

Character-net에서 배역비중의 분류와 커뮤니티 클러스터링

박승보*, 조근식**

Role Grades Classification and Community Clustering at Character-net

Seung-Bo Park *, Geun-Sik Jo **

요 약

동영상으로부터 원하는 정보를 검색하려는 다양한 연구들이 있어왔다. 하지만 기존의 연구들은 동영상의 스토리에 대한 고려 없이 특정한 사물의 인식이나 사물간의 관계 정보만을 추출하여 검색에 이용하였다. 동영상에서 정확한 정보를 검색하기 위해서는 스토리의 주축이 되는 등장인물과 등장인물의 커뮤니티에 기반을 둔 연구가 반드시 필요하다. 따라서 본 논문은 등장인물에 기반을 둔 동영상 정보검색 방법을 기술한다. 등장인물들은 서로 대화를 통해 관계를 맺으며 스토리를 진행시킨다. 등장인물들 간의 관계는 배역의 비중을 분류하고 등장인물들이 이루는 커뮤니티를 구분하여 스토리에 형성되어 있는 관계를 분석할 수 있다. 이를 위해 본 논문에서 등장인물들의 관계를 묘사할 수 있는 Character-net을 제안하고 Character-net으로부터 등장인물들의 배역의 비중을 분류하고 등장인물들이 이루는 커뮤니티를 클러스터링 하는 방법에 대해 기술하고 실험을 통해 그 효용성을 입증한다.

Abstract

There are various approaches that retrieve information from video. However, previous approaches have considered just object information and relationship between objects without story information to retrieve contents. To retrieve exact information at video, we need analyzing approach based on characters and community since these are body of story proceeding. Therefore, this paper describes video information retrieval methodology based on character information. Characters progress story to form relationship through conversations. We can analyze the relationship between characters in a story with the methods that classifies role grades and clusters communities of characters. In this paper, for these, we propose the Character-net and describe how to classify role grades and cluster communities at Character-net. And we show this method to be efficient.

- ▶ Keyword : 멀티미디어 정보검색(multimedia information retrieval), 캐릭터넷(Character-net), 배역비중(role grade), 커뮤니티 클러스터링(community clustering), 스토리기반의 정보검색(story-based information retrieval), 사회연결망(social network)

• 제1저자 : 박승보

• 투고일 : 2009. 10. 06, 심사일 : 2009. 10. 21, 게재확정일 : 2009. 11. 20.

* 인하대학교 박사과정 ** 인하대학교 컴퓨터정보공학부 교수

※ 본 연구는 한국과학재단을 통해 교육과학기술부의 세계수준의 연구중심대학육성사업(WCU)으로부터 지원받아 수행되었습니다. (R33-2008-000-10109-0)

I. 서론

동영상에서 원하는 정보를 검색하거나 중요한 장면을 알아내는 것과 같은 멀티미디어 정보검색 분야는 영상 분석에서 중요한 분야이며 다양한 연구가 이루어져 왔다. 멀티미디어 정보검색의 분야는 사물의 단순한 정보를 파악하는 방식과 사물간의 관계를 추출하는 상황기반의 분석을 하는 방식이 있다. 또한 최근 들어 스토리 기반의 정보검색을 시도하는 연구들이 이루어지고 있다. 스토리 기반의 정보검색은 스토리 진행을 실행하는 개체를 추출하고 개체들 간의 관계를 알아내어 동영상의 스토리 요소를 묘사하여 분석하는 방식이다.

기존의 연구는 특정한 사물의 인식이나 사물간의 관계 정보만을 추출하여 검색에 이용하거나 동영상 축약과 같은 특수한 목적을 위한 스토리 모델을 제안하여 연구를 진행하였다. 따라서 스토리 기반의 동영상 정보 검색을 위해 범용적인 스토리 모델을 작성할 필요가 있다. 또한 스토리 모델링을 위해 동영상에 등장하는 객체들 간의 관계를 정의하고 추출할 필요가 있다. 이를 위해 본 연구에서는 등장인물 정보를 이용한 스토리 기반의 정보검색 방법을 제안한다. TV 드라마나 상업용 영화와 같은 동영상의 경우 등장인물들에 의해 스토리가 진행된다. 동영상에 등장하는 동영상의 등장인물은 사건의 주체이고 등장인물들 간의 관계로 사건의 진행이 이루어진다. 동영상에서 등장인물들의 관계는 대화를 통해 주로 이루어지므로 대화를 기반으로 관계를 설정할 수 있다. 즉 등장인물이 누구인지 찾아내고 등장인물들 간의 대화에 대한 정보를 누락한다면 동영상의 스토리에 관한 분석을 할 수 있다는 것이다. 이렇게 동영상에서 등장인물을 추출하고 대화를 기반으로 등장인물들 간의 관계를 묘사한 그래프를 Character-net이라 한다[1]. 본 논문에서는 동영상을 Character-net으로 표현하는 방법을 제안하고, 이를 이용해서 동영상의 스토리 정보를 검색하는 방법을 제안한다. 동영상 스토리 정보검색은 Character-net에서 등장인물들의 배역비중을 분류하고 배역을 중심으로 한 커뮤니티의 클러스터링을 통해 이루어진다. 동영상에서 자주 만나는 등장인물들끼리 서로 그룹을 이루면서 스토리가 진행되며 이 그룹을 커뮤니티(community)로 하고 커뮤니티를 분류하는 것을 본 논문에서 커뮤니티 클러스터링(community clustering)이라 지칭한다. 스토리 기반의 정보검색 방법은 기본적으로 영상에 존재하는 객체들의 정보를 추출하거나 인식할 수 있다는 가정 하에 진행이 되고 있다. 따라서 본 연구는 현재에도 일정정도 실행되어 있는 얼굴인식 기술과 화자인식 기술을 이용하여 등장인물을 파악할 수 있다는 가정 하에서 연구를 진행하였다[2].

