

신경망에 의한 테두리를 보존하는 영상압축

정회원 박 동 철*, 우 영 준**

Edge Preserving Image Compression with Weighted Centroid Neural Network

Dong-Chul Park*, Young-June Woo** *Regular Members*

요 약

무지도 경쟁학습을 이용하여 압축된 영상의 복원 후에 나타나는 테두리부분의 손상을 최소화하기 위한 영상압축 방법이 제안되었다. 제안된 영상압축방법은 영상데이터에서 테두리부분에 해당하는 데이터의 기하학적인 특징을 이용하는데, 영상데이터의 통계학적인 특성을 함께 이용하여 기존의 Centroid Neural Network을 일반화시키는 무지도 경쟁학습에 의하여 자동적으로 더욱 많은 code vector를 테두리부분에 배정함으로써 압축된 영상의 복원 후에 나타나는 테두리부분의 손상을 최소화하게 한다. 실험 결과, 기존의 SOM, M-SOM, M/R-CNN등과 비교하여 제안된 방법에 의해 압축된 영상의 복원된 테두리 부분에서 PSNR이 약 2db정도 향상된 결과를 보여줄 수 있었다.

ABSTRACT

A new image compression method to preserve edge characteristics in reconstructed images using an unsupervised learning neural network is proposed in this paper. By the unsupervised competitive learning which generalizes previously proposed Centroid Neural Network(CNN) algorithm with the geometric characteristics of edge area and statistical characteristics of image data, more codevectors are allocated in the edge areas to provide the more accurate edges in reconstructed image. Experimental results show that the proposed method gives improved edge characteristics in reconstructed images when compared with SOM, Modified SOM and M/R-CNN.

1. 서 론

벡터양자화(Vector Quantization, VQ)는 변환부호화같은 방법에서는 얻기 어려운 높은 압축률을 상대적으로 적은 블록크기에서 얻을 수 있기 때문에 높은 압축률을 요구하는 응용에서는 많이 사용되는 부호화방법이다^[1]. 테두리는 영상에서 작은 부분을 차지하기 때문에, 평균자승오차 (Mean Square Error: MSE)같은 일반적인 왜곡산정법을 이용한 이전의 영상압축 알고리즘에서는 코드북의 일부분만이 테두리영역의 데이터를 위해 할당되며, 그 결과로 테두리 부분의 손상이 압축된 영상의 복구 시에 중

중 목격되었다^[2-4]. 그러나, 테두리영역은 영상정보의 매우 중요한 정보를 가지고 있기 때문에, 테두리영역에서의 영상의 손상은 테두리가 아닌 영역에서의 영상의 손상보다 시각적으로 나쁜 영향을 미친다. 이러한 의미에서, MSE는 테두리부분의 부호화를 고려할 때 좋은 선택이 아닐 수도 있다.

복구된 영상에서 테두리를 충실하게 보존하는 문제에 대해 몇 가지의 접근이 이루어져 왔는데^[3-6], Gersho와 Ramamurthi는 영상을 테두리부류와 테두리가 아닌 부류처럼 몇가지의 부류로 나누고, 각각의 부류에 따로따로 벡터양자화를 적용함으로써 이 문제를 해결하려 하였다^[2]. 그들은 나중에 학습데이터를 테두리와 테두리가 아닌 부류로 분류하여 각각

* 명지대학교 정보제어공학과-정보통신교육센터(parkd@wh.myongji.ac.kr), ** 명지대학교 정보제어공학과(june@icrl.myongji.ac.kr)

