

KSOM을 이용한 다중생체 인식시스템에 관한 연구

Implementation of Embedded System for Multi-modal Biometric Recognition using KSOM

김재완, 이상배

¹ 부산시 영도구 한국해양대학교 전자통신공학과
E-mail: glitter@bada.hhu.ac.kr, leesh@mail.hhu.ac.kr

요 약

본 논문은 생체인식시스템에서 단일시스템의 각각의 특징을 바탕으로 신뢰성을 증가시키는 것에 있다. 간단하면서 높은 인식률을 가지는 지문과 개개인의 음성을 다중생체인식에 활용하여 다중생체인식 시스템을 구현 하였다. 화자인식부에서는 DSP를 이용하여 화자인식을 수행하고, 이후 지문인식부에서 지문 특징점을 추출하여 KSOM신경망 알고리즘을 이용하여 인식을 수행하였다. 그리고 각 인식부의 전체적인 제어는 ATmega16L을 사용하였다. 또한 인증결과를 PC에 MFC로 디스플레이 한다.

Key Words : KSOM, 지문, 화자인식, DSP, AVR, 임베디드 시스템

1. 서 론

사회가 점차 고도화 되고 복잡화되면 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

최근 널리 사용되고 있거나 중점적으로 연구되고 있는 생체특징으로는 얼굴, 지문, 화자, 손 모양, 음성, 홍채 및 망막 패턴, 서명등 생리학적 특징과 행동적 특징을 이용한 방법들이 있다. 이 모든 것이 최적일 경우이나 어떠한 단일 생체특징도 절대적으로 우수한 성능을 나타내지 않는다. 그러므로 안정성을 증가시키기 위하여 다중생체 인식을 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.

이러한 단일인식시스템들의 특징들을 활용하여 신뢰성을 증가시킬 수 있는 간단하면서도 높은 인식률을 가진 지문과 개개인이 가지고 있는 독특한 음성을 이용하여 임베디드형 다중생체 인식시스템을 구현 하였다. 전체시스템은

크게 세부분으로 지문인식부, 화자인식부, 제어부로 나뉜다. 각 인식부에 대한 전체적인 제어는 ATMEL사의 Atmega16L을 사용하였다. 화자인식부는 기존의 화자인식 알고리즘 중에서 실시간 처리가 가장 적합한 DTW (Dynamic Time Warping) 알고리즘을 적용한 방법을 이용하였고, 지문인식부 인공지능 기법인 KSOM 신경망 및 여러 가지 지문처리 알고리즘을 이용하였다.

2. 화자인식 알고리즘

실시간 화자인식을 위해 간단하면서도 성능이 뛰어난 DTW 알고리즘을 적용하였으며, 전체 화자인식 알고리즘은 그림1 과 같다.

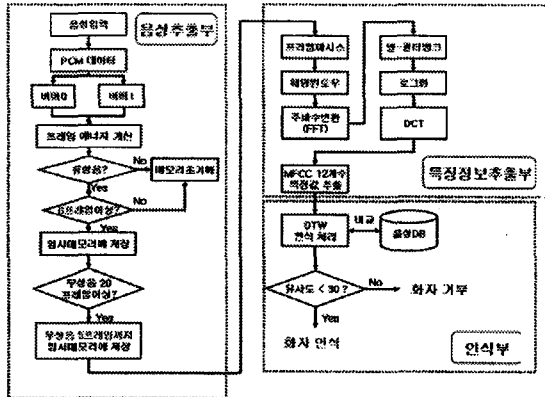
그림1. 화자인식 알고리즘의 흐름도

2.1 음성추출

가장 간단한 방식인 절대에너지, 아래식을 사용하여 음성 감지에 사용하였다. 여기서 말하는 음성의 감지는 유성음이나 무성음을 구별하는 방법이다.

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

$x(n)$ 은 마이크에서 입력받은 음성데이터, N 은 프레임의 길이, E 는 절대에너지를 나타낸



다. 각 프레임의 절대 에너지가 설정된 값보다 크면 이 구간을 실제 음성으로 간주한다.

2.2 음성특징 정보추출

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)를 수행하여 주파수영역으로 변환한다. 변환된 값들을 미리 설정된 MFCC (멜 캡스트럼) 대역필터뱅크를 통과시키고, 로그화,DCT(Distance Cosine Transform)을 거치면 한프레임 당 12개의 계수 값들을 구할 수 있다. 이 값들이 음성의 특징정보들이다.

