

## 얼굴 검출과 인식 및 모션추적에 의한 증강현실 구현

이 희 만\*

### Implementing Augmented Reality By Using Face Detection, Recognition And Motion Tracking

Heeman Lee \*

#### 요 약

내추럴 유저 인터페이스 기술(NUI)의 발달은 게임이나 컴퓨터 사용 및 기타 전자기기 사용법에 일대 변화를 가져오고 있다. 본 논문에서는 스마트 폰 사용자의 얼굴을 식별하고 또한 얼굴 움직임 동작을 감지하여 증강현실을 구현한다. 모바일 기기의 전면 카메라를 이용하여 얼굴의 위치를 Viola - Jones 방법을 사용하여 검색하고 검색한 얼굴의 식별 및 움직임 검출을 Eigenface 알고리즘을 사용하며, 식별된 사람에 대응되는 그래픽스 오브젝트를 GPS나 가속도 센서 등의 정보와 후면 카메라의 영상과 함께 합성하는 증강현실 시스템을 구현한다. 증강현실 시스템은 모바일 기기의 성능에 적합한 알고리즘 및 방법으로 구현하여야 한다.

▶ Keyword :증강현실, 고유얼굴, 아다부스트 알고리즘

#### Abstract

Natural User Interface(NUI) technologies introduce new trends in using devices such as computer and any other electronic devices. In this paper, an augmented reality on a mobile device is implemented by using face detection, recognition and motion tracking. The face detection is obtained by using Viola - Jones algorithm from the images of the front camera. The Eigenface algorithm is employed for face recognition and face motion tracking. The augmented reality is implemented by overlapping the rear camera image and GPS, accelerator sensors' data with the 3D graphic object which is correspond with the recognized face. The algorithms and methods are limited by the mobile device specification such as processing ability and main memory capacity.

▶ Keyword :Augmented Reality, Eigen Face, Adaboost Algorithm

---

• 제1저자 : 이희만

• 투고일 : 2011. 11. 14, 심사일 : 2011. 11. 27, 게재확정일 : 2011. 12. 19.

\* 서원대학교 멀티미디어학과(Dept. of Multimedia Engineering, Seowon University)

## I. 서론

본 논문은 스마트폰을 이용한 증강현실 구현에 관한 것이다. 모바일 통신 기기는 단순히 통화를 하던 단계에서 데이터 통신이 가능한 스마트 폰으로 진화되고 있으며, 스마트 폰의 보급율도 점차 확대되고 있다. 스마트 폰의 기능은 3D 그래픽스 뿐만 아니라 카메라 영상 처리 및 GPS 수신 등, 이제 점차 손안의 움직이는 컴퓨터 통신 기기로 변모되고 있다.

증강현실은 현실세계에 디지털 정보를 덧붙여 사용자의 활동을 지원하는 유저 인터페이스 기술이다[1]. 최근 증강현실에 대한 학계의 관심도 매우 높는데 최근의 연구 논문들로는 다음과 같은 내용이 있다. 한글의 색상 인식과 자소 패턴을 분리하여 실세계 영상에 추가적인 이미지 및 영상을 출력할 수 있는 방법[2], 복수개의 카메라를 이용하여 객체를 추적하고 인식하여 추적된 객체에 대해 검색인식가능한 영역에 대한 정보를 증강현실로 나타내는 시스템[3], 복잡한 장면이나 또는 손으로 가려지는 영상에서도 마커를 인지하고 위치를 추정하여 실세계 영상에 3차원 객체를 표현 할 수 있는 방법 [4], 마커 기반의 뮤직 마커와 뮤직 보드로 구성하여 뮤직 보드상에 배치된 뮤직 마커의 내용을 악기의 종류와 미리 정의된 재생 트랙과 매핑하고 2차원 평면상에서 상대적인 뮤직 마커의 위치 값을 측정하여 이를 미디 트랙에서 공간 효과와 관련된 파라미터 값으로 설정하여 실시간으로 음악을 편집하고 이를 재생할 수 있는 시스템[5], 지식기반의 증강현실 시스템 구축을 위한 프레임워크 개발[6] 등 증강현실 관련 연구가 현재 활발히 진행 중이다.

