

심전도 및 맥파신호 기반의 감정인식 시스템에 관한 연구

홍윤정, 황윤경, 신동규, 김동현*, 신동일
세종대학교 컴퓨터공학과, 디지털콘텐츠학과*

e-mail : { alemidy, sian84 }@gce.sejong.ac.kr, { shindk, mustache*, dshin }@sejong.ac.kr

Research on the Emotion Recognition System based on Electrocardiograph and Pulse Signals

Yoonjung Hong, YunKyung Hwang, Dongkyoo Shin, Donghyun Kim*, Dongil Shin
Sejong University Computer Engineering, Digital Contents*

요 약

본 논문은 생체 신호들 중 데이터 획득이 간편한 심전도와 맥파를 실시간으로 취득하여 기계학습 기법인 SVM(Support Vector Machine)알고리즘과 클러스터링 기법인 k-NN(Nearest Neighbor)알고리즘을 적용한 인간의 감정을 분석하는 시스템에 대한 연구결과를 제시한다.

1. 서론

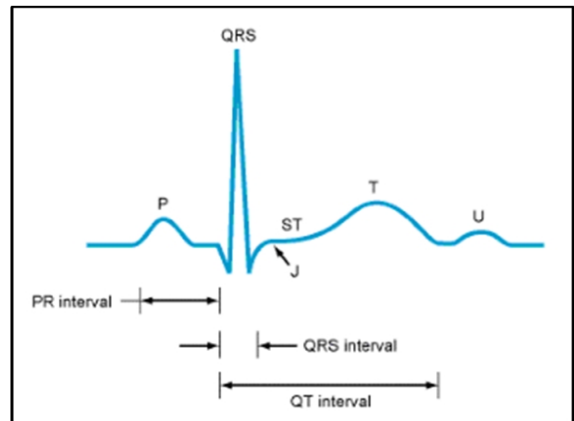
일상에서 감정은 음성, 얼굴 표정 그리고 생체 신호 등과 같은 다양한 방법으로 표현되며 인간 두뇌의 현 상태를 나타내는 기본적인, 필수적 그리고 결정적인 역할을 한다. 따라서 감정 상태나 태도를 파악하기 위해 인간의 음성, 얼굴표정, 손과 몸동작 그리고 생체 등을 통해 제공되는 다양한 신호를 처리, 추출하고 분석하는 것이 필요하다. 현재 감정인식 연구는 여러 분야에서 연구되고 있다. 특히 얼굴 표정이나 음성을 통해 감정인식을 하는 연구가 많이 진행되고 있지만 주변 환경에 영향을 많이 받아 문제가 많다. 반면에 생체 데이터의 경우는 특정한 감정 분류가 어렵고, 작은 움직임에도 변화가 크지만 주변 환경에 영향을 덜 받고 전극, 적외선 센서 등을 이용하여 간단하게 추출하며 실시간으로 감정변화에 대한 정보를 얻을 수 있다. 또한 사용자에 따른 사회적, 문화적인 차이에 덜 민감한 장점들이 있다 [1]. 생체 신호 처리 기술은 각종 재활분야, 건강검진 등의 의료 분야에도 응용될 수 있어 앞으로의 그 활용성 및 상용성에 무한한 잠재력을 가지고 있다 [3]. 하지만 너무 많은 생체 신호를 추출하여 감정을 인식하는 경우는 여러 센서들을 사용자에게 장착시켜야 하기 때문에 불편함을 줄 수 있다. 따라서 우리는 사용자에게 센서 장착의 불편함을 덜어 주면서 주변 환경에 영향을 덜 받는 심전도와 맥파를 이용한 감정인식 시스템에 관하여 연구하였다.

2. 기본개념

2.1 심전도

심전도란 신체 내에 흐르는 전기, 즉 심박동과 함께 발생하는 전위차를 적당한 부위에서 일정한 방법

으로 유도 증폭하여 심전계(electrocardiograph)에 의해 그림으로 나타낸 것으로 ECG(electrocardiogram) 또는 EKG 로 약기한다 [2].



