論文
5-7-1

핑거프린팅을 이용한 디지털 원문 보호시스템 설계 (Design of Digital fulltext Protection System using Digital Fingerprinting)

김상국(Sang-kuk Kim)[1] 이재광(Jae-kwang Lee)[2]

Abstract

It is one of important techniques to protect illegal uses and forgeries for text databases when the internet has become widespread. Lots of studies and system developments has been progressed in order to protect text databases, important documents of the companies, and blueprints of the research institutes. Given this project securing the KISTI (Korea Institute of Science and Technology Information) digital contents, we are to design a watermarking and fingerprinting system protecting document images, and a agent notifying the traitors' location.

1) 정회원 : 한국과학기술정보연구원

2) 정회원 : 한남대학교 컴퓨터공학과 교수

논문접수 : 2004. 7. 6.
심사완료 : 2004. 7. 25.

1. 서 론

멀티미디어 기술의 발달과 기반 통신 시설의 발달에 따라 네트워크를 통한 디지털 데이터의 유통이 빈번해 지고 있다. 하지만 네트워크를 기반으로 한 콘텐트 시장이 활성화되기 이전에 콘텐트가 의도된 목적과 콘텐트에 대한 합법적이고 정당한 권리를 산 사용자에 의해서만 유통되어야 하는 메커니즘이 먼저 확고하게 놓여야 한다. 그러한 메커니즘을 위해서 최근 들어 DRM(Digital Right Management)이나 디지털 워터마킹 같은 기술들이 많이 연구되고 있다. 디지털 핑거프린팅은 디지털 워터마킹 기술을 기반으로 한 새로운 형태의 멀티미디어 저작권 보호의 한 방법이다.

이런 디지털 콘텐트를 보호하기 위한 기술적인 방법은 크게 3가지로 나눌 수 있다. 첫 번째는 권리가 없는 사용자에게 콘텐트 접근을 막는 접근제어방법, 두 번째는 암호화나 비밀키, 공개키를 이용하여 정당하지 않은 사용자의 콘텐트 사용을 막는 방법, 그리고 마지막으로 콘텐트 내용에 보이지 않는 정보를 심어 저작권을 보호하고 불법적인 복제를 막는 내용제어이다. 오래 전부터 연구되어 오던 것 중 하나는 암호화를 통한 방법이나 이는 일단 암호가 풀리게 되면 본문의 내용은 공격에 그대로 노출되게 되므로 저작권이나 진품여부를 판명하기에는 불충분하다. 워터마킹 기술은 내용제어에 해당하는 것으로 이미지, 오디오, 동영상 등 디지털 콘텐트에 사람이 인식하지 못하는 신호를 넣어서 소유권을 주장하고자 하는 특정 데이터를 보장해주고 임의로 콘텐트에 대해 조작을 못하게 하는 기술이다. 워터마킹 기술은 원본 데이터를 인증하고 무결성을 증명하며 무엇보다도 디지털 콘텐트의 최대 약점인 다수의 복사물에 대해 출처를 밝히고 진품여부와 원본의 소유자가 누구인지를 밝혀낼 수 있는 기술이다.

동일한 워터마크(저작권 정보)를 콘텐트에 삽입하는 것과 달리 핑거프린팅 기법은 사용자마다 각기 다른 정보를 콘텐트에 삽입함으로써 불법 복제 및 유통 행위가 발견되었을 때 불법 배포자를 추적하고자 하는 기술이다. 즉, 동일 데이터에 대해 저작권자만이 인지할 수 있는 일련번호를 비밀리에 부여하고, 문제 발생시에 삽입된 일련번호를 제 삼자에게 검증할 수 있는 기능을 제공하는 것이다. 디지털 데이터의 불법복제를 방지하기 위하여 데이터에 2진 마크를 삽입한다는 점에서 핑거프린팅은 기본적으로 워터마킹과 유사하지만, 데이터에 대해 삽입되는 마크의 내용이 모두 다르기 때문에 동일한 데이터 각각을 식별할 수 있는 추가적인 기능이 제공된다는 점에서 디지털 워터마킹과 구별된다. 즉, 디지털 데이터의 소유권에 대한 인증기능만을 제공하는 것이 워터마킹 기술이라면, 소유권 인증 기능뿐만 아니라 식별기능까지 제공하는 것이 핑거프린팅 기술이다. 워터마킹 기술이 저작권자(판매자) 측면에서 자신의 저작권을 증명하기위한 기술이라 한다면 핑거프린팅 기술은 구매자의 측면에서 불법적 구매자를 검출하는 기법이라고 볼 수 있다.[1]-[5]

본 논문을 통하여 위에서 기술한 워터마킹과 핑거프린팅 기술을 이용하여 이미지 형태의 원문 데이터를 보호하는 시스템을 제안한다.

2. 디지털 핑거프린팅 기술

1) 대칭형 핑거프린팅

대칭형 핑거프린팅은 핑거프린팅 프로토콜, 구매자 판별 프로토콜의 두 가지 알고리즘과 구매기록을 위한 데이터베이스로 구성된다. 판매자가 핑거프린팅된 콘텐트를 생성할 수 있다. 핑거프린팅 할 콘텐트와 구매한 사용자의 식별자, 현재까지 판매된 리스트를 입력으로 핑거프린팅을 한다. 그에 대한 결과물로 핑거프린팅된 콘텐트와 구매 레코드가 생성된다. 만약 핑거프린팅된 콘텐트가 어떤 구매자에 의해서 불법 복제 및 배포가 되었다면 판매자는

발견된 복사본과 핑거프린팅 되기 전의 콘텐트, 그리고 구매기록을 입력으로 발견된 복사본의 원구매자를 찾아내게 된다. 하지만 대칭 핑거프린팅은 분배자와 구매자 둘 다 핑거프린팅된 데이터를 알 수 있으므로 불법적 행위시 책임 소재가 모호하다. 대칭적인 특성은 부인방지를 하지 못한다. 그러므로 판매자는 구매자의 잘못을 제삼자에게 증명할 수 없다는 문제점이 있다.

