논문-01-6-2-07

상관성 분석에 기반한 신뢰성있는 워터마크 검출 방법

강 현 수*, 최 재 각**, 이 시 웅***, 안 치 득*, 홍 진 우*

An Improved Watermark Detection Method Through Correlation Analysis

Hyun-Soo Kang*, Jae Gark Choi**, Si-Woong Lee***, Chieteuk Ahn* and Jin-Woo Hong*

요 약

워터마크란 저작권 보호를 위해 영상과 같은 멀티미디어에 심어두는 보이지 않는 신호를 말한다. 많은 경우에 워터마크 검출 기법은 상관성을 구한 후, 그 값을 정해진 임계값과 비교하는 과정으로 이루어지는데 상관성을 구할 때 원영상의 사용 유무에 따라 2가지 부류로 나누어진다. Type1검출기는 원영상(original image)과 테스트할 입력영상과의 차신호를 상관기의 입력으로 사용하는 경우이고, Type2검출기는 입력영상을 직접 상관기의 입력으로 사용하는 경우이다. Type1검출기는 올바른 저작권을 증명하는데 어려움을 가진다. 반면, Type2 검출기는Type1검출기와 비교할 때, 올바른 저작권 증명에 있어서는 유리하지만 Type1검출기가 가지지 않은 문제점을 안고 있다. 그 문제점은 워터마크와 워터마킹할 원영상 사이의 상관성이 워터마크의 에너지로 정규화할 때 그 값이 무시할 수 없을 정도로 크다는 점이다. 따라서 본 논문에서는 그 상관성이 워터마크 검출의 성능에 미치는 영향을 분석하면서 그 영향을 최소화하는 새로운 기법을 제안하고 제안된 기법의 성능을 분석한다.

Abstract

A digital watermark is a perceptually unobtrusive signal embedded in some multimedia asset such as an image for copyright protection. In many cases watermark detection amounts to thresholding a correlation value between a watermark and a received image. Watermarking detection schemes can be classified into two types. Type 1 is based on a correlation process that is applied to the difference between an original image and an input image to be tested. Type 2 is based on a correlation process that is directly applied to an input image. The type 1 fails to prove the rightful ownership, while type 2 has an advantage in terms of rightful ownership compared with type 1. However, type 2 has a problem that doesnt appear in type 1. The problem is that correlation between a watermark and an original image to be watermarked is too significant to be ignored, when it is normalized by watermarks energy. In this paper, based on the analysis of the correlation, we propose a novel watermarking scheme to minimize the effect and also verify the performance of the proposed scheme by experiments.

I. 서 론

컴퓨터의 급속한 발달과 인터넷 같은 컴퓨터망의 확산으로 많은 음성, 영상, 비디오 데이터들이 디지털화되고 있다. 이런 디지털 데이터들은 기존 아날로그 데이터와 비교하여 데이터의 저장과 편집이 매우 용이한 장점을 갖는다. 하지만 데이터를 디지털화 함으로써 생기는 장점들은 또한 누구나 디지털 데이터의 내용을 쉽게 변형하고 복제할 수 있는 단점이 되기도 한다. 특히 디지털화된 데이터는 원본과 복사본의 구분이 불가능하여 소유권의 보호문제가 심각하게 대두되고 있다[1][2][3].

워터마크란 영상의 시각적인 특성을 이용하여 영상에

* 한국전자통신연구원 무선방송연구소
Radio & Broadcasting Technology Laboratory, Electronics and Telecommunications Research Institute

** 동의대학교 컴퓨터응용공학부
Division of Computer Application Engineering, Dong-Eui University

***한발대학교 정보통신컴퓨터공학부
School of Information Comm. and Computer Eng., Hanbat National University

직접 임의의 변화를 주고, 저작권의 확인이 필요할 때는 워터마크를 검출함으로써 영상의 저작권을 주장할 수 있는 방법이다. 따라서 디지털 워터마킹(digital watermarking)은 저작권 보호 문제를 해결하기 위한 하나의 기술적인 해결책이 되고 있다. 워터마크는 압축이나 보통의 신호처리를 포함한 의도적이거나 비의도적인 공격(attack)에 의해 쉽게 지워지지 않도록 충분히 커야 한다. 한편, 워터마크는 시각적으로 눈에 잘 띄지 않도록 충분히 작은 신호여야 한다[4]. 많은 경우에 워터마크 검출은 그림1에서 보여진 것처럼 상관성(correlation)을 구한 후, 그 값을 주어진 임계값(threshold)과 비교하는 과정으로 이루어진다. 즉, 입력영상과 워터마크의 상관성을 구하기 위해 상관기(correlator)를 통과한다. 이때, 그 값이 주어진 임계값 이상이면 워터마크가 존재한다고 판단하고 그렇지 않으면 워터마크가 존재하지 않는다고 판단한다.

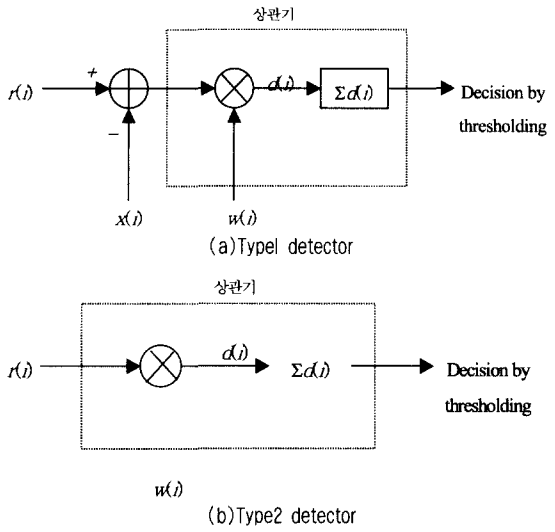


그림 1. 워터마크 검출기의 종류
Fig. 1. The types of watermark detector

워터마킹 기법들은 그림1에서 보여지는 것처럼 워터마크 검출기(detector)의 형태에 따라 두가지 부류로 나누어진다[7]. Type1 검출기는 correlation process가 Cox의 방식[5]처럼 원영상(original image)과 워터마킹된 영상과의 차신호에 적용되는 경우이다. Cox의 방식이 Type 1의 한 예이다. Type2 검출기는 correlation process가 워터마킹된 입력영상에 직접적으로 적용되는 경우이다. Dugad의 방식[6]이 Type2의 한 예이다. Type1은 참고문헌[7]에서 지적한 것처럼 올바른 저작권을 증명하는데 어려움을 가지므로 제안된 워터마크 검출방식은 Type2 방식을 적용한다. 비록 Type2가 Type1과 비교하여 올바른 저작권의 증명에 있어서 유리하지만 Type2는 Type1이 가지지 않은 문제점을 안고 있다. 그 문제점은 워터마크와 워터마킹될 원영상 사이의 상관성이 워터마크 에너지로 정규화(normalization)할 때 그 값이 무시할 수 없을 정도로 크다는 점이다. 따라서 본 논문에서는 그 상관성이 워터마크 검출기의 성능에 미치는 영향을 분석하면서 그 영향을 최소화하는 새로운 방법을 제안하고 또한 제안된 방법의 성능을 분석한다.

