

인간에게 친밀한 한글 인식 및 편집 지원시스템

Human Friendly Recognition and Editing Support System of Korean Language

손영선

Young-Sun Sohn

동명대학교 정보통신공학과

요 약

본 논문에서는 책 또는 논문을 읽을 때 중요한 부분 또는 정리해야 할 부분의 영역을 선택하면, 해당영역의 문자들을 문서 편집기에 순차적으로 출력시킴으로써 수정·저장 가능하게 하여 공부한 부분을 정리하는 시스템을 구현하였다. 이 시스템은 사용자가 원하는 영역을 손가락으로 선택하면, 손 인식 알고리즘을 적용하여 손가락 움직임을 검출하여 선택된 영역을 인식한다. 선택된 영역의 가로, 세로 거리를 펄스 수로 변환한 후 모터를 제어하여 그 위치만큼 카메라를 이동시킨다. 문자 인식이 가능하도록 줌을 확대/축소하고 조정된 줌에 맞는 초점으로 근접 제어한 후, 더욱 선명한 영상을 얻기 위해 명암차이를 이용하여 미세조정을 하였다. 획득된 영상에 문자인식 알고리즘을 적용하여 문서로 변환시켜서 중요 부분들을 정리케 하는 한글 인식 및 편집지원 시스템을 구현하였다.

Abstract

In this paper we realized a system, if a user selects the area of the important parts or the arrangement parts when he reads the books or the papers, which amends, stores and readjusts the characters that are included in the selected area by outputting the characters to the word processor in sequence. If a user selects what he wishes for with his finger, the system detects the movement of the finger by applying the hand recognition algorithm and recognizes the selected area. The system converts the distance of the width and the length of the selected area to the number of the pulse, and controls the motor to move the camera at the position. After the system scales up/down the zoom to be able to recognize the character and controls the focus to the regulated zoom closely, it controls the focus in detail to get more distinct image by using the difference of the light and darkness. We realize the recognition and editing support system of korean language that converts the obtained images to the document by applying the character recognition algorithm and arrange the important parts.

Key words : 한글, 퍼지추론, 영상처리, 지능, 문자인식

1. 서 론

컴퓨터 기술이 발전함에 따라 사용자들이 접근하기 쉬운 인터페이스가 필요하게 되었으며, 이에 대한 다양한 연구가 진행되고 있다[1]. 그 중, 영상 입력 장치를 통하여 문서를 자동으로 입력시켜 처리해주는 시스템은 사용자에게 편리성을 주는 사용자 인터페이스의 하나로서 일반화되어 가고 있다[2]. 그러나 기존의 시스템은 사용자가 시스템에 대한 사전 지식 및 절차를 숙지해야 하는 부담감도 갖게 된다.

본 논문에서는 CCD 카메라를 통해 획득한 전체 문서 내에서 사용자의 손동작으로 원하는 영역내의 문자들만을 선택적으로 발췌한 후, 이를 텍스트 형태로 문서 편집기에 나타냄으로써, 공부한 부분을 정리해주는 한글 인식 및 편집지원 시스템을 구현하였다.

2. 시스템 알고리즘

그림 1의 시스템 흐름도에서 알 수 있듯이, 구현된 시스템

은 흑백 CCD 카메라를 이용하여 일반적인 문서를 입력 받고, 카메라 영역 내에 손이 들어오면 손 인식 알고리즘을 적용하여 손가락의 끝점을 찾는다.