2) 비대칭형 핑거프린팅

대칭형 핑거프린팅 방식의 문제점을 보완하기 위하여 제시된 방법이 Pfitzmann과 Schunter[6]에 의해 소개된 비대칭 핑거프린팅(Asymmetric fingerprinting) 방법이다. 이 방법은 오직 구매자만이 핑거프린트된 데이터를 알 수 있기 때문에 책임소재를 분명히 할 수 있다. 만약 분배자(판매자)가 불법 유통을 확인하였을 경우에 분배자는 불법적 구매자를 검출하고 제삼자(신뢰기관)에게 이를 증명할 수 있다. 이 방식은 네 가지(키 생성, 핑거프린트, 식별, 분쟁) 프로토콜로 구성된다.

- 키 생성 프로토콜 : 구매자는 공개키 pk_B(구매자의 식별정보)와 비밀키 sk_B를 생성하고, 인증 센터를 통해 공개키 pk_B를 공개 디렉토리에 등록한다.
- 핑거프린트 프로토콜 : 핑거프린트 입력으로, 판매자는 cover-object, 구매자 식별정보(pk_B), 데이터설명 문자열(str), 동일 그림에 대한 이전의 판매리스트를 입력한다. 구매자는 구매설명 문자열(str)과 자신의 비밀키(sk_B)를 입력한다. 구매자의 출력은 핑거프린트를 갖는 fingerprinted object이다(or failed). 구매자는 판매후의 논쟁을 위해 판매기록 recordB를 얻어 저장한다. 판매자의 출력은 판매기록 recordM 또는 failed이다.
- 식별 프로토콜 : 식별 알고리즘에 cover-object, 판매 리스트, stego-object를 입력한다. 식별 알고리즘의 출력은 "failed" 또는 구매자 pk_B 식별정보, 구매자 서명 문자열(proof)이다. 이 프로토콜은 분배자(distributor), 중재자(arbiter), 구매자(buyer) 사이의 two-party 또는 three-party 프로토콜이다. 고발된 구매자가 분배자와 중재자의 분쟁 프로토콜에 대하여 이의를 신청할 경우에 분쟁 프로토콜에 참가한다. 판매자와 중재자는 pk_B와 str을 입력하고 판매자는 추가로 proof를 입력한다. 중재자를 위한 출력은 구매자를 고발 / 무혐의 처리하는 Boolean 값이다. 비대칭 핑거프린트는 다음 3가지 요구사항을 만족하여야 한다.
  1)불법 복사를 하려는 구매자를 구별할 수 있어야한다.(effective)
  2)판매자는 구매자들이 속이는 것으로부터 보호받아야 한다.
  3)구매자는 판매자가 속이는 것과 다른 구매자들로부터 보호 받아야 한다.

3) 익명 핑거프린팅

익명 핑거프린팅은 Chaum[7]이 제안한 은닉서명의 한 응용분야이다. 신뢰할 수 있는 third party와 비대칭 scheme을 결합한 형태이다. 은닉 서명은 데이터의 내용을 서명자에게 보여주지 않고, 서명자로부터 데이터에 대한 서명을 얻게 만드는 것이 가능한 서명 방식이다. 판매자는 인증센터의 도움 없이 부정자를 단독으로 구별할 수 없다. 또한 판매자는 인증센터를 사용함으로써 핑거프린트를 구매자와 연결시키는 상세한 기록들을 보관할 필요가 없게 된다. 익명 핑거프린팅은 Pfitzmann과 Waidner[8]에 의하여 TTP에 기반한 방식이 제안되었다. 이 방식은 구매자들이 어떤 제품을 샀는지를 모르게 하는 익명성을 제공한다는 이유에서 주목(개인 프라이버시 보호)할만하고, 또한 구매자가 익명으로 산 정보가 불법적으로 재분배된 경우에만 구매자를 식별할 수 있기 때문에 EC(Electronic commerce)에 적용할 수 있는 방식이다.

- 키 생성 프로토콜 : 구매자는 등록 센터로부터 익명 서명을 얻기 위하여, 임의의 키 쌍(sk_b, pk_b)을 생성하고, 생성한 키 쌍에 대하여 책임을 지기 위하여 구매자의 진짜 식별

정보(S_b) 하에서 서명한다. 등록 센터로부터 익명 키 쌍에 대한 인증서 $cert_b$를 얻는다. 이 인증서를 가지고 등록 센터는 이 익명 키 쌍을 선택한 구매자의 식별정보를 알게 된다.

