

자연 영상에서의 문자 패턴 추출

김형균 · 최원호
전자공학과

<요 약>

본 논문에서는 자연 영상에서 문자의 일반적인 특징을 이용하여 문자 패턴을 추출하는 방법을 구현하였다.

문자의 크기, 형태, 위치, 농도, 문자 종류, 서식, 그리고 조명의 조건이 일정하지 않은 자연 영상을 대상으로 다음의 단계를 거쳤다. 먼저, 조명의 조건에 크게 영향을 받지 않는 국소적 2치화 방법을 이용하여 배경 부분과 문자 영역의 영역 분할을 행하였으며, 문자와 인접해 있는 배경 부분은 농도 차가 크다는 성질을 이용하여 비문자 영역으로 제거하였다. 또한, 입력 영상에서 문자의 기본적인 성질, 즉 위치, 크기, 폭, 종횡비 등을 만족치 못하는 비문자 영역을 제거하였으며, 문자 내의 구멍의 수에 기초하여 비문자 영역을 제거함으로써 문자 패턴 후보를 추출하였다. 다음으로 문자를 둘러싸고 있는 테두리 부분을 제외한 문자 패턴 후보에 대하여 투영 방법을 적용하여 문자열을 검출하고, 이 문자열에서 각각의 문자 패턴을 추출하였다.

여러 가지 서로 다른 상황의 입력 영상을 대상으로 이 방법을 적용한 결과 대부분의 문자 영역 추출이 가능하였으며, 비문자 영역을 효과적으로 제거할 수 있었다.

A Method of Character Pattern Detection from Natural Scenes

Hyoung Gyun Kim · Won Ho Choi
Dept. of Electronic Engineering

<Abstract>

In this paper, an effective algorithm for the character pattern detection using fundamental features of natural scene images is proposed. The character sizes, forms, positions, intensities, kinds of characters, text styles, and illuminations are not uniform

in natural scene images. An image is segmented into background regions and character regions by the adaptive local thresholding method, which is independent of the illumination. The backgrounds of input images are roughly removed by using the characteristic that the gray levels of adjacent regions around characters and those of backgrounds are very different.

For the finer removal of backgrounds and non-character regions, the basic characteristics of character regions are used, that is, the position, size and the aspect ratios of circumscribing rectangles are almost uniform in character regions, and the number of holes of that regions are small, are applied. After the removal of backgrounds, the projection profile method is used to detect character strings. Each character pattern is then detected from the character strings.

The experiments are carried out for several natural scene images, and character patterns are detected clearly with the proposed algorithm.

I. 서 론

최근의 문자 인식 기술은 눈부시게 발전했으며, 한글, 영어, 숫자 등의 인쇄체, 필기체까지 인식 가능한 OCR이 실용화되고 있다[1]-[7]. 문서 인식 시스템 외에도 컴퓨터의 영상 처리 방법을 이용한 유료도로나 주차장의 요금소를 통과하는 자동차의 번호판의 차량번호를 인식해 기록하는 시스템[8], 그리고 화차의 번호를 자동으로 컴퓨터에 입력하여 전산 처리하는 시스템[9] 등이 연구되고 있다.

그러나, 이러한 시스템들은 모두 대부분 사전 지식이나 몇 가지 제약 조건들을 가지고 있으며, 현재까지 개발된 OCR의 인식 대상은 제한된 형식의 문서를 대상으로 하고 있다. 신문과 문서의 문자 추출을 위한 OCR은 흑화소를 위주로 문자열을 추출함과 동시에 문자열의 피치(pitch)에서 문자 크기를 확정하여 문자를 추출하는 것이 대부분이다[1]-[7]. 그러나 이와 같은 OCR의 기존 기술은 문자 크기와 농도 등의 문자에 관한 사전 지식이 없는 경우와 문자열을 구성하고 있지않은 문자에 대해서는 적용이 곤란하며, 그림이 포함되어 있는 문서의 경우 그림 영역 내에 존재하는 문자의 추출이 불가능하다.

차량번호나 화차번호를 인식하는 시스템의 경우에는 차량번호판이나 화차번호의 위치를 문자와 인접한 주변 환경을 이용하여 파악한 후 문자를 추출하고 있다[9][10]. 즉, 문자 주변의 직선 성분을 우선적으로 추출하여, 직선 성분의 위치 정보를 바탕으로 문자의 위치 정보를 추출하고 있다. 이러한 경우 문자 주변의 특성 정보를 사전에 알고 있어야만 하며, 다른 경우에 대해서는 적용이 곤란하다.

이 외에도, 부품 분류를 위한 각인 문자 인식[11] 등의 시스템이 있으나, 위의 예들과 마찬가지로 문자의 농도, 문자의 위치 등 주목하는 문자의 사전 지식을 이용하고 있다.

위와 같이, 특정 분야에서 문자를 추출하는 경우의 연구는 활발히 진행되어 왔으나, 이와는 달리 자연 영상에서 문자를 추출하는 경우의 연구는 그다지 활발하지 않았으며, A. Shio 등이 자연 영상에서 문자를 추출, 인식하는 방법을 연구하였다[12]. 그러나, A. Shio의 방법은 문자 패턴을 추출함에 있어서, 많은 비문자 영역을 포함하고 있으며, 인식의 단계에서 문자 패턴을 추출하고 있다.

본 논문에서는 여러 가지 조명 조건의 자연 영상에 있어서 크기, 위치, 문자 종류, 문자의 농도, 서식이 자유로운 문자에 대해 사전 지식이 없이 문자 패턴을 추출하며, 인식의

이전 단계에서 비문자 영역을 최대한 제거하고, 문자 패턴을 가급적 정확하게 추출하고자 한다. 단, 장식용 문자와 같은 한 문자 영역 내에 여러 가지 색이 있는 것, 그래픽 문자 등, 복잡한 문자에 대해 본 논문에서는 적용하지 않는다.

본 논문에서의 알고리즘은 문자 부분과 그 배경 부분의 농도 콘트라스트를 이용한 동적 2치화 방법, 문자 부분에 해당하는 영역의 종횡비, 한 문자 내의 구멍의 갯수, 입력 영상에서 한 문자에 해당하는 화소의 최소 수, 문자의 폭, 입력 영상에서의 문자의 위치 등의 문자의 특성에 의한 문자 패턴의 추출을 행함으로써 문자 패턴을 자동적으로 추출할 수 있다.

2장에서는 문자 패턴 추출을 위한 기본적인 고찰을 다루었고, 3장에서는 전체 알고리즘 및 동적 2치화 방법을 이용한 영역 분할, 문자 후보 영역 추출, 문자 패턴 후보 추출, 문자 패턴 추출 방법을 설명하였으며, 4장에서는 본 알고리즘을 적용한 실험 및 결과를 다루었다. 마지막으로 5장에서는 결론의 순으로 구성하였다.

II. 기본적인 고찰

실제 공간상의 문자는 일반적으로 문자의 서식, 크기, 위치, 자형, 문자의 농도와 색이 여러 가지 형태로 존재한다. 또한, 실제 공간상의 문자를 자연 영상으로 입력할 경우 조명 조건의 변동이 불안정 요인으로 더해진다. 3장에서 서술한 2치화 처리 방법은 이러한 불안정 요인에 대해 사전 지식이나 제한을 완화시키고 있다. 본 논문에서 취급하는 문자의 보편적인 성질로서는 문자 영역과 그 배경 부분의 농도 콘트라스트가 높으며, 국소적으로는 2치적이라는 것이다[13]. 또한, 많은 문자들은 문자를 둘러싸는 사각형의 비율이 정사각형에 가깝고, 문자 내의 구멍(hole)은 일정 갯수 이하라는 것이다.

