

MPEG-7 시각서술자와 Multi-Class SVM을 이용한 불법 및 유해 멀티미디어 분석 시스템 구현

*최병철, *김정녀, **류재철
 *한국전자통신연구원 정보보호연구본부, **충남대학교 컴퓨터공학과
 e-mail : *(corea, jnkim)@etri.re.kr, **jcryou@home.cnu.ac.kr

Implementation of Illegal and Objectionable Multimedia Retrieval Using the MPEG-7 Visual Descriptor and Multi-Class SVM

*Byeongcheol Choi, *Jungnyeo Kim, and **Jeacheol Ryou
 *Information Security Research Division, ETRI, Korea
 **Division of Electrical and Computer Engineering, Chungnam National University, Korea

Abstract

We developed a XMAS (X Multimedia Analysis System) for analyzing the illegal and objectionable multimedia in Internet environment based on Web2.0. XMAS uses the MPEG-7 visual descriptor and multi-class SVM (support vector machine) and its performance (accuracy on precision) is about 91.6% for objectionable multimedia analysis and 99.9% for illegal movie retrieval.

I. 서론

최근 웹 포털에서는 음란성 유해 멀티미디어(이미지/동영상)으로 많은 사고가 발생하고 있으며, P2P 및 웹디스크 기반의 공유사이트를 통해서 불법 동영상들이 많이 유포되고 있는 현실이다. 이러한 상황에서 불법 및 음란성 유해 멀티미디어에 대한 분석 및 차단 기술 개발이 시급한 현실이다. 이 중에서 음란성 유해 이미지 분석의 경우 J.Z.Wang[3]에 의해 제안된 WIPE(Wavelet Image Pornography) 방법을 이용한 유해 이미지 분석 방식 있으며, S.W.Han[4]에 의해 제안된 MPEG-7 시각서술자 및 SVM을 이용한 방식 등 다양한 방법으로 이 분야에 대한 연구가 시도되고 있으며,

본 연구에서는 S.W.Han[4]에서 기반 기술로 채용한 MPEG-7 시각서술자[1]와 SVM[2]을 이용하지만, 음란성 멀티미디어 분석 뿐만 아니라 멀티미디어 지재권 보호를 위해 Multi-Class SVM을 이용한 Multi-Model 기반의 분석 방법을 추가하여 다양한 응용 분야에 사용할 수 있도록 설계하여 개발하였다.

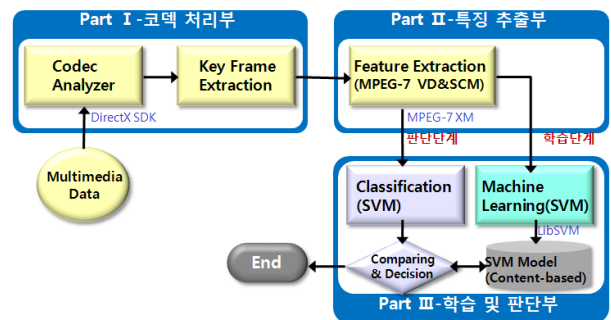


그림 1. XMAS 핵심 엔진 구조

II. 본론

본 연구에서 개발한 XMAS의 핵심 엔진은 그림 1과 같이 코덱 처리부, 특징 추출부, 학습 및 판단부로 구성되어 있다. 코덱 처리부에서는 코덱 분석 및 키프레임 추출을 수행하며, 특징 추출부에서는 이미지 단위로 특징값을 추출하게 된다. 마지막으로 학습 및 판단

부에서는 이미지 학습 모델을 생성하고 판단 단계에서 해당 모델을 사용하여 음란성에 따른 유/무해 및 해당 동영상 여부 판단 등의 기능을 수행할 수 있게 된다.

본 연구에서는 k(class 개수)가 2와 10인 MPEG-7 시각서술자 CLD(Color Layout Descriptor)와 CSD(Color Structure Descriptor) 특징값을 사용한 SVM 모델을 사용하였다. 2-Class SVM을 이용하여 음란성 기반의 멀티미디어 유무해 판단 기술에 사용하였으며, Multi-Model & Multi-Class SVM을 이용하여 불법 동영상 검색 기술에 활용하였다. SVM의 경우 RBF 커널 함수를 사용하였으며, C(cost or error)와 γ (gamma)의 최적화 값을 구하여 사용하였다. 본 연구에서 LIBSVM에서 제공하는 grid.py를 이용하여 C는 32, γ 는 0.125를 산출하여 사용하였다.[2]

