+ 김형중 · 고려대학교 정보경영공학부 교수

TECH&TREND

# 콘텐츠보호 기술의 현황과 전망 : 워터마킹 기술

7월호에서는 콘텐츠보호 비즈니스 모델에 대해 살펴보았다. 아날로그 시대의 저작권보호 개념을 디지털 시대에 그대로 적용하려 하기 때문에 문제가 있음을 지적했다. 당연히 저작권보호 분야에서도 새 술을 담을 새 부대가 필요하다. 지금까지 알려진 저작권보호 기술들이 나름대로 제 기능을 수행했지만 새로운 환경에 더 적합한 기술을 부단히 개발할 필요가 있다. 세상은 더욱 디지털적으로 진화하고 있으므로 다가올 미래에 대해 정확히 예측하고 거기에 맞는 새로운 기술 개발을 모색해야 한다.

## 3D 영화의 시대

소형 캠코더 덕분에 극장에서 몰래 영화를 녹화할 수 있게 됐다. 오디오는 더빙한다. 그러면 한 편의 2D 영화는 간단히 복제된다. 다소 질이 떨어지기는 하지만 그렇게 복제된 영상이라도 돌려보기 원하는 사람이 있다. 어차피 큰 돈 내고 보는 게 아니므로 화질이 다소 떨어지는 것까지 탓하지 않는다. 이런 불법복제가 영화를 만드는 제작자의 입장에서는 커다란 골칫거리였다.

그런데, 3D 영화는 다르다. 왼쪽 눈과 오른쪽 눈에 보이는 영상을 약간씩 다르게 만듦으로써 3D 영화가 만들어진다. 그래서 3D 영상을 보통의 2D 캠코더로 찍으면 영상이 보기 흉하게 된다. 보기에 역겹고 피로감을 느끼게 된다. 그래서, 굳이 그렇게 복제된 영화를 볼 사람은 많지 않다. 할리우드는 3D 영화에 큰 기대를 걸고 있는 것이 그런 이유이기도 하다. 제임스 카메론이 감독한 아바타의 성공으로 3D 영화의 가능성을 확인해 한껏 고무된 탓도 있겠지만 2D 캠코더로 불법복제하기 쉽지 않다는 점도 긍정적인 요인으로 받아들여지는 분위기이다.

3D 영화는 불법복제를 어렵게 만들 수 있는 새로운 패러다임을 제시하고 있다. 기존의 불법복제 방지 기술과 달리 새로운 패러다임이 요구되던 시점에서 3D 영화는 새로운 가능성을 열었다고 볼 수 있다. 많은 불법복제물의 출처가 캠코더를 이용한 무단 녹화와 CD 또는 DVD에서 리핑(ripping)하는 것인데, 3D 영화를 복제할 수 있는 새로운 캠코더가 발명되기 전까지는 유요한 해법이 될 수 있다. 다만, 언젠가는 그런 캠코더도 만들어질 것이라는 불길한 예측 때문에 안심하고 있을 겨를이 없다. 또한, 모든 영상이 3D로 만들어지지도 않을 것이다. 컬러 카메라가 나온 지 오래 되었으나 여전히 흑백 사진이 만들어지고 있듯이 2D는 2D 나름의 가치가 있기 때문이다. 게다가 3D 영화를 굳이 2D 영화로 만드는 공격도 있다. 불법복제 방지가 생각보다 어렵다는 것은 바로 이런 집요한 공격에 쏟는 열정과 노력 때문이다.

