

상황 정보를 활용한 동영상 문맥 광고

이봉준¹, 우현욱¹, 이정태¹, 임해창^{2*}

¹고려대학교 컴퓨터·전파통신공학과, ²고려대학교 컴퓨터·통신공학부

Contextual In-Video Advertising Using Situation Information

Bong-Jun Yi¹, Hyun-Wook Woo¹, Jung-Tae Lee¹ and Hae-Chang Rim^{2*}

¹Department of Computer and Radio Communications Engineering, Korea University

²Division of Computer and Communications Engineering, Korea University

요약 동영상 데이터 서비스가 나날이 증가함에 따라 특정 동영상 장면에 적합한 광고를 보여주거나 추가적인 정보를 제공하려는 요구가 커지고 있다. 장면에 적합한 광고를 보여주기 위하여 동영상의 영상이나 음성 정보를 직접 이용하는 방법은 현재의 기술력으로 한계가 있고, 제목, 카테고리 정보, 요약 등의 메타데이터도 계속해서 변화하는 장면의 내용을 반영하지 못한다. 본 연구는 동영상의 대본·자막에서 추출한 장면의 상황 정보를 이용하여 주어진 동영상 장면에 적합한 광고를 자동으로 부착해 주는 새로운 동영상 문맥 광고 시스템을 제안한다. 대본·자막에서 추출한 상황 정보를 광고 검색에 이용했을 때 높은 성능 향상을 확인할 수 있었고, 이를 이용하여 사용자에게 더 적합한 광고를 보여줄 수 있다.

Abstract With the rapid growth of video data service, demand to provide advertisements or additional information with regard to a particular video scene is increasing. However, the direct use of automated visual analysis or speech recognition on videos virtually has limitations with current level of technology; the metadata of video such as title, category information, or summary does not reflect the content of continuously changing scenes. This work presents a new video contextual advertising system that serves relevant advertisements on a given scene by leveraging the scene's situation information inferred from video scripts. Experimental results show that the use of situation information extracted from scripts leads to better performance and display of more relevant advertisements to the user.

Key Words : Video, Scene, Contextual advertising, Situation information

1. 서론

인터넷 공간에서 영화, TV 프로그램, UCC 등 동영상 자료의 양이 기하급수적으로 증가하고 있고, 이러한 동영상의 장면에 적합한 광고를 보여주려는 요구가 커지고 있다. 하지만 동영상 각 장면의 내용을 반영하기 위하여 음성이나 영상 정보를 직접 분석하는 방법은 많은 처리량과 시간이 필요하고, 아직까지 그 정확도 면에서도 실제 적용하기에 충분하지 못하다. 그리고 동영상 내용을

분석하기 위해 간접적으로 사용할 수 있는 제목, 카테고리, 요약 정보 등의 메타데이터는 계속해서 변화하는 장면의 내용을 충분히 반영하지 못한다. 따라서 인터넷에서 동영상과 함께 쉽게 구할 수 있고, 장면의 내용을 잘 반영하고 있는 대본이나 자막 데이터를 이용하는 것이 장면의 내용을 반영한 광고를 위한 실질적인 대안이 될 수 있을 것이다.

동영상은 특정한 사건이 벌어지는 여러 개의 장면들로 이루어져 있기 때문에 장면 별로 그에 적합한 광고를 찾

본 연구는 교육과학기술부 한국연구재단(KRF-2007-361-AL0013) 및 2단계 BK21사업의 지원을 받아 수행되었음.

*교신저자 : 임해창(rim@nlp.korea.ac.kr)

접수일 10년 06월 24일

수정일 10년 07월 27일

게재확정일 10년 08월 10일

[표 1] 항목별 관련 연구 비교

	[1] UFMG	[2] MS	[3] Yahoo	[4] KU	본 연구
광고 부착 대상	웹페이지	웹페이지	웹페이지	동영상 장면	동영상 장면
광고 키워드 출력	X	O	X	O	O
최종 광고 출력	O	X	O	X	O
의미 고려 여부	X	X	O	O	O
외부 리소스	X	X	자체구축 온톨로지	상황 클래스	상황 클래스

는 것이 합당하다. 그리고 각 장면에는 상황이라는 정보가 존재한다. 등장인물의 행동이나 배경에 의해서 결정되는 상황 정보는 식사 하는 장면, 자동차를 타고 이동하는 장면 등 그 장면을 나타내는 중요한 의미적 요소이다.

예를 들어 고급 레스토랑에서 식사를 하는 상황에서 화면에는 잘 보이지 않지만 ‘꽃’이란 소품이 광고 키워드로 추출될 수 있다. 이 경우에 ‘꽃’에 관련된 꽃 배달 광고 보다는 실제 식사라는 상황에 알맞은 음식점 광고가 더 적합할 것이다. 이렇게 상황이라는 정보를 광고 검색 시 이용한다면, 적합하지 않은 광고 키워드로 검색되는 광고를 줄이고, 그 상황에 더 알맞은 광고를 얻을 수 있을 것이다.

