

휴대폰 카메라로 획득한 저해상도 영상에서의 전화번호 인식

박현일^o 김수형

전남대학교 전산학과

songgot@lip.chonnam.ac.kr^o, shkim@chonnam.ac.kr

Recognition of Phone Number from Low Resolution Cellular Phone Images

Hyun Il Park^o Soo Hyung Kim

Dept. of Computer Science, Chonnam National University

요 약

현재 휴대폰을 이용한 다양한 콘텐츠 산업이 부흥기를 누리고 있으며, 휴대폰 제조 회사도 휴대폰에 소비자의 욕구를 충족시킬 수 있는 다양한 기능을 구현하고 있다. 본 논문은 현재 출시되고 있는 휴대폰에 부착된 카메라로 획득된 저해상도의 자연 영상에서 전화번호를 인식시켜 사용자로 하여금 보다 간편하게 원하는 전화번호를 실시간으로 휴대폰에 저장시키는 기능을 구현하고자 한다. 본 논문에서는 저해상도의 실영상에서 전화번호 부분이 추출되었다는 가정 아래, 전화번호를 인식하기 위해 배경과 전화번호 부분을 분리하기 위한 다양한 이진화 실험을 하였으며, 범용 인식기인 아르미6.0을 이용하여 휴대폰으로 획득된 저해상도의 전화번호를 인식하기 위한 최선의 전처리 알고리즘을 제시하고자 한다.

1. 서 론

현대에는 다양한 미디어 기술을 이용하여 인간에게 필요한 정보를 제공하고 있으며, 인간은 좀더 편하게 필요한 정보를 얻기를 원하고 있다.

현재 우리나라는 전 세계적으로 가장 네트워크 망이 발달된 나라 중에 하나이며, 국민의 대다수는 이러한 환경을 이용하여 공간적 제약성에서 벗어나 인터넷을 통하여 필요한 정보를 실시간으로 획득하고 있다. 정보를 획득하는데 필요한 단말기는 예전엔 컴퓨터 시스템으로 제안이 되어 있었지만 현재는 다양한 멀티미디어 시스템을 이용하여 필요한 정보를 얻고 있다. 그 중에 하나가 바로 휴대폰이다.

현재의 휴대폰은 단순히 일 대 일 통화를 위한 기능뿐만 아니라, 다양한 정보를 송수신 하기 위한 여러 기능이 포함되어 있기 때문에 휴대폰을 이용한 다양한 콘텐츠 산업이 부흥기를 누리고 있다.

휴대폰 제조 회사는 경쟁적으로 휴대폰에 좀더 많은 기능을 부여하여 소비자가 휴대폰을 좀더 편리하게 이용하기를 원하고 있고 자사의 제품을 사용하기를 바라고 있다. 이러한 상황에서, 최근 국내에 출시되는 거의 모든 휴대폰은 카메라를 내장하고 있다.

현재의 휴대폰은 음성인식 분야로는 어느 정도의 기능이 구현되어 있지만, 영상처리를 이용한 정보획득 분야에는 그다지 진전이 없다. 따라서 휴대폰 카메라로 획득된 영상에서 상호명과 전화번호를 인식하는 시스템을 구현하려고 했으며 1단계로 전화번호를 인식하는 시스템을 만들게 되었다 -그림 1 참조.

현재 자연 영상에서 문자영역을 추출하는 다양한 연구가 진행되고 있다. 때문에 본 논문에서는 저해상도 영상에서 전화번호 부분이 추출되었다는 가정 하에, 전화번호를 인식하기 위해 배경과 전화번호 부분을 분리하기 위한 이진화 방법을 제시하고자 한다.

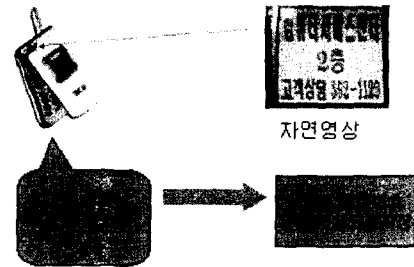


그림 1 자연영상에서 획득된 정보를 저장하는 과정

이진화 성능을 객관적으로 분석하기 위해 범용 인식기인 아르미를 사용한다. 본 실험에서는 72~92dpi의 저해상도의 이미지를 다양한 알고리즘을 이용하여 이진화를 한 후에 아르미6.0에 인식 시켜서 최적의 전처리 알고리즘을 찾는데 초점을 두었다.

