

IPTV 및 스마트폰을 위한 감성 메신저의 개발

성민영 남궁찬 백선욱
상명대학교 대학원 컴퓨터정보통신공학과
mysung@smu.ac.kr

안성혜
상명대학교 디지털콘텐츠 전공
ramsuny@smu.ac.kr

Development of an Emotional Messenger for IPTV and Smart Phone

Minyoung Sung, Chan Namkung, Seon-uok Paek, and Seonghye Ahn
Sangmyung University

요 약

사용자의 감정을 자동으로 인식하고 3D 캐릭터 애니메이션을 통해 표현한다면 기기를 통한 통신에 더 풍부한 감성을 부여하여 의사 소통의 효과를 높일 수 있다. 본 논문에서는 IPTV와 스마트폰 기기에서 구동되는 감성 메신저의 개발에 대해 기술한다. 이를 위해 문장 및 음색 분석을 통한 감정 인식, 영상 속의 얼굴 표정 추적, 그리고 개인화된 3D 캐릭터의 표정 및 몸동작 애니메이션을 통해 감정을 전달하는 감성 메신저를 제안하고 그 효과를 서술한다. Naive Bayes 알고리즘을 이용한 채팅 문장에서의 자동 감정 인식이 개발되었으며 실험을 통해 성능 및 효과를 검증한다.

1. 서론

사용자 감성 디자인, 혹은 UX (User eXperience)는 현대 소프트웨어 설계의 대표적인 경향 중 하나이다. 감성 디자인은 애니메이션, 음향 효과 등을 통해 사용자 경험을 풍부히 하고 이를 통해 정보 전달의 효과를 높이기 위한 것으로 이제는 PC 뿐 아니라 다양한 가전 및 멀티미디어 기기에 대해 널리 적용되고 있다. 특히, 최근 주목을 받고 있는 스마트폰, 웹 패드, 스마트 TV 등은 3D 가속기, 카메라, 진동 모터, 마이크, 그리고 각종 센서들을 내장하고 있어서 다양한 방식의 감정 표현 및 인식을 가능하게 하고 있다.

본 논문에서는 스마트폰 및 TV 환경에서 사용자 감정 전달을 극대화하기 위한 감성 메신저의 개발에 대해 기술한다. 메신저 상의 감정 전달을 위해서는 이미 이모티콘 (emoticon; emotional icon), 아바타 등이 사용되고 있으나 여전히 감정 전달 효과나 편의성 측면에서는 한계를 갖고 있다. 많은 경우 수동 입력에 의존하고 있으며 개인별 편차를 고려하지 않는 획일적 감정 표현에 그치고 있다.

본 연구에서는 텍스트 채팅, 음성 및 화상 통신을 지원하는 메신저에서 감정 및 표정을 자동으로 인식하고, 이를 개인화된 캐릭터 애니메이션으로 표현하는 감성 메신저를 제안한다. 입력된 텍스트와 음성을 분석하여 사용자의 감정을 정확하게 자동 인식한다면 감정 전달의 편의성을 크게 개선할 수 있다. 그리고 이렇게 인식된 감정을 개인화된 3D 캐릭터의 표정 및 몸 동작을 통해 표현한다면 감정

전달의 효과를 크게 개선할 수 있을 것이다. 화상 통신의 경우 얼굴 표정을 인식하여 캐릭터가 이를 따라하도록 함으로써 메신저의 감성을 풍부히 하는데 기여할 수 있다.

감성 메신저 개발의 첫 번째 단계로서 구글 안드로이드를 탑재한 스마트폰과 IPTV 셋탑박스에서 구동하는 메신저를 구현하였으며 텍스트에서 자동으로 감정을 인식하여 3D 아바타의 애니메이션을 통해 표현하도록 하였다. 향후 음성 통신 중 음색 분석을 통한 감정 인식 그리고 화상의 표정 인식 및 추적을 지원하도록 하여 감성 메신저를 완성할 계획이다.