## 정보보호 기술과 CIA

정보보호 분야에서는 CIA가 매우 중요한 역할을 수행한다. CIA는 미국의 중앙정보국을 의미하는 약자이기도 하지만 기밀성(confidentiality), 무결성(integrity), 가용성(availability)을 줄인 말이기도 하다. 기밀성이란 정보의 내용 자체를 알지 못하게 만드는 성질을 의미한다. 암호 기술이 주로 여기에 이용된다. DES, AES 등이 널리 사용되고 있다. 컴퓨터가 빨라져 DES는 쉽게 공격을 당할 수 있다는 이유로 점점 현장에서 사라지고 있다. 무결성은 정보의 내용이 변경되지 않게 만드는 성질을 말한다. 디지털 해쉬(digital hash) 기술이 여기에 주로 이용된다. SHA-1, MD5 등이 널리 이용되고 있다. MD5도 생각보다 쉽게 깨질 수 있다는 중국 연구자의 보고가 있어 조심스럽게 대하고 있다. 가용성이란 정보가 필요할 때 바로 이용할 수 있게 만드는 성질을 의미한다. 요즘 인구에 회자되는 분산 서비스거부(DDoS) 공격이 필요할 때 서버 이용을 막음으로써 가용성을 낮춘다. 서버의 고장이나 접속 폭주에 대비해 미러 사이트를 만드는 것도 가용성을 높이는 방법 가운데 하나이다.

방송 시스템도 궁극적으로는 서버에 의존하는 온라인 시스템으로 진화할 가능성이 높다. 다운로드 서비스, 온라인 쇼핑, 소셜 네트워킹 등이라면 더욱 그렇다. 그렇다면 당연히 CIA에 대해 더욱 고려해야 한다. 고객 정보의 기밀성을 유지하는 것은 당연하다. 고객의 신상정보는 물론이고, 고객의 취향, 고객의 접속 이력이나 다운로드 이력을 보호하지 않으면 고객은 미련 남기지 않고 바로 떠난다. 새로 만든 드라마 영상의 기밀성을 보호하는 것도 당연하다. 힘들여 만든 영상이 해킹으로 뚫려 무단으로 복제되도록 할 수는 없는 노릇이다. 또한, 중요한 데이터의 무결성을 유지하는 것도 필요하다. 영상의 태그나 영상의 일부가 변조되도록 방치할 수 없다. 지금까지는 그런 공격이 없었다고 안심할 수는 없다. 가용성도 높여야 한다. 오프라인 시스템에서도 CIA가 중요한데 온라인 시스템이 된다면 그 중요성은 더 커진다. 그러나, CIA를 고려한 설계는 필연적으로 비용문제 및 시스템 성능과 직결된다.

영상 원본은 강력한 암호 기술을 이용해 보호해야 한다. 그렇다고 해도 영상은 암호가 풀린 후 달리 보호할 방법이 없다. 암호가 풀린 영상을 악의적으로 불법복제해서 유포해버리면 그 이후로는 원본의 보호가 사실상 기술적으로 불가능하다. 물론, 법률적 보호 장치가 남아있지만 소 잃고 외양간 고치는 격이 될 수 있다. 그래서, 불법복제한 곳을 추적하기 위해 핑거프린팅(fingerprinting) 기술을 적용하기도 한다. 핑거프린팅 기술은 기본적으로 원본 배포처마다 조금씩 다른 영상을 주고 그 영상의 지문에 해당하는 정보를 보관하고 있다가 불법복제 영상이 유포되었을 때 거기서 다시 지문을 채취해 대조해봄으로써 불법으로 유포한 곳을 찾기 위한 증거로 사용하는 기술을 말한다. 오늘날 포렌직(forensic) 워터마크라고 부르는 기술들이 주로 핑거프린팅 기술에 의존하고 있다.

그러나, 온라인에서 서비스하는 영상까지 모두 강력하게 보호할 필요는 없다. 일반적으로 온라인 서비스용 영상은 원본보다 화질이 떨어지고 크기도 작게 만든다. 아무리 네트워크 속도가 향상되었다고 해도 여전히 전송속도는 문제가 된다. 그래서, 영상을 압축하게 되면 파일 크기가 작아져 전송시간을 크게 줄일 수 있다. 대신 화질이 떨어지는 것은 감수해야 한다. PC나 모바일 단말에서 즐길 영상이라면 화면 크기도 원본처럼 클 필요가 없다. 물론, 온라인 버전도 불법으로 유출될 경우 피해가 클 수 있지만 원본만큼 심각하지는 않다. 그래도 핑거프린팅 기술이나 암호 기술을 적용할 수 있다.

