

# 누적 히스토그램과 랜덤 포레스트를 이용한 머리방향 추정

(Head Pose Estimation with Accumulated Histogram and Random Forest)

문성희\*, 이철우\*\*

(Sung Hee Mun, Chil woo Lee)

## 요약

스마트 환경 구축이 보편화됨에 따라 사람과 컴퓨터 사이의 상호작용(HCI)에 관한 연구가 활발히 진행되고 있다. 인간-컴퓨터 상호작용에서 사람의 얼굴과 시선 방향을 안다는 것은 그 사람의 의도나 관심의 대상을 파악하는데 중요한 정보를 제공할 뿐만 아니라 신체 구조를 이해하는데도 하나의 기준이 될 수 있으므로 중요한 연구 테마이다. 본 논문에서는 랜덤 포레스트를 이용하여 얼굴 방향을 미리 정해놓은 각도로 분류하는 방법을 제안한다. 먼저 영상은 전처리를 거친 뒤 회전 정보를 얻기 위하여 평균 정면 얼굴과의 차영상을 이용하여 회전정보를 추출한다. 캐니에지 검출법을 이용하여 얼굴의 특징을 검출하고 이를 이용하여 에지 영상을 구한 뒤, 이 영상에 대해 가로 세로축 각각에 대해 픽셀 수를 누적하여 히스토그램을 작성한다. 누적히스토그램을 특징으로 랜덤 포레스트를 생성하였으며, 랜덤 포레스트의 학습과 테스트에는 CAS-PEAL-R1 데이터를 사용하여 80.6%의 인식률을 얻었다.

■ 중심어 : 머리방향 ; 누적히스토그램 ; 랜덤포레스트

## Abstract

As smart environment is spread out in our living environments, the needs of an approach related to Human Computer Interaction(HCI) is increases. One of them is head pose estimation. it related to gaze direction estimation, since head has a close relationship to eyes by the body structure. It's a key factor in identifying person's intention or the target of interest, hence it is an essential research in HCI. In this paper, we propose an approach for head pose estimation with pre-defined several directions by random forest classifier. We use canny edge detector to extract feature of the different facial image which is obtained between input image and averaged frontal facial image for extraction of rotation information of input image. From that, we obtain the binary edge image, and make two accumulated histograms which are obtained by counting the number of pixel which has non-zero value along each of the axes. This two accumulated histograms are used to feature of the facial image. We use CAS-PEAL-R1 Dataset for training and testing to random forest classifier, and obtained 80.6% accuracy.

■ keywords : Head Pose ; Accumulated Histogram ; Random Forest

## I. 서론

인간과 컴퓨터의 상호작용에 있어서 사용자의 머리방향은 많은 정보를 제공해 준다. 특히 머리방향과 관련된 시선방향은 사용자의 ID 판별 뿐만 아니라 그 사람의 의도, 대화내용의 추정, 관심대상의 구별 등에 꼭 필요한 정보를 담고 있다. 따라서 대부분의 지능형 인터페이스의 구현에 있어서 머리방향인식은 중요한 전제 조건의 하나로 인식되고 있다.

머리방향 추정을 위한 연구는 크게 지역적 접근법과 전역적 접근법으로 나누어진다. 지역적 접근법은 대부분 그림1에서 보이는 눈, 코, 입과 같은 얼굴 요소를 이용하여 머리방향을 추정

한다. 이 방법은 얼굴 기관의 위치는 표정이 변하더라도 대체적으로 큰 차이를 보이지 않는다는 가정에서 출발한다. 이렇듯 자유도가 낮은 얼굴 요소의 구조적 특징 정보를 이용하여 머리의 방향성을 추정하는 방법이 지역적 접근법의 대표적인 예이다.

Geel[3]는 얼굴 내의 눈과 입의 위치를 이용해 머리방향을 추정할 수 있는 방법을 제안하고 있다. Horprasert[4]는 눈의 양 끝점과 코의 위치를 이용하여 특징 벡터를 정의함으로써 보다 정밀하게 머리방향을 추정하였다. Wang[5] 또한 마찬가지로 눈의 양 끝점과 입의 양 끝점의 위치를 바탕으로 머리방향을 추정하였다. Fadda[6]는 눈과 코의 위치를 기반으로 레오나르도 다빈치의 인체 황금비례를 응용하여 머리방향을 추정하였다.

얼굴 요소 외에 지역적 특징 사이의 관계를 이용하기도 한다.

