

논문 2006-43CI-5-4

Otsu의 방법을 개선한 멀티 스레쉬홀딩 방법

(A Multi-thresholding Approach Improved with Otsu's Method)

이 철 학*, 김 상 운**

(Zhe-Xue Li and Sang-Woon Kim)

요 약

스레쉬홀딩(thresholding)은 영상 화소의 군집이나 강도를 이용하여 영상을 분할하는 기본 기술이다. Otsu의 스레쉬홀딩 방법에서는 정규화 된 히스토그램을 이산 밀도함수로 보아 화소의 클래스 간 분산을 최대화시키는 판별식을 이용한다. 그러나 Otsu의 방법에서는 여러 객체로 이루어진 영상에서 최적의 스레쉬홀드를 찾기 위하여 그레이레벨 전 구간에 대해 모든 가능한 분산 값을 반복적으로 계산해 보아야 하기 때문에 계산 시간이 길게 걸리는 문제가 있다. 본 논문에서는 Otsu의 방법을 개선하여 간단하지만 고속으로 멀티-레벨의 스레쉬홀드 값을 구할 수 있는 방법을 제안한다. 전체 그레이 구간 영역에 대하여 Otsu의 방법을 적용시키기 보다는 먼저 그레이 영역을 작은 부분-구간으로 나눈 다음 Otsu의 방법을 적용시키는 처리를 반복하여 원하는 개수의 스레쉬홀드를 구하는 방법이다. 본 제안 방법에서는 맨 처음 대상 영상의 그레이 구간을 2부류로 나눈다. 이 때, 분할을 위한 스레쉬홀드는 전 구간을 대상으로 Otsu의 방법을 적용하여 구한다. 그 다음에는 전체 구간이 아닌 분할된 부분-구간을 대상으로 Otsu의 방법을 적용하여 두 부류를 4부류로 나눈다. 이와 같은 처리를 원하는 개수의 스레쉬홀드를 얻을 때 까지 반복한다. 세 종류 벤치마크 영상과 50개 얼굴영상에 대해 실험한 결과, 제안 방법은 대상 영상을 특성에 맞게 고속으로 잘 분할하였으며, 패턴 매칭이나 얼굴인식에 이용될 수 있는 가능성을 확인하였다.

Abstract

Thresholding is a fundamental approach to segmentation that utilizes a significant degree of pixel popularity or intensity. Otsu's thresholding employed the normalized histogram as a discrete probability density function. Also it utilized a criterion that maximizes the between-class variance of pixel intensity to choose a threshold value for segmentation. However, the Otsu's method has a disadvantage of repeatedly searching optimal thresholds for the entire range. In this paper, a simple but fast multi-level thresholding approach is proposed by means of extending the Otsu's method. Rather than invoke the Otsu's method for the entire gray range, we advocate that the gray-level range of an image be first divided into smaller sub-ranges, and that the multi-level thresholds be achieved by iteratively invoking this dividing process. Initially, in the proposed method, the gray range of the object image is divided into 2 classes with a threshold value. Here, the threshold value for segmentation is selected by invoking the Otsu's method for the entire range. Following this, the two classes are divided into 4 classes again by applying the Otsu's method to each of the divided sub-ranges. This process is repeatedly performed until the required number of thresholds is obtained. Our experimental results for three benchmark images and fifty faces show a possibility that the proposed method could be used efficiently for pattern matching and face recognition.

Keywords: 영상 분할, 스레쉬홀딩, Otsu의 스레쉬홀딩법, 패턴 매칭, 얼굴 인식

I. 서 론

패턴인식이나 문서분석 등을 위한 영상처리에서 영

상분할은 기본 단계이다. 지금까지 영상분할을 수행하기 위한 스레쉬홀딩 기술에 대해 많은 연구가 진행되어 왔다. 최근 Sankur 등^[1]은 영상을 스레쉬홀딩할 때 이용하는 정보에 따라 스레쉬홀딩 방법을 다음과 같이 여섯 가지로 분류하였다. 즉, 히스토그램을 이용하는 방법^{[2],[3]}, 측정공간의 군집화에 의한 방법^{[4]-[6]}, 엔트로피를 이용하는 방법^{[7],[8]}, 객체의 속성을 이용하는 방법^[9], 공간 상관에 의한 방법^[10]과 국부적 그레이 표면을 이용하는 방법^{[11],[12]} 등 여섯 가지이다.

* 학생회원, ** 정회원, 명지대학교 컴퓨터공학과
(Dept of Computer Engineering, Myongji University)

※ 이 논문은 2005년 정부(교육인적자원부)의 재원으로
한국학술진흥재단의 지원을 받아 수행된 연구임
(MOEHRD-KRF-2005-042-D00265).

접수일자: 2006년4월25일, 수정완료일: 2006년8월11일

먼저, Chang 등^[2]은 히스토그램의 모양을 이용하여 스톱레쉬홀드를 결정하였고, Tsai 등^[3]은 가우시안 커널 스무싱으로 히스토그램의 정상과 골짜기를 찾아 스톱레쉬홀드를 정하였다. 또한, 군집화에 의한 방법에서는 화소의 그레이 레벨이 몇 개의 클래스로 이루어 졌다고 생각하고, 어떤 판단 조건에 의해 화소를 클러스터링하는 방법이다^{[4]-[6]}. 특히, Otsu^[5]는 클래스사이의 분산을 최대화시키는 방법으로 최적의 스톱레쉬홀드를 찾는 방법을 제안하였다. 이 때, 어떤 클래스에 화소 수가 집중되어 있을 경우 이 방법은 비교적 좋은 효과를 보이고 있다.

엔트로피에 의한 방법은 배경과 객체의 엔트로피 또는 원 영상과 이진화된 영상사이의 엔트로피를 이용하여 스톱레쉬홀딩하는 방법이다^{[7],[8]}. Wu 등^[7]은 그레이레벨과 지역평균 그레이레벨로 이루어진 2D 히스토그램을 이용하여 2차원 엔트로피를 최대화시키는 방법을 이용하였고, Li 등^[8]은 입력 그레이 영상과 이진 영상사이의 교차 엔트로피를 최소화시키는 방법을 이용하였다.

