

# 스토리기반 저작물에서 감정어 분류에 기반한 등장인물의 감정 성향 판단

백 영 태\*

## Detection of Character Emotional Type Based on Classification of Emotional Words at Story

Yeong Tae, Baek\*

### 요 약

본 논문에서는 등장인물이 대사에서 사용한 감정어를 이용하여 등장인물의 감정 유형을 분류하는 방법을 제안하고 성능을 평가한다. 감정 유형은 긍정, 부정, 중립의 3 종류로 분류하며, 등장인물이 사용한 감정어를 누적하여 3 종류의 감정 유형 중에 어디에 속하는지를 파악한다. 대사로부터 감정어를 추출하기 위해 WordNet 기반의 감정어 추출 방법을 제안하고 감정어가 가진 감정성분을 벡터로 표현하는 방식을 제안한다. WordNet은 영어 단어 간에 상위어와 하위어, 유사어 등의 관계로 연결된 네트워크 구조의 사전이다. 이 네트워크 구조에서 최상위의 감정 항목과의 거리를 계산하여 단어별 감정량을 계산하여 대사를 30 차원의 감정 벡터로 표현한다. 등장인물별로 추출된 감정 벡터 성분들을 긍정, 부정, 중립의 3가지 차원으로 축소하여 표현한 후, 등장인물의 감정 성향이 어떻게 나타나는지를 추출한다. 또한 감정 성향의 추출 성능에 대해 헐리우드 영화 4개의 영화에서 12명의 등장인물을 선정하여 평가하여 제안한 방법의 효율성을 측정하였다. 대사는 영어로 이루어진 대사만을 사용하였다. 추출된 감정 성향 판단 성능은 75%의 정확도로 우수한 추출 성능을 나타내었다.

▶ Keywords : 등장인물, 감정 성향, 스토리, 영화, 감정어

### Abstract

In this paper, I propose and evaluate the method that classifies emotional type of characters with their emotional words. Emotional types are classified as three types such as positive, negative

•제1저자 : 백영태

•투고일 : 2013. 8. 19, 심사일 : 2013. 8. 29, 게재확정일 : 2013. 9. 3.

•김포대학교 멀티미디어과(Dept. of Multimedia, Kimpo College)

※ 이 논문은 2013학년도 김포대학교의 연구비 지원에 의하여 연구되었음.

※ 이 논문은 2013년 한국컴퓨터정보학회 하계학술대회에서 발표한 논문("감정어 추출을 통한 등장인물 성향 가시화 연구")을 확장한 것임.

and neutral. They are selected by classification of emotional words that characters speak. I propose the method to extract emotional words based on WordNet, and to represent as emotional vector. WordNet is thesaurus of network structure connected by hypernym, hyponym, synonym, antonym, and so on. Emotion word is extracted by calculating its emotional distance to each emotional category. The number of emotional category is 30. Therefore, emotional vector has 30 levels. When all emotional vectors of some character are accumulated, her/his emotion of a movie can be represented as a emotional vector. Also, thirty emotional categories can be classified as three elements of positive, negative, and neutral. As a result, emotion of some character can be represented by values of three elements. The proposed method was evaluated for 12 characters of four movies. Result of evaluation showed the accuracy of 75%.

▶ Keywords : Character, Emotional Type, Story, Movie, Emotion Word

## I. 서 론

영화나 드라마와 같은 스토리 기반의 저작물에서 등장인물은 스토리를 진행하는 핵심 주체이다. 등장인물들은 상호간에 관계를 맺으면서 사건이나 갈등을 발생시키기도 하며, 우호적인 조연을 받아서 문제를 해결하기도 한다. 영화의 장르에 따라서 등장인물들은 상호간에 상처가 되는 대사를 쉽없이 할 수도 있으며, 사랑의 밀어를 나누기도 한다. 이렇듯 등장인물이 표출하는 감정은 스토리의 갈등이나 심리묘사를 하는 중요한 도구이다. 등장인물의 감정 정보는 등장인물 간의 관계를 묘사하거나 등장인물의 긴장상태 등을 표현하는 정보로써 스토리 기반의 저작물을 분석하는데 매우 유용한 정보이다. 또한 스토리의 전개나 서사구조를 파악하는 데에 활용될 수도 있다.

