

논문 2010-47SP-5-16

# 블록 비교를 이용한 내용기반 동영상 복사 검색용 영상 지문

## (Image Fingerprint for Contents based Video Copy Detection Using Block Comparison)

나 상 일\*, 진 주 경\*\*, 조 주 희\*\*, 오 원 근\*\*\*, 정 동 석\*\*\*\*

(Sang-il Na, Ju-Kyoung Jin, Ju-Hee Cho, Weon-Geun Oh, and Dong-Seok Jeong)

### 요 약

동영상 복사 검출을 위해서 사용되는 내용기반 정보는 크게 공간적인 정보와 시간적인 정보로 나눌 수 있다. 이 중에 공간적인 정보는 내용기반 영상 지문(fingerprint)을 의미한다. 이러한 영상 지문이 갖추어야 할 바람직한 특성으로는 추출이 간단해야 하며, 서로 다른 영상에서는 어느 정도 거리를 갖는 독립성(independence)을 가져야 하되, 변형된 영상과 원본 영상 사이에서는 짧은 거리가 나와야 하는 강인성을 가져야 하며 또한 정합 속도가 빨라야 한다. 본 논문에서는 내용기반 동영상 복사 검출을 위해 공간 정보를 사용한 새로운 영상 지문을 제안한다. 제안하는 방법에서는 블록 영역의 평균과 1차, 2차 미분 값을 사용하는데 이러한 방법은 특징 추출 시 덧셈과 뺄셈만을 사용하므로 빠른 추출이 가능하고, 강인성과 독립성을 모두 가진다. 또한 사용된 특징 비교를 통해 2진화하고, 거친 지문과 미세 지문의 단계적 처리 방법을 사용하여 빠른 정합이 가능하다. 본 논문에서는 MPEG-7의 VCE7실험 조건을 사용하였으며 다수의 변형된 영상을 이용하여 제안된 방법의 성능을 검증 하였다.

### Abstract

Two types of informations are used for content-based video copy detection: spatial information and temporal information. The spatial information means content-based image fingerprint. This image fingerprint must have following characteristic. First, Extraction is simple. Second, pairwise independence for random selected two images. At last, Robust for modifications. This paper proposed image fingerprint method for contents based video copy detection. Proposed method's extraction speed is fast because this method's using block average, first order differentiation and second order differentiation that can be calculated add and minus operation. And it has pairwise independence and robust against modifications. Also, proposed method feature makes binary by comparisons and using coarse to fine structure, so it's matching speed is fast. Proposed method is verified by modified image that modified by VCE7's experimental conditions in MPEG7.

**Keywords:** video copy detection, image fingerprint, content based copy detections

### I. 서 론

인터넷 서비스와 디지털 기기의 발달은 각종 영화 과

일 및 교육용 동영상 같은 디지털 멀티미디어 콘텐츠의 제작과 유통을 활발하게 만들었다. 또한 출판 및 교육, 영화 및 방송에 대한 콘텐츠가 인터넷을 통해 제공되고 있다. 이러한 디지털 콘텐츠는 완전 복제가 가능하고 인터넷을 통한 다량배포도 쉽다는 특징을 가지고 있다. 이런 문제점은 콘텐츠 저작권자와 서비스 업체에 재정적 손해를 끼쳐 산업 활성화에 걸림돌이 된다.

현재 이러한 불법복제물을 여과하고 감독하기 위하여 두 가지 기술이 연구되고 있는데 바로 워터마킹(watermarking)과 내용기반 복사검출(Contents based copy detection) 기술이다. 워터마킹은 정보를 콘텐츠에

\*\* 학생회원, \*\*\*\* 정회원, 인하대학교 전자공학과 (Dept. of Electronic Engineering, Inha University)  
 \* 학생회원, \*\*\* 정회원, 한국전자통신연구원 콘텐츠연구본부 (ETRI)  
 ※ 본 연구는 지식경제부, 문화체육관광부 및 한국산업 기술평가관리원의 IT산업원천기술개발사업의 일환으로 수행하였음. [Rich UCC 기술 개발, 2010]  
 접수일자: 2009년12월21일, 수정완료일: 2010년5월19일

직접 삽입하는 기술이다. 이 정보를 복제된 콘텐츠로부터 추출하여 저작권을 증명한다. 내용기반 복사 검출 방법은 콘텐츠에서 직접 정보를 추출하여 확인하는 방식으로 콘텐츠에 변형을 가하지 않고 유통되는 콘텐츠를 확인한다. 워터마킹에 비해 내용기반 복사검출 방법이 갖는 장점은 특정 추출을 콘텐츠가 유통 된 후에도 할 수 있어 서비스 제공 사업자가 불법 유통 콘텐츠를 사전에 차단할 수 있다는 것이다.

내용기반 동영상 복사 검출 기법은 크게 두 가지로 나누어진다. 시간적 정보를 사용하는 기술과 공간적 정보를 사용하는 기술이다.

