

다중 SVM 알고리즘을 이용한 스트레스 지수에 따른 생체 감성 인식에 관한 연구

김 태 연* · 서 대 웅* · 배 상 현**

The Study of Bio Emotion Cognition follow Stress Index Number by Multiplex SVM Algorithm

Tae-Yeun Kim · Dae-Woong Seo · Sang-Hyun Bae

요 약

본 논문은 사용자의 생체 정보(맥박, 이완기 혈압, 수축기 혈압, 혈당)를 무선 센서들을 통하여 획득한 후 스트레스 지수에 따른 감성을 인식하여 대응되는 컬러와 음원을 분류하는 시스템으로서, 맥박 센서, 혈압 센서, 혈당 센서 등의 입력치를 받아 데이터베이스에 저장한 후 다중 SVM(Support Vector Machine) 알고리즘을 이용하여 스트레스 지수에 따른 감성을 분류한다. 2,000개의 데이터 집합을 사용하여 다중 SVM 알고리즘을 학습한 결과 약 87.7%의 정확도를 가졌다.

ABSTRACT

In this paper, it's a system which recognize the user's emotions after obtaining the biological informations(pulse sensor, blood pressure sensor, blood sugar sensor etc.) about user's bio informations through wireless sensors in accordance of previously collected informations about user's stress index and classification the Colors & Music. This system collects the inputs, saves in the database and finally, classifies emotions according to the stress quotient by using multiple SVM(Support Vector Machine) algorithm. The experiment of multiple SVM algorithm was conducted by using 2,000 data sets. The experiment has approximately 87.7% accuracy.

Keywords : bio emotion recognition, stress index, support vector machine, emotion modeling, biological informations

1. 서 론

최근 과학기술의 발달로 인간의 생활과 사고방

식이 변하고 있으며, 대부분의 시스템이 자동화되고 인간과의 의사소통방법을 요구하고 있다. 이 점에서 필요로 하게 된 기술 중 하나가 사용자와

* 조선대학교 컴퓨터통계학과

** 교신저자 조선대학교 컴퓨터통계학과 (shbae@chosun.ac.kr)

접수일자 : 2012년 1월 12일, 수정일자 : 2012년 1월 25일, 심사완료일자 : 2012년 1월 31일

의 감성교류를 위한 감성정보처리기술이다[1-3]. 감성 인식 기술은 감성정보(emotional information)를 통하여 사용자를 인식하고 각종 표정이나 몸짓, 동작 등에서 정보 추출을 통하여 적절한 행동을 취할 수 있는 지능형 의사결정의 한 방법이다. 이와 같이 컴퓨터가 학습과 적응을 통하여 인간의 감성을 처리할 수 있는 감성인지 능력을 갖는 것은 보다 효율적인 인간과 컴퓨터의 상호 작용을 가능하게 한다. 감성정보 중 시각과 청각 정보인 색상과 음악은 짧은 시간에 형성되고 기억에 오랫동안 지속되기 때문에 인간의 정서를 이해하고 해석하는데 있어서 매우 중요한 역할을 한다[4].

본 논문에서는 개인의 감성에 따라 나타나는 반응을 학습하고 패턴화 함으로써 사용자의 감성을 인식하고 스트레스 지수와 매칭하는 생체 감성을 분류하고자 한다[5].

이에 본 논문에서는 감성 분류를 위해 다수의 센서(맥박, 혈압, 혈당)를 이용하여 생체 데이터를 획득 한 후 다중 SVM 알고리즘을 이용하여 스트레스 지수와 매칭된 감성정보를 분류한다. 일반적으로 의사결정지원 시스템에 대해 적용할 수 있는 기존의 다양한 학습 알고리즘이 있지만 본 논문에서 이용한 데이터는 맥박, 혈압, 혈당, 스트레스 지수 데이터로써 비선형 데이터 구조로 이루어져 있다.

데이터가 비선형 데이터인 경우 판별 문제를 해결하기 위해 다층 퍼셉트론 구조로 비선형 판별 문제를 해결할 수 있는 다중 SVM 알고리즘을 이용한다.

