

통계적 분석 기반 불법 비디오 영상 감식 방법

*김지은 *조혜정 *오승준

*광운대학교 VIA-멀티미디어 센터

*jekim@media.kw.ac.kr

A Novel Video Copy Detection Method based on Statistical Analysis

*Ji Eun, Kim *Hye Jung, Cho *Seoung Jun, Oh

*Kwangwoon University VIA-Multimedia Center

요약

본 논문은 공간영역 상에서 다양하게 변형된 복제 영상과 원본 영상간의 통계적 특성을 이용하여 그 유사도를 측정하고 복제 여부를 판단하는 계층적 구조의 불법 비디오 감식 방법을 제안한다. 영상의 대표적 특성인 명암도에 따라 순위를 매김으로써 공간적 변형에 영향을 받지 않도록 하였으며, 데이터베이스에 저장된 방대한 양의 영상들에 대한 검색 시간과 계산량을 줄이기 위해 순위 표본 프레임(키프레임)을 이용하여 유사한 후보 영상군을 추출한다. 이러한 후보 영상군을 대상으로 일반적인 불법복제 비디오의 형태를 감안하여 각 프레임의 중앙 영역을 포함하여 통계 검정을 이용함으로써 복제 여부를 판단한다. 실험 결과 제안하는 방법은 기존 방법과 유사한 정확도를 보이며 동시에 선택된 순위 표본 프레임 수는 약 50% 가량 적게 추출되어 검색 시간과 계산량이 감소하였다. 또한 영상의 화질 열화, 대비 변형, 확대 및 축소, letterbox 등 다양한 공간적 변형에도 포괄적으로 복제 여부를 판단할 수 있음을 실험을 통해 확인하였다.

1. 서론

인터넷과 멀티미디어 기술의 발달로 우리는 다양한 콘텐츠를 손쉽게 접할 수 있게 되었다. 그러나 최근 무분별한 불법콘텐츠들의 유통으로 인해 저작권 침해가 심각한 사회적 문제로 대두되고 있는 실정이다. 이에 정부와 저작권 보호단체는 불법콘텐츠의 단속과 처벌을 강화하는 등의 노력을 기울이고 있다[1].

기존의 불법콘텐츠를 감식하는 대표적인 방법으로 워터마킹 등이 대안으로 연구되고 있으나, 미디어 파일의 변형이나 이미 배포된 미디어에 대해서는 보호할 수 없는 단점이 있다. 반면에 내용기반 복사 검출(content-based copy detection)은 콘텐츠 자체가 충분히 복사본을 검출할 수 있는 특징 정보를 보유하고 있어 변형되거나 이미 배포된 콘텐츠에 대해서도 감식할 수 있는 장점이 있다.

본 논문은 내용 기반의 순위 정합(ordinal measurement)방법을 사용하여 공간적인 변화에 강인하도록 구성하였다. 영상 흐름의 변화에 따라 순위 표본 프레임(keyframe)을 선택하고, 이를 기반으로 원본 영상과 불법복사본 사이의 상관관계를 측정함으로써 인터넷 상의 많은 콘텐츠 중 유사한 영상만(candidate sequence)을 추출하여 계산량을 줄이고 처리 속도를 향상 시켰다. 또한 다양한 종류의 변형으로 쉽게 변질되는 휘도, 색상 등의 특성을 사용해 유사도를 비교하는 데 한계가 있으므로 영상 특징에 대해 통계적인 검정 방법을 적용하여 포괄적인 범위의 변형에도 복사 여부를 검출 할 수 있다.

논문의 구성은 다음과 같다. 2장에서는 기존의 내용 기반의 복사 검출 연구를 설명하고, 3장에서는 불법콘텐츠를 감식하는 시스템 구조와 구체적인 방법을 제안한다. 4장에서는 제안한 방법에 대한 실험 결과와 분석 제시하고, 5장에서 결론을 맺는다.

