WPS와 장갑 장치 기반의 동적 제스처 인식기의 구현

김 정 현† · 노 용 완† · 홍 광 석††

요 약

차세대 PC를 위한 WPS(Wearable Personal Station)는 정보 처리 및 네트워크 기능을 포함하며 새로운 정보의 획득에 있어 공간적 한계성을 극복할 수 있는 '유비쿼터스 컴퓨팅'의 핵심 단말기로 정의할 수 있다. 기존의 유선 통신 모듈을 이용한 데스크톱 PC 기반의 제스처 인식기는 영상 또는 장갑 장치로부터 사용자의 의미 있는 동적 제스처 데이터를 획득함에 있어 공간상의 제약성 및 이동상의 한계성, 배경 및 음영 조건의 변화에 따른 인식 성능의 변화 등 몇 가지 제약사항을 가지고 있다. 따라서 본 논문에서는 이러한 문제점들을 극복하고 해결하고자 제스처 입력모듈을 차세대 PC 플랫폼 기반의 유비쿼터스 환경으로 확대, 적용시켜 제스처 데이터 입력 모듈로부터 새로운 정보의 획득에 있어 한계성을 극복하고 효율적인 데이터 획득 방안을 제시한다. 또한 퍼지 알고리즘과 신경망 이론을 이용하여 독립적인 제스처 인식 시스템을 구현하고 개별 시스템의 성능을 비교, 분석함으로써 차세대 PC를 위한 보다 효율적이며 합리적인 제스처 인식 시스템을 제안한다. 제안된 제스처 인식시스템은 동적인 손의 움직임을 입력데이터로 처리하는 제스처 입력모듈과 입력된 데이터로부터 의미 있는 제스처를 분리하기 위한 관계형 DBMS모듈, 그리고 인식의 확장성과 연속된 동적 제스처 중에서 의미 있는 제스처를 인식하기 위한 퍼지 인식 모듈 및 신경망 인식 모듈로 구성되어 있다. 30인의 피실험자에 대하여 15회의 반복 실험을 수행하였으며 사용자의 동적 제스처 인식 실험결과 퍼지 제스처 인식 시스템에서는 98.8%, 신경망 제스처 인식시스템에서는 96.7%의 평균 인식률을 도출하였다.

키워드 : 제스처인식, 차세대PC, 퍼지 및 신경망 알고리즘, WPS

An Implementation of Dynamic Gesture Recognizer Based on WPS and Data Glove

Jung-Hyun Kim† · Yong-Wan Roh† · Kwang-Seok Hong††

ABSTRACT

WPS(Wearable Personal Station) for next generation PC can define as a core terminal of 'Ubiquitous Computing' that include information processing and network function and overcome spatial limitation in acquisition of new information. As a way to acquire significant dynamic gesture data of user from haptic devices, traditional gesture recognizer based on desktop-PC using wire communication module has several restrictions such as conditionality on space, complexity between transmission mediums(cable elements), limitation of motion and incommodiousness on use. Accordingly, in this paper, in order to overcome these problems, we implement hand gesture recognition system using fuzzy algorithm and neural network for Post PC(the embedded-ubiquitous environment using blue-tooth module and WPS). Also, we propose most efficient and reasonable hand gesture recognition interface for Post PC through evaluation and analysis of performance about each gesture recognition system. The proposed gesture recognition system consists of three modules: 1) gesture input module that processes motion of dynamic hand to input data, 2) Relational Database Management System(hereafter, RDBMS) module to segment significant gestures from input data and 3) 2 each different recognition module: fuzzy max-min and neural network recognition module to recognize significant gesture of continuous / dynamic gestures. Experimental result shows the average recognition rate of 98.8% in fuzzy max-min module and 96.7% in neural network recognition module about significantly dynamic gestures.

Key Words : Gesture Recognition, Next Generation PC, Fuzzy and Neural Network, Wearable Personal Station

1. 서 론

차세대 PC는 유비쿼터스 환경의 구현을 위한 정보 처리 및 네트워크 기능을 포함하는 핵심 단말 장치로서 착용할 수 있어야 하며 보다 개인화 되어지는 정보 서비스 환경에서 인간의 오감 정보 등 휴먼인터페이스를 지원하는 인간중심의 컴퓨터로 발전될 것으로 전망된다[1]. 또한 기존의 PC에서는 성능이나 속도 등이 중심이었다면, 차세대PC는 특화된 목적에 따른 기능이나 사용의 편의성이 강조된 '인간 중심의 컴퓨팅 기기'를 의미하며 유비쿼터스 환경을 지향하는 WPS(Wearable Personal Station)기반의 차세대 PC에서는 제스처 데이터 처리를 위한 햅틱 인터페이스 장치를 활용하

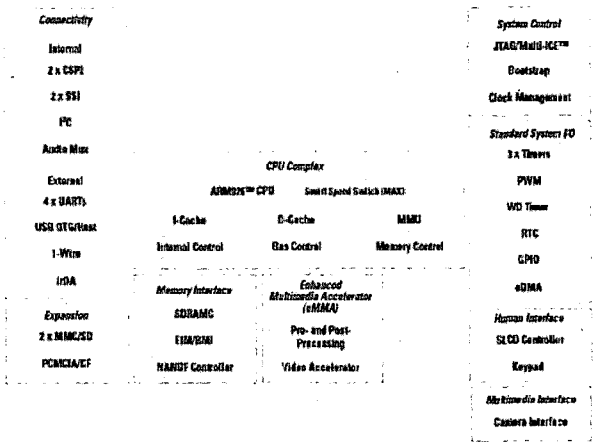
※ 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음. IITA-2005-(C1090-0501-0019).
† 준 회 원 : 성균관대학교 정보통신공학부 박사과정
†† 종신회원 : 성균관대학교 정보통신공학부 교수
논문접수 : 2005년 11월 2일, 심사완료 : 2006년 9월 5일