제안된 방법론의 설명을 위해 본 논문은 2장에서 동영상의 내용 분석 방법에 관련된 연구와 배경지식에 대해 설명하고 3장에서 Character-net 표현 방법과 정보검색 방법에 대해 설명한다. 그리고 4장에서 제안된 방법들을 실험을 통해 그 성능을 평가하고 마지막 결론에서 본 논문의 의의와 향후 과제에 대해 기술한다.

II. 관련 연구

1. 동영상의 구조와 동영상 정보검색 방법

동영상은 이미지들이 연속적으로 나열되어 구성된다. 이때 한 장의 이미지를 프레임(frame)이라고 한다. 프레임들이 모여 샷(shot)을 이루게 되는데 카메라가 끊이지 않고 한 번에 촬영한 프레임들의 집합을 샷이라 한다. 샷은 대상물의 동작이나 배경에 대한 단순한 묘사를 할 때 사용되며 장면(scene)은 한 장소에서 촬영된 연속된 시간의 흐름을 갖는 샷들의 모임이다. 샷들은 장면을 이루기 전에 소규모로 모여서 한 개의 대상물이나 행동을 표현하게 되는데, 이 소규모의 샷들의 집합을 그룹이라 한다. 예를 들어 회의를 하는 장면이 있고 회의를 구성하는 여러 개의 대화가 있을 때 회의 전체가 장면에 해당하고 각각의 대화는 그룹에 해당한다[3].

멀티미디어의 한 종류인 영화나 TV 드라마와 같은 동영상에서 원하는 정보의 검색방법은 내용기반의 정보검색 방법과 상황정보 검색 방법과 스토리기반의 정보검색방법으로 나누어질 수 있다.

첫 번째의 내용기반의 정보검색 방법은 동영상의 이미지나 소리에 나타나는 신호적인 특성을 가지고 멀티미디어에 존재하는 객체정보를 추출하는 방식이다. 연속된 영상에서 색상이나 명암과 휘도의 변화값이 큰 부분을 추출하여 분석하거나 텍스트와 음성신호에서 특정 패턴을 찾아내어 원하는 정보를 추출하는 방식으로 유사한 장면을 찾거나 운동선수의 움직임 추적하거나 얼굴인식과 등장인물과 같은 객체 정보를 추출하는 연구들이 여기에 해당한다[2][5][6][7]. 이 연구방법은 멀티미디어 정보검색의 가장 기초가 되는 연구 분야로 멀티미디어에 존재하는 다양한 객체들을 추출하고 인식하는 연구이다.

두 번째의 상황정보 추출방법은 사물간의 관계를 추출하여 분석에 사용하는 방식이다[8][9]. 동영상에서 상황정보는 사물 간에 동적으로 만들어지는 관계를 의미한다. 이 상황정보를 추출하기 위해서는 먼저 사물들에 대한 객체 인식이 이루어져야하며, 사물들 간의 관계를 파악할 수 있어야 한다. 이러한 과정을 거쳐 추출되는 상황정보는 등장인물들의 대화의 형태가 될 수도 있고[2][8] 사물들 간의 위치적 변동 상황일 수도 있다[9].

마지막의 스토리기반의 정보검색 방법은 최근에 연구가 이루어

지고 있는 분야로 스토리 요소를 추출하고 이를 통해 주요한 스토리 정보를 검색 하는 방식이다. 동영상의 스토리 정보검색 연구는 먼저 스토리에 대한 모델을 구성하고 구축된 모델에서 필요한 정보를 분석하거나 검색하는 방법을 취한다. Byunghee Jung 등 [10]은 그들의 연구에서 장면들의 의미를 장면 내에서 벌어지는 등장인물들의 대화나 액션(action) 상황으로 정의한 후 장면간의 상관관계를 계산하여 그래프로 표현하는 스토리 모델을 제안하였다. 그리고 전체 그래프에서 하위그래프를 찾는 방식으로 동영상 축약을 제작하였다. 하지만 장면이 모델링을 위한 그래프의 최소 단위가 되어 중요한 인물이나 슷을 찾기에 부적합한 방식이다.

동영상에 나타나는 정보들을 중요도에 따른 판단을 하기 위해서는 스토리에 대한 모델을 구축할 필요가 있고 이에 대한 구축 방법 연구가 Character-net 연구이다[1]. 그리고 이를 발전시켜 Character-net으로부터 배역들의 비중과 스토리의 구성원인 커뮤니티의 클러스터링과 같은 스토리 정보검색 방법에 대한 연구를 본 논문을 통해 설명한다. 위에서 설명한 동영상 정보검색 방법을 정리하면 표 1과 같다.

표 1. 동영상 정보검색 방법
Table 1. Video Information Retrieval

	내용기반	상황정보	스토리기반
사용기술	신호특성분석	객체 인식	스토리 모델링
사용목적	객체 정보 추출, 유사장면 검색, 슷 경계 추출	객체 정보 추출, 객체 간의 관계 파악	주요장면 또는 주요인물 추출, 스토리 요소(인물, 시간, 배경) 추출

III. Character-net에서 배역 비중의 분류와 커뮤니티 클러스터링

3.1 Character-net

Character-net은 등장인물 간의 대화 그래프를 누적하여 사회연결망 형태로 표현한 것으로 스토리기반의 정보검색 방법이다[1]. 사회연결망 형태의 Character-net에서 정점(vertex)은 등장인물을 표현하며 간선(edge)은 등장인물간의 대화정도를 의미한다. Character-net에서 간선의 형태는 가중치를 갖는 방향성 화살표를 사용한다. 이는 대화의 특성상 화자와 청자가 존재하여 화자로부터 청자에게 향하는 방향성을 갖고 대화 시간이 존재하기 때문이다. Character-net은 등장인물들과 대화정도를 의미하는 정점들과 간선들로 구성된 그래프 $G(V, E)$ 를 의미한다. 영화나 TV드라마와 같은 동영상에서 등장인물은 스토리를 진행하는 중심축이다. 따라