그런데, 온라인 서비스용 영상의 경우 스트림 전체에 암호 기술을 적용하는 대신 영상의 일부에만 암호 기술을 적용하는 부분암호(partial encryption) 기술을 선호하는 경향이 있다. 영상 전체에 암호 기술을 적용할 경우 스트림암호 기술도 있으나 널리 채택되고 있지 않아 일반적으로 블록암호 기술을 선호한다. 이때 MPEG 스트림의 일부인, 예를 들어 128비트마다 블록암호 기술을 적용하게 된다. 블록암호 기술을 적용하게 되면 입력 스트림 길이와 출력 스트림의 길이가 같아지는 장점이 있다. 그러나, 블록암호의 경우 전송 도중 한 비트라도 에러가 생기면 그 블록은 해독이 불가능하게 된다. 그리고, 블록마다 서로 연계시킬 경우 한 비트의 에러가 자체 블록에만 영향을 미치는 게 아니라 다른 블록으로도 전파되어 스트림의 상당한 분량을 쓸 수 없게 만들 우려가 있다. 블록암호에서는 이전 블록의 비트 일부를 가져다가 다음 블록을 암호로 변환할 때 씨드(seed)로 사용하므로 에러가 전파될 수밖에 없다. 그런데, 스트림암호의 경우 에러 전파는 없고 해당 비트의 에러만 에러로 남는 장점이 있다.

## 부분암호(partial encryption) 기술

부분암호는 일부 신호에 대해서만 암호 기술을 적용하므로 우선 계산시간이 줄어드는 장점이 있다. 또한, 영상의 중요한 부분만 보기 어렵게 만들고 전체적으로 영상의 개론은 알 수 있게 한다. 영상의 내용을 완전히 알 수 없게 만들면 영상을 보고자 하는 호기심 자체가 말살시킬 수 있다. 그런데, 부분암호가 걸리면 영상 내용은 대충 알 수 있으나 화질이 떨어져 볼 수 없기 때문에 호기심이 도져 그 영화를 선택할 가능성이 높아진다. 암호 기술을 적용했다 해서 스트림 길이가 늘어나면 대역폭의 크기가 정해진 채널에서는 전송 문제가 생길 수 있으므로(즉, 스트림의 일부는 전송할 수 없거나 전송이 지연되는 현상이 발생하므로) 스트림의 길이가 늘어나지 않게 해야 한다. 부분암호 기술 역시 스트림의 길이를 일정하게 유지할 수 있다.

예를 들어, MPEG-2에서 양자화된 DC 계수를 코딩할 경우 -1과 1은 같은 카테고리에 속하므로 여기에 속한 계수 값은 일정한 규칙에 의해 바꿔치기할 수 있다. 또한, 양자화된 계수 -3, -2, 2, 3도 같은 카테고리에 속하므로 일정한 규칙에 의해 교환할 수 있다. 이런 카테고리는 여럿이 있으므로 동일한 카테고리 안에서의 계수를 바꿔치기하는 것은 얼마든지 가능하다. 이런 방법을 쓰면 스트림 길이가 늘어나지 않게 하면서 영상도 변화시킬 수 있다. 원래 영상으로 복구하는 데도 많은 시간이 걸리지 않는다. 그렇지만 이런 점은 공격자에게도 좋은 소식이 된다. 공격자도 약간의 노력만으로 원본을 복구할 수 있기 때문이다. 그래서, 더 정교한 부분암호 방법이 만들어지고 있다. 정교한 기술이 만들어진다는 것은 공격을 어렵게 하는 장점이 있으나 동시에 디코더에서의 구현도 복잡해짐을 의미한다. 공학에서 절대로 공짜 점심이 없음을 알아야 한다. 또한, 블록암호 기술에서 에러가 다른 블록으로 전파되는 현상이 문제가 되는데 부분암호 기술에서는 에러 전파를 완전히 방지하거나 또는 부분적으로 억제할 수 있다.

## 워터마킹 기술의 현황

워터마킹 기술에서는 공격에 견딜 수 있는 강인성(robustness) 정도를 가지고 기술의 우수성을 따진다. 공격의 종류는 매우 다양한데 영상에 잡음을 강제적으로 집어넣거나, 필터링으로 일부 정보를 솎아 내거나, 영상의 크기를 줄이거나, 영상을 약간 회전시키거나, 색상을 약간 바꾸거나, 압축하는 등 여러 형태의 공격을 고려하고 있다. 이미 앞에서도 밝혔듯이 SDMI가 실패한 이유 가운데 하나는 2000년 당시의 워터마킹 기술이 신통치 않았기 때문이다. 그렇다면 지금 기술은 그때에 비해 크게 나아졌을까? 그렇다는 사람도 있을 것이고 아니라는 사람도 있을 것이다. 기술은 약간 나아졌을지 모르나 상황은 더 나빠졌다고 할 수 있다.