여 제스처 데이터를 획득하고 분석하는 연구 분야의 중요성이 더욱 부각되고 있다[2]. 즉, 기존의 키보드, 마우스, 모니터를 대체할 초소형 키보드 및 디스플레이 등을 포함하여 양손의 사용을 자유롭게 하는 입출력 장치와 음성, 시각, 촉각, 후각, 미각 등 오감 정보처리 기술을 위한 차세대 사용자 인터페이스 등을 주요 연구 대상으로 보고 있다. 또한 최근 공학 기술의 발전으로 인간이 느낄 수 있는 색상, 소리, 향기, 맛, 촉감을 디지털 신호로 바꾸어 이를 이용하는 오감 기술이 개발되고 있으며, 시·청각에만 의존하던 컴퓨터 인터페이스 기술은 공간과 시간의 제약을 극복하고 인간 친화적이며 실감화를 지향하는 차세대 휴먼 인터페이스를 제공하기 위해서 3D 비디오/오디오, 촉각, 후각, 미각 등 새로운 미디어로 제공하는 인터페이스 기술로 진화하고 있다[3, 4]. 제스처 인식 및 표현을 위한 관련 연구는 건설현장의 크레인 조종을 위한 수신호 인식에서부터 수화(sign language)인식, 손동작 인식을 이용한 이동 로봇 및 아바타의 제어, 가상현실에서의 손동작 인식 등을 대상으로 미국, 일본 등에서 활발히 추진 중[5]이며 특히 휴대용 단말기에서의 이동성을 보장하기 위한 사용자 인터페이스 관련 연구는 입는 컴퓨터의 접촉형 인터페이스를 구현한 카네기 말론 대학의 웨어러블 컴퓨터 시스템스 연구소[6]와 3차원 공간에서 손의 움직임을 감지하여 사용자의 동작을 인지하는 P5 글로브를 만들어 웹브라우저나 게임기로의 응용을 추구하고 있는 Essential Reality사[7]를 대표적인 사례로 들 수 있으나 이들 연구는 대부분 데스크톱 PC기반의 제어 및 손동작 인식에 목적을 두었고 이를 가상현실이나 데스크톱 PC상의 조이스틱과 마우스의 기능을 대체할 수 있도록 응용하고 있으므로 공간상의 제약성 및 이동상의 한계성이라는 문제점을 가지고 있다.

따라서 본 논문에서는 이러한 기존의 제약사항 및 문제점을 개선하고 PC의 진화에 따른 무선 단말기의 이동성을 보장하기 위하여 차세대 착용형 PC 플랫폼-WPS를 기반으로 퍼지알고리즘 및 신경망이론을 이용한 동적 제스처 인식기를 제안하고 구현하며, 이들 개별 시스템의 성능을 비교 분석한다.

2. WPS(Wearable Personal Station)

현재 차세대 PC 기반의 플랫폼 구현 및 핵심기술은 휴대형 PC와 착용형 PC로 구분되어 전개되고 있다. 휴대형 PC는 웹패드, 무선 핸드헬드, PDA, e-Book 등 인간이 쉽게 휴대할 수 있는 형태이며, 착용형 PC는 시계형, 의류형 등 주로 인간의 몸에 부착하는 형태로서 현재 차세대 PC 산업은 폰 중심의 휴대형과 PC 중심의 착용형으로 구분되어 전개되고 있는 과도기적 상황에 있다. 따라서 본 논문에서는 Motorola사에서 개발한 i.MX21 ADS Board를 개량, 이를 활용하였으며 i.MX21 ADS Board는 현재 정보주도하에 개발 중인 착용형 차세대 PC의 프로토타입으로서 이를 기반으로 손목시계형태의 착용형 컴퓨터 개발이 진행 중에 있다. 따라서 본 논문에서 구현된 동적 제스처 인식기는 손목시계형-착용형 컴퓨터의 개발이 완료시 별다른 수정 없이 이식이 가능할 것으로 판단되며 기존의 데스크톱 PC와의 가장 큰 차이점은 PC의 진화에 따른 소형화, 경량화 및 이동성을 고려하였다는 점을 들 수 있다. 그러나 이러한 착용형 컴퓨터의 기본적인 특성을 고려시 사용자 인터페이스 또한 기존의 키보드나 모니터 위주에서 벗어난 다양한 형태의 인터페이스를 제공해야 함은 분명하며 이러한 한계성을 극복하기 위하여 Blue-tooth 모듈과 HMD를 연동하여 명령어 인식을 위한 장갑장치 기반의 제스처 인식 시스템을 구현하였다. (그림 1)에서는 현재 개발 중인 차세대 착용형 PC에서 사용되어지는 i.MX21의 기능별 블록도[8]를 보여준다.

