

유전자 알고리즘을 이용한 인증 워터마킹 방법

이상광 · 호요성

본 논문에서는 유전자 알고리즘(genetic algorithm)을 이용한 인증 워터마킹 방법을 제안한다. 수신된 영상이 원본 영상과 동일한 특성을 유지하는지를 검사하기 위해 원본 영상 내의 블록들에 관한 윤곽선 정보를 이용한다. 또한 이 윤곽선 정보는 유전자 알고리즘을 이용한 메시지 부호화의 적응도(fitness) 값으로 사용된다. 유전자 알고리즘을 통해 진화된 비트열은 비트 기반의 배타적 논리합(exclusive OR) 연산을 사용하여 이진 워터마크와 조합된 후 원본 영상의 최하위 비트(least significant bit)에 삽입한다. 제안한 방법은 변조된 곳의 위치를 표시할 수 있다. 즉, 워터마크가 삽입된 영상의 변조는 검출된 워터마크 신호에 반영된다. 기존의 방법들과는 달리 워터마킹에 이용되는 매개 변수들을 이용하여 강인성의 정도를 조절할 수 있으며, 키 값에 대한 보안성을 높일 수 있다. 실험 결과는 제안한 방법이 JPEG 압축, 비례 축소(scaling), 최하위 비트 변조 및 잘라냄(cropping) 등과 같은 원본 영상의 변조를 검출할 수 있다는 것을 보여주었다.

주제어 : 인증 워터마킹, 블록 기반 윤곽선 분류, 유전자 알고리즘, 메시지 부호화

I. 서론

멀티미디어 콘텐츠의 디지털 표현은 원본 콘텐츠와 동일한 여러 복사본을 쉽고 널리 배포할 수 있다는 장점을 가지고 있다. 하지만 원본과 동일한 형태의 수많은 복사본이 불법적으로 생산될 수 있다는 사실은 콘텐츠 소유권에 대한 심각한 위협이 된다. 워터마킹 기술은 지적 재산권을 보호하기 위해 방송 감시(broadcast monitoring), 소유권 증명(proof of copyright), 거래 추적(transaction tracking), 콘텐츠 인증(content authentication), 복사 제어(copy control), 그리고 디바이스 제어(device control)와 같은 다양한 응용 분야에 적용될 수 있다[1]. 최근 디지털 콘텐츠의 보호를 위해 다양한 종류의 워터마킹 알고리즘들이 제안되고 있다[2],[3].

콘텐츠 인증의 주요 목적은 원본 데이터 내의 변조를 검출하는 것으로 데이터 압축과 같은 콘텐츠의 변형에 깨지기 쉬운 워터마크(fragile watermark)로 가능하다. 모든 워터마킹 응용들 중에 인증 워터마크는 가

장 낮은 수준의 강인성을 요구한다[4].

Friedman은 각각의 기록된 정지영상에 디지털 서명(digital signature)을 삽입하는 신뢰성 있는 디지털 카메라를 제안하였다. 디지털 서명을 통해 정지영상이 변조되었는지는 물론, 심지어 정지영상을 찍은 디지털 카메라도 확인할 수 있다[5]. Yeung과 Mintzer는 정지영상의 무결정성(integrity)을 보호하기 위해 의사랜덤 시퀀스(pseudo-random sequence)와 수정된 오류 확산(error diffusion) 방법을 이용한 인증 워터마크를 제안하였다[6]. Lin과 Chang은 JPEG 부호화 과정에서 DCT 계수에 인증 데이터를 삽입하여 JPEG 압축에 대해 어느 정도 탄력을 갖는 방법을 제안하였다[7]. Wong과 Memon은 정지영상 인증을 위한 비밀 및 공개 키 워터마킹을 제안하였다[8].

본 논문에서는 유전자 알고리즘을 이용한 인증 워터마킹 방법을 제안한다. 이 방법은 수신된 영상이 원본 영상과 동일한 특성을 유지하는지를 검사하기 위해 블록 기반의 윤곽선 분류를 이용한다. 윤곽선 정보는 유전자 알고리즘을 이용한 메시지 부호화의 적응도 값으

