

논문 2008-45SP-6-9

# 영역정보기반의 유전자알고리즘을 이용한 텍스트 후보영역 검출

(Detection of Text Candidate Regions using Region Information-based Genetic Algorithm)

오준택\*, 김옥현\*

(Jun-Taek Oh and Wook-Hyun Kim)

## 요약

본 논문은 화소 단위의 정보가 아닌 분할된 영역들의 정보를 기반으로 유전자 알고리즘을 이용한 텍스트 후보영역 검출방안을 제안한다. 먼저, 영상분할을 수행하기 위해 색상별 화소분류와 비동질적인 군집의 감소를 위한 영역 단위의 재분류 알고리즘을 수행한다. 색상별 화소분류에 이용되는 EWFCM(Entropy-based Weighted Fuzzy C-Means) 알고리즘은 공간정보를 추가한 개선된 FCM 알고리즘으로써, 잡음에 강건한 특징을 가진다. EWFCM 알고리즘에 의해 분류된 화소들의 군집정보를 기반으로 수행되는 영역 단위의 재분류는 화소나 군집 단위의 재분류에 비해 효과적으로 영상에 존재하는 비동질적인 군집들을 감소시킬 수 있다. 그리고 텍스트 후보영역 검출은 분할된 영역들로부터 추출한 방향성 에지 성분에 대한 분산값 및 에너지, 크기, 개수 등의 정보를 기반으로 유전자알고리즘에 의해 수행된다. 이는 화소 단위의 정보를 이용한 방법보다 더 명확한 텍스트 영역정보를 획득할 수 있으며, 향후 자동문자인식에서 좀 더 손쉽게 이용될 수 있다. 실험 결과 제안한 분할방법은 기존 방법이나 화소나 군집 기반의 재분류보다 좋은 결과를 보였으며, 텍스트 후보영역 검출에서도 화소 단위의 정보를 이용한 기존 방법보다 더 좋은 결과를 보여 제안방법의 유효성을 확인하였다.

## Abstract

This paper proposes a new text candidate region detection method that uses genetic algorithm based on information of the segmented regions. In image segmentation, a classification of the pixels at each color channel and a reclassification of the region-unit for reducing inhomogeneous clusters are performed. EWFCM(Entropy-based Weighted C-Means) algorithm to classify the pixels at each color channel is an improved FCM algorithm added with spatial information, and therefore it removes the meaningless regions like noise. A region-based reclassification based on a similarity between each segmented region of the most inhomogeneous cluster and the other clusters reduces the inhomogeneous clusters more efficiently than pixel- and cluster-based reclassifications. And detecting text candidate regions is performed by genetic algorithm based on energy and variance of the directional edge components, the number, and a size of the segmented regions. The region information-based detection method can singles out semantic text candidate regions more accurately than pixel-based detection method and the detection results will be more useful in recognizing the text regions hereafter. Experiments showed the results of the segmentation and the detection. And it confirmed that the proposed method was superior to the existing methods.

**Keywords :** Text detection, EWFCM, Genetic algorithm, Image segmentation

\* 정회원, 영남대학교 전자정보공학부  
(School of Electrical Engineering and Computer Science, Yeungnam University)

※ 본 연구는 2008년도 영남대학교 학술연구조성비에 의한 것임.(This research was supported by the Yeungnam University research grants in 2008.)  
접수일자: 2008년1월3일, 수정완료일: 2008년10월10일

## I. 서론

영상에 인위적으로 삽입되거나 자연적으로 포함되어진 텍스트들은 영상의 내용을 함축적이고 구체적으로 표현하는 중요한 수단으로, 현재 텍스트 검출 및 인식

에 대한 많은 연구가 진행되고 있으며, 이는 시각장애인을 위한 보조시스템, 자동주행시스템, 영상 및 동영상의 자동검색시스템, 전자도서관 등의 다양한 분야에 널리 활용할 수 있다. 그러나 자동문자인식시스템에서 이용되는 제한적인 영상과는 달리 복잡한 배경 및 다양한 텍스트 모양 등을 포함하고 있어 상당히 어려운 문제로 인식되고 있다.

영상에서 텍스트를 추출하기 위한 대부분의 기존연구들은 화소의 에지정보를 기반으로 지역적인 텍스처 정보나 텍스트의 밀도, 크기, 너비, 높이 등과 같은 구조적인 정보를 많이 이용한다. 화소의 에지정보를 기반으로 한 텍스트 영역검출방법으로, C. Liu<sup>[1]</sup> 등은 소벨연산자를 이용하여 수평, 수직, 대각성 성분에 따른 에지 정보를 추출한 후 주어진 크기의 윈도우 내 통계적 특징정보를 기반으로 한 K-Mean 알고리즘과 경험적 실험기준에 의해 텍스트 후보영역을 검출하였다. D. Chen<sup>[2]</sup> 등은 작은 윈도우 영역에서 Canny 에지 검출기를 이용하여 추출한 에지정보에 대해서 형태학적 연산과 경험적 정보를 활용하여 텍스트 영역과 배경 영역을 분리한 후 가보 필터를 이용하여 텍스트 영역만을 검출하였다. X. Liu<sup>[3]</sup> 등은 수직 방향에 대한 에지성분을 추출한 후 형태학적 연산을 이용하여 텍스트 영역을 검출하였다. 그리고 유사한 방법으로 X.-S. Hua<sup>[4]</sup> 등은 소벨 연산자를 이용하여 추출한 수평 및 수직 방향에 따른 에지정보와 함께 코너정보를 이용하여 텍스트 후보영역을 검출하였다. 이러한 에지정보를 기반으로 한 방법들은 기본적인 방향성 에지를 추출한 후 에지정보에 따른 군집화와 군집영역의 분석에 의해 텍스트 영역을 검출한다. 그러나 에지의 구조적 정보를 이용하지 않으므로 다양한 텍스트 모양과 크기를 갖는 영역에 대해서 정확한 검출을 수행하지 못하며, 텍스트와 유사한 특징을 가지는 비텍스트 영역도 검출되는 단점을 가진다.