(그림 1) 심전도의 기본 파형

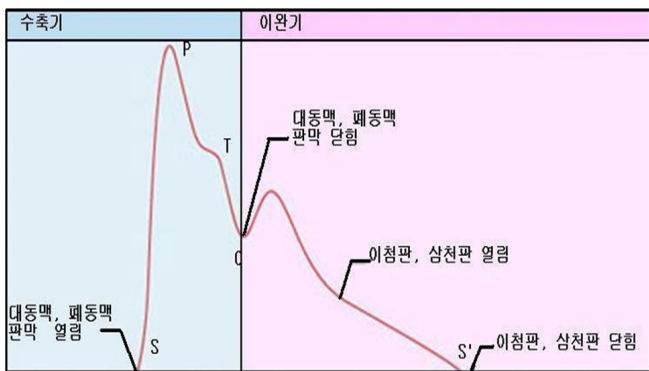
심장의 전기적 활성화 단계는 크게 심방 탈분극, 심실 탈분극, 심실 재분극 시기로 나뉘며, 이러한 각 단계는 P, QRS, T 파라고 불리는 몇 개의 파의 형태로 반영된다. P 파는 심방의 탈분극기로서 폭은 0.11 초 이하이고 높이는 0.25mV 이하의 부드러운 곡선이며, QRS 파는 심실의 탈분극기로서 폭은 0.10 초 이하이고 예리한 파를 이루는데 보통 P 파나 T 파보다 크다. T 파는 심실의 재분극기로서 폭과 높이는 일정하지 않으나, P 파보다 크고 부드러운 곡선이고 약 0.2 초간 지속된다.

2.2 맥파

맥파(pulse)란 심장 박동에 의해 생성되는 흥벽 및

동맥의 박동을 파형으로 기록한 것을 말한다. 혈관의 상태나 심장의 상태에 따라 민감하게 반응하는 특징이 있다. 맥과 신호는 심장에서 발생된 박동에 따른 혈압의 변화가 파를 이루며 혈관을 통해 전달되어 그 영향이 체표에 도달하게 되는 데 이 변화를 관혈, 비관혈적으로 측정하는 것이다.

맥파의 분석은 각 맥파 형을 찾아 각 파형의 위치와 최대 점의 위치 및 크기를 찾아내어 이로부터 임상적 의미를 찾아내는 것이다. 앞서 언급한 것처럼 맥과 신호는 심혈관계의 물리적 요인에 기인한 결과로 생성되므로 심전도보다 복잡하고 다양한 파형을 갖는다. 그러므로 각 신호의 파형 특성을 나타내주는 특징점 추출이 선행되어야 한다.

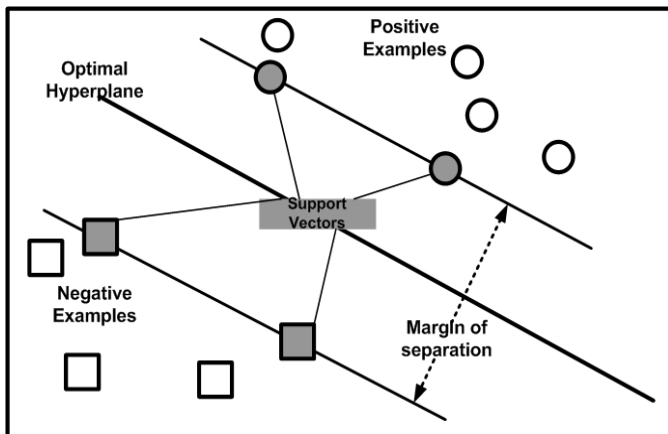


(그림 2) 맥의 파형

2.3 감정 상태 분석을 위한 패턴인식 방법

2.3.1 SVM (Support Vector Machine)

SVM은 1979년 Vapnik에 의해 제안된 통계적 학습 이론(SLT; Statistical Learning Theory)에 기반을 두고 있다. 일반적인 통계 학습 방법에서의 경험적 리스크 최소화(ERM; Empirical Risk Minimize)에 기반한 Nearest-Neighbor과 Neural Network와 같은 이전의 비모수적 기술들과는 다르게, SVM은 일반화 에러에서 상한을 최소화 하는 구조적 리스크 최소화(SRM; Structural Risk Minimize)를 통해 오류를 최소화 시키는 방법을 이용한다 [4].



(그림 3) 선형(Linear) SVM

Vapnik과 Chervonenkis는 통계적 학습 이론인 모델 복잡도의 평균을 측정하는 VC(Vapnik Chervonenkis) dimension을 소개했다. 또한 다음의 학습 모델 경계 조건을 제시했다.

$$R(w) \leq R_{emp}(w) + \sqrt{\frac{h(1/n \sum_{i=1}^n |d_i - F(x_i, w)|) - 1/n \frac{\eta}{4}}{I}}, \forall w \in \mathcal{W}$$

where

$$R(w) = \int |d - F(x, w)| dF_{x, D}(x, d)$$

$$R_{emp}(w) = \frac{1}{I} \sum_{i=1}^I |d_i - F(x_i, w)|$$

여기서 h 는 VC dimension 이고, I 은 학습 데이터의 수이다. 그리고 $R(w)$ 는 기대 위험이며 $R_{emp}(w)$ 는 경험적 리스크이다.

