Journal of Korea Multimedia Society Vol. 7, No. 9, September 2004(pp. 1282-1293)

공개키 기반 구조를 이용한 비대칭 워터마킹

전영민[†], 양선옥[††], 김계영[†††]

요 약

본 논문에서는 공개키 기반 구조를 이용한 비대칭 워터마킹 방법을 제안한다. 제안하는 방법의 특징은 서로 상이한 암호화 기술과 워터마킹 기술을 각 기술 간의 특성에 기초하여 인증 기술로 연계한 점이다. 디지털 콘텐츠에 삽입하는 워터마크는 저작권 정보를 분배자 혹은 저작권자의 개인키로 암호화한 디지털서명과 인증코드로 구성된다. 디지털 콘텐츠에 대한 소유권 판단 시 인증 기술은 인증코드에 근거하여 디지털 콘텐츠의 데이터 무결성을 검사하며 만족하는 경우와 만족하지 않는 경우로 구분하여 저작권을 판단한다. 전자의 경우는 디지털 콘텐츠에 삽입한 워터마크를 추출하고 워터마크를 구성하는 암호화된 저작권 정보를 분배자의 공개키로 복호화하여 복호화 된 저작권 정보와 사용자가 제시하는 저작권 정보를 서로 비교함으로 소유권을 판정하는 암호화에서의 복호화 방법을 사용하며, 후자의 경우는 디지털 콘텐츠에서 추출한 워터마크로부터 분리한 암호화된 저작권 정보와 사용자가 제시하는 워터마크로부터 분리한 암호화된 저작권 정보 간의 유사도를 비교함으로 소유권을 판정한다. 제안하는 방법은 워터마크 암호화키를 알아내거나 제거하려는 시도로부터 안전성을 제공한다.

Asymmetric Watermarking Using Public Key Infrastructure

Young-Min Jun[†], Sun-Ouk Yang[††], Gye-Young Kim[†††]

ABSTRACT

This paper proposes an asymmetric watermarking system using Public Key Infrastructure. The distinguishing characteristic of the proposed method connects between the two different techniques, cryptography technique and watermarking technique, by using the authentication technique. The connection between the two techniques are established based on the special qualities of each technique. Watermarks that are inserted into the digital contents consist of a digital signature described as an encrypted copyright information with the private key of a distributor or a copyright holder, and an authentication code. In the situation where the ownership of the digital contents has to be decided, authentication technique examines the data integrity of the digital contents based on an authentication and decides the ownership of the digital contents by examining whether it satisfies or not satisfies the integrity test. The formal case uses decryption method which compares the user defined copyright information, and the decrypted copyright information extracted from the watermark in the digital contents that are decrypted by distributors' public key. The latter case determines the ownership by comparing the similarity between encrypted copyright information separated from the watermark that are extracted from the digital contents, and the user defined encrypted copyright information that are separated from the watermark. The proposed method provides protection from the assault which attempts to identify or erase the encoding key.

Key words: Digital Watermarking(디지털 워터마킹), Public Key Infrastructure(공개키 기반 구조), Cryptography(암호화), Authentication(인증), Digital Signature(디지털 서명)

※ 교신저자(Corresponding Author) : 전영민, 주소 : 서울시 동작구 상도 5동 1-1(156-743), 전화 : 02)825-1087, FAX : 02)825-1087, E-mail : ymjun@vision.ssu.ac.kr
접수일 : 2003년 9월 22일, 완료일 : 2004년 3월 16일
[†] 정회원, 숭실대학교 대학원 컴퓨터학부 박사과정
[††] 정회원, 숭실대전산원 전임강사
(E-mail : soyang@comist.soongsil.or.kr)
[†††] 숭실대학교 컴퓨터학부 교수
(E-mail : gykim@computing.ssu.ac.kr)
※ 본 연구는 숭실대학교 교내연구비 지원으로 이루어졌음

1. 서 론

디지털 워터마킹은 저작권 정보인 워터마크를 디지털 콘텐츠에 직접 삽입함으로 디지털 콘텐츠에 대한 저작권 보호와 비인가 된 접근 또는 조작을 방지할 수 있는 기술이다[1-3]. 저작권 정보가 삽입된 디지털 콘텐츠가 복제되면 삽입한 저작권 정보도 같이 복제되므로 추출된 워터마크와 제시된 워터마크의 비교를 통하여 그 권리의 진위를 판단할 수 있다. 따라서 디지털 워터마킹은 저작권 보호와 디지털 콘텐츠 산업의 육성을 위해서 중요한 연구 분야이다.

일반적으로 워터마킹에 요구되는 요건들에는 인간시각에 노출되지 않게 워터마크를 삽입하는 시각적 무감지성, 다양한 변형 뒤에도 워터마크를 추출할 수 있는 강인성, 의도적인 공격에 대한 안전성 등이 고려되고 있다. 우리는 참조 논문 [4]에서 워터마크의 시각적 무감지성과 강인성 요건과 관련된 성능 향상 방안에 대해서 연구하였다. 본 논문에서는 워터마크의 안전성 요건에 연구의 중심을 둔다. 일반 사용자가 워터마크 추출기를 통해 자유롭게 워터마크 정보를 확인하는 일반적인 응용환경에서 안전성에 대한 문제는 중요하고 어려운 문제이다. 안전성은 워터마크 삽입과 추출에 사용되는 키와 밀접한 관련이 있다.

워터마크 키의 대칭성 측면에서 워터마크 방식을 분류해 보면, 거의 모든 방식이 삽입기와 추출기 양측에서 동일한 키를 사용하는 대칭 워터마킹 방식이다. 이는 대칭 암호화 방식과 유사하며 키가 안전하게 관리되고 분배된다는 가정 하에서만 시스템의 안전성이 보장된다. 그러나 키의 안전한 관리와 분배는 매우 어렵다. 일반적인 대칭 워터마킹 방식의 워터마크의 삽입과 추출을 식(1)과 같이 표현 할 수 있다.

$$c = E_K(x, m), \quad m = D_K(c) \tag{1}$$

식(1)에서 $E(\cdot)$는 삽입함수이고, 메시지 m은 대칭키 K에 의해 워터마크 w로 만들어져 입력영상 x에 삽입되어 워터마크가 삽입된 영상 c를 생성한다. 영상 c에 삽입된 메시지는 역시 대칭키 K를 사용하여 추출함수 $D(\cdot)$에 의해 추출된다. 그러나 이 방식과 같이 공개적으로 워터마크 정보를 추출하는 환경에서 공개된 워터마크 추출기에 대한 공격이나 실수에 의해 내부 키 정보가 유출될 경우, 유출된 정보가 삽입기의 정보와 동일하기 때문에 시스템 전체의 안전성에 위협이 된다. 이러한 문제점을 해결하기 위해 워터마크의 삽입과 추출에서 다른 키를 사용하는 비대칭 워터마킹 방법이 현재 대두되고 있다[5-7]. 이는 추출기에서 내부 키 정보가 유출되어도 삽입키의 정보는 안전하게 보존됨으로써 전체 시스템의 안전성을 향상시킬 수 있다.