2. 기존의 내용 기반의 복사 검출 연구

내용 기반 비디오 복사 검출은 원본 영상으로부터 시간 또는 공간 영역에서의 특징을 추출하고 이러한 특징을 실시간으로 전송되는 목표 영상으로부터 추출된 특징과 정합하여 복사여부를 검출한다.

Naphade[2]는 MPEG 비디오에서 DC 영상만을 이용하여 프레임 특징의 순서를 사용하는 방법을 제안하였으며 Bhat[3]은 정지영상 검색을 위해 순위 비교 방법을 제안하였고, 이를 Mohan[4]이 영상 기반에서 순위 비교 방법을 제안하였다. 또한 Kim[5]은 순위 비교 방법을 공간적인 특징뿐만 아니라 시간적인 특징 정보를 추가한 방법을 제안하여 보다 색상 변화에 강인성을 향상시켰다.

위에 언급한 기존의 내용 기반 복사 검출 방법들은 실제 대응량의 영상 데이터에 적용할 경우 처리속도 측면에서 많은 시간이 소요된다. 동영상의 편집 소프트웨어가 발달됨에 따라 온라인상에서 일반사용자가 불법복제 영상물을 업로드 하는 양이 방대해졌으며 이러한 환경에서 불법복제 판단 여부를 빠르게 판단 할 수 있는 방법들이 많이 연구되어져야 한다. 이에 Chih-Yi Chiu[6]는 DOG(Difference Of Gaussian)을 통해 순위 표본 프레임을 추출하여 이를 기반으로 후보 영상군을 선택하고 선택된 영상군과 목표 영상간의 순위 정합 방법 기반의 매칭(matching)그래프를 이용하여 복제 여부를 판단하는 방법을 제안하였다.

3. 제안하는 시스템 구조 및 방법

일반적으로 불법콘텐츠를 접할 수 있는 온라인 사이트는 많은 양의 원본 영상(Query)을 보유하고 있기 때문에 하나의 목표 영상(Target)을 방대한 데이터베이스에 저장된 모든 원본 영상과 비교하는 데 한계가 있다. 따라서 본 논문은 영상의 특징인 시간적 중복성을 이

용하여 유사한 영상을 빠르게 추출함으로써 검색시간과 계산량을 감소시켰다. 그림 1에 나타난 것과 같이 제1 데이터베이스에 원본 영상들이 저장되어 있고, 빠른 검출을 위해 오프라인 상에서 원본 영상들로부터 순위 표본 프레임을 추출하여 제2의 데이터베이스에 따로 저장해 둔다. 이후 원본 영상을 복제한 목표 영상(Target)이 발생했을 때 두 영상의 통계적 특성을 이용하여 일치도를 측정한다.

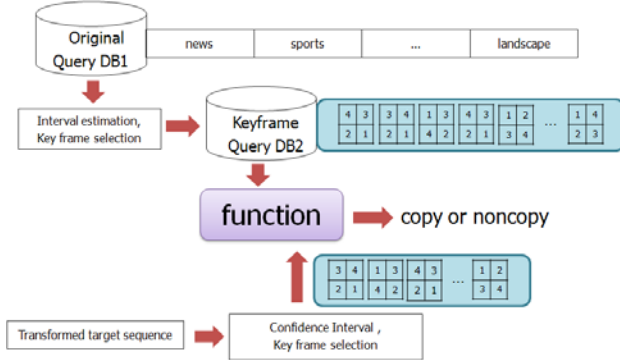


그림 1. 시스템 구성도

그림 1에서의 제1 데이터베이스에 저장된 원본 영상과 복제된 목표 영상을 식(1)과 같이 표기한다.

$$Q = \{Q_1^1, \dots, Q_i^m\}, T = \{T_1, \dots, T_i\} \quad (1)$$

여기서 m 은 원본 영상의 개수이며 i 는 프레임(frame)을 의미한다.

