

# 소셜 네트워킹을 위한 감성-기반 인터페이스의 설계 및 구현

이민기\*, 이기영\*\*

\*\*\*을지대학교 의료IT마케팅학과

e-mail: ohy2ah@hanmail.net

## Design and Implementation of Emotion-based Interface for Social Networking

Min-Ki Lee\*, Ki-Young Lee\*\*

\*\*\*Department of Medical IT and Marketing, Eulji University

### 요 약

최근 세계적으로 소셜 네트워킹에 대한 관심과 스마트폰의 기술이 발전함에 따라 사용자를 위한 다양한 인터페이스가 요구되고 있다. 항상 사용자의 손이 닿는 곳, 바로 옆에 위치한 모바일 단말기에서 다양한 센서를 집약하여 감정을 인식하고 인식된 정보와 위치 정보를 결합하여 세계적으로 관심이 높은 소셜 네트워킹의 인터페이스를 설계하고 구현하였다.

### 1. 서론

최근 컴퓨터 네트워크 기술의 급격한 발전과 정보화 시대가 도래 하면서 사용자들은 특정 사물에 대한 정보나 사용자의 위치와 같은 정보를 줄 수 있는 여러 디바이스와 콘텐츠들로 인해 다양한 서비스를 제공 받고 있다.

세계적으로 소셜 네트워킹에 대한 관심이 매우 높아지고 있으며, 다양한 사용자간의 네트워크가 형성되고 있다. 하지만 소셜 네트워킹은 웹에서 가장 많이 이루어지고 있고, 현실에서는 매우 한정적이기 때문에 현실에서의 밀접한 소셜 네트워킹을 위하여, 스마트폰과 같은 휴대용 디바이스의 다양한 인터페이스가 요구되고 있다.

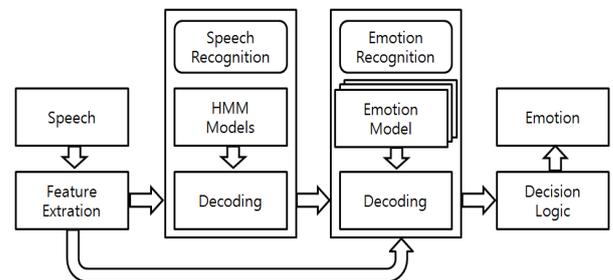
지금까지의 기술은 정확성, 속도, 그리고 효율성을 향상시켜 왔지만, 이제는 패러다임이 바뀌고 있는 추세이며, 감성과 교감, 소통이 중요한 역할을 하고 있다. 최신 스마트폰에는 6~7가지 이상의 센서가 집약돼 있으며, 각 센서의 조합으로 사용자와 단순한 인터랙션이 아닌 스마트폰과 사용자간의 교감이 생성된다.

정확한 위치 정보를 위하여, 스마트 폰의 GPS정보와 현재 주변의 건물이나 간판들을 스마트 폰의 캠으로 영상을 인식하여, 사용자의 위치 좌표에 가까운 장소를 찾아 향상된 위치 정보를 획득한다[1]. 또한, 스마트폰의 캠을 이용하여 얼굴의 표정을 인식하고, 청각센서로 사용자의 음성을 분석하고 글 입력에 대한 내용을 감지하여 사용자의 감정을 추정하여 특정위치나 사물에 대한 사용자의 감정을 자동적으로 포스팅하도록 구현 하는데 목적이 있다.

### 2. 사용자의 감정 인식

음성을 통한 감정 인식을 위해서는 각각의 감정이 음성에 어떠한 변화를 만들어 내는가를 정확하게 규명해야 하며, 특징 파라미터 추출 및 패턴 인식 알고리즘 선택에 있어서 차이가 있다.