등장인물은 대화를 통해 감정과 관련된 정보를 표출한다. 대화 속의 감정이나 표정변화와 어투, 행동과 같은 다양한 정보를 통해 등장인물의 심리 상태나 감정변화가 표출되게 된다. 표정변화나 어투와 같은 요소들로부터 감정을 파악하기에는 기술적으로 해결해야 할 여러 가지 문제점이 존재한다. 영상이나 음성 신호와 같은 비정형 데이터로부터 원하는 감정정보를 추출하기 위해 이미지 프로세싱이나 음성 인식과 같은 난이도 있는 작업이 선행되어야 한다. 원하는 객체나 음성을 추출하고 그로부터 감정요소를 파악하는 것은 기술적 복잡성 뿐만 아니라 낮은 정확도가 문제가 된다[1]. 하지만 이에 비해 대화 문장에서 감정어를 추출하는 것은 현실적으로 활용될 수 있는 실현 가능한 접근 방법이다[2]. 대화 속에 존재하는

감정어들을 추출하면 등장인물이 드러내는 감정 상태를 파악할 수 있고, 스토리 정보로도 활용할 수 있을 것이다. 특히, 감정어를 이용하여 등장인물의 성향을 파악할 경우 스토리의 갈등 구조를 파악하는 데 많은 도움이 될 것이다. 이를 통해 영화 축약이나 브라우징과 같은 영화 정보 검색에 등장인물 정보를 활용할 수 있을 것이다.

이를 위해 본 논문에서는 등장인물의 대화에 존재하는 감정어를 추출하기 위해 WordNet을 이용한 감정어 추출방법을 제안하고, 추출된 감정어들을 긍정, 부정, 중립의 3종류로 분류하여 등장인물이 사용한 감정 성향이 무엇인지 표현하고 추출하는 연구를 진행한다. WordNet은 단어 간에 상위어와 하위어, 유사어 등의 관계로 연결된 네트워크 구조의 유의어 사전이다. WordNet에서 감정과 관련된 최상위의 감정 항목을 30개를 추출한 후에 네트워크 구조에서 최상위의 감정 항목과의 의미적 거리를 계산하여 단어별 감정량을 측정하게 된다. 추출된 감정어는 30차원의 감정 벡터로 표현되고, 감정어를 사용한 등장인물 또는 인물들 간의 관계에 따라 누적 연산되어 인물 또는 관계의 특징 감정 벡터로 활용되게 된다. 30차원의 감정 벡터를 그대로 사용하여 인물의 감정 성향을 표현할 경우 시각적으로 이해되기 힘들기 때문에 대표적인 3가지 요소인 긍정적 요소, 부정적 요소, 중립적 요소로 분류하여 누적하여 표현하였다.

이러한 방법을 통해 표현된 특징 감정 벡터를 이용하여 등장인물 또는 인물들 간의 관계에서 나타난 감정 성향을 표현하고 판단하였으며, 이의 효율성을 판단하기 위해 실험을 통해 감정 성향 판단 성능을 평가하였다.

이를 위해 본 논문에서는 2장에서 감정어 추출 방법과 스토리 가시화를 위한 스토리 모델링과 관련된 연구들을 알아보

고, 3장에서 제안하는 감정이 분류에 기반한 등장인물 감정 성향 판단 방법에 대해 설명한다. 그리고 4장에서 실험을 통해 평가된 감정 성향 추출 방법의 성능에 대해 기술한 후, 마지막으로 5장에서 결론을 맺는다.

## II. 관련 연구

등장인물 또는 인물 간의 감정 성향을 표현하고 판단하기 위해서는 감정이 추출 방법에 대한 이해와 스토리를 가지적으로 보여주는 방법에 대한 이해가 필요하다. 먼저 대사로로부터 감정을 추출하기 위해 활용되고 있는 관련 연구들에 대해 설명한 후에, 스토리 모델을 통해 스토리를 표현해주는 최근의 연구들에 대해 다음에서 설명한다.

### 1. 텍스트 기반의 감정이 추출 방법

어휘의 연관성을 측정하는 방법은 키워드 검색이 가지는 문제점들을 개선한 방법론이다. 어휘의 연관성을 이용해서 감정을 추출하거나 감정의 크기를 측정하는 것이 가능한데, 특정 단어가 어떤 감정을 포함하는 지에 대한 확률적 연관 정도를 측정할 수 있다. 어휘의 연관성 측정 방법에 기반하여 감정을 추출하고 감정량을 측정하기 위해 WordNet, SentiWordNet, WordNet-Affect를 활용한 다양한 시도들이 존재해 왔다[3,4,5].