다중 SVM 알고리즘은 다른 신경회로망에 비해서 결과 해석이 용이하고, 적은 학습 자료만으로 신속하게 분별 학습을 수행할 수 있으며, 인공 신경망과 비슷한 수준의 높은 예측력을 나타낼 뿐만 아니라 인공 신경망의 한계점으로 지적되었던 과대적합, 극소 최적화와 같은 한계점들을 완화할 수 있다.

분류된 스트레스 감성은 감성컬러와 감성음원과 매칭한 후 그 컬러와 음원 값에 해당하는 데이터를 분류한다.

본 논문은 다음과 같이 구성되었다. 2장에서는 시스템 구성 및 설계, 3장은 성능 평가 및 실험결과에 대해 기술한다. 4장에서는 결론과 향후 연구

방향을 제시한다.

II. 시스템 구성 및 설계

본 논문에서는 감성 분류를 위해 다수의 센서(맥박, 혈압, 혈당)를 이용하여 생체 데이터를 획득한 후 다중 SVM 알고리즘을 이용하여 스트레스 지수와 매칭된 감성정보를 분류한다.

분류된 감성은 HP의 ‘The Meaning of Color’에서 정해놓은 20개의 컬러 감성 모델에 따라 분류된 감성에 대응하는 색상 값으로 분류되며, 감성 음악은 감정과 분위기가 다르고 기분에 따라 틀려지기 때문에 음악치료의 자료에 근거를 두어 정리하여 클래식 곡들을 수집, 분류하여 활용하였다. 음악에 관한 내용은 “삼성 이데아”에 의해 제시된 음악치료에 관한 근거로 수집된 자료를 토대로 구성하였다[6][7].

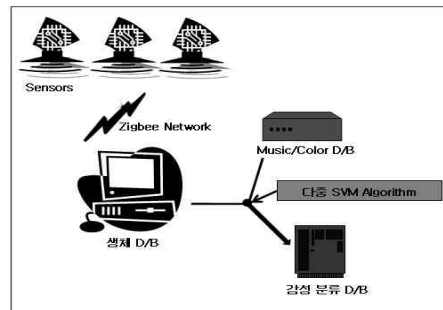


그림 1. 시스템 구성도
Fig. 1. System configuration

그림 1은 센서를 통해 획득된 생체 데이터를 이용하여 스트레스 지수에 따라 감성을 분류하는 시스템의 구성도이다.

1. 생체 데이터 측정

본 논문에서는 사용자의 생체 신호를 측정하기 위해 Zigbee 무선 센서 네트워크 시스템을 구현하였다. IEEE 802.15.4 Zigbee는 저속, 저가, 저 전력 소모를 필요로 하는 응용에 주안점을 둔 근거리 무선 통신 기술이다.

본 실험에서는 맥박, 혈압, 혈당 센서가 통합된

센서 모듈을 사용하며 프로세서 보드는 Telos 플랫폼 계열을 사용하였으며 MSP430의 MCU와 CC2420 Radio Chip을 사용하여 측정하였다.

맥박, 혈압, 혈당 값을 각각 하나의 패킷으로 만든다면 추가적인 트래픽의 발생과 데이터 전송에 따른 에너지 소모가 일어나기 때문에 하나의 패킷으로 묶어서 데이터베이스에 전송하고 센서로부터 5개의 입력 데이터(맥박, 수축기혈압, 이완기혈압, 혈당(공복수치), 혈당(2시간 후 수치))가 전송되도록 하였다.

그림 2는 센싱된 생체 데이터의 구조이다. MSG 타입은 생체 데이터의 타입을 맥박, 수축기 혈압, 이완기 혈압, 혈당 중 형태가 어떤 것인지를 알려준다. GroupID는 센서의 정보를 나타내며 한 센서당 하나의 GroupID를 갖게 된다. timestamp는 센서에서 데이터를 측정한 시간이다. reading은 실제 들어온 데이터의 값을 16진수 2byte로 표현한다.