시간적 정보를 사용하는 기술은 동영상 프레임간의 상관관계를 이용하는 방법이다. 그러나 이 기술은 동영상의 시간적 정보에 대한 공격이 가해지면 사용할 수 없다. 후자는 동영상의 프레임에서 정보를 추출해 사용하는 방법이다. 일반적으로 영상은 다른 영상과 구분할 수 있는 고유한 정보를 갖고 있다. 이를 영상 지문이라고 한다.

이러한 내용기반 복사 검출을 위한 영상 지문은 다음 조건을 만족해야 한다<sup>[1~2]</sup>.

강인성 : 콘텐츠에 의도적/비의도적인 변형이 가해졌을 때도 변형된 콘텐츠의 지문은 원본의 지문과 매우 유사해야 한다.

독립성 : 두 개의 서로 다른 영상은 서로 다른 영상 지문을 가져야 한다. 즉, 서로 다른 영상의 지문을 정합하였을 때는 어느 정도 이상의 거리가 나와야 한다.

고속 정합 : 효율적인 검색을 위하여 매우 빠른 속도의 영상 지문정합 알고리즘을 적용할 수 있어야 한다.

본 논문에서는 동영상 복사본 검출을 위하여 위 조건을 만족하는 새로운 영상 지문을 제안 한다. 제안하는 영상 지문에서는 블록들의 평균과 블록 간의 1차 2차 차분 값을 특징으로 사용한다.

## II. 기존 연구 동향

동영상 지문에 대한 다양한 알고리즘들은 최근까지도 활발히 연구되고 있다. 기 발표된 알고리즘의 대부분은 프레임에서 정보를 추출하여 정합을 하고 대부분의 경우 화소 값 기반의 특징을 사용 한다<sup>[2~8]</sup>.

Cheung과 Zakhor는 히스토그램 기반의 알고리즘을 제안하였다<sup>[2]</sup>. Kim 과 Vasudev<sup>[3]</sup>, Ostveen et al<sup>[4]</sup> 등은 평균값 기반의 특징을 제안하였다. Kim과 Vasudev는 블록의 평균값을 사용하는 ordinal 방법을 제안하였다<sup>[3]</sup>. Ostveen et al. 은 블록 평균값의 차분을 사용하였다<sup>[4]</sup>. 기존에 제안된 이러한 방법들은 Ordinal의 경우, 한 블록의 평균값이 영향을 받으면 모든 블록의 순위가 바뀌는 문제가 있고, Ostveen 의 방법은 서로 이웃한 블록의 평균값만을 비교하여 값들 사이의 상관관계가 큰 문제를 갖고 있다. Lowe 등은 기울기 기반 특징을 사용하였다<sup>[5~8]</sup>. Lowe는 물체 인식을 위하여 기울기 방향 히스토그램을 사용하였다<sup>[5]</sup>. Hampapur와 Bolle는 dominant 칼라와 gradient 크기를 사용하였다<sup>[6]</sup>. Lee와 Yoo는 정규화 된 gradient의 크기 가중치를 적용한 gradient 중심 방향 방법을 제안 하였다<sup>[7~8]</sup>. 또한 Kim은 영상 지문으로서 DCT(Discrete정합 속도 비교 (단위 ms) Cosine Transform)을 사용한 ordinal 방법을 제안하였다<sup>[9]</sup>. [5~9]에서 제안된 방법들은 영상 지문의 차원이 높고, 자막/로고의 삽입 등에 의한 영상

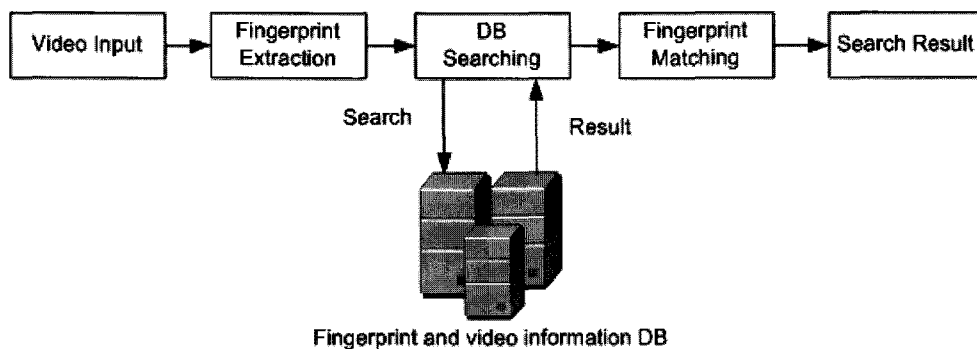


그림 1. 일반적인 동영상 복사 검출 시스템  
 Fig. 1. General video copy detection system.

에지 성분의 변화에 민감한 단점을 갖고 있다.

본 논문에서는 영상의 블록 평균값을 사용한 1차, 2차 미분 값을 응용한 영상 지문 방법을 제안한다.

### III. 블록 비교 기반 영상 지문

그림 1은 일반적인 동영상 복사 검출 시스템을 나타내고 있다. 알려지지 않은 동영상이 입력으로 들어왔을 때 동영상 지문을 추출하여 미리 추출되어 DB에 저장돼 있는 동영상 지문과 비교해서 복사여부를 검출하거나 원하는 정보를 제공해 준다. 전체 시스템의 성능은 많은 요소들에 의해 결정되나, 가장 중요한 요소는 사용된 영상 지문의 성능이다.