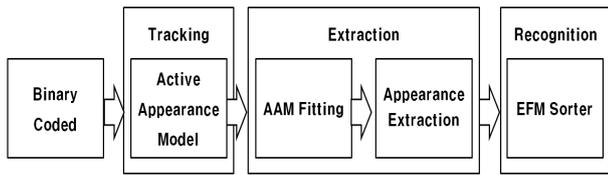
운율적 정보만을 이용하여 감정 인식 시스템을 구성하였던 것과는 달리, 은닉 마르코프 모델(HMM)의 구조를 변형하여 운율적 정보와 음향학적 정보를 결합하여 연구하였다.



(그림 1) 음성인식을 이용한 감정 추정

그림 1. 과 같이 음성을 통해 감정을 인식하기 위해서는 먼저 전처리 과정을 거치고 음성 구간 검출을 통해 음성 구간과 묵음 구간을 구별하고 특징 벡터를 구한다. 다음으로 수행 과정에서 은닉 마르코프 모델(HMM)과 가우시안 혼합 모델(GMM)기반의 인식 단계를 거친다[2].

사용자가 입력하는 글을 분석하여 감정을 추정하는 기술의 핵심은 목표 대상에 대한 긍정 또는 부정의 여부를 문서에 할당해 감정을 분류하는데 있다. 같은 감정의 단어라도 일반적으로 사용될 때와 특정 영역에서 사용될 때는 그 의미의 차이가 있기 때문에 이러한 현상을 반영하는 표현들을 PMI수치를 통해 추정할 수 있다[3]. 전처리 과정을 통해 내용어를 추출하고 추출된 내용어로 문장 벡터들을 구성한 후 일반적 + 영역 의존적 의미지향성 추정으로 명사, 형용사, 동사, 부사만을 고려하여 문장의 의미를 추정하고 감정의 상태를 파악한다.

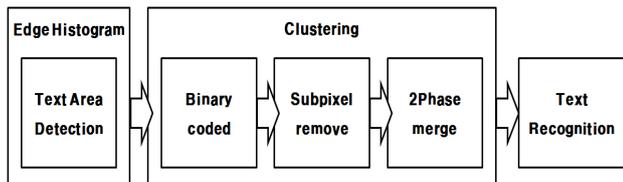


(그림 2) 표정인식을 이용한 감정추정

표정을 인식하여 감정을 추정하는 방법은 그림2. 와 같이 이진화된 데이터에서 AAM을 활용해 영상에서의 얼굴 위치를 추적한 후 Appearance값을 가지고 EFM에 적용하여 표정을 분류하는 기법을 사용하였다[4].

**3. 동영상에서 간판의 텍스트 검출 및 인식**

실외에서 획득한 영상으로부터 텍스트를 분석하고 가공하여 디지털화 하기위한 방법들과 알고리즘들이 다양한 분야에서 연구되고 있다. 모바일 시스템에서의 간판 영상의 텍스트, 주요 정보를 제공하는 상호 명을 자동적으로 검출 및 인식하는 시스템을 연구하였다.



(그림 3) 간판 인식 시스템 구조도

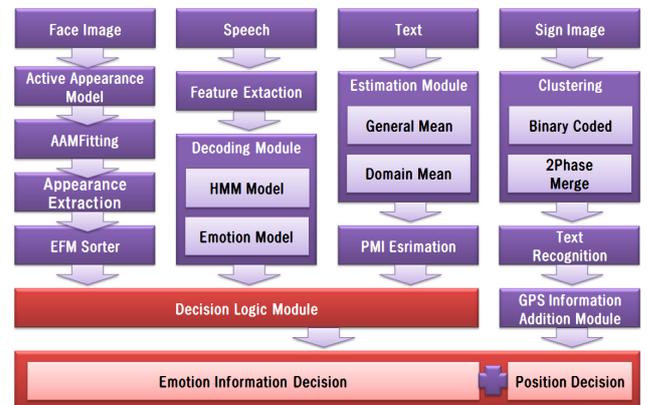
그림 3. 와 같이 간판 영상에서의 텍스트 영역 검출을 위하여 영상의 그레이 스케일 성분을 이용하고 에지 검출을 수행하여 수직 및 수평 성분 분석에 의해 텍스트 영역을 검출한다. 검출된 텍스트 영역의 이진화를 수행하기 위해 클러스터링 알고리즘을 적용 하였고, 텍스트 영역과 배경 영역을 나누기 위해 이진화된 영상을 잡영 제거 후 2단계 병합을 한다. 2단계 병합은 수직축과 수평축의 객체의 상관성을 고려하여 수행하며, 불규칙한 조명의 변화

에 퍼지 클러스터링 방법을 적용하였다[5].