논문번호 : 99091-0308, 접수일자 : 1999년 3월 8일

※ 본 연구는 정보통신부지원 정보통신 우수 시범학교 지원 사업에 의해 수행되었음

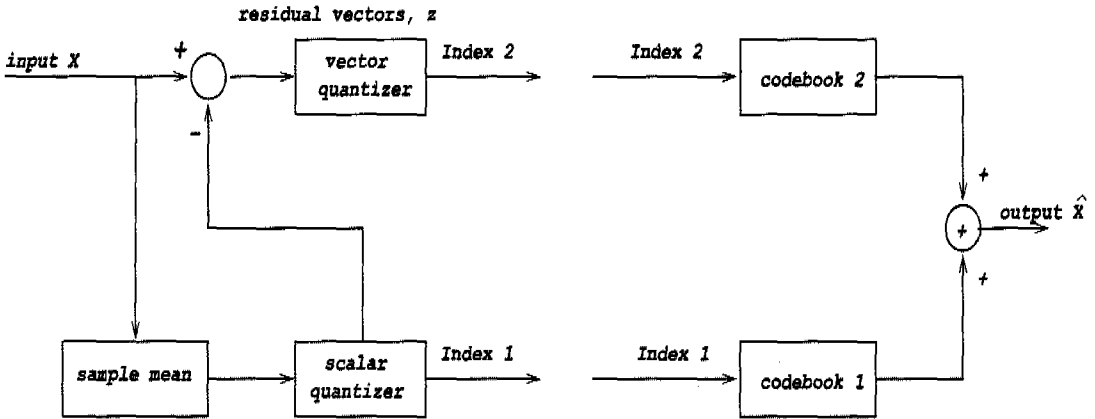


그림 1. 평균/나머지 벡터양자화(M/RVQ) 시스템

의 부류에 대하여 하위 코드북이라 불리는 각각의 코드북을 구성하는 이러한 접근법을 분류벡터양자화 (Classified VQ:CVQ)의 개념으로 확장시켰다^[3].

Baker와 Gray는 각각의 블록에서 평균을 낸 후, 평균과 나머지에 따라 따로 벡터양자화를 적용함으로써 이 문제를 풀려는 시도를 하였다^[4]. 이렇게 벡터의 평균과 각각의 블록에 대하여 나머지벡터를 고려하는 생각은 M/RVQ를 만들었다^[5].

테두리영역에서의 영상정보의 손상을 줄이는 방법에 대한 다른 접근은 수정된 형태의 SOM(Self-Organization Map)이 있다^[6]. 이 방법에서 각각의 블록에 대하여 가중치 인자로서 사용되는 블록의 변화량을 측정하는 데, 이렇게 측정된 가중치 인자는 Kohonen의 SOM^[7]에서 각 데이터 블록의 학습이득으로 사용된다. 이 방법은 벡터양자화에서 테두리영역에서의 영상정보의 손상을 줄이는 매우 간단한 방법이지만 테두리영역에서의 영상정보의 손상을 줄이는 것은 SOM의 학습이득을 간단히 조정하는 것만으로는 부족하다.

국소적으로 최적화된 벡터양자기를 구성하는 것에 대해서는 여러 가지의 알고리즘이 제안되어 왔다. 1957년에 Bell연구소의 논문에서 Lloyd는 1차원 데이터에 대하여 최적의 양자화를 구성하는 두 가지의 방법을 제안하였고^[8], Linde, Buzo, Gray는 Lloyd의 방법을 k 차원의 블록양자화로 일반화시켰는데^[1], 이러한 방법이 유명한 LBG알고리즘으로, LBG알고리즘을 이용함으로써 최소한 국소적으로 최적화된 양자화를 구성할 수 있다.

또한, 양자화에 있어서, 왜곡률의 기대치의 상위

한계를 정하기 위한 다른 종류의 연구들이 진행되어 왔는데, 1948년에 Bennett은 companding 모델을 이용하여 1차원 데이터의 경우에 최적의 왜곡률을 계산하였고^[9], Gersho는 k 차의 왜곡산정법을 이용하여 Bennett의 공식을 k 차의 벡터양자화로 일반화시켰다^[10]. 또한 Gersho는 “최적의 양자기에서는 partition의 각각의 영역들은 왜곡률에 동등한 영향을 미친다”라는 유명한 근사적 결과를 유도해냈고^[11], Gray는 Gersho의 결과를 더욱 일반적인 왜곡률 산정법으로 일반화시켰다^[11].