본 논문에서 이용하는 문자의 성질에 대한 자세한 내용은 다음과 같다.

(1) 문자 후보 영역의 추출

입력된 자연 영상에 대한 전 처리로서, 문자 영역은 국소적으로 2치화상을 구성한다는 성질을 바탕으로, 국소적 2치화 방법에 따라 문자 영역과 배경 부분의 영역 분할을 행하며, 문자 영역과 배경 부분의 농도 콘트라스트를 이용하여, 인접 부분의 농도차 평가에 따라 문자 후보 영역을 추출한다.

(2) 문자 패턴 후보의 추출

(1)에서 추출된 영역 내에는 문자 영역이 아닌 영역도 포함되어 있으므로 문자 영역 이외의 영역을 제거하여야 한다. 본 논문에서는 입력 영상에서 문자가 가지는 성질, 즉 영역 분할 후 얻어진 2치화 영상에서 문자의 종횡비는 20:1을 넘지 않는다는 성질[4], 문자에 해당하는 영역의 폭은 2화소 이상이며, 사람이 인식할 수 있는 한 문자의 최소 크기는 8 화소 이상이라는 성질, 그리고 문자 영역은 입력 영상의 테두리 부분과 인접하지 않는다는 성질을 이용하여 테두리 부분과 인접해 있는 문자 영역은 문자 후보 영역에서 제외된다. 또한 서로 접하여 두 문자가 하나의 영역으로 되어있는 경우, 그 영역 내의 구멍의 수는 일정 갯수 이하라는 성질을 이용하여 비문자 영역을 제거함으로써 문자 패턴 후보를 추출한다.

(3) 문자 패턴의 추출

(2)까지의 처리에서 얻어진 영역에 대해 문자 영역을 둘러싸는 사각형을 만들었을 경우 문자에 해당하는 영역은 사각형 내부에 존재하거나, 접해있는 사각형은 일정 갯수 이하라는 성질을 적용하여 문자 패턴을 둘러싸는 테두리 영역을 제거한다. 그리고, 나머지 영역에 대해 투영을 하여 문자열을 추출하고, 이 문자열을 바탕으로 개개의 문자 패턴 영역을 추출하도록 한다.

Ⅲ. 전체적인 알고리즘의 구성

본 문자 패턴 추출 알고리즘은 동적 2치화 방법을 이용한 영역 분할, 문자 후보 영역 추출, 문자 패턴 후보 추출, 문자 패턴 추출로서 크게 4단계로 구성되어 있다. 그림 3.1은 전체 알고리즘의 흐름도를 보여주고 있다.

3.1 동적 2치화 방법을 이용한 영역 분할

3.2 문자와 배경 영역 이진화 분류를 위한 임계치 계산

자연 영상에서 기존의 방법에 의한 문자 후보 영역 검출 방법으로는 크게 나누어 미분과 영역 분할이 있다. 그러나 미분에 의한 방법은 잡영을 강조하거나 문자 후보 영역의 끊김을

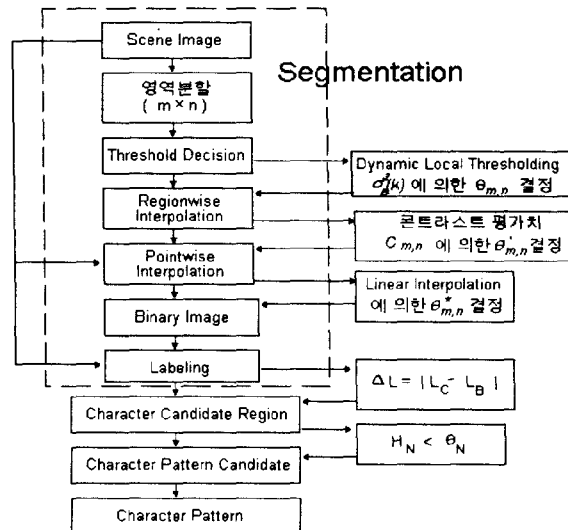


그림 3.1 전체 블록 다이어그램
Fig 3.1 Block diagram of the proposed algorithm

생기게 하는 문제가 있으며, 영역 분할에 의한 방법은 임계치(threshold)를 자동적으로 영상에서 결정하는 것이 곤란하며 주어진 임계치가 결과에 큰 영향을 미치는 문제가 있다. 이들의 문제는 문자 인식의 중대한 오류 요인이 된다.

따라서, 본 논문에서는 자연 영상에 있어서 문자 후보 영역과 그 배경은 국소적으로 2치 화상을 구성한다고 하는 성질에 기초하여, 영상에서 적응적으로 임계치의 결정이 이루어지는 2치화 방법[13][14]을 사용하였다. 이 방법은 최적화 파라메타가 없는 어떤 영상에 대해서도 자동적으로 임계치가 결정되는 특징이 있으며, 2치화한 영상은 다음과 같은 특징을 갖는다.

- 문자 후보 영역에서 확실하게 문자와 배경을 구분할 수 있게 한다.
- 문자 주변의 잡음을 최소화한다.
- 조명이나 문자의 크기 등에 영향을 거의 받지 않는다.

그림 3.2는 국소적 2치화 방법의 원리를 나타낸다. 원 영상($M \times N$ 화소)을 부분 영역($m \times n$, $m \times n = i \times j$ 화소, $m < M$, $n < N$)으로 분할하고, 각 부분 영역(m, n)에 대해 2치화를 위한 임계치를 결정한다. 이와 같이 얻어진 임계치는 전체 영상에 대해서 부분 영역들이 불연속이므로 그림 3.2(a)와 같은 결과가 생긴다. 이러한 불연속을 없애기 위하여 각 부분 영역의 임계치를 부분 영역의 중앙에 위치시켜 선형 보간을 한다. 그러면 그림 3.2(b)와 같이 연속적인 2치화 결과를 얻을 수 있다.