본 논문에서 제안하는 동영상 복사검출용 영상 지문은 블록 영역 평균값의 비교와 차분만으로 이루어져 있어 추출이 간단하고, 고속 정합이 가능하다. 이를 위하여 coarse-to-fine 의 2단계 구조를 갖는 영상 지문을 제안한다. 그림 2는 제안된 알고리즘의 사용 시나리오를 보여준다. 각각의 동영상이 입력으로 들어왔을 때, 각각의 동영상을 프레임별로 나눠서 영상 거친 지문과 미세 지문을 추출한다. 정합 과정에서는 우선 두 동영상에서 추출된 거친 지문을 비교하고, 이 값이 기준 문턱치보다 작다면 미세 지문을 비교함으로써 최종 거리를

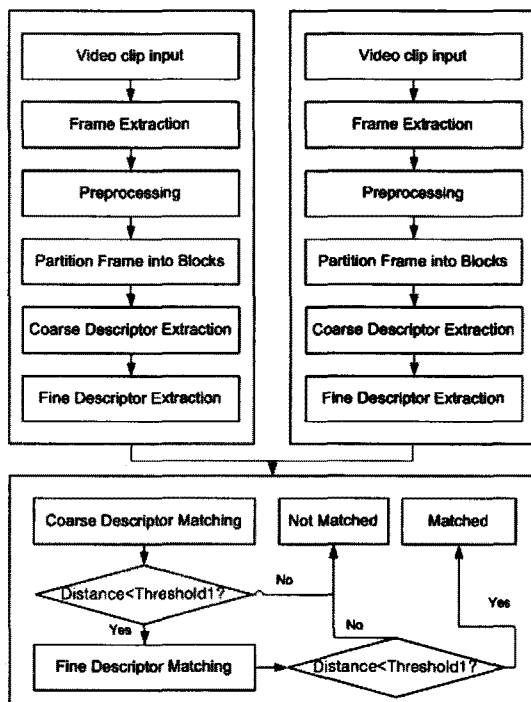


그림 2. 제안된 알고리즘의 사용  
Fig. 2. Use of proposed algorithm.

구하게 된다.

#### 1. 기존 블록 평균 값 기반 알고리즘의 문제

그림 3은 Ordinal 방법에 실제 잡음이 들어간 예를 보이고 있다. 원본영상 (a)는 (b)와 같은 밝기 순위를 갖는데, 중앙 우측 위치에 검은색 자막/로고가 들어감으로, (c)와 같은 형태로 바뀌게 된다. 그림에서 보듯이 하나의 블록에 잡음이 생기면, 그 블록의 순위가 바뀌며 전체 블록의 순위가 바뀌게 된다. 그림에서 (b)와 (c)의 숫자는 밝기 순위를 나타내고 있다.

Oostveen 등의 알고리즘은, 블록의 평균값을 서로 이웃한 블록과 비교하여 영상 지문을 만들었다. 이러한 경우, 서로 이웃한 블록을 비교하기에 값들 간의 상관관계가 커지게 된다. 예를 들어, 첫 번째 블록 값이 두 번째 블록보다 크고, 세 번째 블록이 네 번째 블록보다 크다면 두 번째 블록의 값이 세 번째 블록의 값보다 클 확률이 높게 된다.

본 논문에서는 이러한 단점을 해결하기 위해, 다양한 특징과 비교 쌍을 구성했다. 또한, 각 블록의 하나의 특징은 하나의 비교 방법에서 한번만 사용을 하여 각 비트간의 상관관계를 줄였고, 하나의 비교를 이진화 하여 각 비트를 생성함으로써 잡음에 의한 비트의 외곡이 다른 비교에 영향을 미치지 않도록 했다.

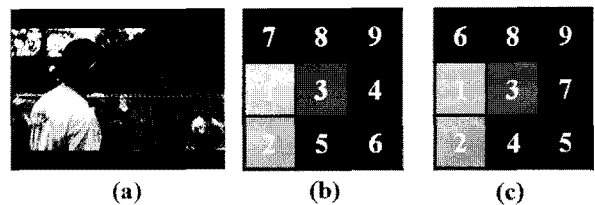


그림 3. Ordinal 방법의 문제 (a)원본 영상, (b)원본 영상의 Ordinal, (c)변형 영상의 Ordinal  
Fig. 3. The problem of ordinal method. (a)original image, (b)ordinal of original image, (c) ordinal of modified image

#### 2. 블록 비교 기반 특징

그림 4는 제안된 알고리즘의 전체 개략도이다. 제안된 알고리즘은 프레임 기반 알고리즘으로, 입력된 동영상 클립으로부터 프레임을 분리한다. 분리된 프레임은 전처리로 영상을 단색조로 바꾸고 100×100 크기로 정규화 한다. 본 논문에서 사용한 특징은 블록 평균값 기반 1차, 2차 미분 값이다. 제안된 영상 지문을 구하기 위하여 전 처리 된 프레임을 그림 5와 같이 5×5 블록

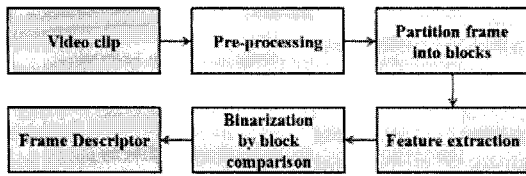


그림 4. 제안된 영상 지문의 개관도  
Fig. 4. Overview of proposed fingerprint.

