 값은 지정 거리가 가까운 포인트의 범위를 결정하기 때문에 매우 중요하다. 지정 거리를 매우 작게 하면 모든 포인트가 잡음 포인트가 될 수 있고, 반대로 지정 거리를 매우 크게 하면 모든 포인트가 단 하나의 클러스터에 속하게 된다. 따라서 다양한 색상들의 군집화 결과에 있어서 적절한

군집 개수와 군집의 구성을 보여주는 지정 거리 값인 0.5로 지정하여 군집화 과정을 진행했다.

<Figure 5>는 ‘붉은색’을 검색했을 때, Google 이미지 검색 후 크롤링 된 RGB값들을 DBSCAN 클러스터링으로 클러스터링 한 결과이다. 1, 0, -1은 클러스터를 의미하고, 각 클러스터 뒤의 1195, 562, 16은 클러스터 별 포인트 수를 보여준다. 클러스터링을 통해서 보다 더 정확하고 나은 결과 값을 얻을 수 있었다.

색상명: 붉은색

Counter({1: 1195, 0: 562, -1: 16})

<Figure 5> DBSCAN Cluster result

3.3 색상명 정렬

앞서 언급한 두 가지 중 한 가지 방식으로 지정된 검색한 색상명의 대표 RGB값을 기반으로, 색상 데이터베이스 내의 모든 색상의 R, G, B 값들과 비교한다. 각 R 값, G 값, B 값에 있어서 ± 50 범위 내의 색상들을 출력하게 된다. ± 50 을

기준으로 하였을 때, 가장 적절한 색상 개수를 출력한다고 판단하였다. 범위를 넓게 하여, 많은 색상명을 출력하여도 다음 단계에서 출력되는 색상명의 개수는 다섯개로 범위 내의 차이가 큰 경우에는 결과적으로 출력되지 않을 것이기 때문에, 범위를 넓게 잡는 것은 불필요하다고 판단했다. 반면 범위를 줄여 ±50 미만의 범위로 설정했을 때는, 다음 과정을 거치기도 전에 출력되는 동의어 또는 유의어 색상명이 5개가 되지 않는 경우가 존재하여 적합하지 않은 좁은 범위라고 판단하였다. 앞선 예에 이어 <Figure 6>은 ‘붉은색’의 대표 RGB 값을 기준으로 하여 데이터베이스의 색상들을 비교하여 ±50 범위 내의 모든 색상들을 출력한 결과이다.

```

색 입력: 붉은색
색 추출값 202.0 10.0 18.0
[['빨강' 178 44 46]
 ['빨간' 178 44 46]
 ['빨간색' 178 44 46]
 ['레드' 178 44 46]
 ['심홍색' 220 20 60]
 ['카민' 176 46 57]
 ['토마토색' 176 49 43]]
    
```

<Figure 6> Result of listing colors by RGB range

생성된 색상 목록 안에서 검색한 색상의 대표 RGB값 간의 유클리디안 거리 유사도(Euclidean distance)를 활용한다. 비슷한 RGB값을 구하는 방식에 대한 선행 연구(Mahama et al., 2016)를 참고하여, 유클리디안 거리 유사도가 가장 높은 색상부터 최대 5개까지 출력하게 하며 <Figure 7>과 같은 최종적인 색상 목록이 출력된다. 색상명, R값, G값, B값, 유클리디안 거리 값 순으로 출력을 하며, ‘붉은색’의 RGB값과 가장 가까운 색상부터 출력되는 것을 확인 할 수 있다.

색 입력: 붉은색

색 추출값 202.0 10.0 18.0

```

[['심홍색' 220 20 60 46.776062254106]
 ['빨강' 178 44 46 50.15974481593781]
 ['빨간' 178 44 46 50.15974481593781]
 ['빨간색' 178 44 46 50.15974481593781]
 ['레드' 178 44 46 50.15974481593781]]
    
```

<Figure 7> Result of Color Synonym Solution

5개를 초과하여 유의어나 동의어를 출력하게 하는 것은 보다 많은 색상명의 출력이 되어 오히려 지나치게 확장된 검색 결과를 출력할 것이라고 판단했다. 반면 3개 이하의 색상명을 출력하는 것은 <Figure 5>와 같은 예시를 보아도 알 수 있듯이 ‘빨강’, ‘빨간’, ‘빨간색’, ‘레드’ 모두 같은 의미의 색상명 임에도 불구하고 출력이 안될 수 있는 경우가 발생할 수 있다. 따라서 적정 개수로 판단되는 최대 5가지까지의 색상명이 출력하게끔 설정했다.