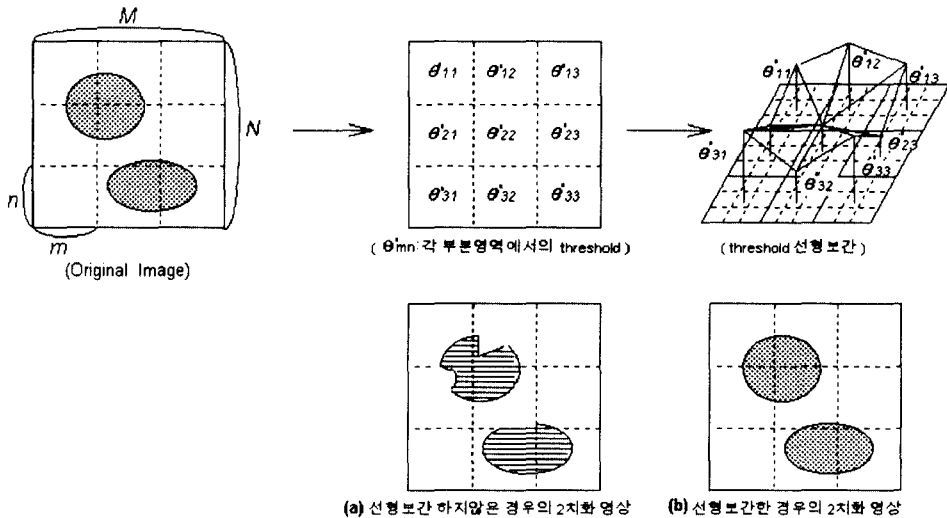


그림 3.2 국소적 2치화 처리
Fig 3.2 Local thresholding for the binarization

부분 영역에서의 임계치를 결정하기 위하여 흑, 백의 2클래스로 나누고 클래스 간 분산이 최대가 되는 농도로 임계치를 결정한다. 즉, L 은 전체 농도레벨수, k 는 농도 레벨($1 < k < L$), $\sigma_B^2(k)$ 를 농도 레벨 $1 \sim k$ 와 $k+1 \sim L$ 의 2클래스로 분리한 경우의 클래스 간 분산이라 했을 때, 임계치는 다음의 식과 같다.

$$\sigma_B^2(k) = \frac{|\mu_T \omega(k) - \mu(k)|^2}{\omega(k)|1 - \omega(k)|} \quad (1)$$

$$\sigma_B^2(k^*) = \max\{\sigma_B^2(k)\} \\ 1 \leq k \leq L-1 \quad (2)$$

식(1)에서 μ_T 는 전체 평균 농도 레벨이며, $\omega(k)$ 와 $\mu(k)$ 는 각각 농도 레벨 $1 \sim k$ 까지의 농도 히스토그램의 0차 및 1차모멘트를 나타낸다. 그리고 식(2)에서와 같이 $\sigma_B^2(k)$ 를 최대로 하는 농도치 k^* 를 부분 영역의 임계치 $\theta_{m,n}$ 으로 결정한다.

3.1.2 임계치의 영역 간 보간

부분 영역마다 얻어진 임계치 $\theta_{m,n}$ 는 문자와 그 배경에서 구성되고 있는 영역에서는 적정치가 주어지지만 그 이외의 영역에서는 반드시 적정치가 주어지는 것은 아니다. 그래서 콘트라스트 평가치 $C_{m,n}$ 에 의해 그 적정도를 평가하고, 인접 부분 영역 간의 부적당한 임계치의 보정을 행함으로써, 배경 잡음의 발생을 억제한다.

$$C_{m,n} = \frac{\sigma_{m,n}}{\mu_{m,n}} \quad (3)$$

$\sigma_{m,n}$: 부분영역(m, n)에서의 분산
 $\mu_{m,n}$: 부분영역(m, n)에서의 평균

먼저 각 부분 영역에 대해, 식(3)을 이용하여 얻어진 콘트라스트 평가치에 의해 보정결정치 $W_{m,n}$ 을 구할 수 있다.

$$W_{m,n} = \begin{cases} 0 & ; C_{m,n} < \phi \\ 1 & ; \text{otherwise} \end{cases} \quad (4)$$

식(4)에서 ϕ 는 콘트라스트 평가치에 대한 임계치이고, 문자 영역의 정보를 유지하면서 배경 잡음이 극히 적게 되도록 실험적으로 결정한다.

$W_{m,n}$ 이 1인 경우에는 $\theta'_{m,n} = \theta_{m,n}$ 이 된다. 다음으로 보정결정치 $W_{m,n}$ 이 0인 부분 영역에 대해, 임계치 $\theta_{m,n}$ 과 인접하는 8근방 영역의 임계치와 보정결정치에 의해 식(5)에 따라 $\theta'_{m,n}$ 을 구한다.

$$\theta'_{m,n} = \frac{\sum_{p=m-1}^{m+1} \sum_{q=n-1}^{n+1} W_{p,q} \theta_{p,q}}{\theta_{m,n}} \quad (5)$$

단, $1 \leq m \leq M-1, 1 \leq n \leq N-1, \theta_{m,n} = \sum_{u=m-1}^{m+1} \sum_{v=n-1}^{n+1} W_{u,v}$

만일 $W_{u,v} = 0$ 이면 $\theta_{m,n} = 1$

$\theta_{m,n}$ 은 부분 영역(m,n)의 8근방의 보정결정치의 합을 나타낸다.

3.1.3 임계치의 화소 간 선형 보간

3.1.2장에서 구한 부분 영역의 임계치 $\theta'_{m,n}$ 을 그림3.2(b)와 같이 각 부분 영역(m,n)의 중심에 위치시켜, 영상 전체에 대하여 선형 보간(linear interpolation) 함으로써 최종 임계치 $\theta^*(i,j)$ ($0 \leq i \leq M, 0 \leq j \leq N$)을 구한다.

3.1.4 자연 영상의 2치화 처리

3.1.3장에서 얻어진 $\theta^*_{i,j}$ 을 이용하여 다음 식에 따라 2치화 처리를 행한다. 이 때, $f(i,j)$ 는 원영상에서 각 화소(i,j)의 농도값(gray level)을 나타낸다.

$$B(i,j) = \begin{cases} 0 & ; f(i,j) \geq \theta^*(i,j) \\ 1 & ; \text{otherwise} \end{cases} \quad (6)$$

3.2 문자 후보 영역 추출

카메라에 잡힌 자연 영상에서 문자 영역에 해당하는 부분에 대해 문자와 배경 부분의 농도 콘트라스트는 높다[12]는 특징을 도입하여 실제 문자 영역에 해당되지 않는 부분을 제거할 수 있다.

농도차 평가용 패턴의 예를 그림 3.3에 나타내며, 위의 문자 영역 특성을 기초로 한다.

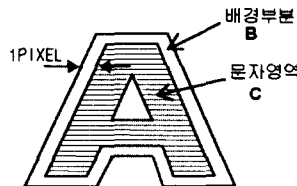


그림 3.3 농도차 평가용 패턴
Fig 3.3 A pattern to evaluate the gray-level difference

위의 그림에서와 같이 배경 부분 B와 문자 영역 C의 농도를 L_B, L_C 로 하면

$$\Delta L = |L_C - L_B| \quad (7)$$

이며, ΔL 을 파라메타로 사용한다.

각 영역에서

$$\Delta L \geq \theta_L \quad (8)$$

(θ_L : threshold)

를 만족하는 영역은 문자 후보 영역으로 남고, 식(8)을 만족하지 않는 경우는 잡음 영역으로 배경 부분에 결합된다.

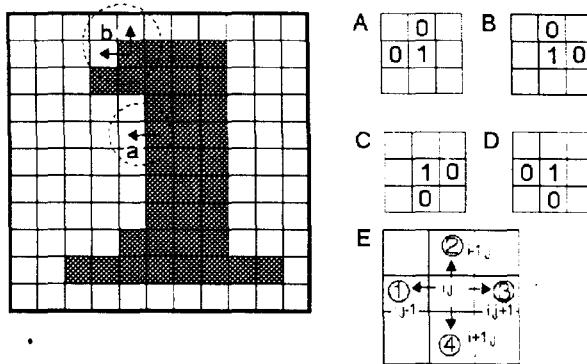


그림 3.4 농도차 평가

Fig 3.4 The evaluation of the gray-level difference

그림 3.4를 참고로 하여 그림 3.5와 같은 알고리즘에 따라 주목 영역을 배경 부분에 결합시킨다.

