

# 운송 컨테이너 영상의 효율적인 문자인식을 위한 전처리에 관한 연구

최 재영\*, 김 낙빈\*

\*경원대학교 전자계산학과

e-mail:jy1001@cs.kyungwon.ac.kr

## A Study on Preprocessing for Efficient Character Recognition of Shipping Container Image

Jae-Young Choi\*, Nak-Bin Kim\*

\*Dept. of Computer Science, Kyungwon University

### 요 약

본 논문은 운송 컨테이너 식별자의 자동화 처리를 위한 문자 인식의 단계중 최종 문자 인식 전단계 까지의 처리 과정을 컨테이너의 특성에 맞게 제안하였으며, 이러한 전처리 과정은 문자 인식 시스템의 성능에 중요한 영향을 미친다. 제안한 방법은 먼저 입력된 컨테이너 컬러 영상을 명암 영상으로 바꾸고 전체 영상중 인식에 필요한 식별자 영역만을 경계선 검출과 형태학적 연산을 이용하여 추출한다. 이어서 다양한 배경색과 문자색을 판단하여 일반 문서와 같이 일관성 있게 통일한 후, DCT를 이용한 명암도별 이진영역으로 분할한 후에 Otsu방법과 새로운 이진화방법을 자동으로 선택하여 효율적인 이진화가 이루어지도록 하였다. 이렇게 얻어진 이진 영상은 문자인식 단계로 넘어갈 수 있도록 개별 문자로 분할한다. 이 방법은 컨테이너 영상의 불균등한 배경색과 잡음으로 인하여 문자인식에 오류가 생기는 단점을 보완하였으며 컨테이너 특성을 최대한 반영함으로써 효과적인 전처리 결과를 얻을 수 있었다. 또한, 제안한 방법의 응용은 컨테이너 이외의 다른 상황에서도 매우 효과적으로 사용될 수 있으리라 본다.

### 1. 서 론

해로를 통한 수출입은 항만에서의 신속하고 정확한 물류처리를 요구하고 있으며 현재까지는 이를 만족시키기 위한 바코드 시스템이 있지만 주로 화상처리 기반의 컨테이너 식별자 자동 인식 시스템을 이용하여 자동화 하고 있다.

따라서 문자인식 기술은 식별자 자동인식 시스템의 성능을 좌우하는 중요한 기반이 되고 있으며 특히 컨테이너 식별자의 경우 일반적인 문서에서의 문자인식과는 달리 외부 환경과 자체 굴곡에 의하여 변형되고 특별히 정형화 되어 있지 않으며 불 균일한 문자의 크기와 간격, 일정하지 않은 바탕색과 문자색등 여러 가지 어려운 요인을 포함하고 있기 때문에 기술의 개선이 더욱 필요하다. 이에 본 연구에서는 식별자의 인식 성능을 향상시키기 위해 영상 입력 후 식별자 영역 추출에서부터 식별자의 개별 문자를 정확하고 빠르게 분리하는 과정까지의 전처리 방법을 제안한다.

또한, 제안한 방법의 효용성을 증명하기 위하여

matlab[1]을 사용하여 여러 컨테이너 입력 영상들을 실험하였다. 그 결과 제안한 방법이 컨테이너 특성에 맞게 효과적으로 적용됨을 확인할 수 있었으며 컨테이너 영상 이외의 다른 상황에서도 매우 효과적으로 사용될 수 있으리라고 확신한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 관련 연구를 기술하고, 제 3장에서는 식별자 영역 추출, 문자색 구분, 이진화, 문자 분할등의 전처리 알고리즘을 제시한다. 제 4장에서는 제시한 알고리즘의 실험 및 결과를 설명하고, 제 5장에서는 결론을 맺는다.

### 2. 관련 연구

이 장에서는 본 연구의 실험에 사용할 컨테이너 식별자 영역의 ISO규격, 그리고 식별자 영역 추출및 문자색 구분을 위하여 쓰이는 형태학적 연산에 대하여 알아보고 마지막으로 이진화 단계에서 적용될 DCT, Otsu 이진화 방법에 대하여 차례로 설명한다.