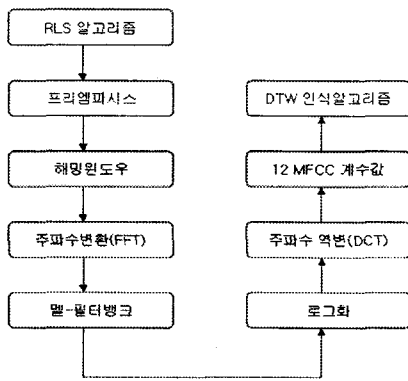


그림 2. 특징추출 알고리즘

2.3 화자인식 알고리즘

DTW(Dynamic Time Warping) 입력패턴과 참조 패턴 사이의 거리를 측정해서 그 유사도를 측정하는 방법이다. 다시 말하면, 제한된 경로 내에서 단조 증가를 통해서 가장 가까운 거리를 판별, 유사도를 측정한다. 예를 들어, 길이가 M인 입력음성 패턴을 $T = T(1), T(2), \dots, T(M)$, 길이가 N인 기준 패턴을 $R = R(1), R(2), \dots, R(N)$ 이라고 하면 두 패턴 간의 유사도 d 는 다음 식과 같이 누적거리로 표현 된다

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

이 때 $D(R(n), T(W(n)))$ 는 R의 n번째와 T의 W(n)번째의 국부적 유사도(Local Distance)이며, DTW는 두 패턴간의 누적거리 최적화 하는(m,n) 평면의 최적 경로 $m=W(n)$ 를 찾는 방법이다.

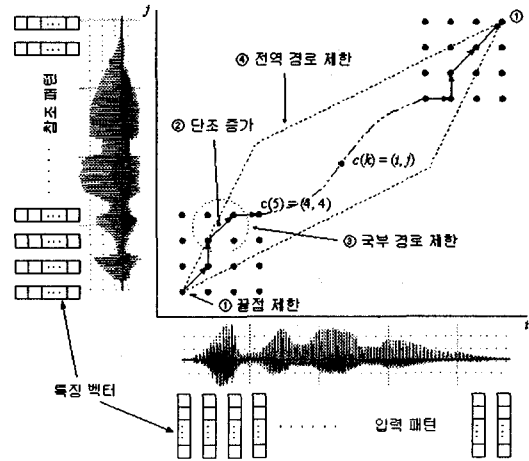


그림 3. DTW의 제약된 조건

3. 지문인식 알고리즘

특징 추출과정 알고리즘을 이용하여 특징값을 추출하였으며, 인식과정을 인공지능기법인 자율신경망 KSOM (Kohonen Self Organizing Maps)을 이용하여 실시간 지문인식 시스템을 구현하였다.

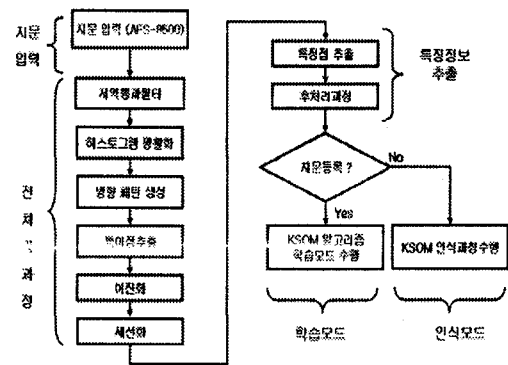


그림 4. 지문인식 시스템 알고리즘

3.1 전처리

입력된 지문영상에서 특징점 추출을 용이하게 하기 위하여 영상을 향상하고 난 뒤에 지문영상의 방향성 및 이진화, 세션화

처리를 하게 된다.

3.2 지문 특징 정보추출

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

3.3 지문인식 알고리즘

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 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 입력 벡터가 통상 0에서 1사이의 정규 (normalized) 된 값을 사용한다. 이런 두 가지 요인은 KSOM 신경망에 있어서 매우 중요하다. 지문에서 추출된 특징점들을 KSOM 신경망의 입력패턴으로 사용하였고 입력층 512개, 출력층 100개의 신경망을 구성하였고 연결강도와 입력 패턴과의 거리는 유클리드 거리법을 사용 하였다.

4. 실험 및 결과

4.1 다중생체인식 시스템

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 오픈된다.