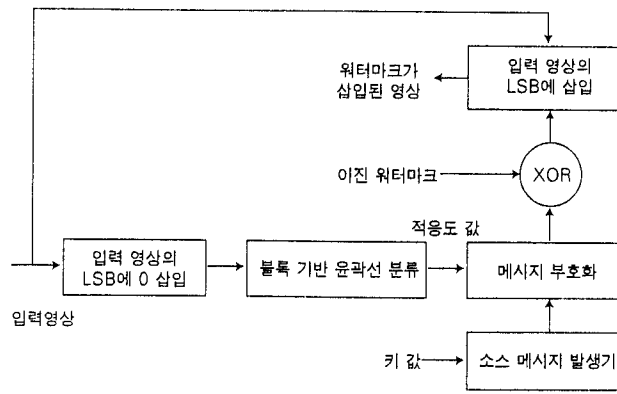


그림 1. 워터마크 삽입 알고리즘

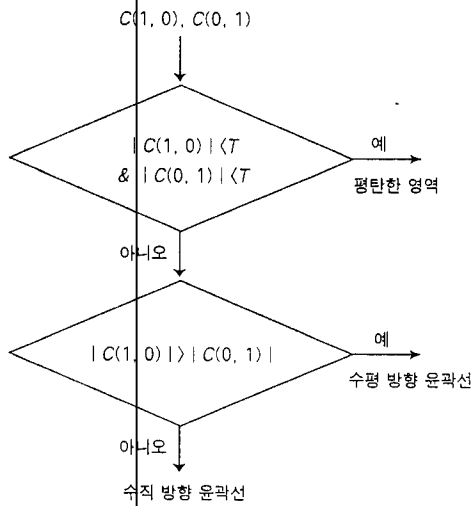


그림 2. 블록 기반 윤곽선 분류 알고리즘

중 워터마킹 방법에 대한 실험 결과를 보여준 후, V장에서 결론을 맺는다.

II. 워터마크 삽입

제안한 워터마킹 방법의 워터마크 삽입 과정은 그림 1과 같다. 워터마크 삽입 알고리즘은 블록 기반 윤곽선 분류(block-based edge classification), 메시지 부호화(message coding), 그리고 워터마크 삽입의 세 개의 주요 함수 블록들로 구성되며, 입력은 원본 영상, 소스 메시지(source message), 워터마크 키(watermark key) 값이다.

1. 블록 기반 윤곽선 분류

정지영상의 블록 윤곽선 특성을 이용하기 위해, 몇 개의 DCT 계수들에 블록 기반 윤곽선 분류 알고리즘을 적용한다. 예를 들면, 블록의 크기가 8x8일 때 64개의 DCT 계수들 중에 C(0, 1)과 C(1, 0)를 선택할 수 있으며 이는 각각 수직 및 수평 윤곽선 요소를 나타낸다. 두 계수는 다음과 같이 계산된다.

$$C(0, 1) = \frac{\sqrt{2}}{8} \sum_{m=0}^7 \sum_{n=0}^7 x(m, n) \cos \frac{\pi}{16} (2n+1)$$

$$C(1, 0) = \frac{\sqrt{2}}{8} \sum_{m=0}^7 \sum_{n=0}^7 x(m, n) \cos \frac{\pi}{16} (2m+1) \quad (1)$$

여기서 $x(m, n)$ 은 크기가 8x8인 영상 블록의 2차원 행렬을 나타낸다. 두 계수의 값을 기반으로 각 블록 내의 윤곽선 방향을 추정할 수 있다.

로 이용된다. 적응도 값의 개수는 워터마크의 강인성에 중요한 역할을 하기 때문에 제안한 알고리즘은 강인성의 정도를 조절할 수 있다. 유전자 알고리즘을 통해 진화된 비트열은 비트 기반의 배타적 논리합 연산을 사용하여 이진 워터마크와 조합된 후 원본 영상의 최하위 비트에 삽입한다. 제안된 워터마킹 방법은 워터마크가 삽입된 영상의 어떠한 변조도 검출할 수 있으며, 변조된 영역을 표시할 수 있다. 정지영상의 변조에 관한 위치 정보는 변조 동기 및 변형의 적법성을 추측하는 근거가 될 수 있으므로 유용하게 사용될 수 있다[1]. 또한, 일부 매개 변수들은 워터마크 키 값과 조합하여 부가적인 키 값으로 사용될 수 있어 키 값에 대한 강인한 보안성을 갖는다.