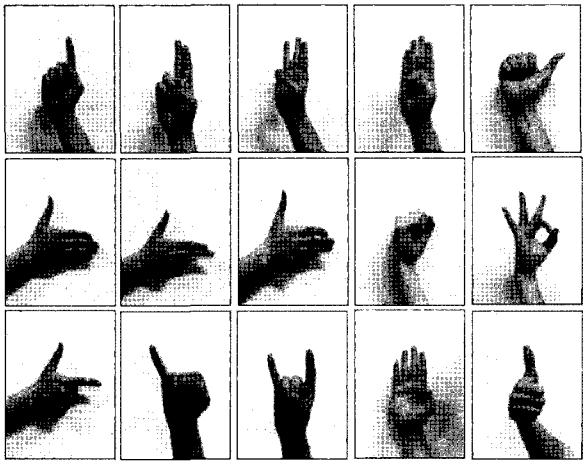


(그림 1) i.MX21의 기능별 블럭도

3. 제스처의 분류 및 제스처 입력 모듈

인간과 컴퓨터의 상호작용을 위한 도구로서 마우스, 키보드, 펜, 음성, 영상 등 다양한 방법들이 제안되고 있다. 본 연구에서는 인간이 가장 빨리 의사를 전달 할 수 있고 다양한 표현이 가능한 동적 손 제스처를 이용하여 입력모듈을 구성하고 이에 따른 사용자의 의도를 분석하여 적절한 응답을 취할 수 있는 방안에 대하여 그 목표를 두었다. 본 논문에서는 사용자의 의미 있는 동적 제스처를 실시간으로 인식하는 시스템을 구현하기 위해 "한국어 표준 수화 교본[9]"을 분석해 15가지의 기본 손 제스처를 선정 후 이들 중 중복되는 9가지 제스처에 대해서는 굽힘과 뒤틀림의 정도로 분류하여 모두 24가지의 의미 있는 제스처 인식모델을 설정하였다. 본 연구에 있어 기본 손 제스처를 분류는 (그림 2)와 같다.

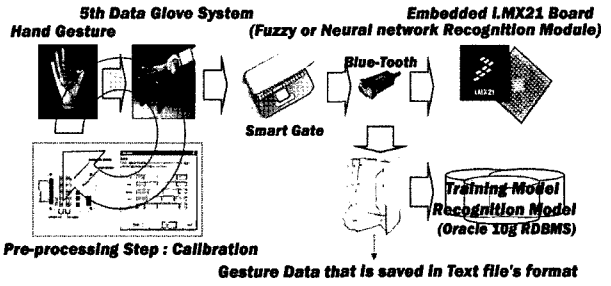
손 제스처의 데이터를 감지하는 장치로 햅틱 응용분야에서 주요한 입력장치로 사용되어지고 있는 5DT사의 장갑장치를 이용하였다. 햅틱 장치로부터 사용자의 의미 있는 동적 제스처 데이터를 획득하는 방법에 있어, 기존의 유선통



(그림 2) 기본 손 제스처의 분류

신모듈 기반의 제스처 데이터 입력 모듈은 공간상의 제약성, 이동상의 한계성, 사용상의 불편함 그리고 케이블 요소의 복잡성 등 몇 가지의 제약사항 및 문제점을 내포하고 있다. 따라서 이러한 문제점을 해결하고 개선하기 위하여 embedded i.MX21 board를 기반으로 blue-tooth모듈을 이용하여 제스처 데이터 입력 방식의 효율성을 향상시키고 사용자의 편의를 도모할 수 있는 임베디드-유비쿼터스 환경을 구현하였다. 5DT사의 장갑 장치는 fiber-optic flex sensor를 사용하여 사용자의 손가락 굽힘 정도를 파악하고 이로부터 데이터를 획득할 수 있는 기본적인 제스처 인식 장치로서 pitch 및 roll sensor를 자체 내장하고 있으므로 손목의 굽힘과 뒤틀림에 대한 측정도 별다른 장비 없이 가능하다는 특징을 가진다. 장갑 장치의 출력 데이터는 9bytes로 header, f1, f2, f3, f4, f5, pitch, roll, checksum의 구조를 가지며 손가락 움직임 정보는 f1=thumb, f2=index, f3=middle, f4=ring, f5=little finger의 순서로 출력되어지고 각각의 데이터 영역은 0에서 255까지 설정되어, 0에 가까운 값이면 손가락의 상태는 펴진 상태이고 255에 가까운 값이면 굽힌 상태를 나타낸다[10]. 제스처 입력모듈은 설정된 제스처 인식모델에 대하여 사용자로 하여금 장갑장치를 이용하여 의미 있는 손 제스처를 입력하고 입력된 데이터를 분석하는 전처리부로 calibration control function을 포함하도록 설계되어졌다. 이러한 calibration control function은 사용자의 손의 크기나 혹은 손 제스처의 행위에 있어 대부분의 경우 개별적인 특성을 가지므로 이에 따른 불필요한 노이즈와 왜곡을 제거하고 사용자에 따른 개별적인 데이터를 효율적으로 수용할 수 있는 과정으로서 가장 적절하고 유효한 데이터를 획득하기 위한 방안으로 사용되어진다. S/W calibration의 원리는 dynamic range를 최대 출력값(value_max)과 최소 출력 값(value_min)의 관계를 수식(1)로 정의하고 최소 출력 값이 40, 최대 출력 값이 206 이라면 dynamic range는 166이 된다. 이 값을 255의 범위로 스케일 변화하면 수식(1) 을 적용하여, 40은 0으로 스케일 변화하고 206은 255로 스케일이 변화되어진다. 수식(1)은 S/W calibration의 기본 원리를 수식으로 표현하였다.

$$Dynamic\ _range = value\ _max - value\ _min$$
$$Value\ _scaled = (value\ _measured - value\ _min) * \left(\frac{255}{dynamic\ _range}\right) \quad (1)$$