객체 속성을 이용하는 방법은 객체 모양의 유사성이나 윤곽선 일치, 원 영상과 이진화 영상의 유사성 등을 이용하여 스톱레쉬홀딩하는 방법이다^[9]. 속성에는 에지 매칭 형식, 집중 정도, 그레이 레벨 모멘트, 연결성, 텍스처, 객체의 안전성 등이 있다. 또한 퍼지 값을 이용하거나 누적 확률 분포의 유사도를 이용하여 그레이 영상과 이진 영상의 유사성을 비교하기도 한다.

공간 상관을 이용하는 방법은 확률분포와 화소사이의 상관관계를 이용하는 스톱레쉬홀딩 방법이다^[10]. 이 때, 그레이 값의 분포 정보를 이용할 뿐만 아니라 상관함수, 동시 발생 확률, 픽셀의 지역 선형 의존 모델, 2D 엔트로피 등을 활용한다. 국부적 그레이 표면을 이용하는 방법은 국부적인 영상 특성에 따라 화소들의 스톱레쉬홀드 값을 적용하는 방법^{[11][12]}으로, 각 화소의 통계정보(분산, 레인지 등)에 근거하여 스톱레쉬홀드를 계산한다.

이상의 스톱레쉬홀딩법은 2-레벨 스톱레쉬홀딩 방법에 대한 연구이었으며, 최근 이를 멀티-레벨로 확장하려는 연구가 활발히 진행되고 있다^{[2]-[4],[7]}. 특히, Otsu의 스톱레쉬홀딩 방법^[5]은 패턴인식이나 문서분석 등과 같은 영상처리에서 우수한 성능의 방법으로 알려져 있다^[1]. 그러나 Otsu의 방법은 클래스사이 분산을 최대화 하는 값을 찾기 위하여 전 구간에서 반복적으로 검색해야 하기 때문에 영상의 객체 클래스 수가 클 경우 계산시간이 크게 증가한다. Liao 등^[4]은 알고리즘의 계산시간 복잡도를 줄이기 위하여 수정된 (클래스사이 분산) 계산식을 제안하였다. 저자들은 기준 값을 찾는 과정에서 분산 계산식에 대

한 계산이 반복된다는 점에 착안하여, 각 구간에 대한 분산 값을 계산한 다음 이를 테이블에 저장해 놓았다가 다시 이용하는 방법을 사용하였다. 그러나 Liao 등의 방법도 스톱레쉬홀드 개수가 늘어나면서 순환이 증가하기 때문에 실제 처리에서 큰 영향을 주는 시간 복잡도($O(L^M)$, 여기서 L 은 그레이레벨 수이고 M 은 스톱레쉬홀드 수이다.) 문제는 여전히 해결하지 못하였다.

또한, Wu 등^[13]은 Otsu의 2-레벨 분할방법의 확장으로 판별계수를 정의하고, 자동으로 스톱레쉬홀드 수를 정하는 방법을 연구하였다. 이 방법에서는 Otsu의 방법을 이용하여 먼저 원 영상의 히스토그램을 스톱레쉬홀딩하여 두 클래스로 나누고, 정의된 클래스 사이의 분산에 대한 판별계수가 정해진 값보다 크게 될 때까지 스톱레쉬홀드를 하나씩 증가시키는 방법을 이용하였다.

본 논문에서도 Otsu의 2-레벨 분할방법을 이용한다. 먼저 대상 영상에서 히스토그램을 작성한 다음, 이를 이용하여 전체 그레이레벨 구간을 2개로 나눈다. 분할된 2구간을 다시 4 구간으로 나누는 과정을 반복하고, 필요한 분할개수에 따라 모자라는 부분은 그 다음 분할에서 분산비에 근거하여 선택적으로 취하는 방법을 이용한다. 이 방법은 비교적 양호한 분할 효과를 유지하면서 탐색 영역을 제한시켜 처리시간을 축소할 수 있다.

본 논문의 구성은 다음과 같다. 제 II장에서는 본 연구의 기초가 되는 Otsu의 방법, Liao의 방법, Wu의 방법 등을 간단히 소개한다. 제 III장에서는 제안 방법을 설명하고, 앞의 세 방법과 계산 복잡도를 이론적으로 비교한다. 제 IV장에서는 기존 방법과 제안 방법의 실험 결과를 고찰하고, 제 V장에서는 결론을 맺는다.

II. 기존의 방법들

1. Otsu의 방법^[5]

Otsu의 방법은 주어진 영상의 히스토그램을 작성하고, 이를 이용하여 유사 밝기 값을 갖는 객체들을 분리 추출하는 방법이다. 즉 Otsu의 방법은 영상 밝기에 대하여, 분할되는 두 클래스의 분리를 클래스사이의 분산을 최대화시키는 스톱레쉬홀드를 찾는 방법이다.

주어진 영상이 L 개 그레이레벨 $[1, 2, \dots, L]$ 의 화소로 되어 있다고 하자. 밝기 값이 i 인 화소의 수를 f_i , 영상의 총 화소 수를 N 이라고 할 때, $N = \sum_{i=1}^L f_i$ 이고, 화소가 i 밝기 값을 갖을 확률은

$$p_i = f_i/N \tag{1}$$

이 된다. 그리고

$$\sum_{i=1}^L p_i = 1, p_i \geq 0 \tag{2}$$

이다. 입력 영상이 배경과 객체 두 개로 나누어 졌다면, 화소들은 그레이레벨 $[1, \dots, t]$ 에 속하는 클래스 C_1 과 그레이레벨 $[t+1, \dots, L]$ 에 속하는 C_2 클래스로 나눌 수 있다. 따라서 두 클래스에 대한 그레이레벨 확률분포는 C_1 : $p_1/\omega_1(t), p_2/\omega_1(t) \dots, p_t/\omega_1(t)$ 와 C_2 : $p_{t+1}/\omega_2(t), p_{t+2}/\omega_2(t), \dots, p_L/\omega_2(t)$ 이 된다. 여기서

