VQ 인덱스의 상관도를 이용한 향상된 웨이브렛 영상 압축

황 재 호†·홍 충 선††·이 대 영†††

요 약

본 논문에서는 웨이브렛 영역에서 인접한VQ 인덱스들의 상관도를 이용하는 웨이브렛 영상 압축 기법을 제안한다. 각 서브 코드북에 있는 코드 워드들은 인덱스들의 상관도를 증가시키기 위해서 에너지 크기 순으로 오름차순으로 정렬된다. 그리하여 벡터 양자화 후에 생성되는 인덱스는 비적응적DPCM/Huffman 방법에 의해 더 압축될 수 있다. LBG 알고리즘과 k-d 트리를 이용한 빠른 PNN 알고리즘이 다중해상도 코드북 생성에 사용된다 실험 결과를 통하여 제안된 기법이 저 비트율에서 일반적인 웨이브렛 벡터 양자화 기법과 JPEG 보다 뛰어남을 보여준다.

Improved Wavelet Image Compression Using Correlation of VQ index

Jae Ho Hwang†·Choong Seon Hong††·Dae Young Lee†††

ABSTRACT

In this paper, a wavelet image coding scheme exploiting the correlation of neighboring VQ indices in the wavelet domain is proposed. The codewords in each sub-codebook are re-ordered in terms of their energies in order to increase the correlation of the indices. Then, the generated indices after VQ can be further encoded by non-adaptive DPCM/Huffman method LBG algorithm and a fast-PNN algorithm using k-d trees are used for generating a multiresolution codebook Experimental results show that our scheme outperforms the ordinary wavelet VQ and JPEG at low bit rates

1. Introduction

Discrete wavelet transform (DWT) has been often used for data compression, because of their close relationship with subband coding. Images can be easily decomposed into multiresolution images using a simple quadrature mirror filter (QMF) bank which consists of analysis and synthesis filters [1, 2]. DWT-based coding schemes provide two main advantages compared to the DCT-based ones such as JPEG. One is the progressive transmission in which significant bits in low scales (high frequencies) are sent to a receiver prior to insignificant bits in high scales(low frequencies) so that an approximated image can be reconstructed in a very short time period The other is the coding efficiency from which the different amount of bits can be allocated onto each subband so that a fractional bit rate can be accomplished.

Vector quantization (VQ) [3] has been preferred by many researchers compared to the scalar quantization (SQ). According to Shannons distortion theory, VQ

† 정 회 원 극동정보대학 멀티미디어과 교수
†† 종신회원 경희대학교 전자정보학부 교수
††† 정 회 원 경희대학교 전자정보학부 교수

논문접수 2000년 4월 20일, 심사완료 : 2000년 6월 14일

overwhelms SQ in coding efficiency, even if the dependency between data doesnt exist. In VQ-based coding schemes, each input block is mapped to the index of the closest codeword in a codebook in terms of a minimum distortion criterion (mainly MSE) in an encoding process The VQ indices instead of image vectors themselves are transmitted to a decoder in which images can be reconstructed by looking up the codebook, the same as in an encoder.

VQ can be divided into two categories, or memoryless VQ and memory VQ. In case of memoryless VQ, tree-structure VQ (TSVQ) [4] and classified VQ (CVQ) [4] has been researched in order to reduce the search complexity possibly at the expense of image quality. Memory VQ schemes generally require more computational complexity and memory size than memoryless ones, even though the bit rate can be more or less reduced. Various memory VQ schemes [4] such as predictive VQ (PVQ), multistage VQ (MVQ), adaptive VQ, and finite-state VQ (FSVQ) have been developed to further increase coding efficiency.

Since W-VQ [5] has been presented, a number of image coders [6] using both VQ and wavelet decomposition have been emerged in order to further improve coding efficiency. In W-VQ, 2-D biorthonormal wavelet filters are used for image decomposition and the ordinary VQ is applied to the wavelet coefficients.

In this paper, we propose a wavelet image coding scheme which exploits the redundancy of VQ indices generated from W-VQ Instead of increasing vector dimensions or codebook sizes, this scheme can increase the coding efficiency by using DPCM and entropy coding to exploit the correlation among VQ indices in a subband This requires that the values of the neighboring indices be close each other. It assumes that the neighboring vectors of wavelet coefficients in a subband would have similar energies Therefore, the indices in a codebook need to be rearranged according to the amount of the energy of their codewords so that the values of the neighboring indices can be close each other.

