

지능형 멀티미디어 검색 기술 동향

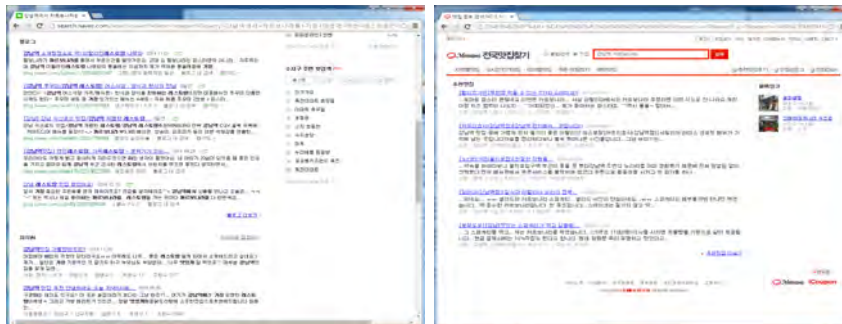
이용환 (극동대학교), 안효창(단국대학교),
조한진·이준환 (극동대학교), 이상범 (단국대학교)

목차	1. 서론
	2. 지능형 모바일 검색 기술
	3. 지능형 모바일 멀티미디어 검색 기술 사례
	4. 결론 및 향후 전망

1. 서론

“강남역에서 카르보나라를 가장 맛있게 하는 레스토랑은?” 이라는 질의로 맛집을 검색하면, 일반적으로, 입력 키워드인 “강남역”, “카르보나라” 등으로 검색하여 결과를 돌려준다(그림 1). 인터넷에서 누구나 비슷한 경험을 해봤으리라 생각되지만, 이러한 맛집 검색을 수행하면, 수십 개의 링크를 열어봐야 사용자가 원하는 검색 서비스를 받게 된다. 이러한 불편함을 해결해 주는 지능형 서비스 지

원으로, 최근 시맨틱 검색(Semantic Search) 기술이 많은 주목을 받고 있다. 핵심 기술인 지능형 검색은 문서, 사진, 동영상 등 다양한 멀티미디어 콘텐츠에 대해 컴퓨터가 상황(Context)을 이해하여 검색어의 의미에 맞는 결과를 제공하는 기술이다^[1]. 기존의 검색 엔진에서 제공하는 서비스는 키워드 기반 검색을 통하여, 방대한 웹 문서에 대해 입력된 검색어를 포함하는 검색 결과를 통계적으로 제공하는 반면에, 시맨틱 웹으로 최적화된 정보를 상호 연결하여 의미를 분석하고 논리적인 추론을 수행하여 사

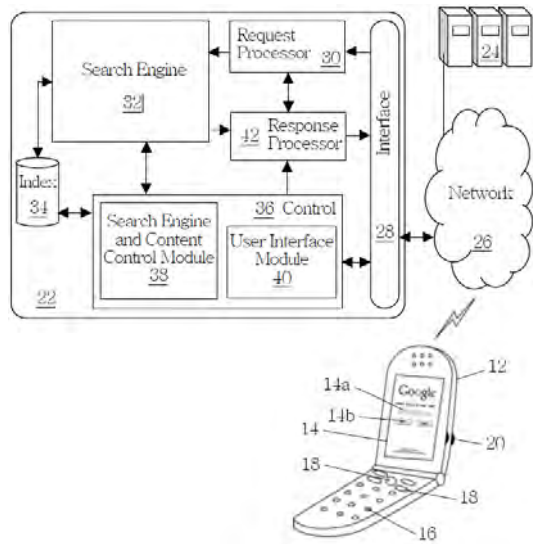


(그림 1) 인터넷 맛집 검색 사례 [출처: www.naver.com과 www.menupan.com]

용자의 검색 의도에 맞는 콘텐츠를 제공하는 핵심 기술을 포함한다.