\* 학생회원, 전남대학교 전자컴퓨터공학과

\*\* (교신저자)정회원, 전남대학교 전자컴퓨터공학과

본 연구는 교육과학기술부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임

접수일자 : 2015년 11월 25일

수정일자 : 2016년 03월 27일

게재확정일 : 2016년 03월 29일

교신저자 : 이철우 e-mail : leecw@jnu.ac.kr

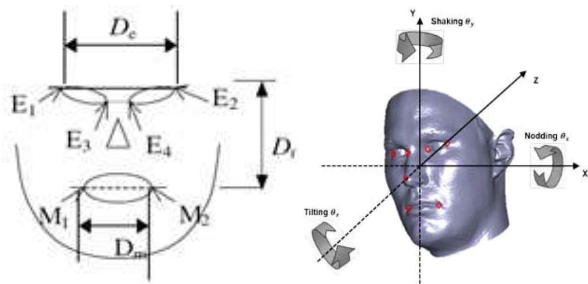


그림 1. 얼굴방향 추정에 사용되는 얼굴의 특징과 3차원 얼굴 모델의 예

Maurer[7]은 얼굴 요소의 구조적 정보를 사용하지 않고 영상에서 Gabor-based Jet 을 이용하여 추출한 특징 벡터를 기반으로 머리방향을 추정하였다.

이들과는 달리 3D 얼굴 모델을 구성하고 특징점과 모델과의 대응관계를 이용하여 머리방향을 구하기도 한다. Kong[8]은 3D 얼굴 모델을 구성하고 눈의 양 끝점과 코끝, 입의 양 끝점과 모델의 대응관계를 이용하여 머리방향을 추정하였다.

지역적 접근법들은 얼굴의 특징 벡터가 정확히 검출되어야 한다는 전제를 바탕으로 방향 추정이 이루어진다. 그러나 얼굴의 개인차나 조명에 따른 얼굴의 음영 변화 등 여러 요인들로 인해 정확한 특징 벡터 검출에 어려움이 있으며, 검출된 특징 벡터에 따라 결과가 민감하게 변화한다는 문제점이 있다.

전역적 접근법은 머리방향 추정에 얼굴 영상 전체를 이용하는 방법이다. 템플릿 영상과 비교하여 머리의 방향을 추정하는 방법을 가장 대표적인 방법으로 들 수 있다.

Niyogi[9]는 TSVQ(Tree Structured Vector Quantization)과 저해상도 얼굴 영상 템플릿을 사용하여 머리의 방향을 추정하는 시스템을 작성하였다. Lanitis[10]는 윌콕슨을 기반으로 하는 Flexible Model을 템플릿으로 사용하여 얼굴 요소의 위치정보를 얻고 머리의 방향을 추정했다. Sumi[11]는 얼굴 요소의 개별적 템플릿을 사용하여 위치를 찾고 이들의 위치 정보를 기록한 템플릿과의 정합을 통하여 대략적인 얼굴의 방향을 추정했다. 템플릿 기반 방법은 신체의 개인적 차이나 배경이나 조명 등에 의하여 영향을 받아 달라질 수 있기 때문에 획득할 수 있는 특징의 정밀도에 한계를 가질 수 있다.

이 외에 다중 검출기를 이용하여 머리방향을 추정하는 방법도 있다. Huang[12]은 여러 개의 분류기가 트리 구조를 이루도록 조합한 부스팅 방법을 사용하여 머리방향을 추정하였다. 또, 비선형 회귀방법을 이용하여 머리방향을 추정하는 방법이 있다. Gourier[13]는 색차 기반의 특징을 이용하여 얼굴을 검출하고 정규화하여 imagettes를 얻고, 이를 Widrow-Hoff 학습법을 이용해 각 포트로 학습시킨 linear auto-associative memory 를 이용하여 머리방향을 추정하였다.

