

순차적 칼라 클러스터링에 기반한 휴대폰 카메라 영상에서의 숫자열 인식

박현일⁰ 김수형

전남대학교 전산학과

songgot@iip.chonnam.ac.kr⁰, shkim@chonnam.ac.kr

Recognition of Digit Strings from Cellular Phone Image by Sequential Color Clustering

Hyun Il Park⁰, Soo Hyung Kim

Dept. of Computer Science, Chonnam National University

요 약

자연영상에서 획득된 문자를 인식하는 연구는 대부분 디지털 카메라나 캠코더를 이용하여 획득된 고해상도 영상을 입력영상으로 사용하고 있다. 본 논문에서는 휴대폰 카메라로 획득된 저해상도 영상을 입력영상으로 사용하였다. 저해상도의 영상은 적은 수의 픽셀로 정보를 표현하고 있기 때문에 기존에 제시되었던 다양한 이진화 방법으로는 문자와 배경을 깨끗하게 분리해 낼 수 없다. 본 논문은 입력영상의 이진화를 위해 K-Means 알고리즘을 이용하여 칼라 클러스터링을 하였으며, 이진화 성능을 향상시키기 위해 지능형 주파수 필터를 사용하였다. 이진화된 영상을 파이프라인 구조의 인식 시스템에 인식시킴으로써 기존의 제안 방법들에 비하여 인식 성능을 향상 시킬 수 있었다.

1. 서 론

우리가 사용하는 휴대폰은 전화기처럼 단순히 일 대일 통화를 위한 기구는 아니다. 다양한 형태의 정보의 송수신뿐만 아니라, 사용자가 보다 편리한 생활을 할 수 있도록 여러 가지 콘텐츠를 제공하고 있다. 휴대폰 제조업체나 이동통신 사업자는 휴대폰에 좀더 많은 기능을 부여하여 소비자로 하여금 자사의 서비스나 제품을 사용하도록 유도하고 있다. 이러한 상황에서 최근 국내외에 판매되는 휴대폰은 대부분 내장형 및 외장형 카메라를 장착하고 있다. 현재 휴대폰은 음성인식, 통신, 금융, 게임 등의 분야로는 어느 정도의 기능이 구현되어 있지만 영상처리를 이용한 정보획득 분야에는 진전이 없다. 따라서 휴대폰 카메라로 획득된 저해상도 영상에서 상호영과 전화번호를 인식하여 휴대폰에 저장하는 시스템을 구현하고자 하였으며 1단계로 전화번호를 인식하는 시스템을 제안한다.

전화번호를 인식하기 위해서는 입력영상을 이진화 하여야한다. 하지만 영상자체가 72dpi의 저해상도 영상이기 때문에 기존의 다양한 이진화알고리즘으로는 전화번호를 깨끗하게 배경과 분리해 낼 수 없다[1][2]. 그래서 본 논문에서는 칼라 정보를 이용하여 K-Means 클러스터링 알고리즘을 이용한 이진화 방법과 지능형 주파수 필터를 사용함으로써 이진화 성능을 향상 시킬 수 있었으며 기존의 이진화 알고리즘에 의해 이진화된 영상 보다 인식 성능을 높일 수 있었다. 그리고 입력 영상을 파이프라인 구조의 이진화 및 인식 시스템에 인식시킴으로써, 단단계 칼라 클러스터링을 이용한 이진화 및 인식 시스템[3] 보다 성능이 좋았다. 이진화 성능을 측정하기 위해서 범용 인식기인 아르미를 사용하였다. 특정한 영상의 효과적인 이진화를 할 수 있는 알고리즘을 찾기 보다는 다양한 입력영상의 이진화 성능을 전체적으로 높일 수 있는 알고리즘을 찾는데 초점을 두었다.

2. 제안하는 이진화 및 인식 시스템

어느 특정한 영상의 이진화 성능을 높일 수 있는 알고리즘이 있다면 그 알고리즘을 다른 영상에 실험을 하면 좋지 못한 결과를 얻는다. 따라서 다양한 배경색과 전경(전화번호)색을 가진 입력영상들 모두의 이진화 성능을 높이기 위해서 그림 1과 같은 시스템을 제안한다.

