

개인방송에서 채팅 트래픽을 기반으로 한 하이라이트 검출

김은율, 이계민

서울과학기술대학교 나노IT디자인융합대학원 정보통신미디어공학전공

{keyiulm, gyemin}@seoultech.ac.kr

Highlights Detection based on Chat Traffic in Personal Broadcasting

Eun-Yul Kim, Gyemin Lee

Seoul National University of Science and Technology

요약

최근 개인방송에 대한 관심이 증가하면서 개인 방송 제작자와 콘텐츠의 수가 급증하고 있다. 이에 시청자의 선택이 용이하도록 방송 내용 중 흥미를 끌 만한 장면을 모아 하이라이트를 제공하는 서비스에 대한 요구가 커지고 있다. 본 논문에서는 전문 편집자나 편집 기술 없이 자동으로 하이라이트를 검출하는 방법을 제시한다. 제안하는 알고리즘은 개인 방송의 채팅 정보를 이용하며, 시간에 따른 채팅의 빈도 변화를 통해 하이라이트 구간을 검출한다. 검출 알고리즘은 시청자 선호도가 높은 게임 방송에 적용하여 그 성능을 확인하였다.

1. 서론

‘V앱’, MBC 예능 ‘마이 리틀 텔레비전’의 등장에서부터 알 수 있는 것처럼 최근 개인방송에 대한 사람들의 관심이 커지고 있다[1]. 방송을 위한 플랫폼인 Youtube, Facebook, Twitch, AfreecaTV, 카카오톡 등이 사용하기 쉬운 인터페이스를 제공하고 있기 때문에 일반 사람들도 쉽게 방송을 제작하고 송출하는 것이 가능해졌다.

지상파나 케이블과 같은 일반 방송과 개인방송의 차이점으로 크게 두 가지가 있다. 첫째는 소통의 방법이다. 일반방송의 경우, 방송 연출자와 편집자로부터 시청자로의 일방적인 정보 전달이 이루어지지만 개인방송은 채팅이라는 수단을 통해 양방향으로 정보전달이 가능하다. 두 번째 차이점으로 개인방송은 방송 길이에 대한 제약이 없어 짧게는 몇 분, 길게는 열 시간이 넘는 방송도 쉽게 찾을 수 있다.

하지만 이렇게 긴 방송시간은 직접 방송을 보고 시청을 판단하는 시청자들이 새로운 방송으로 유입되는 것에 걸림돌이 된다. 이에 제작자들은 직접 혹은 전문 편집자를 통해 시청자들이 흥미를 느낄만한 부분을 편집하여 하이라이트 영상으로 제공하는 경우도 있지만, 시간과 비용이 많이 드는 문제가 있다. 이러한 문제를 해결하기 위해 본 논문은 개인방송의 채팅 정보를 이용하여 하이라이트를 자동으로 검출하는 알고리즘을 제안한다. 이를 위해 우리는 대다수 시청자가 흥미를 느끼는 부분에서 채팅이 많이 이루어지는 점에 주목하여 채팅 트래픽이 높거나 급증하는 부분을 하이라이트로 검출한다.

영상에서 하이라이트를 검출하는 방법에 관한 연구는 최근 활발히 진행되며 대부분의 연구가 영상의 시각적 정보나 내용을 기반으로 하이라이트를 검출하는 방식을 이용하고 있다.[2,3,4] 하지만 본 논문은 채팅 트래픽의 변화를 이용하여 시청자들이 흥미를 느끼는 구간들을

