

영상기기의 프로파일 분석 기반 패턴추적에 의한 비디오 프레임의 위변조탐지

심재연, 천인혁, 김성환
서울시립대학교 컴퓨터과학부
e-mail : simpo@uos.ac.kr, swkim7@uos.ac.kr

Automatic Detection of Forgery in Video Frames using Analysis of Imaging Device Profile based Pattern Trace

Jae-Youen Shim, In-Hyuk Chon, and Seong-Wan Kim
Dept. of Computer Science, University of Seoul

Abstract

본 논문은 HD (High-definition) video, SD (Standard-definition) video, low quality video, handset video, 4 가지 Imaging Device 에 대한 프로파일 분석을 통해 비디오 프레임 상에 나타나는 위 변조를 검사하는 방법을 제안한다. High-definition video, Standard-definition video, low quality video, handset video 에 대한 분석을 하고 각 영상의 특이 점을 파악 하여 분류한 클래스에 대한 프로파일검사를 통해 EM Algorithm 을 이용하여 영상의 위 변조를 검사 하고 영상의 신뢰성을 높인다.

1. 서론

현대는 정보통신 기술의 발달로 손쉽게 휴대폰과 Smartphone 과 같은 다양한 휴대기기를 통해서도 손쉽게 디지털 비디오 영상 찍히고 촬영 할 수 있다. 또한 Photoshop 과 Premier 와 같은 소프트웨어를 통해 이미지 및 비디오를 손쉽게 조작 할 수 있다. 비디오 영상의 위 변조는 영상에 대한 증거로서의 신뢰성을 떨어트리며 이러한 위 변조 영상에 대한 검사는 법의학 적으로도 큰 관심사 중에 하나이다. 위 변조의 판독을 하는 기술로는 Watermark 와 Fingerprint 와 같은 암호화된 신호를 영상에 삽입하여 Watermark 와 Fingerprint 의 손상을 확인하여 영상의 위 변조 상태를 확인 하는 방법 [1]과 CMOS 와 CCD 와 같은 디지털 카메라 이미지 센서의 고유 특성과 같이 영상을 구성하는 고유의 신호를 파악하여 영상의 위 변조를 파악 하는 방법이 있다. Watermarking 과 Fingerprinting 과 같은 후처리 기술을 영상에 Noise 를 생성 할 수 있어 영상의 질을 떨어트리며 Collusion 과 같은 공격을 받았을 시에 신뢰도가 떨어지는 단점이 있다. 영상 고유의 성질을 파악하여 영상에 대한 위 변조를 검사하는 방법은 영상의 직접적인 성질을 파악하고 영상의 상태에 따른 정보와 특성을 검사하므로 위 변조를 파악한다. 영상의 고유의 데이터로는 이미지 센서의 CFA (Color Filter Array)의 Interpolation 에 따른 특성을 파악 하는 방법[2]과 이미지 센서의 특성을 반영하는 PRUN (Photo Response Non-Uniformity)와 같은 정보를 이용하는 방법[3]이 있다.

본 논문에서는 HD (High-definition) video, SD (Standard-definition) video, low quality video, handset video, 4 가지 Imaging Device 에 대한 프로파일 분석을 통해 비디오 프레임 상에 나타나는 위 변조를 검사하는 방법을 제안한다. 본 논문의 2 장에서는 각 영상에 대한 정보를 소개하고 분석 하녀, 3 장에서는 프로파일 분석을 통한 영상의 찾는 기법에 대하여 설명 하며, 4 장에서는 위 변조된 영상에 대한 검증을 실험과 결과 제시하고 5 장에서는 제안 된 기법의 개선 사항에 대한 토의로 결론을 맺는다.

2. 관련연구

본 장에서는 HD video, SD video, Low quality video, Handset video 로 영상에 대하여 클래스를 나누고 각각의 클래스에 대한 소개를 한다.