표 1. XMAS 시스템 개발 사양

항 목	사양
시스템	IBM Z-pro (Window XP)
컴파일러	Visual C++ (Visual Studio 2005)
라이브러리	CxImage Lib.
	DirectX SDK for DirectShow
	Window Platform SDK
	LibSVM Lib.
	MPEG-7 XM Lib
Encoder	Umile Encoder (변형 공격용)

표 2. XMAS 시스템 성능

기능 구분	모델크기 (MByte)	정확도 (%)	속도 (s/f)
음란성 멀티미디어 유/무해 판단 (Single-Model using 2-Class SVM)	27 ~ 30 (약 40,000장)	약91.6 (이미지) 약 91.9 (동영상)	약0.078 (이미지) 약0.088 (동영상)
불법 동영상 검색 (Multi-Model using Multi-Class SVM)	모델별 8 ~ 10 (5000장)	약99.9 (93.5% 신뢰도)	

III. 구현

XMAS 구현에 사용된 시스템 및 개발 도구는 표 1과 같이 구성되어 있으며, 동영상에서 키프레임(이미지)를 추출, 이미지 특징추출, 동영상 특징추출, 이미지/동영상 학습, 이미지/동영상 유/무해 판단, 불법 동영상 판단 등의 기능을 수행할 수 있도록 GUI 및 Batch 프로그램으로 구현되어 있다.

XMAS 시스템의 성능은 표 2와 같이 음란성 멀티미

디어 유/무해 판단의 경우 2-Class SVM 학습 모델을 사용하였으며, 모델 크기는 약 28MBytes (유해 약2만장, 무해 약2만장), 이미지 유/무해 판단 정확도가 약 91.6% (유해 약89.9%, 무해 약93.3%), 동영상 유/무해 판단 정확도가 약 91.9%, 처리 속도는 이미지의 경우 평균 약 0.078 sec/frame이고 동영상의 경우 키프레임 추출 시간 평균 약0.01초를 포함해서 평균 약 0.088 sec/frame으로 나왔다. 동영상 유/무해 판단의 경우 M of N 확률값으로도 판단 가능하고, 동영상 자체의 200 프레임 단위의 이미지 유/무해 판단 확률값으로 모델링한 2-Class SVM 학습모델을 사용할 수도 있다. 불법 동영상 검색의 경우 Multi-Class SVM 학습 모델 (여기에서는 10개의 클래스를 사용, 즉 동영상 10개 단위로 학습 모델 생성)을 사용하였으며, 모델 크기는 약 9MBytes (각 클래스별 500장 사용해서 총 5000장 단위로 모델 생성), 불법 동영상 판단 정확도는 약 99.9% (개별 프레임 판단 정확도 93.5%의 신뢰도를 가지고 임계값 90%를 적용하여 판단함)로 나왔다.

IV. 결론 및 향후 연구 방향

MPEG-7 시각서술자와 SVM은 멀티미디어 분석 및 판단 기술로 활용되고 있다. 본 연구에서는 최근 불법 및 음란성 멀티미디어에 의한 여러 문제점을 해결하기 위해서 상기의 기술을 이용하여 XMAS라는 시스템을 개발하게 되었으며, 이 시스템은 지재권 모니터링 업체를 위한 불법 동영상 검색 솔루션, 안전한 웹기반 멀티미디어 서비스를 위한 음란성 멀티미디어(이미지, 동영상)의 분석 및 차단 솔루션, PC용 유해정보 차단 및 분석 솔루션 등으로 활용할 수 있는 기술이다. 향후 다양한 특징 추출 기술 개발 및 Multi-Class SVM 학습 모델 엔진의 정제 등의 정확도를 높이는 것에 대한 연구를 계속할 것이다.

참고문헌

- [1] B.S. Manjunath, P. Salembier and T. Sikora, Introduction to MPEG-7 : Multimedia Content Description Interface, John Wiley & Sons Ltd., 2002.
- [2] C.W. Hsu, C.C. Chang and C.J. Lin, A Practical Guide to Support Vector Classification, <http://www.csie.ntu.edu.tw/~cjlin>, Dec. 20, 2007
- [3] J. Z. Wang, G. Wiederhold and O. Firschein, "System for Screening Objectionable Images," Computer Communications, Vol.21, 1998.
- [4] S.W. Han, C.Y.Jeong and T.Y. Nam, "Multi-layer Objectionable Video Classification System Using Local-Global Information", WSEAS CSCC2005, July 2005