단어의 감정을 분석하는 것은 대사에서 감정을 인식하기 위한 가장 기본이 되는 작업이다. 키워드 검색 방법은 이용 용이성이나 경제성으로 인해 단어의 감정 분석을 위해 사용되는 가장 범용적인 방법이다. 하지만 키워드 검색 기반의 감정이 분석 방법은 정해져 있는 감정이만을 추출하기 때문에 대상이 되는 감정이 수가 매우 적고, 직접적으로 감정을 표현하는 단어들을 대상으로 하게 된다. 즉 직접적 감정이만을 감정이로써 취급하게 되는데, 기존의 연구들로부터 취급되어진 감정의 개수는 대략적으로 700 단어 정도로 정리될 수 있다. Elliott는 198개의 감정을 제안하여 연구에 활용하였고[6], Salway는 WordNet을 이용하여 627개의 감정을 추출하는 연구를 진행하였다[7]. 이 기술들은 미리 정해진 감정에 속하는 감정이만을 추출하는 방법론으로 명확한 감정을 표현한다는 면에서는 명료성을 띄어서 좋지만, 내재적인 감정을 표현하는 단어들을 추출하기에는 한계가 있는 방법이다. 예를 들어 "I dreamed a nightmare"라는 대사가 있을 경우, 감정이 포함되어 있지만 키워드 기반의 검색 방법으로는 직접적인 감정을 포함하고 있지 않기 때문에 감정이나

감정량을 측정할 수 없게 된다.

키워드 기반의 감정이 추출 방법이 갖는 한계를 극복하기 위해 단어와 감정이와의 어휘의 연관성을 측정하여 내재적인 감정을 추출하는 연구가 시도 되었다[8]. 예를 들어 "nightmare"의 경우 직접적이진 않지만 내재적으로 "fear" 또는 "ambivalence" 감정과 관련을 갖는다. 특히나 단어의 감정의 정도는 어휘의 연관성을 측정하는 기법을 통해 측정될 수 있다. WordNet-Affective의 경우 특정 단어의 감정 정도를 측정하기 위해 어휘 연관성 측정 방법을 활용하였다. 어휘 연관성은 WordNet에 존재하는 특정 단어의 전체 의미들 (sense) 가운데 감정을 포함하는 의미들의 비율로 계산되어질 수 있다. 예를 든다면, 특정 단어의 전체 의미가 3개인데, 그중에 2개의 의미가 감정과 관련된 의미일 경우 단어의 감정과 관련된 어휘 연관성은 2/3이 된다. 따라서 어휘 연관성 방법에 의해 측정된 값은 0과 1사이의 값을 갖게된다. 하지만 어휘 연관성을 따지기 위해 의미에 감정이 포함되어 있는지를 판단해야 하는 문제와 함께 의미들이 바로 직접적으로 미리 정해진 감정에 연결되어 있을 때에만 유효한 방법론이다. 즉 여러 단계를 거쳐 감정에 연결되어 있는 단어들은 감정이로 추출하지 못하는 문제가 존재하게 된다. 예를 들어, "assaulting"이라는 단어의 경우 감정을 포함하고 있음에도 불구하고 "abivalence"에 3단계가 떨어져서 연결되기 때문에 기존의 어휘적 연관성을 통해서만 감정이로 추출되기 힘들다. 이의 개선을 위해 의미별로 여러 단계에 걸쳐 감정 항목을 검색하는 방법에 대한 연구가 필요하다.

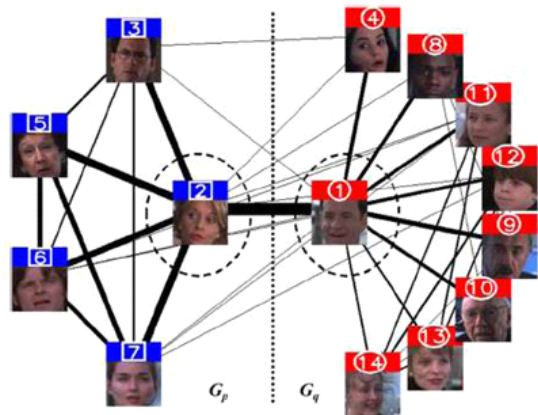


그림 1. 영화 'You've got mail'의 RoleNet  
Fig. 1. RoleNet of Movie 'You've got mail'

## 2. 소설 네트워크 기반의 스토리 모델링

최근 들어 소설 네트워크 방법론에 기반을 둔 영화의 스토리를 모델링하는 연구들이 진행되고 있다. 그중의 대표적인 연구가 그림 1과 2의 RoleNet과 Character-net에 대한 연구이다(9,10,11,12,13).

