논문 02-27-11C-10 한국통신학회논문지 '02-11 Vol.27 No.11C

지역 평균값 제거를 통한 개선된 워터마킹 방법

정회원 강 현 수*, 홍 진 우**, 김 광 용**

Enhanced watermarking scheme based on removal of local means

Hyun-Soo Kang*, Jin-Woo Hong**, and Kwang-Young Kim** *Regular Members*

요 약

본 논문은 원신호의 지역 평균값을 등가적으로 제거함으로써 검출 오류확률을 감소시키는 워터마킹 기법을 제안한다. 우선 원신호의 지역 평균값의 제거가 검출 오류확률을 감소시킴을 보인다. 여기서 원신호의 지역 평균값의 제거는 원신호를 변형하게 되므로 원신호를 변형하지 않고 등가적으로 원신호의 지역 평균값을 제거하는 방식을 필요로 하게 된다. 그 방식은 워터마크 신호의 지역 평균값을 제거한 워터마크를 삽입함으로써 워터마크 검출 시 원신호의 지역 평균값을 제거한 것과 같은 효과를 가짐에 기초한다. 이러한 결과들이 이론적으로 증명되며, 다양한 워터마크 신호를 적용한 실험을 통해 검증한다..

ABSTRACT

This paper presents a new watermarking scheme to reduce the detection error probability through removal of local mean values of an original signal equivalently. At first, we show that the error probability is reduced by the removal of the local mean values. In the removal process, we are in need of a method that equivalently removes the local mean values without modification of the origianl signal since the process changes the original signal. The method is based on the principle that as the watermark with zero local mean values is embedded, the local mean values of the original signal is equivalently removed in detection of the watermark. The principle are analytically proven, and the superiority of the proposed method is verified by experiments for variety of watermarks

Ⅰ. 서 론

널리 알려진 바와 같이, 원영상의 사용 여부에 따라 워터마킹 방식은 blind 방식과 non-blind 방식으로 나뉘어진다. 일반적으로 Cox의 방식[1]과 같은 non-blind 방식은 원영상과 테스트 영상의 차신호와 워터마크 사이의 상관값에 기초하는 반면, Dugad 방식[2]과 같은 blind 방식은 테스트 영상과 워터마크 사이의 상관값에 기초한다. 주목할 점은 [3]에서 보고된 것과 같이 non-blind 방식은 올바른 저작권 판별에 문제가 있다는 것이다. 한편, 비록 blind 방식이 올바른 저작권 판별에 문제가 없지만, 검출 신뢰도에 있어서 상대적으로 약점을 가진다. 이 약점은 워터마크 신호와 원신호 사이의 간섭(interference) 때문이다[4].

이러한 간섭을 감소시키는 일반적인 방법은 테스트 신호가 상관기로 입력되기 전에 필터를 통과시키는 방법이다[5]. 그러나, 주파수 성분 간에는 상관성이 거의 없으므로 필터링은 주파수 영역에 워터마크를 삽입하는 방법에 적용하기에는 부적합하다. 많은 워터마킹 기법들이 주파수 영역에서 고안되는 점을 감안한다면 이에 대한 개선 방안이 필요하다. 본 논문은 그 방안 중의 하나이다.

* 중앙대학교 첨단영상대학원 영상공학과 (hskang@cau.ac.kr),
** 한국전자통신연구원 방송미디어부 방송컨텐츠연구팀 ({jwhong, kwangyk}@etri.re.kr)
※ 본 논문은 2002년 4월 JCCI 학술대회에서 우수논문으로 선정되어 게재 추천된 논문입니다.