그후, Gersho와 Gray의 결과를 이용하여, Ramamurthi와 Gersho는, P_i 는 벡터가 i 번째 부류에 포함될 확률이고, N_i 는 i 번째 부류에서의 코드 벡터의 최적화된 개수 일 때,

$$\frac{P_i D_i^*}{N_i} = \text{상수} \quad (1)$$

라는 유용한 공식을 유도해냈는데^[3], 모든 부류가 충분히 작고, 오직 하나의 코드벡터만을 가진다 ($N_i=1$)라고 가정할 때, 식 (1)은 Gersho의 결과와 같은 의미를 가지고, 아래와 같이 쓸 수 있는데,

$$D_i^* \propto \frac{1}{P_i} \quad (2)$$

식 (2)가 의미하는 것은 각각의 부류의 최적화된 오차는 각각의 입력벡터가 그 부류에 들어갈 확률의 역에 비례한다는 것이다.

위에서 설명한 두 가지의 접근법에서 우리는 중

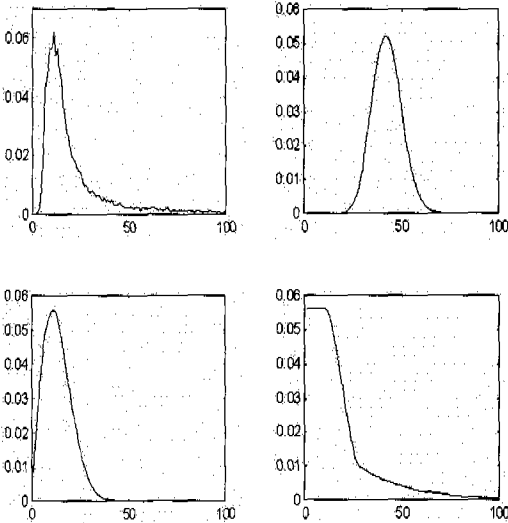


그림 2. (a) 실제 영상 데이터의 r 에 대한 확률 밀도 그래프 ($\sigma = 10.85$). (b) $k=16$ 인 경우의 식(4)의 그래프 (c) $k=2$ 인 경우의 식(4)의 그래프 (d) 식 (5)의 그래프

요한 사실을 알 수 있는데, 그것은 두 가지의 접근법 모두가 왜곡물의 기대값을 최소화시키는 데 초점을 두고 있다는 것이다. 그러나 이러한 최소화는 높은 확률의 데이터 벡터로의 코드벡터의 집중을 야기할 수 있으며, 낮은 확률의 데이터 벡터를 표시하는 코드벡터의 부족은 결과적으로 테두리영역에서의 손상을 일으킨다.

II. 본 문

1. 영상정보의 기하학적 구조

본 논문에서 제안하는 영상압축을 위한 WCNN (Weighted CNN) 알고리즘은 그림1의 M/RVQ (Mean/Residual VQ) 구조를 기본형으로 사용한다^[5]. M/RVQ에서, 영상은 블록으로 나누어지고 블록의 평균값을 먼저 구한 후, 입력벡터에서 스칼라 양자화된 블록의 평균값을 빼면 나머지벡터가 벡터양자기로 보내어지는 데, 이에 의해 평균을 양자화 하는 경우의 오차를 나머지벡터에 포함시킬 수 있다^[4].

벡터, z ,와 평균, b ,는 아래와 같이 정의되는 데,

$$b = \frac{1}{k} \sum_{i=1}^k \sum_{j=1}^p x_{i,j}, \quad z_{i,j} = x_{i,j} - b$$

이때, $x_{i,j}$ 는 $p \times p (k = p^2)$ 인 블록 x 의 (i,j) 번째 인자를 나타내고, $z_{i,j}$ 는 $p \times p$ 인 블록 z 의 (i,j) 번째 인자를 나타내며, b 는 b 의 스칼라양자화에 의한 대표값이다.

