한국 퍼지 및 지능시스템학회 논문지 1999, Vol. 9, No. 4. (1999)444~450

칼라 문서에서 문자 영역 추출 및 문자 분리

The Character Area Extraction and the Character Segmentation on the Color Document

김 의 정

Eui-Jeong Kim

공주대학교 컴퓨터교육과

요 약

본 논문에서는 칼라로 입력된 문서 영상에서 문자 영역 추출을 위하여 k-means을 이용한 클러스터링 알고리즘을 제안하였다. 칼라 영상의 클러스터링을 위해서 HIS 좌표계에 적합한 거리 함수를 제안하였다. 또한 이를 인식하기 위한 전처리 단계인 문자 분리(segmentation)방법은 연결 화소를 이용한 개별 문자 추출 알고리즘을 제안하였다. 본 알고리즘 에서는 문자 분리 방법에서 접촉 문자(touching character) 또는 겹친 문자(overlapped character)등과 같이 분리가 곤란한 문자를 개별 문자로 분리하는 방법이다. 기존의 문자 분리 방법에서는 투영(projection)에 의한 방법과 외곽선(edge) 추적에 의한 방법 등을 사용하여 왔으나, 제안된 방법은 문자열 추출 후 한번의 투영으로 연결 화소를 이용하여 개별 문자를 추출한다. 문자 영역과 비 문자 영역을 구분하여 개별 문자 추출을 한 결과 단순한 이진 영상이 아닌 칼라 영상에서의 문서 처리가 큰 의의가 있고, 기존의 문서 처리기 보다 향상된 알고리즘인 것을 확인하였다.

ABSTRACT

This paper deals with several methods: the clustering method that uses k-means algorithm to abstract the area of characters on the image document and the distance function that suits for the HIS coordinate system to cluster the image. For the prepossessing step to recognize this, or the method of characters segmentate, the algorithm to abstract a discrete character is also proposed, using the linking picture element. This algorithm provides the feature that separates any character such as the touching or overlapped character. The methods of projecting and tracking the edge have so far been used to segment them. However, with the new method proposed here, the picture element extracts a discrete character with only one-time projection after abstracting the character string. it is possible to pull out it, dividing the area into the character and the rest (non-character). This has great significance in terms of processing color documents, not the simple binary image, and already received verification that it is more advanced than the previous document processing system.

1. 서 론

최근 들어 칼라 스캐너와 칼라 비디오 카메라 등의 칼라 영상 입출력 장치가 대중화되고 있으며, 컴퓨터의 속도 및 기억용량의 증대와 멀티미디어 붐으로 인해 칼라 영상을 사용해 정보를 표시하려는 요구가 증가되었다. 이러한 요구에 부응하여 칼라영상에서 문자를 인식할 수 있는 연구의 필요성이 증대하고 있으나, 지금까지의 문자인식은 이진영상을 대상으로 하여 전처리 및 인식연구가 수행되어 왔으며 칼라영상에 대한 전처리 연구[1]는 이제 시작되고 있는 단계이다.

칼라 영상의 각 화소의 색은 빨강, 초록, 파랑 등의 3개의 기본색을 합한 색으로 표시된다. 이 RGB 값을 가지고 영상처리를 할 경우에는 인간의 눈에는 밝기와 색 등이 RGB값 모두에 관련되어 있기 때문에 RGB 각각을 인간의 눈의 감각과 같이 독립적인 처리가 매우 어려우므로, 본 논문에서는, RGB값으로 표현되는 칼라 영상을 입력받아 인간의 눈의 감각과 같은 HSI(heu, saturation, intensity)값으로 좌표계를 변환시킨 후, 주어진 영상에서 다양한 색상을 잡음에 강하면서도 다른 영역들 즉, 배경이나 이웃한 문자 또는 비문자 영역들과의 명확한 구분을 통한 문자영역의 추출을 위해, k-means알고리즘을 이용하여 한 영역에 나타나는 다양한 색을 단일한 색으로 변환하여 이웃한 영역과의 분리 작업을 수행한다.

또한 칼라 문서에서 추출된 문자 영상을 인식하기 위해서는 문자 분리(segmentation)를 하여야 한다. 기존의 문자를 분리하는 방법으로는 투영에 의한 방법

본 논문은 99년도 공주대학교 과학교육연구소의 학술연구비에 의하여 연구되었음.