본 논문의 구성은 다음과 같다. 먼저 제2절에서 워터마크 신호와 워터마킹할 원영상 사이의 상관성이 워터마크 검출에 미치는 영향을 분석한다. 그리고 워터마크 검출의 신뢰성 향상을 위한 제안방법을 제3절에서 기술하였다. 제4절에서 제안방법에 따른 분류기 설계가 제안된 방법에 미치는 효과를 분석하고, 또한 잡음을 포함한 입력신호의 경우 제안된 워터마크 검출 방법에 미치는 효과를 분석한다. 제5절에서 워터마크 검출의 신뢰성 향상을 위해 제안된 방법의 모의실험 결과를 제시하였다. 마지막으로 제6절에서 결론을 맺는다.

2. 워터마크와 원영상사이의 상관성이 워터마크 검출에 미치는 효과 분석

워터마킹될 $N_1 X N_2$(총 $N = N_1 N_2$개의 화소) 크기의 원영상을 $\zeta(n, m)$라고하고 $\zeta(n, m)$의 변환(transformation) 영상을 이라고 하자. 여기서 변환은 워터마크가 삽입될 영역에 따라, DCT(Discrete Cosine Transform), DFT(Discrete Fourier Transform), DWT(Discrete Wavelet Transform), 기타 가능한 변환들 중의 어느 것일 수 있다. $\zeta(n, m)$또는 (u, v)가 순차주사(raster scanning) 또는 지그재그 주사(zigzag scanning)를 사용하여 1차원적인 신호의 형태로 나열한 신호를 $x(i)$라고 하자. 여기서 $i = (n, m)$, $(0 \le n < N_1 - 1, 0 \le m < N_2 - 1)$이다. 그리고 $w(i)$와 $r(i)$는 각각 워터마크 신호 및 워터마크 검출기에 들어오는 입력신호라고 하자. 그때, 워터마킹된 영상은 $x(i) + w(i)$가 된다. 만약 워터마크가 어떠한 공격도 받지 않았을 경우, $r(i) = x(i) + w(i)$이다. 분석의 목적상, $x(i)$와 $w(i)$는 서로 독립이고(independent), 정상적(stationary), 에르고딕(ergodic)이라 가정하고, 워터마크 $w(i)$는 평균이 0이고 분산(variance)이 σ_w^2인 백색잡음(white noise)이라고 가정한다.

False negative(워터마크가 존재하지만 검출기가 워터마크를 검출 못하는 경우)의 확률을 조사하기 위하여, $r(i) = x(i) + w(i)$경우에 대한 상관응답(correlation response)의 분포를 얻는다. 여기서 상관응답 z는 다음과 같이 정의된다.

$$z \equiv \left(\sum_{i=0}^{N-1} r(i)w(i)\right) / E_w = \left(\sum_{i=0}^{N-1} x(i)w(i)\right) / E_w + 1 \qquad (1)$$

여기서 E_w는 $w(i)$의 에너지, 즉, $E_w = \sum_{i=0}^{n-1} w^2(i)$를 나타내고 $\sum_{i=0}^{n-1} x(i)w(i)$는 $x(i)$와 $w(i)$사이의 시간축 상호상관성(cross-correlation)에 해당된다. 시간축 상호상관성이 0인 이상적인 경우 (즉, $\sum_{i=0}^{n-1} x(i)w(i) = 0$), 응답 $z=1$이 기대되는 결과이다. 그러나 실제로 그 값이 1이 아니라 단지 그것의 기대치(expected value)가 1일뿐이다. 따라서 $w(i)$의 에너지에 의해 정규화된 상호상관성 (즉, $\sum_{i=0}^{n-1} x(i)w(i)/E_w$)의 분포에 의해 검출오차의 확률이 결정된다. 그리고 상호상관성은 1과 비교하여 작지 않은 값일 수 있다. 그것은 $x(i)$의 에너지가 $w(i)$의 에너지에 비해 훨씬 크기 때문이다. 이러한 에너지의 차이는 워터마크가 시각적으로 눈에 띄지 않게 삽입되어야 한다는 사실에 기인한다. 결론적으로 $\sum_{i=0}^{n-1} x(i)w(i)/E_w$가 무시할 수 없이 큰 값을 가질 수 있기 때문에 워터마크가 존재하는지 그렇지 않은지를 판단하는데 어려운 점이 있다. 즉, 아무런 공격이 없었음에도 불구하고, 그것이 양수이면 z는 1이상의 값을 가지고 그렇지 않으면 1 이하의 값을 가진다. 이러한 사실은 잘못된 판정을 내릴 확률을 증가시킨다. 따라서 $|z-1|$을 성능평가지수(performance measure)로 사용할 수 있는데 $|z-1|$의 큰 값은 검출오차의 확률이 높음을 나타낸다.

앞에서의 가정으로부터 $w(i)$의 자기상관성(auto-correlation)은 다음과 같다.

$$R_w(m,n) = E[w(m)w(n)] = \sigma_w^2 \delta(m-n) \qquad (2)$$

여기서 $E[\cdot]$는 평균을 나타낸다. 그리고 $E[z]$는 다음과 같이 주어진다.

$$\begin{aligned} E[z] &= E\left|\sum_{i=0}^{N-1} x(i)w(i)\right| / E_w + 1 \\ &= \left(\sum_{i=0}^{N-1} E[x(i)]E[w(i)]\right) / E_w + 1 = 1 \end{aligned} \qquad (3)$$

z의 분산 σ_z^2은 다음과 같다.

$$\begin{aligned} E[(z-E[z])^2] &= E\left[\left(\sum_{i=0}^{N-1} x(i)w(i)\right)^2\right] / E_w^2 \\ &= E\left[\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} x(i)x(j)w(i)w(j)\right] / E_w^2 \\ &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} R_x(i,j)R_w(i,j) / E_w^2 \end{aligned} \qquad (4)$$