서 등장인물을 알아내는 것은 매우 중요한 작업이며 얼굴인식을 통해 등장인물을 식별할 수 있다[2][8]. 대화 정도는 등장인물 간의 1회 대화를 누적하여 계산한다. 본 논문에서 1회 대화의 정의는 대화 자막이 화면에 나타났다가 사라진 시간 동안의 영상을 의미하며 이 시간 동안에 화면에 나타난 모든 인물을 대화 참여자로 규정한다. 1회 대화는 3가지 요소로 구성되는 데, 화자와 청자와 대화량으로 구성된다. 이는 그림 1과 같이 2개의 정점(C_1, C_2)과 1개의 가중치(w_{12})로 구성된 그래프와 같다.

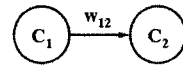


그림 1. 1회 대화의 그래프 표현
Fig. 1. Graph of a Conversation

그림 1의 화살표가 나가는 C_1 은 화자가 되며, 그 화살표가 도달하는 C_2 는 청자가 된다. w_{12} 는 C_1 과 C_2 사이의 대화량으로 대화 횟수와 대화 시간이라는 2가지 종류가 있다. 대화 횟수는 1회 대화마다 w_{12} 값을 1로 설정하는 방식이고, 대화 시간은 1회 대화가 이루어지는 시간을 w_{12} 값으로 할당하는 방식이다. 대화시간이란 동영상에 자막이 나타난 시간을 의미한다. 즉 자막이 2초간 유지되었다면 가중치 w_{12} 는 2라는 값을 갖는다.

대화 횟수 방식은 대화 1회가 1개의 정보를 전달한다는 관점에서 대화 횟수가 많을수록 많은 정보가 전달된다는 것을 의미하지만 모든 1회 대화의 정보량이 동일하지 않기 때문에 단순 누적할 경우 적절한 대화의 양을 대변하지 못하는 단점이 있다. 대화 시간을 사용할 경우 대화시간이 많으면 주요인물이라는 것을 파악할 수 있으나 인물에 따라 말 빠르기가 다르므로 정보량을 정확히 지시하지는 못하는 단점이 있다.

먼저, Character-net을 설명하기 위해 본 논문에서 사용하는 표기들을 정리하면 표 2와 같다.

표 2. 제안하는 방법의 표기
Table 2. Notations for the Proposed Method

기호	설명
g_h	h번째 그룹
Cg_h	h번째 그룹의 등장인물의 집합
$cg_{h,i}$	h번째 그룹의 i번째 등장인물
C	동영상의 등장인물 집합
c_i	동영상의 i번째 등장인물
w_{ij}	동영상의 i번째 등장인물에서 j번째 등장인물에게 설정되는 대화 가중치
DC_i	동영상의 i번째 등장인물의 연결정도 중심성

동영상에서 1회 대화는 여러 개의 샷들로 표현되며, 대화와 같은 행동을 표현하는 샷들은 그림 2처럼 1개의 그룹으로 묶여질 수 있다. 본 논문에서는 자막이 나타나는 1회 대화를 구성하는 샷들의 집합을 1개 그룹으로 정의하였고 자막이 나타나는 1회 대화 동안에 인식된 모든 얼굴들은 대화 참여자에 포함된다.

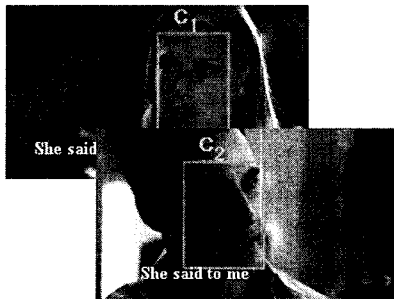


그림 2. 샷들의 그룹
Fig. 2. A Group of Shots

대화참여자인 화자와 청자는 얼굴인식과 화자인식 기술을 이용하여 판단할 수 있다. 얼굴인식 기술을 이용하여 현재 검출된 얼굴이 누구인지를 알아내고 입의 움직임이 있는지를 판단하여 화자인지 아닌지를 판단한다. 그리고 화자가 아닌 다른 얼굴은 청자로 판정한다[2][8]. 1회 대화에 해당하는 얼굴들 중에서 반복되는 동일얼굴은 동일한 등장인물로 결정하게 된다. 인식된 얼굴들은 입의 움직임에 의해 화자와 청자로 분리가 되어 그림 1처럼 정점과 화살표로 표현된다. 이때 화살표 상의 가중치인 대화량 w_{ij} 가 대화시간 또는 대화횟수로 설정된다.

그룹의 화자와 청자 및 대화의 조합에 따라 그래프는 그림 3과 같이 5가지 형태로 그려진다.

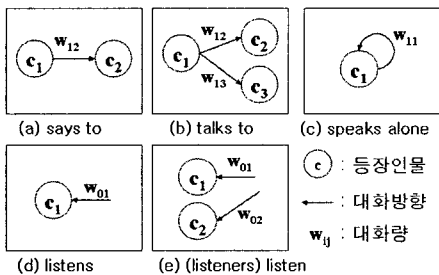


그림 3. 대화의 5가지 유형 그래프
Fig. 3. The Five Types of Conversation Graph

그림 3의 (a)는 화자와 청자가 각각 1명씩 인식된 상황으

로 화자가 청자에게 말하는 경우이다. (b)는 화자 1명이 청자가 다수인 경우로 연설장면이나 친구들 다수가 모여서 대화를 할 경우에 나타난다. (c)는 그룹 내에 화자 혼자 독백을 하는 경우이고, (d)는 화자가 검출되지 않은 상태로 자막은 존재하지만 청자만 인식된 경우로 전화 받는 장면이 여기에 해당한다. (e)는 화자가 검출되지 않은 상태에서 청자만 여러 명이 인식된 경우로 TV를 시청하는 상황이 여기에 해당한다.

