

대표적 얼굴 특징점 추출 방법에 대한 비교분석

신길수^o 김용국
 세종대학교 컴퓨터공학과
 {gislover^o@naver.com, ykim@sejong.ac.kr}

Comarition between Two Facial Feature Detection Methods

Gilsu Shin^o Yong-Guk Kim
 Department of Computer Science
 Sejong University

요 약

이 논문에서는 커널 에지 방식의 얼굴의 특징점을 추출하는 방법과 Adaboost를 이용한 얼굴의 특징점을 추출하는 방법에 대해서 비교 한다. 커널 에지를 이용한 방법은 10개의 커널을 이용하여 추출된 에지를 이용하여 얼굴의 특징점을 추출해 낸다. 커널의 개수를 줄여 사용하면 실시간에 가능하고, 정확성을 높이기 위해서는 이미지의 전처리 단계에서 자극적인 효과를 준다면 정확성 또한 높아 질 것이다. 반면에 Adaboost를 이용한 방법은 각각의 특징점들을 오프라인 상에서 학습을 하고 온라인상에서 실시간으로 특징점을 추출하는 방법을 사용하였다. 각 각의 학습과정에 있어서 positive, negative 이미지를 더 많이 사용하면 정확성이 더 높아질 것이다. 한 가지 주목할 만 한 점은 입과 같은 특징점을 추출하기 어려운 영역에서도 높은 정확성을 보였다.

1. 서 론

자동으로 얼굴의 찾기(Detection)와 인식(Recognition)은 지난 수십 년 간 많은 연구진들의 주요 연구과제였다. 얼굴과, 표정인식에서 가장 중요한 것은 얼굴을 찾고 그 얼굴에 대한 특징점들(Features)을 찾는 것이다.[1]

그 특징점들은 조명, 포즈, 표정과 크기와 같이 다양한 변화들 자동으로 인지하는 데 있어 매우 중요한 인자가 된다. Yacoob et al.[2]은 얼굴 인식을 템플릿(Template)과 특징점 기반의 기술을 갖고 비교를 하였지만 두 가지 방법 모두 그리 좋은 성능을 보여주진 못하였다.

이 논문에서는 커널 에지(Edge)를 기반으로 한 특징점 추출 방법[3]과 Adaboost를 이용한 특징점 추출 방법[4]에 대해서 논하도록 하겠다. 두 가지 방법 모두 훌륭한 평가를 받고 있어서 더욱 더 비교의 가치가 있을 것이라 판단된다.

간략하게 요약해 보면, 2장에서는 이 논문에 사용된 Database를 서술하고, 3장에서는 에지 기반의 특징점 추출 방법과 Adaboost 기반의 특징점 추출 방법에 대하여 기술하고, 4장에서는 그 두 방법의 결과를 말하며, 5장에서는 결론을 맺는다.

2. DATABASE

데이터베이스는 남성 21명의 사람들로 정면 3장, 총 63장으로 구성되어져 있다. 즉 한 사람당 3장의 이미지로 이루어져 있다. 각 이미지들은 그 어떠한 전처리없이, 실제 웹 캠(Webcam)을 이용하여 320X240의 해상도로 만들었다. <표 1>

Database	
Capturing Device	web camera
Image Type	2D still image
Image Number	63
Subject Number	21
Session Number	3
Pose	Frontal
Illumination	same condition (220~260lux)
Expression	neutral

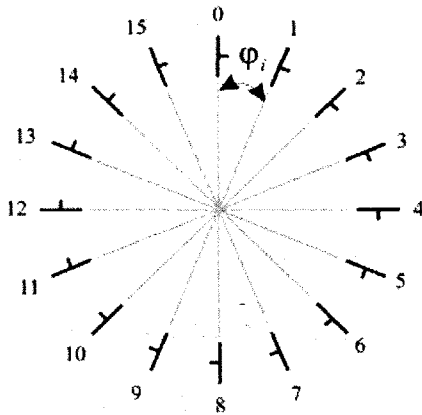
<표 1> DB구성을 위한 환경

3. FACIAL FEATURES DETECTION

여기서 우리는 두 가지 방법으로 각 특징점을 찾으려고 한다. 한 가지 방법은 논문 [3]에서 묘사한 방법으로 커널에지를 이용한 특징점 검출방법이다.