화소의 에지정보를 이용한 텍스트 검출방법과 유사하게 통계적 특성을 이용한 방법으로, H. Li<sup>[5]</sup> 등은 다양한 크기의 텍스트 영역을 검출하기 위해 다해상도의 웨이블릿 변환을 이용하여 텍스트의 획 특징을 추출한 후 신경망을 이용한 학습으로 비디오 프레임에서의 텍스트 패턴을 분류하는 방법을 제안하였다. 그리고 K. I. Kim<sup>[6]</sup> 등은 SVM과 CAMSHIFT(Continuously Adaptive Mean Shift)를 이용한 텍스트 추출방법을 제안하였다. 화소 자체의 밝기값을 SVM에 입력하여 학습함으로써 특징정보를 추출할 필요가 없으나 입력된 특징정보의 차원이 크고 특징정보의 정확성이 떨어진다는 단점을

가진다. K.-Y. Jeong<sup>[7]</sup> 등은 뉴스 비디오 프레임을 대상으로 윈도우 내 화소정보를 입력정보로 하여 신경망 기반의 텍스트 위치 검출 방법을 제안하였으나 텍스트의 크기와 폰트의 크기 및 방향에 제한점을 가진다.

본 논문은 화소 기반의 검출이 아닌 영역 기반의 검출방식으로, EWFCM(Entropy-based Weighted Fuzzy C-Means) 알고리즘<sup>[8]</sup>과 영역단위의 재분류방안에 의해 분할된 영역들의 정보를 기반으로 유전자 알고리즘을 이용하여 텍스트 후보영역들을 검출하는 방법을 제안한다. 먼저, 영상에 존재하는 비동질적인 군집을 최적의 군집수만큼 재분류함으로써 실제적인 응용분야에서도 분할결과를 활용할 수 있다. 그리고 분할된 영역들의 특징정보를 이용함으로써 보다 정확하게 텍스트영역을 검출할 수 있을 뿐만 아니라 텍스트 분할영역의 경계정보를 보다 정확하게 유지할 수 있으므로 차후 문자자동 인식시스템에 보다 용이하게 이용할 수 있다.

## II. 영역정보기반의 유전자알고리즘을 이용한 텍스트 후보영역 검출

본 논문은 그림 1과 같이 영상분할과 텍스트 영역검출로 구성된다. 먼저, 영상분할은 EWFCM 알고리즘을 이용하여 화소들에 대한 분류를 수행한 후 여전히 영상에 존재하는 비동질적인 군집들을 제거하기 위해 최적 군집수에 따른 영역단위의 재분류를 수행함으로써 EWFCM 알고리즘을 이용한 분할결과보다 더 나은 결과를 획득할 수 있다. 그리고 분할된 영역들의 특징정보를 기반으로 유전자알고리즘을 이용한 텍스트 후보영역 검출을 수행한다.

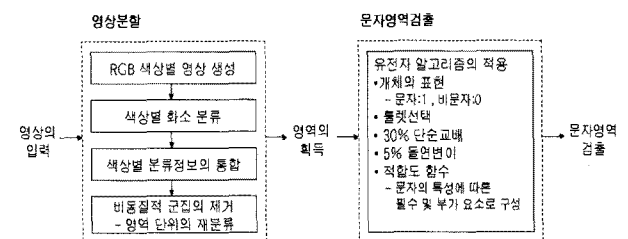


그림 1. 제안한 방법의 처리과정  
Fig. 1. Flowchart of the proposed method.

### 2.1 EWFCM 알고리즘에 의한 화소 분류

먼저, EWFCM 알고리즘을 이용하여 화소들의 색상별 분류정보를 통합한 분할영상을 획득한다.

