

GoF 특징을 이용한 유해 동영상 자동 분류

이승민, 이호균, 남택용
한국전자통신연구원 정보보호연구단 개인정보보호연구팀
e-mail : toptom@etri.re.kr

Automatic Classification of Objectionable Videos Based on GoF Feature

Seungmin Lee, Hogyun Lee, Taekyong Nam
Privacy Protection Research Team, Information Security Division, ETRI

본 논문은 유해한 동영상을 실시간으로 분석하고 차단하기 위하여, 동영상의 비주얼 특징으로서 그룹 프레임(Group of Frame) 특징을 추출하여 SVM 학습모델을 활용하는 유해 동영상 분류에 관한 것이다. 지금까지 동영상 분류에 관한 연구는 주로 입력 동영상을 뉴스, 스포츠, 영화, 뮤직 비디오, 상업 비디오 등 사전에 정의한 몇 개의 장르에 자동으로 할당하는 기술이었다. 그러나 이러한 분류 기술은 미리 정의한 장르에 따른 일반적인 분류 모델을 사용하기 때문에 분류의 정확도가 높지 않다. 따라서, 유해 동영상을 실시간으로 자동 분류하기 위해서는, 신속하고 효과적인 동영상 내용분석에 적합한 유해 동영상 특화의 특징 추출과 분류 모델 연구가 필요하다.

본 논문에서는 유해 동영상에 대하여 신속하고, 정확한 분류를 위하여 유해 동영상의 대표 특징으로서 그룹프레임 특징을 정의하고, 이를 추출하여 SVM 학습 모델을 생성하고 분류에 활용하는 매우 높은 성능의 분석 방법을 제시하였다. 이는 최근 인터넷 뿐만 아니라 다양한 매체를 통하여 급속도로 번지고 있는 유해 동영상 차단 분야에 적극 활용될 수 있을 것으로 기대된다.

1. 서론

21 세기는 인터넷 시대이다. 정보화 혁명을 이끄는 인터넷을 이용하여, 사람들은 시간과 공간을 뛰어 넘어 언제, 어디서든 손쉽게 유용한 정보를 획득할 수 있게 되었다.

하지만 인터넷은 유용한 정보와 손쉬운 활용이라는 순기능과 더불어 사회적으로 통제를 필요로 하는 유해한 정보 역시 인터넷을 이용하는 이용자들에게 무차별적으로 제공함으로써 역기능을 발생시키고 있다. 즉, 사회적인 보호를 받아야 하는 청소년을 비롯한 판단력과 절제력이 부족한 인터넷 이용자들이 인터넷의 유해 정보를 아무런 제재 없이 접근할 수 있게 되어서 개인 뿐만 아니라 사회적인 문제가 되고 있다.

인터넷을 통해 유해 정보를 접하는 경로는, 유해 사이트, 게시판이나 블로그의 유해 정보, 유해 정보를 포함한 스팸 메일 수신, P2P 를 이용한 유해 정보 공유 등 여러 가지가 있다. 이러한 다양한 경로를 통한 유해 정보는 인터넷 통신 속도와 미디어 저장 용량의

증가로 인하여 최근 유해 콘텐츠의 미디어 타입이 문자와 이미지에서 동영상으로 급속히 바뀌고 있는 추세이다. 동영상의 유해물은 문자와 이미지에 비교하여 훨씬 자극적이기 때문에 보다 심각한 사회문제를 야기시킬 수 있다.

따라서, 다양한 방식으로 빠르고 쉽게 접할 수 있는 유해 동영상을 차단하기 위하여 유해 동영상에 대한 자동 분류 기술의 확보가 시급한 실정이다.

그러나, 유해 동영상의 분류 기술 연구는 아직 초기 단계이며, 이를 상용화하여 제품화 하기 위해서는 처리속도와 정확도 등에서 해결해야 할 문제가 많다고 하겠다.