7E 00 0A 7D 10 00 00 02 00 00 00 01 00 02 00 EE D3 FF FF 55 00										
1	2	3	4	5	6	7	8	9	10	11
7E 00	0A	7D	10	00 00	02 00	00 00	01 00	02 00	EE D3 FF FF	55 00
1:Address 2:MSG Type 3:GroupID 4:Data Length 5:Source address 6:Origin address 7:Sequence number 8:Hop Count 9:address 10:timestamp 11:reading										

그림 2. 센싱된 데이터 구조
Fig. 2. Sensed data structure

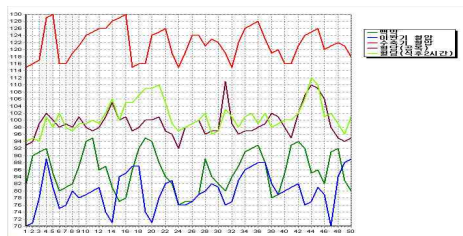


그림 3. 생체 데이터 처리
Fig. 3. Process of bio data

그림 3은 맥박, 수축기혈압, 이완기혈압, 혈당(공복수치), 혈당(식후2시간)의 측정 데이터 그래프이다.

2. 제안한 다중 SVM 알고리즘

생체 데이터를 스트레스 지수로 매칭하기 위해서 기존의 다양한 학습 알고리즘이 있지만 본 실

험에서 이용한 데이터는 맥박, 혈압, 혈당 데이터로써 비선형 데이터 구조로 이루어져 있어서 다층 퍼셉트론 구조로 비선형 판별 문제를 해결할 수 있는 다중 SVM 알고리즘을 이용한다. SVM 분류는 두 그룹을 잘 분리시키는 분류 초평면을 찾는 방법이다[8].

SVM은 기존의 선형 분류방법보다 확장성이 좋고 학습 시마다 성능이 달라지는 신경망 분류방법과는 달리 항상 일정하게 우수한 성능을 보여준다[9]. SVM의 기본 원리는 선형 분리가 가능한 문제에서부터 출발한다. d-차원에서 입력데이터 X_i 가 주어졌을 때 학습데이터의 출력으로 -1과 +1처럼 이진 값으로 구분되는 문제를 고려한다.

두 집합을 분류하기 위한 모델을 정의하기 위하여 그림 4와 같은 선형 식별함수인 초평면(hyperplane)을 정의할 수 있다. 여기에서 Support Vector란 분류 규칙을 결정짓는 경계와 밀접한 연관이 있는 표본을 의미한다.

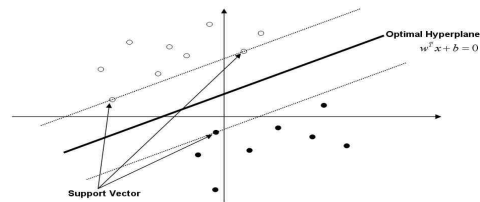


그림 4. 최적화 초평면과 서포트 벡터
Fig. 4. Optimal hyperplane and support vector

본 실험 데이터처럼 선형 분리가 불가능한 데이터인 경우에는 비선형 사상 ϕ 를 이용하여 입력 벡터의 차원보다 높은 선형분류가 가능한 차원으로 변형한 후 선형 분류를 하게 된다.

비선형 사상은 kernel 함수를 이용하여 N차원의 입력공간의 데이터를 고차원의 특징 공간(Q차원)으로 변환함으로써 선형적으로 구별할 수 있으며 식(1)은 kernel 함수와 결정함수이다.

$$K(x, y) = \phi(x) \cdot \phi(y)$$

$$f(x) = \sum_{i=1}^n a_i y_i K(x, x_i) + b \quad (1)$$

SVM은 단순히 분류 평면을 찾는대거나 표본 에러를 최소화하는 작업을 하는 것이 아니라 분류 여백(Margin)을 최대화함으로써 새로운 데이터에 대해 분류 정확도를 높인다.

SVM은 이진 분류를 위해 개발되었기 때문에 실제 환경에서 여러 클래스를 가지는 문제들을 해결하기에는 많은 어려움이 있다. 때문에 이러한 문제점들을 해결하기 위해 One-against-all 기법과 One-against-one 기법이 제시 되었다.