EWFCM 알고리즘은 화소  $i$ 에 이웃하는 화소들이 군

집  $c$ 에 많이 분류되어 있을 경우 화소  $i$ 와 군집  $c$ 간에 높은 가중치를 부여하는 개선된 FCM 알고리즘으로서, 군집컬러 영상의 색상별 성분에 대한 화소들의 분류를 위해 이용된다. EWFCM 알고리즘의 개선된 소속행렬  $U^*$ 와 군집 중심  $V^*$ 는 각각 식 (1) 및 식 (2)와 같다.

$$u_{ij}^* = w_{ij} \times u_{ij} = w_{ij} \times \frac{1}{\sum_{k=1}^c \left( \frac{\|v_i - x_j\|}{\|v_k - x_j\|} \right)^{\frac{2}{m-1}}} \quad (1)$$

$$v_i^* = \frac{\sum_{j=1}^n (u_{ij}^*)^m x_j}{\sum_{j=1}^n (u_{ij}^*)^m} \quad (2)$$

$w_{ij}$ 는 군집  $i$ 에 대한 화소  $j$ 의 가중치를 나타내며, 화소  $j$ 에 이웃하는 화소들이 군집  $i$ 에 많이 포함되어 있을 경우 화소  $j$ 도 군집  $i$ 로 분류되어질 확률이 높기 때문에  $w_{ij}$ 에 높은 값을 부여함으로써 군집  $i$ 로 분류되어진다.  $w_{ij}$ 는 식 (3)에 의해서 계산된다.

$$w_{ij} = 1 - \frac{e_i}{e_i + e_k} = 1 - \frac{p_i \log(p_i)}{p_i \log(p_i) + p_k \log(p_k)} \quad (3)$$

$p_i$ 와  $p_k$ 는 모든 이웃 화소들에 대해서 현재 화소  $x_j$ 와 함께 군집  $i$ 로 분류 또는 분류되지 않은 이웃화소들의 비율을 각각 의미한다. 이는 현재 화소에 대한 이웃화소들의 분류정보만을 이용함으로써 군집들의 밝기분포에 민감하지 않으며 비교적 빠른 시간 내에 효과적인 성능을 얻을 수 있다.

분류된 각 색상별 화소들을 조합한 후 비동질적 군집에 존재하는 영역들을 유사한 나머지 군집들로 재분류함으로써 비동질적인 군집이나 영역들을 감소할 수 있다. 컬러영상의 색상별 화소들에 대한 분류정보를 기반으로 조합된 군집정보는 식 (4)에 의해 획득된다.

$$c_j = r_j \text{level}^0 + g_j \text{level}^1 + b_j \text{level}^2 \quad (4)$$

$c_j$ 는  $j$ 번째 화소에 대한 조합된 군집번호를 나타내며,  $(r_j, g_j, b_j)$ 는 EWFCM 알고리즘에 의해 분류되어진 각 색상별 군집에 대한 번호를 의미한다. 그리고  $\text{level}$ 은 각 색상별 영상에 적용된 군집수를 나타낸다.  $\text{level}$ 이 증가할수록 분할영역의 수는 급격히 증가하며, 이에 실제적인 응용분야에서 활용하기에는 어려움이 있다.

## 2.2 비동질적인 군집의 제거

영상에 존재하는 비동질적인 군집 및 영역들을 제거하기 위한 방안으로, 본 논문에서는 분류된 군집과 군집 내 분할영역들을 대상으로 분산과 거리값에 의한 재분류 방안을 제안한다.

먼저 EWFCM 알고리즘에 의해 분류된 군집들을 대상으로 식 (5)에 의해 재분류할 군집을 선택하며, 이는 영상에 존재하는 가장 비동질적인 군집으로 정의된다.

$$\max_{k \in \text{all\_cluster}} \left( \frac{\text{var}_k}{1.0 + \text{var}_{\text{all}}} \times \frac{\text{size}_{\text{all}} - \text{size}_k}{\text{size}_{\text{all}}} \right) \times \left( \frac{\text{max\_dis}_k - \text{min\_dis}_k}{\text{max\_dis}_k} \right) \quad (5)$$

$\text{var}_k$ ,  $\text{size}_k$ ,  $\text{var}_{\text{all}}$ ,  $\text{size}_{\text{all}}$ 는  $k$ 번째 군집과 영상에 대해서 분산값과 크기를 각각 나타낸다. 그리고  $\text{max\_dis}_k$ 와  $\text{min\_dis}_k$ 은  $k$ 번째 군집과 나머지 군집들간의 외부 거리로써 최대 및 최소외부거리값이다.

식 (5)에 의해 선택된 군집에 존재하는 영역들은 식 (6)에 의해 가장 유사한 군집으로 재분류된다.

$$\min_{c \in \text{all\_cluster} - k_{\text{index}}} \left( \frac{\text{var}_{r,c}}{\text{var}_r + \text{var}_c} \right) \quad (6)$$

$\text{var}_r$ 과  $\text{var}_c$ 는 분류할 군집( $k_{\text{index}}$ ) 내  $r$ 번째 영역과 나머지 군집들( $\text{all\_cluster} - k_{\text{index}}$ ) 중  $c$ 번째 군집에 대한 분산값이며,  $\text{var}_{r,c}$ 는  $r$ 번째 영역과  $c$ 번째 군집을 결합한 후의 분산값을 나타낸다.

