

# 한글 문서에서의 낱자 분리 알고리즘

이 현 표, 양 순 성, 황 교 철, 이 균 하  
인하 대학교 전자계산학과

## An Algorithm for Character Segmentation on the Korean Documents

Hyun Pyo Lee, Soon Sung Yang, Kyo Cheul Hwang, Kyoona Ha Lee  
Dept. of Computer Science, Inha University

### 요 약

본 논문에서는 한글 문자 인식 시스템을 이용하여 한글 문서를 자동 판독하기 위한 낱자 분리 알고리즘을 제안하였다. 제안된 알고리즘은 문자의 심미성 및 가독성을 높이기 위하여 상호 접촉되고 있는 낱자들의 유형을 조사 분석한 후 문서의 수직 및 수평 투영법에 기본을 두어 행과 열을 분리하고, 분리된 문자군을 이루고 있는 문자들의 기하학적 속성들을 분석하여 낱자를 분리하였다.

### I. 서 론

컴퓨터의 주기능에 속하는 중앙 처리 장치와 기억 장치는 하드웨어 기술의 발전으로 크게 진보되었으나, 정보의 입력은 키보드를 이용하는 수작업에 대부분 의존하기 때문에 병목 현상을 초래하고 있다. 또한, 기존의 정보가 거의 대부분 문서로 작성되어 있기 때문에 이를 자료화하는데 보다 효율적이고 신뢰성 있는 방법이 요구되어 1960년대 이후 문자 인식에 관한 연구가 활발하게 진행되었다. 현재 영문의 경우에는 인쇄체 영문서나 제한된 형태의 필기체를 인식하는 시스템이 실용화되었다[2,4]. 우리나라에서도 최근에야 비로서 한글 인식에 관한 연구가 활발히 진행되어 현재에는 실용화를 위한 문자 인식 시스템의 개발을 목표로 연구가 진행중이다[8-11].

일반적으로 문자 인식 시스템은 스캐너 또는 카메라를 통해서 얻어진 문서에 대한 영상을 입력으로 하고, 문서 영상으로부터 문자 단위의 영상을 추출하는 단계를 거쳐 추출된 영상을 인식하는 단계로 이루어지는 것이 기억 용탕이나 처리 시간 및 절차면에서 간편하다. 따라서 문서에서 문자 단위의 영상을 추출하는 작업이 필요하며

전처리 과정의 커다란 과제로 대두되었다.

일반적으로 문서에서 문자열의 방향에 따라 종서 혹은 횡서로 나누나 현대 문서의 대부분은 횡서로 되어 있는 문서가 주종을 이루므로 횡서로 된 문서에서 낱자를 분리하는 작업이 필요하다. 횡서로 되어있는 문서에서 개개의 낱자를 분리해 내는 작업은 우선 문자행을 검출한 후에 대상 문자행 내에서 문자 단위의 낱자를 분리하는 방법이 일반적인 방법으로 이용되고 있다. 국정 교과서와 같이 문자 사이의 간격이 뚜렷한 경우는 단순한 수직 및 수평 투영법을 사용하면 쉽게 낱자를 분리할 수 있다.

사람의 시각으로도 판별이 어려울 만큼의 경계가 모호한 문서에서는 문자 단위의 낱자 추출이 어려운 경우가 많이 있다. 더우기 일상적인 정보 전달의 수단인 일반대중잡지, 서적류등의 한글에서는 심미성 및 가독성을 높이기 위해서 낱자들이 서로 접촉 또는 밀착되어 있는 경우가 두드러지게 많아서 낱자 분리를 어렵게 하고있다. 이러한 문제들을 해결하고 성공적인 문서 인식 시스템을 구현하기 위해 경계가 모호한 낱자들의 경계를 구분하여 문서 영상으로부터 문자 단위의 낱자를 추출하는 작업이 필요하다.

영문의 경우에는 서로 접촉된 알파벳을 분리하는

알고리즘이 많이 연구되어 왔다[2,3]. 영문자는 기본 자소의 조합 방법에서 일차원적 직선 배열로 조합되나, 한글 문자는 이차원적 배열로 조합되므로 영문에서 이용된 알파벳 분리 알고리즘을 한글에 적용시키는 것은 적합하지가 않았다. 조합 방법에서 이차원적 구조를 갖는 한자의 경우 한자의 특징을 면밀히 조사해서 해결한 연구 발표가 있으나[4-7], 한글과는 또 다른 면을 갖고 있어 한글에 적용할 수가 없었다. 따라서 한글 문자를 대상으로한 구체적인 연구가 필요하다.

