웨이블릿 계수의 쿼드트리를 이용한 영상 압축

Image coding using quad-tree of wavelet coefficients

김성탁, 추형석, 전희성, 이태호, 안종구

Sung Tak Kim, Hyung Suk Chu, Hee Sung Jun, Tae Ho Lee, Chong Koo An

요 약

EZW(Embedded Image Coding Using Zerotrees of Wavelet Coefficients) 알고리즘의 경우는 해상도를 높이기 위해 임계값을 낮출 때 중요계수의 위치를 나타내기 위해 필요 없는 zero 값을 많이 코딩을 해야 하는 문제가 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 쿼드 트리(quad tree) 방법을 이용하여 위치정보를 나타내었다. 그리고 그 계수들의 크기는 EZW 에서 사용한 방법과 유사한 메트릭스(matrix) 방법을 사용하였다. 본 논문에서 제안한 알고리즘으로 영상을 압축하게 되면 EZW와 MRWD(Morphological Representation of Wavelet Data)의 방법을 동시에 이용하는 효과가 있어 압축 효율이 증가함을 알 수 있었다.

Abstract

EZW(Embedded coding using Zero-trees of Wavelet coefficients) decreases symbol-position information using zero-trees, but threshold value fall for raising resolution, then coding cost of significant coefficients is expensive. To avoide this fact, this paper uses quad-tree representing coefficient-position information. a magnitude of significant coefficient is represented on matrix used at EZW. the proposed algorithm is hoped for raising a coding cost.

Keywords : Wavelet coefficients, quad-tree, EZW

I. 서 론

웨이블릿 변환(wavelet transform)은 기저함수의 스케일(scale)을 변화시켜 기저함수의 길이가 가변적인 특성을 가지도록 함으로 시간영역 또는 주파수영역에서 함수의 분해능을 바꿀 수 있도록 하는 변환법을 일컫는다. 웨이블릿 변환에서 사용되는 기저함수는 원형 웨이블릿이라 불리는 기본 스케일의 함수를 확장(dilation) 또는 압축(contraction)시켜 스케일을 변화시키고 그 함수를 이동(shifting)시킴으로써 얻을 수 있다. 이런 웨이블릿은 1990년 Haar에 의해 처음으로 언급되었고, 1930년대의 과학자 Paul Levy[1]가 Haar의 웨이블릿 기저함수가 랜덤하고 비정상(non-stationary)의 특징을 가지는 신호를 분석할 경우 푸리에의 기저함수보다 우수한 성능을 나타냄을 검증하였다. 이렇게 계속 연구를 해오다, 1985년 Mallat[2]이 그의 논문에서 컴퓨터 비젼에서 많이 이용되던 MRA(Multi-Resolution Analysis)와 웨이블릿이 아주 밀접한 관계가 있다는 것을 발견, 웨이블릿을 공학적 측면에서 바라 볼 수 있게 되었고, 디지털 신호처리에 웨이블릿을 이용할 수 있게 되었다. 이렇게 공학적으로 이용하기 시작한 웨이블릿은 시간-주파수 국소성(spatial-frequency localization), 에너지 집중성(energy compaction), 중요계수들의 군집성(clustering of significant coefficients), 그리고 서브밴드(subband)간의 계수크기의 쇠퇴성(decay of magnitude of wavelet coefficients across subbands) 등의 장점이 있기 때문에 이미지 압축에 많이 이용되고 있다. 위의 특징 중에서 서브밴드간의 계수의 쇠퇴성을 이용한 EZW의 경우는 POS(Positive Significant), NES(Negative Significant), IZ(Isolated Zero), ZTR(Zero-Tree Root), Z(Zero)의 기호를 사용하여 중요계수들의 위치정보를 나타내었다. 그런데 EZW의 경우는 ZTR가 많이 나타나야 코딩효율이 좋은데 해상도를 높이기 위해 임계값을 낮게 하면, ZTR이 많이 나타나지 않아 코딩 효율이 떨어지는 단점이 있다. 이러한 단점을 극복

하기 위해 제안한 알고리즘에선 이런 중요계수의 위치정보를 이미지의 분할정보를 나타내는 쿼드 트리를 변형하여 나타내었다. 사용한 기호는 S(Split), N(Non-split), C(significant Coefficient) 이다. 그리고 중요계수의 크기를 결정하는 방법은 3-2절에 설명하였다.