위의 과정을 통해 1회 대화에 대한 그룹의 대화 그래프가 그려지면 그래프들을 누적하여 동영상 전체의 등장인물들 간의 그래프를 구축한다.

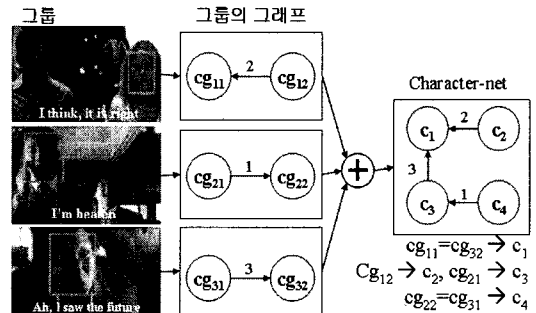


그림 4. Character-net의 구축
Fig. 4. The Construction of the Character-net

예를 들어 그림 4와 같이 그림의 좌측의 그룹들에서 화자와 청자를 판단하여 중간의 해당하는 그룹들의 대화 그래프들을 그린다. 그리고 대화 그래프를 연합하면서 누적하여 우측에 보이는 Character-net을 구축한다. 이때 얼굴인식을 이용하여 등장인물이 동일한 인지를 판단하여 대화 그래프의 정점들이 연결되도록 하고 중복되어 나타나는 등장인물 간의 가중치는 값을 누적시킨다. 이 과정을 거쳐 동영상은 그림 5와 같은 Character-net으로 변환된다. 그림 5에서 정점들은 등장인물을 의미하고 간선과 가중치는 대화의 방향과 등장인물 간의 누적 대화량을 의미한다.

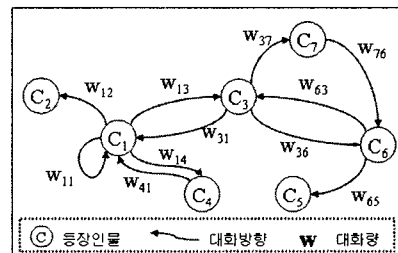


그림 5. Character-net
Fig. 5. The Character-net

3.2 배역비중의 분류

Character-net에서 주변의 등장인물들과 대화량이 많은 등장인물은 중요 인물로 분류할 수 있고 주변 인물과 대화량이 적은 등장인물은 중요하지 않은 인물로 분류할 수 있다. 이러한 방식으로 Character-net으로부터 등장인물의 비중을 측정할 수 있다. 이렇듯 동영상에서 등장인물의 중요도는 Character-net의 정점들과의 관계에서 추출할 수 있다. 중요도 분석 방법은 사회연결망의 연결정도 중심성(DC, Degree Centrality) 개념을 사용하였다. 연결정도 중심성 개념이란 다른 정점과 연결되어 있는 정도로 연결망에서 얼마나 중심의 위치에 있는지를 파악하는 개념이다[11]. Character-net에서 정점별 연결정도 중심성의 계산은 식 1에 의해 이루어진다.

$$DC_i = \left(\sum_{j=1}^{\alpha} w_{ji} + \sum_{k=1}^{\beta} w_{ik} \right) / \left(2 \times \sum_{j=1}^n \sum_{k=1}^n w_{jk} \right) \dots\dots\dots (1)$$

- n : 정점(등장인물)의 개수
- α : i 번째 정점으로 들어오는 정점의 개수
- β : i 번째 정점에서 나가는 정점의 개수

식 1은 i 번째 정점으로 들어오는 정점들의 가중치들과 나가는 정점들의 가중치들의 총합에서 모든 가중치들의 합으로 나누어 정규화 시켜주는 것을 의미한다.

영화나 드라마의 배역비중은 중요도에 의해 주요배역과 보조배역과 단역의 3가지로 분류된다. 주요배역은 극을 이끌어가는 역할로 다수의 등장인물과 만나서 다양한 교류를 하게 된다. 보조배역은 주요배역을 보조하여 극에 재미를 더하거나 사건의 실마리를 풀어나가는데 도움을 주는 역할이다. 단역은 특정 사건에만 출연하는 배역이다. 따라서 배역별로 이러한 특징을 적용하여 다음의 기준과 같이 배역비중을 분류한다.

- 주요배역 : 정점들 중에서 DC 평균값 이상이면서 정점 DC값들의 최고 간극 이상의 정점
 - 보조배역 : 정점들 중에서 DC 평균값이상이면서 정점 DC값들의 최고 간극 이하의 정점
 - 단역 : DC 평균값 이하의 DC값을 갖는 정점
- 배역비중을 분류하기 위해서 DC 평균값을 계산해야 하는데 DC 평균값은 식 2에 의해 구해진다.

$$Average_{DC} = \left(\sum_{i=1}^n DC_i \right) / n \dots\dots\dots (2)$$

주요배역들은 보조배역의 대화 빈도에 비해 월등히 높은

대화 빈도를 갖는다. 따라서 등장인물 별로 DC값을 계산하고 DC값을 기준으로 정렬하면 DC값이 가장 크게 변화되는 등장인물이 주요배역과 보조배역을 나누는 기준이 된다. 그림 6을 예를 들어 설명하면 먼저 DC 평균값(0.126)을 기준으로 이하의 값을 갖는 정점들 $C_5 - C_9$ 을 단역으로 분류한다. 그리고 DC 평균값 이상을 갖는 정점들 중에서 최대 간극을 갖는 C_2 와 C_3 를 주요배역과 보조배역을 나누는 기준으로 삼는다. 따라서 그림 6의 경우 C_1, C_2 는 주요배역이 되며, C_3, C_4 는 보조배역이 되고, $C_5 - C_9$ 는 단역으로 분류된다.