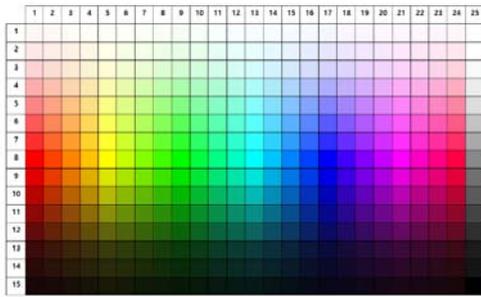
4. 실험 설계 및 결과

4.1 실험 내용

제시한 방안의 유용성을 평가하기 위해 실험을 진행했다. 피설문자들이 전자상거래에서 검색할 만한 제품에 해당하는 색상명과 그 색상에 가장 적합한 색이라고 판단되는 색상표의 색을 선택하게 했다. 자세한 설문문의 내용은 아래 <Figure 8>에서 제시된 것처럼, ‘티셔츠’, ‘페인트’, ‘텀블러’, ‘꽃’, ‘핸드폰 케이스’를 구매한다면 검색하고 싶은 제품 별 색상명과 그 색상명에 적합한 색상표 내의 한가지 색을 선택하게 했다.

1-1. "티셔츠"를 구매할 경우 검색하고 싶은 색상명을 기입해주세요. ex)검정색*

내 답변



1-2. 1-1에 기입한 색상명에 해당되는 색상의 행과 열을 기입해주세요. ex) 15, 25*

내 답변

<Figure 8> First question of the experiment

본 설문은 총 60명의 대학생, 대학원생들을 대상으로 진행했다. 위 설문을 통해서, 한 피설문자 당 제품 별로 한 색상씩 응답을 하여 5개의 색상명을 받아 총 300개의 색상명과 그에 따른 색상 값을 얻을 수 있었다. 설문으로 얻은 HTML 색상표 값을 HTML 값으로 변환 후 두 번의 변환 과정을 거쳐 RGB값을 구했다.

최종적으로 변환한 RGB값이 소비자가 생각하는 색상명에 해당하는 색상 RGB값이라고 판

단하여 이를 기준 값으로 두었다. 피설문자들의 답변 값과 색상 처리에 활용될 수 있는 네 가지 방안을 통해 얻어지는 색상명에 따른 RGB값들 간의 오차율을 비교해 보았다. 색상명에 해당 RGB값 할당 방식에 대한 평가에 사용된 네 가지 방안은 아래 <Table 2>와 같다. C1은 데이터베이스 내 유무에 따라서 존재하는 색상의 경우 색상명의 해당 RGB값을 할당하는 방안이고 존재하지 않는 색상은 랜덤 값을 할당하는 방안이다. C2는 C1의 방안에서 데이터베이스 내 존재하지 않는 색상의 경우에는 WordNet이라는 동의어사전 웹사이트를 통해 얻은 동의어 색상명을 C1방안에 다시 적용하여 더 많은 RGB 값을 할당하는 방안이다. C3은 데이터베이스 내에 존재하는 색상은 해당 RGB값을 부여하고, 데이터베이스 내 존재하지 않는 색상에 대해서는 RGB값을 이미지 크롤링만을 통해 추출해 내는 방안이다. C4는 이미지 크롤링과 DBSCAN 클러스터링을 함께 사용하여 RGB값을 할당하는 방안이다. 각 방안에 있어서 RGB값이 할당되지 않은 색상명에 있어서는 각 R, G, B 값에 0에서 255 내의 랜덤 값을 부여하여 오차율 결과 값을 구할 수 있게끔 설정했다.