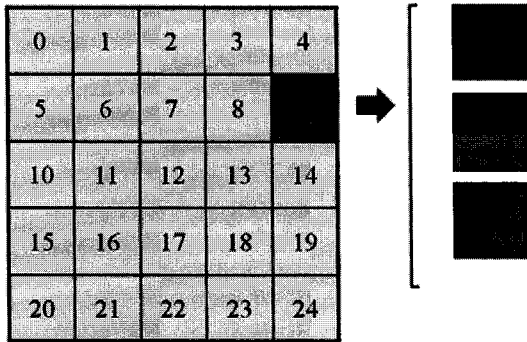


그림 5. 영상 5×5구획 및 사용된 형태  
Fig. 5. 5×5 image divide and using pattern.

으로 나눈다. 그림 5의 우측 3가지 그림은 계산된 특징을 나타낸다. 위쪽 그림은 영역 전체의 평균, 중앙은 블록 위쪽 평균과 아래쪽 평균의 차분, 하단은 블록 왼쪽 평균과 오른쪽 평균의 차분을 의미한다. 사용된 특징 값은 다음 식 (1)~(3)과 같이 구해진다. 각각의 프레임이 100×100이므로, 각 블록은 20×20 화소의 영역이므로, M=20이다. 이때  $B_n(i, j)$ 은 n번째 블록의 (i, j)번째 화소 값을 나타낸다.  $Average[n]$ 은 n번째 영역의 평균값을 나타내고,  $DiffX[n]$ ,  $DiffY[n]$ 은 각각 영역의 x방향과 y방향의 미분 값을 의미한다. 식 (4), (5)는 사용된 2차 미분 특징을 나타낸다.

$$Average[n] = \frac{1}{M \times M} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} B_n(i, j) \quad (1)$$

$$DiffY[n] = \sum_{i=0}^{M-1} \sum_{j=0}^{\frac{M-1}{2}-1} B_n(i, j) - \sum_{i=0}^{M-1} \sum_{j=\frac{M}{2}}^{M-1} B_n(i, j) \quad (2)$$

$$DiffX[n] = \sum_{i=0}^{\frac{M-1}{2}-1} \sum_{j=0}^{M-1} B_n(i, j) - \sum_{i=\frac{M}{2}}^{M-1} \sum_{j=0}^{M-1} B_n(i, j) \quad (3)$$

$$Diff2X[n] = Average[n-1] + Average[n+1] - 2 \times Average[n] \quad (4)$$

$$Diff2Y[n] = Average[n-5] + Average[n+5] - 2 \times Average[n] \quad (5)$$

### 3. 거친 지문(Coarse Fingerprint)의 구조

동영상 복사 검출을 위한 영상 지문의 중요한 특징 중 하나가 고속 정합이다. 이를 위해서는 영상 지문이 이진화 되어 있는 것이 유리하다. 본 논문에서는 제안된 5가지 특징을 서로 비교하여 식 (6)을 사용하여 영상 지문을 이진화 하고, 거친 지문과 미세 지문을 분리하여 빠른 정합이 가능하도록 하였다.

수식 (2)~(5)의 특징은 각각의 값이 음수/양수에 따라 식 (6)을 사용하여 이진화가 가능하다.  $Value[x]$ 는 이진화된 x특징의 값을 얘기한다. x는 수식 (2)~(5)의 값과 각 특징의 비교를 이용하여 구해진다.

$$Value[x] = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

블록 12를 기준으로 비교 쌍을 만들어 12의 비교 횟수가 적고, 외각 영역보다 중심 영역에 더 많은 정보가 있으므로 중심 부분을 더 많이 사용한다. 이런 방법의 비교를 통해 전체 72가지 값을 만들어 이를 9Byte의 용량에 저장한다. 전체 72가지 값은 비교를 통해 만들고, 식 (6)으로 이진화 한다. 전체 72비트는 다음과 같다.