$$\omega_1(t) = \sum_{i=1}^t p_i \tag{3}$$

$$\omega_2(t) = \sum_{i=t+1}^L p_i \tag{4}$$

이고, 클래스 C_1 과 C_2 의 평균은 각각

$$\mu_1(t) = \sum_{i=1}^t ip_i/\omega_1(t), \tag{5}$$

$$\mu_2(t) = \sum_{i=t+1}^L ip_i/\omega_2(t) \tag{6}$$

이다. 또한, 전체 영상의 평균을 μ_T 라고 하면

$$\omega_1\mu_1 + \omega_2\mu_2 = \mu_T, \tag{7}$$

$$\omega_1 + \omega_2 = 1 \tag{8}$$

이다. 클래스 C_1 과 C_2 에 대하여 클래스내의 분산 (within-class variance)을 v_w^2 , 클래스사이의 분산 (between-class variance)을 v_B^2 , 그리고 전체 구간에서의 분산(total variance)을 v_T^2 라고 하면, 최적의 스레쉬홀드는 다음의 판별기준식들 가운데서 하나를 최대화하는 t 로 선택할 수 있다.

$$\lambda = v_B^2/v_w^2, \tag{9}$$

$$\chi = v_T^2/v_w^2, \tag{10}$$

$$\eta = v_B^2/v_T^2 \tag{11}$$

여기서

$$v_w^2 = \omega_1\sigma_1^2 + \omega_2\sigma_2^2, \tag{12}$$

$$v_B^2 = \omega_1(\mu_1 - \mu_T)^2 + \omega_2(\mu_2 - \mu_T)^2, \tag{13}$$

$$v_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p(i) \tag{14}$$

이다. 여기서 σ_1^2 과 σ_2^2 은 클래스 C_1 과 C_2 의 분산이다. 판별기준 식 (9)-(11)에서 v_T^2 는 t 의 선택과 무관하므로 η 의 계산이 가장 간단하다. 따라서

$$v_B^2(t^*) = \arg \max_{1 \leq t < L} v_B^2(t) \tag{15}$$

를 만족하는 t^* 를 최적의 스레쉬홀드로 결정할 수 있다. 위의 2-레벨 스레쉬홀딩법은 쉽게 멀티레벨 스레쉬홀딩법으로 확장할 수 있다. 즉 $M-1$ 개의 스레쉬홀드 $T = \{t_1, t_2, \dots, t_{M-1}\}$ 에 의해 원 영상은 M 개의 클래스: $C_1[1, \dots, t_1], C_2[t_1 + 1, \dots, t_2], \dots, C_i[t_{i-1} + 1, \dots, t_i], \dots, C_M[t_{M-1} + 1, \dots, t_L]$ 로 나누고, 다음의 v_B^2 로 최적의 스레쉬홀드 $T^* = \{t_1^*, t_2^*, \dots, t_{M-1}^*\}$ 를 얻을 수 있다.

$$\{t_1^*, \dots, t_{M-1}^*\} = \arg \max_{1 \leq t_1 < \dots < t_{M-1} < L} \{v_B^2(t_1, \dots, t_{M-1})\} \tag{16}$$

여기서

$$v_B^2 = \sum_{k=1}^M \omega_k(\mu_k - \mu_T)^2 \tag{17}$$

이고

$$\omega_k = \sum_{i \in C_k} p_i, \tag{18}$$

$$\mu_k = \sum_{i \in C_k} ip_i/\omega_k \tag{19}$$

이다.

2. Liao의 방법^[4]

Otsu의 방법에서는 최적의 스레쉬홀드 $T^* = \{t_1^*, t_2^*, \dots, t_{M-1}^*\}$ 를 구하기 위해 그레이레벨 값 $[1, 2, \dots, L]$ 범위 내에서 모든 가능한 스레쉬홀드 $T = \{t_1, t_2, \dots, t_{M-1}\}$ 을 취하고, 각각의 경우에 대하여 식 (17)의 클래스사이의 분산을 계산해야 한다. 이

과정에서 바로 식 (17)를 이용할 경우 같은 구간에 대한 클래스사이의 분산을 반복적으로 계산하게 된다. Liao 등^[4]은 이런 반복계산을 고려하여 식 (17)을 간단한 식으로 수정하였으며, 또한 각 구간에서의 값을 테이블에 저장해 놓았다가 다시 참조하는 방법을 이용하였다.

먼저 영상의 그레이레벨을 어떻게 M 개 클래스로 나누든지 그 클래스들의 확률의 합은 항상 1이므로

$$\sum_{k=1}^M \omega_k = 1 \quad (20)$$

이고,

$$\mu_T = \sum_{k=1}^M \omega_k \mu_k \quad (21)$$

이다.

식 (20)과 (21)를 이용하면 클래스사이의 분산 식 (17)은 다음과 같이 쓸 수 있다.

$$v_B^2(t_1, t_2, \dots, t_{M-1}) = \sum_{k=1}^M \omega_k \mu_k^2 - \mu_T^2 \quad (22)$$

또한, 식 (22)에서 등호 오른쪽의 두 번째 항은 스톱밴드 $\{t_1, t_2, \dots, t_{M-1}\}$ 의 선택과 무관하므로 최적의 스톱밴드 $\{t_1^*, t_2^*, \dots, t_{M-1}^*\}$ 선택은 식 (22)의 등호 오른쪽 첫 번째 항(이를 수정된 클래스사이의 분산이라고 한다.) $(v_B')^2$ 이 최대값을 취하는 최적의 스톱밴드 $\{t_1^*, t_2^*, \dots, t_{M-1}^*\}$ 를 선택하는 것과 같다. 즉,

$$\{t_1^*, \dots, t_{M-1}^*\} = \underset{1 \leq t_1 < \dots < t_{M-1} < L}{\operatorname{argmax}} \{(v_B')^2(t_1, \dots, t_{M-1})\} \quad (23)$$

여기서

$$(v_B')^2 = \sum_{k=1}^M \omega_k \mu_k^2 \quad (24)$$

이다. 식 (24)를 이용하면 최적의 스톱밴드 계산 시간을 반으로 줄일 수 있다. 그런데 최적의 스톱밴드를 구하는 판별기준인 식 (16)과 (23)으로 부터 최대의 v_B^2 또는 $(v_B')^2$ 를 구하기 위해서 구간 $[1, L]$ 에서 모든 가능한 $T = \{t_1, t_2, \dots, t_{M-1}\}$ 에 대하여 하나하나 계산할 경우 $C_M^L = L!/M!(L-M)!$ 번 계산이 필요하게 된다.

