

특집논문 (Special Paper)

방송공학논문지 제23권 제2호, 2018년 3월 (JBE Vol. 23, No. 2, March 2018)

<https://doi.org/10.5909/JBE.2018.23.2.218>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

채팅 트래픽 분석을 통한 개인방송 하이라이트 검출 : 게임 콘텐츠를 중심으로

김 은 율^{a)}, 이 계 민^{a)†}

Highlight Detection in Personal Broadcasting by Analysing Chat Traffic : Game Contests as a Test Case

Eunyu Kim^{a)} and Gyemin Lee^{a)†}

요 약

최근 개인방송 콘텐츠의 수가 급증함에 따라 시청자의 선택이 용이하도록 방송 내용 중 흥미를 끌 만한 장면을 모아 하이라이트를 제공하는 서비스에 대한 요구가 커지고 있다. 본 논문에서는 채팅 트래픽 정보가 하이라이트 검출에 유용함을 보이고 채팅 트래픽을 이용하여 하이라이트를 검출하는 방법을 제시한다. 또한, 하이라이트 검출에 있어 채팅 트래픽 사용의 효용성을 평가하기 위해 평가 방법을 제안한다. 검출 알고리즘은 시청자 선호도가 높은 게임 방송에 적용하여 그 성능을 확인하였다.

Abstract

As the number of personal broadcasting contents is rapidly increasing, the demand for a service that provides highlights is growing. A highlight, a collection of interesting scenes, can improve the quality of viewing experience. In this paper, we propose a method to automatically detect highlights using only chat traffic information. We also propose evaluation methods the effectiveness of using chat traffic in highlight detection. We apply the detection algorithm to game broadcasting, which has larger audience, and demonstrate its performance.

Keyword : highlight, chat traffic, personal broadcasting, video summarization

a) 서울과학기술대학교 전자IT미디어공학과(Dept. of Electronic and IT Media Engineering, Seoul National University of Science and Technology)

† Corresponding Author : 이계민(Gyemin Lee)

E-mail: gyemin@seoultech.ac.kr

Tel: +82-2-970-6416

ORCID: <http://orcid.org/0000-0001-6785-8739>

※ 이 논문의 연구결과 중 일부는 “한국방송·미디어공학회 2017년 추계학술대회”에서 발표한 바 있음.

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

※ This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government (MSIT) (No.2016-0-00099, Personal Broadcast Technology Development for Production Convenience and Maximum Viewing Experience)

· Manuscript received December 29, 2017; Revised March 8, 2018; Accepted March 16, 2018.

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. 서론

최근 개인방송에 대한 사람들의 관심이 커짐에 따라 개인방송 제작자들과 콘텐츠의 수가 급증하고 있다^[1]. 개인방송은 길이에 대한 제약이 없기 때문에 1시간 이상 방송하는 경우도 흔하게 존재한다. 이처럼 긴 방송은 새로운 시청자의 유입에 대한 걸림돌이 되기도 하며 네트워크 전송 대역 낭비가 될 수 있어 콘텐츠 제작자는 하이라이트 영상을 편집하여 제공하기도 한다. 하지만 이 과정은 편집에 걸리는 시간과 전문 편집 기술이 필요해 이에 따른 비용이 많이 소모되는 문제가 있다.

이러한 문제를 해결하기 위해 본 논문은 개인방송 하이라이트를 자동으로 검출하는 알고리즘을 제안한다. 특히 개인방송 플랫폼의 채팅 내역 정보가 하이라이트 검출에 유용함을 보이고자 한다. 지상파나 케이블과 같은 일반 방송과 달리, 개인방송은 채팅을 통해 시청자와 콘텐츠 제작자가 실시간으로 소통할 수 있으며 대체로 대다수 시청자가 흥미를 느끼는 이벤트가 있을 때 채팅이 많이 이루어진다. 우리는 이러한 점에 주목하여 채팅 트래픽이 높거나 급증하는 부분을 시청자가 흥미를 느끼는 이벤트가 발생한 부분으로 간주하고 이를 하이라이트로 검출한다.

영상에서 하이라이트를 검출하는 방법에 관한 연구는 최근 활발히 진행되고 있다. 대부분의 연구는 영상이나 음향을 기반으로 하이라이트를 검출하는 방식을 이용하고 있다. 영상 정보만을 이용하는 경우, 영상의 low-level 시각적 특징(color histogram, HOG)에 기반한 이벤트 통계를 추출하여 해당 클립의 하이라이트 여부를 판단할 수 있다^[2]. 또는 영상을 짧은 길이의 세그먼트로 분할한 후, 각 세그먼트가 하이라이트에 포함될 점수를 매겨 하이라이트 여부를 판단하는 방법도 제안되었다^[3,4,5]. 반면 스포츠 콘텐츠의 경우에는 주요 장면에서 관객들의 함성소리가 크다는 점을 이용해 음향정보를 이용하기도 한다^[6]. 하지만 개인방송 플랫폼의 주요 특징 중 하나인 채팅 정보만을 이용하는 연구는 우리가 조사한 바로는 선행된 것이 없다.