그림 3.5의 알고리즘에서 start point는 해당 영역의 시작좌표점-1을, 그리고 end point는 해당 영역의 마지막 좌표점+1을 나타내며, O_s 와 B_s 는 각각 해당 영역의 농도값(gray-level)과 이와 인접한 배경 영역의 농도 값을 나타낸다.

이 알고리즘을 그림 3.4 E와 같이 문자 후보 영역의 4방향에 대해 모두 행한다. (1)이 ②일 때 (2)=A와B, (3)= $f(i-1, j)$, (1)이 ③일 때 (2)=B와C, (3)= $f(i, j+1)$, (1)이 ④일 때 (2)=C와D, (3)= $f(i-1, j)$ 가 된다. 그리고 (4)는 그림 3.4의 a와b의 경우의 크기를 비교하는 것으로서, 분리하는 이유는 3.1장에서 얻어진 2치화 영상에서, 같은 문자 후보 영역 내의 a와 b 부분의 농도 값이 크게 차이가 생겨 ΔL 계산에 커다란 영향을 미치므로 분리하여 계산한다.


```

FOR(i=start point;i<end point;i++)
{ FOR(j=start point;j<end point;j++)
  { IF (①=배경 부분)&&((i,j)=문자 영역) ← (1)
    { IF( A and D ) ← (2)
      { Os1 = Os1 + f(i,j)
        Bs1 = Bs1 + f(i,j-1) ← (3)
        C1++
      }
      ELSE { Os2 = Os2 + f(i,j)
             Bs2 = Bs2 + f(i,j-1)
             C2++
          }
    }
  }
}
IF( C2 >= C1 ) ← (4)
{ Lc = Os2 / C2
  Lb = Bs2 / C2
  IF( (|Lc-Lb|) <  $\theta_L$  ) Object는 배경 부분에 결합
}
ELSE { Lc = Os1 / C1
       Lb = Bs1 / C1
       IF( (|Lc-Lb|) <  $\theta_L$  ) Object는 배경 부분에 결합
}
}

```

그림 3.5 콘트라스트 평가에 의한 결합 알고리즘

Fig 3.5 Merging algorithm by the evaluation of contrast difference

3.3 문자 패턴 후보 추출

3.2장까지의 처리에서 문자 후보 영역을 검출할 수 있는데, 실제 문자 영역 이외에도 잡 영성분이 포함되어 있다. 그러므로 다음과 같은 특성을 고려하여 잡영성분을 제거한다.

- (1) 문자는 일정 갯수 이하의 구멍을 가진다.
- (2) 문자의 종횡비는 20:1 이하이다.
- (3) 한 문자의 최소 크기는 8화소 이상이다.
- (4) 입력 영상에서 문자의 폭은 2화소 이상이다.
- (5) 입력 영상에서 문자 영역은 테두리와 인접하지 않는다.

위의 특성에 의해, 먼저 테두리와 인접한 문자 영역을 제거한다. 3.1장에서 얻어진 2차화 결과를 검토하여 보면 원영상 중에서 부분 영역(m,n)의 크기 혹은 2차화에 사용된 윈도우의 크기(template size)의 1/2정도의 테두리 부분($0 \leq i \leq m/2, M-m/2 \leq i < M, 0 \leq j \leq m/2,$

$M-m/2 \leq j < M$)은 백 또는 흑의 영역으로 되어있다. 결국 이 테두리 부분과 인접한 문자 후보 영역은 문자가 아닌 배경 부분에 속한다고 볼 수 있으며, 이러한 영역은 배경 부분으로 결합시킨다. 또한, 그림 3.7과 같이 문자의 가로와 세로의 비는 분리 문자의 경우('기'의 '丨')를 고려하여 20:1을 벗어나지 않는다고 가정한다[4]. 본 알고리즘을 실험한 결과 2치화 한 영상에서 하나의 문자에 해당하는 문자의 폭은 2치화 처리에서의 문자 영역 소실을 최소한으로 줄이는 국소적 2치화 처리 방법에 의해 최소 2화소 이상이 된다. 또한 입력 영상에서의 인식할 수 있는 한 문자의 최소 화소 수는 8화소 이상이라고 본 논문에서는 가정한다.

다음은 위의 특성을 바탕으로 비문자 영역을 제거하는 방법을 설명한다.

(1)을 기초로 4장까지의 처리에서 얻어진 문자 후보 영역에서 각 문자 후보 영역의 구멍의 갯수 H_N 를 계산한다. 그림 3.6은 그 예를 나타내고 있다.

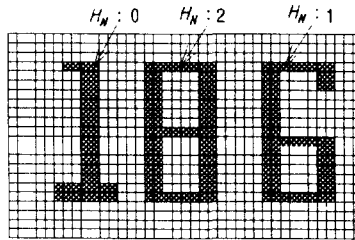


그림 3.6 문자 영역의 구멍의 수
Fig 3.6 The number of holes of a character region

위의 계산에서 $H_N > \theta_H$ 이면 비문자 영역으로서 문자 영역에서 제외된다. 다음은 문자 후보 영역을 둘러싸는 사각형의 종횡비를 계산한다.

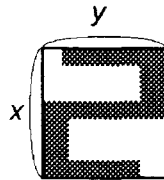


그림 3.7 문자의 종횡비
Fig 3.7 The x-y ratio of the circumscribed rectangle of a character pattern

위의 계산에서 중횡비가 20:1를 넘어서면 문자 영역에서 제외된다.

다음은 하나의 문자 영역에서 인식할 수 있는 주목 화소의 수는 최소 8화소 이상이라는 성질에 기초하여 해당 문자 영역이 가지는 화소 수가 8화소 이하의 화소를 가지는 경우 배경 부분에 결합된다. 그리고, 문자 영역의 최소 폭을 2화소로 하여 그림 3.8에서와 같이 x 나 y 가 1이면 배경 부분에 결합된다. 단, 사선과 같은 경우는 제거되지 않는다.

이상의 결과로 배경 부분에 속하는 부분은 모두 제거되고 문자 영역에 해당하는 부분만 남게 되어 문자 패턴 후보로 검출되게 되며, 다음의 처리로 문자 패턴을 추출할 수 있다.

3.4 문자 패턴의 추출

이 장에서는 3.3장까지의 처리 결과를 대상으로 투영법을 적용하여 문자 패턴을 추출하는 것을 검토한다.

문자 패턴을 추출하기 전에 3.3장까지의 처리 후에도 여전히 문자를 둘러싸고 있는 테두리 영역은 문자의 성질과 동일하기 때문에 문자 패턴 후보로 존재하게 되는데 이러한 영역을 제거하기 위하여, 그림 3.7와 같이 각 문자 영역을 사각형으로 에워싼 후, 이 사각형에 포함되거나 접해 있는 사각형 수의 임계치 θ_m 를 적용하여, 이 임계치보다 많은 문자 영역을 포함하면 비문자 영역으로 제거한다. 그런 다음 다음의 단계에 의해 문자 패턴을 추출한다.

