

工學碩士 學位論文

人工知能 技法을 利用한 임베디드형
多重生體 認識시스템 具現에 관한 研究

A Study on the Implementation of Embedded System for
Multi-Modal Biometrics Recognition using Artificial Intelligence

指導教授 李 尙 培

2005年 2月

韓國海洋大學校 大學院

電子通信工學科

張 元 一

工學碩士 學位論文

人工知能 技法을 利用한 임베디드형
多重生體 認識시스템 具現에 관한 研究

A Study on the Implementation of Embedded System for
Multi-Modal Biometrics Recognition using Artificial Intelligence

指導教授 李 尙 培

2005年 2月

韓國海洋大學校 大學院

電子通信工學科

張 元 一

本 論 文 을 張 元 一 의
工 學 碩 士 學 位 論 文 으 로 認 准 함

委 員 長 朴 東 國 印

委 員 孫 慶 洛 印

委 員 李 尙 培 印

2005年 2月

韓 國 海 洋 大 學 校 大 學 院

電 子 通 信 工 學 科

목 차

Abstract	i
제 1 장 서 론	1
제 2 장 화자인식 알고리즘	3
2.1 음성 추출	4
2.2 음성특징정보 추출	5
2.3 화자인식 알고리즘	7
제 3 장 지문인식 알고리즘	10
3.1 전처리	11
3.2 지문특징정보 추출	17
3.3 지문인식 알고리즘	24
제 4 장 임베디드형 다중생체 인식시스템의 구성	26
4.1 하드웨어의 구성	26
4.2 소프트웨어의 구성	30
제 5 장 실험결과 및 고찰	34
5.1 음성특징 추출	34
5.2 지문특징 추출	35
5.3 인식률 테스트	37
제 6 장 결 론	43
참 고 문 헌	44

Abstract

The established single-biometric recognition systems which are fingerprint, face, speech, iris, and what not have many problems of FAR(False-Acceptance Ratio) and FRR(False-Rejection Ratio). This research studied embedded multi-modal biometrics recognition system using unique features of voice and fingerprint in order to increase reliance and complete that problems.

The whole system that was implemented divided into three parts, fingerprint recognition part, speaker recognition part, control part. First, voice that is inputted from microphone extract features of voice through pre-emphasis, hamming window, MFCC(Mel-Frequency Cepstrum Coefficient) after voice area was detected by short time energy. The extractive features of voice area are stored to database and achieves recognition process by DTW(Dynamic Time Warping) algorithm using voice features at speaker recognition process. Second, fingerprint image that is inputted from AFS-8500 that is a semi-conductor fingerprint sensor achieves LPF(Low Pass Filter), histogram equalization, binarization, thinning process and makes direction pattern and detects singular point of fingerprint. After this process is completed, it achieves feature points detection of fingerprint and post-process. It is trained feature points and singular points of abstracted fingerprint using autonomic neural network KSOM(Kohonen Self Organizing Maps) that is one of Artificial Intelligence techniques. If training process of all feature data is completed, KSOM achieves

recognition process about fingerprint. Third, the control part controls whole action of fingerprint recognition part and speaker recognition part.

As a result of this research, the performance of this research dropped quantity in whole recognition rate than single recognition system. However, the performance is displaying that improved much about FAR that is the most important urea of recognition system. Finally we may have to develop recognition system that uses various biometrics information and increase stability more for research direction hereafter.

제 1 장 서 론

사회가 점차 고도화 되고 복잡화되면서 일상 업무뿐만 아니라 보안을 유지해야 하는 정보들의 필요성이 증가하고 있다. 특히 전자 상거래와 같은 온라인상에서 개인 인증 번호(Personal Identificaton Number)나 비밀번호>Password)를 이용하여 사용자를 인증하는 분야에서는 이와 관련된 보안사고가 빈번히 발생하고 개개인의 중요한 정보가 타인에 의해 도용되거나 파괴되는 심각한 보안문제가 제기되고 있다. 이의 해결방법으로 많은 정보 보호 기술이 개발되고 있으며, 그 중 하나의 방법인 신체의 특성을 이용하는 생체인식기술이 매우 중요한 분야로 대두되고 있다.

생체인식기술은 안정성 측면에서 다른 개인 인증 기법보다 훨씬 높은 기술적 우위를 가지고 있다. 그러므로 21세기는 “생체인식의 시대”라는 말이 나올 정도로 생체인식산업은 인터넷을 기반으로 하는 사업이 확대됨에 따라 기존의 단순한 물리적 보안장비에서 생체인식기술을 이용한 인식시스템 구축 및 이에 대한 라이선스 비즈니스 형태로 발전되고 있으며 향후에는 인터넷과 이동통신을 이용한 무선인증방식으로 더욱 발전될 것이다.

생체정보의 종류에는 지문, 얼굴, 음성, 서명, 홍채, 손 혈관, 걸음걸이 등 생리학적 특징과 행동적 특징을 이용한 방법들이 있다. 하지만 여기서 언급한 어떤 생체정보도 보편성, 유일성, 영구성, 정확성 등을 모두 만족한다고 할 수 없으며, 각 생체특징마다 환경이나 기타 요소들에 의하여 취약점을 가질 수 있다. 그러므로 안정성을 증가시키기 위하여 다중생체인식을 생각해 볼 수 있는데 지문과 얼굴, 홍채와 지문, 손 정맥과 홍채 등 두 가지 이상의 생체 정보를 함께 인식함으로써 오인식에 대한 에러율을 감소시킬 수 있다. 또한 지문과 음성의 경우 가장 보편적인 생체인식

기술로서 거의 유일한 특징을 가지고 있다.

본 논문에서는 이러한 단일인식시스템들의 특징들을 활용하여 신뢰성을 증가시킬 수 있는 간단하면서도 높은 인식률을 가진 지문과 개개인이 가지고 있는 독특한 음성을 이용하여 임베디드형 다중생체 인식시스템을 구현하였다.

전체시스템은 크게 세부분으로 지문인식부, 화자인식부, 제어부로 나뉜다. 화자인식부에서는 TMS320C32 DSP를 사용하여 화자인식을 수행하고, 이 후 지문인식부의 TMS320VC5509 DSP에서 특징점을 추출 후 인공지능 기법의 하나인 KSOM(Kohonen Self Organizing Map)^{[11]-[13]}신경망 알고리즘을 이용하여 인식을 수행하였다. 각 인식부에 대한 전체적인 제어는 ATMEL사의 Atmega16L을 사용하였다.

본 논문은 다음과 같이 구성되어 있다.

2장에서는 기존의 화자인식 알고리즘 중에서 실시간 처리가 가장 적합한 DTW(Dynamic Time Warping)^{[4]-[6]}알고리즘을 적용한 방법을 이용하여 직접 설계된 내용을 서술하며, 3장에서는 지문인식을 수행하기 위한 지문 인식 알고리즘을 제안하는데 있어 인공지능 기법인 KSOM 신경망 및 여러 가지 지문처리 알고리즘을 설계 순으로 설명을 한다.

4장에서는 본 논문에서 제안된 다중생체 인식시스템인 화자/지문 인식 시스템의 설계된 내용을 자세히 다루며, 이어서 5장에서는 4장에서 제안된 여러 가지 이론의 실험 및 결과를 보여준다. 끝으로 6장에서 결론으로 이 논문의 끝을 맺는다.

제 2 장 화자인식 알고리즘

기존의 화자인식 알고리즘은 HMM(Hidden Markov Models), DTW, VQ(Vector Quantization), GMM(Gaussian Mixture Model) 등 여러 가지 알고리즘을 이용하여 개발되었으며, 성능면에서는 GMM이 가장 좋은 방법으로 검증되었다^[4]. 그러나 GMM은 훈련과정을 통해서만 인식함으로써 실제 시스템에 적용하기 위해서는 많은 문제점이 있다.

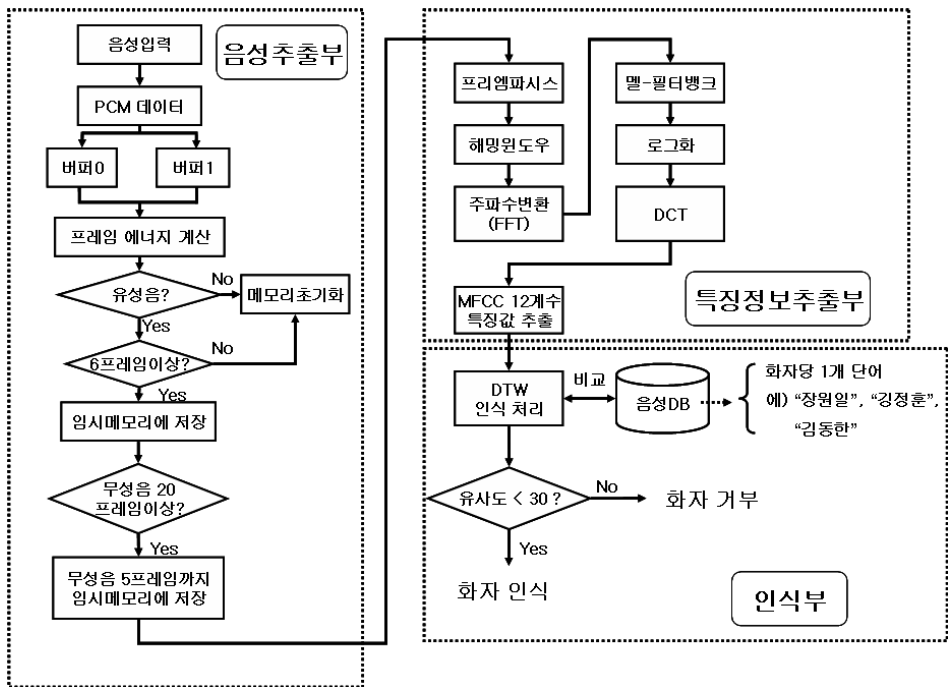


그림 2.1 화자인식 알고리즘의 흐름도

본 논문에서 제안하는 화자인식 알고리즘은 실시간 화자인식을 위해 간단하면서도 성능이 뛰어난 DTW 알고리즘을 적용하였으며, 전체 화자인식

알고리즘은 그림 2.1과 같다. 각 절에서는 그림 2.1의 부분적인 알고리즘을 설명한다.

2.1 음성 추출

음성의 검출은 화자인식 성능에 큰 영향을 미친다. 이것은 마이크를 통해서 들어온 음성데이터로부터 실제 음성영역만 검출하는 것이다. 여기에는 검출점으로부터 영교차율을 측정해서 교차율이 높으면 유성음으로, 교차율이 낮으면 무성음으로 판별하는 방식인 영교차율(ZCR : Zero Crossing Rate) 방식과 음성 신호의 구간 당 에너지를 계산해서 일차적으로 유성음 부분만 검출하는 절대 에너지 방식(Short Time Energy)이 있다^{[4][6][7]}.

본 논문은 실시간으로 음성을 처리하기 때문에 실시간에서 비교적 성능이 떨어지는 영교차율을 적용하지 않았고, 절대에너지 방식을 사용하고 있다. 절대 에너지는 무성음보다 유성음부분이 크다는 이론을 바탕으로 하고 있다. 한 프레임의 절대에너지는 식 (2.1)과 같이 정의된다.

$$E = \sum_{n=0}^{N-1} [x(n)]^2 \quad (2.1)$$

$x(n)$ 은 마이크에서 입력받은 음성데이터, N 은 프레임의 길이, E 는 절대에너지를 나타낸다. 각 프레임의 절대 에너지가 설정된 값보다 크면 이 구간을 실제 음성으로 간주한다.