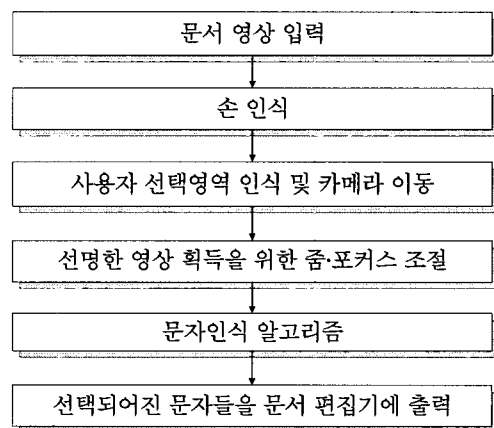


그림 1. 시스템 흐름도
Fig. 1. System flow

접수일자 : 2006년 12월 26일
완료일자 : 2007년 7월 30일

손의 이동을 통해 문서 내에서 인식하고자 하는 영역을

선택하고, 선택된 영역의 시작위치로 카메라를 이동시킨다. 줌과 포커스를 조절하여 문자 인식이 가능한 크기의 선택된 영상과, 이 영상에 문자 인식 알고리즘을 적용하여 인식된 문자는 문서 편집기로 나타내어 편집 가능한 문서를 사용자에게 제공한다.

3. 손가락 이동에 의한 영역 선택[3]

3.1 손 인식

배경 문서로서 A4 두 장 크기의 텍스트로 구성된 문서를 저장한 후, 문서 영역을 선택하기 위해 배경문서 안으로 손이 들어오면 배경 영상과 손이 포함되어 변화된 영상을 화소간 뿔셈 연산하여 배경으로부터 손을 분리하여 인식하였다.

3.2 손가락 인식

손 영상에서 손 내부에 내접하는 최대의 원을 이동하여 도형에서 선 정보를 추출할 수 있는 방법인 최대 원형 이동법을 이용하여 손의 무게중심점을 찾는다. 찾아진 무게중심점에서 2도 간격으로 샘플링한 원형 패턴 벡터 알고리즘을 적용함으로써 아래 그림 3과 같이 구하여진 거리 스펙트럼 데이터를 이용하여 손가락 끝부분을 찾는다.

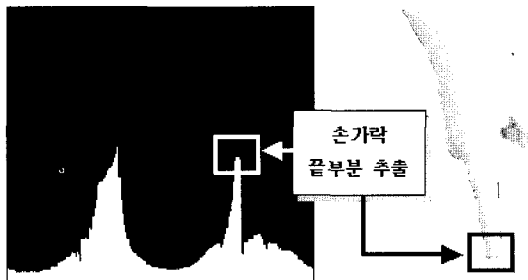


그림 2. 원형 패턴 벡터 알고리즘을 이용한 손가락 끝부분 추출

Fig.2. Finding out the tip of finger by applying the circular pattern vector algorithm

3.3 손가락 이동에 의한 영역 선택

영역을 선택하기 위해 시작 위치를 손가락으로 잠시 가리킨 후, 원하는 영역을 손가락으로 그려서 다시 초기 위치에 도착하면 그 영역을 추출한다.

4. 카메라 이동 및 줌, 포커스 조절

4.1 선택된 영역의 시작 위치로 이동

초기의 카메라는 문서 영상 전체가 보이도록 문서 영역의 중앙에 위치한다. 손가락에 의해 영역이 선택되면, 초기위치와 선택된 위치와의 차이 픽셀 값만큼 선택된 영역의 시작 위치인 좌측 상단으로 카메라를 이동시킨다.

4.2 줌과 포커스 조절

선명하지 못한 영상은 전처리 단계에서 정보를 많이 잃어버리므로, 정확한 포커스 조절이 필요하다[4,5]. 줌과 포커스는 그림 3과 같이 현재배율과 확대할 배율을 이용해 자동 조

절된다. 조절된 영상에서 문자 크기가 인식 가능한 크기가 아니면, 확대할 배율을 재조정하여 그에 대한 줌과 포커스를 조절한다. 여기서 배율에 따른 줌과 포커스 데이터는 실험에 의해 구하였다.