- 1) 수평 방향으로 투영을 하여 문자열을 추출한다.
- 2) 추출된 문자열을 수직 방향으로 투영하여 각 개별 문자 패턴의 수직 결합을 행한다.
- 3) 수직 결합된 개별 문자 패턴들을 바탕으로, 문자 폭이 가장 큰 개별 문자 패턴을 제외한 나머지 개별 문자 패턴 가운데에서 이웃한 개별 문자 패턴의 문자간 간격 B_{cc} 가 평균 문자 폭 A_{cs} 의 절반보다 크지 않으면 서로 결합한다. 즉, 1)에 의해 얻어진 문자열에서, 단어 영역(word region)이 검출된다.
- 4) 단어 영역 내에서 이웃한 두 개별 문자 패턴을 결합한 폭($C_{1w} + C_{2w} + B_{cc}$)이 평균 문자 폭 A_{cs} 의 1.2배보다 작으면, 이웃한 두 개별 문자 패턴은 하나의 문자 패턴으로 결합된다.
- 5) 만일 개별 문자 패턴의 폭 C_{1w} 가 평균 문자 폭 A_{cs} 의 1.2배 보다 크거나 같으면, 개별 문자 패턴으로 남겨진다.
- 6) 3)에서 얻어진 단어 영역 내에 있는, 문자 패턴으로 추출된 패턴들 중에서 가장 큰 패턴의 크기 C_{BASE} 를 기준으로 개별 문자 패턴 C_{1w} 가 C_{BASE} 보다 작으면 이웃한 영역과 합한 크기와 비교하여, 결합의 여부를 판별하고, 또한 C_{1w} 가 C_{BASE} 보다 1.8배 이상이면 $C_{1w}/2$ 를 기준으로 개별 문자 패턴을 두개의 문자 패턴으로 분리한다.
- 7) 추출된 문자열에서 수직 결합한 개별 문자 패턴이 하나만 존재할 경우는 개별 문자 패턴이 문자 패턴으로 존재하게 된다.

위와 같은 알고리즘에 의해 모든 문자 패턴의 추출이 가능하며, 3.3장까지의 처리 후에도 여전히 남아있는 잡영 영역도 개별 문자 패턴으로 존재하여, 이상의 처리에 의해 문자 패턴으로 존재하게 된다. 이러한 경우의 잡영 성분은 자연 영상에서의 문자의 존재 조건과 유사한 상태의 잡영 성분으로서 이러한 성분은 인식에 의해 처리되어야 할 것으로 생각된다.

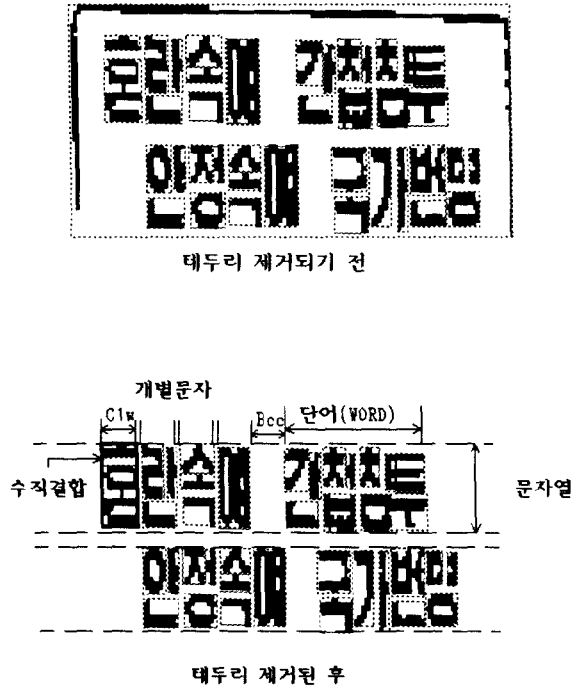


그림 3.8 문자 패턴 추출
Fig 3.8 The detection of character patterns

IV. 실험 및 결과

본 논문에서는 카메라 및 스캐너로 입력한 8-bit gray-level의 자연 영상을 입력 영상으로 사용하였으며, IBM-PC 486에서 제안한 알고리즘을 C언어로 구현하였다.

입력된 자연 영상들에서 문자 패턴을 추출하는 실험을 수행하였다. 실험 결과 문자의 위치, 크기, 문자 종류 등의 사전 지식이 없이 문자 패턴 영역을 비교적 정확하게 추출할 수 있었다.

그림 4.1은 카메라 및 스캐너로 입력된 gray-level 영상을 나타내며, 그림 4.2는 2치화 예를 나타내고 있다. 그리고 그림 4.1(c)의 입력 영상에 대하여 국소적 2치화, 문자의 자연 영상에서의 존재 특성 등을 이용하여 문자 패턴 추출을 하였다.

4.1 부분 영역(m, n)의 크기

국소적 2치화를 하여 그림 3.2에 나타난 부분 영역의 크기 $m \times n$ ($m < 256$, $n < 256$)을 어느 정도로 결정하면 되는가를 실험하였다. 시각적으로 평가한 결과 그림 4.2와 같이 (a)4 \times 4인 경

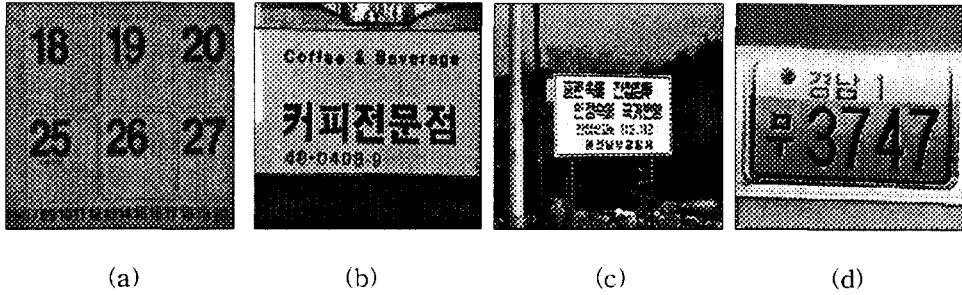


그림 4.1 입력 자연 영상

Fig 4.1 Natural scene images used in the experiments



그림 4.2 부분 영역의 크기에 따른 2치화 영상($C_{mri}:0.1$)

Fig 4.2 Binary images for various sizes of sub-regions($C_{mri}:0.1$)

우는 문자 내부에 배경 부분이 생겨 문자 정보가 소실되는 경우가 생겼으며, (c)16×16 인 경우는 배경 부분과 인접한 부분의 문자 영역의 정보가 소실되거나 문자 영역의 윤곽이 뚜렷하지 않은 형태가 되는 경우가 많고, 문자의 크기가 작은 경우 문자 영역이 사라지는 경우가 있었다. 이러한 이유는 부분 영역의 크기가 작은 경우는 문자 영역에 부분 영역이 포함되어 그 결과 2치화를 강제적으로 하는 관계로 배경 부분이 문자 내부에 생기는 예가 발생한다고 생각되며, 이와는 반대로 부분 영역의 크기가 큰 경우는 임계치(threshold)가 배경 농도의 영향을 받으므로, 문자 영역의 세부 검출이 잘되지 않는다고 생각된다. 반면, 크기가 작은 경우는 문자의 세밀한 부분까지 검출이 가능하고, 크기가 큰 경우는 배경 부분의 잡음 영역의 발생이 억제되는 성질도 있었다. 이에 본 논문에서는 (b)8×8의 부분 영역을 가지고 임계치를 결정하였으며, 비교적 양호한 2치화 결과를 얻을 수 있었다.