본 논문의 II장에서는 제안한 방법의 워터마크 삽입 과정에서 이용되는 블록 기반의 윤곽선 분류 방법, 유전자 알고리즘 및 삽입 알고리즘을 설명하고, III장에서는 워터마크 복원 방법을 설명한다. IV장에서 제안된 인

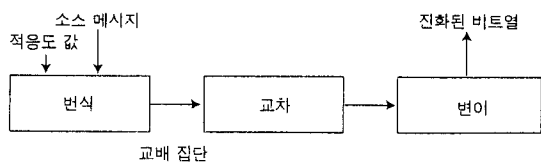


그림 3. 유전자 알고리즘

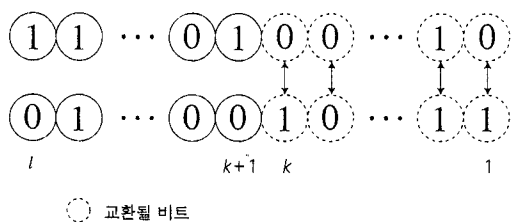


그림 4. 교차 과정

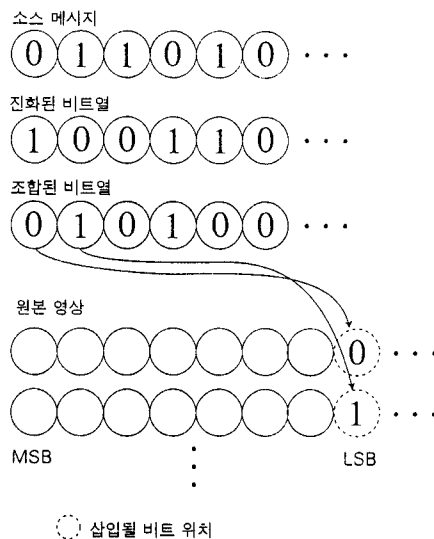


그림 5. 비트 단위의 워터마크 삽입 과정

그림 2는 율곽선 분류 알고리즘을 보여주고 있다. 여기서 T 는 율곽선 분류의 민감도(sensitivity)를 조절해주는 작은 양의 값이다[9]. 또한, T 는 워터마크 키 값과 더불어 부가적인 키 값으로 사용될 수 있다. 수신 영상의 인증 과정에서 인증 데이터의 삽입으로 인해 변형된 부분이 원본 영상의 특성을 검출하는데 영향을 주기 때문에, 원본 영상의 각 화소의 최하위 비트를 0으로 설정한 후 블록 기반의 율곽선 분류를 수행한다. 블록 율곽선 특성은 메시지 부호화의 적응도 값으로 이용된다.

2. 메시지 부호화: 유전자 알고리즘

그림 3은 유전자 알고리즘을 이용한 메시지 부호화 과정을 보여준다. 부호화 입력은 적응도 값과 소스 메시지이다. 소스 메시지는 유전자 알고리즘의 초기 비트열로 이용되며, 삽입 알고리즘의 비밀 키 값의 역할을 한다. 본 논문에서 이용하는 간단한 형태의 유전자 알고리즘은 단지 비트열 복사와 일부 비트열 교환으로 구성된다.

그림 3과 같이 유전자 알고리즘은 번식(reproduction), 교차(crossover), 그리고 변종(mutation)의 세 단계로 구성된다. 번식은 각 비트열이 그에 해당하는 적응도 값에 따라 복사되는 과정으로 큰 적응도 값을 갖는 비트열이 다음 세대에 하나 혹은 그 이상의 자손을 발생시킬 확률이 높다.

번식 과정이 끝나면 간단한 교차 과정이 두 단계로 진행된다. 먼저 교배 집단(mating pool) 내에 새롭게 번식된 비트열들은 무작위로 쌍을 이룬다. 그런 후 각 비트열 쌍은 다음과 같이 교차 과정을 거치게 된다. 그림 4는 교차 과정을 나타내고 있다.

- $[1, l-1]$ 의 범위에서 무작위로 정수 위치 k 를 선택한다. 여기서 l 은 비트열의 길이를 나타낸다.
- $[k+1, l]$ 의 범위에 있는 모든 비트들을 교환하여 두 개의 새로운 비트열을 만든다.

비록 번식과 교차가 유전자 개념들을 효과적으로 반영할지라도, 때로는 지나치게 과열하여 특정 위치의 잠재적으로 유용한 유전자 요소를 손실하므로 변종 과정이 필요하다. 인공 유전자 시스템(artificial genetic system)에서 변종 과정은 그러한 회복할 수 없는 손실을 막아준다. 본 논문에서 이용한 유전자 알고리즘에서 변종은 어느 비트열 위치의 비트 값을 간헐적으로 무작위 교체하는 것으로 수행된다[10].