본 실험에서 성능 평가를 위해 CIELAB ΔE^* 값을 활용했다. 대체적으로 색 거리 정의는 공

<Table 2> Explanation of 4 methods in experiment

Method	For the Color in Database	For the Color not in Database
C1	Assign RGB values in database	Assign random RGB values
C2	Assign RGB values in database	For the color with WordNet's synonym results, follow the same procedure of C1
C3	Assign RGB values in database	Assign RGB values by crawling from Google images
C4	Assign RGB values in database	Assign RGB values by crawling and clustering from Google Images

간 내의 거리로 거리를 결정하는 표준 방법은 유클리드 거리이다. 유클리드 거리를 활용하는 CIELAB ΔE^* 은 국제 조명 위원회 CIE(Commission Internationale de l'Eclairage)에서 푸른색 영역에서 색 공간이 불균일한 문제 때문에 새롭게 제안한 색상 거리 결정 표준이다. 피설문자들에게 가장 색상명과 유사해 보이는 색상으로의 선택을 부탁한 것과 같은 맥락으로, 인간의 눈이 다른 색보다 특정 색에 더 민감한 것을 반영하여 색상 간의 눈에 띄는 차이를 찾아주는 측정 기준인 CIELAB ΔE^* 값이 실험에 가장 적합한 평가식이라 판단했다.

CIEDE2000은 목적에 따라 조절이 가능한 세 개의 파라미터(k_L , k_C , k_H)의 초기값은 1로 되어 있으나, CIEDE2000이 색 패치를 이용한 실험 데이터를 기반으로 개발이 되었으므로 CIE에서는 디지털 이미지의 색 차를 구하기 위해서 k_L 값에 있어서 조정이 필요하다고 주장한다(CIE Central Bureau et al., 2011). 푸른색 영역의 보정을 위하여 CIELAB ΔE^* 값을 계산하기 전에 L^* , a^* , b^* 값을 이용하는 계산식이 추가되었으며, 실험에 활용한 최종 CIEDE2000의 색차식은 아래와 같다.

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2} + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H} \quad (2)$$

(R_T ; A hue rotation term dealing with the problematic blue region, S_L ; Compensation for lightness, S_C ; Compensation for chroma, S_H ; Compensation for hue)

4.2 실험 결과

각 방법 별 구해진 CIELAB 값을 토대로 대응 표본 t-test를 하여 방법에 따른 색 차가 유의한지

를 검정하였다. 네가지 경우에 대한 실험 결과는 아래 <Table 3>와 같다. 색상명에 해당되는 RGB 값을 데이터베이스를 통해서만 예측하여 CIELAB값으로 변환한 첫번째 경우는 색 간의 거리 평균 값이 약 30.88이고, 데이터베이스 내에 없는 색상명에 대한 RGB값 처리는 불가능하기 때문에 161개의 색상은 처리되지 않았다. 이보다 조금 더 발전된 동의어처리가 이루어진 두번째 경우는 데이터베이스 내의 색상명 해당 RGB값을 이용하거나 없는 경우에는 워드넷에서의 동의어처리 후의 색상명을 데이터베이스에서 유무를 재확인하였다. 하지만 데이터베이스 내에 존재하지 않았던 색상명 중 워드넷에서 동의어 결과가 출력된 색상은 ‘흰색’, ‘보랏빛’, ‘취색’, ‘연초록색’ 뿐이었다. 위 네 개의 색상명에 대하여 동의어를 처리한 색상명들이 데이터베이스 내에 존재하는 경우, 그 색상명에 해당하는 데이터베이스 내 RGB값을 이용했다. 따라서 앞서 데이터베이스만을 이용한 방안에서 피설문자들이 작성한 색상명 17개(‘흰색’ 13개, ‘보랏빛’ 1개, ‘취색’ 2개, ‘연초록색’ 1개)의 RGB값 할당이 추가되어 총 156개의 색상명에 대해 RGB값을 가질 수 있었다. 이어서 CIELAB값으로 변환하여 색 차를 구해본 결과, 전 방안보다 낮은 유클리드 거리의 평균 값을 얻을 수 있었다.