결과 영상(b)에서 문자 '커피'의 좌측에 비문자 영역 '■'이 문자 영역처럼 발생한 이유는 그림 4.1(b)의 원래의 입력 영상에서 알 수 있듯이 '■'이 발생한 지점은 배경 부분 내의 농도 차가 크게 나타나므로 이러한 경우가 발생한 것이다. 그림 4.3는 이러한 예를 나타낸다.

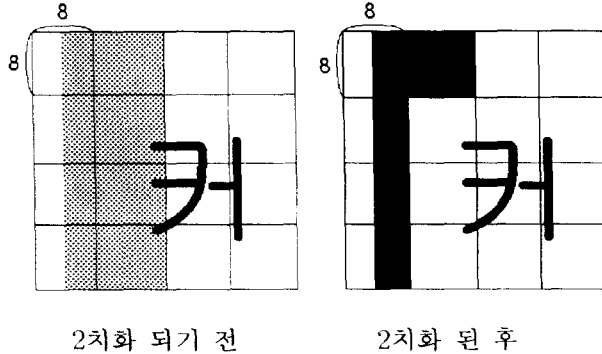


그림 4.3 2치화 영상
Fig 4.3 The binarization of an image

위와 같은 결과는 각 부분 영역의 임계치(threshold)의 영향으로 인하여 주목하는 부분 영역의 임계치가 주위와 인접해 있는 부분 영역의 임계치의 영향을 받기 때문에 생기는 결과이다. 그러나 이러한 결과는 3.2장에서 의 문자 후보 영역 간의 농도 콘트라스트를 평가함으로써 배경 부분으로 결합될 수 있다.

그림 4.4는 (a)0.01, (b)0.05, (c)0.1의 농도 콘트라스트 임계치 C_{mn} 로 실험한 결과이다. 농도 콘트라스트의 임계치 C_{mn} 를 결정함에 있어서, 농도 콘트라스트 임계치가 0.01인 경우는 배경 잡음이 많이 발생하였으며, 0.1인 경우는 배경 잡음이 억제되어 본 논문의 처리를 간소화시킬 수 있었다. 또한 입력 영상에서 시각적으로 배경 부분으로부터 문자를 쉽게 분리 가능한 정도를 생각해 볼 때, 임계치는 0.1정도가 적당할 것으로 생각된다. 하지만 0.1보다 큰 임계치를 설정할 경우는 문자 영역이 배경 부분으로 결합되는 경우가 발생하여, 오히려 문자 영역을 소실하는 경우가 발생하므로 본 논문에서는 농도 콘트라스트 임계치를 $C_{mn} = 0.1$ 로 설정하였다.

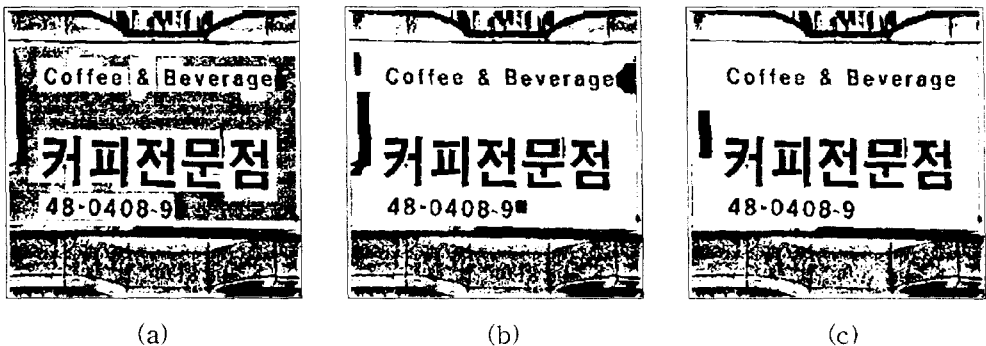


그림 4.4 농도 콘트라스트 임계치($m \times n=8 \times 8$)
(a) $\phi=0.01$ (b) $\phi=0.05$ (c) $\phi=0.1$

Fig 4.4 Images of various thresholds of gray-level contrast differences($m \times n=8 \times 8$) (a) $\phi=0.01$ (b) $\phi=0.05$ (c) $\phi=0.1$

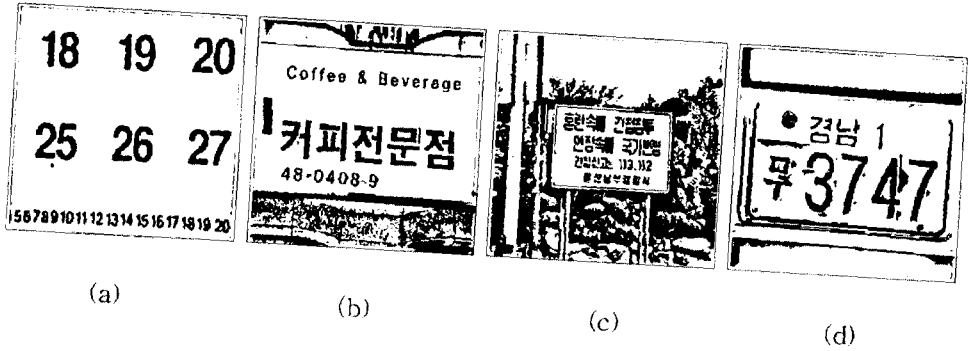


그림 4.5 국소적 2치화 방법에 의한 2치화 영상($m \times n=8 \times 8, \phi=0.1$)
 Fig 4.5 Binary images generated by the local thresholding method ($m \times n=8 \times 8, \phi=0.1$)

4.2 문자 후보 영역의 추출

3.2장의 처리에 의한 결과를 그림 4.6에 나타낸다. 본 논문에서 ΔL 의 임계치는 $\theta_L=18$ 로 설정하였으며, 그 결과는 그림 4.7과 같다. 이 값은 실험에 의하여 얻어진 최적의 임계치이며, 입력 영상에서 쉽게 문자를 추출할 수 있는 경우의 값을 설정한 것이다. 만일 이 임계치의 값이 이 보다 작은 경우는 농도의 변화가 심한 배경 부분에서 3.1장까지의 처리에서 제거되지 않은 비문자 영역이 제거되지 않았다. 또한 이 보다 큰 값을 설정하였을 경우는 오히려 문자 영역이 소실되는 결과가 발생하였다.

본 논문에서는 그림 3.5의 알고리즘을 적용하여 비문자 영역을 제거하였는데, 이러한 방법을 적용한 이유는, 기존의 방법[12]에서 사용한 영역 간의 평균 농도의 차이를 사용할 경우 커다란 효과를 기대할 수 없었으며, 이 방법은 그림 3.4의 농도차 평가용 패턴에서 a, b부분과 문자 후보 영역이 테두리 부분의 값의 영향을 많이 받게 되므로, 배경부분과 문자의 경우는 배경과 결합되는 경우는 잘 발생하지 않지만, 발생하게 되면 시각적으로

도 그 부분에 대해서 문자로 인식하기는 힘들다.