일반적인 비대칭 워터마킹 방식의 구조는 그림 1과 같으며, 워터마크 삽입과 추출을 식(2)와 같이 표현 할 수 있다.

$$c = E_{K_S}(x, m), \quad m = D_{K_P}(c) \tag{2}$$

식(2)에서 $E(\cdot)$는 삽입함수이고, 메시지 m은 개인키 K_S에 의해 워터마크 w로 만들어져 입력영상 x에 삽입되어 워터마크가 삽입된 영상 c가 생성된다. 영상 c에 삽입된 메시지는 공개키 K_P와 추출함수 $D(\cdot)$에 의해 추출 된다.

본 논문에서는 디지털 콘텐츠 산업의 활성화를 위해 암호화 기술을 기반으로 하는 전자상거래시스템에 쉽게 응용 가능한 비대칭 워터마킹 방법을 제안한다. 논문의 구성은 다음과 같다. 2 절에서 관련 연구를 소개하고, 3 절에서 제안한 비대칭 워터마킹 방법에서 워터마크의 삽입과 추출, 인증, 저작권 판정 방법을 기술한다. 4 절에서 본 논문에서 제안한 방법에 대한 실험결과를 보이고, 마지막으로 5 절에서 결론을 맺는다.

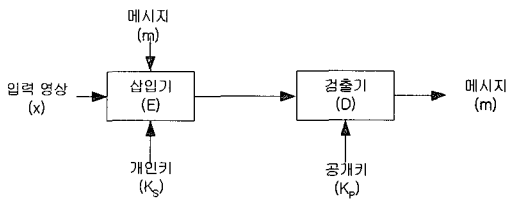


그림 1. 비대칭 워터마킹

2. 관련 연구 동향

기존 대칭 워터마킹 방법의 안전성 문제를 해결하기 위해 [5-7]과 같은 비대칭 워터마킹 방법들이 제안되었다. [5]에서는 길이 N인 *Legendre* 수열 a을 이용한 비대칭 워터마킹 방식을 제안하였다.

$$G_{DFTa} = A_1 a^* \tag{3}$$

G_{DFTa}는 *Legendre* 수열 a의 DFT(이산 푸리에 변환) 행렬이고, A_1은 복소수를 그리고, a^*는 켤레 복소

수 수열을 의미한다. 즉, *Legendre* 수열은 식 (3)과 같이 그 수열을 푸리에 변환하면 같은 수열의 켤레 형태를 얻을 수 있다. *Legendre* 수열 a를 비밀 워터마크 키로 사용하고, 수열의 길이 N을 공개 워터마크 키로 사용한다. 추출기에서는 식 (4)와 같이 전송된 신호와 전송된 신호의 푸리에 변환된 신호의 상관도를 이용하여 워터마크의 삽입 여부를 판단한다.

$$c = r^T G_{DFTr} / N \quad (4)$$

여기서, r^T는 r의 켤레 전치(*Transpose*)를 의미한다. [6]에서는 $N \times N$ 변환행렬 G의 고유 벡터 w와 일치하는 고유값 λ_0 사이에 $Gw=\lambda_0 w$의 관계를 이용한 비대칭 워터마킹 방식을 제안했다. 이 방식에서는 고유 벡터 w를 비밀 워터마크 키로, 변환 행렬 G를 공개 워터마크 키로 사용하여 삽입된 워터마크를 추출한다. 즉,

$$c = r^T G_r / N \quad (5)$$

이 되고 상관도 값을 기준값과 비교하여 추출기에서는 비밀 워터마크 정보를 사용하지 않고 추출이 이루어진다. [7]에서는 원시키 u를 먼저 생성한 뒤, 선형 랜덤 변환 행렬 A와 역행렬 A^{-t}를 이용한 비대칭 워터마킹 방식을 제안하였다. 여기에서는 $s=Au$를 비밀키로, $p=A^{-t}u$를 공개키로 사용하여 삽입된 워터마크를 추출한다. 단, 공개키는 워터마크를 제거하기위한 충분한 정보를 제공하지 않아야 한다.

$$c = p^t \times y = u^t A^{-1} x + u^t u \quad (6)$$

여기서, x 는 워터마크 삽입 대상 콘텐츠, y 는 워터마크가 삽입된 콘텐츠이다. 워터마크 추출시 공개키 p는 공개되고, 선형 랜덤 변환 행렬 A와 원시키 u는 공개되지 않는다.

[5] 방식에서의 문제점은 생성 가능한 워터마크 키의 양이 작아서 공격자가 삽입된 수열을 쉽게 찾아낼 수 있다는 것이다. 길이 N인 *Legendre* 수열은 $N-2$ 개만이 존재하므로 가능한 모든 수열을 추출기 입력으로 주고, 추출 결과를 확인함으로써 삽입된 수열을 찾아낼 수 있다. [6] 방식은 *Legendre* 수열을 이용한 방식의 연장으로 볼 수 있지만, 이러한 문제에 대해서는 좀더 안전하다고 할 수 있다. 즉, 기하학적 중복도가 매우 큰 고유 값을 갖는 고유 벡터를 워터마크로 사용하면 고유 값에 해당하는 고유벡터가 유일하게 정의되지 않고 중복도에 따라 기하급수적으로 증가하게 되므로, 공격자가 시도해 보아야 할 워터마크의 수를 증가시킬 수 있고 이러한 공격에 대해서 안전성을 높일 수 있다.

워터마크 추출 과정을 분석해 보면, [5-7] 방식들은 모두 전송된 신호와 그 신호를 변환시킨 신호와의 상관도를 이용하는 형태이다. 그러므로 이들 방식들은 추출기에서 워터마크의 존재 여부만을 알 수 있으며 어떤 워터마크가 삽입되어 있는지는 확인 할 수 없다. 즉, 이들 방식에 의한 워터마킹의 용량은 1 비트로 볼 수 있으며 이러한 작은 용량으로 가능한 매우 제한적인 응용에만 국한되어 사용할 수 있다. 기존 비대칭 방식들의 비교 내용은 위의 표 1과 같다.