따라서 식 (22)보다 식 (24)를 이용할 경우 많은 계산 양을 줄일 수 있다.

또한 Liao 등^[4]은 분할 가능 구간에서 반복되는 계산을 줄이기 위해 식 (24)에서 $\omega_k \mu_k^2$ ($k=1, 2, \dots, M$)의 계산을 구간 별로 미리 계산하여 저장하였다가 필요한 값을 검색하여 이용하였다. (구체적인 저장 방법은 문헌 [4]를 참조할 수 있다.)

3. Wu의 방법^[13]

Wu 등^[13]의 연구에서는 Otsu의 방법^[5]을 단순하게 확장하였다. 먼저 분할된 클래스로부터 판별계수를 구하고 판별계수가 미리 정한 값 보다 작을 때, 나누어진 구간들에서 분산 값 $\omega_k \sigma_k^2$ ($k=1, \dots, M$)이 큰 구간에 대하여, Otsu의 방법으로 그 구간을 다시 두 구간으로 나눈다. 초기 상태에서는 원 영상의 1부터 L 까지의 전체 그레이레벨 구간에 대하여 시작하여 하나씩 스톱밴드 값을 증가하면서 내부 분산 값이 판별계수 보다 크게 될 때 까지 진행한다.

식 (17)에 상응하는 클래스 내의 분산 v_w^2 , 전체 구간의 분산 v_T^2 및 평균 μ_T 는 각각 다음과 같다.

$$v_w^2 = \sum_{k=1}^M \omega_k \sigma_k^2, \quad (25)$$

$$v_T^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 p_i, \quad (26)$$

$$\mu_T = \sum_{i=0}^{L-1} i p_i \quad (27)$$

여기서

$$\sigma_k^2 = \sum_{i \in C_k} p_i (i - \mu_k)^2 / \omega_k \quad (28)$$

이다. 식 (25)와 식 (26) 그리고 식 (17)에 대하여, 즉 총 분산과 클래스사이의 분산 그리고 클래스내의 분산 간에는 다음의 등식이 성립된다.

$$v_T^2 = v_B^2(T) + v_w^2(T) \quad (29)$$

또한, 판별계수(separability factor) SF 를 다음과 같이 정의한다.

$$SF = v_B^2(T) / v_T^2 = 1 - v_w^2(T) / v_T^2 \quad (30)$$

여기서 $T = \{t_1, t_2, \dots, t_{M-1}\}$ 은 정해진 스레쉬홀드 집합이다.

이미 분할된 클래스에서 판별계수 SF 를 구하여 미리 정한 값 보다 크면 종료하고, 아니면 다시 모든 클래스 $C_i (i = 0, 1, \dots, M)$ 대하여 클래스의 분산 σ_k^2 를 구하여 그 중에서 가장 큰 가중 분산 값($\omega_k \sigma_k^2$)을 가지는 C_p 를 선택한다. Otsu의 방법으로 C_p 를 두 클래스로 나누어 얻은 스레쉬홀드를 T 에 추가하여, $M+1$ 개 클래스로 분할하는 M 개의 새로운 스레쉬홀드를 구성한다. 이와 같은 과정을 식 (30)이 미리 정한 값보다 클 때 까지 반복한다.

III. 제안 방법

본 논문에서 제안한 방법도 Otsu방법의 확장으로서, 여기서는 먼저 영상의 전체 그레이레벨 구간 $[1, L] = \{1, 2, \dots, L\}$ 에 대해 Otsu의 방법을 이용하여 2개 클래스로 나누고, 2개 클래스를 다시 4개, 8개, 16개 등으로 나눈다. 그리고 원하는 스레쉬홀드의 개수에 따라 마지막으로 얻은 스레쉬홀드들 중에서 분산비율을 고려하여 모자라는 부분을 선택적으로 취하는 방법을 이용한다.

예를 들어, 히스토그램을 6개 클래스로 나눌 경우(이 때에는 5개의 스레쉬홀드 값이 필요하다.), 먼저 두 번에 걸쳐 전체 그레이레벨을 4개의 클래스로 나누고, 여기서 얻어진 스레쉬홀드 3개를 저장한다. 그리고 다시 각 클래스를 2개로 나누어 8개 클래스를 얻는다. 이 때, 각 클래스사이 분산과 그 클래스 전체 분산의 비율을 계산하고, 비율이 높은 클래스 2개를 선택한다. 그리고 선택된 클래스에서 얻은 스레쉬홀드 값 2개를 이미 얻어진 3개의 스레쉬홀드에 추가하여 원하는 5개의 멀티레벨 스레쉬홀드 값을 구성한다.