본 연구는 동영상의 대본·자막에서 추론한 상황 정보를 반영하는 실질적인 동영상 문맥 광고 시스템을 제안하고, 한국어 동영상을 대상으로 실제 시스템을 구축한다. 동영상 문맥 광고 시스템에서 필요한 정보들과 모듈들을 설명하고, 동영상 각 장면에 존재하는 상황 정보를 광고 검색에 이용함으로써 높은 성능 향상을 보여준다.

2. 관련연구

지금까지 문맥 광고에 관한 연구는 주로 웹페이지를 대상으로 어휘 정보를 분석하여 그에 알맞은 광고를 보여주는 것이었다. 각 관련연구의 항목별 비교는 표 1과 같다. UFMG 대학의 Ribeiro-Neto 교수팀이 수행한 연구 [1]은 웹페이지와 광고에 있는 단어들의 유사성을 고려하여 적합한 광고를 찾는 방법을 보여주고, Microsoft Research에서 수행한 연구 [2]는 웹페이지의 모든 단어가 아닌, 광고에 적합한 광고 키워드를 찾는 데 초점을 맞춘다. Yahoo Research에서 수행한 연구 [3]은 단지 웹페이지와 광고에 존재하는 어휘만을 고려하는 방법에서 벗어나 웹페이지와 광고의 의미적 요소를 고려하기 위하여, 6,000개의 노드를 가지고 각 노드는 평균 100개의 질의어로 이루어진 의미 분류 트리를 구축한다. 구축된 트리에 대해 웹페이지와 광고를 특정 노드에 할당하고, 노드

간의 거리로 그 유사도를 측정한다. 하지만 이런 형식의 지식 베이스를 구축하는 것은 많은 시간과 비용이 드는 작업이다. 본 연구에서는 이와 같은 온톨로지를 구축하지 않고 미리 정의한 상황 클래스로 동영상 장면을 분류하여 광고 검색에 이용하는 방법을 제안한다.

고려대학교에서 수행한 연구 [4]는 광고를 게시하는 대상이 동영상이라는 차이점이 있으며, 동영상에 존재하는 상황이라는 정보를 이용하여 대본에서 광고 키워드를 추출한다. 하지만 컴퓨터가 아닌 사람이 직접 분석한 상황 정보를 사용하였고, 광고 키워드 추출에만 상황 정보를 이용하였다. 본 연구는 [4]의 내용을 더욱 발전 시켜 동영상 장면의 상황을 자동으로 분석하여 키워드 추출에 이용하고, 분석된 상황 정보를 광고 검색에도 활용하여 최종적으로 동영상 장면에 가장 적합한 광고를 찾는 방법을 제안한다.

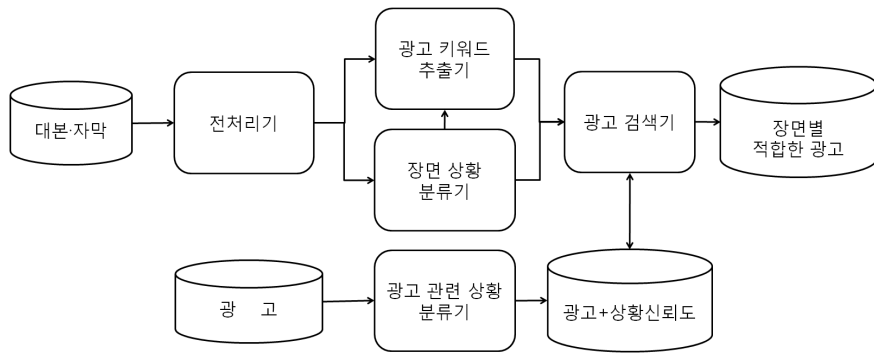
3. 동영상 문맥 광고 시스템

제안하는 시스템(그림 1)은 다음과 같은 부분들로 구성되어 있다. 입력 데이터로써 처리되지 않은 동영상의 자막과 대본이 존재하며, 검색 대상은 광고 데이터이다. 검색할 광고는 웹에서 수집한 약 40만 개의 광고 데이터를 사용한다.

먼저 가공되지 않은 동영상의 자막과 대본을 전처리하여 각 장면으로 분리한다. 전처리 된 장면에서 상황 정보를 추출하고, 광고 데이터에 대해서는 관련 상황을 예측해 낸다. 또한, 전처리 된 장면에서 광고에 적합한 키워드를 추출해 광고 데이터를 대상으로 초기 검색을 수행한다. 장면과 광고에 대해 추출된 상황 정보는 초기 검색 결과를 재순위화 하는데 사용되며, 이 결과 가장 적합한 광고를 사용자에게 보여준다.

