

## 수식 표현의 인식에 관한 연구

이 영 교\* · 김 영 포\*\*

### *Experimentation on The Recognition of Arithmetic Expressions*

Lee, Young Kyo · Kim, Young Po

#### 〈Abstract〉

The formula contains up between the text and the structural information, as well as their mathematical symbols. Research on-line or off-line recognition formula is underway actively used in various fields, and various forms of the equation are implemented recognition system. Although many documents are included in the various formulas, it is not easy to enter a formula into the computer. Recognition of the expression is divided into two processes of symbol recognition and structural analysis. After analyzing the location information of each character is specified to recognize the effective area after each symbol, and to the structure analysis based on the proximity between the characters is recognized as an independent single formula. Furthermore, analyzing the relationship between the front and back each time a combination of the position relationship between each symbol, and then to add the symbol which was able to easily update the structure of the entire formula. In this paper, by using a scanner to scan the book formula was used to interpret the meaning of the recognized symbol has a relative size and location information of the expression symbol. An algorithm to remove the formulas for calculation of the number of formula is present at the same time is proposed. Using the proposed algorithms to scan the books in the formula in order to evaluate the performance verification as 100% separation and showed the recognition rate equation.

Key Words : Image Processing, Patterns Matting, Mathematical Expression Recognition, Visual Inspection, Location Information

## I 서론

문자인식 기술과 그에 연관된 기술들은 여러 가지 입력장치들이 폭넓게 보급되면서 더욱 더 중요시 되

고 있다. 최근 주변에서 접할 수 있는 다양한 문자 인식장치들의 개발로 인하여 사람이 직접 타이핑을 하지 않더라도 텍스트로 입력되는 애플리케이션인 OCR은 큰 호응을 얻고 있다. 문자인식은 자유롭지만, 수식분리, 수식인식, 수식연산 과정은 사용자가 의도하는 수식의 결과 값을 도출해 내기 위하여 복잡

\* 부천대학교 정보통신과 부교수

\*\* 한국항공대학교 정보통신공학과 박사과정(교신저자)

한 수식을 입력하는 방법과 그 수식을 해석하는 방법은 일반 텍스트와는 다른 방식을 요구한다[1].

기존 수식의 구조 분석에 관한 연구로는 구문론적인 방법, 투영을 이용한 방법, Graph Rewriting, 순차적으로 코드화된 수학 구문이 있다.

본 논문에서는 구문론적인 방법은 문법을 수학적인 표기 구조로 만든다. 구문론상의 목표를 가지고 top(윗자리)과 down(아랫자리)을 구문 분석하여 심벌에 적용하여 수식의 구조를 탐색하기 위해 구조 명세도를 제안하였다.

2차원 공간상에 쓰인 수식으로부터 심벌 단위 인식이 완벽하게 되었다고 가정 하에 수식의 구조를 해석하는 순차적인 구조 분석 방법을 사용한다.

제안 방법을 통해 빠른 탐색으로 수식 인식을 수행할 수 있음을 보이기 위해, 여러 곳이 책에 있는 수식 수식으로부터 스캔을 받아 수식을 측정하였다[2]. 본 논문에서 제안한 탐색과 최상 우선 탐색 간의 탐색 반복 횟수 및 탐색 시간을 비교한 그래프로써, 많은 수의 심벌을 사용하는 경우에도 빠르고 정확한 인식 결과를 내어준다[3].

2장에서는 제안된 시스템을 통한 영상의 전처리 과정과 3장에서는 순차적인 구조 분석 방법을 소개하고 4장에서는 실험 및 결과에 대해 기술하였다. 최종 5장에서는 결론을 기술한다.

## II. 영상의 전처리

본 장에서는 수식을 분할하기 전 해야 할 전처리 과정에 대하여 알아본다[4].

### 2.1 이미지 획득

영상 처리 시스템은 처리해야 하는 데이터가 많기

때문에 입력 영상이 좋으면 더욱 정확한 검사가 가능하고 영상 처리 알고리즘이 간단해지며 시스템의 제작비용도 낮출 수 있다. 그러므로 입력 영상의 질을 좌우하는 시각 인식부의 설계는 매우 중요하다. 영상의 획득은 대상 물체의 제한 조건에 알맞게 설계되어야 한다.