Ⅱ. 기존 방법의 오류 확률

워터마크의 검출은 다음 식과 같이 표현되는 상관기 출력 z를 검사함으로써 이루어진다.

$$z = \sum_{n=0}^{N-1} r(n)w(n) \tag{1}$$

여기서 $r(n)$, $w(n)$는 테스트 신호와 워터마크이다. 공격이 없는 경우, 워터마크가 존재하는 테스트 신호는 $r(n) = x(n) + w(n)$ 표현된다. 여기서 원 신호 $x(n)$은 [5]에서와는 달리 결정적신호(deterministic signal)로 간주될 것이다. 중앙극한이론(central limit theorem)에 의해, z는 Gaussian 분포를 가지며, 평균과 분산은 다음과 같이 주어진다.

$$\mu_z = E[\sum_{n=0}^{N-1} (x(n) + w(n))w(n)] = E_w, \tag{2}$$

$$\sigma_z^2 = E[(\sum_{n=0}^{N-1} x(n)w(n))^2] = \sigma_w^2 E_x \tag{3}$$

여기서 E_w, E_x 는 $w(n)$, $x(n)$의 에너지이며, 워터마크는 σ_w^2의 파워를 가지는 백색잡음(white noise)로 가정되었다.

만약 테스트 신호에 워터마크가 존재하지 않는다면, $r(n) = x(n)$로서, 위와 유사한 과정에 의해 $\mu_z = 0$, $\sigma_z^2 = \sigma_w^2 E_x$의 Gaussian 분포를 가진다. 결과적으로 positive/negative false 확률(P_+, P_-)는 다음과 같다.

$$\begin{aligned} P_+ = P_- &= Q(T/\sigma_z)|_{T=E_w/2} \\ &= Q(1/2 \cdot \sqrt{E_w/(m_x^2 + \gamma_x^2)}) \end{aligned} \tag{4}$$

여기서 T, m_x, γ_x^2는 문턱값, $x(n)$의 시간평균값, 시간분산값이다. 즉, $m_x = N^{-1}\sum_{m=0}^{N-1} x(n)$, $\gamma_x^2 = N^{-1}\sum_{n=0}^{N-1} (x(n) - m_x)^2$이다. 비록 $x(n)$이 결정적 신호로 간주되었지만, 그 결과는 [5]에서와 흡사하다. 식(4)의 m_x는 $w(n)$의 시간 평균을 0으로 강제함으로써, 즉 $\sum w(n) = 0$, 쉽게 제거될 수 있다. 즉,

$$\begin{aligned} \sum_{n=0}^{N-1} x(n)w(n) &= \sum_{n=0}^{N-1} [m_x + x_{ac}(n)]w(n) \\ &= \sum_{n=0}^{N-1} x_{ac}(n)w(n) \end{aligned} \tag{5}$$

여기서 $x_{ac}(n)$는 $x(n)$의 A.C. 성분, m_x는 D.C. 성분에 해당된다. 여기서 $x_{ac}(n)$의 에너지는 $N\gamma_x^2$ 임에 주목하자. 식(5)를 식(3)에 대입함으로써, 오류확률을 크게 개선하는 효과를 주는 감소된 분산값을 다음과 같이 얻을 수 있다.

$$\sigma_z^2 = N\sigma_w^2\gamma_x^2 \tag{6}$$

식(6)은 워터마크의 시간 평균값이 0일 때, 상관기 출력값의 분산이 작아짐을 보이고 있다. 그래서, 워터마크의 삽입 시, D.C. 제거과정은 적극 추천되며, 본 논문에서, 기존의 방법이라고 함은 D.C. 제거과정이 적용된 방법을 일컫는다.

Ⅲ. 제안된 방법

식(6)에서 보여준 D.C. 제거의 원리가 본 논문에서 일반화된다. 제안된 방법은 워터마크를 M개의 부워터마크로 나뉘고, 부워터마크의 시간평균을 제거함으로 σ_z^2를 감소시키는 방법이다. 제안된 방법은 다음과 같은 과정을 따른다.

(1) $w(n)$을 얻는다.

(2) $w(n)$를 M개의 부워터마크로 분류한다. 즉 $w_i(n)$, $i = 0, 1, \cdots, M-1$, 여기서 $w_i(n)$의 길이는 N_i, $N = \sum_{i=0}^{M-1} N_i$이다.