본 논문은 스마트 폰 사용자의 얼굴과 얼굴 움직임 동작을 감지하여 증강현실에 적용하는데 목적이 있다. 동작인식 기술의 발달은 손을 안대고 컴퓨터를 사용하거나 전화기를 받는 등 사용자의 기기 사용법에 일대 변화를 가져오고 있다. 기기 조작방법을 별도로 익힐 필요가 없는 '내추럴 유저 인터페이스' 기술을 활용한 제품도 활발히 연구되어 곧 출시될 전망이다. 사람의 자연스런 동작을 활용한 내추럴 유저 인터페이스는 이미 닌텐도 Wii, 마이크로소프트(MS) 엑스(X)박스 키넥트 등의 게임이나 카메라의 얼굴 인식과 웃음 감지 촬영 및 검색엔진의 음성 인식 기술 등으로 이미 부분적으로 활용되고 있는 상태이다. 그래픽 유저 인터페이스(GUI)가 컴퓨터를 대중화시키고 멀티터치가 태블릿 피서를 확산시킨 것처럼 내추럴 유저 인터페이스(NUI)는 새롭게 주목받는 기술이 될 것이다.

본 논문에서는 스마트 폰 사용자의 얼굴을 식별하고 또한

얼굴 움직임 동작을 감지하여 증강현실을 구현한다. 본문의 구성은 다음과 같다. 2장에서는 본 논문에서 구현한 증강현실 시스템을 설명하고 각 모듈에서 사용된 기술에 대하여 논의하며, 3장에서는 아이폰4와 아이패드2를 사용한 구현시스템의 실험결과를 논한다.

## II. 증강현실시스템

증강현실이란 실제환경에 컴퓨터로 구현한 가상의 사물이나 정보를 중첩하여 사람이 감지할 수 있는 정보 이상의 증강된 환경을 구현하는 것이다. 증강현실은 가상현실의 한 분야로 가상의 공간과 사물만이 아닌 현실세계의 기반위에 부가적인 정보를 보강하여 제공할 수 있는 기술이다.

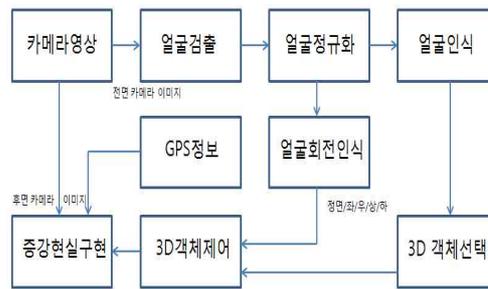


그림 1. 증강현실 시스템 구성도  
Fig 1. Augmented Reality System Diagram

본 논문에서의 제안한 증강현실 시스템은 <그림 1>과 같다. 증강현실 시스템은 모바일 전면 카메라의 실시간 영상으로 부터 얼굴을 검출하고 얼굴 인식을 한다. 얼굴 인식을 하는 이유는 사용자에 따라 각기 다른 캐릭터 사용 및 개인 맞춤 정보를 표현하기 위함이다. 또한 모바일 기기의 GPS정보와 기타 스마트폰의 다양한 센서 정보와 함께 증강현실 화면 표출을 위하여 3D 그래픽스 객체를 모바일 후방 카메라 영상과 함께 합성한다. 현재의 위치는 GPS정보로 부터 받고 동서남북의 방향 정보는 스마트 폰의 지자기 센서로 부터 받을 수 있다. 위치기반의 정보는 데이터베이스 또는 서버에 저장된 정보에서 가져올 수 있으며 사용자 얼굴의 방향을 인식하여 상호 작용이 가능한 정보를 표출하는 것이다.

### 1. 얼굴검출 모듈

사람의 얼굴을 인식하거나 얼굴의 동작을 인식하기 위해서는 얼굴의 검출을 선행하여야 한다. 얼굴의 검출에 대한 많은 연구 결과가 문헌상으로 발표되었다.

규격기반 얼굴의 검출 방법은 사람 얼굴의 구성요소인 눈,

코 및 입의 위치에 규칙이 있음을 이용한 방법이다. 미리 작성된 규칙에 따라 대상 이미지에서 얼굴 후보 이미지를 검출한다[7]. 이러한 지식 기반의 얼굴 검출 방법은 여러명이 등장하는 복잡한 배경을 가진 이미지 등에서는 오류가 높은 문제점이 있다.