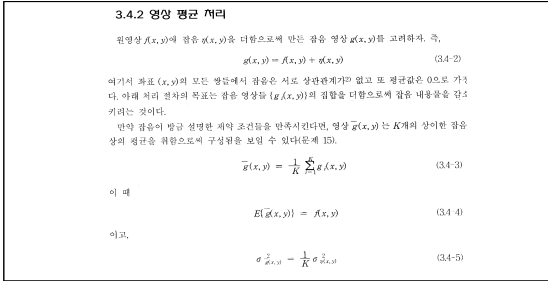
컴퓨터 비전은 대상의 크기, 회전, 조명과 같은 다양한 변화에 따라 같은 영상이라도 서로 다른 상황이 될 수 있다. 특히 크기가 서로 다른 부분을 검사할 경우 동일한 객체 일 경우라도 다른 객체로 인식할 문제를 가지고 있다. 영상은 한글과 워드 그리고 수식이 들어있는 책으로 실험을 하였다.

영상을 획득하기 위한 장치로는 일반적으로 CCD(Charge Coupled Device)소자가 2차원으로 배열된 스캐너를 통하여 300dpi의 해상도로 수식을 읽어 들었다. 스캐너는 따로 조명이 필요치 않고, 노이즈의 영향을 덜 받는다는 장점이 있다. 하지만 촬영 대상이 움직이거나 넓은 영역을 검사해야 하는 경우, 그리고 기술의 발달과 더불어 더욱 세밀한 영상 처리 기술을 필요로 하는 분야로 확대됨에 따라 고해상도 영상의 획득이 요구되고 있다.

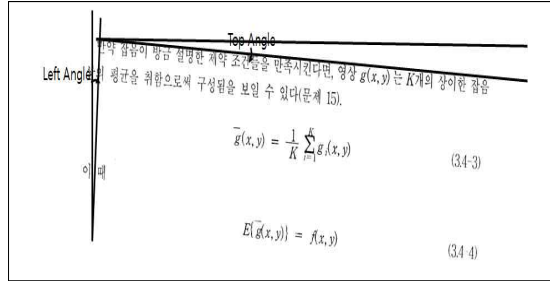
영상의 정확한 위치를 찾기 위한 영상의 정렬은 자동 시각 검사 분야에서 매우 중요한 부분이다. 정확한 검사를 하기 위하여 영상 매칭 알고리즘을 이용하여 두 영상의 좌표축을 정확히 일치시켜야 한다. [그림 1]은 본 논문에서 사용된 이미지의 일부이다.

영상 정렬은 제품의 모서리나 절단면, 화살표 등의 윤곽선이 선명한 부분들을 기준으로 선택하여 지정된 정보를 가지고 정렬한다. 가급적이면 검사할 영상은 정렬 영역에 포함시키지 않는 것이 좋다. 검사할 영상을 정렬 영역에 포함시켜 검사할 경우 정렬 영역이 불량이면 영상정렬에 영향을 미치게 되어 정렬이 잘못될 수도 있기 때문이다[5].

문서의 영상의 전처리는 영상 획득 과정에서 기술



<그림 1> 실험 이미지



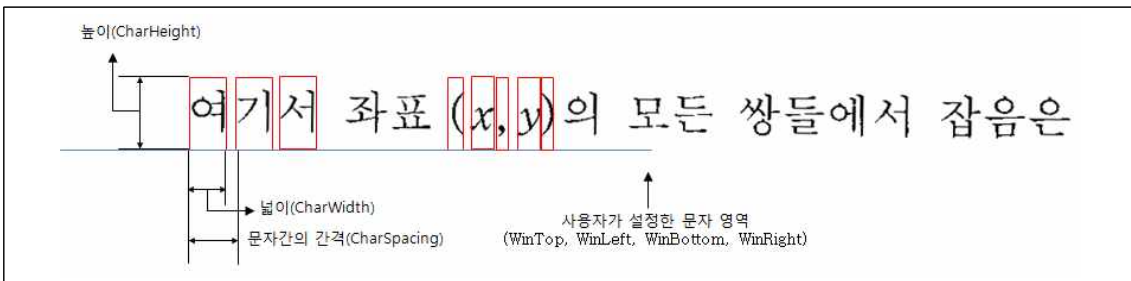
<그림 2> 이미지 보정

어짐이 발생할 수 있다. 영상의 왜곡이 심해지면 문서 인식이 불가능하기 때문에 기울어진 영상의 교정이 필요하다. 영상의 기울어진 정도를 [그림 2]와 같이 위쪽(Top Angle)과 왼쪽(Left Angle)의 기울기를 계산하여 각도를 구한다.

