

# 지능형 홈을 위한 HTM 기반의 얼굴 이미지 인식 시스템의 개발

배선갑<sup>†</sup>, 이대한<sup>\*\*</sup>, 조건화<sup>\*\*\*</sup>, 남해보<sup>\*\*\*\*</sup>, 김성진<sup>\*\*\*\*\*</sup>, 강현석<sup>\*\*\*\*\*</sup>, 배종민<sup>\*\*\*\*\*</sup>

## 요 약

본 논문에서는 Hopkins가 제안한 인간두뇌의 신피질(neocortex)의 동작 원리를 컴퓨터에 접목시킨 HTM 이론을 적용하여 가족 구성원들의 얼굴을 인식하는 지능형 홈을 위한 얼굴 이미지 인식 시스템인 FRESH(Face image REcognition System for intelligent Home)를 개발하였다. 개발 결과를 확인하기 위해 실제 사진을 촬영하여 깨끗한 이미지와 왜곡된 이미지에 대하여 실험해 본 결과 깨끗한 이미지뿐만 아니라 실제 환경에서 흔히 입력되는 다양하게 왜곡된 얼굴 이미지에 대한 인식의 정확도가 높았다. 또한, 기존의 연구들은 얼굴영역 인식에 초점이 맞추어져 있어서 지능형 홈을 위한 시스템에 적용하기가 어려우나 본 연구는 얼굴 이미지와 그 사람의 이름을 일치시킨 것으로 지능형 홈에 적합한 시스템으로 개발하였다.

## Development of An HTM Based Face Image Recognition System for Intelligent Home

Sun-Gap Bae<sup>†</sup>, Dae-Han Lee<sup>\*\*</sup>, Jian-Hua Diao<sup>\*\*\*</sup>, Hai-Bao Nan<sup>\*\*\*\*</sup>,  
Sung-Jin Kim<sup>\*\*\*\*\*</sup>, Hyun-Syug Kang<sup>\*\*\*\*\*</sup>, Jong-Min Bae<sup>\*\*\*\*\*</sup>

## ABSTRACT

This paper developed *FRESH*(Face image REcognition System for intelligent Home), proposed by Hopkins, which was to recognize the face of family members applying for HTM system that integrates a principal movement of human brain's neocortex with computers. For the purpose of confirming a result of the works, we took real photographs of both clear and distorted images, but the result produced higher accuracy about the distorted images providing by a real environment as well as clear images. In addition, the existing studies were focused on the recognition of facial zone that it was difficult to adapt for System for intelligent home. However, this thesis developed a system suitable for intelligent home which corresponds a facial image to his or her name.

**Key words:** Face image recognition(얼굴 이미지 인식), HTM(계층형 시간적 메모리), neocortex(신피질)

\* 교신저자(Corresponding Author) : 배종민, 주소 : 경상남도 진주시 가좌동 900번지(660-701), 전화 : 055)751-5995, FAX : 055)762-1944, E-mail : jmbae@gnu.ac.kr

접수일 : 2008년 12월 19일, 완료일 : 2009년 4월 20일  
<sup>†</sup> 준회원, 경상대학교 대학원 컴퓨터학과 박사과정 (E-mail : bsgap@hanmail.net)

<sup>\*\*</sup> 준회원, 경상대학교 대학원 컴퓨터학과 재학(석사) (E-mail : neokshacker@nate.com)

<sup>\*\*\*</sup> 준회원, 경상대학교 컴퓨터학과 박사과정

(E-mail : diaoxiaoxiao@hotmail.com)

<sup>\*\*\*\*</sup> 준회원, 경상대학교 컴퓨터과학과 석사과정 (E-mail : nanhaibao666@hotmail.com)

<sup>\*\*\*\*\*</sup> 정회원, 연암공업대학 컴퓨터공학과 교수 (E-mail : sjkim@yc.ac.kr)

<sup>\*\*\*\*\*</sup> 중신회원, 경상대학교 컴퓨터과학과 교수 (E-mail : hskang@gnu.ac.kr)

<sup>\*\*\*\*\*</sup> 정회원, 경상대학교 컴퓨터과학부 교수

## 1. 서 론

최근 편안하고 안전한 홈 환경을 만들기 위한 지능형 홈의 개발이 많이 이루어지고 있다. 지능형 홈을 위한 시스템이 되기 위해서는 얼굴 인식 기능의 개발이 필요하다. 얼굴 인식은 영상 기술을 바탕으로 많은 연구가 진행되고 있다. 그러나 기존의 얼굴 인식 연구들의 대부분은 얼굴영역 인식에 초점이 맞추어져 있어서 특정한 사람의 이름과 얼굴을 일치시키는 기능이 필요한 지능형 홈을 위한 시스템에 사용하기에는 부적합한 기술들이다. 기존의 얼굴인식 연구들의 방법과 적용된 이론들을 살펴보면 대부분 얼굴인식의 첫 단계로 얼굴 영역 및 얼굴 특징 추출 방법을 사용해 왔다. 그 예로서 색 정보를 이용하는 방법, 주성분 분석법(Principal Component Analysis)에 의한 고유얼굴(eigenface) 이용 방법, 색 정보와 동영상 이용 방법, 신경망 이용 방법 등이 사용되고 있다[1].

