콘텐츠 불법 사용자 추적을 위한 디지털 비디오 핑거프린팅

한국과학기술원 강인구 · 이흥규

1. 개 요

인터넷의 발달과 디지털 멀티미디어 기술의 발달로 동영상, 이미지, 오디오 등의 디지털 멀티미디어 콘텐츠를 쉽고 빠르게 접할 수 있게 되었다. 그러나 디지털 콘텐츠의 특성상 원본과 동일하게 복사할 수 있다는 점과 사이버 세계의 익명성을 이용해 불법 복제 및 불법 유통이 빈번히 일어나고 있다. 콘텐츠의 무분별한 불법 배포는 콘텐츠 유통 사업의 활성화와 시장 확대를 막을 뿐 아니라 청소년들을 유해환경으로부터 보호할 수 없는 상황을 만들어가고 있다. 이러한 문제점을 해결하기 위해 저작권 보호를 위한 DRM과 디지털 워터마킹 등의 기술이 활발히 연구되고 있다. 디지털 핑거프린팅은 워터마킹 기술을 기반으로 하는 콘텐츠 보호 기술로써 디지털 콘텐츠에 구매자의 정보인 핑거프린트를 삽입 및 추출하는 기술이다. 구매자의 정보가 삽입된 콘텐츠가 불법복제 또는 배포를 통해 발견됐을 경우, 콘텐츠 안에 삽입된 핑거프린팅 정보를 검출함으로써 불법 배포자를 찾아낼 수 있다[1].

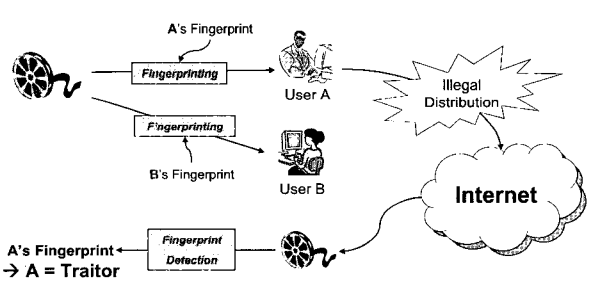

그림 1 콘텐츠 불법 사용자 추적을 위한 디지털 핑거프린팅

그림 1에서 보다시피 판매자는 구매자를 고유하게 구별할 수 있는 핑거프린트 코드를 동일한 콘텐츠 속에 삽입하여 판매한다. 향후 구매자가 콘텐츠를 불법 경로를 통하여 유통하거나, 다수의 콘텐츠 구매자들이 서로 공모하여 콘텐츠를 복사 및 생성을 하다가 적발되는 경우에, 콘텐츠에 삽입되어 있는 핑거프린트 코드를 추출하여 어느 구매자로부터 유통이 이루어졌는지를 알아낼 수 있다.

본 고에서는 이러한 특성을 지니는 디지털 핑거프린팅 기술에 대하여 알아보고자 한다. 우선, 콘텐츠 불법 유통자 추적을 위한 비디오 핑거프린트 시스템의 이해를 돕기 위하여, 비디오 핑거프린팅의 근간을 이루는 기술 요소들을 소개한다. 다음으로, 기반 기술 요소들을 바탕으로 평균화 공격에 강인한 비디오 핑거프린팅 시스템 구성에 대하여 살펴본다.

2. 비디오 핑거프린팅 기술 요소

본 고에서 제시하는 비디오 핑거프린팅 기법은 공모 보안 코드 기법을 바탕으로 하는 기술로써 그의 근간을 이루는 기반 기술 요소들은 크게 세 가지로 구분할 수 있다. 첫 번째는, 핑거프린팅 기술에 있어서 가장 중요한 핑거프린트 코드의 디자인이다. 두 번째는, 핑거프린트 코드를 콘텐츠에 삽입하기 위한 디지털 워터마킹 기술이다. 디지털 워터마킹의 한 응용 분야인 디지털 핑거프린팅은, 핑거프린트의 삽입 및 추출을 위하여 대체로 디지털 워터마킹의 기법을 사용한다. 마지막으로, 핑거프린트가 삽입된 콘텐츠에 대한 여러 가지 공격에의 강인성이다. 불법 사용자들은 핑거프린트 코드가 삽입된 콘텐츠로부터 핑거프린트 코드를 제거하려는 공격을 할 수 있다[2]. 이러한 공격들에 대하여 살펴보도록 한다.

2.1 핑거프린트 코드

디지털 핑거프린팅 시스템에 있어서 핑거프린트 코드의 디자인은 매우 중요하다. 핑거프린트 시스템의 예상되는 공격에 대하여 견고하면서도 효율적인 코드의 디자인이 요구된다. 핑거프린트 코드의 디자인 시 고려해야 할 사항들을 살펴보면 다음과 같다.

