

표정인식 시스템을 위한 얼굴 특징 영역 추출

김상준*, 이성오*, 박귀태**
 고려대학교 전기공학과 대학원 지능시스템리서처연구실*,
 고려대학교 전기전자전파공학부**

FacialFeaturesExtractionforRecognitionSystemofFacialExpression

Sang-jun Kim*, Sung-oh Lee*, Gwi-tae Park**
 Intelligent System Research LAB., Department of Electrical Engineering, Korea University*
 Department of Electrical · Electronic · Radio Engineering, Korea University**

Abstract - 표정인식은 컴퓨터 비전 분야에서 중요한 부분을 차지하고 있으며, 현재 꾸준히 연구가 진행되고 있다. 표정인식 시스템은 크게 얼굴 영역 추출과 표정인식 부분으로 나눌 수 있으며, 얼굴 영역 추출은 전체 인식 시스템의 성능에 큰 영향을 미친다. 특히 표정인식 시스템은 일반 얼굴인식 시스템과 다르게 부분적으로나 전체적으로 형태의 변화가 큰 얼굴에 대해서 정확한 얼굴 영역이 확보되지 않으면 높은 인식성능을 기대하기 어렵다. 따라서 표정인식 시스템은 얼굴 영역 추출이 비중한 부분을 차지하고 있다. 본 논문에서는 영상에서 실시간으로 얼굴 영역을 찾아내고, 그 영역에서 얼굴의 특징적인 눈과 입의 위치를 검출하고, 이를 바탕으로 얼굴의 정확한 영역을 확정하는 일련의 과정을 서술한다.

1. 서 론

표정인식은 컴퓨터 비전 분야에서 중요한 부분을 차지하고 있으며, 현재 꾸준히 연구가 진행되고 있다.[4][5][6] 표정인식 시스템은 크게 얼굴 영역 추출과 표정인식 부분으로 나눌 수 있으며, 얼굴 영역 추출은 전체 인식 시스템의 성능에 큰 영향을 미친다. 특히 표정인식 시스템은 일반 얼굴인식 시스템과 다르게 부분적으로나 전체적으로 형태의 변화가 큰 얼굴에 대해서 정확한 얼굴 영역이 확보되지 않으면 높은 인식성능을 기대하기 어렵다. 따라서 표정인식 시스템에서 얼굴 영역 추출이 비중한 부분을 차지하고 있다.[3][4] 본 논문에서는 영상에서 SVM을 이용하여 얼굴 영역을 찾아내는 일련의 과정을 서술한다.

표정인식시스템에서는 주로 일정한 크기로 정규화된 이미지를 사용하는데, 일반적으로 얼굴의 특징되는 부분의 위치 정보를 바탕으로 얼굴 영역을 정규화한다. 눈과 입의 상호 위치는 얼굴의 특징을 나타내는 중요한 정보가 되는데, 이러한 눈과 입의 위치를 찾아내는 것은 서로 다른 얼굴의 표정을 일정한 크기로 정규화 하는데 중요한 부분을 차지한다. 특히 표정인식 시스템에서는 표정에 따라 급격히 변화되는 다양한 눈과 입의 패턴을 정확하게 인식하여 그 위치를 알아내는 것이 중요하다고 하겠다.

SVM은 두 개의 class를 구분하는데 매우 유용하게 사용되는 기법 중에 하나이다. 가령, 얼굴의 특정 패턴을 구분한다고 할 때, 눈과 눈이 아닌 데이터, 또는 입과 입이 아닌 데이터를 구분하는데 쓰일 수 있다. 또한 SVM은 다양한 패턴을 가지는 데이터와 그러한 패턴을 가지지 않는 데이터에 대해 확률적인 방법을 통하여 최적의 하이퍼

플레인(hyperplane)을 구할 수 있기 때문에 표정이 있는 얼굴에서 눈과 입을 잘 나타낼 수 있고 따라서 표정이 있는 얼굴에서 눈과 입을 잘 찾아낼 수 있다. 본 논문에서는 이러한 SVM의 특징을 이용하여 눈과 입의 위치를 알아내고, 이러한 위치 정보를 바탕으로 얼굴을 정규화 하는 방법을 설명하겠다.