본 논문에서는 일반적인 한글 문서인 잡지류 및 서적에서 서로 접촉된 문자의 유형을 조사 분석하고, 접촉된 문자의 구조적인 속성을 검출하여 낱자 분리에 관한 알고리즘을 제시하였다.

II. 낱자 분리에 관한 고찰

낱자 분리의 일반적인 방법은 문자의 구조적인 특성상 수평 방향으로 투영(projection)하여 문서상의 문자행을 검출하고, 검출된 문자행 내에서 수직 투영법을 이용하여 개개의 낱자로 분리해 내는 방법이 일반적으로 많이 쓰이고 있다[2-7].

1. 문자행의 검출

사람의 시각으로 문자행을 검출하는 과정은 연속되어 있는 낱자의 흐름으로 문자행을 검출하게 된다. 이 경우에 문자행과 문자행 사이에 행간이라는 공간이 있다는

- 모형물을 조립하는 당사자는 그 작업이 그저 장난에
- 그치는 것이 아니라, 그 나름대로 여가를 활용하는 방
- 법이며, 건전한 취미생활이고, 그저 공부만 하기를 바
- 라는 환경 속에서 받게 되는 스트레스를 해소하는 하나
- 의 탈출구가 된다는 주장이다.

(a)

- 모형물을 조립하는 당사자는 그 작업이 그저 장난에
- 그치는 것이 아니라, 그 나름대로 여가를 활용하는 방
- 법이며, 건전한 취미생활이고, 그저 공부만 하기를 바
- 라는 환경 속에서 받게 되는 스트레스를 해소하는 하나
- 의 탈출구가 된다는 주장이다.

(b)

그림 1 문서에서 문자행의 검출

- (a) 성공적인 문자행 검출
- (b) 문자행 검출의 실패

일반적인 상식에 기초하여 문자행을 식별하게 된다. 따라서 문서에서 문자행의 검출은 행간에 기초하여 문서 영상을 수평 방향으로 투영하여 문자행을 검출할 수 있다. 그림 1의 (a)는 수평 투영법을 이용하여 성공적으로 검출된 문자행과 수평 투영절을 보여주고 있다. 그림 1의 (b)는 문자행 검출이 안된 경우를 보여주고 있다. 이와 같은 경우에 일반적인 문서에서 문자행 검출이 가능한 입력 각도는 행간과 문자행의 폭에 따라 달라지나, 실험적으로 최대 허용 가능한 입력 각도는 1° 이내이고 문자행의 폭이 큰 경우에는 A4용지를 기준으로 하였을 경우 0.7° 이내의 허용 각도를 갖고 있다. 그림 1의 (b)와 같은 경우에 입력된 원시 문서 영상을 회전하는 방법도 생각할 수 있으나, 처리 시간 및 원시 영상의 손상을 갖는 문제점이 있다.

2. 문자들의 검출

분리된 임의의 문자행에 대해 수직 방향으로 투영하면 그림 2에서와 같이 대부분의 문자는 자간이라는 문자와 문자 사이에 공간이 존재하여 기본 문자들을 분리해낼 수 있다. 분리된 기본 문자들은 문자행내에 포함되어야 한다. 그림 2에서 임의의 문자들에 대해서 투영폭( $W_k$ )는 식 (1)에 의해 구해지며,  $F_{x1}$ 은 문자들의 X-축 좌측 좌표이고  $F_{x2}$ 는 우측 좌표이다.

$$W_k = (F_{x2} - F_{x1} + 1) \text{ ————— (1)}$$

$F_{x1}, F_{x2}$  : X-축 좌표

$k$  : 점자

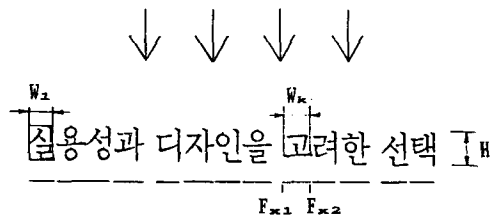


그림 2 기본 문자를 검출

보통 인쇄체 문자의 경우 조형적인 변형을 주지 않는 한 문자의 가로와 세로의 비는 거의 1:1.2를 갖고 있다. 따라서 각 문자행의 세로 길이(H)의 1/1.2배 값을 문자 크기(S)로 추정할 수 있다. 각각의 문자행에 대해 하나의 문자 크기를 결정한다는 것은 문자 크기가 다른 문자행에 대해서도 적용될 수 있도록 해준다.