2.3.2 k-NN (k-Nearest Neighbor)

Nearest Neighbor는 분류하고자 하는 class의 종류에 대해서는 알고 있지만 sample 들 각각에 대한 확률밀도함수(probability density function)를 알지 못하는 상태에서 사용한다. 굳이 각 sample에 대한 확률적 파라미터들을 구하지 않고 sample의 값을 그대로 좌표에 표시하여 reference set에서 가장 유사(similar)하거나 거리상으로 가까운(nearest) class에 속하는 것으로 분류하는 방법이다.

k-NN이라는 명칭은 k개의 최 근접하는 이웃을 이용한다는 의미에서 붙여졌다. 이 방법은 학습 데이터 집합에 있는 표본들 간 유사도에 따라 라벨이 붙여져 있지 않은 표본들을 분류하는 매우 직관적인 방법이라 할 수 있다. 새로운 데이터의 가장 가까운 이웃 중에 k개를 뽑아서 그 중에 많은 범주로 새로운 데이터를 분류한다.

Nearest의 의미는 smallest Euclidean distance, absolute difference, maximum distance, Minkowski distance 등으로 나눌 수 있다.

• Euclidean distance

n-차원의 feature space로 구성된 경우라 하면 두 개의 포인트 $a = (a_1, \dots, a_n)$ 와 $b = (b_1, \dots, b_n)$ 사이의 기하학적 거리는 다음과 같이 구해질 수 있다. 이것은 피타고라스 정리를 n-차원으로 확장한 것이다.

$$d_e(a, b) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2}$$

이 방법은 가장 흔하게 사용되는 거리 측정 방법이지만 항상 최고의 방법만은 아니다. 각 차원에서 합계되기 전에 제곱을 하기 때문에 계산이 복잡하고 차이점(dissimilarity)이 큰 경우가 강조될 수 있다.

• Absolute difference

두 포인트의 차이를 그대로 표현하여 계산하기가 쉽다. City block distance, Manhattan metric, Taxi-cab distance 라고도 표현된다.

$$d_{abs}(a, b) = \sum_{i=1}^n |b_i - a_i|$$

• maximum distance

feature 들 중에서 가장 유사하지 않은 (거리가 많이 떨어진) 부분이 강조되는 것이다.

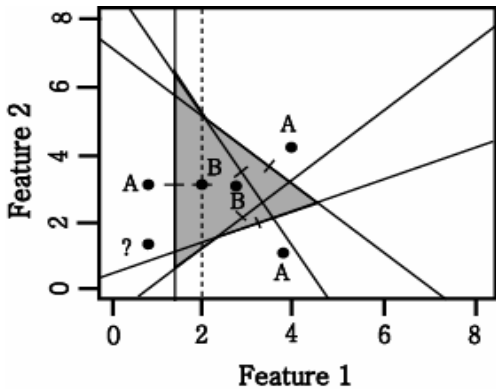
• Minkowski distance

거리를 측정하는 위의 방법들을 종합한 방법이다. 여기서 r 은 조정 가능한 파라미터로서 그 값이 1 이면 absolute difference 와 같고, 2 이면 Euclidean distance 와 같다.

$$d_r(a, b) = \left[\sum_{i=1}^n |b_i - a_i|^r \right]^{\frac{1}{r}}$$

그림 4 는 class A 에서 3 개, class B 에서 2 개의 sample 을 가지는 경우의 feature space 이다. 어떤 class 에 속하는지 알려지지 않은 sample 이 (1,1) 좌표에 있을 경우 Euclidean distance 의 방법을 써서 가장 가까이 있는 class 는 (1,3)에 위치한 class A 이다. 따라서 class A 에 속하는 것으로 한다.