워터마킹 기술은 영상에 존재하는 잉여정보(redundancy)를 활용해서 부가정보를 숨기는 기술을 총칭한다. 영상이 영상일 수 있는 것은 인접하는 픽셀들 사이에 비슷한 부분이 많아서 부드러운 영역이 다수 존재하기 때문이다. 즉, 인접한 픽셀들과의 상관관계(correlation) 값이 높기 때문에 영상은 부드러움을 유지한다. [그림 1]의 상단 네모가 좋은 예이다. 인접 픽셀과의 상관관계가 매우 낮다면 그것은 잡음에 가까운 영상이기 때문에 영상으로서의 가치가 없다. 물론, 영상에는 윤곽선도 존재하고 경계도 존재하기 때문에 항상 상관관계가 높은 것은 아니다. 그럼에도 불구하고 영상에는 충분히 높은 상관관계가 유지되어 영상으로서의 가치와 존재의의를 지닌다.


[그림 1] 공간영역에서 영상의 상관관계와 잉여정보와의 관계

[그림 1]의 하단 네모처럼 윤곽선이나 경계가 존재하는 영역에 주로 워터마크 정보를 숨기려고 한다. 상관관계가 높은 곳에 정보를 은닉하는 것이 기술적으로는 쉽고 안전하지만 그런 영역의 정보는 필터링에 의해 쉽게 제거되는 약점을 지니고 있다. 물론, 필터링에 의해 윤곽선이나 경계도 어느 정도 뭉갤 수 있으나 완전히 뭉개버리면 영상으로서의 가치가 소멸되므로 적절한 수준에서 필터링이 이루어지게 된다. 따라서, 이런 영역에 정보를 숨기게 되면 윤곽선과 마찬가지로 숨긴 정보도 잘 소멸되지 않는 장점이 있다. 게다가 경계가 분명할수록 큰 값의 정보를 숨길 수 있고 또한 정보은닉 여부도 육안으로 구별하기 어려운 장점이 있다. 역으로 [그림 1]의 상단과 같이 평탄한 공간에다 큰 값의 정보를 숨길 경우 확연히 육안으로 구별되는 약점이 드러난다. 그래서, 가능하다면 정보는 경계 부근에다 숨긴다. 즉, 너무 높지도 너무 낮지도 않은 적정한 상관관계가 존재하는 영역에 보통 워터마크 정보를 은닉하려고 한다.

워터마킹도 기본적으로 영상에 존재하는 잉여정보를 활용해서 정보를 숨긴다. 공간영역에 존재하는 잉여정보이든 시간영역에 존재하는 잉여정보이든 잉여정보가 존재해야 정보를 숨길 수 있다. 잉여정보가 없으면 정보를 숨기기가 사실상 불가능하다. 그런데, 한편에서는 잉여정보를 최대한 제거하려는 시도가 이루어지고 있다. 영상압축이 그렇다. 2000년 이전에는 MPEG-2와 MPEG-4가 압축표준의 대세를 이루었다. 그 이후 H.264가 출현했다. 다시 현재는 H.265라 불리는 영상압축 표준화가 이루어지고 있다. H.264 이후의 압축표준은 그나마 남아있던 한 방울의 잉여정보까지 쥐어짤 태세로 덤비고 있다. 그래서, 2000년 당시의 워터마킹 기술을 지금 압축방법으로 공격하면 성능이 한참 뒤진다. 동일한 양의 정보를 숨기고 지금의 H.264나 앞으로 만들어질 H.265에서 압축한 후 성능을 검증하면 찾아낼 수 있는 정보의 양이 크게 줄어든다는 뜻이다. 달리 말하면 숨길 정보의 양을 크게 줄여야 그나마 안심하고 워터마킹으로서의 제 기능을 수행할 수 있게 된다는 뜻이다.