3.1 전처리기

인터넷에서 구할 수 있는 대본과 자막은 데이터와 형식이 각각 다르다. 따라서 대본과 자막의 정보들을 합치



[그림 1] 동영상 문맥 광고 시스템 구성도

고 시스템에서 사용하는 형식으로 변환하는 것이 필요한데, 전처리기가 그 기능을 담당한다. 대본은 지문과 대사로 이루어져 있으며, 자막은 대사와 화면에 출력되는 시간 정보를 포함한다. 특히 대본 내 지문은 장소, 소품, 등장인물의 행동 등 동영상의 의미 분석에 필요한 많은 정보들이 포함되어 있기 때문에 상대적으로 중요도가 높다. 또한 대본은 장면으로 구분되어 있으므로 쉽게 장면 분할이 가능하다. 자막 내 시간 정보는 장면의 시작과 끝을 추정하거나, 광고를 게시하는 시간을 설정하는 데 사용된다.

본 시스템은 한국어 동영상을 대상으로 하기 때문에 형태소 분석이 필요하다. 대본·자막 전처리기에서 이 역할을 수행하며, 데이터 접근성을 높이기 위해 대본·자막 정보 및 형태소 분석 정보를 XML 형식으로 저장한다. 실제 입력되는 대본과 전처리된 대본의 예제는 그림 2, 그림 3과 같다.

썬17 썬 북악스카이웨이, 낮.
한결 조깅하며, 이어폰으로 핸드폰 받는,
한 결 : (밝은, 웃고) 서울이야, 만나자.

[그림 2] 실제 대본

```
<SCENE>
<DIRECTION>썬/NNCG ㄴ/PX 17/SCD
썬/NNCG ㄴ/PX 북악스카이웨이/NN? ./SS,
낮/NNCG ./SS. </DIRECTION>
<DIRECTION>한결/AA 조깅/NNCG
하며/PN ./SS, 이어폰/NNCG 으로/PA
핸드폰/NNCG 받/VV 는/EFD
./SS.</DIRECTION>
<LINE TIME=1538720/1540825>한/DU
결/NNCG ./SS:
<DIRECTION>/SS( 밝/VJ 은/EFD ./SS,
웃/VV 고/EFC )/SS</DIRECTION>
서울/NNP 이야/PV ./SS, 만나/VV 자/EFF./SS.
</LINE>
</SCENE>
```

[그림 3] 전처리된 대본

3.2 장면 상황 분류기

동영상 장면에서 존재하는 상황 정보를 사용하기 위해서는 장면이 어떤 상황인지 판단하는 분류기가 필요하다. 동영상에 존재하는 모든 상황을 미리 정의하고 분류하는 것은 실질적으로 어렵기 때문에 20개의 상황 클래스를 정의하여 분류에 사용하였다[4]. 또한, 하나의 장면은 여러 개의 상황을 가질 수 있다고 판단하여 분류기가 개별 상황에 대한 신뢰도를 기반으로 복합 상황 판별이 가능하도록 구축하였다.

20개의 상황 클래스는 아래와 같은 기준으로 정의되었고, 정의된 클래스는 표 2와 같다.

- ① 본 연구는 장면의 상황을 판별하여 그 장면에 더 적합한 광고를 부착하는데 의의가 있으므로 상황 클래스는 광고에 적합한 일반적인 클래스로 정의한다.
- ② 인터넷을 통해 수집된 다량의 상품 카테고리에 대해 연구원(4명)들이 일반적인 상황 클래스를 제안한다.
- ③ 제안된 상황들을 군집화하여 최종적으로 19개의 상황 클래스를 정의한다.
- ④ 도입부와 같이 특정 상황이라고 정의할 수 없는 장면, 복잡한 상황 때문에 판별하기 어려운 긴 장면을 ‘기타’로 정의한다.

[표 2] 정의된 상황 클래스 (순서 없음)

가정생활, 공부, 일, 쇼핑, 통화, 교류, 여가, 여행, 사고, 건강관리, 육아, 미용, 금융거래, 식사, 연애, 결혼, 갈등, 이동, 종교활동, 기타

본 연구에서는 동영상 메타데이터인 대본·자막을 이

용한다. 대본·자막은 텍스트 데이터이므로 일반적인 텍스트 분류 관점에서 연구를 진행하였다. 한 장면을 최대 3가지의 복수 상황으로 사람이 직접 태깅하여 학습 데이터를 구축하고, 지도 학습 기법 중 최대 엔트로피 모델[6]을 사용하여 분류기를 구현하였다. 장면의 텍스트 데이터를 x 라 가정했을 때, x 에 대한 상황이 y 일 조건부 확률은 최대 엔트로피 모델을 이용하여 수식 (1)과 같이 구할 수 있다.