본 논문에서는 음성프레임 길이를 100으로 설정하고, 데이터의 손실을 방지하기 위해서 프레임을 중복시키는 방식인 프레임 블록킹을 30으로 적용하고 있다. 음성구간추출은 절대 에너지가 설정된 값보다 큰 프레임이 6개 이상이 되면 그 부분은 유성음 구간으로 간주하여 메모리에 저장이

되고 이 후 절대 에너지가 설정된 값보다 작은 프레임이 20개 이상이 되는 무성음 구간까지 음성의 끝점을 검출하여 저장하였다. 아울러 끝점이 검출되면 20프레임의 무성음 구간 중 15프레임은 버리고, 나머지 5프레임만을 음성영역으로 사용하였다.

2.2 음성특징정보 추출

음성 데이터양은 인식 알고리즘을 실시간으로 처리하기에는 상당히 크므로 이 데이터를 효율적으로 줄여야 한다. 이러한 처리를 위해 특징정보 추출 과정이 필요하다. 본 논문에서는 MFCC(Mel-Frequency Cepstrum Coefficient) 방식을 사용하여 음성의 특징정보를 추출하였다. MFCC의 처리 절차는 그림 2.1의 특징정보추출부에 나타난 것과 같다. 프리엠퍼시스(Pre-emphasis) 처리 후, 해밍 윈도우(Hamming Window)를 씌우고 FFT(Fast Fourier Transform)을 수행하여 주파수영역으로 변환한다. 변환된 값들을 그림 2.2와 같이 미리 설정된 MFCC(멜 캡스트럼) 대역필터뱅크를 통과시키고, 로그화, DCT(Distance Cosine Transform) 을 거치면 프레임당 12개의 계수값들을 구할 수 있다. 이 값들이 음성의 특징정보들이다.

2.2.1 프리엠퍼시스

검출된 음성영역의 저주파 성분을 약화시키고 고주파 성분만을 강조시켜 마이크에서 많이 발생된 음성신호의 DC성분을 제거하는 방식이다. 프리엠퍼시스는 식 (2.2)와 같다^{[4][7]}.

$$\bar{S}(n) = S(n) - aS(n-1), \quad 0.9 < a < 1.0 \quad (2.2)$$

$S(n)$ 은 검출된 음성데이터이고, 프리엠파시스 계수 a 는 0.95를 사용했을 때 최적의 결과가 나왔다.

2.2.2 해밍 윈도우

프리엠파시스 처리를 한 음성데이터는 각 프레임의 시작과 끝에서 신호의 불연속이 발생하는데 이것을 최소화시키기 위해 해밍윈도우를 사용한다. $x(n)$ 은 프리엠파시스 처리된 음성신호이고 해밍윈도우 $w(n)$ 은 식 (2.4)와 같이 표현된다. 여기서 N 은 프레임의 길이이다^{[4][7]}.

$$\tilde{x}(n) = x(n) \cdot w(n) \quad (2.3)$$

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad (0 < n < N-1) \quad (2.4)$$

2.2.3 MFCC

캡스트럼 분석 방법 중 인간의 청각 특성을 이용한 것으로 멜 캡스트럼이 있다. 멜은 톤 신호의 인지된 피치 또는 주파수 측정치를 나타내는 단위로서, 인간의 청각 시스템이 피치를 선형적으로 인지하지 않는 것처럼 톤 신호의 물리적인 주파수에 선형적으로 대응하지 않는다. Stevens와 Volkman은 임의로 1000Hz를 1000mel로 선정하고, 사람들에게 인지된 피치가 기준 주파수의 두 배가 되도록 변화시켜 이를 2000mel로 표기하였다. 이러한 방법으로 실제 물리적인 주파수와 인지된 주파수사이의 대응 관계를 구하였다. 이 대응 관계를 보면 1kHz 이하에서는 선형적으로 1kHz 이상에서는 대수적(log scale)으로 대응됨을 볼 수 있다. 멜과 주파수사이의 대응 관계를 근사적으로 나타낼 수 있는데 식 (2.5)와 같다. 여기서 F_{mel} 는 각각 근사식에 의해 구해진 인지된 주파수이고, F_{Hz} 는 실제주파수

를 나타낸다. 그림 2.2은 멜 캡스트럼 삼각대역 필터를 나타낸다.^{[4][6][7][8]}.

$$F_{mel} = 2595 \log_{10} \left(1 + \frac{F_{Hz}}{700} \right) \quad (2.5)$$

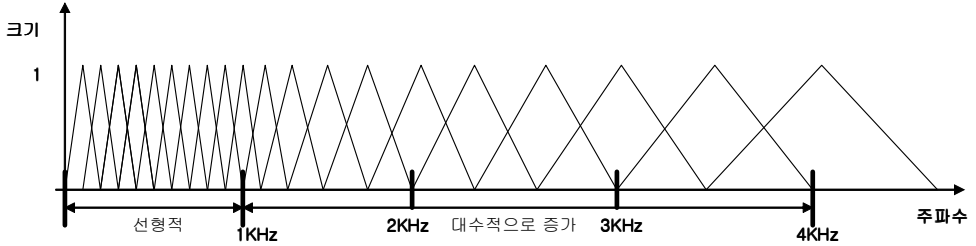


그림 2.2 멜 캡스트럼 삼각대역 필터

해밍 윈도우 처리가 된 음성데이터를 FFT를 통해 주파수영역으로 변환 한 뒤 그림 2.2의 멜 캡스트럼 삼각대역 필터를 통과한다. 통과한 신호는 로그화를 거친 후 식 (2.6)을 통해 프레임당 n개의 멜 캡스트럼 계수 값이 나오게 된다.

$$C_n = \frac{1}{20} \sum_{k=1}^{20} X_k \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{20} \right] \quad (2.6)$$

여기서 X_k 는 로그화를 거친 음성신호이고 K는 필터의 개수이며 n는 멜 캡스트럼 계수의 차수이다. 본 논문에서는 12차의 멜 캡스트럼 계수 C_n 을 구하였다.

2.3 화자인식 알고리즘

DTW(Dynamic Time Warping) 는 입력패턴과 참조 패턴 사이의 거리

를 측정해서 그 유사도를 측정하는 방법이다. 다시 말하면, 제한된 경로 내에서 단조 증가를 통해서 가장 가까운 거리를 판별, 유사도를 측정한다. 예를 들어, 길이가 M 인 입력 음성 패턴을 $T = T(1), T(2), \dots, T(M)$, 길이가 N 인 기준 패턴을 $R = R(1), R(2), \dots, R(N)$ 이라고 하면 두 패턴 간의 유사도 d 는 다음 식 (2.7)과 같이 누적거리로 표현된다^{[4][5][7]}.

$$d = \sum_{n=1}^N D(R(n), T(W(n))) \quad (2.7)$$

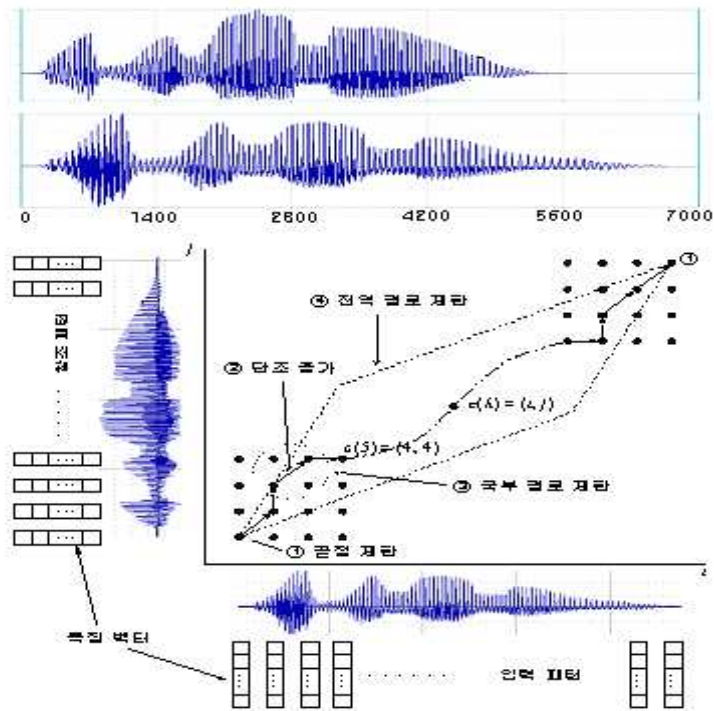


그림 2.3 DTW의 제약된 조건

이 때 $D(R(n), T(W(n)))$ 는 R 의 n 번째와 T 의 $W(n)$ 번째의 국부적

유사도(Local Distance)이며, DTW는 두 패턴간의 누적 거리 최적화를 하는 (m, n) 평면의 최적 경로 $m = W(n)$ 를 찾는 방법이다. 이 방법은 음성 신호의 특성을 고려해서 최적 경로 탐색에 다음과 같은 제약조건을 가한다. 이것은 그림 2.3에 잘 나타나 있다.

- 끝점 제한 (Endpoint Constraints)
- 단조 조건 (Monotonic Condition)
- 국부 경로 제한 (Local Path Constraints)
- 전역 경로 제한 (Global Path Constraints)
- 기울기 가중치 (Slop Weighting)

본 논문에서는 지역거리는 유클리드 거리법을 사용하고 있으며, 지역 제약방법은 ITAKURA방식^[4]을 사용하고 있다.

DTW의 경우 기준 모델 집합의 작성은 간단하다. 인식하고자 하는 명령어들을 발음하고 분석한 후 연속된 프레임들의 특징 벡터들을 저장하고 있으면 된다. 인식 시에는 입력된 음성을 분석해 특징 벡터를 추출한 후 이들 기준 모델 집합의 구성원과 개별적으로 DTW를 수행하여 가장 적은 누적 거리를 주는 구성원을 찾으면 된다. 보통 음성인식에서는 DTW 알고리즘을 이용하여 고립단어 인식에 많이 사용하지만, 화자인식에 사용하기 위해서 본 논문에서는 패턴 매칭의 유사도(Distance)값만 이용하여 구현하였다.

제 3 장 지문인식 알고리즘

특징점을 이용한 지문 인식 알고리즘은 지문인식 센서로 획득한 지문 이미지에서 특징점을 추출하여 특징정보를 정의하는 추출과정(Feature Extraction)과 정의된 특징정보로 두 지문이미지를 비교하여 유사도를 판정하는 정합과정(Minutiae Matching)으로 이루어진다. 본 논문에서는 특징 추출과정은 규격화 되어 있는 알고리즘을 이용하여 특징값을 추출하였으며, 인식과정을 인공지능 기법인 자율신경망 KSOM(Kohonen Self Organizing Maps)을 이용하여 실시간 지문인식시스템을 구현하였다.

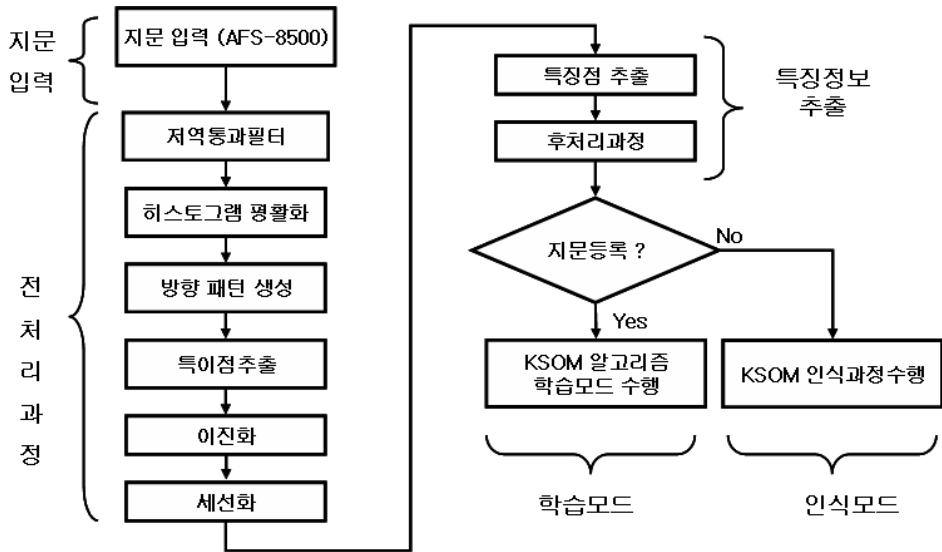


그림 3.1 지문인식 알고리즘의 전체 흐름도

그림 3.1에서는 본 논문에서 사용한 실시간 지문인식시스템에 구현된 알고리즘을 나타낸 것이며, 전처리 과정부, 특징정보 추출부, 인식부로 나눌 수 있다. 각 절에서는 그림 3.1의 부분적인 알고리즘을 설명한다.