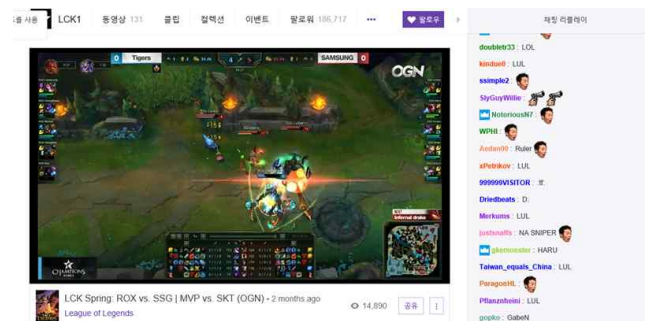


그림 1. 개인방송 화면 구성

하이라이트로 검출한다는 점에서 차이를 보인다. 본 논문에서는 개인 방송 중 상대적으로 시청자 선호 비중이 높은 게임 방송을 중심으로 기술하고[5], 제안한 알고리즘을 적용한다.

2. 방법

일반적인 개인방송 플랫폼은 그림 1과 같이 영상이 나오는 부분과 채팅 내용이 나오는 채팅창 부분이 함께 존재한다. 채팅창에 표시되는 채팅 내역은 채팅이 달릴 때마다 업데이트가 된다. 방송이 진행되면서 시청자가 흥미를 느끼는 부분에서는 그렇지 않은 부분에 비해 시청자들은 대체로 많은 채팅을 남기게 된다. 본 논문에서는 이러한 특징을 이용하여 개인방송에서 하이라이트를 검출하는 방법을 제시한다.

하이라이트 검출은 크게 네 가지 단계로 진행된다. 처음 data extraction 단계를 통해 개인방송 플랫폼으로부터 채팅 내역을 추출하고, 이 추출된 데이터는 denoising 단계를 거치게 된다. 그다음, 시간

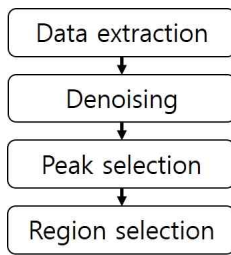


그림 2. 본 논문에서 제안한 하이라이트 검출 알고리즘

에 따른 채팅 내역의 양과 변화량을 이용하여 하이라이트 검출에 사용될 피크를 선정하고 선정된 피크를 기준으로 하이라이트 구간을 검출한다.

2.1 Data extraction & Denoising

하이라이트 검출 알고리즘의 첫 번째 단계에서는 개인방송 플랫폼의 채팅창에 존재하는 채팅 내역들을 텍스트 형태로 추출한다. 추출된 채팅 정보에는 채팅이 달린 시간, 채팅 작성자의 아이디, 채팅 내용이 포함되어있다. 채팅의 트래픽을 알기 위해 수집한 정보 중 시간 정보를 이용하여 시간에 따른 채팅 수의 변화를 그려보면 그림 3의 raw data와 같은 결과를 얻을 수 있다. 이렇게 그려진 raw data는 채팅 트래픽을 보다 용이하게 분석하기 위해 가우시안 필터를 사용하여 노이즈를 제거한다.

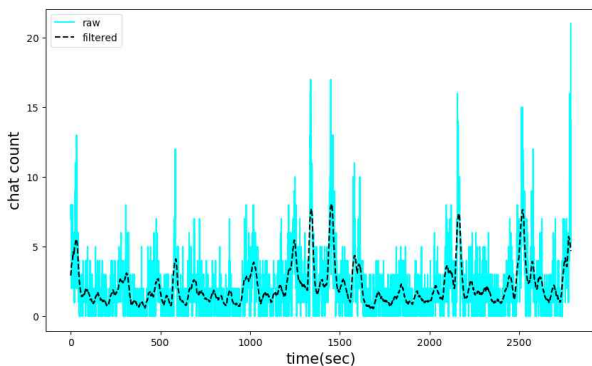


그림 3. 시간에 따른 채팅 트래픽의 변화에 대한 raw data와 raw data에 가우시안 필터링을 적용한 결과

2.2. Peak selection

필터링 적용 후, 하이라이트 구간 선정을 위하여 채팅 트래픽이 집중된 피크를 선택한다. 그림 3에서 확인할 수 있듯이 채팅 트래픽은 다수의 local maximum을 갖는다. 따라서 양질의 하이라이트 검출을 위해서는 적절한 피크를 선택하는 것이 중요하다.