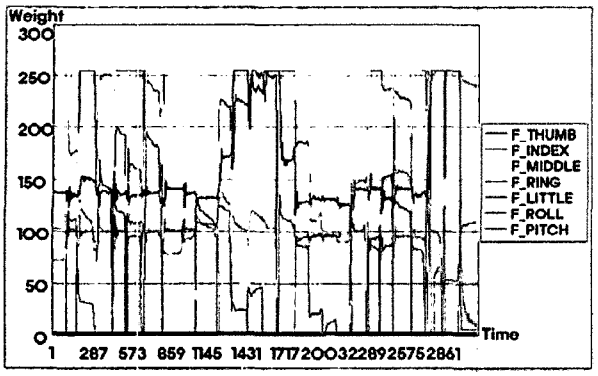


(그림 3) 제스처 입력모듈의 개요

입력된 사용자의 동적제스처 데이터는 스마트게이트 역할의 노트북 컴퓨터를 통하여 WPS(embedded i.MX21 board) 및 서버(Oracle 10g & RDBMS)로 전송한다. WPS로 전송된 제스처 데이터는 의미 있는 제스처 인식을 위한 퍼지 및 신경망 인식모듈의 입력변수로 사용되어지며, 서버로 전송된 제스처 데이터는 text형태의 document로 저장되어 Oracle 10g RDBMS의 SQL analytic function에 의해 세그먼테이션 되어지고 의미 있는 제스처 인식 모델을 형성한다. 이러한 제스처 입력 모듈의 개요를 살펴보면 (그림 3)과 같다.

4. RDBMS를 이용한 학습 및 인식모델

K-means 클러스터링, QT(Quality Threshold) 클러스터링, Fuzzy c-means 클러스터링, Self-Organizing Map (SOM)과 같은 통계적 클러스터링 알고리즘들은 기존의 unsupervised 패턴 인식 및 기계 훈련, 데이터마이닝, 이미지 분석 등의 분야에서 보편적으로 이용되어져왔다. 그러나 이러한 기법들의 적용에 있어서는 다차원적 특징에 따른 복잡한 수학적 계산 알고리즘이 필요하고 분산처리환경에서는 적용하기가 매우 힘들며 패턴(데이터)의 크기에 따라 계산 비용이 결정되어지고 메모리 교환과 할당의 최소화를 위한 여러 가지 제약사항 및 문제점들을 고려하여야만 한다[11-13]. 따라서 본 논문에서는 입력되어진 동적 제스처 데이터에 대하여 의미 있는 유효 제스처 구간과 동작간의 천이구간을 분류하고 효율적인 분석 및 클러스터링을 위하여 RDBMS(Relational Database Management Systems)를 채택하였고 SQL(Structured Query Language)기반의 analytic function을 적용함으로써 별도의 계산 및 비용이 필요 없이 사용자에 의한 동적 제스처의 학습 및 인식모델을 형성할 수 있었다. 학습모델의 구축 과정은 장갑 장치에서 입력된 데이터 분석을 통해 의미 있는 제스처와 각 동작간의 천이구간의 데이터 구간을 분류하는 세그먼테이션 과정이라 볼 수 있다. 입력 제스처 데이터의 현재행과 과거 3개의



(그림 4) SQL에 의한 데이터 분류 및 세그멘테이션

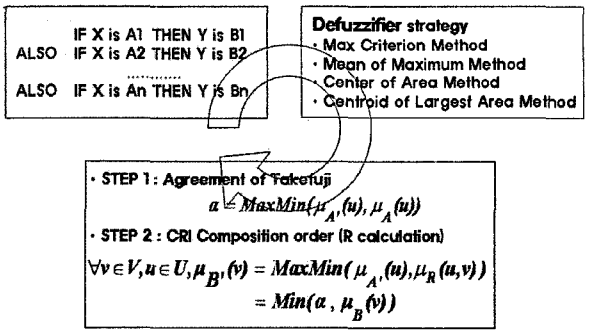
행의 평균값을 비교하여 그 차이가 5 이상이면 천이구간의 레코드로 간주하고 대상 데이터의 다섯 손가락 값 중 하나의 손가락이라도 5 이상의 차이가 날 경우 천이 구간의 레코드로 간주한다. 연속된 천이구간과 유효구간의 구분을 위하여 위의 과정에서 유효하다고 판정된 구간일 지라도 연속적인 값이 5 이상이 차이나지 않을 경우 천이 구간으로 간주한다. 인식모델의 구축 또한 유효 제스처 구간에 대한 개개인의 개별적인 평균값을 결정하고 계산되어진 각각의 통계량으로부터 각 제스처에 대한 통합 평균과 분산비를 재계산하여 이를 규정된 24가지 기본 손 제스처 분류 기준에 따라 비교 분석 후 인식모델로 데이터베이스에 입력하고 저장한다. 사용자의 연속적인 데이터는 SQL의 analytic function에 의해 의미 있는 유효데이터 구간과 천이구간을 신뢰도 98%의 범주로 세그먼테이션 하였으며 (그림 4)는 SQL의 Analytic Function에 의한 입력 데이터의 세그먼테이션 결과를 보여준다.

5. Fuzzy와 신경망을 이용한 제스처 인식 모듈

5.1 Fuzzy Max-Min 인식 모듈

퍼지 논리는 임베디드 시스템에서의 제어 및 정보 처리와 관련된 수많은 애플리케이션들에 대하여 강력한 문제해결도구로 사용되어질 수 있으며 모호하거나 부정확한 정보로부터 일정한 결론을 도출할 수 있는 단순한 방법론을 제공한다[14, 15]. 또한 퍼지논리시스템은 퍼지집합, 퍼지 규칙기반, 퍼지 추론엔진, fuzzifier, defuzzifier로 구성된다. 이와 같은 퍼지 논리 시스템은 함수의 근사화 관점에서 볼 때 비선형 시스템을 모델링하는데 사용될 수 있다. 이는 Universal Approximation 정리를 통해 퍼지 논리 시스템에서 만약 전체집합 U가 밀집하다면 U위로 어떠한 비선형 함수도 균등하게 근사화 할 수 있는 능력이 있음을 보여준다. 이러한 퍼지관계는 통상적 관계의 일반화이므로 합성이 가능하고 본 논문에서 제안된 퍼지관계 또한 다음 (그림 5)와 같이 정의된다.