2. 텍스트 기반 감정인식 관련 연구

Ma 외는 텍스트에서 감정을 유추하여 적절하게 캐릭터 애니메이션하고 음향 효과를 발생시키는 메신저를 개발하였다 [1]. 감정 판단을 위해 keyword spotting 이라는 방법을 이용하였는데 이는 단어별로 각 감정에 대한 기여치를 갖는 벡터를 데이터베이스에 구축해 두고, 문장을 구성하는 단어들의 벡터를 합한 후 그 결과로부터 감정을 판단하는 방법이다 [2]. 문현구 외는 Naive Bayes 알고리즘을 이용한 텍스트 기반 감정 추출을 한 연구하였다 [3]. 방대한 채팅문장들 샘플에 대해 각 문장에 실린 감정을 수동으로 분석하였다. 각 감정이 나타날 확률을 구하였으며 감성과 연관된 단어 각각에 대해 각 감정에 대한 조건부 확률을 계산하였다. 이렇게 계산된 확률 값들은 Naive Bayes 알고리즘을 통해 온라인 텍스트에 대한 감정 판단에 사용되었다. 실험을 통해 약 80% 이상의 감정 인식률

을 갖는 것을 보였다. 또한 접속사 등으로 연결된 문장 집합에 대해 인식률이 낮을 것을 개선하기 위해서는 HMM (Hidden Markov Model)을 사용하는 것이 효과적임을 제안하였다 [4].

3. 감성 메시저의 구현

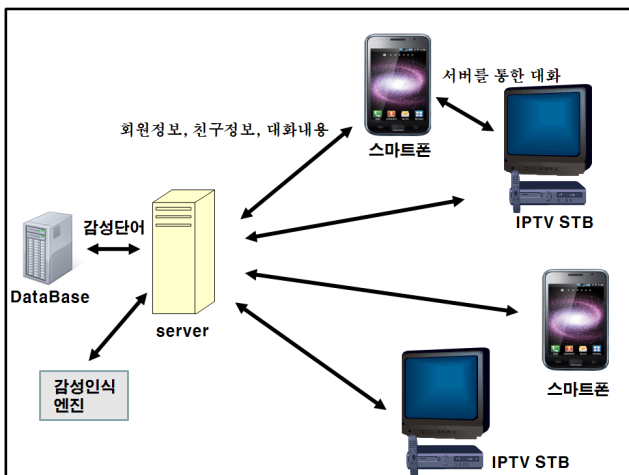
3.1 텍스트 기반 감정 인식

본 논문에서 제안하는 텍스트 기반 감정 인식 기법은 Naive Bayes 기법을 기반으로 한다. 다음은 감정 결정을 위한 수식이다.

$$e_s = \max_{e_j \in E} P(e_j) \prod_i P(s_i | e_j)$$

E 를 전체 감정의 집합이라고 하면 $e_s \in E$ 는 문장 s 에 대해 결정된 감정이다. $P(e_j)$ 는 감정 $e_j \in E$ 가 나타날 확률이며 $P(s_i | e_j)$ 는 e_j 감정이 인식되었을 때 단어 s_i 가 나타날 확률이다. 즉 감성과 관련된 단어들 s_i 를 포함하는 문장 s 에 대해 $P(e_j) \prod_i P(s_i | e_j)$ 를 최대화하는 감정 e_j 를 이 문장에 대한 감정으로 판단한다. 감정과 연관된 단어는 명사와 동사로 분류되며 동사는 그 기본형을 다루도록 하였다. 다만, 관련 연구 [3]의 모델과는 달리 제안한 방식에서는 “:”, “(”, “~”, “ㅋㅋ” 등 메시저에서 흔히 사용되는 이모티콘 및 그림문자를 명사에 포함하도록 하여 감정 인식을 개선할 도모하였다. 접속사로 연결된 문장이나 혹은 단순히 열거된 문장들 집합이 입력되면 전체에 대해 감정을 판단하기보다는 문장별로 감정을 정하고 이들 감정을 순차적으로 표현하도록 하였다.