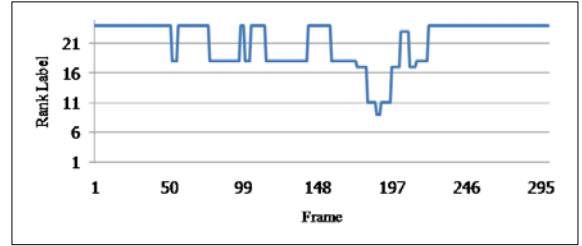
일치도를 측정하기 위해 가장 먼저 Q 와 T , 각 영상에 대해 순위 정합 방법에 따라 각 프레임의 특징을 추출하고, 이 프레임들의 흐름에서 Spearman의 순위 상관 계수(Rank Correlation Coefficients, RCC)를 이용하여 변화가 일어난 것으로 판단되는 프레임을 표본 프레임으로 선택한다. 이렇게 추출된 표본 프레임을 기반으로 원본 영상과 목표 영상간의 상관도를 Spearman의 순위 상관 계수를 이용하여 측정함으로써 후보 영상군을 선택하게 된다. 이 후보 영상들과 목표 영상 사이의 일치도를 t -검정(t -test)을 통해 판별하는데, 이 때 사용될 값은 위에서 나누었던 $N_x \times N_y$ 로 나눈 블록과 같은 크기의 중앙 영역 블록을 추가하여 t_0 값을 계산한다. t_0 값은 t -검정표(t -table)를 참고하여 정해진 임계값과 비교함으로써 복제 여부를 최종적으로 판단하게 된다.

영상의 흐름상 급격한 휘도 변화가 나타나는 프레임을 기반으로 공간적(spatial) 특징 정보를 활용한다. 본 장에서는 이러한 특징을 추출하는 방법과 이를 이용한 유사도 및 복사 여부를 판단하는 자세한 방법에 대해 설명한다.

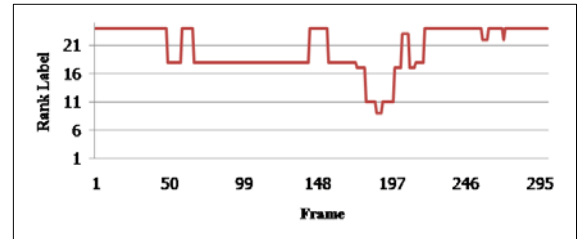
3.1. 순위 표본 프레임(keyframe) 추출

복제된 불법 영상은 다양한 방법으로 변형된 형태가 일반적이다. 이러한 변형에도 영상의 대표적인 특징을 나타낼 수 있는 것이 휘도이며, 이 휘도값도 대비 변화나 잡음 등에 의해 그 값이 달라진다. 따라서 본 논문에서는 여러 가지 방법의 변화에도 영상을 특징화 할 수 있으면서 동시에 강인성을 제공하는 순위 정합 방법을 사용한다. Q 와 T 의 영상의 i 번째 프레임의 순위 정합 방법, $Rank(\cdot)$ 을 이용해 얻어진 순위 프레임을 $Q_{i,j}^m, T_{i,j}$ 로 표기한다.

그림 2는 foreman의 원본 영상과 변형된 목표 영상과의 순위 상관에 따른 변화를 도시한다. 한 영상 내에 순위 정합 방법을 통해 얻을 수 있는 각 순위조합의 경우에 대해 원본 영상(a)과 변형된 영상(b)의 라벨링(labeling)한 결과이다. 그림 2에서와 같이 시간에 따른 순위의 중복성이 뚜렷하기 때문에 순위의 변화가 발생하는 위치를 표본 프레임으로 추출함으로써 영상 전체를 비교하는 것 보다 처리시간을 크게 절약할 수 있다. 또한 변형된 영상에 대해서 순위 변화가 원본 영상과 차이가 크지 않기 때문에 두 영상간의 유사한 순위 표본 프레임 도출이 가능하다.