2. 본 론

2.1 SVM (Support Vector Machines)

SVM은 확률적인 방법으로 구분된 학습데이터들을 학습하여 최적의 인접벡터(Support Vectors)를 찾아내고 데이터를 구분할 하이퍼플레인(hyperplane)을 얻어내는 방법이다.[1][2] 예를 들어 (1)과 같은 학습데이터가 있을 때 그림 1과 같은 인접벡터(SVs)와 하이퍼플레인을 찾는 것이 SVM의 목적이라 하겠다.

$$(x_1, y_1), (x_2, y_2), \dots, (x_a, y_a)$$

$$x \in R^n, y \in \{-1, 1\}$$
(1)

그림 1은 선형인 하이퍼플레인으로 데이터를 구분할 수 있는 경우를 보여주고 있는데, 데이터가 비선형적인 특성을 갖는 경우에는 커널함수(kernel function)를 사용하여 비선형적인 특성을 갖는 데이터를 선형적인 특성을 갖는 고차원의 데이터로 변환한 후 그림 1과 같이 하이퍼플레인을 얻어낼 수 있다.[1][2]

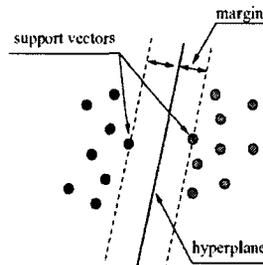


그림 1. 인접벡터(Support Vectors)와 하이퍼플레인(hyperplane)

2.2 인식 시스템

표정인식 시스템은 인식 방법에 따라 다양하게 나타낼 수 있지만, 대부분의 인식 시스템은 그림 2와 같이 얼굴 영역 검색, 정규화, 인식의 3단계로 구성된다. 여기서 정규화 과정은 다양한 크기나 자세를 가지는 얼굴 이미지를 일정한 크기의 얼굴 이미지로 변환하는 과정을 말

한다.



그림 2. 인식 시스템

일반적으로 인증을 위한 얼굴 인식 시스템에서는 두 눈의 위치만으로 얼굴을 정규화 하는데 반해 표정 인식 시스템에서는 두 눈의 위치만으로 정규화 하기 어렵다. 표정 인식 시스템에서는 얼굴의 특징 부위인 눈, 코, 입 등의 변화를 인식하는 것이 매우 중요하기 때문이다. 따라서 표정 인식 시스템에서는 각 부위의 위치가 정규화 하였을 때 일정한 위치에 놓이는 것이 인식률을 높일 수 있다. 그림 3에서는 검색된 얼굴 영역에서 두 눈의 위치와 입을 위치를 검출하고 이 정보를 바탕으로 정규화 이 이미지로 변화하는 예를 보여주고 있다.

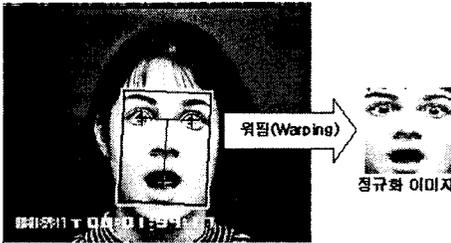


그림 3. 두 눈과 입의 위치를 이용한 얼굴 이미지의 정규화

2.3 얼굴 영역의 추출

잘못된 얼굴 영역은 전체 인식 결과를 저하 시키는 주요한 원인이 되기 때문에 얼굴 영역을 찾아내는 과정은 매우 중요하다. 특히 주위 조명 환경이나 얼굴의 다양한 변화에 매우 강인하게 대처해야 할 필요가 있다.

얼굴 영역 추출은 색도공간을 이용한 방법과 흑백이미지(gray scaled image)를 이용한 방법이 있다.[3] 색도공간을 이용한 방법은 얼굴 피부색을 색도공간에서 정의하고 이미지의 칼라 값에서 피부색을 찾아내어 얼굴의 영역을 추출하는 방법이다. 이 방법은 빠른 처리 능력을 가지고 있지만, 주위 조명 변화에 매우 민감하고, 조도가 낮은 환경에서는 사용할 수 없다는 단점을 가지고 있다. 반면, 흑백이미지를 이용하는 경우 처리속도가 느리지만 주위 조명 변화에 둔감하고, 낮은 조도에서도 사용할 수 있다. 여기서는 흑백이미지를 가지는 환경에서 SVM을 이용하여 얼굴의 영역을 추출하는 방법에 대해 서술하고자 한다.