3.1 전처리

전처리 과정은 입력된 지문영상에서 잡음을 제거하여 영상을 보정하고 난 뒤에 지문영상의 방향성 및 이진화, 세션화처리를 하게 된다.

3.1.1 영상의 보정

입력된 지문영상은 고주파 잡음과 융선이 끊어지는 오차와 서로 다른 평행한 융선이 하나로 겹쳐지는 오차를 보정하기 위해서 두 단계의 필터 처리를 해준다. 첫 번째는 256레벨에서의 영상처리로 지문에서 잡음제거와 두 번째는 영상 전체에서 256레벨이 균일하도록 조정하는 것이다.

1) 저대역 통과 필터

주파수 공간에서 고주파와 저주파의 성분을 조작하여 고주파 성분을 분리해내고 저주파 성분만을 얻어내는 필터이다. 저대역 통과 필터는 전체적으로 영상의 급격한 변화를 완화시키는 역할을 한다. 그러므로 검은색 선(융선)과 흰색 선(골)의 사이에 들어가 있는 잡음을 어느 정도 제거할 수 있다. 식 (3.1)이 저주파 통과 필터 수식으로써 D_0 는 원점으로부터의 거리로서 차단 주파수이고, n 이 커지면 주파수 반응의 정도가 커지게 된다^{[9]-[11]}.

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)}{D_0} \right]^{2n}} \quad (3.1)$$

$$D(u, v) = \sqrt{(u^2 + v^2)} \quad (3.2)$$

여기서, $D(u, v)$ 는 주파수 평면의 점 (u, v) 로부터 원점까지의 거리이

고, u 와 v 는 주파수 평면의 좌표를 말한다.

2) 히스토그램 평활화

지문을 입력할 때 주변의 환경과 입력하는 사람의 손가락 특성, 눌러지는 압력에 따라 256레벨의 분포가 다르게 나타나게 된다. 즉, 지문 채취시 손가락의 힘이 영상에 전체적으로 골고루 분포되지 않기 때문에 부분별로 밝기가 달라진다. 그러므로 지문 이미지의 명암 값의 분포가 한쪽으로 치우치거나 균일하지 못한 영상을 히스토그램 균일화 처리에 의해 명암 값의 분포가 균일화되어 영상이 향상될 수 있다. 히스토그램 평활화를 구현하려면 다음과 같은 3단계를 거쳐야 한다.

- ① 히스토그램을 이용한 명도 값의 빈도수를 계산한다.
- ② 구한 빈도수를 이용하여 축적 히스토그램 값을 구하고 정규화한다.
- ③ 정규화된 축적 히스토그램을 그레이 스케일 사상 함수로 이용하여 그레이 레벨 값을 매핑한다. 축적 히스토그램의 정규화 수식은 식 (3.3)와 같다.

$$k_i = \frac{g_{\max}}{n_i} H(i) \quad (3.3)$$

여기서 n_i 는 지문영상에서의 픽셀의 총 개수이고, g_{\max} 는 명도의 최대값이며, $H(i)$ 는 위에서 구한 축적 히스토그램이다^{[9][11][12]}.

3.1.2 방향성 추출

지문영상의 방향성을 추출하기 위한 방법으로는 Mehtre 방법, Ratha 방법, 창틀 마스크 방법, Sobel 연산자 방법이 있다^{[10][12]}. 본 논문에서는 4

가지 방법 중 시간 대 성능비가 가장 우수한 Sobel 연산자 방법을 사용하였다.

용선의 방향성 추출과정은 지문영상을 16×16의 블록으로 나눈 다음 각 블록에서의 각 화소에 대한 방향을 Sobel 마스크에 의해 구하여, 그림 3.2와 같이 8방향으로 양자화 시킨 후에 그림 3.3에 나타낸 것과 같이 가장 많이 추출되는 방향을 각 블록의 대표방향으로 결정한다^{[9][11]}.

지문의 방향성을 계산하기 위한 Sobel 마스크의 크기는 3×3으로 구성했다. 3×3마스크는 그림 3.4와 같으며, 각 마스크에서의 기울기의 크기인 S와 방향각인 θ 는 식 (3.4)과 (3.5)에 나타내었다. 여기서 G_x 는 가로방향의 차분값이고, G_y 는 세로방향의 차분값이다.

$$G_y = (g + 2h + i) - (a + 2b + c) \quad (3.4)$$

$$G_x = (c + 2f + i) - (a + 2d + g)$$

$$S = \sqrt{G_x^2 + G_y^2} \quad (3.5)$$

$$\theta = \tan^{-1}(G_y / G_x)$$

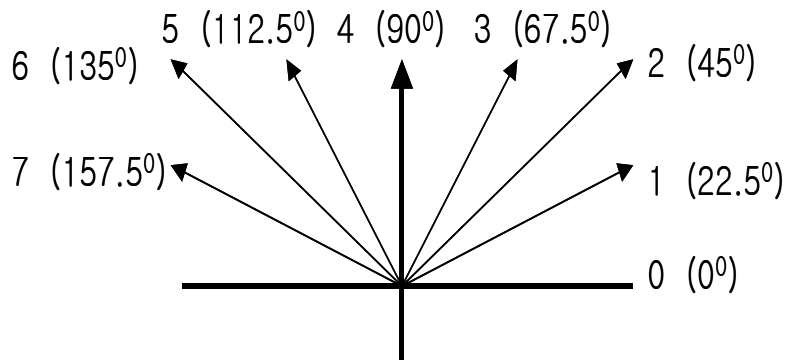


그림 3.2 8-방향성 패턴

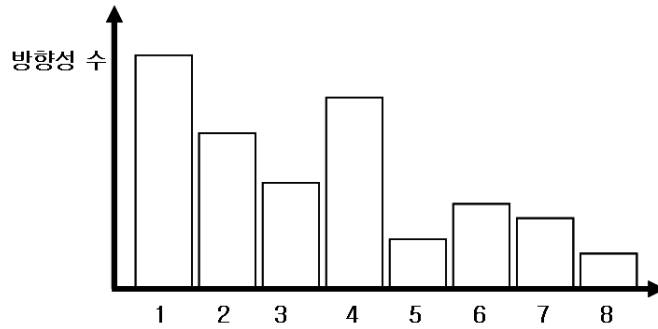


그림 3.3 Sobel 연산에 의해 구해진 블록의 대표 방향 설정

a	b	c
d	e	f
g	h	i

그림 3.4 3×3 크기의 Sobel 마스크

3.1.3 방향성 평활화

지문의 방향성을 구하면 지문의 품질이 좋은 영역에서는 실제 용선의 흐름과 일치하는 방향성이 생기는 반면에 지문의 정보가 손상된 영역에서는 용선의 흐름과 무관한 방향성이 생성된다. 그래서 지문의 방향성 평활화(smoothing)를 수행하여야 한다. 지문의 방향성을 평활화하기 위해서는 다음과 같은 두 단계의 필터처리를 해주어야 한다^[3].

- ① 3×3 창안에서 중심 블록에 대한 이웃하는 여덟 블록의 방향성 중 다섯 개 이상의 방향이 같은 방향일 때, 중심 블록의 방향성을 다수의 방향성으로 치환하게 되는 필터이다.
- ② 3×3 창안에서 중심 블록과 이웃하는 여덟 블록의 방향성의 평균값

을 구하여 중심블록의 방향성으로 치환하는 필터이다. 방향성의 평균을 구할 때는 항상 두 방향성이 이루는 각 중에 작은 각의 평균을 구해야 한다.

3.1.4 이진화

이진화는 디지털 영상을 검정색과 흰색 두 개의 값으로만 표현하는 기법으로 명도 값으로 표현할 경우 0 또는 255의 값만을 갖는 것이고, 일반적인 표현으로는 0 또는 1의 값만을 갖는 것을 말한다. 이진화 처리를 위해서는 히스토그램을 이용하여 어떤 임계값을 정하고 픽셀의 명도값이 임계값보다 작으면 명도 값 0(검정색)으로, 임계값보다 크면 명도 값 255(흰색)로 만드는 과정이다. 이때 가장 중요한 것이 임계값의 선택 문제이다. 일반적으로 지문 센서에 의한 입력된 지문영상은 손가락의 압력과 조명상태에 따라서 명암이 크게 차이를 보인다. 그러므로 전체적인 임계값을 이용하여 이진화를 할 경우 정확한 이진화 처리가 되지 않는다. 이를 위하여 지역적인 임계값을 구해서 지역적으로 분할해서 수행하여야 한다. 임계값이 임의의 값을 가질 경우 배경 영역이 큰 이미지에서는 배경을 효과적으로 구분하지 못하는 경우가 발생하기 때문에 임계값의 범위를 지정해 주어야 한다. 일반적으로 임계값의 범위는 0에서 255 사이이다^{[9][11]}.

본 논문에서는 실험을 통해 임계값을 78로 설정했을 때, 지문영상의 이진화가 가장 잘 처리되었다.

3.1.5 세션화

지문영상의 세션화는 효과적인 지문의 특징점 추출을 위해 지문 융선의 두께를 1픽셀로 고정시켜주는 작업을 의미한다. 지문영상의 경우는 거의

일정한 폭과 완만한 경사를 지닌 용선의 흐름으로 구성되어 있으므로 두께를 1픽셀로 고정시켜 표시하기가 매우 적합한 형태라고 볼 수 있다. 세선화에 대한 효율성은 기존의 문자인식 등의 응용에 이미 증명된 바 있으며, 본 논문에서는 특징점을 추출하기 위하여 세선화 알고리즘을 선택하였고, 세선화 알고리즘 기법은 다음과 같다.^{[9][11][13]}.

- 1) 2개 이상, 6개 이하 이웃 픽셀이 있어야 한다. 이 조건은 끝점이나 영역 내부의 점은 지워서는 안 된다는 것을 의미한다.
- 2) 두 개 영역을 연결하는 픽셀은 지워서는 안 된다. 이 픽셀을 지운다면 영역 개수가 달라지게 된다.

그림 3.5는 세선화 알고리즘에 사용하는 픽셀들의 배열을 마스크 형식으로 나타내고 있다. 첫 번째 단계는 그림 3.5에 정의된 세선화 마스크로부터, 다음의 조건을 만족하면 외곽선 점 M을 삭제하기 위해 표시해 놓는다.