## 2.2 문자 인식

문자인식 방식에는 크게 온라인 인식과 오프라인 인식 방식이 있다. 온라인(On-Line) 인식 방식은 사용자가 필기하는 동안에 컴퓨터가 필기를 인식하는 것을 의미하는데 “실시간” 또는 “동적”이라는 용어가 온라인 대신 사용되기도 한다. 여기서 “실시간”이라는 것은 사용자가 느끼기에 거의 즉각적이라는 의미로써 온라인 필기 인식 시스템의 속도는 사용자의 필기 속도에만 뒤지지 않으면 된다.

다른 문자인식 방식 중의 하나인 오프라인 (Off-Line)인식 방식은 사용자의 필기에 대한 인식이 필기한 이후에 행해지는 것을 말하는데, 그 기간은 하루, 한 달 또는 수년이 될 수도 있다. 이러한 오프라인 인식에는 광학 스캐너 혹은 정지 영상을 촬영할 수 있는 카메라 등이 이용된다[6]. 온라인 인식과 오프라인 인식의 또 다른 차이점은 필기 데이터의 획득 방법에 있다. 온라인 방식에서는 필기의 시간적 공간적인 동적 정보를 얻을 수 있는데 획수, 획순의 각 획에 대한 필기 방향과 각 획 내에서의 필기 속도 등이 이러한 정보의 예이다. 이러한 정보들은 오프라인 인식에서는 찾기 힘들거나 아예 찾을 수 없는 아주 중요한 것들이다. 그러나 스캐너나 카메라 등을 통하여 얻어진 오프라인 데이터는 윤곽선을 추출하거나 골격 화를 위해 많은 계산량을 필요로 하고 불완전한 전처리를 해야 하는 차이점이 있다. 또한, 온라인 데

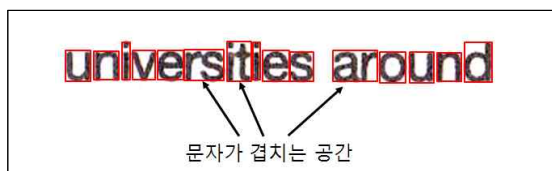


<그림 3> 문자 인식 영역

이터는 사용자가 쓰고 있는 그 순간에만 얻을 수 있는 반면에, 오프라인 데이터는 글 쓴 이후에는 언제라도 다시 데이터를 얻을 수 있는 차이점이 있다. 사용자 필기체 입력방식의 특성상 오프라인 인식보다는 온라인 인식방식이 적합하다. 문자 인식을 하기 위해서는 이미지의 어느 위치에 문자들이 있는지 알아야만 한다. [그림 3]은 문자별로 인식영역을 나누는 과정이다[7].

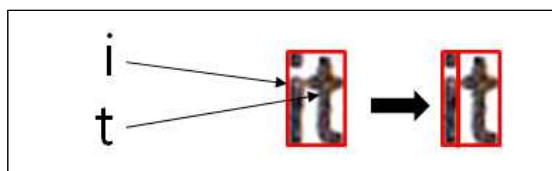
그래서 OCR에서는 문자들을 인식할 영역을 설정해주면 자동으로 문자들의 위치, 높이, 넓이, 문자 간의 간격들을 설정해 준다.

문자들의 높이, 넓이, 문자 간의 간격, 사용자가 설정한 문자 영역은 전역 변수인 Parameters PM에 저장되고, 위치(첫 글자의 위치)는 Left, Top 전역변수에 저장된다[8].