(a) foreman 순위 정보의 시간적 변화



(b) Letterbox로 변형된 foreman 순위 정보의 시간적 변화

그림 2. 순위 정보의 시간적 변화

공간상에서의 주어진 순위를 기반으로 프레임 간의 순위 변화에 대한 상관관계를 분석하기 위해 RCC 을 이용하여 식(2)의 조건에 만족하는 순위 표본 프레임을 추출한다.

$$\hat{Q} = \{\hat{Q}_{1,1}^1, \dots, \hat{Q}_{i,j}^n | n \leq m \text{ and } RCC(Q_{i-1}, Q_i) < 1\}, \quad (2)$$

$$\hat{T} = \{\hat{T}_{1,1}, \dots, \hat{T}_{k,j} | RCC(T_{i-1}, T_i) < 1\}$$

RCC 값은 -1에서 1 사이의 값을 갖게 되는데, 1에 가까울수록 두 프레임 사이에 상관도가 높다는 것을 의미하고, 1의 값을 갖는 경우 두 요소가 완벽하게 일치한다고 할 수 있다. RCC 을 구하는 방법은 식(3)과 같으며 $Diff(U, V)$ 는 U 와 V 각 영상의 한 프레임 내에서 (x, y) 번째 블록의 순위의 차를 의미한다.

$$RCC(U, V) = \frac{6 \times Diff(U, V)}{N \times (N^2 - 1)} \quad (3)$$

$$Diff(U, V) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} Rank(U_{x,y}) - Rank(V_{x,y})$$

여기서 RCC 값이 1의 값을 갖지 않을 경우 현재 프레임과 이전 프레임 사이에 변화가 있음을 의미하므로 변화가 있는 위치의 프레임은 순위 표본 프레임으로 정의한다.

3.2. 순위 표본 프레임에 따른 후보 영상군 선택

앞서 선택된 \hat{Q} 와 \hat{T} 를 기반으로 Q 와 T 의 영상간의 유사도를 측정하여 후보 영상군을 선택한다.

$$H = \{Q^1, \dots, Q^m | \pi \geq l \times \alpha\} \quad (4)$$

$$\pi = \sum_{l=0}^l \sum_{i=0}^k F(\hat{Q}_{i+v}, \hat{T}_i)$$

$$F(U, V) = \begin{cases} 1, & RCC(U, V) \geq \beta \\ 0, & otherwise \end{cases}$$

식(4)과 같이 \hat{T} 을 \hat{Q} 에 스캔하면서, RCC 값이 β 이상을 갖는 프레임 수가 \hat{Q} 의 총 프레임 수인 l 의 일정 수치(α)보다 많을 경우 그 프레임이 속한 Q^m 영상을 선택한다. β 는 원본 영상과 변형된 복제 영상의 RCC 값의 분포에 대해 95%의 신뢰구간을 갖는 χ^2 -검증을 사용하였다. 그 결과 RCC 값이 0.8 이상인 경우에 두 영상이 유사하다고 판단할 수 있었다. 또한 α 는 좀 더 정확한 검증을 위해 두 가지 경우의 수를 두었다. 먼저 영상의 변화도가 거의 없어 1개의 \hat{T} 의 프레임이 추출된 경우, 필요 이상의 후보 영상군이 선택되기 때문에 유사도 판정에 있어서 좀 더 엄격한 수치로 α 에 0.8을 적용한다. \hat{T} 의 프레임 수가 2개 이상인 경우에는 α 에 0.5를 적용한다.

3.3. 통계적 특성을 이용한 복사본 검출 방법

일반적으로 통계학에서 두 모집단의 데이터 유사성을 검정하기 위한 방법으로 평균을 이용한 가설검증 기법을 사용하고 있다. 본 논문에서도 이러한 가설검증을 이용하여 최종적으로 두 영상 간의 복제 여부를 판정한다.

변형된 복제 영상의 편집 위치를 보면 대부분 화면 중앙을 제외한 모서리나 상/하단 영역이며 이는 영상의 많은 정보를 담고 있는 가운데 영역을 훼손하지 않기 위해서이다. 이러한 복제 영상의 편집 특성을 고려하여 제안한 방법은 그림 3과 같이 중앙 블록을 추가하여 검증에 사용한다.