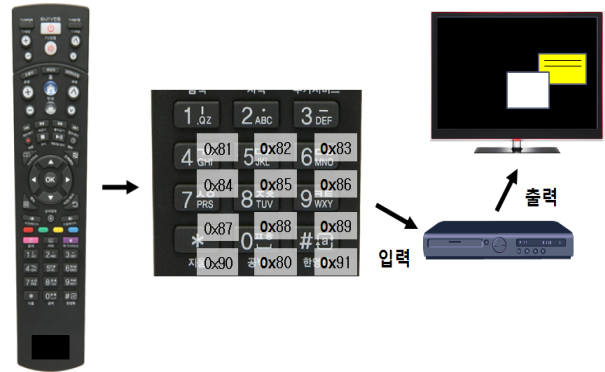
3.2 감성 메시저 시스템



(그림 1) 감성 메시저 시스템의 구성

그림 1은 감성 메시저 시스템의 전체 구성을 보인 것이다. 크게 통신을 중계하는 서버와 메시저 클라이언트 용

을 탑재한 스마트폰, IPTV 셋탑박스로 구성된다. 서버는 채팅 문장들을 중계함과 동시에 문장의 구문을 분석하고 감성 인식 모듈을 구동하여 감정을 판단하여 이를 텍스트에 태깅한다. 감정이 태깅된 채팅 문장들은 대화에 참여 중인 모든 클라이언트들에게 전달되고 클라이언트는 수신된 감정 정보를 이용하여 해당 감정의 아바타 애니메이션을 수행하게 된다.



3.3 IPTV용 메시저의 구현

(그림 2) 리모콘을 통한 문자입력

IPTV 메시저는 상용 셋탑박스에 널리 사용되는 SoC인 Sigma Design사의 SMP8634를 내장한 셋탑 테스트베드 상에서 개발하였다. 이 SoC는 300MHz로 동작하는 MIPS 코어를 내장하고 있고 Linux 운영체제가 탑재되어 있어서 메시저와 같은 간단한 응용 개발에 적합한 환경을 제공하였다.

TV 리모콘을 이용한 한글 및 영문 입력을 지원하기 위해 ‘0’에서 ‘9’까지의 숫자 버튼에 한글 자음 및 모음을 배치하였으며 이를 처리하는 오토마타를 구현하였다. 또한 ‘*’ 버튼은 한글, 영문, 숫자간의 모드 변환에 사용하도록 하였다. 그림 2는 버튼들의 스캔코드, 한글, 영문 자소 배치를 보이고 있다.

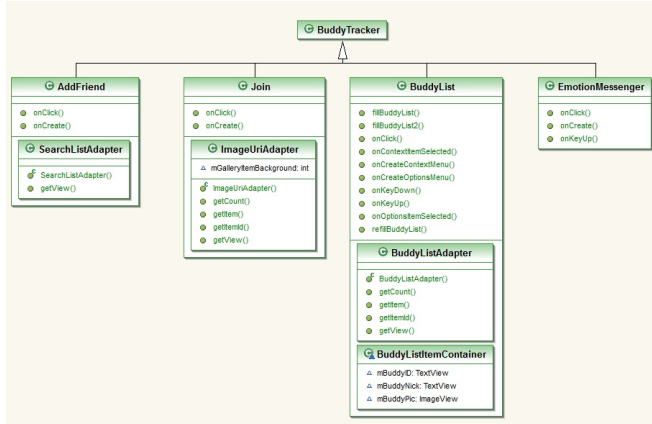
메시저의 GUI, 폰트출력, 입력처리를 위해서 공개 소스 멀티미디어 라이브러리인 SDL (Simple DirectMedia Layer)와 DirectFB를 포팅하여 사용하였다. 특히, 고성능의 그래픽 연산 지원을 위해 비트블릿 (bit-blit), 사각형 이동 및 채움 등 호출 빈도가 높고 성능에의 영향이 큰 DirectFB의 함수들은 셋탑박스 SoC에 내장된 그래픽 가속 함수들을 이용하도록 수정하였다.

IPTV에서의 3D 캐릭터 애니메이션은 동영상으로 처리하도록 설계하였다. 즉, 각 감성 상태에 대한 캐릭터의 표정 및 몸동작을 미리 렌더링하여 동영상으로 저장해 두고 인식된 감정에 따라 해당 동영상을 재생하는 방식으로 감정을 표현하도록 하였다. 이는 비록 최신 셋탑박스용 미디어 프로세서들에 3D 그래픽 가속 하드웨어가 탑재되는 추세이지만, 아직 다수의 IPTV용 셋탑박스에서는 2D 그래픽만을 지원하고 있는 현실을 고려한 것이다. 동영상은 최신 압축 표준인 H.264 코덱으로 인코딩하였으며 MPEG2

TS (Transport Stream) 포맷으로 저장하였다.