지역적 접근법은 특징 벡터 추출이 잘 이루어진다면 템플릿

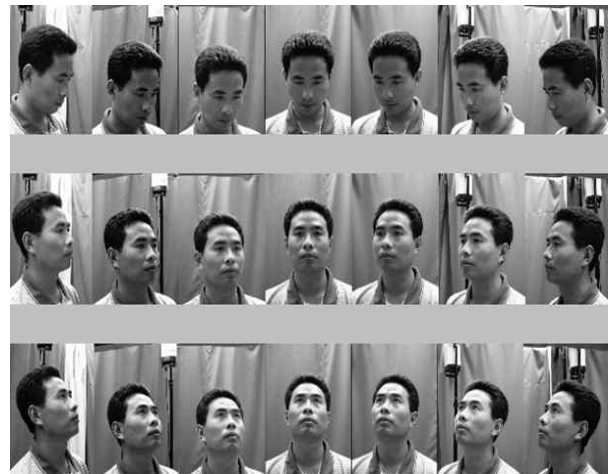


그림 2. 본 연구에서 사용한 CAS-PEAL 얼굴영상 Dataset

기반 방법보다 정밀하게 방향을 추정할 수 있다. 그렇기 때문에 본 논문에서는 얼굴 특징을 추출하기 위해 템플릿 기반 대신 지역적인 특징을 충분히 고려할 수 있는 히스토그램 기반 방법을 사용하였다. 또한 특징 벡터의 생성에 무작위성을 부여하고 신뢰도 있는 결과를 얻기 위해 랜덤 포레스트 알고리즘을 사용하였으며, 본 논문의 목적이 머리방향의 상세한 추정이 아니므로 미리 정해진 방향으로의 분류를 수행하였다.

본 논문에서는 얼굴 영상을 이진 에지 영상으로 변환하고 평균 정면 에지 얼굴과의 차영상으로부터 누적 히스토그램을 작성하고 특징으로 사용하였다. 또한, 수학적 이론인 ‘큰 수의 법칙’에 근거하여 신뢰도 있는 결과를 보장해주는 알고리즘인 랜덤 포레스트를 사용하여 분류를 수행하였다. 본 논문의 구성은 2장에서는 얼굴 영상을 가공하여 전처리 과정을 거쳐 누적 히스토그램 특징을 추출하는 방법과 랜덤 포레스트 분류기의 학습 방법을 소개한 뒤, 본 논문의 실험 내용을 기술하고, 3장에서는 실험 결과 및 향후 연구 방향에 대해 기술한다.

## II. 얼굴영상 특징 벡터의 추출

### 1. 얼굴영상 전처리 및 정규화

실세계에서 얼굴영상은 촬영방향과 거리, 또 조명 조건에 따라 다양한 형태로 얻어지게 된다. 이런 영상을 수학적 데이터로 활용하기 위해서는 표준형의 의미를 갖는 크기와 밝기로 정규화가 이루어져야 한다. 또 불필요한 노이즈와 배경을 제거하기 위한 전처리 과정이 필요하다.

본 연구에서는 다음과 같은 전처리과정을 거쳐서 정규화된 영상을 얻는다(그림3 참조). 먼저 그림2와 같은 영상이 입력되면 그림1과 같이 얼굴의 특징을 포함한 머리 영역을 추출하고 불필요한 잡음을 제거하기 위해 가우시안 필터링을 수행한다. 얼굴영역을 찾는 것은 기존의 알고리즘[4]을 이용할 수 있지만

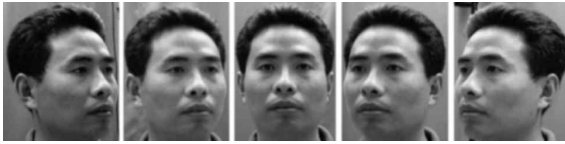


그림 3. 정규화 과정을 거친 입력 영상

본 연구에서는 얼굴영상을 집적한 데이터베이스를 이용하므로 머리영역이 비교적 완전하게 얻어진 것으로 가정한다. 아울러 얼굴영역의 크기를 기준으로 얼굴전체의 사이즈를 정규화하면 얼굴이 회전하였음에도 불구하고 전술의 얼굴의 지역적 특성에 의해 비슷한 범위 안에서 특징을 얻을 수 있다는 장점을 가진다.

## 2. 에지 누적히스토그램에 의한 특징 벡터의 추출

앞 절에서 기술한 바와 같이 얼굴영역의 특징, 즉 입, 눈, 코 등의 위치관계를 분석하면 개략적인 얼굴방향을 알 수 있다. 그러나 이 방법의 단점은 얼굴 각 기관을 정확히 추출할 수 없느냐에 따라 인식결과가 현저하게 달라진다는 점이다. 특히 얼굴은 다양한 조명 조건하에서 얻어지는 경우가 많으므로 얼굴 특징을 정확히 찾아낸다는 것은 거의 불가능한 일이다.