3. 제안한 비대칭 워터마킹 방법

본 논문은 사용자 자신의 개인키와 공개키 인증서에 대한 사회적 책임을 전제로 디지털 콘텐츠의 소유권을 보호할 수 있는 비대칭 워터마킹 방법을 제안한다. 제안하는 방법의 이슈는 공개키 기반 구조의 개

표 1. 기존 비대칭 워터마킹 방식 비교

방식	비밀키	키용량	공개키	추출기 비용	용량	추출 형태	신호 간섭	추출성능	행렬 공개	공개키 공격
Schynde[5]	Legendre 수열	소	수열 길이	소	1	y^tAy	고	$D_{sym}/10$	공개	가능
Egger[6]	고유 벡터	중	변환 행렬	중	1	y^tAy	고	$D_{sym}/10$	공개	가능
변환키[7]	비밀 수열	대	공개 수열	대	$\log_2 N$	y^tAw	저	$D_{sym}/2$	비공개	가능
Furon	잡음 수열	대	PDS 모양	대	$\log_2 N$	y^tAy	고	$D_{sym}/10$	공개	가능
Picard	랜덤 수열	대	투사 함수, 공개 수열	대	$\log_2 N$	y^tAw	저	$D_{sym}/\sqrt{2}$	공개	가능
Smith	랜덤 수열	대	수열 길이	소	1	y^tAy	고	$D_{sym}/10$	공개	가능

념에 기초하여 워터마킹 기술과 암호화 기술을 연계하는 방법의 제안이다. 각 기술 간의 연계를 위하여 다음과 같이 두 기술 간의 차이와 특성을 충실히 고려한다. 먼저 워터마킹 기술은 워터마킹된 데이터가 원본과 크게 다르지 않아야하고, 일반적인 신호처리와 고의적인 공격을 허용하므로 삽입한 워터마크와 추출한 워터마크 간의 일대다 관계를 가지며, 저작권 판정방식 측면에서 유사성을 비교하는 상관도 비교 방식을 사용한다. 그러나 암호화 기술은 암호문 그 자체로의 의미가 부여되지 않고, 복호기에 전송되는 암호문이 변경되지 않음을 전제로 하며, 메시지와 암호문이 일대일 관계를 가진다. 본 논문에서는 두 기술 간의 근본적인 기술의 차이를 활용하여 두 기술을 자연스럽게 연계하기 위한 접근방안으로 인증을 사용하며, 인증결과에 따라 추출 방식을 달리하는 비대칭 워터마킹 방법을 제안한다. 인증의 목적은 두 데이터 간의 데이터 무결성을 검사하는 것이며, 인증의 대상은 초기 원 영상을 분석하여 생성한 인증코드와 배포된 워터마크가 삽입된 영상(*Watermark Inserted Image*, 이하 *WII*로 표기)으로부터 추출한 인증코드'이다.

제안하는 방법의 개요는 그림 2와 같다. 사용자가 분배자에게 제공하는 저작권 정보를 분배자의 개인키로 암호화하여 디지털서명을 생성하고, 원 영상을 분석하여 인증코드를 생성한다. 디지털서명과 인증코드로 워터마크를 구성하고, 워터마크를 원 영상에 삽입하여 *WII*를 생성하고 배포한다. 만일 *WII*가 신호처리와 고의적인 공격에 의해 왜곡된다면 *WII*에 삽입된 워터마크도 영향을 받으므로 인증과정에서 왜곡 유무를 검증할 수 있다. 차후에 배포된 *WII*(이하 *WII'* 로 표기)에 대한 소유권을 판정할 때, *WII'* 로부터 워터마크를 추출하고, 워터마크를 구성하는 인증코드와 디지털서명을 분리한다. *WII'* 로부터 추출한 인증코드'와 사용자가 제시하는 워터마크로부터 추출한 인증코드 간의 비교를 통해 워터마크의 데이터 무결성을 검사하고, 이를 만족하는 경우와 만족하지 못하는 경우로 구분하여 저작권 판정방식을 달리한다. 전자의 경우는 *WII'* 로부터 추출한 워터마크에서 디지털서명을 분리한 후, 이를 분배자의 공개키로 복호화하여 복호화된 저작권 정보'와 사용자가 제시하는 저작권 정보를 서로 비교함으로서 저작권을 판정하는 기존 암호화에서의 복호화 방식 *A*를 따르고, 후자의 경우는 분리한 디지털서명'과 사용자가 제시하는 워터마크로부터 분리한 디지털서명

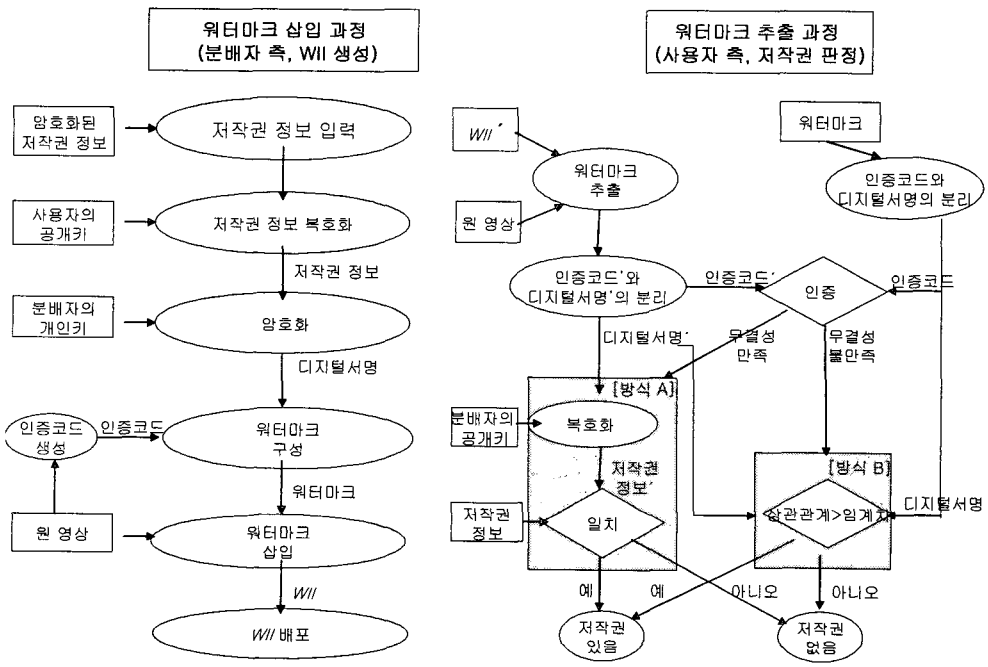


그림 2. 제안하는 비대칭 워터마킹 방법의 개요도

을 복호화 하지 않고, 디지털서명 간의 비트단위로 유사성을 비교하는 상관도 비교 방식 B로 저작권을 판정한다.

3.1 워터마크 삽입과 추출

워터마크 삽입은 [4]에서 제안한 방법을 사용하며 그 개요는 그림 3과 같다. 원 영상을 8×8 블록으로 분할하고 DCT 하여 워터마킹 정보획득 모듈과 적응적 워터마킹 모듈에 입력한다. 워터마킹 정보획득 모듈에서는 입력된 블록에 대해 텍스쳐 분석을 통해 워터마크의 삽입 위치를 결정하고 휘도와 대비 분석을 통해 삽입하는 워터마크의 강도를 결정한다. 적응적 워터마킹 모듈에서는 DCT 된 블록과 워터마킹 정보획득 모듈에서 결정된 워터마크의 삽입위치 및 삽입강도와 워터마크를 입력받아 워터마크를 삽입한다. 적응적 워터마킹 모듈은 워터마크 삽입기, 워터마크의 시각적 무감지성 검사기, 제어기로 구성되며, 삽입한 워터마크의 시각적인 무감지성과 워터마크를 제거하기 위한 공격에 대한 강인성 조건을 고려하여 두 조건을 동시에 최대한 만족하도록 워터마크를 삽입한 후 시각적 무감지성 검사결과에 따라서 제어기에 의해 적응적으로 삽입된 워터마크의 강도를 재조정한다. 워터마크 삽입 과정은 다음과 같다.