본 논문의 구성은 다음과 같다. 2절에서는 웨이블릿을 이용한 영상의 분리 및 복원, 3절에선 제안한 알고리즘인 쿼드트리를 이용한 이미지 압축, 4절에서 실험 및 결과, 마지막으로 5절에서 결론과 앞으로 연구할 내용으로 끝을 맺었다.

II. 웨이블릿을 이용한 영상의 분리 및 복원

이미지에 대해 웨이블릿변환을 수행하는 과정은 다운 샘플링(down-sampling)과 필터 뱅크(filter bank)로 구성되어 있고, 복원하는 과정은 업 샘플링(up-sampling)과 필터뱅크로 구성되어 있다. $h_0(n)$은 신호의 분리를 위한 저역통과 필터이고, $g_0(n)$은 신호의 합성을 위한 저역통과 필터이다. $h_1(n)$는 신호의 분리를 위한 대역통과필터이며, $g_1(n)$은 신호의 합성을 위한 대역통과 필터이다. 이런 직교 쌍교성 필터(biorthogonal filter)는 에이리어싱(aliasing)의 발생을 방지하기 위하여 아래의 식(1)[3]을 만족해야 한다.

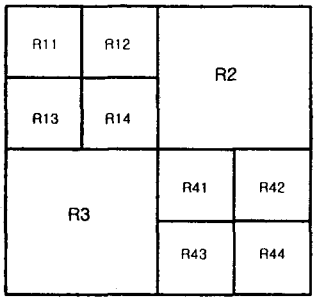
$$\begin{aligned} g_1(n) &= (-1)^{n+1} h_0(n) \\ h_1(n) &= (-1)^n g_0(n) \end{aligned} \quad (1)$$

III. 쿼드 트리를 이용한 압축

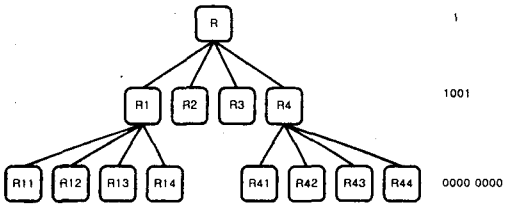
본 논문에서 제안하는 알고리즘은 기존의 쿼드트리 정보를 구하는 것을 변형하여 주어진 임계값에서 중요계수의 위치를 구하고, EZW[4]에서 제안한 방법으로 중요계수의 절대값을 결정, 그리고 부호를 결정하게 된다.

1. 중요계수의 위치정보 추출

제안한 알고리즘에서 중요계수들의 위치정보를 나타내기 위해 사용하는 쿼드 트리[5]는 이미지 분석법의 하나인 영역방향성 분할기법(region-oriented segmentation)중에서 이미지를 분할할 때, 분할 정보를 나타내는 것이라 볼 수 있다. 그림 1에 간단한 이미지분할에 대한 쿼드트리를 나타내었다.



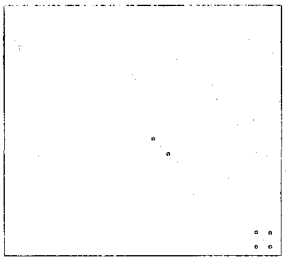
(a) 분할된 이미지



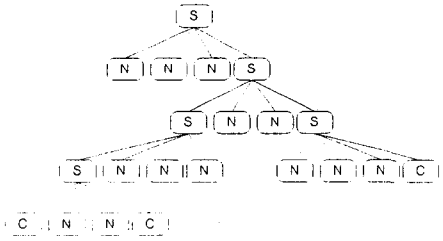
(b) 상응하는 쿼드트리

그림 1. 분할 영상과 쿼드트리 표현
Fig. 1 Partitioned image and Definition of quad-tree

그림 1에서 쿼드 트리 정보를 비트 스트림(bit stream)으로 나타내면 1 1001 0000 0000으로 나타낼 수 있다. 1은 분할이 가능함을 0은, 더 이상 분할이 불가능함을 나타낸다. 이와 같은 쿼드 트리를 사용하여 중요계수의 위치를 나타내는 정보로 바꿀 수 있다. 이미지를 분할하는 조건은 분할 시 중요계수가 있으면 분할하고 없으면 분할하지 않았다. 이런 사실을 이용하여 위치정보를 나타내기 위해 본 논문에선 S(split), N(non-split), 그리고 C(significant coefficient)라는 기호를 사용하였다. 그림 2에 간단한 예를 16×16크기의 이미지에 대해 나타내었다.