- 핑거프린팅 프로토콜 : 구매자는 판매자의 지식 없이 구매를 식별하는 문자열에 서명을 한다. 구매자는 구매할 때 sig := sign(sk_b, text)한다. 그리고 구매된 데이터 항목 안에 정보 emb := (text, sig, pk_b, $cert_b$)를 삽입한다. 구매자는 이 값을 bit commitment 안에 숨기고 영지식으로 인증서와 commitment를 판매자에게 보낸다. 여기서도 비대칭 핑거프린팅 방식과 같이 2-party 평가를 사용하게 된다.
- 식별 프로토콜 : 불법적으로 데이터를 재분배한 구매자를 식별할 때, 판매자는 불법 복사본으로부터 emb를 추출하고 구매자 서명 문자열인proof := (text, sig, pk_b) 를 등록센터에 보내고, 등록센터에게 이 proof에 대한 식별을 요청한다. 등록센터는 판매자에게 이 proof에 대응하는 구매자의 서명(sig)을 돌려보낸다.
- 분쟁 프로토콜 : 등록센터로부터 받은 서명을 통해, 판매자는 자신이 갖고 있는 sig값들을 검증할 수 있고 구매자를 고발하는 증거를 갖게 된다. 분쟁 프로토콜은 비대칭 방식과 같이 판매자와 중재자 사이의 2-party 프로토콜과 등록 센터의 부정을 확인하기 위하여 등록 센터를 포함하는 3-party 프로토콜 또는 구매자가 판매자, 중재자, 등록 센터의 고발에 이의를 제기할 경우에는 구매자까지 포함하는 4-party 프로토콜로 구성될 수 있다.

4) 공모 보안 코드(Collusion Secure Code)

핑거프린팅된 콘텐트간의 차이점을 이용한 공모공격은 적은 수의 콘텐트만으로도 핑거프린트를 완전히 제거할 수 있기 때문에 상당히 위협적인 공격이다. 이러한 공모공격에 강인하도록, 삽입하는 핑거프린트 자체를 공모가 어렵도록 설계하는 연구가 수행되고 있다. 일반적인 핑거프린트는 구매자마다 완전히 다른 코드, 즉 무작위순열을 사용한다[8]. 무작위순열은 그 길이에 따라 어느 정도 공모공격에 강인성을 갖는다[15]. 하지만 공모자가 많아질수록 필요한 코드의 길이가 기하급수적으로 증가한다. 그래서 구매자마다 다른 위치에서 공통된 부분을 갖도록 하여 무작위순열보다 효율적인 코드를 설계할 필요가 있다. 코드의 공통된 부분은 공모공격을 해도 제거되지 않으므로 이 부분의 위치 정보가 공모에 참여한 구매자의 정보를 나타내게 된다. [9][12]-[15].

3. 디지털 원문 보호시스템의 설계

1) 디지털 원문 보호시스템 구성

KISTI가 구축하여 무료 서비스중인 국내 학회지 및 연구보고서의 원문 데이터베이스에 대하여 부정한 사용자가 상업적인 목적으로 유포 및 판매를 근절시키기 위한 대책이 필요하다. 이를 위하여 핑거프린팅 기법을 이용하여 불법한 사용자가 콘텐트를 다운로드 받지 못하게 하는 불법 복제 방지, 핑거프린트 정보(다운로드 받는 사용자 정보)를 콘텐트에 삽입하는 핑거프린팅, 인터넷 웹상에서 콘텐트를 불법적으로 유통시키는 자(traitor)에 대한 콘텐트 사용자 위치 추적 시스템이 필요하다.

KISTI에서 제공하는 원문은 각 학회지별로 스캔 작업을 통해 원문이미지(tiff)로 변환되어 디렉토리별로 Image Achive에 저장된다. 각각의 서지정보는 Oracle RDB에 저장된다. Web Client가 Web Browser를 통하여 원문에 대한 검색 요청을 하면 Web Server는 웹 서비스 프로그램 모듈을 통해 DBMS와 상호 연동하여 그 원문에 해당하는 서지정보를 찾아 사용자에게 결과를 보여주게 된다. 만약 사용자가 원문 다운로드 요청을 하게 되면 Image Archive에서 그 서지정보에 해당하는 원문을 호출하여 Client에게 다운로드 서비스를 하게 된다.

KISTI에서는 원문을 File로 받아서 이미지로

변환하는 방법보다 Scanner를 사용하여 학회지에 실려 있는 문서를 스캔(scan)하여 tiff 이미지를 생성하는 방법을 더 많이 사용하고 있다. 본 연구에서는 KISTI에서 현재 서비스하고 있는 원문 데이터베이스 시스템에 다음과정을 추가하여 원문의 불법 복제 및 배포를 방지하기 위한 체계를 설계하였다.

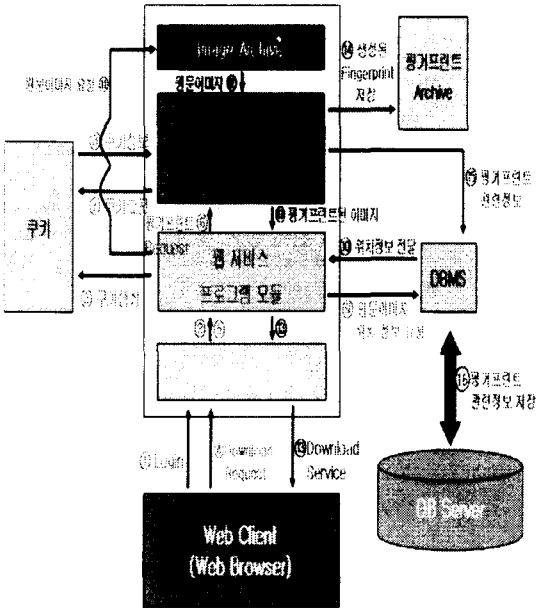


그림 1. 디지털 원문보호시스템 구성도

그림 1은 KISTI의 원문 서비스 시스템에 불법 복제 방지와 불법 사용자 추적을 할 수 있는 핑거프린트를 생성·삽입하는 핑거프린팅 모듈을 추가한 시스템 구성도이다.