그러나 본 연구에서 제시한 방안은 사전 처리 방식으로는 RGB값을 할당 받지 못하는 색상명들이 여전히 많다는 한계를 보완한다. 공인되거나 확실한 한글 색상명이 정의되어 있지 않아 자주 갱신되지 못하는 데이터베이스와 국문 위주의 동의어 처리가 이루어지는 워드넷 사이트만으로는, 전자상거래에서 많이 검색되는 외래어 색상명 및 새로운 색상명에 대한 동의어 처리가 이루어질 수 없다. 반면에 색상명이 데이터베이

〈Table 3〉 Paired T-Test Result for 3 Cases

Methods	N	Mean	Std. Deviation			
C1	300	30.8858	0.2223			
C2	300	30.3785	0.2138			
C3	300	13.8791	0.1025			
C4	300	13.8515	0.0999			
Pair	Paired Differences			T	df	Sig. (1-tailed)
	Mean	Std. Deviation	Std. Error Mean			
C1 - C2	0.5073	0.1804	0.0104	0.4862	299	0.6272
C1 - C3	17.0067	0.2281	0.0132	12.8894	299	0.0000***
C1 - C4	17.0342	0.2276	0.0132	12.9384	299	0.0000***
C2 - C3	16.4993	0.2210	0.0128	12.9115	299	0.0000***
C2 - C4	16.5269	0.2191	0.0127	13.0423	299	0.0000***
C3 - C4	0.0276	0.0154	0.0009	0.3104	299	0.7565

*** p < 0.01 / ** p < 0.05 / * p < 0.1

스 내 없는 경우에는 Google 이미지 검색 크롤링 또는 크롤링과 군집화를 통해 RGB값을 할당해주는 C3과 C4는 300개의 모든 색상명에 대하여 RGB값을 할당하였다. 하지만 C4의 경우 C3에 DBSCAN 군집화 과정을 추가하여 차이가 통계적으로 유의하지는 않지만, CIELAB값으로 변환하여 구한 색 차는 보다 낮은 값을 보였다. 따라서 피셀문자가 의도한 색상명에 따른 색상 RGB 값과 가장 유사한 색상들의 RGB값을 가지면서, 다른 방안보다 낮은 13.85의 색상 간의 평균 유클리드 거리 값을 보이는 것을 확인할 수 있다. 또한 대응표본 t-test를 통한 검증 결과 본 연구에서 제안한 방안들이 C1, C2에 비해서 유의하다는 것을 확인할 수 있었다.

5. 결론

본 연구에서는 전자상거래 사이트 내 검색 시스템 향상을 위해 상품 검색어 내 색상 특징의 동의어 처리 방안을 제시했다. 기존 동의어 처리 방식인 사전 방식이 지닌 한계에서 벗어나기 위해, 사전 방식에 새로운 색상명에 대한 실시간 동의어 처리 방식을 결합한 RGB값 기반의 새로운 색상 동의어 처리 방안을 고안하였다. 텍스트만을 기준으로 형성되었던 기존 기술에서의 색상 동의어 사전과는 달리, 본 연구에서는 색상 데이터베이스 내 색상명과 그 색상명에 따른 RGB값을 함께 저장했다. 구매자가 색상명을 검색하게 되면, 검색된 색상명이 데이터베이스 내

에 존재하는지 확인하여 두 가지 방식으로 대표 RGB값을 지정했다. 색상 데이터베이스 내에 존재하는 경우는 해당 색상의 RGB값을 대표 RGB값으로 지정하고, 존재하지 않는 경우는 인터넷상의 이미지 검색을 통한 이미지 크롤링으로 대표 RGB값을 지정하였다. 앞서 지정된 검색한 색상의 대표 RGB값을 기준으로 색상 데이터베이스의 모든 색상의 RGB 값들과 비교하여 유사한 색상을 찾아낸다. 또한 본 연구에서는 제시한 방안의 유용성을 확인하기 위한 실험을 수행하였다. 서로 다른 방식으로 300개의 색상명에 RGB값을 할당하였고, 각각의 RGB값을 할당된 RGB값과 피셜문자가 색상명에 적합하다고 판단한 색상값을 CIELAB값으로 변환하여 두 색간의 거리를 구하였다. 이를 대응 표본 t-test를 통해 평가한 결과, 단순히 색상 사전만을 기반으로 하거나, 사전과 한글 동의어 사전을 활용하는 방법보다, 사전과 실시간으로 RGB값을 추출하는 것이 좀 더 RGB값의 정확도를 높이는데 효과적임을 확인하였다. 본 연구에서 제시하는 방안은 기존의 색상 동의어 처리 방식에 비해 사용자의 색상 검색 의도를 보다 정확하고 신속하게 파악하여 색상의 동의어 및 유의어를 제시한다. 제시된 색상 질의 처리 방안은 기존의 단순 사전식 텍스트 매칭에서 발전된 질의 확장 방안으로 전자상거래 검색 시스템에 큰 기여가 될 수 있을 것이다. 게다가 전자상거래 검색 시스템 중 영역이 넓어지고 있는 이미지 검색에서, 이미지 속 상품의 색상에 대한 검색은 본 연구에서 제시한 방법을 통해 정확한 검색 결과를 얻을 수 있을 것이라 기대한다.

