

# 인간 시각 시스템의 공간 지각 특성을 이용한 개선된 이진트리 벡터양자화

유 성 필<sup>†</sup> · 곽 내 정<sup>†</sup> · 안 재 형<sup>††</sup>

## 요 약

칼라 양자화는 칼라 카메라로부터 입력된 자연색의 칼라 영상을 컴퓨터 모니터에 표현하기 위해 더 적은 칼라 팔레트로 구성하는 것이다. 이것은 자연색 중에서 인간이 인식할 수 있는 색은 수백가지 정도일뿐 아니라 실제 칼라 영상에 나타나는 색은 이 보다는 훨씬 적다는 것을 바탕으로 한다. 본 논문에서는 인간의 시각 특성중 공간 지각 특성을 고려한 개선된 고유벡터를 이용한 이진 트리 벡터양자화 기법을 제안한다. 제안 방법은 고유벡터를 이용한 이진 트리 벡터 양자화의 두 노드로 분할하는 과정에 영상의 블록 내의 칼라의 변화에 따른 인간의 시각의 반응 정도를 가중치로 결합하여 양자화를 하였다. 제안 방법은 기존의 방법보다 색상이 선명해지며 유사한 영역의 분할에 뛰어난 성능을 보여 주었으며 주관적인 화질 평가와 PSNR에서도 좋은 결과를 보였다.

## The Improved Binary Tree Vector Quantization Using Spatial Sensitivity of HVS

Soung-Pil Ryu<sup>†</sup> · Nae-joung Kwak<sup>†</sup> · Jae-hyeong Ahn<sup>††</sup>

## ABSTRACT

Color image quantization is a process of selecting a set of colors to display an image with some representative colors without noticeable perceived difference. It is very important in many applications to display a true color image in a low cost color monitor or printer. The basic problem is how to display 256 colors or less colors, called color palette. In this paper, we propose improved binary tree vector quantization based on spatial sensitivity which is one of the human visual properties. We combine the weights based on the responsibility of human visual system according to changes of three primary colors in blocks of images with the process of splitting nodes using eigenvector in binary tree vector quantization. The test results show that the proposed method generates the quantized images with fine color and performs better than the conventional method in terms of clustering the similar regions. Also the proposed method can get the better result in subjective quality test and WSNR.

키워드 : Vector Quantization, Binary Tree, MTF

### 1. 서 론

칼라 카메라로부터 입력되는 자연색 영상은 한 화소를 표현하기 위하여 적색, 녹색 및 청색의 세 채널을 필요로 하며 각 채널에 대해 8비트로 2<sup>24</sup>가지의 색을 가진다. 그러나 사람의 눈은 일정한 크기의 화소에 대하여 수십 가지의 휘도와 수 백 가지의 색상만을 느낄 수 있으며 실제 영상에 있어서도 사람이 느끼는 색상의 수는 표현할 수 있는 화소수보다 훨씬 적기 때문에 2<sup>24</sup>가지의 색을 수백가지의 색상으로 양자화 하더라도 인식되는 화질은 크게 변화하지 않는다. 이러한 특성으로 인해 영상을 다루는 대부분의 디

지털기기는 영상 부호화를 하여 사용한다. 즉, 원영상의 화질을 충실히 보존하면서 그에 필요한 데이터양을 최대한 줄이는 것이다. 이러한 영상 부호화의 방법 중에 널리 사용되는 것이 벡터 양자화이다[1]. 벡터 양자화의 성능은 코드북을 얼마나 효율적으로 설계하느냐에 따라 결정된다. Linde, Buzo, Gray[2]는 양자화 오차가 최소가 될 때까지 알고리즘을 반복하면서 양자화 칼라 벡터를 개선시켜 주는데, 이 때문에 양자화 오차는 최소이지만 반복과정에 있어서 모든 입력 칼라 벡터를 탐색하여 양자화 칼라를 개선시키므로 양자화에 요구되는 시간이 매우 길다. 그 후 많은 연구자들에 의해 변형된 방법이 소개되었지만, 이 방법들은 한번에 하나의 벡터가 작용하므로 오랜 학습 시간이 걸리며, 영상의 특징에 기인하는 코드북의 크기를 미리 결정해