(a) 16×16의 중요계수 맵



(b) 상응하는 쿼드트리

그림 2. 16×16의 중요계수 맵에 대한 예
Fig. 2 The example of 16×16 significant coefficient map

그림 2의 결과를 심벌로 나타내면, S NNNS SNNS SNNN NNNC CNNC로 나타낼수 있다. S는 쿼드 트리 정보의 1과 같이 분할이 가능함을, N은 0과 같이 분할이 불가능함을, 그리고 C는 중요계수를 나타낸다.

2. 중요계수의 값 결정

여기서 설명할 계수들의 값을 결정하는 방법은 EZW에서 소개한 것과 같다. 간단히 설명하면, 어떤 임계값(T)의 범위에 들어가는 계수의 값을 결정하는 것이데 이때 임계값을 T라고 하면 T와 2T사이에 있는 계수들을 중요계수라 두고, 이 계수가 2T와 (2T+T)/2 사이에 있으면 1, (2T+T)/2에서 T 사이에 있으면 0이라 둔다. 이렇게 구한 비트 스트림을 가지고 복호기(decoder)에선 메트릭스를 구성하여 업데이트 된 인터벌(interval)과 비교해서 값을 결정하는 것이다. 이런 방법으로 계수의 값을 결정하게 되면 임계값을 낮출수록 복원값이 압축전의 값과 가까워져서 해상도가 높아진다. 그리고 마지막으로 부호비트(sign-bit)를 보내면 값을 결정할 수 있다. 결과를 보면 임계값 T를 낮출수록 압축전 값과 가까워짐을 알 수 있다. 이미지에 대해서도 같은 방법으로 수행 한다. 코딩은 적응산술부호화(adaptive arithmetic coding)[6]를 하였다. 그리고 효율을 조사하기 위해 비트율(bit-rate)과 PSNR(Peak Signal to Noise Ratio)을 조사하였다. PSNR은 왜곡의 정도를 표시하고 식(2)와 같다.

$$PSNR = 10\log_{10}\frac{255^2}{D} \qquad (2)$$

여기서 D는 원래 영상과 압축 복원한 영상간의 MSE(Mean Squared Error)를 의미하며, 255는 화소의 최대값이다.

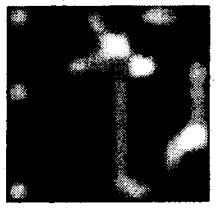
Ⅳ. 실험 및 결과

본 논문에서 수행한 실험의 목적은 웨이블릿 변환으로 얻어진 계수들을 보다 효율적으로 압축하는데 있다. 본 논문에서 제안한 알고리즘은 기존의 EZW의 알고리즘과는 중요계수의 위치정보를 나타내는 방법만 다르고 나머지는 같다. EZW의 경우는 중요계수의 위치 정보를 제로 트리라는 기호를 이용하여 코딩효율을 높였지만, EZW의 경우는 해상도를 높이기 위해 임계값을 내리면 제로 트리 구조가 효율이 떨어지는 현상이 있다. 그래서 본 논문에선 그런 현상을 피하기 위해서 쿼드 트리로 나타내었다.

실험에선 512×512 크기의 그레이 레벨 영상인 "LENA"영상과 "BARBARA"영상을 사용하여 실험을 하였다. 여기서 "LENA"영상은 일반적으로 고주파성분이 적은 영상에 속하고, "BARBARA"영상은 고주파성분이 많은 영상에 속한다. 웨이블릿 변환에서는 정규 쌍직교 필터를 가지고 스케일 레벨 6의 이산 웨이블릿 변환을 수행하였다. 이렇게 구한 계수를 임계값을 변화시키면서 압축을 수행한다. 압축 시 필요한 정보는 중요계수의 위치를 나타내는 쿼드 트리 정보, 중요계수의 크기를 나타내는 정보, 그리고 중요계수의 부호를 나타내는 정보가 필요하다. 그리고 코딩은 적응산술부호화를 이용하여 비트 스트림으로 전송하였다. 실험은 임계값을 2^{13}에서 2^5까지 변화를 시키면서 각각의 임계값에서의 비트율 과 PSNR을 조사하였다.