- 공격에 대한 강인성(Robustness): 악의적인 구매자들은 불법 콘텐츠의 유통을 위해서 핑거프린트가 삽입된 콘텐츠에 대하여 여러 가지 조작을 할 수

있다. 핑거프린트 코드는 공격이 가해진 콘텐츠에서도 원래의 정보를 잃지 않아야 한다.

- 핑거프린트 코드의 효율성(Efficiency): 핑거프린트 코드의 효율성이란, 핑거프린트 하나의 코드 길이에 대해 배포할 수 있는 전체 코드의 수를 의미한다. 예를 들어, 10명에게 부여할 수 있는 핑거프린트 코드 집합이 있고 각각의 코드의 길이가 10비트라면, 이러한 코드의 효율성은 1이라고 할 수 있다. 따라서 효율성이 높은 코드일수록, 같은 길이의 핑거프린트 코드로도 더 많은 구매자들에게 코드를 부여를 할 수 있게 되므로 핑거프린트 시스템의 성능이 좋아진다.
- 구매자들을 분별할 수 있는 성질(Uniqueness): 각각의 구매자들에게 배포되는 핑거프린트 코드는 구매자들을 유일하게 분별해야 한다. 서로 다른 구매자들에게 같은 핑거프린트 코드를 부여하면, 추후에 일어날 수 있는 불법 유통에 대하여 불법 유통의 원인지를 알 수 없으며, 나아가서는 불법 유통에 관여하지 않은 구매자들이 불법 유통에 관련되었다고 오해를 받는 일이 일어날 수 있다. 따라서 핑거프린트 코드는 각각의 구매자를 고유하게 분별할 수 있어야 할 뿐만 아니라, 추후에 일어날 수 있는 불법 유통 및 공격에 대해서 관련되지 않은 구매자를 불법 유통 및 공모의 요인으로 보고하지 않아야 한다.

2.2 비디오 워터마킹 기술

디지털 핑거프린팅은 디지털 워터마킹의 한 응용 분야이다. 삽입하는 정보에 따라서 저작권 정보를 삽입하면 워터마킹으로, 구매자 개개의 정보를 삽입하면 핑거프린팅으로 구분한다. 삽입하는 정보와 응용 시스템의 목적에 따라서 약간의 기술적 차이와 그에 따라 요구되는 성능의 차이는 있지만, 워터마킹 기술이나 핑거프린팅 기술의 기본 틀은 동일하다. 그러면 비디오 워터마킹의 기술에 대하여 살펴보자. 비디오 핑거프린팅에서의 정보의 삽입과 추출 방법도 비디오 워터마킹 기술과 유사하다. 비디오 워터마킹 기술에서의 삽입/추출 방법은 워터마크의 삽입/추출이 일어나는 영역의 압축 여부에 따라서 크게 두 가지로 구분된다. 비디오를 로우 프레임(Raw frame)으로 디코딩하여 각각의 로우 프레임에 워터마크를 삽입하고 다시 압축하는 비압축 영역(Uncompressped domain)에서의 삽입 방식과, 디코딩 과정을 거치지 않고 직접 비디오의 비트 스트림 속에 워터마크를 삽입하는 압축 영역(Compressed domain)에서의 삽입 방식으로 구분된다.

다시 비압축 영역에서의 워터마크 삽입/추출 기술은 워터마크의 삽입 영역에 따라서 공간 영역(Spatial domain) 삽입 방식과 주파수 영역(Frequency domain) 삽입 방식으로 분류된다. 공간 영역의 삽입 방식은 로우 프레임 데이터의 RGB, YCbCr, 또는 YUV 영역에 워터마크를 삽입하는 방식이다. 이 방식은 인지 시각 모델에 따라 RGB의 B 채널이나 YCbCr 또는 YUV의 Y 성분 (휘도, Luminance) 영역 등 사람의 시각에 민감하지 않은 부분에 워터마크를 삽입하는 방식이다. 이 방식은 삽입 방식이 간단하긴 하지만, 여러 가지 공격이나 압축 같은 신호처리에 약한 단점이 있다.