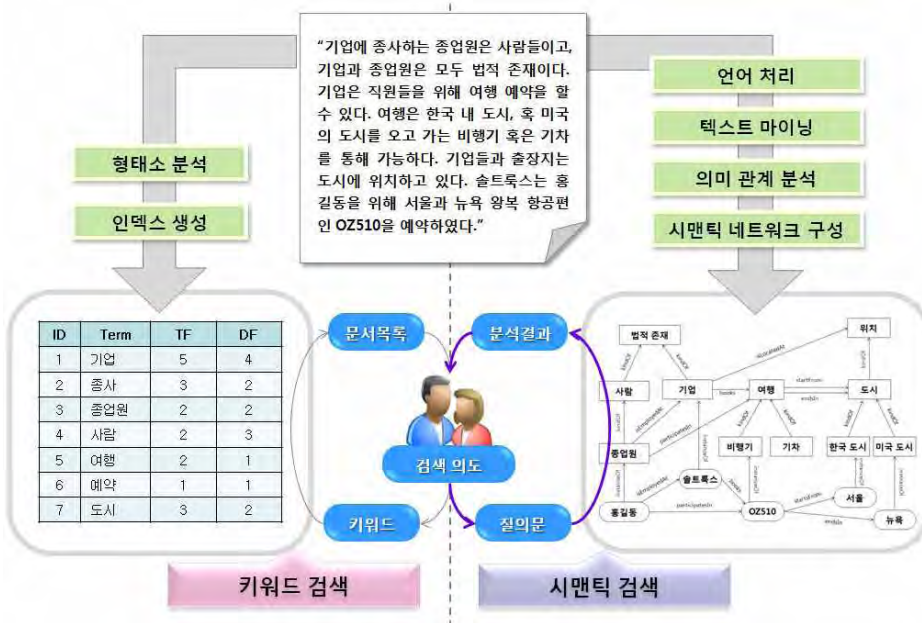
이러한 검색 기술의 발전은 정보 관리의 패러다임의 변화를 의미하며, 초기 야후(Yahoo)와 같이 정의된 디렉토리를 기반으로 분류하고 검색어를 통하여 단순한 메타데이터 검색 방식에서 인덱싱, 저장 구조 변형과 랭킹 검색 등으로 대표되는 구글(Google) 검색이 시장 지배력을 갖는 변화를 가져오게 되었다. 최근에는 시맨틱 웹 검색과 같은 기계학습 기술을 이용하고 정보간의 상호 관계를 활용하여 사용자 의도에 부합되는 정보를 찾아주는 방식으로 발전하고 있으며, 이는 정보간의 구조를 표현하는 온톨로지(Ontology), 자연어 처리와 텍스트 마이닝(Mining)을 포함한 학습 및 추론 등의 인공지능 기법을 통합하여 적용한다^[2].

근래에, 스마트 디바이스의 발전에 따라 이러한 검색 기술은 사용자 요구를 충족시키는 모바일 환경으로 확장되고 있으며, 정보 분석과 검색이 모바일



(그림 3) 구글 모바일 검색 서비스 구성도^[4]

일 플랫폼에 적용되는 변화가 나타나고 있다^[3]. (그림 3)은 구글에서 모바일 검색(Mobile Search)으로 특허 출원한 서비스 구성도이다. 사용자의 입력에 따라 적절한 정보를 제공하는 검색 엔진의 핵심 역



(그림 2) 기존의 키워드 기반 검색과 시맨틱 검색의 차이^[2]

할은 유선이든 무선이든 동일하지만, 모바일 플랫폼에서의 검색은 모바일 환경에 맞는 제어(Control) 모듈과 디바이스의 편리성을 제공하는 사용자 인터페이스(User Interface) 모듈에서 차이점을 보인다. 다시 말해, 모바일 검색은 사용자 개인이 항상 휴대하고 이동하면서 사용한다는 휴대성과 이동성을 감안하여 사용자 검색 요청에 적절한 맞춤형 검색을 지원해야 하며, 모바일 단말기의 작은 화면을 고려하여 개별화된 최적의 검색 결과를 제공해야 한다는 특징을 가진다.

본 고에서는 스마트폰으로 시작된 모바일 시대에 사용자 수요가 가장 많은 모바일 플랫폼 상에서의 지능형 멀티미디어 검색 기술에 대해 소개하고, 현재 이슈가 되고 있는 관련 기술 동향과 핵심적인 모바일 검색 기술의 표준화 동향을 알아본다.