[9-13]과 외곽선 추적에 의한 방법[14]이 있다. 투영에 의한 방법의 장점은 두 번의 투영에 의해 문자열과 문자행을 분리해 내는 빠른 속도와 문서 영상의 분리 결과를 문자 인식에 줄 수 있는 정보를 얻을 수 있다는 장점도 있다. 그러나 투영에 의한 방법은 겹친 문자를 분리해 낼 수 없고, 자소 분리를 할 수도 없다. 따라서 분리가 곤란한 경우는 강제 분리[9]를 한다. 외곽선 추적에 의한 방법은 겹친 문자를 분리할 수 있지만 속도면에서 늦고, 인식에 이용할 획득된 정보가 적다. 그러므로 투영에 의한 방법은 이탤릭체나 신문 영상에서처럼 문자들이 서로 겹치는 경우 문자 분리에 문제를 발생시키고, 외곽선 추적에 의한 방법은 정보를 얻기 위해서 다시 투영을 해야만 한다.

본 논문에서는 투영에 의해서 연결 화소를 추적하는 방법이다. 따라서 투영에 의한 방법과 외곽선 추적에 의한 방법의 장점을 이용한 본 방법을 실험한 결과, 효율이 증대됨을 확인하였다. 문자추출을 위한 입력 데이터는 도서의 칼라 영상, 신문, 논문지, 기타 잡지 등을 대상으로 하였다.

논문의 전개는 2장에서 칼라 문서 처리, 3장에서는 개별문자 추출, 4장은 실험 및 결과, 5장 결론 순으로 논문을 기술한다.

2. 칼라 문서 처리

2.1 RGB 칼라 모델

칼라의 삼원색, 빨강, 파랑, 초록 등을 컴포넌트로 하여 직교 좌표계에 표시된다(그림 1). RGB 칼라 모델을 사용하는 영상들은 각 주요 색상에 따라서, 세 개의 독립된 평면을 구성한다. RGB 칼라 모델의 가장 유용한 경우는 항공 위성 다중 스펙트럼 이미지 데이타에 적용할 때이다. 예를 들면, LANDSAT 이미지는 다른 스펙트럼 영역에서 작동하는 센서로부터 영상을 얻게 된다. 즉, 가시영역에서의 데이타, 적외선이나 자외선 센서를 이용한 데이타 입력을 통하여 3차원의 영상 데이타가 생성되며, 각 칼라 평면은 실질적인 의미를 지니게 되고, 각 평면의 조합을 통하여 실제 디스플레이 장치에 보여지게 된다

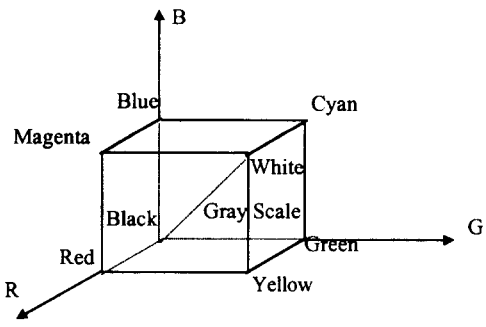


그림 1. RGB 칼라 좌표계

칼라 모델의 목적은 어떤 기준에 의해서 칼라를 명시하기 위함이다. 칼라 모델은 3차원 좌표계로 표현되며, 각 칼라는 좌표계의 한 점으로서 표현된다. RGB를 HSI 모델로 변환[5, 6]하는 방법은 다음과 같이 표현된다.

2.2 RGB를 HSI 모델로 변환

$$I = 1/3(R+G+B)$$

$$S = 1-[3/(R+G+B)]*\min(R,\ G,\ B)$$

$$H = \cos^{-1}\left[\frac{1/2\{(R-B)+(R-G)\}}{\{(R-G)^2+(R-B)(R-G)\}^{1/2}}\right] \quad (1)$$

R, G, B 값은 [0, 1]의 범위를 가지며, I, S도 역시 [0, 1], H는 $[0, 2\pi]$의 범위를 갖게 된다. Hue 값이 0도, 360도의 경우는 빨간색이고, 120도는 녹색, 240도는 청색이 된다. 같은 Hue값을 가지는 영상에서 Saturation값이 감소할수록 백색 성분에 가깝게 된다. 그리고, Intensity값이 감소될수록 검정색 성분에 가까워지게 된다. 예외적으로, Intensity값이 0인 경우에는 Hue와 Saturation값은 정의되지 않으며, Saturation이 0인 경우에도 Hue값은 정의되지 않는다[8, 9]. 또한, Hue 경우에 B/I가 G/I보다 크면 $H = 2\pi - H$가 된다.

