

# 웨이블릿 영역에서 통계적 판정법을 이용한 고신뢰 워터마크 검출 알고리즘

권성근<sup>\*</sup> · 김병주<sup>\*\*</sup> · 이석환<sup>\*\*</sup> · 권기규<sup>\*\*</sup> · 김영춘<sup>\*\*\*</sup> · 권기룡<sup>\*\*\*</sup> · 이건일<sup>\*\*\*\*</sup>

## 요 약

멀티미디어에 삽입된 워터마크의 검출은 저작권 보호 및 인증 분야에서 매우 중요한 역할을 한다. 최근 워터마크의 검출에 많이 사용되는 유사도 기반 알고리즘은 대상 영상의 분포 특성을 이용하지 않기 때문에 검출 성능이 떨어지는 단점을 가진다. 따라서 본 논문에서는 웨이블릿 변환 영역에서 효율적인 워터마크 검출 알고리즘을 제안하였다. 제안한 워터마크 검출 알고리즘은 통계적 판정법에 근거하여 Bayes 판정 이론, Neyman-Pearson 정리, 및 웨이블릿 계수들의 분포 모델을 이용한다. 따라서 제안한 검출 알고리즘에서는 주어진 오류 검출 확률에 대하여 간과 검출 확률을 최소화할 수 있는 장점이 있다. 제안한 검출 알고리즘의 성능 평가는 견고성 측면에서 수행되었고, 실험 결과로부터 제안한 알고리즘이 유사도 기반 알고리즘에 비하여 우수한 성능을 나타냄을 확인하였다.

## Highly Reliable Watermark Detection Algorithm using Statistical Decision Method in Wavelet Domain

Seong-Geun Kwon<sup>\*</sup>, Byung-Ju Kim<sup>\*\*</sup>, Suk-Hwan Lee<sup>\*\*</sup>, Kee-Koo Kwon<sup>\*\*</sup>,  
Young-Choon Kim<sup>\*\*\*</sup>, Ki-Ryong Kwon<sup>\*\*\*</sup> and Kuhn-Il Lee<sup>\*\*\*\*</sup>

## ABSTRACT

Watermark detection has a crucial role in copyright protection and authentication for multimedia. Because the correlation-based algorithm which has widely been used in the watermark detection doesn't utilize the distributional characteristics of cover image to be marked, its performance is not optimum. So a new detection algorithm is proposed which is optimum for multiplicative watermark embedding. By relying on statistical decision method, the proposed method is derived according to the Bayes decision theory, Neyman-Pearson criterion, and distribution of wavelet coefficients, thus permitting to minimize the missed detection probability subject to a given false detection probability. The superiority of the proposed method has been tested from a robustness perspective. The results confirm the superiority of the proposed technique over classical correlation-based method.

**Key words:** watermark, Bayes decision theory, Neyman-Pearson criterion

---

접수일 : 2002년 9월 14일, 완료일 : 2002년 11월 5일  
본 연구는 정보통신부의 정보통신기술기초연구지원사업  
(정보통신연구진흥원)으로 수행한 결과임.  
<sup>\*</sup> 정희원, 삼성전자 무선 통신 사업부  
<sup>\*\*</sup> 정희원, 경북대학교 전자공학과 박사과정  
<sup>\*\*\*</sup> 영동대학교 정보전자공학부 조교수  
<sup>\*\*\*\*</sup> 중신회원, 부산외국어대학교 컴퓨터전자공학부 부교수  
<sup>\*\*\*\*\*</sup> 경북대학교 공과대학 전자공학과 교수

## 1. 서 론

디지털 데이터는 아날로그 데이터에 비하여 품질이 우수하고, 편집 및 복사가 용이하며, 복사 시 원 데이터에 대한 충실도가 떨어지지 않는다. 그러나, 이는 인터넷 등의 네트워크에 의하여 빠르고 넓게 배포되므로 저작권 보호 (copyright protection)에 어려움이 따른다. 따라서 디지털 데이터의 소유자 또는 분배자의 개인 정보 즉 워터마크 (watermark)를 데이터 내에 삽입시켜 저작권을 보호하는 디지털 워터마킹 (digital watermarking)에 관한 연구가 최근 급속히 이루어지고 있다.

일반적으로 워터마킹 알고리즘은 워터마크의 삽입 영역 및 삽입 방법에 따라 분류될 수 있다. 워터마크의 삽입 영역으로는 공간 영역 (spatial domain)[1, 2] 및 주파수 영역[3,4], 그리고 이들의 혼성 영역[5, 6]으로 나눌 수 있다. 공간 영역 및 주파수 영역의 알고리즘에서는 영상의 회색 준위 및 주파수 변환 계수들을 각각 변형시켜 워터마크를 삽입하고, 혼성 영역의 알고리즘에서는 공간 성분 및 주파수 성분을 동시에 고려하는 이산 웨이블릿 변환 (discrete wavelet transform, DWT) 등의 변환 계수들을 변형시켜 워터마크를 삽입한다.

워터마킹 알고리즘을 삽입 방법에 따라 합의 형태 (additive) 알고리즘 및 곱의 형태 (multiplicative) 알고리즘으로 분류할 수 있다. 합의 형태 워터마킹에서는 대상 원 신호에 워터마크를 가산하며, 이 경우 원 신호의 크기에 관계없이 워터마크를 삽입한다. 이 방법은 수행 과정은 간단하나, 워터마크가 삽입될 계수들의 크기에 따라 삽입 강도가 결정되는 신호 적응적 워터마킹과 인간 시각 시스템 (human visual system)의 특징들을 이용하는 워터마킹에는 부적합하다. 따라서 대상 신호의 크기에 비례하여 워터마크를 삽입하여, 비가시성과 견고성을 동시에 만족시킬 수 있는 곱의 형태 워터마킹이 일반적으로 사용된다.

워터마킹 알고리즘의 필요 조건들은 이의 적용 분야에 따라 다소 다르지만, 삽입된 워터마크의 고신뢰 검출은 모든 분야에서 필요하다. 가장 널리 사용되는 워터마크 검출 방법으로는 워터마크가 삽입된 계수들과 워터마크와의 유사도 (correlation)를 구하는 방법[7,8]이 있다. 그러나 이 방법은 검사 대상 신호 및 워터마크의 통계적 분포 특성을 이용하지 않은 유사

도의 설정 및 워터마크를 제거하려는 공격을 고려하지 않은 판정 문턱값의 설정으로 인하여, 워터마크 검출 성능이 열화되는 단점을 가진다. 따라서 고신뢰 워터마크 검출 척도의 개발과 효과적인 판정 문턱값의 설정에 대한 연구가 필요하다.