제안하는 파이프라인 구조의 이진화 및 인식 시스템은 입력영상을 그림 1의 파이프라인으로 마치 물처럼 흘러 보낸 다음 각 단계에서 이진화 및 인식을 수행하고 오인식일 경우 수문을 열고 아랫방향으로 다시 영상을 흘러 보내는 시스템이다.

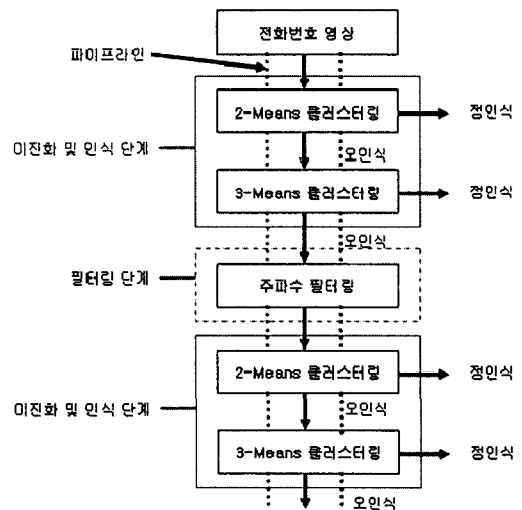


그림 1 제안하는 이진화 및 인식시스템

2.1 입력영상의 특징

실험에 사용된 총 96개의 칼라영상은 30만 화소의 휴대용 카메라로 획득된 72dpi의 저해상도 영상으로서 배경과 전경의 색조함을 최대한 다양하게 하여 획득하였다. 입력영상은 전화번호부문을 확대하여 찍은 영상이 아니라 그림 2와 같이 전체적인 정보를 얻을 수 있는 작은 사이즈의 영상 내에서 전화번호부문을 추출하여 인식하였다.

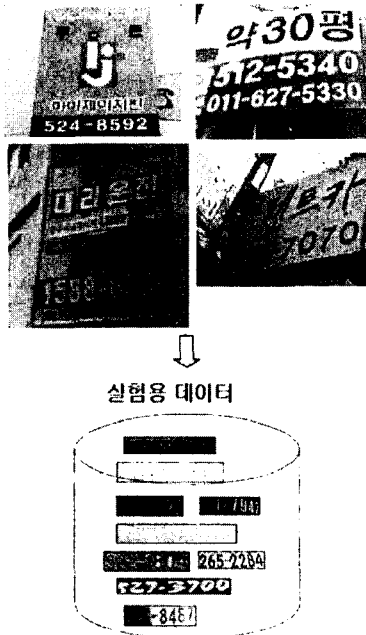


그림 2 휴대폰으로 획득된 영상 및 실험용 데이터

2.2 이진화 및 인식단계

제안하는 시스템은 영상을 이진화하기 위해서 2-Means 클러스터링 [4]과 3-Means 클러스터링을 사용하며 지능형 주파수 필터와 파이프라인 구조의 이진화 및 인식 시스템을 사용한다. 72dpi의 입력영상을 이진화하기 위해서는 기존의 다양한 이진화 알고리즘에 보다는 주파수 필터로 전처리를 수행한 후에 2-Means 클러스터링 방법과 3-Means 클러스터링 방법을 이용하여 이진화하는 것이 성능도 좋고 인식률도 높다 [3].

2.2.1 2-Means 클러스터링

2-Means 클러스터링을 사용하는 이유는 입력 영상이 대부분 크게 배경과 전경(전화번호)의 두 가지 색으로 구분되기 때문이다. 따라서 입력되는 픽셀 P_{ij} 들을 전경색의 군집과 배경색의 군집으로 칼라정보(R, G, B)를 이용하여 클러스터링하게 된다. 클러스터링 과정을 간단하게 수식으로 나타내면 다음과 같다.