- ① $2 \leq N(m_1) \leq 6$
- ② $S(m_1) = 1$
- ③ $m_2 \cdot m_4 \cdot m_6 = 0$
- ④ $m_4 \cdot m_6 \cdot m_8 = 0$

여기서, $N(m_1)$ 은 m_1 의 0이 아닌 값을 가진 이웃의 수이다. 즉,

$$N(m_1) = m_2 + m_3 + \dots + m_8 + m_9$$

이고 $S(m_1)$ 은 $m_2, m_3, \dots, m_8, m_9$ 의 정렬된 순서에서 0에서 1로 천이의 회수이다. 천이의 회수라는 말은 0과 1이 바뀌는 수를 말한다. 두 번째 단계는 ③, ④의 조건이 다음과 같이 바뀐다.

$$\textcircled{3} \quad m_2 \cdot m_4 \cdot m_8 = 0$$

$$\textcircled{4} \quad m_2 \cdot m_6 \cdot m_8 = 0$$

m ₉	m ₂	m ₃
m ₈	m ₁	m ₄
m ₇	m ₆	m ₅

그림 3.5 세션화 마스크

첫 번째 단계는 고려 대상의 2진 영역 내에 있는 모든 경계 화소에 대해서 적용된다. 만약 조건 ①에서 ④중에 한 개 또는 그 이상이 조건을 만족하지 않으면, 그 대상 점에 대한 값은 변화되지 않는다. 만약 모든 조건이 만족되면 그 점은 삭제를 위해 표시되어 진다. 그러나, 그 점은 모든 경계점들에 대한 처리가 수행되기 전까지는 삭제되지 않는다. 이 지연은 알고리즘이 수행되는 동안에 데이터 구조가 변화되지 않도록 한다. 첫 번째 단계의 모든 경계점에 대해서 적용된 후, 삭제로 표시되었던 점들이 삭제된다. 그 다음, 두 번째 단계는 첫 번째 단계와 똑같은 방법으로 이 결과 데이터에 적용되어진다.

3.2 지문특징정보 추출

입력된 지문영상의 특징정보인 특이점과 특징점을 추출하는 과정으로 구성된다. 특이점을 추출하고 단점, 분기점으로 구성되어지는 특징점을 구한 뒤에 잘못 추출된 특징점을 후처리 과정을 통해서 보정하게 된다.

3.2.1 특이점 추출

지문영상의 방향성을 이용하여 특이점(singular point)을 추출할 수 있으며, 특이점을 추출하는 목적은 데이터베이스에 등록된 지문의 방대함에 서 특정한 범주로 분류하여 정합에 사용할 대상 지문의 개수를 최소화하는데 있다. 이러한 특이점은 Poincare 지수^{[13]-[15]}를 계산함으로써 보통점, 중심점, 삼각주로 추출할 수 있다. Poincare 지수를 구하는 방법은 그림 3.6과 같이 특정 블록에서 이웃하는 여덟 블록의 방향을 시계 반대 방향으로 돌면서 방향각이 0° 일 경우 보통점, 180° 일 경우 중심점, 그리고 -180° 일 경우 삼각주라고 한다. 이러한 방향각을 계산하기 위한 Poincare 지수를 그림 3.7에 나타내었으며, 식 (3.6)에 의하여 계산한다.

$$\text{방향각} = \sum_{k=1}^9 \{P_k - P_{k+1}\} \quad (3.6)$$

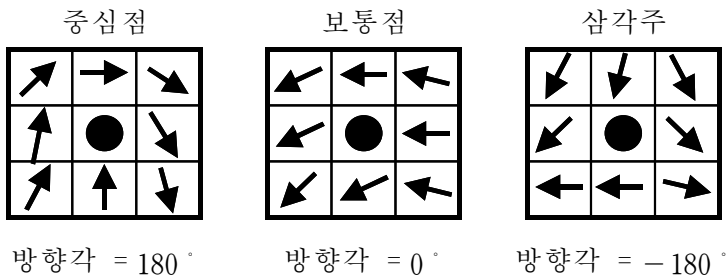


그림 3.6 특이점에 대한 방향각

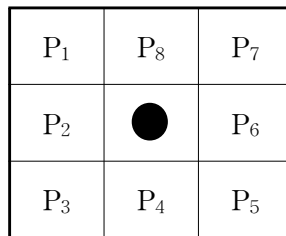


그림 3.7 Poincare 지수

3.2.2 특징점 추출

지문영상을 구성하고 있는 용선은 일정한 흐름을 깨는 비연속점들로 구성되어지고, 이를 특징점이라고 부른다.








	단점
	분기점
	고림(lake)
	짧은용선
	점
	갈라진 분기(spur)
	교차점

그림 3.8 특징점의 종류

특징점은 개개인마다 다르며 타인의 지문과 구별할 수 있는 중요한 요소가 된다. 일반적인 지문인식시스템에서 많이 쓰이는 특징점들로는 단점(ending point)과 용선이 갈라지는 분기점(bifurcation point)이 사용된다. 이러한 특징점의 형태를 그림 3.8에서 나타내었다. 지문의 특징점을 구하기 위해 그림 3.9와 같은 3×3 마스크를 사용한다.

M(1)	M(2)	M(3)
M(8)	S	M(4)
M(7)	M(6)	M(5)

그림 3.9 특징점 구분을 위한 3×3 마스크

이러한 특징점들은 식 (3.7)에 의해 구해진다.

$$FE = \frac{1}{2} \sum_{i=1}^8 |M(i) - M(i+1)| \quad (3.7)$$

표 3.1 특징점의 속성

FE	속성
0	고립점(isolated point)
1	단점(ending point)
2	융선(ridge point)
3	분기점(bifurcation point)
4	교차점(crossing point)

여기서 M 은 마스크를 나타내고, i 는 그림 3.9의 순서로 식 (3.7)을 이용하면 $M(9) = M(1)$ 이다. 즉, 선을 이루는 화소의 값을 '1'로, 배경 화소의 값을 '0'으로 하고 화소의 값이 '1'인 중심점 S 를 기준으로 해서 마스크연산을 할 경우 점, 단점, 분기점은 각각 '0', '1', '3'의 FE(Feature Extraction) 값을 갖는다. FE의 속성은 표 3.1과 같이 나타낼 수 있다. 표 3.1에서 보여진 속성을 이용하며 단점과 분기점은 특징점에 해당하며, 고립점(isolated point)과 교차점(crossing point)은 후처리과정을 통해 보정되어야 하는 의사 특징점에 해당된다^{[16][17]}.

3.2.3 지문영상의 후처리과정

특징점 추출 과정에 의해 추출된 특징정보에는 손상된 지문 이미지와

이미지 처리 오류에 의한 가짜 특징점들이 포함되어 있다. 이러한 가짜 특징점들은 후처리과정을 통해 제거되어야 한다. 먼저, 용선 사이의 평균 거리를 λ 라고 하고 이것을 계산한다. 계산된 용선 사이의 평균 거리 λ 는 잡음 제거 과정에서 사용되는 임계치를 기준으로 하여 알고리즘에 지문의 다양한 변화에 대한 유연성을 준다. 후처리과정의 알고리즘 순서는 다음과 같다^[13].

1) 끊어진 용선을 연결하는 방법은 다음과 같다.

- ① 평균 용선거리 λ 에 의한 $\lambda \times \lambda$ 창을 구성한다.
- ② 단점이 두 개 이상인 영역을 탐색하여 찾는다.
- ③ 각 단점의 방향성을 구한다.
- ④ 각 단점의 좌표사이의 거리를 구한다.
- ⑤ 시작점(starting point)을 정한 후 3×3 블록을 이용하여 이웃의 픽셀 중 목적점과 가장 가까운 곳을 찾는다.
- ⑥ 찾은 곳으로 이동한 후 ⑤과정에 의해 가장 가까운 목적점(object point)을 찾은 후 그 곳으로 이동한다.

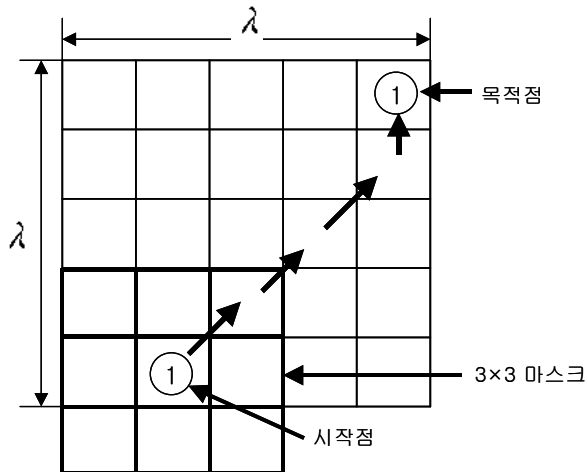


그림 3.10 끊어진 용선 연결

2) FE(Feature Extraction)가 0인 픽셀을 모두 제거한다(점 제거).

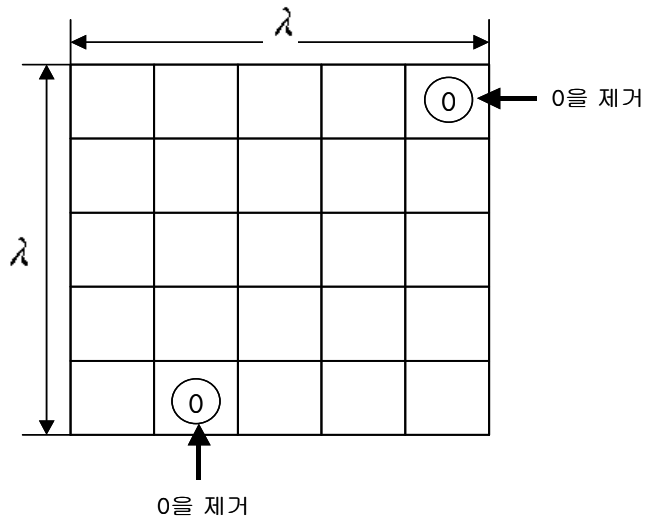


그림 3.11 점 제거

3) 용선의 길이가 $2 \times \lambda$ 보다 짧은 용선을 모두 제거한다.

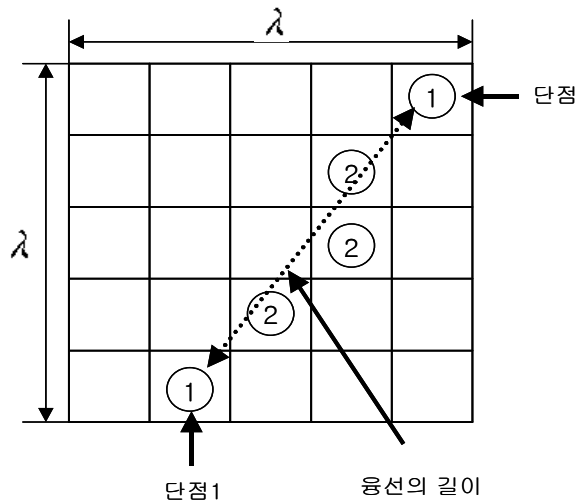


그림 3.12 짧은 용선 제거

4) 분기점의 가지 길이가 λ 보다 짧고 가지의 끝이 단점일 때 해당 가지를 제거한다(갈라진 분기점 제거).

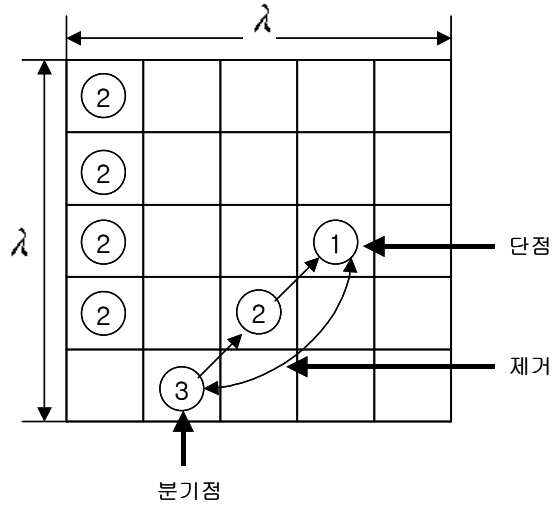


그림 3.13 갈라진 분기점(spur) 제거

5) 분기점의 가지 길이가 λ 보다 짧고 가지의 끝이 분기점일 때 해당 가지를 제거한다(고립점을 제거).