본 논문은 채팅 정보가 하이라이트 검출에 있어서 유용할 수 있음을 보이고, 채팅 트래픽을 이용하여 하이라이트를 검출하는 알고리즘과 더불어 제안한 알고리즘의 성능을 평가하는 방법을 제안한다. 이를 위해, 본 논문에서는 개인

방송 중 상대적으로 시청자 선호 비중이 높은 게임 방송을 중심으로 제안한 알고리즘을 적용하여 성능을 보인다^[7].

II. 하이라이트 검출 알고리즘

일반적으로 개인방송 플랫폼은 그림 1과 같이 영상과 함께 채팅창이 존재한다. 시청자들은 영상을 시청하면서 채팅을 통해 방송 제작자나 다른 시청자와 의견을 교환한다. 이때 시청자들은 흥미를 느끼는 부분에서 더 많은 채팅을 남기는 경향이 있다. 일반적으로 하이라이트가 시청자들이 흥미를 느낄 만한 부분들을 모아서 만든 영상이라는 것을 고려한다면, 채팅이 많이 이루어지는 부분은 영상의 하이라이트에 속할 가능성이 높다고 볼 수 있다. 본 논문에서는 이러한 특징을 이용하여 하이라이트 구간을 검출하는 알고리즘을 제시한다. 제안하는 알고리즘은 그림 2와 같이 채팅 트래픽 추출 및 노이즈 제거, 피크 선정, 구간 설정의 네 단계로 진행된다.

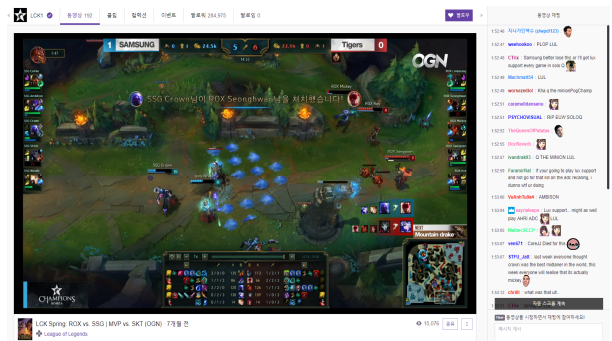


그림 1. 개인방송 화면 구성

Fig. 1. An example of personal broadcasting platform (from Twitch^[7])

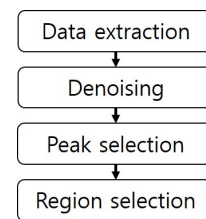


그림 2. 제안하는 하이라이트 검출 알고리즘의 구성

Fig. 2. The flow of the proposed highlight detection algorithm

1. 채팅 트래픽 추출 및 노이즈 제거

개인방송 플랫폼으로부터 획득할 수 있는 채팅 정보 중에서 채팅의 트래픽 변화를 알기 위해 채팅이 달린 시각에 대한 정보를 추출한다. 획득된 데이터를 이용하여 시간에 따른 채팅 빈도의 변화를 그려보면 그림 3의 미처리 데이터(raw) $C_{raw}(t)$ 와 같다. 그림 3에서 볼 수 있듯이 미처리 데이터는 순간적인 변화나 노이즈가 많아 전반적인 빈도 변화를 파악하기 어렵다. 따라서 이후 단계에서 데이터를 효율적으로 이용하고 검출 오류를 낮추기 위해 노이즈를 줄일 필요가 있다. 이를 위해, 잡음 제거 필터를 이용하여 노이즈를 감소시킨다.

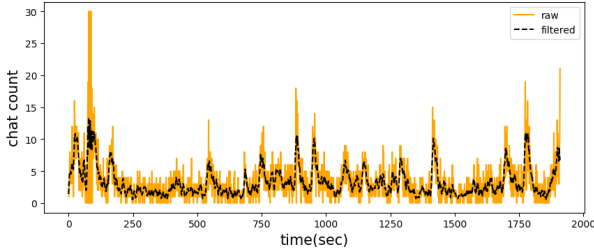


그림 3. 시간에 따른 채팅 트래픽 데이터
Fig. 3. The change of chat traffic over time

$$C(t) = C_{raw}(t) * H(t) \quad (1)$$

여기서 $H(t)$ 는 잡음 제거 필터를, *는 합성곱(convolution)을 의미한다. 잡음 제거 필터를 사용하여 노이즈를 감소시킨 채팅 트래픽 $C(t)$ 는 그림 3의 검은 점선으로 표시되어 있다.