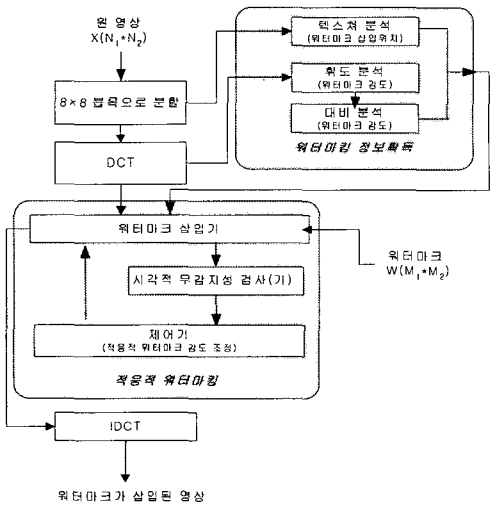


그림 3. 워터마크 삽입방법의 개요도

[step 1] 영상을 재배열한다.

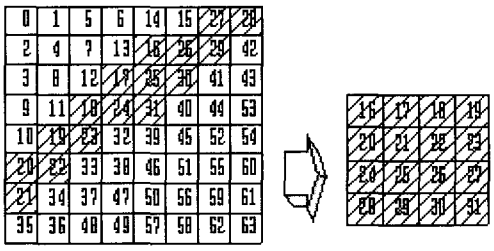
워터마크 삽입영역을 부분으로 제한하는 이유는 관심대상의 지역화(*localization*)로 연산의 효율성을 향상하기 위함이다. 영상 재배열 과정은 *8×8 DCT* 블록의 *Zig-zag* 스캔 인덱스 8번에서 31번 계수들 중 텍스쳐 분석결과에 따라 그림 4와 같이 *4×4* 블록을 만드는 과정이다. 텍스쳐 분석은 각 블록 단위에 대해 복잡도 검사를 수행하여 세 클래스로 분류하며, 분류된 블록들 가운데 단순한 블록 *Class1*에 속하는 블록들에는 삽입한 워터마크가 노출될 수 있기 때문에 워터마크를 삽입하지 않고 *Class2*와 *Class3*로 분류된 블록들에 워터마크를 삽입하며 각각의 경우 워터마크 삽입위치는 그림 4와 같다.

[step 2] 극성(*polarity*)을 계산한다.

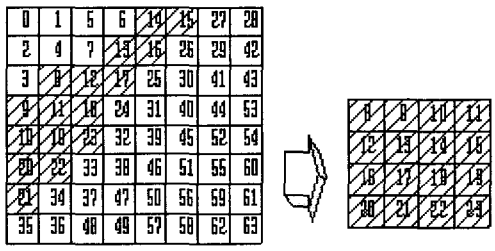
극성은 복잡도가 보통인 블록 *Class2*와 복잡도가 높은 블록 *Class3*의 각 경우를 구분하여 계산하며, 먼저 식 (7)을 만족하는 블록의 쌍을 구한다. 즉, n개의 재배열된 블록들의 집합을 R이라하고 R의 두 부분집합을 $R_a=\{a_1, a_2, ..., a_{n/2}\}$, $R_b=\{b_1, b_2, ..., b_{n/2}\}$라 할 때, 두 부분집합에 속하는 블록 간의 쌍을 이루는 방법은 식 (7)의 사상함수 $\boldsymbol{M}$에 기초한다.

$$R_a=M(R_b) \text{ 단, } R_a \cap R_b=\varnothing,\ R_a \cup R_b=U \qquad (7)$$

블록 간의 쌍을 이루는 재배열된 두 개의 *4×4* 블록



(*a*) *Class2*: 복잡도가 보통인 블록의 경우



(*b*) *Class3*: 복잡도가 높은 블록의 경우

그림 4. 영상 재배열

을 각각 A, B 라고 할 때, 두 개의 블록으로부터 만들어지는 4×4 크기의 극성은 식 (8)에서 보는 것과 같이 블록 A, B 간의 정합되는 계수의 차로부터 계산된다.

$$P_{i,j,k}=\begin{cases}1 & \text{if } A_{i,j,k}\geq B_{i,j,k}\\ 0 & otherwise\end{cases}$$
$$(\text{단, } 0\leq i,j<4, 0<k\leq n/2) \quad (8)$$

워터마크 삽입 시 극성을 이용함으로 얻을 수 있는 장점은 삽입하는 워터마크의 시각적 무감지성과 강인성을 향상시킬 수 있으며, 식(7)의 사상함수 M에 의한 워터마크의 보안성을 강화할 수 있기 때문이다.

[step 3] 워터마크를 생성한다.

워터마크는 인증코드와 디지털서명으로 구성된다. 인증코드는 데이터 무결성 검사 즉, 인증을 위한 단서로 식 (7)의 사상함수 $\boldsymbol{M}$에 의해 서로 사상되는 두 블록 간의 정합되는 DCT 계수들 간의 비교로 생성된다. 비교하는 대상계수의 위치는 8×8 DCT 블록의 지그재그 스캔 인덱스 0번에서 7번 계수이다. 그 이유는 저주파 영역의 계수 값이 에너지 집중 때문에 고주파수의 것들 보다 일반적으로 크고 $JPEG$ 손실 압축 전후에도 변화가 적기 때문이다. k 번째 서로 쌍을 이루는 두 개의 블록을 각각 p, q 라 할 때, k 번째 블록 쌍에 대한 인증코드 S_k은 식 (9)에서 보는 것과 같이 블록 p, q 간의 정합되는 DCT 계수 간의 차로부터 계산된다. 여기서 i 는 8×8 DCT 블록 p, q 상의 지그재그 스캔 인덱스이고 N 은 8×8 DCT 블록의 개수이다.

$$S_k(i)=\begin{cases}1 & \text{if } \Delta F_{p,q}(i)\geq 0\\ 0 & otherwise\end{cases},$$
$$(\text{단, } 0\leq i<8, 0<k\leq(N-1)/2) \quad (9)$$

[step 4] 워터마크를 삽입한다.