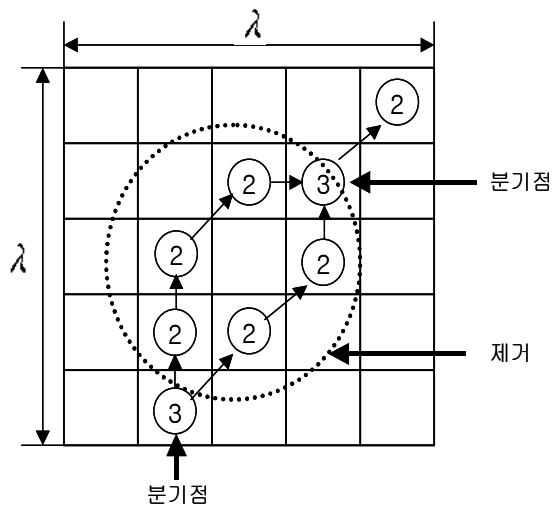


그림 3.14 고립점(lake) 제거

3.3 지문인식 알고리즘

본 논문은 인공지능 기법의 하나인 자율신경망 KSOM (Kohonen Self Organizing Maps)을 이용하여 지문의 분류 및 인식 알고리즘을 사용하였다. KSOM 신경망은 음성 인식, 문자 인식, 구문 분석 등 다양한 분야에 응용되는 자율 학습 신경망으로, 입력층과 출력층으로만 구성되는 순방향 단층 신경망 구조이다. KSOM 신경망의 학습 알고리즘은 그림 3.15와 같다.

KSOM 신경망의 학습에서 각 뉴런은 연결강도벡터와 입력벡터가 얼마나 가까운가를 계산한다. 그리고 각 뉴런들은 학습할 수 있는 특권을 부여받으려고 서로 경쟁하려는데 거리가 가장 가까운 뉴런이 승리하게 된다. 이 승자(winner)뉴런이 출력신호를 보낼 수 있는 유일한 뉴런이다. 또한 이 뉴런과 이와 인접한 이웃뉴런들만이 제시된 입력벡터에 대하여 학습이 허용된다. 이것은 학습에 있어서 전혀 새로운 접근 방식이다. 이 모델이 있기 이전에는 신경망에 있는 모든 뉴런들이 반복되는 훈련 과정에서 연결강도를 조정한다.

KSOM 신경망을 만들 때 다른 신경망들에서는 일반적으로 필요하지 않는 두 가지 일을 해야 한다. 하나는 층 내의 뉴런의 연결강도 벡터가 임의값을 가지면서 적합하게 초기화되어야 한다. 다른 하나는 연결강도 벡터와 입력벡터가 통상 0에서 1사이의 정규화(normalized) 된 값을 사용한다. 이런 두 가지 요인은 KSOM 신경망에 있어서 매우 중요하다.

본 논문에서는 지문에서 추출된 특징점들을 KSOM 신경망의 입력패턴으로 사용하였고 입력층 512개, 출력층 100개의 신경망을 구성하였고 연결강도와 입력패턴과의 거리는 유클리드 거리법을 사용하였다.

Step 1 : 연결강도 초기화

$w \leftarrow$ random 값 입력

Step 2 : 반경 r 및 학습률 α 입력

$r \leftarrow$ 정수값 입력

$\alpha \leftarrow 0 < \alpha < 1$

Step 3 : stop 조건이 만족하지 않을 시

do Step 4 - 8

Step 4 : 각 패턴 x 동안

do Step 5 - 8

Step 5 : 거리 계산 (유클리드 거리법)

$$D(j) = \sum_i (w_{ji} - x_i)^2$$

Step 6 : 승자 뉴런 y_j 를 검색

Step 7 : 반경안의 연결강도 갱신

$$w_{ji}^{k+1} = w_{ji}^k + \alpha [x_i - w_{ji}^k]$$

Step 8 : 학습률 및 반경 감소

Step 9 : stop 조건 테스트

그림 3.15 KSOM 신경망의 학습 알고리즘

제 4 장 임베디드형 다중생체 인식시스템의 구성

4.1 하드웨어의 구성

그림 4.1에서 보는 바와 같이 구현된 전체 시스템은 제어부, 지문인식부, 화자인식부로 3부분으로 구분되어진다. 지문인식부와 화자인식부에서는 미리 등록된 사용자 정보(지문패턴 및 음성패턴)와 입력된 음성 및 지문을 비교하여 그 인식여부를 제어부에 전송하면 제어부에서는 지문인식부와 화자인식부로부터 받은 인식정보를 비교하여 두 곳 모두에서 인식이 성공되면 그 정보를 LCD에 디스플레이하며 도어락이 열린다.

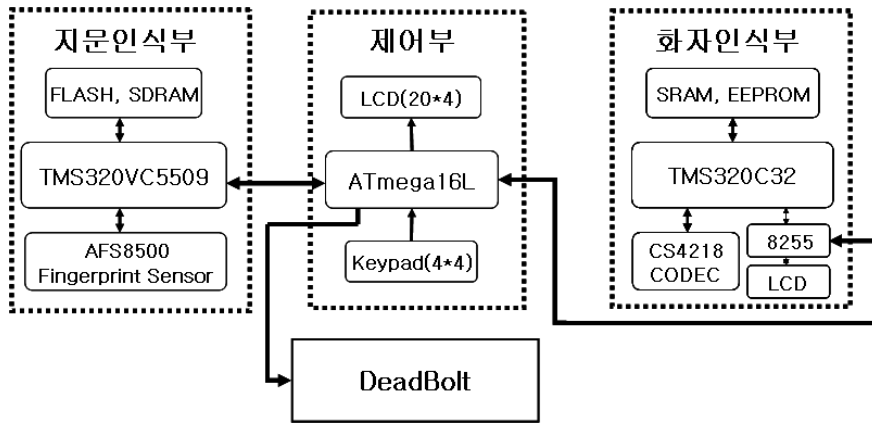


그림 4.1 전체 시스템의 구성

4.1.1 화자인식부의 하드웨어 구성

음성은 그 특성상 많은 데이터를 가지고 있기 때문에 특징정보를 추출하기 위해서는 빠른 산술연산이 필요하다. 본 논문에는 이런 점을 감안해

서, TI사의 플로팅 포인터 DSP인 TMS320C32(30MIPS)와 SRAM 61C256 * 4(128Kbyte)을 연결하여 많은 음성데이터를 연산하는데 충분한 메모리용량을 설계하였으며, 용량이 큰 음성프로그램과 음성패턴을 저장하기 위해 64Kbyte ROM인 27C512를 사용하고 있다^{[18][19][20][21]}. 프로그램의 다운로드에는 MPSD-PP 에뮬레이터를 통한 J-TAG 방식을 사용하고 있다. 시리얼 포트를 통해서 16비트 분해능을 가지고 있는 코덱(CS4218)과 연결, 음성 데이터의 입/출력을 담당하고 있으며, XF0와 XF1의 I/O포트를 통해서 CPU의 동작상태를 확인하고 있다. DSP의 외부인터럽트 단자 INT2와 INT3를 통해서 저장모드와 인식모드를 설정할 수 있게 했다. 그림 4.2는 화자인식보드의 전체적인 구성도를 나타내고 있다.

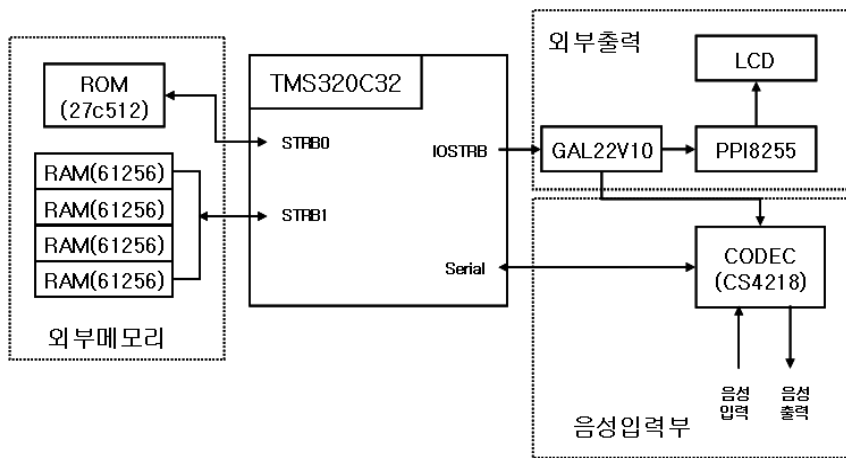


그림 4.2 화자인식보드의 구성

마이크로부터 입력된 음성데이터는 코덱을 거쳐 시리얼로 TMS320C32 DSP에 전달한다. 전달된 음성 데이터는 특징정보 추출을 위해 DSP내부에서 MFCC를 처리한 후 미리 저장된 음성패턴과 DTW 인식 알고리즘으로 유사도를 계산하여 화자인식과정을 수행한다. 인식결과는 LCD에 디스

플레이되며 전체시스템의 제어부에 그 결과를 전송한다. 그림 4.3은 실제 구현된 화자인식보드의 사진이다.

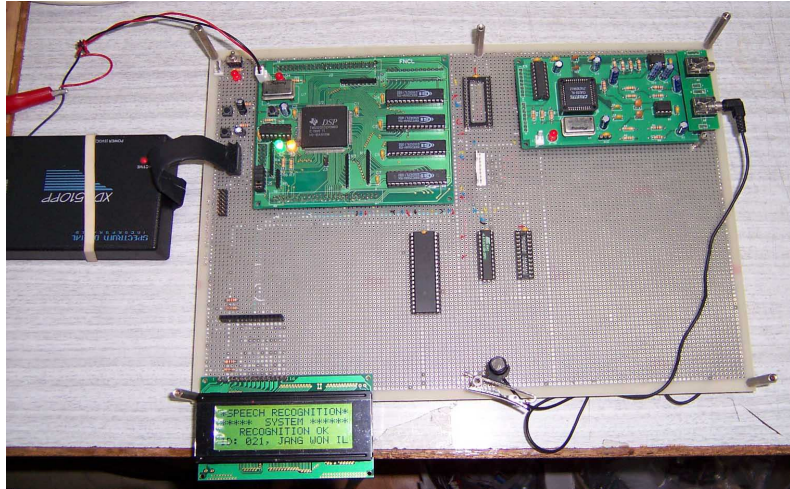


그림 4.3 구현된 화자인식보드

4.1.2 지문인식부의 하드웨어 구성

지문 영상의 실시간 처리를 위해 사용된 CPU(Central Processing Unit)는 크기와 기능을 고려해 TI사의 TMS320VC5509를 사용 하였다.

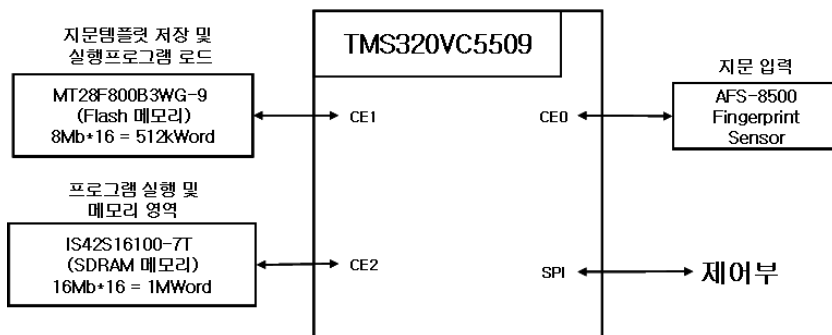


그림 4.4 지문인식보드의 구성

그림 4.4는 지문인식보드의 전체적인 구성을 나타내었다. 지문인식보드는 DSP에 지문영상을 입력받는 지문센서부와 지문영상의 저장과 처리를 위한 플래쉬 메모리와 SDRAM을 각각 연결하여 설계하였다.

총 256개의 지문템플릿을 저장할 수 있도록 플래쉬 메모리의 용량은 1MB로 구성하였고, 지문영상을 처리하기 위해 많은 메모리공간이 필요로 하는데 2MB의 SDRAM을 연결하여 충분한 메모리 공간을 확보하였다.