본 연구의 한계점으로는 기본 색상 데이터베이스 생성은 불가피하며, 데이터베이스 내의 추가할 수 있는 공인된 색상명과 RGB값이 부족하

다는 것이다. 전자상거래에서 사용되는 트렌드에 따른 새롭고 복잡한 색상명들에 비해 출력되는 색상명들은 원색적이고 단순한 경우가 대다수다. 또한 이미지 크롤링 과정에서 Google 이미지 검색의 이미지는 상시 갱신되어 검색하는 시기에 따라 다른 결과가 나와 추출되는 RGB값이 일정하지 못하다. 하지만, 연구 기간 내 같은 검색어에 있어서의 다른 구글 이미지 검색 결과에도 불구하고 대체적으로 유사한 RGB값을 출력한 것으로 보아 본 방식은 유의미하다는 것을 보여준다. 그리고 향후 연구 과제로 RGB값을 이미지 내에서 추출하는 방식에 있어서, 현재 제시한 방안들보다 향상된 방법에 대한 모색이 필요하다.

참고문헌(References)

- Apple Inc. (2015). Patent No.10-2015-7004968. Washington, DC: U.S. Patent and Trademark Office
- Apple Inc. (2017). Patent No.10-2017-0069606. Washington, DC: U.S. Patent and Trademark Office
- Cao Y., Ju T., Xu J., Hu S-M. (2016). Extracting Sharp Features from RGB-D Images. Computer Graphics Forum. 36. 10.1111/cgf. 13069.
- Chakraborty S., Nagwani N-K., Dey L. (2014). Performance comparison of incremental k-means and incremental dbscan algorithms. arXiv preprint arXiv:1406.4751.
- Cho Y., Kim Y. (2011). Color Expression by Information Extraction. Proceedings of KIIT Summer Conference, 618-620.
- ClickZ Intelligence. "Seven Ways Artificial

- Intelligence Can Be Used for Marketing.” ClickZ, May 31, 2013.clickz.com/seven-ways-artificial-intelligence-can-be-used-for-marketing/96572,(accessed September 2018).
- Davis, B. “15 Examples of Artificial Intelligence in Marketing.” Econsultancy, April 19, 2016. econsultancy.com/blog/67745-15-examples-of-artificial-intelligence-in-marketing (accessed September 2018).
- Ester M., Kriegel H. P., Sander J., Xu X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd (Vol. 96, No. 34, pp. 226-231).
- Google Image Search <https://www.google.co.kr/imghp?hl=ko&tab=wi&authuser=0> (accessed September 2018).
- Google LLC. (2017). U.S. Patent No.10-2017-7031186. Washington, DC: U.S. Patent and Trademark Office
- Gómez-Pérez A., Fernández-López M., Corcho O. (2006) Ontological Engineering: With Examples From the Areas of Knowledge Management, e-Commerce and the Semantic Web. Springer, London. <https://doi.org/10.1007/b97353>.
- HTML chart. <https://html-color-codes.info/> (accessed September 2018).
- Hwang M.-N., Lee S., Cho M., Kim S.-Y., Choi S.-P., Jung H. (2012). Ontology Construction of Technological Knowledge for R&D Trend Analysis, Journal of the Korea Contents Association, 12(12), 35-45
- Kim S., Kim G. (2012). Ontology-based User Customized Search Service Considering User Intention, Journal of Intelligence and Information Systems, 18(4), 129-143
- Kim T., Yang J., Lee J., Son J., Jeong Y. (2005). Efficient production of Ontology for Intelligent E-Commerce. Journal of Intelligence and Information Systems, 273-279.
- Lei Y., Uren V., Motta E. (2006) SemSearch: A Search Engine for the Semantic Web. In: Staab S., Svátek V. (eds) Managing Knowledge in a World of Networks. EKAW 2006. Lecture Notes in Computer Science, vol 4248. Springer, Berlin, Heidelberg.
- Lin S., Hanrahan P. 2013. Modeling how people extract color themes from images. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, 3101-3110. DOI: <https://doi.org/10.1145/2470654.2466424>
- Mahama S., A. T., Dossa, A. S., Gouton, P. (2016). Choice of distance metrics for RGB color image analysis. Electronic Imaging, 2016(20), 1-4.
- Naver Corp. (2007). Patent No.10-2007-0115690. Seoul: Republic of Korea Patent and Trademark Office
- Qu, Y., Cheng, G. (2011). Falcons concept search: A practical search engine for web ontologies. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 41(4), 810-816.
- Rose D., Levinson D. 2004. Understanding user goals in web search. In Proceedings of the 13th international conference on World Wide Web (WWW '04). ACM, New York, NY, USA, 13-19.
- Sudeepthi, G., Anuradha, G., Babu, M. S. P. (2012). A survey on semantic web search engine. International Journal of Computer Science Issues (IJCSI), 9(2), 241.
- Tran T., Cimiano P., Rudolph S., Studer R. (2007) Ontology-Based Interpretation of Keywords