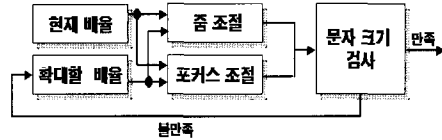


그림 3. 줌과 포커스 조절 블록도

Fig.3. The control block of the zoom and the focus

4.2.1 줌 조절

현재 배율에서 확대할 배율로 조절할 경우, 배율과 DC모터 구동시간 관계는 그림 4와 같고 이 관계를 이용하여 그림 5와 같은 멤버십 함수를 구축하여 줌 모터를 조절하였다[6].

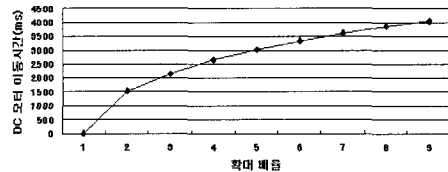


그림 4. 배율과 DC모터 구동시간과의 관계

Fig.4. The relation between the magnification and the driving time of the DC motor

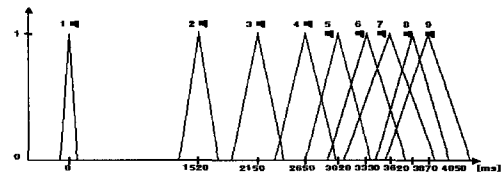


그림 5. 줌 조절 멤버십 함수

Fig.5. The membership function of the zoom control

4.2.2 배율 조절

포커스는 피사체에서 카메라까지의 거리와 배율에 따라 조절된다. 본 논문에서는 거리가 일정하게 고정되므로 배율만을 고려해 포커스를 추론 하였으며[6], 배율에 대한 멤버십 함수를 그림 6과 같이 구축하였다.

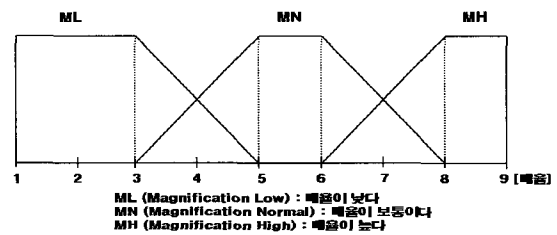


그림 6. 배율에 대한 멤버십 함수

Fig.6. The membership function of the magnification

실험을 통하여 현재 배율과 조절될 배율에 대한 포커스 조절 값을 표 1과 같이 구성하였으며, 그림 6의 배율에 대한

멤버십 함수와 표1의 퍼지추론 규칙을 사용하여 그림7과 같은 포커스 조절에 대한 멤버십 함수를 구하였다.

표 1. 퍼지 추론 규칙
Table1. The fuzzy inference rule.

	작대 배음	ML	MN	MH
현재 배음				
ML		N	LN	LL
MN		RN	N	LS
MH		RL	RS	N

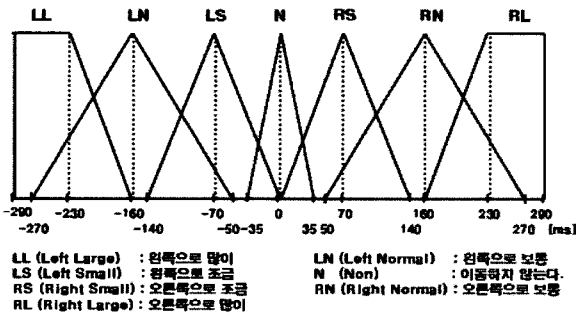


그림 7. 포커스 조절의 멤버십 함수
Fig.7. The membership function of the focus control

4.2.3 포커스 조절

입력 영상이 문서이므로 흑백 비율이 높은 영상이 더욱 선명함을 알 수 있다. 퍼지 추론에 의해 근사 포커스 값까지 조절한 후, 명암 대비법을 적용하여 가장 선명한 포커스로 조절한다. 포커스 조절 알고리즘으로 가우시안 분포를 가지는 모델이 적합하므로[7], 본 논문에서는 가우시안 분포를 나타내는 식(1)을 가중치 함수로 사용하였다.