(3) 각 부워터마크에 대해 시간평균을 제거하여 $w_i'(n)$를 얻는다. 즉, $w_i'(n) = w_i(n) - m_{w_i}$, 여기서 m_{w_i}는 $w_i(n)$의 시간평균값이다.

(4) 마지막으로, $w_i'(n)$를 조합하여, 새로운 워터마크 $w'(n) = \cup_{i=0}^{M-1} w_i'(n)$를 얻고, $w'(n)$에 목표 분산값 σ_w^2 이 나오도록 스케일링을 가한다.

제안된 방법의 분석을 위해, 공격이 없음을 가정한다. 즉, $r(n) = x(n) + w(n)$. 이때, 평균값은 $\mu_z = E_w$로 쉽게 없어진다. 이제 분산값에 대해 조

사한다. 우선, 다음 관계식에 주목하자.

$$\sum_{n=0}^{N-1} x(n)w'(n) = \sum_{i=0}^{M-1}\sum_{n=0}^{N_i-1} x_i(n)w_i'(n) \quad (7)$$

여기서 $x_i(n) = x(n+T_i)$, $T_i = \sum_{j=0}^{i-1} N_j$이다. 식(7)에서, $x_i(n)$는 $w_i'(n)$가 곱해지는 $x(n)$의 일부분으로 정의된다. 즉 $x(n)$, $T_i \leqq n < T_{i+1}$ 이다. $x_{i,ac}(n)$를 $x_i(n)$의 A.C. 성분으로 정의할 때,

$$\sum_{i=0}^{M-1}\sum_{n=0}^{N_i-1} x_i(n)w_i'(n) = \sum_{i=0}^{M-1}\sum_{n=0}^{N_i-1} x_{i,ac}(n)w_i'(n)$$

$$\sigma_z^2 = E\left(\left(\sum_{i=0}^{M-1}\sum_{n=0}^{N_i-1} x_{i,ac}(n)w_i'(n)\right)^2\right)$$

$$= \sum_{i=0}^{M-1} E\left(\left(\sum_{n=0}^{N_i-1} x_{i,ac}(n)w_i'(n)\right)^2\right)$$

$$= \sum_{i=0}^{M-1}\sum_{n=0}^{N_i-1} x_{i,ac}(n)\sigma_w^2 = \sigma_w^2\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2 \quad (9)$$

이다. 식(6)과 식(9)에서, $\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2 \leqq N\gamma_x^2$이면, 제안된 방법의 분산은 기존의 방법의 분산보다 작다. $\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2 \leqq N\gamma_x^2$를 증명하기 위해, 다음 식의 관계를 이용한다.

$$E_x = \sum_{n=0}^{N-1} x^2(n) = N(\gamma_x^2 + m_x^2) \quad (10)$$

$$E_x = \sum_{n=0}^{N-1} x^2(n) = \sum_{i=0}^{M-1} N_i(\gamma_{x_i}^2 + m_{x_i}^2) \quad (11)$$

식(10)과 식(11)에서, $\sum_{i=0}^{M-1} N_i m_{x_i}^2 \geqq N\gamma_x^2$이면, $\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2 \leqq N\gamma_x^2$이다. 이는 다음 관계식에 의해 쉽게 증명된다.

$$\sum_{i=0}^{M-1} N_i m_{x_i}^2 - 2m_x\sum_{i=0}^{M-1} N_i m_{x_i} + m_x^2\sum_{i=0}^{M-1} N_i$$

$$= \sum_{i=0}^{M-1} N_i(m_{x_i} - m_x)^2 \geqq 0 \quad (12)$$

식(12)는 $Nm_x = \sum_{i=0}^{M-1} N_i m_{x_i}$와 $N = \sum_{i=0}^{M-1} N_i$에 의해 다음과 같이 다시 쓸 수 있다.

