인간의 감정을 자동 인식하는 전자메일 클라이언트의 설계 및 구현

김나영† · 이상곤††

요 약

현대에는 인터넷과 휴대폰의 대중화로 인해 일반 국민들도 전자메일 클라이언트를 통해 통신이 매우 자유롭다. 전자메일의 사용은 개인적인 일은 물론 기업 사무, 광고 메일, 뉴스, 광고 메일 등 폭넓게 이용되고 있으나 여러 문제점들이 나타나고 있다. 이제는 전자메일 클라이언트에 자연언어 처리 기술의 고급 기능의 탑재와 컴퓨터 바이러스 혹은 스팸 메일을 방지하는 고기능이 임베디드 되어야 한다. 이 기술을 이용하여 전자메일의 내용에 포함된 작성자의 “즐거움”, “성남”, “슬픔” 등 인간의 감정 표현을 여섯 가지의 감정 속성(전달 내용, 말투 표현, 불쾌 표현, 기쁨 표현, 슬픔 표현, 상대방 입장)으로 자동 매핑 하여 메일 클라이언트의 송신과 수신 프로세서에서 보다 정밀한 사람의 감정을 이해하고자 한다. 본 논문에서는 메일의 송신 전에 문서의 내용 중에 불쾌한 표현이 존재하면 제거하도록 문서 교정을 독려하는 지능적인 전자메일 클라이언트 시스템을 설계하고 구현하였다.

주제어 : 감정인식, 전자메일, 메일 클라이언트, 불쾌 레벨, 상대방 입장

Design and Implementation of E-mail Client based on Automatic Feeling Recognition

Na-young Kim† · Sang-kon Lee††

ABSTRACT

Modern day people can easily use an e-mail client for general communication, because of using Internet and cellular phone. The mail client for the purpose of private and business affair, advertisement, news searching, and business letter is widely used and has side effects. People could send an important document via an electronic mail client. It is important to support an e-mail client intelligent. We think that many kinds of techniques of natural language processing must be provided in the client with human's emotion. We consider to design a new mail client with six kinds of senders' emotional information; delight, angry, sad feeling and message to express, manner of talking, a discomfort index etc. Before sending an e-mail, we suggest a user to correct a bad word because we do not want to feel bad to a receiver. We present a proper process of sending/receiving for users with a new designed e-mail clients.

Keywords : Feeling Recognition, Electronic Mail, E-mail Client, Discomfort Index

† 정 회 원: 전주대학교 교육대학원 컴퓨터교육 전공
†† 종신회원: 전주대학교 컴퓨터공학과 부교수(교신 저자)
논문접수: 2008년 07월 09일, 심사완료: 2009년 03월 25일

1. 서 론

현대에는 인터넷과 휴대폰의 보급으로 인하여 전자메일이 우리의 일상생활에서 없어서는 안 될 중요한 통신 수단이 되었다. 전자메일은 통신망을 통하여 전달되기 때문에 시간에 관계없이 송수신 할 수 있으며, 상대가 부재중에도 전달하고자 하는 내용을 전달할 수 있는 장점이 있다[1]. 또한 전자메일은 우체국 우편의 기록성과 전화의 즉시성이 결합된 매우 효과적인 통신 방법이다. 2002년 한국리서치 미디어 인덱스에서 성인 남녀 2,350명을 대상으로 한 앙케이트 조사에 의하면 인터넷을 사용하는 이유에 대해 (복수 응답 선택을 허용하였을 때) 전자메일을 목적으로 사용하는 경우가 72.2%, 재미를 위해 64.4%, 파일을 다운로드 하기 위해 51.1%, 물건/서비스 정보의 조회가 48.0%로 조사되었다[2]. 이와 같이 인터넷 사용의 큰 목적이 전자메일 클라이언트를 사용하기 위해서임을 알 수 있다. 이와 같이 메일 클라이언트는 전자메일의 신속성과 사용자의 흥미 측면에서도 밀접한 관계를 갖는 통신 수단이다.

현대사회에서 전자메일, 인터넷 메신저, 휴대폰의 문자 메시지(SMS) 등은 커뮤니케이션 도구의 대표 주자이다. 이 도구들은 전화와 편지의 영역을 순식간에 잠식하며 현대인의 일상에 아주 빠르게 침투하였다. 하지만 전자메일을 사용하면서 겪는 혼란과 불쾌감도 아울러 증대되고 있다. 상대방에게 전자메일을 보낼 경우 상대의 반응을 보며 사용하는 문구를 수정할 수 없기 때문에 직접 면대면으로 대화하는 것보다 훨씬 소통이 어렵다. 전자메일의 심각성에 대하여 살펴보면, 첫째 작성자가 화가 난 상태에서 직접 종이에 글을 쓰거나 말을 하는 것 보다 전자메일로 쓴 경우에는 작성자의 감정을 더욱 통제하기 힘들다. 전자메일은 속도가 빠른 탓에 당장의 화를 참지 못하고 자신의 감정을 쏟아내는 메일을 보내버린 경우 상대방과의 관계가 회복할 수 없는 상태가 되는 일도 있다. 또한 부주의한 조작으로 개인적인 메일 내용이 많은 사람들에게 순식간에 퍼질 수 있다. 화가 난 상태에서 전자메일을 작성했다면 보내기 버튼을 클릭하기 전, 전달하고자 하는 내용이 제대로 기술되었는지 반드시 생각해 보아야 한다. 둘째, 전자메일 작성 시, 남녀 간에 생각의 차이를 염두에 두고 작성하여야 한다. 남성들은 재미있다고 생각할 수 있는 무례한 언어, 성적인 농담 등이 내용에 포함된 경우, 여성들은 불쾌감을 느낄 수 있다[3]. 셋째, 현대인은 하루에도 수십 통의 전자메일을 주고 받는다. 수신자는 한꺼번에 많은 전자메일을 읽을 경우 애매모호한 제목을 써서 스팸메일과 혼동을 일으킬 수도 있으며, 발신자가 긴급 메일 기능을 습관적으로 사용하는 경우에는 업무에 지장을 줄 수 있다. 전자메일 쓰기의 매뉴얼이 되다시피 한 시플리 & 슈발브 (2008)의 두 저자 데이비드와 윌은 “제발 ‘보내기(send)’를 누르기 전에 한 번 더 생각하라”고 충고한다[4].

요약하면 전자메일은 개인적인 일, 기업 내의 업무 연락, 메일 뉴스, 상업용 광고 메일 등 여러 분야에 매우 폭넓게 이용되고 있다. 이와 같이 현대에는 전자메일이 일상생활에서 없어서는 안 되는 중요한 도구가 되었으며, 사용자는 단순히 송/수신 기능 이외에도 별도의 다양하고 유용한 기능이 추가적으로 탑재된 새로운 개념의 전자메일 클라이언트의 개발을 원하고 있다. 이에 부응하듯 산업계에서는 메일 클라이언트가 컴퓨터 바이러스를 자동 제거하거나 스팸 메일을 자동 인식하여 필터링하는 등 고기능성 메일 클라이언트[5][6][7][8][9]에 대한 연구가 활발하다.

국내에는 전자메일 클라이언트에 고도의 자연언어처리(NLP; Natural Language Processing) 기술이 이식된 연구[10][11]가 진행되고 있으나 여전히 부족한 실정이다. 지금까지 이러한 기술이 도입된 시스템은 매우 高價이며, 일반 사용자까지 이러한 고가의 기능을 이용할 수 없는 실정이다. 또한 한국어 정보처리 분야에서의 접근 방법 개발이나 문제 해결 방법에 대한 연구 또한 여전히 미흡한 실정이다. 일반적인 형태의 전자메일을 살펴보면 다음과 같은 문제점이 있다. 전자메일은 보낸 사람의 감정 표현(즐거움, 슬픔, 성남, 두려움 등)이 포함되어 있으며, 수신자는 메일을 읽으면서 작성자의 의도와 비슷한 감정을 일으킬 수 있지만 작성자의 감정이 상대방에게 정확하게 전달되지 않아 작성자의 의도와는 반대되는 감정을 느끼는 상황이 발생할 수 있다.

인류 역사상 人間의 감정은 오랜 기간에 걸쳐 정신 의학, 심리학, 생물학 등 여러 학문 분야에서 꾸준히 연구되어 온 주제이다. 최근에는 감정 컴퓨팅(Affective Computing)이라는 키워드로 컴퓨터와

인간의 상호작용을 보다 윤택하게 설계하기 위해 시각 디자인 분야와 인공지능을 비롯한 전산학의 여러 분야에서도 감정을 표현하는 얼굴 표현, 감성 로봇, 감정이 표현된 음성 합성, 감정 예측 등의 여러 감정에 관한 요소 기술이 개발되고 있다[12].