주파수 영역의 삽입 방식은 이미지를 DFT, DCT 또는 DWT 등의 변환을 사용하여 주파수 영역으로 변환한 뒤 적절한 주파수 밴드를 선택하여 워터마크를 삽입한 후 다시 역변환하여 원래 영상을 얻는 방식이다. 이 방식은 연구 초기에는 DCT나 DFT를 이용한 방법이 많이 연구되었지만 현재에는 DWT를 이용한 방법이 주로 연구되고 있다. 이 방식에서는 워터마크를 삽입하는 주파수 대역(Frequency band)에 따라 조금씩 다른 결과를 보인다. 저 주파수 대역(Low frequency band)에 삽입할 경우 압축이나 신호 처리에는 강하지만, 사람의 눈에 잘 띄어서 콘텐츠의 품질을 저하시키는 요인이 될 수 있다. 고 주파수 대역(High frequency band)에 워터마크가 삽입될 경우에는 비 인지성이 높은 반면, 워터마크의 강인성이 저하되는 성질이 있다. 따라서 두 주파수 대역의 특징과 시스템의 응용 목적에 따라서 삽입 대역을 결정하는 것이 중요하다.

MPEG-2 비디오의 경우, 비트 스트림은 헤더(Header)와 부가 정보(Side information), 모션 벡터(Motion vector)와 DCT 계수 블록으로 이루어진다. 압축 영역에서의 워터마크 삽입 방법은 비디오의 비트 스트림 중 DCT 계수를 나타내는 부분만을 채취하여 워터마크 신호를 삽입하여 새로운 DCT 계수 블록을 생성하고, 나머지 부분들은 그대로 복사하여 워터마크가 삽입된 새로운 비디오 비트 스트림을 생성하는 방법이다. 이 방법은 비디오 스트림을 로우 프레임으로 디코딩하고 인코딩하는 과정이 필요 없기 때문에, 워터마크를 삽입하는 계산 복잡도가 낮다는 장점이 있는 반면, 여러 가지 신호 처리 공격에 약하다는 단점이 있다.

2.3 디지털 핑거프린팅에서의 공격

디지털 핑거프린팅은 동일한 콘텐츠에 서로 다른 구매자의 정보를 삽입하기 때문에, 여러 구매자들이 공모하여 콘텐츠 간의 특성을 비교하여 서로 다른 점을 찾아 이를 제거하려는 공격을 할 수 있다. 이러한 공격을 공모

공격이라고 한다. 공모 공격에는 선형(Linear) 공격에 해당하는 평균화 공격(Averaging Attack)과 비선형(non-linear) 공격에 해당하는 최대최소 공격(Min-Max Attack), 중간값 공격(Median Attack) 및 상관계수 음수화공격(Negative-modified Attack) 등이 있다[2]. 최대최소 공격의 경우, 공모를 하는 콘텐츠에서 최소값과 최대값을 구한 후 평균값으로 새로운 콘텐츠를 생성하는 방법이고, 중간값 공격은 공모를 하는 콘텐츠에서 중간값을 구한 후, 그 값으로 새로운 콘텐츠를 생성하는 공격방법이다. 상관계수 음수화 공격은 상관계수를 이용하여 핑거프린팅 정보를 추출한 후, 이 값을 음수로 만들어 공모자 추출을 어렵게 만드는 공격이다. 삽입된 핑거프린트 정보는 이러한 공격 후에도 견고성(robustness)을 유지하여 **최소한 한 명 이상**의 공모자들 또는 배포자를 찾을 수 있어야 한다. 비선형 공격의 경우 대부분 선형 공격과 유사한 성능을 나타내므로 본 고에서는 선형 공격인 평균화 공격에 초점을 맞추었다[3].

평균화 공격은 공모자들이 가장 빠르고 효과적으로 핑거프린트 정보를 제거 할 수 있는 공모 공격이다. 일반적으로 핑거프린트 정보는 -1 또는 1 (혹은 0과 1)의 값을 갖는 난수열의 집합이 되며, 평균화 공격은 삽입된 핑거프린트 신호를 약하게 함으로 핑거프린트를 검출하지 못하도록 한다[4].

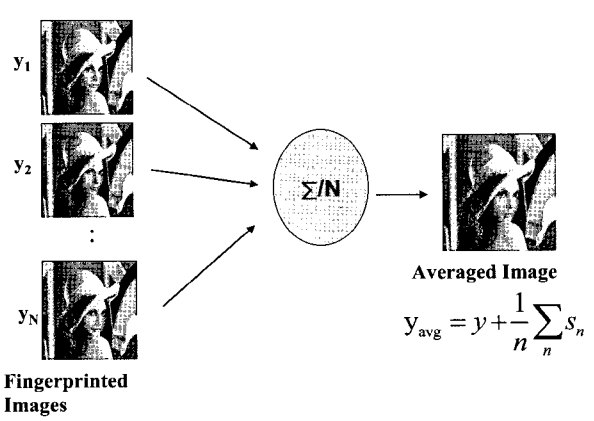

그림 2 평균화 공격

그림 2는 평균화 공격에 대한 모델이다. 여기서 y_i는 원본 콘텐츠이고, y_{avg}는 평균화 공격된 콘텐츠이다. n 명의 공모자들이 평균화 공격에 가담한 경우, 원본 콘텐츠의 값은 그대로 임에 반해 삽입된 핑거프린트 정보 s_n의 크기값은 $1/n$로 줄어들게 된다. 따라서 공모자의 수가 증가할수록 핑거프린트 신호의 크기가 줄어들어 원본 콘텐츠와 비슷하게 된다.