위의 설명과 같이 Otsu의 방법에서 $M=2$ 일 때 식 (26)은 다음과 같이 쓸 수 있다.

$$\begin{aligned} v_T^2 &= v_w^2(T) + v_B^2(T) \\ &= [\omega_1 \sigma_1^2 + \omega_2 \sigma_2^2] \\ &\quad + [\omega_1 (\mu_1 - \mu_T)^2 + \omega_2 (\mu_2 - \mu_T)^2] \end{aligned} \quad (31)$$

먼저, 영상이 두 개의 그레이레벨 값만으로 이루어졌다고 생각해 보자. 이 때, 두 그레이레벨사이의 한 값을 스레쉬홀드로 선택했다면 이 스레쉬홀드로 나누어지는 두 클래스의 분산(σ_1^2 과 σ_2^2)은 모두 0이 된다. 그러나 전체 분

산은 0이 아니다. 반대로 두 클래스가 큰 분산 값 σ_1^2 과 σ_2^2 을 가질 경우 전체 분산 값 역시 큰 값을 가질 것으로 보인다. 사실 식 (31)에서 볼 수 있는 바와 같이 영상을 두 개 클래스로 이루어진 것으로 생각하면 전체 분산을 두 클래스의 각각의 특성을 나타내는 클래스 내부 분산 v_w^2 과 그 두 클래스 사이의 관계를 나타내는 클래스사이 분산 v_B^2 으로 표시할 수 있다. 그리고 또 식 (31)로부터 $1 = v_w^2/v_T^2 + v_B^2/v_T^2$ 임을 알 수 있다. 따라서 작은 v_w^2/v_T^2 값 또는 큰 v_B^2/v_T^2 값을 가지면 영상이 두 클래스로 잘 군집화 될 수 있음을 알 수 있다. v_B^2/v_T^2 은 영상의 분할정도(또는 필요성)을 의미하며, 필요에 따라 그 비율을 참조하여 클래스를 두 개로 나눌 수 있다. 여기서 v_B^2/v_T^2 비율을 R 로 표기한다.

이상의 고찰로부터 입력 영상의 멀티레벨 스레쉬홀드 값 찾는 제안 방법은 다음과 같이 정리할 수 있다. 여기서 C_0, C_1, \dots, C_q 는 분할하여 얻은 클래스, m 은 구하려는 스레쉬홀드 수, n 은 스레쉬홀드를 구하기 위한 순환 회수($n = 1, 2, \dots$), T_0 은 매 순환에서 얻은 스레쉬홀드의 집합이고, q_0 은 T_0 의 요소 수이다. 그리고 $T = \{t_1, t_2, \dots, t_q\}$ 는 n 번 순환하여 얻은 전체 스레쉬홀드의 집합이고, $q = 2^n - 1$ 은 T 의 요소 수이다.

Step 1: 입력 영상의 그레이 레벨의 도수(히스토그램)를 작성하고, 모든 화소를 하나의 클래스에 할당한다. 초기 값으로 $q_0 = 0; q = 0; n = 1$ 으로 한다. T_0 과 T 는 공집합이다.

Step 2: 스레쉬홀드 집합 T 를 이용하여 얻은 각 클래스에 대하여 식 (13)과 식 (14)으로 새로운 스레쉬홀드를 계산하여 T_0 라하고 T_0 의 스레쉬홀드의 개수를 q_0 라 한다.

Step 3: $m > q + q_0$ 이면 T_0 를 T 에 추가하여 새로운 T 를 얻는다. $q = q + q_0; n = n + 1$ 로 하고 Step 2로 간다.

$m = q + q_0$ 이면 T_0 를 T 에 추가하여 새로운 T 를 구하여 $q = q + q_0$ 로 하고 Step 4로 간다.

$m < q + q_0$ 이면 T_0 의 각 스레쉬홀드를 얻을 때 비율 R 값이 가장 큰 것으로 $m - q$ 개를 취하여 T 에 추가 한다(이 때, $q = m$ 가 된다). 그리고 Step 4로 간다.

Step 4: 현재의 T 를 최종 스톱스레쉬홀드로 출력하고 종료한다.

1. 기존의 방법과 제안 방법의 비교

영상의 그레이레벨 $[1, 2, \dots, L]$ 이 스톱스레쉬홀드 $T = \{t_1, \dots, t_i, \dots, t_{M-1}\}$ 에 의하여 구간 $[1, t_1], \dots, [t_{i-1} + 1, t_i], \dots, [t_{M-1} + 1, L]$ 으로 나누어지고, i 번째 구간의 그레이레벨이 $t_{i-1} + 1, t_{i-1} + 2, \dots, t_{i-1} + k = t_i$ 이라고 하자. 그러면 클래스 C_i 에서 누적확률 ω_i 를 계산하는데 k 번 덧셈과 평균 μ_i 를 구하는데 k 번의 곱셈과 덧셈 계산, 그리고 총 합산을 계산하는데 1번의 뺄셈과 2번의 곱셈, 따라서 모두 $3k + 3$ 번 계산이 필요하다. 이 계산을 각 클래스 별 진행되고 매개 클래스의 그레이 값 개수 합이 L 개이므로 클래스사이 분산식 (17)를 계산하는 데는, 전체 평균 μ_T 를 계산하는데 $2L$ 번과 M 개 클래스의 분산 값 계산하는데 k 번, 모두 $3L + 3 + 2L + M - 1 = 5L + M + 2$ 번의 계산이 필요하다. 그런데 이런 계산을 모든 가능한 $T = \{t_1, \dots, t_i, \dots, t_{M-1}\}$ 에 대하여 수행해야 하므로 Otsu 방법의 계산 복잡 도는 $(5L + M + 2)L!/M!(L - M)!$ 이다. 그런데 실제로 M 이 L 에 비해 매우 작기 때문에 계산 복잡 도는 $O(L^{M+1})$ 이 된다.

다음의 컴퓨터 실험 결과에서도 알 수 있는 바와 같이 M 이 5이상일 때에는 계산 시간이 현저히 증가한다. 이러한 점에 주목하여 Liao의 방법은 스톱스레쉬홀드의 최적성을 판별하기 위하여 Otsu방법의 식 (17) 대신에 식 (24)를 사용하였다. 따라서 계산 복잡도를 $(5L + M + 2)$ 에서 $(3L + M)$ 으로 줄여, 계산 시간을 원래의 3/5 정도로 단축하였다. 그러나 순환으로 인한 계산 복잡도는 여전히 $O(L^M)$ 으로 남아 있게 된다.