소설 네트워크는 인물들을 노드로 하고 인물들 간의 관계가 간선과 가중치로 표현되는 네트워크 구조를 지칭한다(9). RoleNet은 영화 내에서 장면을 기본단위로 하여 동시에 등장한 인물들을 간선으로 연결하고, 이를 누적하여 소설 네트워크로 표현하는 방법이고, Character-net은 영화 내에서 등장인물들을 노드로 하고 그들의 대화를 간선으로 하는 소셜 네트워크로 표현된다. 하지만 대사에서 감정어를 추출하여 등장인물과 인물 간의 관계에서 나타난 감정 성향을 판단하고 표현하는 연구에서는 장면을 기본 단위로 하는 것보다는 대사를 기본단위로 하는 것이 훨씬 정교하고 정확한 표현이 가능할 것이다.

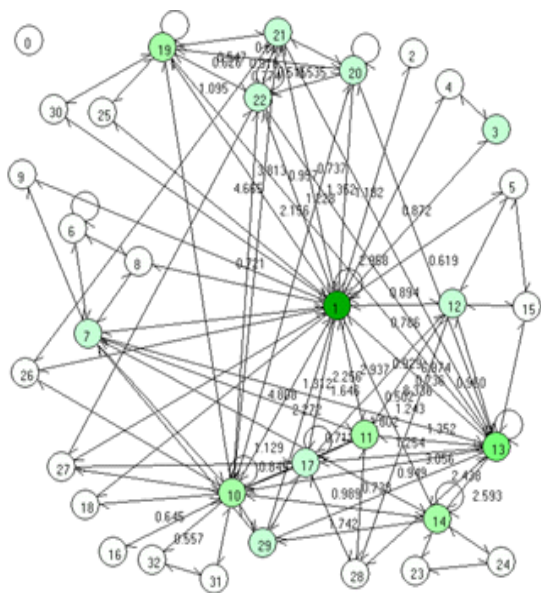


그림 2. 영화 'Avatar'의 Character-net  
Fig. 2. Character-net of Movie 'Avatar'

대본이나 자막을 이용하여 등장인물들 간의 대화를 누적하여 Character-net을 표현하는 연구를 진행하였다(10,11,12,13). 영화 전체에 대해 대화 그래프들을 누적하여 Character-net을 구축하였다. 대화를 기본 단위로 사용하여 인물 간의 관계를 표현하기 때문에 장면보다는 정확한 성능을 추출 성능을 나타내게 된다.

## III. 감정어 분류와 등장인물 성향 판단

스토리의 진행이나 변화를 표현하기 위해서는 등장인물들 간의 감정 변화에 주목해야 한다. 영화의 진행에 따른 등장인물들의 감정 변화 또는 누적량과, 등장인물들 간의 감정 변화를 보여주고 표현할 필요가 있다. 또한 등장인물의 감정 성향을 판단 하게 될 경우 스토리의 중요한 요소인 인물에 대한 심층적인 분석이 가능해 질 것이다. 이를 위해 WordNet 기반의 감정어 추출 방법을 제안하고 감정어 기반의 등장인물 감정 성향 판단 방법과 표현방법에 대해 설명한다.

### 1. WordNet 기반의 감정어 추출 방법

WordNet은 단어들을 상위어(hyperonym), 하위어(hyponym), 유사어(synonym) 등의 관계로 연결되어진 사전이다(14). 감정과 관련된 최상위의 용어는 'emotional state', 'emotion', and 'feeling'의 3개이다. 3개의 용어의 하위어들을 추출한 후, 유사한 용어들을 그룹으로 모아서 표 1과 같이 총 30개의 감정항목을 추출하였다. 이 30개의 감정항목은 감정어가 가질 수 있는 독립적인 감정 차원이므로 감정어는 식 1처럼 30차원의 벡터로 표현할 수 있다.