$$y = \left(1 - \frac{n}{255}\right)^2 \quad (1)$$

여기서, n : 그레이 레벨 값(0~255)
y : 레벨 값에 대한 가중치

4.3 카메라 이동

문자 이미지를 확대하였을 때, 사용자가 선택한 영역이 한 화면에 표현되지 못하는 경우에는 카메라를 이동시켜 가면서 선택된 영역 전체를 받아들인다. 카메라의 이동 시 사용된 모터는 50펄스 당 1mm를 이동하며, 초기화된 영상에서 1픽셀 당 실제 거리는 0.61mm 이므로 중심에서 사용자가 확대를 원하는 위치로의 이동에 관한 step 모터의 펄스 입력은 식(2)과 같이 계산되어 진다.

$$y = 0.61 \times 50 \times x \quad (2)$$

x : 중심에서의 픽셀차이 값
y : 모터에 입력될 펄스 수

5. 문자 인식

그림 8의 문자인식 흐름도에서 알 수 있듯이 문서영상을 입력 받아 문자단위로 분할한 후, 한글의 구조적 특징을 이용하는 경계선 기울기 방법을 적용하여 수평, 수직 모음의 유무를 판별함으로써 6가지 유형으로 분류하였고, 문자의 윤곽선 위상 그래프를 이용하여 자음과 모음을 분리할 수 있었다. 분리된 모음은 최대 길이 투영법을 사용하여 직선성분을 검사함으로써 인식할 수 있었고, 자음은 경계선 기울기 방법을 사용하여 윤곽선의 각도를 특징점으로 추출한 후 표준패턴과 비교해서 인식하였다. 인식된 자소는 KS 완성형 한글 코드를 사용하여 조합함으로써 완성된 한글을 만들었으며, 완성된 문자들은 문서편집기로 출력하였다.

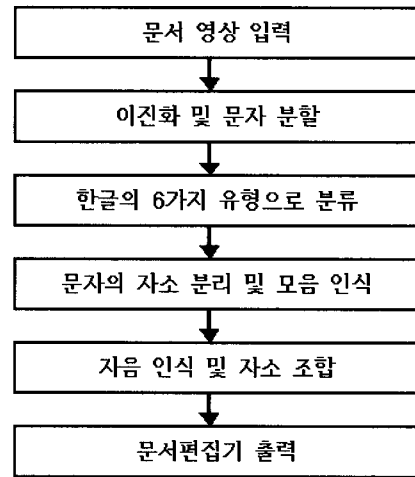


그림 8. 문자인식 흐름도
Fig.8. The flow of the character recognition

5.1 문자 분할

문서 영상이 입력되면, 그림 9와 같이 가로 히스토그램을 사용해 문서를 라인별로 나누고, 각 라인마다 세로 히스토그램을 사용하여 문자 단위로 분할 한다.

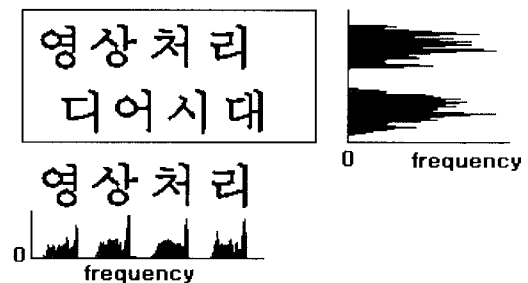


그림 9. 문자영상에 대한 문자 분할
Fig.9. The character division for the character image

5.2 문자의 유형 분류

한글은 구조적 특성상 그림 10과 같이 초성, 중성, 종성의 3가지 종류의 자소로 구성되어 있으며, 모음의 형태와 종성의 존재 유무에 따라 6가지 유형으로 분류된다[8].