$$p(y|x) = \frac{1}{Z} \exp\left(\sum_i \lambda_i f_i(x, y)\right) \quad (1)$$

수식 1에서 $f_i(x, y)$ 는 i 번째 자질이고, λ_i 는 i 번째 자질 함수이며, Z 는 모든 가능한 y 에 대하여 $p(y|x)$ 의 합이 1이 되게 하는 정규화 상수이다. 분류기 구현에 사용한 자질은 단어의 tf-idf 값과 단어와 상황간의 PMI(Pointwise Mutual Information) 값의 평균이다. 여기서 단어의 tf-idf 값은 문서 분류에서 널리 사용되는 방법이기 때문에 기본 평가 척도로 삼았다. 단어·상황 간의 PMI 값은 특정 상황을 잘 나타내는 단어들이 있을 것이라는 가정 하에 적용하였다. 단어 t 와 상황 클래스 c 와의 PMI값 계산은 [5]에서 이용한 방법과 유사하며, 수식 (2)와 같다:

$$PMI(t, c) = \log_2 \frac{p(t, c)}{p(t)p(c)} = \log_2 \frac{\frac{\text{count}(t \in c)}{\text{count}(\text{word} \in c)}}{\frac{\text{count}(t)}{\text{count}(\text{word})} \cdot \frac{\text{count}(\text{word} \in c)}{\text{count}(\text{word})}} \quad (2)$$

[표 3] 상황 클래스와 단어의 PMI 값

단어 \ 상황클래스	여가	건강관리	식사
꽃	6.1184	0	3.1998
식당	4.3595	4.2327	5.2481
포크	0	0	5.2481

표 3은 이러한 계산식을 이용하여 정답 상황이 태깅되어 있는 10개의 드라마 대본에서 ‘여가’, ‘건강관리’, ‘식사’ 상황과 관련 있는 단어의 PMI 값을 계산한 결과이다. 여가 생활(꽃꽂이)과 관련 있는 꽃이라는 단어는 ‘여가’ 상황 클래스와 PMI값이 높지만 식사하는 테이블에 꽃이 있을 수도 있으므로 ‘식사’ 상황 클래스와도 관련이 있다. 그리고 식당은 차를 마시며 여가 생활을 즐길 수도 있고, 건강 관리를 위해 특별한 음식을 먹을 수도 있기 때문에 세가지 상황 클래스와 모두 관련이 있으나 ‘식사’ 상황

클래스와 가장 관련성이 높다. 포크는 ‘여가’, ‘건강관리’ 상황 클래스와는 관련이 없고, ‘식사’ 상황 클래스만 관련이 있다. 이러한 직관적인 사실들을 단어와 상황 클래스와의 PMI값으로 명확하게 확인할 수 있다.

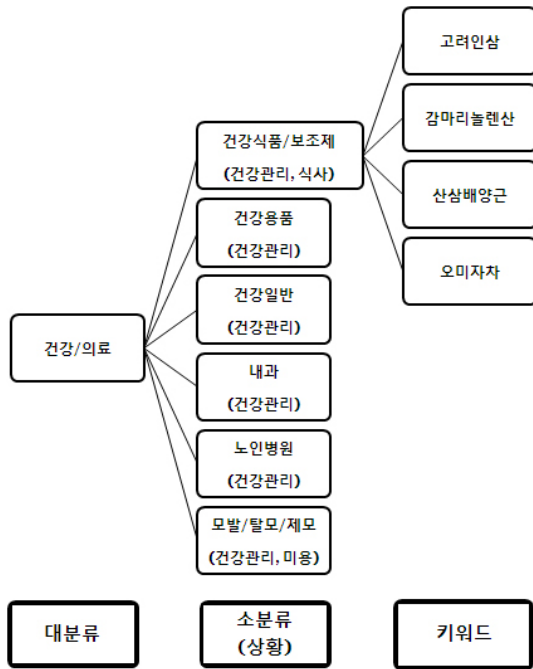
하지만 각 단어와 20가지 클래스와의 PMI값을 직접 자질로 사용하는 방법은 자료 부족 문제를 더 부각시키기 때문에 장면에 존재하는 모든 단어들과 상황간의 PMI 값을 상황별로 평균 내어 학습 자질로 사용하였다. 그리고 ‘기타’ 클래스는 여러 가지 상황이 섞여 있는 클래스이기 때문에 이런 경향성을 찾을 수 없어, 19개의 상황으로만 일단 분류한 후, 그 신뢰도가 특정 임계치 미만인 장면을 ‘기타’ 상황으로 분류하였다.

3.3 광고 관련 상황 분류기

상황 정보를 광고 검색에 활용하기 위해서는 광고의 상황 정보 또한 필요하다. 하지만 광고가 특정 상황이라고 말하기는 어려우므로 광고와 관련 있는 상황으로 광고를 분류하는 광고 관련 상황 분류기를 구축하였다. 광고의 상황 정보로 광고 재순회화를 수행하기 위하여 각 상황 클래스의 분류 신뢰도를 출력한다.