2.3 HSI를 RGB 모델로 변환

i) $0^o < H \le 120^o$

$$b = 1/3(1-S)$$

$$r = \frac{1}{3}\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right]$$

$$g = 1-(r+b) \quad (2)$$

ii) $120^o < H \le 240^o$

$$H = H - 120^o$$

$$r = 1/3(1-S)$$

$$g = \frac{1}{3}\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right]$$

$$b = 1-(r+g) \quad (3)$$

iii) $240^o < H \le 360^o$

$$H = H - 240^o$$

$$g = 1/3(1-S)$$

$$b = \frac{1}{3}\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right]$$

$$r = 1-(g+b) \quad (4)$$

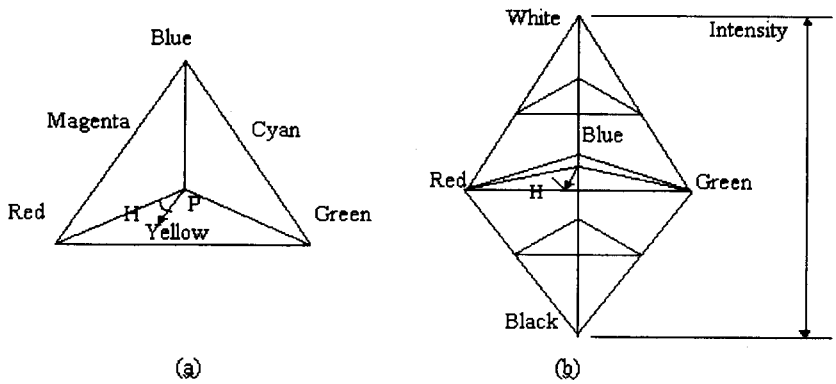


그림 2. HSI 좌표계 (a) HSI color triangle; (b)HSI color solid

각 r, g, b는 [0, 255]의 범위를 갖도록 R=3Ir, G=3Ig, B=3Ib를 계산할 수 있다.

Niblack[10]는 다음과 같은 보다 간단한 HSI와 RGB간의 변환 모델을 이용하였다. 본 논문에서도 이의 변환 모델을 이용하였다.

$$\begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{6} & -2/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix}$$

$$H = \tan^{-1}\begin{bmatrix} V_2 \\ V_1 \end{bmatrix}$$

$$S = [(V_1)^2 + (V_2)^2]^{1/2}$$

$$V_1 = S\cos(H)$$

$$V_2 = S\sin(H)$$

$$\begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} = \begin{bmatrix} 4/3 & -2\sqrt{6}/9 & \sqrt{6}/3 \\ 2/3 & -\sqrt{6}/9 & -\sqrt{6}/3 \\ 1 & \sqrt{6}/3 & 0 \end{bmatrix} \begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix}$$

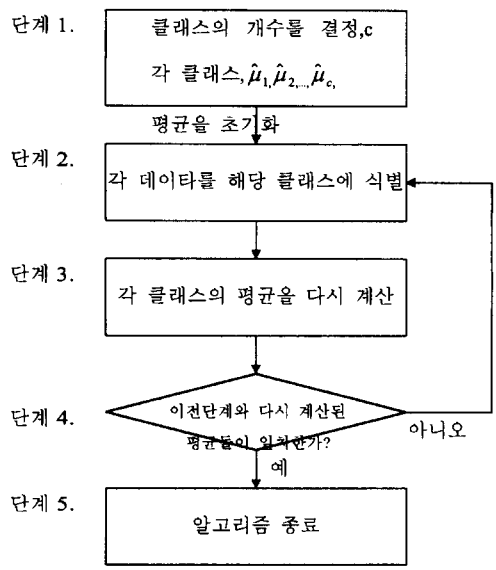


그림 3. k-means알고리즘 단계별 분류

2.4 칼라 문서 클러스터링

본 논문에서는 문자 영역의 색상을 추출하기 위하여 k-means 알고리즘을 이용하였다[7]. k-means 알고리즘을 적용하기 위해, 본 논문에서는 주어진 칼라 영상의 칼라 인덱스만을 이용하여 클러스터링을 수행하였다. 즉, 256개의 칼라 영상의 칼라 인덱스와 임의의 주어진 클래스 개수를 갖는 칼라 클래스 인덱스를 가지고 알고리즘을 적용하였다.