## 2.1 컨테이너 식별자

운송 컨테이너 영상은 그림 1에서 보는 바와 같이 컨테이너 식별을 위한 식별자 영역을 측면에 표기하는데, 식별자영역의 ISO 규격[2]은 4개의 영문자로 구성된 운송회사 코드(shipping company code), 아라비아 숫자 6개로 구성된 일련번호(serial number), 1자리 숫자인 검사 숫자(check digit), 컨테이너 유형 코드(container type code)의 4개의 영역으로 구성되어 있다. 운송회사 코드는 마지막(4번째) 문자가 반드시 "U" 자 이어야 한다는 규칙이 있다. 또한, 검사숫자까지 11자(영문자 4개 + 숫자 7개)는 규정 사항이므로, 컨테이너를 식별하기 위해서는 컨테이너 식별자의 처음 11자만을 인식하면 된다.



그림 1. 운송 컨테이너 영상

## 2.2 형태학적 연산

형태학적 연산(morphology)이란 AND와 OR, NOT 등과 같은 논리연산을 사용하여 형태소(structuring element)와 해당 화소들을 처리한다[3]. 형태학적 연산에서 기본이 되는 팽창연산(dilation)과 침식연산(erosion)은 형태처리연산에서 가장 기본적인 2가지 연산이며 팽창연산은 배경에 대해 물체의 크기를 확장한다. 형태소 B에 의한 원 영상 A의 팽창연산은 식(1)과 같이 정의 된다.

$$A \oplus B = \{x \mid [(\hat{B})_x \cap A] \subseteq A\} \quad \dots(1)$$

침식연산은 배경에 대해 물체의 크기를 축소하며 식(2)와 같이 정의된다.

$$A \ominus B = \{x \mid (B)_x \subseteq A\} \quad \dots(2)$$

열기연산(opening)과 닫기연산(closing)은 2차적인 연산으로 팽창과 침식연산을 이용하여 구현되며 팽창과 침식연산과 달리 원영상의 크기를 유지시킨다. 열기연산은 식(3)과 같이 침식연산 다음에 팽창연산을 행하는 것으로서, 그 효과는 침식과 마찬가지로 고립된 화소들과 같은 잡영을 없애는 것이다.

$$A \circ B = (A \ominus B) \oplus B \quad \dots(3)$$

닫기연산은 식(4)와 같이 팽창연산 다음에 침식연산을 수행하며 객체의 작은 틈새를 메우는데 이용된다.

$$A \cdot B = (A \oplus B) \ominus B \quad \dots(4)$$

## 2.3 DCT

DCT는 직교 변환의 하나로서 통계적으로 최적변환으로 알려진 KLT(Karhunen-Loeve Transform)에 가장 근접한 성능을 가지고 있으며, 영상 정보 특유의 인접 화소 간 높은 상관성을 이용하여 영상 신호를 특정 좌표축으로 집중시켜 영상 신호의 중복성을 줄여 주므로서 영상 및 음성 정보의 부호화에 많이 사용되고 있다[4].

$$C(u, v) = a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad \dots(5)$$

$$a(u) \text{ or } a(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u, v=0 \\ \sqrt{\frac{2}{N}} & \text{for } u, v=1, 2, \dots, N-1 \end{cases}$$

,where

DCT 변환 행렬의 각 성분은 코사인 함수로 표현하며, N차 (행렬 크기 N×N) 변환 행렬 C에서 2차원 DCT는 식(5)와 같다.

## 2.4 Otsu 이진화 방법

본 연구에서는 영상의 임계값을 설정하고 이진화하기 위하여 새로운 이진화 방법을 제안하는데, 이때에 같이 사용될 알고리즘은 광역임계법의 하나인 Otsu 이진화 방법으로서 식이 간단하고 효과적이며 매우 만족스러운 성능을 가지고 있으므로 일반적으로 자주 사용되고 있다[5].

