

논문 2011-1-16

모바일정보검색 서비스를 위한 문자 인식

Image Processing for Mobile Information Retrieval Service

임명재*, 현성경**, 박지은**, 이기영***

Myung-Jae Lim, Sung-Kyung Hyun, Ji-Eun Park, Ki-Young Lee

요 약 현대 사회는 정보화에 대한 인식이 크게 보편화되면서 정보 통신 기술에 대한 개발이 급속화 되고 있는 추세이다. 특히 모바일 기기에 대한 기술 또한 빠르게 진행됨에 따라 자신이 원하는 정보를 언제 어디서나 제공받기를 기대한다. 이에 따라 편리한 정보검색을 위한 이미지 검색이 보편화 되고 있다. 하지만 일반적인 이미지 검색은 이미 지에서의 문자영역 추출의 부정확성과 추출된 문자로 상세한 정보를 얻는데 어려움이 따른다. 따라서 본 논문에서는 촬영한 영상에서 문자인식을 통해 관광지를 포함한 여러 상점들의 간판을 인식하여 사용자에게 편리한 정보제공을 목적으로 하고 있으며, Top-Hat이라는 문자추출 방법과 서버를 함께 연동시키므로 시킴으로써 보다 정확하고 상세한 정보를 얻을 수 있다.

Abstract The modern society with the wide spread recognition of the importance of informatics and for the development of information and communication technology is rapidly processing. Especially the rapid development in mobile technology boost up the general expectation that one can get the information he wants anytime anywhere. Accordingly image search for the convenient information retrieval is becoming common. However general image search has difficulties because inexactitude extracting character in the image and getting the detail information in extracted character. Therefore these paper make character recognition through the images that I photographed a sightseeing resort, a signboard of a lot of stores to a smart phone camera, so information offer to be convenient to users is a purpose. A user can get detailed information, by character extraction way called top-hat algorithm and connect to a server.

Key Words : Image processing, Character recognition, LBS, Information retrieval

1. 서 론

현대 사회는 정보화에 대한 인식이 크게 보편화되면서 정보 통신 기술 및 뉴미디어의 개발에 노력을 기울이고 있다. 특히 모바일 기기에 대한 기술이 빠르게 진행되고 있고, 관련 기술 또한 속출하고 있다. 현재 전 세계적으로 모바일 기기에 카메라 장착은 필수 사항이 되어가고 있고, 모바일 폰은 각종 휴대기기들과 결합하여 가장

급격한 발전을 보이고 있다. 단순히 전화통화라는 본래의 기능을 넘어서 PC에서 행해지던 작업이나 다른 정보 기기를 통해서만 이루어지던 많은 기능들이 모바일 폰을 통하여 가능하게 되었다^[1].

이 중에서도 모바일 폰의 카메라로 촬영한 자연영상을 통해 사용자가 다양하고 풍부한 정보를 활용하는 것이 가능해 졌다. 촬영된 영상에서 일정한 패턴을 분석하여 그것을 통해 다양한 정보를 제공 받을 수 있다. 영상 내에서도 이미지 프로세싱을 이용한 문자인식을 통해 정보검색이나 위치정보 서비스 등 보다 다양한 정보를 제공 받을 수 있다. 현재 전 세계적으로 모바일 폰의 사용자는 꾸준히 증가하고 있고, 그 성장의 원동력은 스마트

*중신회원, 을지대학교 의료IT마케팅학과

**준회원, 을지대학교 의료전산학과

***중신회원, 을지대학교 의료IT마케팅학과(교신저자)

접수일자: 2011.1.19 수정일자: 2011.2.1

게재확정일자: 2011.2.11

폰 출현에 있다. 현재 구글에서는 ‘고글스’ 라는 서비스를 제공하고 있으며, 여행, 출장 등 다양한 목적으로 해외를 방문하게 되는 경우 길거리의 간판, 안내문, 표지판 등 즉각적인 번역을 필요로 하는 곳에서 모바일 폰 카메라로 촬영한 영상을 통해 원하는 언어로의 번역을 제공한다^[2].