한편, DWT는 HVS와 매우 유사한 다해상도 특성을 나타내고, 공간 및 주파수 영역에서 동시에 국부성 (locality)을 가짐으로써, 비정상 (non-stationary) 신호를 처리하는데 매우 효과적이다. 따라서 이를 워터마킹 분야에 이용하면 워터마크를 효과적으로 삽입 및 검출할 수 있을 것으로 판단된다.

본 논문에서는 DWT 영역에서의 통계적 판정법 (statistical decision method)을 이용한 고신뢰 워터마크 검출 알고리즘을 제안하였다. 제안한 워터마크 검출 알고리즘에서는 통계적 판정법에 근거하여 Bayes 판정 이론[9], Neyman-Pearson 정의[10], 및 DWT 계수들의 분포 모델을 이용한다. 즉 정확한 판정 규칙의 설정에 Bayes 판정 이론에 따른 우도비 (likelihood ratio)를 이용한다. 이때 우도비에는 DWT 계수들의 확률 밀도 함수 (probability density function, PDF)가 필요한데, 제안한 방법에서는 실제 이와 가장 유사한 분포를 나타내는 가우스 분포 (Gaussian distribution)로 이를 모델링한다. 또한 우도비 값과 비교되어 워터마크의 존재를 판정하는데 사용되는 판정 문턱값을 설정하기 위하여 최대 허용 허위 검출 확률 (false alarm probability)에 대하여 간과 검출 확률 (missed detection probability)을 최소화하는 Neyman-Pearson 정리를 이용한다.

본 논문에서 제안한 고신뢰 워터마크 검출 알고리즘의 성능을 평가하기 위하여 다양한 정지 영상 및 동영상에 대하여 컴퓨터 모의 실험을 행하였다. 이 실험에서는 제안한 알고리즘과 기존의 유사도 기반 알고리즘의 워터마크 검출 성능에 대하여 평가하였다. 실험 결과로부터 제안한 고신뢰 워터마크 검출 알고리즘은 기존의 유사도 기반 알고리즘에 비하여, 다양한 공격을 당한 영상들에 삽입된 워터마크를 정확하게 검출하는 것을 확인할 수 있었다.

## 2. 기존의 유사도 기반 알고리즘

일반적으로 사용되는 곱의 형태 워터마크 삽입 방법은

$$y_i = x_i(1 + \alpha w_i^*) \quad (1)$$

와 같고, 여기서  $y_i$ ,  $x_i$ ,  $\alpha$ ,  $w_i^*$ , 및  $i$ 는 각각 워터마크가 삽입된 DWT 계수의 크기, 원 DWT 계수의 크기, 워터마크 삽입 강도, 워터마크의 크기, 및 삽입 위치를 나타낸다.

유사도 기반 알고리즘[7,8]에서는 대상 영상의 계수들과 워터마크와의 유사도를 구한 후, 이 유사도 값을 판단 문턱값과 비교하여 주어진 워터마크  $w^*$ 의 존재 유무를 결정한다. 대상 영상에 대한 워터마크의 존재 유무를 확인하기 위하여 아래와 같이 두 가지 가정을 둘 수 있다.

$H_0$ : 대상 영상이  $w^*$ 을 포함하지 않는다.

$H_1$ : 대상 영상이  $w^*$ 을 포함한다.

$w^*$ 와 대상 영상에 대한 DWT 계수들의 벡터  $y$ 의 유사도는

$$\rho(w^*, y) = \frac{w^* \cdot y}{N} = \frac{1}{N} \sum_{i=1}^N w_i^* \cdot y_i \quad (2)$$

와 같고, 이 값을 판단 문턱값  $T_\rho$ 과 비교함으로써 대상 영상에 대한 워터마크  $w^*$ 의 존재 유무를 판단한다. 여기서 중심 제한 정리 (central limit theorem) [11]에 의하여  $\rho(w^*, y)$ 는 가우스 분포를 따르게 되고, 위에서 설정한 두 가지 가정에 대한  $\rho$ 의 평균 및 분산은 각각

$$m_\rho = \begin{cases} 0, & \text{if } H_0 \text{ is in force} \\ \frac{\alpha m_x \sigma_w^2}{N}, & \text{if } H_1 \text{ is in force} \end{cases} \quad (3)$$

$$\sigma_\rho^2 = \frac{\sigma_w^2 \sigma_x^2 (1 + \alpha^2 \sigma_w^2)}{N} \quad (4)$$

와 같다.

이러한 유사도 기반 검출 알고리즘은 유사도의 계산 과정이 간단하지만, 대상 영상 및 워터마크의 통계적 특성을 고려하지 않기 때문에 이용하지 않기 때문에 검출 성능이 떨어진다는 단점을 가진다.

### 3. 제안한 워터마크 검출 알고리즘

제안한 워터마크 검출 알고리즘에서는 디지털 통

신 시스템에서 사용되는 검파 이론을 이용하여 대상 신호에 삽입된 워터마크를 검출한다. 통신 시스템 및 워터마킹 시스템의 일반적인 모델은 그림 1에서와 같다. 통신 시스템은 메시지를 전송하는 송신기 (transmitter), 전송 잡음 (channel noise), 및 메시지를 복원하는 수신기 (receiver)로 구성된다. 이와 유사하게 워터마킹 시스템은 대상 신호에 워터마크를 삽입하는 삽입기 (embedder), 워터마크를 제거하려는 각종 공격, 및 워터마크를 추출하는 검출기 (detector)로 구성된다. 따라서 제안한 워터마크 검출 알고리즘에서는 워터마크를 메시지로 모델링하고, 대상 신호를 워터마크를 전송하는 전송자 (carrier)로 모델링하며, 워터마크를 제거하려는 각종 공격을 전송 잡음으로 모델링한다.

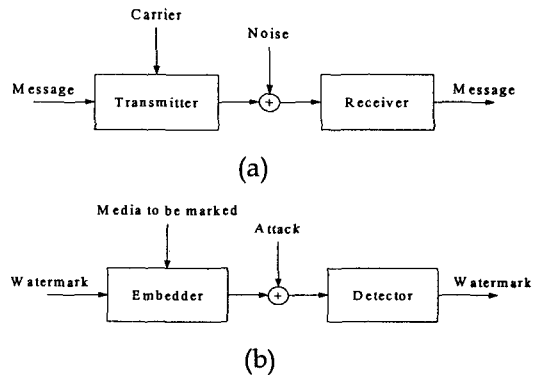


그림 1. (a) 통신 시스템 및 (b) 워터마킹 시스템의 표준 모델들

#### 3.1 판정 규칙

제안한 검출 알고리즘에 대한 판정 규칙의 흐름도는 그림 2에서와 같다. 먼저 Bayes 판정 이론을 이용하여 대상 신호  $y$ 에 대한 우도비  $l(y)$ 를 설정한다. 이 우도비에는 DWT 계수의 PDF가 필요한데, 제안한 방법에서는 이를 가우스 분포로 모델링한다. 이때  $\lambda$ 는 우도비와 비교되는 판정 문턱값을 나타낸다.