광고 관련 상황 분류기는 장면 상황 분류기와 같은 방법으로 구축하였다. 하지만 광고 텍스트는 대본·자막 텍스트에 비해 상대적으로 적은 양의 단어들로 구성되어 있기 때문에 학습을 위한 데이터가 매우 부족하다. 이를 보완하기 위해 광고에 대한 대량의 학습 데이터를 수작업으로 구축하는 것은 많은 시간과 비용이 든다. 따라서 본 연구에서는 계층화된 광고 키워드를 이용하여 반자동으로 대량의 학습 데이터를 구축하는 방법을 사용하였다. 모든 키워드의 관련 상황을 태깅하는 것은 많은 비용이 소요되기 때문에, 소분류 계층의 키워드를 관련 상황으로 태깅하여 하위 키워드에 대한 연관 상황을 결정한다. 그림 4의 예를 보면, “건강식품/보조제”와 같은 소분류 키워드는 “건강관리”, “식사” 라는 관련 상황을 가지며, 하위의 “고려인삼”, “감마리놀렌산”, “산삼배양근”, “오미자차”와 같은 하위 키워드로 검색된 광고들은 모두 “건강관리”, “식사”의 관련 상황을 가진다.

이러한 학습 데이터 확장에는 정제된 키워드와 광고 검색 엔진이 필요하다. 키워드의 경우, 기업주들을 위해 광고 업체에서 제공하는 공개된 광고 키워드(오버추어 키워드팩)를 사용한다. 이러한 키워드팩은 조회수, 가격 등의 정보를 가지고 있으므로 정제된 광고 키워드 리스트를 얻을 수 있다. 광고 검색은 검색 엔진에 키워드를 입력하여 키워드와 관련 있는 광고를 수집하였다.



[그림 4] 계층화된 광고 키워드

3.4 광고 키워드 추출기

시스템 내의 광고 키워드 추출은 장면 내에서 광고가 가능한 키워드를 추출하는 것이다. 웹 페이지를 대상으로 하는 일반적인 광고 키워드 추출기의 동작은 tf-idf, POS 패턴 등의 빈도수에 의존적인 어휘 자질을 사용하여 문서 내에서 광고에 적합한 키워드를 결정한다. 이러한 방법은 충분한 텍스트를 가진 웹 페이지에서 좋은 성능을 보여준다.

그러나 오류가 존재하고 텍스트의 양이 상대적으로 부족한 동영상 대본에서 어휘 자질만으로 좋은 성능을 보이는 어렵다. 특히, 한국어 대본에서는 키워드의 중의, 형태소 분석 오류와 같은 추가적인 문제점이 발생한다. 이를 해결하기 위하여 광고 키워드 추출에도 동영상의 특성을 반영하는 상황 자질을 사용한다[4].

기존의 연구[2]는 키워드 추출을 분류 문제로 접근하여 해결한다. 여기에서 분류 기준은 문서 내 후보 키워드의 광고에 대한 적합성이다. 본 연구는 학습 데이터 내의 여러 가지 자질로 최대 엔트로피 모델[6]을 이용하여 후보 키워드의 적합성 여부를 판단한다. 기존 연구에서 사용된 자질들은 표 4와 같다.

[표 4] 기존 연구에서 사용된 자질

빈도 기반	TF, DF 통계
질의 로그	존재 여부, 클릭수, 키워드 가격
위치 기반	상대적 위치, 출현 위치
길이 기반	단어 길이, 장면 길이
어휘 기반	단어 패턴, 형태소 패턴

빈도 기반 자질은 정보 검색 분야에서 사용되는 단어의 통계 자질로써 텍스트 분류에 일반적으로 사용된다. 질의 로그는 문맥 광고에 대한 사용자 피드백을 의미한다. 그러나 현재 서비스 하고 있는 시스템이 없는 상황에서 사용자 피드백을 직접 구할 수는 없다. 이를 대체하기 위해 오버추어 코리아에서 제공되는 키워드 팩을 이용하여 질의 로그 자질과 유사한 정보를 구축하였다. 위치 기반 자질은 후보 키워드가 대본 내에 나타나는 상대적 출현 위치와 대사·지문과 같은 대본의 특수한 위치 정보를 사용한다. 길이 기반 자질은 후보 키워드의 단어 길이와 장면 전체의 길이를 사용하며, 어휘 자질은 단어 패턴, 형태소 패턴에 대한 빈도 정보를 사용한다.

상황과 관련된 자질들은 표 5와 같다. 상황 유형은 장면 상황 분류기에 의해 분류된 상황 클래스 자체를 나타낸다. 단어-상황 연관도는 단어와 장면 상황과의 PMI 값을 의미한다. 상대적 연관도는 단어에 대한 평균 PMI 값과 장면 상황과 일치하는 단어-상황 연관도의 차를 의미한다. 이러한 연관 척도를 자질로 사용하여 장면의 상황과 관련 있는 광고 키워드 후보가 상위에 위치할 수 있도록 한다.