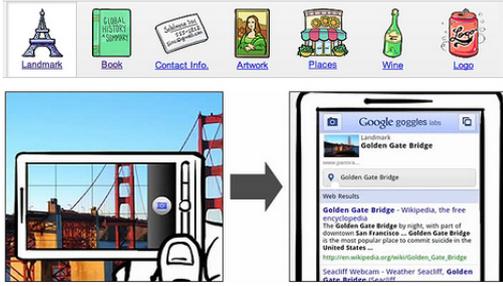


그림 1. 구글의 고글스 서비스
Fig. 1. Google Goggles service

현재 전 세계적으로 이와 같은 서비스를 제공하고자 하는 연구가 활발히 진행되고 있다. 본 논문에서도 모바일 폰을 통해 정보를 습득하고자 하는 사용자에게 보다 간편한 방법으로 다양한 계층에서 정보를 습득할 수 있도록 이미지프로세싱 방법을 통한 문자인식 정보검색에 대한 연구를 제안하고자 한다.

본 논문의 구성은 2장에서 관련연구, 3장에서 모바일 정보검색을 위한 문자인식, 4장에서 실험 및 평가, 5장에서 결론 및 향후 연구과제를 제시한다.

II. 관련 연구

1. 영상처리 및 문자인식

문자인식(Character Recognition)이란 시각(Vision)정보를 통하여 문자를 인식하고 의미를 이해(Understanding) 하는 인간의 능력을 컴퓨터로 실현하려는 패턴인식 (Pattern Recognition)의 한 분야로서, 광학 문자인식 (Optical Character Recognition), 우편물 자동 분류, 문서인식, 도면인식 등의 분야에서 실용화가 이루어지게 되었으며, 최근에는 인공지능의 최신기법인 신경망 (Neural Network), 퍼지 (Fuzzy), 유전알고리즘 (Genetic Algorithm) 등의 응용과 자연어처리, 심리학, 생리학, 인

지과학 등 관련 학문과의 접목에 의해 연구 개발되고 있다^[7]. 일반적으로 문자인식은 공간 정보만을 이용하는 오프라인 문자 인식 (광학문자인식 Optical Character Recognition) 과 공간 정보 외에 시간 정보도 함께 이용할 수 있는 온라인 문자 인식 (필기체인식 (Handwriting Recognition) 인식으로 구분 된다^{[2][3]}.

2. 모바일 정보검색

피쳐폰, 스마트폰, MID(Mobile Internet Device) 등 휴대용 단말기를 통해 단말기 내, 인터넷 웹, 파일 등의 필요한 정보를 찾는 정보검색을 말한다. 모바일 검색의 종류로는 일반 웹 검색(General Web Search), 휴대폰 내부 검색(On Device Search), 지역 검색(Local Search), 이동통신사 내 콘텐츠 검색(On Portal Content Search), 기타 포털 내에 콘텐츠 검색(Off Portal Content Search)이 있다. 현재 모바일 정보 검색은 증강현실(Augmented Reality)을 활용한 검색과 공간지리정보와의 결합을 통한 위치 기반 서비스가 유비쿼터스 환경에 적합한 형태로 제공되고 있다^[8].

3. LBS (Location Based Service)

휴대폰 속에 기지국이나 GPS와 연결되는 칩을 부착해 위치추적 서비스, 공공안전 서비스, 위치기반정보 서비스 등 위치와 관련된 각종 정보를 제공하는 서비스를 일컫는다. 즉 유선·무선 통신망을 통해 얻은 위치정보를 바탕으로 여러 가지 서비스를 제공하는 것이 위치기반 서비스이다^{[4][6]}.



그림 2. 스마트폰에서의 위치기반 서비스
Fig. 2. Local based service in smart phone

LBS를 구현하기 위해서는 위치측위기술, 서비스 및 콘텐츠, 통신기술, 단말기 등이 복합적으로 구현되어야

한다^[5]. 위치측위기술이란 사용자의 위치를 파악하는 기술로 이동통신 기지국이나 GPS를 이용하는 방식이 대표적이다. 서비스 및 콘텐츠는 획득된 위치정보를 실시간으로 처리하고 사용자 위치를 지도에 표시하는 등 사용자의 요청처리, 정보저장을 수행한다.