여기서 워터마킹 기술의 딜레마가 시작된다. 압축 기술은 가능한 한 거의 모든 잉여정보를 제거하려 하고, 워터마킹은 잉여정보를 최대한 활용하려고 한다. 그래서, MPEG-2를 상대로 압축에 대한 강인성을 평가하려고 하면 H.264나 H.265에서는 살아남기 어렵게 된다. 워터마킹 기술이 혁신적으로 바뀌어야 할 이유가 바로 여기에 있다. 지금 강인하다고 해서 5년 후 새로운 압축방법에 강인하다고 말할 수 없고 10년 후에는 더더구나 그런 주장을 하기 어렵다.

그럼에도 불구하고 이제 워터마킹 기술은 충분히 개발했으니 더 이상 연구할 필요 없다고 주장하는 것은 기술진보에 대해 몰라도 한참 모르고 하는 말이다. 그리고, 바로 그 점이 강인한 워터마킹의 한계인지도 모른다. 아무리 강인한 워터마킹 기술을 개발한다고 해도 더 뛰어난 압축방법이 개발될 것이므로 압축 기술이 한계에 도달할 시점을 학수고대하며 기다릴 수밖에 없다. 게다가 현재의 공격방법은 복합공격이 아닌 단일공격 위주이며 강인성은 단일공격에 대한 평가에 불과하다. 두 가지 공격만 복합적으로 적용해도 강인성은 매우 낮아진다. 한 가지 공격에도 잘 견디지 못하는데 복합공격을 가하면 상황은 더욱 심각해진다. 게다가 압축은 여러 공격방법 가운데 하나에 불과할 뿐이다.

### 확산대역(spread-spectrum) 워터마킹

워터마킹 기술로 가장 널리 쓰이는 것이 확산대역(spread-spectrum) 기법이다[1]. 이 기술이 Cox 등에 의해 1997년 발표된 이후 많은 곳에서 채택하고 있다. 기본 원리는 스펙트럼을 확산시키는 것이다. [그림 1]의 평평한 지점의 한 픽셀에다 50이라는 큰 값을 더하면 누구라도 그 부분이 이상하다는 것을 안다. 정보를 숨긴 사람 관점에서는 그 곳의 값이 주변 픽셀보다 월등히 높기 때문에 정보를 은닉했음을 알 수 있고 숨긴 정보도 알아낼 수 있다. 주변 픽셀 값보다 50 정도 크면 숨긴 정보가 1이고 50 정도 적으면 0이라고 정했다고 치자. 문제는 공격하는 사람 입장에서도 주변 픽셀보다 값이 월등히 크거나 작은 영상을 찾으면 의심하게 되고 그래서 필터링해버리면 그런 정보는 사라져버리게 된다. 애써 정보를 숨겼는데 간단한 공격으로 숨긴 정보가 사라져버리는 것이다. 그래서, 누구 눈에나 확연히 드러나게 정보를 숨기는 것은 어리석은 일이다.
대안은 50이라는 큰 값. 즉, 큰 에너지를 확산시키는 것이다. 확산시키는 방법은 이렇다. 길이가 50쯤 되는 이진 난수를 발생시켜 이것을 영상에다 숨기는 것이다. 여기서는 편의상 50개의 난수가 아닌 21개의 난수로 예를 들었다. 예를 들어, 이 난수가

1 1 −1 1 −1 −1 1 −1 1 −1 −1 −1 1 −1 1 1 1 1 −1 1 −1 (1)

이라고 하자. 50이라는 큰 값 대신 길이가 21인 난수를 발생시켜 에너지를 확산시킨 것이다. 에너지를 확산시킴으로써 영상에 대한 변형을 최소화할 수 있게 된 셈이다. CDMA 통신에서 사용하는 스펙트럼 확산과 비슷한 개념을 이용하는 셈이다. 원래 CDMA 기술이 출현한 것도 스펙트럼 에너지를 확산시켜 통신내용의 감청을 어렵게 하려는 의도에서였다. 바로 이 점이 확산대역 워터마킹이 지닌 장점 가운데 하나이다. 이 난수를 영상의 원본 픽셀 값

50 51 50 49 50 50 51 52 49 48 60 60 61 59 62 60 61 62 61 60 59 (2)

에다 더한다. 그러면 정보를 은닉한 영상의 픽셀 값은

1) I. J. Cox, J. Kilian, F. T. Leighton, T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Transactions on Image Processing, vol. 6, no. 12, pp. 1673–1687, 1997.