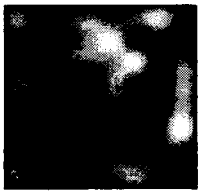
그림 3과 표 1는 "LENA"이미지에 대한 실험결과를 나타내었다

(a) Original image

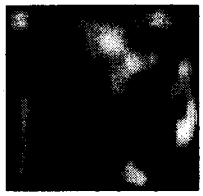


(b) T = 2^{13}

(a)"LENA" image (b)0.00057 bit/pixel



(c) T = 2^{12}



(d) T = 2^{11}

(c) 0.00125 bit/pixel (d) 0.00202 bit/pixel

(e) T = 2^{10}

(f) T = 2^{9}

(e) 0.00401 bit/pixel (f) 0.00803 bit/pixel

(g) T = 2^{8}

(h) T = 2^{7}

(g) 0.0164 bit/pixel (h) 0.0332 bit/pixel

(i) T = 2^{6}

(j) T = 2^{5}

(i) 0.0686 bit/pixel (j) 0.1371 bit/pixel

그림 3. 압축전과 압축후의 비트율과 PSNR ("LENA"영상)

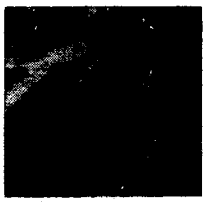
Fig. 3 Bit-rate of original image and compressed image ("LENA"image)

표 1. "LENA"영상에 대한 비트율과 PSNR

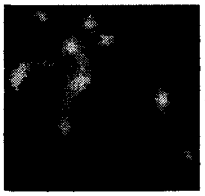
Table 1. Bit-rate and PSNR of "LENA" image

bit-rate (bpp)	PSNR (dB)	bit-rate (bpp)	PSNR (dB)	bit-rate (bpp)	PSNR (dB)
(T= 2^{13}) 0.00057	9.43	(T= 2^{10}) 0.00401	17.92	(T= 2^{7}) 0.0332	25.69
(T= 2^{12}) 0.00125	14.71	(T= 2^{9}) 0.00803	20.33	(T= 2^{6}) 0.0686	28.79
(T= 2^{11}) 0.00202	16.20	(T= 2^{8}) 0.0164	22.82	(T= 2^{5}) 0.1371	32.10

그림 4 와 표 2 는 "BARBARA"이미지에 대한 결과를 나타내었다.



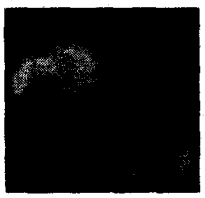
(a) Original image



(b) T = 2^{13}

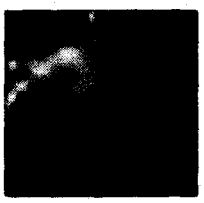
(a) "BARBARA"image (b) 0.00052 bit/pixel

(c) $T = 2^{12}$

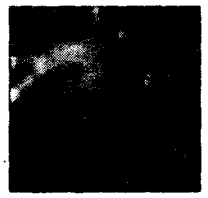


(b) $T = 2^{11}$

(c) 0.00130 bit/pixel (d) 0.00200 bit/pixel

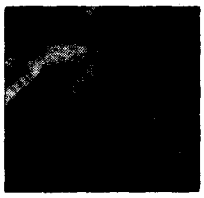


(e) $T = 2^{10}$

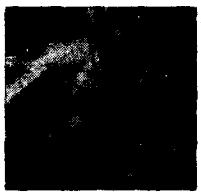


(f) $T = 2^{9}$

(e) 0.00340 bit/pixel(f) 0.00630 bit/pixel



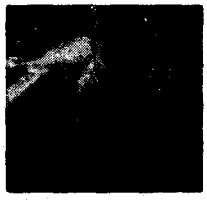
(g) $T = 2^{8}$



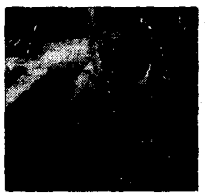
(h) $T = 2^{7}$

]

(g) 0.0130 bit/pixel (h) 0.0290 bit/pixe



(i) $T = 2^{6}$



(j) $T = 2^{5}$

(i) 0.0870bit/pixel j) 0.2120bit/pixel

그림 4. 압축전과 압축후의 비트율 ("BARBARA"영상)

Fig. 4 Bit-rate of original image and compressed image ("BARBARA"image)