일반적으로 방송에서 흥미로운 부분이 진행될 때는 시청자들이 채팅을 통해 더욱 많은 관심을 표현하는 경향이 있으므로 채팅 수가 많거나 급격히 증가할 것으로 기대할 수 있다. 이러한 점을 고려하여 제안하는 알고리즘은 시간에 따른 채팅 트래픽이 높은 피크 지점과 변동 폭이 큰 지점을 선정한다. 이러한 기준은 다음과 같이 구체적으로 기술할 수 있다.

우선, 채팅 트래픽 $C(t)$ 에서 local maximum과 local minimum 지점을 모두 찾아 그 지점들의 집합 T 를 만든다.

$T = \{t: C(t) \text{ is a local max \& local min}\}$
 그러면 하이라이트 검출을 위한 피크 P 는 다음과 같이 선택할 수 있다.

$$P = P_1 \cup P_2$$

where $P_1 = \{t_i \in T: C(t_i) > th_1, C(t_i) \text{ is a local max}\}$
 $P_2 = \{t_i \in T: C(t_i) - C(t_{i-1}) > th_2\}$

여기서 P_1 은 th_1 보다 큰 모든 피크를 나타내고, P_2 는 채팅 트래픽이 갑자기 증가한 피크들을 나타낸다. 그림 4에서 th_1 보다 높은 부분에 선택된 점들은 P_1 을 만족하는 피크이며, th_1 보다 낮은 부분에 선택된 점들은 P_2 의 조건을 만족하는 피크이다.

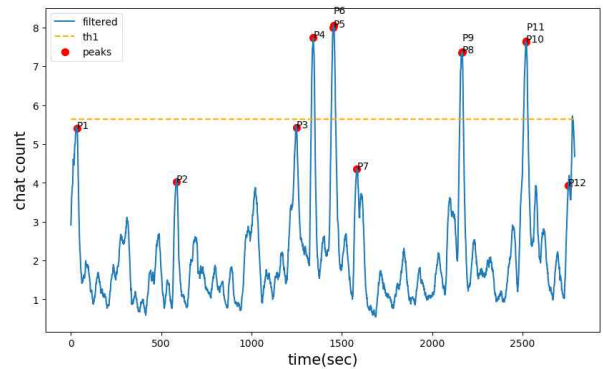


그림 4. 채팅 트래픽 분석을 통해 선정된 피크

2.3. Region selection

Region selection 단계에서는 앞에서 구한 P 를 이용해 적절한 하이라이트 구간을 검출한다. 개인방송에서의 채팅은 시청자들이 방송 영상을 보고 그에 반응하여 채팅 글을 입력하기까지의 지연과 입력한 글이 서버에 전달되어 화면에 보이기까지 네트워크 딜레이를 포함한다. 따라서 이를 고려하여 선정된 피크의 앞쪽에 하이라이트 구간의 시작과 끝을 지정함으로써 하이라이트 구간 H 를 검출할 수 있다.

$$H = \{(t_i - s_{offset}, t_i - e_{offset}) : t_i \in P\}$$

위의 수식에서 s_{offset} 은 하이라이트 구간의 시작으로부터 피크까지의 거리(초)를 의미하고, e_{offset} 은 하이라이트 구간의 끝으로부터 피크까지의 거리(초)를 의미한다.

이렇게 검출한 구간들은 한 구간을 기준으로 가까운 거리에 다른 구간이 존재한다면, 이들은 방송 흐름을 보면 같은 내용에 해당하는 경우가 많다. 따라서 검출 알고리즘은 이들을 연결하여 하나의 구간으로 만든다.