가우스 분포로 모델링된 DWT 계수의 PDF를 포함하는 우도비는 지수 함수의 형태이기 때문에, 이를 쉽게 해석하기 위하여 우도비 및 판정 문턱값에 로그를 취하여 로그 우도비  $\ln l(y)$  및 로그 문턱값  $\ln \lambda$ 를 생성한다. 그리고  $\ln l(y)$ 를 더욱 더 간단히 표현하여  $z(y)$ 를 유도하고,  $z(y)$ 와 비교되는 판정 문턱값  $\lambda'$ 을 설정하기 위하여 Neyman-Pearson 정리를 이

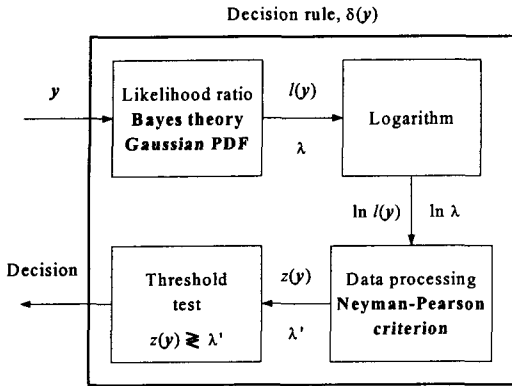


그림 2. 제안한 판정 규칙의 흐름도

용한다. 따라서 제안한 위터마크 검출 알고리즘에서는  $z(y)$ 와  $\lambda'$ 을 비교함으로써 대상 신호  $y$ 에 특정 위터마크의 존재를 판정한다.

대상 신호  $y$ 에 대하여 특정 위터마크  $w^*$ 의 존재 유무를 확인하기 위하여 아래와 같이 위터마크 존재에 대한 두 가지 가설  $H_0$  및  $H_1$ 을 설정할 수 있다.

$H_0$ : DWT 계수  $y$ 가 특정 위터마크  $w^*$ 를 포함하지 않는다.

$H_1$ : DWT 계수  $y$ 가 특정 위터마크  $w^*$ 를 포함한다.

Bayes 판정 이론에 따른 판정 규칙  $\delta(y)$ 는

$$\delta(y) = \begin{cases} 1, & y \in R_1 \text{ (} H_1 \text{ is true)} \\ 0, & y \in R_0 \text{ (} H_0 \text{ is true)} \end{cases} \quad (5)$$

와 같고, 여기서  $R_1$  및  $R_0$ 은 각각 가설  $H_1$  및  $H_0$ 에 대한 수락 영역이며,  $R_1$ 은

$$R_1 = \left\{ y : \frac{f_y(y|W_1)}{f_y(y|W_0)} > \frac{p_0 L_{10}}{p_1 L_{01}} \right\} \quad (6)$$

로 표현된다. 여기서  $f_y(y|w)$ 는 사상(event)  $w$ 에 대한  $y$ 의 조건부 PDF이고,  $W_1$  및  $W_0$ 은 검사할 위터마크  $w^*$ 를 나타내는 공간 및 이를 포함하지 않는 공간을 나타낸다. 그리고  $p_0$  및  $p_1$ 은 각각 가설  $H_0$  및  $H_1$ 이 발생할 확률이고,  $L_{10}$  및  $L_{01}$ 는 각각  $L[W_0, \delta(y) = 1]$  및  $L[W_1, \delta(y) = 0]$ 을 나타내며, 이는 위터마크 존재를 잘못 판정했을 경우의 손실(loss)을 의미한다. 따라서 식 (6)을 식 (5)에 적용하

면  $\delta(y)$ 는

$$\delta(y) = \begin{cases} 1, & \frac{f_y(y|W_1)}{f_y(y|W_0)} > \frac{p_0 L_{10}}{p_1 L_{01}} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

와 같게 된다. 이때  $\frac{f_y(y|W_1)}{f_y(y|W_0)}$  및  $\frac{p_0 L_{10}}{p_1 L_{01}}$ 를 우도비

$l(y)$  및 판정 문턱값

$\lambda$ 로 정의한다.

위터마크 성분들은  $[-1, 1]$  구간에서 균일하게 분포한다고 가정한다면  $f_y(y|W_0)$ 는 전체 확률 정리(total probability theorem)[11]에 의하여

$$f_y(y|W_0) = \int_{[-1,1]^N - w^*} f_y(y|w) f_w(w) dw \quad (8)$$

와 같이 나타낼 수 있다.  $W_0$ 은 전체 위터마크 공간  $W$ 에서  $w^*$ 를 뺀 공간이므로, 식 (8)의 적분 영역은  $[-1, 1]^N - w^*$ 와 같다. 이 식의 적분 영역 중  $w^*$ 는 연속적인 공간  $[-1, 1]^N$ 에서 하나의 포인트이기 때문에 측정 이론(theory of measure)<sup>[12]</sup>을 이용하면

$$f_y(y|W_0) = \int_{[-1,1]^N} f_y(y|w) f_w(w) dw \quad (9)$$

을 얻을 수 있다.

또한  $w$  및  $y$ 의 성분들이 서로 독립적으로 분포한다면 이들의 PDF  $f_w(w)$  및  $f_y(y|w)$ 는 각각

$$f_w(w) = \prod_{i=1}^N f_{w_i}(w_i) \quad (10)$$

$$f_y(y|w) = \prod_{i=1}^N f_{y_i}(y_i|w_i) \quad (11)$$

와 같다. 그리고  $w$ 는  $[-1, 1]$ 에서 균일하게 분포한다고 가정하였으므로 식 (10)은

$$f_w(w) = \frac{1}{2^N} \quad (12)$$

와 같고, 결과적으로 우도비  $l(y)$ 는

$$l(y) = \frac{f_y(y|w^*)}{\frac{1}{2^N} \int_{[-1,1]^N} f_y(y|w) dw} = \frac{\prod_{i=1}^N f_{y_i}(y_i|w_i^*)}{\frac{1}{2^N} \prod_{i=1}^N \int_{-1}^1 f_{y_i}(y_i|w_i) dw_i} \quad (13)$$

와 같게 된다.