2.1. HD (High-definition) video

HD video 는 HDTV 를 위한 비디오 Format 을 지칭한다. 720p 와 1080i 이상의 대역을 커버 할 수 있는 고 화질의 영상들이 이 범주에 들어가며 720p 는 720 선, Progressive, 30 fps 를 적용 하며 1080i 는 1080 선, Interlaced, 30 fps 또는 초당 60 필드를 적용 한다. 기존의 영상과의 차이점은 종횡비가 16:9 로 기존의 4:3 의 종횡비보다 비해 가로길이가 1/3 정도 더 길다. 또한 Progressive 방식을 사용하므로 기존의 Interlaced 방식이 움직이는 객체에 대하여 톱니모양의 경계선이 발생하고 수평 경계에서 깜빡임을

발생시키는 것을 방지 하였다. 디지털 영상의 표준 양식으로는 표 1 과 같다[4].

<표 1> HD Video 의 표준 형식분석

Number of active per line	Number of active line	Aspect ratio	Picture rate
1,920	1,080	16:9	60I 30P 24P
1,280	720	16:9	60P 30P 24P

2.2. SD (Standard-definition) video

SD Video 는 SDTV 를 상영하기 위한 비디오 영상의 Format 지칭한다. SD 비디오 방식에 는 NTSC 비디오와 PAL 비디오 방식이 있다. NTSC 방식은 주로 북미와 일본에서 사용되며 4:3 종횡비, 525 스캔선수, 29.97fps 를 가진다. YIQ 컬러 모델을 사용하며 총 60MHz 의 대역폭을 가진다. NTSC 비디오에서 Blank 정보는 매 필드의 시작부분에 제어정보를 위해 배정된 20 개의 선에 놓이게 되어 각 프레임당 비디오 선수는 485 개가 된다. 종종 화소는 주사선 사이에 놓이게 되는데, 이 때문에 Interlaced 가 아닌 경우에도 340 개의 건만을 보여주게 된다. 485 개의 활성 선의 약 70%에 해당하는 수이며 Interlaced 의 경우 50%이하가 될 수도 있다. PAL 비디오 방식은 독일에서 고안된 방식으로 프레임당 625 개의 주사선, 초당 36 프레임, 4:3 의 종횡비, Interlaced 방식을 기본 구조로 한다. PAL 방식은 8MHz 채널을 가지는 YUV 컬러 모델을 사용하며 Y에 5.5MHz, U와 V에 각각 1.8MHz 의 대역폭을 할당한다. 표 2는 SD 비디오 Format 에 대한 시스템 비교표이다[4].

<표 2> SD Video 의 표준 형식분석

System	Frame rate(fps)	Number of scan lines	Total channel width	Bandwidth allocation		
				Y	I,U	Q,V
NTSC	29.97	525	6.0	4.2	1.6	0.6
PAL	25	625	8.0	5.5	1.8	1.8

2.3. Low quality video

본 논문에서 정의 하는 Low quality video 란, 휴대용 Webcam 을 통해 촬영된 저 사양 영상을 지칭하며, 인터넷 화상통화 및 인터넷 방송 등에 사용 되고 있으며, YouTube 와 같은 동영상공유 웹사이트를 통해 상영 될 수 있다. YouTube 의 비디오 codec 은 H.263 사용시 Maximum standard quality 로 320 X 240 을 지원 하고 H.264 사용시 480 X 360 을 지원 한다. 또한 보안 및 감시를 위해서도 Webcam 은 사용된다. 초기의 Webcam 은 320 x 240 를 제공 했으며 16bit grayscale 촬영 시 초당 60frames 를 제공하고, 256bit color 촬영 시 초당 15frames 를 제공

했다. 현재 대부분의 Webcam 은 320 x 240 의 영상을 초당 30 frames 에 제공하나 일부 제품들은 초당 120frames 의 영상을 제공 하기도 한다.