z 의 각각의 인자가 같은 값의 표준편차, σ ,와 평균값 0을 가지는 다음과 같은 정규분포를 따르는 독립적인 확률변수의 연속이라고 할 때,

$$P_{z_d}(z_d) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-z_d^2/2\sigma^2}$$

z 의 확률밀도함수는 아래와 같이 주어진다^[12]:

$$f_z(z) = f_z(z_1, z_2, \dots, z_k) = \prod_{d=1}^k P_{z_d}(z_d) \quad (3)$$

이제 원점과 나머지벡터사이의 거리에 관한 거리 변수 r ,를 아래와 같이 정의할 수 있는 데,

$$r_i \equiv \|z\| = \left(\sum_{d=1}^k z_d^2 \right)^{1/2}$$

여기서 r 이 크면 클수록 그 블록데이터는 더욱 테두리의 특성을 나타낸다고 할 수 있다. 한편, r 의 확률밀도함수는 아래와 같이 구해질 수 있는 데,

$$f_R(r) = \frac{r^{k-1}}{2^{k/2} \Gamma(k/2) \sigma^k} e^{-r^2/2\sigma^2} \quad (4)$$

이것은 k 차의 자유도를 가지는 chi분포이다^[12].

하지만 실제영상에서 입력벡터 x 의 각 차원은 통계적으로 독립적이지 않고 높은 상관성을 가지고 있기 때문에, 그림2-(a)에서 보여지는 실제영상에 대한 r 의 확률분포 그래프는 그림2-(b)에서 보여지는 식 (4)의 그래프와 다르다. 사실 실제 영상의 확률 밀도함수는 각각의 인자들 사이의 높은 상관성 때문에 2차나 3차의 자유도를 가지는 chi분포와 비슷해 보이는 데, 이를 토대로 확률밀도함수는 다음과 같이 수정되고, 수정된 식의 그래프는 그림2-(d)와 같다.

$$P_r(r) = \begin{cases} \frac{\sigma_r}{\sigma_z^2} e^{-r^2/2\sigma_z^2}, & 0 \leq r < \sigma_r \\ \frac{r}{\sigma_z^k} e^{-r^2/2\sigma_z^2}, & \sigma_r \leq r < 2.5\sigma_r \\ 0.0299e^{-0.04r}, & 2.5\sigma_r \leq r < 10\sigma_r \\ 0.0299e^{-0.4\sigma_r}, & 10\sigma_r \leq r \end{cases} \quad (5)$$

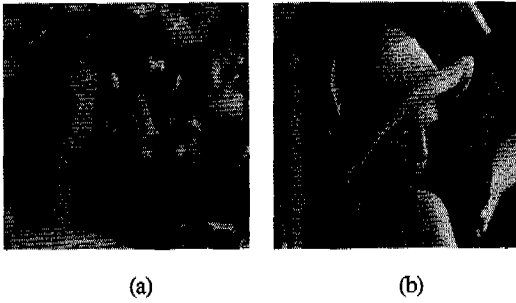


그림 3. (a) 학습영상, Peppers (b) 실험 영상, Lena

이러한 근사화의 기본이 되는 생각은 1) 그림2-(c)에서 보듯이 $r \geq 2.5\sigma$ 인 영역에서의 너무 작은(거의 0에 가까운) 확률값은 앞으로 사용할 확률값의 역수에서 연산에 관계된 문제를 일으킬 수 있고, 2) 그림2-(c)에서 보듯이 $r < \sigma$ 인 영역은 테두리가 아닌 영역으로 간주되는 데도 낮은 확률값을 가진다는 점이다. 따라서 높은 확률값을 이 영역에 할당함으로써, 결과적으로 앞으로 쓰이게 될 $1/P_r(r)$ 이 값이 작아지게 되어 우리가 원하는 것인 적은 수의 code vector가 이곳에 할당되게 한다.

2. Weighted CNN(WCNN)

실제적인 영상에서 약 70% 정도의 블록은 테두리가 아닌 블록이고 30% 정도의 블록만이 테두리 블록이기 때문에^[3] MSE같은 일반적인 왜곡률 산정법을 적용할 경우 대다수의 코드 벡터들은 테두리가 아닌 영역에 존재하고 약간의 경우만이 테두리 영역에 존재하는 데, WCNN 알고리즘은 이러한 문제를 해결하기 위해서 역확률을 이용하여, 이미 존재하는 Centroid Neural Network (CNN)^{[13]-[15]} 알고리즘의 최적화에 관한 조건을 수정하였다.