<sup>†</sup> 준 회원 : 충북대학교 대학원 정보통신공학과

<sup>††</sup> 정 회원 : 충북대학교 전기전자공학부 교수

논문접수 : 2003년 4월 10일, 심사완료 : 2003년 12월 11일

야하는 단점을 갖고 있다. 이를 개선하기 위해서 C.Chan과 C.Ma[3-4]은 영상을 하나의 군집으로 시작하여 최적의 분할경계에 의해서 군집을 2개로 계속해서 분할해 나가는 것을 반복하는 최대 하강법(maximum decent method)을 제안하였다. 그러나 이 방법 역시 최적의 분할 경계를 설정하는 과정에서 많은 시간이 요구되는 단점이 있다. 이러한 단점을 보완 하고자 Orchard와 Bouman[5]은 고유벡터를 이용한 이진 트리 기법에 의하여 벡터의 차원을 줄여서 계산하는 방법을 제안하였다. 이 양자화 방법은 속도가 빠른 장점이 있지만, 낮은 레벨로 양자화를 하였을 때 색상의 열화가 발생하는 단점이 있다. 또한, 강대성 등[6]은 주성분 해석 기법에 기반한 양자화를 제안하여 각 분할단계에서 오차를 최소화 하는 방법을 제안했다. 그러나 이 방법은 원영상과 복원영상의 오차를 개선하였으나 시간이 많이 걸리고 인간의 시각 특성을 고려하지 않은 단점이 있다. 김경만 등[7]은 인간 시각 특성을 고려한 양자화 방법을 제안하였다. 이 방법은 칼라 히스토그램을 이용하여 양자화 하는 방법으로 [2]의 방법에 비해 속도는 개선되었고 높은 레벨로 양자화 할 경우 [2]와 비슷한 결과 영상을 얻을 수 있지만 낮은 레벨로 양자화 하였을 때 양자화 오차가 현저하게 커지는 단점을 가지고 있다.

본 논문에서는 고유벡터를 이용한 이진 트리 벡터양자화의 분할 과정에 인간의 공간 지각 특성을 이용한 가중치를 부여함으로써 화질을 개선한다. 가중치는 영상의 명암이 심하게 변하는 곳이나 그 부분에서 발생하는 양자화 오차는 인간의 시각에 잘 띄지 않는다는 성질을 이용하여 구한다. 먼저 영상을 일정한 크기의 블록으로 나누고, 각 블록의 밝기 성분을 이용하여 각각의 블록을 인간이 지각하는 세 시각대역으로 나누어 시각 대역별로 R, G, B 세 칼라성분별로 가중치를 다르게 준다. 또한 주관적 화질 평가 방법과 PSNR(Peak to Peak Signal to Noise Ratio)을 이용하여 영상의 화질을 평가하였다.

본 논문의 구성은 다음과 같다. 2장에서는 고유벡터를 이용한 이진 벡터 양자화의 기법을 소개하고, 3장에서는 인간의 공간 지각 특성을 고려한 고유벡터를 이용한 벡터 양자화 방법을 제안한다. 4장에서는 시뮬레이션을 통하여 기존 방법과 제안한 방법을 비교하여 성능을 분석하였고, 5장에서 결론을 맺었다.

## 2. 고유벡터를 이용한 이진 트리 벡터 양자화

고유벡터를 이용한 이진 트리 벡터 양자화는 고유벡터를 이용해 영상을 이진 트리 구조의 노드로 분할하는 방법이다. 트리의 각 n번째 노드는 전체 영상의 분할된 클러스터

$C_n$ 으로 표기하며, 부모 노드는 고유벡터를 이용해 두 개의 자식 노드인  $C_{2n}$ 과  $C_{2n+1}$ 로 분할하게 되며 각 노드마다 각 노드에 속하는 모든 화소들을 대신할 하나의 대표값  $q_n$ 을 갖는다. 하나의 노드를 구성하는 입력 벡터들은  $x_s = [x_s^0, x_s^1, x_s^2]^t$ 로 표기하며, 여기서  $s = (i, j)$ 로 화소 위치를 나타내고, 0, 1, 2는 칼라성분(R, G, B)의 인덱스이다.  $C_n$ 은 n번째 노드의 화소 집합이고, 각 클러스터에 대한 대표값  $q_n$ 은 다음의 식으로 계산한다.

$$m_n = \sum_{s \in C_n} x_s \tag{1}$$

$$N_n = |C_n| \tag{2}$$

$$q_n = \frac{m_n}{N_n} \tag{3}$$

식 (1)의  $m_n$ 은 각 칼라성분의 합으로 구성된 벡터를 나타내며, 식 (2)의  $N_n$ 은 노드에서  $C_n$ 의 화소의 개수를 나타낸다.