One-against-one 기법은 k개의 클래스가 입력 되었을 때 $k(k-1)/2$ 개의 SVM으로 구성되며 각각의 학습데이터는 두 개의 소속을 나타내는 데이터로만 구성되고 각 학습에 사용되는 학습 데이터의 수가 적기 때문에 학습 속도가 빠른 것으로 나타났다.

본 논문에서는 학습의 성능 향상을 위해서 One-against-one 기법을 이용하여 실험을 수행하였으며, 그림 5와 같이 SVM 알고리즘을 구성하였다.

Algorithm: SVM
 학습을 위한 데이터의 개수 : N
 Inputs: sample x to classify 데이터 셋 : I_i
 I_{i1} : 온도, I_{i2} : 조도, I_{i3} : 습도, I_{i4} : 음성
 Output: decision $y \in \{-1, 1\}$
 Classify using SVM, get the result in the form of a real number.

그림 5. 알고리즘 구성
 Fig. 5. Algorithm configuration

제안한 다중 SVM 알고리즘은 주어진 트레이닝 데이터의 각 Feature에 대해 최대 여분(margin)이 많이 생성되는 hyper-plane을 생성하며, 테스트 단계에서는 트레이닝 단계에서 생성된 hyperplane에 의해 분할된 다차원 공간에 매핑하여 새로운 데이터를 분류한다.

3. 감성 컬러 · 음원 매칭

감성은 외부 환경에 따라 표현될 수 있는 감성

의 종류가 다양함으로, 사용할 감성 컬러와 음원은 미리 정의하여 사용하는 것이 효과적이다. 본 논문에서는 감성컬러의 분류를 위해 HP의 'The Meaning of Color'에서 정해놓은 20개의 컬러 감성 모델을 대표 요소로 선정하였다.

HP사의 컬러 테이블과 본 논문에서 조사한 감성 어휘 매칭의 결과를 보면, 공통된 감성어휘를 찾을 수 있는데, 이들 공통된 감성 어휘를 바탕으로 표 1과 같이 분류할 수 있었다.

표 1. 공통된 감성 어휘 분류
 Table 1. Duplicated emotion word classification

감성어휘	컬러	감성어휘	컬러
순수	light blue	고요	blue
자연스러움	bright red	위험	bright red
우아함	burgundy	활력	bright yellow
따뜻함	beige, orange, terra-cotta	강함	navy
고전	beige, olive green, neutral gray	안정	blue, green, brown, terra-cotta
희망	green, bright yellow	공격	bright red
젊음	bright yellow	고유	teal blue
균형	orange	신비	purple
부드러움	light blue, light pink, beige	현명	fuchsia
안락	brown	부드러움	light pink

감성을 이해하고 파악하기 위해 중요한 감성정보의 하나인 컬러와 음은 파동이라는 공통된 특성을 갖는다. 따라서 파동이란 색과 음을 연결하기 위한 근본적인 실마리를 제공한다. 파장과 진동수는 물리, 수학적으로 역비례 관계에 있어 수학적으로 상호 변환이 가능한 물리량이다. 도를 기준으로 미와 솔의 파장 관계는 4/5, 2/3이 되고, 파장 비율은 1

: 4/5 : 2/3이 된다. 이 비율은 삼원색 빨강색, 초록색, 파랑색의 각 파장 650nm, 520nm, 433nm의 파장 비율과 일치한다. 도, 미, 솔은 빨강색, 초록색, 파랑색 빛을 적절히 섞어 수없이 다양한 색을 만들어 낼 수 있는 컬러의 삼원색과 많이 닮아 있다. 따라서 평균율에 의한 12음계의 파장 비율을 삼원색의 조합으로 만들 수 있는 컬러들과의 주파수를 순차적으로 대응시키면 음원과 컬러를 연결할 수 있다.