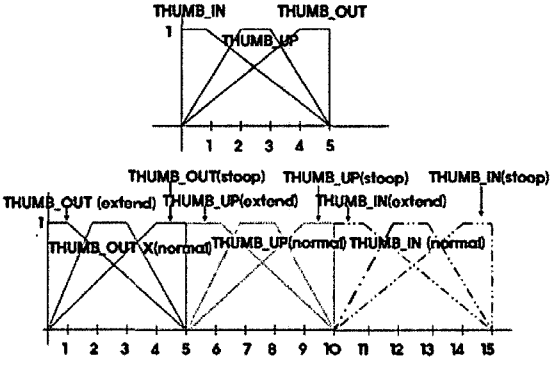
따라서 본 연구에서는 손 제스처에 대한 반복 실험을 통하여 얻어진 제스처 입력 데이터들의 특징들을 언어적 변수



(그림 5) Fuzzy Max-Min CRI (Direct Method)

로 표현하는 규칙을 만들고 표현된 변수에 대하여 가중치를 부여하여 RDBMS 모듈에 의해 구현된 인식 모델과의 비교를 통해 fuzzy max-min 추론을 하도록 설계하였다. 제안하는 퍼지 집합의 소속 함수는 다음의 수식(2)와 같이 정의할 수 있으며 제안되어진 시스템의 fuzzy 입력 변수는 손의 모양과 인식 후보에 따라 매우 다양하게 나타나므로 본 연구에서 사용되어진 fuzzy 소속 함수를 부분적으로 표현하면 (그림 6)과 같다. 예를 들어 "한국 표준 수화사전"에서의 영문 "S" 는 '가위-바위-보' 게임에서 바위에 상응하는 제스처 행위로서 "S"를 표현하는 제스처 타입을 'hand motions'으로 규정하고 3가지 타입으로 정의하였다. 첫 번째 경우, 나머지 모든 손가락은 굽힌 채 엄지손가락을 검지와 중지손가락 위에 위치하는 주먹에 해당하는 제스처로 "THUMB_UP"으로 정의하였고 두 번째 경우에는 나머지 모든 손가락은 굽힌 채 엄지손가락을 검지손가락 위에 위치시켜 이를 "THUMB_UP"으로 정의하였다. 마지막 경우로, 4손가락을 모두 구부린 상태에서 엄지손가락을 그 사이로 넣어 두는 바위의 형태로 이를 "THUMB_IN"으로 정의하였다. 그리고 엄지손가락의 굽힘과 펼침 정도에 따라 'hand actions'에 해당하는 "extend-stoop-normal"로 분류, 정의하였다.

$$\mu_{tz} = \begin{cases} \frac{1}{(s-p)}(x-s)+1, & p<x\le s \ ; \ Slopes\text{-}up \\ 1, & s<x\le t \ ; \ Horizontal\ ity \\ -\frac{1}{(q-t)}(x-t)+1, & t<x\le q \ ; \ Slopes\text{-}down \end{cases} \quad (2)$$



(그림 6) 퍼지 소속함수(The fuzzy membership functions)

5.2 신경망(Neural Network) 인식 모듈

신경망은 생명체의 신경조직에서 착안하여 모델화한 정보처리 시스템으로서 단순한 소자들의 병렬, 분산 연결 구조를 가지고 있으며, 외부로부터 받아들이는 입력에 대하여 동적반응을 일으킴으로써 필요한 출력을 생성시킨다[16]. 생물학적 뉴런들 간의 정보전달에 있어 시냅스가 중요한 역할을 담당하고 있듯이, 처리 요소들 간의 연결 강도(weight)를 반영하기 위해 신경회로망에서는 가중치 연결(weighted-connection)를 사용하는데, 각 처리요소들은 전달받은 입력값과 연결가중치를 사용하여 입력 값을 계산한 후 이를 이용하여 출력 값을 결정한다. 신경망 기술의 가치는 불완전하고 잡음이 많은 입력의 해석뿐만 아니라 패턴인식(pattern recognition), 학습, 분류, 일반화, 추상화 등을 위한 활용성에 있다[16, 17]. 또한 신경망 구조는 기본적으로 병렬로 연산을 처리해내기 때문에 복잡한 자극에 대해서도 반응이 빨라 병렬연산장치에 응용될 수 있으며, 한 부분의 신경망을 통해서도 여러 가지 처리를 해낼 수 있고, 관계된 일들을 연관해서 기억할 수 비교적 저용량의 기억장치에 응용가능하다. 본 논문에서 제안되어진 퍼셉트론 학습 및 가중치 처리에 대한 일반적인 델타 규칙은 다음의 수식(3), (4)와 같이 정의할 수 있다.

$$y(t) = F(W_i \cdot A_i - \theta)$$

where,
- $y(t)$: output of R layer's PE
- F : fucntion that using in R layer's PE
- W_i : weights of A layer's & R layer's PE
- A_i : input of A layer's PE
- θ : critical value of R layer's PE

(3)

$$W_i(t+1) = W_i(t) + @ \cdot (d_i - y_i) \cdot A_i(t)$$

where,
- $W_i(t+1)$: weight after tarining
- $W_i(t)$: weight from input i (time is t)
- $@$: training rate
- $(d_i - y_i)$: error
- $W_i(t)$: input pattern

(4)

RDBMS에 의한 인식모델에 상응하는 의미 있는 제스처 데이터는 신경망의 입력 변수로 사용되어지며 서로 다른 가중치가 부여된다. 신경망으로부터의 출력 결과는 가중치의 합으로 결정되며 이를 제스처 인식결과로 그래픽과 음성으로 사용자에게 제공한다. 신경망을 이용한 제스처 인식 시스템의 기본개념을 (그림 7)과 같이 표현할 수 있다.