하나의 노드를 두개의 노드로 분할하는 것은 두개의 새로운 양자화 레벨을 결정하는 것이다. 이때 부모노드의 화소들을 최적의 양자화 레벨을 갖는 두개의 노드로 분할하는 최적의 평면을 결정하는 것이 중요하다. 고유벡터를 이용한 양자화 방법은 클러스터의 분산이 최대인 방향을 결정하여 이 방향에 수직이 되고 노드의 평균을 지나는 평면을 최적의 분할 평면으로 결정한다. 이 방법은 Gaussian 분포를 갖는 경우에 최적의 효율을 나타낸다. 노드의 공분산  $\widetilde{R}_n$ 은 식 (5)의  $R_n$ 을 이용하여 식 (1)과 식 (2)를 이용해 식 (5)와 같이 구한다.

$$\widetilde{R}_n = R_n - \frac{1}{N_n} m_n m_n^t \tag{4}$$

이때  $R_n$ 은 다음과 같다.

$$R_n = \sum_{s \in C_n} x_s x_s^t \tag{5}$$

공분산  $\widetilde{R}_n$ 의 고유값을 구하기 위해, 식 (6)과 같이 단위 벡터를 곱한다.

$$\sum_{s \in C_n} (x_s - q_n)^t e_n = e^t \widetilde{R}_n e \tag{6}$$

여기서  $e$ 는 식 (4)을 최대로 하는 단위 벡터이며, 이 식의 해는  $\widetilde{R}_n$ 의 가장 큰 고유값  $\lambda_n$ 의 고유벡터  $e_n$ 이며, 식 (6)은 고유값  $\lambda_n$ 을 이용해 다음과 같이 나타낼 수 있다.

$$\sum_{s \in C_n} ((x_s - q_n)^t e_n)^2 = \lambda_n \quad (7)$$

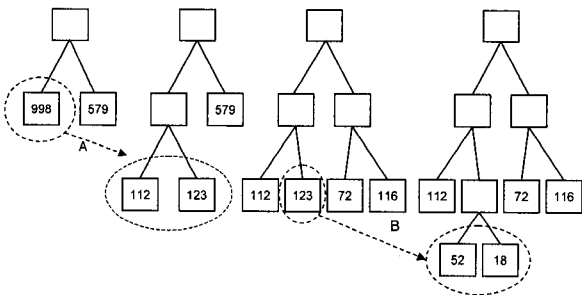
고유벡터  $e_n$ 이 결정되었으므로 노드  $C_n$ 의 화소들은 다음과 같은 방법으로 두 노드  $C_{2n}$ 과  $C_{2n+1}$ 으로 식 (8)과 식 (9)와 같이 분할된다.

$$C_{2n} = \{s \in C_n : e_n^t x_s < e_n^t q_n\} \quad (8)$$

$$C_{2n+1} = \{s \in C_n : e_n^t x_s > e_n^t q_n\} \quad (9)$$

즉, 한 노드의 분할은 공분산  $\widehat{R}_n$ 의 가장 큰 고유값  $\lambda_n$ 의 고유벡터  $e_n$ 을 결정하여 이를 기준으로 두 노드로 분할하는 것이다. 이는 양분 하였을 때 고유값이 가장 크게 되는 곳을 양분하는 것을 의미하며 이는 노드의 분산이 가장 작게 되는 방향으로 노드를 분할하는 것이다.

이진 트리에서 분할할 노드는 종단 노드중 노드의 화소들의 고유값을 구하여 가장 큰 고유값을 갖는 노드를 선택한다. (그림 1)은 이진 트리에서 노드의 분할 과정을 설명한 것이다. (그림 1)에서 A의 과정은 첫 번째 분할 후 두개의 자식 노드중 좌측 노드의 고유값(998)이 우측 노드의 값(579)보다 크므로 이를 다음 분할노드로 선택하며 이를 분할하게 된다.