<그림 4> 문자 분할의 잘못된 사례

문자가 정상적으로 분할 됐는지를 평가에서는 분류에서 텍스트로 설정된 영역들에 대해 학습된 골든 이미지 문자와 비교하여 비정상적으로 분할된 문자들을 파악하여 [그림 4]에서 [그림 5]와 같이 재수정된다[9].



<그림 5> 수정된 문자 영역

### III. 영상의 전처리

다양한 수식인식 방식 중의 하나인 블록모델을 이용한 필기체 수식 인식 방식은 구문론적 분석 없이 심벌의 위치 관계만으로 수식을 인식한다. 심벌을 순차적으로 추가할 때마다 추가한 심벌과 이미 추가된 심벌간의 관계만을 분석하여 전체 수식의 구조를 쉽게 갱신할 수 있다. 심벌별로 분할 및 인식된 수식의 심벌을 하나씩 순차적으로 추가해가면서 탐색 트리를 생성한다[10].

#### 3.1 수식 분리

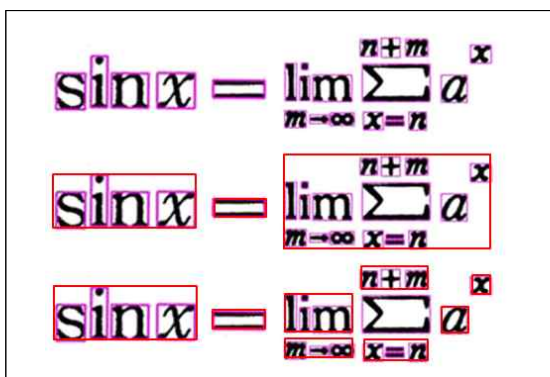
이진 영상에서 수식의 영역을 획득하기 위해 Labeling을 수행한다. 영상을 탐색하여 Intensity가 Maximum인 Pixel을 찾아서 해당 위치에서 인접한 픽셀을 동일한 영역에 포함한다. 인접한 백색 픽셀은 4-Neighbor 집합(상하좌우)에서 검색한다.

하나의 문장으로 되어 있는 여러 개의 수식을 인식하기 위해서는 각 수식을 연결하는 연산자를 정확히 인식할 수식을 객체별로 인식을 할 수 있어야 한다.

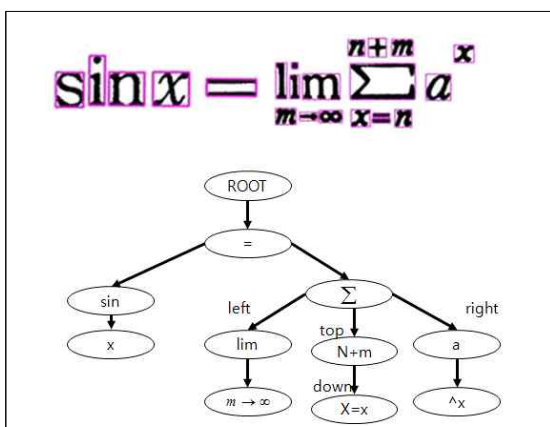
예를 들어 문자 집합에서 각각의 연산자를 연산이 발견했을 경우 크기나 위치에 따라 수식과 수식 사이를 분리할 것인가, 아니면 계산 과정인가를 인식하여야 한다. [그림 6]은 수식을 분리하는 과정을 보여주고 있다[11].

#### 3.2 수식의 분류

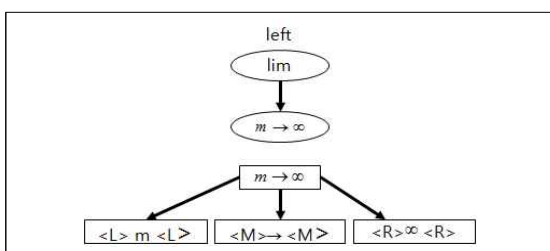
인식된 수식은 하나의 문자 집합 단위로 구성하며 구성된 단위는 하나의 수식으로 인식하였다. 하나의 인식된 수식은 [그림 7]과 같이 하나의 Root 노드와 각 연산자와 수식에 따른 Sub 노드를 가진다. 각 연산자마다 규칙에 의해 해당 연산자를 트리에 넣는



<그림 6> 수식 분리 과정



<그림 7> 1차 수식 트리



<그림 8> 2차 수식 트리

다. 하나의 수식 인식이 끝나면 발견되는 다음 수식에 다시 인식할 수 있게 하기 위하여 재귀적으로 구성한다[12].