3. 문자의 접촉

그림 2에서 '용성'과 같이 일반적인 지식에 기초하여

사람의 시각으로는 문자의 구별이 용이하나, 문자의 접촉으로 인해 수직 투영을 하였을 경우 두개의 문자가 하나의 투영 문자로 나타나게 된다. 수직 투영만으로는 문자의 완전한 분리가 불가능하다.

영문의 경우에도 문자와 문자가 접촉하는 유사한 현상이 있다. 영문 인쇄체에서 접촉된 문자를 분리하는 연구는 다음과 같이 두개의 범주로 나눌 수 있으며, 그림 3은 LAG에서 분리하는 예[2]를 보여주고 있다. 그림 3의 (a)는 전형적인 세리프(serif) 접합의 예이고, (b)는 더블-0 접합의 예이다. 그림 3에서 중앙에 위치한 그래프는 수직 투영법을 실시한 히스토 그래프이며, 아래 부분에 위치한 그래프는 절단 함수  $V(x-1) - 2 \times V(x) + V(x+1)$ 를 적용한 그래프로서 절단 함수를 적용하였을 때의 절단점을 보여주고 있다. 영문자와 같이 조합 방법이 일차원적 직선 배열로 되어있고 수직 성분끼리의 가벼운 세리프 접촉일때 절단점 함수를 이용 분리 가능하다. 그러나 한글 문자의 조합 방법은 영문자의 경우와는 달리 이차원적 구조로 조합되며, 그림 2에서 '용성'의 경우와 같이 수평 성분끼리의 접촉이기 때문에 적용할 수 없다.

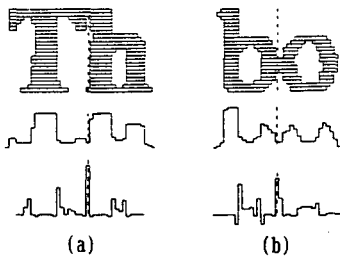


그림 3 LAG에서의 접합 문자 분리  
(a) 전형적인 세리프 접합의 분리  
(b) 더블 - 0 접합의 분리

영문서에서 합자 처리의 또 다른 예는 원형 비교 방법(template matching method)으로 그림 4에서와 같이 합자군을 조금씩 나누어서 원형 비교하는 것으로 남아있는 부영역이 존재하지 않고, 원형 비교가 이루어 질때까지 순환적으로 비교하는 방법이다. 영문의 경우 대,소문자 52개만을 비교하여야 하는데 반해 한글의 경우에는 조합 문자라는 특성상 비교의 횟수가 방대하다. 더우기 문자에 조형적인 변형을 가할 경우 또 다른 처리를 하여야 하는 문제점이 있어 순환적 원형 비교 방법을 그대로 적용할 수는 없었다.

### III. 한글 문자의 접촉 유형 및 분석

한글 낱자의 접촉 유형 조사를 위해서 일반 대중 잡지 5종류 85면 89,780자의 문서 표본을 선택하였으며, 이미지 스캐너를 통하여 문자 입력을 하였다. 접촉된 문자 유형 및 문자수는 컴퓨터를 이용하여 조사 분석하였다.