특징기반 얼굴 검출 기법은 얼굴의 피부색이나 텍스처 등의 정보를 활용하는 기법이다[8]. 특징이란 얼굴 성분들의 크기와 모양, 상호 연관성, 얼굴의 색상과 질감(texture) 혹은 이러한 요소들의 혼합된 형태의 정보를 말한다. 이 방법은 검출처리 시간이 빠르다는 장점을 가지고 있다. 하지만, 조명 및 밝기 변화에 따라 피부색과 유사한 배경에서 오인식 할 수 있고 얼굴의 색상 및 질감 정보가 다르거나 얼굴의 기울어진 정도에 따라 얼굴의 특징 성분들을 검출하지 못할 수 있다. 기울어진 얼굴의 검출을 위해 윈도우 내의 에지 포인트들로부터 대칭축을 추출하여 검출된 대칭축 방향에 대해서만 얼굴 검출을 적용하는 보완방법도 제안 되었다[9].

템플릿 방식은 얼굴 영역의 표준 이미지를 이용하여 영상 전영역에서 일치하는 부분을 찾아내는 방법으로 특정 함수에 의해 계산된 얼굴에 대한 표준 패턴 정보를 만들고 입력된 얼굴 이미지와의 상관 정도를 측정하여 검출을 하는 방식이다. 이 방식은 비교적 구현이 쉽다는 장점이 있으나 다양한 환경이나 얼굴 이미지의 크기, 모양, 포즈 등의 변화에 효율적으로 대처하지 못하는 단점이 있어 다양한 가변 템플릿을 사용한다. 템플릿은 얼굴 인식 뿐만아니라 물체의 위치 추적등에 유용하게 활용되고 있다[10].

외형 기반 얼굴 검출 방법은 복수개의 얼굴 이미지에서 학습한 얼굴 이미지의 특징을 이용하여 얼굴을 검출하는 기법으로 AdaBoost[11], Eigenfaces[12], LDA[13], 뉴럴네트워크[14], hidden Markov[15]방법 등이 있다.

Haar-like feature와 AdaBoost 학습 알고리즘은 비교적 성공적인 얼굴패턴 검출알고리즘으로 Haar-like feature는 특징 값을 추출하는 방법이며 AdaBoost 학습 알고리즘은 얼굴패턴 인식률을 높이는 방법이다[16]. Haar-like features<그림 2>는 Haar wavelets과 외형적으로 유사하여 붙인 이름이다. Papageorgiou[17]는 이미지의 휘도 정보만을 이용하는 것보다는 Haar 웨이블릿을 특징기반으로 사용하는 방법이 효과적임을 연구하였다. Viola and Jones는 Haar 웨이블릿 개념에서 아이디어를 도입하여 Haar-like feature를 사용한다.

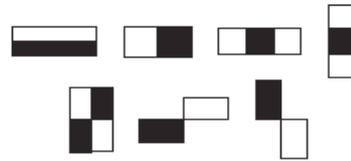


그림 2 Haar-like feature  
Fig 2 Haar-like feature

Haar-like feature는 탐색 윈도우 안의 이웃하는 직사각형 영역을 설정하여 각각 이미지 휘도의 합을 구하여 서로 값의 차이를 구한 것이다. 이 차이는 이미지 영역을 구분하는데 사용된다. 예를 들어 사람의 얼굴 영상에서 통상적으로 볼의 영역이 눈의 영역보다 더 밝다. 그러므로 이 영역이 얼굴 부분으로 잠정 간주될 수 있다. Viola - Jones는 입력 영상에서 얼굴 크기의 윈도우를 스캔하여 Haar-like feature를 계산하고 얼굴이 아닌 다른 이미지일 때의 값과 비교하여 얻은 학습 임계치로 판단한다. Haar-like feature는 계산속도가 빠르다는 장점이 있다. 또한 누적 영상값을 사전 계산하여 사용함으로써 어떤 윈도우 크기라도 같은 시간에 특징 값을 계산할 수 있다.