문자인식은 두 가지의 특징인 구조 및 통계적 방법에 기반 하여 최소거리분류기(MDC) 기법을 사용하였다.

**4. 시스템 설계 및 구현**

사용자마다 각각의 정보를 따로 저장하도록 DB를 구성하고, 사용자가 입력하는 글과 표정, 음성 데이터를 입력 받아 감정을 인식하고 세 가지의 정보를 조합하여 사용자의 감정을 기쁨, 즐거움, 평온, 스트레스, 노여움의 5가지의 감정상태중 하나로 결정하도록 구현하였다. 또한, 카메라로 촬영한 간판의 정보를 인식하여 GPS정보의 좌표와 결합하여 지도에서의 특정 위치를 보다 정확하게 인식하도록 구현하였다.



(그림 4) 시스템 구조도

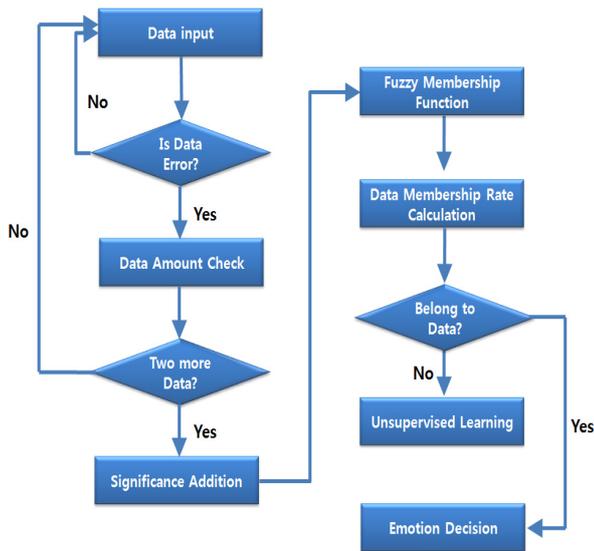
그림4. 와 같이 각각의 Input 모듈은 얼굴 영상과 음성 데이터, 텍스트와 간판이미지를 입력받는다.

얼굴 영상의 처리는 먼저 영상에서의 얼굴의 위치를 AAM을 이용하여 인식하고, 인식한 부분의 표정 정보의 특징 점은 Appearance Extraction 모듈에서 추출한다. 추출된 데이터는 EFM분류기를 이용하여 데이터베이스에 저장되어 있는 데이터와 비교를 통해 표정의 감정 상태를 추정한다. 추정된 데이터의 형태는 감정의 단계별 수치로 표현되며, 수치에 해당하는 가까운 감정을 결정하게 된다.

음성 데이터 처리는 전처리 과정 후 은닉 마르코프 모델(HMM)과 감정모델(EM)로 운율적 정보와 음향학적 정보를 결합하여, 감정의 단계를 추정한다.

텍스트 데이터 처리는 문장을 입력받아 단어들의 일반적 + 영역 의존적 의미지향성 추정과 PMI 수치를 이용하여 감정의 단계를 추정한다.