3. 평균화 공격에 강인한 비디오 핑거프린팅

이 장에서는 앞 장에서 설명한 비디오 핑거프린팅 시스템의 요소들에 대한 구체적인 예를 제시하고, 더불어 실제 비디오 핑거프린팅 시스템을 구성하여 본다. 이 장에서 예제로 제시하는 시스템은, 2.3절에서 설명한 여러 가지의 공격 중 평균화 공격에 견고한 비디오 핑거프린팅 시스템이다.

3.1 핑거프린트 코드 생성

본 절에서는 평균화 공격에 강인한 핑거프린트 코드를 생성한다. 평균화 공격에 강인한 핑거프린트 코드란 코드가 평균화 공격을 받은 후에도 공격에 참여한 공모자들을 추적할 수 있는 코드이다. 본 고에서는 평균화 공격에 강인한 핑거프린트 코드를 생성하기 위해, 조합설계이론(Combinatoril Design Theory)의 한 분야인 GD-PBIBD(Group Divisible Partially Balanced Incomplete Block Design)를 이용하여 코드를 설계하였다. 조합설계이론은 수학 이론 분야의 하나로써, 오류 정정 코드(Error Correction Code)의 설계와 통계적 실험 계획법(Statistical Design of Experiments)의 근간을 이루는 학문이다. GD-PBIBD는 크기가 v인 모집합 X의 원소들을 크기가 k인 b개의 부분집합으로 주어진 조건에 맞게 분류한다[5].

위의 정의를 이용하여 구체적인 핑거프린트 코드를 생성하고, 생성된 핑거프린트 코드가 어떤 방식으로 동작하는가를 알아보자.

1부터 9까지의 정수를 원소로 하는 모집합 X={1,2,3,4,5,6,7,8,9}가 있다. 주어진 모집합에 대하여 (9,9,

[정의] (v, b, r, k, λ_1, λ_2)-GD-PBIBD

크기가 v 인 모집합 $X=\{x_1,x_2,.....x_v\}$의 원소들은, 아래의 조건을 만족하는, 크기가 k 인 b 개의 부분집합 $B=\{s_1, s_2,.....,s_b\}$으로 분류된다:

1. v 개의 원소를 n 개의 원소를 가지는 m 개의 그룹으로 나눈다.
2. 모든 원소들은 적어도 한 번은 부분집합에 포함된다.
3. 모든 원소들은 총 r 번씩 부분집합에 포함된다.
4. 서로 같은 그룹에 있는 원소들끼리는 λ_1 번씩 부분집합에 포함되면, 서로 다른 그룹에 있는 원소들끼리는 λ_2 번씩 부분집합에 포함된다.

정의 1 GD-PBIBD 정의

3,3,0,1) GD-PBIBD를 생성할 수 있다. 우선, 모집합 X에 대하여 m=3, n=3의 조건으로 그룹핑을 하면,

$$G=\{\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8, 9\}\}$$

와 같은 결과를 얻을 수 있다. 이렇게 그룹으로 분류된 부분집합에 대하여 GD-PBIBD 생성 조건 2~4번을 적용하면 최종적으로 다음과 같은 부분집합의 결과를 얻을 수 있다.

$$B=\{\{1, 4, 7\}, \{2, 5, 8\}, \{3, 6, 9\}, \{1, 5, 9\},$$
$$\{2, 6, 7\}, \{3, 4, 8\}, \{1, 6, 8\}, \{2, 4, 9\}, \{3, 5, 7\}\}$$

최종적으로 위와 같이 분류된 부분집합 s_1~s_9에 대하여, 다음의 과정을 통하여 핑거프린트 코드 행렬 M을 생성한다.

$$m_{ij}=\begin{cases}0 : \text{만약에 } i \text{ 번째 원소가 } j \text{ 번째 블록에 포함될 때,}\\ 1 : \text{그렇지 않을 때.}\end{cases}$$

그림 3-(a)의 행렬 M에서 각 행은 모집합 X의 원소를 나타내고, 각 열은 구매자에게 배포할 수 있는 핑거프린트 코드이며, 위의 코드 행렬은 두 명의 공모자들이 평균화 공격을 하였을 때에도 공모자들을 추적할 수 있는 핑거프린트 코드이다. 위의 코드를 살펴보면, 각각의 구매자에게 배포되는 코드는 서로 같은 것이 없어서 구매자들을 구별할 수 있다. 또한, 코드 생성 과정 중에서 최종 부분집합에 포함되어 있는 세 개의 원소들 중, 두 원소의 쌍은 집합 b_i에 반드시 한 번씩만 포함된다. 이러한 특성이 코드를 유일하게 구분할 수 있고, 또한 평균화 공격을 받은 후에도 공격에 참여한 코드를 추적할 수 있는 원리이다.