2.4. Handset video

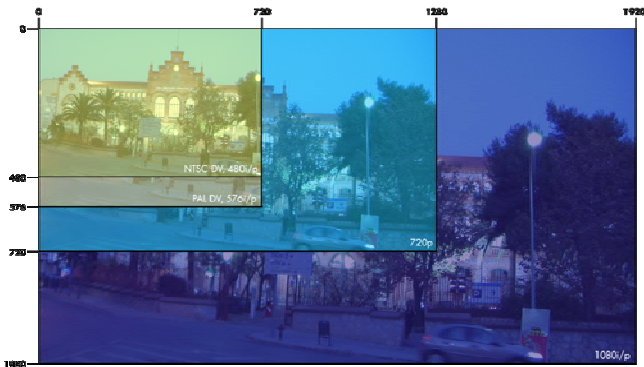
본 논문에서 지칭하는 Handset video 는 휴대용단말기와 같은 휴대 기기를 통해 촬영된 영상을 지칭한다. 휴대폰과 같은 기기에 부착되어 있는 다양한 카메라 장치를 통해 손쉽게 영상을 얻을 수 있다. 본 논문의 Handset video format 은 3gp 파일 바탕으로 한 MPEG4 형의 format 을 사용하는 k3g 파일로 선택했으며 이는 영상을 176 x 144(QCIF)로 설정하고 초당 15 frames 로 하는 것을 원칙으로 한다. 휴대폰 카메라의 이미지 센서로 대부분 CMOS 방식을 사용한다. 일반 디지털 카메라에 사용되는 CCD 방식에 비해 Noise 가 많지만 공정이 간단하고 가격이 저렴하며 전력 소모가 작다. CMOS 가 휴대폰 카메라에 많이 사용되는 것은 가격적인 측면도 있지만 전력 소모가 적다는 장점이 크게 작용한다. 본 논문에서 사용되는 Handset video 는 최신 Smartphone 의 고화질의 동영상이 아닌 일반적인 휴대폰에서 사용되는 CMOS 를 이용하여 만들어진 영상을 사용하였다. 다음 표 3 은 CCD 와 CMOS 방식의 장단점을 비교하였다.

<표 3> CCD & CMOS 이미지센서 비교

구분	CCD	CMOS
장점	좋은 조명감도, 풍부한 색상, 선명한 이미지	간단한 생산과정, 저렴한 가격, 적은 전력소모 (CCD 대비 약 1/100)
단점	높은 생산비용, 복잡한 제조과정	낮은 조명감도, 높은 Noise 율, 상대적으로 큰 센서 칩의 크기

3. 프로 파일을 통한 영상분석

HD video, SD video, Low quality video, Handset video 로 정의된 영상의 클래스에 대한 프로파일 링을 통해 각 영상의 특징을 찾아 그에 대한 분석이 필요하다. 영상을 분류하는 보편적인 값으로는 주사 방식과 영상 크기의 차이를 들 수 있다. HD video 은 720p 와 1080i/p 를 사용 하며 SD video 에서는 NTSC 비디오 방식을 사용 할 때 480i/p 를 PAL 방식을 사용할 때 576i/p 를 사용 한다. Low quality video 에 대해서는 240p 혹은 288p 를 지원 하며 Handset video 에 대해서는 특별히 정의 된 기준은 없지만 본 논문에서는 176 x 144(QCIF)를 기반으로 한다. 그림 1 은 SD 비디오와 HD 비디오에 대한 디지털 비디오 해상도를 나타내고 있다. HD 영상과 SD 영상의 차이를 시각적으로 확인 해 볼 수 있다.



(그림 1) 영상에 대한 해상도

각각의 영상 클래스의 특징을 찾기 위하여 영상을 EM Algorithm 을 통해 분석 한다. EM Algorithm 을 통해 영상이 생성된 시점에서의 해상도를 추정하여 비디오 Frame 에 영상이 위 변조 되었을 시에 그 영상에 대해 프로파일링 된 정보를 토대로 영상의 위 변조를 판단 한다.

본 논문에서는 프로파일 인식을 위해 EM 알고리즘을 사용하였다. EM 알고리즘은 두 단계의 반복 알고리즘으로 구성되어 있다. 먼저 첫 번째 E 단계는 샘플이 어느 가우시안에 속하는지 추정 하는 단계로 소속 정도를 확률로 표현하는 연성 소속을 이용한다. E 단계의 소속 값을 추정한 후 매개변수 집합을 Θ 를 추정 하는 단계를 M 단계라고 한다. Θ 에는 성격이 다른 두 가지 매개 변수가 있는데 K 개의 요소와 혼합 계수 벡터를 계산한다. 이들 E 와 M 단계를 번갈아 가며 반복하다가 수렴 조건이 만족되면 연산을 중지한다. EM Algorithm 은 다음과 같이 표현 될 수 있다 [5].