삽입할 워터마크를 극성 크기 4×4와 같이 변형하고 이를 W_k라 하고, 워터마크 삽입 기준 정보를 $\hat{p}_k$라 하면, $\hat{p}_k$은 식 (10)과 같이 극성 P_k와 W_k간의 배타적-OR($\oplus$)연산에 의해 구한다. 그림 6은 식 (10)의 과정을 보여준다.

$$\hat{P}_k = P_k \oplus W_k$$
$$\hat{P}_k = \hat{P}_{i,j,k}\ 0\leq i,j<4 \text{ and } 0<k\leq n/2) \quad (10)$$

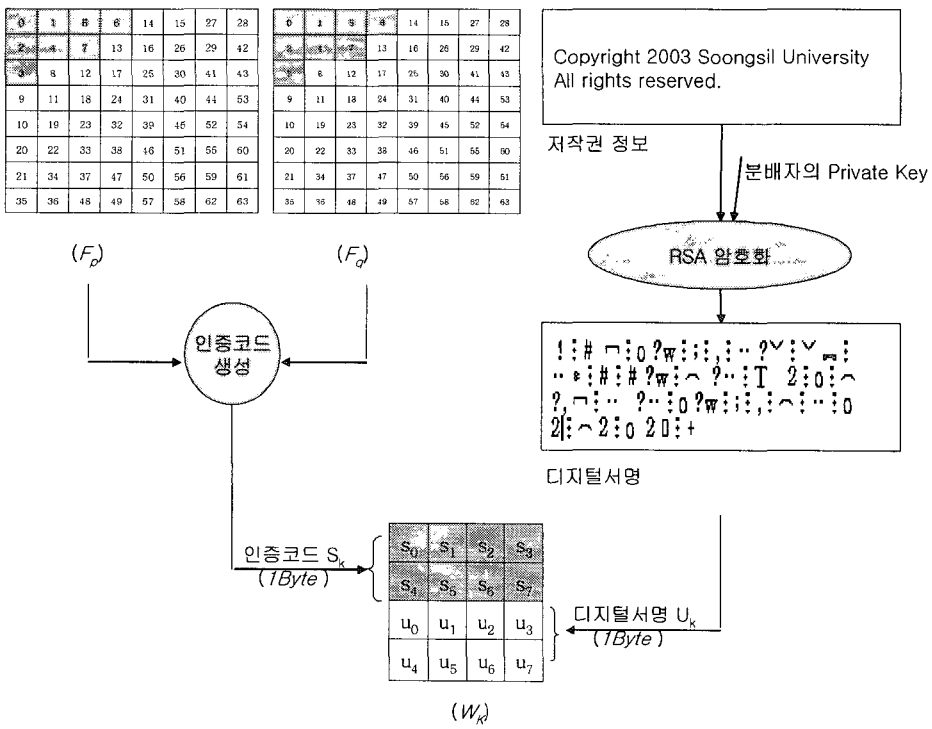


그림 5. 워터마크의 생성

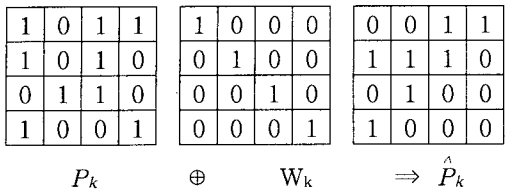


그림 6. 워터마크 삽입 기준 정보 $\hat{P}_k$ 계산

워터마크 삽입과정은 워터마크 추출 시 식 (10)에 의해 계산되는 워터마크 삽입 기준 정보인 $\hat{P}_k$가 정확히 구해지도록 기존의 *DCT* 계수를 조정하는 과정이다. 워터마크 삽입은 식 (7)에 의해 선택된 두 개의 블록을 A_k와 B_k라 하면, 단계 3에서 구해진 $\hat{P}_k$를 이용하여 기존의 A_k와 B_k의 *DCT* 계수를 조정함으로 이루어 진다. 계수 조정 후의 B_k 블록을 $\bar{B}_k$라 하면, B_k의 *DCT* 계수를 조정하는 방법은 식 (11)과 같다. 즉, $\hat{P}_k$의 값이 1이면 $(\bar{B}_k - B_k) \geq WS_k$가 되도록, $\hat{P}_k$의 값이 0이면 $(\bar{B}_k - B_k) < WS_k$가 되도록 계수 값을 변경한다.

$$\bar{B}_k = B_k + (2\hat{P}_k - 1) \times WS_k \quad (11)$$

픽셀값 $\bar{B}_k$와 *DCT* 계수 B_k 값 간의 차에 대한 절대값의 크기 $|\bar{B}_k - B_k|$ 는 식 (11)에서 워터마크 강도 WS_k이며, WS_k의 산출은 참조 [4]의 방법을 따른다. WS_k 값이 클수록 워터마크의 강인성과 추출기에서의 워터마크 추출 정확도가 향상된다.

[step 5] 영상을 역으로 재배열한다.

단계 5는 단계 1의 역 과정으로, 워터마크가 삽입된 *4×4* 블록상의 각 계수 값을 사상하는 *8×8* 블록상의 기존 계수 값으로 갱신하는 과정이다. 단계 5이후 *IDCT* 를 수행하여 워터마크가 삽입된 블록을 생성한다.

이상의 단계1에서 5까지의 과정을 반복 수행함으로 워터마크가 삽입된 영상을 생성하며, 이 알고리즘은 이론적으로 모든 워터마크 비트를 삽입할 수 있다. 그러나, 실제적으로는 워터마크의 일부 비트가 잡음에 의해 소실될 수 있으므로 이를 보완하기위해 동일한 워터마크를 원 영상에 비중첩으로 반복적하여 삽입한다.

워터마크 추출 과정은 워터마크 삽입과정과 유사하다. 배포된 *WII'*로부터 워터마크를 추출하기 위하여 블록 쌍을 결정하는 사상함수 M과 관련된 키와 워터마크 삽입위치에 관한 두 개의 사전정보가 필요하다. 워터마크 W'_k는 원 영상으로부터 극성 P_k와 *WII'*로부터 워터마크 삽입 기준 정보인 $\hat{P}_k$를 구한 후, P_k와 $\hat{P}_k$를 식 (12)에 적용하여 추출한다.

$$W'_k = P_k \oplus \hat{P}_k$$
$$W'_k = \{W_{i,j,k} \; 0 \leq i,j < 4 \text{ and } 0 < k \leq (N-1)/2\} \quad (12)$$

단, 여기서 N은 *8×8 DCT* 블록의 개수이다.

3.2 인증

최근 포토샵과 같은 강력한 편집 소프트웨어들이 많이 개발되어 영상을 인위적으로 변경하거나 위조하기가 용이해졌다. 워터마크 추출 대상에 대한 인위적 조작 여부는 워터마크 추출 성능에 큰 영향을 미친다. 인위적인 조작여부 즉, 데이터 무결성을 검증하는 기술이 인증이며, 인증의 대상은 배포된 *WII'*에 삽입된 워터마크로부터 추출한 인증코드 S'와 사용자가 제시하는 워터마크로부터 분리한 인증코드 S이다.

인증과정은 그림 7과 같으며 S를 기준으로 S와 S'와 비교하여 상이한 비트부분이 있다면 인위적인 조작이 이루어진 것으로 보고 변경된 비트부분을 표시한다. 만일 S를 구성하는 인증 코드를 $s(i)$ 그리고 S'를 구성하는 인증 코드를 $s'(i)$라 하면, 인증 평가함수(이하 *AAF*로 표기)는 식 (13)과 같이 논리적 배타연산(*Exclusive OR*)으로 정의 할 수 있다. 만일 데이터 무결성을 만족한다면 *AAF*의 값은 1이다.