분리된 수식은 XML 방식으로 [그림 8]과 같이 다시 한 번 2차 분리과정을 거치게 된다. 트리구조로 저

장된 수식의 노드를 List에 저장한 후 수식유형에 대한 분석을 한다. 각각의 연산자는 태그로 된 감싸진 스트링으로 연산자에 맞게 설계되었다[13-14].

트리구조로 저장된 수식을 터미널 노드를 List에 저장한 후 이를 이용하여 수식유형에 대한 분석을 시작한다. 여기서 각 마크업과 연산자를 나타내는 마크업으로 감싸진 스트링을 분석하여 각 마크업과 연산자에 맞는 계산과정의 알고리즘을 호출하도록 설계하였다. 현재 단계에서 주로 쓰이는 마크업의 유형은 다음과 같다. 계산이 많을 경우 그에 맞는 마크업 유형을 생성하여 저장한다.

#### IV. 실험결과

제안 방법을 통하여 수식 인식의 정확성을 파악하기 위해 한글과 워드의 수식 그리고 곤잘레스의 디지털 영상처리 2판과 사디쿠의 전자기학5판의 수식을 스캔 받아 실험을 하였다. 탐색하여 정확한 수식결과가 인식되는 것을 확인 하였다.

<표 1> 수식 인식시간

수식	평균 인식시간
$5 + 3 + 2$	30m/s
$x^2 + 2 \times \frac{3}{2}$	50m/s
$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$	인식불가
$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$	110 m/s
$(x^2 + 24x + 6)^3$	58 m/s
$\int_{-\infty}^{\infty} 3x^2 + 4x + 1 dx$	110 m/s
$\sum_{i=1}^{10} 3x^3 + 2x^2 + 1$	150m/s
$\lim_{x \rightarrow 0} \frac{1}{x}$	80m/s

테스트에 사용된 학습 데이터는 총 150장으로 하나의 수식당 평균 5여 개의 서로 다른 학습 데이터를 가지고 있다. 그러나 이러한 환경에서 인식속도는 수식과 문자마다의 차이로 인하여 서로 다르나 C언어의 프로세서 시간 체크 한계상 최저 30m/s 이하는 측정이 되지 않으며 [표1]은 C언어의 타이머로 측정한 수식의 인식시간이다.

Blur 된 영상은 Labeling을 수행함에 있어서 하나의 문자영역을 정확히 탐지하기 힘들다. 문자와 배경 사이의 경계를 구분하기 힘들기 때문인데, Focusing 기능으로 어느 정도 해결할 수 있지만 이미 입력되어 버린 흐린 영상의 보정은 한계가 있었다. 또한, 해상도가 크지 않아서 작은 위첨자나 아래첨자는 피라미드 알고리즘을 통하여 확장하여 모든 문자의 크기를 정규화 하여 동일한 인식을 할 수 있도록 처리 하였다.

알고리즘을 설계할 때는 모든 수식을 인식을 하려고 계획을 했지만 모든 수식에 대한 인식은 구현하지 못하였다. 근호가 포함되어있는 수식은 Labeling을 수행했을 때 하나의 Label이 또 다른 Label을 포함하고 있어 원활한 인식에 문제가 되고 있으며 횡적으로 구성된 수식은 위치와 관계를 쉽게 정의할 수 있지만, 행렬이나 병렬식은 경계가 되는 위치가 모호할 경우 인식이 불가능한 경우가 발생 하였다.

알고리즘의 최적화가 진행되어 있지 않아 만족할 만한 수준의 인식 속도는 나오지 않지만 150개의 수식 중 행렬이나 병렬식이 포함된 5개의 수식은 인식 불가 하여 평균 97%의 인식률을 보이고 있다.