### 3.2 DWT 계수들의 분포 특성

우도비  $l(\mathbf{y})$ 을 구하기 위하여서는 워터마크  $w_i$ 가 삽입된 DWT 계수  $y_i$ 의 PDF  $f_{y_i}(y_i|w_i)$ 가 필요하다. 일반적으로 많이 사용되는 워터마크 삽입 방법은 곱의 형태 방법으로, 이는

$$y_i = x_i(1 + aw_i) \tag{14}$$

와 같다. 이 방법에 의하여 워터마크가 삽입된 DWT 계수  $y_i$ 의 PDF  $f_{y_i}(y_i|w_i)$ 는

$$f_{y_i}(y_i|w_i) = \frac{1}{1+aw_i} f_{x_i}\left(\frac{y_i}{1+aw_i}\right) \tag{15}$$

와 같다. 따라서  $f_{y_i}(y_i|w_i)$ 를 구하기 위하여서는 원 DWT 계수들의 PDF  $f_{x_i}(x_i)$ 가 필요하다. 다양한 영상에 대한 DWT 계수들의 실제 분포는 가우스 분포와 유사하다. 따라서 제안한 알고리즘에서는 DWT 계수들의 분포를 가우스 분포로 설정한다.

가우스 분포를 따르는 DWT 계수  $x_i$ 의 PDF는

$$f_{x_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(x_i - m_i)^2}{2\sigma_i^2}\right] \tag{16}$$

와 같이 표현되고, 여기서  $m_i$  및  $\sigma_i^2$ 은 각각  $x_i$ 가 속한 부대역의 DWT 계수들의 평균 및 분산을 나타낸다. 그러므로 곱의 형태 삽입 방법에 의하여 워터마크가 삽입된 DWT 계수  $y_i$ 의 PDF는

$$f_{y_i}(y_i|w_i) = \frac{1}{1+aw_i} \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{\left(\frac{y_i}{1+aw_i} - m_i\right)^2}{2\sigma_i^2}\right] \tag{17}$$

와 같게 된다.

### 3.3 우도비에 DWT 계수의 분포 특성 적용

워터마크가 삽입된 DWT 계수  $y_i$ 의 PDF  $f_{y_i}(y_i|w_i)$ 를 우도비  $l(\mathbf{y})$ 에 적용하면, 우도비의 분

모 (denominator)  $\frac{1}{2^N} \prod_{i=1}^N \int_{-1}^1 f_{y_i}(y_i|w_i) dw_i$ 는

$$\frac{1}{2^N} \prod_{i=1}^N \int_{-1}^1 \frac{1}{1+aw_i} \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{\left(\frac{y_i}{1+aw_i} - m_i\right)^2}{2\sigma_i^2}\right] dw_i, \tag{18}$$

와 같게 되고, 여기서  $\frac{y_i}{1+aw_i}$ 를  $t_i$ 로 치환하면 식 (18)은

$$\prod_{i=1}^N \frac{1}{2a\sqrt{2\pi\sigma_i^2}} \int_{\frac{y_i}{1+a}}^{\frac{y_i}{1-a}} \frac{1}{t_i} \cdot \exp\left[-\frac{(t_i - m_i)^2}{2\sigma_i^2}\right] dt_i \tag{19}$$

와 같이 간단히 표현된다. 이때  $a$ 가 1에 비하여 작다고 가정하면 식 (19)의 적분 구간은  $y_i$ 를 중심으로 매우 작으므로, Taylor 정리에 의한 선형 근사화 (linear approximation)에 의하여 식 (19)에서의 적분될 함수는

$$\frac{1}{t_i} \cdot \exp\left[-\frac{(t_i - m_i)^2}{2\sigma_i^2}\right] \approx \frac{1}{y_i} \cdot \exp\left[-\frac{(y_i - m_i)^2}{2\sigma_i^2}\right] + \mu_0(t_i - y_i) \tag{20}$$

와 같이 표현되고, 이때  $\mu_0$ 는

$$\mu_0 = \left. \frac{d}{dt_i} \left( \frac{1}{t_i} \cdot \exp\left[-\frac{(t_i - m_i)^2}{2\sigma_i^2}\right] \right) \right|_{t_i=y_i} \tag{21}$$

을 나타낸다. 식 (19)에서의 적분 구간  $\left[\frac{y_i}{1+a}, \frac{y_i}{1-a}\right]$ 에 대한 선형 근사화에 의한 적분될 함수의 오차는 그림 3에서와 같다. 이 그림으로부터 삽입 강도  $a$ 가 증가함에 따라 선형 근사화에 의한 오차도 증가하는 것을 알 수 있고, 특히  $a$ 가 0.3 이상 일 때는 매우 큰 오차가 발생함을 알 수 있다. 따라서 제안한 알고리즘에서는 선형 근사화에 의한 적분될 함수의 오차, 비가시성, 및 견고성을 고려하여 삽입 강도  $a$ 를 0.3으로 결정한다.

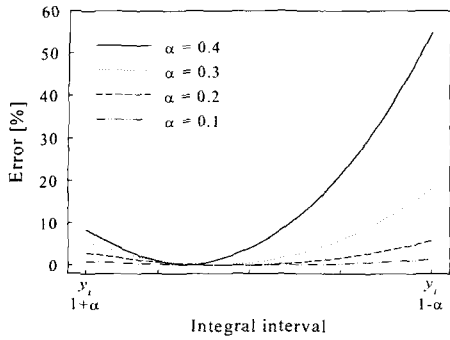
식 (20)을 식 (19)에 적용하면

$$\prod_{i=1}^N \frac{1}{2a\sqrt{2\pi\sigma_i^2}} \int_{\frac{y_i}{1+a}}^{\frac{y_i}{1-a}} \left( \frac{1}{y_i} \exp\left[-\frac{(y_i - m_i)^2}{2\sigma_i^2}\right] + \mu_0(t_i - y_i) \right) dt_i \tag{22}$$

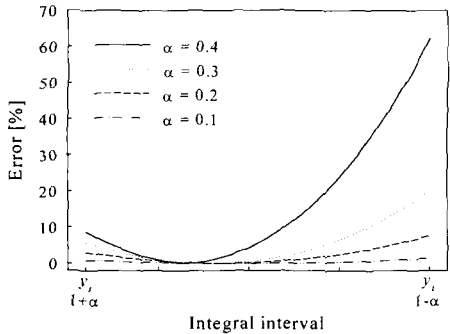
을 얻을 수 있고, 두 번째 항의 적분값은 거의 0에 가까우므로 식 (18)은

$$\prod_{i=1}^N \frac{1}{2ay_i\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{(y_i - m_i)^2}{2\sigma_i^2}\right] \cdot \left( \frac{y_i}{1-a} - \frac{y_i}{1+a} \right) \tag{23}$$

와 같게 된다. 그리고 이를 정리하면



(a)



(b)

그림 3. (a)  $y_i = 20$  및 (b)  $y_i = 25$  일 때 선형 근사화에 의한 적분될 함수들의 오차들

$$\prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{(y_i - m_i)^2}{2\sigma_i^2}\right] \cdot \left(\frac{1}{1 - \alpha^2}\right) \quad (24)$$

와 같고, 이때  $\alpha$ 가 1에 비하여 작으므로 우도비의 분모는

$$\frac{1}{2^N} \prod_{i=1}^N \int_{-1}^1 f_y(y_i | w_i) dw_i \approx \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{(y_i - m_i)^2}{2\sigma_i^2}\right] \quad (25)$$

와 같이 유도된다.