$$\sum_{i=0}^{M-1} N_i m_{x_i}^2 \geqq Nm_x^2 \quad (13)$$

식(13)은 D.C. 성분 에너지(Nm_x^2)이 지역 D.C. 성분 에너지의 합($\sum_{i=0}^{M-1} N_i m_{x_i}^2$)이 항상 작거나 같음을 의미한다. 이는 $N\gamma_x^2$ 가 항상 $\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2$ 보다 크거나 같음을 의미하기도 한다. 그림1은 이러한 사실을 좀 더 명확하게 보여준다. 그림1에서 빗금쳐진 부분은 분할된 부분의 D.C. 성분을 나타낸다. 그림1은 $Nm_x^2 = 0$ 일지라도, $\sum_{i=0}^{M-1} N_i m_{x_i}^2 \neq 0$ 로서, 분할된 부분의 D.C. 성분의 에너지가 항상 신호 전체의 D.C. 성분의 에너지보다 크거나 같음을 명확하게 보여준다. 결과적으로 다음과 같은 관계가 성립한다.

$$Q\left(\frac{1}{2}\sqrt{E_w \Big/ \sum_{i=0}^{M-1}\frac{N_i}{N}\gamma_{x_i}^2}\right) \leqq Q\left(\frac{1}{2}\sqrt{\frac{E_w}{\gamma_x^2}}\right) \quad (14)$$

그림2는 $x(n)$의 하나의 극단적인 예이다. 여기서 $N_i = L_i$, M=3의 경우 $\gamma_{x_i}^2 = 0$이므로, 제안된 방법의 오류확률은 0인 반면, 기존의 방법은 $\gamma_x^2 \neq 0$ 이므로 오류확률이 0이 아니다. 이 예에서 보듯이, 제안된 방법은 주어진 M에 대해 지역분산값 $\gamma_{x_i}^2$이 적은 경우 매우 효과적이다. 다른 예로서, $\gamma_{x_i}^2$이 모든 i에 대해 동일한 경우, $\gamma_{x_i}^2 = \gamma_x^2$가 되어, 제안된 방법은 기존의 방법과 동일한 성능을 가진다. 위의 두 가지 예와 식(14)를 통해, 우리는 $\gamma_{x_i}^2$가 가능한 작을수록 (이것은 i에 따라 $m_{x_i}^2$의 변화가 심할수록 $\gamma_{x_i}^2$가 작아진다) $\sum_{i=0}^{M-1} N_i\gamma_{x_i}^2$가 더욱 감소됨을 알 수 있다.

Blind 방식에서, 어떤 M에 대해, 최적의 N_i는 삽입

부에서 얻어질 수 있지만, 검출부에서는 $x(n)$을 알 수 없으므로 최적의 N_i를 알 수 없다. 따라서, 가능한 $\gamma_{x_i}^2$이 작도록 미리 결정할 필요가 있다.

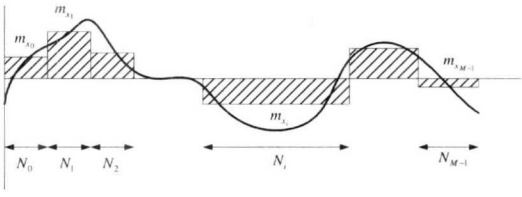


그림 1. 워터마크 신호의 분할 영역별 DC성분

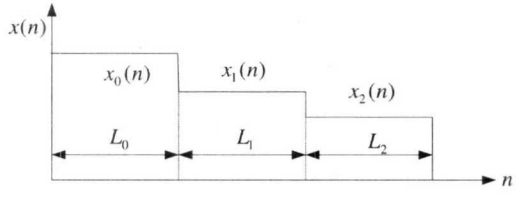


그림 2. 워터마크 신호($x(n)$)의 극단적인 예

Ⅳ. 제안된 방법

우선, blind방법이고 원영상 $I(x,y)$의 DFT된 영상 $I(u,v)$의 중간대역의 크기 성분에 워터마크를 삽입하는 주파수 영역 워터마킹 기법을 사용한다고 가정하자. 그림3에서 보여진 바와 같이, 다이아몬드 형태의 대역에 워터마크를 삽입한다. 즉, $a \leq |u| + |v| < b$, 여기서 u와 v는 x와 y방향의 주파수를 나타낸다. 그 때, $x(n)$은 정해진 스캐닝 규칙에 따라 $I(u,v)$의 삽입 대역의 값을 스캐닝함으로써 구성될 수 있다.