Wu^[13]는 새로운 스톱스레쉬홀드를 구할 때 마다 그 전의 스톱스레쉬홀드를 기준으로 한 구간씩 선택하여 그 클래스에서 얻은 스톱스레쉬홀드를 기존의 것에 하나씩 추가하는 방법을 사용하였다. 따라서 매번의 순환에서 기존의 클래스들에 대해 클래스사이의 분산, 즉 식(17) 또는 식(22)을 구하게 되는데, 그 계산 복잡도는 k 개 레벨로 이루어진 클래스에서 $3k + 3$ 번이고, 전체 구간에 대해서는 $(5L + M + 2)$ 번 계산이 된다. 따라서 M 단계 스톱스레쉬홀드를 구하는데 $(M - 1)(5L + M + 2)$ 번 계산이 필요하므로 결국 $O(L)$ 의 계산이 된다.

한편, 본 논문에서 제안한 방법을 이용할 경우에 Wu

의 방법보다 시간을 더 줄일 수 있다. Wu의 방법에서 M 단계 스톱스레쉬홀드, 즉 $M + 1$ 개 스톱스레쉬홀드를 구하는데 $M - 1$ 번의 순환을 해야 하는데 반하여, 제안 방법에서는 $\log_2[M]$ 번 순환하면 되기 때문이다. 또한, 실험 결과에서도 알 수 있는 바와 같이, 일부 영상에 대해서는 분할 효과가 Wu의 방법보다 더 잘 분할되었으며 실험 시간도 더 적게 걸리었다.

IV. 실험

제안 방법의 분할 효과를 고찰하기 위하여 세 벤치마크 영상과 얼굴영상 데이터베이스를 이용하여 실험하였다. 프로그램은 펜티엄 4 컴퓨터에서 MATLAB 7.1 버전을 이용하여 구현하였다.

1. 수행시간과 영상분할 효과의 비교실험

실험에서 방법들을 비교하기 위하여, Wu의 방법을 구현할 때 판별계수를 이용한 스톱스레쉬홀드 수 자동선택 방법을 사용하지 않고, 요구에 따라 반복수를 고정하여 실험하였다. 그리고 수행시간을 측정할 때, 100번 수행한 시간을 평균하였다.

그림 1은 실험에 이용한 영상 "lena", "house", "auroras"이고, 표 1은 실험에서 얻은 분할 시간(sec)이다. 먼저, 영상 분할에서 4-레벨 분할과 5-레벨 분할사이의 실행시간을 비교하기 위해 스톱스레쉬홀드 개수를 4개로 정하여 실험하였다. 실험 결과 Otsu의 방법에서 5-레벨 분할 방법은 4-레벨 분할 방법보다 놀라울 정도로 많은 실행시간이 걸리었다. 또 Liao의 방법은 Otsu의 방법보다 수행시간은 짧았다. 그러나 본 논문의 제안 방법(표에서 "Ours"로 표기한 방법)은 이들 방법 중 가장 적은 실행시간이 걸리었다.

또한, 멀티-레벨 스톱스레쉬홀드를 이용한 영상 분할 효과를 비교하기 위하여 본 논문에서는 세 실험 영상에 대해 4-레벨 3-스레쉬홀드를 실험하였다. 즉, 세 실험 영상에

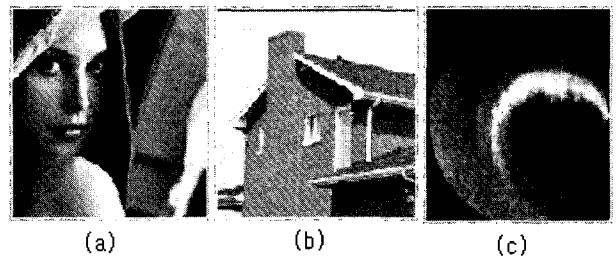


그림 1. 실험 영상. (a) "lena", (b) "house", (c) "auroras"
Fig. 1. Experimental images. (a) "lena", (b) "house", (c) "auroras".

대하여 기존의 Otsu의 방법, Liao의 방법, Wu의 방법 그리고 본 논문의 제안 방법을 이용하여, 먼저 영상의 세 개의 스레쉬홀드를 구하고, 이 세 스레쉬홀드를 이용하여 입력 영상을 4개의 클래스로 분할하였다. 분할된 영상 객체 각각에 대해 그레이 레벨 1, 86, 171, 256을 할당하여 표시한 후 얻어지는 영상에 대하여 zero-crossing 방법으

표 1. 스레쉬홀딩을 수행한 실험 결과의 비교
Table 1. Comparison of thresholding experimental results.

Images	Threshold levels	Methods	Selected thresholds	Times (sec)
lena	5	Otsu	50, 93, 130, 170	23651
		Liao	50, 93, 130, 170	20.33
		Wu	43, 81, 141, 176	0.034
		Ours	43, 81, 141, 176	0.031
	4	Otsu	52, 98, 146	305.2
		Liao	52, 98, 146	0.734
		Wu	43, 81, 141	0.017
		Ours	43, 81, 141	0.016
house	4	Otsu	83, 113, 158	438.1
		Liao	83, 113, 158	1.203
		Wu	67, 96, 147	0.026
		Ours	96, 147, 205	0.024
auroras	4	Otsu	22, 65, 137	455
		Liao	22, 65, 137	1.203
		Wu	64, 136, 188	0.029
		Ours	22, 64, 136	0.025

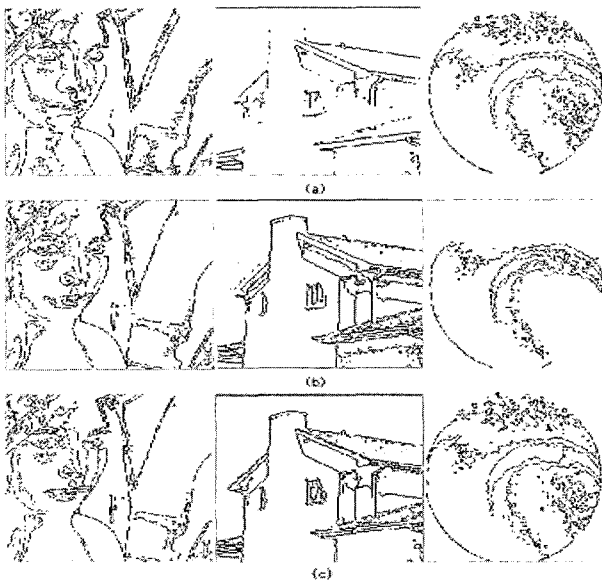


그림 2. 분할된 영상의 에지. (a) Otsu(또는 Liao)의 방법, (b) Wu의 방법, (c) Ours의 방법
Fig. 2. Edge of segmented images. (a) Otsu's(or Liao's) method, (b) Wu's method, (c) Ours method.