SVM을 이용하여 얼굴 영역을 추출하기 위해서는 먼저 학습과정을 거쳐 하이퍼플레인 데이터를 만들어야 한다. 얼굴인 이미지와 얼굴이 아닌 이미지를 준비하고, 모든 이미지를 (1)과 같이 똑같은 크기를 가지는 벡터를 만들고, 학습하는 과정을 거치게 된다. 주위의 조도에 강인하게 하기 위하여 히스토그램 평준화(Histogram Equalization)를 전처리 과정으로 삽입하였다.

학습 할 이미지는 그림 5와 같이 얼굴인 이미지의 경우에 다양한 표정과 자세를 골고루 포함시키도록 인위적으로 선정하고, 얼굴이 아닌 이미지의 경우에 얼굴 영역이 없도록 인위적으로 선정한다.

학습이 끝나면 하이퍼플레인 데이터가 얻어지고, 얼굴 영역을 알아내기 위해서 그림 6(a)과 같이 입력된 전체

이미지에서 전 영역을 각각 학습한 이미지와 같은 크기의 블록(block)영역으로 떼어내고 이것을 하이퍼플레인 과 비교하여 얼굴인 부분과 얼굴이 아닌 부분을 그림 6(b)과 같이 구분(classification)해 낸다. 그림 6(b)에서 점으로 나타난 부분은 일정 격자(grid)간격으로 검색하였을 때 얼굴이라고 구분한 영역의 중심점을 표시한 것이다.

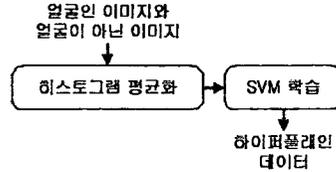


그림 4. SVM 학습 과정



그림 5. 얼굴에 대한 SVM 학습 이미지



그림 6. 얼굴 검색 방법과 검색 결과

그림 6(b)에 격자 간격으로 얻어진 얼굴이라고 구분된 영역의 중심점들은 라벨링(labeling)기법을 이용하여 중심점들을 그룹화(grouping) 하고 각 그룹의 중심점들의 평균값을 얻어낼 수 있다. 이 평균값은 얼굴 영역을 대표하는 값이라 할 수 있다. 이렇게 얻어진 평균값은 얼굴의 중심점이라 할 수 있고, 중심점으로 입체 영역을 정의 하던 이것을 얼굴 영역이라 할 수 있다.

2.4 얼굴 이미지의 정규화

앞 절에서는 얼굴 영역을 추출하는 방법에 대해 알아보았다. 이제 그림 2의 과정에 따라서 확보된 얼굴 영역에서 두 눈과 입의 위치를 구하고 정규화 된 이미지를 얻어 내는 과정을 알아보겠다.

얼굴 영역이 확보 되면 그림 7의 검색 영역 모델과 같이 각 부위 별로 검색 영역을 정의한다. 각 검색 영역에서 두 눈과 입의 위치를 구할 때는 2.3절의 SVM을 이용하여 얼굴의 중심점 위치를 구하는 방법과 같은 방법을 사용한다. 즉, 먼저 두 눈과 입의 학습이미지를 정의하고 SVM학습을 한 다음, 각 부위별 검색 영역에서 그림 6과 같은 방법으로 각 블록 영역별로 눈과 입의 중심점을 구하게 된다. 그림 8은 학습이미지의 일부분을 나타내고 있다.

두 눈과 입의 중심점을 구해내면 이것으로 그림 3과 같이 이미지를 워핑(Warping)하여 정규화 된 이미지를 구해 낸다. 두 눈과 입의 중심점에 대한 정확도는 인식 시스템의 인식률에 많은 영향을 미친다.



그림 7. 각 부위별 검색 영역 모델

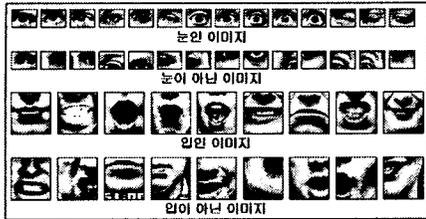


그림 8. 눈과 입의 SVM 학습 이미지

3. 실험

실험에 사용된 이미지는 모두 Kanade DB를 사용하였다. Kanade DB는 그림 9와 같이 다양한 표정을 가진 얼굴 이미지 DB이다.



그림 9. Kanade 얼굴 이미지 DB

실험은 입력된 이미지를 다운샘플링(down sampling)하여 160×120 픽셀의 이미지를 사용하였다. 얼굴의 학습 이미지는 40×45 픽셀의 이미지를 사용하였고, 눈의 학습 이미지는 20×15 픽셀, 입의 학습 이미지는 25×25 픽셀의 이미지를 사용하였다. 밝기의 차이에 대한 영향을 줄이기 위해 모든 이미지를 히스토그램 평균화 하였다.