이들 방법들은 얼굴 영역을 검출하고, 주변 환경의 영향에 기초해 얼굴 영역을 크기 변화, 회전 변화 및 밝기 변화 등으로 정규화를 수행한다[2]. 그리고 난 후 이 얼굴 영역에 대한 특징들을 추출하고, 여러 가지 변환들을 적용하여 연산을 수행한 후에 인식 과정을 거친다. 그리고 인식 과정에서 인공신경망(Artificial Neural Network), 최근접 분류기(Nearest Neighbor Classifier), SVM(Support Vector Machines) 등이 이용되어 왔다[3].

그러나 인공신경망은 상호 연결된 뉴런에 의해 연상 기억, 필터, 변환, 분류, 최적화 등의 기능을 수행할 수 있지만 자료가 가지는 시계열 모형의 특징을 기억 구조로서 인공신경망의 학습구조에 도입한 연구는 매우 드물고, 네트워크 학습을 위해서는 많은 시간이 필요하며, 이런 일련의 학습 과정은 새로운 자료가 추가되면 전체 과정을 모두 다시 수행해야 하는 단점을 가지고 있다[4]. 최근접 분류기는 입력 벡터가 주어지면 유사도 함수에 기반하여 참조 집합에서 입력 벡터와 가장 유사한 K개의 벡터를 선택한다. 이 K개의 벡터 이웃들의 레이블과 유사도 값의 분포에 따라 입력 벡터의 클래스를 결정하게 된다. 이는 일반적으로 인식할 대상의 수가 적을 경우 잘 분류되나 인식할 대상의 수가 많아지면 클래스간 특징들의 유사도가 증가하여 인식에 어려움이 따른다[5]. SVM은 선형 분리 가능한 샘플들에 대하여 클레

스간의 거리가 최대가 되도록 함으로써 비교적 우수한 성능을 나타낸다. 하지만 얼굴 영상의 다양성과 복잡성 등의 특징을 효과적으로 인식할 수 있는 방법에 대한 개발이 요구된다[6]. 따라서 본 논문에서는 이들이 가지고 있는 문제점을 해결하기 위하여 지능형 홈에 새로운 HTM 이론을 적용하였다.

HTM 이론은 Jeff Hopkins가 최근 제안한 계층형 시간적 메모리(Hierarchical Temporal Memory (HTM)) 이론[7]이다. 이는 인간 두뇌의 외부를 둘러싸고 있는 얇은 막인 신피질(neocortex)의 동작 원리를 흉내낸 것으로 인간 두뇌를 구성하는 뉴런들과 시냅스들이 계층적으로 시간적 메모리 구조를 갖고 있다고 보는 이론이다. 그리고 이 이론은 기존 인공신경망 이론 보다 노드의 역할이 좀 더 복잡한 연산을 수행하는 방법을 사용하며, 실제세계에 대한 시공간적인 패턴을 계층적으로 기억한다. 그런데 지능형 홈을 위한 얼굴 인식 시스템이 되기 위해서는 조명이나 잡음(noise), 얼굴의 회전(rotation) 또는 이동(shift) 등의 다양한 가변 환경에서도 인간이 얼굴을 인식할 때와 같은 수준의 인식률이 요구되며 각 가정마다 가족 구성원들에 대해 쉽게 주문형으로 개발이 이루어질 수 있어야 한다. 따라서 각 가정마다 다양한 가변 환경에서 사용 가능한 얼굴 이미지 인식 시스템을 개발하기 위해서는 인간 두뇌의 인식 기능에 비견할 수 있는 적합한 기술이 필요하다. 인간의 얼굴 이미지 인식 능력은 주로 두뇌의 신피질의 기능에서 비롯된다고 보고 있으므로 이 HTM 기술이 지능형 홈을 위한 시스템에 적합하다.

본 논문은 이 HTM 기술을 이용하여 지능형 홈을 위한 얼굴 이미지 인식 시스템 FRESH(Face image REcognition System for intelligent Home)를 개발하였다. 우리는 FRESH의 개발을 위해 가정에서 활동하는 사람들의 얼굴을 다양한 환경에서 카메라 촬영을 통해 이미지를 추출해 그룹화하고, 적절한 HTM 네트워크 구조를 결정한 후, 다양한 이미지들을 훈련시키는 실험 결과를 통해 적합한 환경을 설정하는 방법을 사용하였다. 그리고 우리는 개발한 시스템을 다양한 잡음(noise), 크기 변환(Scale), 줄 삽입 또는 폐색(Line or Occlusion), 이동 변환(Translation) 등이 있는 새로운 얼굴 이미지 데이터들에 대해 추론하여 인식하는 실험을 수행하는 방법으로 평가하였다. 그 결과 이러한 영향들이 있는 실제 상황에서도

인식 능력이 우수한 것으로 확인되었다.

논문의 구성은 다음과 같다. 2장에서는 FRESH 개발에 사용한 HTM 이론의 전반적인 내용을 기술하고 3장에서는 우리가 개발한 FRESH의 개념을 설명한다. 4장에서는 이의 구현과 그 결과를 실제 환경에 적용해 분석한 후, 5장에서는 기존 연구들과 비교한다. 마지막 6장에서는 결론과 향후 과제를 논한다.

## 2. HTM 이론

이 장에서는 FRESH 개발에 사용된 HTM 이론의 전반적인 내용을 기술한다.