본 논문에서는 “유해”를 “음란”의 의미로 제한하고, 유해 동영상에 대한 자동 분류를 위하여, 그룹프레임 특징을 사용한 분류를 시도하였으며, 구체적으로 2 절에서 관련 연구 동향을 살펴보고, 3 절에서 특징 추출 및 학습 방법과 분류 과정을 자세히 설명하고, 4 절에서 분류 실험 결과를 살펴보기로 한다. 마지막으로 5 절에서 결론을 맺기로 한다.

2. 관련 연구

동영상의 내용을 분석하여 자동 분류하는 연구가 최근까지 활발히 진행되어왔다[1].

지금까지 연구된 동영상 분류 방식은 대부분 입력 동영상을 뉴스, 스포츠, 영화, 뮤직 비디오, 상업 비디오 등 사전에 정의한 몇 개의 장르에 자동 할당하는 기술이다[2][3].

이와 같이 사전에 몇 개의 장르로 구분하는 경우는 장르의 객관성이 약하다. 즉, 동영상의 자동 분류를 위하여 적용한 알고리즘에 따라 동일한 영화 장르라 하더라도, 동영상의 내용 아닌 동영상의 길이, 프레임 크기, 비트율 등에 따라 사전에 다른 장르로 구분하는 경우가 많다[3]. 예를 들어, 같은 영화 장르라 하더라도 액션 영화와 애정 영화를 다른 장르로 구분하고, 코미디 동영상과 애정 영화를 같은 장르로 구분하는 경우이다. 따라서 실험결과에 관계없이 현실성이 떨어진다고 볼 수 있다.

그리고, 입력 동영상을 사전에 정의한 장르로 자동 분류하기 위하여, 일반적인 분류 모델을 생성하기 때문에 정확도가 높지 않다. 따라서, 본 논문과 같이 입력 동영상을 정확도가 높게 특정 장르로 분류하기 위해서는, 특정 장르에 적합한 특징 추출과 분류 모델 생성 방법이 필요하다.

또한, 기존의 연구는 오프라인에서 내용을 분석하여 해당 장르의 특징을 검색하는 방식이 대부분이기 때문에, 신속하고, 효과적인 내용 분석 기술이 필요하다.

그리고, 실험결과의 정확도를 보장하기 위해서는 다양한 분야의 많은 양의 데이터를 사용하여 분류 실험을 수행할 필요가 있다.

본 논문에서는 유해 동영상의 자동 분류를 위하여, 모두 800 건의 동영상을 대상으로 5 개 장르를 가지고 실험을 수행하였으며, 분류 모델을 생성하기 위하여 휴먼 디텍션에 적합한 HSV 스펙트럼으로부터 그룹 프레임의 특징을 추출한 후, 이를 사용하여 SVM 학습 모델을 생성하였다.

3. GoF 를 이용한 유해 동영상 분류

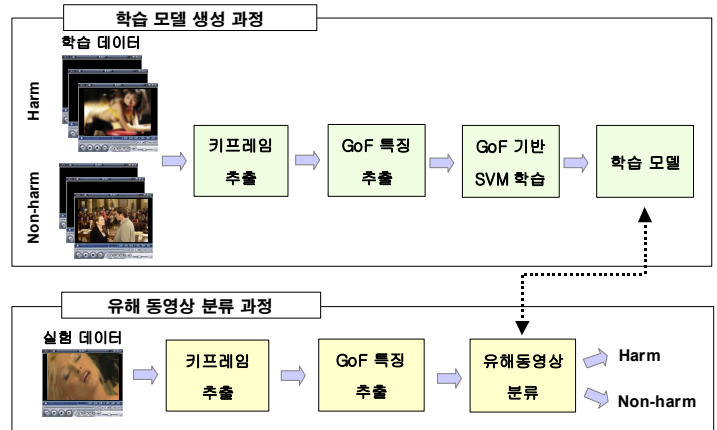
본 절에서는 유해 동영상 자동 분류의 전체적인 과정과 데이터 수집, 특징 추출 및 모델 생성, 유해 동영상 분류에 대하여 상세하게 설명하기로 한다.