[단계 1]

k-means알고리즘은 분류해야 할 데이타를 가장 유사한 클래스로 분류하는 간단하면서도 매우 적절한 방법이다. 그러나, 분류할 클래스의 개수를 적절히 결정해야 하며, 초기 클래스의 평균값을 설정해 주어야 하는게 큰 단점이다. 즉, 이런 초기치 설정이 분류 결과에 영향을 줄 수 있다. 이런 문제를 해결하기 위한 스케일-스페이스 필터링 방법 등이 제안되고 있으나 스케일 상수의 조정이 필요하므로 궁극적인 해결책이 될 수 없다.

[단계 2]

입력된 데이타를 주어진 클래스로 사상시키기 위해 칼라 좌표계의 거리함수를 이용하여 유사도를 측정하여 분류를 한다.

계산된 c개의 클래스와의 거리 중 가장 작은 값을 갖는 클래스로 주어진 색상을 분류한다.

[단계 3, 4]

모든 데이타의 분류가 끝난 후 분류된 클래스의 평균 좌표값을 구한다. 이전 단계에서 구해진 평균 좌

표값과 비교하여 같거나 일정한 차이 이내에 들어오면 분류를 끝낸다. 그렇지 않으면 분류 조건을 만족할 때까지 위의 작업을 반복한다.

2.5 HSI 좌표 관계성을 이용한 거리함수

칼라 좌표계의 칼라 성분들의 유사도는 대표적으로 유클리디언 거리함수를 사용한다. 그런데, 이런 수치적인 계산 방법은 사람의 눈으로 판단하는 것과는 거리가 있다. 그래서 본 논문에서는 HSI 좌표계의 특성을 고려하여 거리함수를 제안하였다.

2.5.1 RGB좌표계의 거리함수

$$D = \sqrt{(R_i - R_c)^2 + (G_i - G_c)^2 + (B_i - B_c)^2} \quad (6)$$

R_i, G_i, B_i : i번째 칼라 인덱스의 RGB 좌표

R_c, G_c, B_c : c번째 클래스의 평균 칼라 인덱스 RGB 좌표

2.5.2 HSI 좌표계의 거리함수

$$D=\sqrt{w_h*(H_i - H_c)^2 + w_s*(S_i - S_c)^2 + w_i*(I_i - I_c)^2} \quad (7)$$

H_i, G_i, I_i: i번째 칼라 인덱스의 HSI 좌표

H_c, G_c, I_c: c번째 클래스의 평균 칼라 인덱스의 HSI 좌표

w_h, w_s, w_i: 각 밴드에 대한 가중치 상수

2.3.3 본 논문에서 제안한 거리함수

$$D = e^{\Delta I + \max(S_i,\ S_c)/\Delta S} \mid H_i - H_c \mid \quad (8)$$

단, $|H_i - H_c| > 60^o$ 이면, $D = \infty$

$\Delta I = |I_i - I_c|$: 명도의 변화량

$\Delta S = |S_i - S_c|$: 채도의 변화량

H_i, S_i, I_i: i번째 칼라 인덱스의 HSI 좌표

H_c, S_c, I_c: c번째 클래스의 평균 칼라 인덱스의 HSI 좌표

제안된 거리함수는 명암과 채도 그리고 색상 정보를 고려한 함수이다. 명암의 차이가 클 경우, 비록 비교하는 칼라의 채도와 색상이 같을지라도 사람의 눈으로 볼 때는 엄연히 구분된다. 따라서, 큰 차이가 있을 경우 다른 클래스로 분류를 해야 하므로, 명암 변화량을 지수함수에 고려하였다. 또한, 채도는 비교하고자 하는 색상의 값이 1에 가까워지면 좌표계의 특성상 색상이 보다 명확해지므로, 채도보다는 색상에 더 비중을 갖도록 거리를 계산해야 한다. 따라서, 이를 반영하기 위하여 비교하고자 하는 채도들의 최대값을 변화량으로 나누었다. 그리고, 두 색상의 거리가 60도 이상이 될 경우, 실제로는 전혀 다른 색상을 비교한다. 따라서, 수치적 계산을 할 경우 불합리한 거리가 되므로, 이를 보안하기 위해서 거리값을 아주 크게 줌으로써, 분류시 다른 클래스로의 사상을 방지할 수 있다. 이렇게 클래스화한 2진화된 문자영역을 인식의 전처리 부분인 개별문자 추출을 행하게 된다.