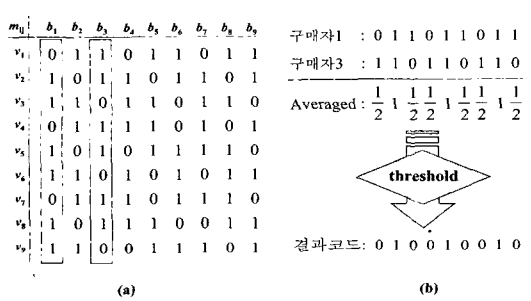

그림 3 9x9 핑거프린트 코드(a)와 평균화 공격된 코드(b)

3.2 평균화 공격과 공모자 추적

앞선 3.1절에서 생성한 핑거프린트 코드를 이용하여 평균화 공격이 일어나는 과정과 공격 후에 공모자들의 코드를 추적하는 방법에 대하여 알아보자[6]. 예를 들어, 구매자 1과 구매자 3이 평균화 공격을 하였다고 가정하자.

그림 3-(b)에서 보듯이 구매자1의 코드는 (0,1,1,0,1,1,0,1,1)이고, 구매자3의 코드는 (1,1,0,1,1,0,1,1,0)이다. 두 구매자들의 코드가 평균화 공격을 받으면 그 결과 코드는 (0,1,0,0,1,0,0,1,0)이 된다.

공모에 가담한 구매자를 찾기 위해서 결과 코드 중에서 1의 위치를 이용한다. 위의 결과에서 두 번째, 다섯 번째, 여덟 번째 자리에 1의 값을 가지고 있다. 이러한 자리의 위치와 핑거프린트 코드 행렬 M을 이용하여 공모자들을 추적한다. 코드 행렬 M에서 각각의 행 중 두 번째, 다섯 번째, 여덟 번째의 행에서 값이 모두 1인 열은 구매자1의 코드와 구매자3의 코드 밖에 없다는 것을 알 수 있다. 따라서 평균화 공모에 참여한 구매자는 1번과 3번이라고 결론을 내릴 수 있다. 이러한 방법으로 평균화 공격된 결과 코드를 알면, 그 공모에 참여한 원래 구매자를 추적할 수 있다. 위에서 언급하였듯이, 코드 행렬 M의 핑거프린트 코드는 아홉 명의 구매자의 코드 중 최대 두 명까지 평균화 공모 공격을 하더라도 원래 구매자를 찾아낼 수 있다.

3.3 비디오 핑거프린팅 시스템

본 절은 3.2절에서 설명한 방법에 의해 생성한 핑거프린트 코드를 실제로 비디오에 삽입 및 추출하는 과정에 대하여 설명한다. 본 절에서 사용하는 핑거프린트 코드는 (72,89,9,8,0,1) GD-PBIBD로써, 코드 하나의 길이는 72 비트이고, 총 89 명의 구매자에게 배포가 가능한 코드이다. 또한 이 코드는 7명까지의 평균화 공모 공격에서도 공모자들을 추적할 수 있는 코드이다. 이 코드의 생성 방법은 3.2절에서 설명한 방법과 동일하다. 본 절에서는 이렇게 생성된 코드를 CIF(352x288 픽셀) 크기의 비디오에 삽입하고, 추출하는 방법에 대하여 알아본다.

3.3.1 코드 변조와 핑거프린트 삽입

3.1절에서 설명한 것과 같이 공모 방지 코드는 0 과 1의 행렬 형태로 표현할 수 있고, 각각의 열이 사용자에게 할당되는 핑거프린트 코드이다. 그러한 코드의 삽입은 각 사용자의 핑거프린트 코드에서 0에 해당하는 직교 신호(Orthogonal signal)를 콘텐츠에 삽입한다. 디지털 워터마킹이나 핑거프린팅에서 직교 신호는, 평균이 0이고 분산이 1인 가우시안 분포($N(0, 1)$)를 따르는 잡음 형식의 신호로써, 삽입하고자 하는 정보를 사람의 눈에 보이지 않는 방식으로 변조하여 콘텐츠에 삽입할 때 쓰인다. 또한 직교 신호를 생성할 때 신호 생성의 보안을 위하여 키(key) 값을 이용한다. 핑거프린트를 삽입하는 과정은 알고리즘 1로 요약할 수 있다.

