

칼라 군집화 및 비등방성 확산필터를 이용한 저해상도 영상에서의 숫자열 인식

박현일*, 김수형
전남대학교 전산학과

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

Recognition of Digit String from Low Resolution Image by using Color Clustering and Anisotropic Diffusion

Hyun-Il Park*, Soo Hyung Kim

*Dept of Computer Science, Chonnam National University

요 약

자연영상에서 문자를 인식하는 연구는 활발히 진행되고 있지만 대부분 디지털 카메라나 캠코더 등으로 획득한 고해상도의 영상에서의 연구에 국한되어 있다. 휴대폰 카메라로 획득된 저해상도의 영상은 아주 적은 수의 픽셀로 정보를 표현하기 때문에 기존의 이진화 알고리즘으로는 문자와 배경을 깨끗하게 분리해 낼 수 없다. 본 논문은 영상의 칼라정보를 K-Means 클러스터링을 이용하여 전경과 배경으로 이진화 하였으며, 이진화 성능을 향상시키기 위해 지능형 주파수 필터와 비등방성 확산 필터를 사용하였다. 또한 입력영상을 파이프라인 구조의 이진화 및 인식 시스템에 인식시킴으로써 인식 성능을 향상시켰다.

1. 서론

현재 휴대폰은 음성, 통신, 금융, 네트워크, 게임 등등 여러 분야에 걸쳐서 기술개발이 이루어지고 있다. 하지만 영상처리를 이용한 정보획득 분야에는 기술 개발의 진전이 없다. 요즘 판매되는 대부분의 휴대폰은 카메라를 장착하고 있다. 휴대폰 카메라로 획득된 영상에서의 문자인식 연구는 이루어지지 않았다. 따라서 휴대폰 카메라로 획득된 저해상도 영상에서 상호명과 전화번호를 인식하여 휴대폰에 저장하는 시스템을 구현하고자 하였으며 1단계로 전화번호를 인식하는 시스템을 만들게 되었다.

입력된 영상의 인식성능을 높이기 위해서는 전화 번호를 잘 분리해 낼 수 있는 이진화 알고리즘이 필요하다. 하지만 휴대폰 카메라로 획득된 영상들은 72dpi의 저해상도 영상이기 때문에 기존에 사용되었던 다양한 이진화 방법으로는 영상을 깨끗하게 이진화 할 수 없었다[1][2]. 그래서 본 논문은 입력 영상의 칼라정보를 K-Means 클러스터링을 이용하여 이진화 하였으며 잡음은 제거하면서 유효한 정보는 부

각시키는 비등방성 확산필터와 지능형 주파수 필터를 사용하여 이진화 성능을 개선시켰다. 이것은 기존의 다양한 이진화 방법보다 이진화 성능이 좋다. 전화번호의 인식은, 입력영상을 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 이진화 성능을 높일 수 있는 알고리즘 보다는 다양한 입력영상의 이진화 성능을 전체적으로 높일 수 있는 알고리즘을 찾는 데 주안을 두었다.

2. 제안하는 이진화 및 인식 시스템

전화번호 인식시스템을 상용화 한다면 특정한 패턴을 가진 영상의 이진화 성능을 높일 수 있는 알고리즘보다는 다양한 배경색과 전경색을 가진 영상들 모두의 이진화 성능을 높일 수 있는 알고리즘을 선택해야 한다. 따라서 본 논문에서는 그림 1과 같은 시스템을 제안한다. 제안하는 파이프라인 구조의 이

진화 및 인식 시스템은 입력영상을 그림 1의 파이프라인으로 마치 물처럼 흘러보내며 각 단계에서는 이진화 및 인식을 수행하고 오인식일 경우 수문을 열고 아랫방향으로 영상을 흘러 보내는 시스템이다.