- for Semantic Search. In: Aberer K. et al. (eds) The Semantic Web. ISWC 2007, ASWC 2007. Lecture Notes in Computer Science, vol 4825. Springer, Berlin, Heidelberg.
- Turban E., Outland J., King D., Lee J.K., Liang TP., Turban D.C. (2018) Intelligent (Smart) E-Commerce. In: Electronic Commerce 2018. Springer Texts in Business and Economics. Springer, Cham.
- Wikipedia color name chart. https://ko.wikipedia.org/wiki/%EC%83%89_%EB%AA%A9%EB%A1%9D (accessed September 2018).
- Woo S., Kim K., Kim C. (2005). User Category - Based Intelligent E-Commerce Meta - Search Engine. Journal of Intelligence and Information Systems, 346-355.
- WordNet. <http://www.wordnet.co.kr/> (accessed September 2018).

Abstract

Color-related Query Processing for Intelligent E-Commerce Search

Hong Jung A* · Koo Kyo Jung** · Cha Ji Won* ·
Seo Ah Jeong*** · Yeo Un Yeong**** · Kim Jong Woo*****

As interest on intelligent search engines increases, various studies have been conducted to extract and utilize the features related to products intelligently. In particular, when users search for goods in e-commerce search engines, the ‘color’ of a product is an important feature that describes the product. Therefore, it is necessary to deal with the synonyms of color terms in order to produce accurate results to user's color-related queries. Previous studies have suggested dictionary-based approach to process synonyms for color features. However, the dictionary-based approach has a limitation that it cannot handle unregistered color-related terms in user queries. In order to overcome the limitation of the conventional methods, this research proposes a model which extracts RGB values from an internet search engine in real time, and outputs similar color names based on designated color information.

At first, a color term dictionary was constructed which includes color names and R, G, B values of each color from Korean color standard digital palette program and the Wikipedia color list for the basic color search. The dictionary has been made more robust by adding 138 color names converted from English color names to foreign words in Korean, and with corresponding RGB values. Therefore, the final color dictionary includes a total of 671 color names and corresponding RGB values.

The method proposed in this research starts by searching for a specific color which a user searched for. Then, the presence of the searched color in the built-in color dictionary is checked. If there exists the color in the dictionary, the RGB values of the color in the dictionary are used as reference values of the retrieved color. If the searched color does not exist in the dictionary, the top-5 Google image search results

* School of Business, Hanyang University

** School of Finance, Hanyang University

*** School of Computer Software, Hanyang University

**** School of Business Informatics, Hanyang University

***** Corresponding Author: Kim Jong Woo

School of Business, Hanyang University

222 Wangshimni-ro, Seongdong-gu, Seoul 04763, Korea

Tel: +82-2-2220-1067, Fax: +82-2-2220-1169, E-mail: kjw@hanyang.ac.kr

of the searched color are crawled and average RGB values are extracted in certain middle area of each image. To extract the RGB values in images, a variety of different ways was attempted since there are limits to simply obtain the average of the RGB values of the center area of images. As a result, clustering RGB values in image's certain area and making average value of the cluster with the highest density as the reference values showed the best performance.