51 52 49 50 49 49 52 51 50 47 59 59 62 58 63 61 62 63 60 61 58 (3)

을 얻게 된다. (2)와 (3)을 비교해보면 알 수 있듯이 변형된 값이 원본과 거의 차이가 나지 않는다. 따라서, 워터마크 정보가 숨겨져 있는지의 여부를 알기조차 어렵게 된다. (2)의 신호를 보통 x, (1)을 w, (3)을 y라고 부른다. 이 식을 다시 쓰면 y = x + aw로 나타낸다. 여기서 a는 워터마크 신호의 강도를 나타내는 계수이며, 이 예에서는 그 값이 1이다.

그렇다면 과연 이렇게 해서 숨긴 정보를 알아낼 수 있을까? 1997년 Cox 등이 알아낸 방법은 원본픽셀 값 (2)와 워터마크 정보 (1)을 가지고 있다는 전제 아래 워터마킹 픽셀 값 (3)을 가지고 있을 경우 (3)에서 (2)를 빼면 (1)이 구해진다는 것이다. 그렇다면 (1)을 가지고 있으므로 (3)에서 (2)를 빼 얻은 (1)과 비교해보면 자신이 숨긴 워터마크인지 아닌지 알 수 있다. 당연히 (1)과 (1)을 비교하면 같다는 것을 금방 알 수 있다. 여기서는 (1)의 길이가 21이므로 (1)과 (1)의 내적을 구하면 21이 된다.

이 방법에서는 워터마크의 진위여부를 파악하기 위해 (1), (2), (3)을 다 가지고 있어야 한다는 모순이 생긴다. 여기서 (2). 즉, 원본을 가져야 판별할 수 있으므로 비효율적이라 해서 2000년대 이후에는 크게 의미를 부여하지 않는다. 1997년 이론이 처음 발표된 이후 잠시 스포트라이트를 받았으나 역사의 뒤안길로 묻혀가고 있다. 그럼에도 불구하고 아직도 워터마킹을 처음 시작하는 초보자들은 여전히 이 방법을 신주단지 모시듯 한다.

여기서 문제를 약간 비틀어보자. 디코더는 (1)과 (2)를 가지고 있는데 전송 도중 (3)에 변형이 가해졌다고 가정하자. 예를 들어, 악의적인 공격이 가해져 워터마크를 제거하기 위해 공격자도 (3)에다가 자신이 만든 이진 난수를 더했다고 하자. 그러면 디코더는 (3) 대신 변형된 (3*)를 지니게 된다. 디코더는 (3*)에서 (2)를 뺀 (1*)을 얻게 된다. 확산대역 방법의 진가는 여기서 발휘된다. 이제 (1)과 (1*)는 분명히 다르다. 그런데 놀라운 것은 (1)과 (1*)의 내적을 구하면 대부분 크기가 21에는 미치지 못하지만 큰 값이 나온다는 사실이다. 만일 디코더가 자신이 워터마크를 넣지 않은 영상의 (3*)를 가지고 워터마크를 찾으려 한다면 무슨 현상이 생길까? 당연히 (3*)에서 (2)를 빼 얻은 (1*)와 (1)의 내적을 구하면 놀랍게도 이 경우에는 0 근처의 값이 얻어진다는 점이다. 여기서 중요한 사실은 (1)이 진정한 난수라야 한다는 점이다. 진정한 난수는 다른 난수와 상관관계가 수학적으로 0이다. 그런데, 현실에서는 약간의 에러가 포함된다. Cox 등의 논문이 중요한 이유는 그들이 확산대역 워터마킹에 필요한 통계적 이론을 확립했다는 데 있다. 길이가 긴 난수를 사용해도 난수가 완전한 난수가 아닐 수 있고 공격자가 삽입한 난수의 영향을 받을 수도 있어 통계학적으로 유의한 수준에서 판단을 내릴 수 있는 기초를 제공한 점이 그들의 공로라고 할 수 있다.