이에 비해 얼굴특징들이 제공하는 에지 정보와 이를 누적한 히스토그램을 이용하면 쉽게 얼굴특징들의 영역을 찾아낼 수 있다. 이 방법은 덧셈과 뺄셈의 단순한 연산으로 이루어지므로 처리 속도가 빨라 실시간 검출이 필요한 응용에 적합한 방식이다. 그러나 누적 히스토그램을 이용하여 모델을 학습시키기 위해서는 머리의 위치가 일치해야하며 배경에 의한 잡음 또한 최소화해야한다. 따라서 앞 절에서 기술한 학습영상의 전처리과정이 필요하게 된다.

특징벡터를 구하기 위해 정규화된 영상에 캐니 에지 검출기를 적용하여 이진 에지 영상으로 변환하고 평균 정면 에지 영상과의 차영상을 계산한 후, 수평·수직 방향으로 에지 픽셀을 합한 누적 히스토그램을 작성한다. 평균 정면 에지 영상은 전처리 과정을 거친 학습데이터에서 각 축에 대한 회전량이 0인 영상의 평균을 구한 뒤, 얼굴의 특징이 잘 드러나도록 임계값을 조절하여 이진화한 영상이다. 평균 정면 에지 영상과의 차영상을 이용하면 정면 에지 얼굴의 얼굴 요소와 위치가 다른 부분의 에지를 검출할 수 있기 때문에 누적 히스토그램을 구하면 얼굴 회전에 따른 변화 정보를 얻어낼 수 있다.

차영상에서 얻은 두 누적 히스토그램은 다시 한 번 히스토그램 스무딩 및 정규화 과정을 거친다. 이 과정은 그림4에서 보이는 바와 같이 일정 pixel 간격으로 히스토그램 구간을 분할하고 각 구간은 해당 구간의 평균값으로 값을 대체한다. 정규화 과정에서는 히스토그램 특징값의 범위를  $[0, 1] \in R$  사이의 값으로 재조정한다.

## III. 랜덤 포레스트에 의한 얼굴 방향 추정

랜덤 포레스트[2]는 2001년에 L. Breiman이 제안한 알고리즘으로 전통적인 결정트리(decision tree) 기법을 다수 개의 나무 구조로 확장시킨 결정트리 메타학습(meta-learning) 기법이다. 랜덤 포레스트를 구성하는 트리는 일반적인 결정트리에서 다루기 힘든 가지치기 과정을 생략하고 성장과정만을 거쳐 생성되기 때문에 분류 모델 디자인이 더욱 용이하다. 이는 '큰수의 법칙'에 근거하여, 충분히 많은 서로 다른 랜덤결정트리의 결과를 조합함으로써 가지치기를 한 단일 결정트리의 성능을 능가하는 정확도를 얻을 수 있기 때문이다.

랜덤 포레스트의 임의성을 높일수록 각 결정트리간의 연관성이 낮아지는데, 이는 분산 감소와 직결되어 알고리즘의 정확도 향상에 기여한다. 임의성을 부여할 수 있는 요소로는 각 트리의 학습 데이터 선택과 각 노드에서의 최적화 과정이 있다. 이를 고려한 랜덤 포레스트 학습과정은 다음과 같다.

먼저, 학습 집합으로부터 생성하고자하는 결정트리 수만큼의 부트스트랩을 구성한다. 부트스트랩이란 학습 집합으로부터 임의 복원추출을 반복하여 원학습 집합과 크기가 동일하게 생성한 부분집합이다. 부트스트랩을 이용하여 각 트리를 생성하면 학습영상에 차이가 생기므로 각 나무의 첫 번째 노드의 분기합수부터 변화가 생기게 되고, 이로 인해 다양한 형태의 나무들을 얻을 수 있다. 즉, 부트스트랩은 학습영상의 선택에 임의성을 부여하는 장치이다.

부트스트랩 생성 이후, 각 부트스트랩으로부터 결정트리를 성장시킨다. 일반적인 결정트리와 달리 랜덤포레스트를 구성하는 나무는 각 노드에서 학습영상의 모든 특징값 또는 학습영상의 특징값의 모든 차원을 고려하지 않고 사전에 지정된 개수만큼 임의로 선출된 것만을 고려하여 분기함수를 생성한다. 이로 인해 같은 학습영상에 주어지더라도 다른 형태를 갖는 나무로 성장시킬 수 있다. 즉, 특징값의 개수나 차원에 임의성을 부여하는 방법은 노드 최적화에 임의성을 부여하는 장치로 쓰인다. 이 경우 각 개별 랜덤결정트리의 정밀도는 떨어질 수 있으나, 앞서 설명한 바와 같이 이들을 조합하여 예측을 수행하게 되는 랜덤포레스트의 정확도와 안정성은 높아지게 된다.