3.1 Facial Landmarks Extraction

눈을 포함한 눈썹, 코, 입을 얼굴의 주요 특징점이라 하여 그 특징점들을 찾는다. 커널에지를 이용한 특징점을 찾는 방법은 이미지의 에지의 위치에 따라 맵을 구성하는 방법이 있었다. 이 방법은 22.5도씩 45° 에서 135° 와 225° 에서 315° 까지 이동하여 에지의 위치를 이용하여 맵을 구성하였다. (그림 1)



<그림 1> 에지 추출을 위한 위치 템플릿

이 방법은 Preprocessing, Edge Map Construction, Orientation Matching과 같은 3가지 단계로 이루어져 있다.

각 각의 단계는 다음에서 간략하게 기술한다.

3.1.1 Preprocessing

전처리하는 과정은 입력 이미지를 흑백이미지로 만들고, 재귀 가우시안을 이용하여 이미지를 좀 더 부드럽게 처리한다. 이렇게 부드러워진 이미지는 얼굴의 특징점을 찾는데 활용된다.

3.1.2 Edge Map Constructing

10개의 커널(Kernel)을 이용해 이미지에 덧씌우면 각각의 커널들은 10개의 커널에 민감한 반응을 보인다. 아래의 수식들을 이용하여 Edge Map을 만든다.

$$G_{\psi k} = \frac{1}{Z} (G_{\psi k}^- - G_{\psi k}^+)$$

$$Z = \sum_{p,q} (G_{\psi k}^- - G_{\psi k}^+), G_{\psi k}^- - G_{\psi k}^+ > 0$$

$$G_{\psi k}^- = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(p - \sigma \cos \psi k)^2 + (q - \sigma \sin \psi k)^2}{2\sigma^2}\right),$$

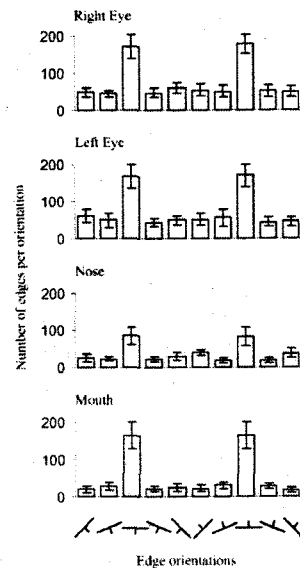
$$G_{\psi k}^+ = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(p + \sigma \cos \psi k)^2 + (q + \sigma \sin \psi k)^2}{2\sigma^2}\right),$$

$$g_{ij\psi k} = \sum_{p,q} b_{(i-p, j-q)}^{(l)} G_{\psi k},$$

- σ 가우시안 분포의 제곱 평균 편차
- ψk $k \cdot 22.5$
- k 2, 3, 4, 5, 6, 10, 11, 12, 13, 14
- p, q -3, -2, -1, 0, 1, 2, 3
- $l = 0 \div W - 1$
- (i, j) $j = 0 \div H - 1$
- (W, H : Width, Height)
- l 1, 2

3.1.3 Orientation Model

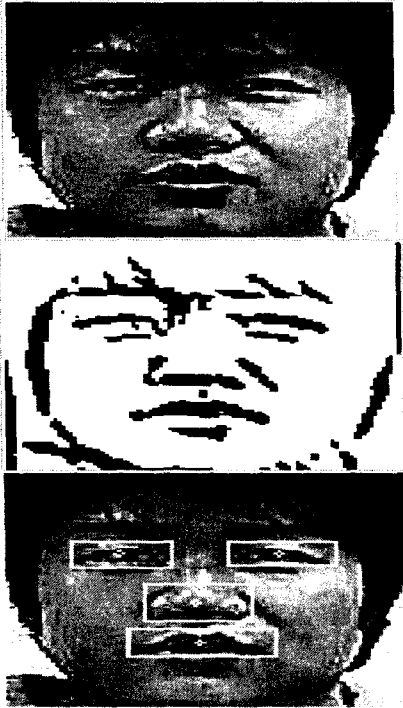
마지막으로 4가지의 특징점들 중 각각의 특징점의 평균적인 위치를 구한다. 평균 위치를 유지하며 각각의 특징점에 대한 패턴을 갖고 대조를 한다. (그림 2)



<그림 2> 평균 특징점의 위치

3.1.4 Facial Landmarks Extraction Result

이 방법은 실시간으로 실행하기엔 많은 문제점이 있고, 조명변화에 굉장히 민감하다. 이러한 단점들을 극복하는 것은 비교적 다른 알고리즘에 비해 간단한 알고리즘을 갖고 구현하는 것이 이 논문의 장점이 되겠다. (그림 3)



<그림 3> 에지를 이용한 특징점 검출

3.2 Adaboost를 이용한 얼굴의 특징 검출

커널 에지를 이용한 얼굴 특징 검출의 방법과 대조를 이루며 실시간 검출이 가능하고 성능 또한 더 좋은 방법에 대해서 기술한다.