지문 영상을 입력 받기 위해서 사용된 센서는 최근 많이 쓰이고 있는 반도체 방식의 센서로서 AuthenTec사의 AFS-8500 모델을 사용하였다. 이 센서는 DSP와 패러럴 통신방식으로 96×96 지문영상 데이터를 전송한다. 실제 구현된 지문인식보드와 제어부의 사진은 그림 4.5와 같다.

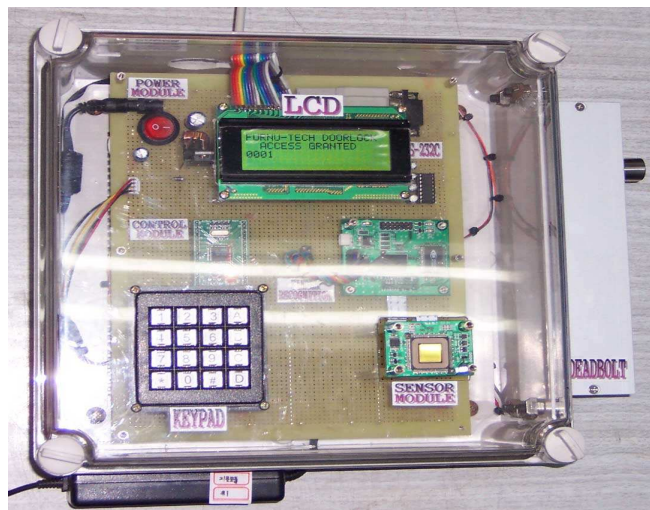


그림 4.5 구현된 제어부와 지문인식보드

4.1.3 제어부의 하드웨어 구성

제어부의 구성하기 위해 8bit 고성능 Microprocessor ATmega16L을 사용하였다. 이 프로세서는 내부에 SRAM과 EEPROM, FLASH를 내장하고

있으며 2.7V-5.5V의 다양한 구동전압을 사용할 수 있기에 일반 산업현장에서 많이 쓰이고 있다.

그림 4.6에 제어보드의 전체적인 구성을 나타내고 있다. ATmega16L에 사용자가 원하는 동작의 입력을 위해 4×4 키패드와 동작 상태를 디스플레이하기 위해 LCD를 연결하였다. 제어부에서는 지문인식부와 SPI통신, 화자인식부와는 패러럴통신으로 연결되어 있으며 이 라인을 통해 인식여부에 관한 정보를 수신하고 키패드를 이용하여 각 모듈의 제어를 위한 데이터를 송신한다. 그리고 시각적인 효과를 위해 인식성공 데이터가 두개의 보드로부터 수신이 되면 도어락이 자동적으로 열리도록 구성하였다.

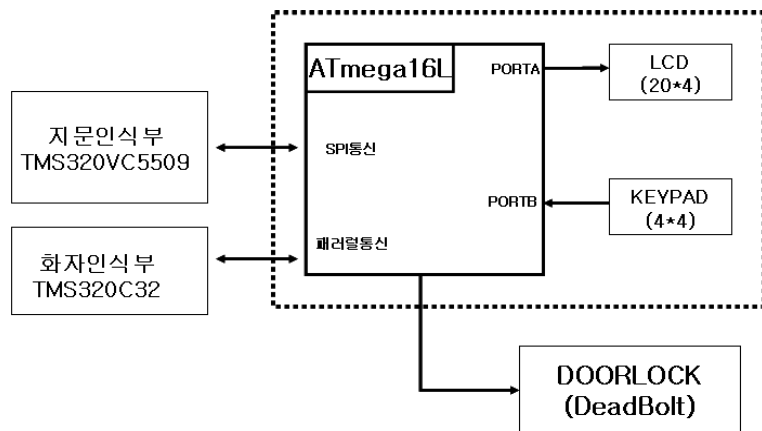


그림 4.6 제어보드의 구성

4.2 소프트웨어의 구성

본 논문의 소프트웨어 구성은 학습모드와 인식모드 두 부분으로 나뉘는데, 학습모드의 전체 흐름도는 그림 4.7과 같다. 학습모드는 음성이 입력되면 유/무성음을 판단하여 음성영역을 검출한 뒤, 특징정보추출(MFCC)

을 통해서 프레임당 12개의 특징정보를 생성한다. 이 후 화자인식보드의 메모리에 음성 패턴을 저장하고, LCD에 그 정보를 표시한다. 음성이 저장되고 난 뒤 지문인식보드는 지문저장모드로 전환하게 되는데, 새롭게 입력된 지문은 전처리, 특징정보추출의 과정을 수행 후 추출된 특징정보를 PC로 전달하게 된다. 이 특징정보는 KSOM 신경망의 입력층에 입력되어 오프라인 상에서 학습을 수행하게 된다. 신경망의 학습이 완료되면 갱신된 연결강도 값을 지문인식부에 전달하면 지문인식부는 인식모드로 전환한다.

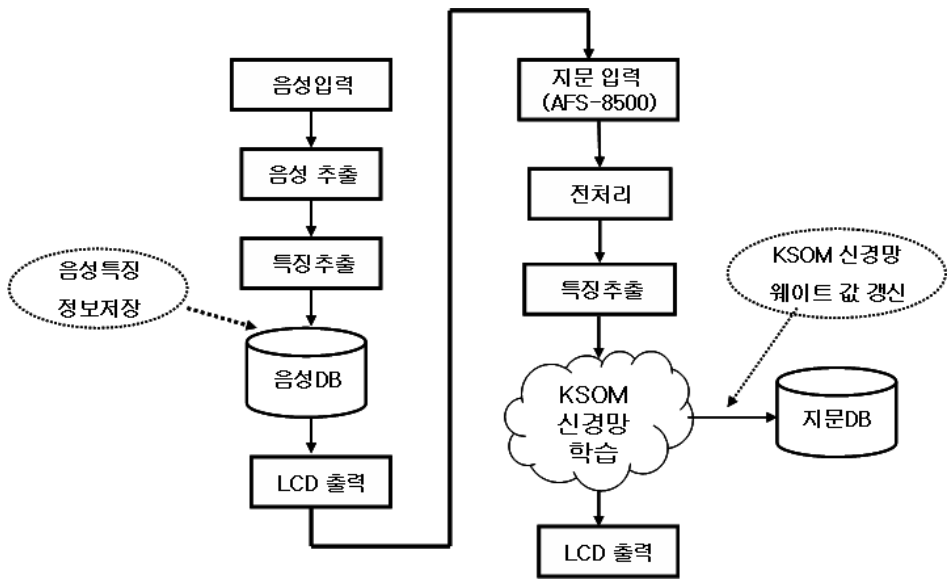


그림 4.7 전체 소프트웨어의 학습모드 블록도

그림 4.8은 전체 소프트웨어 인식모드의 블록도를 나타낸 것이다. 우선 화자인식부는 음성이 입력되면 음성영역을 추출하고, 특징정보추출(MFCC)을 한다. 추출된 특징정보는 화자인식을 수행하게 되는데, DTW 알고리즘을 이용하여 기존에 저장되었던 음성 패턴과 유사도를 구하여 가

장 유사한 음성패턴을 인식한다. 만약 유사도가 30보다 작으면, 화자로 인식하고, 30보다 크면 등록되지 않은 사람으로 판단되어 인식거부를 하였다. 유사도 값은 실험을 통해 적절한 임계값을 30으로 결정하였다.

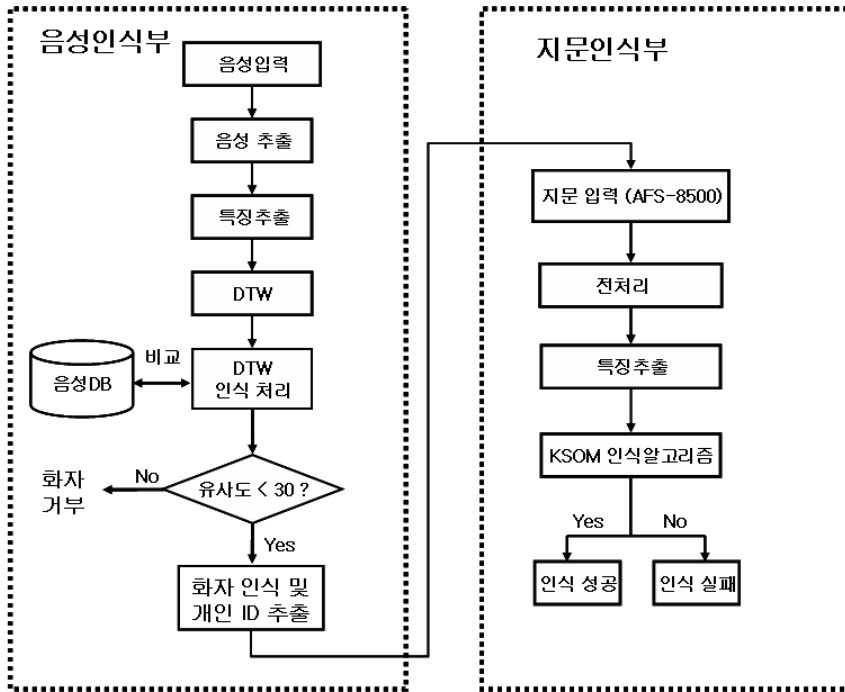


그림 4.8 전체 소프트웨어의 인식모드 블록도

화자인식을 통해 개인 ID를 추출한 후 지문 인식부에서 인식단계를 수행하게 된다. 센서로부터 들어온 지문을 통해서 전처리, 특징정보추출 과정을 통해서 지문의 특징정보를 추출한 뒤, 기존에 오프라인에서 학습된 KSOM 신경망의 입력층에 입력하게 되면 연결강도와 입력값과의 연산을 통해 가장 유사한 지문을 인식하게 된다.

본 논문에서는 2가지 인식 알고리즘인 화자 및 지문인식 알고리즘을 수행해서 모두 인식이 되면, 최종 인식을 하게 되지만, 만약 1개라도 거부되

면 인식이 수행되지 않는다. 그림 4.9는 실제 구현한 다중생체 인식시스템의 사진을 나타내고 있다.

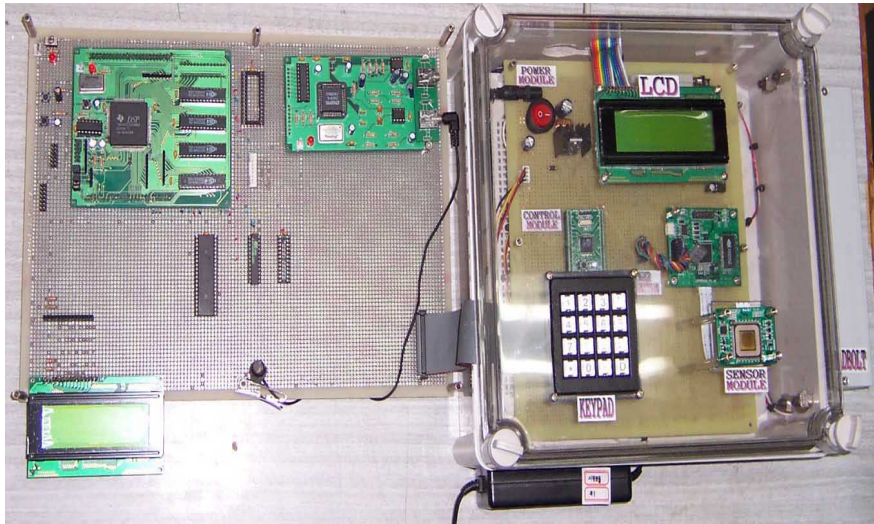


그림 4.9 구현된 다중생체 인식보드

제 5 장 실험결과 및 고찰

본 논문의 실험은 총 3단계로 구성되어 있다. 첫 번째로 음성의 특징정보 추출과정을 통해 실제 처리된 결과를 나타내고, 두 번째로 지문의 특징정보 추출과정에 의해 처리된 결과를 나타냈다. 마지막으로 본 논문에서 제시한 다중생체 인식시스템과 단일인식시스템의 인식률 테스트를 비교하여 그 결과를 보여준다.