$$d_m(a, b) = \max_{i=1}^n |b_i - a_i|$$



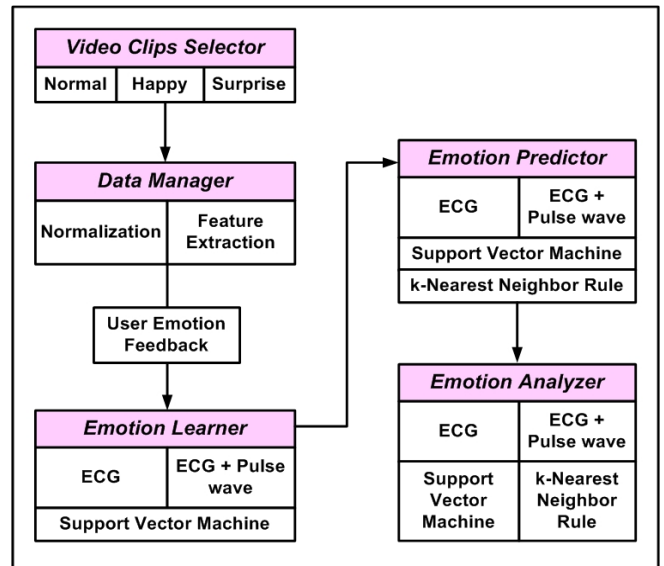
(그림 4) Nearest Neighbor 규칙의 예

3. 감정인식 시스템의 설계 및 구성

3.1 감정인식 시스템의 설계

그림 5 는 감정인식 시스템의 전체 구조를 보여준다. Video Clips Selector 는 감정 인식 실험을 위한 준비 단계로서 실험에 참가하지 않는 사람들을 대상으로, 실험에 사용할 비디오 클립들을 선택한다. Data Manager 는 사용자가 Video Clips Selector 에서 선택된 비디오 클립을 시청하는 동안 심전도 데이터와 맥파 데이터를 측정하고, 정규화 한다. 정규화 된 데이터를 통해 특징점을 구하고, 이 특징점의 값들을 데이터베이스에 저장한다. 이 때 사용자로부터 자신이 실제 느낀 감정을 선택하게

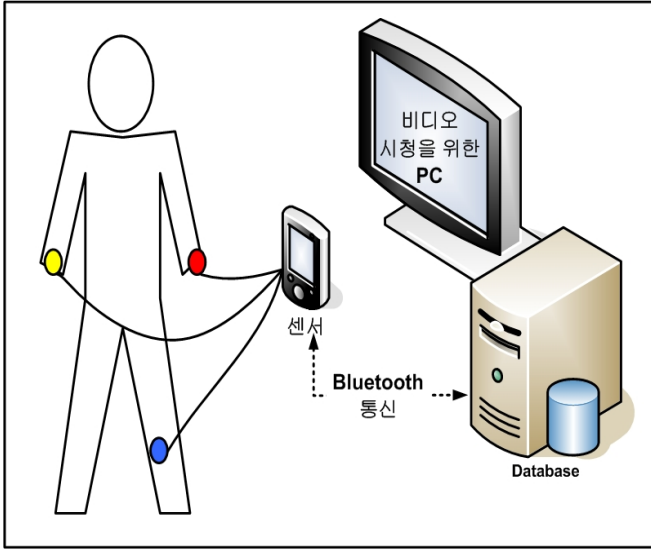
함으로써 feedback 을 얻는다. Emotion Learner 는 학습을 담당하는데, 학습 단계는 두 단계로 이루어진다. 먼저 데이터베이스에 저장된 특징점의 값들 중 심전도 데이터와 사용자의 feedback 으로부터 얻은 감정을 학습하고, 다음으로 심전도 데이터와 맥파 데이터, 그리고 사용자의 feedback 감정을 함께 학습한다. 학습에 사용된 알고리즘은 교사 학습(Supervised Learning) 방법인 SVM 과 클러스터링 방법인 k-NN 이다. Emotion Predictor 는 학습 단계에서 사용했던 비디오 클립을 제외한 다른 비디오 클립들을 사용자에게 시청하게 하고, 학습 단계에서와 마찬가지로 Data Manager 의 데이터 정규화 과정 통해 추출된 특징점의 값들을 데이터베이스에 저장한다. 데이터베이스에 저장된 특징점의 값들 중 심전도 데이터만을 SVM 과 k-NN 에 적용하고, 또한 심전도 데이터와 맥파 데이터를 함께 SVM 과 k-NN 에 적용하여 감정을 예측한다. Emotion Analyzer 는 예측 단계에서의 감정 예측률을 비교 분석한다.