3.4 스마트폰을 위한 감성 메신저

스마트폰용 감성 메신저는 Google Android 버전 2.1 환경에서 개발하였다. 그림 3은 설계된 메신저의 클래스 다이어그램과 화면 구성을 보인 것이다.



(그림 3) 스마트폰 기반 메신저의 구조

안드로이드 응용은 액티비티 (activity)와 서비스 (service) 로 구성된다. 액티비티란 독립 실행 단위로서 상호작용, 질의 등을 통해 이들 액티비티 간의 전환이 이루어진다. 개발된 메신저는 로그인, 친구목록, 대화창 등의 액티비티들로 구성된다. 서비스는 주로 백그라운드로 실행되는 프로세스로 이해될 수 있으며 메신저에서의 서비스는 주로 텍스트 입력 및 출력, 캐릭터 애니메이션 등을 위한 이벤트들로 정의된다.

3D 캐릭터의 애니메이션은 OpenGL ES 버전 1.0에 기반하여 구현되었다. 감 감성에 대한 캐릭터 매쉬와 애니메이션 정보는 MD2 포맷으로 저장하여고 메신저에서 이를 관리하고 로딩하여 애니메이션하도록 구현하였다.

4. 시연 및 성능 실험

메신저의 기본 기능의 개발이 완료되어 스마트폰 및 IPTV 장치들간의 채팅이 원활히 이루어짐을 확인하였다. 그림 4는 개발된 메신저의 구동 모습을 보이고 있다. TV 메신저의 경우 리모콘을 통한 문자 입력이 다소 불편한 것으로 판단되어 추후 스마트폰용 리모콘 응용을 개발하여 스마트폰의 가상 키보드를 이용한 문장 입력을 지원하

도록 할 예정이다.



(그림 4) 스마트폰 및 TV 메신저의 구동 화면

텍스트 기반 감성 인식은 크게 평상, 기쁨, 슬픔, 화남의 4가지 감정 상태를 다루도록 정의하였다. Naive Bayes를 이용한 감정 판단 모듈의 개발은 완료되었으며 채팅 문장 샘플의 수집 및 확률 분석이 조만간 완료되면 실험을 통해 성능을 검증할 계획이다.

5. 결론 및 향후 계획

사용자의 감정을 자동으로 인식하고 3D 캐릭터 애니메이션을 통해 표현한다면 기기를 통한 통신에 더 풍부한 감성을 부여하고 궁극적으로 의사 소통의 효과를 높일 수 있을 것이다. 본 논문에서는 문장 및 음색 분석을 통한 감정 인식, 영상 속의 얼굴 표정 추적, 그리고 개인화된 3D 캐릭터의 표정 및 몸동작 애니메이션을 통해 감정을 전달하는 감성 메신저의 개발을 제안하였다. 첫 번째 단계로서 스마트폰과 TV환경에서 채팅 텍스트에서 자동으로 감정을 추출하고 이를 3D 캐릭터의 애니메이션으로 표현하는 메신저를 개발하였다.

향후 연구에서는 텍스트 기반 감정 인식을 개선하여 20가지 이상의 감정 상태에 대해 85% 이상의 감정 인식률을 갖도록 할 예정이다. 또한 음성을 기반으로 한 감정 분류와 얼굴 영상을 이용한 표정 추적 기능을 포함하도록 하여 감성 메신저의 완성도를 높여나갈 계획이다.

참고문헌

[1] C. Ma, et al., "A Chat System Based on Emotion Estimation from Text and Embodied Conversational Messengers", In Proc. Active Media Technology, 2005.
 [2] A. Osherenko, "Modeling Emotions Using a Shallow Natural Language Processing Technique," Master's thesis, Humboldt Univ. 2004.
 [3] 문현구, 장병탁, "채팅 텍스트로부터의 화자 감정상태 학습," 한국정보과학회 춘계학술대회, 2001.
 [4] 문현구, 장병탁, "HMM을 이용한 채팅 텍스트로부터의 화자 감정상태 분석," 한국정보과학회 추계학술대회, 2001.