영역단위의 재분류는 사전에 정의된 군집의 수가 될 때까지 반복 수행되며, 초기 군집들이 2개의 군집으로 재분류되는 과정에서 군집들에 대한 평가값이 최소가 되었을 때 최적의 군집수로 결정한다. 최적의 군집수는 식 (7)에 의해 결정된다.

$$\text{opt}_{\text{cluster}} = \min_{\text{all\_cluster}} \left( \frac{\sum_{i=0}^{\text{size}_{\text{all\_cluster}}} wd_i}{\text{size}_{\text{all\_cluster}}} \right) \quad (7)$$

$$wd_i = \frac{\sqrt{\sum_{j=0}^{\text{size}_i} ((r_j - m_{r_i})^2 + (g_j - m_{g_i})^2 + (b_j - m_{b_i})^2)}}{\text{size}_i}$$

$\text{opt}_{\text{cluster}}$ 는 최소평균내부거리값이며,  $\text{size}_{\text{all\_cluster}}$ 는 군집수를 나타낸다.  $wd_i$ 와  $\text{size}_i$ 는  $i$ 번째 군집에 대한 내부거리값과 크기를 각각 나타낸다.  $(m_{r_i}, m_{g_i}, m_{b_i})$ 와  $(r_j, g_j, b_j)$ 는  $i$ 번째 군집의 평균 RGB 색상값과  $i$ 번째 군집 내  $j$ 번째 화소의 RGB 색상값을 나타낸다.

2.3 유전자알고리즘에 의한 텍스트 후보영역 검출

유전자알고리즘<sup>[9]</sup>은 어떠한 문제에 대하여 가능한 해들의 집합을 정해진 형태의 자료구조로 표현하고 이를 반복적으로 변형함으로써 최적의 해를 찾는 방법이다. 본 논문에서는 분할영역들에 대해서 텍스트 후보영역을 분류하기 위한 방안으로 유전자알고리즘을 이용한다.

본 논문에서의 개체는 그림 2와 같이 이진데이터로 표현된다.  $I$ 는 개체 내 위치로써 분할된 영역에 대한 번호이며,  $N$ 은 개체의 길이로써 분할된 영역의 개수를 나타낸다. 그리고  $V$ 는 각 분할된 영역에 대한 분류값으로써, 1이면 텍스트 영역을 나타내며, 0이면 비텍스트 영역을 의미한다. 그리고 룰렛 방법에 의해 개체를 선택하며, 개체군의 30%에 해당하는 개체들의 단순 교배와 개체 내 5%의 유전자에 대해 돌연 변이를 수행한다. 그리고 적합도 함수( $F$ )는 식 (8)와 같이 텍스트의 특성을 고려하여 필수적인 요소( $es$ )와 부가적인 요소( $non\_es$ )로 구성되며,  $F$ 의 값이 클수록 텍스트 영역이 잘 분류되었음을 의미한다.

$$F = es \times non\_es \tag{8}$$

$$es = \frac{energy_{text}}{m\_size_{text}} \times \frac{hisSim_{text}}{hisSim_{non\_text}} \times \frac{e\_size_{text}}{m\_size_{text}} \tag{9}$$

$es$ 는 필수적인 성능평가 요소로써, 텍스트 영역의 방향성 에지성분을 기반으로 한다. 먼저, 그림 3과 같은 3×3 크기의 마스크를 이용하여 분할된 영역의 경계부분에 대해서 각 색상별 성분의 방향성 에지를 추출한 다음 이들의 최대크기값을 이용한다.

식 (9)에서  $text$ 와  $non\_text$ 는 개체에서 유전자가 1과 0인 분할 영역들의 집합으로, 각각 텍스트와 비텍스트 영역들을 나타낸다.  $energy_{text}$ 는 텍스트 영역들의

I	1	2	3	4	5	...	N
V	0	1	1	0	1	...	

그림 2. 개체 표현  
Fig. 2. Representation of the individual.

0 0 0	0 0 1	0 1 0	1 0 0
-1 0 1	0 0 0	0 0 0	0 0 0
0 0 0	-1 0 0	0 -1 0	0 0 -1

그림 3. 방향성 에지 성분 추출을 위한 3×3 크기의 마스크  
Fig. 3. 3×3 masks for extracting directional edge components.

최대 에지성분으로부터 추출한 텍스처 성분 중 하나로서, 식 (10)을 이용한다.

$$\sum_{c=0}^{C_{num}-1} \sum_{i=0}^{E_{num}-1} \max(e_{c,i})^2 \tag{10}$$

$C_{num}$ 은 텍스처 집합에 포함된 분할영역의 개수이며,  $E_{num}$ 은  $c$ 번째 영역 내 경계선 부분에서의 에지개수이다. 그리고  $\max(e_{c,i})$ 는  $c$ 번째 영역 내  $i$ 번째 에지의 최대 성분값을 의미한다. 텍스트 영역은 비텍스트 영역에 비해 더 많은 방향성 에지성분을 가지며, 배경으로부터 시각적으로 뚜렷히 나타나기 위해 비교적 강한 에지성분을 가진다. 그리고  $hisSim_{text}$ 과  $hisSim_{non\_text}$ 은 각각 텍스트와 비텍스트 영역에서 추출한 최대 에지성분에 대해서 주어진 임계치보다 큰 에지성분들로 구성된 히스토그램 내 각 분포들에 대한 평균분산값을 나타낸다. 이는 텍스트 영역에 대한 에지분포가 보다 불규칙적이라는 특성을 이용한 것으로, 텍스트 영역이 비텍스트 영역에 비해 방향성 에지히스토그램의 분산값이 비교적 높음을 이용한다. 마지막으로,  $e\_size_{text}$ 와  $m\_size_{text}$ 는 각각 텍스트 영역들에 대해서 주어진 임계치보다 큰 에지 성분들의 개수와 평균 크기를 나타낸다. 그리고 방향성 히스토그램의 각 분포들에 대한 분산값은 배경영역에 비해서 높은 값을 가진다.