3. 워터마크 삽입

본 논문에서는 이진 로고 영상을 워터마크로 이용하고 원본 영상의 각 화소의 최하위 비트에 인증 데이터를 삽입한다. 워터마크 W 와 유전자 알고리즘에 의해 진화된 비트열 S 의 크기가 원본 영상 X 의 크기에 비해 상대적으로 작기 때문에 X 와 동일한 크기가 되도록 W 와 S 를 복사해 붙인다. 배타적 논리합 함수의 세계의 요소들 중에 두 개의 입력을 알고 있다면 다른 세번째 요소를 유도할 수 있다는 특성을 이용하여 진화된 비트열은 비트 기반의 배타적 논리합 연산을 이용하여 이진 워터마크와 조합된다.

$$E = S \oplus W \tag{2}$$

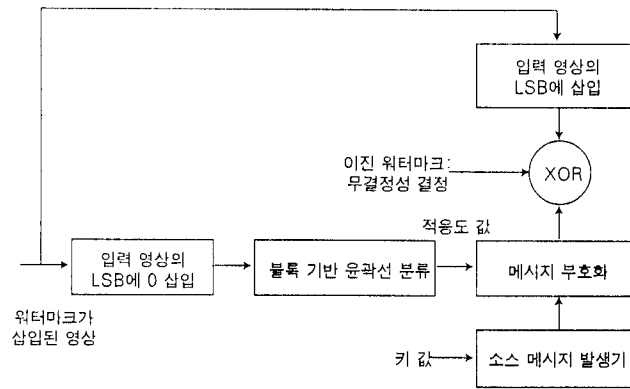


그림 6. 워터마크 검출 알고리즘



그림 7. 변조된 영상들

여기서 \oplus 는 배타적 논리합을 나타낸다. 마지막으로 E 는 원본 영상의 최하위 비트들에 삽입된다. 그림 5는 비트 단위의 워터마크 삽입 과정을 보여준다.

합 연산을 수행한다.

$$W_r = S_d \oplus D \tag{3}$$

III. 워터마크 검출

그림 6은 워터마크 복원 방법을 보여준다. 입력은 워터마크가 삽입된 영상, 워터마크 키 값, 그리고 소스 메시지이다. 이진 워터마크를 복원하기 위해 진화된 비트열 S_d 와 검출된 비트열 D 에 비트 기반의 배타적 논리

워터마크가 삽입된 영상이 변조되지 않았을 때 적절한 이진 워터마크 영상이 검출될 것이다. 그림 7과 같이 화소 값 또는 크기가 변조되거나 워터마크 검출 과정에서 올바르게 않은 키 값이 사용되면 검출된 워터마크 신호는 왜곡될 것이다.