검색 블록은 학습 이미지와 동일한 크기로 정의하였다. 얼굴 검색은 4픽셀의 그리드(grid) 단위로 검색하였고, 눈과 입은 2픽셀의 그리드 단위로 검색하였다.

얼굴, 눈, 입에 대한 각각의 학습 이미지로 학습하고, 600개의 이미지에 대해 정확도를 테스트하였다. 정확도는 ROC분석법(receiver operating characteristic Analysis)을 사용하여 나타내었다.

사용자가 미리 지정해 놓은 정확한 얼굴, 눈, 입의 위치와 테스트 결과의 얼굴, 눈, 입의 위치의 차이를 인식 오차라 하고, 이 인식 오차가 일정한 경계범위(threshold range)안에 들어오게 되면 검출되었다고 판단한다. 얼굴, 눈, 입이 검출된 이미지의 개수를 각각 총 테스트 이미지의 개수로 나누면 정확도가 된다. 여기서 경계범위를 변화해 가면서 인식율을 테스트하면 인식 결과를 ROC커브(receiver operating characteristic curve)로 나타낼 수 있는데, ROC커브는 검색의 정확도를 나타내는 지표가 된다. 그림 10는 얼굴인 학습 이미지의 개수와 얼굴이 아닌 학습 이미지가 200, 100, 50 일때 각각에 대한 ROC커브이고, 그림 11은 눈에 대한 각각의 ROC커브, 입에 대한 ROC커브, 눈과 입에 대한 ROC커브이다. ROC커브에서 가로축은 오차의

경계범위에서 검색 영역의 넓이를 나눈 값을 나타내었다.

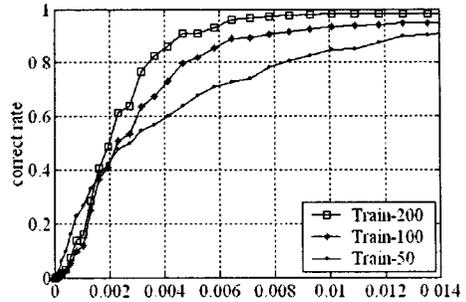


그림 10. 얼굴에 대한 ROC커브

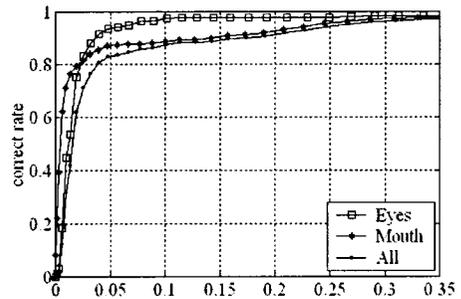


그림 11. 눈, 입에 대한 ROC커브

4. 결 론

본 논문에서는 표정인식 시스템에서 얼굴의 영역을 찾아내고, 얼굴의 특징 영역을 정규화 하는 방법에 대해 설명하였다. SVM을 사용하여 얼굴과 특징점을 찾아내는 방법으로 얼굴의 특징영역을 추출하였고, 실험을 통하여 높은 검색 성공률을 가지고 있음을 ROC커브를 통하여 나타내었다.

[참 고 문 헌]

- [1] Vladimir N. Vapnik, "An Overview of Statistical Learning Theory", IEEE Transactions on Neural Networks, VOL.10, NO. 5, pp.213~236, september 1999.
- [2] Hyeran Byun, Seong-Wan Lee, "Applications of Support Vector Machines for Pattern Recognition: A Survey", SVM 2002, LNCS 2388, pp.213~236, 2002
- [3] Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and machine intelligence, VOL.24, NO.1, page.34~58, January 2002.
- [4] Ebine, H., Shiga, Y., Ikeda, M., Nakamura, O., "The recognition of facial expressions with automatic detection of the reference face", Electrical and Computer Engineering, 2000 Canadian Conference on , Volume: 2, Page(s): 1091 -1099 vol.2, 2000.
- [5] Y. Tain, T.Kanade and J.Cohn, "Recognizing action units for facial expression analysis", IEEE PAMI, VOL.23, NO.2, page.97~115, 2001
- [6] T.Kanade, J.Cohn, and Y. Tian, "Comprehensive database for facial expression analysis", Proc. Int'l. Conf. Face and Gesture Recognition, page.46~53, 2000.