1. 알고리즘1

$M = b - a$ 이고, M개의 부워터마크가 다이아몬드 형태($|u| + |v| = c$, $a \leq c < b$)의 크기가 증가하는 순서대로 삽입된다. 즉, $w_0'(n)$, $w_1'(n)$, $\cdots$, $w_{M-1}'(n)$ for $c = a\,a+1, \cdots, b-1$, 여기서 $w_i'(n)$는 D.C. 성분이 없는 신호이다. 그림4는 본 알고리즘을 그림으로 보여준다. 주파수 영역에 있어서 이러한 분할은 앞 절의 조건을 대체로 잘 만족한다. 이는 고정된 (u,v) 에 대해 c를 변화시킴으로써 겪는 주파수 성분의 변화량이 고정된 c에 대해 (u,v) 를 변화시킴으로써 겪는 주파수 성분의 변화보다 크기 때문이다. 즉, 후자의 경우, $\gamma_{x_i}^2$가 더 작고, i에 따른 $m_{x_i}^2$의 변화가 심함을 의미한다.

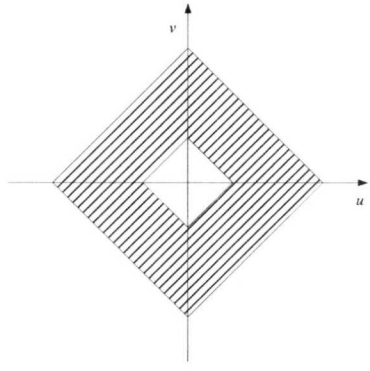


그림 3. 워터마크 신호의 삽입 대역

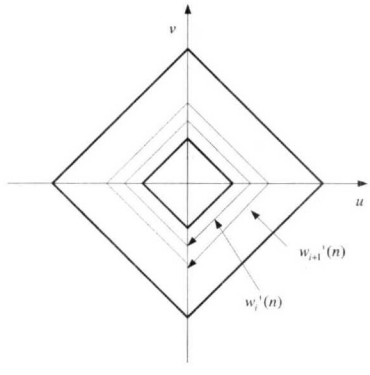


그림 4. 알고리즘1의 설명

2. 알고리즘2

앞 절의 분석은 M이 증가함에 따라 오류확률이 감소하는 경향이 있음을 보인다. 그래서, 제안된 방법의 성능을 극대화하기 위하여, M이 최대값을 갖는 경우를 생각해 보자. 즉 $M=N/2$, 여기서 N은 짝수로 가정한다. $M=N/2$인 이유는 $N_i \geq 2$이기 때문이다. 이는 $N_i = 1$의 경우 D.C. 제거 과정은 부워터마크 자체가 D.C.에 해당하므로 부워터마크가 완전히 사라지도록 하기 때문이다. 모든 i에 대해 $N_i = 2$인 경우, averaging attack은 워터마크를 제거하는데 매우 효과적이다. 그림5는 이 경우를 보여주는 것으로서, 홀수 좌표값에 위치한 값들이 짝수 좌표값에 위치한 값의 부호반전의 형태로 나타난다. 이는 $N_i = 2$의 부워터마크의 평균이 0이어야 하기 때문이다. 그래서, averaging attack에 강인하기 위해서는, 이러한 공격 시 영상의 화질이 크게 저하되도록 워터마크를 샘플단위로 재배치될 필요가 있다. 그림6에서 보듯이, 이러한 공격을 피하는 하

나의 방법은 1사분면과 3사분면에 $w'(2n)$을 삽입하고, 2사분면과 4사분면에 $w'(2n+1)$를 삽입하는 것이다. 이것은 $w'(2n)$와 $w'(2n+1)$가 삽입된 신호에 대한 averaging attack이 화질을 크게 저하시키므로 워터마크를 보호할 수 있게 한다.