결과적으로 다음과 같이 주어진다.

$$\sigma_z^2 = N(\mu_x^2 + \sigma_x^2)\sigma_w^2 / E_w^2 \qquad (5)$$

N이 충분히 크고 $w(i)$가 에르고딕이면, 다음 식이 성립한다.

$$E[w^2(i)] = \frac{1}{N}\sum_{i=0}^{N-1} w^2(i) \qquad (6)$$

최종적으로 z의 분산은 다음과 같다.

$$\sigma_z^2 = \frac{\mu_x^2 + \sigma_x^2}{N\sigma_w^2} = \frac{\mu_x^2 + \sigma_x^2}{E_w} \qquad (7)$$

한편, central limit theorem에 의해 N이 충분히 크면 $\sum_{i=0}^{N-1} x(i)w(i)$는 가우시안 분포를 갖는다. 그러므로 z는 평균이 1이고 분산이 $(\mu_x^2+\sigma_x^2)/E_w$인 가우스 분포를 갖는다고 말할 수 있다. 여기서 우리는 $\mu_x=0$일 때, z의 분산이 $x(i)$의 분산에 선형적으로 비례한다는 것을 알 수 있다. 그러므로 잘못 판단할 확률은 $x(i)$의 분산에 비례한다.

이번에는 false positive(워터마크가 존재하지 않지만 존재하는 것으로 판단되는 경우)의 확률을 계산하기 위하여 입력영상이 워터마크를 포함하지 않은 경우를 생각해 보자. 즉, $r(i)=x(i)$경우이다. $r(i)=x(i)+w(i)$의 경우와 비슷한 과정에 의해 z는 다음과 같은 평균과 분산을 갖는 가우스 분포를 갖는다는 것을 알 수 있다.

$$E[z]=0, \quad \sigma_z^2 = \frac{\mu_x^2 + \sigma_x^2}{E_w} \qquad (8)$$

그림 2는 $r(i)=x(i)+w(i)$와 $r(i)=x(i)$에 대한 상관응답 z의 분포를 보여준다. 결과적으로 그림 2에서 볼 수 있듯이, 0.5가 최적 문턱값이고 false positive(P_+)와 false negative(P_-)의 확률은 다음과 같이 주어진다.

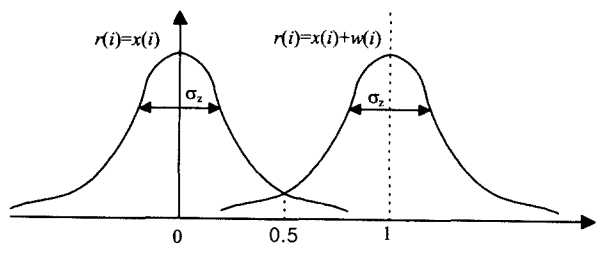


그림 2. 워터마크와 입력신호사이의 상관성 분포: 입력신호가 워터마크를 포함하는 경우와 포함하지 않는 경우

Fig. 2. The distribution of the correlation response z for $r(i)=x(i)+w(i)$ and $r(i)=x(i)$

$$P_+ = P_- = \frac{1}{2} erfc\left(\sqrt{\frac{E_w}{8(\mu_x^2+\sigma_x^2)}}\right) \quad (9)$$

이 식에서 $r(i)$가 어떠한 공격도 받지 않았음에도 불구하고 잘못 판단할 확률이 존재함을 알 수 있다. 이것은 첫째, 비록 시간축 상호상관성($\sum_{i=0}^{n-1} x(i)w(i)$)의 기대치는 0이지만 시간축 상호상관성은 0이 아니고 둘째, 의 에너지가 의 에너지에 비해 훨씬 크기 때문에 그 값이 무시할 수 없는 값이 되기 때문이다.

3. 제안된 워터마크 검출의 신뢰성 향상 방안

식 (9)에서 나타난 것처럼 오차 확률은 E_w가 가능한 크고 이 가능한 작을 때 최소화된다. 그러나 E_w의 증대는 실질적인 한계가 있다. 즉, E_w는 워터마크가 invisibility를 만족하게 하기 위하여 충분히 크게 만들어질 수 없다. 그리고 N은 워터마킹할 원영상의 크기에 비례하는 거의 고정된 값이다. 따라서 본 논문에서는 오차 확률을 줄이기 위해 $(\mu_x^2+\sigma_x^2)$가 작은 값을 가지도록 하는 방법들을 제안한다. 첫 번째 제안은 μ_x^2 을 제거하는 것이고 두 번째 제안은 σ_x^2의 크기 작도록 만드는 것이다.

우선 첫 번째 제안을 먼저 $(\mu_x^2+\sigma_x^2)$ 다룬다. μ_x^2는 $w(i)$의 직류성분(DC 값)이 0이 될 때(즉, $\sum_{i=0}^{n-1} w(i)=0$), 쉽게 제거할 수 있다. 즉, $\sum_{i=0}^{n-1} w(i)=0$일 때, 다음 식이 성립한다.

$$\sum_{i=0}^{N-1} x(i)w(i) = \sum_{i=0}^{N-1}(m_x + x_{ac}(i))w(i) = \sum_{i=0}^{N-1} x_{ac}(i)w(i) \quad (10)$$

여기서 m_x와 $x_{ac(i)}$는 $x(i)$의 DC 성분과 AC 성분을 나타낸다. 따라서 에르고딕의 가정 아래 다음 식이 성립한다.

$$\sigma_z^2 = E\left[\left(\sum_{i=0}^{N-1} x(i)w(i)\right)^2\right] / E_w^2 = E\left[\left(\sum_{i=0}^{N-1} x_{ac}(i)w(i)\right)^2\right] / E_w^2 = \sigma_x^2 / E_w \quad (11)$$

결과적으로 일 때, $\sum_{i=0}^{n-1} w(i)=0$오차 확률은 다음 식처럼 감소한다.

$$P_+ = P_- = \frac{1}{2} erfc\left(\sqrt{\frac{E_w}{8\sigma_x^2}}\right) \quad (12)$$

본 논문에서 제시된 첫 번째 제안은 주어진 N에 대하여를 $\sum_{i=0}^{n-1} w(i)=0$만족하도록 $w(i)$를 얻는 것이다. 이것은 의사난수(pseudo-random number) 발생 알고리즘에 의해 얻어진 $w(i)$의 DC 성분을 제거함으로써 실현될 수 있다. 즉, $w'(i)=w(i)-\frac{1}{N}\sum_{i=0}^{N-1} w(i)$이다. 다음부터 언급되는 $w(i)$는 첫 번째 제안을 만족하는 $w'(i)$를 나타낸다.