- $DiffX[n]$ ,  $DiffX[n]$ 의 값 자체.  $DiffX[n]$ 의 경우 영상 좌우측 최외각 블록은, 필러 박스의 영향을 많이 받으므로 제외한다. 같은 이유로  $DiffY[n]$ 의 상하 최외각 블록은 레터박스의 영향을 많이 받아 제외한다. (30 비트)

- 12번 블록을 기준으로 원점 대칭, 평균, 1차 미분 값 비교. 위와 같은 이유로,  $DiffX[n]$ ,  $DiffX[n]$ 의 비교는 7가지 씩 사용을 했다. 평균값 역시 상하 최외각 블록은 비교에서 제외한다. (21비트)

- 12번 블록을 기준으로 원점 대칭  $DiffX[n]$ ,  $DiffX[n]$ 의 합산 값을 비교하였다. 위와 같은 이유로 중심 영역에서만 비교를 한다. (4비트)

- 큰 영역을 비교하기 위하여, 좌우측(직사각형) 블록들의 평균값을 비교하고, 위 아래 영역을 비교했다. (비교 영역 : 0~21 vs. 3~24, 0~9 vs. 15~24) 또한, 12번째 블록과 그 주변(6, 7, 8, 11, 13, 16, 17, 18 블록)의 평균값을 비교한다. (3비트)

- 6, 7, 8, 11, 13, 16, 17, 18 블록의 평균값과 12번 블록을 제외한 중심부 블록(6, 7, 8, 11, 13, 16, 17, 18 블록)의 값을 비교한다. (8비트)

- 식 4, 5에 의한 값 자체를 식 6을 사용하여 이진화 하였다. 식의 값은 필러, 레터 박스의 영향을 받지 않는

중심부 영역만을 사용한다. (6비트)

4. 미세 지문의 구조

앞에서 설명한 영상 지문의 요구 조건 중 하나가 쌍 간 독립성(pairwise independence)이다. 위에 설계된 9 Byte로는 영상 지문이 충분한 변별력을 갖지 못한다. 충분한 변별력을 가질 수 있도록 그림 6과 같이 영상을 4 영역으로 나눠 각각의 영역에 대한 거친 지문을 구해 전체 36Byte의 미세 지문을 구성 한다.



그림 6. 미세 지문을 위한 영상 구획  
Fig. 6. Image divides for Fine fingerprint.

5. 정합

제안된 지문은 각 비트가 이진화 되어 있다. 그러므로 해밍 거리(Hamming distance)로 거리를 구할 수 있다. 우선, 거친 지문의 거리가 문턱치보다 작다면 복사본일 가능성이 높기 때문에 2단계인 미세 지문의 거리를 구함으로 최종 거리를 구할 수 있다. 이에 대한 도식은 그림 2에 나와 있다.

2단계의 정합을 하기 위해서는 1단계 정합의 문턱치가 필요하게 된다. 1단계 정합에서는 거친 지문을 사용하므로 전체 72bit의 값을 갖는다. 이상적인 경우, 0과 1의 분포는 각각 0.5의 확률을 갖는다. 이를 바탕으로 BER(Bit Error Rate)을 구하였다. 이상적인 경우 BER은 정규 분포를 갖는다. 0과 1의 확률이 0.5로 같으므로,  $p = \mu = 0.5, \sigma = \sqrt{p(1-p)/n}$  이 된다<sup>[10]</sup>. 전체 양성 거짓(False Alarm)은 다음 식 (7)과 같이 구할 수 있다.

$$P_{FA} = \int_{-\infty}^T \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \tag{7}$$

$$= \frac{1}{2} \operatorname{erfc}\left(\frac{\mu-T}{\sqrt{2}\sigma}\right)$$

여기서 T는 문턱치 값으로, 본 논문의 구현에서는 0.35를 사용하였다. 이 값을 식 (7)에 대입하면,

$P_{FA} \approx 0.01$ 이 된다. 서로 다른 영상을 1% 정도 통과시키게 되는 것이다. 이 값을 통과한 영상만을 대상으로 미세 지문을 비교하여 최종 결과를 계산한다.

IV. 실험 결과

1. 실험 조건

본 논문은 동영상 복사 검출을 위한 새로운 영상 지문을 제안하였다. 현재 MPEG 7의 Visual Core Experiment 7(VCE-7)에서는 동영상 식별자(Video Signature)에 대한 표준화가 진행 중이다<sup>[11]</sup>. 본 논문의 실험에서는 VCE-7의 실험 조건을 사용하였다. 표 1은 VCE-7에서 사용한 변형의 종류 및 변형 정도를 나타낸다. 표에서 자막/로고 삽입(Text/logo overlay)의 수치는 전체 영역 중 얼마의 비율에 자막/로고가 삽입되어 있는지의 수치를 의미한다. 그리고 카메라 촬영