본 논문의 목적은 전자메일의 기본 기능인 송신과 수신의 각 프로세스에서 감정 표현을 자동으로 이해하는 클라이언트의 개발에 있다. 감정 표현 이해를 통해 전자메일의 내용에 포함된 불쾌한 표현을 제거하여 문서 교정을 독려하는 새로운 개념의 메일 클라이언트 시스템을 제안하고자 한다. 덧붙여 말하면 본 연구의 목적은 송신자가 작성한 전자메일에서 감정 표현을 자동으로 이해하는 것에 있다. 전자메일의 내용에 포함된 작성자의 “즐거움”, “성남”, “슬픔” 등 여러 복잡한 감정 표현을 송신과 수신 단계에서 이해하여 감정 표현을 추출하고, 메일 내용에 포함된 불쾌한 표현을 제거하여 수신자 입장에서 기분을 상할지도 모르는 표현들을 제거하도록 독려하는 시스템을 구현하였다.

감정 표현 이해 모듈을 개발하여 상대에게 “즐거움”, “성남”, “슬픔” 등의 표현을 추출하고 문서의 전달 내용, 말투 표현, 불쾌 표현, 기쁨 표현, 슬픔 표현 등 다섯 가지의 감정 속성과 상대방과 자신의 관계를 고려하는 상대방 입장에 대한 감정 속성을 추출하고, 어느 문장이 어느 정도의 감정인가를 나타내는지 이해하고자 한다.

이하 본 논문의 구성은 다음과 같다. 제 2장에서는 관련 연구에 대해 설명하고, 제 3장에서는 본 논문에서 제안하는 시스템의 전체 개요와 감정 표현의 이해 방법, 감정 레벨(Emotional Level)의 결정과 감정 해석 방법에 대하여 기술하였다. 제 4장에서는 본 논문에서 제안하는 전자메일 클라이언트를 구현하고, 구현된 결과를 제시한다. 마지막으로 제 5장에서는 결론과 향후의 연구 과제에 대해 논의한다.

2. 관련 연구

본 장에서는 인간의 감정 표현을 보다 효과적으로 이해하기 위하여 감정 인식의 여러 연구 방법과 감정 정보의 처리 시스템에 대해 관련 연구를 조사하였다. 본 논문의 연구 주제와 유사한 열 한 개의 참고 문헌을 조사한 결과 감정 인식 방법을 지식베이스로 이용한 연구, 음성 정보를 이용한 연구, 감정 인식을 적용한 모델 등 크게 세 가지로 분류할 수 있었다.

2.1 감정의 지식베이스화에 대한 연구

2.1.1 감정 요소를 사용한 정보검색

한국어의 감정에 기반 한 정보검색 시스템을 구축하는 것을 목표로 하여 ECRAS의 영어 단어 시소러스를 기반으로 한국어에서 자주 출현하는 행복에 관한 1,054 단어, 슬픔 745 단어, 노여움 571 단어, 공포 575 단어, 혐오 720 단어 등으로 감정 시소러스를 구축하였다.

이 연구에서 개발된 감정 요소 추출기를 통해서 다섯 가지 감정 요소인 (1) 행복, (2) 슬픔, (3) 노여움, (4) 공포, (5) 혐오 등을 대상으로 하는 문서에서 각종 형태소 사전을 이용해 원문을 벡터화시킨다. 벡터화 된 문서들은 상관 계수 방법, 단순 베이지안, K-Nearest Neighbor 기법 등을 사용하여 최종 분류하고 사용자가 원하는 키워드와 함께 감정 요소 값에 의해 검색을 수행하는 시스템이다.

감정 단어들을 색인하고 검색하는 기술을 통하여 한국어의 문서에서 나타난 감정을 검색하는 시스템 구축을 목적으로 하였다. 이 연구는 인터넷 환경에서 웹 문서, 뉴스, 소설, 영화 등의 사이트를 분류, 검색에 적용 가능하다[13]. 이 연구의 기반과 본 논문의 기반은 같은 맥락에서 출발한다. “대표하는 감정은 감정 단어의 빈도에 비례한다”는 정의[13]에서 본 연구와 공통점을 찾을 수 있다.

2.1.2 신문 기사의 감정 추출 방법

기사의 검색에도 인간의 주관적인 느낌이나 생각을 표현하는 감정 요소를 동반하여 질의어에 적합한 정보를 검색할 수 있는 방법이 요구됨에 따라 인간의 감정이 내재된 신문 기사를 검색하는 방법을 제안하였다. 감정 표현 방법의 제 1단계는 기사 내용의 구문 분석과 품사 태깅 절차를 진행하고, 2단계에서는 감정 동사를 추출하고, 마지막 3단계에서는 추출된 감정 동사를 추출하고 그들 간의 관계를 이용하여 신문 기사의 감정을 최종적으로 추출하였다.

추출된 감정 동사들을 이용하여 신문 기사의 감

정 정보를 분석하기 위해서 OWL/RDF(S)를 기반으로 한 온톨로지[29]를 구축하고, 구축된 온톨로지를 참조하여 추출된 감정 동사들은 자신의 위치를 파악하고, 최상위 클래스와의 에지를 기반으로 하여 유사도를 측정하여 감정 정도를 백분율로 제시하였다[14].

2.1.3 감정 자질 추출을 이용한 감정 분류 시스템

이 연구에서는 한국어 문서에 대한 감정을 분류하고자 하였다. 감정 분류란 저자가 작성한 문서가 주제에 대해 긍정 감정이냐 혹은 부정 감정이냐에 초점을 두어 감정을 분류하는 방법이다. 이 감정 분류는 상품에 대한 고객 평가의 글에 들어 있는 고객의 감정을 분류하고, 그 내용을 요약, 공공 의견 조사, 고객 성향 분석 등 여러 분야에 대해 폭넓게 적용하고자 하였다.

이 연구는 문서의 분류를 위한 효과적인 감정 자질 추출 방법을 제안하였다. 한국어 감정의 분류 기준은 문서에서 나타나는 단어의 형태가 아닌, 긍정과 부정 즉, 단어의 의미에 기반함을 제시하였다. 감정 자질 추출 방법은 영어 단어 시소러스[15] 정보를 이용하고 영한사전을 이용하여 자질들을 번역하여 추출한다. 추출된 감정 자질들은 지지벡터 기계(SVM; Support Vector Machine)를 이용하여 문서 내 감정을 판단하였다[16].

2.2 음성 정보를 이용한 감정 인식 연구

2.2.1 감정 인식용 음성 신호 비교 분석

이 연구는 인간의 감정을 전달하는 요소인 음성 신호로부터 감정의 특징을 나타내는 요소를 찾아내는 것에 대해 주안점을 두었다. 여섯 가지의 기본 감정(행복, 슬픔, 분노, 증오, 놀람, 두려움 등)을 바탕으로 하여 감정에 따른 음성의 특징을 분석하였다. 아홉 명의 실험자를 대상으로 감정 요소가 포함된 이야기책을 읽게 한 후, 시간의 흐름에 따라 감정을 느끼는 시점에서 그 감정 단어를 녹음하여 평상시, 기쁠 때, 화날 때, 슬플 때의 음성 데이터를 추출하였다. 음성 데이터를 추출한 결과에서 강도(intensity)에 대한 측정 결과는 인간의 감정이 화가 났을 때보다 기쁠 때의 강도 파형이 훨씬 높게 나타남을 제시하고 있다[17].

이 연구는 음성 신호에서 감정을 인식하는 것이 아니고, 화자의 감정 상태에서 나타나는 음성 신호를 비교 분석하였다. 이러한 맥락에서 본 연구는 감정 상태에 따라 프로그램이 제시하는 정보를 변화하도록 하는 감정을 인식하는 메일 클라이언트로 확장할 수 있다고 생각한다.

2.2.2 억양의 강조에 의한 인식

인간에게 음성은 가장 쉽고 자연스럽게 의사를 교환할 수 있는 수단이며, 억양은 감정 상태를 음성으로 표현하는데 매우 중요한 수단이 된다. 한국어의 억양에 대한 음운론적 지식을 감정 인식 시스템에 적용하여 감정 인식률을 최대로 향상시키고자 한 연구이다. 이 연구의 감정 인식 시스템은 (1) 음성 문장 검출, (2) 음성 문장에서 피치 곡선 추출, (3) 감정 특징의 계산, (4) 감정 학습 및 인식 등 네 가지 단계를 통하여 처리한다. 이 시스템에서 구축한 감정 데이터베이스는 기쁨, 슬픔, 분노, 중성 등 네 가지로 분류된 감정이며, 감정을 인식하는데 문미 억양이 미치는 영향을 확인하기 위해서 세 가지의 실험 조건을 설정하여 실험한 결과 문미 억양을 포함하는 미부에 중요도를 부여했을 때의 조건이 일반 조건일 때보다 4%의 감정 인식률의 향상을 얻는 실험 결과를 제시하고 있다[16].