2. 피크 선정

본 논문에서 제안하는 알고리즘은 최종적으로 하이라이트 구간을 검출하는 것이 목표이며 이를 위해 구간의 기준이 될 지점을 찾는 것은 중요하다. 일반적으로 방송에서 흥미로운 부분이 진행될 때는 시청자들이 채팅을 많이 남기는 경향이 있으므로 채팅의 빈도수가 많을 것이라 예상할 수 있다. 따라서 시간 t 부근에서 채팅 트래픽 $C(t)$ 가 가장 큰 값을 가질 때의 t 를 모두 찾아 피크 위치 후보군 T 를 형성한다.

$$T = \{t : C(t) \text{ is a local max}\} \quad (2)$$

지역 최대점(local max)은 이전, 이후에 비해 트래픽이 큰 지점을 의미한다. 즉, $C(t) > C(t+1)$ 이고 $C(t) > C(t-1)$ 이면 t 는 지역 최대점이다.

이때 개인방송의 콘텐츠 종류 중 일부의 경우에 -특히 게임 방송의 경우- 영상의 시작 부분은 중요하거나 흥미를 끌

input raw chat traffic $C_{raw}(t)$, length of a region s_{offset} , time offset t_0 , target # of region n , target highlight length B_0
output highlight B

1. $C(t) \leftarrow filter_H(C_{raw}(t))$
2. $T = \{t > t_0 : C(t) \text{ is a local max}\}$
3. $i \leftarrow 1, B \leftarrow \emptyset$
4. **while** $i \leq n \times 1.1$ and $|B| \leq |B_0| \times 1.1$ **do**
5. select next largest peak point t_i from T
6. $B_i = \{(t_i - s_{offset}, t_i)\}$
7. **if** $B_i \cap B = \emptyset$ **then**
8. $i \leftarrow i + 1$
9. **end if**
10. $B \leftarrow B \cup B_i$
11. **end while**

그림 4. 제안하는 하이라이트 검출 알고리즘
Fig. 4. The proposed highlight detection algorithm

만한 내용이 적어 채팅 빈도수가 높더라도 하이라이트에 속한다고 보기 어렵다. 하지만 영상의 시작에서는 시청자들의 기대로 인해 채팅 트래픽이 높은 경향이 있어 영상의 시작 부분이 피크로 선택되는 경우가 존재한다. 이러한 이유로 영상이 시작되고 일정 시간 t_0 가 지난 이후부터 피크를 선택하면 오검출을 줄일 수 있다. 이에 위에 제안한 피크 위치 후보 T 는 다음과 같이 변형할 수 있다.

$$T = \{t > t_0 : C(t) \text{ is a local max}\} \quad (3)$$

위의 방법으로 정해진 피크 후보 T 에서 $C(t_i)$ 의 크기가 큰 순서대로 t_i 가 구간 설정을 위한 피크로 사용된다.

3. 구간 선정

구간 선정 단계에서는 앞서 구한 피크를 기준으로 최종 하이라이트로 검출될 적절한 구간을 채택한다. 개인방송 플랫폼에서의 채팅은 시청자가 영상을 보고 채팅을 입력하기까지의 시간과 입력된 채팅이 서버에 전달되어 화면에 보이기까지의 시간에 의해 어느 정도 시각이 지연되어 나타난다. 이러한 지연 시간을 고려하여 피크의 앞부분에 구간 B_i 를 만들고, 만들어진 B_i 들을 모아 검출 하이라이트 B 를 형성한다.

$$B_i = \{(t_i - s_{offset}, t_i) \text{ for } t_i \in T\} \quad (4)$$

$$B = \bigcup_i B_i \quad (5)$$

위의 수식에서 s_{offset} 은 설정한 하이라이트 구간의 시작 점으로부터 피크까지의 길이(초)를 의미한다. 만일 검출 구간이 겹치는 경우, 두 구간은 하나의 구간으로 합친다. 이 과정을 반복하여 하이라이트 B 를 만든다. 이때, 검출 하이라이트의 길이와 구간의 수가 학습 데이터의 하이라이트 B_0 와 유사하게 만들기 위해 B 의 길이나 구간의 수가 B_0 보다 10% 이상 커지면 위 과정을 멈춘다.

위에 설명한 전체 알고리즘을 의사 코드(pseudo code)를 이용해 그림 4와 같이 나타낼 수 있다. 입력으로 미처리 채팅 데이터 $C_{raw}(t)$, 하이라이트 검출 구간의 수 n , 검출 구

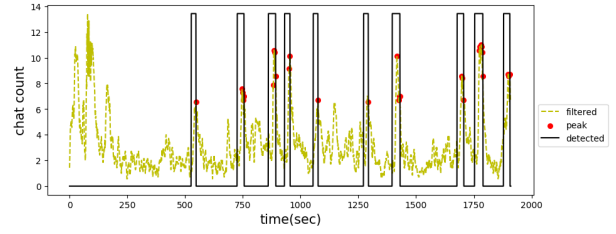


그림 5. 검출된 하이라이트 구간과 기준 피크

Fig. 5. The selected peak points(red dots) and the detected highlight regions(black solid line)

간의 길이 s_{offset} , 검출 시작 파라미터 t_0 를 넣어주면 노이즈 제거, 피크 및 구간 선정 단계를 거쳐 검출된 하이라이트 B 가 나오게 되고 결과는 그림 5와 같다. 그림 5에서의 점선은 필터링된 채팅 트래픽 $C(t)$, 빨간 점은 선정된 피크를 나타낸다. 더불어 실선은 검출된 하이라이트를 의미하며 위에서 설명한 알고리즘을 이용하여 만들어진 10개의 하이라이트 구간을 보여주고 있다. 각 검출 하이라이트 구간의 길이가 다른데, 이는 선택된 피크의 거리가 가까워 여러 구간이 합쳐졌기 때문이다.