(그림 1)은 유해 동영상 자동 분류를 위한 학습 모델 생성과 분류과정을 도시한 것이다. 즉, GoF(Group of Frame) 특징을 이용한 유해 동영상 분류 과정을 구성하는 기능을 나타낸 것으로, 학습 모델 생성과정과 유해 동영상 분류과정으로 구성된다.

학습모델 생성과정은, 동영상 데이터에서 동영상별 정지영상의 키 프레임(n 개)을 추출하는 기능, 추출한 프레임으로부터 히스토그램 특징을 추출하여 동영상별 그룹 프레임의 비주얼 특징인 GoF 를 생성하는 기

능, 그리고 유무해 동영상에 대하여 GoF 를 사용하여 SVM(Support Vector Machine) 학습과정을 통하여 분류 모델 생성 기능으로 구성된다.

유해 동영상 분류과정은, 실험 동영상으로부터 정지영상의 키 프레임을 추출하는 기능, 추출한 프레임으로부터 히스토그램 특징을 추출하여 실험 동영상에 대한 GoF 를 생성하는 기능, 그리고 실험 동영상의 GoF 를 학습모델 생성과정에서 구한 분류모델과 비교하여 유무해를 판별하는 기능으로 구성된다.



(그림 1) GoF 를 이용한 유해 동영상 자동분류 과정

3.1 데이터 수집

동영상의 분류의 정확도를 보장하기 위하여, 학습 및 실험 데이터 수집은 매우 중요한 부분이다. 학습 모델 생성과 분류 실험을 위하여 유해 동영상과 무해 동영상 모두 800 건을 수집하였으며, 무해 동영상의 경우에는 데이터의 다양성을 위하여 영화, 다큐멘터리, 스포츠, 뮤직 비디오 등의 장르를 대상으로 수집하였다. 특히, 사람이 등장하는 내용의 동영상을 주로 수집함으로써 유해 동영상 분류의 선별력을 높이고자 했다.

<표 1> 학습 및 실험 데이터

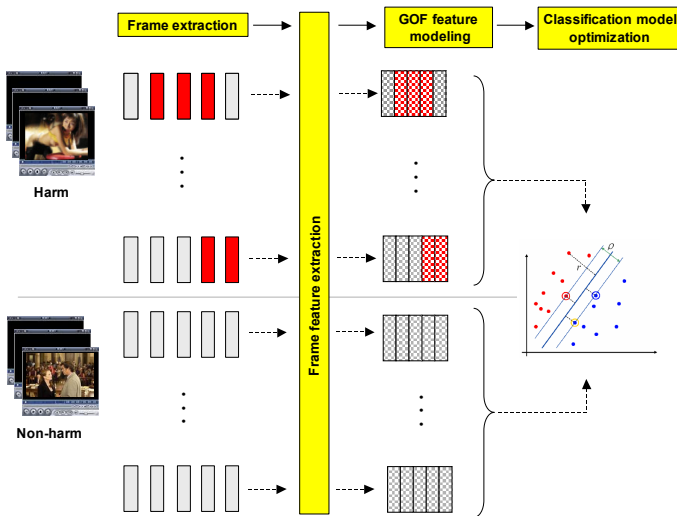
동영상		정지영상	
유해 동영상	영화	F _{1,1} F _{1,60}	:
	다큐멘터리	F _{100,1} F _{100,60}	
무해 동영상	영화	F _{1,1} F _{1,60}	:
	다큐멘터리	F _{100,1} F _{100,60}	
	스포츠	F _{1,1} F _{1,60}	
	뮤직비디오	F _{100,1} F _{100,60}	

<표 1>은 학습 모델 생성과 분류 실험 동영상 800

건을 나타낸 것으로서, 유해 동영상 400 건과 영화, 다큐멘터리, 스포츠, 뮤직 비디오 등 4 가지 장르 각각 100 건의 무해 동영상으로 구성하였다. 그리고, GoF 특징 추출을 위하여 각각의 동영상에 대하여 60 개의 키프레임을 일정간격으로 추출하여 정지영상을 수집하였다.