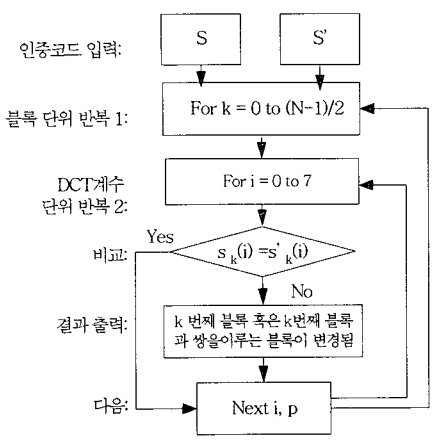


그림 7. 인증 과정

$$AAF(S,S') = \frac{1}{L}\sum_{i=1}^{L} s(i) \oplus s'(i) \quad (13)$$

단, 여기서 L은 인증코드의 비트 수이다.

본 논문에서 사용하는 인증 방법은[8]에서 제안한 디지털 서명 방법과 유사하나 [8]과는 다르게 *JPEG* 압축 전후에 불변하는 8×8 *DCT* 블록의 인덱스 0~7번 계수에 해당하는 저주파 영역의 계수를 이용하여 디지털 서명을 생성하기 때문에 *JPEG* 손실압축에도 정보손실이 없다.

3.3 저작권 판정

인증결과 데이터 무결성을 만족하는 경우와 만족하지 못하는 경우로 구분하여 저작권 판정방식을 달리한다. 전자의 경우는 암호화 기술의 특징 즉, 암호문 그 자체로는 의미가 부여되지 않고 복호기에 전송되는 암호문이 변경되지 않음을 전제하며 메시지와 암호문이 일대일 관계를 갖는 것을 만족하므로 암호화에서의 복호화 방식을 따르고, 후자의 경우는 워터마킹 기술의 특징 즉, 워터마크가 삽입된 데이터가 원본과 크게 다르지 않아야하고 일반적인 신호처리와 고의적인 공격을 허용하므로 삽입한 워터마크와 추출한 워터마크간의 일대다 관계를 가짐으로 워터마킹 기술에서의 저작권 판정방식인 상관도 비교방식으로 소유권을 판정한다.

전자의 경우는 *WII* ′로부터 워터마크를 추출하고 추출한 워터마크에서 디지털서명(*Encoding* Copyright *Information*, 이하 *ECI*로 표기)을 분리한 후 *ECI*를 분배자의 공개키로 복호화하여 복호화된 저작권 정보와 사용자가 제시하는 저작권 정보를 서로 비교함으로 저작권을 판정하고, 후자의 경우는 *WII* ′로부터 추출한 워터마크에서 분리한 *ECI* ′와 사용자가 제시하는 워터마크에서 분리한 *ECI* 간의 비트단위로 유사성을 비교하는 상관도 비교 방식으로 저작권을 판정한다. 상관도 비교 식은 (14)와 같고, 구해진 상관도 *Corr*값이 사전에 정의된 임계값 이상이면 저작권을 인정하며 임계값 보다 작을 경우 상관관계가 없다고 간주해서 저작권을 인정하지 않는다.

$$Corr = \frac{\sum_i ECI_i \cdot ECI_i'}{\sum_i [ECI_i]^2} \quad (14)$$

4. 실험 결과

실험 환경은 인텔 *Pentium 4 CPU 1.7GHz, 256 MB RAM* 시스템과 윈도우 *XP*운영체제에서 *Visual C++ 6.0*으로 프로그램을 작성하였다. 실험에서는 제안하는 공개키 기반 구조를 이용한 비대칭 워터마킹 방법의 성능을 검증하기 위하여 *512×512* 크기의 *Lena*와 *Babara* 영상을 사용하였다.

먼저 실험에서는 식 (7)의 사상함수에 의해 사상되는 두 *DCT* 블록, 그림 8 (*a*)의 오른쪽 눈 영역에 해당하는 2033, 2034 번째 *DCT* 블록을 예로 들어 2034 번째 블록에 워터마크를 삽입하는 과정을 기술한다. 실험에서는 *128 bits* 블록크기의 *RSA*암호화 알고리즘에 의해 저작권 정보를 암호화하여 디지털 서명을 생성한다. 이때 사용되는 분배자의 개인키는 {*p*=17, *q*=23, *D*=8191}이고 저작권 정보는 "*Copyright 2003 Soongsil University All rights reserved.*"와 같은 55 개의 문자로 구성된 문자열이다. 디지털서명 즉, 암호화된 저작권 정보를 *Hex Code*로 표현하면 "00 21 01 23 00 b6 00 1a 01 6f 00 c1 77 01 3b ..."와 같으며 저작권 정보를 구성하는

(a)

727	-188	-3	-28	-16	-4	-6	-1
51	-77	22	6	11	1	2	3
31	-52	-28	-8	5	5	10	7
12	31	-21	-7	1	-13	-2	-2
19	14	-23	-17	4	2	2	-1
21	15	-2	-17	-5	2	0	-1
15	16	13	1	2	6	-2	0
-1	-3	-6	-12	-6	-1	1	3

(b)

456	91	-65	-91	-47	-1	14	0
140	41	44	6	-6	-12	-6	-4
43	108	-54	5	16	13	-9	0
-148	-21	84	34	22	0	-12	6
4	-13	-2	-32	8	5	5	12
-28	-4	1	-1	-8	1	2	0
1	10	-14	4	6	-1	-1	-6
-8	-10	14	3	-1	-2	-2	-3

(c)

그림 8. (a) Lena 영상, (b) a 영상의 2033 번째 DCT 블록, (c) a 영상의 2034 번째 DCT 블록

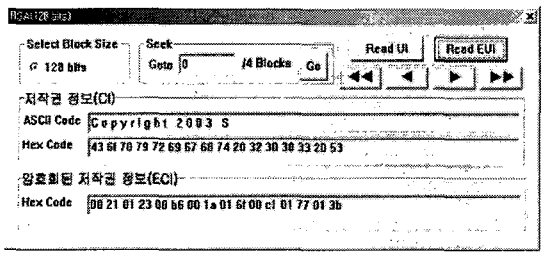


그림 9. 분배자의 개인키에 의한 저작권 정보 암호화

각 문자 단위의 *Hex Code*는 1 *byte*의 이진코드로 변환되어 워터마크의 구성 정보로 사용된다. 2033, 2034 번째 *DCT* 블록에 문자 "C"를 삽입한다면, 문자 "C"에 대한 *Hex Code*는 43이며 *Hex Code* 43을 *RSA* 암호화 한 후 획득되는 Hex Code는 00이다. Hex Code 00을 1 *byte* 이진코드로 변환하면 $00000000_{(2)}$이다.

데이터 무결성을 검증하기 위한 인증코드를 생성하기 위해 2033, 2034 번째 *DCT* 블록을 식 (9)에 대입하면 산출된 인증코드 $s_0, s_1, s_2, s_3, s_4, s_5, s_6, s_7$은 1, 0, 0, 0, 0, 1, 1, 1이다.

산출된 1 *byte*의 디지털서명과 1 *bytes*의 인증코드로 *4×4* 크기의 워터마크, W_k를 구성하면 그림 10과 같다.