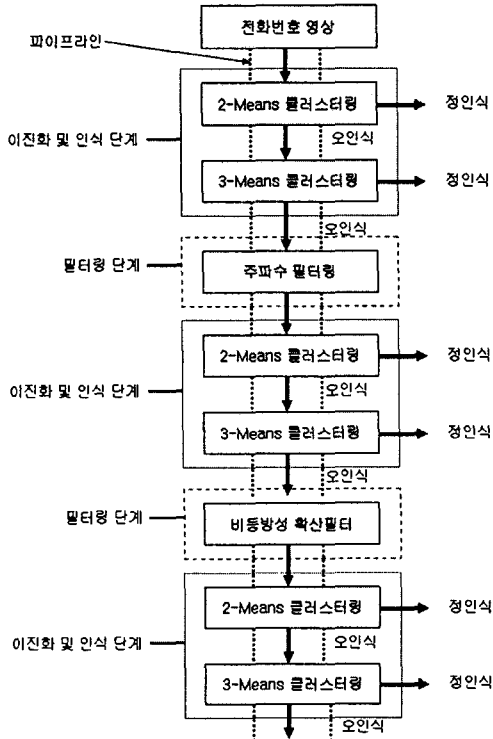


그림 1 제안하는 이진화 및 인식시스템

2.1 입력영상의 특징

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

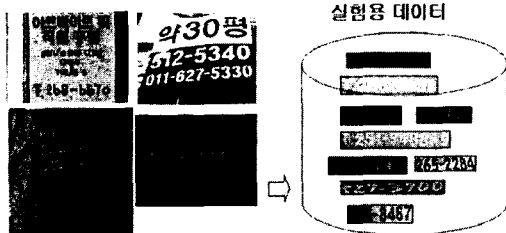


그림 2 휴대폰으로 획득한 영상 및 실험 데이터

2.2 이진화 및 인식단계

제안하는 시스템은 영상을 이진화하기 위해서 2-Means 클러스터링 과 3-Means 클러스터링을 사용하며 지능형 주파수 필터와 비등방성 확산필터, 그리고 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 인식율도 높다[3]. 본 논문에서는 [3]에서 사용한 주파수 필터를 좀더 개선시켰으며, 영상의 잡음을 제거하고 윤곽선정보는 부각시키는 비등방성 확산필터를 추가로 사용하여 이진화 성능을 높였다.

2.2.1 2-Means 클러스터링

입력 영상은 대부분 크게 배경과 전경의 두 가지 색으로 구분되기 때문에 2-Means 클러스터링[4]을 주 이진화 방법으로 사용한다. 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.

$$\text{군집1의 중심} : R1, G1, B1$$

$$\text{군집2의 중심} : R2, G2, B2$$

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

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

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

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

$$\text{else } P(i, j) \in \text{군집2}$$

2.2.2 3-Means 클러스터링

본 논문에서 3-Means 클러스터링을 이용한 이진화 방법은 2-Means 클러스터링을 이용하여 이진화 하였을 경우에 발생하는 단점을 보완하기 위해 사용한다. 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 배경군집(흰색)을 제외한 숫자부분에 2개의 군집(빨강,검정)이 나타나므로

3-Means 클러스터링 방법은 2-Means 클러스터링의 단점을 보완하는 방법으로 사용하는 것이 좋다.

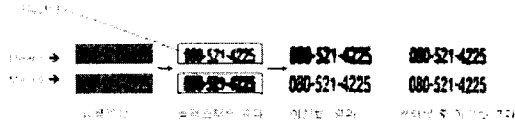


그림 3 3-Means 클러스터링의 잡음영역 해결기능

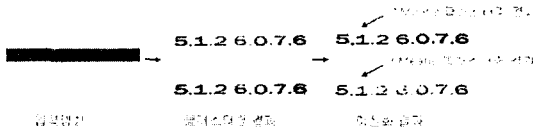


그림 4 3-Means 클러스터링의 단점

2.3 필터링

저해상도의 입력영상의 이진화능을 향상시키기 위해서 영상내의 잡음은 제거하고 윤곽선 정보는 강화시키는 주파수 필터[5]와 비등방성 확산 필터[6]를 사용한다. 그림 5의 위쪽은 주파수 필터를 사용하지 않은 것이고 중간은 비등방성 확산필터, 아래쪽은 주파수 필터를 사용 했을 때의 결과이다.

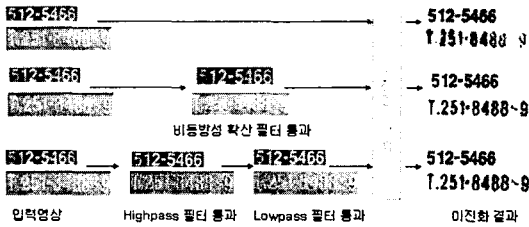


그림 5 주파수 필터와 비등방성 확산필터의 효과

2.3.1 주파수 필터

입력영상에 고주파를 통과시키면 명도값 I의 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 명도값이나, 적은 명도값을 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.

I_{Max} : 98%의 픽셀들중 가장 높은 밝기 값

I_{Min} : 가장 낮은 밝기 값

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

Scale(1눈금 간격) : Interval/256

변환식: $I = (\text{입력되는 명도값} + (|I_{Min}|) / \text{Scale}) - 0.5$

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

2.3.2 비등방성 확산필터

본 논문에서는 저해상도 영상의 윤곽선 정보를 유지하면서 영역 내부를 부드럽게 하여 영상의 잡음을 제거하기 위해서 아래식과 같은 비등방성 확산 필터를 사용하였다[6].

$$I_0 = \text{div}(\text{div}(I, 0, 0) \nabla I) = \text{div}(I, 0, 0) \Delta I + \nabla \text{div} \nabla I$$

여기서 div 는 발산연산, ∇ 는 그래디언트, Δ 는 라플라시안을 의미한다. 즉, $c(x, y, t) \Delta I$ 는 영상을 부드럽게 하는 역할을 수행하며, $\nabla c \nabla I$ 는 그래디언트가 높은 경계선 정보를 유지하는 역할을 한다. Perona등[6]이 제안한 비등방성 확산필터는 동, 서, 남, 북 네 방향의 픽셀정보를 이용한 확산필터로 아래와 같은 식으로 표현한다.