3. 개별문자 추출 [8]

문자행 추출 시 줄 단위의 가로 방향으로 투영을 하면서 흑화소의 누적 분포를 판단하여 문자행의 시작점과 끝점을 결정한다. 문서 영상에서는 백화소일 때 0, 흑화소일 때 1을 나타내는 함수 f(x, y)로 표현한다.

$$f(x, y) = \begin{cases} 0 & white\ pixel \\ 1 & black\ pixel \end{cases}$$

3.1 투영에 의한 연결화소 추출

우선 가로 방향 줄단위 문자행 추출을 한다. 추출된 문자행을 가지고 연결화소를 추출하기 위해서는 각 문자열을 세로로 투영하여 추적하는데, 우선 찾아진 점에서 부터 이웃화소들과 연결되었는지를 알아야 한다. 추출된 행에 대해서 세로 방향으로 투영을 한다.

(1) 가로 방향 투영

$$P(y) = \sum_{x=1}^{M} f(x, y) \quad (9)$$

[$P(y)$는 가로 방향으로 투영한 흑화소의 합]

(2) 문자행 추출

i번째 열의 시작점과 끝점은 $P(y) \geq l$일때

$$LStart(i) = y \quad \text{if } P(y-1) < l$$
$$LEnd(i) = y \quad \text{if } P(y+1) < l \quad (10)$$

[l은 잡영으로 추측될 수 있는 임계치]

3.1.1 연결화소 추출법

(1) $f(x, y) = 0$이면, $L(x, y) = 0$

$f(x, y) = 1$이면, 연속 흑화소(Black Run)구간을 구한다.

세로 방향 연속 흑화소 구간은

$P \leq y \leq Q$에서 $f(x, y) = 1$ (11)

(2) 구해진 연결 흑화소 구간(P, Q)의 왼쪽 이웃열의 흑화소의 레이블 값을 참조한다.

$$L(x, y) = \begin{cases} L(x-1, y) \text{if} L(x-1, y) > 0 \\ 0 \quad \text{if } L(x-1, y) = 0 \end{cases} M = \{ L(x, y) / P \leq y \leq Q \} \quad (12)$$

(3) M의 원소가 0 하나이면 구간 (P, Q)에 *NewLable*을 부여한다.

x = 0일 때, *NewLable* = 0이고 $L(x, y) = 0$ (13)

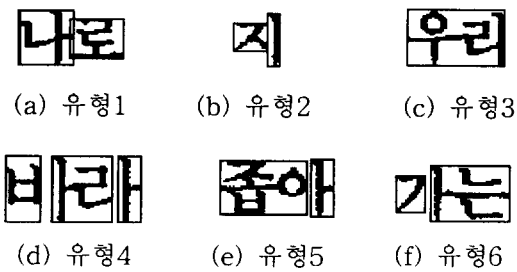
(4) M이 0이외 두개 이상의 원소를 가지면 하나의 원소를 선택하고 나머지는 합성한다.

3.2 연결 화소 블록의 수직 합성

투영에 의한 연결 화소 추적에 의해 추출된 문자 블록은 하나의 문자가 될 수도 있지만, 대부분의 문자는 조합되지 않은 자소들로 분리되어 있다. 따라서 개별 문자 단위로 추출하기 위해서는 이 자소들을 한 문자로 조합해야 한다.

한글 또는 한자의 경우 문자의 형태가 조합 문자로 이루어져 있기 때문에 이러한 문자들은 자소들의 수직적 조합에 의해 단일 문자를 이루는 경우와 수평적 조합에 의해서 단일 문자를 형성하는 경우가 있다. 전자의 경우는 "을", "를", "는" 등과 후자의 경우는 "이", "가", "에" 등을 예로 들 수 있다. 한자의 경우 역시 "示", "言", "𠂤" 등과 "心", "化", "小" 등으로 구분될 수 있다. 연결 화소가 하나의 문자 블록으로 추출되려면 수직적인 결합과 수평적인 조합이 필요하다. 수직 조합을 우선으로 하여 가능한 연결 화소의 문자블록을 추출한 후 여러개의 문자 블록을 하나의 문자로 합성하는 데에는 각 블록들의 겹친 정도를 사용한다. 즉 세로축으로 겹침이 심한 블록들을 하나의 블록으로 합성하게 된다. 이렇게 합성된 블록은 하나의 문자를 이루게 되고 분리된 문자나 겹친 문자들은 완전한 개별 문자로 추출하게 된다. 그림 4(b)에서 연결된 화소의 블록을 세로로 합성하여 그림 4(c)와 같이 수직축으로 합성을 한다.