In Chapter 2, we illustrate the wavelet transform and multiresolution analysis. Codebook re-ordering is discussed in Chapter 3 And the proposed coding scheme is described in Chapter 4.

Chapter 5 depicts experimental environments and results. Finally, a conclusion is given in Chapter 6.

2. Wavelet transform and multiresolution analysis

Wavelets are functions generated from one single function Ψ by dilations and translations,

$$\psi^{a,b}(t) = |a|^{-1/2}\Psi\left(\frac{t-b}{a}\right) \qquad (1)$$

The mother wavelet Ψ has to satisfy

$$\int \Psi(x)\,dx = 0 \qquad (2)$$

and have at least some oscillations.High frequency wavelets correspond to $a<1$ or narrow width, while low frequency wavelets have $a>1$ or wider width.

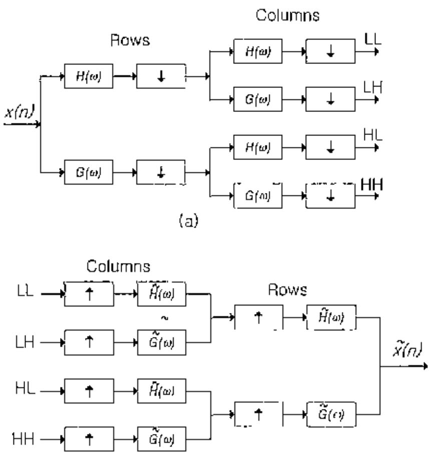
The basic idea of the wavelet transform is to represent any arbitrary function f as a superposition of wavelets. The wavelet transform decomposes f into different scale levels, where each level is then further decomposed with a resolution adapted to the level. The signal f can be depicted as an summation over a and b of $\Psi^{a,b}$ with appropriate weighting coefficients by the following equationf nd

$$f = \sum c_{m,n}(f)\,\Psi_{m,n}, \qquad (3)$$

where the coefficients $c_{m,n}$ can be defined as

$$c_{m,n}(f) = \int \Psi_{m,n}(t) f(t)\,dt. \qquad (4)$$

The wavelet transform can also be viewed as a special case of multi-rate filter bank with dyadic tree decomposition It can be implemented by quadrature mirror filter (QMF) banks using wavelet filters. When wavelet filters are used for image application, the perfect reconstruction of decomposed images is needed An exact reconstruction condition corresponds with the orthonormal wavelet basis, since a number of portions of images are smooth except for small edge area [2]. Linear phase is also needed to avoid border distortion and a short length of a filter required for fast computation. Unfortunately, there are no FIR filters satisfying the orthonormality and linear phase so that biorthonormal wavelet filters are often used for quadrature mirror filter (QMF) banks. A QMF bank which consists of an analysis and a synthesis filter is shown in (Figure 1)



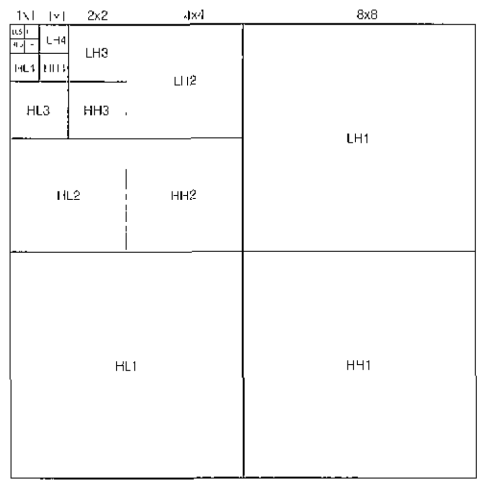
(Figure 1) A QMF bank (a) analysis filters, (b) synthesis filters.

In (Figure 1(a)) , $x(n)$ denotes an input image, $H(\omega)$ a low-pass filter, $G(\omega)$ a high-pass filter, and denotes down-sampling by a factor of two. The signal $x(n)$ splits into a smooth and a detail image by applying $H(\omega)$ and $G(\omega)$ to the rows of the image, respectively. The low-pass filtered image is down-sampled by a factor of two keeping one column out of two. The subsequent filters and down-sampling are then applied to the resulting columns In this way we can obtain four subband images LL, LH, HL, and HH according to whether the rows and columns of the image is low- or high-pass filtered

Exactly the opposite process is done in reconstruction phase shown in (Figure 1b) where $\widetilde{H}(\omega)$ and $\widetilde{G}(\omega)$ indicate a low-pass and a high- pass filter, respectively, ↑ upsampling, and and $\widetilde{x}(n)$ indicates a reconstructed image $\widetilde{H}(\omega)$ and $\widetilde{G}(\omega)$ must be specially adapted to the analysis filters $H(\omega)$ and $G(\omega)$ to satisfy a perfect reconstruction.