그림 4는 사용자 3번의 핑거프린트 코드를 비디오에 삽입할 수 있는 신호로 변조하여 핑거프린트 신호를 생

1. 핑거프린트 코드의 길이(비트 수 M)만큼 직교 신호(u_1~u_M)들을 생성한다.
2. 각 사용자의 핑거프린트 코드에 대해 0의 위치에 해당하는 직교 신호들(u_i)를 선택한다.
3. 비디오 신호를 디코딩(Decoding)하여 로우 프레임(Raw frame)을 추출한다.
4. 각각의 로우 프레임에 대하여,
   A. 32×32 블록(Block)으로 분할하고 DCT 변환을 한다.
   B. 선택된 직교신호를 각각의 블록에 견고성과 비인식성을 고려하여 DCT의 중간 주파수 영역(Middle frequency)에 삽입한다[7],[8].
5. 각 블록을 Inverse DCT 변환을 통해 원영상으로 복원한다.
6. 로우 프레임을 인코딩(Encoding)하여 비디오 신호로 변환한다.

알고리즘 1 핑거프린트 삽입 알고리즘

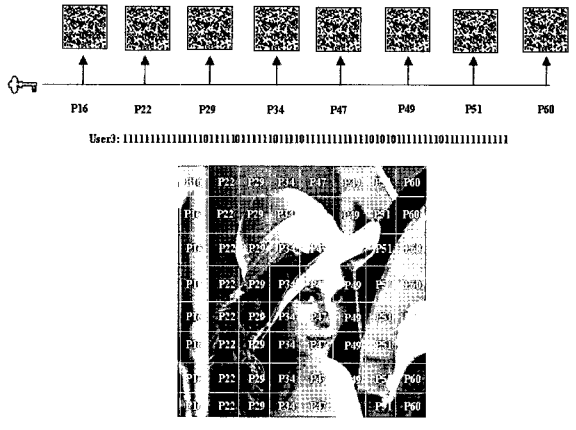

그림 4 사용자 3번의 핑거프린트 신호 생성 과정

성하는 과정을 나타낸 것이다. (72, 89, 9, 8, 0, 1) GD-PBIBD 공모 방지 코드에서 사용자 3번의 코드를 살펴보면 16^{th}, 22^{nd}, 29^{th}, 34^{th}, 47^{th}, 49^{th}, 51^{st}, 60^{th} 값이 0이고 나머지는 모두 1인 코드이다. M 비트의 코드 중에서 0의 위치에 해당하는 직교 신호를 생성하여 핑거프린트 신호를 만든다. 0의 위치에 해당하는 신호만을 생성하여 삽입하는 이유는, 3.2절에서 살펴보았듯이 평균화 공격에 대한 결과 코드는 0의 코드 비트에 의해서 좌우되기 때문에, 핑거프린트 신호를 삽입하고 추출하는 과정에서 0의 위치를 표시하기 위해서이다.

본 고에서는 비디오 핑거프린팅 시스템이므로 압축 공격에 견고한 영역인 DCT (Discrete Cosine Transform) 영역에 신호를 삽입한다. 그림 5는 핑거프린트 신호들을 비디오 영상에 삽입하는 과정을 나타낸다. 디코딩된 CIF 포맷(352x288)의 로우 프레임(Raw frame) 영상에 핑거프린트 신호들을 삽입하기 위하여, 각각의 로우 프레임 영상을 32x32 블록으로 분할하고, 분할된 각각의 영역에 대하여 DCT 변환을 한 다음, 해

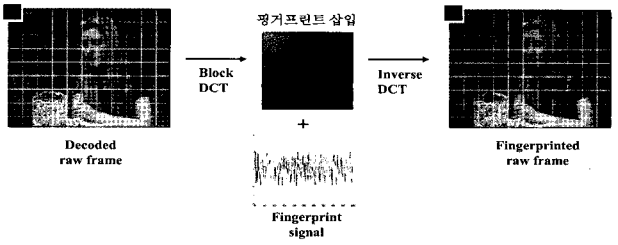

그림 5 핑거프린트 코드 삽입과정

당하는 직교 신호를 각 블록의 DCT 계수(Coefficients)에 더한다. 핑거프린트 신호가 더해진 DCT 계수를 다시 역 DCT 변환을 하여 원래 블록 영역으로 복원하고, 모든 블록에 대하여 이와 같은 처리가 끝난 다음, 다시 비디오 파일로 인코딩(Encoding)을 하면 핑거프린트 신호의 삽입 과정이 완성된다.