5.1 음성특징 추출

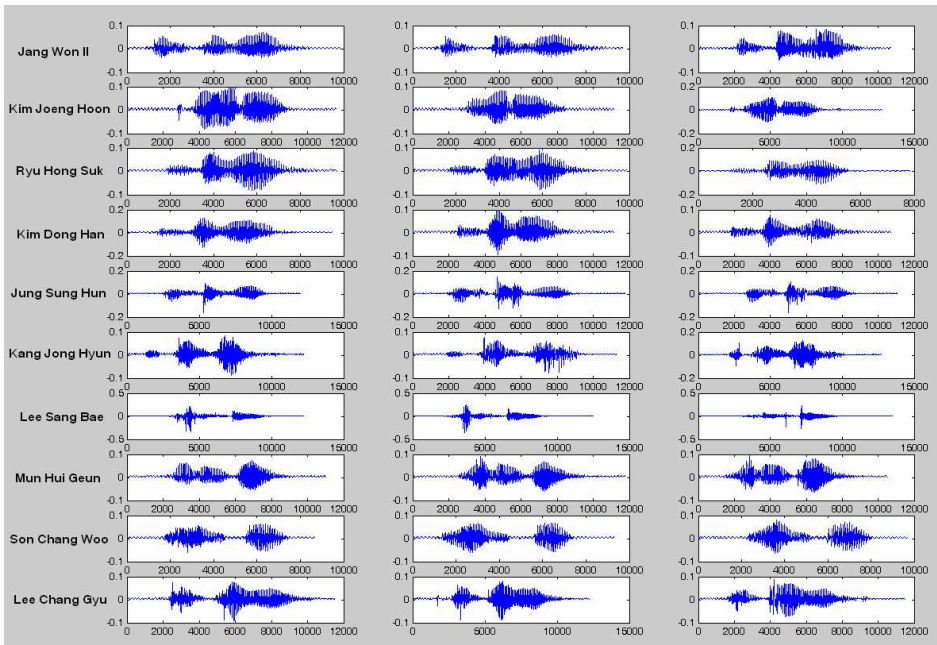


그림 5.1 화자인식에 사용된 음성 데이터베이스

화자인식에 사용된 전체 음성 데이터베이스는 그림 5.1과 같으며, 모두

10명의 사람이 각각 3번씩 발성하여 총 30개의 음성 파일을 데이터베이스에 저장시켰다. 저장시킨 패턴의 음성은 본인의 이름을 발성하였고, 30개의 음성에 대한 특징정보들을 화자인식시스템의 메모리에 저장시켰다. 특징정보 추출과정(MFCC)을 거친 음성의 특징정보 패턴은 그림 5.2와 같이 나타났다.

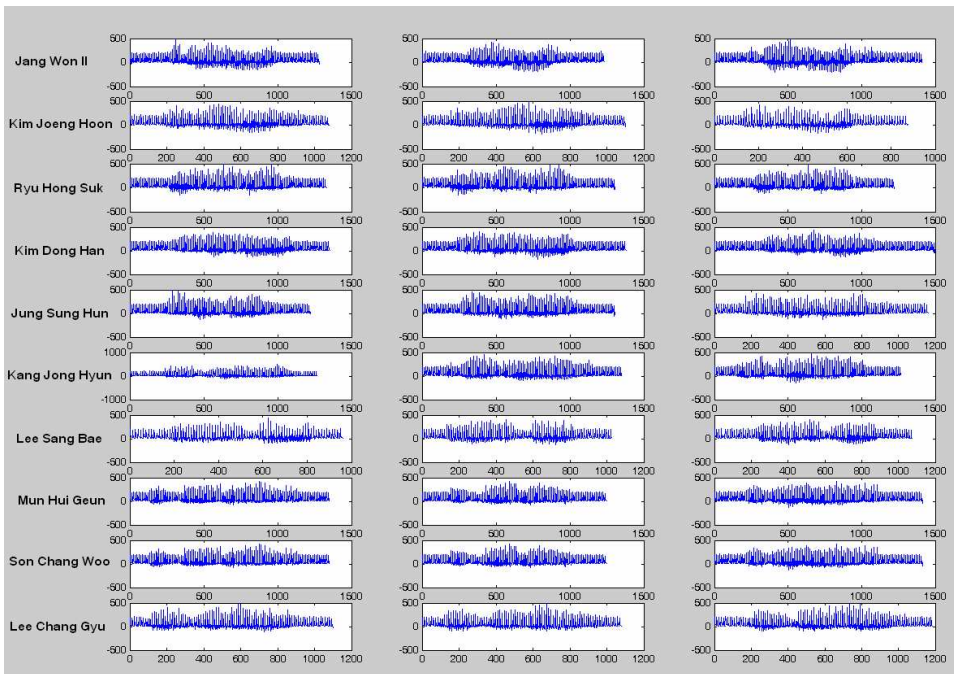


그림 5.2 음성특징정보 데이터베이스

5.2 지문특징 추출

KSOM 신경망을 이용한 지문인식을 수행하기 위해서는 우선 학습 알고리즘을 수행해야 한다. AuthenTec 사의 AFS-8500 반도체 지문센서로

부터 한 사람당 한 지문을 3번씩 입력하여 총 10명의 지문을 저장하였다. 그림 5.3는 인식하고자 하는 사람의 지문 데이터베이스이며, 총 30개의 지문을 KSOM 신경망에 학습시키는데 사용하였다.

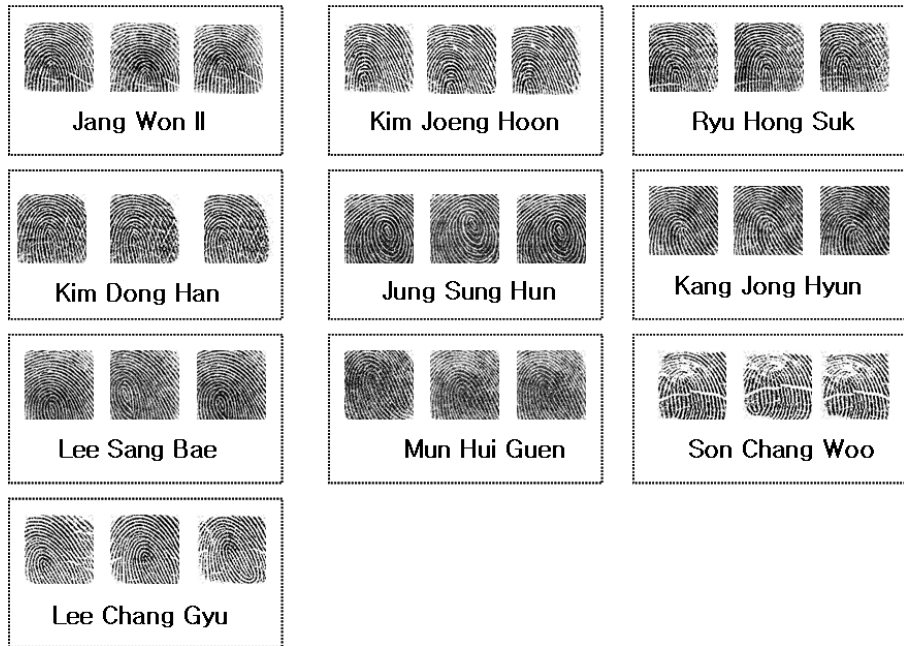


그림 5.3 지문인식에 사용된 지문 데이터베이스

이 지문 데이터베이스는 지문 인식 시스템에 저장시키기 위해 그림 3.1과 같이 전처리, 특징정보 추출과정을 통해 각각의 지문에 대한 특징값들을 추출하였다. 그림 5.4는 추출된 특징정보의 패턴이다. 이 특징정보들을 이용하여 KSOM 신경망 학습을 Pentium 4 3.0GHz PC에서 Visual C++ 6.0을 이용하여 오프라인 상에서 학습과정을 처리하였다.

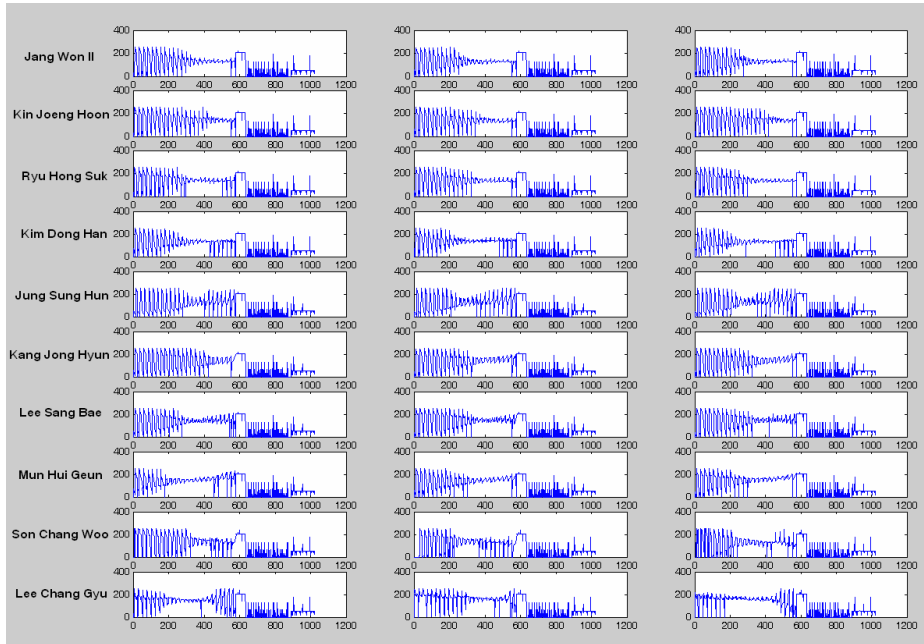


그림 5.4 지문특징정보 데이터베이스

5.3 인식률 테스트

인식률을 평가하는 기준으로는 2가지가 있다. 그 첫째는 접근이 허락되지 않은 사람을 통과시키는 오류, 즉 타인수락율(FAR-False Acceptance Ratio)이다. 이 비율이 1%라면 100번의 인식이 수행되는 동안 1번의 오인식이 발생하였다는 것이다. 즉 타인수락율이 높을수록 시스템에서는 치명적이 된다. 일반적인 보안을 목적으로 하는 시스템에서 그 기술의 우위를 가늠하는 대표적인 척도로 대개 이 FAR이 꼽힌다. 두 번째는 출입이 허락된 사람을 통과시키지 않는 오류, 즉 본인거부율(FRR-False Rejection Ratio)이다. 이 오류율이 높다면 당연히 출입해야 할 사람인데도 출입불가가 되는 것이다. 따라서 현실에서 상용화되기 위해서는 이 비율을 최소화

시키는 것이 곧 시스템의 보안성을 높여 안정성 면에서 신뢰를 얻을 수 있다.

본 절에서는 지문과 음성에 대해 각각 FRR과 FAR을 테스트하였다.

5.3.1 화자인식 테스트

표 5-1 화자 인식 FRR 테스트

단어명	인증횟수	거부횟수	평균 Distance
장원일	29	1	23.5
김정훈	28	2	25.9
류홍석	25	5	35.1
김동한	28	2	24.7
정성훈	28	2	25.2
강종윤	27	3	30.2
이상배	29	1	22.8
문희근	24	6	40.1
손창우	29	1	20.8
이창규	28	2	24.4

화자인식만 단독으로 인식률을 테스트한 결과는 표 5-1과 같다. 화자인식 FRR 테스트는 총 10명이 30번씩 발성 테스트를 한 결과를 보여준다. FRR은 8.3%로 나타났다. 본 논문에서는 Distance 값이 30이하일 때만을 화자인식에 성공하고, 30이상일 때는 등록되지 않은 음성이라 판단하여 화자인식을 거부하게끔 동작한다. 표 5-2에서는 타인이 발음하여

인식을 시도한 경우의 FAR 테스트 결과를 나타낸 것이다. 실험은 각 단어별로 본인을 제외한 나머지 사람들로 구성하여 랜덤하게 테스트를 수행하였다. FAR은 2.3%로 나타났다.