2. 지능형 모바일 검색 기술

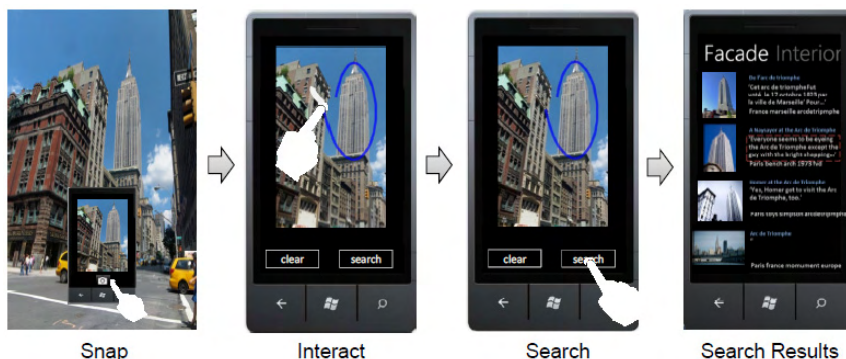
2.1 모바일 비주얼 검색

모바일 비주얼 검색(Mobile Visual Search)은 스마트 디바이스 환경에서 사용자 질의 영상의 특징을 추출하고, 주변 영역의 정보를 이용하여 추출한 기술자(Descriptor)를 기반으로 미리 구축

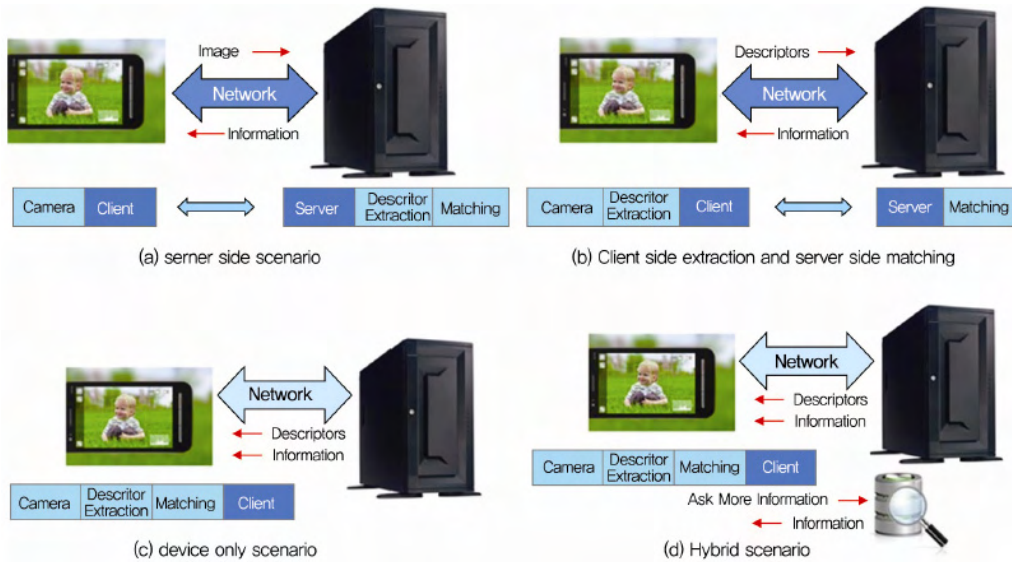
된 영상 정보 데이터베이스와 비교하여 검색하는 기술이다⁵⁾. 이러한 기술은 사용자가 키워드를 모를 경우에도 검색의 편의성을 만족시킬 수 있기 때문에, 효과가 매우 큰 것으로 조사되었으며, 세계 각국의 IT 기업들은 영상 기반의 정보검색 기술을 차세대 성장산업으로 보고, 기술 개발과 서비스 지원에 집중하고 있다.

(그림 4)는 모바일 비주얼 검색 기술에 대한 개념도이다. 사용자가 검색하고자 하는 대상을 영상으로 받아 입력 질의로 검색을 요청하면, 이미 구축되어 있는 데이터베이스에서 관련 영상 및 해당 정보들을 검색하여 사용자에게 모바일 환경에 적합한 인터페이스에 따라 검색 결과를 반환한다. 예를 들어, 건물을 촬영하고 이를 DB에 검색하여, 해당 건물에 대한 여러 정보를 얻을 수 있다. 혹은 상품 이미지를 촬영하여 입력하면, 해당 상품에 대한 가격, 기능 및 판매처 등의 부가 정보를 제공받을 수 있다.