3.3.2 핑거프린트 신호 검출 및 공모자 추적

삽입한 핑거프린트 신호를 검출하기 위해서는 콘텐츠에서 핑거프린트 신호를 추정하고, 삽입 과정에서 사용된 직교 신호와의 유사성 비교(Similarity comparison)을 통해서 해당 신호의 존재 여부를 판단한다. 핑거프린트 신호를 검출하는 과정은 알고리즘 2로 요약할 수 있다.

1. 삽입 과정에서 사용한 직교 신호(u_1~u_M)들을 생성한다.
2. M 비트의 결과 코드 비트 값을 1로 초기화한다.
3. 비디오 신호를 디코딩(Decoding)하여 로우 프레임(Raw frame)을 추출한다.
4. 각각의 로우 프레임에 대하여,
   A. 32x32 블록(Block)으로 분할하고 DCT 변환을 한다.
   B. 중간 주파수 영역에 삽입된 직교 신호를 추출한다.(추정된 신호)
5. 핑거프린트로 추정된 신호와 삽입과정에서 사용된 신호 사이의 정규상관관계(Normalized Correlation)를 통하여, 상관 계수가 임계값(Threshold)을 초과하면 결과 코드 비트 값을 0으로 수정한다.
6. 검출된 코드에 대하여 공모자를 추적한다. (3.2절)

알고리즘 2 핑거프린트 검출 알고리즘

그림 6은 공모자 추적을 위해 실제로 콘텐츠에서 공모 코드를 검출한 결과값이다. 검출된 결과값을 보면, 72 비트의 상관관계계수 값 중 16^{th}, 22^{nd}, 29^{th}, 34^{th}, 47^{th}, 49^{th}, 51^{st}, 60^{th}에서의 상관관계 값이 임계값(Threshold)보다 높은 값을 나타낸다. 즉 해당 위치에 0의 값이 포함되어 있음을 의미한다. 따라서 72 비트의 공모 코드 값 중 16^{th}, 22^{nd}, 29^{th}, 34^{th}, 47^{th}, 49^{th}, 51^{st}, 60^{th} 값은 0의 값이 되고 나머지는 모두 1의 값을 갖는다. 검출된 공모 코드를 이용하여 3.2절에서 설명한 방식으로 공모자를 추적하면 구매자3번의 코드임을 알 수 있다.

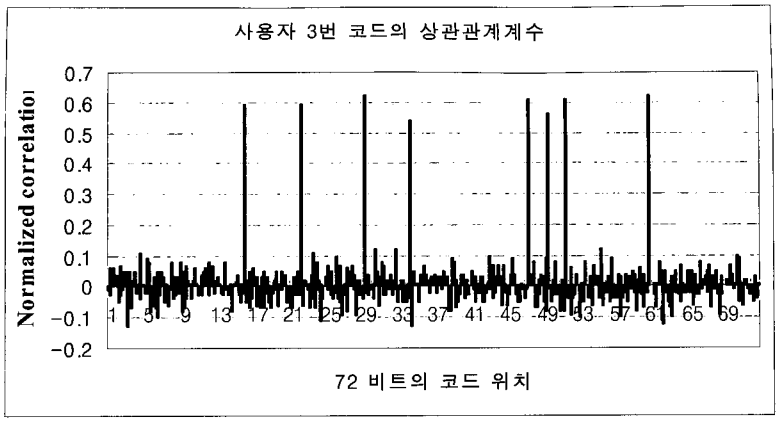

그림 6 핑거프린트 신호의 상관관계계수 추출 값

표 1 공모자 수 증가에 따른 평균화 공격 후 공모자 검출률

공 모	검출된 공모자 수						
	1	2	3	4	5	6	7
1	100	-	-				
2		100					
3			100				
4				100			
5		2	3	3	92		
6			1	3	5	91	
7		4	1	4	1	3	87

3.4 비디오 핑거프린팅의 공격 및 실험 결과

본 실험에서는 공모자 수 증가에 따른 공모자 검출률을 분석한다. 표 1은 100개의 비디오 데이터에 대해 공모자 수(1, 2, 3, 4, 5, 6, 및 7 명)에 따라 평균화 공격을 수행한 뒤 공모자들을 추출하고, 실제 사용된 공모자의 수와 비교를 통해 얻은 결과이다. 결과를 살펴보면 4명의 공모자까지는 해당 공모자들을 모두 검출했으나 공모자의 수가 5명에서 7명으로 증가할수록 검출률이 떨어지는 것을 알 수 있다. 공모자의 수의 증가에 따라 검출이 저하되는 이유는 공모자의 수가 증가할수록 평균화 공격에 의해 상관관계 값의 차이가 줄어들고, 특히 MPEG-4 동영상 압축을 통한 데이터 손실이 발생하였기 때문이다. 그러나 7명의 경우에도 최소 2명의 공모자는 찾아낼 수 있으며 85% 이상 모든 공모자를 찾는 결과를 볼 수 있다. 즉, 핑거프린팅 시스템의 요구사항인 최소 한 명 이상의 공모자를 올바르게 검출할 수 있음을 확인할 수 있다.