Web Client가 KISTI의 홈페이지에 로그인을 하면 KISTI의 웹 서비스 프로그램 모듈에서 쿠키(세션)를 생성한다. Client가 원문 서비스를 받고자 할 때 원문의 다운로드 요청 시 Web Server를 통해 웹 서비스 프로그램 모듈을 거쳐 핑거프린트 모듈에 핑거프린트 Request를 요청한다. 그런 후 로그인시 생성된 쿠키정보를 가져와서 그 Client의 유일한 핑거프린트를 생성한다. 핑거프린트 모듈은 DBMS에게 원문의 위치 정보를 요구하게 되고 DBMS는 DB Server(Oracle RDB)에서 서비스할 원문의 위치정보를 찾아내어 핑거프린팅 모듈에게 전달한다. 핑거프린팅 모듈은 Image Archive로부터 웨이블릿 변환된 원문 이미지를 호출하여 생성된 핑거프린트를 삽입한 후 웨이블릿 역변환을 통하여 Client에게 다운로드 서비스를 한다. 그 후 생성된 핑거프린트는 핑거프린트 Archive에 저장하고 그에 관련된 정보를 DBMS를 통해 DB Server (Oracle DB)에 저장한다.

표 1 핑거프린트 상세처리 과정

단계	처 리 과 정
1	Client Log-in
2	Web Server -> 웹 서비스 프로그램 모듈
3	쿠키 생성
4	Client의 Download Request
5	Web Server -> 웹 서비스 프로그램 모듈
6	웹 서비스 프로그램 모듈의 핑거프린트 Request
7	쿠키정보 요청
8	생성된 쿠키정보를 받아
9	핑거프린트 모듈에서의 DBMS에 대한 원문 이미지 위치정보 요청
10	DB Server(Oracle RDB)에서 찾은 원문 이미지 위치정보 전달
11	Image Archive에 원문 이미지를 요청
12	웨이블릿 변환된 원문 이미지를 호출
13	생성된 핑거프린트를 웨이블릿 역변환후 웹 서비스 프로그램 모듈과 Web Server를 통하여 Client에게 다운로드
14	생성된 핑거프린트를 핑거프린트 Archive에 저장
15	핑거프린트 관련정보를 DBMS에 전달
16	핑거프린트 관련정보를 DB Server(Oracle RDB)에 저장

2) 핑거프린트 생성 및 삽입

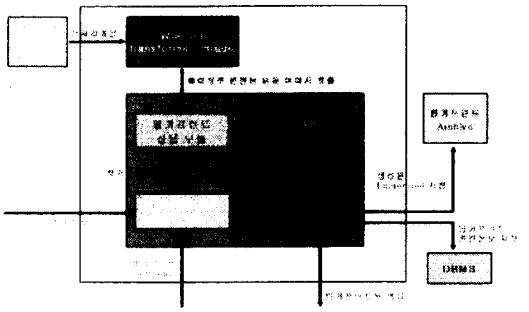


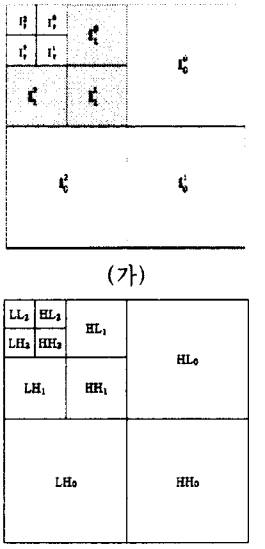
그림 2. 핑거프린팅 모듈의 상세구조

그림 2는 핑거프린팅 모듈의 상세 구조이다. 다운로드 시 시간을 줄이기 위해서 전처리 과정으로 원문이미지를 미리 웨이블릿 변환을 한 후 Image Archive에 저장한다. Download Request가 들어오면 생성된 쿠키정보를 통해 그 Client에 해당하는 중복되지 않는 유일한 핑거프린트를 생성하게 되고 생성된 핑거프린트를 웨이블릿 변환된 원문 이미지에 삽입한 뒤 웨이블릿 역변환을 통해 핑거프린트된 영상을 전달한다. 한편, 생성된 핑거프린트는 핑거프린트 Archive에 저장을 하고 핑거프린트 관련정보(핑거프린트와 Client 계정과의 매핑 관계)는 DBMS를 통해 DB Server(Oracle RDB)에 저장한다.

표 2. 핑거프린팅 모듈 처리과정

단계	처 리 과 정
1	웹 서버에 저장된 tiff 이미지를 웨이블릿 변환(전처리 과정)
2	핑거프린트 요구(Request)
3	쿠키정보 요청
4	생성된 쿠키정보를 받아옴
5	핑거프린트 생성
6	웨이블릿 변환된 이미지 요청
7	호출된 이미지에 생성된 핑거프린트 삽입
8	웨이블릿 역변환
9	Client에게 핑거프린팅된 영상 다운로드
10	생성된 핑거프린트를 핑거프린트 Archive에 저장
11	핑거프린트 관련정보를 DBMS에 전달함
12	핑거프린트 관련정보를 DB Server(Oracle RDB)에 저장함

핑거프린트에 대한 삽입 방법은 DWT변환을 이용하여 중간 주파수 밴드에 특정계수를 선택하고 휴먼 비쥬얼 시스템을 적용하여 워터마크(핑거프린트)를 삽입하는 일련의 과정을 살펴본다.



(가) (나)

그림 3. DWT에 의한 각 밴드의 표기법 및 변환

본 논문에서는 각 밴드의 표기법을 그림 3(가)와 같이 한다[16]. 즉, 3레벨의 DWT 영역상에서 해상도 레벨이 $l=0,1,2$ 방향이 $\theta\in\{0,1,2,3\}$인 서브밴드를 I_l^{θ}로 칭한다. 회색으로 칠해진 부분은 워터마크가 삽입될 밴드이다. 그림 3 (나)는 영상을 DWT변환을 했을 시 고주파, 저주파 밴드를 일반적으로 나타낸 것이다.