Otsu의 이진화 방법은 쌍봉성의 히스토그램에서 임의의 기준값 T를 기준으로 T보다 작거나 같은 화소값들의 집합 C<sub>1</sub>과 큰 화소값들의 집합 C<sub>2</sub>가 있다고 가정하면 C<sub>1</sub>과 C<sub>2</sub> 각각의 분산은 작아야 하며 C<sub>1</sub>의 평균값의 위치와 C<sub>2</sub>의 평균값의 위치의 차는 클수록 좋다는 기본 개념을 가지고 있으며 이를 수식화하면 식(6)과 같다.

$$T^* = \arg \max S(T), 0 < T < L-1 \quad \dots(6)$$

$$S(T) = \frac{P_1(T)P_2(T)[m_1(T) - m_2(T)]^2}{P_1(T)\sigma_1^2(T) + P_2(T)\sigma_2^2(T)}$$

,where  $P_1(T) = \Pr\{C_1\} = \sum_{z=0}^T h(z)$ ,  $P_2(T) = 1 - P_1(T)$

$$m_1(T) = \frac{\sum_{z=0}^T z \Pr\{z \mid C_1\}}{\sum_{z=0}^T \Pr\{z \mid C_1\}} = \frac{1}{P_1(T)} \sum_{z=0}^T zh(z)$$

$$m_2(T) = \frac{1}{P_2(T)} \sum_{z=T+1}^{L-1} zh(z)$$

$$\sigma_1(T) = \frac{\sum_{z=0}^T [z - m_1(T)]^2 \Pr\{z \mid C_1\}}{\sum_{z=0}^T \Pr\{z \mid C_1\}} = \frac{1}{P_1(T)} \sum_{z=0}^T [z - m_1(T)]^2 h(z)$$

$$\sigma_2(T) = \frac{1}{P_2(T)} \sum_{z=T^*+1}^{L-1} [z - m_2(T)]^2 h(z)$$

, where  $\sum_{z=0}^{L-1} h(z) = 1, 0 \leq z \leq L-1$

식 (6)에서  $S(T)$ 가 최대가 되는  $T$ 를 임계값  $T^*$ 로 선택한다. 여기서  $z$ 는 입력영상에서 화소의 명암값이며  $L$ 은 입력영상의 명암도의 수이다.

### 3. 전처리 알고리즘

전처리 과정은 그림 2와 같이 컨테이너 측면의 컬러 영상으로부터 문자인식 단계 전까지 이루어지는 일련의 과정이다.

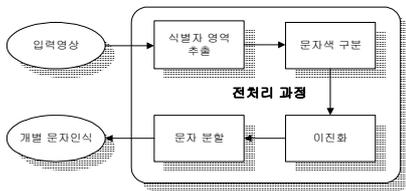
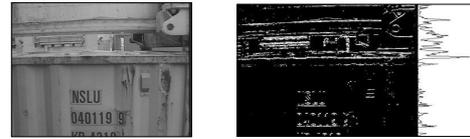


그림 2. 전처리 과정 흐름도

먼저 처리 속도를 높이기 위해 입력된 컨테이너영상인 컬러 영상을 명암(grey scale) 영상으로 변환하고, 수평 경계선을 이용하여 식별자를 포함한 부분의 영상을 얻는다[6]. 이렇게 얻어진 영상을 문자색과 배경색을 구분한 후에 DCT를 이용하여 명암도별로 이진화할 영역을 분할하고 각 분할된 영역별로 이진화를 수행한다. 마지막으로 이진화된 문자를 수직 경계선을 이용하여 개별 문자를 분할한다.

#### 3.1 식별자 영역 추출

명암영상으로 변환된 입력영상에서 식별자 영역 즉, 컨테이너를 식별하기 위해 인식해야하는 11문자가 있는 영역의 추출을 위하여 그림 3의 (a) 영상을 Sobel 마스크(mask)를 사용하여 수평방향의 경계선이 검출된 그림 3의 (b)와 같은 이진 영상을 얻는다. 그림 3(b)의 오른쪽 부분은 수평방향으로 각 행에 대한 흰 화소의 빈도수를 그린 것으로서, 흰 화소수가 가장 많은 곳이 컨테이너의 상단 모서리가 된다. 이 곳을 기준으로 영상의 밑 부분을 얻으면 그림 4의 (a)와 같이 식별자를 포함하는 부분의 순수한 컨테이너 영상을 얻는다. 이렇게 함으로써 컨테이너 영상이 아닌 다른 영상에 의한 오 인식 요소가 제거되고, 영상의 크기가 축소되어서 처리 시간이 단축된다.