4. 요약 및 결론

최근 들어 인터넷 환경의 발달 및 고화질의 콘텐츠 공급과 함께 콘텐츠 보호의 중요성이 증대하고 있으며 향후에 더 많은 보안 기술이 발달할 것으로 예상된다. 이는 콘텐츠 보호가 더 이상 있으면 불편하고 없으면 그만인 수동적인 요소가 아니라 반드시 적용하여야 할 적극적인 기술이라는 의미이다. 콘텐츠 시장이 점점 고화질, 고품질의 제품을 생산하는데 반해 아직까지 보호 기술은 그에 미치지 못하고 있다. 최근 들어, 디지털 시네마의 도래와 함께 이에 대한 심각성은 더욱 부각되고 있는 현실이다.

디지털 핑거프린팅 기술은 콘텐츠의 불법적인 사용 및 유통을 막기 위한 기술이자 콘텐츠의 지적 재산권을 보호하기 위한 적극적인 기술이다. 콘텐츠 속에 콘텐츠를 구매하는 구매자 고유의 정보를 삽입하여 추후에 일어날 수 있는 불법 유통이나 불법 복사, 의도되지 않은 사용(Unintended usage)을 방지하는 기술이다.

본 고에서는 그러한 디지털 핑거프린팅에 필요한 기술 요소에 대하여 알아보았고, 구체적으로 평균화 공격에 강인한 핑거프린트 코드를 사용하여 비디오 핑거프린팅 시스템의 구성에 대하여 살펴보았다. 디지털 핑거프린팅에 관한 기술이 최근까지 많은 연구가 이루어졌으나, 여러 가지 불법 유통 요인을 방지하는 기술이나 핑거프린트가 삽입된 콘텐츠에 대한 의도적인 공격에 대하여 완벽한 해결책이 제시되지 않는 등 여러 가지 풀어야 할 문제점들을 가지고 있다. 그러나 이러한 문제들과 관련해서 계속적인 연구가 진행되고 있으며, 따라서 이에 대한 해결책이 곧 제안될 것으로 기대한다.

참고문헌

[1] B.Chor, A. Fiat, M.Naor, and B.Pinkas, "Tracing traitors," IEEE Trans. Inform. Theory, vol. 46, pp. 893-910, May, 2000.

[2] H. S. Stone, "Analysis of attacks on image

watermarks with randomized coefficients," NEC Res. Inst., Tech. Rep. 96-045, 1996.
[3] H. Zhao, M. Wu, Wang Z.J., Liu, K.J.R., "Nonlinear collusion attacks on independent fingerprints for multimedia," Proc. ICME 2003, vol. 1, 2003.
[4] I.Cox, J.Kilian, F.Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing, vol.6, pp. 1673-1687, Dec. 1997.
[5] Willard H. Clatworthy , "Tables of two-associate-class partially balanced designs," National Bureau of Standards, Washington D.C, U.S. 1973.
[6] W. Trappe, M. Wu, Zhen Wang, and K.J.R. Liu, "Anti-collusion Fingerprinting for Multimedia," IEEE Trans. On Signal Processing, vol. 51, 2003, pp. 1069-1087.
[7] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT domain system for robust image watermarking," Signal Processing, vol. 66, 1998, pp. 357-372.
[8] C.Podilchuk and W.Zeng, "Image adaptive watermarking using visual models," IEEE J. Select. Areas on Communications, vol 16, pp.525-540, May 1998.

강 인 구

2002 건국대학교 컴퓨터공학과(학사)
2004 한국과학기술원 전산학과(공학석사)
2004~현재 한국과학기술원 전산학과 박사과정
관심분야 : 콘텐츠 정보 보호(DRM, 디지털워터마킹, 디지털핑거프린팅 등)
E-mail : ikkang@mmc.kaist.ac.kr

이 흥 규

1979 서울대학교 전자공학과(학사)
1981 한국과학기술원 전산학과(공학석사)
1984 한국과학기술원 전산학과(공학박사)
1984~1986 Univ. of Michigan 연 구원
1998~1999 정보통신부 바이러스해킹센터 보안기술 연구부장
1999~현재 한국과학재단지정 첨단정보기술연구센터 부소장
1986~현재 한국과학기술원 전산학과 교수
관심분야 : 디지털워터마킹/핑거프린팅, 정보은닉, Digital Right Managements
E-mail : hklee@mmc.kaist.ac.kr