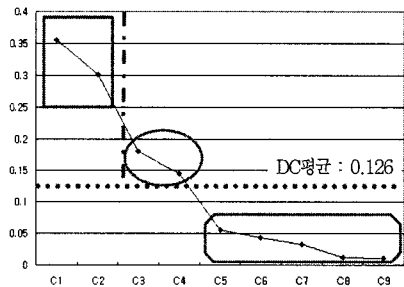


그림 6. 배역 비중의 분류
Fig. 6. The Classification of Role Grades

3.3 커뮤니티 클러스터링

동영상은 자주 만나는 등장인물들끼리 서로 그룹을 이루면서 스토리가 진행된다. 이렇게 형성된 그룹을 본 논문에서는 커뮤니티(community)로 지칭한다.

```

Input : V, E //V: Vertice, E: Edges
Output : C //C: Communities
MaCs = ExtractMajorCharacters(V,E) //주요배역 추출
MiCs = ExtractMinorCharacters(V,E) //보조배역 추출

CommunityCluster(V, E) //커뮤니티 클러스터링
begin
for ( i = 0 to n-1 )
begin
//i번째 정점이 주요배역 또는 보조배역 중에 있으면
if (V[i] in MaCs) or (V[i] in MiCs) then
begin
for ( j = 0 to n-1 )
begin
//j번째 정점이 i번째 주요배역과 연결되어 있으면
if V[j] is connected to V[i] then
V[j] is added to C[i] //커뮤니티 C[i]에 포함됨
//j번째 정점이 i번째 주요배역과 한방향선만을 가지면
if V[i] has only one direction between V[i]
V[j] is erased at C[i] //커뮤니티 C[i]에서 제거됨
end
//i번째 보조배역 주변에 형성된 커뮤니티가
if V[i] in MiCs) then
//기존 커뮤니티들 중에 1개 커뮤니티이라도 포함되면
if at least, one of C include C[i] then
C[i] is erased at C //i번째 커뮤니티는 삭제됨
end
return C
end
    
```

그림 7. 커뮤니티 클러스터링 알고리즘
Fig. 7. Community Clustering Algorithm

서로 자주 만나는 등장인물들은 스토리 중에 특정 사건이

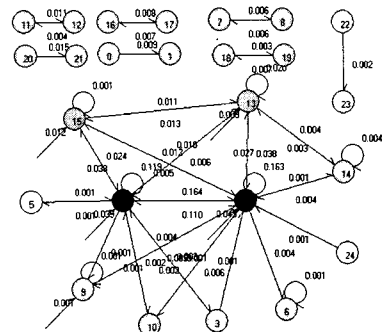
나 하위 스토리를 공유하게 되고 상호간에 영향을 주는 관계가 형성되어 양방향으로 방향선이 연결되는 커뮤니티 형태를 나타내게 된다. 커뮤니티끼리는 연결된 스토리를 공유하게 되므로 Character-net에서 커뮤니티의 발견은 하위 스토리의 개수와 스토리를 공유하는 등장인물들을 발견하는 것과 같다. 만약 동영상에서 2개의 커뮤니티가 찾아졌다면 스토리는 2개의 축으로 진행된다는 것을 의미한다. 커뮤니티는 주요배역 또는 보조배역과 밀접한 관계를 가진다. 주·보조배역은 그 주변의 등장인물들과 관계를 맺으며 스토리를 진행하므로 커뮤니티의 중심에 위치하게 된다. 따라서 커뮤니티를 분류하기 위해 먼저 주·보조배역을 찾는다. 그리고 직접 만나지 않은 사람이 스토리상으로 관련성을 갖지 않기 때문에 커뮤니티는 주·보조배역으로부터 직접 연결된 정점만을 범위로 한다. 그리고 주·보조배역과 양방향 연결을 갖는 정점만을 대상으로 하는데 그 이유는 단방향 연결은 일방적인 관계로 스토리 공유의 가능성이 낮기 때문이다. 설정된 커뮤니티 분류 규칙을 정리하면 그림 7과 같다.

IV. 실험 및 토의

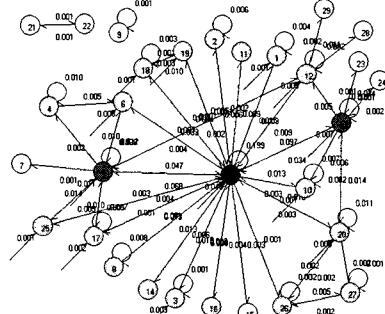
앞 절의 Character-net을 구축하기 위해 얼굴인식, 화자인식, 사회연결망 표현 등의 기술 요소들이 필요하다. 하지만 얼굴인식 기술은 다양한 얼굴 각도에 대해 신뢰할 수 있을 정도의 성능을 보이지 못하여 본 논문에서는 화자인식이 되었다는 가정 하에 수동으로 대화상황을 누적하였다. 그리고 Delphi 7.0으로 분석 툴을 개발하여 Character-net을 구축하고 분석하였다. Character-net의 구축과 분석을 위해 동영상 5개를 선택하여 실행하였다. 동영상은 영화중에서 '해리가 쉘리를 만났을 때(When Harry Met Sally...)'와 '에이스 벤츄라(Ace Ventura)'를 선택하였고 TV 시리즈로는 '히어로즈(Heroes)' 시즌1의 에피소드 4와 23과 '프렌즈(Friends)' 시즌3의 에피소드 1을 선택하였다. 이 동영상들은 대중적으로 많이 시청이 되었던 것들로 제안한 방법론이 일반적인 동영상의 스토리 정보검색을 하는데 적합하다는 것을 보이기 위해 선택하였다.