3. 실험 결과

앞서 제시한 알고리즘을 사용해 채팅 트래픽을 기반으로 하이라이트 구간을 검출한다. 본 논문에서는 Twitch에서 방송한 게임 영상을 대상으로 실험을 진행하였다. 이 영상들은 ‘2017 LoL 챔피언스 코리아 서머’ 대회 게임 영상으로, 전체 게임 영상과 더불어 대회 주최 측에서 제작한 게임 하이라이트(이하 대조군으로 표기)를 영상 끝에 포함하고 있다. 위에서 제안된 알고리즘을 이용해 검출한 하이라이트 구간을 방송 영상에 포함되어있는 하이라이트와 비교하여 그 성능을 검증한다.

	영상 길이 (초)	채팅 수 (개)	초당 채팅 수 (개/초)	대조군 길이(초)	대조군 비율(%)	대조군 구간 수(개)	대조군 구간당 평균 길이(초)
Game1	2,789	6,143	2.20	142	5.09	8	17.75
Game2	2,616	4,922	1.88	93	3.56	10	9.30
Game3	2,480	4,597	1.85	162	6.53	7	23.14
Game4	2,669	6,897	2.58	92	3.45	6	15.33
Game5	2,552	6,126	2.40	94	3.68	5	18.80
Game6	1,344	5,278	3.93	68	5.06	8	8.50
Game7	1,893	7,062	3.73	54	2.85	6	9.00
평균	2,334.71	5,860.71	2.65	100.71	4.32	7.14	14.55

표 1. 사용된 데이터 정보

각 영상은 평균 38분 55초이며 5861개의 채팅이 포함되어 있어 초당 2.65개의 채팅이 생성된다고 볼 수 있다. 또한, 영상에 포함된 하이라이트는 평균 1분 41초이며 평균적으로 7개 구간으로 이루어져 있고, 각각의 구간은 평균 14.5초 정도의 길이를 가지고 있다. 획득한 7개의 영상에 대한 정보는 표 1에 정리되어 있다.

3.1. 대조군과 채팅 트래픽

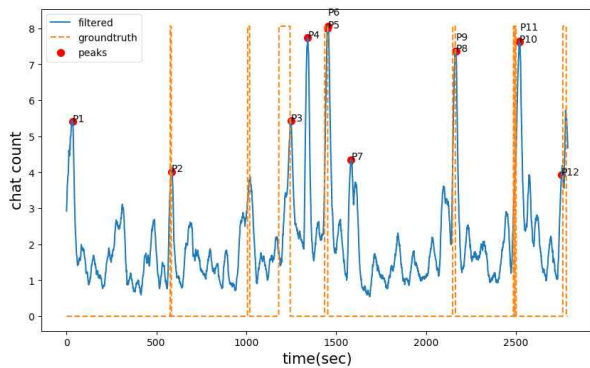


그림 5. 필터링이 적용된 채팅 트래픽과 대조군 구간

우선 peak selection 단계를 통해 선택된 피크 P 가 대조군에 얼마나 상응하는지 살펴본다. 이때 그림 5에서 볼 수 있듯이 하나 이상의 피크가 하나의 하이라이트 구간에 대응할 수 있다. 전체 7개의 영상에서 선택된 피크의 총 개수는 88개이며 그중 31개가 대조군 구간에 속하고, 전체 50개의 대조군 구간 중 24개가 피크를 포함하고 있다. 영상별 관계는 표 2에 정리되어 있다.

	A 검출된 피크 수(개)	B 대조군 구간 수(개)
Game1	12 (6 + 6)	8 (5 + 3)
Game2	10 (5 + 5)	10 (6 + 4)
Game3	9 (2 + 7)	7 (2 + 5)
Game4	17 (5 + 12)	6 (3 + 3)
Game5	13 (5 + 8)	5 (3 + 2)
Game6	14 (6 + 8)	8 (6 + 2)
Game7	13 (2 + 11)	6 (3 + 3)

표 2. A. 검출된 피크 수(개) (대조군에 속하는 피크 수(개) + 속하지 않는 피크 수(개)) B. 대조군 구간 수(개) (피크를 포함하는 구간 수(개) + 포함하지 않는 구간 수(개))