표 5-2 화자인식 FAR 테스트

단어명	인증횟수	거부횟수	평균 Distance
장원일	1	29	97.5
김정훈	0	30	120.4
류홍석	0	30	130.8
김동한	2	28	70.1
정성훈	1	29	92.3
강종윤	0	30	100.3
이상배	0	30	111.7
문회근	2	28	99.3
손창우	1	29	120.2
이창규	0	30	140.2

5.3.2 지문인식 테스트

본 절에서는 KSOM 신경망 알고리즘을 이용하여 지문 데이터베이스를 분류한 것과 분류된 신경망의 연결강도 값을 이용하여 인식테스트를 수행하였다. 표 5-3에서는 KSOM 신경망을 이용하여 입력된 지문을 분류한 결과이다. 10명의 지문을 3번씩 입력하여 총 30개의 지문 데이터베이스를 분류한 것이며, 출력층의 개수는 총 100개로 구성하였다.

표 5-3 KSOM 지문 분류표

지문명	첫 번째 지문	두 번째 지문	세 번째 지문
장원일	45 Winner	6 Winner	6 Winner
김정훈	47 Winner	4 Winner	47 Winner
류홍석	64 Winner	45 Winner	33 Winner
김동한	27 Winner	31 Winner	66 Winner
정성훈	95 Winner	87 Winner	87 Winner
강종윤	59 Winner	91 Winner	79 Winner
이상배	71 Winner	25 Winner	25 Winner
문희근	83 Winner	51 Winner	83 Winner
손창우	99 Winner	0 Winner	12 Winner
이창규	39 Winner	39 Winner	19 Winner

표 5-4에서는 등록되지 않은 10명의 사람이 각자의 지문을 3번씩 입력하여 FAR 테스트를 한 결과이고, 표 5-5에서는 지문을 등록한 10명이 각자의 지문을 3번씩 입력하여 FRR 테스트를 한 결과이다.

FAR은 6.7%, FRR는 3.3%로 나타났다.

표 5-4 지문인식 FAR 테스트

총지문입력수	인식 횟수	거부 횟수
30	2	28

표 5-5 지문인식 FRR 테스트

등록자명	첫 번째 지문	두 번째 지문	세 번째 지문	인식횟수
장원일	45 Winner	6 Winner	6 Winner	3
김정훈	47 Winner	4 Winner	47 Winner	3
류홍석	64 Winner	45 Winner	33 Winner	3
김동한	27 Winner	31 Winner	66 Winner	3
정성훈	95 Winner	87 Winner	87 Winner	3
강종윤	59 Winner	91 Winner	79 Winner	3
이상배	71 Winner	25 Winner	25 Winner	3
문희근	83 Winner	51 Winner	83 Winner	3
손창우	99 Winner	0 Winner	12 Winner	3
이창규	38 Winner	39 Winner	19 Winner	2

5.3.3 다중생체인식 테스트

총 20명의 사람이 지문인식과 음성인식을 병행하였을 때의 실험결과이다. 그 중 10명은 기존 시스템에 음성과 지문이 저장되어 있으며, 나머지 10명은 등록되지 않은 사람들이다. 표 5-6에서는 다중생체 인식테스트의 결과와 지문과 음성을 단독으로 테스트 한 결과를 비교하여 나타내고 있다. FRR 테스트는 미리 저장되어 있는 사람 10명이 각자 3번씩 음성과 지문을 입력하여 테스트를 하였고 FAR 테스트는 저장되어 있지 않은 사람 10명의 지문과 음성을 각자 3번씩 입력하여 테스트를 하였다. 다중생체 인식시스템과 단일인식시스템과의 인식률을 비교해보면 다중생체 인식

시스템이 화자인식 8.3%, 지문인식 3.3%로 나타났지만 본인거부율(FRR)은 10%로 높아졌다. 그러나 타인수락율(FAR)은 화자인식 2.3%, 지문인식 6.7%로 나타났지만 본 논문에서 제안한 다중생체 인식시스템에서는 한건도 발생하지 않아 안정된 보안 시스템임을 입증하였다.

표 5-6 다중생체인식과 단일생체인식의 비교

	FAR	FRR
화자 인식	2.3%	8.3%
지문 인식	6.7%	3.3%
복합 인식	0%	10%

제 6 장 결 론

본 논문에서는 기존의 단일생체 인식시스템이 가지고 있는 오인식률(FAR)을 보완하기 위해 생체 인식시스템에서 많이 사용되고 있는 지문과 음성을 이용한 다중생체 인식시스템을 구현하여 오인식률(FAR)을 감소시키는데 그 목적을 두었다. 인식시스템은 각각 임베디드 형태로 구현하기 위해, TI사의 DSP TMS320C32와 TMS320VC 5509를 사용하여 실제 제작하였으며, 실시간 구현을 위해 인공지능 알고리즘인 KSOM 신경망을 채택하여 구현하였다.

본 논문에서 구현한 다중생체 인식시스템을 실험한 결과 지문과 음성을 각각 이용한 단일인식시스템보다 인식률(FRR)은 2~7%정도 떨어졌지만, 인식시스템에서 가장 중요한 요소인 오인식률(FAR)은 전혀 발생하지 않음을 확인하였다.

앞으로 연구방향으로는 음성과 지문뿐만 아니라 홍채와 얼굴 등 좀 더 많은 생체정보를 이용한 실시간 생체인식시스템을 개발하여 시스템의 안정성을 더욱 더 증가시켜야 할 것이며 아울러 화자인식시스템에서는 여러 가지 잡음 환경을 고려하여 잡음 환경에 강한 화자인식 알고리즘에 대한 연구가 있어야겠다.

참 고 문 헌

- [1] Laurene V. Fausett, "Fundamentals of Neural Networks", Prentice Hall, 1993.
- [2] 오창석 "뉴로 컴퓨터 개론". 내하출판사 2000.
- [3] 이상배, "퍼지-뉴로제어 시스템", 교학사, 1999.
- [4] Lawrence Rabiner and Biing Hwang Juang, "Fundamentals of Speech Recognition", Prentice Hall, 1993
- [5] 정익주, 정훈, "TMS320C32 DSP를 이용한 실시간 화자 종속 음성인식 하드웨어 모듈(VR32)의 구현", 한국음향학회, vol. 17, no. 4 pp.14-22, 1998.
- [6] 오경환, "음성언어정보처리", 홍릉과학출판사, 1997
- [7] 김정훈, " 음성인식처리용 임베디드 시스템의 설계 및 구현에 관한 연구 ", 한국해양대학교 공학석사학위 논문, 2003
- [8] 유강주, "DHMM을 이용한 숫자음 인식의 Data Fusion에 관한 연구", 한국해양대학교 공학석사학위 논문, 1998
- [9] GONZALEZ & WOODS, "Digital Image Processing", Addison Weley Longman
- [10] 이상배, L.C.Jain, R.K.Jain "하이브리드 지능 시스템", 도서출판 그린, 1997.
- [11] 이문호, "Visual C++ 실용 영상 신호처리", 대영사, 2001.
- [12] L.C.Jain, U.Halici, I.Hayashi, S.B.Lee, S.Tsutsui, "Intelligent Biometric Techniques in Fingerprint and Face Recognition", CRC Press
- [13] 이재현, "人工知能 技法을 利用한 指紋 認識 알고리즘 및 시스템 具現

- 에 관한 研究”, 한국해양대학교 박사학위 논문, 2002.
- [14] H.C Lee and R.E. Gaensslen, *Advances in Fingerprint Technology*, Elsevier, 1991
- [15] R. Brunelli and D. Falavigna, "Personal Identification Using Multiple Cues," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp.995-966, Oct. 1995.
- [16] M.M.S. Chong et al., "Geometric Framework for Fingerprint Image Classification," *Pattern Recognition*, vol. 30, no. 9, pp.1475-1488, 1997.
- [17] N. Ratha, S. Chen, and A.K. Jain, "Adaptive Flow Orientation Based Feature Extraction in Fingerprint Images," *Pattern Recognition*, vol. 28, no. 11, pp.1,657-1,672, 1995
- [18] " *TMS320C32 General Purpose User's guide* ", Texas Instrument
- [19] 윤덕용, " *TMS320C32 마스터* ", Ohm사, 1999.
- [20] 류홍석, "ART2 적용 임베디드 음성인식 시스템의 설계 및 구현에 관한 연구", 한국해양대학교 공학석사학위 논문, 2004
- [21] 이태한, "범용 DSP를 이용한 차량 항법용 음성 인식 시스템의 성능 개선", 연세대학교 공학석사학위 논문, 1999

감사의 글

본 논문이 있기까지 세심한 배려와 관심으로 지도해주신 이상배 교수님께 진심으로 감사드립니다. 그리고 바쁘신 가운데도 논문 심사과정에서 부족한 점들을 조언해 주신 박동규 교수님, 손정락 교수님과 많은 가르침을 주신 임재홍 교수님께 깊은 감사드립니다.

학부 4학년 처음 연구실 생활을 하면서 영상처리와 프로그래밍에 대해 많은 가르침을 주신 동명정보대학교 이응주 교수님과 학부과정에서 다양한 전공지식을 접할 수 있게 해주신 최대우 교수님, 권순량 교수님, 최영복 교수님, 안형근 교수님, 손영선 교수님, 김민성 교수님께 진심으로 감사드립니다.

같이 생활한 기간은 짧았지만 부족한 후배에게 항상 많은 조언과 격려를 아끼지 않으신 김관형, 이재현, 강성인, 손창우, 공석민, 강재명, 문희근, 이창규 선배님에게 감사드리며, 같이 생활하고 프로젝트를 하면서 후배를 먼저 배려해주신 이해심 넓은 김정훈, 김동한, 김영탁 선배님과 사회에서 열심히 연구개발에 전념하면서도 후배의 취업과 진로에 대해 따뜻한 조언과 격려를 아끼지 않으신 강종윤, 김수정 선배님에게 감사하다는 말을 전합니다. 그리고 동기이자 동생인 정성훈, 후배 박주원에게도 고맙다는 말을 전합니다.

학부 때부터 대학원까지 항상 저에게 힘이 되어준 선배이자 친구인 홍석과 언제나 최선을 다하는 일식과 학부 때 같은 영상처리 연구실에 있었던 정훈형, 영일형, 정락형, 수현형, 주성형, 창희형, 순영, 영수, 상아와 소모리 모임 회원 경태형, 승현형, 일원형, 필모형, 종학형, 용수형, 진영에게 감사드리며 모두들 하시는 일 잘 되시길 바랍니다.

중학교 때부터 항상 곁에서 힘이 되어준 영원한 친구 현채, 명주, 준형, 상환, 대기, 정필, 정환, 선진과 언제나 반갑게 맞아주는 명관, 현철, 회진, 영남희 계원들과 군대동기 영철, 장호에게 진심으로 고맙다는 말을 전합니다.

제가 학업에만 전념할 수 있게 항상 믿고 지켜 봐 주신 아버지와 어머니의 아낌없는 사랑에 진심으로 고개 숙여 감사하다는 말을 드립니다. 그리고 형을 잘 따라준 동생 명일에게도 고맙다는 말을 전하며 우리 가족의 영원한 화목과 건강을 바랍니다.

끝으로 미처 적지 못한 _____에게도 고맙다는 말을 전하며 그 동안 저를 아끼고 사랑해주신 모든 분들에게 이 논문을 바칩니다.