Algorithm: EM Algorithm

1. 매개 변수집합 Θ 를 초기화 한다.
2. repeat{
3. E 단계 : Θ 를 이용하여 샘플 별로 K 개의 가우시안에 속 할 확률을 추정 한다.
4. M 단계 : E 단계에서 구한 소속 확률을 이용하여 Θ 를 추정 한다.
5. } until (멈춤 조건이 만족);

EM Algorithm 을 통해 SD video, HD video, Low quality video, Handset video, 각 영상의 클래스에 대한 프로 파일을 얻어 위 변조된 비디오 프레임을 확인 할 수 있다.

4. 실험

실험에 쓰인 영상은 HD Video 는 Sony HVR-Z1E 을 SD Video 는 Canon IXY70 을 통해 저장 된 것이며 Low quality video 는 Logitech QuickCam® Pro 5000 를 Handset video 로는 EVER W300 을 통해 촬영된 영상을 사용 하였다. 실험 환경으로는 Intel Xeon CPU E5506, 8GB RAM, Window7 64bit 의 사양에 MATLAB R2009 를 통한 이미지 분석을 실행 하였다.

4.1. Sample video frame image

영상의 클래스에 대한 video frame 샘플을 설정 샘플을 통해 얻어진 프로파일을 통하여 합성 여부를 판단 한다. 실험에 쓰인 영상은 그림 2 와 같다.

HD Video	SD Video
Low quality video	Handset video

(그림 2) 영상 클래스에 대한 샘플영상

4.2. 위 변조 검색

영상에 대한 위 변조를 하기 위하여 각각의 video frame 에 다른 영상의 video frame 를 삽입 하여 확인 하였다. 실험을 통해서 프로파일 된 영상의 정보를 토대로 다른 클래스의 영상간의 특성 패턴을 추적 하여 영상간의 위 변조가 일어 났을 때 위변조에 대한 확인을 할 수 있었다. 그림 3 은 검사에 사용된 HD 영상에 SD 영상을 삽입 한 영상이다.



(그림 3) HD 영상에 SD 영상을 삽입 한 영상

본 논문에서 제안 하는 프로 파일분석을 통하여 각 영상에 EM Algorithm 을 적용 하여 생성되는 값에 FFT (Fast Fourier transform)을 수행 하였을 시 생성되는 최대 피크 값을 분석 하였다. 동일 한 크기의 블록 마스크를 각 영상에 대한 적용을 하였을 시 각 영상에 따른 최대 피크 값의 분포가 클래스에 따라 다른 것을 확인 하였다. 실험을 통해 256 x 256 크기의 블록 마스크로 영상을 분석 하였을 시 HD Video 클래스는 26000 이상의 피크 값을 SD Video 는 15000 이하의 최대 피크 값을 나타내며 Low quality

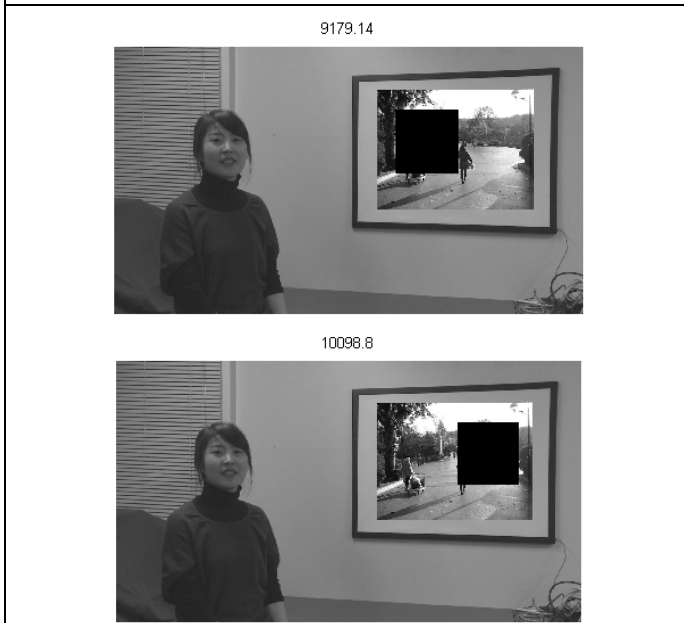
video 는 20000 에서 25000 사이의 값을 가지게 되는 것을 확인 하였다. Handset video 영상에 대해서는 영상의 크기로 인해 256 x 256 상에서의 최대 피크 값을 찾지는 못하였지만 64 x 64 블록 마스크의 크기로 SD Video 와 Handset video 영상에 대한 차이를 확인 해본 결과 최대 피크 값이 SD Video 는 800 을 넘지 못하는 것에 비해 Handset video 는 1300 에서 3300 의 분포를 보였다. 이는 각 영상 클래스에 대한 픽셀 간의 밀집도와 상관 관계가 다르기 때문에 이러한 성질을 보인다. 그림 4 는 그림 3 의 영상에 대하여 블록 마스크를 통한 최대 피크 값을 보여준다. 그림 위에 나타난 숫자를 블록의 최대 피크 값을 표현 한다.