실험을 진행한 7개의 영상에 대해 제안한 알고리즘을 이용해 검출된 피크들 중 대조군에 속하는 피크의 비율은 평균적으로 35.48%이며, 검출된 피크를 포함하는 대조군은 평균 55.15%이다. 또한, 대조군에 속하는 피크를 가장 많이 찾은 영상은 game1과 game2로 50%의 비율로 검출된 피크가 대조군에 속함을 확인하였지만, game7은 검출된 피크 중 15.38%만이 대조군에 속했다. 마찬가지로 검출된 피크를 포함하는 대조군은 game6이 75%로 가장 높고, game3은 28.57%의 대조군만이 검출된 피크를 포함한다. 검출된 피크 중 대조군에 속하는 피크의 비율과 검출된 피크를 포함하는 대조군의 비율을 평균을 내서 비교한 결과 game6이 58.93%로 가장 높은 성능을 보인 반면, game3은 25.40%로 가장 낮은 성능을 보였다.

3.2. 대조군과 검출된 하이라이트

본 논문에서 제안한 알고리즘으로 검출한 하이라이트와 대조군을 비교하기 위해 leave-one-out 방법으로 실험을 진행하기 위해 획득한 7개의 영상 중 6개의 영상을 이용해 s_{offset} 과 e_{offset} 을 지정해준 후, 남은 한 개의 영상으로 실험 결과를 확인하였다. 각 실험의 결과는 그림 6과 같다.

실험을 진행한 7개의 영상에 대해 제안된 알고리즘을 이용해 검출된 하이라이트 구간 중 대조군과 일치하는 구간의 비율은 42.73%이며, 대조군이 검출된 비율은 50.39%이다. 검출 구간을 기준으로 대조군과 가장 많이 일치하는 영상은 game2로 66.67% 일치하지만, 가장 일치하지 않는 영상은 game4로 27.27% 만 일치한다. 마찬가지로 대조군을 많이 검출한 영상은 game6으로 75% 검출하였으나 검출이 잘 안된 영상은 game3으로 28.57%만 검출하였다. 검출된 하이라이트 구간 중 대조군과 일치하는 구간의 비율과 검출된 대조군의 비율을 평균내서 비교한 결과 game2가 63.34%로 가장 높았고, game4가 30.30%로 가장 낮은 성능을 보였다.

4. 결론

본 논문에서는 영상을 보는 다수의 사람이 흥미롭게 여기는 구간을 이용하여 하이라이트 구간을 검출하기 위해 채팅 트래픽을 이용하는 방법을 제시하였다. 실험 결과 하이라이트 구간에서 채팅 수가 증가하는 것을 확인하였고, 이를 이용하여 하이라이트 구간을 자동으로 검출할 수 있음을 보였다.