III. 검출된 하이라이트 구간의 평가

본 논문에서는 Jaccard 인덱스와 유사 재현율(pseudo recall), 유사 정밀도(pseudo precision), F1점수(F1-score)를 이용하여 검출 알고리즘의 성능을 평가하고자 한다.

1. Jaccard 인덱스

제안하는 알고리즘을 통해 검출된 하이라이트는 평가 영상에 대한 실제 하이라이트(ground truth)가 있을 때 유사도 계산으로 알려진 Jaccard 인덱스를 사용하여 정량적으로 평가할 수 있다. 실제 하이라이트를 A , 검출 구간의 집합을 B 라 하면, Jaccard 인덱스는 다음과 같이 표기된다.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

이 값이 1에 가까울수록 A 와 B 가 일치함을 나타내며 이

는 실제 하이라이트와 검출된 하이라이트 중에 겹치는 부분에 대한 비율을 높다는 것을 뜻한다. 따라서 실제 하이라이트와 검출 하이라이트가 정확히 일치하지 않으면 같은 내용의 장면을 검출하여도 Jaccard 인덱스는 작은 값이 나오게 된다. 이런 특성은 채팅 정보가 하이라이트 검출에 유용하다는 것을 보이는 데 어려움이 있으므로, 재현율(recall)과 정밀도(precision)와 유사한 평가지표를 제안하여 이용한다.

2. 유사 재현율과 유사 정밀도, F1점수

유사 재현율 $R(A,B)$ 는 재현율과 비슷하지만 단순화된 형태를 가지며, 검출된 하이라이트 B 와 일치하는 부분이 존재하는 실제 하이라이트 A 의 구간을 제대로 찾은 구간으로 생각한다. 이를 이용하여 유사 재현율은 실제 하이라이트의 구간들 중 잘 찾아진 구간들의 길이의 비율을 의미한다.

$$R(A,B) = \frac{\left| \bigcup_{i: A_i \cap B_j \neq \emptyset, \forall j} A_i \right|}{\left| \bigcup_i A_i \right|} \quad (7)$$

유사 재현율과 마찬가지로, 유사 정밀도 $P(A,B)$ 는 정밀도와 비슷한 형태를 갖는다. 제안한 알고리즘을 통해 검출된 구간 중 실제 하이라이트와 일치하는 부분이 있는 구간의 비율을 나타낸다.

$$P(A,B) = \frac{\left| \bigcup_{j: A_i \cap B_j \neq \emptyset, \forall i} B_j \right|}{\left| \bigcup_j B_j \right|} \quad (8)$$

위의 방법으로 구해진 유사 재현율과 유사 정밀도를 이용하여 다음과 같이 F1점수를 구할 수 있다.

$$F(A,B) = 2 \frac{R(A,B) \cdot P(A,B)}{R(A,B) + P(A,B)} \quad (9)$$

본 논문에서는 제안하는 평가방법을 통해 하이라이트 검출에 있어 채팅 트래픽의 유효성을 검증한다.

IV. 실험 및 결과

앞서 제안한 하이라이트 검출 알고리즘을 평가하기 위해 시청자 선호 비중이 높은 게임 영상^[6]을 이용해 실험을 진행하였다. 실험에 사용된 영상은 Twitch에서 중계한 'League of Legends' 대회 영상들이며^[8], 게임 전문 방송국인 OGN에서 제작한 하이라이트가 존재한다. 각각의 하이라이트 영상은 방송국에서 전문 편집자가 정해진 기준에 의해 제작한 영상이므로 실제 하이라이트(ground truth)로 보기에 적합하다고 볼 수 있으므로 이를 기준으로 실험을 진행하였다.

사용한 데이터의 수는 총 36개이며, 2017년에 개최된 대회 중 4개 -LoL 올스타 2017, LoL 2017 월드 챔피언십, 2017 LoL 챔피언스 코리아 서머, 2017 LoL 챔피언스 코리아 스프링, EM 월드챔피언십 카토비체 2017- 에서 진행된 게임 가운데 일부 게임을 사용하였다. 이 영상들은 평균적으로 36분 25초 길이며 9,122개의 채팅을 포함하고 있다. 또한, 각 영상에 해당하는 실제 하이라이트는 평균적으로 4분 2초 길이로, 평균 20초의 길이를 갖는 12개 구간으로 이루어져 있다. 사용 데이터의 정보는 표 1에 요약되어 있다.