수평적 결합은 한 문자가 수평적으로 떨어진 두개 이상의 자소로 분리되어 그들이 다른 자소들과 접촉되어 있는 경우가 있으므로 순서상 2.3절에서 자세히 설명하기로 한다.



(a) 유형1 (b) 유형2 (c) 유형3

(d) 유형4 (e) 유형5 (f) 유형6

그림 4. 6가지 접촉 유형

3.3 접촉 문자의 분리

연결 화소 블록의 합성 후 분리되지 않은 접촉 문자들이 남게된다. 접촉 문자를 분리하기 위해서는 분리할 블록을 선택해야 한다.

3.3.1 분리할 블록의 선택

분리할 블록을 선택하는 것은 분리를 하는 방법보다 더욱 중요하다. 어떤 블록이 분리할 블록인가를 알아야 정확한 분리가 가능하다. 만일 분리하지 않아도 될 블록을 분리 한다든지, 접촉 문자를 분리하지 않아도 되는 블록으로 결정하게 되면 아무리 좋은 분리 알고리즘이라도 정확한 분리를 할 수 없기 때문이다.

본 논문에서 제안한 분리 블록을 선택하는 경우는 합성 단계에서 만들어진 블록들을 문서 영상의 한 행씩에 대해 평균 문자 블록 가로폭의 최빈수(mode)를 기준 블록 폭이라고 한다. 그 폭에 일정한 임계치를 주어서 그 이상이 되는 블록을 분리할 블록으로 판단하게 된다. 한글이나 한자의 문서는 한 행에서는 문자의 크기 변화가 심하지 않으므로 기호 문자와 같은 극단적인 값에 영향받지 않는 최빈수로 문자의 대체적인 크기를 알 수 있다. 최빈수 값이 두 개 일때, 즉 쌍등형(bimodal)일 때는 문자행의 세로 높이에 가까운 최빈수를 선택하게 된다.

3.3.2 문자 분리

한글 문서에서의 문자 간 접촉 형태는 6가지의 유형[7]으로 대변할 수 있다. 그림 1에서는 일반적인 문서에서 문자 간의 6가지 접촉 유형을 설명한 것이다. 유형 1은 자소가 겹쳐 있으나 연결화소 추적에 의하여 분리 할 수 있다. 유형 2는 문자 단위로 추출하기 위해서는 자소 블록을 가로로 합성해야 한다. 그 나머지 3, 4, 5, 6 유형을 분리할 대상으로 삼았다.

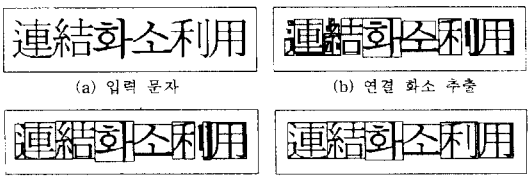
3, 5, 6 유형은 최빈수 문자폭 보다 크기 때문에 분리 대상으로 선정되지만 4 유형은 정교한 규칙이 필요하다. 유형 4는 최빈수 문자폭에 가까운 문자폭을 가진 블록에 대해서 그 특징에 맞는 블록을 선정한다.

유형 3은 대략 양쪽 끝에, 유형 5는 왼쪽 끝에, 그리고 유형 6은 오른쪽 끝에 최빈 블록폭의 분리 위치가 있다. 여러 문자가 붙은 경우 그 크기는 최빈 블록폭의 2배 이상이 된다. 이러한 경우는 유형 3으로 보고 그에 따른 분리를 한다.

각 유형별 특징에 맞는 범위에 대해서 세로로 투영하여 얻은 흑화소수들 중 최소 흑화소수를 문자간 분리 위치로 선정하고,분리된 블록들에 대해서 다시 분리될 블록의 조건에 맞는지를 확인한다.