2.3 감정 정보의 적용 모델에 관한 연구

2.3.1 시각 디스플레이어에서 감성 모형 개발

동적 시각 디스플레이에서 움직임의 속성과 감성 간의 관계를 이해하고, 움직임에 의한 감성 모형을 개발하기 위해 다음과 같이 세 가지의 연구를 시행하였다. 첫째, 움직임을 설명할 수 있는 감성 어휘를 잡지 분석, 사전 분석, 설문 조사를 통해서 수집하고, 수집된 감성 어휘에 대해 적절성을 평가하였다. 시각의 움직임을 제시하면서 자유 연상에 의해 감성 어휘를 보고하여 70개의 주요 감성 어휘를 제시하였다. 아홉 개의 요인 분석에서 상관관계가 높은 각 요인을 대표하는 19개의 기본 감성 어휘를 추출하고, 기본 감성 어휘의 움직임

패턴을 관찰하였다. 관찰된 패턴은 다차원 척도법을 이용하여 감성 공간에 표현하였다. 감성 차원으로 수평축은 “적극적이다-소극적이다” 차원이며, 수직축은 “밝다-어둡다” 차원으로 설명하였다. 둘째, 감성 차원에 배치된 결과를 통해서 감성 차원과 관련되는 물리적 속성들을 다양하게 변화시키면서 감성을 평가한다. 움직임을 결정하는 물리적 속성 중 속도, 움직이는 방향, 곡선적 움직임의 주기, 진폭 등이 감성에 영향을 끼친다고 연구자는 말한다. 셋째, 움직임 속성에 색 정보를 추가할 때 감성의 변화를 살펴보고, 단색에서의 감성 차원은 색에 따라 “밝다-어둡다” 차원 뿐 아니라 “적극적이다-소극적이다” 차원에서 차이가 있었지만 색과 움직임이 동시에 적용되었을 때 움직임에 의한 감성 효과는 “적극적이다-소극적이다”의 경우 보다 강하며, 색에 의한 감성 효과는 “밝다-어둡다” 차원이 더 강하게 작용한다는 연구 결과를 제시하고 있다[19].

2.3.2 포털 사이트에서 아이콘 디자인

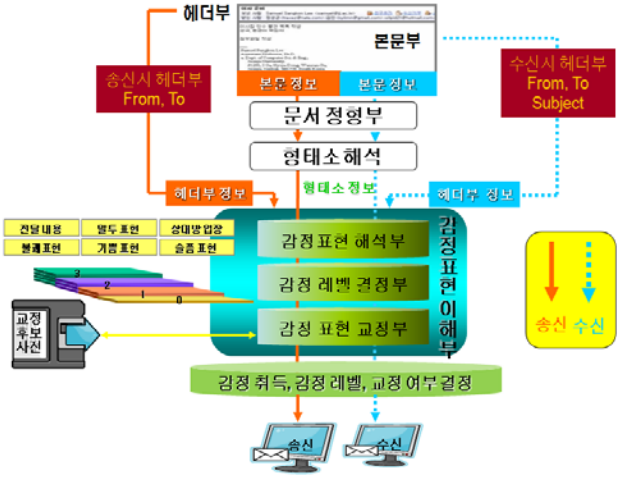
국내 인터넷 활용이 증가함에 따라 이용하는 연령층이 확대되면서 초등학생의 인터넷 이용률도 증가하고 있다[20]. 이 연구에서는 어린이들에게 친숙한 인터넷 환경을 이용하여 숙제 해결, 게임, 자료 검색 등에 사용하였다[21]. 아이콘은 사용자가 클릭을 통하여 다른 곳으로 이동하는 방향성이 있는 지시자 역할로 사용자의 추가적인 행동을 유도한다[22]. 인간은 그림에 대한 기억과 처리 능력이 우수하다. 특히 어린이들은 그림에 대한 인지 능력이 매우 높기 때문에 아이콘을 제작할 때 전달하려는 내용을 빠르게 이해할 수 있도록 디자인되어야 한다[23]. 본 연구의 연구자는 어린이들이 주로 이용하는 웹 사이트에 아이들의 감성을 잘 파악하고 그들의 감성에 잘 맞는 아이콘 제작이 이루어져야 한다고 주장한다. 국내에서 가장 많이 이용되고 있는 포털 사이트인 네이버(Naver), 야후(Yahoo), 다음 커뮤니케이션(Daum Communication)의 KIDS 서비스의 아이콘 샘플을 초등학생들에게 ‘선호’ 혹은 ‘비선호’에 대해 문의하여 인간의 감성에 적당한 아이콘 디자인에 대해 연구하였다[24].

2.3.3 교사 감정 모델을 통한 감정 캐릭터 개발

컴퓨터를 통한 학습량이 증가하면서 그래픽, 사운드, 동영상 등 풍부한 멀티미디어 자료와 플래시, 자바 등을 이용한 학습용 콘텐츠의 개발도 다양화 되고 있다. 그러나 학습 프로그램에는 정적인 캐릭터나 단순 동작의 캐릭터가 대부분이기 때문에 학습자의 지속적인 교육이 될 수 있도록 학습 촉진자의 역할을 최소화 하여야 한다. 이에 학습자의 지속적인 관심과 흥미를 유발하고 학습자가 컴퓨터 학습에 대해 거부감을 없애고 학업 성취도 향상을 위해 교사의 감정 캐릭터를 사용하였다. 본 연구에서는 현직 교사들이 캐릭터에 적용할 교사 감정 모델을 설정하고 수업의 진행 상황에 따라 감정이 어떻게 변화하는지에 대해 설문을 조사하고, 학습자들이 컴퓨터를 통해 대화형 학습이 가능하도록 캐릭터를 개발하였다. 캐릭터의 교사 감정 모델은 기쁨, 슬픔, 놀람, 화남 등 기본적인 네 가지 감정을 중심으로 시스템을 설계하였다[25].

<그림 1> 전자메일 문서의 예

이상과 같이 현대에는 여러 형태의 감정 요소를 이용하여 인간의 감정을 인식하고 모델링하는 연구가 다양하게 이루어지고 있다. 지식베이스 기반이나 텍스트를 기반으로 한 감정 인식에 대한 연구 이외에도 음성 정보를 이용한 감정 인식의 필요성이 대두되고 있다. 지식베이스를 이용한 연구는 본 논문의 아이디어와 동일하게 감정 단어 구축에 기반을 두고 있어 문서 내에서의 감정의 추출을 목적으로 하고 있으므로 한국인에 적합한 감정 시소러스를 보강하는 연구자의 노력이 절실하다. 이러한 견지에서 본 논문의 연구에서 구축하는 감정 지식의 데이터베이스가 추후의 연구자에 도움이 될 것으로 확신한다.



<그림 2> 전체 시스템 개요도

감성 모형을 개발하여 적용한 포털 사이트에서 아이들이 선호하는 아이콘 디자인, 교사 감정 모델을 이용한 감정 캐릭터 개발들은 감정을 교육적인 개념으로 접근하고 있다. 이러한 연구들의 후속 과제로써 본 논문에서는 불쾌한 단어를 교정할 수 있는 기회를 제공함으로써 청소년 언어 순화의 교육적 목표를 실현하는 새로운 개념의 전자메일 클라이언트를 제안할 수 있으리라 기대한다.

<표 1> 감정의 개념

논자	감정의 개념
William James (1884)	감정이란 육체의 감각적 변화에 의하여 나타나는 일종의 정신 상태이다.
Maisonneuve Pascal (1948)	감정은 일종의 직관 능력으로서 감각적 지각을 초월하여 파악하는 능력이다.
Philip Johnson Laird (1991)	감정은 의식 경험과 다양한 신체적, 내분비적 변화, 특징적인 표현과 행동을 함께 수반한다.
Petrovsky (1993)	감정은 개인의 평가에 의해서 결정되는 것으로 행동에 주체의 태도나 가치 부여가 수반되는 것이다.
조정옥 (1999)	감정은 즉각적으로 그리고 직접적, 자동적으로 사물을 파악하는 기능이다.

3. 감정 인식 시스템의 설계

3.1 전체 시스템의 개요

본 논문에서 제안하는 감정 인식용 전자메일 시스템의 전체적인 구성을 송신과 수신 프로세스 두 가지로 나누어 설명한다. 사용자가 작성한 메일 문서는 아래의 <그림 1>과 같이 헤더부(Header)와 본문부(Body)로 나누어지는데, 헤더부에서는 송신자와 수신자의 주소 정보를 취득하며, 본문부는 메일의 내용을 포함하고 있다. 헤더부와 본문부의 인식 과정은 김보미(2005)의 연구를 주로 참고하였다[27].