[표 5] 상황 자질

상황 유형	장면 상황 분류기(3.2)에 의해 분류된 상황
단어-상황 연관도	후보 키워드가 가지는 단어-상황 PMI 값
상대적 연관도	평균 PMI와 단어-상황 연관도의 차

3.5 광고 검색기

위의 여러 모듈에 의해 장면의 상황, 광고의 관련 상황 분류 신뢰도, 광고 키워드라는 정보를 추출하였다. 광고 검색기는 위의 정보를 이용하여 장면에 가장 적합한 광고를 부착하는 모듈이다. 데이터를 이용하기 위하여 기존의 검색 모델을 사용한 초기 검색이 필요하다. 초기 검색은 현재 사용되고 있는 검색 모델 중 어떤 것이라도 사용

이 가능하지만, 본 연구에서는 원활한 실험 및 데이터 적용을 위하여 Indri 검색 엔진[10]을 사용한다.

장면에 존재하는 광고 키워드들로 초기 검색을 수행하여, 키워드 하나당 상위 100개씩의 광고를 추출한다. 이 광고들을 합쳐서 후보 광고로 정의하고 장면의 상황, 광고의 상황 분류 신뢰도를 이용하여 재순위화 한다. 초기 검색 시 얻을 수 있는 검색 신뢰도와 상황 정보를 여러 가지 방법으로 결합할 수 있지만, 본 연구에서는 상황 정보가 광고 검색에 미치는 영향을 알아보기 위하여 순수하게 상황 정보만을 이용하여 재순위화 하였다. 후보 광고가 가지고 있는 각 상황의 신뢰도 중, 장면의 상황과 일치하는 상황의 신뢰도 값들만 더하여 이 값으로 최종 순위를 정한다. 이렇게 재순위화 된 광고 중, 상위 몇 개의 광고만 선택하여 사용자에서 보여준다. 재순위화를 위한 점수 계산식은 수식 (3)과 같다:

$$score = \sum_{s \in sit} r_s \quad (3)$$

sit = 동영상 장면의 상황들
 r_s = 광고의 s 상황에 대한 분류 신뢰도

3.6 전체 과정 예제

그림 5는 시스템이 대본을 입력받아서 그에 적합한 광고를 찾는, 전체 과정에 대한 실제 예제이다. 그림과 같이 대본이 입력되면 전처리기를 거쳐서 그 장면의 상황이 분석된다. 예제에서는 차를 타고 도착한 상황이기 때문에 ‘이동’ 상황으로 분류되었다. 광고 DB에 있는 광고들도 어떤 상황인지 미리 분류하여 분류 신뢰도와 함께 저장

한다. 대본에서 광고에 적합한 광고 키워드를 추출하면 차, 전화, 벨트, 핸드백 등이 나오고, 일정 신뢰도 이상의 광고 키워드만을 사용하기 때문에 차, 전화, 벨트, 3가지 키워드로만 광고를 검색한다. 이렇게 검색된 광고를 ‘이동’이라는 장면에 적합한 광고인가를 기준으로 재순위화 하여, 최종적으로 ‘이동’에 관련된 광고들이 상위에 위치하게 된다.

4. 실험

4.1 실험 데이터 및 환경

실험에 사용된 동영상은 드라마 10편이다. 각 드라마 당 1화부터 5화까지 총 50화의 대본과 자막을 직접 인터넷에서 수집하여 실험하였다. 전처리 후 생성된 장면은 모두 3,281개이며, 3명의 언론학과 학생이 각 장면에 복수개의 상황 클래스와 광고 키워드를 태깅하여 학습 데이터를 구축하였다. 장면 상황 분류와 광고 키워드 추출은 모두 최대 엔트로피 모델[6]을 사용하여 분류기를 구축하였으며, 결과 검증을 위하여 10-fold 교차 검증을 수행하였다. 광고 관련 상황 분류 실험은 대량의 정답 데이터를 구축하는 데 많은 시간과 비용이 필요하기 때문에 간접적으로 광고 검색 성능을 비교하는 것으로 대신하였다.

4.2 장면 상황 분류

분류기 구축 시 단어의 tf-idf값을 자질로 사용한 것을



[그림 5] 전체 과정 예제

베이스라인으로 하였다. 실험 설정은 1. 상황별 PMI 평균 값(pmi avg.) 사용, 2. tf-idf 값과 PMI 평균값을 함께 사용(tf-idf + pmi avg.), 3. tf-idf 값에 자질 선택(feature selection)을 적용한 것과 PMI 평균값을 함께 사용(tf-idf + fs + pmi avg.)한 것으로 설정 하였다. 복합 상황으로 분류한 결과를 평가하기 때문에 평가 척도는 precision(수식 (4)), recall(수식 (5)), F-measure(수식 (6))을 사용하였다.