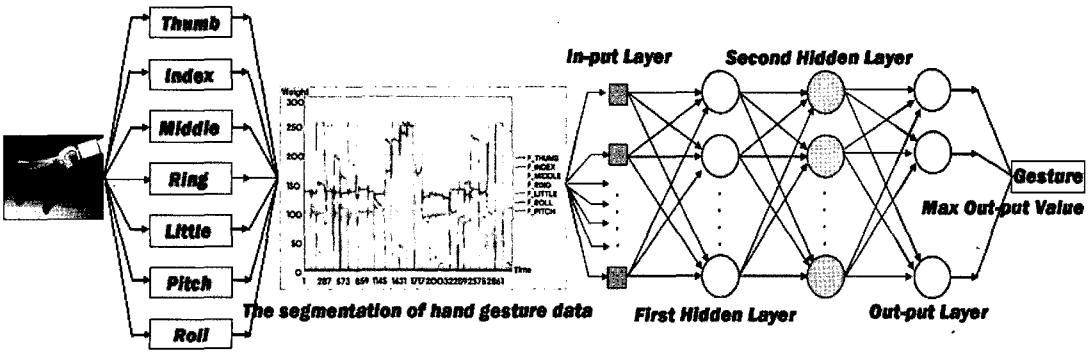
6. 실험 및 결과

제안된 퍼지 및 신경망 제스처 인식 시스템은 제스처 입력모듈로서 임베디드- 유비쿼터스 환경을 구현하기 위하여 무선통신 모듈인 blue-tooth모듈을 적용하였으며, RDBMS 모듈을 구현하기 위하여 LINUX기반의 워크스테이션을 사용하였고 Oracle 10g RDBMS를 이용하였다. 장갑 장치는 19,200[bps]의 전송 속도로 7 종류의 데이터(5개의 제스처 데이터, 2개의 pitch 및 roll data)를 embedded i.MX21 board로 전송한다.

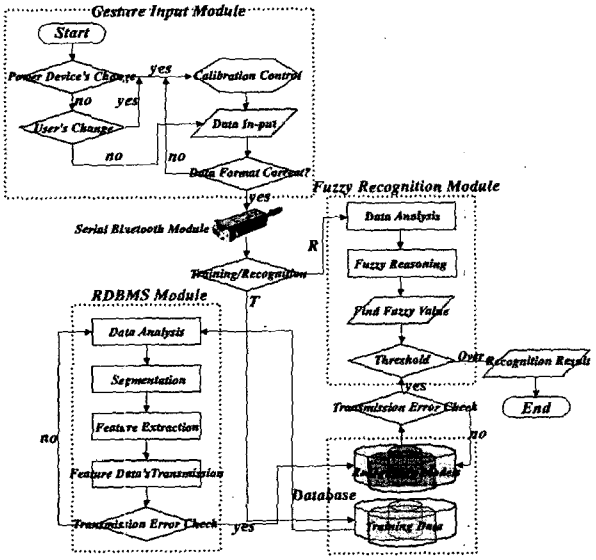
6.1 Fuzzy Max-Min 인식 모듈

제안되어진 퍼지 제스처 인식 모듈의 전체 시스템 크기는 141 Kbytes 이며 WPS - embedded i.MX21 board에서는 초당 200개의 샘플에 대하여 처리 및 연산이 가능하다.

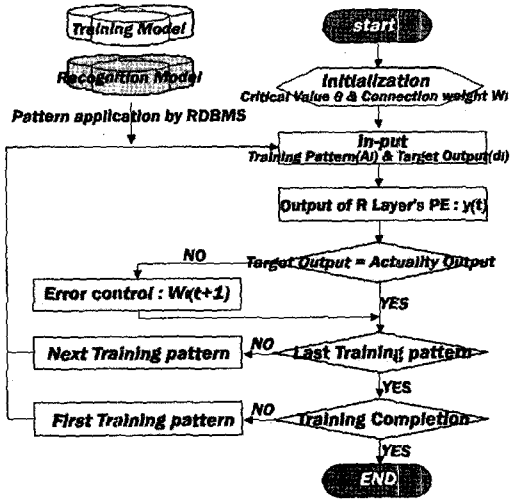
손 제스처 인식 시스템은 3 단계로 구성되어진다. 1 단계로서의 제스처 데이터 입력모듈에서는 사용자가 장갑 장치를 이용하여 규정된 손 제스처를 실행하고 정확한 사용자의 손 제스처를 획득하기 위하여 calibration control 과정을 수행한다.(장갑 장치는 전원 및 사용자의 변화에 따라 손 제스처의 측정값이 상당히 변화되어진다.) 또한, 장갑 장치에서는 0~255에 해당하는 손 제스처 측정값이 출력되어지기 때문에 입력된 제스처 데이터를 분석하여 전송되어지는 데이터 포맷의 정확성을 판단한 후 사용자의 제스처 데이터를 입력받아 데이터의 특징들을 언어적 변수(입력변수)로 변환시킨다. 입력된 손 제스처 데이터는 학습 및 인식모델을 형성하기 위하여 RDBMS 모듈로 전송되어지고 인식모델이 형성된 이후에는 손 제스처 인식을 위한 Fuzzy MAX-MIN 인식 모듈로 손 제스처 데이터(입력변수)가 전송되어진다. 2