### 2.1 HTM 이론

기존의 경험적 탐색을 기본으로 하는 인공지능(AI)이나 단순 뉴런들의 연결로 보는 인공신경망(ANN)과는 달리, HTM 이론에서는 시스템을 일정한 공동 연산을 수행하는 메모리 노드들을 계층 네트워크(hierarchy network)로 구성한다[8]. 이때, 네트워크의 기본 단위인 노드(node)는 6개 층으로 이루어진 신피질(neocortex)의 기본 단위인 신피질 칼럼(neocortex column)에 해당하는 것으로 본다. 이렇게 구성된 HTM 네트워크는 세계의 시공간적 패턴 정보(특히, 시간 정보)를 이용해 효율적으로 지능적 판단을 할 수 있게 한다. 또한, HTM은 일종의 메모리 시스템으로 HTM에서는 프로그램 되지 않으며 문제마다 다른 알고리즘을 수행하는 것이 아니라 문제를 해결하는 방법을 학습한다. 그리고 학습 과정에서 감각 신호의 특성들 사이의 시공간적 관계성이 계층적 메모리 구조에 형성된다. 이를 위해 외부 감

각 데이터에 자신을 노출시키는 방법으로 훈련을 하게 되며, 모델링된 방식의 레벨을 통해 학습된 정보를 저장한다. 학습이 끝나면 새로운 패턴을 인식하게 되는데, 인식 과정은 저장된 표현 중에서 그것을 최고로 잘 예측할 수 있는 것을 선택하는 것이라 볼 수 있다.

### 2.2 기본 알고리즘

HTM에서 네트워크는 노드들의 계층으로 구성된다. 데이터는 계층의 바닥으로부터 감각(sensor)을 통해 입력되고, 입력된 데이터는 노드들의 계층 구조를 통해 위로 올라가면서 처리 되며, 이때 계층의 각 노드는 학습과 추론의 두 단계로 이루어진 같은 알고리즘을 사용한다[9].

학습 단계에서는 네트워크가 훈련 패턴에 노출되어 범주들에 패턴들을 사상시키는 모델을 구축하고, 학습이 완료되면 네트워크는 추론 단계로 넘어간다. 추론 단계에서는 네트워크가 이전에 본 일이 없는 테스트 패턴에 대해 범주들에 대한 간단한 분포를 생성한다. 노드들의 개략적인 구조를 그림 1에 나타내었다. 그림에서 보는 것과 같이 HTM 네트워크는 Zeta1, Zeta1Top, Sensor, Effector 등 몇 가지 유형의 노드들로 구성된다[10]. 이들 중 가장 핵심적인 것이 Zeta1 노드로서 계층의 대부분을 구성하게 된다. 여기서는 이 Zeta1 노드에 대해 간략하게 알아본다.

각 Zeta1 노드는 정보를 그림 2와 같이 공간적 풀러(spatial pooler)와 시간적 풀러(temporal pooler)의 두 가지 구성요소로 처리하게 된다. 공간적 풀러는 데이터의 공간적 패턴을 분류하는 역할을 수행한다. 즉, 입력 벡터들의 분산된 표현들을 분류하는 방법으

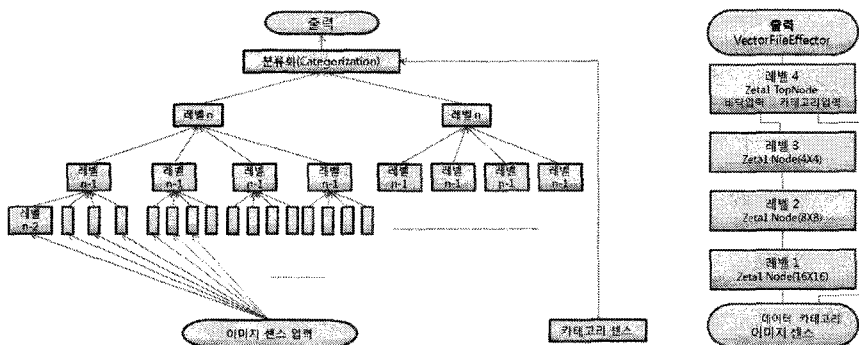


그림 1. HTM 시스템의 계층 구조

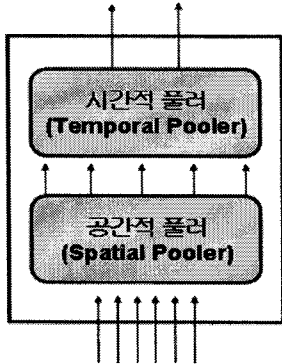


그림 2. Zeta1 노드의 내부 구조

로 학습한다. 시간적 풀러는 데이터의 동시발생을 시간적 패턴을 그룹화 하는 알고리즘을 수행한다. 이는 추론 결과가 이전 입력 값뿐만 아니라 최근의 입력 값을 기초로 하여 계산하는 시간 기반의 알고리즘이다. 보다 상세한 것은 [10]을 참조하기 바란다.

### 2.3 학습과 추론

HTM 네트워크의 동작은 크게 학습과 추론 단계로 이루어진다[10]. 학습은 HTM 네트워크에서 레벨(level) 별로 이루어진다. 즉, 학습 단계에서 공간적 풀러는 입력 패턴을 정량화하여 정량화 센터(quantization center)에 기억한다. 이 공간적 풀러의 출력은 시간적으로 동시발생되는 패턴들을 알기 위한 시간적 풀러의 입력으로 들어간다. 시간적 풀러는 시간적으로 동시발생되는 패턴들을 그룹 지음으로써 학습을 완료된다.