이러한 모바일 비주얼 검색은 특징 추출과 검색 단계로 이뤄지며, 응용 서비스를 구축함에 있어, 특징 추출과 검색을 수행하는 주체가 서버 또는 사용자 단말기에 따라, 4가지의 시나리오가 가능하다(그림 5). (그림 5-a)은 질의 영상을 사용자 단말기를 통해 입력받아 서버로 전송되고 특



(그림 4) 모바일 비주얼 검색 개념도 [6]



(그림 5) 특징 추출과 검색을 수행하는 주체에 따른 모바일 비주얼 검색 서비스 시나리오^[5]

징 추출과 검색이 모두 서버 측에서 수행되는 경우이며, (그림 5-b)는 입력된 질의 영상의 특징 추출은 사용자 단말기에서 수행되고, 추출된 특징만이 서버로 전송되어 검색은 서버 측에서 수행되는 경우이다. 특징 추출과 검색이 적절하게 서버와 사용자 단말기에서 수행되는 하이브리드 방식도 있다(그림 5-d).

Extraction), 데이터베이스 검색(Database Search)과 정합/검증(Matching and Verification)으로 이뤄진다 (그림 6). 이러한 각 단계의 기술 요소를 요약하면, <표 1>과 같다.

<표 1> 모바일 비주얼 검색의 핵심 기술 요소

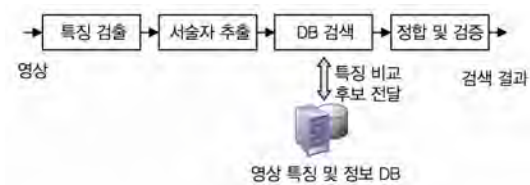
요소기술	내용
특징 검출	이미지에서 뷰포인트, 환경변화 등의 왜곡에 강인한 영역 또는 특징점을 검출하는 기술
기술자 추출	검출된 특징 영역 또는 주변 영역을 기반으로, 콘텐츠를 구분하는 특징 벡터를 추출하는 기술
데이터베이스 검색	기존에 구축된 데이터베이스에서 추출된 특징 벡터를 기반으로 적절한 콘텐츠를 선별하는 기술
정합/검증	선별된 콘텐츠에 대해, 정합 유무를 판별하고 정합이 이뤄진 콘텐츠에 대해 기하학적 검증을 수행하는 기술

2.2 모바일 비주얼 검색의 핵심 요소 기술

모바일 비주얼 검색기술은 기술자 기반이 특징 추출 방법을 사용하며, 핵심 기술은 특징 검출 (Feature Detection), 기술자 추출(Descriptor

2.3 요소 기술들의 기술 동향

각 요소 기술에 대한 다양한 방법들이 연구 개발되고 있으며, 본 절에서는 주요 기술에 대한 동향을 살펴본다.



(그림 6) 모바일 비주얼 검색 요소 기술^[5]

가장 핵심적인 요소기술 중의 하나인 특징 검출은 이미지 내에서 시점이나 환경 변화에 따른 조명 등 외부환경에 따라 영향을 미치는 요소들에 강인한 불변의 요소를 검출하는 기술이다. 이는 환경에 대한 강인성뿐만 아니라, 사용자 단말기에서 수행되는 경우(그림 5-b)에 대비하여 계산의 효율성도 함께 고려해야 한다.

이러한 특징 검출의 대표적인 기술로는, 해리스 검출기(Harris Detector)^[7], 헤시안 검출기(Hessian Detector)^[8], 해리스-라플라스 검출기(Harris-Laplace Detector)^[9], SIFT 검출기(Scale Invariant Feature Transform Detector)^[10], MSER 검출기(Maximally Stable External Regions Detector)^[11], SURF 검출기(Speeded Up Robust Features Detector)^[12]와 FAST 검출기^[13] 등이 있다.

시점과 환경 요소에 강인한 영역을 찾은 후, 이미지의 영역 정보를 이용하여 구별을 하는 기술자를 추출한다. 이러한 기술자 추출 방법으로는 대표적으로 SIFT 특징, SURF 특징, 바이너리 특징을 이용한 LBP^[14], BRIEF(Binary Robust Independent Elementary Feature)^[15], BRISK(Binary Robust Invariant Scalable Key Point)^[16]와 LIOP(Local Intensity Order Pattern for Feature Description)^[17] 등이 있다.