또한 우도비의 분자 (numerator)  $\prod_{i=1}^N f_y(y_i | w_i^*)$ 는

$$\prod_{i=1}^N f_y(y_i | w_i^*) = \prod_{i=1}^N \frac{1}{1 + \alpha w_i^*} \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{\left(\frac{y_i}{1 + \alpha w_i^*} - m_i\right)^2}{2\sigma_i^2}\right] \quad (26)$$

와 같다. 이 결과들을 우도비에 적용하면 우도비  $l(\mathbf{y})$ 는

$$l(\mathbf{y}) = \prod_{i=1}^N \frac{1}{1 + \alpha w_i^*} \cdot \exp\left[\frac{(y_i - m_i)^2}{2\sigma_i^2} - \frac{\left(\frac{y_i}{1 + \alpha w_i^*} - m_i\right)^2}{2\sigma_i^2}\right] \quad (27)$$

와 같게 된다. 식 (27)을 간단히 표현하기 위하여 이 식의 양변에 로그를 취하면 로그 우도비  $\ln l(\mathbf{y})$ 가

$$\ln l(\mathbf{y}) = \sum_{i=1}^N \left[ \frac{(y_i - m_i)^2}{2\sigma_i^2} - \frac{\left(\frac{y_i}{1 + \alpha w_i^*} - m_i\right)^2}{2\sigma_i^2} \right] - \sum_{i=1}^N \ln(1 + \alpha w_i^*) \quad (28)$$

와 같이 생성된다. 따라서 이 로그 우도비와 비교되는 판정 문턱값은  $\ln \lambda$ 가 된다. 이때  $z(\mathbf{y})$  및  $\lambda'$ 을

$$z(\mathbf{y}) = \sum_{i=1}^N \frac{1}{2\sigma_i^2} \left[ (y_i - m_i)^2 - \frac{\left(\frac{y_i}{1 + \alpha w_i^*} - m_i\right)^2}{2\sigma_i^2} \right] \quad (29)$$

$$\lambda' = \ln \lambda + \sum_{i=1}^N \ln(1 + \alpha w_i^*) \quad (30)$$

와 같이 설정한다면, 최종 판정 규칙은

$$\delta(\mathbf{y}) = \begin{cases} 1, & z(\mathbf{y}) > \lambda' \\ 0, & \text{otherwise} \end{cases} \quad (31)$$

이다.

### 3.4 판정 문턱값 설정

판정 문턱값의 설정은 위터마크 검출 알고리즘의 성능에 매우 중요한 역할을 한다. 일반적으로 Bayes 판정 이론에서는 Bayes 위험, 즉 판정 오류 확률  $P_e$ 를 최소화하는 우도비 값을 판정 문턱값으로 선택한다. 이때  $P_e$ 는

$$P_e = p_0 P_{FA} + p_1 P_{MD} \quad (32)$$

와 같고,  $P_{FA}$  및  $P_{MD}$ 는 각각 오류 검출 확률 및 간과 검출 확률로서

$$P_{FA} = P(\delta(\mathbf{y}) = 1 | W_0) = P(z(\mathbf{y}) > \lambda_1 | W_0) \quad (33)$$

$$P_{MD} = P(\delta(\mathbf{y}) = 0 | W_1) = P(z(\mathbf{y}) < \lambda_1 | W_1) \quad (34)$$

와 같이 표현될 수 있다. 그러나 위터마크를 제거하기 위한 공격을 당한 신호에 대하여  $P_e$ 를 최소화하는  $z(\mathbf{y})$ 값이 판정 문턱값으로 선택된다면,  $P_{MD}$ 가

$P_{FA}$ 보다 매우 커지게 되므로 실제 위터마크가 삽입되었지만 그렇지 않다고 판정할 확률이 그 반대의 경우보다 훨씬 커지게 된다. 이는 기존의 유사도 기반 검출 알고리즘에서 제기되었던 문제점이다.

따라서 이러한 문제를 해결하기 위하여 제안한 검출 알고리즘에서는  $P_e$ 를 최소화하는 대신, 위터마크 검출 알고리즘에서 허용할 수 있는 최대 허용 허용 검출 확률  $\overline{P_{FA}}$ 에 대하여  $P_{ML}$ 를 최소화할 수 있는 판정 문턱값의 설정을 위하여 Neyman-Pearson 정리를 이용한다.

$\overline{P_{FA}}$ 와 판정 문턱값  $\lambda'$ 의 관계는

$$\overline{P_{FA}} = P(z(\mathbf{y}) > \lambda' | W_0) = \int_{\lambda'}^{+\infty} f_{z(\mathbf{y})}(z(\mathbf{y}) | W_0) dz(\mathbf{y}) \quad (35)$$

와 같고,  $f_{z(\mathbf{y})}(z(\mathbf{y}) | W_0)$ 는  $W_0$  공간에 대한  $z(\mathbf{y})$ 의 조건부 PDF를 나타낸다. 이때  $f_{z(\mathbf{y})}(z(\mathbf{y}) | W_0)$ 는  $\mathbf{y} = \mathbf{x}$ 일 때  $z(\mathbf{y})$ 의 PDF와 동일하므로,  $\mathbf{y} = \mathbf{x}$ 일 때의  $z(\mathbf{y})$ 를  $z(\mathbf{x})$ 라고 한다면,  $z(\mathbf{x})$ 는

$$z(\mathbf{x}) = \sum_{i=1}^N \frac{1}{2\sigma_i^2} \left[ (x_i - m_i)^2 - \left( \frac{x_i}{1 + aw_i^*} - m_i \right)^2 \right] \quad (36)$$

와 같고, 따라서  $z(\mathbf{x})$ 의 PDF가 필요하다.