(그림 5) 감정인식 시스템의 전체구조

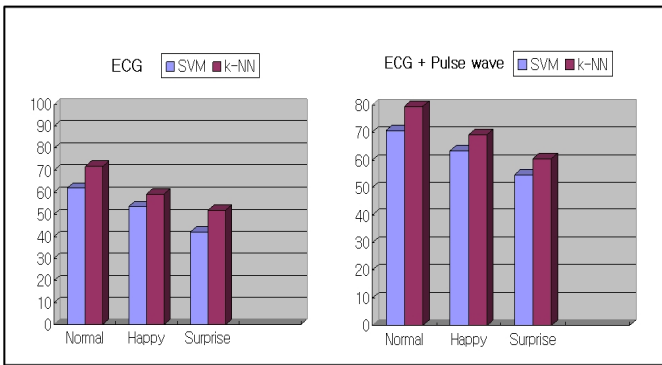
3.2 감정인식 시스템을 이용한 실험 방법

감정 인식 시스템을 이용한 전체 실험 시스템은 그림 6 과 같이 구성된다. 사용자는 생체 센서 디바이스로부터 심전도 및 맥파를 획득하기 위해 양쪽 팔목과 오른쪽 발목에 태그를 부착한다. 생체 센서 디바이스는 감정인식 에이전트와 블루투스를 이용하여 데이터를 전송하며, 감정인식 에이전트는 전송받는 모든 데이터의 패킷을 분석한다. 사용자의 감정을 분석할 때, 감정인식 에이전트는 사용자에게 감정을 유발시키기 위하여 감정유발 콘텐츠를(비디오 데이터)를 제공하며, 감정유발 콘텐츠를 감상하는 동안 사용자로부터 특정 감정(Normal, Happy, Surprise)이 발생했다는 피드백이 전달되면, 감정인식 에이전트는 그 순간의 생체 신호를 생체 센서로부터 전송을 요청하고, 데이터베이스에 저장한다



(그림 6) 감정인식을 위한 실험방법

4. 결론



(그림 7) 감정인식 시스템의 실험결과

본 논문은 생체 신호들 중 다른 생체 신호들 보다 많은 장치가 필요 없어 데이터 획득이 간편한 심전도와 맥파를 이용하여 인간의 감정을 분석하는 시스템을 제시하였다. 실험은 건강에 별 다른 이상이 없는 20 대 남녀 대학생 10 명을 대상으로 하였고, 사용자의 감정 분석을 위하여 생체 신호(심전도, 맥파)를 실시간으로 획득하였다. 실험 대상자들은 생체 신호를 측정하기 위해 양쪽 손목과 왼쪽 다리에 센서를 부착하고 다양한 종류(사용자가 느낄 수 있는 감정 - Normal, Happy, Surprise)의 비디오를 감상하면서 그에 따라 발생하는 생체 신호의 패턴을 분석하였다. 비디오를 시청하는 동안 센서로부터 생체 신호를 획득하여 추출된 특징 점(심전도: R 파와 R-R 간격, 맥파: P 파와 P-P 간격)의 값을 데이터베이스에 저장하였다. 이 때 학습을 위해 사용자가 비디오를 감상이 끝난 후 자신이 실제 느낀 감정을 선택함으로써 감정을 피드백하였다. 예측 단계에서도 마찬가지로 비디오를 시청하는 동안 발생하는 생체 신호의 패턴과 비교하여 감정을 인식하였다. 생체 신호의 패턴 분석은 기계학습 기법인 SVM(Support Vector Machine) 알고리즘과

클러스터링 기법인 k-NN(Nearest Neighbor) 알고리즘을 이용하여 패턴 분석의 모델을 확립하며, 심전도 데이터만을 이용하였을 경우와 심전도 데이터와 맥파 데이터를 함께 이용하였을 경우의 인식률을 비교, 분석하였다. 두 알고리즘을 비교한 결과 k-NN 이 SVM 보다 다소 나은 결과를 보였으며, 단일 생체 신호를 이용하였을 경우보다 다중 생체 신호를 적용하였을 때 감정 인식률이 더 뛰어났다.

참고문헌

[1] Kollias, S., Karpouzis, K., "Multimodal Emotion Recognition and Expressivity Analysis," Multimedia and Expo, 2005. ICME 2005, IEEE International Conference on, pp. 779-783, July 2005
 [2] 이승한, "ECG 신호의 R 파 검출 및 PVC 판별 알고리즘", 명지대학교, 2001
 [3] 김종성, 김홍기, 정혁, 김기홍, 임선희, 손옥호, "생체신호 기반 사용자 인터페이스 기술", 전자통신동향분석, 제 20 권, 제 4 호, pp. 67, 2005
 [4] Vapnik V. N. "The nature of statistical learning theory", NewYork, Springer-Verlag, 1995