III. 모바일 정보검색을 위한 문자인식

1. 정보검색을 위한 문자인식

간판인식은 관광지를 포함한 여러 상점들의 간판 내 문자를 인식하여 사용자에게 편리한 정보제공을 하는 것을 목적으로 하고 있다. 간판은 가게를 나타내는 일반적인 표시판으로, 간판내의 문자가 검색어 역할을 하여 해당 가게에 대한 자세한 내용을 제공할 수 있다는 장점을 갖고 있다. 이러한 간판인식은 기존 DB저장되어 있는 간판과 새로 촬영하여 들어온 입력영상을 비교하여 간판의 문자영역과 그 주변 영역을 검출하여 문자 정보를 검색하는 방법으로 결정한다. 이 연구에서는 객체(문자)를 늘이고(팽창) 줄이는(침식) 방법을 연속적으로 반복하는 Top-Hat연산을 사용한다. 즉, 촬영된 영상 내에서 문자 주변의 잡음을 침식연산의 반복 실행으로 제거 한 후 문자 외곽선은 팽창연산을 반복적으로 실행하여 추출해낸다. 즉, Top-Hat 연산의 적용으로 영상내의 잡음은 최대한으로 줄이고 문자의 외곽선을 보다 뚜렷하게 하여 문자영역을 추출하게 된다.



원본 Top-Hat 결과 이진화 레이블링 통한 잡영 제거

그림 3. Top-Hat연산
Fig. 3. Top-Hat operation

$$(f \ominus b) = \min_{z+x \in D_f, z \in D_b} \{f(z+x) - b(z)\} \quad (1)$$

$$(f \oplus b) = \max_{z-x \in D_f, z \in D_b} \{f(z-x) - b(z)\} \quad (2)$$

위 수식에서 D_f 와 D_b 는 각각 f 와 b 가 속해있는 영역을 나타낸다. D_b 의 유효한 범위는 $-k \sim +k$ 까지 이다.

여기서 k 는 함수 b 의 값이 0으로 떨어지는 곳의 x 좌표를 뜻한다. 수식 (1)은 침식연산으로 최소값 필터, 수식 (2)는 팽창연산으로 최대값 필터의 역할을 한다.

본 논문의 연구 방법은 사용자가 모바일 폰의 카메라로 간판을 촬영하여 간판 내 문자영역의 외곽선을 추출하고 문자 자체를 인식하여 인식된 문자가 나타내는 정보를 제공해준다. 또한 간판의 상세한 정보 제공을 원하면 웹 서버에 연동하여 서버에 저장되어 있는 기존의 DB 간판영상과 비교하여 간판에 해당하는 상점의 정보를 확인 할 수 있다.

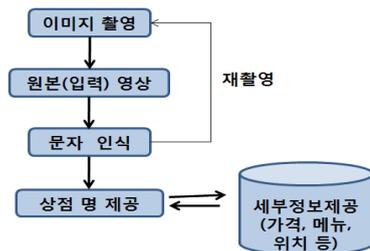


그림 4. 시스템 흐름도
Fig. 4. System flow chart

2. 정보검색 환경 및 서버 DB

본격적인 정보화 시대가 도래함에 따라 정보 검색 수요 또한 증가하였다. 최근에는 스마트폰 사용자가 급증하면서 모바일 정보검색 서비스도 다양한 방향으로 발전하고 있다. 본 논문에서는 휴대폰 카메라로 촬영한 간판을 인식하여 그것을 통해 사용자에게 필요한 정보를 제공하는 방식을 적용하였다.



그림 5. 모바일 웹 검색의 예
Fig. 5. Example of mobile web search

스마트 폰의 기존 웹 검색으로는 일반적인 검색 엔진

을 통하여 PC에서 검색하는 것과 같은 형태로 검색하게 된다. 반면 본 논문에서는 2개의 서버와 연동하여 보다 상세한 정보를 쉽게 제공하고자 한다.

사용자가 촬영한 간판 영상을 통해 정보를 검색하고자 할 때 촬영한 영상을 PHP1 서버로 전송한다. PHP1 서버에서는 촬영된 영상 내에서 문자를 추출하여 문자에 대한 기본적인 번역과 정보를 제공한다. 즉, 추출된 문자가 검색어 기능을 하게 된다. 보다 다양한 정보를 검색하기 위해서는 추출된 문자정보를 PHP2 서버로 전송한다. 이 때 저장되어 있는 간판DB정보와의 매칭을 통해 메뉴판, 가격, 위치 등 사용자가 원하는 정보를 제공하게 된다.