2. 제안방법

실험에 사용할 데이터를 획득하기 위해 최대한 다양한 배경과 전경을 가진 영상과 다양한 잡음을 가진 영상을 획득하려고 노력하였다. 그리고 이진화 성능을 객관적으로 평가하기 위해 전경과 배경, 잡음이 서로 다른 22개의 영상을 선택하여 다양한 이진화 알고리즘[1]을 적용하여 아르미를 이용하여 성능 평가를 하였다.

실험에 사용된 영상은 전화번호만을 확대해서 휴대폰 카메라로 찍은 것이 아니라 그림 2와 같이 전체적인 정보를 얻을 수 있는 작은 사이즈의 영상 내에서 전화번호를 추출하여 이진화 실험을 하였다.

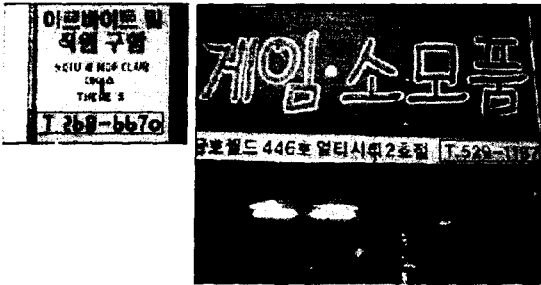


그림 2 휴대폰 카메라로 찍은 영상

2.1 Gray Image로 변형 후 인식

22개의 실험 영상을 Gray Image로 변환 시킨 후 다양한 이진화 알고리즘 [1]을 이용하여 실험하였지만 72dpi의 저해상도 이미지가기 때문에 글자와 배경 부분이 정확히 분리되지 않았으며 글자가 많이 번져서 좋은 인식률을 얻지 못했다.

911-1428 07-620-0477

그림 3 Gray Image를 이진화한 영상

표 1 Gray Image를 이용한 이진화 인식 결과

알고리즘	1		2		3		4	
True / False	T	F	T	F	T	F	T	F
총 167 자	83	84	67	60	82	85	97	90
문자별 인식률(%)	49.7	50.3	52.1	47.9	49.2	50.8	58	42
필드 인식률(개수)	2개		2개		2개		1개	

- 1:3Means Clustering+Moving Average Horizontal+FBC1x1
- 2:Laplacian(에지픽셀이용) + Moving Average Horizontal+ FBC1x1
- 3:Laplacian(에지픽셀이용)+ Moving Average Vertical+FBC3x5
- 4:Otsu가 제안한 알고리즘
- * FBC NxM FBC 방법에서 파라미터가 N, M임을 의미

2.2 칼라 정보를 이용한 이진화 후 인식

2.2.1 Color Reduction에 의한 이진화

영상에 잡음이 많기 때문에 Color Reduciton[2]을 실행하면 그림 4와 같이 잡음이 많이 없어지고 영상에 존재하는 색깔의 수가 현저하게 줄어들기 때문에 이진화 성능이 좋을 것이라 생각하였다. 그러나 실험 결과는 2.1의 실험에 비하여 성능이 많이 떨어졌다. 이때 상위 2비트를 남긴 것 보다 3비트를 남긴 것이 성능이 좋았다.