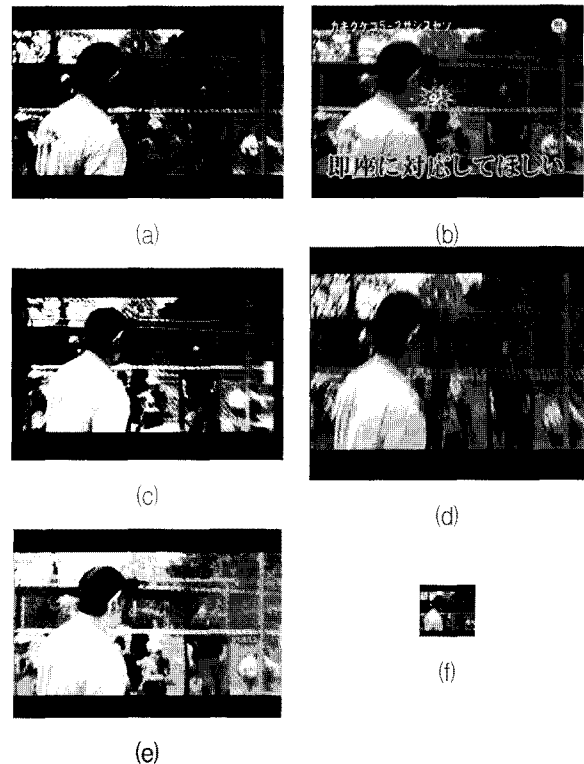


그림 7. 영상 변형(변형정도 Heavy)의 예 (a) 원본, (b) Text/logo overlay, (c) Capturing on Camera, (d) Analog VCR recording & recapturing, (e) Brightness change, (f) Compression  
Fig. 7. Example of modified image(Heavy level) (a) Original, (b) Text/logo overlay, (c) Capturing on Camera, (d) Analog VCR recording & recapturing, (e) Brightness change, (f) Compression.

표 1. MPEG 7의 VCE-7 변형과 변형 정도 (9가지 변형에 총 22가지 종류)

Table 1. Modifications and modification levels for VCE7 in MPEG 7 (9 modifications, 22 category).

Modifications	Levels		
	Heavy	Medium	Light
Text/logo overlay	30%	20%	10%
Severe compression	64kbps	256kbps	512kbps
Resolution reduction	N/A	QCIF	CIF
Frame-rate reduction	4fps	5fps	15fps
Capturing on camera	3 times 10%	2 times 5%	1 time 0%
Analog VCR recording&recapturing	3 times	2 times	1 time
Color to monochrome conversion	N/A	N/A	$I = 0.299 \times R + 0.587 \times G + 0.144 \times B$
Brightness change	+36	-18	+9
Interlaced Progressive conversion	N/A	N/A	P->I->P, I->P

(Capturing on Camera)의 % 수치는 카메라로 찍었을 때의 여백 정도의 수치이다. 그림 7은 사용된 영상의 변형 예시를 보여주고 있다.

본 논문에서 제안한 영상 지문의 성능을 비교하기 위하여 각각의 동영상 클립에 변형을 취하고, 각 프레임을 정지 영상으로 저장해 강인성과 독립성을 구했다. 또한, 이를 바탕으로 ROC(Region Of Conversions) 곡선을 구했다. 이 때 전체 사용된 동영상 클립은 500여 개이고, 원본 동영상에서 추출한 약 6000장의 정지 영상을 사용하여 전체 50000여장의 시험영상을 만들었다. 여기서 사용된 변형은 표 1의 프레임 율 변화(Frame Rate Reduction)를 제외한 모든 변형을 포함한다. 이 때 강인성은 원본 영상과 변형된 영상을 비교했을 때 특정 문턱치보다 작은 비율을 나타내고, 독립성은 원본과 서로 다른 영상을 비교 했을 때 특정 문턱치보다 작은 비율을 나타낸다.

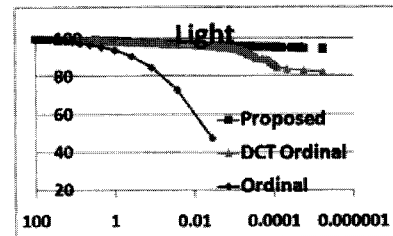
본 논문에서는 변형정도 Light, Medium, Heavy의 평균 성능을 비교하고, 지면 관계상 같은 성능의 독립성에서 가장 성능이 떨어지는 성능만을 표2에 나타냈다. 또한, 각 변형 정도에 대한 평균을 사용하여 ROC 곡선을 구하고 그 성능을 비교 했다.

2. 실험 결과

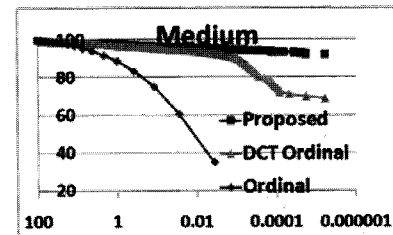
그림 8은 제안된 알고리즘과 비교 알고리즘의 ROC 곡선을 보여주고 있다. 그림의 (a), (b), (c)는 각각 변형 정도 Light, Medium, Heavy의 강인성을 평균하여 나타냈다. 그림의 x축은 독립성의 %를 log 단위로 나타낸 것이고, y축은 강인성의 %를 나타내고 있다.

그림에서 보듯이, 제안된 방법이 같은 독립성에서 강인성의 성능이 위쪽으로 붙어 있고, 독립성이 낮아질 때 완만한 속도로 강인성이 떨어지는 것을 볼 수 있다. 이는 독립성이 떨어질 때 강인성에 받는 영향이 다른 방법들에 비해 크지 않음을 의미한다. 비교에 사용한 방법은 Ordinal 방법<sup>[3]</sup>과, DCT-Ordinal<sup>[9]</sup> 방법이다.