다음은 두 번째 제안에 대해 기술한다. 식(12)는 오차 확률이 σ_x가 감소함에 따라 비례하여 줄어줌을 보여준다. 여기서 σ_x를 줄이는 것이 가능한 것인지 의문이 제기될 수도 있다. 즉, σ_x는 원영상으로부터 얻어지기 때문에 $x(i)$를 변경하지 않는 한 σ_x를 감소시키는 것이 불가능해 보일지 모른다. 그러나, 다행히도 $x(i)$를 $w(i)$의 크기에 따라 여러 개의 클래스로 나눔으로써 σ_x를 줄이는 것이 가능하다. 즉, 각각의 i에 대해 $x(i)$는 $x(i)$의 크기가 비슷한 것들끼리 같은 클래스를 가지도록 분류한다. 이것은 그 클래스에 속한 신호들의 분산이 작아짐을 의미한다. 이러한 분류 작업 후, 각 클래스에 속한 신호들의 DC 성분은 첫 번째 제안에 의해 쉽게 제거된다.

클래수의 총 수를 M, $x(i)$의 최대값과 최소값을 각각 $x_{\max}$과 $x_{\max}$이라고 하자. 그 때, 분류(classification) 과정은 다음 식에 의해 수행된다.

$$class_{x(i)} = \min\left\{\left\lfloor \frac{M(x(i)-x_{\min})}{x_{\max}-x_{\min}} \right\rfloor + 1, M\right\}, \quad 1 \le class_{x(i)} \le M \quad (13)$$

여기서 $[\cdot]$는 부동소수점의 값(floating point value)을 정수 값으로 버림(truncation)함을 의미한다. $x(i)$가 분류화 간격(step size) $\triangle = (x_{\max} - x_{\max})/M$으로 $[I, M]$범위의 정수로 양자화 된다고 생각할 수 있다. 그림 3에 어떤신호에 대한 분류 과정의 한 예를 나타내었다. 클래스 k

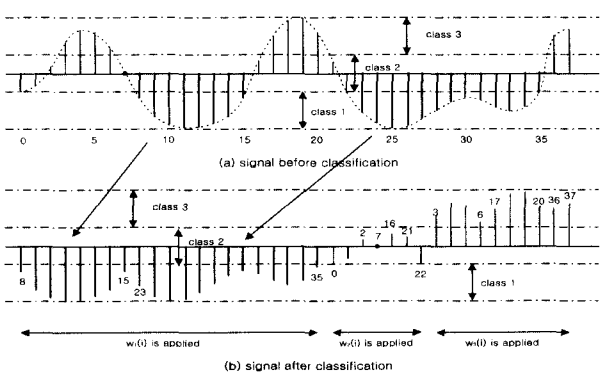


그림 3. 입력신호의 분산을 줄이기 위해 클래스로 분류화된 신호의 예
Fig. 3. An example of the classification for a signal to reduce its variance

에 속한 신호들의 집합을 다음과 같이 정의한다.

$$S_k = \{x(i): k = class_{x(i)}\},\ 1 \le k \le M \tag{14}$$

여기서 $x_k(i)$의 활성영역(dynamic range)은 $x(i)$의 활성영역보다 1/M배 작아지게 된다. 이때 $x_k(i)$는 S_k에 속한 모든 원소들을 일렬로 나열한 신호이다. 즉, $x_k(i)$의 분산 $\sigma_{x_k}^2$이 $x(i)$의 분산에 비해 작은 값을 가진다. 예를 들어 $x(i)$가 $[-L/2,\ L/2]$의 구간에서 균일 분포(uniform distribution)를 가질 때, $\sigma_x^2 = L^2/12$이고 $\sigma_{x_k}^2 = \triangle^2/12 = \sigma_x^2/M^2$이다. 이 예에서, $\sigma_{x_k}^2$가 σ_x^2와 비교하여 M^2배 감소함을 알 수 있다. 따라서 워터마크 삽입을 위해 워터마크 $w(i)$를 클래스 단위로 순차적으로 분할함으로써 부워터마크(sub-watermarks) $W_k(i)$를 얻는다. 즉,

$$w_k(i) = w(j),\ 0 \le i < N_k,\ \sum_{p=0}^{k-1} N_p \le j < \sum_{p=0}^{k} N_p \tag{15}$$

여기서 N_k는 S_k에 속한 원소들의 개수이다. 각 부워터마크에 대해 첫 번째 제안에서 설명한 것처럼 DC 성분 제거과정이 수행된다. 그 때, 오차의 확률은 현저하게 줄어들 수 있다. $x(i)$가 균일 분포를 가지는 경우에, 오차 확률은 다음과 같이 주어진다.

$$P_+ = P_- = \frac{1}{2} erfc\left(M\sqrt{\frac{E_w}{8\sigma_x^2}}\right) \tag{16}$$

식 (16)으로부터 M이 증가함에 따라 오차의 확률이 현저히 감소함을 알 수 있다. 일반적으로 $x(i)$가 불균일분포(non-uniform distribution)를 가지기 때문에 k번째 클래스 S_k의 원소 개수 $N(S_k)$가 모든 k에 대해 일정하지는 않다. 극단적인 경우 $N(S_k)=0$일 수 있다. $N(S_k)$가 충분히 크지 않는 경우, 워터마크의 백색성질(whiteness property)이 저하될 수 있다. 예를 들어 $N(S_k)=2$이면 첫번째 제안 $\sum_{i=0}^{1} w(i)=0$에 의해 $w(0) = -w(1)$이 된다. 이 경우 $w(0)$ 와 $w(1)$는 상관성이 매우 높다. 이처럼 $N(S_k)$는 너무 작은 값을 가지지 않도록 해야 한다. 그래서 위와 같은 극단적인 경우를 피하기 위해서 $N(S_k)$를 모든 k에 대해 일정하도록 하는 방법이 있을 수 있다. 즉, $N(S_k) = N/M = constant$, for $allk$. 이것은 $x(i)$의 히스토그램을 사용하여 식 (17)처럼 간단히 구현할 수 있다.

$$Class_{x(i)} = k,\quad for\ \frac{(k-1)N}{M} < \sum_{y=0}^{x(i)} h(y) \le \frac{kN}{M} \tag{17}$$

여기서 $h(x)$는 $x(i)$의 히스토그램이며, N은 $x(i)$의 길이, 그리고 M은 클래스의 수이다. 이러한 과정은 histogram equalization과 매우 유사하다.