부가적인 요소인  $non\_es$ 는 식 (11)과 같이 정의된다.

$$non\_es = \sqrt{\frac{size_{non\_text}^2}{size_{text}^2 + size_{non\_text}^2}} \times \frac{num_{text}}{num_{non\_text}} \tag{11}$$

$non\_es$ 은 영상에 존재하는 텍스트 영역들의 크기와 개수 등에 의한 부가적인 요소로써, 일반적으로 텍스트 영역은 배경영역에 비해서 개수는 많지만 그에 비해 크기는 작다는 사실에 기반한다. 그리고  $size_{text}$ ,  $size_{non\_text}$ ,  $num_{text}$ , 그리고  $num_{non\_text}$ 는 텍스트 영역과 비텍스트 영역들의 크기와 개수를 각각 나타낸다. 이러한  $non\_es$ 에 의해 크기가 큰 영역들을 제거하거나 작은 영역들의 개수를 증가할 수 있다.

III. 실험

실험을 위해 구현한 시스템은 윈도우 XP 환경 하에 자바 SDK 1.4.1로 구현하였다. 그리고 영상분할에 대한 성능평가함수( $Q$ )로는 M. Borsotti<sup>[10]</sup>의 함수를 이용하였으며, 식 (12)와 같다.

$$Q(I) = \frac{1}{10000(N \times M)} \sqrt{R \sum_{i=1}^R \left[ \frac{e_i^2}{1 + \log(A_i)} + \left( \frac{R(A_i)}{A_i} \right)^2 \right]} \quad (12)$$

$I$ 는 분할된 영상을 의미하며,  $N$ 과  $M$ 은 영상의 너비와 높이를 나타낸다.  $R$ 은 영역의 개수를 의미하며,  $A_i$ 는  $i$  번째 영역의 크기를 나타낸다.  $e_i$ 는 평균 색상 오류로써 입력 영상에서 영역  $i$ 에 대한 평균 밝기값과 화소들의 밝기값간의 총 거리값을 의미한다. 그리고  $R(A_i)$ 는  $A_i$ 와 같은 크기를 가지는 영역의 개수를 의미한다. 성능평가  $Q$ 는 잘 분할된 영상일수록 작은 값을 가진다.

표 1은 색상별 성분에 대해서 EWFCM 알고리즘에 의해 2개의 군집으로 분류하고 조합한 뒤 비동질적인 군집의 수를 감소하기 위한 재분류과정에서 영역, 군집, 그리고 화소 단위를 이용하였을 때의 성능평가와 영역의 개수를 보여준다. 화소나 군집 기반의 재분류는 영역 기반의 재분류와 동일한 과정에 의해 처리된다. 그리고 표 안의 \* 부분은 최소평균내부거리에 의해 결정되어진 군집에 따른 성능평가값을 나타낸다. 이러한 최소평균내부거리는 단지 군집의 평균밝기와 소속된 화소의 밝기값간의 차이값에 의존한다. 그러므로 표 1에서

볼 수 있듯이 최소평균내부거리에 의해 선택된 군집이 M. Borsotti이 제안한 함수에 따른 성능평가에서 최적의 성능평가를 보이지는 않았지만 좀 더 빠른 재분류과정을 수행하기 위해서 본 논문에서는 간단한 처리를 요구하는 최소평균내부거리를 최적의 군집을 결정하기 위한 함수로 이용하였다. 최적의 군집을 선택하기 위해 비동질적인 군집들을 감소하는 과정에서 영역 기반의 재분류가 군집 개수별 영역의 개수가 적음에도 불구하고 가장 좋은 성능평가를 보임을 알 수 있다. 반면에 군집 및 화소 기반의 재분류는 군집의 수를 감소되었음에도 불구하고 영역의 개수가 동일하거나 증가함을 보였다. 이는 비동질적인 군집 또는 군집 내 화소들이 하나의 독립된 영역으로써 나머지 군집들 중 하나에 결합됨을 의미하는 것이며, 이것은 영상분할결과에는 아무런 효과를 주지 않음을 의미한다. 그리고 재분류 과정에서 적정한 군집의 수보다 더 많이 감소를 수행하면 분할성능평가가 커짐을 알 수 있다. 이는 영역의 개수가 적어지는 반면 그 영역 내 오류는 증가하기 때문이다.

그림 4는 제안한 분할방법과 Y. Du 방법<sup>[11~12]</sup>간의 성능평가 및 분할영상을 보여준다. Y. Du는 N. Otsu 방법<sup>[13]</sup>에 의해 색상별 성분들을 2개의 군집으로 분류하

표 1. 영역, 군집, 그리고 화소 기반의 재분류에 따른 성능평가의 비교

Table 1. Comparison of performance evaluations by region-, cluster-, and pixel-based reclassification.