Lewis and Knowles가 고안한 HVS는 눈, 망막, 신경망 등 사람의 눈이 가지는 여러 가지 특성을 고려하여 워터마크를 삽입할 때 인간이 지각할 수 없을 수준의 워터마크를 삽입하기 위한 시스템이다. 이전에는 두 신호 사이에 사람이 탐지할 수 있는 최소한의 차이를 나타내는 JND(Just Noticeable Difference)를 적용하였다.

그러나 이는 흑백 영상에만 적용이 가능하다는 단점이 있다. HVS은 다음과 같은 인간의 시각적인 특성을 고려하였다[16].

(1) 인간의 눈은 높은 해상도 밴드와 45° 방향의 밴드, 즉 θ=1인 밴드 에서는 노이즈에 대하여 덜 민감하다.

(2) 인간의 눈은 영상의 밝거나 어두운 곳에서는 노이즈에 대하여 덜 민감하다.

(3) 인간의 눈은 높은 텍스쳐 지역에서 노이즈에 대하여 덜 민감하나 에지 근처의 지역에서는 노이즈에 더욱 민감하다.

이러한 특성을 기반으로 하여 사람의 눈에 덜 민감한 부분에 큰 크기를 갖는 워터마크를 삽입할 수 있고 이를 통해 여러 견고한 워터마킹 알고리즘을 설계할 수 있다.아래와 같이 3개의 항의 곱으로써 각각의 계수에 대한 양자화 과정을 계산한다.

$$q^{\theta}_{l}(i,j) = \Theta(l,\theta)\Lambda(l,i,j)\Xi(l,i,j)^{0.034} \quad (1)$$

우선 첫 번째 항은 높은 해상도 밴드와 45° 방향의 대각선 밴드에 덜 민감한 특성에 대하여 아래와 같이 수식적으로 워터마크의 강도(Strength)를 조정할 수 있다. 고 해상도 부분인 레벨 1에서는 1을 레벨 2에서는 0.32, 레벨 3에서는 0.16을 각각 대각선 부분 즉 HH_0와 HH_1 그리고 HH_2에서는 $\sqrt{2}$의 값을 곱하고 나머지 부분에 대해서는 1을 곱해 가중치를 준다.

$$a(\theta,l) = \begin{Bmatrix} \sqrt{2} & \text{if } \theta = 1 \\ 1 & \text{otherwise} \end{Bmatrix} \cdot \begin{Bmatrix} 1.00 & \text{if } l = 0 \\ 0.32 & \text{if } l = 1 \\ 0.16 & \text{if } l = 2 \\ 0.10 & \text{if } l = 3 \end{Bmatrix} \quad (2)$$

여기서 θ는 수직에지, 수평에지, 대각선에지를 나타내는 밴드위치를 나타내고 θ=1일 땐 대각선 위치를 나타낸다. l은 해상도를 나타낸다. 아주 밝고 어두운 성분에 대해서 인간의 눈은 노이즈에 덜 민감한 특성이 있는데 이는 가장 낮은 주파수 대역인 I^{3}_{2}을 기준으로 계산하고자 하는 계수의 특성을 파악한다. 이는 웨이블릿 해석의 특징을 이용한 것으로 각 레벨의 같은 위치에 있는 계수 값은 서로 밀접한 연관관계가 있는 것에 근거한다.그림 3은 워터마크 삽입과정을 전체적으로 도식화 한 것이다.

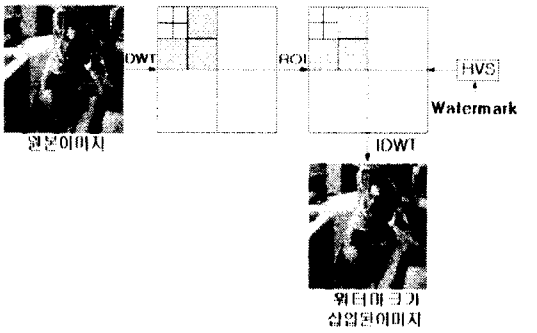


그림 4. 워터마크 삽입 과정

[단계 1] 첫 번째로 영상에 워터마크를 삽입하기 위하여 그림 4와 같이 세 개의 레벨로 DWT과정을 통하여 분해한다. 본 연구에서 사용한 웨이블릿 필터는 9/7 탭을 가지는 Antonini[17]가 제안한 방법을 이용했다.

회색으로 채워진 중간주파수 밴드 즉, I^{0}_{1}, I^{1}_{1} 그리고 I^{2}_{1}가 워터마크가 삽입되는 주파수밴드이다. 이렇게 중간 주파수 밴드에 워터마크를 삽입하는 이유는 중간 주파수 밴드에 삽입한 워터마크는 어떠한 영상 처리 과정 후에도 높은 비율의 추출을 보였으며, 워터마크가 삽입된 영상의 왜곡정도도 다른 대역보다 상대적으로 적게 나타났기 때문에 중주파 대역에 워터마크를 삽입하였다.

[단계 2] 견고한 워터마크의 삽입을 위해 시각적으로 중요한 계수 값을 선택하게 되는데 이때 MTWC의 원리를 이용하여

ROI를 찾아 삽입하게 된다[18]. 시각적으로 중요한 계수 값은 보통 큰 값을 가지고 있는데 일반적으로 이러한 계수 값은 압축과 같은 영상 처리 후에도 변화가 적다. 만약에 이러한 계수 값들이 변화한다면 역 DWT과정을 거친 영상은 원본 영상과 매우 다를 것이다. 시각적으로 중요한 계수 값을 찾는 과정은 다음과 같다.