3.2.1 Adaboost를 사용하는 이유와 방법

대다수의 특징점 추출 알고리즘은 뉴럴넷(Neural Net)[7] 혹은 SVM(Support Vector Machine)[8]과 같은 학습기반의 기술을 사용하였다. 이 훌륭한 성능을 발휘하는 알고리즘들은 실시간으로 적용하기에는 많은 어려움이 있다.

그래서 실시간으로 얼굴의 특징점을 찾기 위하여 Adaboost[4][5]를 사용하게 되었다. Adaboost 알고리즘은 오프라인 상에서 패턴을 학습하고 분류하여 온라인 상에서 특징점을 찾을 수 있는 알고리즘이다. 오프라인

상에서 Haar functions[5][6]은 positive 이미지에 그 에 해당하는 특징점이 있고 negative 이미지에 그 특징점이 없어야 학습을 할 수 있다. 이렇게 오프라인 상에서 학습되어진 패턴들의 연결을 이용하여 실시간으로 온라인상에서 특징점들을 검출하게 되는 것이다.

3.2.2 Adaboost 학습과 흐름도

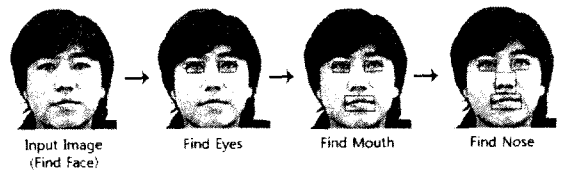
이 논문에서는 눈, 코, 입 4개의 특징점을 찾으려 한다. 우리는 positive와 negative 47,000개의 이미지를 이용하여 학습을 시켰고, 각 특징점에 대한 것은 표 2에서 기술한다.

Features	# of positive	# of negative
eyes	7,000	14,000
nose	4,000	9,000
mouth	4,000	9,000

<표 2> 각 특징점들을 추출을 위한 학습개수

또한 10개의 Haar-Features와 20개의 Weak Classifiers들의 집합으로 패턴들을 학습하였다.

빠르고 좀더 정확한 추출을 위한 시나리오는 다음과 같다



<그림 4> 전체 흐름도

1. 전체의 이미지 중 얼굴의 이미지를 자동 검출한다.
2. 자동 검출된 얼굴의 영역 내에서 두 눈을 찾으려 하고 순차적으로 코와 입을 찾으려 한다. (그림 4)

3.2.3 Adaboost를 사용하여 특징점 추출 결과

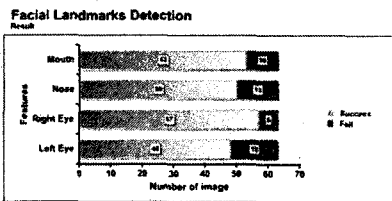
2D 상의 모든 이미지들을 일일이 매뉴얼로 특징점을 찍는 것보다 훨씬 빠르고 다른 알고리즘에 비해 좀 더 정확하다는 것을 알 수 있다. 그 결과로 데이터베이스에서 Adaboost를 이용한 몇 개의 결과를 볼 수 있다. (그림 5)



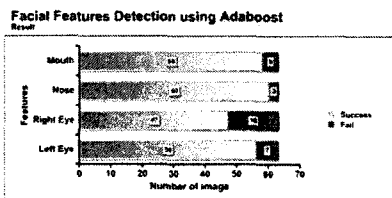
<그림 5> 특징점을 찾은 결과 예제

4. RESULTS

두 가지 방법의 특징점 추출은 다음과 같은 결과를 보여주었다. 에지를 이용한 추출방법(그림 6)은 입과 코를 많이 못 찾는 결과를 보여주고 있고, Adaboost를 이용한 추출방법(그림 7)은 오른쪽 눈을 제외한 모든 특징점에서 상대적으로 뛰어난 성능을 보여준다. 실시간에서 에지를 사용하게 되면 아래의 그림 6에서 보는 것과 같은 성능을 발휘하기는 힘들게 된다. 또한 실시간이라고 말을 붙이기 무색할 정도의 성능을 보인다. 그러나 Adaboost를 이용한 방법은 실시간과 아래의 그림 7에서 보는 것과 유사한 성능을 발휘하게 된다.