표 1. 30개의 감정항목  
Table 1 Emotion categories to be grouped

happiness	liking	dislike	anxiety
pride	expectation	unhappiness	ecstasy
gratitude	humility	apathy	desire
anger	ungratefulness	astonishment	glow
soul	sentiment	shame	gravity
calmness	fearlessness	humor	sympathy
ambivalence	embarrassment	pain	faintness
sensitivity	devastation		

예를 들어 joy의 경우 happiness에 synonym 관계로 연결되어 식 1과 같이 첫 번째 감정항목에만 1이라는 값을 갖는 30차원의 벡터로 표현된다.

$$EV(joy) = \{e_1, e_2, e_3, \dots, e_{30}\} = \{1, 0, 0, \dots, 0\} \quad (1)$$

감정어는 WordNet에서 상위어와 유사어를 검색하다보면 30개의 감정항목 중에 1개에 도달하게 된다. 그리고 감정항목에 도달하기까지의 거치는 단어수가 경로의 길이가 된다. 경로의 거리를 역수로 취한 것이 기본적인 감정어의 크기가 된다. 또한 도달한 감정항목에 크기값이 적용되어 감정벡터로 표현된다. 경로는 중간에 2~3개로 분기하여 서로 다른 감정항목에 도달할 수 있기 때문에 중간 용어가 갖는 총 의미(sense)들 개수 중에 감정항목에 연결된 의미의 개수로 나누어 주어 여러 감정항목으로 분기되는 감정어는 분기의 비율로 크기를 감소시켜준다. 이러한 개념들이 반영되어 계산되는 감정어 크기의 계산은 식 2와 같이 표현된다.

$$e_i = \frac{1}{D(w \rightarrow EC_i)} \times \frac{m}{t} \quad (2)$$

$w$ : 단어,  $EC_i$ :  $i$ 번째 감정항목  
 $D(w \rightarrow EC_i)$ : 단어( $w$ )부터 감정항목( $EC_i$ )까지 거리  
 $m$ : 감정항목( $EC_i$ )까지 연결된 의미(sense)의 개수  
 $t$ : 단어( $w$ )의 총 의미 개수

예를 들어 “crazy”라는 단어는 1번째 감정항목인 “happiness”와 10번째 항목인 “anger”에 연결된다. 그리고 감정항목까지의 거리는 4이고, 전체 의미의 개수는 6이고, 각 감정항목에 연결된 의미의 개수는 1이다. 따라서 감정벡터 중 1번째와 10번째인  $e_1$ 과  $e_{10}$ 의 값은 각각 1/24가 된다. 이렇게 계산된 “crazy”의 감정벡터는 식 3과 같이 표현된다.

$$EV(crazy) = \{1/24, 0, 0, 0, 0, 0, 0, 0, 0, 1/24, \dots, 0\} \quad (3)$$

감정어의 누적을 위해서는 감정어들 간의 합 연산이 가능해야 한다. 감정어는 벡터로 표현 가능하므로 식 4와 같이 벡터의 합 연산을 그대로 적용할 수 있고, 등장인물의 모든 감정어를 누적할 경우 식 5와 같이 표현된다.

$$EV_m = EV_a + EV_b = \{e_{a1} + e_{b1}, e_{a2} + e_{b2}, \dots, e_{a30} + e_{b30}\} \quad (4)$$

$$EV_k = \sum_{i=1}^n EV_i(character_k) \quad (5)$$

$n$ : 등장인물 'k'의 감정어의 개수

2. 인물 또는 인물관계의 감정성향 가시화 방법  
 영화들에 대해 그림 10의 Character-net을 구축하였다

[10]. 장면 내의 등장인물 명(character name)과 대사(dialog)을 이용하여 대화 그래프를 표현하고, 대사의 글자수를 가중치로 사용하였다. 표현된 대화 그래프들은 누적되어 Character-net을 구축하게 된다.

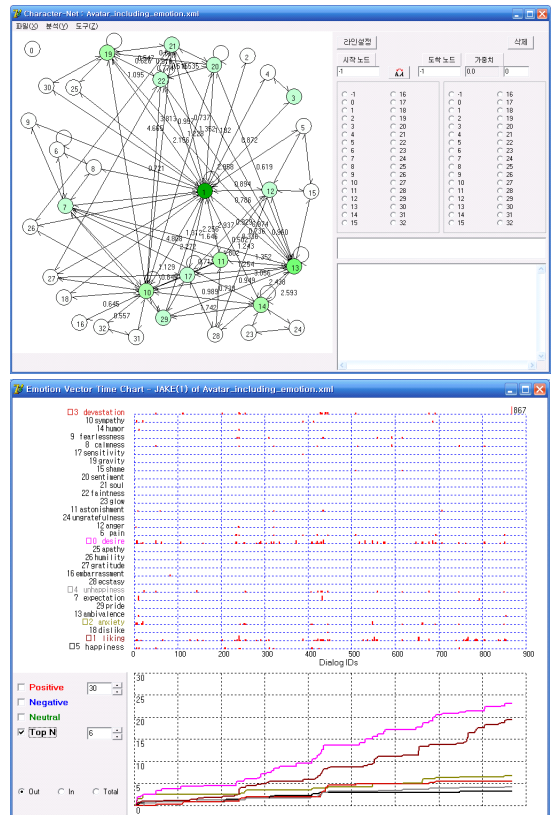


그림 3. 영화 'AVATAR'의 등장인물 'Jake'의 감정 벡터 그래프  
 Fig. 3. Emotion Vector Graph of Character 'Jake' in a Movie 'AVATAR'