4.1 Character-net 구축

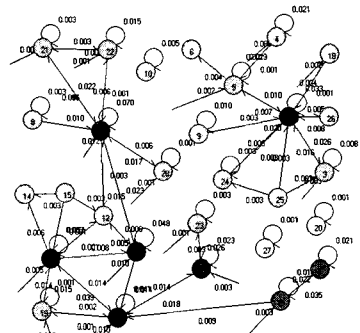
5개의 동영상에 대해 1회 대화 별로 그래프를 누적하여 그림 8과 같이 Character-net을 구축하였다. 가중치는 대화 시간을 이용하였고 Character-net의 정점별 등장인물명은 표 3에 표시하였다.



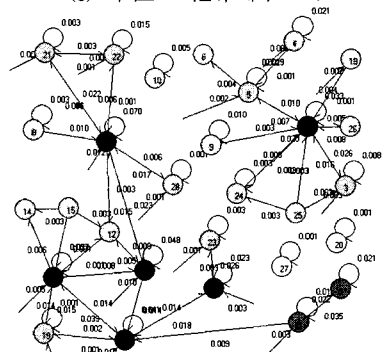
(a) 해리가 쉘리를 만났을 때



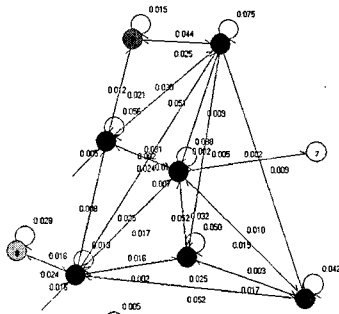
(b) 에이스 벤츄라



(c) '히어로즈' 시즌1, 에피소드 4



(d) '히어로즈' 시즌1, 에피소드 23



(e) '프렌즈'의 시즌 3의 에피소드1
 그림 8. Character-net
 Fig. 8. Character-nets

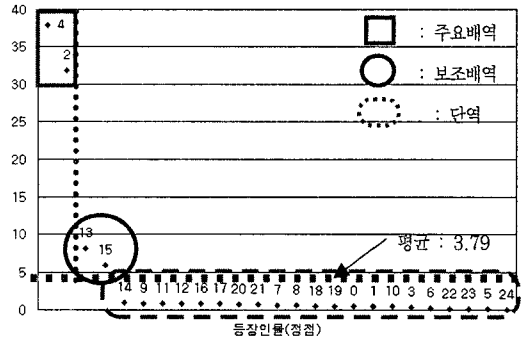


그림 9. '해리가 샬리를 만났을 때'의 배역비중 분류
 Fig. 9. Role Grades of 'When Harry Met Sally...'

표 3. 동영상의 정점번호별 등장인물 이름
 Table 3. Character Name of Each Vertex of Videos

번호	이름	번호	이름
해리가 샬리를 만났을때			
4	Harry	7	Niki Sanders
2	Sally	13	Mohinder Suresh
13	Marie	16	Peter Petrelli
15	Jess	2	Claire Bennet
14	Joe	11	Nathan Petrelli
9	Alice	17	Simone Deveaux
에이스 벤츄라			
0	Ace Ventura	1	Ando Masahashi
13	Lt. Lois Einhorn	5	Noah Bennet
5	Melissa Robinson	3	Brody Mitchum
12	Emilio	21	Tina
20	Dan Marino	히어로즈 S1, E23	
4	Riddle	0	Peter Petrelli
17	Roger Podacter	13	Noah Bennet
6	Woodstock	17	Claire Bennet
프렌즈 S3, E1			
4	Dr. Ross Geller	16	Angela Petrelli
0	Joey Tribbiani	1	Mohinder Suresh
1	Chandler Bing	2	Molly Walker
2	Monica Geller	5	Hiro Nakamura
5	Rachel Green	18	Charles Deveaux
3	Phoebe Buffay	8	Niki Sanders
6	Janice Litman	3	Matt Parkman
8	Jack Geller	15	Sylar

4.2 배역 비중의 분류

Character-net에서 배역비중의 분류는 3장에서 제안한 주요배역과 보조배역과 단역의 3단계로 하였으며 그 과정을 영화 '해리가 샬리를 만났을 때'에 적용하면 결과적으로 그림 9의 그래프가 그려진다.

그림 9를 위해 정점별로 DC값을 식 3을 사용하여 계산한 후 표 4와 같이 내림차순으로 정렬하였다.

표 4. '해리가 샬리를 만났을 때'의 정점별 DC값
 Table 4. DC Values at 'When Harry Met Sally...'

정점	등장인물	DC
4	Harry	37.855
2	Sally	31.907
13	Marie	8.097
15	Jess	5.895
14	Joe	0.977
9 & 기타	Alice & etc	0.852 이하

그리고 전체 정점들에 대한 평균값 3.79를 구하여 평균값 이하의 정점들을 단역으로 구분하였다. 단역의 정점들은 그림 10의 정점 간극 그래프를 그렸다. 간극 그래프는 정렬된 후 다음 정점과의 DC값 차이를 표시한 그래프로 주요배역과 보조배역 그룹 간에 현격한 차이가 나타나게 된다. 따라서 최고 간극을 보이는 정점을 기준으로 최고 간극 이상의 정점들은 주요배역으로, 이하의 정점들은 보조배역으로 분류하였다.

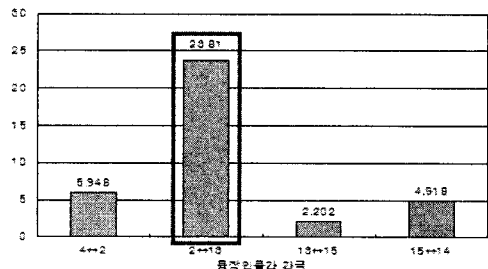


그림 10. '해리가 샬리를 만났을 때'의 정점별 간극
 Fig. 10. The Vertex DC Difference of 'When Harry Met Sally...'