모바일 비주얼 검색기술이 응용 서비스로 활용되기 위해서는 특징 검출과 기술자 추출뿐만 아니라, 대용량, 다량의 콘텐츠 레포지토리(Repository)에서 효율적으로 검색을 수행하는 방법이 필요하다. 대표적으로, Vocabulary Tree^[18], Bag of Words, Residual Enhanced Visual Vector^[19], Vector of Aggregated Local Descriptors^[20]와 Fisher Vector^[21] 등을 기반으로 한 검색 구조 등이 있다.

정합/검증 단계는 기존에 구축된 데이터베이스에서 적절한 콘텐츠를 검색하고, 해당 후보 콘텐츠

들에 대해 실제로 일치하는 정보가 맞는지를 검증하는 기술이다. 콘텐츠가 포함하는 위치 정보와 기술자를 통해 추출한 특징 벡터를 기반으로 기하학적인 관계를 추정한다. 대표적으로 RANSAC(Random Sample Consensus) 방법이 있으며^[22], 최근에는 확률 통계 모델 기반의 정합쌍과 비정합쌍을 구분하는 방법이 연구되고 있다.

모바일 비주얼 검색 기술에 대한 표준으로, 현재 MPEG-7 CDVS(Compact Descriptors for Visual Search)가 국제표준화가 진행 중에 있으며, 다양한 목적과 응용 환경에 강인하고 효율적인 기술자 추출과정을 중심으로 표준을 지정하고 있다. 이는 모바일 검색 응용 서비스를 위한 설계, 최고의 성능과 응용 환경에 대한 호환성 보장, 최소 크기 및 효율적 구현 검색, 추출된 기술자의 비트 스트림과 상호호환성 지원 등을 포함하고 있다. 현재 110회까지 진행된 CDVS는 기존의 영상 특징을 압축하고, 다양한 환경을 지원하기 위해 검출된 특징 선별, 기술자의 양자화 과정을 포함하며, 대용량 검색을 위한 효율적인 검색 구조를 실험하고 특징 검출에 있어 기존의 다양한 특징 검출기로부터의 특허 회피 검출 방법을 채택하여 검증하고 있다.

3. 지능형 모바일 멀티미디어 검색 기술 사례

유선 인터넷 검색 시장은 압도적인 시장 점유율을 기록하는 구글을 중심으로 시장이 형성되어 있지만, 모바일 검색 시장에서도 이러한 현상이 지속되지는 않는다. 모바일 검색 기술과 시장은 검색 엔진을 탑재한 스마트 디바이스를 중심으로 IT기업에서 전략적으로 접근하고 있다. 본 절에서는 모바일 검색 관련 서비스를 중심으로, 기업의 서비스 동향을 살펴본다.

<표 2> 모바일 비주얼 검색의 응용 서비스

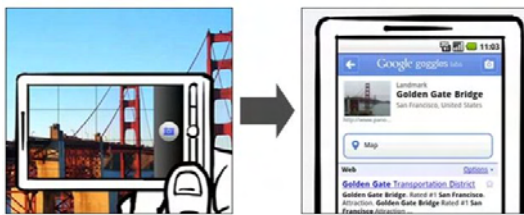
응용분야	응용 서비스
정보검색	사용자 검색 대상을 촬영하여 해당 객체에 대한 다양한 정보 및 유사 객체 검색 서비스를 제공
증강현실	모바일 비주얼 검색 기술을 적용한 실제 영상과 3D 가상 객체간의 연동을 통하여 증강현실 정보 서비스를 제공
비전	로봇 및 머신 비전에 적용하여 사물 및 객체를 인식하고 다양한 비전 서비스를 제공
방송	동영상을 포함한 방송 콘텐츠에서 상품 및 기타 객체(예, 건물 또는 유적)를 검색하는 서비스를 제공
자동차	스마트 자동차 분야에 적용되며, 자동차 주행 중에 객체를 탐지하고 표지판 식별, 위험물 감지 등의 서비스를 제공
영상 감시	CCTV에서 객체를 검출, 인식, 감지하며 발생하는 이벤트를 감지하는 서비스를 제공
콘텐츠 관리	멀티미디어 콘텐츠 내의 객체 검색 및 분류, 관리 서비스를 제공

모바일 비주얼 검색 기술을 활용하는 다양한 응용 서비스를 요약하면 <표 2>와 같다.