CNN알고리즘은 “어떤 부류의 데이터의 무게중심은 최소한의 에너지를 가지는 점이다.”라는 이론에 근거하여, 주어진 출력 뉴런을 위한 연결강도는 그 부류에 포함되는 벡터들의 거리를 최소화시키는 것을 선택한다. 즉,

$$w_j = \min_w \sum_{i=1}^N \|x_j(i) - w\|$$

$$= \frac{1}{N_j} \sum_{i=1}^N x_j(i)$$

이며, 이때, N_j 는 j 번째 군집에 포함되는 데이터들의 개수이다. CNN에 관한 자세한 설명은 [13]-[15]

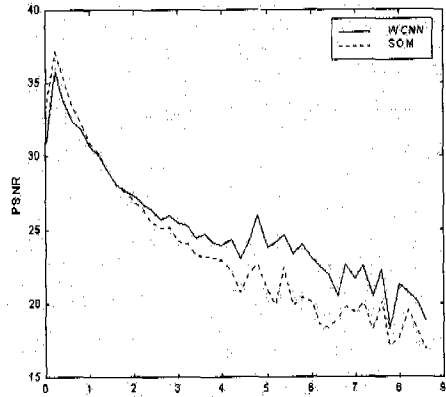


그림 4. SOM(점선)과 WCNN(실선)의 r에 대한 PSNR의 비교 (r축의 단위는 σ .)

에서 찾을 수 있다.

앞에서 언급하였던 테두리 영역에서의 영상정보의 손상을 막기 위하여 Gersho의 가정의 역인 “벡터양자화를 이용한 영상압축에서는 영상조각의 각 영역은 같은 오차를 가진다”에 근거하여, WCNN 알고리즘에서는 아래와 같이 i 번째의 데이터 블록의 역확률, $P_r^{-1}(r_i)$,을 CNN알고리즘의 연결강도의 갱신법칙(weight updating rule)에 가중치로 준다^[6].

$$w_j = \min_w \sum_{i=1}^N P_r^{-1}(r_i) \|x_j(i) - w\|$$

$$= \frac{\sum_{i=1}^N P_r^{-1}(r_i) x_j(i)}{\sum_{i=1}^N P_r^{-1}(r_i)} \tag{6}$$

주어진 데이터를 원하는 개수의 군집으로 나눌 때, WCNN은 각 군집의 각 데이터의 역확률로 주어지는 가중치평균을 그 연결강도로 구해낸다.

3. 실험결과

WCNN알고리즘을 이용한 영상압축방법의 성능을 확인하기 위한 실험에서 그림3의 Peppers영상을 학습에 사용하였고, Lena영상을 테스트에 사용하였다. 제안된 영상압축방법의 성능을 검사하기 위하여 SOM, M/R-CNN, WCNN, 수정된 SOM^[6]을 비교하였다.

M/RVQ구조는 두 가지 종류의 코드북을 가지기 때문에 아래와 같은 식이 부호기의 픽셀당 비트율 구량(bpp)을 계산하기 위하여 사용된다.

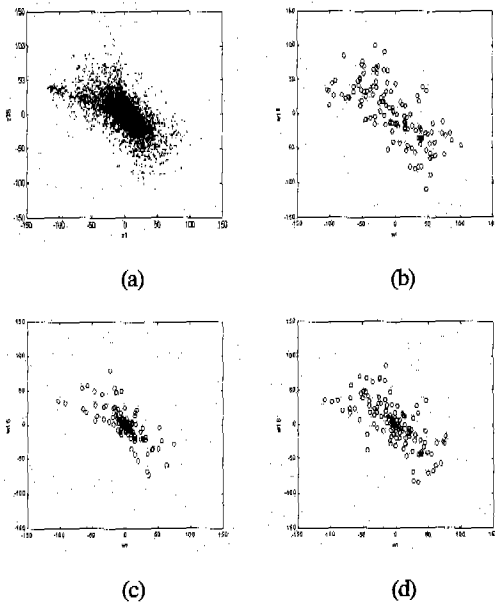


그림 5. (a) 학습영상, Peppers,의 2차원 출력. (b) WCNN의 연결강도의 학습 후 분포 (c) M/R-CNN의 연결강도의 학습 후 분포 (d) M/RVQ의 연결강도의 학습 후 분포