실험은 leave-one-out 방식으로 진행하였다. 36개의 영상 중 하이라이트를 검출하고자 하는 하나의 영상(예, Game 5)을 제외하고 나머지 35개 영상으로 훈련 집합을 만든다. 그 후, 훈련 집합의 실제 하이라이트 시작 위치, 구간 수, 구간 당 길이의 평균으로 t_0 , n , s_{offset} 을 각각 정한다. 이 값과 제안한 알고리즘을 이용하여 Game 5의 하이라이트를 검출한다. 전체 36개 영상의 하이라이트를 검출하는 데 소요된 시간은 0.46초(Intel i7-4770 CPU, 32G RAM)로 상당히 빠른 시간에 수행되었다.

그림 6은 검출된 하이라이트의 결과를 실제 하이라이트와 비교하여 보여주고 있다. 그림 6(a)는 Game 5의 실제 하이라이트 중 640초 근처의 구간과 800초, 1260초 근처의 구간을 제외한 나머지 구간이 검출 알고리즘을 통해 찾아졌음을 보인다. 이 중 640초부터 시작되는 구간과 가까운 위치에서 검출된 구간은 이어지는 장면이지만, 영상 정보를 알지 못하기 때문에 이 구간을 검출하지 못하였음을 영상을 보며 확인하였다. 또한 검출 하이라이트 구간 중 850

표 1. 사용된 데이터 정보

Table 1. The summary of 36 video data

	Video length (sec)	Number of chats	Number of chats per second	Number of regions in ground truth	Length of ground truth highlights (sec)	Average length of each highlight region (sec)
mean	2,185.33	9,122.83	4.35	12.72	242.69	19.50
$\pm se$	± 120.70	± 541.87	± 0.25	± 0.59	± 11.06	± 0.62
max	4,790.00	18,826.00	8.21	20.00	480.00	29.14
min	1,314.00	3,561.00	1.61	7.00	130.00	12.17

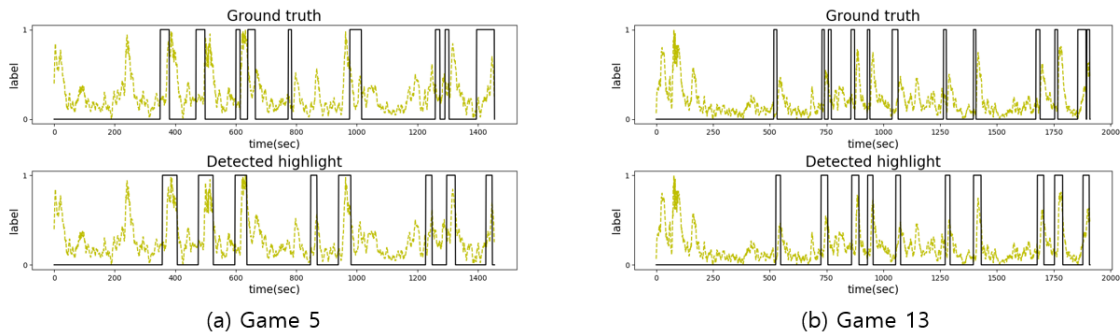


그림 6. 제안한 알고리즘을 이용한 하이라이트 검출 결과. 위: 실제 하이라이트(Ground truth) 아래: 검출된 하이라이트(Detected highlight)
점선은 채팅트래픽 $C(t)$, 실선은 하이라이트 여부를 의미.

Fig. 6. The result of the proposed highlight detection algorithm on Game 5 and Game 13.

초와 1230초 근처에서 만들어진 두 구간을 제외한 나머지 구간은 모두 실제 하이라이트 구간을 찾음을 보였다. 위 실험의 결과를 정량적으로 평가하면 Jaccard 인덱스 $J(A,B)$ 는 0.222, 유사재현율 $R(A,B)$ 는 0.784, 유사정밀도 $P(A,B)$ 의 값은 0.844, 마지막으로 F1 점수 $F(A,B)$ 는 0.813으로 계산된다. 이와 같은 결과는 그림 6(a)를 통해 볼 수 있듯이 실제 하이라이트의 78.4%가 검출되었으며, 검출된 하이라이트 270초 가운데 228초가 실제 하이라이트 구간과 중첩된다는 것을 의미한다.

마찬가지로 Game 13의 하이라이트를 검출하고 결과를 그림 6(b)에 나타내었다. Game 13의 실제 하이라이트 구간은 750초 부근의 한 구간을 제외한 나머지 구간이 전부 검출되었으며, 검출 하이라이트의 모든 구간은 실제 하이라이트를 찾았음을 보인다. Game 13은 $J(A,B)$ 의 값이 0.322, $R(A,B)$ 의 값은 0.930, $P(A,B)$ 의 값은 1, 그리고 $F(A,B)$ 의 값은 0.966이었다. 이 결과도 그림 6(b)를 통해 확인한 결과인 실제 하이라이트 구간의 93%가 검출되었으며, 검출된 하이라이트는 전부 실제 하이라이트와 겹치는

구간임을 나타낸다.