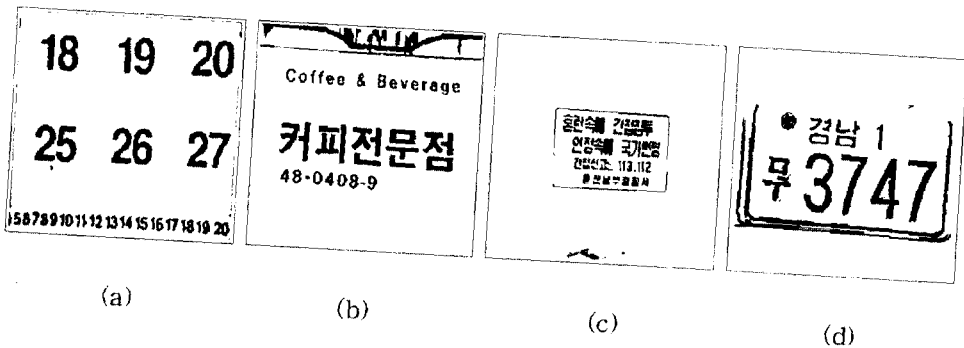


그림 4.6 농도 콘트라스트 평가의 결과 ($\theta_L=18$)
 Fig 4.6 The result images of the gray-level contrast difference evaluation($\theta_L=18$)

4.3 문자 패턴 후보 추출

3.3장에서 제시한 몇 가지 제약 조건을 만족하는 영역들을 배경 부분 영역에 결합시킴으로써 다음과 같은 결과를 얻을 수 있었다. 집합 문자의 경우를 고려하여 문자 영역 내의 구멍(hole)수의 임계치를 $\theta_H=4$ 로 결정하였다.

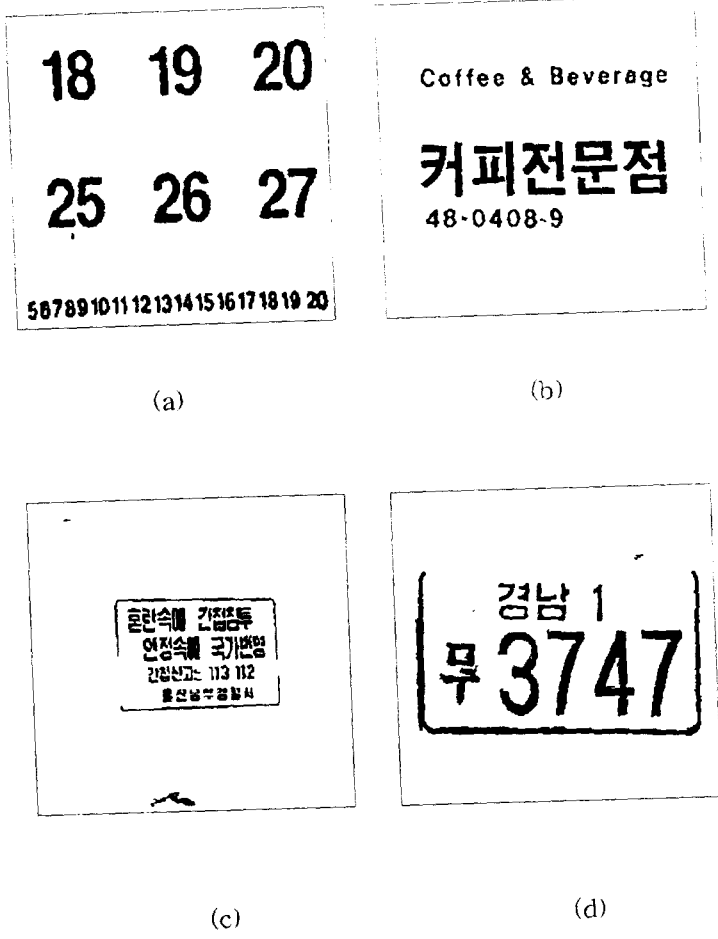


그림 4.7 문자 패턴 후보
Fig 4.7 Character pattern candidates

4.4 문자 패턴의 추출

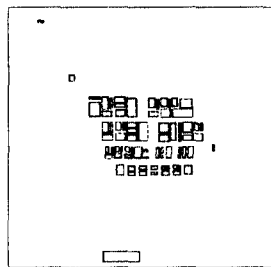
3.4장의 투영 방법을 적용하여 다음과 같은 결과를 얻었다. 수평 방향으로 투영을 하여 문자열을 추출한 후, 그림 4.8(a)와 같이 문자 영역을 둘러

싸는 사각형을 바탕으로, $\theta_m=5$ 를 적용하여 태두리 부분을 제거하였다. 이 결과에 대해 수직으로 다시 투영을 함으로써, 그림(b)와 같이 한글 폰트와 같은 문자 영역에 대해 수직 결합을 수행하였다. 여기에서 ‘침투’, ‘번영’, ‘간첩’과 같은 영역은 문자가 서로 연결되어 있거나, 또는 문자 영역을 에워싼 사각형이 서로 접하는 경우로서 하나의 영역으로 결합된다. 이러한 영역은 마지막 단계에서 분리가 가능하다.

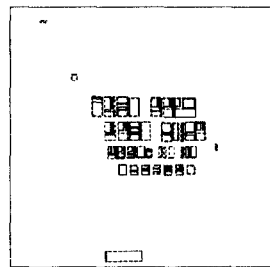
다음으로 한 문자열에서의 평균 글자 폭 Acs의 1/2보다 크지 않은 문자 간격 Bcc를 가지는 이웃한 개별 문자 패턴끼리 결합을 함으로써, 그림 (c)와 같은 단어 영역을 추출한다. 단어 영역을 추출한 이유는 문자열 가운데에 주변의 문자보다 그 크기가 큰 문자가 있을 경우, 이러한 문자는 대부분 이웃 문자와 결합되었다고 생각될 수 있으며, 이러한 문자는 주변의 문자 크기와 비교하여 분리할 필요가 있다.

개별 문자 패턴의 폭이 평균 문자 폭 보다 1.2배 이상의 문자 패턴, 즉 그림(c)의 ‘침투’, ‘번영’의 경우와 같이 서로 영역이 결합되어 있거나 접해 있는 경우, 주위의 문자들과 비교하여 그 크기가 1.2배 이상이 되므로 단어 영역으로 결합되지 않고 개별 문자 영역으로 분리된다. 그림(c)에서 ‘울’에 해당하는 영역이 단어 영역에서 제거되어있다. 이것은 해당 문자열에서 그 크기가 가장 크기 때문에 3.4장의 3)단계에서 제거된 것이다.

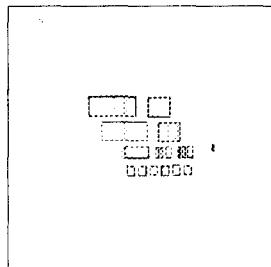
마지막으로 서로 이웃한 개별 문자끼리 결합한 크기가 해당 문자열에서의 평균 문자 크기와 비교하여 1.2배를 넘지 않으면 하나의 문자 패턴으로 결합되며, 개별 문자의 크기가 평균 문자 크기의 1.8배 이상이면 이 문자의 크기는 가운데 지점을 기준으로 분리가 이루어진다. 그 결과를 그림 (d)에 나타내고 있다.