레벨별로 학습이 완료 되면 입력 데이터들에 대해 추론하는 추론 단계로 이행된다. 즉, 추론(inference) 단계에서는 매 레벨마다 학습이 끝나면 위 레벨로 보낼 데이터를 결정하기 위해 추론 단계를 수행한다. 추론은 아래로부터의 입력 패턴에 대해 학습 단계에서 학습한 동시발생에 대한 그룹(시간적 풀러에 의한)들에 대한 범주의 분포를 발생(공간적 풀러에 의한)시킨다.

이렇게 하여 네트워크의 모든 노드들이 전체적으로 학습이 완료되면 훈련이 끝난 HTM 네트워크가 완성된다. 이제 이를 이용하여 새로운 입력 데이터를 인식할 수 있게 되는데, 이것은 최상위 레벨에서 추론을 하는 것을 의미한다. 그리고 그 결과는 입력 데이터의 패턴 범주들에 대한 분포를 나타낸다.

## 3. HTM 기반의 얼굴 이미지 인식 시스템

이 장에서는 우리가 개발한 지능형 홈을 위한 HTM 기반의 얼굴 이미지 인식 시스템 FRESH (Face image REcognition System for intelligent Home)의 개념을 설명한다.

### 3.1 시스템의 수행 구조

HTM 기술을 이용한 얼굴 이미지 인식 시스템의 개발은 그림 3과 같이 Numenta 사에서 추천하는 절차[9]에 몇 가지를 추가하는 방법을 통해 이루어졌다. 그림에서 보는 것과 같이 우선 노드 생성과 노드의 연결을 통해 네트워크를 형성한 후 실행 네트워크를 생성한다. 이때 HTM 엔진에 학습시킬 훈련 데이터는 사전에 수집된 데이터를 활용하며 HTM 네트워크는 이들을 이용하여 훈련시킨다. 훈련(학습과 추론 기능 이용) 과정을 거친 네트워크 파일이 형성 되면 새롭게 인식하고자 하는 테스트 데이터를 입력 데이터로 받아서 인식하고 이를 바탕으로 분석하여 적절한 동작을 수행하게 된다. 그런데 좋은 HTM 응용이 되기 위해서는 이러한 새로운 데이터의 다양한 인식 실험을 바탕으로 HTM 네트워크의 인식 성능 향상을 위해 여러 가지 파라미터 조정 작업을 거쳐야 한다. 우리는 이를 위해 조정된 HTM 네트워크가 만족한 레벨에 이를 때까지 파라미터를 조정하고 실험을 반복하는 방법을 추가하였다.

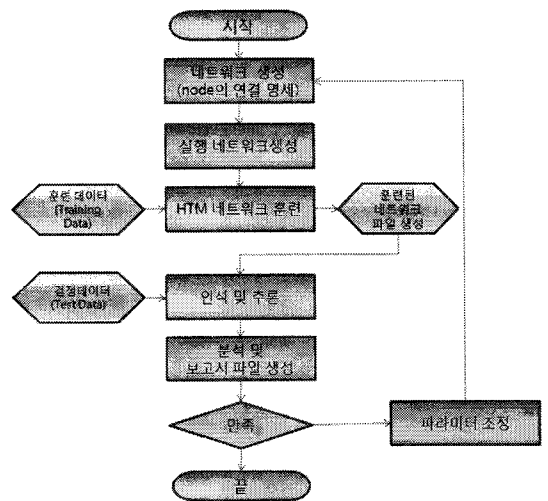


그림 3. 얼굴 이미지 인식 시스템의 수행 구조

### 3.2 HTM 네트워크 구성

그림 4와 같이 우리는 FRESH를 위한 네트워크로 총 4개의 레벨을 구성하였다. 여기서 각 노드는 하나의 사각형으로 표현되어 있다. 가장 아래의 센서 노드는 이미지 센스로서 외부의 데이터를 입력 받게 되고, 이것은 앞장에서 본 Zeta1 노드들로 구성된 가장 아래 레벨 1로 입력된다.

네트워크에서 가장 높은 레벨의 노드는 한 개로 이루어져 있으며, 이는 Zeta1 노드와 유사하지만 외부로 출력을 보내는 효과(effector) 노드와 연결시키기 위해 특수 기능을 담당하는 Zeta1Top 노드가 된다. 이렇게 하여 HTM 네트워크는 레벨별로 1개, 16개, 64개, 256개의 노드로 구성하였다. 센서 노드를 통해 들어오는 입력 데이터는 크기는 128×128 픽셀로 이루어져 있다. 레벨 1의 16×16개 노드 각각에 64(8×8)픽셀로 반영되어 사상(mapping)하게 하였다.

높은 레벨에서의 각 노드는 낮은 레벨의 여러 노드들로부터 입력을 받게 된다. 그림 4의 레벨 2의 노드는 레벨 1의 4(2×2)개의 결과들로부터 입력을 받는 것을 나타낸 것이다. 이렇게 자식 노드들의 결과를 입력으로 받아들이는 과정을 거치면 가장 높은 레벨에서의 노드는 전체 이미지 영역을 한 번에 인식하게 된다.