워터마크의 삽입위치는 그림 8 (*b*), (*c*)의 음영부분이 워터마크 삽입 후보영역이 되며 본 실험에서 실제 워터마크 삽입위치는 그림 8 (*c*)의 음영부분이다. 그림 8 *Lena* 영상의 2033과 2034번째 *DCT* 블록에 대해 식 (8)에 따라 극성정보 P_k를 구하고 구해진 P_k와 워터마크 W_k를 식 (10)에 적용하여 워터마크 삽입 기준 정보 $\hat{P}_k$를 계산한다.

그림 8 (*c*) 블록을 *4×4* 크기로 재배열한 블록의

1	0	0	0
0	1	1	1
0	0	0	0
0	0	0	0

그림 10. 워터마크 W_k

0	1	1	1
0	1	1	0
1	0	0	1
1	1	1	0

P_k ⊕

1	0	0	0
0	1	1	1
0	0	0	0
0	0	0	0

W_k ⇒

1	0	0	0
0	1	1	1
1	0	0	0
1	1	1	1

$\hat{P}_k$

그림 11. 워터마크 삽입 기준 정보, $\hat{P}_k$ 계산

(i, j) 위치에 워터마크의 시각적 무감지성을 유지하며 할당 가능한 최대 워터마크 강도 값, $WS_{i,j}$ 그림 12와 같다. $WS_{i,j}$의 산출은 참조 [4]의 방법을 따른다. 계산된 $WS_{i,j}$와 $\hat{P}_k$를 워터마크 삽입 식 (11)에 대입하여 워터마크를 삽입하고 다시 *8×8 DCT* 블록으로 계수들을 재배열하여 워터마크를 삽입한 결과는 그림 13과 같다. 이상의 3.1절에서 제시한 워터마크 삽입 알고리즘을 모든 블록에 반복하여 그림 14와 같은 워터마크가 삽입된 *DCT* 주파수 공간을 획득하고 이를 *IDCT*하여 그림 15의 (*c*)와 같은 결과영상을 획득한다. 실험에서 워터마크의 시각적 무감지성을 고려하지 않고 워터마크를 삽입할 경우, *512×512* 크기의 영상을 구성하는 8×8 DCT 블록 쌍에 대해 한 문자를 삽입하므로 저작권 정보 "*Copyright 2003 Soong-*

25.35	34.56	5.54	9.74
14.87	12.49	8.84	2.56
4.82	3.75	27	8.61
11.89	3.85	5.74	2.21

그림 12. 최대 워터마크 강도값, $WS_{i,j}$

486	91	-66	-91	-8	-3	14	0
140	41	44	47	-3	1	-6	-4
43	81	39	1	56	13	-9	0
-108	-12	56	-9	22	0	-12	6
14	-9	-4	-32	8	5	5	12
-11	-3	1	-1	-8	1	2	0
8	10	-14	4	6	-1	-1	-6
-8	-10	14	3	-1	-2	-2	-3

그림 13. 워터마크가 삽입된 그림 8의 (c) 블록

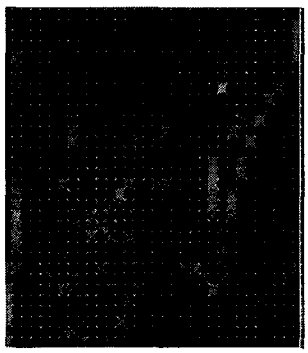


그림 14. 워터마크가 삽입된 *DCT* 주파수 공간

sil University All rights reserved."는 비중첩으로 중복하여 37회 삽입 된다. 그러나 워터마크의 시각적 무감지성을 고려할 경우, 텍스쳐 분석결과 복잡도가 낮은 *Class1*에 속하는 블록들을 제외하므로 저작권 정보는 18회 삽입된다.

인증과정에서는 분배된 워터마크가 삽입된 영상, *WII'*으로부터 워터마크를 추출하고 추출한 워터마크로부터 분리한 인증코드 S'와 사용자가 제시하는 워터마크로부터 분리한 인증코드 S에 대해 그림 7의 방식으로 데이터 무결성을 검사한다.

*WII'*가 공격 받지 않았을 경우, 데이터 무결성을 만족하므로 워터마크 추출이후 저작권 판정 방식은 암호화에서의 복호화 방식을 따른다. 모든 블록에서 추출한 저작권 정보를 저작권 정보의 길이 단위로 분배자의 개인키 $\{p=17, q=23, D=8191\}$에 대응되는 공개키는 $\{N=391, E=63\}$로 복호화한 결과는 결국 원 영상에 워터마크로 구성하여 삽입한 저작권 정보와 일치하여 저작권인정 판정을 받는다.

반면에 *WII'*가 공격 받았을 경우, 데이터 무결성을 만족하지 않으므로 워터마크 추출이후 저작권 판정 방식을 유사성을 비교하는 상관도 비교 방식을 사용한다. 실험에서는 본 논문에서 제안한 방법으로 워터마크를 삽입한 그림 15의 (c)와 그림 16의 (c)에 대해 공격의 유형에 따른 실험결과를 제시하며, 공격의 유형은 *JPEG* 손실압축, 잡음 추가, 대비변화, 밝기 변화, 부분 편집이다. *Lena*와 *Babara* 영상에 대한 공격에 대한 실험 결과를 그림 17에서 보이고 방법 [5]와 제안한 방법간의 성능 비교를 표 3에서 보인다.

(a) *Lena* 영상 (b) [5]의 방법으로 워터마크가 삽입된 영상 (c) 제안하는 방법으로 워터마크가 삽입된 영상

그림 15. 워터마크의 시각적 무감지성 비교

(*a*) *Babara* 영상 (*b*) [5]의 방법으로 워터마크가 삽입된 영상 (*c*) 제안하는 방법으로 워터마크가 삽입된 영상

그림 16. 워터마크의 시각적 무감지성 비교

표 3. 기존방법[5]과 제안방법의 성능 비교

공격유형		워터마크 검출율(%)			
		방 법[5]		제안된 방법	
		Lena	*Babara*	*Lena*	*Babara*
영상 강화	대비값 5증가	96.3	95.2	100	100
	대비값 10 증가	96.2	95.0	100	100
	밝기값 20 감소	98.3	98.1	100	100
	밝기값 20 증가	98.4	99.5	100	100
노이즈 추가	노이즈 비율 10%	95.3	97.3	99.2	99.9
	노이즈 비율 20%	90.2	89.6	96.2	98.0
	노이즈 비율 30%	82.5	85.4	95.9	93.7
JPEG 압축	압축율 9.1	54.9	53.1	67.6	70.1
	〃 8.4	58.2	57.4	74.2	73.6
	〃 7.1	65.6	70.6	98.4	95.5
	〃 5.9	78.9	75.2	99.5	98.1
	〃 4.7	91.2	80.3	99.7	99.5
	〃 3.5	98.4	99.2	100	99.5
	〃 2.5	99.9	100	100	100