3. Codeword reordering

We use the 5-scale wavelet decomposition shown in (Figure 2) where LL5 denotes the lowest frequency band and LH(i), HL(i), and HH(i) denotes level-i horizontal directional band, vertical directional band, and diagonal directional band, respectively The dimensions of the subbands are 1x1 (scale-5), 1x1 (scale-4), 2x2 (scale-3), 4x4 (scale-2), and 8x8 (scale-1). The subbands of scale-5 and -4 are scalar-quantized because it has very significant information. The other bands are vector-quantized according to how important information they have. This means that in case of lower frequency (or higher scale) band more bits are assigned and higher vector dimensions are used, and in case of higher frequency (or lower scale) band the opposite rule is applied to.



(Figure 2) 5-scale wavelet decomposition

The real PDF of wavelet coefficients in any directional band (horizontal, vertical, or diagonal band) can be approximated by a generalized Gaussian law [5]. And the blocks of the coefficients also have the high redundancy so that the correlation of the interblock in a subband needs to be exploited However, the arbitrary order of the codebook generated from LBG (Linde-Buzo-Gray) algorithm may result in lowering the interblock correlation. Subsequently, the entropy increases so that the correlation is not efficiently exploited.

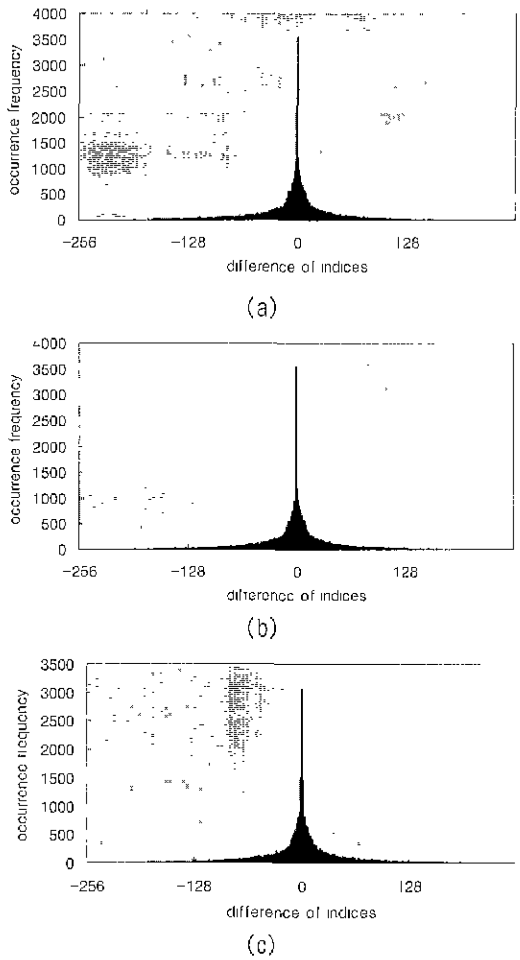
Palriwal and Ramasubramanian [7] have studied the effect of ordering the codevectors on the computational complexity of an encoder. Explicit ordering of codevectors allows the minimum possible values to be found as early as possible so that premature exit condition can be satisfied to reduce computational time In their paper codevectors has been sorted according to the sizes of their corresponding clusters They concluded that an explicit ordering of the codevectors is advantageous for deriving more computational savings from the partial distance method.

In this paper codevectors, or codewords in each subband are sorted in an ascending order in terms of their energies Energy E_i in an ith codevector is defined,

$$E_i = \sum_{j=1}^{k} c_{ij}^2, \qquad (5)$$

where c_{ij} denotes a wavelet coefficient in codevectors, k is the dimension of the codevectors and the reordering scheme is based on the fact that neighboring indices in a subband have similar energy values. It provides two main advantages : 1) fast codebook search algorithms can be used, 2) entropy coding can be applied for the VQ indices to explore the interblock correlation.

(Figure 3) shows the histograms of the occurrence frequency against the difference between the current index and its left-side index in decomposition scale-2 (LH2, HL2, and HH2 in (Figure 2) for each directional orientation In all three bands, the occurrence frequency of the difference near 0 is very high so that entropy coding for the difference of neighboring VQ indices can be used to further reduce the bit rates.