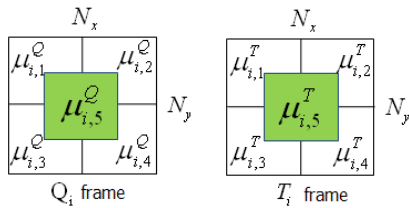


그림 3. 검증에 사용될 5개의 블록

각 블록의 평균 휘도값을 μ 라 하고, 유사성을 검증하기 위한 가설을 아래와 같이 정한다.

$$\begin{aligned} H_0 : \mu_{i,0}^T &= \mu_{i,0}^Q, \mu_{i,1}^T = \mu_{i,1}^Q, \mu_{i,2}^T = \mu_{i,2}^Q, \mu_{i,3}^T = \mu_{i,3}^Q, \mu_{i,4}^T = \mu_{i,4}^Q \\ H_1 : Not H_0 \end{aligned} \quad (5)$$

여기서 H_0 을 귀무가설(null hypothesis), H_1 을 대립가설(alternative hypothesis)이라고 하며 이 것은 두 블록의 표본 화소 그룹이 같은 모집단으로부터 온 것인지를 판단한다. 만일 설정된 귀무가설이 받아들여지면, 후보 영상과 목표 영상이 매우 유사한 것으로 판단하여 목표 영상을 불법영상으로 판단한다.

$\bar{X}_{Q_{i,j}^m}, \bar{X}_{T_{i,j}}$ 을 각각 $Q_{i,j}^m$ 과 $T_{i,j}$ 내 화소의 표본 평균이라고 하고, 식(5)에 대해 가설 검증을 실시한다. 평균에 대하여 귀무가설을 검증하기 위해 식(6)을 이용해 자유도 $v = n_Q - n_T - 2$ 을 가진 검증통계량(test statistics) t 을 사용한다.

$$t_0 = \frac{\bar{X}_{Q_{i,j}^m} - \bar{X}_{T_{i,j}}}{S_P \sqrt{\frac{1}{n_Q} - \frac{1}{n_T}}} \quad (6)$$

여기서 n_Q 와 n_T 는 각각 가설 검증을 하고자 하는 현재 후보 영상과 목표 영상의 표본의 수를 나타낸다. 따라서 통계적 가설 검증을 수행하기 위해 사용된 표본의 수는 충분히 큰 표본의 수를 가지기 때문에 평균 검증을 위해 t 분포를 이용하는데 아무런 문제가 없다.

한편 식(8)에서 공분산 S_P^2 은 다음과 같이 정의한다.

$$S_P^2 = \frac{(n_Q - 1)S_C^2 + (n_T - 1)S_C^2}{n_Q + n_T - 2} \quad (7)$$

S_P^2 은 두 영상의 표본 집단을 대표하는 분산으로 공분산이라고 한다. 일단 t 검증통계량이 계산되면 γ 을 갖는 신뢰구간(confidence interval) $-t_{\gamma/2, v} \leq t_0 \leq t_{\gamma/2, v}$ 을 설정하여, 신뢰구간 내에 t 검증통계량 값이 존재하면, 귀무가설을 받아들이고 그렇지 않을 경우는 기각한다. 본 논문에서는 자유도 v , 유의성 0.05에서 t -분포표를 참고하여 $t_{\gamma, v}$ 을 1.96으로 정의하고 실험하였다.

최종 검출 여부는 위와 같이 검증을 수행한 5개의 각블록에 대해 귀무가설이 받아들여진 블록의 개수가 전체 영상의 모든 블록 개수의 $\tau\%$ 이상이면 복제 영상으로 필터링 하게 된다. 여기서 τ 는 후에 실험 결과의 precision-recall률을 조절할 변수이다.