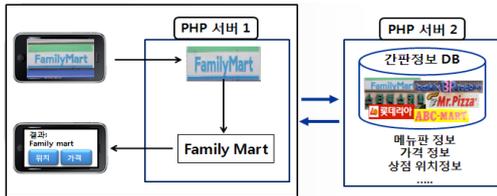


그림 6. 웹 서버 연동을 통한 간판정보 검색
Fig. 6. Sign information search through web server link

예를 들어, 'Family Mart'라는 간판을 스마트폰으로 촬영하면 처음에 PHP 서버1로 보내지게 된다. 보내진 영상은 배경과 객체(문자)를 구분하여 배경에서 'Family Mart' 문자를 추출하고, 이에 해당하는 기본정보를 가져오게 된다. 1차적으로 제공된 정보에서 상세한 정보(가격, 메뉴, 위치 등)를 알고자 하면 PHP 서버2로 보내 'Family Mart'의 위치와 해당 품목의 가격 등을 제공한다.

모바일기기에서는 언제 어디에서도 빠르고 정확성이 높은 정보검색 방법을 연구하고 있다. 이에 따라 추출된 텍스트와 패턴을 비교하는 중 일치하지 않은 문자가 발생했을 시 이전의 일치했던 문자들은 다음 비교과정에서 제외함으로써 효율성을 높이는 KMP 알고리즘을 사용하고자 한다. 이 알고리즘은 기존의 정보검색에 많이 사용되고 BF(Brute Force)알고리즘과는 다르게 이미 수행했던 텍스트와 패턴의 비교들을 기억할 수 있는 수단을 제공한다^[5].

예를 들어 STARBUCKS라는 문자가 추출되어 들어오면 기존의 간판DB에 입력되어있는 간판텍스트와 'S' 문자열부터 하나씩 비교를 시작하게 된다. 비교하는 도

중 다섯 번째 문자열인 'B'에서 일치하지 패턴을 찾지 못한다면 인덱스는 'B'를 건너뛰어 'U'부터 재검색을 시작하도록 한다. 특히 가장 많은 패턴 매칭이 이루어지는 간판텍스트를 찾아 이 구조 아래에 있는 정보까지 검색이 가능하도록 한다.

```

Kmpsearch(text, n, pat, m) /* search pat[1..m] in text[1..n] */
int n, m;
{
    int i, k, resume;
    int next[MAX_PATTERN_SIZE];
    pat[m+1] = CHARACTER_NOT_IN_THE_TEXT;
    initnext(pat, m+1, next);
    resume = next[m+1];
    j = k = 1;
    do { /* Search */
        if(j==0 text[k]==pat[j])
        {
            k++; j++;
        }
        else j=next[j];
        if(j > m)
        {
            Report_match_at_position(k-m);
            j = resume;
        }
    } while(k <= n);
    pat[m+1] = END_OF_STRING;
}
    
```

그림 7. Knuth-Morris-Pratt 알고리즘
Fig. 7. Knuth-Morris-Pratt algorithm

또한 영상에서 추출된 문자를 통해 해당 상점에 대한 정보를 검색한다. 이 때 보다 다양하고 실시간적인 정보 제공을 위해 위치기반서비스인 LBS를 적용하였다. 사용자의 현재 위치로부터 목적지까지의 상대적 거리를 표시하여 원하는 상점의 위치 및 상점의 정보를 파악 할 수 있다.

LBS를 활용한 서비스로는 프랜차이즈, 백화점 등의 매장에서 고객으로부터 위치정보 제공 및 쿠폰·광고 전송 동의를 받아 중계사업자들에게 광고를 요청해주는 서비스가 가능하다.

IV. 실험 및 평가

실험은 스마트폰의 카메라로 고딕체의 문서와 간판을 촬영 할 경우 템플릿 매칭 과정으로 문자 인식의 결과를 확인한다.



그림 8. 스마트 폰에서 image processing 결과
 Fig. 8. Result of image processing in smart phone

스마트폰 카메라로 촬영된 이미지는 image processing을 통해 문자만 추출된다. 이 값은 기존의 표준 패턴값의 픽셀과 새로 입력된 영상 픽셀과 비교하여 텍스트로 나타난다. 이렇게 텍스트로 나타난 문자는 이후 상세한 정보검색으로 편리성을 제공한다.

탭플릿 매칭의 결과 흰색바탕에 검은색 고딕체의 문자를 100회 촬영하였을 경우 모두 정확하게 100%의 인식률을 보장한다. 반면 문자의 크기가 일정하지 않고 배경이 복잡한 간판의 경우 100회 촬영 시 70회 정도만 정확한 문자를 추출해내어 약70%의 인식률을 나타 내어 고딕체 문서보다 인식률이 낮아진다는 것을 볼 수 있다. 따라서 간판의 정확한 인식률을 높여 정보검색의 효율성을 위해 배경(잡음)처리와 여러 글씨체를 인식률이 높아 지도록 하는 연구과정에 있다.