(Figure 3) The histograms of the occurrence frequency of the difference between the current index and its left-side index for decomposition level-2, 4x4 vector dimension, and 256 codewords : (a) LH2 band, (b) HL2 band, and (c) HH2 band

4. The proposed coding scheme

4.1 Multiresolution Codebook design

Each codebook in accordance with each subband shown in (Figure 2) should be designed The codebooks called a multi-resolution codebook are designed by LGB [8] except for scale -5 and -4. The scalar codebooks of scale-5 and -4 band are designed by Lloyd and Max algorithm [4] which uses two conditions : the nearest neighbor condition and the centroid condition. The best encoder for a given decoder satisfies the nearest neighbor condition which requires that the ith region of the partition cells should consist of values closer to the ith codeword, y_i than to any other output level.

For a given set of output alphabets, the partition cells satisfy

$$R_i \subset \{x: d(x, y_i) \le d(x, y_j);\ all\ j \ne i\} \qquad (6)$$

Inversely, the best decoder for a given encoder satisfies the centroid condition. Given a partition cell R_i, the optimal codebook for a random variable X is given by

$$y_i = E[X \mid X \in R_i] \qquad (7)$$

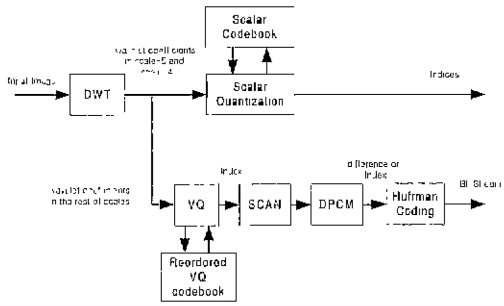
The algorithm operates in the iterative manner and stops when the average distortion error in the process of codebook generation is less than the predefined threshold value.

LBG algorithm allows an iterative and converging capability so that local minimum can be accomplished. The problem resides in the fact that its performance is highly dependent on the initial codebook. In order to obtain the good codebook using this algorithm, it is required to obtain good initial codebook before LBG is applied. To generate initial value for LBG we use a fast PNN algorithm [9, 10] using k-d trees.

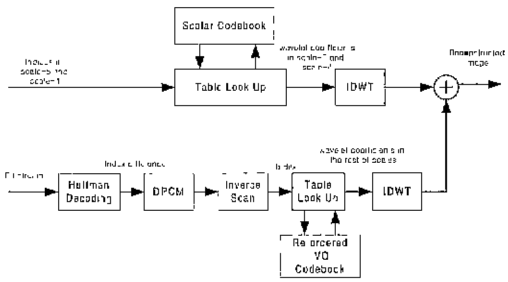
4.2 Encoder and decoder design

The encoder and the decoder in the proposed coding scheme have the same set of the reordered codebooks corresponded to each wavelet subband. The functional block diagram of the encoder is shown in (Figure 4). The input image is decomposed into the 5-scale wavelet subband by using DWT, or QMF banks. The subbands in the scale-5 and -4 are scalar-quantized. The rest of the bands called are vector-quantized by reordered codebooks.

The architecture of the proposed decoder is an opposite process of the encoder, as shown in (Figure 5). The wavelet coefficients in scale-5 and -4 are obtained by looking up the scalar codebook from the input indices. The bit streams from an input bit stream are Huffman decoded. Then VQ indices are obtained by DPCM and an inverse scan. The wavelet coefficients in the rest of scales emerge by searching the reordered vector codebook from the VQ indices to emerge the difference of the indices. The wavelet coefficients from which their gray level values are obtained by inverse DWT are summed to result the reconstructed image.

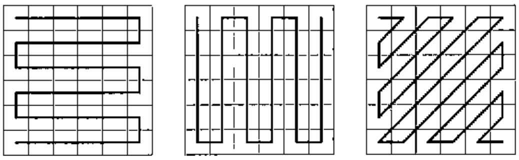


(Figure 4) The functional block diagram for the proposed encoder.



(Figure 5) The functional block diagram for the proposed decoder.

The scanned VQ indices are further encoded by DPCM and Huffman method. This means that the relative address of the codebook is transmitted to the receiver instead of the absolute address. The generated VQ indices are scanned in terms of the orientation of its band shown in (Figure 6).



(a) (b) (c)

(Figure 6) The scanning pattern corresponding to each directional orientation : (a) LH-band, (b) HL-band, and (c) HH-band.