$$precision = \frac{|정답\ 상황들 \cap\ 분류된\ 상황들|}{|분류된\ 상황들|} \quad (4)$$

$$recall = \frac{|정답\ 상황들 \cap\ 분류된\ 상황들|}{|정답\ 상황들|} \quad (5)$$

$$F-measure = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (6)$$

표 6을 보면 일반적으로 텍스트 분류에 사용하는 tf-idf 자질보다 본 연구에서 제안한 PMI 값의 평균을 자질로 사용했을 때 모든 평가 척도에서 높은 성능 향상을 이룬 것을 확인할 수 있다(실험 1). tf-idf값과 PMI 평균값을 함께 사용했을 때 낮은 성능 향상을 보이는 것은 tf-idf 자질이 전체 자질 공간에서 차지하는 크기(13,003개)가 PMI 평균값(19개)에 비해 상대적으로 크기 때문인 것으로 분석된다(실험 2). 이를 확인하기 위하여 문서 분류 시 간단하지만 좋은 성능을 보이는 자질 선택 방법인 문서 빈도 수(document frequency)[7]를 기준으로 선택한 tf-idf 자질과 PMI 값의 평균을 함께 자질로 사용해 보았다(실험 3). 그러나 학습 데이터가 충분히 많지 않은 상황(약 3,000개)에서 자질 선택을 적용하였기 때문에 실험 2의 경우보다 더 낮은 성능을 보였다.

[표 6] 장면 상황 분류 성능

	Precision	Recall	F-measure
baseline (tf-idf)	0.4554	0.4347	0.4448
실험 1 (pmi avg.)	0.6225 (+36.69%)	0.6070 (+39.64%)	0.6147 (+38.18%)
실험 2 (tf-idf + pmi avg.)	0.5075 (+11.44%)	0.5111 (+17.58%)	0.5093 (+14.50%)
실험 3 (tf-idf + fs + pmi avg.)	0.5046 (+10.80%)	0.5126 (+17.92%)	0.5086 (+14.34%)

4.3 키워드 추출

키워드 추출 실험에는 3명의 언론학과 학생들이 정답 광고 키워드를 태깅한 평가 데이터를 사용하였다. 상황 자질을 반영하는 평가를 수행하기 위하여 장면의 상황에 대한 세 가지 단계의 적합성(1=비적합, 2=관련있음, 3=적합)을 정답 광고 키워드에 부여하였으며, 본 논문에서는 3으로 태깅된 키워드를 정답으로 간주하여 평가하였다.

평가 척도로는 Top-n 척도[2]를 사용한다. 여기서 n은 후보 키워드 중 상위 순위에 순위 부여된 n개를 의미하며, 본 연구에서는 Top-1, Top-5, Top-10의 척도를 사용한다. Top-n 척도는 수식 7과 같이 계산 할 수 있다.

$$Top-n = \frac{\sum_i |K_i \cap A_i|}{\sum_i \min(|A_i|, n)} \times 100\% \quad (7)$$

여기서, K_i 는 장면 i 에 대하여 시스템이 출력한 상위 n 개의 키워드 집합이며, A_i 는 장면 i 의 정답 키워드 집합이다.

표 7을 보면 기존 연구[2]에 대비하여, 추가적인 자질로 상황 자질을 사용하였을 때, 전체적인 성능 향상이 있음을 알 수 있다. 특히, Top-1의 성능 향상이 두드러지는데, 이는 텍스트 내에서 중요도를 가지는 것과 동시에 장면의 상황을 잘 반영하는 광고 키워드를 추출할 수 있음을 의미한다.

[표 7] 키워드 추출 성능

	Top-1	Top-5	Top-10
Baseline	41.91	65.72	81.33
Baseline+ Situation	48.51 (+15.7%)	70.08 (+6.63%)	84.12 (+3.43%)

4.4 광고 검색

검색에 사용한 광고 데이터는 1,083,938개의 광고를 수집하여 사용하였고, 그 중 약 60%의 광고를 광고 상황 분류기를 위한 학습 데이터로 사용하였고, 나머지 약 40%의 광고를 관련 상황으로 분류하여 실제 광고 검색에 사용하였다.

3,281개의 장면 중 임의로 선정된 75개 장면을 실험에 사용하였고, 광고 키워드 추출기를 사용하여 실제 광고 키워드를 추출한 결과, 장면 당 광고 키워드는 평균 3.25개였다. 추출한 광고 키워드로 Indri 검색 엔진을 사용하여 초기 검색을 수행한 결과에 대해 장면이 적합한 광고는 2점, 아주 적합하지는 않지만 장면과 관련 있는 광고

는 1점, 전혀 관련 없는 광고는 0점으로 나누어 광고 검색 정답 집합을 구축하였다.

각 장면에서 추출한 광고 키워드로 Indri 검색 엔진을 이용하여 초기 검색을 수행한 결과를 baseline으로 하였고, 앞에서 제안한 PMI 평균을 이용하여 분류한 장면의 상황과 광고의 상황 신뢰도만을 이용하여 재순위화 한 결과를 nDCG@1, @3, @5를 이용하여 비교하였다. 동영상 문맥 광고의 특성상 광고를 보여줄 수 있는 장소가 한정되어 있기 때문에 상위 광고 몇 개의 성능을 평가하는 것이 의미가 있을 것이라고 생각하였고, nDCG는 검색된 순위와 정확도를 모두 반영하는 평가 척도이기 때문에 선택하였다[8]. nDCG@k의 계산식은 수식 (8)과 같고, Z는 nDCG의 최대값이 1이 되게 하는 정규화 상수이며 s(p)는 p 순위 문서의 적합도이다.