## 원본이 필요 없는 워터마킹 기술

비록 Cox 등이 제안한 기술이 통계학적으로 중요한 의미를 지닌다고 해도 원본이 필요하다는 약점이 있었다. 그래서, 원본이 필요 없는 기술이 필요했다. 원본이 없으므로 (3*)에서 (2)를 빼는 일은 할 수 없다. 그래서, 등장한 방법이 아예 직접 (1)과 (3*)의 내적을 구하는 방법이다. 만일, (3*)가 평활한. 즉, 상관관계가 높은 픽셀들로 구성되어 있다면 내적 값은 양수가 될 것이라는 가정에서 출발한다. 디코더가 얻은 신호는 $y^* = x + aw^*$인데 여기서 전송도중 공격을 받았다면 w 대신 w*가 얻어진다. 여기에 직접 w와 내적을 구하면 $y^*w' = xw' + aw^*w'$가 얻어진다. 그런데 이론적으로 xw'는 서로 상관관계가 낮으므로 0에 가까운 값이 얻어질 것이고 w*w'는 큰 양수가 될 것이라는 가정에 따르면 y*w'의 값은 a의 부호에 의해 결정될 것이라는 가정이 성립한다. 이것이 Hartung과 Girod의 연구성과[2]이다. 만일, (3*)가 평활하지 않다면? 필터링을 통해 불필요한 정보를 제거하고 내적을 구하면 된다. 그래서, 이후 불필요한 정보를 제거하는 기술이 많이 개발되었다. 불필요한 정보를 제거하는 것보다 쉬운 길은 상관관계가 높은 평활한 공간에 정보를 숨기는 것인데, 달리 말하면 낮은 주파수에 정보를 숨기는 것과 같은 효과가 있다. 오늘날 대부분의 워터마킹 기술은 낮은 주파수 영역에 정보를 숨긴다.

오늘날 사용되는 거의 모든 워터마킹 기술은 원본을 요구하지 않는다. 원본이 없기 때문에 진위여부 판단에 다소 오류가 발생할 수 있으나 그 오류도 크게 낮출 수 있게 되었다. 그렇지만 공격에 대한 강인성을 높이는 방법을 강구하는 것은 여전히 풀기 어려운 숙제이다. 내일 새로운 공격방법이 개발된다면 거기에 견딜 수 있는 새로운 방어 기술을 개발하는 데 상당한 시일이 소요되기 때문이다. 마치 제로데이(zero day) 공격과 같은 현상이 나타난다. 운영체제의 허점을 보완하는 패치가 나오기 전 공격하는 것을 제로데이 공격이라 하는데 보안에서는 어디서나 있는 일이다.

워터마킹 기술은 확산대역 방법 외에도 여러 가지가 있다[3]. 특별히 오디오에서는 에코를 집어넣는 방법이 있고 켑스트럼(cepstrum)을 이용해서 찾아내는 방법이 있다. 두 집합의 차이를 이용해서 정보를 숨기고 찾는 방법이 있다. 패치워크(patchwork) 방법이 대표적인 예이다. 히스토그램을 이용해서 숨기는 방법도 있다. 다만 이 공격은 히스토그램 평활화 공격에 약하다는 단점이 있고 숨길 수 있는 용량도 높지 않다. 그러나, 거의 대부분 여전히 확산대역 기법을 널리 쓴다. 성능평가에서는 강인성도 평가하지만 그보다 우선 실제로 인간의 인지능력을 활용한 실전 테스트를 중시한다. 매우 섬세한 골든이어(golden ear) 또는 골든아이(golden eye)에게 원본과 차이가 있는지를 평가하도록 하는데 매우 민감한 음악이나 영상에서는 대부분 워터마크로 인한 잡음이 쉽게 확인되어 워터마크 무용론의 빌미를 제공하곤 한다. 그러나, 보통의 영상에서는 워터마킹 여부가 잘 감지되지 않는다.