로 에지를 구하였다. 그림 2는 그림 1의 영상에서 얻은 에지 영상이다. 여기서 Otsu의 방법과 이 방법을 고속화한 Liao의 방법은 세 영상 모두에서 동일한 스레쉬홀드 값을 얻어, 에지를 구한 결과가 같기 때문에 하나만 표시하였다.

그림 2에서 "lena" 영상으로 얻은 결과는 세 가지 방법이 모두 비슷한 분할 결과를 얻었다. 그러나 "auroras" 영상에 대해서는 Otsu의 방법과 제안 방법이 원 영상의 특성을 잘 표현하였다. Wu의 방법으로 얻은 결과는 윗 부분이 잘 나타나지 않았으며 영상을 잘 표현하지 못하였다. 그리고 "house" 영상에 대해서는 제안 방법은 지붕의 모서리며 집의 구조를 잘 표현하였다. 이러한 결과 고찰로부터, 제안 방법이 비교적 좋은 결과를 얻을 수 있음을 보인다.

2. Hausdorff거리를 이용한 얼굴영상 매칭실험

두 객체 에지 사이의 거리 계산을 기반으로 하는 Hausdorff distance^[14]는 영상처리에서 객체 모양을 비교하는 방법에서 많이 사용하는 한 가지 방법이다. Hausdorff 거리의 주요한 장점은 비교되는 두 집합 점들 사이에 명시적으로 대응시켜 계산하지 않는다는 점이다. 두 점들의 유한 집합을 $A = \{a_1, \dots, a_p\}$ 와

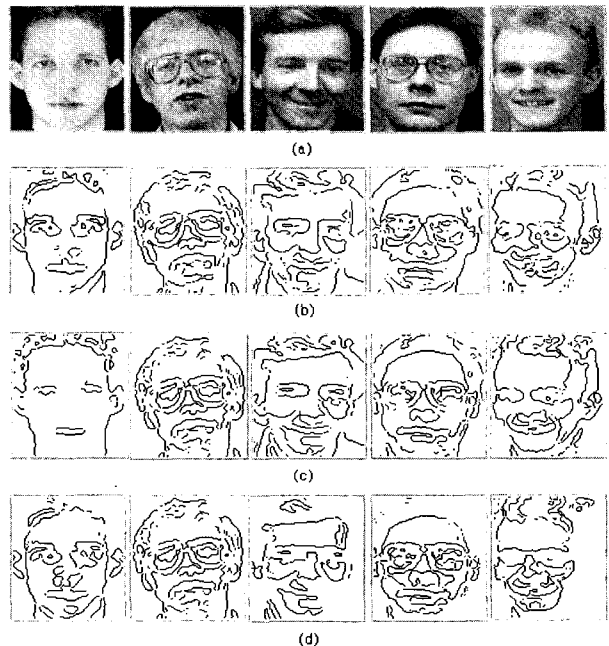


그림 3. 실험 얼굴 영상과 추출한 에지 예. (a) 얼굴 영상, (b) Otsu(또는 Liao)의 에지, (c) Wu의 에지, (d) Ours의 에지
Fig. 3. Examples of experimental faces and extracted edges. (a) faces, (b) Otsu's(or Liao's) edge, (c) Wu's edge, (d) Ours edge.

$B = \{a_1, \dots, a_q\}$ 라고 하면 Hausdorff 거리는 다음과 같이 정의 된다^{[14][16]}.

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (31)$$

여기서 $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$ 이고, $\|\cdot\|$ 는 점집합 A 와 B 에서의 놈(norm)이다. Hausdorff 거리는 두 점 집합에서 서로 매칭 되지 않는 점들을 찾는 방법으로서 최근 영상 모양 비교측정에서 많이 이용되는 방법이다^[14].

본 실험에서는 제안 방법의 멀티 스톱레쉬홀딩 방법의 효율성을 고찰하기 위하여 얼굴 영상 데이터베이스 AT&T^[15]에서 1번부터 5번 사람까지의 50개 얼굴 영상에 대해 매칭 실험을 하였다. 그림 3은 이들 다섯 사람의 얼굴 영상과 각 방법으로 추출한 에지 영상의 예이다.

먼저, 데이터베이스의 50개 얼굴 영상들에 대하여 세 가지 방법으로 영상을 3-레벨로 분할한 다음 zero-crossing 방법으로 에지를 구하였다. 그리고 구한 에지 영상에 대하여 영상들 사이의 Hausdorff 거리를 계산하여 50×50 규모의 비유사도(dissimilarity) 행렬^{[16][17]}을 작성한 다음, 사람단위로 10×10 부분행렬을 평균하여 5×5 비유사도 행렬을 작성하였다. 그런데, 얼굴 영상의 비유사도는 동일한 영상의 경우 0의 값을 갖게 되고, 서로 다를 경우 비유사도에 비례한 큰 값을 갖게 된다. 따라서 비유사도 행렬은 0부터 255사이의 그레이 값을 갖는 영상으로 나타낼 수 있다. 비유사도 행렬 요소 0은 그레이 영상에서 검은 색(black)으로 나타나고, 행렬 요소 255는 밝은 색(white)으로 표시된다.