Boosting 알고리즘은 교사학습 머신러닝 방법으로 약-학습기의 결합으로 강-학습기를 만들어 내는 기법이다. 약-학습기라함은 정확한 분류를 못하지만 랜덤한 것 보다 확률이 높게(일반적으로 최소 50% 이상의 확률) 분류할 수 있는 분류자를 말하며 약분류자의 결합에 의해 높은 확률의 강한 분류자를 만들어 내는 것이다. Schapire[18]의 논문은 boosting 알고리즘에 대한 연구에 크게 기여하였다. 부스팅 알고리즘을 실전에 응용될 수 있도록 개량한 것이 Adaboost(Adaptive Boosting) 알고리즘이다[19]. 그 외에도 LPBoost, TotalBoost, BrownBoost, MadaBoost, LogitBoost 등의 알고리즘 등이 개발되었다. Adaboost 알고리즘은 약-분류기를 반복해서 호출하면서 정확히 분류한 분류기는 가중치를 줄이고 잘못 분류한 분류기는 가중치를 늘려 주어 점차 강한 분류기를 만들어 낸다. Adaboost 알고리즘은 노이즈에 민감한 반응을 하지만 비교적 성공적인 분류기로 사용된다. 알고리즘은 다음과 같다[19].

[단계1]  $m$  개의 학습데이터를 준비한다.

$$(x_1, y_1), \dots, (x_m, y_m) \quad x_i \in X, \quad y_i \in Y = \{-1, 1\}$$

가중치 초기화

$$D_1(i) = \frac{1}{m}, \quad i = 1, \dots, m \dots\dots\dots (식 1)$$

[단계2] 다음을  $t = 1 \dots T$ 회 반복한다.

여러개의 약-학습기 중에서 가중치  $D_t$ 를 사용하여 분류하였을 때 가중치 에러가 최소값이 되는 약한 분류기  $h_t$ 를 찾아낸다.

$$h_t = \operatorname{argmin} \epsilon_t \quad \dots\dots\dots (\text{식 } 2)$$

$$\epsilon_t = \sum_{i=1}^m D_t(i)[y_i \neq h_t(x_i)]$$

1. 계산값  $\epsilon_t \geq 0.5$  이면 정지한다.

2.  $\alpha_t \in R$ 를 계산한다.

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \epsilon_t}{\epsilon_t} \quad \dots\dots\dots (\text{식 } 3)$$

3. 가중치를 변경한다.

$$\frac{D_{t+1}(i)}{Z_t} = D_t(i) \exp(-\alpha_t y_i h_t(x_i)) \quad \dots\dots\dots (\text{식 } 4)$$

$Z_t$ 는  $D_t$ 가 확률 분포가 되도록 정규화를 위한 값이다. 가중치 계산은 다음과 같이 한다.

$$y_i h_t(x_i) = \begin{cases} +1 & y(i) = h_t(x_i) \\ -1 & y(i) \neq h_t(x_i) \end{cases} \quad \dots\dots\dots (\text{식 } 5)$$

[단계3] 학습한 분류기의 출력은 다음과 같다.

$$H(x) = \operatorname{sign} \left( \sum_{t=1}^T \alpha_t h_t(x) \right) \quad \dots\dots\dots (\text{식 } 6)$$

Viola - Jones는 Haar-like feature와 AdaBoost알고리즘을 이용하여 비교적 정확도 높은 얼굴 패턴을 찾아 낼 수 있었다. 얼굴 크기의 윈도우를 입력영상 위에 이동하면서 Haar-like feature를 계산한다. 각 계산값은 학습한 임계값과 비교하여 얼굴이 아닌 부분을 분류한다. Haar-like feature는 약-분류기이므로 많은 수의 Haar-like feature가 필요하다. Viola - Jones는 <그림 3>과 같이 단계적(Cascade) 분류기를 사용하여 검출속도의 향상과 더불어 강-분류기를 만든다. 얼굴로 판단되는 데이터는 계속 진행하지만 배경이나 얼굴이 아닌 부분에 대하여는 초기 단계부터 배제시킴으로서 빠르게 연산을 할 수 있다[20].