#### 1. 접촉 문자의 식별

수직 투영을 하여 얻은 문자열의 폭이 문자의 높이 즉, 2장 2절에서 얻은 문자의 크기(S)보다 크다고 하면 두개 이상의 문자가 서로 접촉 혹은 밀착되어 있다는 것을 알 수 있다. 또한, 문자군의 폭이 문자 크기의 약 1/2정도이면 좌,우 문자군에서 접촉 혹은 분리 문자가 있는 것으로 판별되며 다음과 같은 알고리즘으로 분류하였다.

```

if ( Wk > S )
    if ( Wk+1 < S/2 )
        두개 이상의 문자 접촉 and 뒷문자의 자소 분리
    else 두개 이상의 문자 접촉
if ( Wk ≈ S/2 )
    if ( Wk+1 > S )
        앞자소 분리 and 뒷자소와 다음 문자군의 접촉
    else if ( Wk+1 ≈ S/2 )
        분리된 문자
        else 문장 부호
else , 잡음(noise)
    
```

#### 2. 한글 문자의 접촉 및 분리 유형

접촉 혹은 밀착된 유형의 조사는 투영 방식을 이용하여 문서에서 문자행을 검출하고, 검출된 문자행에서 투영폭(W<sub>k</sub>)을 문자 크기(S)와 비교하여 접촉 유,무를

Inout Pattern	Windowed Input	Matching Prototype 1	Residue	Matching Prototype 2
mm	m	m	l	
	n	n	n	
	n	o	m	
	r	r	m	m

그림 4 순환적 원형 비교 방법에서의 접합 문자 분리

판별하였다. 표 1은 한글 문자의 자획이 서로 접촉 혹은 밀착된 유형을 파악하기 위하여 조사한 결과이다. 네모를 공간에서 다소 벗어나는 자획 성분을 갖고 있는 'ㄷ', 'ㅌ' 등의 수직 모음과 'ㅇ', 'ㅛ', 'ㅠ', 'ㅠ', 'ㅡ' 등의 수평 모음에 중점을 두어 조사하였다. 중성이 있는 경우에는 자획의 변형에 따라 접촉의 유형이 달라지므로 중성유, 무로 나누어 조사하였다. 표1에서와 같이 수평 성분이 있는 문자에서의 접촉이 많았으며, 특히 앞문자의 'ㄷ'선분과 뒷문자의 'ㅡ'선분의 접촉이 가장 많았다.

표 1 한글 문자의 유형 조사 비교표

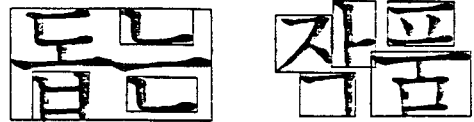
한글자 중성유	접촉 유형												소성자획 ㄱ, ㅋ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ, ㆁ
	중성없는 모음						중성있는 모음						
	ㄱ	ㅋ	ㆁ	ㆁ	ㆁ	ㆁ	ㄱ	ㅋ	ㆁ	ㆁ	ㆁ	ㆁ	
ㄱ	중성유 320	15	142	45	54	101	10	95	5	815		738	
	중성무 141	25	28	0	111	63	13	95	0	525		626	
ㅋ	중성유 3	0	0	0	0	1	0	11	0	4		18	
	중성무 17	0	1	0	8	11	2	3	1	29		31	
ㆁ	중성유 53	4	7	0	28	9	0	8	1	41		71	
	중성무 15	8	6	1	18	8	4	6	0	88		37	
ㆁ	중성유 10	1	5	1	8	1	0	8	33	30		29	
	중성무 2	0	1	0	1	1	0	1	1	6		5	
ㆁ	중성유 33	5	24	1	3	23	1	31	0	57		45	
	중성무 10	6	19	0	9	22	2	10	0	116		40	
ㆁ	중성유 0	0	0	0	1	3	0	2	0	6		4	
	중성무 0	0	1	0	0	0	0	0	0	7		1	
ㅡ	중성유 202	1	3	1	7	4	0	6	0	49		40	
	중성무 22	0	8	0	13	10	5	6	0	144		88	

한글 문자는 기본 자소가 최소한 하나 이상의 자음과 모음으로 구성되어 문자를 이루고 있다. 기본 자소가

수직 모음이고 받침이 없는 경우, 수직 투영을 하면 하나의 문자가 두개의 영역으로 나누어지게 된다. 수직 투영을 하여 분리되는 문자와 두개의 문자가 접촉되는 문자 영역의 형태에 따라 6가지 형태로 분리하여 그림 5에 나타내었다.

IV. 날자 분리 알고리즘

한글 문자의 구조는 이차원적 구조로 구성되어 중성이 없고 기본 모음이 'ㄷ', 'ㅌ', 'ㄷ' 등을 포함하지 않는 한 수직적으로 중첩되어 상호 연관 관계를 형성하므로, 접촉된 문자에 대해 문자의 자획을 그림 6에서와 같이 상호 연결된 선분에 대해서 연결획 평면을 만든다.