(a)

(b)

그림 3. 입력 영상과 수평 경계선 검출

여기서 실질적으로 문자들이 위치하는 수평부분만을 얻고자 본 논문에서는 순수한 컨테이너 영상에 Sobel 연산을 사용하여 수평 및 수직 방향의 경계선을 검출한 다음 경계선이 검출된 영상에 1x15의 형태소로 단기연산을 행한 후에, 3x3의 4-이웃 형태소로 침식연산을 차례로 적용하였다. 그림 4의 (b)는 그 결과를 보여준 것으로 단기연산에 의하여 문자들의 수직 경계선들이 서로 연결되어 문자 영역은 확실해지고, 침식연산으로 기타 불필요한 작은 선들은 제거된 것을 알 수 있다.



(a)

(b)

그림 4. 형태학적 연산 결과

또한, 그림 4(b)의 오른쪽 부분은 수평방향으로 각 행에 대한 흰 화소의 빈도수를 그린 것으로서, 문자열의 위치를 쉽게 구별할 수 있음을 알 수 있다.

이와 같이 문자열이 존재한다고 판단한 부분 즉, 컨테이너의 식별자 영역들을 추출한 결과는 그림 5와 같이 문자들만 포함하는 2개의 문자열 영상이 된다.



그림 5. 분할된 문자열 영상

#### 3.2 문자색 구분

추출된 식별자 영역 영상에서 먼저 문자색과 배경색을 구분한다. 이러한 이유는 컨테이너의 문자색과 배경색은 일반 문서의 경우(배경:흰색, 문자:검은색)와 달리 여러 가지 경우가 있으므로 한가지색 문자(문자: 밝은색)로 통일을 하여 차후에 있을 연산에 일관성을 유지시킨다.

문자색과 바탕색을 판별하기 위하여 본 연구에서 제안하는 알고리즘은 다음과 같다.

크기  $m \times n$ 의 입력 영상을  $f(x,y)$ 라고 하면, 입력 영상에 Sobel의 수직 방향 마스크를 이용하여 수직 경

계선을 검출한 뒤 이를 이진화 시키면 그림 6의 (a)와 같은 이진 영상  $T(x,y)$ 를 얻을 수 있다.



그림 6. 수직 경계선이 검출된 이진 영상

다음은 그림 6의 (a)영상에서와 같이 글자 높이의 중심을 지나가는 스캔라인을 잡는다. 그런다음 스캔라인을 따라 수평방향으로 주사를 해 나가면서 세로 경계선이 검출된, 즉 화소값이 1인 좌표들 사이의 거리  $T_{dist}$ 를 구한 후  $T_{dist}$ 의 median 값을 구하면 문자의 대략적인 굵기인  $M$ 을 구할 수 있다. 문자의 굵기를 구하면 문자의 굵기만큼의 너비를 가진  $1 \times M$ 의 형태소를 사용하여 영상  $T(x,y)$ 를 팽창연산하므로써 그림 6의 (b)와 같은  $T_{dila}(x,y)$ 를 구한다. 이렇게 얻은  $T_{dila}(x,y)$ 은 문자의 폭을 가진 형태소로 팽창연산을 하였기 때문에 문자의 수직 경계선의 양쪽이 다시 흰색으로 칠해지게 되어 글자부분을 이루게 된다.