### 3.3 파라미터 결정

네트워크를 구성하고 생성한 후 네트워크 훈련과정에서 파라미터를 조정하여 가장 정확도가 높은 값

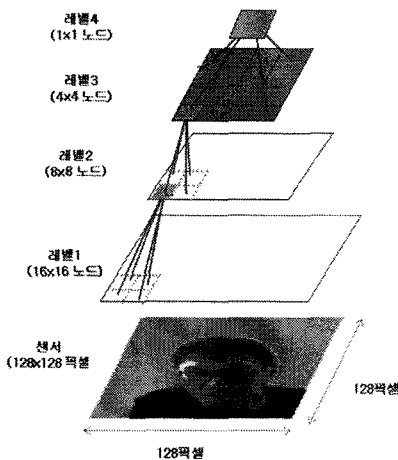


그림 4. HTM 네트워크 구성

으로 결정하게 되는데 여러 파라미터들 중에서 최대 거리와 시그마의 값이 정확도에 가장 큰 영향을 미치므로 이 두 파라미터를 우선적으로 실험하여 결정해야 된다.

최대거리(max distance)는 하나의 기준 센서로부터 원본 이미지와 비교 대상인 입력 이미지의 벡터 사이의 유클리디안(Euclidean) 거리의 최대값이다. 이는 학습하는 동안에 입력된 벡터들이 한 기준점에서 최대 거리 값을 설정하면 입력되는 이미지 센스의 벡터들의 거리를 각각 계산하여 이것과 비교하게 된다. 최대 거리의 값보다 작을 경우 같은 동시발생으로 취급하고, 이 최대거리 보다 클 경우에 새로운 동시발생으로 취급한다. 따라서 최대 거리를 크게 설정해 줄 경우, 속도는 빠르지만 정밀한 비교는 어렵다. 대신 최대 거리를 작게 설정할수록 학습 속도는 느리지만 좀 더 정밀한 비교를 할 수 있다. 얼굴과 같이 뚜렷이 구분이 되는 이미지를 인식하는 응용에서는 비교적 최대거리 값이 커야 높은 정확도를 나타낸다.

시그마(sigma)는 추론 과정에서 필요한 파라미터로서 동시발생 사건의 정규 분포의 표준편차이다. 이는 동시에 발생하는 경우의 수의 범위를 지정한다. 이 값이 작은 것은 동시발생 사건이 밀집되어 있음을 의미하며, 큰 값은 넓게 분포함을 뜻한다. 대개 추론시 잡음이 많은 이미지일 경우는 이 값을 크게 설정하고, 깨끗한 이미지일 경우는 이 값을 작게 설정하면 정확도가 높게 나타난다.

그런데 우리가 추론 과정에서 사용한 이미지 데이터에는 크기변화(Scale), 선 삽입 또는 폐색(Line or Occlusion), 위치이동(Translation) 형태의 잡음들이 포함되어 있다. 따라서 최대의 정확도를 얻기 위해 적당한 값으로 조정을 하는 것이 중요하다. 이들 파라미터를 조정하여 실험한 결과들은 다음 5장에서 보다 상세하게 기술한다.

### 3.4 훈련 데이터와 테스트 데이터

FRESH의 훈련(Training)에 사용한 데이터는 본 연구를 수행한 팀원들을 가족 구성원으로 가정하고 웹 카메라의 이미지로 만들었다. 그 이미지는 128×128 픽셀 크기의 그레이스케일의 jpg 포맷이다. 이 시스템의 인식 정확도를 확인하기 위하여 훈련 데이터를 크기 변화(Scale), 줄 삽입 또는 폐색(Line or Occlusion), 위치 변화(Translation) 등의 왜곡





		이미지 데이터 예
훈련 데이터 (원본)		
테스트 데이터 (왜곡)	줄 삽입 또는 폐색	
	크기 변화	
	위치 변화	

그림 5. 훈련 및 테스트 데이터의 실제 예

(distorted) 효과를 주어서 테스트 데이터로 사용하였다. 그림 5는 이러한 데이터들의 실제 예들이다.

#### 4. 실험 결과 및 분석

이 장에서는 개발된 FRESH의 구현 결과를 보이고, 그것의 여러 가지 실험 결과와 분석에 대하여 설명한다.

##### 4.1 구현 결과

개발된 FRESH의 구현 환경은 크게 3가지로 이루어져 있다. Windows XP를 기본 운영 체제로 구성하였고, HTM 엔진인 NuPIC 1.6.1과 Vitamin D 툴킷 [11]을 이용하였으며, Python 2.5 언어를 사용하였다. 그림 6은 FRESH의 인식 결과 화면이다. 오른쪽

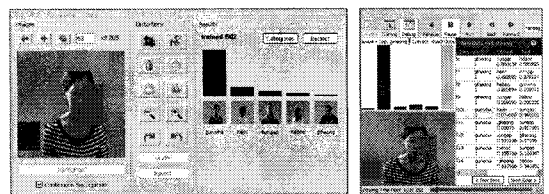


그림 6. FRESH 결과 화면

그림은 입력된 테스트 이미지 데이터를 학습된 이미지의 추론결과를 그래프로 나타내며, 왼쪽 그림은 오류를 검정하는 과정으로 파라미터를 수정한 후 추론 결과를 보이는 화면이다. 추론결과가 막대그래프로 표현되어 인식률을 쉽게 확인할 수 있게 했다.