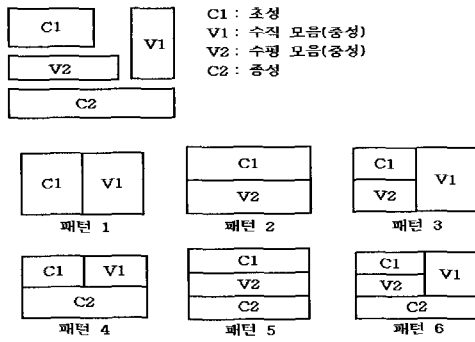


그림 10. 한글의 6가지 유형 분류
Fig.10. The six pattern classification of the Korean

5.2.1 경계선 기울기 방법

경계선 기울기 방법은 다양한 정보를 포함하고 있는 윤곽선의 다차원 정보를 1차원 정보로 나타내는 것이다[9]. 그림 11은 물체 윤곽선의 트레이싱 과정을 나타낸 것이다.

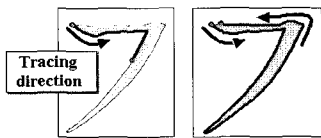


그림 11. 경계선 추적
Fig.11. Boundary Tracing

5.2.2 문자의 유형 분류

문자의 모음 형태를 파악하기 위하여 서로 다른 위치에서 경계선 기울기방법을 두 번 수행한다. 그림 12(a)에서는 문자의 최고점을 시작으로 트레이싱하여 x축, y축 기울기 변화량을 우측 그래프로 표현하였다. 그림 12(b)에서는 문자의 가장 좌측에 존재하는 지점부터 트레이싱을 시작하여 x축, y축 기울기 변화량을 그래프로 나타내었다.

문자의 x축과 y축 변화량을 위상으로 변환하기 위하여 식(3)을 사용하였고, 문자의 수평모음의 윤곽선 위상 특성을 고려하고, 다양한 문자에 대하여 트레이싱한 결과를 이용하여 식(4)를 구하였다. 그림 13은 그림 12(b) 및 식(3)에 의하여 구하여진 위상 그래프와 식(4)에 의하여 구하여진 축에 대칭하는 필터를 나타내고 있다. 그림 14로부터, 위상 그래프와 필터가 교차하는 곳의 위치를 검출하여 수평모음의 존재 유무를 파악하였다.

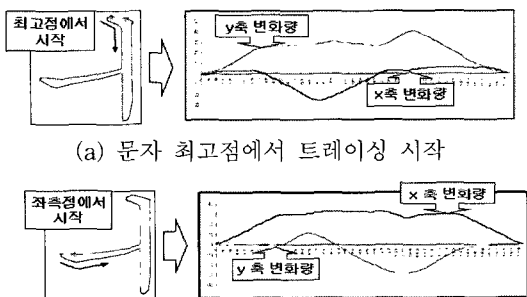


그림 12. 트레이싱 시작점에 따른 결과 그래프
Fig.12. The result graph according to the tracing start point

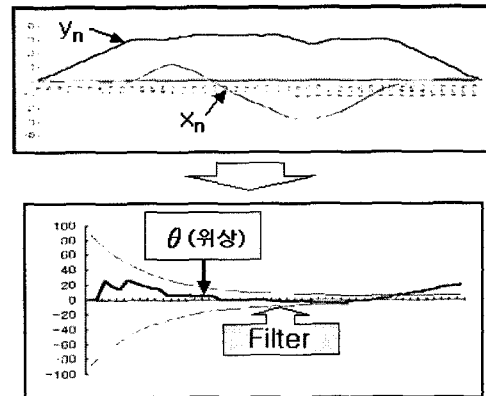


그림 13. 그림 5(b)의 위상 그래프 및 필터함수
Fig.13. The filter function and the phasor graph of the Fig.5(b)

$$\theta = \tan^{-1}(y_n / x_n) \quad (3)$$

$$\text{Filter} = (5 + 85 \cdot e^{(-4n/\text{width}^2)}) \quad (4)$$

한글에서는 수평모음이 존재하지 않으면, 수직모음은 항상 존재한다. 그러나 수평모음이 존재하면 수직모음의 존재 유무를 판별해야 한다. 그림 14에서는 그림 12(a)의 그래프를 옮겨 그린 후 점진 사각형 부분이 나타내고 있는 역 방향으로부터 탐색하여 x축 방향의 변화량과 y축 방향의 변화량을 검사하여 수직모음이 존재하는 것을 확인하였다.