$z(\mathbf{x})$ 는 랜덤 변수들의 합으로 구성되어 있으므로, 이는 중심 극한 정리에 의하여 가우스 분포를 가진다. 이때  $v_i$ 를

$$v_i = \frac{1}{2\sigma_i^2} \left[ (x_i - m_i)^2 - \left( \frac{x_i}{1 + aw_i^*} - m_i \right)^2 \right] \quad (37)$$

와 같이 설정하면,  $z(\mathbf{x})$ 의 평균 및 분산은 각각  $v_i$ 의 평균 및 분산의 합과 같다. 이때  $v_i$ 는 가우스 분포를 가지는  $x_i$ 의 2차 함수이므로,  $v_i$ 의 평균 및 분산은 각각

$$m_{v_i} = a_i(m_i^2 + \sigma_i^2) - b_i m_i \quad (38)$$

$$\sigma_{v_i}^2 = [\sigma_i(2a_i m_i - b_i)]^2 \quad (39)$$

와 같다. 여기서  $a_i$  및  $b_i$ 는 상수로서 각각

$$a_i = \frac{aw_i^*(aw_i^* + 2)}{2\sigma_i^2(aw_i^* + 1)^2} \quad (40)$$

$$b_i = \frac{am_i w_i^*}{\sigma_i^2(aw_i^* + 1)} \quad (41)$$

이다. 그러므로  $z(\mathbf{x})$ 의 평균 및 분산은 각각

$$m_{z(\mathbf{x})} = \sum_{i=1}^N m_{v_i} = \sum_{i=1}^N [a_i(m_i^2 + \sigma_i^2) - b_i m_i] \quad (42)$$

$$\sigma_{z(\mathbf{x})}^2 = \sum_{i=1}^N \sigma_{v_i}^2 = \sum_{i=1}^N [\sigma_i(2a_i m_i - b_i)]^2 \quad (43)$$

와 같다. 따라서  $\overline{P_{FA}}$ 와  $\lambda'$ 의 관계는

$$\begin{aligned} \overline{P_{FA}} &= \frac{1}{\sqrt{2\pi\sigma_{z(\mathbf{x})}^2}} \int_{\lambda'}^{+\infty} \exp \left[ -\frac{(z(\mathbf{x}) - m_{z(\mathbf{x})})^2}{2\sigma_{z(\mathbf{x})}^2} \right] dz(\mathbf{x}) \\ &= \frac{1}{2} \operatorname{erfc} \left( \frac{\lambda' - m_{z(\mathbf{x})}}{\sqrt{2\sigma_{z(\mathbf{x})}^2}} \right) \end{aligned} \quad (44)$$

와 같고, 만약  $\overline{P_{FA}} = 10^{-9}$ 로 설정한다면 판정 문턱값  $\lambda'$ 은

$$\lambda' = 4.24\sqrt{2\sigma_{z(\mathbf{x})}^2} + m_{z(\mathbf{x})} \quad (45)$$

와 같이 설정된다. 최종적으로 식 (29)에서의  $z(\mathbf{y})$ 와 식 (45)에서의 판정 문턱값  $\lambda'$ 을 비교함으로써 대상 신호  $\mathbf{y}$ 에 특정 위터마크  $w^*$ 의 존재를 판정한다.

지금까지 설명한 바와 같이, 제안한 위터마크 검출 알고리즘에서는 위터마크 존재에 대한 변별력이 뛰어난 우도비를 Neyman-Pearson 정리에 따른 판정 문턱값과 비교함으로써 위터마크를 검출한다. 그리고 대상 신호 및 위터마크의 분포 특성을 정확하게 모델링하여 이를 우도비에 적용하였으므로 판정 오류를 더욱 더 최소화할 수 있다. 따라서 제안한 검출 알고리즘은 부적절한 판정 규칙 및 판정 문턱값의 설정으로 인하여 큰 간과 오류 확률이 발생하는 기존의 유사도 기반 알고리즘에서의 단점을 개선할 수 있다.

#### 4. 실험 결과 및 고찰

본 논문에서 제안한 위터마크 검출 알고리즘의 성능을 평가하기 위하여 컴퓨터 모의 실험을 행하였다. 실험에서 사용된 영상으로는 512×512 화소 크기의 CABLECAR 및 LENA 등의 정지 영상과 352×240

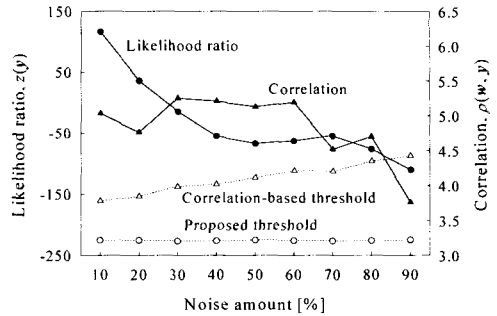
화소 크기의 SIF (source input format) 급 동영상 FOOTBALL 및 MOBILE 등을 각각 30 프레임 씩 사용하였다. 이들의 웨이블릿 변환에는 쌍직교 필터 (biorthogonal filter)를 사용하였다.

제안한 알고리즘과 기존의 유사도 기반 알고리즘의 성능 평가 척도로는 위터마크를 제거하려는 다양한 공격에 대한 각 알고리즘의 검출 응답을 사용하였다. 이 때 두 알고리즘의 객관적인 비교를 위하여 동일한 방법으로 위터마크가 삽입된 영상에 대하여 공격을 가한 후, 이에 대하여 위터마크 검출 실험을 행하였다. 또한 기존의 유사도 기반 알고리즘에서의 판정 문턱값 선택에도 제안한 알고리즘에서 사용한 Neyman-Pearson 정리를 적용함으로써 더욱 객관적인 실험을 행하였다.

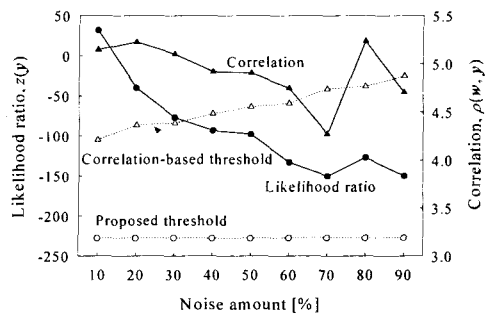
비가시성과 견고성을 동시에 고려하기 위하여, 4-레벨로 분해된 다해상도 영상에서 기저 대역과 최저 레벨의 부대역을 제외한 중간 주파수 성분을 나타내는 부대역에서 크기가 큰 5000개의 계수들에 대하여 위터마크를 삽입하였다. 이때 삽입 강도  $\alpha$ 로는 선형 근사화에 의한 적분될 함수의 오차, 비가시성, 및 견고성을 고려하여 0.3을 사용하였고, 최대 허용 허위 검출 확률  $\overline{P_{FA}}$ 로는 위터마크가 삽입되지 않았는데 삽입되었다고 판정할 확률을 매우 작게 설정하기 위하여  $10^{-9}$ 을 사용하였다. 위터마크가 삽입된 정지 영상 및 동영상의 PSNR은 모두 40 dB 이상으로 나타남으로써 화질의 열화는 없는 것을 알 수 있다.