3.2 특징 추출 및 모델 생성

(그림 2)는 학습모델 생성과정을 도시한 것으로서, 학습 동영상에서 키프레임을 추출하여 256 개의 특징값을 갖는 GoF 를 생성한다. 이 특징값을 가지고 기계학습 방법의 하나인 SVM 을 이용하여 학습모델을 생성한다[4].



(그림 2) 학습모델 생성 과정

동영상별 GoF 값을 생성하기 위하여, 우선 동영상별로 추출한 60 개의 키 프레임 각각에 대하여 HSV(Hue, Saturation, Value) 칼라 히스토그램값을 구한 후, 256 coefficient 의 SCD(Scalable Color Descriptor)를 계산한다[5]. 즉, 픽셀마다 HSV 값을 구한 후, 256(H:16, S:4, V:4) 경우로 할당하여 하나의 키프레임에 대하여 SCD 값들을 계산한다. 그리고 256 의 경우로 할당된 픽셀의 수를 전체 픽셀수로 나누어 정규화를 수행한다. 그리고 60 개의 키프레임의 평균을 GoF 로 정의한다.

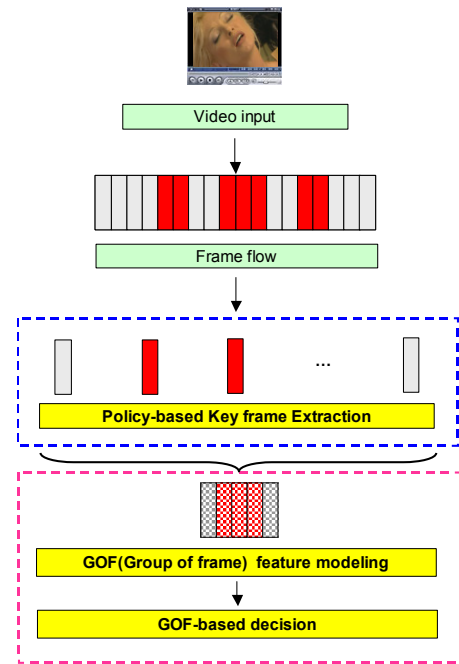
SVM 학습은 동영상별로 생성된 GoF 파일을 입력 데이터로 사용하여, 유무해 학습모델 생성을 위한 최적화 과정을 수행하게 된다.

3.3. 유해 동영상 분류

3.2 에서 생성한 학습모델을 사용하여, 입력 동영상을 대상으로 분류 실험을 수행한다. (그림 3)은 이를 도시한 것으로서 SVM 학습기능을 제외하고는 3.2 의 학습 모델 생성 과정과 유사하다.

입력되는 실험 동영상에 대하여 키 프레임을 추출하여 GoF 를 생성한다. 키 프레임의 추출간격 및 개수는 학습모델 과정에서 추출한 것과 동일할 필요는 없다. 추출개수가 많으면 분류의 정확성은 높을 수 있

나, 소요시간이 더 걸릴 수 있다.



(그림 3) 유해 동영상 분류 과정

4. 실험 결과

본 절에서는 지금까지 살펴본 유해 동영상 분류를 위한 학습 모델 생성과 분류에 대한 실험결과를 보여 준다. 학습모델 생성과 분류를 위하여 SVM^{light} 를 이용하였으며, SVM 커널은 RBF(Radial Basis Function)를 사용하여 최적화를 수행하였다[6]. 학습 모델 생성을 위하여 <표 1>에서 유해 동영상 200 건, 무해 동영상 200 건을 사용하였다. 무해 동영상의 경우 영화, 다큐멘터리, 뮤직 비디오, 스포츠 분야에 대해 각각 50 건씩을 사용하였다. 학습 모델은 (그림 2)의 과정을 수행하여 SVM 학습 모델을 생성하였다. <표 2>는 학습 모델 최적화에 사용된 데이터의 학습 실험 결과이다.