각 결정트리의 성장을 멈추는 기준은 다음과 같다. 첫째, 해당 노드에 도달한 학습 데이터의 수가 어떤 값보다 작은 경우, 해당 노드에서 데이터를 다시 분할하기엔 너무 적은 데이터를 다룸으로써 일반성을 잃을 수 있으므로 성장을 멈추고 잎노드를 생성한다. 둘째, 결정트리의 깊이 값이 일정 값에 도달한 경우, 해당 깊이까지 완전히 성장시킨 트리의 경우 깊이 값  $d$ 가 커질수록 총 계산량은 식 1과 같이 기하급수적으로 증가하므로 시간적 측면에서 효율이 매우 떨어지게 되므로 성장을 멈추고 잎

노드를 생성한다. 셋째, 데이터를 분할하여 얻을 수 있는 정보 획득량이 일정 값 이하인 경우, 해당 노드가 트리의 정확도에 끼치는 영향력이 매우 적으므로 성장을 멈추고 잎노드를 생성한다.

학습 데이터  $X = \{(x_1, y_1), \dots, (x_N, y_N)\}$ 가 특징 기술자  $x_i = [x_i^1, \dots, x_i^D]^T$ 와 부류 라벨  $y_i \in \{1, \dots, M\}$ 로 이루어져 있을 때, 특징의 각 차원에 대해 특징 값의 최대값과 최소값을 구하고 그 범위 안의 임의의 임계값을 기준으로 데이터를 나누고 나누어진 데이터들을 기준으로 해당 노드와 자식 노드의 불순도를 계산한다.

정보획득량은 아래의 식들을 이용하여 구한다. 식 (2)는 엔트로피 지수를, 식 (3)은 지니 지수, 식 (4)는 오분류 불순도를 각각 나타내며 식 (5)는 식 (2), (3), (4) 중 하나를 대입하여 정보획득량을 계산한다. 여기에서  $\omega_i$ 는  $M$ 개의 부류 중에서  $i$ 번째를 나타내며  $P(\omega_i|T)$ 는  $T$ 노드에서  $\omega_i$ 가 발생할 확률이다.

$$im(T) = \sum_{i=1}^M (y_i - \bar{y})^2 \tag{1}$$

$$im(T) = - \sum_{i=1}^M P(\omega_i|T) \log_2 P(\omega_i|T) \tag{2}$$

$$im(T) = 1 - \sum_{i=1}^M P(\omega_i|T)^2 = \sum_{i \neq j} P(\omega_i|T) P(\omega_j|T) \tag{3}$$

$$im(T) = 1 - \max_i P(\omega_i|T) \tag{4}$$

$$I(T) = im(T) - \frac{|X_L|}{|X_T|} im(T_L) - \frac{|X_R|}{|X_T|} im(T_R) \tag{5}$$

식(5)에서  $L$ 은 왼쪽 가지를  $R$ 은 오른쪽 가지를 나타내는 첨자이다. 정보획득량  $I(T)$ 의 값이 클수록 분류가 잘된 것으로 판단할 수 있으므로, 이를 최대로 만드는 임계값과 차원을 구하여 해당 노드의 분기함수로 저장하고, 해당 노드가 정지기준을 만족할 때 까지 같은 방법을 되풀이하여 나무를 성장시킨다. 결정트리의 실험과정은 부류 라벨을 알 수 없는 특징기술자를 입력 받아 각 노드의 분기함수의 만족여부에 따른 특정 자식노드로의 이동을 반복하여 잎 노드에 도달하면 저장된 값에 따라 분류결과나 회귀결과 값을 얻는다.

이 알고리즘은 ‘큰 수의 법칙’에 근거하므로 랜덤포레스트의 크기가 충분히 커야 일반화 능력이 우수한 모델을 얻을 수 있기 때문에 실험적으로 알고리즘의 성능을 유지하면서도 계산량을 최소화하는 최적의 나무 개수를 찾아야한다. 또한, 의사결정나무의 성장에 영향을 끼치는 나무 깊이의 최댓값, 임의성의 정도와 최적의 수치도 마찬가지로이다. 이와 같이 랜덤 포레스트 학습에 영향을 미치는 매개변수들은 학습영상과 특징에 따라 최적 값

이 변화하기 때문에 사용자가 실험적으로 알맞은 값을 찾아야 한다.