$$R = \frac{\log_2(M_1) + \log_2(M_2)}{k}$$

이때, M_1 은 스칼라 코드북의 크기이고, M_2 는 나머지 벡터의 코드북의 크기이며, k 는 블록의 크기이다. 본 실험에서는 $M_1 = 32, M_2 = 128, k = 16$ 으로 R 은 0.75bpp이다.

그림 4에서 우리는 PSNR을 구하기 위하여 블록들을 그들의 r 에 따라 선택하였다. 점선은 Kohonen의 SOM을 적용한 결과이고, 실선은 WCNN을 적용한 결과이다. 비록 WCNN을 이용하여 복구된 영상의 전체 오차가 SOM을 이용하여 복구된 영상의 오차와 비슷하고 블록의 분산이 1.5 σ 보다 작은 지역에서 WCNN을 적용한 경우의 오차가 SOM을 적용한 결과보다 크지만, 블록의 분산이 1.5 σ 보다 큰 지역(테두리에 해당하는 지역)에서의 오차는 WCNN을 적용한 경우가 일반적인 SOM을 적용한 경우보다 작은 데, 이러한 현상은 WCNN을 이용하는 방법이 영상압축에서 발생할 수 있는 테두리영역의 손상과 테두리가 아닌 영역으로의 코드벡터의 집중을 막을 수 있다고 생각되어질

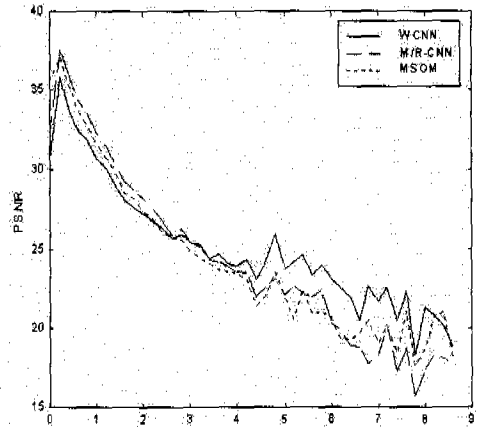


그림 6. M/R-CNN(점선), 수정된 SOM(큰점선), WCNN(실선)의 r 에 대한 PSNR비교(r 축의 단위는 σ .)

수 있다.

그림5는 각 알고리즘을 학습영상, Pepper, 에 적용시켰을 때 학습 후에 나타나는 각 알고리즘에 의한 신경망에서 첫 번째 입력과 16번째 입력(실험에서 4×4 블록의 데이터가 입력으로 쓰였음)과 연결된 연결강도의 2차원 평면상에서 보여주는데, SOM과 WCNN간의 구조적 차이 때문에 SOM대신 나머지 벡터의 벡터양자화에 SOM을 사용하는 M/RVQ가 사용되었다. 그림5-(d)에서 보듯이 M/RVQ의 연결강도들은 원점 가까이 존재하려는 경향이 있는데, 이 영역들은 발생확률이 매우 높고, 또한 테두리가 아닌 영역으로 생각되어 질 수 있는 영역이다. 그림5를 비교해보면 일반적인 벡터양자기들은 Gersho의 최적화조건을 따르고, 제안된 WCNN알고리즘을 이용하는 방법은 Gersho의 최적화 조건의 역을 따름을 알 수 있다. 구조적 차이점으로부터 오는 혼선을 피하기 위하여 WCNN과 M/R-CNN의 비교하는 실험이 여러차례 진행되었는데, M/R-CNN이란 WCNN과 같은 구조를 가지지만 최적화 조건을 수정하지 않은 것으로, M/RVQ알고리즘에서 나머지벡터의 벡터양자화로 CNN알고리즘을 사용한 것이다.