실험 영상, 재분류방법		군집수						
		8	7	6	5	4	3	2
Lena	영역	2212.94 (668)	2210.04 (666)	2180.28 (648)	2118.67 (610)	1954.05 (499)	2578.95* (254)	5100.74 (61)
	군집	2212.94 (668)	2212.94 (668)	2212.94 (668)	2207.95 (665)	2100.11 (600)	2769.56* (300)	5821.96 (77)
	화소	2212.94 (668)	2212.94 (668)	2240.69 (684)	2099.99 (599)	2383.83 (740)	3523.63* (507)	6092.48 (89)
House	영역	313.51 (427)	311.15 (420)	307.95 (411)	306.71 (407)	241.05 (212)	245.31* (198)	2307.59 (42)
	군집	313.51 (427)	313.13 (426)	307.88 (409)	310.82 (407)	255.88 (230)	263.60 (223)	1586.2* (18)
	화소	313.51 (427)	313.88 (428)	310.52 (419)	310.23 (417)	311.11 (419)	267.65* (246)	1791.46 (24)
Peppers	영역	1218.44 (607)	1204.16 (591)	1153.80 (531)	1124.80 (481)	1005.71* (286)	5918.45 (194)	3669.37 (57)
	군집	1218.44 (607)	1218.44 (607)	1201.89 (590)	1164.88 (519)	1456.82 (313)	7143.45 (253)	3464.15* (52)
	화소	1218.44 (607)	1220.44 (607)	1235.15 (618)	1209.18 (557)	1676.14 (443)	5347.67 (254)	6349.21* (83)
Baboon	영역	3024.97 (1234)	3022.63 (1232)	2997.96 (1207)	3755.4 (1022)	3365.51* (749)	6733.96 (502)	31568.29 (216)
	군집	3024.97 (1234)	3024.97 (1234)	3020.02 (1230)	3855.68 (1095)	3479.11 (817)	6470.09* (440)	16550.87 (55)
	화소	3024.97 (1234)	3037.55 (1244)	3194.79 (1329)	3274.64 (1340)	3377.96* (1356)	5435.70 (507)	34327.23 (660)

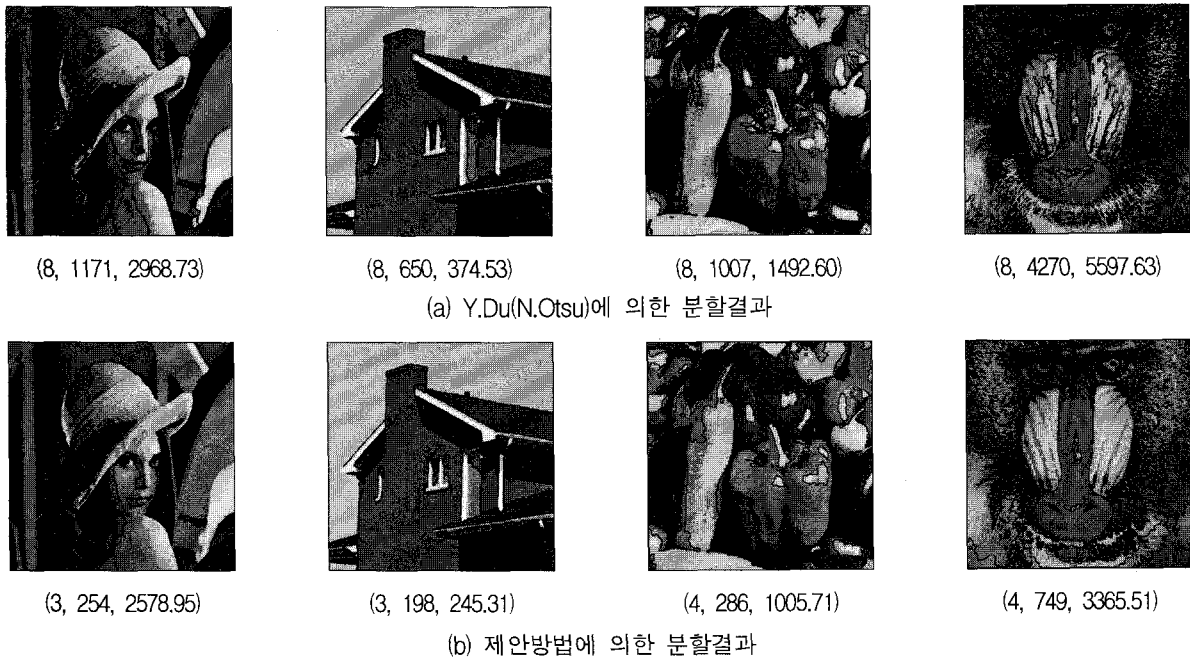


그림 4. 기존방법과 제안방법에 따른 분할 결과(군집의 개수, 영역의 개수, 평가값)

Fig. 4. Segmented results by the existing method and the proposed method(# of clusters, # of regions, evaluation value)  
 (a) Results segmented by Y.Du(N.Otsu) (b) Results segmented by the proposed method.



그림 5. 유전자알고리즘에 의한 텍스트 후보영역 검출

Fig. 5. Text candidate regions detected by genetic algorithm (a) Segmented images (b) Text candidate regions detected by the essential factor (c) Text candidate regions detected by both essential and additional factors.