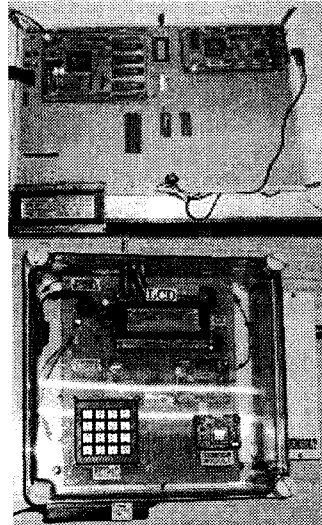


그림 5. 구현된 시스템

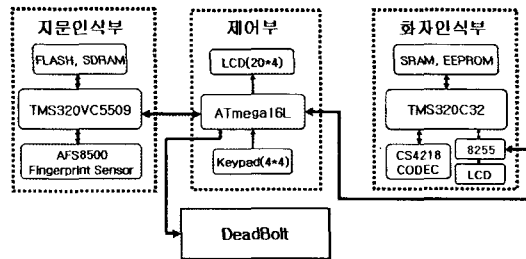


그림 6. 전체 시스템의 구성

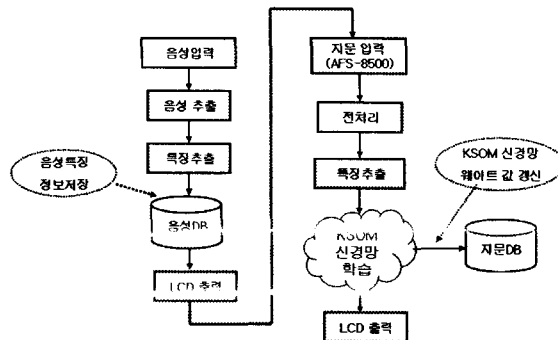


그림 7. 전체 소프트웨어의 학습 모드

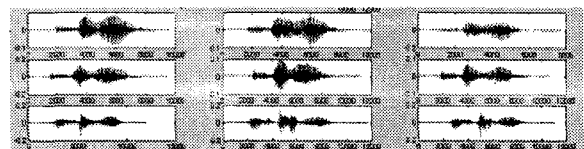


그림 8. 음성 데이터베이스



그림 9. 지문특징정보 데이터베이스

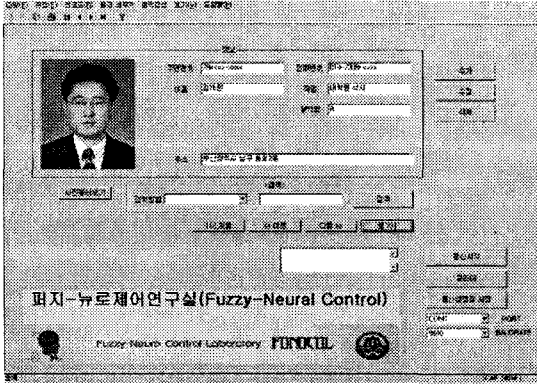


그림 10. 인증결과를 PC에서 디스플레이

	FAR	FRR
화자 인식	2.7%	9.3%
지문 인식	6.5%	3.5%
복합 인식	0%	10%

표 1. 다중생체인식과 단일생체인식의 비교

5. 결론

기존의 단일생체 인식시스템이 가지고 있는 오인식률(FAR)을 보완하기 위해 생체 인식 시스템에서 많이 사용되고 있는 지문과 음성을 이용한 다중생체 인식시스템을 구현하여 오인식률(FAR)을 감소시키는데 그 목적을 두었다.

인식시스템은 각각 임베디드 형태로 구현하기 위해, TI사의 DSP TMS320C32 TMS320VC5509를 사용하여 실제 제작하였으며, 실시간 구현을 위해 인공지능 알고리즘인 KSOM 신경망을 채택하여 구현하였다.

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 연구하고자 한다.

참 고 문 헌

[1] Lawrence Rabiner and Biing Hwang Juang, "Fundamentals of Speech Recognition", Prentice Hall, 1993
 [2] Steven L.Gay, Jacob Benesty, "Acoustic Signal Processing for Telecommunicatio

n", Kluwer Academic Publishers, 2000.
 [3] 정익주, 정훈, "TMS320C32 DSP를 이용한 실시간 화자 중속 음성인식 하드웨어 모듈 (VR32)의 구현", 한국음향학회, vol. 17, no.4 pp.14-22, 1998.
 [4] 오경환, "음성언어정보처리", 홍릉과학출판사, 1997
 [5] GONZALEZ & WOODS, "Digital Image Processing", AddisonWiley Longman
 [6] 이문호, "Visual C++ 실용영상신호처리", 대영사, 2001.
 [7] L.C.Jain, U.Halici, I.Hayashi, S.B.Lee, S.Tsutsui, "Intelligent Biometric Techniques in Fingerprint and Face Recognition", CRC Press
 [8] Vinay K. Ingle, John GProakis, "Digital signal Processing Using MATLAB", Brooks/Cole Publishing Company, 2000