$$chflg = \begin{cases} 1 \text{ (white)} & \text{if } T_{mean} > F_{mean} \\ 0 \text{ (black)} & \text{if } T_{mean} < F_{mean} \end{cases} \quad \dots(7)$$

$$\text{where } T_{mean} = \frac{\sum T_{ch}}{N_{T_{ch}}}, F_{mean} = \frac{\sum f(x,y)}{N_f}$$

$$, T_{ch} = \{f(x,y) | T_{dila}(x,y) = 1\}$$

$$\text{for } 0 \leq x \leq m-1, 0 \leq y \leq n-1$$

이제 문자색의 판별은  $T_{dila}(x,y)$ 에서 흰색을 나타내는 화소들의 위치에 있는 원 영상  $f(x,y)$ 의 화소값들의 평균  $T_{mean}$ 과  $f(x,y)$ 의 전체 평균  $F_{mean}$ 을 비교하여 문자색이 밝은 색인지 또는 어두운 색인지를 식 (7)과 같이 판별하게 되며 차후 연산의 일관성을 위하여 만약 문자가 어두운 색이라면 밝은 색으로 반전시켜준다.

### 3.3 이진화

문자색과 바탕색을 판별한 후에 입력영상을 두가지 방법을 써서 각각 변환하는데 첫 번째는  $W=3 \times 3$  마스크를 이용한 평균 필터링(mean filtering) 방법이고 두 번째 방법은 히스토그램 평활화(histogram equalization) 방법인데, 평균 필터링의 역할은 입력영상에서 잡음에 해당하는 즉, 주변에 비해 높은 화소값이나 낮은 화소값을 평균화하므로써 잡음을 제거해주는 효과가 있다. 또한 히스토그램 평활화 방법은 빈약한 명암값 분포를 가진 영상을 일정한 분포를 가진 히스토그램으로

재구성해 주므로써 열화된 영상을 전체적으로 개선시켜 주는 효과가 있다.

본 연구에서 평균 필터링을 통하여 얻은 영상  $f_m$ 은 임계값을 적용하여 이진화하는 단계에서의 입력 영상으로 쓰이며 히스토그램 평활화를 통하여 얻은 영상  $f_h$ 는 이진화할 영역 분할을 위한 입력 영상으로 쓰인다.

두 가지의 영상처리 작업이 끝나면  $f_h$  영상의 각 세로줄(column)별로 DCT값을 구한다. 일반적으로 DCT를 이용하여 얻은 행렬  $H_{uv}$ 에서 행렬의 원점인  $H_{00}$ 는 직류(dc)성분을 가지고 있고 나머지는 교류(ac)성분을 가지고 있다. dc계수는 입력영상의 전체 평균을 의미하기도 하며 ac계수는 원점에서 멀어질수록 저주파에서 고주파의 성질을 가진다.

본 연구에서는 이와 같은 성질을 이용하여 영상의 각 세로줄별로 얻은 DCT값의 dc계수의 분산(variation)인  $dc_{var}$ 를 구하여 일정값  $T_{var}$ 이하이면 입력 영상의 컨테이너 식별자 영역이 일반 종이문서와 같은 배경의 명암이 균일한 형태이고,  $T_{var}$  이상이면 컨테이너 식별자 영역이 세로로 굴곡이 있거나 빛과 기타 다른 외부조건에 의해 왜곡된 영상이라고 판단을 하여, 전자인 경우 기존에 문서영상 등에 쓰였던 방법과 마찬가지로 광역 임계법의 하나인 Otsu 이진화 방법을 사용하여 영상 전체를 이진화 하고, 후자인 경우는 제안한 알고리즘을 통하여 빛의 명암도에 따라 영역을 분할하여 각 분할된 영역의 상황에 적합한 임계값을 설정하여 이진화하는 방법을 제안하였다. 여기서  $T_{var}$ 는 실험값에 의해 2.23으로 설정하였다.

만약 원 영상을 히스토그램 평활화한 크기  $M \times N$  입력 영상  $f_h$ 가 전체 광역 임계법에 의해서는 제대로 이진화가 되지 않는다고 판단되면 즉,  $T_{var}$ 값이 2.23이상인 경우에는 전 단계에서 구한 각 세로줄별 DCT값  $H$ 를 이용하여 식 (8) ~ 식 (12)에 따라 영역을 분할한다.