Based on the reference RGB values of the searched color, the RGB values of all the colors in the color dictionary constructed aforetime are compared. Then a color list is created with colors within the range of ± 50 for each R value, G value, and B value. Finally, using the Euclidean distance between the above results and the reference RGB values of the searched color, the color with the highest similarity from up to five colors becomes the final outcome.

In order to evaluate the usefulness of the proposed method, we performed an experiment. In the experiment, 300 color names and corresponding color RGB values by the questionnaires were obtained. They are used to compare the RGB values obtained from four different methods including the proposed method. The average euclidean distance of CIE-Lab using our method was about 13.85, which showed a relatively low distance compared to 3088 for the case using synonym dictionary only and 30.38 for the case using the dictionary with Korean synonym website WordNet. The case which didn't use clustering method of the proposed method showed 13.88 of average euclidean distance, which implies the DBSCAN clustering of the proposed method can reduce the Euclidean distance.

This research suggests a new color synonym processing method based on RGB values that combines the dictionary method with the real time synonym processing method for new color names. This method enables to get rid of the limit of the dictionary-based approach which is a conventional synonym processing method. This research can contribute to improve the intelligence of e-commerce search systems especially on the color searching feature.

Key Words : E-Commerce, Color, Image Crawling, RGB, Synonym

Received : November 2, 2018 Revised : February 27, 2019 Accepted : February 28, 2019

Publication Type : Regular Paper Corresponding Author : Jong Woo Kim

저 자 소개



홍정아

현재 한양대학교 경영학부 학생으로 재학 중이며 컴퓨터공학부를 다중전공 중이다. 빅 인텔리전트 경영교육 사업단에서 주관하는 프로젝트 학기제 인턴 과정에서 BIZ&AI랩에 근무하였다. 주요 관심 분야는 빅데이터, IT 경영, 기계학습 및 딥러닝 기법의 비즈니스 활용이다.



구교정

현재 한양대학교 파이낸스 경영학과 학생으로 재학 중이다. 한양대학교 빅 인텔리전트 경영교육 사업단에서 주관하는 프로젝트 학기제 인턴 과정(Business & AI)을 수료하였다. 주요 연구 관심사는 컴퓨터 비전, 텍스트 마이닝, 주가 예측 등이다.



차지원

현재 해군 재정장교 복무중이며, 한양대학교 경영학과를 졸업하였다. 한양대학교 (주)한양비즈랩에서 주관하는 프로젝트 학기제의 Business AI Lab의 인턴 과정을 수료하였다. 모바일 어플리케이션 개발, 웹 개발, 기계학습 및 딥러닝 등이다.



서아정

현재 한양대학교 컴퓨터공학부 학생으로 재학 중이다. 한양대학교 빅 인텔리전트 경영교육 사업단에서 주관하는 프로젝트 학기제 인턴 과정으로 A.I.랩에 근무하였다. 이후 네이버 Clova에서 인턴으로 근무하였고, 인공지능 관련 학과에 석사과정 진학을 희망하고 있다. 주요 관심 분야는 NLP, Visual Storytelling, 데이터 마이닝 등이다.



여운영

현재 한양대학교 일반대학원 비즈니스인포매틱스학과에서 석사과정으로 재학 중이며, 한국해양대학교 데이터정보학과에서 이학사를 취득하였다. 주요 연구 관심분야는 기계학습 및 딥러닝 기법 응용과 자연어 처리(NLP) 등이다.



김종우

현재 한양대학교 경영대학 경영학부 교수로 재직 중이다. 서울대학교 수학과에서 학사를 마쳤으며, 한국과학기술원에서 경영과학으로 석사학위를, 산업경영학으로 박사학위를 취득하였다. 주요 연구 관심분야는 데이터마이닝 기법과 응용, 기계학습과 딥러닝, 오피니언 마이닝, 상품추천기술, 지능형 정보시스템, 집단지성, 사회 네트워크 분석, 클라우드 컴퓨팅 서비스 등이다.