고 조합한 뒤 군집들의 내부 및 외부거리값에 의해 최적의 군집수를 결정하고 영상을 분할하였다. 그리고 제안한 방법은 최소평균내부거리에 의해 최적의 군집수를 결정하였다. 제안한 방법이 대부분의 실험영상에 대해

서 가장 작은 군집 개수와 영역 개수를 가짐에도 불구하고 더 나은 성능평가를 보였다. 영상을 이루는 최적 군집의 선택을 위해 본 논문에서 이용한 최소평균내부 거리는 Y. Du가 이용한 군집들의 외부 및 내부거리를

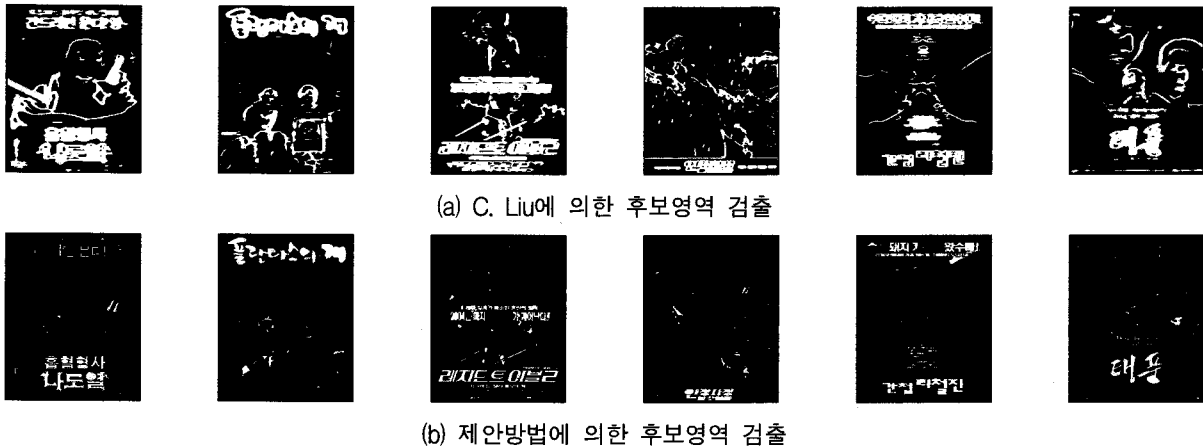


그림 6. 기존방법과 제안방법에 따른 검출결과

Fig. 6. Results detected by the existing method and the proposed method (a) Candidate regions detected by C. Liu (b) Candidate regions detected by the proposed method.

이용하였을 때보다 더 적은 군집의 수에서 최적으로 선택하였다. 이는 영상이 가지는 주요 색상의 수나 영상 분할을 수행한 후 이를 더 적은 수의 색상으로 표현할 수 있음을 의미한다.

텍스트 후보영역 검출에 대한 실험은 다양한 크기와 형식의 텍스트들을 포함하는 한국 영화 포스터 영상들을 이용하였으며, 영상을 분할한 후 영상의 크기에 대해서 5%를 초과하는 영역들이나 영상의 테두리에 연결되어 있는 영역들은 제거하였다. 또한, 시각적으로 인식하기에 어려움이 있을 뿐만 아니라 제안시스템의 목적은 의미있는 텍스트 후보영역들의 검출이기 때문에 영상 크기의 0.001%보다 작은 영역들은 제거하였다. 그림 5는 유전자알고리즘을 이용하여 검출한 텍스트 후보영역들을 보여준다. 그림 5(a)는 제안한 방법에 의해 분할된 영상을 보여주며, 그림 5(b)는 유전자알고리즘의 적합도 함수에서 필수요소만을 이용하여 검출한 텍스트 영역들을 보여준다. 그리고 그림 5(c)은 적합도 함수에 부가적인 요소를 추가하여 검출한 텍스트 영역들을 보여준다. 그림 5(b)에서 볼 수 있듯이 필수요소만으로도 영상에 존재하는 의미있는 큰 텍스트 후보영역들이 잘 검출되어짐을 알 수 있다. 그리고 부가적인 요소의 추가로 인하여 텍스트가 아닌 큰 영역들이 제거되어지면서 작은 텍스트 후보영역들이 더 많이 검출되어짐을 볼 수 있다. 그러나 큰 텍스트 영역도 비텍스트 영역으로 분류되어 제거되어지는 경우도 있었다. 그러므로 부가적인 요소는 영상 내 텍스트들의 특성에 따라 적절히 조절되어질 필요가 있음을 알 수 있었다.

그림 6은 C. Liu의 방법과 제안방법에 따른 검출결과

를 보여준다. 그림 6(b)는 필수적인 요소와 부가적인 요소를 모두 이용하였으며, 영상 크기에 0.001% 이하의 작은 영역들을 제거하지 않고 수행한 결과이다. 그림 6(a)와 같이 화소의 에지정보를 이용할 경우 배경과의 경계가 뚜렷한 비텍스트 영역들도 검출되었으며, 이에 형태학적 연산자나 텍스트의 특성을 이용한 경험적 임계치를 적용하여 후보영역들을 검증하는 후처리과정이 반드시 요구된다. 텍스트 후보영역의 검출 결과만을 비교하였을 때 제안한 방법은 더 나은 결과를 획득하였으며, 그리고 복잡한 처리과정없이 텍스트 후보영역에 대한 검증과정을 수행할 수 있으리라 기대된다. 또한, 텍스트 영역에 대한 명확한 경계정보를 유지하고 있으므로, 자동문자인식시스템에 좀 더 손쉽게 적용 가능하리라 생각된다. 그러나 배경과 유사한 색상의 텍스트 후보영역은 정확히 검출되지 않았으며, 이는 분할알고리즘의 성능에 의해 좌우되어지므로 분할알고리즘의 성능 향상이 요구된다. 또한, C. Liu의 방법과 비교하였을 때 작은 텍스트 후보영역들이 정확히 검출되지 않은 부분도 존재하였다. 본 논문에서는 크기가 큰 의미있는 텍스트 후보영역을 고려하였으며, 향후에는 작은 텍스트 후보영역들도 검출할 수 있는 연구가 기대된다.