4.1 정지 영상에 대한 견고성 평가

정지 영상에 대한 공격으로는 가우스 랜덤 잡음, JPEG (Joint Photographic Experts Group) 부호화, 및 절단 (cropping) 을 위터마크가 삽입된 영상에 가하였고, 이들의 결과는 그림 4, 그림 5, 및 그림 6에 각각 나타내었다. 이때 가우스 랜덤 잡음의 첨가에서는 전체 영상 크기에 대하여 256 밝기값을 가지는 랜덤 잡음이 첨가되는 부분을 변화시켰고, JPEG 부호화에서는 다양한 성능으로 대상 영상을 부호화하였다. 그리고 절단은 영상의 가운데를 기준으로 수평 및 수직 방향으로 동일 크기의 영역을 위터마크가 삽입된 영상으로 구성하고, 나머지 영역을 원 영상으로 구성하는 것을 의미한다.

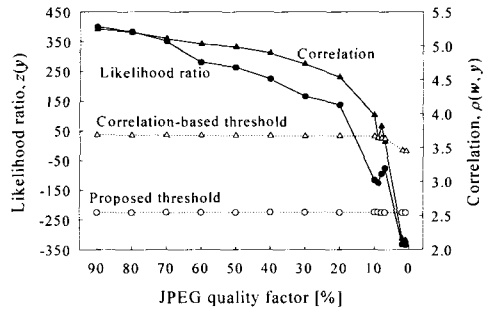


(a)

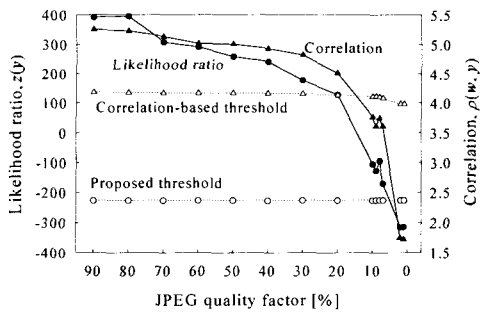


(b)

그림 4. 잡음 첨가에 대한 (a) CABLECAR 및 (b) LENA 영상들의 검출 응답들



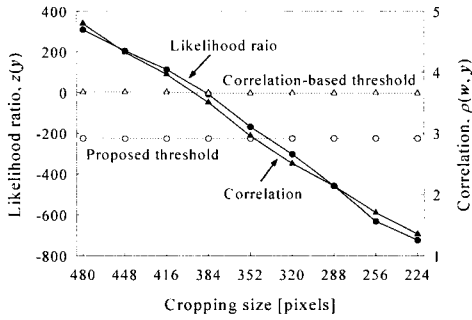
(a)



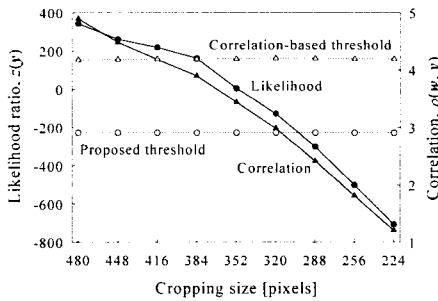
(b)

그림 5. JPEG 부호화에 대한 (a) CABLECAR 및 (b) LENA 영상들의 검출 응답들





(a)



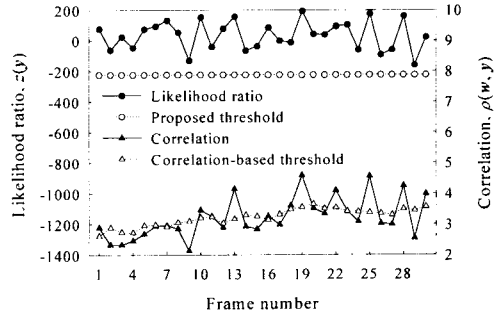
(b)

그림 6. 절단에 대한 (a) CABLECAR 및 (b) LENA 영상들의 검출 응답들

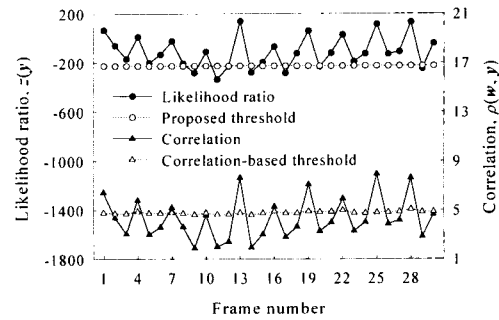
4.2 동영상에 대한 견고성 평가

동영상에 대한 실험으로는 워터마크가 삽입된 영상을 MPEG (Moving Picture Experts Group) TM5 (Test Model 5)에 의하여 다양한 비트율로 부호화한 후, 부호화된 영상에 대하여 워터마크를 검출하였다. 이때 GOP (group of pictures) 길이는 12, I/P 프레임 거리는 3, 및 비월 주사 방식으로 1.5 및 0.5 Mbps의 비트율로 부호화하였다. B 프레임에서는 양방향으로 움직임이 추정되므로 워터마크의 손실 가능성이 크게 된다. 각 비트율에 따른 견고성 결과를 그림 7 및 그림 8에 나타내었다.

이상의 결과들로부터 정지 영상 및 동영상 모두의 경우에 대하여 제안한 검출 알고리즘이 유사도 기반 알고리즘보다 우수함을 알 수 있었다. 이러한 결과들은 제안한 워터마크 검출 알고리즘에서는 가우스 분포로 모델링된 DWT 계수의 PDF를 Bayes 판정 이론에 따른 우도비에 적용한 후, 이 우도비 값과 Neyman-Pearson 정의에 따라 결정된 판정 문턱값을 비교하여 워터마크의 존재 유무를 밝히기 때문이다. 따라서 정확한 판정 규칙 및 판정 문턱값에 기반

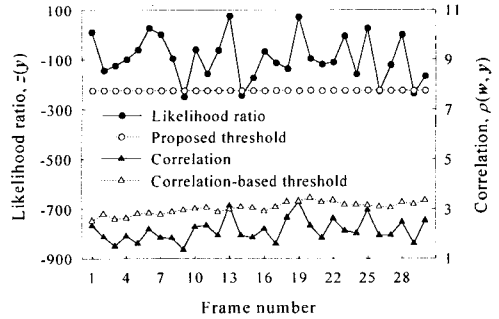


(a)

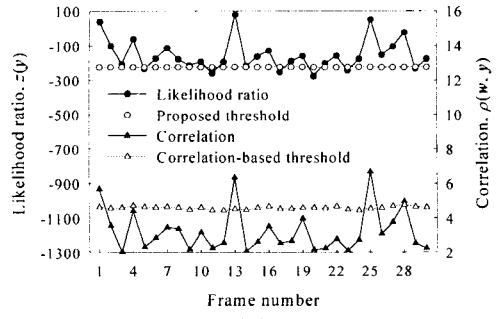


(b)

그림 7. 1.5 Mbps로 부호화된 (a) FOOTBALL 및 (b) MOBILE 영상들에 대한 검출 응답들



(a)



(b)

그림 8. 0.5 Mbps로 부호화된 (a) FOOTBALL 및 (b) MOBILE 영상들에 대한 검출 응답들

한 제안한 검출 알고리즘은 주어진 최대 허용 오류 검출 확률에 대하여 간과 검출 확률을 최소화함으로써 기존의 유사도 기반 워터마크 검출 알고리즘에 비하여 우수한 성능을 나타낸다.