2.3 분리된 문자 블록의 수평 합성

분리된 문자 블록들은 단일 문자를 구성하는 경우



(a) 입력 문자 (b) 연결 화소 추출

(c) 세로축 블록의 합성 (d) 개별 문자 추출

그림 5. 개별 문자 추출 순서

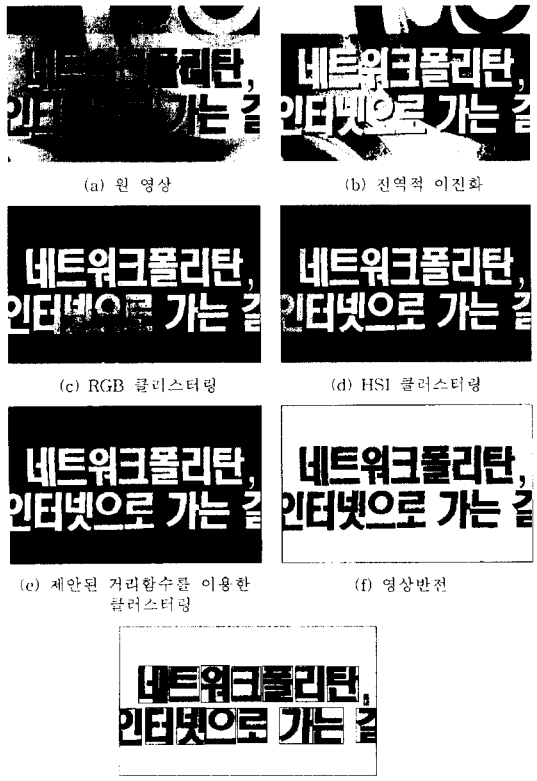
도 있지만 대부분의 문자는 수평적인 결합을 필요로 하게 된다. 3.2절에서 연결 화소 블록의 합성시 제외된 수평적 결합을 하여야 하나의 완전한 문자로 추출된다. 최빈 블록폭을 하나의 문자 폭으로 보고 두 블록의 크기가 최빈 블록폭과 비슷하면 하나의 블록으로 합성해서 개별 문자로 추출한다.

그림 5에서는 문자의 자간 간격 −15의 혼합 문자를 그림 5(b)처럼 연결 화소 추출로 얻은 블록을 그림 5(c)와 같이 수직으로 합성 하였고, 그림 5(d)와 같이 수평 블록의 합성과 함께 개별 문자의 결과를 얻을 수 있다.

4. 실험 및 결과

본 논문에서는 칼라 문서에서 문자 영역 추출을 위하여 k-means 알고리즘을 이용한 거리함수의 성능을 비교, 분석하였으며, 제안된 거리함수를 이용한 클러스터링 결과를 볼 수 있었다. 다음 실험 결과들은 원 영상에 대해 계조영상에서 사용되는 전역적 이진화 방법과 칼라 좌표계에서 클러스터링 방법을 이용한 문자 영역 클러스터링 방법을 비교하였다. 계조영상에서 수행하는 방법보다 칼라 성분을 이용해서 분류하므로, 계조영상과는 달리 많은 정보를 지니게 된다 . 따라서, 계조영상에서는 가질 수 없는 색상 정보를 이용하여 더욱 정확한 칼라 문서의 문자 영역 분류가 가능하다. 그리고, 칼라 좌표계를 이용한 클러스터링 방법에서, RGB 좌표계와 HSI 좌표계의 유클리디언 거리를 이용한 분류의 결과를 보였으며,본 논문에서 제시한 거리함수를 이용한 클러스터링 결과와 비교하여 그림 6(a)~(f)까지의 결과를 얻었다.

이와 같이 클러스터링된 문자 영역을 가지고 연결 화소를 이용한 개별 문자를 추출하였다. 기존에 분리가 곤란했던 접촉문자 및 겹친 문자 등을 정확히 분리하여 문자인식의 성능을 향상 시켰다. 그림 6의 결과들은 원 영상에 대해 계조 영상에서 사용되는 전역적 이진화 방법과 칼라 좌표계에서 클러스터링 방법을 이용한 문자영역 클러스터링 방법을 비교하였고, 추출된 문자 영역으로부터 연결화소를 이용한 개별문자 추출결과를 보였다.