그림 5 는 이러한 영상의 특성을 이용하여 자동으로 위변조 영역을 지정 해 내는 시스템이다. 영상이 경계면에 대해서는 다른 영상 클래스의 간섭을 받아 영상이 확연하게 검출 되지 못 함을 보인다.

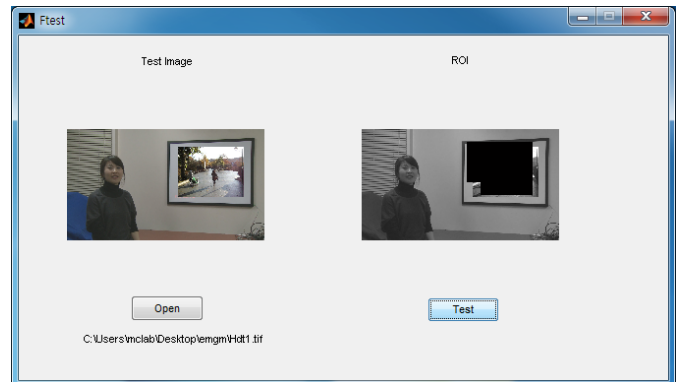
원 영상 부분



위 변조 부분



(그림 4) 위변조 영상 분석



(그림 3) AUTO 위변조 위치 검출 시스템

5. 결론

본 논문은 HD video, SD video, low quality video, handset video 와 같은 4 가지 Imaging Device 에 대한 프로파일 분석을 통해 비디오 프레임 상에 나타나는 위 변조를 검사하였다. 각 클래스에 대한 프로파일을 통해 영상의 특성을 확인 하고 이를 토대로 영상의 위 변조 상황에 대하여 분석 하였다. 이 분석 기법을 통하여 영상의 Imaging Device 의 상태에 따른 영상의 위 변조를 확인 할 수 있었으며 앞으로 압축 강도를 높였을 때의 위 변조 상황에 대한 연구를 할 것이다.

Acknowledgement

본 연구는 서울시정개발연구원 2009 년도 신기술연구 개발사업 과제 <IP-TV 등 고화질 방송 서비스의 콘텐츠 보호를 위한 포렌식 워터마킹 기술 개발> 의 2 차년도 연구결과로 수행되었음

REFERENCES

- [1] S. Katzenbeisser and F. Petitcolas, "Information Techniques for Steganography and Digital Watermarking". Norwell, MA: Artech House, 2000.
- [2] Popescu A.C. and Farid H. " Exposing Digital Forgeries in Color Filter Array Interpolated Images" , IEEE Transactions on Signal Processing, vol. 53(10), pp. 3948- 3959, 2005.
- [3] Jan Lukáš , Jessica Fridrich, and Miroslav Goljan, " Detecting digital image forgeries using sensor pattern noise Proceedings of SPIE. Vol. SPIE-6072, pp. 362-372. 2006.
- [4] Ze-Nian Li, Mark S. Drew, " Fundamentals of Multimedia" . Prentice-Hall, 2004
- [5] 오일석, " 패턴인식" , 교보문고, 2008.