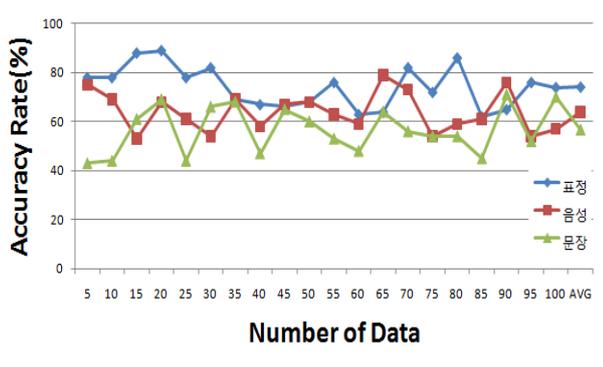
감정을 결정하기 위해서는 Decision Logic 모듈에서 두 가지 이상의 데이터를 비교하여, 결정하도록 구현하였다. 간판 영상은 이진화를 거쳐 텍스트를 추출하고, 특징 점을 2단계 병합을 통하여 단어를 인식하고, GPS의 정보와 결합하여 정확도 높은 위치 정보를 획득할 수 있도록 구현하였다.



(그림 5) 감정 결정 흐름도

그림5. 와 같이 각각의 데이터의 감정의 단계별 수치 데이터가 들어오면 이를 인식한 후 먼저 에러를 체크한다. 다음으로 2개 이상의 데이터를 입력 받았는지 확인하고, 1개의 데이터만 입력 받게 되면 다른 데이터의 입력을 대기한다. 데이터가 음성데이터, 표정데이터, 문장데이터 인지 확인하고, 해당 데이터에 가중치를 곱한다. 결과 값은 구축되어진 데이터베이스와 퍼지 소속 함수를 이용하여 감정별 소속도 값을 결정한다. 각각의 정보들은 소속도 값을 통해 감정인식 결과 값을 출력하게 된다. 데이터가 소속되지 않으면 비지도 학습을 하게 되도록 구현하였다.

5. 실험 결과



(그림 6) 감정별 가중치를 위한 감정 테스트

그림6. 과 같이 표정인식, 음성인식, 문장인식을 통한 감정 추정 정확도를 각각 100개의 데이터로 테스트 한 결과 감정인식 정확도는 표정인식, 음성인식, 문장인식의 순으로 결과가 나왔다. 이러한 결과로 표정인식은 70%의 가중치, 음성인식은 66%의 가중치, 문장인식은 62%의 가중치를 합산하여 데이터의 소속도를 결정하였다.

6. 결론

본 논문에서는 소셜 네트워킹에서 스마트 폰을 이용하여 정확한 위치 정보와 내장되어 있는 센서들을 이용하여 표정인식, 음성인식, 문장인식을 통하여 자신의 감정을 추정하는 시스템을 구현하였다. 최근 현대는 다른 사람과 감각을 공유하고 같이 느끼는 교감의 시대로 접어들었으며, 각각의 센서가 별도로 동작하는 것이 아니라 유기적인 센서의 조합으로 사용자에게 대한 정보를 파악하고 이를 통해 보다 개인화되고, 차별화된 서비스를 제공할 수 있다.

향후 사용자와 스마트 폰의 센서와의 교감을 이용하여 인터페이스, 사용자 인식, 인지, 주변환경 인지 등 여러 부분에서 혁신을 이끌면서 인간과 주변 환경, 인간과 기계, 그리고 인간과 인간 사이의 상호작용을 강화시켜주는 역할을 할 것이다.

참고문헌

[1] 강대기, "위치 정보가 증강된 스마트폰 상의 소셜 네트워킹 서비스", 대한전자공학회, 제33권, 1호, 1463~1465쪽, 2010년.  
 [2] 김원구, "음성 인식 정보를 사용한 감정 인식", Proceeding of KIIS, 제18권, 1호, 425~428쪽, 2008년.  
 [3] C Strapparava, "Learning to identify emotions in text", Proceedings of the 2008 ACM symposium on Applied computing, 2008.  
 [4] RP Hodson, "Emotion recognition in autism: Coordinating faces and voices", Cambridge Univ Press, pp. 122~131, 2009.  
 [5] 박중현, "모바일 시스템 응용을 위한 실외 한국어 간판 영상에서 텍스트 검출 및 인식", 전자공학회, 제46권, 2호, 187~194쪽, 2009년.