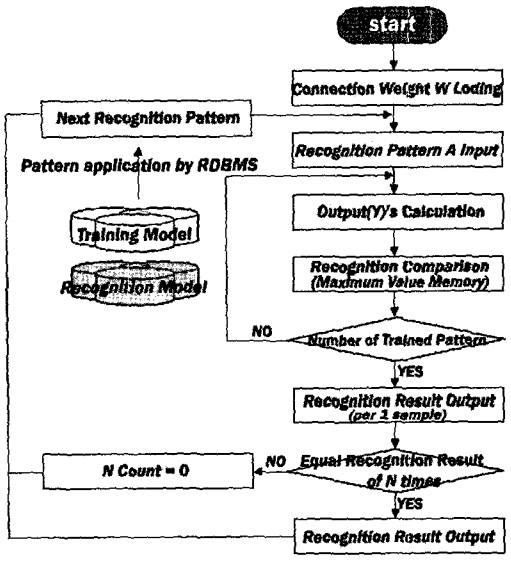
(그림 7) 신경망을 이용한 제스처 인식 시스템의 기본 개념



(그림 8) Fuzzy Max-Min 인식 모듈의 흐름도



(그림 9) 신경망을 이용한 인식 모듈의 흐름도



(그림 10) 신경망을 이용한 인식 모듈의 흐름도

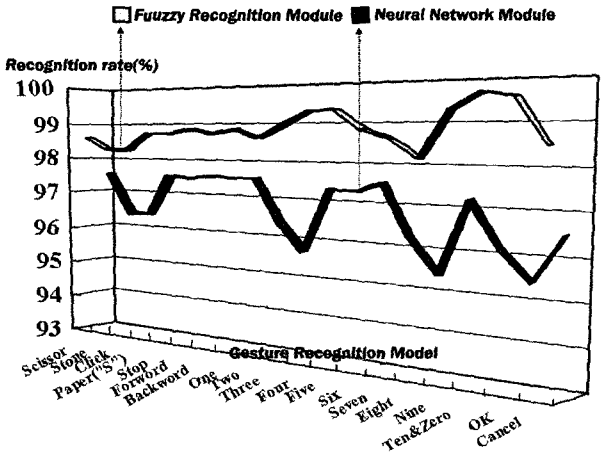
단계는 제스처 데이터의 세그먼테이션을 위한 RDBMS모듈로서 RDBMS의 SQL analytic function에 의해 제스처 데이터를 분석하고 세그먼테이션 처리를 수행하며 그 결과를 바탕으로 입력된 손 제스처의 특징을 추출한다. 손 제스처의 특징 추출 결과는 인식 모델로서 데이터베이스에 전송 및 저장되며 전송 중 에러나 저장 시 에러가 발생하면 이를 체크하여 재전송을 요구한다. Fuzzy MAX-MIN 인식 모듈은 3단계 처리 과정으로 제스처 데이터 입력모듈로부터 전송되어진 입력변수의 값(입력 제스처 데이터)이 Fuzzy MAX-MIN 인식 모듈에서 설정한 제스처 인식후보에 포함되는 정도에 따라 인식후보에 대하여 가중치를 부여하여 퍼지 추론 과정을 수행하며 Fuzzy-value를 산출하게 된다. 산출되어진 Fuzzy-value는 인식모델과의 비교연산을 통해 그 정도에 따라 사용자의 손 제스처 데이터는 최종적으로 인식하고 인식결과는 사용자에게 음성 및 그래픽으로 제공한다. 본 연구에서 제안한 전체 시스템의 흐름도는 (그림 8)과 같다.

6.2 신경망을 이용한 학습및 인식 모듈

신경망 제스처 인식 모듈의 전체 파일 크기는 215 Kbytes 이며 학습 및 인식 연산의 수행정도에 따라 WPS에서는 초당 175개의 샘플에 대하여 처리 및 연산이 가능하다. 신경망 제스처 인식 시스템은 크게 학습 모듈과 인식모듈로 구분되어지며 학습 모델의 처리 과정은 1) RDBMS에 의하여 결정된 입력 패턴과 목표 출력 패턴에 대하여 임계값과 연결가중치를 초기화한 후에, 2) R 계층의 PE의 출력을 계산하고 목표 출력과 현재 출력을 비교하게 된다. 3) 만약 목표 출력과 현재 출력을 비교하여 동일한 패턴이면, 학습을 종료하고 그렇지 않으면, 에러 제어를 수행한 후, 학습 과정을 재 수행한다. 신경망 인식 모듈의 전체 인식 과정은 1)각 계층에서의 연결가중치를 취한 후 장갑장치를 이용하여 인식 패턴을 입력하고, 2) 출력 결과를 계산하며 인식율과 비교하여 최대값을 기억한다. 3) 학습되어진 패턴을 따른다면 이를 인식결과로 그래픽과 음성으로 사용자에게 제공한다. 제안되어진 신경망과 RDBMS를 이용한 제스처 학습 및 인식시스템의 전체 흐름은 (그림 9), (그림 10)과 같다.

6.3 실험 및 결과

본 논문에서는 서버(Oracle 10g RDBMS)와 제스처입력모듈간의 이격거리를 반경 10M의 타원 형태로 구성하고 규정된 위치로 장갑장치를 이동시키면서, 30인의 피 실험자에 의해 규정된 24가지의 제스처에 대하여 15회의 반복 실험을 수행하였다. 실험 결과, 사용자의 의미 있는 동적 제스처 데이터에 대하여 퍼지 제스처 인식모듈에서는 98.8%, 신경망 제스처 인식모듈에서는 96.7%의 평균 인식률을 도출하였다.