(a) 원 영상 (b) 전역적 이진화

(c) RGB 클러스터링 (d) HSI 클러스터링

(e) 제안된 거리함수를 이용한 클러스터링 (f) 영상반전

(g) 개별문자 추출

그림 6. 문자 영역추출 및 개별문자 추출

5. 결 론

본 논문에서는 칼라 문서 영상을 보다 효과적인 클러스터링 방법인 유사도 거리함수를 제안하였다. 또한, 문자영역에 대해서 개별문자 추출하는 방법으로서는 겹친 문자 및 접촉 문자를 분리 할 수 있는 연결 화소를 이용한 개별문자 추출방법을 사용하여 칼라영상을 인식할 수 있는 전처리 방법을 제안하였다. 이러한 결과로 기존의 이진영상에서 문자인식을 해야하는 제약 사항을 탈피하여 각종 칼라 문서에서 문자영역을 추출하고, 인식 할 수 있도록 하는 연구 결과를 얻을 수 있었다.

앞으로 칼라 좌표계의 보다 면밀한 분석이 요구되며, 특히 칼라 영상 데이타의 에지영역 부분의 다양한 색상 변화에 대한 연구가 필요하다. 또한, 누적분포를 이용한 스케일-스페이스 필터링 작업등을 사용해서, k-means 알고리즘의 초기 클러스터 개수와 클러스터 중심의 위치 선정 등의 자동적인 결정 알고리

즘을 연구한다면, 보다 좋은 클러스터링 결과가 나올 것이다.

개별 문자 추출 알고리즘의 경우, 국, 한문 또는 영문 등 문자가 혼용이 되었을 때 정확히 분리 할 수 있는 알고리즘이 제안되어야 할 것이다.

참고문헌

[1] 이문주, 진영근, 이은주, 김태균, "칼라 영상에서 문자와 비문자 영역 분류에 관한 연구," 영상 처리 및 이해에 관한 워크샵 발표 논문집, pp. 200-203, 1994.

[2] S. Tsujimoto and H. Asada, "Resolving Ambiguity in Segmenting Touching Characters," Proceeding 1st, International Conference on Document Analysis and Recognition, pp. 701-709, 1991.

[3] D. Wang, S. N. Srihara, "Classfication of Newspaper Image Blocks Using Texture Analysis," Computer Vision, Graphics, and Image Processing 47, pp. 327-352, 1989.

[4] 장명욱, 천대녕, 양현승, "연결화소를 이용한 문서 영상의 분할 및 인식," 한국정보과학회논문지, Vol. 20, No. 12, pp. 1741-1751, 1993.

[5] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Addison-Wesley, 1992.

[6] William K. Pratt, Digital Image Processing, John Willy&Sons, 1991.

[7] Robert J. Schalkoff, Pattern Recognition: Statistical, Structural and Neural Approaches, John Wiley&Sons, pp. 114-116, 1992.

[8] 김의정, 김태균 "인쇄체 문서 인식을 위한 문자 추출에 관한 연구," 제2회 문자인식 워크샵발표 논문집, pp. 171-179, 1994.

[9] S. Tsujimoto and H. Asada, "Resolving Ambiguity in Segmenting Touching Characters," Proceeding 1st, International Conference on Document Analysys and Recognition, pp. 701-709, 1991.

[10] S. Liang, M. Ahmadi, M. Shridhar, "Segmentation of Characters in Printed Document Recognition," Proceeding 2nd, International Conference on Document Analysys and Recognition, pp. 569-572, 1992.

[11] Y. Lu, "On the Segmentation of Touching Characters," Proceeding 2nd, International Conference on Document Analysys and Recognition, pp. 440-572, 1992.

[12] D. Wang, S. N. Srihara, "Classfication of Newspaper Image Blocks Using Texture Analysis," Computer Vision, Graphics, and Image Processing 47, pp. 327-352, 1989.

[13] 이인동, 권오석, 김태균, "블록영상의 추출 알고리즘," 한국정보과학회논문지, Vol. 18, No. 2, pp. 218-226, 1991.

[14] 장명욱, 천대녕, 양현승, "연결화소를 이용한 문서 영상의 분할 및 인식," 한국정보과학회논문지, Vol. 20, No.12, pp. 1741-1751, 1993.

김 의 정 (Eui Jeong Kim) 정회원

1993년 : 충남대학교 컴퓨터공학과 (공학석사)

1997년 : 충남대학교 컴퓨터공학과 (공학박사)

1997~1998년 : 시스템공학연구소 영상처리연구부(연구원)

1998년~현재 : 공주대학교 컴퓨터교육과(전임강사)

주관심분야 : 패턴인식, 컴퓨터비전, 멀티미디어