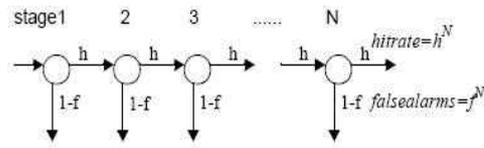


그림 3. 단계적 분류  
Fig 3. Cascade Classifier

## 2. 얼굴인식 모듈

Eigenfaces 알고리즘은 사람의 얼굴을 인식하는 알고리즘으로 PCA (principal component analysis)라고 하는 수학적인 절차를 여러 사람의 얼굴 이미지에 적용하여 기존 데이터 베이스에서 가장 유사도가 높은 얼굴을 찾을 수 있는 알고리즘으로 Sirovich and Kirby[20]에 의해 제안된 성공적인 얼굴인식 알고리즘이다. Eigenfaces는 평균화된 얼굴요소 성분이라고 생각할 수 있으며 많은 얼굴 사진으로 부터 통계적으로 분석하여 얻은 데이터이다. 대부분 인간의 얼굴은 이러한 표준적인 얼굴로 부터 합성할 수 있다는 이론이다. Eigenfaces는 필기체 인식이나 입술읽기, 음성인식 및 손동작 인식 등 다양한 분야에 이용된다. 그러므로 Eigenfaces보다 EigenImages라고 부르는 경우가 많다. 알고리즘은 다음과 같다.

[단계1] 학습단계

1.  $M$ 개의 얼굴 이미지( $n \times n$ )를 각각 크기  $N = n \times n$ 의 벡터  $\Gamma_i$ 로 변환한다. 집합  $S$ 는  $M$ 개의 벡터  $\Gamma_i$ 로 구성된다.

$$S = \{\Gamma_1, \Gamma_2, \dots, \Gamma_M\} \quad \dots\dots\dots (\text{식 } 7)$$

2. 평균이미지를 구한다.

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad \dots\dots\dots (\text{식 } 8)$$

3. 평균 이미지와 각각의 입력이미지와의 차이를 구한다.

$$\Phi_i = \Gamma_i - \Psi \quad \dots\dots\dots (\text{식 } 9)$$

$$A = \{\Phi_1, \Phi_2, \dots, \Phi_M\}$$

4. Covariance 매트릭스를 구한다.

$$C = AA^T \quad (n^2 \times n^2) \quad \dots\dots\dots (\text{식 } 10)$$

그러나 이 매트릭스는 실제 사용할 때 크기가 커서 많은 컴퓨팅 파워를 요구한다. 그러므로 크기가 작은 다음 매트릭스를 계산한다.

$$L = A^T A \quad (m \times m) \dots\dots\dots \text{(식 11)}$$

5.  $L$  매트릭스의 Eigen Vector를  $v_i$ 라하면 다음과 같다.

$$Lv_i = \lambda_i v_i$$

$$ALv_i = \lambda_i Av_i$$

$$AA^T Av_i = \lambda_i Av_i \dots\dots\dots \text{(식 12)}$$

$$CAv_i = \lambda_i Av_i$$

6.  $u_i = Av_i$ 라고 정의하면  $u_i$ 는 Eigen Vector(Eigen Face)가 된다. Eigen Face는 face space의 기저가 된다.

$$Cu_i = \lambda_i u_i \dots\dots\dots \text{(식 13)}$$

7. 유효한  $K$ 개의 Eigen Face를 선택한다. 선택은 경험적으로 결정한다. 여기에서  $K < M$ 이다.

[단계2] 인식단계

얼굴 인식은 인식할 대상의 이미지를 Eigen Face의 기저에 프로젝션 하였을 때 가장 유사도가 큰 기저를 해당 얼굴로 인식하는 원리이다.  $\Phi_i$ 는 여러 Eigen Face의 선형조합이 된다.

$$\Phi_i = \sum_{j=1}^K w_j u_j \dots\dots\dots \text{(식 14)}$$

가중치는  $w_j = u_j^T \Phi_i$ 로 계산한다.