랜덤 포레스트 실험과정에서는 주어진 영상에서 특징을 추출한 뒤 학습된 모든 의사결정나무의 루트노드에 값을 넣어 해당 노드에서 하위 노드로의 이동을 반복한다. 각 나무마다 최종적으로 도달한 잎노드에 저장된 값들을 일반적으로 평균이나 voting 방법을 이용하여 최종 결과를 도출한다. voting 방법은 승자독식방식으로 랜덤포레스트를 이루는 모든 나무의 결과를 합산하였을 때, 이들 중 가장 많은 득표를 얻은 클래스로 분류 결과를 예측하는 방식이다. 이외에도 각 트리의 분류 정확도를 해당 결과의 가중치로 이용하여 합산함으로써 좀 더 정확한 분류 결과를 도출할 수도 있다.

#### IV. 실험 및 분석

본 논문에서 제안한 알고리즘의 타당성을 확인하기 위해서 그림 2와 같이 CAS-PEAL Dataset[1]을 입력 영상으로 사용하여 실험을 실시하였다. 이 데이터는 남성 595명과 여성 445명 총 1040명의 인물에 대해 포즈변화, 표정변화, 장신구 착용, 조명 조건 변화 등의 다양한 가변요소를 포함한 영상들로 구성되어 있다.

머리방향 인식을 위한 데이터베이스는 개인별로 21(7\*3)개의 다른 영상 그룹으로 구성되어있다. 이 중 101명은 수직축에 대한 회전이 (-67°, -45°, -22°, 0°, 22°, 45°, 67°) 939명은 (-45°, -30°, -15°, 0°, 15°, 30°, 45°)이고 수평축에 대한 회전은 (-30°, 0°, 30°)으로 동일하다. 본 실험에는 이들 중 후자인 939명에 대한 포즈영상만을 이용하였으며 학습에 700명의 포즈영상을 사용하고 나머지 239명의 포즈영상으로 실험을 진행하였다.

본 논문에서 이전에 설명한 정규화 과정을 거친 입력 영상은 다음 그림 3과 같으며 160x190 크기의 영상이다. 정규화된 영상은 캐니 에지 검출법을 이용하여 에지 영상으로 변환한다. 이 영상의 수직·수평 누적 히스토그램은 다음 그림 4와 같이 구할 수 있으며, 본 실험에서 실제 특징으로 사용되는 누적 히스토그램은 입력영상의 에지영상과 평균얼굴에지영상의 차영상으로부터 구한다.

히스토그램은 디지털 영상의 점검과 화질을 개선하는데 사용될 뿐만 아니라 데이터의 특징을 한눈에 알아볼 수 있도록 그래프 형태로 표현되는 장점이 있다. 본 논문에서는 이 히스토그램을 에지의 특징을 추출하는 용도로 사용하였으며 밝기값이 아닌 영상 특징의 분포를 x축과 y축에 누적시키는 히스토그램으로 표현하였다. 히스토그램과 각 회전 영상들에서 히스토그램을 통해 얻을 수 있는 특징 정보의 예는 다음 그림 5와 같다.

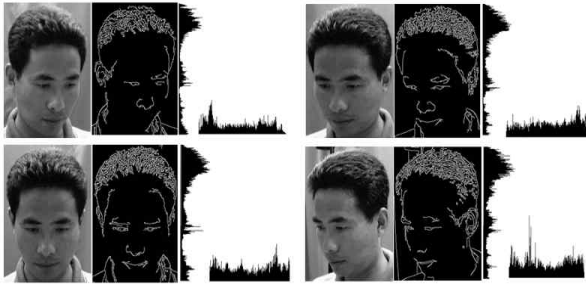


그림 4. 입력영상과 에지영상 및 에지영상에 대한 누적 히스토그램

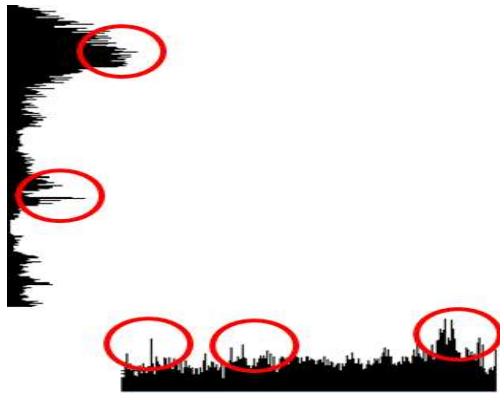


그림 5. 히스토그램의 회전 특징 정보. 회전에 의해 발생한 변화량이 히스토그램에 누적되어 있다.