### 5. 결 론

본 논문에서는 최근 워터마킹 분야에서 많이 연구되고 있는 DWT 영역에서 통계적 판정법을 이용한 고신뢰 워터마크 검출 알고리즘을 제안하였다. 제안한 워터마크 검출 알고리즘에서는 신뢰성 있는 검출 척도로 Bayes 판정 이론에 의한 우도비를 사용하고, 판정 오류를 최소화할 수 있는 판정 문턱값의 설정을 위하여 Neyman-Pearson 정리를 이용한다. 이때 우도비에는 DWT 계수의 분포가 필요한데, 제안한 방법에서는 이를 가우스 분포로 모델링한다.

제안한 고신뢰 워터마크 검출 알고리즘의 성능을 평가하기 위한 실험 결과로부터 통계적 판정법을 이용한 제안한 검출 알고리즘이 기존의 유사도 기반 알고리즘에 비하여 우수한 검출 성능을 나타냄을 확인할 수 있었다. 즉 제안한 워터마크 검출 알고리즘은 워터마크를 제거하려는 다양한 공격을 당한 검사 대상 신호에서, 실제 이에 삽입된 워터마크를 기존의 유사도 기반 알고리즘보다 정확하게 검출하였다.

### 참 고 문 헌

[1] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," *Proc. of the SPIE Conf. on Storage and Retrieval for Image and Video Databases III*, vol. 2420, pp. 164-173, San Jose, California, Feb. 1995.

[2] I. Pitas, "A method for watermark casting on digital images," *IEEE Trans. on Circuits System Video Technology*, vol. 8, pp. 775-780, Oct. 1998.

[3] I. Cox, J. Kilian, T. Leighton, and T. Shamoan, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. on Image Processing*, vol. 6, no. 12, pp. 1673-1687, 1997.

[4] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "DCT-based watermark recov-

ering without resorting to the uncorrupted original image," *Proc. of the IEEE Int. Conf. Image Processing*, Santa Barbara, California, vol. 1, pp. 520-527, Oct. 1997.

[5] X. Xia, C. Boncelet, and G. Arce, "A multiresolution watermark for digital images," *Proc. of the Int. Conf. Image Processing*, Santa Barbara, California, vol. 1, pp. 548-551, Oct. 1997.

[6] W. Zhu, Z. Xiong, and Y. -Q. Zhang, "Multiresolution watermarking for images and video," *IEEE Trans. on Circuits and Systems for video technology*, vol. 9, no. 4, pp. 545-550, June 1999.

[7] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, pp. 357-372, May 1998.

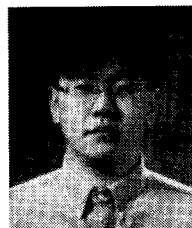
[8] A. Piva M. Barni F. Bartolini V. Cappellini, "Threshold selection for correlation-based watermark detection," *Proc. of the COST254 Workshop*, L'Aquila, Italy, Apr. 1998.

[9] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[10] J. Franco and W. Rubin, *Radar Detection*. Norwood, MA: Artech House, 1980.

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Process*. New York: McGraw-Hill, 1991.

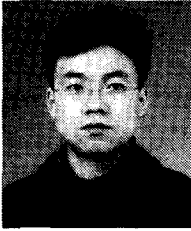
[12] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*. New York: Dover, 1975.



권 성 근

1996년 2월 경북대학교 전자공학과 졸업  
 1998년 2월 경북대학교 전자공학과 석사  
 2002년 8월 경북대학교 전자공학과 박사  
 2002년 11월 ~ 현재 삼성전자 무선 통신 사업부

관심 분야 : 영상처리, 영상통신, 정보보호



**김 병 주**

1997년 2월 경북대학교 전자공학과 졸업  
 1999년 2월 경북대학교 전자공학과 석사  
 2000년 3월~현재 경북대학교 전자공학과 박사과정  
 관심 분야 : 영상처리, 영상통신, 정보보호



**이 석 환**

1999년 2월 경북대학교 전자공학과 졸업  
 2001년 2월 경북대학교 전자공학과 석사  
 2001년 3월~현재 경북대학교 전자공학과 박사과정  
 관심 분야 : 영상처리, 영상통신, 정보보호



**권 기 구**

1997년 2월 인제대학교 전자공학과 졸업  
 2000년 2월 경북대학교 전자공학과 석사  
 2000년 3월~현재 경북대학교 전자공학과 박사과정  
 관심 분야 : 영상처리, 영상통신



**김 영 춘**

1991년 2월 경북대학교 전자공학과 졸업  
 1993년 2월 경북대학교 전자공학과 석사  
 1997년 2월 경북대학교 전자공학과 박사  
 1998년 3월~현재 영동대학교 정보전자공학부 조교수

관심 분야 : 영상처리, 영상통신



**권 기 룡**

1986년 2월 경북대학교 전자공학과 졸업 (공학사)  
 1990년 2월 경북대학교 대학원 전자공학과 졸업 (공학석사)  
 1994년 8월 경북대학교 대학원 전자공학과 졸업 (공학박사)

2000년 7월~2001년 8월 Univ. of Minnesota, Post-Doc. 과정  
 1996년 3월~현재 부산외국어대학교 컴퓨터전자공학부 부교수  
 2002년~현재 한국멀티미디어학회 총무이사  
 관심분야 : 멀티미디어 정보보호, 멀티미디어 통신, 웨이블릿 변환



**이 건 일**

1965년 2월 경북대학교 물리학과 졸업  
 1970년 2월 경북대학교 물리학과 이학석사  
 1984년 2월 부산대학교 물리학과 이학박사  
 1969~1984년 경북대학교 공과대학 전자공학과 조교

-부교수  
 1982~1983 미국 RPI 공과대학 교환교수  
 1984~현재 경북대학교 공과대학 전자공학과 교수  
 1995~현재 산업표준심의회 전자부회 전자관 (IEC/TC39) 전문위원회 위원  
 관심 분야 : 비디오 및 오디오 공학, 전자계측, 의용전자 공학

**교신저자**

권 성 근 (702-701) 대구광역시 북구 산격동 1370번지  
 경북대학교 대학원 전자공학과