그림 10에서 2번과 13번간 정점이 최고 간극을 형성하므로 최고 간극 이상인 2, 4번 정점을 주요배역으로 정하고 이하에 위치하는 13, 15번을 보조배역으로 정하였다. 이 과정들을 거쳐 그림 9와 같은 3가지 배역비중 분류가 이루어졌으며 실험에 사용된 동영상들에 대해 분류 결과를 정리하면 표 5와 같다. 분류 결과가 정확한지 판별하기 위해 영화의 경우 네이버 영화 사이트(12)의 분류 결과와 비교하였으며 TV 드라마의 경우 사용자 평가 결과와 비교하였다. 사용자 평가는 10명의 시청자를 대상으로 드라마 시청 후 주요배역과 보조배역과 단역을 분류하여 표시한 결과들을 평균하여 사용하였다.

표 5. 배역 비중의 분류
Table 5. Role Grades Classification

동영상	평가	주요배역	보조배역	단역
해리가 샐리를 만났을 때	제안	4,2	13,15	14,9,...
	네이버	4,2	13,15,	14,9,...
에이스 벤추라	제안	0	5,13,12,20	17,6,...
	네이버	0	13,5,12,20,4	17,6,...
히어로즈 S1, E4	제안	7,13,16,2,11	17,0,1,5	3,21,...
	사용자	16,2,7,0,13	11,1,5,17	3,21
히어로즈 S1, E23	제안	0	13,17,7,16,1,2,5,18,8	3,15,...
	사용자	0,13,7,17,5	8,3,2,1	16,18,...
프렌즈 S3, E1	제안	4,0,1,2,5,3	-	6,8,...
	사용자	4,0,1,2,5,3	6	8

배역비중의 분류 시 영화의 경우 네이버 영화의 배역분류와 유사한 결과를 얻을 수 있었다. 하지만 드라마의 경우 사용자 평가와 차이가 발생하였다. 이는 영화의 경우 내용의 진행이 기승전결이 뚜렷하고 배역비중이 등장시간에 비례하는데 반하여 드라마의 경우 연결되는 앞뒤의 에피소드에서 중요한 배역이 현재의 에피소드에서 대화 횟수가 적지만 극에 영향을 주는 배역으로 설정되는 경우가 많다. 따라서 드라마의 경우 배역 비중을 분석하기 위해 에피소드별로 비교하기 보다는 앞뒤의 에피소드와 연계하여 분석할 필요가 있으며, 대화량에 의한 분석뿐만 아니라 대화 내용에 대한 분석이 필요하다.

4.3 커뮤니티 클러스터링

3장의 3절에서 제시한 방법론에 따라 '해리가 샐리를 만났을 때'의 Character-net에서 커뮤니티를 클러스터링한 결과는 그림 11과 같다. 그림 11의 상단의 정점들은 포함되지 않았는데 영화의 중간마다 인터뷰하는 노인 부부들의 정점으로 극의 진행과 재미에 일조하는 부분은 있으나 주·보조배역과 관계성이 없는 등장인물이다.

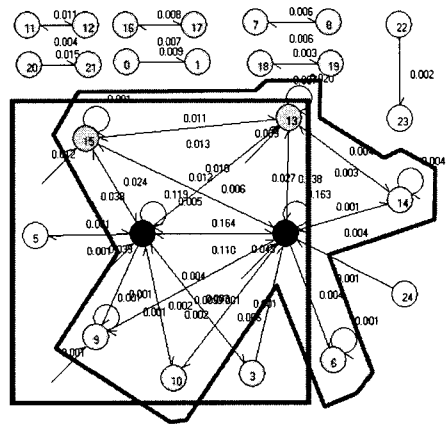


그림 11. '해리가 샐리를 만났을 때'의 커뮤니티
Fig. 11. Communities of 'When Harry Met Sally...'

3장의 3절의 방법으로 동영상들에 대해 커뮤니티 클러스터링을 한 결과를 정리하면 표 6과 같다.

표 6. 커뮤니티 클러스터링
Table 6. Community Clustering

동영상	커뮤니티	
	중심	해당 정점들
해리가 샐리를 만났을 때	4	2,13,14,6,10,9,15
	2	4,3,10,9,5,15,13
에이스 벤추라	0	5,6,18,19,2,11,1,12,13,10,20,26,15,16,3,14,8,17,25
	5	4,6,0,17,25,7
	13	0,12,23,24,20,10
	12	0,5,29,28,13
	20	0,13,27,26
히어로즈 S1, E4	7	21,22,28,11,12
	13	14,12,11,16,19
	16	19,13,17
	2	5,18,26,3,24
	11	12,7,16
	17	16,23
히어로즈 S1, E23	0	1,16
	5	6,4,2
	0	15,5,8,13,17,7,19,18
	13	0,3,1,2,12,17,7,16
	16	0,13,17,7,18
	1	3,9,8,11,2,13
프렌즈 S3, E1	5	0,15,6,14
	8	0,11
	4	1,7,3,5,2
	0	6,3,5,4,1
	2	1,0,4,5,3,8

이렇게 형성된 커뮤니티의 개수를 동영상에 따라 정리하면 표 7과 같이 나타난다. 결과가 정확하지 비교하기 위해 사용자 평가 결과와 비교하였다. 사용자 평가는 10명의 시청자를 대상으로 동영상 시청 후 표시한 결과들을 평균하여 사용하였

다. 제안된 방법의 결과와 사용자평가 결과가 비교적 유사하지만 0.87정도의 커뮤니티 개수 차이가 발생한다. 이유는 '해리가 샬리를 만났을 때'의 경우 제안된 방법이 동영상 전체에 해당하는 Character-net을 분석하였기 때문이다. 즉 동영상 중간에 형성되는 커뮤니티를 고려하여야 한다. 그리고 다른 동영상들의 차이는 커뮤니티 구성원들이 거의 유사한 경우 동일한 스토리를 공유하여 동일한 커뮤니티가 되었지만 알고리즘에는 그러한 내용이 반영되어 있지 못한 이유 때문으로 판단된다. 따라서 유사한 커뮤니티의 경우 지워줄 수 있는 추가적인 기준이 필요하다.