히스토그램을 영상의 특징을 이용하여 랜덤포레스트를 생성한다. 랜덤포레스트의 학습과정은 앞 서 설명한 바와 같이 일반적인 의사결정트리와 다르게 노드 최적화 시 임의성을 부여한다. 본 실험에서 사용한 특징은 각 영상에서 2개의 특징값을 얻으며 그 차원의 수가 높기 때문에 트리 생성 시 각 노드에서 학습 데이터의 특징 값의 모든 차원을 고려하지 않고, 임의로 선택된 차원의 특징 벡터 중 가장 높은 정보획득량을 얻는 값으로 분리함수를 구성한다. 그림 6은 랜덤포레스트를 구성하는 결정 트리를 생성할 때, 특징 영상 내에 존재하는 임의의 차원의 값을 이용하여 결정트리를 생성하는 예를 보여준다.

랜덤포레스트는 노드의 최적화 시 고려하는 차원의 수가 많을수록 임의성이 낮아지므로 다른 결정트리와의 연관성이 높아져 대부분이 비슷한 형태의 나무로 생성되며, 이러한 경우 단일 결정트리와 비슷한 성능을 갖는다. 따라서 고려하는 차원의 수를 적게 지정하여 임의성을 높이면 다른 결정트리와의 연관성이 감소하므로 다양한 형태의 나무들이 생성되어 이를 결합하면 단일 결정트리보다 정확한 결과를 얻어낼 수 있다.

표 1은 실험 영상을 이용하여 머리방향을 인식한 결과이다. 머리방향을 좌우 회전과 상하 기울임 두 방향의 조합으로 나누고 각 경우에 대해 인식을 한 결과 고개를 들거나 숙이면서 좌

우로 회전을 한 경우 인식률이 가장 낮게 나왔다. 정면을 향한 경우 거의 완전하게 방향을 인식할 수 있었다.

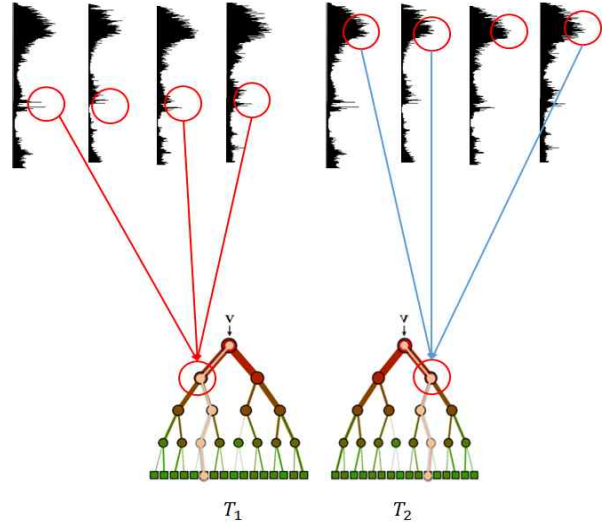


그림 6. 히스토그램을 이용한 랜덤투리의 생성 예

표 1. 머리방향 추정 결과의 정확성

		y축 각도						
		-45°	-30°	-15°	0°	15°	30°	45°
x축 각도	-30°	67%	74%	84%	89%	82%	76%	60%
	0°	73%	85%	96%	98%	94%	87%	75%
	30°	65%	78%	87%	88%	88%	84%	62%

### V. 결론

본 논문에서는 누적히스토그램을 특징으로 이용하여 랜덤포레스트 분류기로 얼굴 방향을 추정하는 방법을 제안하였다. 얼굴 입력 영상으로는 CAS-PEAL face Database를 이용하였으며, 얼굴 영역 전체의 에지 특징을 히스토그램으로 표현함으로써 전역적인 특징을 이용하였다.