3.1 구글

검색엔진의 선두주자인 구글에서는 MWC (Mobile World Congress) 2010에서 “향후 구글의 모든 서비스를 모바일로 제공할 것”이라고 모바일 사업전략을 발표하였고, 안드로이드 운영체제를 탑재한 모든 스마트폰에서는 구글검색(Google Search), 구글크롬(Google Chrome), 지메일

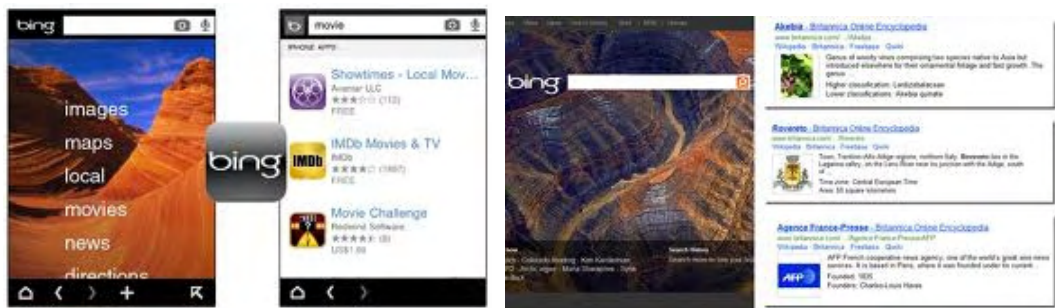
(Gmail), 유튜브(YouTube), 위성사진 지도서비스 구글어스(Google Earth), 인터넷전화 서비스 구글 보이스(Google Voice), 비주얼검색 구글고글스(Google Goggles) 등을 기본적으로 제공하고 있다. 특히, 구글고글스는 스마트폰을 이용하여 이미지를 촬영하고 이를 통해 영상 데이터베이스로부터 검색하는 영상 검색 서비스(그림 7)이며, 현재 특정 카테고리(Text, Landmark, Book, Contact Info, Artwork, Wine)의 특정 사물을 대상으로 한 모바일 영상검색 서비스를 제공하고 있다^[23].



(그림 7) 구글 고글스 서비스^[23]

3.2 마이크로소프트

MS 검색엔진 서비스 Bing(Bing) 또한 모바일 검색 서비스를 출시하였으나(그림 8), 검색 엔진 시장의 경쟁력이 높지 않으며, 이는 모바일 OS 시장 점유와도 연관성을 가진다^[24]. 애플 아이폰 iOS, 구글 안드로이드 OS에 밀려 Windows

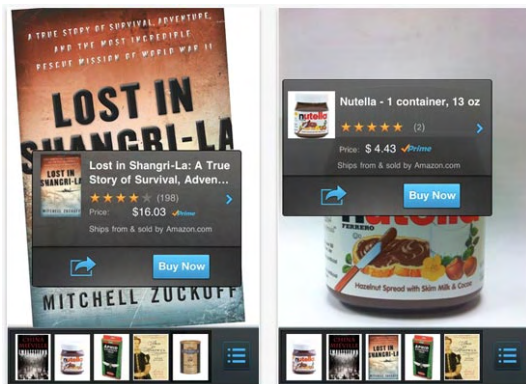


(그림 8) 마이크로소프트 Bing 서비스^[24]

Mobile OS의 시장 장악력이 낮으며, 검색 기술의 후발주자로서 점유율이 높지 않은 상황이다.