4. 실험결과 및 분석

본 논문의 실험에 사용된 원본 영상의 데이터베이스는 풍경, 뉴스, 스포츠 등의 12개 영상으로 구성되어 있으며 구체적인 내용은 다음과 같다.

- 원본 영상 : Akiyo, Carphone, Coastguard, Container, Foreman, Hall, Mobile, Mother and Daughter, News, Soccer, Stefan, Table tennis
- 영상 크기 : 352 X 288(CIF)
- 프레임 수 : 300 X 12개 = 3600

복제 영상은 그림 4와 같이 원본 영상의 6 가지의 단일변형을 통해 이루어진 영상을 이용하여 실험하였다.

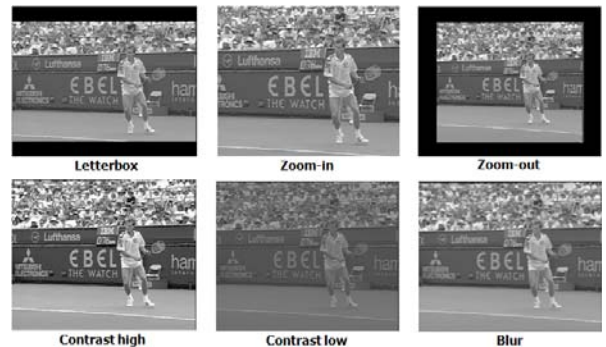


그림 4. 단일 변형된 복제 영상

4.1. 순위 표본 프레임의 정확도 실험

표 1,2에서 제시된 일치율은 원본 영상의 순위 표본 프레임과 변형된 복제 영상의 순위 표본 프레임간의 일치도를 측정된 결과이다. 순위 표본 프레임에 대해 정확도를 측정된 결과 제안하는 방법이 기존 방법에 비해 순위 표본 프레임 수는 원본 영상 3600프레임 중 제안한 방법이 174개의 순위 표본 프레임이 추출되었고, 반면 기존 방법은 388개의 순위 표본 프레임이 추출되어 그 수가 50%가량 감소하였다. 따라서 제안한 방법은 기존의 방법과 일치율이 유사하면서 동시에 적은 순위 표본 프레임을 이용하여 복사 여부를 빠르게 검출할 수 있다.

표1. 제안 방법의 순위 표본 프레임 정확도

	순위표본 프레임 수	일치하는 순위 표본 프레임 수	일치율
Blur	172	144	82.3
Contrast75	175	153	87.4
Contrast128	184	90	51.4
Letterbox	210	64	39.5
Zoom_In	223	59	33.7
Zoom_Out	175	85	48.5
평균	189.8	99.1	56.6

표2. 기존 방법의 순위 표본 프레임 정확도

	순위표본 프레임 수	일치하는 순위 표본 프레임 수	일치율
Blur	386	278	71.6
Contrast75	388	295	76.0
Contrast128	380	207	53.3
Letterbox	391	180	46.3
Zoom_In	424	164	42.2
Zoom_Out	349	168	43.2
평균	386.3	215.3	55.5

4.2. 복사 검출 성능 비교

본 논문에서는 최종적인 복사 검출율을 이전 방법인 Chih-Yi-Chiu와 비교하였다. 검출율의 성능을 평가하기 위해 식(7)의 precision 및 recall 측정하여 PR(precision-recall) 곡선을 이용한다.

$$Recall = TP / (TP + FN) \quad (7)$$

$$Precision = TP / (TP + FP)$$

식(7)에서 TP (True Positive)는 실제 복제 영상을 복제라고 검출하는 것이고, FP (False Positive)는 비 복제 영상을 복제라고 검출하는 것, 그리고 FN (False Negative)는 복제 영상을 비 복제 영상으로 판단하는 것이다. 그림 5는 각 변형된 형태에 따른 제안된 방법과 기존 방법의 성능 비교를 PR 곡선을 통해 나타낸다.