군집1의 중심 : $R1, G1, B1$

군집2의 중심 : $R2, G2, B2$

입력픽셀 : $P(i, j) \rightarrow \{R_{ij}, G_{ij}, B_{ij}\}$

$$d1_{ij} = \sqrt{(R_{ij} - R1)^2 + (G_{ij} - G1)^2 + (B_{ij} - B1)^2}$$

$$d2_{ij} = \sqrt{(R_{ij} - R2)^2 + (G_{ij} - G2)^2 + (B_{ij} - B2)^2}$$

if ($d1_{ij} < d2_{ij}$) then $P(i, j) \in$ 군집1

else $P(i, j) \in$ 군집2

2.2.2 3-Means 클러스터링

2-Means 클러스터링을 이용하여 이진화를 하였을 경우에는 그림 3과 같이 이진화된 전화번호 부분에 잡음이 생기는 경우가 있다. 그래서 3-Means 클러스터링을 이용하면 잡음영역을 제거 할 수 있다. 2-Means 클러스터링 방법보다 3-Means 클러스터링 방법을 주 이진화 방법으로 사용하면 좋다고 생각할 수도 있지만 그림 4와 같이 배경군집을 제외한 숫자부분에 2개의 군집(빨강, 검정)이 나타나므로 3-Means 클러스터링 방법은 2-Means 클러스터링을 보조하는 방법으로 사용하는 것이 좋다.

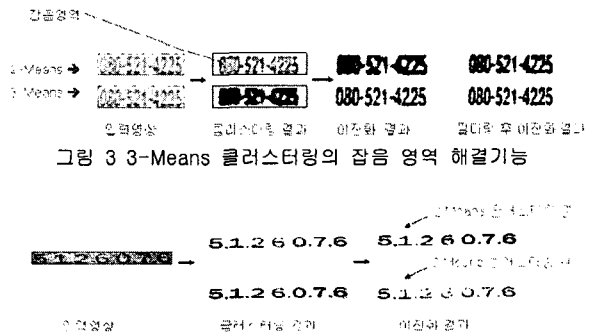


그림 3 3-Means 클러스터링의 잡음 영역 해결기능

그림 4 3-Means 클러스터링의 단점

2.3 필터링 단계

저해상도의 입력영상의 이진화성능을 향상시키기 위해서 주파수 필터 [5]를 사용한다. 입력영상의 윤곽선 부분을 강화하기 위해서 고주파통과 필터를, 고주파통과 필터 사용 후에 생긴 잡음을 제거하기 위해 저주파통과 필터를 사용한다. 그림 5의 위쪽은 주파수 필터를 사용하지 않은 것이고 아래쪽은 주파수 필터를 사용했을 때 이진화 결과이다.

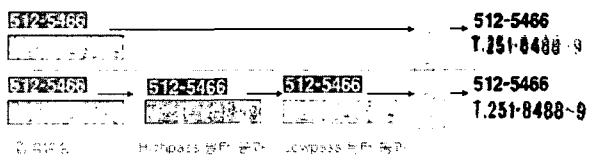


그림 5 주파수 필터링의 효과

2.3.1 지능형 주파수 필터링

입력영상에 고주파를 통과시키면 명도값 1의 범위가 0~255사이의 범위를 벗어나게 된다. 통상적으로 255 이상의 값은 255로 0 이하의 값들은 0으로 변환했다. 제안하는 주파수 필터는 픽셀들 상호간의 명도값의 차를 유지하면서 해당 픽셀의 명도값을 0~255의 범위로 매핑시킨다. 또한 어느 특정한 픽셀이 아주 높은 명도 값이나, 적은 명도 값을 가지면서 그 빈도수는 적을 수 있으므로, 전체 픽셀의 정규분포를 파악하여 98%에 해당하는 픽셀에 대해서 가장 높은 밝기 값과 가장 낮은 밝기 값을 구하여 0~255 사이로 매핑시켰다. 변환식은 다음과 같다.

L_{Max} : 98%의 픽셀들중 가장 높은 밝기 값

L_{Min} : 가장 낮은 밝기 값

$Interval$: $L_{Max} + |L_{Min}|$

$Scale$ (1눈금 간격) : $Interval/256$

변환식: $y = (입력되는\ 명도값 + |L_{Min}|) / Scale) - 0.5$

96개의 칼라영상으로 기존의 주파수 필터를 사용했을 경우와 지능형 주파수 필터를 사용했을 경우, 2-Means 클러스터링과 3-Means 클러스터링으로 이진화후 인식 실험결과 필드인식률과 문자별 인식률이 각각 8.34%, 2.1% 향상되었다.