위와 같은 실험을 각 영상에 대해 36번 반복하여 획득한 결과를 표 2에 요약하였다. 알고리즘에 사용되는 파라미터 시간 오프셋 t_0 는 평균 305.94초로 설정되었고, 검출 구간의 수 n 과 검출 구간의 길이 s_{offset} 는 각각 평균적으로 12개 구간과 20초로 설정되었다. 또한 잡음 제거 필터는 윈도우 크기가 7인 moving average 필터를 사용하였다. 제안한 알고리즘을 통해 찾은 하이라이트는 평균적으로 273.97초의 길이로 실제 하이라이트의 평균 길이인 242.69초와 비슷하다. 반면 구간의 수는 약 9.69개로 평균 12개의 구간으로 이루어진 실제 하이라이트에 비해 적은 구간을 갖는다. 또한 전체 영상에 대한 성능의 평균은 $R(A,B)$ 는 0.603, $P(A,B)$ 는 0.673, $F(A,B)$ 는 0.624로 계산되었다. 이 값들은 실제 하이라이트 구간 중 60.3%가 제안하는 알고리즘을 통해 검출되며, 하이라이트로 검출된 부분 중 67.3%가 실제 하이라이트에 해당한다는 것을 뜻한다. 이러한 결과는 하이라이트 검출에 채팅 트래픽 정보가 유용하게 이용될 수 있다는 것을 의미한다.

표 2. 정량적 평가

Table 2. The result summary

	Jaccard index	Pseudo recall	Pseudo precision	F1-score
mean \pm se	0.213 \pm 0.011	0.603 \pm 0.027	0.673 \pm 0.026	0.624 \pm 0.022
max	0.322	0.934	1.000	0.966
min	0.065	0.245	0.449	0.347

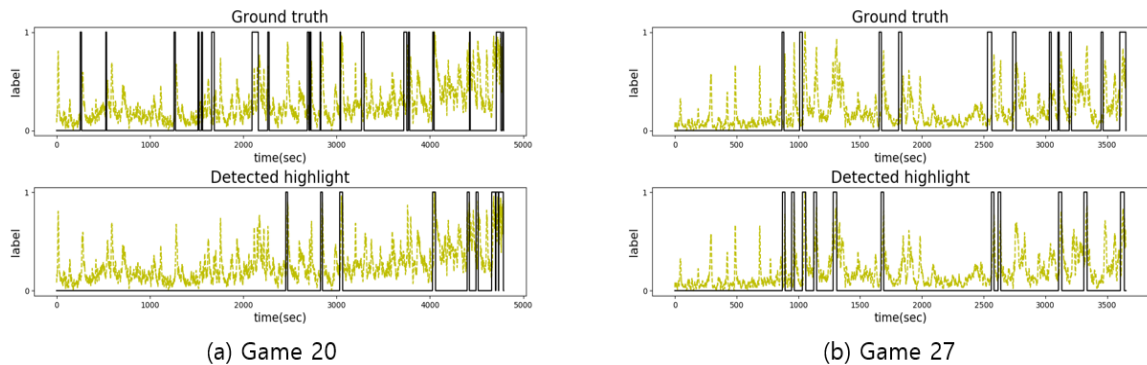


그림 7. 제안한 알고리즘을 이용한 하이라이트 검출 성능이 좋지 않은 경우. 위: 실제 하이라이트(Ground truth) 아래: 검출된 하이라이트(Detected highlight). 점선은 채팅트래픽 $C(t)$, 실선은 하이라이트 여부를 의미

Fig. 7. The result of the proposed highlight detection algorithm on Game 20 and Game 27

반면에 실험 결과 검출 알고리즘의 성능이 좋지 않은 경우도 존재하였다. 이러한 경우를 그림 7을 통해 나타내었다. 그림 7(a)로부터 Game 20의 구간 수는 실제 하이라이트 평균 구간 수인 12개에 비해 훨씬 많은 19개의 실제 하이라이트 구간이 존재함을 알 수 있다. 하지만 검출 알고리즘 9개의 구간으로 이루어졌다. 이와 같은 이유로 검출 하이라이트는 실제 하이라이트의 상당수를 찾지 못하였다. 정량적으로 성능을 확인해본 결과, Game 20의 $R(A, B)$ 는 0.245, $P(A, B)$ 는 0.599로 유사 재현율이 작은 값을 보였다. 이 결과는 실제 하이라이트 중에서 24.5%만 검출되었으며, 검출된 하이라이트의 길이 267초 중 160초 만이 실제 하이라이트와 겹치는 구간의 길이임을 의미한다.