$$dc(j) = H_0(j), 0 \leq j < N \quad \dots(8)$$

$$ac_{max}(j) = \arg \max |H_k(j)|, 0 < k < M \quad \dots(9)$$

$$dst_{ad}(j) = dc(j) / ac_{max}(j) - 4dc(j) \quad \dots(10)$$

$$chflg(j) = \begin{cases} 1 & \text{if } dst_{ad}(j) < 0 \\ 0 & \text{otherwise} \end{cases} \quad \dots(11)$$

식 (11)은 입력된 세로줄 영상안에서 크게 변화되는 화소를 발견하지 못하면 배경으로 인식해서  $chflg$ 를 0으로 주고 그렇지 않으면 세로줄에 글자로 추정되는 화소가 포함되어 있는 것으로 판단하여  $chflg$ 값을 1로 설정한다. 이와 같이  $chflg$ 값이 0으로 결정된 영역은 임계값

설정과정에서 배경으로 판단하여 임계값을 설정하지 않고 다음영역으로 넘어가기 때문에 연산의 시간을 단축시키는 효과도 부수적으로 얻을 수 있다.

$$bgflg = \begin{cases} 1 & \text{if } dc(j) > dc_{mean} \\ 0 & \text{otherwise} \end{cases} \quad \dots(12)$$

, where  $dc_{mean} = \frac{1}{N} \sum_{j=0}^{N-1} dc(j)$

식 (12)는 *chflg*값이 1인 경우 즉, 문자가 포함되어 있다고 판정된 영역에서 과연 이 문자가 빛이나 기타 왜곡으로 인해 다른 영역의 문자보다 어느 정도 명암도 차이를 가지고 있는가를 dc계수들의 전체 평균과 각 세로줄의 dc계수를 비교하여 다른 영역보다 밝은 부분이면 *bgflg*를 1로 설정하고 어두운 경우 *bgflg*를 0으로 하여 영역을 나누게 된다.

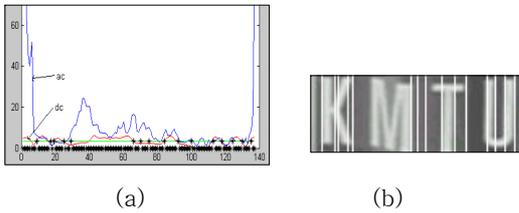


그림 7. 각 세로줄 별 dc계수와 최대 ac계수

그림 7의 (a)는 각 세로줄별 dc계수와 ac계수중 최대값을 히스토그램으로 나타낸 것이다. 입력영상  $f_h$ 의 dc값이 2.8157 이기 때문에  $T_{var}$ 를 넘었으므로 영역분할이 이루어지는데 그림에서 '\*'표로 표시된 것은 앞에서 언급한 영역 분할을 위한 식 (8) ~ (12)에 의해 결정된 영역의 분기점이며 그림 7의 (b)는 분기점에 의해 나누어진 영역들을 표시한 영상이다.

DCT를 이용하여 명암도별로 이진화할 영역을 분할한 후에 각 영역에 적합한 임계값을 설정해 주기 위해 각각의 분할된 영역을 Otsu 방법으로 설정한 임계값과 전체 영상의 명암도를 고려한 임계값중 최대값을 선택하므로서 배경의 잡음에 덜 민감하도록 임계값을 설정하는 방법을 제안하였다.

영역 분할을 위해서 원영상을 히스토그램 평활화한 영상을 사용하였지만, 이진화 과정에서는 원영상을 평균 필터링한 영상  $f_m$ 을 사용하기로 한다. 이는 앞에서도 언급하였듯이 잡음에 강하므로 잘못된 정보로 인해 계산과정에서 생기는 오차를 줄여줄 수 있다.

$$T_G = \max \{ T1_G, T2_G \} \quad \dots(13)$$

$$T1_G = Otsu(f_m)$$

$$T2_G = (M_{G_{mean}} + M_{G_{min}})/2 + (T1_G - M_{f_m})/2$$

식 (13)은 제안한 임계값  $T_G$ 의 설정방법이며, 여기서  $T2_G$ 를 보면 분할된 영역  $G$ 의 각 세로줄별 평균값의 평균과 각 세로줄별 최소값의 평균의 중간값을 선택하는