(a) (b)

그림 6 접촉된 문자의 연결획 평면

접촉된 문자의 연결획 평면 알고리즘은 다음과 같다.

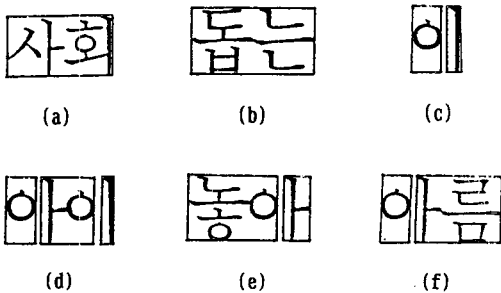


그림 5 접촉 문자의 6가지 형태

- (a) 두개의 문자들의 무영적 접촉
- (b) 두개의 문자들의 접촉
- (c) 하나의 문자들이 두개의 반문자들로 분할
- (d) 서로 다른 두개의 문자를 상호간의 접촉
- (e) 앞문자의 반문자들과 뒷문자들의 접촉
- (f) 앞문자들과 뒷문자의 반문자들의 접촉

- 단계 1. 임의의 문자 자획의 수직 및 수평 선분의 구획 RS(k)을 정한다.
- 단계 2. 임의의 선분 구획 RS(k)와 RS(k+1)의 연결성을 조사하여 상호 연결된 선분을 하나의 연결획 평면 RS(k)로 병합한다.
- 단계 3. 단계 1,2를 모든 선분에 대해 적용한다.
- 단계 4. 연결획 평면 RS(k)에 대해 수직 중첩도를 계산한다.
- 단계 5. 임의의 연결획 평면 RS(k)와 RS(k+1)의 수직 중첩도가 일정한 상수  $C_k$  이상이면, 두개의 연결획 평면을 다시 하나의 연결획 평면으로 병합시킨다.
- 단계 6. 모든 연결획 평면에 대해 단계 5를 적용한다.

연결획 평면 구획을 모든 획에 대해 적용한뒤 두개 이상의 연결획 평면이 존재하면, 그림 5의 (a)에서와 같이

두개의 문자가 투영적 접촉을 이룬 문자이다.

유형 1. 연결획 평면  $RS(k)$ 의 폭  $RW(k)$ 과 문자 크기( $S$ )와의 상호 관계를 고려하여 연결획 평면  $RS(k)$ 와  $RS(k+1)$ 의 합자 처리 또는 연결획 평면  $RS(k)$ 와  $RS(k+1)$ 의 중간 부분을 절단점( $P$ )로 하여 분리한다.

문자획의 연결성을 조사하여 구획  $RS(1)$ 만이 존재하는 단일 구획이면, 자획의 선분이 접촉된 문자로 판명되어 다음의 과정을 적용한다.

유형 2. 연결획 평면  $RS(k)$ 의 폭  $RW(k)$ 이 투영폭( $W_k$ )와 같은 값이고, 연결획 평면의 폭  $RW(k)$ 가 문자 크기  $S$ 보다 작은 경우 그림 5의 (b)와 같은 유형이다.

- 1) 연결획 평면폭  $RW(k)$ 를 문자 크기  $S$ 로 나누어 절단점 후보 위치( $P$ )로 정한다.
- 2) 수평 방향으로  $P \pm d$ 범위를 수직 방향으로 투영하여 히스토 그래프를 작성한다.
- 3) 작성된 히스토 그래프에서 가장 작은 수치의 최우측 값을 갖는 위치를 절단점 위치로 정하여 문자를 분리한다.

유형 3. 투영폭( $W_k$ )이 문자 크기( $S$ )의 약 1/2 정도의 투영폭이 연속해서 나타나면, 그림 5의 (c)와 같이 하나의 문자가 두개의 반문자들로 분리되는 경우이다. 두개의 반문자들을 합성하여 하나의 문자들로 만든다.

유형 4. 연결획 평면폭  $RW(k)$ 가 투영폭( $W_k$ )과 같으며, 문자 크기( $S$ )와 같은 경우 그림 5의 (d)와 같이 서로 다른 문자를 상호간의 접촉이다.