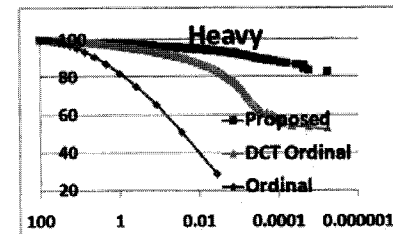
표 2는 제안된 방법과 기존의 방법들을 비교한 것이다. 각각의 독립성의 값을 비슷하게 맞춰 놓고 VCE-7에서 규정해 놓은 변형에 대한 강인성을 비교한 것이



(a)



(b)



(c)

그림 8. 제안된 영상 지문과 비교 방법의 ROC 곡선 (a) 변형정도 Light, (b) 변형정도 Medium, (c) 변형 정도 Heavy

Fig. 8. ROC curves for proposed fingerprint and comparing methods (a) modified level Light, (b) modified level Medium, (c) modified level Heavy.

다. 각각의 경우 비슷한 비트 수를 사용하는 것으로 설정하여 비교하였다. ordinal 방법은 3×3 영역에 대하여 수행하여 9Byte를 사용하였고, DCT-Ordinal 방법은 48개의 영역에 대해 수행하여 48Byte를 사용하였다. 표에서 제시한 제안된 알고리즘의 수치는, 각각 거친 영상 지문, 미세 영상 지문, 2단계 정합을 사용한 제안된 영상 지문이다.

2단계 정합을 사용한 제안된 영상 지문의 결과와 미세 영상 지문의 수치를 비교해보면, 같은 독립성에서 성능의 저하가 전혀 없이 평균 0.1% 정도의 성능 향상이 있는 것을 볼 수 있다. 이 결과는, 거친 영상 지문에서 서로 다른 영상을 잘 걸러줘서 생긴 성능 향상이고, 이는 2단계 정합을 사용해도 성능에 문제가 없음을 의미한다.

거친 영상 지문의 경우, 용량이 비슷한 Ordinal 방법과 비교해 보면, 같은 독립성에서 평균 24%의 높은 성능을 보이고 있다. 이 수치는, 독립성조건이 어려워질수록

표 2. 제안된 지문과 기존 방법의 비교 (단위 %)  
Table 2. Comparison for proposed fingerprint with other method (%).

	Independence	Light	Medium	Heavy	Worst Case
Proposed (Coarse)	0.00224	94.5	91.3	84.9	55.7
	0.01911	96.2	94.0	90.6	69.7
	0.11717	97.5	95.9	94.1	79.9
	0.9173	98.7	97.4	97.0	89.7
	4.55564	99.4	98.3	98.6	95.1
Proposed (Fine)	1.76E-05	94.8	92.5	83.9	65.8
	0.00012	95.5	93.7	88.2	73.0
	0.00106	96.1	94.7	92.0	80.5
	0.00224	96.3	95.0	93.0	82.9
Proposed (Cascade)	0.00588	96.5	95.3	94.3	86.0
	1.76E-05	94.8	92.5	83.9	65.6
	0.00014	95.6	93.8	88.6	73.6
	0.00116	96.2	94.8	92.6	81.3
	0.00186	96.4	95.0	93.2	82.8
Ordinal	0.00512	96.6	95.4	94.4	85.6
	0.00353	47.3	34.9	28.8	2.3
	0.02699	73.0	60.5	51.0	10.5
	0.11838	84.7	75.2	65.5	25.5
	0.99313	93.9	88.8	82.0	58.2
DCT Ordinal	4.2351	96.8	94.1	90.8	75.4
	1.76E-05	83.2	69.9	53.6	1.8
	0.00011	87.0	75.4	57.6	3.9
	0.00108	95.1	90.6	76.0	42.7
	0.00204	95.7	91.8	80.2	54.7
	0.00645	96.5	93.4	86.1	64.3

표 3. 정합 속도 비교 (단위 ms)  
Table 3. Matching speed (unit: ms).

	Proposed	Ordinal	D C T Ordinal
Processing time/Image	0.08	0.09	0.11

더 차이가 나는 경향을 보이며, 이는 제안된 거친 영상 지문이 어려운 조건에서 좋은 성능을 보일 것을 나타낸다. 2단계 정합을 사용한 전체 제안된 알고리즘을 사용한 결과는 DCT-Ordinal과 비교했을 때, 평균 11%이상의 높은 성능을 보이고, 독립성이 어려운 조건 하에서 20% 이상의 큰 차이를 보이고 있다. 또한, 같은 독립성 조건에서 강인성이 가장 떨어지는 변형을 각각 보면, 제안된 알고리즘이 평균 44% 이상의 좋은 성능을 보인다.

표 3은 평균 정합 속도를 보이고 있다. 실험은 서로 다른 영상 6000여장에 대한 정합을 하여 평균 처리 속도를 구하였다. 표에서 보듯이, 2단계 정합을 사용한 제안된 방법이 0.08ms로 Ordinal의 0.09ms, DCT-Ordinal의 0.11ms에 비해 가장 빠른 정합 속도를 보이고 있다.

### V. 결 론

본 논문에서는 내용기반 동영상 복사 검출에 적합한 영상 지문을 제안하였다.

일반적으로, 내용기반 복사 검출에서는 지문의 강인성과 독립성이 중요한 평가 요소이다. 제안된 방법은 영역에 대한 평균값과 1차, 2차 미분 값에 의한 결과 값을 사용하여 동영상 복사 검출에 적합한 새로운 영상 지문을 구성하였다.