데 그 이유는 각 세로줄별 최소값은 문자와 상관없는 배경영역이므로 명암도의 변화를 가장 잘 나타내어 주며 각 세로줄의 평균은 문자와 배경의 정보를 모두 포함한 명암도의 변화를 대변해 주므로 이 두값의 중간값은 각 세로줄이 전체영역에서 위치하는 명암도의 정도를 알 수 있다.  $T1_G$ 는 최초 입력된 전체 식별자 영역  $f_m$ 을 Otsu의 방법으로 얻은 임계값이며,  $M_{f_m}$ 는  $f_m$ 의 각 세로줄별 최소값의 평균이다. 이들의 차이를 임계값에 더해 주는 이유는 앞에서 결정된 중간값  $(M_{G_{mean}} + M_{G_{min}})/2$ 를 임계값 수준으로 올리려는 의도이다.

그림 8의 (a)는 그림 7(b)의 영역별 임계값을  $T1_G$ 로만 주었을 경우의 결과이며 그림 8의 (b)는  $T_G$ 를 임계값으로 선택한 결과이다.



그림 8. 영역별 임계값에 따른 결과

그림 8의 (a)에서 보면 부위별 명암도 차이를 고려하지 않고 전체 임계값으로 이진화 함으로써 문자를 포함하고 있다고 판단한 영역은 문자와 비슷한 밝기의 배경 일부 화소도 문자로 판단하였으나 그림 8 (b)의 경우 전체 영상의 명암도와 비교한 값중에 최대값을 설정하였을 때는 문자만을 정확히 이진화 할 수 있었다.

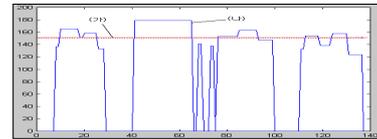


그림 9. 영역별 임계값

그림 9는 그림 7 (b)의 영역별 임계값을 나타낸 것으로서 150정도를 갖는 흐린 직선(가)는 그림 8 (a)의 임계값이며 진한 선(나)는 그림 8 (b)의 임계값인  $T_G$ 값으로서 각 영역별로 적절하게 임계값이 설정됨을 알 수 있다.

### 3.4 문자 분할

마지막으로 이진화된 문자를 3.1절의 식별자 영역 추출 방법과 유사하게 그림 10의 (a)와 같은 수직 히스토그램을 이용하여 문자가 있는, 즉 흰 화소들이 존재하는 영역을 분할하면 그림 10의 (b)와 같이 개별 문자로 분할된다.

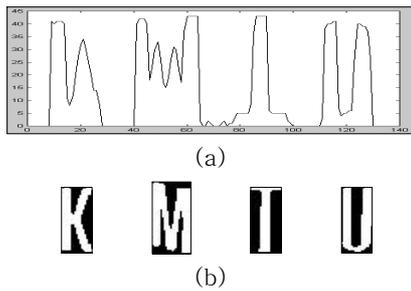


그림 10. 히스토그램과 분리된 개별 문자

#### 4. 실험 및 결과 분석

제안한 이진화 방법의 실험을 위해 실제 항만에서 수집한 다양한 컨테이너 영상을 가지고 수행하였다. 실험 환경은 PentiumIII-500MHz를 주컴퓨터로 사용하였고 메모리 128Mbyte, Windows98 환경에서 Matlab5.0으로 구현하였다. 최초 입력영상은 24비트 컬러 영상이나 처리 속도의 향상을 위하여 256레벨의 명암영상으로 변환하였다.

그림 11의 (a)는 최초 입력된 컨테이너 영상이며 그림 11의 (b)는 입력영상을 앞서 설명한 방법을 사용하여 식별자 영역만을 추출해 낸 것이다.



그림 11. 입력 영상과 추출된 식별자 영역

그림 11(a)의 운송회사 코드(영문 식별자)는  $dc_{var}$  값이 2.6967로서  $T_{var}$ 보다 큰 값이므로 제안한 방법에 의해 처리가 되어야 한다. 반면에 그림 11(b)의 일련번호(숫자 코드)는  $dc_{var}$  값이 1.4153으로서  $T_{var}$ 보다 작으므로 전체 영상을 Otsu의 이진화 방법으로 이진화하였고 그 결과는 각각 그림 12의 (a)와 (b)이다.