위의 그림에서 표시된 바와 같이 메일 본문의 마지막 부분에는 서명(Signature) 정보가 포함되어 있으나, 해석 과정에서는 제외하였다. 다음에 본 논문에서 제안하는 전체 시스템 개요도를 앞의 <그림 2>와 같이 설계하였다.

1) 송신시의 처리

사용자가 작성한 메일 문서는 헤더부와 본문부로 나누어지게 되는데, 헤더부에서 송신자(From)와 수신자(To)의 주소 정보를 취득하며, 본문부는 메일의 내용을 추출할 수 있다. 헤더부의 정보 즉, 송신자와 수신자의 주소 정보는 감정 표현 이해부에서 처리되지만 본문부의 정보는 문서 정형부와 형태소 해석 단계에서 문장을 구성하는 형태소를 인식하여 문장 단위로 분할한다. 감정 표현 이해부에서는 헤더부의 정보와 분할된 본문부의 정보를 이용하여 감정을 취득하고 감정 레벨을 결정한 후, 혹시 내용 중에 포함되어 있을 불쾌한 표현들을 교정한다. 교정이 완료되면 메일은 송신 처리 단계로 진행한다.

2) 수신시의 처리

수신자는 송신자가 작성한 메일의 헤더부에서 송신자(From)와 수신자(To)의 주소 정보와 제목(Subject)을 취득한다. 또한 본문부에서는 메일의 내용을 추출한다. 헤더부의 정보는 감정 표현 이해부로 이동하여 처리되지만, 본문부의 정보는 문서의 정형부 인식, 형태소 해석 등의 과정을 거쳐 처리된다. 감정 표현 이해부에서는 헤더부의 정보와 분할된 본문부의 정보에서 감정을 취득한다.

3.2 감정의 이해와 감정 속성

3.2.1 감정의 개념

인간의 의사결정 방식은 합리적이지 않는 경우도 있다. 즉 인간은 기분대로 결정하기도 한다[28]. 인간의 감정 표현 능력은 자신 혹은 타인이 느끼는 감정이나 자신이 느끼는 기분이 무엇인지를 표현하는 능력이다. 따라서 이러한 능력은 인간 자신이 느끼고 있는 감정과 기분을 파악하고 타인의 목소리나 표정에서 미묘하게 내 비치는 감정의 단서를 통해 상대방의 기분을 인지할 수 있도록 구현되어야 한다. 감정은 생활에서 일어나는 사건이나 활동의 대상을 경험하는 태도로 보이며 개인의 평가에 의해 사물, 사건을 보고 느끼는 마음의 움직임이다[26].

<표 2> 동・서양의 감정의 분류

동양	예기	喜怒哀樂愛惡欲, 喜怒哀懼愛惡慾(기쁨, 분노, 슬픔, 즐거움, 두려움, 사랑, 미움, 욕망)
	춘추좌씨전	喜怒哀樂好惡 (기쁨, 분노, 슬픔, 사랑, 미움)
	한의학	喜怒憂思悲恐驚 (기쁨, 분노, 슬픔, 사랑, 미움, 욕망)
서양	Descartes	기쁨, 슬픔, 사랑, 미움, 욕망, 경이
	Plutchik	기쁨, 분노, 슬픔, 공포(놀람), 수용, 혐오
	Fkman	기쁨, 분노, 슬픔, 공포(놀람), 혐오
	Scherer	슬픔, 즐거움, 공포(불안), 혐오, 수치심

앞의 <표 1>과 같이 감정의 정의에 관한 여러 연구들을 종합적으로 정리해 보면 感情은 “개인의 평가에 의해 결정되며, 육체의 감각적 변화에 의해 나타나는 정신 상태이며, 즉각적이고 직접적으로 사물을 파악하는 기능이다”로 요약할 수 있다. 이러한 감정의 개념에 대한 이해를 토대로 하여 다음 절에서는 인간의 감정 표현에 대한 이해 과정을 제시한다.

3.2.2 감정 표현의 이해

현대에는 인터넷 사용이 빈번하게 일어남으로써 전자메일이나 게시판의 글에도 작성자의 감정 표현이 포함되어 있다. 인간의 감정 표현의 요소는 매우 다양하게 존재한다. 이러한 다양한 인간의 감정을 동・서양에서는 <표 2>와 같이 여러 가지로 분류하고 있다[26].

텍스트로 작성된 문서는 화상이나 음성에 비해 작성자의 뉘앙스나 감정 표현이 충분히 반영되지 못하는 특징이 있다. 그 결과 작성자의 감정이 기록된 전자메일이나 게시판의 문장을 통해 전달하고자 의도대로 전달되지 못하거나 의도와는 반대로 전달되어 상대에게 불쾌감이나 분노를 주는 상황이 발생하기도 한다. 따라서 작성자가 의도한 문서에서 감정을 올바르게 인식하여 문서에 내포된 여러 형태의 감정 레벨을 파악하고, 이 과정에서 만약 문서의 내용에 불쾌한 표현이 존재하면 이를 추출하여 송신 전에 문서의 교정을 충고하여야 한다.

감정 표현 이해부에서는 수신자에게 기쁨(喜), 불쾌(怒), 슬픔(哀), 즐거움(樂)의 표현 어구를 추출하여 작성된 문장이 어떤 감정을 어느 정도의 레벨로 느끼는가를 판단한다. 문장에서의 작성자의 감정 표현을 추출하기 위해서 문서의 성질을 표시하는 속성으로 아래의 <표 3>과 같이 여섯 가지의 감정 속성을 정의하였다.

<표 3> 여섯 가지 감정 속성의 예

순번	심볼	속성 내용
①	α	문장에서의 전달 내용,
②	β	작성자의 말투 표현,
③	γ	상대에게 불쾌 감정을 일으키는 불쾌 표현,
④	δ	상대에게 기쁜 감정을 일으키는 기쁨 표현,
⑤	ϵ	상대에게 슬픔 감정을 일으키는 슬픔 표현,
⑥	ζ	상대방과의 상관관계를 나타내는 상대방 입장

3.2.3 감정 속성

전달 내용, 말투 표현, 불쾌 표현, 기쁨 표현, 슬픔 표현, 상대방 입장 등 여섯 가지로 분류된 감정

속성에 대해 차례대로 설명한다.

1) 전달 내용

전달 내용은 메일 문서에 출현하는 단어를 이용하여 송신자가 어떠한 감정을 표현하고 싶은지를 파악하기 위해 사용되는 정보이다. 이 내용은 메일 문서의 문장이 어떠한 의미를 갖는가를 나타내는 속성이다. 직접적으로 불쾌한 감정을 내포하는 내용은 「중상모략」이 있다. 또한 요구의 의미로 상대에게 지시에 따르도록 「강제 · 강요」, 상대방에게 지시, 의무 등의 실무적인 명령이나 요구를 하는 「명령 · 요구」도 불쾌감을 주는 경우가 있다[27]. 일반적인 연락 사항 및 회의에 사용되는 「임무 · 업무 연락」, 상대에게 용건을 묻는 「의뢰 · 부탁」등은 불쾌감을 직접적으로 주는 것은 아니지만 반복해서 사용되면 불쾌감을 유발할 수 있다. 이상과 같이 「강제 · 강요」, 「명령(지시) · 요구」, 「임무 · 업무 연락」, 「의뢰 · 부탁」 등을 각각 네 가지 속성으로 나누어 살펴본다. 전달 내용에 따른 서술 표현은 아래의 <표 4>와 같다.

<표 4> 전달 내용과 서술 표현

전달 내용	서술 표현의 예
강제 · 강요	명령(지시) · 요구 보다 강한 표현
명령(지시) · 요구	권고, 의무, 명령, 금지
임무 · 업무 연락	고시, 예정
의뢰 · 부탁	의뢰

가)「강제 · 강요」

「강제 · 강요」는 「명령(지시) · 요구」의 경우 보다 더 강하게 상대에게 요구하는 표현이다. 예를 들어 “오라”는 명령문에 강조하는 수식어가 붙어 “반드시 올 것”으로 표현되거나, 오라는 것에 화자의 요구가 더 강해져 거절하기 어려운 요구의 표현이 된다. 「강제 · 강요」의 문장은 「명령(지시) · 요구」의 문장과 같이 문장의 끝(문말)이 명령 구조로 구성된 경우가 많으나, 문말만을 보면 「명령(지시) · 요구」의 문장과 구분되어 있다. 예를 들면 「강제 · 강요」의 의미가 되는 “고지서를 발급 받기 원하면 돈을 가지고 오라”라는 문장에서 문말 표현만을 취득하면 “오라”가 얻어진다. 이 문말 표현에서 문장의 의미를 판정하면 「명령(지시) · 요구」가 된다. 서술 표현 이외의 조건을 추출하여 「강제 · 강요」를 판정한다. 예를 들어 “고지서를 발급 받기 원하면 돈을 가지고 오라”는 문장에 나타나는 “- 하고 싶으면”이란 표현은 어떤 결과를 얻기 위한 조건의 요구 표현이므로 이러한 형태는 「강제 · 강요」의 의미라 추정할 수 있다. 또한 “꼭 돈을 지불할 것!”이라는 문장의 표현 중 “꼭”이라는 단어도 강제의 의미가 내포되어 있음을 추정할 수 있다.