워터마킹 기술은 그럼에도 중요한 기술로 자리잡아가고 있다. 강인한 워터마킹 기술은 저작권을 보호하는데 널리 쓰이고 있고, 핑거프린팅 정보를 숨기는 데도 유용하게 쓰이고 있다. 문서의 진본 여부를 판단하기 위해서도 이미 공문서 출력에 한국에서는 널리 쓰이고 있다. 물론, 여기에 대해서도 할 이야기는 많지만 본질에서 벗어나는 문제라 넘어가기로 하자.

2) F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," Signal Processing, vol. 66, no. 3 (Special issue on Watermarking), pp. 283–301, May 1998.
3) W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," IBM Systems Journal, vol. 35, no.3/4, pp.313–336, 1996.

그런데, 꼭 강인한 워터마킹 기술만 쓸모가 있는 것은 아니다. 가역적 워터마킹 기술도 주목할 필요가 있다. 가역적(reversible) 워터마킹이란 숨긴 정보를 찾아낸 후 훼손된 원본도 원래대로 복구하는 기술을 말한다. 정보를 숨긴다는 것은 원본을 훼손하는 행위로부터 출발한다. 강인한 워터마킹은 숨긴 정보만 찾아내고 훼손된 원본은 그대로 방치하는 기술이다. 그래서, 강인한 워터마킹 기술은 비가역적 워터마킹 기술이라고 부른다. 가역적 워터마킹 역사도 그리 길지는 않다. 최초로 쓸 만한 기술이 출현한 것은 2003년 Tian이 발표한 차이확장(difference expansion) 방법[4]과 히스토그램이동(histogram shifting) 방법[5]이 주류인데 최근에는 이 둘을 통합한 기술[6]이 대세로 자리잡아가고 있다.

영상의 기밀성이 중요하다지만 일단 암호가 풀리면 영상을 더 보호할 길이 없다. 그래서, 영상에 핑거프린팅 정보를 숨길 때 강인한 워터마킹 기술을 적용한다. 무결성을 보장하기 위해 영상의 요약정보(hash)를 계산해서 그 정보를 가역적 워터마킹 기술로 영상에 숨길 수 있다. 요약정보를 뽑아내고 원본도 복구하면 영상이 변조되었는지의 여부도 알 수 있게 된다. 기술적으로 CIA의 C와 I에 해당하는 두 기술은 워터마킹 기술로 보완할 수 있다. A에 해당하는 기술은 잉여 장비로 해결해야 할 문제이므로 워터마킹과는 무관하다.

## 결론

워터마킹 기술은 저작권 보호에 큰 기여를 할 수 있다. 그러나, 강인한 워터마킹 기술은 앞길이 험난하다는 점을 기억해야 한다. 날로 새로운 공격이 개발되고 있고, 특히 더 진보된 압축 기술이 만들어지고 있어 현재의 강인성을 계속 유지하기 어렵다는 점에 주목해 기술개발을 멈추지 말아야 한다. 지난 10년간 워터마킹 기술에 얼마나 많은 돈을 쏟아 부었는데 또 아직도 워터마킹 타령이냐고 말하면 곤란하다. 방패가 얼마나 좋아졌는지 모르면서 아직도 창 타령이나 하고 있냐고 말하는 것과 같은 이치이다.

또한, 가역적 워터마킹 기술이 무결성 보완을 위해 널리 쓰일 가능성이 높아지고 있다. 이전에 고려하지 못했던 새로운 기술영역이 출현한 셈이다. 워터마킹 기술은 여전히 발전 가능성이 높은 연구영역임에 틀림없다. 그러나, 이 기술을 반드시 저작권 보호에만 국한해서 사용할 일도 아니다.

이 글은 문화체육관광부 및 한국콘텐츠진흥원의 2010년도 문화콘텐츠산업기술지원사업의 연구결과로 수행되었다.

4) J. Tian, "Reversible data embedding using a difference expansion," IEEE Transaction on Circuits and Systems for Video Technology, vol. 13, no. 8, pp. 890–896, 2003.
5) Z. Ni, Y. Q. Shi, N.Ansari, and S. Wei, "Reversible data Hiding," IEEE Proceedings of the 2003 International Symposium on Circuits and Systems, vol. 2, pp. II-912–II-915, Thailand, May 2003.
6) D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 721–730, Mar. 2007.