2.4 제안하는 시스템의 인식 성능 평가

총 96개의 실험 데이터를 그림 1의 시스템에 인식시킨다. 인식기는 이진화 성능을 측정하기 위해서 범용 인식기인 아르미를 사용하였다. 실험영상은 단계별로 이진화 및 인식을 하게 되고 오인식된 영상은 파이프라인을 타고 흐르는 물처럼 다음 단계로 넘어가게 된다. 인식결과 그림 7과 같이 총 96개 영상 중 51개(53.13%)의 필드인식률과 총 703자 중 558개(79.37%)의 문자별 인식률이 나왔다. 이것은 박현일등[3]이 제안한 시스템의 인

표 1 이진화 후 인식 성능 비교

96개 영상 703자		
알고리즘	참고문헌[3]	제안시스템
문자별 인식률(%)	77.95% (548자)	79.37% (558자)
필드 인식률(%)	41.66% (40개)	53.13% (51개)

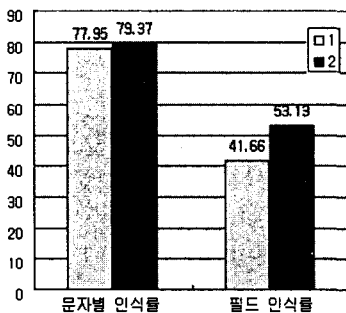


그림 6 성능 비교 그래프

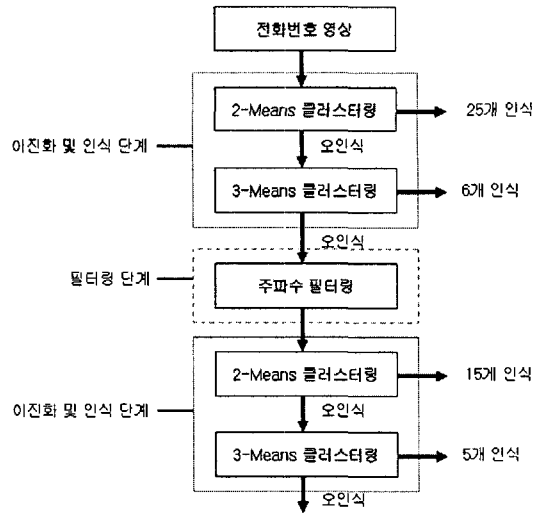


그림 7 단계별 인식 결과

식률에 비하여 필드인식률은 11개(11.46%) 향상되었으며 문자별 인식률은 10자(1.42%) 향상되었다.

3. 결론 및 향후 연구방향

휴대폰 카메라로 획득된 저해상도 영상(72dpi)의 이진화는 영상의 특징 때문에 기존의 다양한 이진화 알고리즘 보다는 본 논문에서 제안한 칼라정보를 이용한 K-Means 클러스터링 방법과 지능형 주파수 필터를 사용하는 것이 이진화 성능이 더 좋았으며, 파이프라인 구조의 이진화 및 인식 시스템을 사용함으로써 인식 성능을 높일 수 있었다.

향후 연구로는 이진화 성능을 더욱 향상시킬 수 있는 알고리즘과 범용인식기가 아닌 전화번호의 특성에 맞는 전용인식기의 개발이 이루어져야 할 것이다.

참고 문헌

- [1] 박상철, 김수형, "주민등록증 영상의 이진화를 위한 다중 결합 알고리즘," 정보과학회 논문지, 2003 (심사중).
- [2] 노명철, 최영우, 이성환, "색 및 명도 정보를 이용한 장면 텍스트 추출," 제 14회 영상처리 및 이해에 관한 워크샵 논문집, pp.515-520, 2002.
- [3] 박현일, 김수형, "휴대폰 카메라로 획득한 저해상도 영상에서의 전화번호 인식", 한국정보처리학회 학술발표논문집 (B), pp. 691-693, 2004.
- [4] 김상운, 식별 알고리즘을 중심으로 한 패턴인식 입문, 홍릉과학출판사, 1995.
- [5] R.C Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.
- [6] 이상훈, 홍충선, 곽윤식, 이대영, "내용기반 영상검색을 위한 칼라 영상 분할," 한국정보처리학회 논문지 제7권 제9호, pp.2994-3001, 2000.9.
- [7] 김지수, 김수형, "명도 정보를 이용한 자연영상에서의 텍스트 영역 추출," 제3회 한국정보처리학회 호남제주지부 학술발표논문집, 제3권 제1호, pp.127-132, 2003.