나)「명령(지시) · 요구」

화자가 의뢰하는 경우나 명령하는 경우는 청자에게 무언가를 요구하기 위해서 사용되는 표현이다. 예를 들어 상대방에게 올 것을 요구하는 경우, 의뢰 문에서는 “와 주세요”, “와 주시기 않겠습니까”가 되고, 명령문에서는 “오라”가 된다. 명령문은 윗사람이나 친한 친구로부터 받으면 불쾌감이 느껴지지 않지만, 아랫사람이나 자신보다 지위가 낮은 사람으로부터 받으면 불쾌감을 느낄 수 있다. 「명령(지시) · 요구」의 문장은 문말의 표현으로 판단되므로 어떤 행동을 요구하는 “권고”, “의무”, “명령”, “금지” 등 네 가지 종류로 분류한다.

다)「임무 · 업무 연락」

업무 상 사무 처리나 회의 혹은 세미나의 날짜, 시간, 장소 등의 업무 일정에 관한 연락은 “이하와 같이 결정되었습니다”, “알립니다”, “비가 올 경우에는 행사를 중지할 예정입니다” 등과 같이 작성되는 경우가 많다. 이러한 경우에는 앞에서 서술한 <표 4>의 “고시”, “예정”의 서술 표현을 적극적으로 이용할 수 있다.

라)「의뢰 · 부탁」

상대에게 요청하는 「의뢰 · 부탁」에는 예를 들어, “오늘 중으로 해 주시기 않으시겠어요”, “내일 5시까지 제출해 주기 바란다” 등의 문장이 있다. 이들 문장의 서술 표현은 모두 “의뢰”에 포함되므

로 <표 4>와 같이 “의뢰” 만으로도 판정이 가능하다.

<표 5> 말투 표현의 예

말투 표현	서술 표현의 예
반말 표현	-했지, -이야,
명령 표현	-해라, -오라,
평체 표현	-이다, -간다,
경체, 존대 표현	-입니다, -그렇습니다,
극존칭 표현	하여 주시기 바랍니다, 말하여 주시도록

2) 말투 표현

같은 의미의 내용을 상대방에게 전달하는 경우라 할지라도 화자나 필자의 말투 표현에 따라 받아들이는 상대방의 감정은 다르다. 말투 표현은 필자의 의도가 어떤 형태로 전해지는지를 표시하는 속성이다. 이 표현은 상대방과의 관계성에 대한 것과 문맥상의 말투 표현과의 조합에 의해서 상대방이 받아들이는 감정의 변화가 서로 다를 수 있다. 예를 들면, 인간은 윗사람으로부터 받는 명령 형태보다는 아랫사람으로부터 받는 명령 형태의 경우에 불쾌감을 더 강하게 느끼게 된다. 같은 의미의 내용이더라도 아래의 <표 5>에서 제시한 바와 같이 말투 표현에 따라 상대의 감정은 다를 수 있다.

존대 표현은 존대어가 포함되어 상대에게 존대의 의미가 전해지는 표현이며, 극존칭 표현[2]은 단어 내에 두 가지 동사가 동시에 존재하는 경우(AND 연산자, &&) ‘극존칭 표현’이라 정의하였다.

<표 6> 불쾌 ·기쁨 · 슬픔 표현의 예시

표현 패턴	표현 예	표현 패턴	표현 예	표현 패턴	표현 예
불쾌 표현	모욕적 언행 (비어 · 속어)	기쁨 표현	칭찬	슬픔 표현	죽 음 · 아픔
	배반 · 배신		격 려 · 지지		이 별 · 실연
	위협 · 협박		축 하 · 이익이 있는 통지		외로움
	무시 · 단점				
	무성의, 예의 없음		희 망 · 기대		실망
	싫음		믿음		실직

3) 불쾌 표현

불쾌 표현에는 바보(나쁜 놈), 죽는다, 이 녀석, 자네는 머리가 좋으니, 공부하지 않아도 되지 않나(비꼬는 투), (시험이 떨어진 사람에게) “떨어지면 창피하지” 등의 표현 어구들을 예시할 수 있는데, 상대방에게 불쾌한 감정을 일으키는 속성이다.

4) 기쁨 표현

문서에서 상대에게 기쁜 감정을 일으키는 속성이다. 기쁨 표현(앞의 <표 6>을 참고)은 그 표현을 기록하는 것만으로도 기쁨을 느끼기 때문에 한 가지 속성값을 갖는다.

5) 슬픔 표현

문서에서 슬픈 감정을 일으키는 속성이다. 슬픔 표현은 <표 5>에 제시하였다. 이러한 속성은 슬픔 표현의 단아를 읽고 난 후에 느끼는 감정이므로 기쁨 표현과 마찬가지로 한 가지 속성값으로 설정하였다.

6) 상대방 입장

우리말에서 높임말은 대화의 상대가 글을 쓰는 화자보다 높은 위치(혹은 지위)에 있을 때 사용된다. 상대에게 실례가 되거나 혹은 위화감을 조성하는 표현을 들었을 경우에 듣는 사람이 불쾌감을 느끼는 경우가 있다. 새로운 감정 속성 인자로서 자기와 상대방 사이의 입장에 대한 정보를 고려할 필요가 있다. 이러한 속성을 상대방 입장이라 정의하였다. 속성값으로 사용되는 상대방 입장은 작성자와 문서를 수신하는 상대방과의 상관관계를 나타내는 속성이다. 아래의 <표 7>과 같이 상대방 입장과 그 관계에 대해 네 가지로 나누어 정의하

2) 극존칭 표현이란 단어 내에 두 가지 동사가 있는 경우, 극존칭 표현이라 한다. 예를 들어 “하여 주시기 바랍니다” 문장에서 “하여 다오”, “바랍니다” 등 두 가지 모두의 성분이 함께 나타나는 경우이다.

였다. 상대방 입장의 취득 메커니즘은 참고 문헌 [10]과 [27]을 참고하면 된다.

<표 7> 상대방 입장의 예

관계	상대방 입장
1	윗사람(상하 관계)
2	아랫사람(상하 관계)
3	대등 관계
4	업무 관계

3.3 감정 레벨과 해석 모듈

<그림 3> 불쾌 레벨의 네 단계

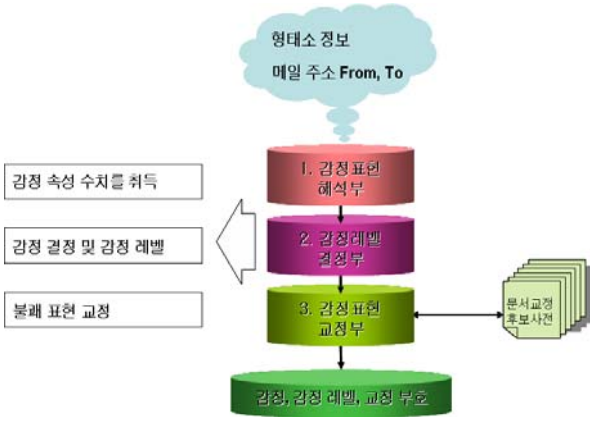
3.3.1 감정 레벨

감정 레벨이란 “성남(불쾌)”, “즐거움”, “슬픔”을 갖는 감정 표현에 대해 그것을 받아들이는 상대방이 어느 정도로 자신의 감정을 느끼는가를 표시하는 정도이다. 여섯 가지의 감정 속성을 기반으로 작성된 문장이 어느 감정인가를 네 단계의 불쾌 레벨로 정의하였다. 기쁨과 슬픔은 그 속성의 有無에 따라 간단히 두 단계로 정의하였다. 이하에서는 불쾌 레벨, 기쁨 레벨, 슬픔 레벨 등 세 가지로 나누어 서술한다.

1) 불쾌 레벨

불쾌 레벨은 숫자의 값이 클수록 불쾌 정도가 많다. 불쾌한 감정은 위의 <그림 3>에 모두 네 단계(0, 1, 2, 3)의 불쾌 레벨로 정의한다.