에지를 히스토그램으로 표현하는 특징인 HOG의 경우 내부의 패턴이 복잡하지 않으면서 윤곽선 정보가 뚜렷한 객체를 식별하는데 좋은 성능을 보이지만 계산량이 너무 많다는 단점이 있다. 따라서 본 논문에서는 HOG특징에 비해 간단하면서도 얼굴의 특징을 표현할 수 있도록 에지 차영상의 히스토그램을 사용하였다.

그러나 본 논문에서 제안한 알고리즘을 이용하기 위해서는 제한된 배경과 환경에서 실험해야하므로 향후에는 이를 보완하



기 위한 방법으로 진처리 과정을 자동화할 수 있도록 수정하고 정확도를 보정하기 위한 부가적인 요소들을 추가, 보완하는 연구를 수행할 계획이다.

## References

- [1] <http://www.jdl.ac.cn/peal/JDL-PEAL-Release.htm>
- [2] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp.5 - 32, 2001.
- [3] A. Gee and R. Cipolla, "Non-intrusive Gaze Tracking of Human Computer Interaction," *Cambridge University*, 1995.
- [4] T. Horprasert, Y. Yacoob and L.S. Davis, "Computing 3-D Head Orientation from a Monocular Image Sequence," *Proc. 2nd Int. Conf on Automatic Face and Gesture Recognition*, pp.242-247, 1996.
- [5] J.G. Wang and E. Sung, "EM Enhancement of 3D Head Pose Estimated by Point at Infinity," *Image and Vision Computing*, vol. 25, no. 12, pp. 1864-1874, 2007.
- [6] G. Fadda, G. L. Marcialis, F. Roli, L. Ghiani, "Exploiting the Golden Ratio on Human Faces for Head-Pose Estimation." *In: Image Analysis and Processing - ICIAP 2013*, Springer Berlin Heidelberg, pp. 280-289, 2013.
- [7] T. Maurer and C. von der Malsburg, "Tracking and Learning Graphs and Pose on Image Sequences of Faces," *Proc. 2nd Int. Conf on Automatic Face and Gesture Recognition*, pp. 176-181, 1996.
- [8] S.G. Kong and Ralph Oyini Mbouna, "Head Pose Estimation From a 2D Face Image Using 3D Face Morphing With Depth Parameters," *IEEE Trans. on Image Processing*, Vol.24, No.6, pp.1801-1808, 2015.
- [9] S. Niyogi and W. Freeman, "Example-Based Head Tracking", *Proc. 2nd Int. Conf on Automatic Face and Gesture Recognition*, pp.374-377, 1996.
- [10] A. Lanitis, C.J. Taylor, T.F. Cootes and T.ahmed, "Automatic Interpretation of Human Faces and Hand Gestures Using Flexible Models", *Proc. IEEE Int. Conf on Automatic Face and Gesture Recognition*, pp.98-103, 1995.
- [11] Y. Sumi and Y. Ohta, "Detection of face orientation and facial components using distributed appearance modeling", *Proc. IEEE Int. Conf on Automatic Face and Gesture Recognition*, pp. 254-259, 1995.
- [12] C. Huang, H.Ai, Y. LI and S.Lao, "High-performance rotation invariant multiview face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.29, No.4, pp. 671-686, 2007.
- [13] N. Gourier, J. Maisonnasse, D. Hall and J.L. Crowley, "Head pose estimation on low resolution images," *Lecture Notes in Computer Science 4122*, pp.270-280, 2007.

## 저자 소개

### 문 성 희(학생회원)

2009년 전남대학교 전기공학과 학사 졸업.  
2016년 전남대학교 전자컴퓨터공학과 석사 졸업.

<주관심분야 : 제스처 인식, 스마트폰, HCI>



### 이 철 우(정회원)

1992년 동경대학교 대학원 전자공학과 박사 졸업.

1992년 4월 ~ 1995년 12월 일본 이미 지정보과학연구소 수석연구원, 오사카대학 기초공학부 협력연구원 및 리즈메이칸대학

정보학부 특별초빙강사.

1996년 1월 ~ 현재 전남대학교 전자컴퓨터 공학부 교수.

2002 1월 ~ 현재 2003년 2월 미국 NC A&T State University 방문교수.

2008 10월 ~ 2009년 8월 KAIST 및 오사카대학 초빙교수.

2006 3월 ~ 현재 전남대학교 문화콘텐츠기술연구소 소장.

<주관심분야 : 컴퓨터 비전, 지능형 휴먼 인터페이스, 디지털 콘텐츠, 컴퓨터그래픽스>