- 1) 연결획 평면폭  $RW(k)$ 를 문자 크기  $S$ 의 1/2로 나누어 절단점 후보 위치( $P$ )로 정한다.
- 2) 유형 2의 2), 3)을 수행하여 반문자들과의 합자 처리를 한다.

유형 5. 연결획 평면폭  $RW(k)$ 가 투영폭( $W_k$ )과 같고, 문자 크기( $S$ )보다 크며, 반문자들이 앞에 존재할 경우 그림 5의 (e)와 같이 앞문자의 반문자들과 뒷문자들의 접촉이다.

- 1) 연결획 평면  $RW(k)$ 를 문자 크기  $S$ 로 나누어 연결획 평면의 시작점 부터의 위치를 절단점 후보 위치( $P$ )로 한다.
- 2) 유형 2의 2), 3)을 수행하여 앞의 반문자들과의 합자 처리를 한다.

유형 6. 연결획 평면폭  $RW(k)$ 가 투영폭( $W_k$ )과 같고, 문자 크기  $S$ 보다 크며, 반문자들이 뒤에 존재할 경우 그림 5의 (f)와 같이 앞문자들과 뒷문자의

반문자들의 접촉이다.

- 1) 연결획 평면폭  $RW(k)$ 를 문자 크기  $S$ 로 나누어 연결획 평면의 끝점으로부터의 위치를 절단점 후보 위치( $P$ )로 한다.
- 2) 유형 2의 2), 3)을 수행하여 뒤의 반문자들과의 합자 처리를 한다.

## V. 실험 및 결과

개별 문자 분리 및 접촉 문자 분리 실험을 위한 문서 영상 입력은 가로,세로 300 dpi의 해상도를 가지는 이미지 스캐너를 사용하였다. 실험에 사용된 문서는 본문 3장에서 한글 문자의 접촉 유형을 조사하는데 사용된 일반 대중 잡지 5종류 85면 89,780자의 문서를 그대로 사용하였으며, 문자 크기는 대체로 대중 잡지 활자 크기인 13급 (3.25mm)이었다. 이미지 스캐너를 이용하여 이진 데이터의 원시 문서 영상을 얻은 후 수평 투영을 하여 문서영상의 수평 투영절을 얻어 문자행을 추출하였다. 추출된 문자행에 대하여 수직 투영을 하여 수직 투영절을 얻었으며, 수직 투영절의 폭과 문자 크기와의 비교를 통하여 본문에 제시된 알고리즘에 따라 접촉 문자를 식별하였다. 접촉 문자 및 분리 문자에 대해서는 그림 7과 같은 계통도에 따라 처리하였다.

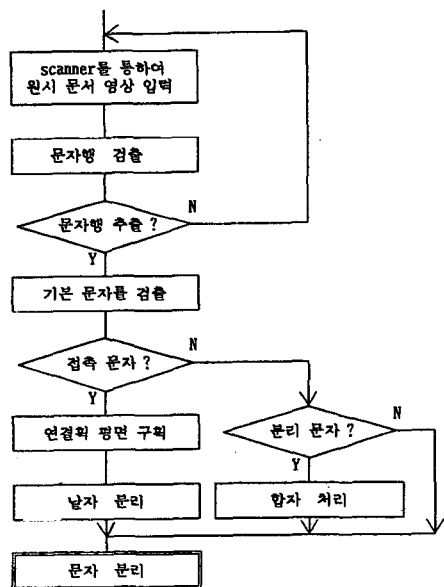


그림 7 날자 분리의 전체 계통도

접촉 문자의 비율은 표본에 따라 다르겠지만 본 실험의 경우에는 총 89,780자중 8,810개가 접촉 문자로 나타나 전체 문자의 9.78%를 차지하였다. 이들중 투영적으로만 접촉을 이룬 경우는 접촉된 경우의 20.77%인 1,830개 이었으며, 문자획에 의한 접촉은 (b), (c), (d), (e), (f)유형이 각각 30.39%, 36.68%, 0.15%, 2.67%, 9.34%로 모두 분리 가능하였다. 그러나 추가 자료에 대한 실험에 의하여 그림 8과 같이 유형 (a)-(f)이외의 특이하게 접촉된 유형이 나타났다. 이런 경우는 유형 (e)와 (f)가 반씩 결합된 유형으로서 유형 (e)가 전체 문자의 0.26%로 나타났고, 유형 (f)가 0.91%로 나타났기 때문에 두개의 유형이 동시에 나타난 결합된 유형은 전체 문자의 0.26%보다 훨씬 적게 나타날 것으로 생각되며 커다란 문제는 아니었다.