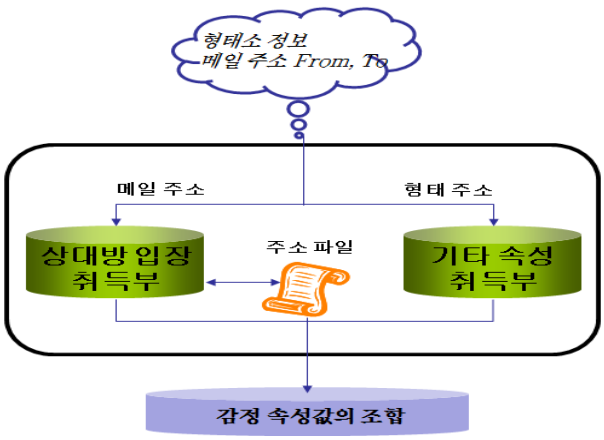
불쾌 레벨 0(불쾌감 없음)은 어떤 불쾌 감정도 느끼지 않는 레벨이다. 불쾌 레벨 1(약간 불쾌함)은 단순히 텍스트를 보는 것만으로는 불쾌 감정을 느끼지 못하고 받아들이는 쪽의 기분과 상태에 따라 가끔씩 불쾌감을 느끼는 정도에 대한 불쾌감이다. 불쾌 레벨 2(불쾌)는 불쾌 감정을 느끼며, 화가 커져 가는 레벨이다. 불쾌 레벨 3(아주 불쾌)은 어떤 표현을 보는 것으로도 매우 크게 불쾌감을 느끼며, 때때로 분노의 감정까지 일으키는 레벨이다. 불쾌 레벨은 여섯 가지의 감정 속성 중 전달 내용, 말투 표현, 불쾌 표현, 상대방 입장 등의 영향에 따라 해당하는 불쾌 레벨이 결정된다.



<그림 4> 감정 표현 이해부의 처리 흐름

2) 기쁨과 슬픔 레벨

기쁨 레벨과 슬픔 레벨은 기쁨 레벨이면 슬픔이 없고, 반대로 슬픔 레벨이면 기쁨이 없기 때문에 그 有無에 따라 두 가지 단계의 레벨로 부여하였다. 다시 서술하면, 값이 ‘0’이면 기쁘고(슬픔 없음), ‘1’이면 슬픔(기쁨 없음)이다. 기쁨(슬픔 없음) 레벨은 감정 속성의 기쁨 표현에 영향을 미친다.



<그림 5> 감정 표현 해석부의 처리 흐름

3.3.2 감정 표현의 해석 방법

감정 표현 이해부의 처리 흐름을 위의 <그림 4>에 설명하였다. 감정 표현 이해부의 감정 표현 해석부에서는 메일 문서에서의 본문부의 형태소 정보와 메일 주소를 이용하여 메일 주소의 참조, 감정 속성 수치를 취득한다. 해석부의 결과를 기초로 감정 레벨 결정부에서는 “즐거움”, “성남”, “슬픔” 등의 감정을 결정한다. 감정 표현 교정부에서는 위의 결과를 이용하여 송신 전에 사용자가

작성한 메일 내용의 불쾌한 표현을 교정 후보 사전의 지식에 근거하여 교정하도록 권유한다.

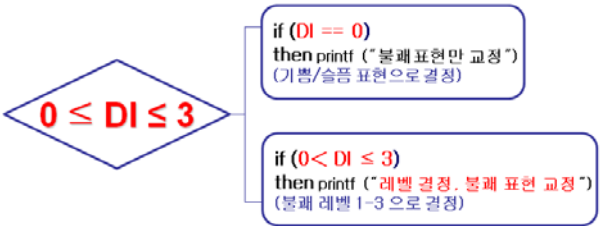
<표 8> 불쾌 표현 단어의 속성 가중치

가중치	해당 단어
1.0	분노의 제1 단계에 속하는 단어와 인간이 읽어서 바로 기분이 나쁜 단어들
1.5	분노의 제2 단계에 속하는 단어와 인간이 읽어서 바로 기분이 나쁜 단어 중 교정할 수 있는 단어들
2.0	분노의 제3 단계에 속하는 단어와 가중치 1.5의 교정할 단어 보다 더 불쾌한 단어들
2.5	분노의 제4 단계에 속하는 단어들
3.0	죽음의 의미와 놈, 년, 욕 등 욕설에 관련된 단어들
3.5	교정할 단어가 없는 비속어에 해당되는 단어들(“제거” 대상 단어)

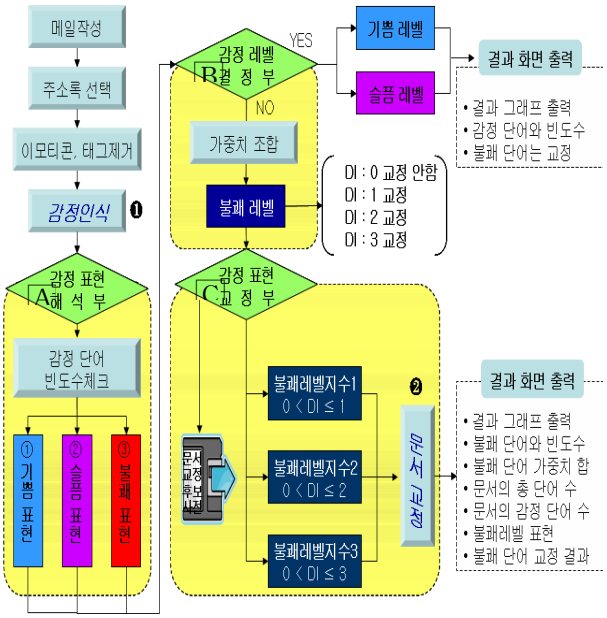
3) 감정 표현 해석부

문서 인식 모듈과 형태소 해석기를 통해 추출된 형태소 정보와 함께 전자메일의 헤더부에서 추출된 “From”과 ”To”의 메일 주소를 상대방 입장 취득부에 넘겨 기존의 주소 파일의 길이와 추출된 정보의 메일 주소를 비교하면 상대방 입장의 속성값을 얻을 수 있다. 그리고 위와 같은 해석 과정을 통해 추출된 형태소 정보를 이용하여 문장에서 얻어진 속성(전달 내용, 말투 표현, 기쁨 표현, 슬픔 표현, 불쾌 표현)을 기타 속성 취득부에서 취득한다.

4) 감정 레벨 결정부



<그림 6> 불쾌 표현의 교정 조건

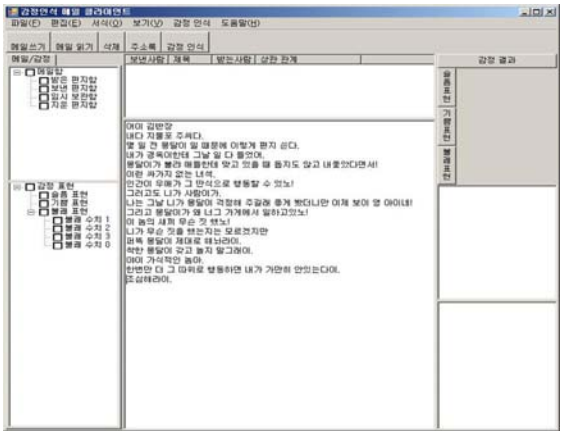


<그림 7> 감정 인식 시스템의 전체 흐름도

감정 레벨의 결정 방법은 불쾌 레벨, 기쁨 레벨, 슬픔 레벨에 따라 서로 다르지만 불쾌 레벨에서는 불쾌 레벨 결정 테이블과 불쾌 표현 속성치의 유무에 따라 결정할 수 있다. 기쁜 레벨과 슬픔 레벨은 기쁨 표현과 슬픔 표현의 유무에 따라 결정된다.

불쾌 표현 단어의 속성 가중치는 <표 6>의 서술 표현의 내용과 같이 구성된 255개의 불쾌 단어를 1.0에서 3.5까지 여섯 단계로 나누어 각각 표현하였다. 김향숙(2003)에서 “분노는 에너지의 증대(사나움, 억셈)로부터 생기며 분노를 열(발열, 끓어오름, 분출, 폭발)”로 표현하고 있다[24]. 본 연구에서는 분노의 제1 단계인 “에너지의 증대(사나움, 억셈)”에 속하는 단어에는 가중치 1.0을 부여하였다. 제2 단계인 “분노는 무기(날카로움, 독성)이다”에 속하는 단어는 가중치 1.5를 주었다. 분노의 3단계 “분노는 불(타오름, 연소)이다”에 속하는 단어는 가중치 2.0을, 분노의 제4 단계 “분노는 열(발열, 끓어오름, 분출, 폭발)”에 속하는 단어를 가중

치 2.5로, 죽음과 욕설에 관련된 단어를 가중치 3.0으로 각각 부여하였다. 마지막으로 3.5는 비속어 단어에 부여하였다.



<그림 8> 전자 메일의 작성 화면

결과적으로 가중치 1.0에서 3.5까지의 여섯 단계를 위의 <표 8>과 같이 요약 정리하였다. 이를 이용하여 불쾌 레벨 지수(DI; Discomfort Index)는 <표 9>와 같이 불쾌 표현 단어들의 가중치의 합을 불쾌 표현 단어의 총 개수로 나누어 계산한다.