표 2. "BARBARA"영상에 대한 비트율과 PSNR
Table 2. Bit-rate and PSNR of "BARBARA" image

bit-rate (bpp)	PSNR (dB)	bit-rate (bpp)	PSNR (dB)	bit-rate (bpp)	PSNR (dB)
($T=2^{13}$) 0.00052	8.99	($T=2^{10}$) 0.00340	18.31	($T=2^{7}$) 0.0290	22.90
($T=2^{12}$) 0.00130	16.02	($T=2^{9}$) 0.00630	19.87	($T=2^{6}$) 0.0870	26.08
($T=2^{11}$) 0.00200	16.64	($T=2^{8}$) 0.0130	21.40	($T=2^{5}$) 0.2120	30.04

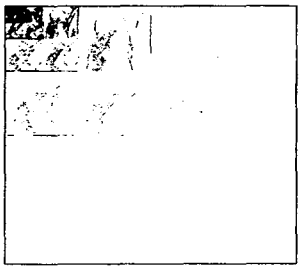
표 3에서는 Shapiro가 제안한 EZW 알고리즘을 이용하여 Lena, Barbara 영상을 코딩한 결과를 나타내었다.

표 3. EZW 알고리즘을 이용한 LENA, BARBARA 영상의 코딩 결과
Table 3. Coding Results for 512×512 LENA and BARBARA using EZW algorithm

Image	LENA	BARBARA
bit-rate(bpp)	PSNR(dB)	PSNR(dB)
1.0	39.55	35.14
0.5	36.28	30.53
0.25	33.17	26.77
0.125	30.23	24.03
0.0625	27.54	23.10
0.03125	25.38	21.94
0.015625	23.63	20.75
0.0078125	21.69	19.54

위의 실험 결과를 비교하여보면 본 논문에서 제안한 알고리즘의 실험 결과에서 임계값이 2^8 이고 비트율이 0.00164bpp 일 때 PSNR은 22.82dB을 보였고, Shapiro의 실험결과에선 0.015625bpp의 비트율에서 23.63dB의 PSNR을 얻었다. 여기에서 임계값이 2^8이상에서는 EZW 알고리즘이 좋은 성능을 보임을 알 수 있고, 임계값을 더 내릴 경우에는 제안한 알고리즘과 비슷한 압축률을 보임을 확인할 수 있었다. 또한 Barbara 영상에 대한 실험에서 제안한 알고리즘은 임계값이 2^9이고 비트율이 0.00630bpp 일 때 PSNR은 19.87dB, EZW 알고리즘을 이용한 경우는 0.0078125bpp 의 비트율에서 19.54dB의

PSNR을 보였다. 그리고 임계값이 2^5 이고 비트율이 0.2120bpp 에서는 30.04dB, EZW의 실험에서는 0.25bpp 비트율을 보일 때 26.77dB의 PSNR을 나타내었다. 그러므로 Barbara 영상에서는 EZW 알고리즘보다는 제안한 알고리즘을 이용하였을 때 보다 향상된 압축률을 보임을 확인할 수 있다. Lena 영상과 Barbara 영상을 비교하였을 때 상대적으로 고주파 성분이 많은 Barbara 영상에서 제안한 알고리즘이 향상된 압축률을 보임을 알 수 있고, 이러한 결과를 보여주는 것은 고주파 성분이 많은 Barbara 영상의 경우는 해상도를 높이기 위하여 임계값을 낮게 하면 서브밴드에 중요계수들이 많이 생기게 되므로 제로 트리 구조가 적게 나타나기 때문이다.
그림 5는 임계값 T가 2^6일 때 "LENA"이미지와 "BARBARA"이미지에 대해 나타내었다.



(a) $T=2^6$일 때 significant coefficient

(b) $T=2^6$일 때 significant coefficient

그림 5. $T=2^b$일 때 "Lena"이미지와 "Barbara"이미지의 중요계수
Fig. 5 Significant coefficient of "Lena" image and "Barbara" image with $T=2^b$

그림 5를 보면 같은 임계값 $T=2^6$에 대하여 중요계수를 보면 "Barbara"이미지일 때 서브밴드에서 계수가 더 많이 나타나 EZW의 제로 트리 구조가 잘 생기지 않을 것을 예상할 수 있고, 제안한 알고리즘의 경우가 효율이 더 높아질 것을 예상할 수 있다. 하지만 상대적으로 고주파 성분이 적은 "Lena"image의 경우는 임계값을 낮게 하여도 서브밴드에 계수들이 많이 나타나지 않으므로 EZW의 제로 트리 구조가 많이 나타나므로 압축효율이 높다.