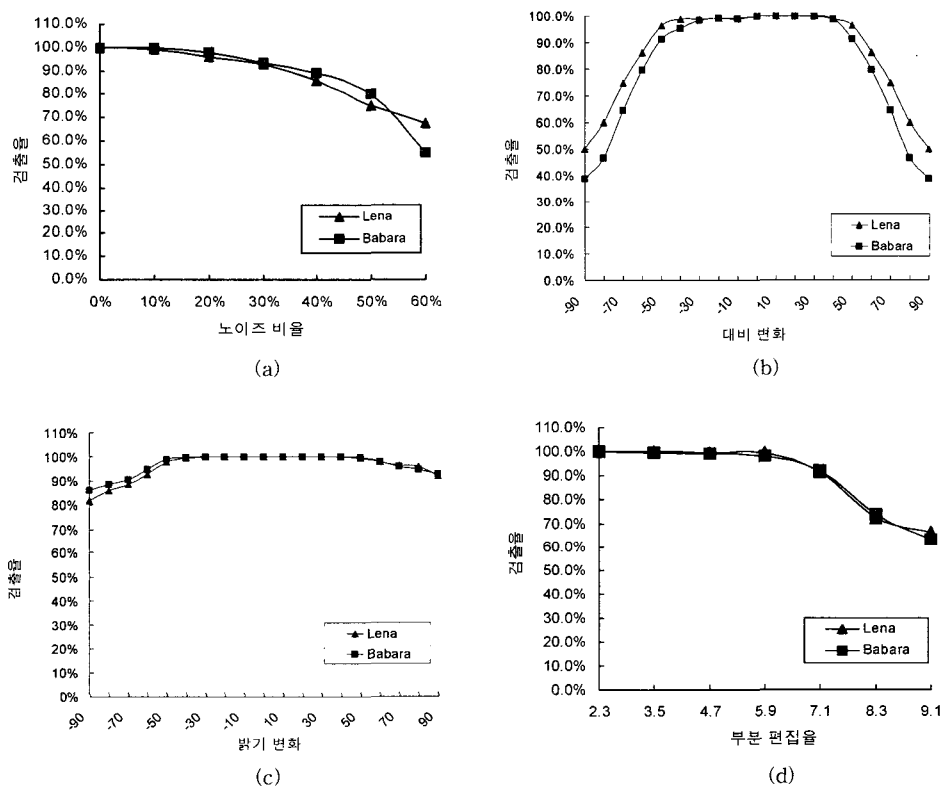


그림 17. 워터마크 추출: (a) 노이즈 비율에 따른 상관도의 변화 (b) 대비변화에 따른 상관도의 변화 (c) 밝기변화에 따른 검출율의 변화 (d) 부분 편집율에 따른 검출율의 변화

*Lena*와 *Babara* 각 영상에 대한 *JPEG*압축률에 따른 상관도 추출 실험결과는 표 3과 같다. 표 3에서 워터마크 검출율(%)은 *Corr*×100이다.

5. 결 론

본 논문에서는 서로 상이한 암호화 기술과 워터마

킹 기술을 각 기술 간의 특성에 기초하여 인증 기술로 연계한 워터마킹 방법을 제안한다. 제안하는 공개키 기반 구조를 이용한 비대칭 워터마킹 방법을 이용하면 저작권자는 자신의 개인키와 공개키 인증서에 대한 사회적인 책임을 전제로 워터마크가 삽입된 디지털 콘텐츠의 저작권을 보호받을 수 있다. 또한 *RSA* 암호화 알고리즘에서 키의 길이를 확장하여 안전성을 더욱 강화할 수 있으며 암호화 기술을 기반으로 하는 전자상거래 시스템에 쉽게 응용 가능하여 디지털 콘텐츠 산업의 활성화에 기여할 수 있을 것으로 사료된다.

참 고 논 문

[1] F. Hartung and M. Kutter, "Multimedia Watermarking Technique," Proc. IEEE, Vol. 87, pp. 1079-1107, 1999.

[2] E. Koch and J. Rindfrey and, J. Zhao, "Copyright Protection for Multimedia Data," In Digital Media and Electronic Publishing. Academic Press, London, pp. 203-213, 1996.

[3] A. Bors and I. Pitas, "Image Watermarking Using DCT Domain Constraints," in Proc. IEEE Int. Conf. on Image Processing, Lausanne, Switzerland, pp. 231-234, September 1996.

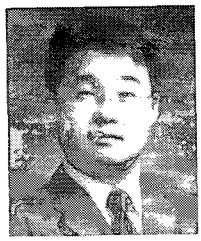
[4] 전영민, 김계영, 최형일, "DCT 기반 워터마킹의 적응적 강인화 방법(A adaptive robust method of DCT-based watermarking," 한국정보처리학회, 논문지 B 제6호, pp. 629-638, 2003. 10.

[5] R. G. van Schyndel, A. Z. Tirkel, and I. D. Svalbe, "Key independent watermark detection", In Proc. of the IEEE Intl. Conf. on Multimedia Computing and Systems, vol. 1, Florence, Italy, June 1999.

[6] J. J. Eggers, J. K. Su, and B. Girod, "Public key watermarking by eigenvectors of linear transforms", In Proc. of European Signal Processing Conf., Tampere, Finland, April 2000.

[7] H. Choi, K. Lee, and T. Kim, "Transformed-key asymmetric watermarking system", in Proc. of SPIE: Security and Watermarking of Multimedia Contents, vol. 4314, pp. 280-289, San jose, USA, Jan. 2001.

[8] G. L. Friedman, "The trustworthy digital camera: Restoring credibility to the photographic image," IEEE Trans. Consumer Electron., vol. 39, pp. 905-910, Nov. 1993.



전 영 민

1997년 2월 군산대학교 컴퓨터과학과 졸업(이학사)
1999년 2월 숭실대학교 대학원 컴퓨터학과 졸업(공학석사)
1999년 2월~현재 숭실대학교 대학원 컴퓨터학과 박사 수료
관심분야 : 컴퓨터 비전, 패턴인식, 디지털 워터마킹, 인터페이스 에이전트 등.

양 선 옥

1991년 2월 숭실대학교 전자계산학과(공학사)
1993년 2월 숭실대학교 전자계산학과 대학원(공학석사)
2000년 2월 숭실대학교 컴퓨터학과 대학원(공학박사)
1997년 9월~현재 숭실대학교 전자계산원 전임교수
관심분야 : 원격교육, 웹에이전트, 멀티미디어 등

김 계 영

1990년 2월 숭실대학교 전자계산학과 졸업(공학사)
1992년 2월 숭실대학교대학원 컴퓨터학과 졸업(공학석사)
1996년 2월 숭실대학교대학원 컴퓨터학과 졸업(공학박사)
1996년 3월~1997년 11월 한국전자통신연구원(Post Doc.)
1997년 12월~2001년 2월 한국전력공사 전력연구원(선임연구원)
2001년 3월~현재 숭실대학교 컴퓨터학부(조교수)
관심분야 : 컴퓨터비전, 형태인식, 생체인식, 증강 현실, 영상 및 신호처리 등