$$I_0^{011} = I_0^0 + \lambda [0_N \nabla_N I + 0_0 \nabla_0 I + 0_E \nabla_E I + 0_W \nabla_W I]$$

I 는 영상을 의미하고 λ 는 전체적인 확산의 영향을 결정하는 상수로 일반적으로 $0 \leq \lambda \leq 1/4$ 값을 사용할 때 안정적으로 이루어진다. ∇I 는 4 방향 그래디언트 값으로 아래의 계산 방법으로 정의된다.

$$\nabla^N I_{i,j} \equiv I_{i-1,j} - I_{i,j}$$

$$\nabla^S I_{i,j} \equiv I_{i+1,j} - I_{i,j}$$

$$\nabla^E I_{i,j} \equiv I_{i,j+1} - I_{i,j}$$

$$\nabla^W I_{i,j} \equiv I_{i,j-1} - I_{i,j}$$

전도 계수(conduction coefficient) C_D 는 (식 4)와 같이 그래디언트에 반비례하게 정의된다.

$$C_D = \text{div}(|\nabla D|) \quad \text{div} D = \{N, S, E, W\}$$

$$\text{div}(|\nabla I|) = \frac{1}{1 + \left(\frac{|\nabla I|}{K}\right)^2}$$

C_D 는 영상의 그래디언트가 큰 값을 가지면 작아지고 그래디언트가 작은 값을 가지면 큰 값을 가진다.

2.4 제안하는 시스템의 인식 성능 평가

제안한 시스템을 이용하여 총 96개의 실험 데이터를 인식시키면 그림 6과 같은 인식 결과가 나온다. 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 총 96개 영상 중 54개(56.25%)의 필드인식률과 총 703자 중 568개(80.80%)의 문자별 인식률이 나왔다. 이것은 기존시스템[3]의 인식률에 비하여 필드인식률은 14개(14.58%) 향상되었으며 문자별 인식률은 20자(2.84%) 향상된 결과이다.

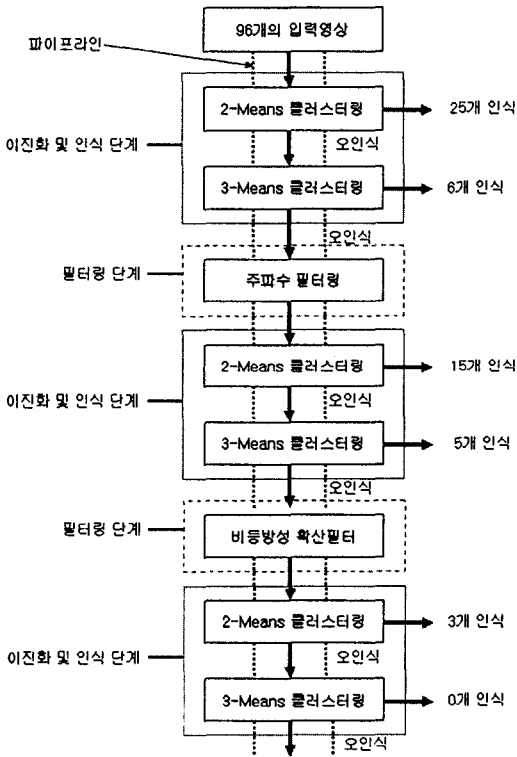
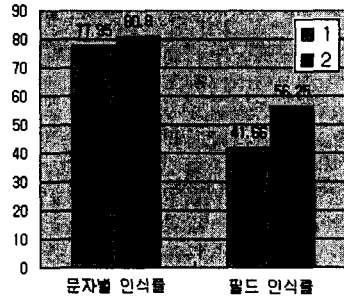


그림 6 제안된 시스템에서의 인식결과

표 1 이진화 후 인식 성능 비교

96개 영상 703자		
알고리즘	참고문헌[3]	제안시스템
문자별 인식률(%)	77.95% (548자)	80.80% (568자)
필드 인식률(%)	41.66% (40개)	56.25% (54개)



1: 참고문헌[3]
2: 제안시스템

그림 7 인식 성능 비교 그래프

4. 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 으로 이진화 하였고 이진화 성능을 향상시키기 위해 지능형 주파수 필터와 비등방성 확산 필터를 사용하였다. 또한 입력 영상을 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.

참고문헌

1 [박상철, 김수형, "주민등록증 영상의 이진화를 위한 다중 결합 알고리즘," 정보과학회 논문지, 2003 (심사중).
[2] 노명철, 최영우, 이성관, "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 추계 학술 발표논문집(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] P.Perona and J.Malik, "Scale Space and Edge Detection Using Anisotropic Diffusion," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.12, no.7, pp.629-639, July 1990.
[7] 이상훈, 홍충선, 박윤식, 이대영, "내용기반 영상검색을 위한 칼라 영상 분할," 한국정보처리학회 논문지 제 7권 제9호, pp.2994-3001, 2000.9.
[8] 김지수, 김수형, "명도 정보를 이용한 자연영상에서의 텍스트영역 추출," 제3회 한국정보처리학회 호남제주지방 학술발표논문집, 제3권 제1호, pp.127-132, 2003.