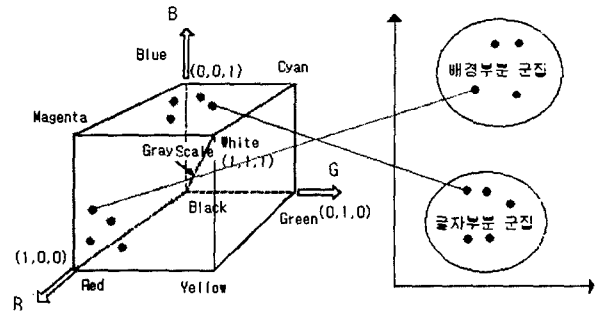
그림 4 Color Image를 Color Reduction한 영상

표 2 Color Reduction을 이용한 이진화 결과

알고리즘	1		2		3		4	
2bit / 3bit	2	3	2	3	2	3	2	3
총 167 자	61	79	60	68	62	57	65	63
문자별 인식률(%)	49.7	50.3	52.1	47.9	49.2	50.8	58	42
필드 인식률(개수)	0개	1개	0개	0개	0개	0개	0개	0개

2.2.2 Color Clustering에 의한 이진화

Color Clustering은 2-Means를 이용하여 RGB 3차원 정보를 2차원 상에 Mapping하여 군집화를 하였다.



019-605-3966 980-521-4225

그림 5 2-Means를 이용한 Color Clustering 방법

2-Means 알고리즘을 적용할 경우 클러스터의 개수가 적은 쪽이 글자영역일 확률이 높다. 하지만 그렇지 않은 경우가 있기 때문에 반전되는 경우가 있다. 따라서 논문에서는 군집 1과 군집 2의 최외각 바운드 박스의 크기를 비교하는 방법을 제안하여 위의 문제를 해결할 수 있었으며 총 96개의 전화번호 영상을 대상으로 실험한 결과 다음과 같은 인식률이 나왔다.

표 3 2-Means를 이용한 인식 결과

알고리즘	2 Means (96개 영상 실험)	
문자별 인식률(%)	True	513자 (총 703자 중) (72.97%)
	False	190자 (총 703자 중) (27.03%)
필드 인식률(개수)	1차 틀린 영상	23개 (23.95%)
	모두 맞춘 영상	25개 (26.04%)

980-521-4225

그림 6 2-Means에서 개선 시켜야 할 영상

그림 6의 빨간선 부분을 3-Means를 이용하여 제거한다면 더 좋은 인식률을 보일 것이라 예상하고 3-Means 실험을 하였지만 그림 7과 같이 배경과, 글자, 그리고 글자주변의 잡음을 정확히 분리하기 어려운 영상이 존재하기 때문에 인식률은 2-Means보다는 떨어졌다.

하지만 2-Means로 인식하지 못한 영상을 3-Means에서 인식하는 경우가 있었다.

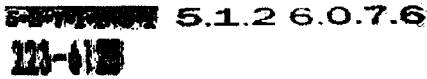


그림 7 3-Means 클러스터링 (검정, 빨강, 흰색)

2.2.3 단단계 Color Clustering에 의한 이진화

2-Means나 3-Means 중 어느 한 개의 알고리즘만을 사용했을 경우 보다 2-Means를 이용하여 인식 시킨 후 인식되지 않았을 때 3-Means를 이용하는 알고리즘을 사용하면 표 4와 같이 2-Means의 성능보다 좋게 나왔다.

표 4 단단계 Color Clustering을 이용한 인식 결과

알고리즘		2-Means 3-Means (96개 영상 실험)	
문자별 인식률(%)	True	535자 (총 703자 중) (76.1%)	
	False	168자 (총 703자 중) (23.9%)	
필드 인식률(개수)	1자 틀린 영상	25개 (26.04%)	
	모두 맞춘 영상	27개 (28.125%)	

2.2.4 Filtering 및 Color Clustering을 이용한 이진화

인식에 사용되는 영상은 휴대폰으로 획득된 저해상도 영상이기 때문에 숫자영역의 윤곽선이 불투명한 영상이 대부분이다. 이런 영상을 단순히 클러스터링 하는 것보다는 윤곽선 부분을 부각시키기 위해 Highpass 필터를 사용하였다. 하지만, 고주파 통과로 인하여 영상의 전체적으로 생긴 점들은 인식률을 떨어뜨리기 때문에 Lowpass 필터를 이용하여 잡음을 없앴다. 이렇게 하면 Original Image 보다는 윤곽선이 잘 나타나고 잡음도 없어지기 때문에 인식률이 더 높아졌다. Lowpass→Highpass→lowpass, Lowpass→Highpass 등 다양한 실험을 해 보았지만 Highpass→Lowpass 조합이 가장 이진화가 잘되었으며, 인식률도 높았다. 표 5는 Filtering 후 2-Means를 이용하여 이진화를 한 후 아르미에 인식 시킨 결과이다.