워터마크 검출의 신뢰성 향상을 위한 첫 번째와 두 번째 제안으로부터 새로운 워터마크 삽입(insertion) 방식은 다음 순서로 구현된다.

(1) 길이 N인 의사난수(pseudo-random number) $w(i)$를 미리 주어진 분산에 맞추어 발생시킨다.

(2) 워터마킹 방식이 어느 영역에서 수행되는지에 따라 $\xi(n, m)$(원영상) 또는 (u, v)(변환된 영상)으로부터 $x(i)$를 얻는다. 예를 들어 Cox는 공간영역에서 워터마크를 적용하는 대신 DCT 영역에 적용하였다. 제안된 방식도 Cox의 방식처럼 DCT 영역에서 워터마크를 삽입하였으나 올바른 소유권을 제대로 증명하기 위하여Cox의 것과는 달리 Type2의 검출방식이 사용되었다.

(3) 식(13) 또는 (17)에 따라 $x(i)$를 상응하는 클래스로 분류한다. 여기서 클래스 개수, M은 사용자에 의해 결정된다. 이때, 클래스로 분류된 신호 $x_k(i)$를 얻는다. 최적의 M을 결정하는데 대한 연구는 진행 중에 있다.

(4) 워터마크 $w(i)$를 클래스의 개수에 따라 순차적으로 M개로 분리함으로써 $w_k(i)$를 얻는다.

(5) 첫 번째 제안에 따라 $w_k(i)$, $k=1,2,\cdots,M$의 DC 성분을 제거한다. 그 때 DC 성분이 제거된 $w_k(i)$를 $w_k'(i)$라고 하자.

(6) 워크마킹된 신호 $y(i)$를 $y_k(i)$의 합으로부터 얻는다. 즉, $y(i) = \bigcup_{k=1}^{M} y_k(i),\ 0 \le j < N, 0 \le i < N_k$. 여기서 $y_k(i) = x_k(i) + \gamma |x_k(i)| w_k'(i)$, $k=1,2,\cdots,M$이고 γ는 워터마크의 에너지를 조절하기 위한 계수이다. 워터마크 검출을 위해 최종적으로 얻어진 워터마크, 즉, $w_f(j) = \bigcup_{k=1}^{M} w_k'(i)$를 저장한다.

(7) 주파수 영역 워터마크 방식이 사용되었다면 역스캐닝(inverse scanning)과 역변환(inverse transformation)을 $y(i)$에 대해 수행함으로써 공간영역에서 워터마킹된 영상 Ω를 얻는다.

그림 4가 위의 과정을 도식적으로 보여준다.

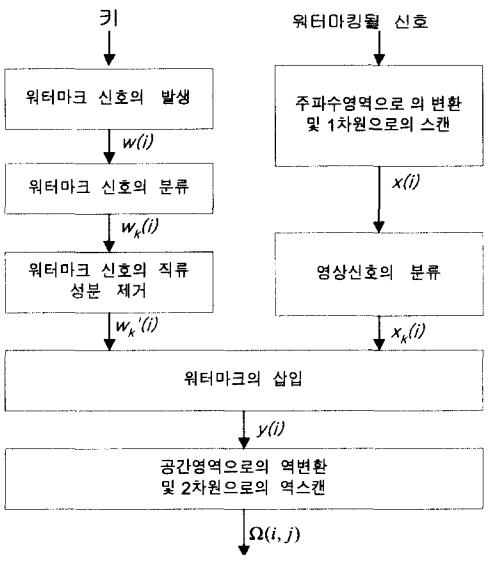


그림 4. 워터마크 검출의 신뢰도 향상을 위한 워터마크 삽입 과정
Fig. 4. The block diagram of the watermark insertion to improve the reliability of the watermark detection

워터마크 검출(detection)은 다음과 같이 이루어진다.

(1) 워터마킹된 영상 $\Omega(i,j)$로부터 상응하는 영역으로의 변환과 스캐닝을 수행함으로써 $r(i)$를 얻는다.
(2) 원영상으로부터 분류(classification)에 대한 정보를 얻는다.
(3) 분류정보와 워터마크 $w_f(i)$를 이용하여 상관응답 $z=(\sum_{k=1}^{M}(r_k(i)w_k'(i))/\sum_{i=0}^{n-1}(w'(i))$을 얻는다.
(4) 워터마크가 존재하는지를 문턱값 0.5로 결정한다.

4. 클래스 분류기 설계 및 잡음을 포함한 입력신호가 제안된 방법에 미치는 효과 분석

본 절에서는 먼저 신호의 분산을 줄이기 위해 클래스로 나누는 분류화 과정의 클래스 개수 M및 분류화 간격 $\triangle_k$가 제안된 방법에 미치는 효과를 분석한다. 그리고 워터마크 검출과정의 입력신호가 압축과 같은 다양한 신호처리에 의해 공격되었을 경우 제안된 방법의 효과를 분석한다.

제안된 방식을 사용함으로써 워터마크 검출의 오차확률이 M만큼 감소함을 수식적으로 나타내었다. 균일분포의 경우 직관적으로 오차확률을 최소화하는 $\triangle_k$($x_k(i)$의 dynamic range)는 최적 양자화기의 설계에서처럼 모든 k에 대해 상수 값을 가진다는 사실을 짐작할 수 있다. 그러나 일반적으로 $x(i)$는 균일 분포를 가지지 않는다. "일반적인 분포를 갖는 $x(i)$에 대해 오차확률을 최소화하는 최적의 $\triangle_k$는 무엇인가?" 하는 문제는 최적 양자화기를 설계하는 문제와 같은 것이다. $x_k(i)$의 평균과 $\triangle_k$는 양자화기를 설계하는데 있어서 reconstruction value와 step size라고 생각될 수 있다. 더욱이, $x_k(i)$의 분산은 양자화 오차에 해당된다. $x_k(i)$의 분산을 최소화하는 것은 오차확률을 최소화하는 것과 같기 때문에 최적 양자화기의 step size로서 최적 분류간격 $\triangle_k$가 주어진다. 만약 M이 충분히 크면 $x_k(i)$의 분포는 부분적으로 균일분포로 근사화될 수 있고 오차확률은 식 (16)을 따른다. 만약 M이 너무 과도하게 크면, N_k는 너무 작은 값이어서 $x_k(i)$가 백색(whiteness)의 특성을 잃어버릴 수 있다. 그러므로 오차확률을 감소시키기 위해서는 가능한 한 큰 값의 M이 선택되어야 하지만 $x_k(i)$는 잡음의 고유한 특성을 잃지 않도록 정해져야 한다.