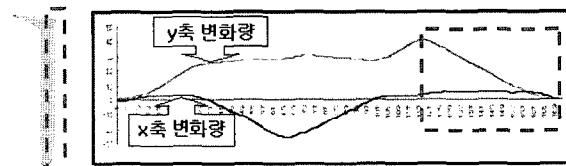


그림 14. 결과 그래프를 역방향 탐색하여 수직 모음의 유무 판별
Fig.14. The distinction of the existence of the vertical vowel by the searching the result graph in opposition

문자의 수평 모음과 수직 모음의 유무를 검사하고 난 후, 그림 15와 같이 문자의 수직모음에서 윤곽선 트레이싱한 결과를 이용하여 받침자음의 존재 유무를 판별한다.

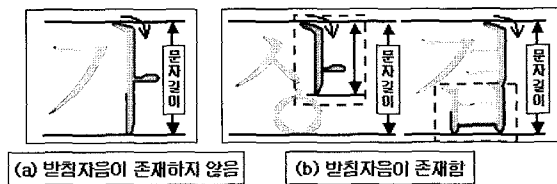


그림 15. 받침자음의 유무 판별
Fig.15. The distinction of the existence of the final consonant

그림 15(a)와 같이 문자의 수직 성분과 문자길이가 일치하고 상대적으로 수평성분이 임계치 이하이면 받침 자음은 존

재하지 않는다. 반대로 그림 15(b)와 같이 수직 성분이 문자 길이보다 적거나, 문자의 최저점에서 시작한 수평 성분이 임계치 이상 검출되면받침 자음은 존재한다. 이와 같은 과정의 결과를 모두 조합하여 입력 문자를 한글의 6가지 유형으로 분류할 수 있었다.

5.3 자소 분리

그림 16의 예에서 보여지듯이, 식(3)에 의한 위상 데이터와 식(4)로 구해진 필터의 교차점 위치를 구함으로써 자음과 수평 모음의 분리위치를 선정하고, 문자를 세로로 투영하여 검은 화소의 개수가 가장 작은 지점을 자음과 수직 모음의 경계선으로 하여 문자의 자소를 분리하였다.

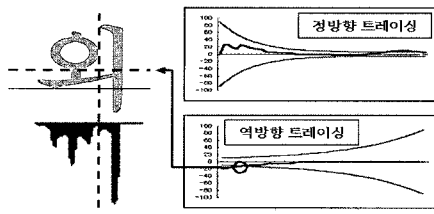


그림 16. 문자의 자소 분리 경계선 추출 예
Fig.16. The example of the abstraction of the border line for the separation of the consonant and vowel of the character

5.4 모음 인식

자소 분리 후, 각 행 또는 열에 존재하는 연속된 흑화소 그룹들 중에서 최대 길이를 히스토그램으로 형성하는 최대 길이 투영법을 적용하여, 그림 17과 같이 수평, 수직 모음을 인식한다[7].

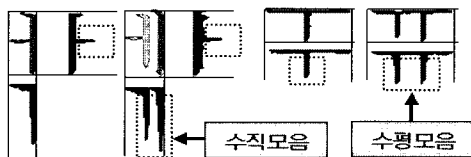


그림 17. 수평모음과 수직 모음 구분
Fig.17. The classification between the horizontal vowel and the vertical vowel

5.5 자음 인식

자음의 패턴을 만들기 위해 다음의 세 단계를 수행한다. 1 단계는 경계선의 위상변화를 재 추출하며, 2단계는 그림 18과 같이 이웃한 위상 변화 데이터들의 평균값을 이용하여 자음의 위상 그래프를 포락선으로 표현한다. 3단계는 동일한 조건에서 자음을 인식하기 위해 각도 변화량이 100도가 되는 지점을 기준으로 전체 위상 그래프를 이동하고, 최종적으로 데이터를 100개로 재 샘플링하여 표준화함으로써 문자의 크기에 구애 받지 않는 패턴이 생성된다. 정규화된 입력 데이터를 미리 저장되어 있는 자음의 특징 점과 비교하여 오차율이 가장 낮은 자음을 선택하도록 하였고[10], 모음과 조합하여 문서 편집기에 나타내었다.