## V. 결론 및 향후연구

본 연구에서는 블록 모델과 이를 분류하는 방법을 통하여 수식을 인식하였다. 제안한 알고리즘은 전역

적 탐색을 하면서도 실용적으로 사용이 가능한 인식 속도를 내어주어, 코드가 많은 수식에 있어서도 인식은 가능하나 효과적으로 분류가 되지 못하여 많은 메모리를 낭비하는 결과를 초래하였다. 아직 만족할 만한 인식속도가 나오지 않기에 알고리즘의 최적화와 수식의 인식 및 분리만 가능한 상태이므로 아직 한글이나 워드 등에 인식된 수식을 재조립하지 못하였다. 향후 한글과 워드의 수식 구조를 파악하여 자동으로 한글과 워드에 인식할 수 있게 할 것이다. 향후 태블릿과 터치스크린 같은 입력장치들의 직접 사람이 입력한 수식의 입력장치에서 사용자의 필기체 수식을 입력 받아 수식을 인식할 수 있게 하여 사용자들에게 한발 더 다가갈 수 있기를 기대한다.

## 참고문헌

- [1] S. H. Park, A Study on the Development Direction of the Digital Textbook for Disabled, KERIS, 2007.
- [2] S. I. Lee, The National Disabled Library Support Center: The Report on Digital Information Service for Disabled, The National Library of Korea, 2007.
- [3] H. G. Lee, S. A. Kim, J. H. Kim, "Methods to Improve Reading Environment for the Disabled," Journal of Special Education, vol. 8, no. 4, 2007, pp. 699- 720.
- [4] R. Zanibbi & D. Blostein, "Recognition and retrieval of mathematical expressions," IJDAR, Oct. 2011.
- [5] R. C. Gonzalez & R. E. Woods, "Digital Image Processing," 3rd Edition, Pearson Education, 2010.

[6] W3C's Document Object Model, <http://www.w3.org/DOM>

[7] amaya - W3C's Editor/Browser, <http://www.w3.org/Amaya/>

[8] 이택현, 김진형, 김기응, "순차적 블록 모델을 이용한 필기체 수식 인식," 한국정보과학회 학술 발표논문집, vol. 34, no. 2(A), 2007, pp. 207-208.

[9] R. T. Chin, and C. A. Harlow, "Automated Visual Inspection: A Survey," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. PAMI-4, No. 6, 1982, pp. 557-573.

[10] K. W. Tobin, "Inspection in Semiconductor Manufacturing," Webster's Encyclopedia of Electrical and Electronic Engineering, Vol. 10, 1999, pp. 242-263.

[11] R. Anderson, "Two-dimensional mathematical notation," in Syntactic Pattern Recognition and Applications, 1977, pp. 147-177.

[12] 이현구, 김동주, "Face Recognition using Extended Center-Symmetric Pattern and 2D-PCA," vol 9, no 2, 2013. 6.

[13] S. Chang, "A method for the structural analysis of two-dimensional mathematical expressions," Information Sciences, vol. 2, no. 3, 1970, pp. 253-272.

[14] 김정재, 류민우, 차시호, 조국현, "Design and Implementation of User Location-based N-Screen Service System," vol. 10, no. 1, 2014. 3.

■ 저자소개 ■



이 영 교  
Lee, Young Kyo

2001년 9월~현재  
부천대학교 정보통신과 부교수  
2002년 2월 한국항공대학교 대학원  
통신정보공학과 (공학박사)  
1995년 8월 한국항공대학교 대학원  
통신정보공학과 (공학석사)  
1992년 2월 한국항공대학교 통신정보공학과  
(공학사)

관심분야 : 유비쿼터스, 유무선네트워크,  
트래픽 제어  
E-mail : yklee@bc.ac.kr



김 영 포  
Kim, Young Po

2011년 2월 한국항공대학교 대학원  
정보통신공학과 (박사수료)  
2005년 2월 한국항공대학교 대학원  
항공전자공학과 (공학석사)  
2001년 2월 국립경상대학교 정보통신공학과  
(공학사)

관심분야 : Mobile IP, Self-organizing  
wireless networks, Image  
Processing  
E-mail : zero-po@hanmail.net

논문접수일: 2014년 11월 20일  
수 정 일: 2014년 12월 5일  
게재확정일: 2014년 12월 10일