표 5 Filtering 후 2-Means 를 이용한 인식

알고리즘		Filtering 후 2-Means(96개 영상실험)	
문자별 인식률(%)	True	530자 (총 703자 중) (75.4%)	
	False	173자 (총 703자 중) (24.6%)	
필드 인식률(개수)	1자 틀린 영상	20개 (20.83%)	
	모두 맞춘 영상	29개 (30.20%)	

그림 8과 같이 Highpass 및 Lowpass Filtering을 거친 후에 2-Means를 이용한 인식 후, 인식되지 않은 영상에 대해서 3-Means를 이용하는 시스템을 사용하면 표 6과 같이 더 높은 인식률이 나왔다.

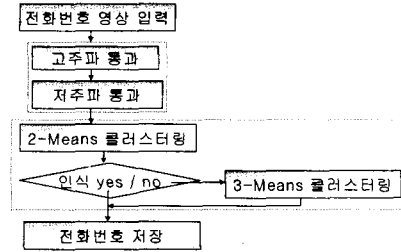


그림 8 저해상도 영상 이진화 알고리즘

표 6 Filtering 및 단단계 Color Clustering을 이용한 인식 결과

알고리즘	Filtering 후 2 Means 3 Means (96개 영상 실험)	Filtering 후 2 Means 3 Means (96개 영상 실험)	
		문자별 True	548자 (총 703자 중) (77.95%)
인식률(%)	False	155자 (총 703자 중) (24.05%)	
필드 인식률(개수)	1자 틀린 영상	26개 (27.08%)	
	모두 맞춘 영상	40개 (41.66%)	

3 결론 및 향후 연구 방향

본 논문은 휴대폰 카메라로 획득한 저해상도 이미지 (72~92dpi)에서 전화번호를 인식하기 위한 전처리 알고리즘을 제안하였다. 본 실험은 전화번호 인식 시스템 중 에서 이진화 단계를 다루었으며 이진화 성능을 판별하기 위한 객관적인 평가를 위해 아르미6.0을 이용했다.



그림 9 아르미 인식결과

그림 9는 이진화가 잘되었지만 아르미가 잘못 인식한 경우이다. 전용 인식기를 개발하여 인식시킨다면 좀더 나은 결과를 얻을 것으로 기대 된다. 앞으로 이진화 성능을 좀더 개선시키고 전화번호를 잘 인식하는 최적의 인식기를 설계하는 것이 향후 연구과제 이다

참고 문헌

- [1] 박상철, 김수형, "주인등복중 영상의 이진화를 위한 다중 결함 알고리즘," 정보과학회 논문지, 2003 (심사중).
- [2] 노명철, 최영우, 이상환, "색 및 명도 정보를 이용한 장면 텍스트 추출," 제 14회 영상처리 및 이해에 관한 워크샵 논문집, pp.515-520, 2002.
- [3] 이상준, 홍충선, 곡은식, 이대영, "내용기반 영상검색을 위한 칼라 영상 분할," 한국정보처리학회 논문지 제7권 제9호, pp.2994-3001, 2000.9.
- [4] 김의정, 정원일, "칼라 문서에서 문자 영역 추출을 위한 클러스터링기법," 제14회 대전산업대학교 논문집, pp. 107-116, 1997.
- [5] 김지수, 김수형, "명도 정보를 이용한 자연영상에서의 텍스트 영역 추출," 제3회 한국정보처리학회 호남제주주부 학술발표논문집, 제3권 제1호, pp.127-132, 2003.
- [6] R.C Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.
- [7] 김상운, 식별 알고리즘을 중심으로 한 패턴인식 이론, 흥릉과학출판사, 1995.