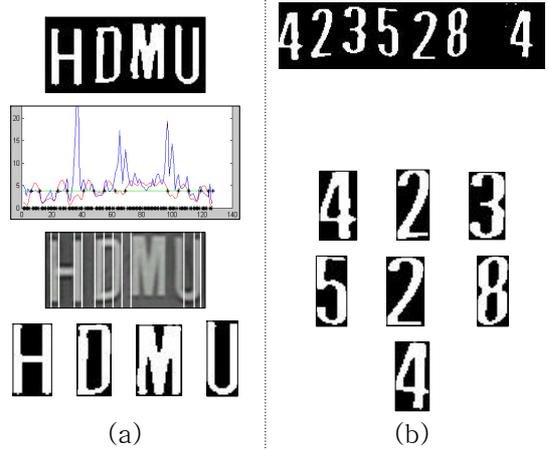


그림 12. 이진화 및 분할 결과

그림 12의 경우를 보면 같은 컨테이너 영상이라 할지라도 컨테이너 표면의 굴곡으로 인한 명암도의 차이등으로 인하여 각각 다른 전처리 방법이 요구됨을 알 수 있었으며 이와 같은 방법으로 실험에 쓰인 대부분의 영상에서 식별자의 문자들을 효율적으로 이진화하여 개별 분리 할 수 있었다.

#### 5. 결론

본 연구는 효율적인 문자 인식이 이루어질 수 있도록 하기 위한 전처리 과정으로서 경계선 검출과 형태학적 연산을 이용하여 운송 컨테이너 영상의 식별자 영역인 문자열 영역만을 추출한 후에 차후 연산의 일관성을 제공하기 위하여 식별자 영역의 문자색과 바탕색을 판별하였다. 이진화 단계에서는 기존의 이진화 알고리즘으로 쉽게 이진화 할 수 있는 영상은 Otsu 알고리즘을 그대로 적용을 하고 그렇지 못하는 영상에 대해서는 제안한 알고리즘을 적용하도록 하였다. 제안한 이진화 알고리즘은 DCT를 적용하여 명암도의 변화에 따른 영역을 효율적으로 분할하여 각각의 영역에 맞는 임계값을 설정할 수 있도록 하였다.

제안한 방법을 사용하였을 경우 속도면에서도 일반적인 이진화 방법으로 처리가 되는 영상들에 대해서는 아무런 속도제약을 받지 않으며 영역분할이 불가피한 영상들에 대해서만 선별적으로 할 수 있다. 따라서 전체적인 처리시간으로 볼 때에 기존의 방법들과 비교하여 속도의 차이를 전혀 고려하지 않고도 훨씬 만족스러운 결과를 얻을 수 있었다.

특히 제안한 방법은 세로 굴곡으로 인하여 빛의 명암도 차이가 심한 경우의 컨테이너 영상들에서는 월등한 전처리 효과를 확인할 수 있었으며 컨테이너 이외의 다른 상황에서도 매우 효과적으로 사용될 수 있으리라고 확신한다.

앞으로의 연구 과제는 분할된 개별 문자의 세선화와 신경망등을 이용한 효율적이고 정확한 인식 방법을 추가하는 것이다.

## 참고 문헌

- [1] D. M. Etter, *Engineering Problem Solving with MATLAB*, Prentice Hall International Editions, New Jersey, 1993.
- [2] *Freight Containers - Coding, Identification and Marking* [ISO 6346 1995(E)].
- [3] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, pp. 518-528, 1992.
- [4] K. R. Rao, P. Yip, *Discrete Cosine Transform*, Academic Press, 1990
- [5] N. Otsu, "A threshold selection method from grey level histograms," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 8, pp.62-66, 1978.
- [6] 최재영, 김낙빈, "운송 컨테이너 식별자 영역의 이진화에 관한 연구," 한국멀티미디어학회 추계학술 발표논문집, 제2권 2호, pp.173-178, 1999.