(그림 11) 평균 인식율의 비교 및 평가

(그림 11)에서는 개별 제스처 인식 시스템의 평균 인식률을 보여준다. 또한 기본인식 모델에 대하여 각 모듈의 인식 시간을 비교, 분석해 본 결과 퍼지 제스처 인식모듈에서는 0.16[sec], 신경망 제스처 인식모듈에서는 0.23[sec]의 평균 인식시간이 소요됨을 알 수 있었다. 인식 중 발생하는 오류는 규정된 손 제스처를 실행함에 있어 동작 미숙과 사용자의 동작 미 숙지, 장갑 장치 굴곡 센서의 물리적인 변형, 주변 실험 환경의 시간적 변화 등에 의한 인식률 저하로 규정할 수 있다.

7. 결 론

본 논문에서는 동적이며 연속적인 사용자의 제스처를 차세대 착용형 PC 폼랫폼인 i.MX21 board를 기반으로 blue-tooth모듈을 적용하여 임베디드-유비쿼터스 환경에서 실시간으로 인식하는 제스처 인식 인터페이스를 구현하는데 그 목적을 두었고, 이를 통해 유선방식에 한정되어 있던 제스처 인식 모듈의 여러 제약사항을 극복하고 해결할 수 있는 하나의 방안을 제시하였다. 또한 퍼지 알고리즘과 신경망을 이용하여 2개의 독립적인 제스처 인식 시스템을 구현하였고 각 시스템에 대한 제스처 인식률, 시스템의 처리 능력, 인식 시스템의 전체 시스템 크기 그리고 인식모델에 대한 평균 인식 시간 등과 같은 성능들의 분석 및 평가를 통하여 차세대 PC에 기반의 임베디드-유비쿼터스 환경에서는 퍼지 알고리즘을 응용한 제스처 인식 시스템이 신경망을 이용한 제스처 인식 시스템보다 보다 효율적이며 합리적임을 증명할 수 있었다. 그리고 본 연구를 통해 퍼스널 헬스케어, 엔터테인먼트 등의 형태로 언제, 어디서나, 누구나 자신만의 디지털라이프스타일을 즐길 수 있는 보편적인 정보통신 이용 환경을 제공할 수 있을 것이며 아울러 장애인의 복지 향상 및 차세대 PC에 대한 새운 연구방향 및 발전 가능성을 제시할 수 있을 것이다.

참 고 문 헌

[1] 한동원, "유비쿼터스 컴퓨팅 시대의 차세대PC", 주간기술동향 통권 1236호, 정보통신진흥연구원.
[2] 정보통신진흥연구원, "IT신성장동력 - 차세대PC"
[3] 박준석, "차세대 휴먼 인터페이스의 오감 정보처리 기술", 주간기술동향 통권 1252호, 정보통신진흥연구원.
[4] 손미숙 외, "착용형 컴퓨터를 위한 햅틱 기술 동향", 전자통신동향분석 제20권 제5호, 2005.
[5] Jong-Sung Kim, Won Jang, Zeungnam Bien, "A Dynamic Gesture Recognition System for the Korean Sign Language (KSL)", Ph.D. Thesis, 1996.
[6] http://www.wearablegroup.org
[7] http://www.essentialreality.com
[8] http://www.freescale.com, "i.MX21 Applications Proce-ssor Brochure"
[9] 김승국, "표준수화교본", 오성출판사, 2000.
[10] "5DT Data Glove 5 Manual", http://www.5dt.com
[11] Richard O. Duda, Peter E. Hart, David G. Stork,"Pattern Classification", 2nd, Wiley, New York, 2001.
[12] Dietrich Paulus and Joachim Hornegger, "Applied Pattern Recognition", 2nd, Vieweg, 1998.
[13] J. Schuermann, "Pattern Classification: A Unified View of Statistical and Neural Approaches", Wiley&Sons, 1996.
[14] W. B. Vasantha kandasamy, "Smaranda Fuzzy Algebra", American Research Press, Seattle, 2003.
[15] Von Altrock C, "Fuzzy Logic and NeuroFuzzy Applications Explained", 2002.
[16] C.H.Chen, "Fuzzy Logic and Neural Network Handbook", 1st edn. McGraw-Hill, New York, 1992.
[17] Abdi, H. "A neural network primer", Journal of Biological Systems, 1994.

김 정 현
e-mail : kjh0328@skku.edu
2000년 대전대학교 전자공학과(학사)
2002년 대전대학교 정보통신공학부 (공학석사)
2002년~2004년 서강전문학교 정보통신과 전임강사
2004년~현재 성균관대학교 정보통신공학부 박사과정
관심분야 : 제스처 인식, 멀티 모달 오감 인식, HCI

노 용 완

e-mail : elec1004@hotmail.com

2001년 남서울대학교 정보통신공학과 (학사)

2003년 성균관대학교 정보통신공학부 (공학석사)

2003년~현재 성균관대학교 정보통신공학부 박사과정

관심분야 : 음성인식, 음성이해, 신호처리

홍 광 석

e-mail : kshong@skku.ac.kr

1985년 성균관대학교 전자공학과(학사)

1988년 성균관대학교 전자공학과 (공학석사)

1992년 성균관대학교 전자공학과 (공학박사)

1990년~1993년 서울보건전문대학 전산정보처리과 전임강사

1993년~1995년 제주대학교 정보공학과 전임강사

1995년~현재 성균관대학교 정보통신공학부 교수

관심분야 : 음성인식, 오감 인식 및 재현, HCI