(가) 처음은 초기화 단계로써 워터마크가 삽입 될 밴드의 초기 임계값(T_s)을 설정한다. 이 때 임계값은 그 밴드의 계수 값 중 절대값이 가장 큰 계수값(C_{max})의 반이다.

즉, $T_s = \frac{C_{max}}{2}$ 이다.

(나) 워터 마크가 삽입될 밴드의 중요한 계수 값을 찾기 위해서 그 밴드의 모든 계수 $C_s(i,j)$와 임계값과 비교하여$C_s(i,j) > T_s$ 이면 중요한 계수로써 선택된다.

(다) 선택된 시각적으로 중요한 계수에 HVS을 적용하여 워터마크를 삽입한다.

(라) 삽입되어야 할 워터마크가 모두 삽입되지 않았을 경우 나머지 워터마크를 모두 삽입하기 위하여 다시 (가)단계부터 위 과정을 반복수행 한다.

[단계 3] 워터마크가 삽입되는 중간주파수 밴드 즉, I_1^0, I_1^1 그리고 I_1^2밴드에서 선택된 시각적으로 중요한 계수는 HVS특성을 적용한다. 워터마크의 삽입은 모든 계수에 대해서 적용하는 것이 아니므로 다른 알고리즘에 비해 계산적인 부하도 적게 들고 영상에 미치는 영향도 적다. 워터마크 삽입은 영상에 워터마크 정보를 삽입하는 과정으로 아래의 식을 따른다.

$$C'_s(i,j) = C_s(i,j) + \alpha \cdot q(i,j) \cdot W_k \quad (3)$$

C_s는 선택된 원래의 계수 값이고, C_s'은 워터마크가 삽입된 계수 값이다. 그리고 α값은 워터마크의 강도를 나타내는 파라미터고 $\alpha \in (0.0, 3.0]$이다. $q(i,j)$는 HVS의 특성을 이용한 weighting 수이다. W_k는 워터마크로써 평균이 0이고 분산이 1인 pseudo-random 가우시안 랜덤 벡터를 사용하였다. 가우시안 랜덤 벡터는 서로 독립적이고 연관성이 없으며 패턴을 추측할 수 없는 장점이 있다. 워터마크는 문자열, 영상, 로고, 난수열 등을 알고리즘의 특성에 맞게 사용할 수 있다.

3) 핑거프린트 추출

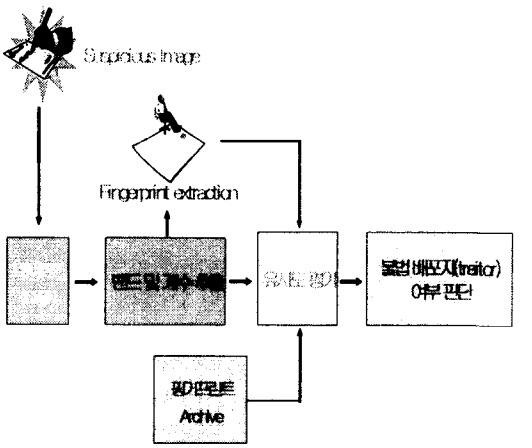


그림 5. 핑거프린트의 추출 과정

그림 5은 의심이 가는 원문 이미지에 대하여 불법 배포자(traitor)를 찾기 위해 핑거프린트의 추출 유무를 판단하는 알고리즘이다. 먼저 이미지를 웨이블릿 변환한 후 각 밴드 및 계수를 추출한 후 원래의 원문 이미지와 비교하여 유

사도 평가를 통하여 핑거프린트를 추출하고 핑거프린트 Archive와 대조하여 불법 배포자를 찾아낸다.(핑거프린트는 로그인시 생성된 쿠키를 바탕으로 만들어지기 때문에, 한 Client에 대한 핑거프린트는 유일하며 그로인해 불법 배포자를 추적할 수 있다.) 핑거프린트 추출 방법은 워터마크가 삽입된 영상에서 워터마크를 추출하여 워터마크의 삽입여부 및 유효성을 판단하는 방법을 기술한다.

표 4. 워터마크 추출 과정

단계	처 리 과 정
1	원본영상과 워터마크가 삽입된 영상을 호출.
2	호출된 두 영상을 DWT변환 과정을 거친 후 3개의 레벨로 각각 주파수 분해.
3	주파수 분해 된 두 영상을 비교하여 워터마크를 추출.
4	Cox의 유사도 측정식을 사용하여 유사도를 측정.
5	측정된 유사도 값과 실험에 의해 측정된 Threshold값과 비교하여 워터마크의 삽입부를 판단.

$$C(W, W^*) = \frac{W \cdot W^*}{\sqrt{W \cdot W^*}} \qquad (4)$$

C는 유사도를 나타내고, W는 삽입되기 전 워터마크이고 W^*는 삽입한 뒤 추출한 워터마크를 나타낸다. 워터마크의 유무는 C값이 Threshold 값 이상이면 워터마크가 있는 걸로 판단하고 그 이하는 워터마크가 없는 것으로 판단한다.