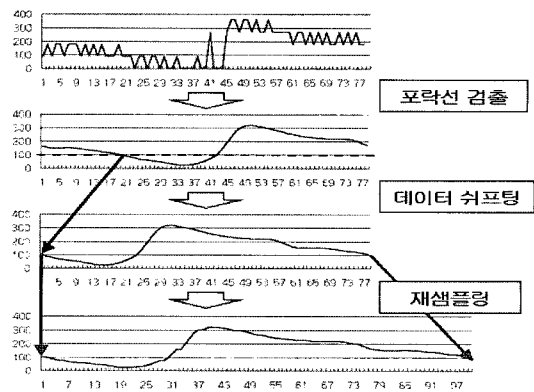


그림 18. 자음의 각도 추출 및 정규화 과정
Fig.18. The regularity process and the angle abstraction of the vowel

6. 하드웨어 구현

6.1 카메라

WV-BP334 흑백 CCD 카메라와 DT-3155 PCI frame grabber를 사용하여 영상을 입력 받도록 하였다.

6.2 PC 인터페이스 카드

PC에 I8255를 이용한 인터페이스 카드를 제작하여 장착함으로써 3bytes의 데이터 버스 라인을 개설하고 이를 이용하여 제어 데이터를 모터에 전송하였다. 각 byte는 step모터, DC모터 및 정지회로에 사용되었다.

6.3 하드웨어 및 step/DC 모터 제어

외관은 step모터 및 DC모터 부분, 카메라 고정부분, 직선형 이동부분과 제어회로 부분을 포함하도록 하였다. 카메라의 위치 이동에는 SLA7024M step모터를 사용하였다. 정전류 방식으로 step모터 회로를 구성하였으며, 2상 여자방식으로 구동하고, 시스템 클럭을 이용하여 가감속 제어하였다 [11]. 줌과 포커스 조절은 H-bridge회로를 이용한 DC모터를 사용하였으며, PWM방식으로 제어하였다[12]. 상기와 같이 구성된 시스템이 그림 19에 보여진다.

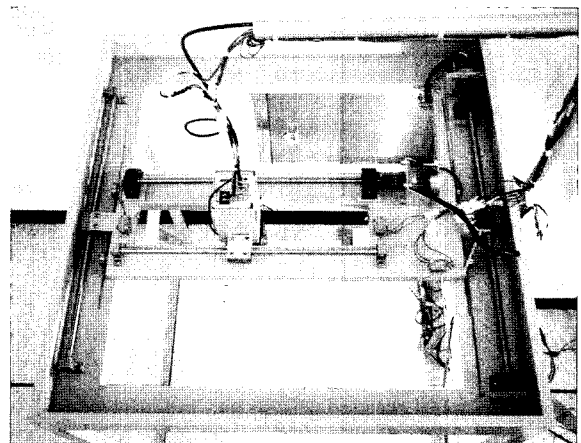


그림 19. 구현된 시스템
Fig.19. The realized system

7. 결론 및 향후 과제

본 논문에서는 지적 인터페이스의 한가지로서, 인간에게 친밀감을 주는 한글 인식 및 편집지원 시스템을 구현하였다. 문서영상을 배경으로 하여 사용자가 선택한 영역을 추출한 후, 퍼지 추론을 사용해 카메라의 줌과 포커스를 조절함으로써, 빠른 속도로 문자 인식이 가능한 크기의 선명한 영상을 획득하였다.