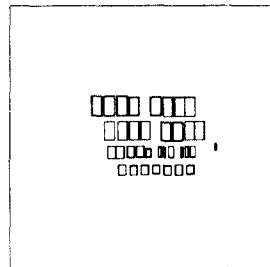
(a)



(b)



(c)



(d)

그림 4.8 그림 4.1(c)에 대한 문자 패턴 추출 결과
 (a) 테두리를 제거한 문자 영역 (b) 수직 결합에 의한 문자 영역
 (c) 단어 영역 (d) 문자 패턴

Fig 4.8 The extraction results of character patterns of the input image of Fig 4.1(c)

- (a) Character regions without outlines
- (b) Character regions formed by the vertical merging
- (c) Word regions (d) Character patterns

다음 그림은 그림 4.1의 입력 영상들에 대하여, 각각의 문자 패턴을 추출한 결과 영상들이다. 그리고 다른 영상에 대한 결과를 부록에 나타내었다. 그림에서 볼 수 있듯이 본 알고리즘은 옥외에서 얻은 자연 영상에서 정확하게 문자 패턴을 추출하고 있음을 알 수 있다.

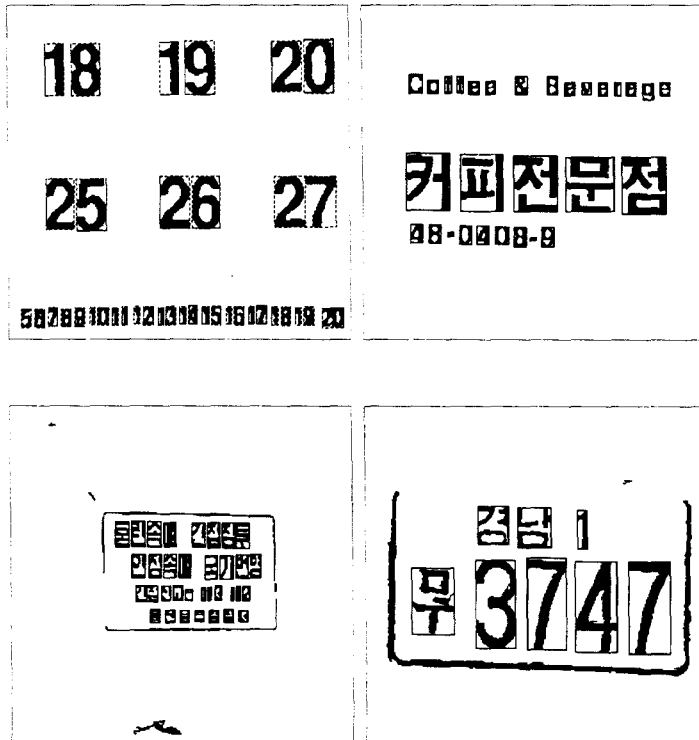


그림 4.9 문자 패턴 추출 결과

Fig 4.9 The resultant images of the detected character patterns

V. 결 론

자연 영상에서 배경 부분과 분리가 가능한 문자들이 가지는 보편적인 성질들을 이용하여 문자 패턴을 추출하였다. 또한, 배경 부분과 문자 영역과의 식별이 가능한 최소 농도 콘트라스트 차이를 이용함으로써 비문자 영역들을 문자 패턴에서 제거할 수 있었으며, 효과적인 농도 콘트라스트 평가 방법을 제시하였다.

본 방법은 범용성을 가지기 위해 제안되었으며, 사전에 주어지는 파라메타가 적다. 본 방법을 앞에서 보여준 종류의 패턴들 외에도 여러 가지 패턴에 적용을 해 보았으며, 다음의 결과를 얻었다.

- (1) 문자 영역과 배경 부분의 영역 분할을 위한 국소적 2치화 방법을 적용함에 있어서, 최적의 부분 영역 크기와 콘트라스트 평가치가 있었다.
- (2) 문자 영역과 배경 부분의 농도 차를 비교함으로써, 국소적 2치화 방법에 의해 발생하는 비문자 영역을 대부분 제거할 수 있었다.
- (3) 위에서 얻어진 결과에 대해, 제거가 되지 않은 비문자 영역에 대하여 몇 가지 입력 영상에서의 문자의 존재 제약 조건을 적용함으로써, 대부분의 비문자 영역을 제거할 수 있었다.

그러나, 본 방법은 굴곡이 있는 문자나, 그래픽 문자에 적용하였을 경우 별로 효과적이지는 못했다.

참 고 문 헌

1. 이성환, "문자 인식," 홍릉과학출판사, 1993
2. 남궁재찬, 류황빈, 남궁연, "한국어 문서로부터 문자 분리 및 도형 추출에 관한 연구," 대한전자공학회논문지, 제25권 9호, pp. 1091-1100, 1988년 10월.
3. T. K. Kim, T. Agui and M. Nakagima, "On-Line Recognition of Korean Characters by Multistage Translation of Directional Codes," 일본전자통신학회논문지, Vol. J680-D, No. 2, pp. 177-184, 1985.
4. Lloyd A. Fletcher and Rangacher Kasturi, "A Robust Algorithm for Text String Segregation from Mixed Text/Graphics Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, pp. 910-918, Nov. 1988
5. L.O'Gorman, "The Document Spectrum for Page Layout Analysis," *IEEE Trans. on PAMI*, vol. 15, no. 11, pp. 1162-1173, Nov. 1993
6. 백영목, 임길택, 김우태, 진상일, "영역레이블링 방식을 이용한 일반 문서의 LAYOUT UNDERSTANDING," 제6회 영상처리 및 이해에 관한 워크샵 논문집, pp. 204-209, 1994
7. 김의정, 김태균, "인쇄체 문서인식을 위한 문자 추출에 관한 연구," 제2회 문자 인식 워크샵, pp. 171-179, 1994
8. 최형진, 오영환, T. Agui, M. Nakajima, "컴퓨터 화상 처리를 이용한 차량번호판 추출 방법," 전자공학회논문지, 제24권 2호, pp. 127-132, 1987년 3월
9. 구정희, 장정훈, 홍기상, "Karhunen-Love" 변환을 이용한 옥외 환경에서의 화차 번호 인식

- 시스템 설계," 제7회 영상 처리 및 이해에 관한 워크샵 발표 논문집, pp. 109-115, 1995
10. 변성일, 구정화, 홍기상, "화차 번호 인식을 위한 영역 분할," 제1회 문자 인식 워크샵, pp. 201-205, 1993
 11. 이승우, 이기성, 구건서, 오해석, "부품 분류를 위한 각인 문자 인식," 제2회 문자 인식 워크샵, pp. 213-227, 1994
 12. J. Ohya, "Recognition Characters in Scene Images," *IEEE Trans. on PAMI*, vol. 16, no. 2, pp. 214-220, Feb. 1994
 13. A. Shio, "A Dynamic Thresholding Method for Character Detection in Complex Scene Images," *Trans. on IEICE, Japan*, vol. J71-D, no. 5, pp. 863-873, May. 1988
 14. N. Otsu, "A Threshold Selection Method from Gray Scale Histogram," *IEEE Trans. on SMC*, vol. SMC-9, no. 1, pp. 62-66, 1979.