$$nDCG@k = \frac{1}{Z} \sum_{p=1}^k \frac{2^{s(p)} - 1}{\log(1+p)} \quad (8)$$

표 8은 nDCG@1, @3, @5 모든 부분에서 인상적인 성능 향상이 있음을 보여준다. Indri 검색 신뢰도, 키워드 추출 시 얻을 수 있는 신뢰도(광고 키워드로서의 적합도), 광고의 상황 신뢰도 등을 이용하여 여러 가지 조합으로 실험해 보았지만, 광고의 상황 신뢰도만을 사용하였을 때 가장 좋은 결과를 보여주었다.

[표 8] 광고 검색

	nDCG@1	nDCG@3	nDCG@5
Indri (baseline)	44.67	45.62	45.66
Rerank with Situation	47.33 (+5.95%)	47.87 (+4.93%)	50.02 (+9.55%)

5. 결론

본 연구는 동영상 기반의 문맥 광고 시스템을 제안하고, 시스템에 필요한 모듈들을 다루며, 동영상 메타 데이터에서 추출한 상황 정보를 이용하는 새로운 문맥 광고 시스템을 개발하고 평가하였다. 즉, 동영상 환경에서 상황이라는 의미 요소를 광고 검색에 반영하여 기존의 문맥 광고 시스템의 성능을 향상하는 방법을 실험하였다.

연구 결과로서 동영상의 장면과 광고의 관련 상황을 자동으로 분류하고, 문맥 광고에서 사용되는 키워드 추출, 광고 검색 모듈의 성능이 새로운 상황 자질을 사용하

여 향상될 수 있음을 보인다. 이는 동영상 장면이 가지는 의미적 요소를 문맥 광고 시스템에 반영한 결과이며, 상황이 동영상 장면의 내용을 반영하는 좋은 자질을 보여준다.

참고문헌

- [1] B. Ribeiro-Neto, et al., "Impedance Coupling in Content-targeted Advertising", Proc. of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 496-503, August, 2005.
- [2] W. Yih, et al., "Finding Advertising Keywords on Web Pages", Proc. of the 15th international conference on World Wide Web, pp. 213-222, May, 2006.
- [3] A. Broder, et al., "A Semantic Approach to Contextual Advertising", Proc. of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 559-566, July, 2007.
- [4] J. Lee, et al., "Finding Advertising Keywords on Video Scripts", Proc. of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 686-687, July, 2009.
- [5] G. Mishne, et al., "Experiments with Mood Classification in Blog Posts", Proc. of ACM SIGIR 2005 Workshop on Stylistic Analysis Of Text For Information Access, August, 2005.
- [6] R. Malouf, "A comparison of algorithms for maximum entropy parameter estimation", Proc. of the 6th conference on Natural language learning, pp. 1-7, September, 2002.
- [7] Y. Yang, et al., "A Comparative Study on Feature Selection in Text Categorization", Proc. of International Conference on Machine Learning, pp. 412 - 420, July, 1997.
- [8] K. Järvelin, et al., "Cumulated Gain-Based Evaluation of IR Techniques", ACM Transactions on Information Systems, Vol. 20, No. 4, pp. 422-446, October, 2002.
- [9] D. Metzler and W. B. Croft, "Combining the Language Model and Inference Network Approaches to Retrieval", Information Processing & Management, Volume 40, Issue 5, pp. 735-750, September 2004.

이 봉 준(Bong-Jun Yi)

[준회원]



- 2009년 2월 : 고려대학교 컴퓨터학과 (공학사)
- 2009년 3월 ~ 현재 : 고려대학교 대학원 컴퓨터·전파통신공학과(석사과정)

<관심분야>

정보검색, 자연어처리, 분산처리, 데이터 마이닝

임 해 창(Hae-Chang Rim)

[정회원]



- 1983년 12월 : University of Missouri-Columbia 대학원 전산학과 (전산학석사)
- 1990년 12월 : University of Texas-Austin 대학원 전산학과 (전산학박사)
- 1991년 9월 ~ 현재 : 고려대학교 컴퓨터·통신공학부 교수

<관심분야>

자연어처리, 정보검색

우 현 욱(Hyun-Wook Woo)

[준회원]



- 2009년 2월 : 고려대학교 컴퓨터학과 (공학사)
- 2009년 3월 ~ 현재 : 고려대학교 대학원 컴퓨터·전파통신공학과(석사과정)

<관심분야>

정보검색, 문맥광고, Community Question Answering

이 정 태(Jung-Tae Lee)

[정회원]



- 2008년 2월 : 고려대학교 대학원 컴퓨터·전파통신공학과 (컴퓨터학 석사)
- 2008년 3월 ~ 현재 : 고려대학교 대학원 컴퓨터·전파통신공학과 박사과정

<관심분야>

정보검색, 자연어처리