워터마킹된 영상이 압축이나 일반적인 신호처리와 같은 공격에 의해 변형되었을 때, 워터마크 검출기의 입력은 $r(i)=x(i)+w(i)+e(i)$로 표현될 수 있다. 여기서 $e(i)$는 공격에 의해 야기된 잡음 신호이다. $e(i), x(i)$와 $w(i)$는 서로 독립이고 $E[e]=0$, $E[e^2]=\sigma_e^2$이라고 가정한다. 그때 상관기(correlator)의 응답 z는 다음과 같다.

$$\begin{aligned} z &= \left(\sum_{i=0}^{n-1} x(i)+w(i)+e(i))w(i)\right)/E_w \\ &= \left(\sum_{i=0}^{n-1} (x(i)+e(i))w(i)\right)/E_w+1 \end{aligned} \tag{18}$$

독립의 가정으로부터 z의 평균은 다음과 같이 쓸 수 있다.

$$E[z]=\left(\sum_{i=0}^{N-1} E[x(i)+e(i)]E[w(i)]\right)/E_w+1=1 \tag{19}$$

식 (4)의 $x(i)$에 $x(i)+e(i)$를 대입함으로써 분산은 다음과 같이 주어진다.

$$E[(z-E[z])^2]=N(\mu_x^2+\sigma_x^2+\sigma_e^2)\sigma_w^2/E_w^2 \tag{20}$$

$w(i)$가 에르고딕하다는 가정과 함께 분산은 다음과 같이 다시 쓸 수 있다.

$$\sigma_z^2=\frac{\mu_x^2+\sigma_x^2+\sigma_e^2}{E_w} \tag{21}$$

첫 번째 제안을 이용하여 는 식 (10)에서 보는 것처럼 쉽게 제거되고 일 때, 오차확률은 다음과 같다.

$$P_+ = P_- = \frac{1}{2} erfc\left(\sqrt{\frac{E_w}{8(\sigma_x^2 + \sigma_e^2)}}\right) \quad (22)$$

두 번째 제안을 균일 분포를 가지는x(i)에 적용하면 오차확률은 다음과 같이 감소한다.

$$P_+ = P_- = \frac{1}{2} erfc\left(\sqrt{\frac{E_w}{8(\sigma_x^2 / M^2 + \sigma_e^2)}}\right) \quad (23)$$

따라서 공격이 존재하는 경우에 대해서도 제안된 워터마크 검출방법이 기존 방법보다 검출오차를 줄일 수 있다는 것을 알 수 있다.

5. 실험결과 및 고찰

제안된 방식의 성능을 평가하기 위해서는, 많은 종류의 신호에 대해 성능을 테스트함으로써 제안된 방식이 기존 방식보다 통계적으로 우수함을 보여야 한다. 즉, 유한한 수의 영상에 대해 기존방식과 제안방식을 비교했을 때, 제안된 방식이 기존방식보다 항상 우수한 결과가 나온다는 보장은 없다. 따라서 가능한 많은 영상에 대해 실험하여야 하는데, 본 논문에서는 먼저 특정 주파수를 가진 합성신호들에 대해 실험하고, 다음으로 실제영상에 대해 실험하였다. 실제영상은 3가지 종류만을 사용하되, 대신 많은 종류의 워터마크 신호들을 이들 영상에 적용하므로 많은 실제 영상에 적용하는 효과를 내도록 하였다. 이를 위하여 상관응답 z의 분산 σ_z^2에 대한 추정치로서, L개의 샘플영상들에 대한 상관응답의 평균분산 $\hat{\sigma}_z^2$를 식 (24)처럼 도입하여 성능평가지수로 사용한다.

$$\hat{\sigma}_z^2 \equiv \frac{1}{L}\sum_{l=0}^{L-1}\left(\sum_{i=0}^{N-1} x_l(i) w_l(i) / E_{w_l}\right)^2 = \frac{1}{L}\sum_{l=0}^{L-1}(z_l - 1)^2 \quad (24)$$

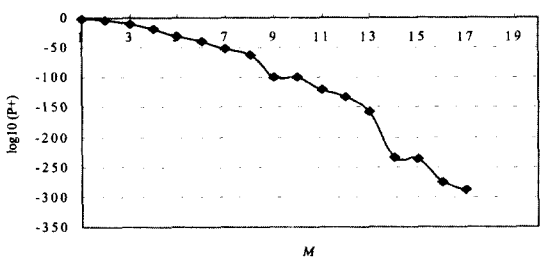
여기서 $x_l(i)$, $w_l(i)$, 및 l는 각각 워터마킹되는 l번째 영상, l번째 워터마크, l번째 워터마크의 에너지를 나타낸다. 식 (24)에 의하여 L개의 샘플 신호에 대한 상관응답의 분산을 계산하고 이들의 평균 $\hat{\sigma}_z^2$을 σ_z^2의 추정치로 사용한다. 그래서 $\hat{\sigma}_z^2$는 성능평가의 하나의 척도가 된다. 즉, 식(16)에서 볼 수 있듯이 오차확률은 σ_z^2에 반비례한다.

먼저 다양한 주파수를 갖는 정현파 신호(cosine function)를 워터마크 삽입을 위한 테스트 신호로 사용한다. 정현파 신호를 선택한 이유는 정현파 신호가 DCT의 기저함수(basis)이기 때문이다. 즉, 이 기저함수들의 조합으로 모든 신호를 만들 수 있으므로 이 기저함수에 대해 적용되어 얻어진 결과는 실제 신호에도 적용될 것으로 예상되기 때문이다. 정현파 함수는 식 (25)처럼 발생되었다.

$$x_f(i) = A\cos\left(\frac{(2f+1)\pi i}{2N}\right), \quad 0 \le i < N \quad (25)$$