- a : 유해를 유해로 분류한 개수,
- b : 유해를 무해로 분류한 개수,
- c : 무해를 유해로 분류한 개수,
- d : 무해를 무해로 분류한 개수라고 할 때,

$$Precision = a / (a+c),$$

$$Recall = a / (a+b),$$

$$Accuracy = (a+d) / (a+b+c+d) \text{ 라고 정의한다.}$$

<표 2> 분류 모델 학습 실험결과

Precision	Recall	Accuracy
97.99%	97.50%	97.75%

그리고, 분류 실험을 수행하기 위하여 사용한 데이터는 학습 모델을 생성에 사용된 데이터와는 다른 데

이터로서 <표 1>에서 유무해 동영상 각각 200 개씩을 사용하였다. 무해 동영상의 세부 장르는 학습모델 생성에 사용된 것과 동일하며, 동영상별 키프레임을 60 개씩 추출하였다. <표 3>은 실험 데이터에 대한 분류 실험결과를 나타낸 것이다.

<표 3> 유해 동영상 분류 실험결과

<i>Precision</i>	<i>Recall</i>	<i>Accuracy</i>
93.91%	92.50%	93.25%

5. 결론

지금까지 유해 동영상에 대한 자동 분류를 위하여 GoF 기반의 특징 추출과 학습모델 생성 및 분류 과정에 대하여 살펴보았다. 본 논문에서 제안한 분류 기술은 기존의 일반적인 동영상의 분류 방법과는 달리, 유해 동영상 분류에 적합한 신속하고 정확도 높은 동영상의 내용 분석 방법이라고 할 수 있다.

제안한 분류 방법은 휴먼 디텍션에 성능이 우수한 HSV 스펙트럼을 이용하여 그룹프레임의 특징을 활용하였다. 이는 유해 동영상이 다른 장르에 비하여 피부색 노출이 많고 촬영 기법의 유사성 등으로 인하여 높은 분류 성능을 보였다.

그러나, 유해 동영상 분류 기술이 직접 상용화 제품에 적용되기 위해서는 현재의 성능보다 높고, 신속한 처리 기술이 요구된다.

이를 해결하기 위하여, 실험 결과에서 약 7% 내외의 에러율의 원인을 분석함으로써 본 논문에서 제시한 분류 모델의 문제점을 보완 할 수 있을 것으로 보인다. 그리고, 유해 동영상 판단의 기준이 되는 유해 이미지 기반의 분류 모델을 개발하여 활용하는 방안과, 유해 동영상 분류에 음성 정보 등 멀티모달 특징을 활용하는 방안을 고려해 볼 수 있다.

향후, 이와 같이 다양한 특징과 깊이 있는 분석을 수행하여 신속하고 정확한 유해 동영상에 대한 분류 연구를 진행할 계획이다. 이는 인터넷 뿐만 아니라 다양한 매체를 통하여 급속도로 번지고 있는 유해 동영상 차단 분야에 적극 활용될 수 있을 것으로 기대된다.

참고문헌

- [1] M.J. Roach, J.S.D. Mason, and M. Pawlewski, "Video genre classification using dynamics," Int. Conf. on Acoustics, Speech and Signal Processing, 2001.
- [2] Ba Tu Truong; Dorai, C., "Automatic genre identification for content-based video categorization", Pattern Recognition, 2000. Proceedings. 15th Int. Conf. on Volume 4, 3-7 Sept. 2000 Page(s):230 - 233 vol.4
- [3] Ye Yuan, Jun-Yi Shen, Qin-Bao Song, "A new rule-based video classification approach", Proceedings of the 2nd Int. Conf. on Machine Learning and Cybernetics, 2003.
- [4] Nello Cristianini, John Shawe-Taylor, "An Introduction to Support Vector Machines," Cambridge University Press

2000.

[5] B.S. Manjunath, "Introduction to MPEG-7", John Wiley & Sons, LTD.

[6] Thorsten Joachims, "SVM^{light} Support Vector Machine", <http://svmlight.joachims.org/>