또한 등장인물 별로 감정벡터와 그래프를 그림 3의 아래 그림과 같이 표현하였다. 그림 3에서 윗 그림은 character-net을 표현한 것이며, 아랫 그림은 감정벡터 중에서 상위 6개의 감정 항목의 대사 진행에 따른 누적 변화량을 그래프로 그린 것이다. 등장인물 별로 사용되는 감정 벡터는 등장인물별 대사들에 대한 감정 벡터들을 누적인 값을 사용하였다. 즉 등장인물이 대사를 통해 갖는 모든 감정 성분을 통합하여 등장인물의 특징 벡터로 사용하였다.

감정 벡터 그래프를 보는 방법은 특정 등장인물을 선택하거나 등장인물들 간의 연결선을 선택하여 관계에서 나타난 감정벡터 그래프를 표현하여 보는 방법을 지원한다. 특정 등장인물의 감정의 변화를 보거나, 인물들 간의 관계 변화를 추적

하기 위해 제공되는 가시화 방법이다.

표 2. 감정항목의 분류  
Table 2. Classification of Emotion Categories

분류	해당 감정항목
긍정	happiness, liking, expectation, pride, ecstasy, gratitude, astonishment, glow, humor, sympathy
부정	dislike, anxiety, ambivalence, unhappiness, embarrassment, apathy, pain, anger, ungratefulness, shame, devastation
중립	humility, desire, faintness, soul, sentiment, gravity, sensitivity, calmness, fearlessness

### 3. 감정어 분류를 통한 등장인물의 감정 성향 표현 및 판단

WordNet으로부터 얻어진 감정어 항목들을 표 2와 같이 긍정, 부정, 중립의 3개로 분류하였다. 이것을 등장인물의 대화에서 추출된 감정어들에 적용하여 등장인물의 성향이 긍정적인지 부정적인지 중립적인지를 분류하였다. 감정 벡터로 표현된 감정어들을 등장인물에 대하여 합산한 후에 감정 벡터의 값들을 긍정, 부정, 중립의 3가지 항목으로 분류한 후 합산하였다.

이것을 그림 4와 같이 긍정, 부정, 중립의 3축의 그래프로 인물의 감정 성향을 표현하였다.

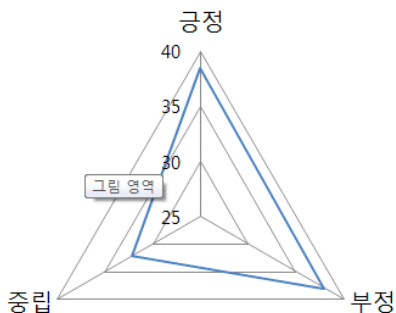


그림 4. 등장인물의 성향 그래프  
Fig. 4. Graph of Character Type

## IV. 실험

제안된 방법의 성능을 평가하기 위해 표 3과 같이 4개의 영화에서 12명의 주요 등장인물을 선정하였다. 구현과 평가

를 위한 소프트웨어는 마이크로 소프트 사의 윈도우 XP 운영 체제 하에서 Delphi (Delphi 7.0, Embarcadero) and JAVA (JAVA 1.7, Eclipse Indigo)를 사용하여 개발하였다. Delphi는 user interface와 Character-net 구축과 분석을 위해 사용되었으며 JAVA는 감정어 추출과 감정벡터 표현을 위해 사용되었다.