획득된 영상내의 문자는 6패턴 분류, 경계선 기울기 방법 등을 적용하여 인식함으로써 원형패턴 벡터 알고리즘을 적용한 경우(44msec)보다 처리 시간이 단축됨(7msec)을 알 수 있었다. 실험에서는 바탕체 한글을 사용하여 106개 중에서 93개를 인식하여 87.7%의 인식율을 얻었으며, 인식된 문자는 문서편집기에 나타내어 편집 및 저장할 수 있게 하였다.

향후 과제로는 그림, 표 및 다양한 문자의 인식, 음성인식 기능을 추가하여 음성명령의 수행, 사용자의 퇴실에 따른 방법 기능 지원, 입실 후의 사용자 행동에 따른 향상된 학습 환경 지원, 예를 들면 더워하는 행동에 대하여 에어컨 가동, 등 더 지능화된 인터페이스의 구현이 고려되어진다.

8. 참고 문헌

- [1] 김미진, "이용자-컴퓨터 인터페이스 측면에서 고찰한 하이퍼 미디어 시스템", 정보관리학회지, Vol.09, No.2, PP.154-180, Dec.1992.
- [2] 김홍배, 송종수 외 3명, "멀티폰트 한글 문서인식 시스템의 개발", 한국 인공지능개발연구조합, 1992.
- [3] 백승복, 손영선, "손가락 이동에 의해 선택된 영역의 인쇄체 한글 영상 문서화", 한국 퍼지 및 지능시스템 학회, Vol.12, No.4, PP.306-310, 2002.
- [4] 이형근, 노경완, 김충원, "인근 픽셀 차이 값의 분산(VOD)을 이용한 자동 초점 조절 알고리즘", 大韓電子工學會, Vol.21, No.2, PP.935-938, Nov.1998.
- [5] 이성희, 김주현, 최병태, 고성제, "미디어 필터의 차이를 이용한 비디오 카메라의 자동초점조절 알고리즘", 韓國通信學會論文誌, Vol.23, No.1, PP.41-49, Jan. 1998.
- [6] 本多中二, 大理有生, "퍼지 공학 입문", 용보출판사, 2000.

- [7] 장서익, 이용만, 정기욱, 김원찬, 이충웅, "자동 초점 조절 기법 및 하드웨어의 구현", 韓國通信學會 論文誌, Vol.20, No.2, PP.502-517, Feb. 1995.
- [8] 이근수, 최형일, "퍼지 추론을 이용한 한글 문자 인식: 최대 길이 투영에 의한 한글 문자 유형 분류", 인지 과학 논문집, Vol.03, No.2, PP.249-170, 1992.
- [9] 김민기, 권영빈, "인쇄체 문자 인식을 위한 펼쳐진 윤곽선의 샘플링 기법", 인지과학회, Vol.6, No.4, PP.99-112, 1995.
- [10] 장석진, 강선미 외, "자소 인식 신경망을 이용한 한글 문자 인식에 관한 연구", 전자공학회 논문지, Vol.31, No.1, PP.81-87, Jan.1994.
- [11] 許旭烈, "스텝모터의 制御", 電氣學會誌, Vol. 36, No. 2, PP.17-24, Feb, 1987.
- [12] 신정환, "C가 미는 로봇1, 2, 3", Ohm사, 2000.

저 자 소 개



손영선(Young-Sun Sohn)
 1981년 : 동아대학교 전자공학과 졸업(공학사)
 1983년 : 동 대학원 졸업(공학석사)
 1990-1998년 : 한국전자통신연구소 선임연구원
 1998년 : 쑤저우대학교 졸업(공학박사)
 1998년~현재 : 동명대학교 정보통신공학과 부교수
 관심분야 : 휴먼인터페이스, 퍼지 속도·적분, 평가
 Phone : 051-610-8374
 Fax : 051-610-8846
 E-mail : yssohn@tu.ac.kr