<그림 6> 특징점 추출 결과(edge)



<그림 7> 특징점 추출 결과(Adaboost)

5. 결 론

커널에지를 사용하여 얼굴의 특징점들을 찾는 방법은 수식이 간단한 만큼 구현해서의 결과물도 평균이상의 성과를 볼 수 있다. 또한 실시간에서 사용을 하기 위해서 몇 가지 제의를 해보자면 얼굴의 특징에 있어 굳이 10개의 커널을 사용하는 것보다는 단 소수개의 커널을 사용하여 얼굴의 에지를 추출하는 게 실시간에 적합할 것이며, 일반 이미지를 전처리 과정에서 좀 더 매끄럽게 다듬어 준다면 기대했던 성능보다 빠르고 좀 더 정교한 성능을 발휘 할 수 있을 것이다.

반면에, Adaboost 방법은 학습시간과 positive, negative 이미지의 수가 실제 사용했을 때 영향을 많이 준다는 것을 알 수 있다. 물론 그 만큼의 시간투자에 대한 대가는 성능으로 돌아온다는 것도 실험을 통해서 알 수 있다.

아래 표 3과 같이 두 가지 방법에 대한 추출율을 보여주는 바와 같이 일반적으로 알려진 입에 대한 추출율도 좋은 성능을 발휘하는 것으로 나타났다.

	Edge	Adaboost
Accuracy	82.25 %	87.25 %
Left eye	76 %	88 %
Right eye	90 %	74 %
Nose	79 %	95 %
Mouth	84 %	92 %

<표 3> 두 가지 방법의 각 특징점에 대한 추출율

이러한 얼굴의 특징점 추출은 앞으로 여러 분야에 적용될 것을 생각된다. 각 각의 피쳐들은 표정인식에 있어 기초자료가 될 수도 있고, 이미지 상에서의 얼굴의 각도를 계산할 수도 있을 것으로 생각된다. 그렇지만, 여전히 두 가지 방법 모두 단점을 갖고 있으며 특히, Adaboost 같은 경우 모든 환경에 맞춘 positive, negative 이미지가 많은 양이 필요 하고, 또 그 많은 이미지들을 다시 학습시키는 데 있어 비록 오프라인 상이지만, 시간과 그 특징점이 맞는지 틀리는지 일일이 사람의 손으로 학습을 시켜야 하기에 많은 노력이 필요하다는 단점을 갖고 있다.

6. 참고문헌

[1] A. Pentland, B. Moghaddam, and T.Starner, "View-based and Modular Eigenspaces for Face Recognition", Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.84-91, June 1994

[2] Y. Yacoub, H.M. Lam and L. Davis, "Recognition faces Showing Expressions", Proc. Int'l

Workshop Automatic Face and Gesture-Recognition,
pp.278-283, June 1995

[3] Y. Gizatdinova, V. Surakka, "Feature-Based
Detection of Facial Landmarks from Neutral and
Expressive Facial Images". IEEE TRANSACTIONS ON
PATTERN ANALYSIS AND MACHINE INTELLIGENCE,
VOL.28, NO.1 JAN 2006

[4] P. Viola and M. Jones, "Rapid Object Detection
Using a Boosted Cascade of Simple Features", In
Proceedings IEEE Conference on Computer Vision
and Pattern Recognition pp. 511-518, 2001

[5] R. Lienhart and J. Maydt, "An Extended Set of
Haar Like Features for Rapid Object Detection", IEEE
ICIP 2002, pp. 900-903, 2002

[6] C. Papageorgiou, M. Oren, and T. Poggio, "A
general framework for Obejct Detection", In
International conference on Computer Vision, 1998

[7] H. Rowley, S. Baluja and T. Kanade, "Neural
Network Based Face Detection", IEEE Transactions
on Pattern Analysis and Machine Intelligence,
20(1):23-38, 1998

[8] E. Osuna, R. Freund and F. Girosi "Training
Support Vector Machines: An Application to Face
Detection", In Proceedings IEEE Conference on
Computer Vision and Pattern Recognition,
pp.130-136, 1997