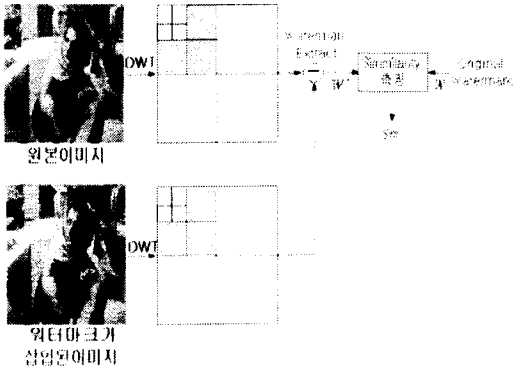


그림 6. 워터마크 추출 과정

4. 향후 연구 과제

본 연구를 통하여 핑거프린팅 및 워터마킹에 대한 원천기술을 확보하고 KISTI의 디지털 원문 보호시스템의 설계를 하였으나 원문서비스에 대한 보안을 위해선 콘텐트에 보이지 않는 정보를 심어 저작권을 보호하고 불법적인 복제를 막는 디지털 워터마킹 기술과 불법 복제 및 유통행위가 발견되었을 시 불법 배포자(traitor)를 추적하는 시스템이 필요하다.

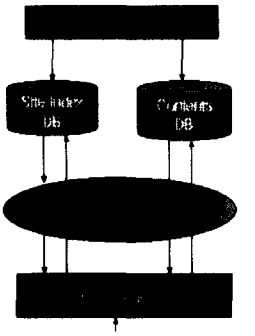


그림 7. 콘텐트의 불법 복제 추적 시스템의 구성도

그림 7.는 인터넷 웹상에서 불법적인 재배포를 추적하는 시스템을 보여주고 있다. 불법 복제된 콘텐트를 추적하기 위해서는 핑거프린팅/워터마킹 기법과 검색 에이전트와의 연동이 필수적이다. 멀티미디어 검색 에이전트를 이용하여 불법적으로 배포되어진 디지털 콘텐트에 대해서 인터넷을 통하여 검색한다. 검색 도중 원본 콘텐트와 유사한 특성 및 크기를 가진 콘텐트를 핑거프린터 추출 서버로 전송하여 서버에서 핑거프린트 추출 모듈이 불법 복제 여부를 판단하고 배포한 원인 제공자의 ID를 추적한다. 검색 에이전트가 콘텐트를 검색하고 선정된 콘텐트를 전송하는 과정에서 수반되는 과도한 네트워크 부하를 줄이기 위해 검색 에이전트로부터 전송되어지는 콘텐트의 작은 부분만으로 불법적인 콘텐트인지를 판단하여야 한다. 서버는 이러한 기능을 포함하여야 하며 실시간 처리가 가능하도록 신속하여야 하며

또한 문제의 소지를 방지하기 위해 정확성을 기해야 한다.

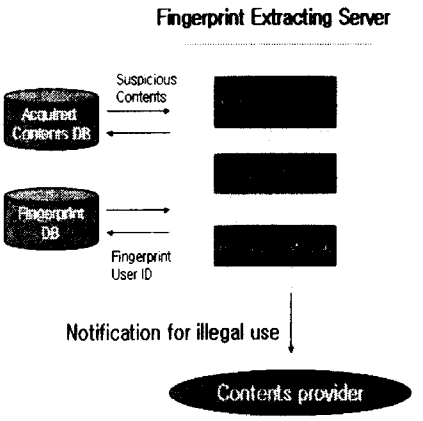


그림 8. 콘텐트 DB에서 소유권 판별 구성도

그림 8은 핑거프린트 추출자가 획득한 콘텐트 DB에서 콘텐트 소유권자의 콘텐트인지를 판별하는 구성도를 나타낸다. 인터넷 웹상에서 저작권자의 콘텐트와 유사한 자료를 찾아 저장된 콘텐트 DB(Acquired Contents DB)에서 핑거프린트 추출서버(Fingerprint Extracting Server)의 추출 모듈(Extraction Module)에서 핑거프린트를 추출하여 결정 모듈(Decision Module)에서는 핑거프린트 DB (Fingerprint DB)에 저장되어 있는 핑거프린트와 추출된 핑거프린트를 비교 검토하여 불법 사용자 여부를 판단한다. 이 결과를 보관 모듈(Recording Module)에 저장하고 이 결과를 콘텐트 제공자에게 통보하게 된다.

5. 결 론

본 논문을 통하여 불법 복제를 방지하는 시스템을 설계하였으며 불법 유통 시 원인 제공자를 추적하는 시스템 개발에 관하여 연구하였다. 또한 원문 데이터베이스를 보호하기 위해서 필요한 현존하는 핑거프린팅을 분석하였다. 핑거프린팅 기술은 원문 데이터베이스를 이용하는 사용자들이 다운로드 받은 원문데이터를 불법 유통시킬 경우 원인 제공자를 추적하는데 필요하게 된다. 핑거프린트는 원문을 다운로드 하는 각 사용자에게 부여하는 원문 복사본 코드를 의미한다. 많은 사용자들이 원문을 다운로드 받아 사용하기 때문에 핑거프린트(워터마크) 크기를 크게 하여 핑거프린트 검출율을 높이고 공격자들의 변형 공격에 대해 워터마크가 살아남을 수 있는 알고리즘을 개발하였다. 이러한 견고성(rubustness)를 높이면서 고화질을 유지하기 위해서 웨이블릿 변환 기술과 HVS (Human Visual System) 기술을 이용하였다.

웨이블릿 신호처리 기술은 우리 인간의 눈이 고주파에 둔감하고 저주파에 민감하기 때문에 원 신호를 고주파와 저주파 성분으로 분해하는데 사용된다. 변환된 웨이블릿 계수에 눈의 특성에 맞게 워터마크를 삽입함으로써 공격자의 공격에 대해 워터마크가 생존할 확률을 높일 수 있게 된다. 본 논문을 통하여 개발된 콘텐트 보안 시스템과 워터마킹 알고리즘을 이용하여 원문 데이터를 효율적으로 보호할 수 있게 되었다. 개발된 보안 체계와 알고리즘을 실제 원문 서비스 시스템에 적용함으로써 효율적으로 원문 데이터베이스를 보호할 수 있을 것으로 판단한다.