##### 4.2 최대거리 변화에 따른 실험 결과

최대거리 파라미터를 변경하여 가장 최적의 수치

를 찾는 실험을 수행하였다. 이 실험을 위해 깨끗한 이미지(Clean)는 훈련을 위해 사용한 훈련데이터이며, 나머지 데이터 3가지는 이와 무관하게 크기 변환(Scale), 줄 삽입 또는 폐색(Line or Occlusion), 이동 변환(Translation)을 한 테스트 데이터로서 이들을 사용하여 최대거리 파라미터를 변경시키면서 실험한 결과를 표 1과 그림 7에 나타내었다.

표 1과 그림 7에서 알 수 있듯이 최대거리 값에 따라 해당되는 정확도가 달라지는 것을 확인할 수 있다. 최대거리 값은 평균적으로 50,000일 때 가장 정확도가 높았다.

4.4 시그마 값에 따른 실험 결과

우리는 앞에서 수행한 최대거리에 따른 실험 결과를 바탕으로 정확도가 가장 높은 50,000으로 적용하여 시그마 값을 변화시켜가면서 실험을 수행한 결과를 표 2와 그림 8에 나타내었다. 표 2와 그림 8에서 알 수 있듯이 시그마 값에 따라 해당되는 정확도가 달라지는 것을 확인할 수 있다. 시그마 값이 110일

표 2. 최적 파라미터의 실험 결과

Accuracy Sigma	Clean	Scale	Line	Translation	Avg
70	100.0	94.1	74.5	56.9	81.38
80	100.0	98.0	76.5	60.8	83.83
90	100.0	100.0	70.6	51.0	80.40
100	100.0	100.0	80.4	52.9	83.33
110	100.0	100.0	88.2	62.7	87.72
120	100.0	96.1	80.4	64.7	85.30
130	100.0	98.0	70.6	41.2	77.45

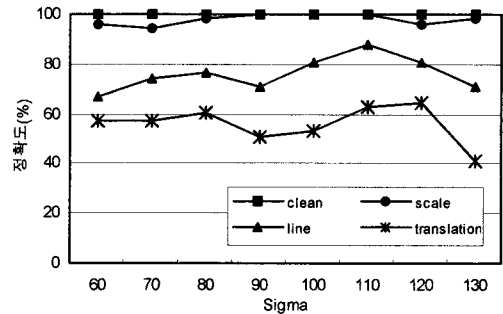


그림 8. 최적 파라미터에 대한 결과

때 가장 정확도가 높았다.

최종적으로 정리하면 FRESH에 대해 가장 정확도가 높은 파라미터는 최대거리의 수치가 50,000인 경우에 전체적으로 정확도가 높았으며, 시그마 값은 110인 경우가 가장 정확도가 높았다. 정확도는 표 2에서 보는바와 같이 깨끗한 이미지(Clean), 크기변화(Scale)에서 100%의 정확도를 보였고, 줄 삽입 또는 폐색(Line or Occlusion)에서는 88.2%의 정확도로 만족할 만한 결과이다. 그러나 위치가 변화된(Translation) 것에서는 62.7%의 정확도로 이에 보완할 여지가 있음을 확인하였다. 이것은 향후 과제로 남긴다.

5. 기존 연구와의 비교

기존의 여러 연구 방식과 우리가 개발한 시스템의 인식률을 비교해 보았다. 기존 연구들 중에 등고선 영역의 투영 벡터를 이용한 연구는 3차원 영상 스캐너를 이용하여 얼굴의 등고선 영역을 추출하고, 이것으로 특징 벡터를 추출하여 얼굴 영역을 검출한다 [4]. 그리고 휴리스틱 평가함수를 이용하는 연구는

표 1. 최대거리 파라미터의 실험 결과

MaxDistance	Clean (%)	Line (%)	Scale (%)	Translation (%)	Avg (%)
30000	100	92.2	84.3	58.8	78.43
50000	100	94.1	76.5	63.7	79.10
70000	100	90.2	64.7	52.9	69.27
90000	100	94.1	68.6	49.0	70.57
110000	100	86.3	70.6	52.9	69.93
130000	100	88.2	72.5	49.0	69.90
150000	100	84.3	64.7	45.1	64.70
170000	100	92.2	64.7	35.3	64.07

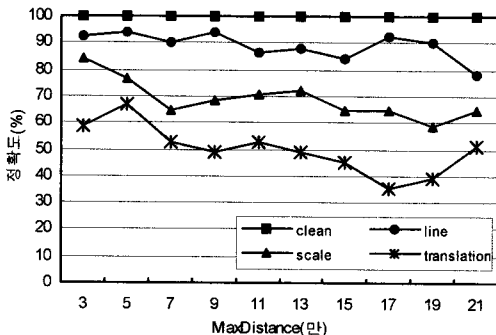


그림 7. 최대거리 결과 그래프

표 3. 기존 연구들의 인식률

연구방식	인식률	특 정
HTM (FRESH)	100%	얼굴인식 성공률(Clean, Scale)
등고선 영역의 투영벡터[4]	94.3%	근접한 5위 내의 인식 성공률
휴리스틱 평가함수[6]	93.8%	눈, 입 인식 성공
GA-BP알고리즘[13]	98.1%	얼굴영역 검출 성공률
YCbCr 색 공간에서 피부색 분할 [14]	96.0%	얼굴영역 검출 성공률
스테레오영상기반 3차원 정보[15]	95.0%	얼굴영역 검출 성공률