모든 훈련용 입력이미지 각각에 대하여 다음을 구한다.

$$\Omega_i = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{bmatrix} \quad i = 1..M \dots\dots\dots \text{(식 15)}$$

인식이 필요한 입력 이미지  $\Gamma$ 에 대하여 평균 이미지와의 차이를 구한다.

$$\Phi = \Gamma - \Psi \dots\dots\dots \text{식 16}$$

$\Phi$ 를 Eigen Face의 기저로 프로젝션을 한다. 이때 가중치 벡터를 구한다.

$$w_i = u_i^T \Phi \dots\dots\dots \text{(식 17)}$$

$$\Omega = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{bmatrix}$$

모든  $M$ 개의 Eigen Face와의 에러를 계산하여 거리가 가장 최소이면서 에러값이 특정값 이하인 경우 해당 이미지와 같은 얼굴이라고 간주한다.

$$\epsilon_r = \min \|\Omega - \Omega_i\| \dots\dots\dots \text{(식 18)}$$

$$\epsilon_r < \theta$$

### 3. 얼굴 회전인식 모듈

사용자의 현재 위치는 스마트폰의 GPS 센서에서 알아내며 스마트폰의 방위방향은 장비의 지자기센서를 이용한다. 그러나 사용자가 스마트폰을 바라보면서 어느 위치를 바라보는 지는 알 수가 없으므로 스마트폰의 지자기센서 정보와 더불어 사용자의 얼굴 회전 정보를 알아내는 것이다. 얼굴 회전 인식은 얼굴인식에 사용한 Eigenfaces 알고리즘을 사용한다. 얼굴의 회전은 정면, 상, 하, 좌 및 우측 방향 만을 인식하는데 스마트폰 화면을 바라보면서 얼굴회전이 크면 화면을 볼 수 없으므로 실제 회전각도는 크지 않기 때문이다. 중간 부분의 방향은 3D 캐릭터가 한 위치에서 현재 인식방향으로 부드럽게 이동하도록 함으로서 해결 가능하다. 얼굴회전인식은 3D캐릭터가 사용자와 상호 인터랙티브하게 반응하면서 실시간으로 정보를 표출할 수 있도록 도움을 준다.

## III. 실험결과

본 시스템의 구현 및 테스트는 아이폰4 및 아이패드2에서 실시하였다. 아이폰4는 1GHZ Cortex-A8 Apple A4 프로세서와 512MB 메인 메모리를 탑재하고 있으며, 아이패드2는 1GHZ 듀얼코어 Apple A5를 사용하며 512MB 메인 메모리를 탑재하고 있다. 두기종 모두 GPS는 물론 가속도 센서 및 전후방 카메라를 갖고 있다. 소프트웨어는 iOS5.0 운영체제와 개발도구 xcode 4.3을 사용하였으며 이미지 라이브러리 OpenCV 2.2를 사용하며 3D 그래픽스 디스플레이를 위해 OpenGL ES 2를 사용하였다.

본 실험에서는 <그림 4>와 같이 4명의 얼굴을 Viola-Jones방법에 의해 정면, 좌측, 우측, 위 및 아래의 시선방향 얼굴을 각각 캡처를 하고 50×50 픽셀 크기로 정규화하여

Eigenface의 훈련 이미지로 저장한다. 저장된 이미지로 부터 평균이미지 및 Eigenface를 추출하여 학습 데이터와 함께 데이터베이스에 저장한다. <그림 5>는 훈련 이미지에 적용된 사람이 전면 카메라 앞에서 얼굴을 상하좌우로 움직일때 Eigenface의 인식 알고리즘을 사용하여 훈련 이미지에서 가장 일치하는 얼굴을 찾아내고 해당 얼굴에 해당하는 3D 그래픽스 오브젝트를 OpenGL ES 2로 렌더링한 화면이다. 전면 카메라에 보이는 사람의 얼굴 움직임을 계속 트래킹을 하며 얼굴 움직임에 연동하여 후면 카메라의 영상과 3D 그래픽스 오브젝트 및 기타 다른 정보와 함께 합성한 증강현실 영상을 보여준다. <그림 5> (a)는 얼굴방향에 대한 인식결과를 각 사진 우측 하단에 작은 글자로 결과를 보여주고 있다. 사진 이미지 안의 C는 정면이미지, L은 좌측을 응시하는 이미지, R은 우측을 응시하는 이미지, U는 위방향 응시하는 이미지, 마지막으로 D는 아래방향을 응시하는 이미지로 인식한 결과이다. <그림 5> (b)는 합성한 증강현실 영상이다.