그림6은 최적화 조건을 수정하지 않은 M/R-CNN이나 수정된 SOM보다 WCNN을 이용하는 방법이 테두리영역이 손상되지 않은 결과를 갖는다는 것 보여준다. 그림7은 제안된 WCNN과 다른 알고리즘

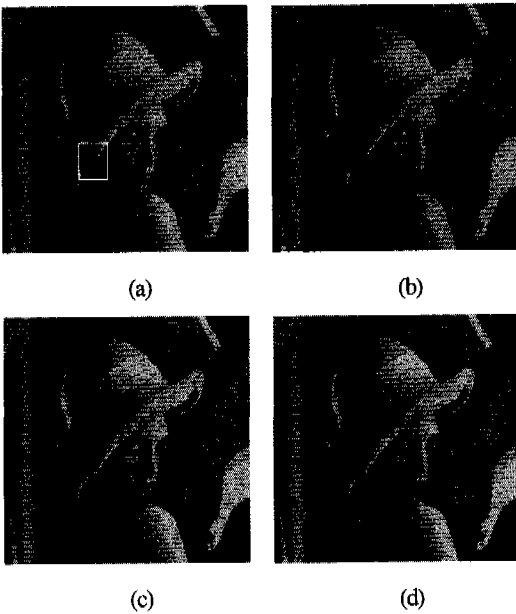


그림 7. 각각의 알고리즘을 적용하여 복구된 영상들 (a) SOM, (PSNR = 30.74) (b) WCNN (PSNR=30.95) (c) M/R-CNN (PSNR=31.64) (d) 수정된 SOM (PSNR=31.29)

들에 의해서 압축 후 복원된 영상들을 보여주며, 그림8은 압축 후 복원된 영상들의 확대된 영상을 보여주는 데, 확대된 부분은 그림7-(a)에서 네모표시를 한 부분이다. 그림7과 8의 각각의 영상들을 비교해 보면 실제적으로도 WCNN을 이용하여 압축·복구된 영상은 다른 알고리즘에 의해 압축·복구된 영상보다 테두리 영역에서의 손실이 심하지 않다는 것을 확인할 수 있다.

III. 결론

본 논문에서 우리는 먼저 일반적인 벡터양자화를 위한 최적화 조건을 설명하고, 그것으로부터 복구된 영상에 나타나는 시각적으로 매우 좋지 않은 영향을 미치는 테두리영역에서의 영상정보의 손실의 원인을 찾을 수 있었으며, 무지도 경쟁학습신경망을 이용하여 테두리영역에서의 영상정보의 손실을 줄일 수 있는 새로운 영상압축방법을 제안할 수 있었다. 제안된 영상압축방법에서 사용하는 WCNN알고리즘은 테두리강도에 대한 확률밀도함수를 먼저 구한 후, 그것을 코드벡터의 할당에 사용하는 것이다.

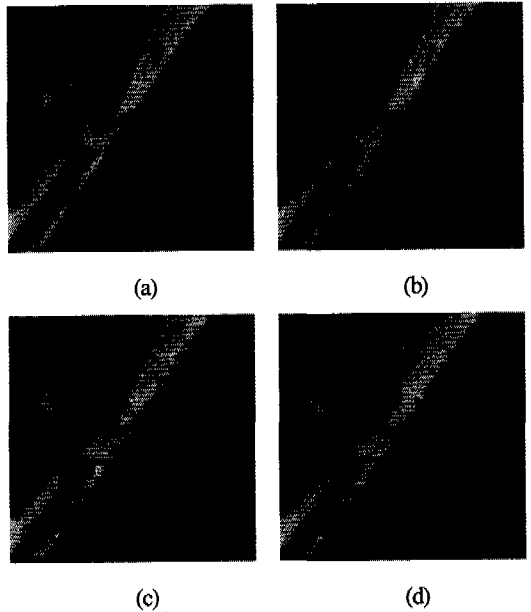


그림 8. 그림7-(a)에서 흰 사각형부분에 대한 결과 비교 (a) 원 영상 (b) WCNN (c) M/R-CNN (d) SOM

SOM, M/RCNN, 수정된 SOM등과 같은 신경망을 이용한 일반적인 벡터양자화와의 비교실험에서 WCNN에 의한 영상압축은 테두리영역에서의 영상정보를 더욱 잘 보존 결과를 얻을 수 있었다.