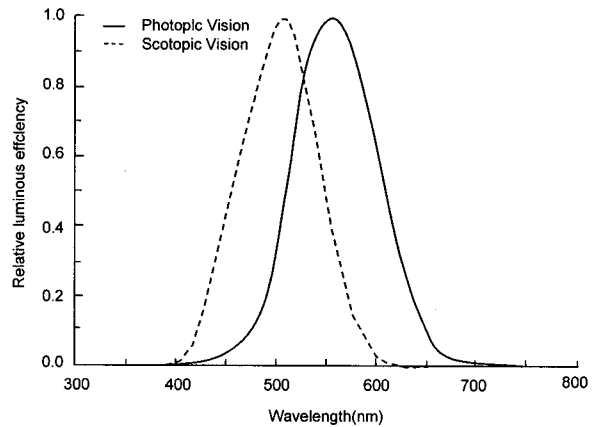
(그림 1) 이진 트리 구조의 노드 분할

분할된 결과는 종단 노드가 세 개로 늘어나며 이 종단 노드 중에서 고유값이 579인 노드가 가장 큰 값을 가지므로 이 노드를 선택하여 분할하고, B의 과정도 고유값이 가장 큰 값을 선택하여 분할하는 것을 보여준다. 이의 과정을 반복하여 영상의 양자화 레벨 수와 이진트리의 종단 노드의 수가 일치 할 때까지 노드를 분할한다.

### 3. 인간의 공간 지각 특성을 고려한 이진 트리 벡터양자화

인간 시각의 인식율은 상대적인 빛의 밝기와 파장에 따라서 달라지는데 파장이 다른 여러 단색광이 모두 같은 강

도의 에너지를 갖고 있다고 가정할 때, 각 단색광에서 느끼는 밝음의 감각을 시감도라고하며, 이 시감도의 차를 상대적으로 나타낸 것을 비시감도라 한다. CIE에서는 이 비시감도를 밝기에 따라 명순응시각(photopic vision)과 암순응시각(scotopic vision)으로 분류한다[8]. (그림 2)는 광학 복사에 대한 인간 시각 시스템의 주파수 민감도를 나타낸다. (그림 2)의 오른쪽 곡선은 상대적으로 높은 밝기에서의 명순응 시각을 나타내고, 왼쪽 곡선은 상대적으로 낮은 밝기에서의 암순응 시각을 나타낸다.



(그림 2) 인간 시각 시스템의 명순응 시각과 암순응 시각

제안 알고리즘은 영상의 명암이 심하게 변하는 곳이나 그 부근에서 발생하는 양자화 오차는 인간의 시각에 잘 띄지 않는 성질을 이용하여 고유벡터를 이용한 이진트리 벡터양자화의 분할 과정에 가중치를 적용한다. 이때 인간의 시각이 밝기에 따라 반응정도가 다를 것을 고려하여 영상의 각 블록에서의 밝기 성분을 상대적으로 높은 부분인 명순응 시각 대역과 상대적으로 낮은 부분인 암순응 시각대역, 그리고 그 사이의 중간 대역으로 분류하여 각 대역에 대해서 가중치를 구하도록 하였다.

먼저 영상을  $4 \times 4$ 블록으로 나누어주고 다음의 식과 같이 인간 시각의 칼라에 대한 활성 척도  $w_s^k$ 를 구한다.

$$w_s^k = \frac{1}{\frac{1}{16} \sum_{s=1}^{16} a^k (x_s^k - b^k)^2} \quad (10)$$

여기서 블록의 평균 칼라 벡터를  $b^k$ 라 하고 블록의 입력 벡터를  $x_s^k$ 라 하고  $k$ 는 각 칼라성분(R, G, B)의 인덱스 값 0, 1, 2를 나타낸다.  $a^k$ 는 각 시각 대역에서 적색, 녹색 및 청색에 대한 가중치로, (그림 2)에서 제시된 적색, 녹색 및 청색에 해당하는 상대적인 민감도가 높은 영역의 민감도를 합하여 1이 되도록 정규화 하여 구하며, 이는 다음과 같다.

$$\begin{aligned} \text{명순응 시각대역} : \alpha^0 &= 0.322, \alpha^1 = 0.628, \alpha^2 = 0.05 \\ \text{중간 시각대역} : \alpha^0 &= 0.299, \alpha^1 = 0.587, \alpha^2 = 0.114 \\ \text{암순응 시각대역} : \alpha^0 &= 0.048, \alpha^1 = 0.547, \alpha^2 = 0.405 \end{aligned} \quad (11)$$

영상의 각 블록에서의 대역을 나누는 방법은 전체 영상의 밝기 성분에 대한 평균값  $I_m$ 과 표준편차  $\sigma$ 로 구분한다. 블록에서 영상의 평균밝기가  $I_m - \sigma$ 보다 작은 영역은 암순응 시각대역으로,  $I_m + \sigma$ 보다 큰 영역은 명순응 시각대역으로 구분하였고 그 외의 부분은 중간 시각 대역으로 설정하여 가중치를 적용하였다.