#### IV. 결 론

본 논문은 영역정보 기반의 유전자알고리즘을 이용한 텍스트 후보영역 검출방법을 제안한다. 먼저, 영상분할에 이용된 EWFCM 알고리즘은 공간정보를 추가한 개선된 FCM 알고리즘으로써, 잡음에 강건한 특징을 가

진다. 또한, 영역 단위의 재분류는 화소 또는 군집 단위의 재분류보다 영상에 존재하는 비동질적인 군집을 보다 효과적으로 감소함으로써 더 나은 분할결과를 획득할 수 있었다. 그리고 텍스트 후보영역의 검출을 위해 이용된 영역정보 기반의 유전자알고리즘은 화소 단위의 정보를 이용한 방법보다 더 나은 검출결과를 획득하였다. 또한, 텍스트 후보영역에 대해서 보다 명확한 경계 정보를 획득할 수 있음에 따라 향후에 수행될 텍스트 후보영역의 검증 및 인식에도 손쉽게 적용될 수 있으리라 생각한다.

### 참 고 문 헌

- [1] C. Liu, C. Wang, and R. Dai, "Text detection in images based on unsupervised classification of edge-based features," Proc. of International Conf. on Document Analysis and Recognition, vol.2, pp.610-614, 2005.
- [2] D. Chen, K. Shearer, and H. Bourlard, "Text enhancement with asymmetric filter for video OCR," Proc. of International Conf. on Image Analysis and Processing, pp.192-197, 2001.
- [3] X. Liu and J. Samarabandu, "An edge-based text region extraction algorithm for indoor mobile robot navigation," Proc. of IEEE Conf. on Mechatronics and Automation, vol.2, p.701-706, 2005.
- [4] X.-S. Hua, L. Wenyin, H.-J. Zhang, "An automatic performance evaluation protocol for video text detection algorithms," IEEE Trans. on Circuits and Systems for Video Technology, vol.14, no.4, pp.498-507, 2004.
- [5] H. Li, D. Doermann, and O. Kia, "Automatic text detection and tracking in digital video," IEEE Trans. on Image Processing, vol.9, no.1, pp.147-156, 2000.
- [6] K. I. Kim, K. Jung, and J. H. Kim, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.25, no.12, pp.1631-1639, 2003.
- [7] K.-Y. Jeong, K. Jung, E. Y. Kim, and H. J. Kim, "Neural network-based text location for news video indexing," Proc. of International Conf. on Image Processing, vol.3, pp.319-323, 1999.
- [8] J. T. Oh, H. W. Kwak, Y. H. Sohn, and W. H. Kim, "Multi-level thresholding using entropy-based weighted FCM algorithm in color image," LNCS 3804, pp.437-444, 2005.
- [9] U. Maulik and S. Bandyopadhyay, "Genetic Algorithm-based Clustering Technique," Pattern Recognition, vol.33, no.9, pp.1455-1465, 2000.
- [10] M. Borsotti, P. Campadelli and R. Schettini, "Quantitative evaluation of color image segmentation results," Patt. Recogn. Lett. vol.19, no.8, pp.741-747, 1998.
- [11] Y. Du, C. Chang, and P. D. Thouin, "Unsupervised approach to color video thresholding," Opt. Eng. vol.32, no.2, pp.282-289, 2004.
- [12] Y. Du, C. I. Change and P. D. Thouin, "An unsupervised approach to color video thresholding," Proc. of IEEE Conf. on Acoustics, Speech and Signal Processing, vol.3, pp.373-376, 2003.
- [13] N. Otsu, "A threshold selection method from gray level histograms," IEEE Trans. Syst. Man Cybern. vol.9, no.1, pp.62-66, 1979.

### 저 자 소 개



오 준 택(정회원)  
1999년 영남대학교 컴퓨터공학과 졸업(공학사)  
2001년 영남대학교 컴퓨터공학과 졸업(공학석사)  
2006년 영남대학교 컴퓨터공학과 졸업(공학박사)

<주관심분야 : 영상처리, 패턴인식>



김 옥 현(정회원)  
1981년 경북대학교 전자공학과 졸업(공학사)  
1983년 경북대학교 전자공학과 졸업(공학석사)  
1993년 일본 쓰쿠바대학 공학 연구과 졸업(공학박사)

1983년~1993년 한국전자통신연구원 선임연구원  
1994년~현재 영남대학교 전자정보공학부 교수  
<주관심분야 : 시각정보처리, 영상처리>