(a)정면 (b)좌측면 (c)우측면 (d)위 (e)아래  
그림 4. Eigenface 훈련용 이미지  
Fig 4. Eigenface Training Images



(a)인식 (b) 증강현실 표현  
그림 5. 증강현실구현을 위한 얼굴 방향 인식  
Fig 5. Face Orientation Recognition for Augmented Reality

<표 1>는 훈련에 참여한 본인의 얼굴을 이용하여 각 방향의 얼굴 분류의 정확도와 훈련에 사용하지 않은 전혀 다른 사람의 얼굴을 사용하여 얼굴회전 인식의 정확도를 각각 분석하여 본 것이다. 본인의 경우 97.5%의 높은 정확도를 유지하였으나 본인 이미지가 아닌 타인 이미지로 훈련한 이미지를 사용하여 분류한 경우에는 많은 편차를 보인다. 실제 응용프로그램에서 얼굴 방향 검출을 위해 본인의 얼굴을 사전에 캡처하여 사용하는 방법과 또는 다양한 훈련 이미지를 확보하고 훈련 이미지와 제일 유사한 사람의 학습 데이터를 사용하는 방법을 사용하면 좋은 결과를 얻을 수 있을 것이다.

<표 2>는 구현한 증강현실 시스템의 각 모듈별 평균 소요 시간을 측정한 것이다. 아이폰4의 경우 카메라 전환시간을 제외한 처리시간은 합계 307.6ms이며 아이패드2의 경우는 188.5ms이다. 그러나 두기종 모두 전면/후면 카메라 전환시간이 매우 많이 소요되어 iOS에서 이를 지원해 주지 않는 한 두 개의 카메라를 동시에 사용하기에는 아직 무리가 있다고 사료된다. 모바일 기기는 퍼스널 컴퓨터 보다 프로세서의 처리 능력이 떨어지고 메인 메모리도 작아 연산량이 많거나 메모리를 많이 소요하는 알고리즘을 사용할 수가 없다. 각 모바일 기기 수준에 적합한 알고리즘을 사용하여야 하며 기기의 사양이 올라가면 그 사양에 적합한 다른 알고리즘 및 방법을 적용해야 할 것이다. 모바일 기기는 성능이 계속 향상되고 있으므로 점점 더 정교한 알고리즘의 적용이 가능할 것이다.

표 1. 이미지 분류 실험  
Table 1. Image Classification Experiments

방향	본인이미지		타인이미지 A		타인이미지 B	
	정분류	오분류	정분류	오분류	정분류	오분류
정면	98.7	1.3	82.5	17.5	10.5	89.5
좌측	97.5	2.5	88.7	11.3	12.3	87.7
우측	99.2	0.8	86.2	13.8	13.6	86.4
상	95.3	4.7	50.6	49.4	20.7	79.3
하	96.6	3.4	59.7	40.3	25.8	74.2
평균%	97.5	2.5	73.6	6.4	16.6	83.4

표 2. 각 모듈의 소요시간(ms)  
Table 2. Elapsed Time of Each Module(ms)

사용기종	전면카메라 영상획득	후방카메라 영상획득	얼굴 검출	얼굴 인식	카메라 전환
아이폰4	66.2	78.3	158.4	4.7	1497.1
아이패드2	65.7	65.5	55.9	1.4	1441.0

#### IV. 결론

동작인식 기술의 발달은 게임이나 컴퓨터 사용 및 기타 전자기기 사용법에 일대 변화를 가져오고 있다. 본 논문은 스마트폰 사용자의 얼굴과 얼굴 움직임 동작을 감지하여 증강현실에 적용하였다. 모바일 기기의 전방 카메라를 이용하여 얼굴의 위치를 검색하고 또한 검색한 얼굴을 인식하며 인식한 사람에 대응되는 그래픽스 오브젝트를 제어한다. 그래픽스 오브젝트는 얼굴의 움직임을 인식하여 다른 추가정보 예를 들면 GPS나 가속도 센서 등과 함께 후면 카메라의 영상과 함께 합성된 증강현실 이미지를 보여 줄 수 있다. 향후 연구 과제로는 모바일 기기의 진화와 함께 각각의 모바일 기기의 성능에 적합한 알고리즘 및 시스템을 개발하여 보다 더 정교하고 안정된 증강현실 시스템을 구현하는 것이다.