3.3 아마존

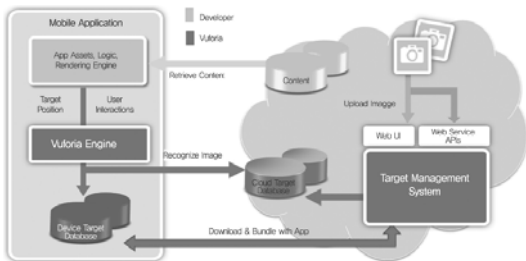
아마존은 자사의 온라인 쇼핑 서비스와 모바일 비주얼 검색 기술을 결합하여 효과적인 서비스 상승효과를 얻고자 하며, 영상 검색 기업인 Snap Tell을 인수하여 검색기술을 적용하였다^[25]. 현재 상품 검색 앱(Flow), 특정 상품 전용 검색 서비스(예, 신발을 대상으로 한 검색 앱Fabulous)를 제공하고 있다(그림 9).



(그림 9) 아마존의 상품 검색 서비스 Flow^[25]

3.4 쉐이크

쉐이크에서는 모바일 비주얼 검색 기술을 활용한



(그림 10) 쉐이크의 Vuforia 서비스 개념도^[26]

응용서비스 개발 플랫폼을 제공하고 있다^[26]. 현재 문자 인식, 3차원 객체 인식, 영상 인식 등의 기술을 지원하며, 개발환경으로 안드로이드와 iOS를 지원하며, 개발 플랫폼으로 개발된 앱이 6,000 여개에 달한다.

4. 결론 및 향후 전망

최근 구글, 마이크로소프트, 아마존, 쉐이크 등 거대 IT 기업들이 모바일 비주얼 검색 기술을 둘러싼 경쟁이 본격화되고 있으며, 모바일 검색 기술 확보 및 제품 검색에 적용하는 다양한 응용 서비스가 부상되고 있다. 본 고에서는 지능형 모바일 비주얼 검색의 기술 동향과 적용 서비스의 사례를 살펴보았다.

스마트폰으로 대표되는 모바일 시장이 본격화되면서 모바일 검색 시장도 활성화되고 있으며, 거대 IT업체들이 앞다투어, 모바일 비주얼 검색 시장으로 진출하고 있다. 현재 기술의 가능성과 시장성, 가치를 판단해 볼 때 모바일 비주얼 검색 기술은 전망 좋은 기술 영역이며, 기술 및 서비스 개발에 많은 노력을 기울여야 할 부분이다. 비전과 인식 분야를 바탕으로 보이는 모든 세상을 대상으로 비주얼 검색이 적용되는 시대가 올 것으로 기대되며, 모바일 비주얼 검색 기술은 인공지능, 정황인식, 개별화 등의 서비스와 결합하고 사용자 경험 인터페이스를 기반으로 지능형 모바일 멀티미디어 검색 시스템으로 발전해야 할 것이다.

참고 문헌

- [1] 행정안전부, “국가정보화백서”, 2011년
- [2] 박세영, 이현재, 조준면, “지능형 검색 서비스 기술 동향 및 산업전망”, 한국방송통신전파진흥원 RM Issue Report, 2013년.