그림 6은 1 : 8/9 : 4/5 : 2/3 : 3/5 : 8/15 : 1/2 이 되는 ‘도레미파솔라시도(한 옥타브 위의 도 포함)’의 순수파장 비율의 비 균일성을 보완하여 완성된 평균율에 의한 12음계의 각 음계 간 파장비율은 1 : 1.0594가 된다. 색상환을 이루는 12지점의 색상 주파수는 12평균율의 파장비율과 똑같이 맞추어진 것을 볼 수 있다.

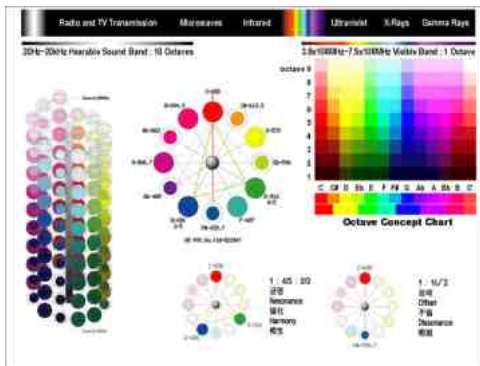


그림 6. 음파 컬러의 변환
Fig. 6. Conversion of music and color.

이렇듯 분석된 컬러 측정값과 지표로 사용된 음원의 음계를 분석 후 매칭하여 추론된 데이터를 분석하여 최적의 음원 매칭 리스트를 결정하게 되며 ‘삼성 이데아’에 의해 제시된 음악치료에 관한 근거로 수집된 자료를 토대로 구성하고 스트레스 지수에 따른 5가지의 감성을 분류하였다.

표 2는 대응하는 감성컬러와 감성음원으로써 사용자의 스트레스를 변화시키기 위한 것이다. 스트레스 지수가 0~30 즉, 1단계인 경우 해당하는 감성은 ‘지침’이고 대응하는 감성컬러는 ‘red’이며, 감성음원은 ‘비발디 사계-봄 외 10곡’이다.

표 2. 대응 감성컬러와 감성음원
Table 2. Correspondence emotion color & emotion music

Stress 지수	감성	대응 감성컬러	대응 감성음악
00-30 [1단계]	지침	red	비발디 사계 - 봄 외 10곡
31-40 [2단계]	우울 (체력저하)	yellow	모차르트 협주곡 1악장 - 알레그로 외 10곡
41-60 [3단계]	차분 (정상)	cerulean	슈베르트 - 자장가 외 10곡
61-70 [4단계]	과잉 (면역력감소)	blue	슈베르트 - 아베마리아 외 10곡
71-100 [5단계]	흥분	green	슈만 - 꿈 외 10곡

III. 성능 평가 및 실험결과

본 논문에서는 3개의 무선 센서(맥박, 혈압, 혈당)를 배치하여 생체를 인식하는 실험을 진행하였다.

센서로부터 획득한 2,000개의 맥박, 이완기 혈압, 수축기 혈압, 혈당(공복수치), 혈당(식후 2시간) 등의 생체 데이터를 이용하여 스트레스 지수와 매칭 후 감성을 분류하였다.

생체 인식에 대한 스트레스 지수 데이터 분류 기준은 표 3과 같이 스트레스 상태에 따라 5가지 상태로 분류하였다.

실험 결과에 대해 학습 데이터로부터의 영향을 최소화하고 신뢰성을 확보하기 위해 10-Fold 교차검증을 수행하였다. 실험 데이터는 평가용 데이터

표 3. 스트레스 상태 분류 기준
Table 3. Classification standards of stress

No	스트레스 상태	맥박	이완기 혈압	수축기 혈압	혈당 (공복)	혈당 (식후2시간)
		단위:수	단위:수	단위:수	mmhg	mmhg
1	1단계 (지침)	60-75	60-70	100-115	70-80	70-80
2	2단계 (우울)	76-90	71-80	116-130	81-90	81-90
3	3단계 (차분)	91-140	81-90	131-149	91-110	91-110
4	4단계 (과잉)	141-180	91-100	150-180	111-120	111-120
5	5단계 (흥분)	180이상	100이상	180이상	120이상	120이상