마찬가지로 성능이 좋지 못한 Game 27은 그림 7(b)를 통해 확인할 수 있다. Game 27은 실제 하이라이트와 검출된 하이라이트의 구간의 수는 일치하나, 실제 하이라이트 구간 중 6개만이 검출되었다. 이 영상의 1124초부터 1564초까지는 게임 진행 중 경기가 중단된 부분으로, 실제 게임영상 대신 관객석이나 해설자들을 비추는 장면으로 이루어져 있다. 하지만 이 부분에서 시청자들의 채팅이 많아져 검출

알고리즘을 통해 하이라이트로 선정이 되었다. 그러나 해당 구간은 경기 장면이 아니어서 실제 하이라이트에는 포함되어 있지 않아 오검출 구간이라 볼 수 있다. 이는 영상과 채팅 내용을 직접 검사하여 확인할 수 있었다. Game 27에 대한 정량적 결과는 유사 재현율의 값이 0.512, 유사 정밀도는 0.449로 구해졌다. 다시 말해, 실제 하이라이트의 51.2%만이 검출되었으며, 검출된 구간의 전체 길이 267초 가운데 136초만 실제 하이라이트와 일치하는 부분이 존재하는 구간의 길이임을 의미한다.

실험을 통해 채팅 트래픽이 하이라이트 검출에 있어 의미를 가진다는 사실을 알 수 있다. 하지만 위의 두 가지 사례와 같은 경우는 검출 알고리즘이 영상의 정보는 전혀 사용하지 않고 채팅 트래픽 정보만을 이용하여 발생하는 한계점이라 볼 수 있다. 이러한 문제점은 차후에 영상과 음향 정보를 결합하여 개선할 수 있을 것이라 기대한다.

그림 6과 7을 통해 채팅 트래픽은 대체로 시작 부분에서 높은 경향이 있음을 확인할 수 있다. 하지만 게임 영상의 경우, 영상의 시작 부분은 하이라이트에 해당하지 않을 때가 많다. 따라서 제안하는 알고리즘은 성능을 개선시키기

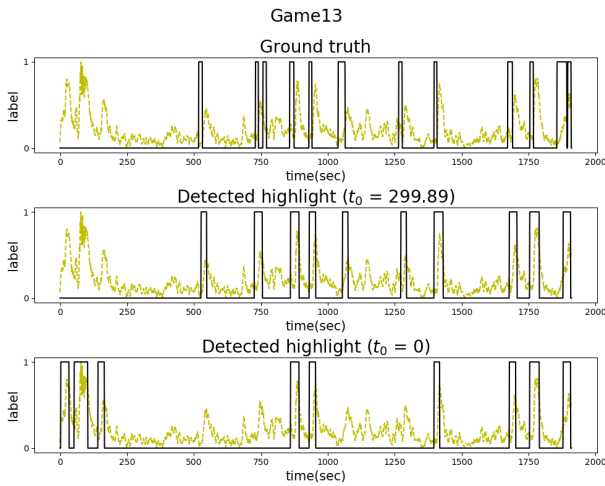


그림 8. 검출 시작 파라미터 t_0 에 따른 하이라이트 검출 성능 비교

Fig. 8. The effect of the time offset t_0

위해 시간 오프셋 t_0 를 고려하였으며, 그림 8은 실제로 성능이 향상된 것을 보인다. Game 13에서 t_0 를 고려하지 않을 때의 F1점수는 60.8%이지만 방송 시작 후 약 5분 후부터 검출을 시작하였을 때의 F1점수는 96.57%로 성능이 35.77% 향상되었다. 마찬가지로 방법으로 전체 영상에서의 F1점수를 비교해보면 시간 오프셋이 없을 때의 F1점수는 58.63%지만 t_0 를 고려했을 때는 62.38%로 F1점수가 3.75% 증가했음을 알 수 있다.

본 논문에서 제안하는 검출 알고리즘은 선정된 피크를 기준으로 하이라이트 구간을 검출하기 때문에 하이라이트 부근에서 피크를 선정하는 것이 중요하다. 따라서 어떤 잡

음 제거 필터를 사용하는가에 의해 성능이 변하게 된다. 이에 필터링을 하지 않은 트래픽과 가우시안 필터, 중앙값 필터, 그리고 moving average 필터를 이용하여 노이즈가 줄어든 트래픽을 이용하여 실험해 보고, 각 필터의 윈도우 크기에 따른 F1점수와 각 필터의 최대값을 그림 9에 나타내었다. 필터의 종류에 관계없이 윈도우의 크기가 너무 커지면 채팅 트래픽이 과하게 단조로워져 상세한 변화 정보를 잃게 된다. 이 때문에 검출 알고리즘의 성능이 낮아지는 경향이 있다. 필터의 윈도우 크기에 상관없이 대체로 moving average 필터가 좋은 성능을 보였으며, 윈도우 크기가 7일 때 F1점수가 가장 높았다.