$w'(2n+1)$이 단순히 $w'(2n)$의 반대 극성으로 나타난다. 즉 $w'(2n+1) = -w'(2n)$. 따라서, 다음과 같은 과정에 의해 쉽게 구현될 수 있다.

(1) $w(n)$를 얻는다.

(2) 미리 정해진 스캐닝 순서에 따라 $|I(u,v)|$의 1사분면과 3사분면에 $w(n)$을 삽입한다.

(3) 과정(2)와 마찬가지로 2사분면과 4사분면에 $-w(n)$를 삽입한다.

위의 과정은 비록 D.C. 제거 과정이 없지만, 반대극성의 워터마크신호가 반복되어 있어, D.C. 제거 효과를 포함하고 있음에 주목하자.

V. 실험 결과

성능 평가를 위해, 다음과 같이 정의된 정규화된 상관값을 측정한다.

$$\rho = \sum_{n=0}^{N-1} r(n)w(n)/E_w \qquad (15)$$

ρ는 1에 가까울 수록 워터마크가 정확히 검출되는 것이라는 점에 주목하자. 즉, ρ의 평균값 μ_ρ이 1이고, 분산값 σ_ρ^2 이 0일 경우가 이상적인 경우다. 그래서 μ_ρ가 1에 얼마나 가까운지 그리고 σ_ρ^2가 얼마나 적은 값을 가지는지를 관찰함으로써 μ_ρ와 σ_ρ^2는 성능평가의 좋은 기준이 될 수 있다.

또한, ρ가 문턱값 이상의 값을 가질 때 제대로 검출되었다고 가정하고, 제대로 검출되는 비율 RDR (right detection ratio)을 측정함으로써 성능 평가를 수행할 수도 있다. 여기서 문턱값은 0.5로 가정하였다. 결론적으로, 성능평가를 위해 RDR, μ_ρ, σ_ρ^2의 세가지 평가기준을 사용한다.

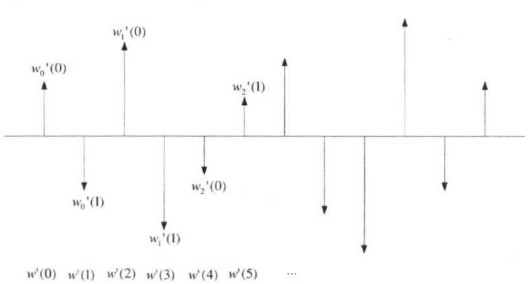


그림 5. 반대극성을 동반한 워터마크 신호의 예

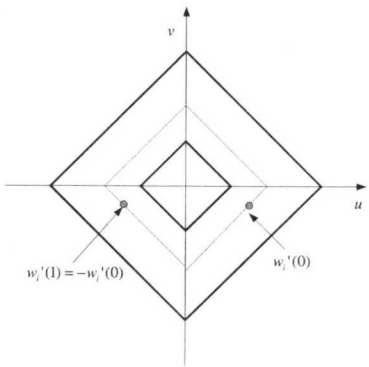


그림 6. 알고리즘2의 설명

표 1. 실험결과 (N=3696, a=16, b=64, JPEG QF = 2).

방법	공격	PSNR	RDR	μ_ρ	σ_ρ^2
기존 방법	NO	45.00	99.2%	0.962507	0.033367
	JPEG	28.75	88.6%	0.722674	0.032563
알고리즘1	NO	45.04	99.9%	0.976155	0.021355
	JPEG	28.76	94.4%	0.735458	0.020715
알고리즘2	NO	45.08	100%	0.985237	0.015544
	JPEG	28.76	97.5%	0.745729	0.015597

256x256 Lena영상에 대해 1000개의 다른 seed로 발생된 워터마크를 이용하여 1000번의 실험을 수행하여 표1을 얻었다. 표1에서 볼 수 있듯이 제안된 방법이 (1) 더 높은 RDR, (2) 좀 더 1에 가까운 μ_ρ, (3) 좀 더 작은 σ_ρ^2의 결과를 보인다. 기대할 수 있는 바와 같이 알고리즘2가 가장 좋은 성능을 가짐을 관찰할 수 있다.