여기서 $A, N,$ 및 f는 각각 $x_f(i)$의 피크값(peak value), 길이, 주파수를 나타낸다. 실험에서 이러한 파라미터의 값을 다음처럼 정하였다. A는 신호의 크기가 워터마크의 크기보다 충분히 크도록 1000으로 하였다. 그리고 워터마크의 길이는 충분히 길도록 N=100,000으로 하였고, 사용된 주파수는 $N/8 \le f < N/2$로 하였다. 따라서 375개의 신호, 즉 $x_{N8}(i)$, $x_{N8+100}(i)$, $x_{N8+200}(i)$...,가 워터마킹되었다. 그리고 워터마크는 평균이 0인 가우시안 분포를 가지며 분산은 64로 정규화되었다. 그 후, 워터마크 검출이 수행되는데 워터마크가 삽입된 375개의 정현파 신호와 이에 대응되는 워터마크 사이의 상관성, 즉 상관응답 z가 계산된다. 그 결과를 그림 5에 z의 분산 추정치 $\hat{\sigma}_z^2$와 오차확

(a) 클래스 개수 대 분산 평균치 $\hat{\sigma}_z^2$



(b) 클래스 개수 대 오차확률

그림 5. cosine 신호($1000\cos(\frac{(2k+1)\pi i}{2L})$, for $N/8 \le k < N/2$, watermark variance $=64.0, L=100,000$)에 대한 분산 평균치 $\hat{\sigma}_z^2$ 및 오차확률률 P

Fig. 5. Results for cosine functions,($1000\cos(\frac{(2k+1)\pi i}{2N})$, for $N/8 \le k < N/2$ watermark variance $=64.0, N=100,000$)

률 P_+로나타내었다. 그림 5로부터 알 수 있듯이 클래스의 개수, M이 1에서 20으로 증가함에 따라, $\hat{\sigma}_z^2$ 및 P_+가 현저히 감소된다. 여기서 $M=1$은 분류화를 행하지 않는 기존 방식에 해당한다. 그림 5(b)에서 M이 증가함에 따라, 오차확률, $P_+ = P_- = \frac{1}{2} erfc(1/\sqrt{8\sigma_z^2})$가 매우 급속히 감소하기 때문에 P_+에 대해 상용로그를 적용하여 $\log_{10} P_+$로 표시하였다. 즉 x축은 M을, y축은 $\log_{10} P_+ = \log_{10}(\frac{1}{2} erfc(1/\sqrt{8\sigma_z^2}))$을 나타낸다. 그림의 결과를 보면 M=1, 2, 3에 대해 오차확률은 각각 $10^{-1.6316} \approx 2\%$, $10^{-4.1205} \approx 0.0075\%$, $10^{-4.1205} \approx 0.0075\%$, $10^{-10.1134} \approx 7.7 \times 10^{-9}\%$ 로 나타난다. 기존방법($M=1$)이 2%정도의 오차확률을 나타내는 반면 제안된 방법은 이보다 훨씬 적은 오차확률을 나타냄을 알 수 있다. 특히 $M>17$에 대해서는 오차확률이 0에 근접하기 때문에 이에 대한 상용로그 값이 무한대이므로 그림에 표시할 수 없었다. 그림 5(a)에서 어떤 M이상의 값에 대해 이 값들이 포화 상태에 도달함을 알 수 있는데, 포화상태가 시작되기 직전의 M을 선택하는 것이 적합하리라 생각된다. 그 이유는 매우 큰 M값은 워터마크 신호의 whiteness 성질의 저하를 유발하기 때문이다.

실제영상(real images)에 대해 실험하기 위해 크기가 256×256인 Lenna, Airplane, Baboon 영상이 사용되었다. $\hat{\sigma}_z^2$를 얻기 위해, 각 영상에 대해 워터마크를 삽입하고 검출하는 과정을 200번 수행하여 그것의 평균 값을 구하였다. 워터마크 삽입은 DCT 영역에서 수행되었고, 중간주파수 대역에 대해 삽입되었다. 그 이유는 고주파 영역에 삽입된 워터마크는 필터링과 같은 조작에 쉽게 제거될 수 있고, 저주파 영역에 삽입된 워터마크는 쉽게 눈에 띄기 때문이다. 본 실험에서는, 먼저 원영상 $\xi(n,m)$으로부터 DCT 변환한 영상 (u,v)을 얻고, 이를 일차원으로 배열한 신호를 얻어 이들의 중간 주파수 대역에 대해 길이가 37,000인 워터마크가 삽입되는데 이를 $x(i)$, $0 \le i < N$로 표시한다. 식 (13) 또는 (17)에 따라 분류화 과정이 수행되고 클래스 k에 해당하는 신호를 $x_k(i)$, $i=0,1,\ldots,N_k-1$로 표현한다. 다른 한편 평균이 0, 분산이 1인 가우시안 분포를 갖는 워터마크 $x_k(i)$, $i=0,1,\ldots,N_k-1$를 발생하고 이를 식 (15)처럼 클래스로 나눈다. 먼저 첫번째 제안에 따라 각 클래스에 대해 직류성분이 제거된 워터마크가 식 (26)처럼 주어진다.

$$w_k'(i) = w_k(i) - m_k, \quad 0 \le i < N_k \tag{26}$$

여기서 m_k는 $w_k(i)$의 직류성분 즉, 이다. 마지막으로 식 (27)처럼 워터마크 $w_k'(i)$가 $x_k(i)$에 삽입된다.

$$y_k(i) = x_k(i) + \gamma | x_k'(i), \quad 0 \le i < N_k \tag{27}$$

여기서 γ는 워터마크 신호의 에너지를 조절하는 인자이다. 역스캔과 역변환을 수행함으로써 마침내 워터마킹된 영상 Ω가 얻어진다. 영상에 삽입되는 최종 워터마크는 $w_k'(i)$의 합집합으로 구성된다. 본 실험에서는 삽입되는 워터마크의 최종 분산 값이 64가 되도록 γ를 결정하였다.

표 1. 실제영상의 경우, 클래스 개수 M의 증가에 $\hat{\sigma}_z^2$ 따른 및 P_+
Table 1. $\hat{\sigma}_z^2$ and error probability for real images

Image	Measure	M=1	M=2	M=4	M=8	M=16	M=32
Lena	$\hat{\sigma}_z^2$	0.006977	0.005898	0.001596	0.000247	0.000086	0.000022
	$Log_{10}P_+$	-8.9684	-10.4267	-35.5136	-221.6868	-Inf	-Inf
Airplane	$\hat{\sigma}_z^2$	0.006036	0.003709	0.0001711	0.000419	0.000109	0.000023
	$Log_{10}P_+$	-10.2114	-15.9562	-33.2125	-131.3505	-Inf	-Inf
Baboon	$\hat{\sigma}_z^2$	0.002923	0.001117	0.000614	0.000202	0.000039	0.000007
	$Log_{10}P_+$	-19.9424	-50.1965	-90.1200	-270.6923	-Inf	-Inf