제안방법은 블록내에서 동일하게 적용되는 식 (10)의 가중치  $w_s$ 를 각 화소별로 확장하여 가중치  $w_s$ 를 구하고, 이를 노드의 공분산  $\widehat{R}_n$ 를 구할 때 적용하여 시각 효과를 가중치로 적용한 고유벡터를 얻는다.

그리고 이 고유벡터를 이용하여 노드를 분할한다.  $w_s$ 를 노드의 공분산을 구하는데 적용하기 위해 가중치를 식 (12)와 같이 바꾼다.

$$W_s = \begin{bmatrix} w_s^0 & 0 & 0 \\ 0 & w_s^1 & 0 \\ 0 & 0 & w_s^2 \end{bmatrix} \quad (12)$$

그리고 이를 공분산을 구하는데 필요한 통계량에 결합하면 다음과 같다.

$$R_n = \sum_{s \in C_n} (W_s x_s)(W_s x_s)^t \quad (13)$$

$$m_n = \sum_{s \in C_n} W_s x_s \quad (14)$$

$$N_n = \sum_{s \in C_n} W_s U \quad (15)$$

여기서  $U = [1 \ 1 \ 1]^t$ 이며, 위 식들로 식 (4)에 적용하여 가중치가 적용된 공분산  $\widehat{R}_n$ 와 노드의 대표값인  $q_n$ 을 구한다.  $q_n$ 은  $m_n$ 과  $N_n$ 의 각각의 열성분을 이용하여  $q_n = [m_n^0/N_n^0 \ m_n^1/N_n^1 \ m_n^2/N_n^2]^t$ 로 구한다.

#### 4. 실험 및 결과

본 논문에서는 제안된 방법의 평가를 위해서 다양한 화소 분포 특성을 가지는  $256 \times 256$  크기의 1600만 칼라 정지 영상을 대상으로 Microsoft Visual C++ 6.0을 이용해 시뮬레이션을 하였다.

(그림 2)(a)는 원영상을 나타내고 (그림 2)(b)는 양자화 레벨수  $L=32$ 인 이진 트리 벡터양자화 영상을 나타내며 (그림 2)(c)는 동일한 양자화 레벨을 갖는 제안방법의 결과 영상이다. (그림 2)(c)에서 볼 수 있는 바와 같이 전체 영상의 색상이 강조되어 보다 선명하게 보이며, 모자의 깃털 부분이 선명해졌음을 알 수 있다. 또한 어깨부분과 배경 부분에서도 원영상에 유사한 색상을 보임을 알 수 있다. 즉, 양자화에서 발생하는 열화현상을 인간의 공간 시각 특성으로 강조하게 되어 선명한 색상의 영상을 얻을 수 있었으며, 유사색상 영역분할에서 더 좋은 성능을 보여주었다.

제안 방법의 성능을 평가 하기 위해 본 논문에서는 PSNR과 주관적 화질 평가 방법을 사용했다. PSNR을 구하는 식은 다음과 같다.

$$RMSE = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N [\hat{f}(i, j) - f(i, j)]^2} \quad (16)$$

$$PSNR = 20 \log_{10} \left( \frac{255}{SE} \right) \quad (17)$$



(a) 원영상

(b) BTVQ

(c) 제안알고리즘

(그림 2) 원영상과  $L=32$ 인 양자화 영상

<표 1>은 양자화 레벨  $L=16, 32, 64$ 일때 BTVQ 방법과 제안방법의 PSNR 결과를 보인다. 제안방법은 BTVQ 보다 1.6-3.2dB정도가 개선되었다.

<표 1> 양자화레벨에 따른 PSNR 비교

(단위 : dB)

레벨	알고리즘	영 상			
		girl	lena	tiffany	pepper
16	BTVQ	29.24	31.44	31.81	29.36
	제안방법	31.34	33.80	33.74	30.98
32	BTVQ	32.05	33.95	34.10	31.74
	제안방법	33.93	36.77	35.97	33.32
64	BTVQ	34.13	35.95	36.25	33.74
	제안방법	36.20	38.97	38.49	35.43

주관적 화질 평가는 비전문가 6인을 대상으로 기존의 방법(A)을 기준으로 하여 제안방법(B)의 화질 개선 정도를 7 단계로 나누어 화질을 비교 평가하였다[9, 10]. 주관적 화질 비교 평가 결과가 <표 2>에 있으며 기존의 방법 보다 화질이 개선되었음을 보였다.