표 1. 스마트 폰의 인식률
 Table 1. recognition rate of smart phone

구분	고딕체 문서	간판
스마트 폰	100회/100회	70회/100회
	100%	70%

V. 결론 및 향후 계획

모바일 기기의 휴대성과 촬영한 영상에 대한 즉각적인 정보제공을 통해 정보를 습득하고자 하는 다양한 계층의 사용자에게 보다 간편한 방법으로 정보를 습득할 수 있도록 하였다. 또한 자체 구축한 PHP 서버와의 연동을 통해 정보 제공을 하는 방법을 적용하였기 때문에 단순히 인터넷 포털사이트에 연동시켜 검색을 할 때의 번

거로움을 줄일 수 있었다, 간판이 검색어 역할을 하여 간판을 통해 상점의 정보를 미리 파악할 수 있도록 하였다.

영상내의 문자 인식률을 높이기 위해 본 논문에서는 Top-Hat 연산을 적용하였다. 그러나 자연영상은 촬영 각도, 조도, 명암 등에 따라 화질 차이가 많이 나기 때문에 문자인식 과정에서 상당한 인식률의 차이를 보인다. 또한 간판내의 문자가 정형화 되어있지 않아 같은 문자임에도 모양이 다양하게 표현된다. 이것을 촬영할 때에도 각도나 방향 등에 따라 인식률에 차이를 보였다. 따라서 다양한 환경에서 촬영한 자연영상에 대해서도 문자인식이 이루어 질 수 있도록 문자 인식률을 높이는 방향으로 연구를 지속할 것이다.

또한 간판정보DB의 정기적인 업데이트를 간판을 통한 정보검색의 정확도를 높이고 정보검색 속도를 향상시킬 수 있는 방법을 연구할 계획이다.

참고 문헌

- [1] 김학수, “모바일 환경에서의 정보 검색”, 2008
- [2] 김병현, 한영준, 한현수, “불규칙 조명 환경에 강인한 번호판 문자 분리 기법”, 한국컴퓨터정보학회, 2009
- [3] 김지훈, 이택현, 김기웅, 김진형, “상향식 접근 방법을 사용한 강인한 간판 인식”, 한국컴퓨터종합학술대회, 34권, 234-235쪽, 2007년
- [4] SeongHun Lee, JinHyung Km, “Complementary combination of holistic and component analysis for recogniton of low-resolution video charcter images”, ScienceDirect, 29, 383-391p, 2008
- [5] 권준식, 디지털 영상처리 이론 및 응용, 홍릉과학출판사, 5-74쪽, 2002년 8월
- [6] 이성호, “부상하는 위치기반 서비스”, 삼성경제연구소, 2007
- [7] 이동진, 이기영, “모델링 기반 시선추적을 이용한 가상모델하우스 시스템”, 한국인터넷방송통신학회논문지, 제10권5호, 223-228p, 2010
- [8] 정용규, 한송이, “유전자 알고리즘을 이용한 웹 검색 랭킹방법”, 한국인터넷방송통신학회논문지, 제10권3호, 91-96p, 2010

※ 본 연구는 지식경제부 지역혁신센터사업중 바이오-메디테크 산업화 지원을 받아 수행된 연구임
(2010-02-10)

저자 소개

임 명 재(중신회원)



- 1998년 중앙대학교 공학박사
- 1992 - 현재 을지대학교
의료 IT마케팅학과 교수

<관심분야: SE 개발방법, HCI, U-Healthcare 등>

현 성 경(준회원)



- 을지대 의료전산학과 학생
- 2010년 iwit 동계 학술대회 우수 논문
상 수상

<관심분야: RFID/USN 기술, 이동무선통신, 영상처리>

박 지 은(준회원)



- 을지대 의료전산학과 학생
- 2010년 iwit 동계 학술대회 우수 논문상
수상

<관심분야: RFID/USN 기술, 이동무선통신, 영상처리>

이 기 영(중신회원) : 교신저자



- 제 10권 1호 참조
- 2009년 - 현재 한국인터넷방송통신학
회 이사
- 1991년 - 현재 을지대학교 의료 IT마
케팅학과 교수

<관심분야: u-Healthcare, 공간데이터베이스, GIS, LBS,
USN, 텔레메틱스 등>