표 1은 Type2 검출기가 사용되었을 때, 3가지 영상에 대해 200 종류의 워터마크를 사용하여 계산한 $\hat{\sigma}_z^2$및 P_+를 나타낸다. 그림 5(b)처럼 표 1에서도 P_+대신 $\log_{10}(P_+)$가 사용되었는데 그 이유는 M이 증가함에 따라 P_+가 급속히 작은 값이 되기 때문이다. 표에서도 수 있듯이 $M \ge 16$에 대해 $\log_{10}(P_+)$가 $-\infty$로 나타남을 알 수 있는데 이는 오차확률 P_+가 0에 가까운 값임을 의미한다. 또한 M이 증가함에 따라 $\hat{\sigma}_z^2$가 현저히 감소함을 알 수 있다. 여기서 $M=1$은 기존방식에 해당한다. 앞에서도 언급한 것처럼 M은 과도하게 크지 않아야 한다. 역으로 M이 1로 접근하면 제안방식의 장점이 사라진다. 그러므로 제안된 방식의 성능을 최대화하도록 M을 선택하여야 한다. 최적의 클래스 개수 M을 선택하는 문제는 현재 연구 중에 있다.

6. 결 론

본 논문에서는 워터마크 검출의 신뢰성을 높이는 방법을 제안하였다. 본문에서 워터마크와 원영상 사이의 상호 상관성이 워터마크 검출에 큰 영향을 미침을 살펴보았다. 이것은 워터마크의 크기가 원영상의 크기에 비해 워낙 작기 때문이다. 즉, 비록 워터마크와 원영상 사이의 상호 상관성이 작더라도 워터마크의 에너지로 정규화되었을 때 얻어지는 그 값은 무시할 수 없을 만큼 크기 때문이다. 여기서 우리가 주목해야 할 점은 비록 워터마킹된 영상이 어떠한 공격을 겪지 않았음에도 불구하고 워터마크와 원영상 사이의 상호 상관성에 의해 오차확률이 존재한다는 점이다. 그래서 본 논문에서는 워터마크와 원영상 사이의 상호 상관성을 분석하고 이를 현저히 감소시킬 수 있는 방안 두 가지를 제안하였다. 이론적 그리고 실험적으로 제안된 방식은 워터마크 검출의 오차 확률을 크게 감소시킴을 알 수 있었다

참 고 문 헌

[1] Ioannis Pitas, "A Method for Watermark Casting on Digital Images," IEEE Transactions on Circuits and Systems for Video Technology, vol. 8, no. 6, pp. 775-780, Oct. 1998.

[2] Chiou-Ting Hsu and Ja-Ling Wu, "Hidden Digital Watermarks in Images," IEEE Transactions on Image Processing, vol. 8, no. 1, pp. 58-68, Jan. 1999.

[3] Christine I. Podilchuk and Wenjun Zeng, "Image-Adaptive Watermarking Using Visual Models," IEEE Journal on Selected Areas in Communications, vol. 16, no. 4, pp. 525-539, May. 1998.

[4] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "Copyright protection of digital images by embedded unperceivable marks," Image and Vision Computing, vol. 16, pp. 897-906, 1996.

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

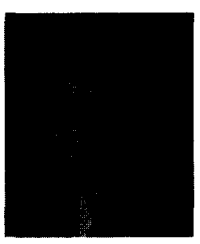
[6] R. Dugad, K. Ratakonda, and N. Ahuja, "A New Wavelet-Based Scheme for Watermarking Images," ICIP'98, pp. 419-423, 1998.

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

저 자 소 개

강 현 수

1969년 5월 18일생
1991년 2월 : 경북대학교 전자공학과 졸업 (공학사)
1994년 2월 : 한국과학기술원 전기및전자공학과 졸업 (공학석사)
1999년 2월 : 한국과학기술원 전기및전자공학과 졸업 (공학박사)
1999년 3월~2001년 4월 : 하이닉스㈜ 선임연구원
2001년 5월~현재 : 한국전자통신연구원 방송미디어부 선임연구원
주관심분야 : 영상처리, 영상부호화, 영상저작권 보호 기술 등



최 재 각

1984년 2월 : 경북대학교 전자공학과 졸업 (학사)
1987년 2월 : 한국과학기술원 전기및전자공학과 (석사)
1997년 8월 : 한국과학기술원 전기및전자공학과 (박사)
1987년 2월~1998년 2월 : 한국전자통신연구원 선임연구원
1998년 3월~2001년 8월 : 경일대학교 제어계측공학과 조교수
2001년 9월~현재 : 동의대학교 컴퓨터응용공학부 조교수
주관심분야 : 영상 및 멀티미디어 통신, 멀티미디어저작권보호

이 시 웅

1991년 2월 : 경북대학교 전저공학과 졸업(공학사)
1993년 2월 : 한국과학기술원 전기및전자공학과 졸업(공학석사)
1997년 8월 : 한국과학기술원 전기및전자공학과 졸업(공학박사)
1995년 5월~2000년 3월 : 삼성전자반도체
2000년 4월~현재 : 한밭대학교 조교수
주관심분야 : 영상처리, 영상압축, 컴퓨터비젼

안 치 득

1980년 2월 : 서울대학교 공과대학 전자공학과 졸업(학사)
1982년 2월 : 서울대학교 대학원 전자공학과 졸업(석사)
1991년 8월 : 미국 University of Florida 대학원 전기공학과 졸업(박사)
1982년 12월~현재 : 한국전자통신연구원 책임연구원 (방송시스템연구부장)
1996년 7월~현재 : MPEG-Korea 의장
1997년 5월~현재 : SC29-Korea 의장

홍 진 우

1978년 3월~1982년 2월 : 광운대학교 응용전자공학과 졸업 (공학사)
1982년 3월~1984년 2월 : 광운대학교 대학원 전자공학과 졸업 (공학석사)
1990년 3월~1993년 8월 : 광운대학교 대학원 전자계산기공학과 졸업 (공학박사)
1998년~1999년 : 독일 프라운호퍼연구소 (교환연구원)
1984년 3월~현재 : 한국전자통신연구원 방송컨텐츠연구팀장 (책임연구원)
2000년 1월~현재 : 한국음향학회 홍보이사, 뉴미디어음향 학술분과위원장, 한국방송공학회 편집위원
1993년 1월~현재 : 정보통신표준화연구단 방송기술위원회 위원
주관심분야 : 오디오 신호처리 및 부호화, 디지털 컨텐츠 보호 및 관리, 디지털 오디오 방송