참고 문헌

- [1] Y. Linde, A. Buzo, and R. Gray, "An Algorithm for Vector Quantization Design," *IEEE Trans. Comm.*, Vol. 28, pp. 84-95, Jan. 1980.
- [2] A. Gersho and B. Ramamurthi, "Image coding using vector quantization," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 428-431, May 1982
- [3] Bhaskar Ramamurthi and Allen Gersho, "Classified Vector Quantization of Images," *IEEE Trans. Comm.*, Vol. COM-34, No. 11, pp. 1105-1115, Nov. 1986
- [4] R. L. Baker and R. M. Gray, "Image compression using non-adaptive spatial vector quantization," *Proc. 16th Asilomar Conf. circuits, Syst., Comput.*, pp. 55-61, Oct. 1982.

[5] B. L. Baker, *Vector quantization of digital images*, Ph.D dissertation, Stanford University, Stanford, CA, June 1984.

[6] Y. K. Kim and J. B. Ra, "Adaptive Learning Method in Self-Organizing Map for Edge Preserving Vector Quantization," *IEEE Trans. Neural Networks*, Vol. 6, No. 1, pp. 278-280, Jan. 1995

[7] T. Kohonen, *Self-Organization and Associative Memory*, 3rd ed. Berlin: Springer-Verlag, 1989.

[8] Lloyd S. P., "Least Squares Quantization in PCM's," *Bell Telephone Laboratories Paper*, Murray Hill, NJ, 1957

[9] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446-472, July 1948

[10] A. Gersho, "Asymptotically optimum block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 4, pp. 373-380, July 1979

[11] Yoshio Yamada, Saburo Tazaki and Robert M. Gray, "Asymptotic Performance of Block Quantizers with Difference Distortion Measures," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 1. pp. 6-14, Jan. 1980

[12] Papoulis, *Probability, Random Variables, and Stochastic Processes*, Tower Press, 1964

[13] 박동철, "새로운 무감독 경쟁 학습 알고리즘의 개발," *한국퍼지 및 지능 시스템학회 논문지*, Vol. 9, No. 2, pp. 196-204, 1999

[14] Dong C. Park and Young June Woo, "Centroid Neural Network Algorithm," *IEEE Trans. on Neural Networks*, (submitted for publication)

[15] Dong C. Park, "Development of a Neural Network Algorithm for Unsupervised Competitive Learning," *Proc. of 1997 Int. Conf. on Neural Networks*, Vol.3, pp. 1989-1993, 1997

[16] Dong-Chul Park and Yong-June Woo, "Weighted Centroid Neural Networks for Edge Preserving Image Compression," *Proc. of 1999 Int. Joint Conf. on Neural Networks*, Washington D.C., July 1999

박 동 철(Dong-Chul Park)

정회원



1980년 2월 : 서강대학교 전자공학
학과 (공학사)

1982년 2월 : 한국과학기술원
전기 및 전자공학과
(공학석사)

1990년 6월 : Ph.D in Electrical
Engineering University
of Washington at Seattle

1990년 8월 ~ 1994년 2월: Florida Int'l Univ. 조교수
(State Univ. of Florida at Miami)

1994년 3월 ~ 현재 : 명지대 정보제어공학과 부교수

1997년 ~ 현재 : IEEE Tr. on Neural Networks Associate Editor

1999년 ~ 현재 : IEEE Senior Member

<주관심 분야> 신경망, 음성인식, 신경망의 통신 시스템에의 응용.

우 영 준(Young-June Woo)

정회원



1997년 8월 : 명지대학교 제어계
측공학과 졸업

1998년 3월 ~ 현재 : 명지대학교
정보제어공학과
석사과정

<주관심 분야> 신경망, 신호처리, 영상처리