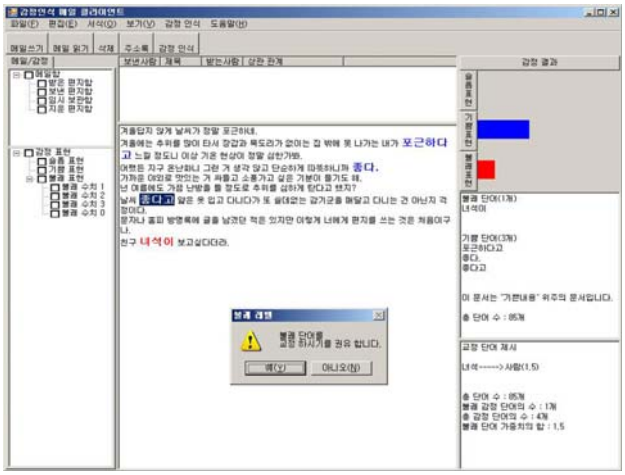
<표 9> 불쾌 레벨의 결정 범위

불쾌 레벨 지수	불쾌 레벨의 범위	불쾌 지수의 계산 방법
0	DI = 0	$DI = \frac{WS}{DW_N}$ DI= 불쾌 지수 WS= 가중치의 합 DW_N= 불쾌 단어의 수
1	0 < DI ≤ 1	
2	1 < DI ≤ 2	
3	2 < DI ≤ 3	

5) 감정 표현 교정부

전자메일을 송신하기 전에 얻어진 감정의 종류가 불쾌한 감정일 경우에는 메일 작성자에게 문서를 교정하도록 알린다. <그림 6>에서와 같이 불쾌 레벨이 1을 초과하고, 3 이하인 경우에는 미리 구축해 놓은 교정 후보 사전의 후보 중에서 하나를 선택하여 불쾌한 문장을 교정하도록 독려한다. 예를 들어 “바보”라는 형태소가 있다면 불쾌 표현 중 하나의 속성이므로 “불쾌 표현 제거” 메시지를 출력하고, 교정 모듈을 실행하여 “녀석”이란 단어로 교정하도록 독려하고 메일을 전송한다.

4. 인식 시스템의 구현



<그림 9> 불쾌 단어 출현의 교정 권유

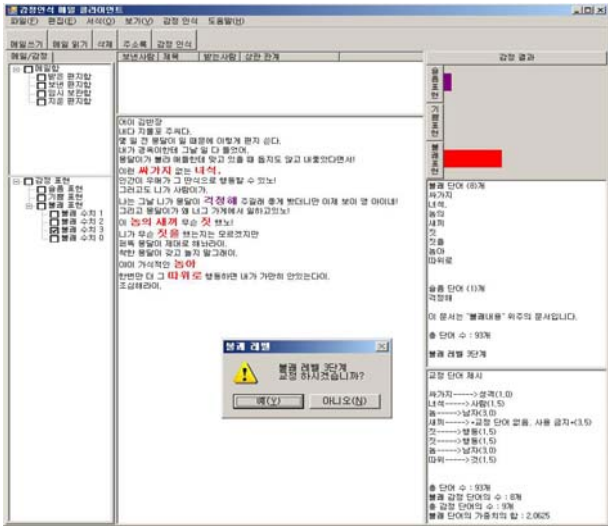
본 시스템의 구현 환경은 주 메모리 1GB를 탑재한 CPU Pentium Ⅵ 3.3 속도를 가진 컴퓨터 시스템에서 Microsoft社의 Visual Studio .Net 2005를 이용하여 구현하였고, 감정 단어를 저장한 사전은 추후에도 용이한 추가/수정/삭제가 가능하도록 우선적으로 텍스트 파일로 구축하였고 추후에는 데이터베이스로 관리할 예정이다.

기쁨 감정을 표현하는 단어는 207개, 슬픔 표현 300개, 불쾌 표현 311개 등 총 820여 개의 감정 단어를 지식베이스로 준비하고, 불쾌 표현의 교정 예상 단어는 162개로 구성하여 문서 교정 후보 사전을 지식베이스로 시스템을 구현하였다.

본 논문에서 설계된 시스템의 전체 흐름을 설명하기 위해 위의 <그림 7>에서 보는 바와 같이 시스템의 전체 구조를 Ⓐ~Ⓒ로 나누어 부분별 기능별로 표시하였다. 세 가지 모듈(Ⓐ~Ⓒ)을 순서대로 수행하면 최종적으로 전자메일의 감정 인식 결과 및 불쾌 단어의 교정 최종 작업을 달성하도록 하였다.

전자메일을 송신하기 위해 메일 작성 창에 위의 <그림 8>에서 제시한 예제와 같은 전자메일을 작성하였다고 가정한다. 전자메일의 내용에는 형식과 표현 방식에 제한이 없기 때문에 사용자의 자유로운 작성이 가능하지만, 시스템의 오류를 줄이기 위해 이모티콘(Emoticon)이나 특정한 기호(Special Character) 및 태그(Tag)는 제거하였다.

송신자가 작성한 메일 문서에는 여러 가지의 감정이 표현되어 있다. 수신자가 느끼게 될 기쁨(喜), 불쾌(怒), 슬픔(哀) 표현을 추출하여 작성된 메일 내용이 어떤 감정을 어느 정도로 느끼는가를 자동으로 인식하기 위해 앞의 <그림 7>에 표시된 전체 흐름도의 ❶ 감정 인식 과정을 수행한다. 단계 Ⓐ의 감정 표현 해석부에서는 문서 내용과 감정 단어 사전을 매칭 하여 ① 기쁨 표현, ② 슬픔 표현, ③ 불쾌 표현에 해당하는 단어의 빈도수를 체크하고, 그 빈도수에 따라 단계 Ⓑ의 감정 레벨 결정부에서 감정의 유형을 결정하도록 하였다.



<그림 10> 불쾌 레벨 정의와 교정 결과

감정 레벨 결정부에서는 문서에 나타난 사용자의 감정 유형을 결정하고, 기쁨이나 슬픔 표현일 경우에는 아래의 <그림 9>의 오른쪽 위와 같이 막대그래프로 감정별 유형을 나타내고, 그 아래에는 기쁨 표현 단어의 수와 슬픔 표현 단어의 수를 계량하여 표시하였다. 덧붙여 본 시스템은 불쾌 표현 단어를 교정하는 것도 연구의 주요 목표이기 때문에 기쁨이나 슬픔 표현으로 감정 유형을 결정하더라도 불쾌 표현 단어가 존재할 경우에는 교정 후보 사전을 기반으로 ❷와 같이 교정하도록 하였다.

감정 레벨 결정부에서 문서의 유형이 불쾌 표현으로 결정된 경우에는 <그림 10>과 같이 불쾌 표현에 대한 그래프와 함께 불쾌 표현 단어들을 표시하고 각각의 가중치의 합(WS; Weight Sum)을 구한다. 가중치는 1.0 ~ 3.5 까지 0.5 단위씩 각 여섯 단계로 비교적 세세하게 나누었다. 또한 문서 내의 총 단어수와 문서의 불쾌 감정 단어 수 (DW_N; Discomfort Word)를 <표 9>에 표시해 주었다. WS와 DW_N을 이용하여 불쾌 레벨 지수(DI; Discomfort Index)를 결정한다. 결정된 불쾌 레벨 지수를 대화 상자를 통해 표시해 주고, 사용자가 확인 버튼을 누르면 Ⓒ와 같이 감정 표현 교정부에서 처리된 불쾌 레벨 지수(DI)가 0 超過, 3 以下인 경우에 한하여 교정하도록 하였다.

5. 결론 및 향후 연구

본 논문에서는 전자메일 클라이언트에서 인간의 감정 데이터를 적절하게 처리하는 절차를 조사하고 이 과정을 통해 인간의 감정 인식 과정을 컴퓨터 프로그램으로 구현하였다. 현대에는 인터넷을 사용하는 연령층이 확대되면서 초/중/고등 학생들의 인터넷 사용 윤리도 중요시 되고 있다. 인터넷을 사람의 감정 표현에 대한 기회가 증가되어 청소년들 사이에서 사용되는 통신 언어의 순화가 필요한 실정이다. 본 논문은 이러한 청소년들의 언어 문화를 선도하고자 하는 일환으로 송신 전에 교정의 기회를 제공하는 신 개념의 전자메일 클라이언트를 제안하였다.