- [3] 장명길, "모바일 지능형 검색 기술", 전자통신연구원, 2009년
- [4] 임수증, 오효정, 류범모, 정호영, 장명길, "모바일 지능형 검색 기술 동향", 전자통신동향분석, 25권, 3호, 2010년.
- [5] 이승재, 이근동, 나성일, 제성관, 정다운, 오원근, 서영호, 손욱호, "모바일 비주얼 검색: 기술과 표준화 동향", 전자통신연구원, Electronics and Telecommunications Trends, 2014.
- [6] Jitao Sang, Tao Mei, Ying-Qing Xu, Chen Zhao, ChangSheng Xu, Shipeng Li, "Interaction Design for Mobile Visual Search", IEEE Transactions on Multimedia, vol15, Issue,7, 2013.
- [7] C. Harris and M. Stephens, "A combined corner and edge detector," Vision Conference, 1988
- [8] Beaudet, "Rotationally invariant image operators," International Joint Conference Pattern Recognition, 1978.
- [9] T. Lindeberg, "Feature detection with automatic scale selection" International Journal of Computing Vision, 1998.
- [10] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints" International Journal of Computing Vision, 2004.
- [11] J. Matas et al., "Robust wide-baseline stereo from maximally stable external regions" Image Vision Computing, 2004.
- [12] H. Bay et al., "Surf: Speeded up robust features" European Conference Computing Vision, 2006.
- [13] E. Rosten et al., "Real-Time Video Annotations for Augmented Reality" Proceeding International Symposium Visual Computing, 2004.
- [14] M. Heikkila, M. Pietikainen, and C. Schmid, "Description of Interest Regions with Center-Symmetric Local Binary Patterns," ICVGIP, LNCS 4338, 2006.
- [15] M. Calonder et al., "Brief: Binary robust independent elementary features," European Conf. Comput. Vision, 2010.
- [16] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints," Proc. IEEE ICCV, 2011.
- [17] Z. Wang, F. Bin, and F. Wu., "Local intensity order pattern for feature description," Proceeding IEEE ICCV, 2011.
- [18] D. Nistér and H. Stewénus, "Scalable recognition with a vocabulary tree," Proceeding Conference Computing. Vision Pattern Recognition(CVPR), 2006.
- [19] D. Chen et al. "Residual enhanced visual vector as a compact signature for mobile visual search," Signal Processing, 2012.
- [20] H. Jegou et al., "Aggregating local descriptors into a compact image representation," Proceeding Conference Computing. Vision Pattern Recognition(CVPR), 2010.
- [21] F. Perronnin, J. S'anchez, and T. Mensink, "Improving the Fisher kernel for large-scale image classification," ECCV, 2010.
- [22] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", Communication ACM, 1981.
- [23] 구글 고글스, <http://www.google.com/mobile/goggles/>
- [24] 마이크로소프트 웹사이트, <http://www.microsoft.com/en-us/download/details.aspx?id=2012>
- [25] Flow <http://a9.com/whatwedo/mobile-technology/>
- [26] Vuforia <http://www.qualcomm.com/solutions/augmented-reality>

저 자 약 력



이 용 환

이메일 : hwany1458@empal.com

- 1993년 단국대학교 전자계산학과 (학사)
- 1995년 단국대학교 전산통계학과 (석사)
- 2007년 단국대학교 전자컴퓨터공학과 (박사)
- 1995년~2000년 (주)한국정보시스템 기술개발연구소 /선임연구원
- 2000년~2003년 (주)이칼로스 팀장
- 2003년~2007년 (주)한국e문화 기술이사
- 2007년~2009년 성균관대학교 박사후연구원
- 2009년~2013년 단국대학교 응용컴퓨터공학과 연구 교수
- 2014년~현재 극동대학교 스마트모바일학과 연구교수
- 관심분야: 멀티미디어 검색, 모바일 멀티미디어 통신, 증강현실



안 효 창

이메일 : youcu92@dankook.ac.kr

- 2003년 상지대학교 전자계산공학과 (학사)
- 2006년 단국대학교 전자컴퓨터공학과 (석사)
- 2012년 단국대학교 전자컴퓨터공학과 (박사)
- 2013년~2014년 단국대학교 미디어콘텐츠연구원 스토리텔링연구소 연구원
- 2014년~현재 단국대학교 응용컴퓨터공학과 연구교수
- 관심분야: 영상처리, 컴퓨터비전, 임베디드시스템, 모바일 멀티미디어 통신



조 한 진

이메일 : hanjincho@hotmail.com

- 1997년 한남대학교 전자계산공학과 (학사)
- 1999년 한남대학교 컴퓨터공학과 (석사)
- 2002년 한남대학교 컴퓨터공학과 (박사)
- 2003년~현재 극동대학교 스마트모바일학과 교수
- 관심분야: 네트워크, 스마트미디어, 모바일 앱



이 준 환

이메일 : rainbow@kdu.ac.kr

- 1994년 단국대학교 전자공학과 (학사)
- 1996년 단국대학교 전자공학과 (석사)
- 2001년 단국대학교 전자공학과 (박사)
- 2001년~현재 극동대학교 스마트모바일학과 교수
- 관심분야: 음성 처리 시스템, 멀티미디어 응용, 스마트 미디어, 모바일 앱



이 상 범

이메일 : sbrhee@dankook.ac.kr

- 1974년 연세대학교 전자공학과 (학사)
- 1978년 서울대학교 전자공학과 (석사)
- 1986년 연세대학교 전자공학과 (박사)
- 1979년~현재 단국대학교 응용컴퓨터공학과 교수
- 2005년 한국정보처리학회장
- 관심분야: 영상 및 음성 신호 시스템, 멀티미디어 응용, 임베디드 시스템,