얼굴 특성을 반영하는 평가함수를 이용하여 얼굴 영역을 인식하는 방법이다[6]. 또한, GA-BP 알고리즘을 이용하는 연구는 역 전파 오류 알고리즘에 유전자 알고리즘을 혼합하는 방법을 사용하여 얼굴 영역을 인식하며[12,13], 얼굴 영역의 YCbCr 색 공간에서 피부색과 윤곽선 정보를 이용하여 얼굴 영역 인식한다[14]. 그리고 두 개 이상의 카메라 영상(스테레오 영상)의 시차를 이용하여 거리를 측정하고, 그 거리 보정을 통하여 얼굴 영역을 검출한다[15]. 이러한 기존 연구들의 대부분은 얼굴 영역 인식에 초점이 맞추어져 있으며, 다양하게 왜곡된 얼굴 이미지를 인식하여 특정 사람의 이름과 일치시키는 연구는 미미하다. 또한, 기존 연구들에 사용한 이미지는 대부분 깨끗한 영상만을 대상으로 하고 있어 우리의 시스템과 목표가 다소 다르다. 무엇보다 우리는 동일한 이미지를 구하기가 힘들어 우리의 시스템과 정확한 비교는 어려웠다. 하지만 여기서는 논문들에 언급한 자신들의 인식률을 근거해 표 3에 정리하였다.

표 3에서 보는 것과 같이 FRESH의 경우는 왜곡되지 않은 경우(Clean)와 크기 변환(Scale)의 경우 다른 기법들에 비해 높은 100% 인식률을 보였다. 뿐만 아니라 표에는 나타내지 않았지만, 위에서 비교한 다른 연구들에서 시도하지 않은 왜곡들(줄 삽입 또는 폐색, 위치 변환 등)에 대해서도 정확도가 비교적 우수한 것으로 판단된다.

### 6. 결론 및 향후 과제

지금까지 우리는 지능형 흡을 위한 HTM 기반의 얼굴 이미지 인식 시스템(FRESH)의 개발과 실제 데이터를 이용한 얼굴 실험 결과에 대해 기술하였다. 우리가 개발한 시스템을 다양하게 왜곡된 이미지에 적용하여 인식률을 확인한 결과 깨끗한 이미지

(Clean)와 크기변환(Scale)에서는 100% 정확도를 보였으며 줄 삽입 또는 폐색(Line or Occlusion)에서는 88.2%의 정확도로 만족할 만한 결과이다. 이는 가변적인 이미지가 흔히 입력되는 실제 환경에서도 높은 인식률을 보인 것으로 얼굴영역 인식에 초점을 맞춘 타 시스템들 보다 지능형 흡을 위한 시스템에 적합하다. 그리고 이 시스템은 각 가정의 출입문에 설치된 카메라를 사용함으로써 적은 비용으로 설치할 수 있으며, 각 가정의 가족과 방문자의 얼굴을 학습시켜서 인식하게 함으로써 각 가정에 맞는 주문형으로 개발을 할 수 있다. 따라서 이 시스템을 침입자 유무를 판단하고 동시에 사용자에게 방문자의 정보를 알려주는 등의 편의성을 제공하는 지능형 흡을 위한 얼굴 인식 분야에 효과적으로 활용할 수 있을 것이라 생각한다.

향후 더 많은 데이터 수집 및 실험으로 왜곡된 얼굴 이미지들의 인식에 대한 정확도를 높이는 작업이 좀 더 필요하며, 일반 사용자가 보다 쉽게 주문형으로 얼굴 이미지 인식 시스템을 구축해 활용할 수 있도록 하는 사용자 인터페이스의 개발도 필요하다.

또한, HTM 기반의 음성 인식을 얼굴 인식 시스템에 접목하는 연구가 이루어 질 경우 더욱 정확도가 높은 인식 시스템이 될 것으로 기대된다.

### 참 고 문 헌

[1] Y. Wang, H. Ai, B. Wu, C. Huang, "Real time facial expression recognition with adaboost," *17th IEEE International Conference on Pattern Recognition*, Vol.3, pp. 926-929, 2004.

[2] M. S. Bartlett, "Face image analysis by unsupervised learning and redundancy reduction," *PhD thesis, University of California, San*

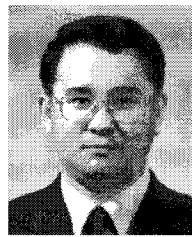


Diego, 1998.