참 고 문 헌

[1] 김태중, 김상국, 송유진, "콘텐트 저작권 보호 및 관련기술 표준화 동향", 정보보호학회지, 제14권 제1호, pp91-106, 2004

[2] 한국과학기술정보연구원, "원문 데이터베이스의 도용방지를 위한관리체계 연구 개발", pp15-72, 2003

[3] 유원영, 서영호, 최재귀, 박지환, "디지털 핑거프린팅과 구매자/판매자 워터마킹의 기술동향", 전자통신동향분석, 제19권 제1호, pp96-105, 2004

[4] 김종원, 신동환, 신승원, 최종욱, "디지털 워터마킹 기술의 산업적 응용", 정보보호학

회지, 제12권 제1호, pp11-18, 2002

[5] 박성득, "지적재산권 보호를 위한 정보은닉 기술 및 표준화 연구" 정보통신부 정보통신 연구개발사업, pp. 103-138, 2000

[6] Pfitzmann, B., and M. Schunter, "Asymmetric Finger -printing", in Advances in Cryptology, Proceeding of EUROCRYPT '96, vol. 1070 of Lecture Notes in Computer Science, Springer -Verlag, pp. 84-95, 1996

[7] Chaum, D. L., "Blind Signatures for Untraceable Payments", in Advances in Cryptology, Proceedng of CRYPTO '82, Plenum Press, pp.199-203, 1983

[8] Pfitzmann, B., and M. Waidner, "Anonymous Fingerprinting", in Advances in Cryptology, Proceedng of EUROCRYPT '97, vol. 1233 of Lecture Notes in Computer Science, Springer-Verlag, pp. 88-102, 1997

[9] W.G. Kim, S.H. Lee, H.W. Jang, "Trend of Fingerprinting Technique for Tracing Illegal Contents", kidbs.itfind.or.kr :8888/WZIN/techtrnd/18-4_082_094.pdf

[10] B. Pfitzmann and A. Sadeghi, "Coin-Based Anonymous Fingerprinting," in Advances in Cryptology, Proc. of EUROCRYPT'99, Vol. 1592, of Lecture Notes in Computer Science, Springer-Verlag, 1999, pp. 150 - 164.

[11] J. Kilian, T. Leighton, L.R. Matheson, T.G. Shamoon, R.E. Tarjan, and F. Zane, "Resistance of Digital Watermarks to Collusive Attacks," Tech. Rep., TR-585-98, Dept. of Computer Science, Prinston University, 1998.

[12] D. Boneh and J. Shaw, "Collusion-Secure Fingerprinting for Digital Data," IEEE Trans. Inf. Theory, Vol. 44, No. 5, Sep. 1998, pp. 1897 - 1905.

[13] J. Dittmann, "Combining Digital Watermarks and Collusion Secure Fingerprints for Customer Copy Monitoring," Proc. IEE Seminar Sec. Image & Image Auth., Mar. 2000, pp. 128 - 132.

[14] J. Domingo-Ferrer and J. Herrera-Joancomartí, "Simple Collusion-secure Fingerprinting Schemes for Images," in IEEE International Conference on Information Technology: Coding and Computing, ITCC'2000, ISBN 0-7695-0540-6, pp. 128 - 132.

[15] Yiwei Wang, John F. Doherty, and Robert E. Van Dyck, "A Watermarking Algorithm for Fingerprinting Intelligence Images," 2001 Conference on Information Sciences and Systems, The Johns Hopkins University, March 21-23, 2001.

[16] M. Barni, F. Bartolini, A. Piva, "Improved Wavelet Based Watermarking Through Pixel-Wise Masking," IEEE Trans. Image Processing, vol. 10, pp.783-791, May. 2001.

[17] M. Antonini, M. Barlaud. P. Mathieu and I. Daubechies, "Image Coding using Wavelet Transform", IEEE Trans. Image Processing, vol. 1. no 2, pp. 205-220, Apr. 1992.

[18] H, J. Wang, C.-C. Jay Kuo, "A multi-Threshold Wavelet Coder For High Fidelity Image Compression," IEEE Image Processing, vol.1. pp.652-655, 1997.

[19] J. D. Villasenor, B. Belzer, and J Liao, "Wavelet Fiter Evaluation for Image Compression", IEEE Trans. Image Processing, vol. 4, no. 8, pp.1053-1060, Aug, 1995.

김상국(Sang-Kuk Kim)

1989년 인천시립대학교 전자공학과(공학사)
1991년 한양대학교 대학원 전자계산학전공 (이학석사)
2001년~현재 : 한남대학교 대학원 컴퓨터공학과 (박사과정)
1989년~1994년 : 시스템공학연구소 연구원
1995년~2000년 연구개발정보센터(KORDIC) 선임연구원
2001년~현재 : 한국과학기술정보연구원 정보 표준화실 선임연구원
관심분야 : 정보 보호, 정보통신, 멀티미디어 DB 등

이재광(JaeKwang Lee)

1984년 광운대학교 전자계산학과 (학사)
1986년 광운대학교 대학원 전자계산학과 (석사)
1993년 광운대학교 대학원 전자계산학과 (박사)
1986년~1993년 군산전문대학 전자계산학과 부교수
1997년~1998년 University of Alabama 객원교수
1993년~현재 한남대학교 컴퓨터공학과 교수
관심분야 : 컴퓨터 네트워크, 정보통신, 정보보호