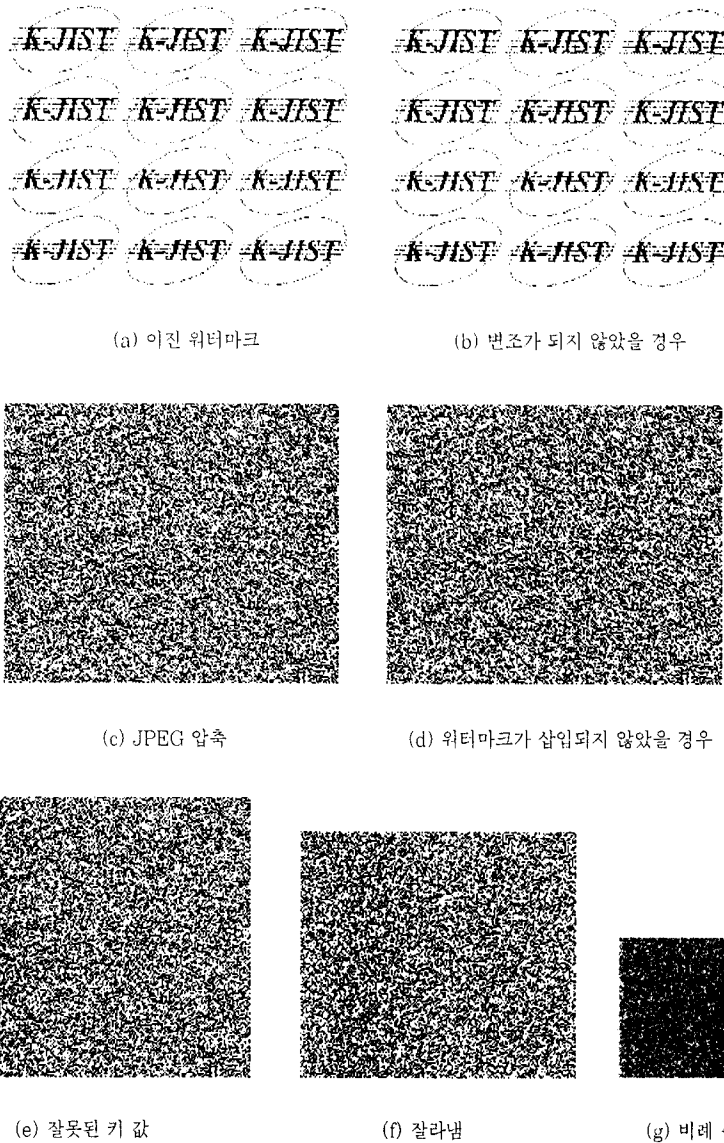


그림 8. 검출된 이진 워터마크

IV. 실험 및 결과

본 논문에서 제안한 인증 워터마킹 알고리즘의 성능을 평가하기 위해 352x288 크기의 실험 영상에 컴퓨터 시뮬레이션을 수행하였다. 그림 7과 같이 워터마크가 삽입된 영상에 대한 다양한 변조를 고려하였으며, 그림 8은 그 결과로 검출된 워터마크 영상을 보여준다.

실험 결과는 제안한 워터마킹 알고리즘이 키 값이 잘못되었는지, 영상에 워터마크가 삽입되어 있는지, 또는 영상이 변조되었는지를 검출할 수 있다는 것을 보여준다. 즉, JPEG 압축, 비례 축소(scaling), 최하위 비트 변조 및 잘라냄(cropping) 등과 같은 원본 영상의 변조에 대해서 잡음과 같은 형태의 잘못된 워터마크가 검출된다.

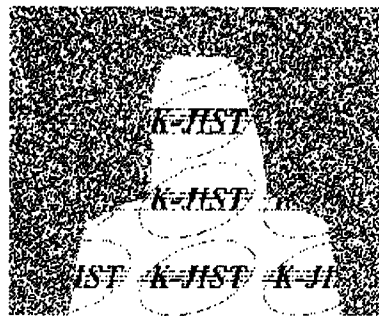
표 1에 실험 결과를 정리하였다. 또한, 워터마크가

삽입된 영상의 일부 최하위 비트들을 변형시켰을 때 변형된 영역이 워터마크 검출 과정의 출력에 반영된다. 그림 9는 워터마크가 삽입된 영상 내에서의 화소 변화가 검출된 워터마크 신호 내에 어떻게 반영되는지를 보여준다.

워터마크가 삽입된 영상에 대한 시각적인 결과를 살펴보기 위해 PSNR(Peak-Signal-to-Noise Ratio) 값을 사용하였으며, 대략 51 dB의 결과를 나타내었다. 이는 워터마크 삽입에 의해 최하위 비트 값이 확률적으로 전체 영상 크기의 절반 가량 바뀌므로 $10 \times \log_{10}(2 \times 255^2) = 51.1411$ dB 근방의 값을 갖는다는 것을 보여주고 있다. 또한, 워터마크 비트가 최하위 비트 값에 삽입되기 때문에 그림 7(a)와 (b)에서 보듯이, 맨눈으로는 원본 영상과 워터마크가 삽입된 영상과의 차이를 구별할 수 없다.



(a) 화소 변화 영역



(b) 검출된 이진 워터마크

그림 9. 변조된 영역의 국부화

표 1. 실험 결과

	무결정성	
	예	아니오
변조되지 않음	○	
JPEG 압축		○
잘려짐		○
비례 축소		○
워터마크를 삽입하지 않음		○
잘못된 키 값 사용		○
최하위 비트 변조		○