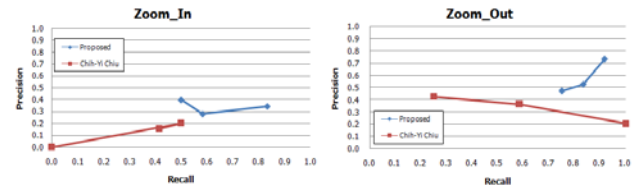
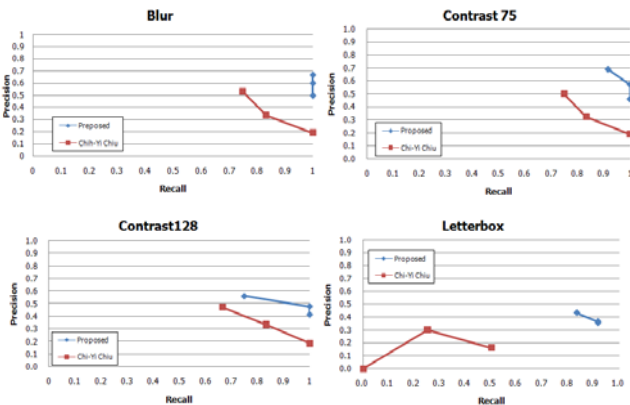


그림 5. 변형된 영상에 따른 PR 곡선

그림 5와 같이 제안한 방법이 기존 방법에 비해 precision과 recall 향상되었고, 기존 방법은 Zoom_In이나 Letterbox와 같은 변형에서 거의 검출이 안 되는 반면 제안한 방법은 다양한 변형에도 복제 여부를 적절히 판별할 수 있음을 알 수 있다. 제안한 방법의 precision은 낮은 편이지만, 상대적으로 Recall이 높기 때문에 웹상에 업로드 되는 불법 복제 영상의 복제 여부를 가리기에는 무리가 없다.

5. 결론

본 논문은 저작권과 불법복제 영상의 성행으로 인한 사회적 문제에 기반을 두어 불법으로 업로드 되는 복제 영상을 기존의 방법에 비해 보다 효과적으로 검출하는 방법을 제안하였다. 실험 결과에서와 같이 순위 표본 프레임의 정확도가 기존 방법과 유사함과 동시에 그 수가 기존 방법보다 50%가량 적어 검출 시간을 감축하여 고속처리가 가능함을 확인하였다. 또한, 통계적인 가설 검증을 통해 포괄적인 변형된 목표 영상에 대해서도 기존 방법에 비해 보다 높은 검출율을 보여 불법 복제 영상물의 복제 여부를 신속하고 정확하게 판단하는데 적합하게 활용될 수 있다.

감사의 글

본 연구는 2008년도 광운대학교 교내 학술연구비[60012008090]과 지식경제부와 한국산업기술재단의 전략기술인력양성사업과 정보통신연구진흥원의 차세대 DTV핵심기술개발사업의 일환으로 수행되었음.

참고 문헌

- [1] 김주섭, 남제호, “불법 복제 콘텐츠 필터링 기술 동향 분석”, 방송공학회지, 제12권, 제4호, pp. 371-381, 2007년 12월.
- [2] M. R. Naphade and M. M. Yeung, "A novel scheme for fast and efficient video sequence matching using compact signatures," in Proc. SPIE Conf. Storage and Retrieval for Media Databases, vol. 3972, pp. 564-572, Jan. 2000.
- [3] D. N. Bhat and S. K. Nayer, "Ordinal measure for image correspondence," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 4, pp. 414-423, Apr. 1998.
- [4] R. Mohan, "Video sequence matching," in Proc. IEEE ICASSP, vol. 6, pp. 3697-3700, May 1998.
- [5] C. Kim and B. Vasudev, "Spatiotemporal sequence matching for efficient video copy detection," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 1, pp. 127-132, Jan. 2005.
- [6] C. Y. Chiu and C. S. Chen, "A framework for handling spatiotemporal variations in video copy detection," IEEE Trans. Circuits Syst. Video Technol, vol. 18, no. 3, pp. 412-417, Mar. 2008