5. Experimental Results

All simulations have been performed with Pentium III (450 MHz) by C++ language. We generated a multi-resolution codebook by using 10 test images (512x512, 256 gray levels) as the training set. In order to prove the coding efficiency of the proposed scheme, the PSNRs have been compared with given bit rates (bit per pixel) for Lena (outside training set), Barbara (outside training set), and Goldhill (inside the training set) image. The 9/7-tab wavelet filter bank [5] shown in <Table 1> is used for the decomposition of test images and a reflection extension method used for minimizing their border distortion.

Mean square error (MSE) and peak signal-to-noise ratio(PSNR) is defined by

$$MSE = \frac{1}{N^2} \sum_i \sum_j (x_{ij} - \tilde{x}_{ij})^2, \tag{8}$$

$$PSNR = \log_{10}\left(\frac{255^2}{MSE}\right) dB, \tag{9}$$

where x_{ij} and $\tilde{x}_{ij}$ denote the gray levels of the original and reconstructed images, respectively.

〈Table 1〉 The coefficients of the wavelet filter used for our proposed scheme.

Filter	Symmetrical filter coefficients				
H_0	0.859	0.377	-0.110	-0.023	0.037
G_0	0.788	0.418	-0.046	-0.064	-

Given several low bit rates, PSNRs for Lena, Goldhill, and Barbara image are shown in <Table 2> <Table 3>, and <Table 4>, respectively. As shown in the tables, JPEG and the proposed scheme show a similar image quality in 0.25 bpp (bit per pixel). But it goes to lower bit rates, our scheme outperforms JPEG. In the same PSNRs, our coder requires less bits than W-VQ.

〈Table 2〉 The comparison of bit rate and PSNR for the proposed coding scheme with JPEG and W-VQ for Lena image

JPEG		W-VQ		Proposed	
dB	bpp	dB	bpp	dB	bpp
28.90	0.247	29.32	0.329	29.32	0.256
28.22	0.233	29.18	0.306	29.18	0.229
27.73	0.217	28.75	0.259	28.75	0.206
24.73	0.185	28.61	0.212	28.61	0.179
18.00	0.155	28.03	0.189	28.03	0.163

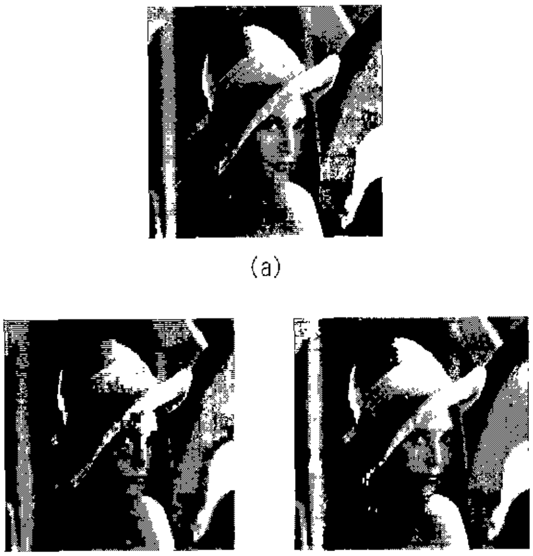
〈Table 3〉 The comparison of bit rate and PSNR for the proposed coding scheme with JPEG and W-VQ for Goldhill image

JPEG		W-VQ		Proposed	
dB	bpp	dB	bpp	dB	bpp
27.95	0.255	27.99	0.274	27.99	0.250
27.36	0.237	27.65	0.259	27.65	0.231
26.56	0.217	27.58	0.236	27.58	0.217
25.62	0.197	27.57	0.212	27.57	0.197
24.32	0.180	27.12	0.189	27.12	0.178

〈Table 4〉 The comparison of bit rate and PSNR for the proposed coding scheme with JPEG and W-VQ for Barbara image

JPEG		W-VQ		Proposed	
dB	Bpp	dB	bpp	dB	bpp
23.19	0.242	23.52	0.274	23.52	0.246
22.47	0.218	23.35	0.259	23.35	0.223
21.72	0.195	23.28	0.212	23.28	0.192
20.22	0.173	23.01	0.189	23.01	0.172
16.97	0.155	22.71	0.142	22.74	0.134

(Figure 7) shows the original image (a) and the reconstructed images from JPEG (b) and from the proposed scheme (c). In case of JPEGs result of (Figure 7b), the well-known blocking artifact appears. The experimental results for Lena image show more or less better image quality than those of Barbara and Goldhill image. It attributes to the fact that DPCM/Huffman method used in our scheme is well adapted to smooth images.