표 7. 커뮤니티의 개수
Table 7. The Number of Communities

동영상	커뮤니티 개수	
	제안(P)	사용자(U)
해리가 샬리를 만났을 때	2	2.82
에이스 벤추라	5	4.09
히어로즈 S1, E4	8	6.27
히어로즈 S1, E23	6	5.18
프렌즈 S3, E1	3	2.91
차이 평균 ($\sum P_i - U_i /5$)	0.87	

V. 결론

본 연구는 동영상 스토리 전개의 기본 요소인 인물과 사건에 대한 분석을 할 수 있는 Character-net 표현 방법을 제안하고 Character-net에서 주요배역, 보조배역, 단역의 3가지 배역비중을 분류하는 방법과 사건이 일어나는 등장인물의 모임인 커뮤니티를 클러스터링 하는 방법에 대해 제안하고 검증하였다. 등장인물의 배역비중을 나누는 방법은 각각의 동영상에 대해 연결정도 중심성의 분포를 이용하여 동적으로 분류할 수 있는 규칙을 제안하고 검증하였다. 또한 동영상 내의 스토리 속의 사건의 기반이 되는 커뮤니티를 Character-net의 연결망의 구조에서 자동으로 추출할 수 있는 규칙을 제안하고 검증하였다.

하지만 본 연구에서 제안한 방법들은 TV 드라마 보다는 영화에서 그 효용성이 더 검증되었는데 그 이유는 1개의 스토리 구조를 갖는 영화에 비해 TV 드라마는 다양한 스토리가 혼재되어 있고, 한편의 에피소드로 이야기가 종결되지 않고 이어지기 때문인 것으로 파악된다.

따라서 향후 연구에서는 Character-net으로부터 장면 단위에서 등장인물 간의 관계의 변화를 추적하는 연구가 필요하며 시간의 흐름에 따라 등장인물 간의 변화나 등장인물의 관

계의 유형 변화에 대한 세밀한 분석 방법론에 대한 연구가 필요하다. 또한 영화의 종류에 따라 스토리 정보를 검색하기 위한 다양한 요인이 존재할 것으로 판단되므로 장르에 따른 스토리 정보검색 방법에 대한 추가 연구가 필요할 것으로 판단된다.

참고문헌

- [1] Seungbo Park, Yoowon Kim, M. N. Uddin, Geunsik Jo, "Character-Net: Character Network Analysis from Video," 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, pp. 305-308, Sep. 2009.
- [2] M. Everingham, J. Sivic, A. Zisserman, "Taking the bite out of automated naming of characters in TV video," Image and Vision Computing, In Press, Corrected Proof, Available online, May 2008.
- [3] Y. Rui, T.S. Huang, S. Mehrotra, "Constructing Table-of-Content for Videos," to appear in ACM Multimedia Systems Journal, Special Issue Multimedia Systems on Video Libraries, Sep. 1999.
- [4] R. Zabih, J. Miller, K. Mai, "A feature-based algorithm for detecting and classifying scene break," ACM International Conference on Multi-media, pp. 189-200, Nov. 1995.
- [5] 김병기, "Feature Extraction for Automatic Golf Swing Analysis by Image Processing," 한국컴퓨터정보학회논문지, 제 11권, 제 5호, 53-58쪽, 2006년 11월.
- [6] 신성운, 표성배, "컬러와 형태에 기반을 둔 상표 영상 검색 시스템," 한국컴퓨터정보학회논문지, 제 11권, 제 3호, 167-172쪽, 2006년 7월.
- [7] S. H. Khatoonabadi, M. Rahmati, "Automatic soccer players tracking in goal scenes by camera motion elimination," Image and Vision Computing, Vol. 27, Issue 4, pp. 469-479, Mar. 2009.
- [8] 박승보, 김유원, 조근식, "얼굴인식을 이용한 동영상 상황 정보 어노테이션," 한국지능정보시스템학회, 2008 추계 학술대회 논문집, 319-324쪽, 2008년 11월.
- [9] L. Liang, G. Haifeng, L. Li, W. Liang, "Semantic event representation and recognition using syntactic attribute

graph grammar," Pattern Recognition Letters, Vol. 30, Issue 2, pp. 180-186, Jan. 2009.

- [10] Byunghee Jung, Taeyeong Kwak, Junehwa Song, Yoonjoon Lee, "Narrative abstraction model for story-oriented video," Proceedings of the 12th annual ACM international conference on Multimedia, Oct. 2004.
- [11] 손동원, "사회 네트워크 분석," 경문사, 25-38, 161, 242-244쪽, 2002년 12월.
- [12] 네이버 영화, <http://movie.naver.com/index.html>

저자 소개



박 승 보

1995: 인하대학교 학사
1997: 인하대학교 석사
1996.12-2002.5: 대우전자
품질경영연구소,
디지털 TV 연구소
2003.9 - 현재 : 인하대학교 정보공학과
박사과정
관심분야: 멀티미디어 정보검색, USN



조 근 식

1982: 인하대학교 학사.
1991: City University of New York
Computer Science 공학박사
1992.3 - 현재: 인하대학교
컴퓨터정보공학부 교수
관심분야: 인공지능, Semantic Web,
지능형 에이전트 시스템