#### 참고문헌

[1] Mantro Press, "Smart Phone And Web Revolution," pp.19-37, ISBN:978-89-6305-064-5, Sep., 2010.

[2] Seong-Yoon Shin, Byung-Seok Choi, Yang-Won-Rhee, "Color Recognition and Phoneme Pattern Segmentation of Hangeul Using Augmented Reality," Journal of The Korea Society of Computer and Information, Vol.15, No.6, pp.29-35, June, 2010.

[3] Hag-hee Kim, "Implementation of augmented reality and object tracking using multiple camera," Journal of The Korea Society of Computer and Information, Vol.16, No.6, pp.89-97, June, 2011.

[4] Hag-hee Kim, "Techniques on Multi-Marker for the Implementation of Augmented Reality," Journal of The Korea Society of Computer and Information, Vol.15, No.11, pp.109-116, Nov., 2010.

[5] Eun-Young Kim, Dong-Yeol Oh, "A Realtime Music Editing and Playback System in An Augmented Reality Environments," Journal of The Korea Society of Computer and Information, Vol.16, No.6, pp.79-88, June, 2011.

[6] Chong Woo Woo, Doohee Lee, "A Development of a Framework for Building Knowledge based Augmented Reality System," Journal of The Korea Society of Computer and Information, Vol.16, No.7,

49-58, July, 2011.

[7] Constantine Kotropoulos, Ioannis Pitas, "Rule-Based Face Detection in Frontal Views," Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 97), Vol. IV, 1997.

[8] Kyung-Shik, Jang, "Facial Features Detection Using Heuristic Cost Function," The KIPS transactions, Part B, Vol.8B, No.2, pp.183-188, 2001.

[9] Bo Whan Won, Ja Young Koo, "Rotated Face Detection Using Symmetry Detection," Journal of The Korea Society of Computer and Information, Vol.16, No.1, pp.53-59, Jan., 2011.

[10] Chulhan Lee, minYi Jung, Jongsun Kim, "A Method of Generating Changeable Face Template for Statistical Appearance-Based Face Recognition," Journal of the Institute of Electronics Engineers of Korea, SP, Vol.44, No.314, pp.27-36, 2007.

[11] Viola, P, Jones, M, "Rapid object detection using a boosted cascade of simple features," IEEE Computer Vision and Pattern Recognition, Vol.1, pp.511-518, Apr. 2001.

[12] A. R. Mohan and N. Sudha, "Fast face detection using boosted eigenfaces," IEEE Symposium on Industrial Electronics & Applications, Vol.2, pp.1002-1006, 2009.

[13] Hyun-Chul Kim, Daijin Kim, "Face Recognition using LDA Mixture Model," Journal of KIISE, Vol.32, No.8, 803-833, Aug., 2005.

[14] C. C. Tsai et al, "Face Detection Using Eigenface and Neural Network," IEEE Intl. Conf. on Systems, Man and Cybernetics, Vol.5, pp.4343-4347, 2006.

[15] Kung-Ah Lee, Dae-Jong Lee, "Face Recognition Using Wavelet Coefficients and Hidden Markov Model," Journal of Fuzzy Logic and Intelligent Systems, Vol.13, No.6, pp.673-678, Dec., 2003.

[16] Viola and Jones, "Rapid object detection using boosted cascade of simple features," Computer Vision and Pattern Recognition, pp.1-9, 2001.

[17] Papageorgiou, Oren and Poggio, "A general framework for object detection," International Conference on Computer Vision, pp.555-562, 1998.

[18] Rob Schapire, "Strength of Weak Learnability,"

Machine Learning, vol.5, pages 197-227. 1990.

- [19] Yoav Freund, Robert E. Schapire, "A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting," Vol.904, pp.23-37, 1995.
- [20] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," Journal of the Optical Society of America A 4: 519-524, 1987.

## 저 자 소 개



### 이 희 만

1984 : 고려대학교 전자공학과 공학사

1986 : 한국 과학기술원 전기 및 전자  
공학과 공학석사

1994년 : Texas A&M Electrical  
Eng. Ph.D

현 재 : 서원대학교 멀티미디어학과 교수

관심분야 : 가상현실 증강현실

Email : hlee@seowon.ac.kr