V. 결 론

본 논문에서는 웨이블릿 변환을 수행한 계수를 코딩할 때 기존의 EZW가 가지고 있는 단점, 즉 해상도를 높이기 위해 임계값을 낮게 할 경우 제로 트리 구조가 잘 생기지 않아 중요계수의 위치를 나타내기 위해 필요 없는 zero를 많이 코딩해야 한다는 점을 보완하고자 실험을 수행하였다. 웨이블릿 변환을 수행 후, 임계값을 만족하는 계수를 코딩할 때 그 계수들의 위치정보를 가장 효율적으로 나타내는 것이 중요하다. EZW에서는 제로 트리를 본 논문에서는 쿼드 트리를 사용하였다. 본 논문에서 제안한 알고리즘은 서브밴드에 계수가 많이 위치하고 있을 땐 그 효과가 좋지만, 그렇지 않을 경우는 EZW에 비해 압축율이 높을 때 그 효율이 떨어짐을 알수 있다. 하지만 임계값을 낮추어서 해상도를 높이게 되면 압축효율이 좋아짐을 알 수 있었다. 여기서 더 좋은 효과를 얻으려면 계수의 크기를 결정하는 방법을 더 연구해서 나타내면 효율을 많이 높일 수 있을 것으로 기대된다.

접수일자 : 2000. 9. 8 수정완료 : 2000. 12. 2

이 논문은 1999년도 정보통신부 지원 정보통신 분야 우수지원사업 지원금에 의하여 연구되었음.

참고문헌

[1] Y. Meyer,"Wavelets: Algorithm and Applications," Society for Industrial and Applied Mathematics, Philadelphia, 1993, pp. 13-31, 101-105.

[2] S. Mallat,"A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Patt. Anal. Machine Intell, vol. 11, no. 7, pp. 674-693, July 1989.

[3] Robin N. Strickland and Hee II Hahn, "Wavelet Transform Methods for Object Detection and Recovery", IEEE Trans. image processing, vol. 6, no. 5, pp. 725, May 1997.

[4] Jerome M. shapiro,"Embedded Image Coding Using Zerotrees of Wavelet Coefficients," IEEE trans. signal processing, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.

[5] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*. Addison-Wesley Publishing Company, Inc.

[6] I. H. Witten, R. Neal, and J. G. Cleary,"Arithmetic coding for data compression," Comm. ACM, vol. 30, pp. 520-540, June 1987.

[7] R. Gray,"Vector Quantization," IEEE ASSP Mag., pp. 4-29, Apr. 1984.

[8] P. Cosman, R. Gray and M. Vetterli,"Vector Quantization of Image Subbands: A Survey," IEEE Trans. image processing, vol. 5. no. 2. Feb. 1996.

김성탁(Sung Tak Kim)
準會員
2000년 울산대학교
전자공학과(공학사)
2000년 울산대학교 대학원
전자공학과
석사과정 재학중
관심분야 : wavelet, 영상신호처리

추형석(Hyung Suk Chu)
準會員
1997년 울산대학교
전자공학과(공학사)
1999년 울산대학교 대학원
전자공학과(공학석사)
2000년 울산대학교 대학원 전자공학과
박사과정 재학중
관심분야 : wavelet, 영상신호처리

전희성(Hee Sung Jun)
正會員
1981년 서울대학교
전기공학과(공학사)
1983년 서울대학교
전기공학과(공학석사)
1992년 Rutgers-The State University of New
Jersey (Ph. D.)
1983년-1986년 금성반도체(주) 주임연구원
1992년-1993년 삼성전자(주) 통신연구소
수석연구원
1993년-현재 울산대학교
컴퓨터 · 정보통신공학부 부교수
관심분야 : 컴퓨터비젼, 영상처리, 가상현실

이태호(Tae Ho Lee)
正會員
1966년 한양대학교
공과대학(공학사)
1969년 서울대학교
(공학석사)
1975년 연세대학교
(공학박사)
1969년~1972년 한국과학기술연구소 연구원.
1973년~현재 울산대학교
전기전자 및 자동화공학부 교수
관심분야 : 음성신호처리, 통신시스템

안종구(Chong Koo An)
正會員
1977년 서울대학교
전기공학과(공학사)
1979년 한국과학기술원
전기및전자공학과
(공학석사)
1989년 Univ. of Texas at Austin(공학박사)
1979-1980년 현대중공업주식회사
1980년~현재 울산대학교 전자공학과 교수
관심분야 : wavelet, higher-order spectrum