Ⅳ. 결 론

본 논문은 지역평균값 제거를 통한 워터마킹 기법의 성능을 향상시키는 방법을 제안하였고, 이론적으로 그리고 실험적으로 검증하였다. 제안된 방법은 워터마크 신호의 파워를 소모하면서도 어떠한 정보도 전달할 수 없는 워터마크의 DC 성분을 제거함으로써 워터마크의 효율을 증대시키는 방법이라고 말할 수 있다. 그 방법으로써 두가지 알고리즘을 제시하였으며, 두 번째 알고리즘인 반대극성의 워터마크를 이용한 알고리즘이 더 우수한 성능을 보임을 알 수 있었다.

본 논문에서는 분할 파라미터인 M과 N_i이 휴리스틱하게 정해졌지만 이에 대한 분석이 더 필요할 것으로 보인다.

감사의 글

본 논문은 정보통신부 지원 "디지털 컨텐츠 관리 기술 개발" 과제의 결과물로서 관계자 분들에게 감사의 글을 드립니다.

참 고 문 헌

[1] I. J. Cox, F. T. Leighton, and T.Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. on Image Processing, vol. 6, pp. 1673-1687, Dec. 1997.

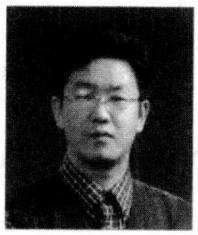
[2] R. Dugad, K. Ratakonda, and N. Ahuja, "A New Wavelet-Based Scheme for Watermarking Images," proc. of ICIP'98, pp. 419-423, 1998.

[3] K. Ratakonda, R. Dugad, N. Ahuja, "Digital Image Watermarking: Issues in Resolving Rightful Ownership," proc. of ICIP'98, pp. 414-418, 1998.

[4] H. S. Kang, J. G. Choi, S. W. Lee, and S. J. Cho, "Reliable watermark detection method based on analysis of correlation," Optical Eng., vol. 39, no. 12, pp. 3308-3316, Dec. 2000.

[5] G. Depovere, T. Kalker and J-P Linnartz, "Improved watermark detection reliability using filtering before correlation," proc. of ICIP'98, pp. 430-434, 1998.

강 현 수(Kang Hyun-Soo) 정회원



1999년 2월 : 한국과학기술원 전기및전자공학과 (공학박사)
1995년 5월~2001년 4월 : 하이닉스반도체 (주) 선임연구원
2001년 5월~2002년 2월 : 한국전자통신연구원 선임연구원
2002년 3월~현재 : 중앙대학교 첨단영상대학원 조교수
<주관심 분야> 영상처리, 부호화, 컨텐츠보호기술 등.

홍 진 우(Jin Woo Hong) 정회원

1993년 8월 : 광운대학교 대학원 전자계산기공학과 (공학박사)
1999년 2월 : 독일 프라운호퍼 연구소(파견연구원)
1984년 3월~현재 : 한국전자통신연구원 방송컨텐츠연구팀장 (책임연구원)
2000년 1월~현재 : 한국음향학회 홍보이사, 뉴미디어 음향 학술분과위원장, 방송공학회 편집위원
1993년 1월~현재 : 정보통신표준화연구단 방송기술위원회 위원
2001년 6월 ~ 현재 : SEDICA 운영위원
<주관심 분야> 오디오 신호처리 및 부호화, 디지털 컨텐츠 보호 및 관리, 디지털 오디오 방송

김 광 용(kwang-yong Kim) 정회원

1993년 2월 : 충남대학교 컴퓨터공학과(공학석사)
1998년 2월 : 충남대학교 컴퓨터공학과(공학박사)
1998년 5월~2000년 3월 : 한국전자통신연구원 Post-Doc
2000년 4월~현재 : 한국전자통신연구원 방송미디어 연구부 선임연구원
<주관심 분야> 컴퓨터공학, 정보처리, 전자공학