V. 결 론

본 논문에서는 인증을 위한 새로운 워터마킹 방법을 제안하였다. 실험 결과는 워터마크가 삽입된 영상이 맨 눈으로는 원본 영상과 구별이 되지 않지만, 워터마크가 삽입된 영상의 어떠한 변조도 검출할 수 있으며 변조된 영역을 표시할 수 있다는 것을 보여주었다. 키 값이 틀리거나, 영상에 워터마크가 삽입되어 있지 않거나, 또는 워터마크가 삽입된 영상이 변조되면, 검출된 워터마크는 손상되며 이러한 모든 영상들은 인증되지 않은 것이라고 말할 수 있다. 또한, 유전자 알고리즘을 이용한 메시지 부호화의 적응도 값의 개수는 워터마크의 강인성에 중요한 역할을 하기 때문에 제안한 알고리즘은 강인성의 정도를 조절할 수 있으므로 반인증(semi-fragile) 워터마킹에 적용될 수 있다. 유전자 알고리즘을 이용한 인증 워터마킹 방법은 영상 뿐만 아니라, 오디오 및 동영상 신호 등과 같은 다른 멀티미디어 콘텐츠에도 적용될 수 있다.

(참고 문헌)

- [1] I. Cox, M. Miller and J. Bloom, *Digital watermarking*, Morgan Kaufmann Publishers, 2001.
- [2] C. Podilchuk and E. Delp, "Digital watermarking: algorithms and applications," *IEEE Signal Processing Magazine*, Vol. 18, 2001, pp. 33-46.
- [3] S. Lee and S. Jung, "A survey of watermarking techniques applied to multimedia," *Proc. IEEE International Symposium on Industrial Electronics*, Vol. 1, 2001, pp. 272-277.
- [4] S. Katzenbeisser and F. Petitcolas, *Information hiding techniques for steganography and digital watermarking*, Artech House, 2000.
- [5] G. Frieman, "The trustworthy digital camera: Restoring credibility to the photographic image," *IEEE Trans. Consumer Electron.*, Vol. 39, 1993, pp. 905-910.
- [6] M. Yeung and F. Mintzer, "Invisible watermarking technique for image verification," *Proc. ICIP* Vol. 2, 1997, pp. 680-683.
- [7] C. Lin and S. Chang, "A robust image authentication method surviving JPEG lossy compression," *Proc. SPIE*, Vol. 3312, 1998, pp. 296-307.
- [8] P. Wong and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," *IEEE Trans. Image Processing*, Vol. 10, 2001, pp. 1593-1601.
- [9] B. Shen and I. Sethi, "Direct feature extraction from compressed images," *Proc. SPIE*, Vol. 2670, 1996, pp. 404-414.
- [10] D. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Addison Wesley, 1989.

A Fragile Image Watermarking Scheme Using A Simple Genetic Algorithm

In this paper, we propose a fragile image watermarking scheme using a simple genetic algorithm. In order to check whether the received image maintains the same attributes as the original image, we exploit characteristics of block edges in the image. We also use the edge information as a fitness measure for message coding in the simple genetic algorithm. The evolved strings are combined with the binary watermark signal by the bit-wise exclusive OR operation and then inserted into the least significant bits of the original image. The proposed scheme can identify specific locations where changes have been made. Pixel changes in the watermarked image are reflected in the extracted watermark, which indicates the attacked area. Experimental results demonstrate that the proposed watermark extraction scheme can detect modifications of the original image by JPEG compression, scaling, pixel changes, or cropping.

Keywords : Fragile watermarking, block-based edge classification, genetic algorithm, message coding



호요성
(Yo-Sung Ho)

1981. 2: 서울대학교 전자공학과 학사
 1983. 2: 서울대학교 전자공학과 석사
 1983. 3~1995. 9: 한국전자통신연구소 선임연구원
 1989. 12: University of California, Santa Barbara
 Dept. of Electrical and Computer Engr.
 박사
 1990. 1~1993. 5: 미국 Philips 연구소
 Senior Research Member
 1995. 9~현재: 광주과학기술원 정보통신공학과
 부교수
 관심분야: 디지털 신호처리, 영상신호처리 및 압축,
 초저속 영상통신, 디지털TV와 고선명 TV
 방식, 멀티미디어 통신
 E-mail: hoyo@kjist.ac.kr
 Tel: +82-62-970-2211
 Fax: +82-62-970-2247



이상광
(Sang-Kwang Lee)

1996. 2: 한국항공대학교 항공전자공학과 학사
 1998. 8: 광주과학기술원 정보통신공학과 석사
 1998. 9~현재: 광주과학기술원 정보통신공학과
 박사과정
 관심분야: 영상 및 오디오 부호화, 멀티미디어 통신,
 디지털 워터마킹
 E-mail: sklee@kjist.ac.kr
 Tel: +82-62-970-2258
 Fax: +82-062-970-2247