기존의 ordinal방법 등과 비교했을 때, 독립성 0.0035% 이하에서 Light 변형에서 평균 47.3%를 통과시키는데 비해, 제안된 방법은 96.3%를 통과시키고 있다. 또한, 제안된 방법보다 더 많은 비트를 사용하는 DCT-Ordinal 방법과 비교했을 때, 독립성 1.76E-05에서 Light 평균 83.2%를 통과시키는 반면, 제안된 방법은 94.8%를 통과시킨다. 특히, 같은 독립성에서 가장 떨어지는 성능을 보이는 변형을 비교 방법은 1.8%를 통과시키는데 반해, 65.6%를 통과시켜 더 좋은 성능을 보이고 있다. 또한, 2단계 정합 방법을 사용함으로써 빠른 정합 속도를 보여준다. 실험 결과에서 보듯이, 본 영상 지문은 동영상 복사 검출에 좋은 성능을 보일 것이다.

## 참고 문헌

- [1] Jin S. Seo, Minho Jin, Sunil Lee, Dalwon Jang, Seungjae Lee and Chang D. Yoo, "Audio Fingerprinting Based on Normalized Spectral Subband Moments," *IEEE Signal Processing Letters*, vol. 13 no. 4, pp. 209-212, Apr. 2006.
- [2] S. C. Cheung and Avideh Zakhor, "Efficient video similarity measurement with video signature," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13 no. 1, pp. 59-74, Jan. 2003.
- [3] Changick Kim and B. Vasudev, "Spatiotemporal sequence matching for efficient video copy detection," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15 no. 1, pp. 127-132, Jan. 2005.
- [4] Job Ostveen, Ton Kalker, and Jaap Haitsma, "Feature Extraction and a Database Strategy for Video Fingerprinting," *Proc. International Conference on Recent Advances in Visual Information Systems*, pp. 117-128, 2002.
- [5] D. G. Lowe, "Object recognition from local scale-invariant features," *Proc.*, vol. 13 no. 4, pp. 209-212, Apr. 2006.
- [6] Arun Hampapur and Rudolf M. Bolle, "VideoGREP: Video Copy Detection using Inverted File Indices," *Technical Report, IBM Research*, 2001.
- [7] Sunil Lee and Chang D. Yoo, "Video Fingerprinting Based on Centroids of Gradients," *Proc. ICASSP 2006, Toulouse, France.*, vol. 2, pp. 401-404, May. 2006.
- [8] Sunil Lee and Chang D. Yoo., "Robust Video Fingerprinting for Content-Based Video Identification," *IEEE Signal Processing Letters*, vol. 13 no. 4, pp. 209-212, Apr. 2006.
- [9] Changick Kim, "Content-based image copy detection," *Signal*, vol. 2, pp. 401-404, May. 2006.
- [10] Jin S. Seo, Jaap Haitsma, Ton Kalker, Chang D. Yoo, "A robust image fingerprinting system using the Radon transform," *Signal Processing: Image Communication*, vol. 19, pp. 325-339, 2004.
- [11] MPEG Video Sub-Group, "Description of Core Experiments in Video Signature Description development", ISO/IEC JTC1/SC29/WG11, w10345, Lausanne, Switzerland, February 2009.



저 자 소 개



나 상 일(학생회원)  
 2002년 인하대학교 전자공학과  
 학사 졸업  
 2004년 인하대학교 전자공학과  
 석사 졸업  
 2004년~현재 인하대학교  
 전자공학과 박사과정

2008년~현재 ETRI 연구원  
 <주관심분야 : 영상처리, 컴퓨터 비전, 패턴인식,  
 내용 기반 검색>



진 주 경(학생회원)  
 2003년 인하대학교 전자공학과  
 학사 졸업  
 2005년 인하대학교 전자공학과  
 석사 졸업  
 2005년~현재 인하대학교  
 전자공학과 박사과정

<주관심분야 : 영상처리, 컴퓨터 비전, 패턴인식,  
 내용 기반 검색>



조 주 희(학생회원)  
 2009년 인하대학교 전자공학과  
 학사 졸업  
 2009년~현재 인하대학교  
 전자공학과 석사 과정  
 <주관심분야 : 영상처리, 컴퓨터  
 비전, 패턴인식, 내용 기반 검색>



오 원 근(정회원)  
 1988년 Osaka 대학교  
 전자공학과 공학 박사  
 1988년~현재 ETRI 책임 연구원  
 <주관심분야 : DRM, 컴퓨터 비  
 전, AI, 멀티미디어 정보처리>



정 동 석(정회원)  
 1977년 서울대학교 전기공학과  
 학사 졸업  
 1985년 Virginia 주립 공과대학  
 전자공학과 공학석사  
 1988년 Virginia 주립 공과대학  
 전자공학과 공학 박사

1988년~현재 인하대학교 전자공학과 교수  
 <주관심분야 : 영상처리, 컴퓨터 비전, 패턴인식,  
 멀티미디어 정보처리>