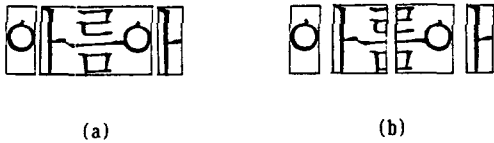


그림 8 잘못된 문자 분리의 예

VI. 결 론

심미성 및 가독성을 높이기 위해서 접촉 또는 밀착되어 있는 문자는 단순 투영에 의한 문자 분리가 곤란하였으며, 이러한 접촉율이 한글 문서에서는 9.78%로 나타났다. 이들중 20.77%인 투영적 접촉을 이루는 문자의 분리는 연결획 평면 구획을 이용하여 쉽게 해결할 수 있었다. 문자획이 접촉하는 문자의 경우는 5가지 유형으로 분류하여 처리 해결할 수 있었다. 문서의 특성상 유형 조사에 사용된 89,780자의 표본에 접촉의 모든 유형이 포함되어 있다고는 볼수 없으나 여기에 포함되지 않은 유형들은 극히 낮은 비율로 추정되므로 큰문제는 아니라고 생각되며 새로운 유형이 발견될 때마다 그에 적절한 알고리즘의 보완으로 해결이 가능하다고 생각된다. 본 알고리즘을 이용하여 한글 가로쓰기의 낱자 분리가 가능하였으며, 한글 문서 인식 시스템에 이용할 수 있으리라 본다. 좀더 일반적인 시스템의 구축을 위하여 한글 전용뿐만 아니라 영문 및 숫자가 혼용된 문서에 대한 연구가 필요하다.

[참 고 문 헌]

- [1] King Sun Fu, Syntatic Pattern Recognition and Applications, Prentice-Hall, Englewood Cliffs, N.J., 1982.
- [2] Simon Kahan, Theo Pavlidis, "On the Recognition of Printed Characters of Any Font and Size," IEEE Conf. on PR&IP, pp. 274-287, 1987.
- [3] R.G. Casewy, G. Nagy, "Recursive Segmentation and Classification of Composite Character Pattern," IEEE Conf. on PR&IP, pp. 1023-1026, 1982.
- [4] H. Kida, O. Iwaki, K. Kawada, "Document Recognition System for Office Automation," IEEE Conf. on PR&IP, pp. 446-448, 1986.
- [5] Nobon BABAGUCHI, Masatoshi TSUKAMOTO, Tsunehiro AIBARA, "Knowledge Aided Character Segmentation from Handwritten Document Image," IEEE Conf. on PR&IP, pp. 573-575, 1986.
- [6] M. Maier, "Separating Characters in Scripted Documents," IEEE Conf. on PR&IP, pp. 1056-1058, 1986.
- [7] Yea Shuan Huang, Ken Wen Lin and Yihung Chen, "Field Segmentation and Character Isolation Method in Free-Format Chinese Printed Document," on 89' International Conf. on Computer Processing of Chinese and Oriental Languages, pp. 151-155, 1988.
- [8] Kyoona Ha Lee, Kie Bum Eom and R. L. Kashyap, "Character Recognition using Attribute Grammar," on IEEE Proc. CV&PR, pp. 418-423, 1988.
- [9] 이 균하, "ADP&PG를 이용한 한글 문서 인식," 인하대학교 기초과학연구소 논문집 제 9 호, pp. 67 - 71, 1988.
- [10] 김 태균외, "숫자와 문장 부호를 포함한 한글 문서의 자동 인식," 한국정보과학회 '88 봄 학술 발표 논문집, 제 15 권 1 호, pp. 143-150, 1988.
- [11] 김 진형외, "한글 문서 인식 시스템 SILNOON의 개발," 한국정보과학회 '89 봄 학술발표논문집, 제 16 권 1 호, pp. 211 - 214, 1984.4.