<표 2> 양자화 레벨수에 따른 주관적 화질 비교 평가 결과 (B/A)

영상 \ 레벨	16	32	64
girl	5.0	4.8	4.5
lena	4.7	5.0	4.8
tiffany	5.7	4.3	5.2
pepper	4.2	4.2	4.8

### 5. 결 론

제안 방법은 공간 마스킹 효과를 적용하여 가중치를 구해 결합한 고유벡터를 이용한 이진트리 벡터양자화를 제안한다. 제안 알고리즘은 영상의 명암이 심하게 변하는 곳이나 그 부근에서 발생하는 양자화 오차는 인간의 시각에 잘 띄지 않는 성질을 이용하여 고유벡터를 이용한 이진트리 벡터양자화의 분할 과정에 가중치를 적용한다.

제안방법을 영상에 적용하였을 때 색상이 더 선명한 양자화 영상을 얻을 수 있었으며 유사색상의 영역분할에서 더 좋은 성능을 보여주었다. 또한 주관적 화질평가의 결과 기존의 방법에 비해 영상을 화질이 개선되었음을 알 수 있었으며 인간의 시각 시스템이 고려된 PSNR 성능에서도 제안한 방법이 기존 방법보다 1.6~3.2dB 정도 더 좋은 성능

을 보여 주었다.

### 참 고 문 헌

- [1] N. M. Nasrbadi and R. A. King, "Image coding using vector quantization : A Review," *IEEE Transaction on Communication*, Vol.36, No.8, pp.604-619, 1988.
- [2] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transaction on Communication*, Vol.28, No.1, pp.84-95, 1980.
- [3] C. K. Ma and C. K. Chan, "Maximum descent method for image vector quantization," *Electronic Letters*, Vol.27, pp.1772-1773, 1991.
- [4] C. K. Chan and C. K. Ma, "A fast method of designing better codebooks for image vector quantization," *IEEE Transactions on Communication*, Vol.42, No.2/3/4, pp.237-242, 1994.
- [5] M. T. Orchard and C.A. Bouman, "Color Quantization of Images," *IEEE Transaction on signal processing*, Vol.39, No.12, pp.2677-2690, 1991.
- [6] 강대성, 서석배, 김대진, "주축의 연속적 분할을 통한 고속 벡터 양자화 코드북 설계", 정보과학회논문지 : 소프트웨어 및 응용, 제27권 제4호, pp.422-431, 2000.
- [7] 김경만, 이양우, 이채수, 하영호, "인간 시각의 칼라 활성 가중 왜곡 척도를 이용한 칼라 영상 양자화", 전자공학회논문지, 제33권 제4호, pp.101-109, 1996
- [8] M. D. Fairchild, *Color Appearance Models*, Addison-Wesley LongmanInc., 1998.
- [9] CCIR, "Methods of the subjective assessment of the quality of televisionpictures," 13th Plenary Assembly, Rec. 500, Vol.11, pp.6-68, 1974.
- [10] Recommendation ITU-R BT. 500-6, "Methodology for the SubjectiveAssessment of the Quality of Television Pictures," ITU-R 1974~1994.



유 성 필

e-mail : cucus@chol.com

2000년 충북대학교 정보통신공학과(학사)

2003년 충북대학교 대학원 정보통신공학과 (공학석사)

2003년~현재 충북대학교 대학원 정보통신공학과(박사과정)

관심분야 : 영상통신, 영상정보처리



### 곽 내 정

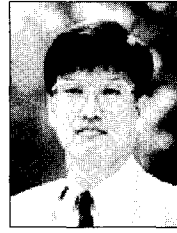
e-mail : knj0125@hanmail.net

1993년 충북대학교 정보통신공학과(학사)

1995년 충북대학교 대학원 정보통신공학과  
(공학석사)

1997년~현재 충북대학교 대학원 정보통신  
공학과(박사 과정)

관심분야 : 영상통신, 영상정보처리



### 안 재 형

e-mail : jhahn@cbucc.chungbuk.ac.kr

1981년 충북대학교 전기공학과(학사)

1983년 한국과학기술원 전기 및 전자  
공학과(석사)

1992년 한국과학기술원 전기 및 전자  
공학과(박사)

1987년~현재 충북대학교 전기전자공학부교수

관심분야 : 영상 통신 및 영상정보처리, 멀티미디어 제작 및 정  
보제공, 인터넷 통신 및 프로그래밍