그림 4는 5×5 비유사도 행렬을 그레이 영상으로 표시한 것이다. 먼저, 그림 3에서 Otsu의 방법과 제안 방법의 영상에서 대각 요소가 고르게 "검은 색"을 띠고 있다. 즉, Otsu의 방법과 제안 방법이 다섯 사람 중 자신의 얼

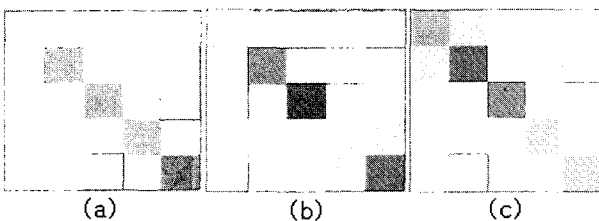


그림 4. 다섯 사람 50개의 얼굴 영상에 대한 Hausdorff 거리 비유사도 행렬. (a) Otsu의 방법, (b) Wu의 방법, (c) 제안 방법으로 얻은 결과

Fig. 4. The dissimilarity matrix using Hausdorff distance for 50 facial images of five individuals. (a) Otsu's method, (b) Wu's method, (c) Ours.

굴 영상에 대하여 서로 다른 얼굴 영상과의 거리보다 모두 고르게 작은 값을 가진다는 것을 보인다. 이는 얼굴 유사성 비교에 있어서 Otsu의 방법과 제안 방법은 좋은 효과를 얻을 수 있음을 보여 준다. 이상의 고찰로부터, 처리 시간과 처리효과로 고려할 때 제안 방법은 실용적으로 이용할 수 있는 비교적 좋은 방법임을 알 수 있다.

V. 결 론

본 논문에서는 Otsu의 방법을 확장하여 멀티-레벨 스톱레쉬홀딩을 빠르게 수행할 수 있는 간단한 방법을 제안하였다. 이 방법에서는 전체 영역을 대상으로 다중 스톱레쉬홀드를 구하는 기존의 방법과는 달리, 간단한 스톱레쉬홀딩법을 구간별 반복하여 적용하는 방법으로 고속화를 시도하였다. 컴퓨터 실험 결과, 가장 짧은 시간으로 입력영상을 효과적으로 분할 처리할 수 있음을 확인하였다. 고속화된 멀티-레벨 스톱레쉬홀딩법을 패턴매칭과 얼굴인식에 포함한 영상인식에 도입하는 것이 앞으로의 과제이다.

참 고 문 헌

- [1] B. Sankur, M. Sezgin, "Survey over image thresholding techniques and quantitative and quantitative performance evaluation", *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146-165, 2004.
- [2] C. -C. Chang, L. -L. Wang, "A fast multilevel thresholding method based on lowpass and highpass filtering", *Pattern Recognition Letters*, vol. 18, pp. 1469-1478, 1997.
- [3] D. -M. Tsai, "A fast thresholding selection procedure for multimodal and unimodal histograms", *Pattern Recognition Letters*, vol. 16, pp. 653-666, 1995.
- [4] P. -S. Liao, T. -S. Chen, P. -C. Chung, "A fast algorithm for multilevel thresholding", *Journal of Information Science and Engineering* vol. 17, pp. 713-727, 2001.
- [5] N. Otsu, "A threshold selection method from gray-level histogram", *IEEE Trans. on System Man Cybernetic*, vol. SMC-9, no. 1, pp. 62-66, 1979.
- [6] J. Gong, L. -Y. Li, W. -N. Chen, "Fast recursive algorithm for two-dimensional thresholding", *Pattern Recognition*, vol. 31, no. 3, pp. 295-300, 1998.

- [7] X. -J. Wu, Y. -J. Zhang, L. -Z. Xia, "A fast recurring two-dimensional entropic thresholding algorithm", *Pattern Recognition*, vol. 32, pp. 2005-2061, 1999.
- [8] C. H. Li, P. K. S. Tam, "An iterative algorithm for minimum cross entropy thresholding", *Pattern Recognition Letters*, vol. 19, pp. 771-776, 1998.
- [9] Y. Shao, M. Celenk, "Higher-order spectra(HOS) invariants for shape recognition", *Pattern Recognition*, vol. 34, pp. 2097-2113, 2001.
- [10] K. -L. Chung, W. -Y. Chen, "Fast adaptive PNN-based thresholding algorithms", *Pattern Recognition*, vol. 36, pp. 2793-2804, 2003.
- [11] P. K. Sahoo, G. Arora, "Image thresholding using two-dimensional Tsallis-Havrda-Charvat entropy", *Pattern Recognition Letters*, vol. 27, pp. 520-528, 2006.
- [12] Q. Huang, W. Gao, W. Cai, "Thresholding technique with adaptive window selection for uneven lighting image", *Pattern Recognition Letters*, vol. 26, pp. 801-808, 2005.
- [13] B. -F. Wu, Y. -L. Chen, C. -C. Chiu, "A discriminant analysis based recursive automatic thresholding approach for image segmentation", *IEICE Trans. Inf. & Syst.*, vol. E88-D, no. 7, pp. 1716-1723, 2005.
- [14] K. -H. Lin, K. -M. Lam, W. -C. Siu, "Spatially eigen-weighted Hausdorff distances for human face recognition", *Pattern Recognition*, vol. 36, pp. 1827-1834, 2003.
- [15] www.uk.research.att.com/facedatabase.html
- [16] E. Pekalska, R. P. W. Duin, *The dissimilarity representation for pattern recognition, foundations and applications*, World Scientific, Singapore, pp. 1-607, 2005.
- [17] 김상운, "Prototype reduction schemes와 Mahala nobis 거리를 이용한 relational discriminant analysis", *대한전자공학회논문지-CI*, vol. 43, no. 1, pp. 9-16, 2006. 1.

 저 자 소 개



이 철 학(학생회원)
 1980년 중국연변대학 학사.
 2002년 한국 명지대학교 대학원
 컴퓨터공학과 공학석사.
 2003년~현재 동 대학원 박사과정
 재학.

<주관심분야: 패턴인식, 미디어처리>



김 상 운(정회원)
 1978년 한국항공대학 공학사.
 1980년 연세대학교 대학원
 공학석사.
 1988년 동 대학원 공학박사.
 1989년~현재 명지대학교 컴퓨터
 공학과 교수.

<주관심분야: 패턴인식, 미디어처리>