- [3] 송근원, "얼굴인식 기술 및 동향," 한국멀티미디어학회지, 제7권 제2호, pp. 1-8, 2003.6.
- [4] 이영학, 심재창, 이태홍, "등고선 영역의 투영 벡터를 이용한 3차원 얼굴 인식," 한국멀티미디어학회논문지, 제6권 제2호, pp. 230-239, 2003.4.
- [5] 정성욱, 김도형, 안광호, 정명진, "실시간 얼굴표정인식을 위한 새로운 사각특징형태 선택기법," 제어·자동화·시스템공학논문지, 제12권 제2호, pp. 130-137, 2006. 2.
- [6] 장경식, "얼굴의 특성을 반영하는 휴리스틱 평가함수를 이용한 얼굴 특징 검출," 한국정보처리학회논문지, 제8-B권 제2호, pp. 183-184, 2001. 4.
- [7] Jeff Hawkins, *OnIntelligence*, New York, Henry Holt, 2005. <http://www.onintelligence.com>
- [8] Jeff Hawkins, Dileep George, "Hierarchical Temporal Memory," Numenta, Inc., March. 2007. [http://www.numenta.com/Numenta\\_HTM\\_Concepts.pdf](http://www.numenta.com/Numenta_HTM_Concepts.pdf)
- [9] Numenta Inc., "Numenta Platform for Intelligent Computing," June. 2008. [http://www.numenta.com/for-developers/software/pdf/nupic\\_plugin\\_guide.pdf](http://www.numenta.com/for-developers/software/pdf/nupic_plugin_guide.pdf)
- [10] Dileep George and Bobby Jaros, "The HTM Learning Algorithms," Numenta Inc., March. 2007. [http://www.numenta.com/for-developers/education/Numenta\\_HTM\\_Learning\\_Algos.pdf](http://www.numenta.com/for-developers/education/Numenta_HTM_Learning_Algos.pdf)
- [11] Vitamin D, Inc., "Vitamin D Toolkit Reference Guide," June. 2008. <http://www.vitamindinc.com/downloads/Vitamin%20D%20Toolkit%20Reference%20Guide.pdf>
- [12] W. W. Bledsoe, "The model method in facial recognition," Panoramic Research Inc., Palo Alto, CA, Rep. PRI-15, Aug. 1996.
- [13] 전상호, 남궁재찬, "혼합된 GA-BP 알고리즘을 이용한 얼굴인식 연구," 한국정보처리학회논문

지, 제7권 제2호, pp. 552-557, 2000. 2.

- [14] 권혁봉, 권동진, 장언동, 윤영복, 안재형, "YCbCr 색공간에서 피부색과 윤곽선 정보를 이용한 얼굴영역 검출," 한국멀티미디어학회논문지, 제7권 제1호, pp. 27-34, 2004.1.
- [15] 박장한, 백준기, "얼굴인식의 향상을 위한 스테레오 영상기반의 3차원 정보를 이용한 인식," 전자공학회논문지, 제43권 CI편 제3호, pp. 30-38, 2006. 5.



**배 선 갑**

1987년 경북대학교 물리학과 졸업(학사)  
 2004년 경상대학교 대학원 컴퓨터과학과 졸업(석사)  
 2006년~현재 경상대학교 대학원 컴퓨터과학과 박사과정

관심분야 : 패턴인식, 지능시스템, 내장형 데이터베이스



**이 대 한**

2004년 경상대학교 컴퓨터과학과 졸업(학사)  
 2007년~현재 경상대학교 대학원 컴퓨터과학과 재학(석사)  
 관심분야 : XML, 데이터베이스 통합, 기계학습



**조 건 화**

2004년 안산과학대학 컴퓨터과학과 졸업(학사)  
 2006년 경상대학교 컴퓨터과학과 졸업(석사)  
 2006년 8월~현재 경상대학교 컴퓨터과학과 박사과정

관심분야 : Intelligence, Embedded DB, XML, MPEG-7



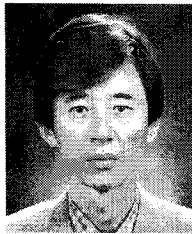
**남 해 보**

2007년 길림북화대학 컴퓨터과  
학과 졸업(학사)  
2007년 9월~현재 경상대학교 컴  
퓨터학과 석사과정  
관심분야 : Intelligence, Image  
Processing, Patten  
Design, JSP



**강 현 석**

1981년 동국대학교 전자계산학  
과 졸업(학사)  
1983년 서울대학교 대학원 전산  
학과 졸업(석사)  
1989년 서울대학교 대학원 전산  
학과 졸업(박사)  
1981년~1984년 2월 한국전자통  
신연구원 연구원  
1984년 3월~1993년 2월 전북대학교 전임강사, 부교수,  
교수  
1993년 3월~현재 경상대학교 컴퓨터학과 교수, 컴  
퓨터정보통신연구소 연구원  
관심분야 : 멀티미디어, 내장형 데이터베이스, 지능시스템



**김 성 진**

1979년 경북대학교 공과대학 전  
자공학과 졸업(공학사)  
1981년 경북대학교 대학원 전자  
공학과 졸업(공학석사)  
1995년 영남대학교 대학원 컴퓨  
터공학전공(공학박사)  
1983년~1985년 삼성전자(주) 시  
스템개발부  
1996년~1997년 University of California, Santa Cruz 연  
구교수  
1985년 3월~현재 연암공업대학 컴퓨터공학과 교수  
관심분야 : 멀티미디어, 병렬처리, 운영체제



**배 종 민**

1980년 서울대학교 수학교육과  
졸업(학사)  
1983년 서울대학교 대학원 계산  
통계학과 졸업(석사)  
1985년 서울대학교 대학원 계산  
통계학과 졸업(박사)  
1982년~1984년 한국전자통신연  
구소 연구원  
1997년~1998년 Virginia Tech. 객원연구원  
1984년~현재 경상대학교 컴퓨터과학부 교수  
관심분야 : XML, 데이터베이스 통합, 기계학습