(a)

(b) (c)

(Figure 7) The result of Lena image (a) Original image, (b) Reconstructed image (24.73 dB) with JPEG at 0.185 bpp, (c) Reconstructed image(28.61 dB) with W-VQ and Proposed scheme at 0.212 bpp and 0.179 bpp, respectively.

6. Conclusion

In this paper, a wavelet image coding scheme using vector quantization and DPCM/Huffman method in wavelet domain is proposed. In order to exploit the correlation of neighboring VQ indices, the codewords in the multiresolution codebook are sorted in ascending order according to their energies. After the codebook re-ordering, the dependency of the indices increase so that we could reduce the required bits by applying DPCM/Huffman method to the indices.

From the experiments, the proposed coding scheme is much superior than JPEG at low bit rates. And our scheme requires less bits than W-VQ at the same bit rate. The distribution of the gray levels in original image is highly correlated with wavelet coefficients so that it influences on the coding efficiency. The research of fast codebook search algorithms in wavelet domain remains for future study.

References

[1] M. Vetterli and C. Herley, "Wavelets and Filter Banks," IEEE Trans. Signal Processing, Vol.40, pp. 2207-2232, Sept. 1992.

[2] G. Strang, T. Nguyen, "Wavelets and Filter Banks," Wellesley-Cambridge Press, pp.365-370, 1997.

[3] N. M. Nasrabdi, Image Coding Using Vector Quantization : A Review," IEEE Trans. Commun., Vol. COM-36, pp.957-971, Aug. 1988.

[4] A. Gersho, R. M. Gray, "Vector Quantization and Signal Compression," Kluwer Academic Publishers, Boston, 1992.

[5] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image Coding using Wavelet Transform," IEEE Trans. Image Processing, Vol.1, No.2, pp. 205-220, Apr. 1992.

[6] P. C. Cosman and R. M. Gray, "Vector Quantization of Image Subbands : A Survey," IEEE Trans. Image Processing, Vol.5, No.2, pp.202-225, Feb. 1996.

[7] K. K. Paliwal and V. Ramasubramanian, Effect of Ordering the Codebook on the Efficiency of the Partial Distance Search Algorithm for Vector Quantization , IEEE Trans. Commun., Vol.37, No.5, pp.538-540, May 1989.

[8] Y. Linde, A. Buzo, and R. M Gray, "An Algorithm for Vector Quantizer Design," IEEE, Trans. Commun., Vol COM-28, pp 84-95, Jan 1980.
[9] J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching," Communications of the ACM, Vol.18, No.9, pp.509-517, Sep 1975.
[10] W. H. Equitz, "A new Vector Quantization Clustering Algorithm," IEEE Trans. On Acoustics, Speech, and Sig., Vol 37, No.10, Oct 1989.

황 재 호
e-mail : mjrin@unitel.co.kr
1986년 경희대학교 전자공학과 졸업(학사)
1991년 경희대학교 전자공학과 (공학 석사)
2000년 경희대학교 전자공학과 (박사수료)
1991년 1998년 한국통신 전송기술 연구소 전임 연구원
1998년~현재 극동정보대학 멀티미디어과 전임강사
관심 분야 영상 및 비디오 압축, 인터넷 서비스, 객체 지향 시스템 등

홍 충 선
e-mail : cshong@nms.kyunghee.ac.kr
1983년 경희대학교 전자공학과 졸업(학사)
1985년 경희대학교 전자공학과 (공학석사)
1997년 게이오대학교 정보통신공학과(공학 박사)
1988년~1999년 한국통신 통신망 연구소 선임 연구원/네트워킹연구실장
1999년~현재 경희대학교 전자정보학부 전임강사
관심분야 : 인터넷 서비스 및 망 관리 구조, 분산컴포넌트관리, IP 멀티캐스트, 멀티미디어스트리밍 등

이 대 영
e-mail dylee@nms.kyunghee.ac.kr
1964년 서울대 물리학과 졸업 (학사)
1971년 캘리포니아 주립대학원 컴퓨터학과 (공학석사)
1979년 연세대학교 전자공학과 (공학 박사)
1971년~현재 경희대학교 전자정보학부 교수
1990년~1993년 경희대학교 산업정보대학원 대학원장
1999년~2000년 한국통신학회 회장
관심분야 : 영상처리, 컴퓨터 네트워크, 컴퓨터 시스템