각 개인의 메일 서버에는 하루에도 막대한 양의 메일이 저장되고, 각각의 메일 문서에는 송신자의 여러 감정 표현이 포함되어 있다. 본 논문에서는 종래의 전자메일 시스템에 자연언어처리 기법을 이식하여 감정 인식 전자메일 시스템을 도입하고 사용자에게 폭넓은 지원을 제공하고자 하였다. 새로운 전자메일 기능을 제안하고 메일 송· 수신시의 각 프로세서에서 감정 표현을 이해하여 송신 전에 불쾌 표현을 인지하여 교정을 요구하며, 수신시에는 송신자의 감정을 자동으로 취득할 수 있는 연구도 병행하였다.

최근 일본에서는 사람의 감정이 얼굴 표정으로 나타나는 장면을 웹 카메라로 포착하고 이를 컴퓨터로 분석하여 사용자의 감정 상태를 부호화 하고, 의료 장비를 통해 상대방의 생각을 파악하는 연구가 진행되고 있다. 국내에서도 고객의 기분에 맞는 상품이나 광고를 보여주는 감성 인식 컴퓨터와 감성 로봇의 개발이 각광을 받을 것이라 예상된다 [10]. 본 논문에서 제시하는 연구 방법을 각종 상품 개발에 응용하면 보다 혁신적인 시스템으로 발전시킬 수 있다고 생각한다. 아울러 보다 더 인간의 감정을 잘 표현할 수 있으리라 기대한다. 또한 본 연구를 심화하면 본 논문의 서두에서 제시한

여러 관련 연구애서와 같이 어린 아이들의 감성적 습관도 형상화 할 수 있을 것으로 기대된다.

향후에는 불쾌 표현 단어를 교정 후에도 원래 화자가 의도한 자연스러운 문맥이 될 수 있도록 교정 후보 단어의 다양한 제시가 필요하다. 역으로 단어 지식 사전을 기반으로 하여 작성자가 그 것을 표현하고 싶은 감정에 따라 감정 단어들을 제시하면 그 감정에 맞는 문장이나 문서를 생성할 수 있는 시스템으로 설계 방향을 확장할 계획이다. 또한 감정 표현 중 “사죄”, “욕망”, “두려움”, “고뇌”, “치욕” 등 본 연구에서 제시하지 못한 감정도 추가하여 처리하고, 수신자의 성별(남자, 여자)을 고려하여 불쾌 표현 교정을 확장하며, 감정 레벨을 취득한 후 감정을 구비한 요약, 수신 메일의 내용에 의한 자동 반송 기능을 추가하는 방향으로 시스템을 확장할 계획이다.

참 고 문 헌

[1] 류제 · 윤성희 · 한광록 (2003). 특정 속성과 Co-training을 이용한 전자메일 분류. **한국정보과학회 봄 학술발표 논문집(B)**, 30(1), 549-554.

[2] 한국 리서치 미디어 인덱스 (2002). **2002년 1R 조사 결과**.

[3] 김윤덕 (2008). **상대방 배려 없는 이메일, 정말 보내실 겁니까.** 2008년 5월 14일자 조선일보 A-21면 기사.

[4] 데이비드 시플리 · 윌 슈발브(2008). **비즈니스는 이메일로 완성된다.** 서울: 랜덤 하우스 코리아.

[5] 현영순 · 정옥란 · 조동섭 (2003). E-Mail 시스템의 첨부 파일 형식별 자동 분류 및 스팸 제거 에이전트 설계. **제 20회 한국 정보처리 학회 추계 학술발표 논문집, 10**(2), 801-804.

[6] 김현준 · 정재은 · 조근식 (2004). 가중치가 부여된 베이지안 분류자를 이용한 스팸 메일 필터링 시스템. **정보과학회 논문지 : 소프트웨어 및 응용, 31**(8), 1092-1100.

[7] 이상호 (2005). 자동 생성 메일 계정 인식을 통한 스팸 필터링. **정보과학회 논문지 : 소프트웨어 및 응용, 32**(5), 378-384.

[8] 김강민 (2005). 사용자의 행동과 점진적 기계학습을 이용한 쓰레기 편지 여과 시스템의 설계. **제 23회 한국 정보처리 학회 춘계 학술발표대회 논문집, 12**(1), 775-778.

[9] 임정택 · 김형준 · 강승식 (2005). 나이브 베이지안 분류자와 메일 주소 유효성 검사를 이용한 스팸 메일 필터링 시스템. **한국 정보과학회 제 32회 추계 학술발표 논문집, 32**(2), 523-525.

[10] 김보미 · 이상열 · 이상곤 (2005). 이메일 문서의 속성값에 기반 한 필터링 시스템의 설계 및 구현. **한국 컴퓨터 종합 학술대회 논문집, 32**(1B), 142-144.

[11] 장정효 · 이상열 · 이상곤 · 조현준 (2006). 한국어 문서의 유형 정보를 이용한 EMFA의 구현. **한국 컴퓨터 종합 학술대회 논문집, 33**(1B), 28-30.

[12] 문애경 · 정인철 · 손강민 · 강태근 · 함호상 (2003). 감성 지능형 컴퓨팅 기술 동향. **한국정보통신 연구진흥원.**

[13] 김명관 · 박영택 (2003). 감정 요소를 사용한 정보검색에 관한 연구. **정보처리학회 논문지 (B), 10**(6), 579-586.

[14] 백선경 · 김판구 (2005). 신문 기사의 감정 추출 방법에 관한 연구. **한국 컴퓨터 종합 학술대회 논문집, 32**(1B), 562-564.

[15] 네이버 영어 단어 시소러스 (2008).

[16] 황재원 · 고영중 (2007). 효과적인 감정 자질 추출을 통한 한국어 문서 감정 분류 시스템 구축. **제 34회 정보과학회 추계 학술발표 논문집, 34**(2), 60-61.

[17] 조동욱 · 김봉현 · 이세환 (2006). 감정 인식을 위한 음성 신호 비교 분석. **한국 정보처리 학회 춘계 학술발표 대회 논문집, 13**(1), 533-536.

[18] 이태승 · 박미경 · 김태수 (2005). 한국어 문미 억양 강조를 통한 향상된 음성 문장 감정 인식. **한국 컴퓨터 종합 학술대회 논문집, 32**(1B), 514-516.

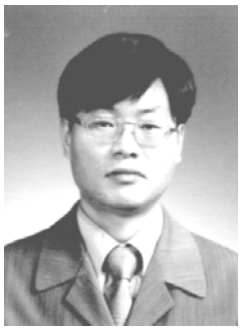
[19] 임은영 · 조경자 · 한광희 (2004). 시각 디스플레이에서의 감성 모형 개발. **한국 인지과학회 논문지, 15**(2), 1-14.

[20] RI Korea (2003). Korean Click 제 6차 인터넷 사용 실태 조사.

[21] 임인수 · 김인숙 · 노규형 · 노영화 · 박성용 · 한창수 · 박태훈 · 류윤상 (2000). **어린이의 디지털 매체 이용 행태 조사**. 한국 인터넷 정보센터.
[22] 이혜구 (1999). **아이콘 개발에 있어 효율적인 휴먼 인터페이스 디자인에 관한 연구.** 석사 학위 논문, 한양대학교.
[23] 이만재 · 이상선 (1998). **멀티미디어 교과서.** 서울: 안그라픽스.
[24] 고인화 (2004). 국내 어린이 포털 사이트 아이콘 디자인에 관한 연구. **한국 정보처리학회 제 22회 추계 학술 발표대회 논문집, 11**(2), 963-966.
[25] 이영희 (2005). 교사 감정 모델을 통한 감정 캐릭터 개발. **한국 컴퓨터교육학회 동계 학술발표 논문집, 9**(1), 9-14.
[26] 김향숙 (2003). **한국어 감정 표현 관용어 연구.** 서울: 한국문화사.
[27] 김보미 (2005). **프로파일 정보에 기반 한 전자메일 클라이언트 시스템의 설계 및 구현.** 석사학위 논문, 전주대학교 교육대학원.
[28] 오군석 · 김판구 역 (2002). **감성 데이터 해석.** 서울: 홍릉과학출판사.
[29] 황도삼 · 최기선 · 김태석 역 (2007). **차세대 웹과 온톨로지의 핵심기술 자연언어처리.** 서울: 홍릉과학출판사.

김 나 영

2001 전주대학교 정보통신공학 전공(학사)
2006~2008 전주대학교 교육대학원 컴퓨터교육 전공 교육학(석사)
관심분야: 한국어 정보처리, 자연언어처리, 컴퓨터교육
E-Mail: ny5991@nate.com



이 상 곤

1994 전주대학교 영어영문학과(학사)
1996 전북대학교 컴퓨터과학과(학사)
1998 전북대학교 전산통계학과(석사)
2001 일본 국립 도쿠시마대학교 지능정보공학과 (박사)
2002 원광대학교 음성정보 기술산업 지원센터 연구원,
2002~현재 전주대학교 컴퓨터공학과 부교수,
관심분야: 한국어 정보처리, 정보검색, 문서분류, 컴퓨터교육
E-Mail: samuel@jj.ac.kr