표 3. 영화 데이터  
Table 3. Movie Data

ID	영화 제목	장르
M1	AVATAR (2009)	Action Adventure Fantasy
M2	A Few Good Men (1992)	Crime Drama Mystery
M3	The Departed (2006)	Crime Drama Thriller
M4	Die Hard (1988)	Action Thriller

영화별로 주연과 중요한 조연을 추출하였다. 모든 등장인물들에 대해 분류작업을 진행하지 못한 것은 조연 이하의 등장인물의 대사가 많지 않아서 추출된 감정어량이 절대적으로 부족하였기 때문이다. 인물별로 감정어를 분석하여 면 <표 4>와 같은 결과를 얻게 된다. <표 4>의 제안된 방법론에 의해 자동측정된 결과와 전문가에 직접 평가한 수동측정의 결과를 비교하였다. 자동측정된 결과의 판정은 자동측정된 3가지 값들이 수동측정된 3가지 값들의 15% 범위 내에 있을 때 올바르게 판단(OK)된 것으로 결정하였다. 범위 밖에 있는 경우에는 잘못되게 판단(NG)된 것으로 결정하였다.

표 4. 실험 결과의 평가  
Table 4. Evaluation of Experimental Result

영화명	인물명	감정성향		판단결과
		제안된 방법에 의한 측정	수동측정	
M1	Jake	+ :29.0 - :21.4, N: 27.8	+ :26.5, - :23.4, N: 25.4	OK
	Grace	+ :15.0 - :17.1, N: 11.1	+ :12.1 - :16.2, N: 8.4	NG
	Quaritch	+ :9.3 - :13.9, N: 12.3	+ :8.5 - :14.4, N: 10.9	OK

M2	Kaffee	+ :54.1 - :55.6, N: 41.9	+ :59.3, - :52.2, N: 40.8	OK
	Jo	+ :12.8 - :19.7, N: 13.0	+ :13.5 - :19.3, N: 8.8	NG
	Jessep	+ :12.2 - :17.8, N: 12.0	+ :12.0 - :16.9, N: 11.0	OK
M3	Billy	+ :24.9 - :26.1, N: 21.2	+ :27.1, - :25.1, N: 21.9	OK
	Dignam	+ :6.9 - :2.5, N: 4.0	+ :7.0 - :2.5, N: 6.6	NG
	Costello	+ :15.1 - :11.4, N: 26.7	+ :14.9 - :12.3, N: 24.0	OK
M4	McClane	+ :45.6 - :57.3, N: 34.9	+ :44.1, - :55.8, N: 34.3	OK
	Powell	+ :11.1 - :9.3, N: 14.0	+ :10.9 - :8.5, N: 14.6	OK
	Hans	+ :15.3 - :19.1, N: 15.1	+ :15.9 - :20.3, N: 14.7	OK
정확도				75%

+ : 긍정(Positive), - : 부정(Negative), N: 중립(Neutral)

12명의 등장인물에 대해 감정 성향을 판단하는 방법론에 대한 실험 결과는 9명의 등장인물을 올바르게 판단하여 75%의 정확도를 나타내는 것으로 평가되었다. 7명에 대해서는 잘못된 감정 성향을 판단하는 것으로 나타났다.

3명은 잘못된 감정 성향 분류의 경우 M1의 Grace, M2의 Jo, M3의 Dignam에서 나타났다. 이들의 경우 기본적으로 대화량이 적고, 추출된 감정의 수가 절대적으로 적어서 추출 성능에 오류가 발생했을 것으로 판단된다. 또한 감정이 추출 시에 잘못된 경로를 추적하여 잘못된 감정이나 감정없이 측정되는 경우가 발생하였다. 이의 개선을 위해 WordNet 기반의 감정의 추출 방법론의 성능을 개선하는 작업과 대사 이외의 감정을 판단하기 위한 추가적인 정보를 추출하는 연구를 진행할 필요가 있다.

추가적으로 기존의 Salway의 방법과 비교를 진행하려고 하였으나, 627개의 감정이만을 미리 선정하여 대사에서 동일한 감정이만을 추출하는 Salway의 방법은 대사에서 발견되는 감정의 수가 너무 희박하여 비교를 진행할 수가 없었다. M1과 M4의 경우 주인공인 Jake와 McClane의 경우 일정량 이상의 감정이 추출되었으나 본 논문에서 제안된 방법에 의해 측정된 방법보다 현저히 적은 수가 발견되어 비교평가를 진행할 수가 없었다.

## IV. 결론

본 논문에서 감정을 이용한 등장인물의 감정 성향을 판단하는 방법론을 제안하고 실험을 통해 그 효율성을 보였다. 4개의 영화 속의 12명의 주요 등장인물들에 대해 감정 성향을 판단하는 실험을 진행하여 75%의 정확도로 평가되어지는 것을 보여주었다. 즉 제안하는 방법론이 효율적으로 등장인물들의 감정 성향을 표현하고 판단할 수 있다는 것을 증명하였다.