V. 결 론

본 논문에서는 채팅 정보만을 이용하여 개인방송 하이라이트를 자동으로 검출하는 알고리즘을 제안하였다. 또한, 이 알고리즘으로 검출한 구간의 하이라이트로서의 유효성을 평가하는 방안을 제시하였으며, 실험을 통해 개인방송에서 하이라이트를 검출하기 위한 채팅 트래픽의 이용 가능성을 확인하였다. 제안하는 알고리즘은 채팅 정보만을 이용하기 때문에 연산량이 적어 빠른 시간 안에 하이라이트를 검출할 수 있는 장점이 있다. 차후에는 개인방송 하이라이트 자동 검출의 효용성을 높이기 위해 음향과 영상 정보를 이용하는 방법에 대한 연구를 진행할 계획이다.

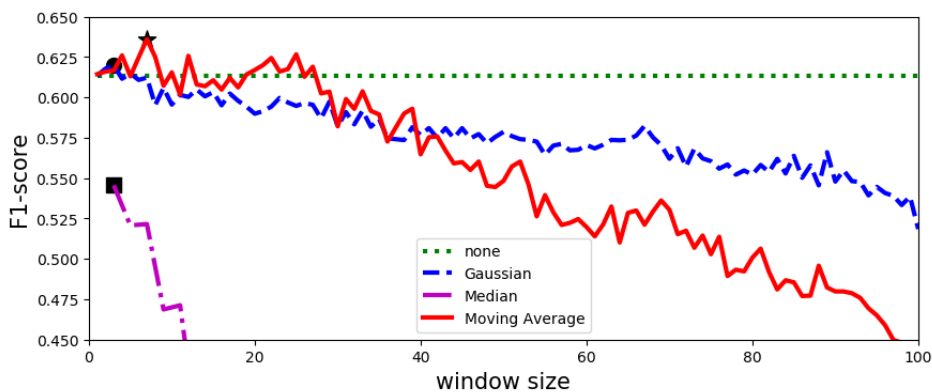


그림 9. 윈도우 크기에 따른 필터별 F1점수 비교

Fig. 9. The F1-score of the each filter

참 고 문 헌 (References)

- [1] DMC Report, "온라인 동영상 시장 현황 및 트렌드(Online Video Market Status and Trend)," 2015.10.
- [2] H. Tang, V. Kwatra, M. E. Sargin, and U. Gargi, "Detecting highlights in sports videos: Cricket as a test case," Multimedia and Expo (ICME), IEEE International Conference on, Barcelona, Spain, pp. 1 - 6, 2011, doi:10.1109/ICME.2011.6012139.
- [3] Y-L. Lin, V. I. Morariu, W. Hsu, "Summarizing While Recording: Context-Based Highlight Detection for Egocentric Videos", 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), pp. 443-451, Santiago, Chile, 2015, doi:10.1109/ICCVW.2015.65.
- [4] M. Gygli, Y. Song, L. Cao, "Video2GIF: Automatic Generation of Animated GIFs from Video", Computer Vision and Pattern Recognition (CVPR) 2016 IEEE Conference on, Las Vegas, NV, USA pp. 1001-1009, 2016.
- [5] M. Sun, A. Farhadi, and S. Seitz, "Ranking domain-specific highlights by analyzing edited videos," European Conference on Computer Vision, Zurich, Swiss, pp. 708-802, 2014, doi:10.1007/978-3-319-10590-1_51.
- [6] Z. Xiong, R. Radhakrishnan, A. Divakaran, T.S. Huang, "Highlights extraction from sports video based on an audio-visual marker detection framework," IEEE International Conference on Multimedia and Expo, Amsterdam, Netherlands, pp. 29-32, 2005.
- [7] Korea Communications Agency, "온라인 개인방송 서비스 이용행태 조사(A Study on the User Behavior of Online Personal Broadcasting Service)," 2014.07.
- [8] Twitch, <https://www.twitch.tv/> (accessed Dec. 28, 2017).

저 자 소 개



김 은 율

- 2017년 : 서울과학기술대학교 전자IT미디어공학과 학사
- 2017년 ~ 현재 : 서울과학기술대학교 정보통신미디어공학과 석사과정
- ORCID : <http://orcid.org/0000-0001-9023-7834>
- 주관심분야 : 머신러닝, 신호처리, 딥러닝



이 계 민

- 2001년 : 서울대학교 전기공학부 학사
- 2007년 : University of Michigan EECS 석사
- 2011년 : University of Michigan EECS 박사
- 2011년 ~ 2012년 : University of Michigan Research Fellow
- 2013년 ~ 현재 : 서울과학기술대학교 전자IT미디어공학과 조교수
- ORCID : <http://orcid.org/0000-0001-6785-8739>
- 주관심분야 : 머신러닝, 신호처리, 의료정보학