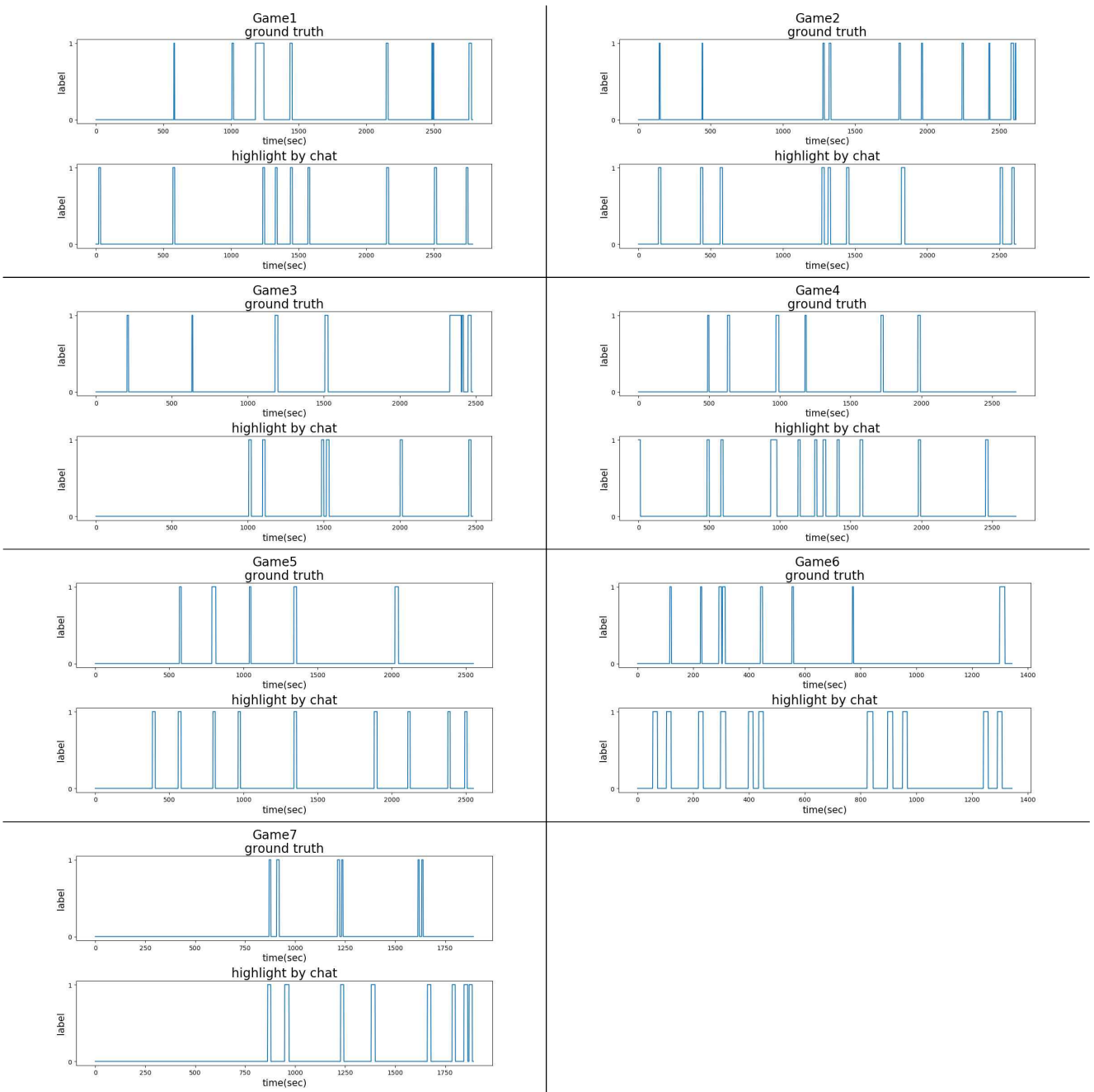


그림 6. 각 영상에 대한 대조군(위)과 제안된 알고리즘으로 검출된 하이라이트(아래)

감사의 말

이 논문은 과학기술정보통신부의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (NO. 2016-0-00099, 제작 권리성과 실감 시청 체험 극대화를 위한 개인방송 제작 기술 개발)

참고문헌

- [1] DMC Report, 온라인 동영상 시장 현황 및 트렌드, 2015.10.
- [2] Y.L.Lin, V.I. Morariu, and W.Hsu. Summarizing While Recording: Context-Based Highlight Detection for Egocentric Videos. In ICCVW, 2015.
- [3] H.Tang, V.Kwatra, M.E.Srgin, and U.Gargi. Detecting Highlights in Sports Videos: Cricket as a Test Case. In ICME, 2011.
- [4] M.Gygli, Y.Song, and L.Cao, Video2GIF: Automatic Generation of Animated GIFs from Video. In CVPR, 2016.
- [5] 한국콘텐츠진흥원, 온라인 개인방송 서비스 이용행태 조사, 2014.07.