또한 영화의 등장인물 관계를 소셜 네트워크 모델인 Character-net으로 표현하고, 영화와 등장인물과 등장인물들간의 대화들에 대해 감정 벡터와 감정벡터의 변화그래프를 보여줄 수 있는 가시화 도구를 구현하였다. 이를 통해 사용자는 영화에서 감정이 집중되어 나타나거나 감정의 분포가 변화하는 부분들을 관측할 수 있을 것으로 판단된다.

감정 성향 판단 결과에서 3명의 판단에서 오류가 발생하였다. 이것은 적은 대사량과 감정량을 갖는 경우로 인해 발생한 것으로 추측된다. 이의 개선을 위해 WordNet 기반의 감정의 추출 방법론의 성능을 개선하는 작업과 대사 이외의 감정을 판단하기 위한 추가적인 정보를 추출하는 연구를 진행할 필요가 있다. 따라서 향후에 WordNet의 감정이 추출 방법론의 개선과 지문이나 상대방의 감정의 반영과 같은 추가적인 연구를 진행할 것이다.

## 참고문헌

- [1] K.-E. Ko, and K.-B. Sim, "Development of Context Awareness Service Inference Method using Multi-Modal Emotion Recognition System," Autumn Conference on Korea Institute of Intelligent System, Vol. 18, No. 2, pp. 261-264, Oct. 2008.
- [2] S.-B. Park, E. You, and J.J. Jung, "Potential Emotion Word in Movie Dialog," Proceedings of the International Conference on IT Convergence and Security 2011, pp. 507-516, Dec. 2011.
- [3] Yassine, M. and Hajj, H., "A Framework for Emotion Mining from Text in Online Social Networks," Data Mining Workshops (ICDMW), 2010 IEEE International Conference on, pp. 1136-1142, 2010.

- [4] Strapparava, C. and Valitutti, A., "WordNet-Affect: an Affective Extension of WordNet," In Proceedings of the 4th International Conference on Language Resources and Evaluation, pp. 1083-1086, 2004.
- [5] Esuli, A. and Sebastiani, F., "SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining," In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06), pp. 417-422, 2006.
- [6] Elliot, C., "The Affective Reasoner: A Process Model of Emotions in a Multi-agent System," PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.
- [7] Salway, A., Graham, M., "Extracting Information about Emotions in Films," In Proceedings of the eleventh ACM international conference on Multimedia (MULTIMEDIA '03), pp. 299-302, 2003.
- [8] Liu, H., Lieberman, H. and Selker, T., "A Model of Textual Affect Sensing Using Real-World Knowledge," In Proceedings of the 2003 International Conference on Intelligent User Interfaces, pp. 125-132, 2003.
- [9] C. Y. Weng, W. T. Chu, and J. L. Wu, "RoleNet: movie analysis from the perspective of social network," IEEE Transaction on Multimedia, vol. 11, no. 2. pp. 256-271, 2009.
- [10] S.-B. Park, K.-J. Oh, and G.-S. Jo, "Social Network Analysis in a Movie using Character-net," Multimedia Tools and Applications. Vol. 59, No. 2, pp. 601-627, 2012. 7.
- [11] J. Kaminski, and M. Schober, "Social networks in movies," COINs Conference, pp. 1-3, 2011.
- [12] S.-B. Park and G.-S. Jo, "Role Grades Classification and Community Clustering at Character-net," Journal of the Korea Society of Computer and Information, vol. 14, n. 11, pp. 169-178, Nov. 2009.
- [13] W. Kim, S.-B. Park, G.-S. Jo, "Improvement of Character-net via Detection of Conversation Participant," Journal of the Korea Society of Computer and Information, vol. 14, n. 10, pp. 241-249, Oct. 2009.
- [14] G.A. Miller, "WordNet: A Lexical Database for English," Communications of the ACM, Vol. 38, No. 11, pp. 39-41, 1995.

### 저 자 소 개



#### 백 영 태

1989 : 인하대학교  
전자계산학과 이학사

1993 : 인하대학교  
전자계산공학과 공학석사

2002 : 인하대학교  
전자계산공학과 공학박사

1993-1998 : 대상정보기술(주)  
정보통신연구소 선임연구원

1998-현재 : 김포대학교  
멀티미디어과 부교수

관심분야 : 멀티미디어 정보검색,  
웹교육시스템, 모바일시스템

Email : hannaek@kimpo.ac.kr