A Robust Transform Image Coder Over Noisy Channel

Chi-Hsi Su, Hsueh-Ming Hang*, and Che-Ho Wei

Dept. of Electronics Eng., National Chiao-Tung University
1001, University Rd., Hsin-Chu 30050, Taiwan, Republic of China
*Fax: (886)-3-5723283, E-mail: hmhang@cc.nctu.edu.tw

ABSTRACT

In this paper, we propose a robust quantizer design for image coding. Because the bits representing the reconstruction levels are transmitted directly to the channel, the proposed quantizer can be viewed as a compound of a quantizer, a VLC coder, and a channel coder. The conventional combined source/channel design produces a source coder designed for a channel with a specific channel noise. Our proposed quantizer is designed within a noise range. In comparison with the ordinary JPEG coder, simulation results show that our proposed scheme has a much more graceful distortion behavior within the designed noise range.

Keywords: combined source/channel coder, source coding, channel coding, VLC, noisy channel, and quantizer

1. INTRODUCTION

Shannon[1] showed that under the asymptotical assumption (very large data blocks), source coding and channel coding problems can be treated separately without sacrificing the overall optimality. This is the so-called Shannon's separation principle. According to this separation principle, the quantized coefficients in a transform image coder are first encoded by a VLC coder to reduce the redundancy and some additional bits are then inserted into the VLC-coded data by a channel coder to reduce channel noise effect, as shown in Figure 1. In Shannon's formulation, the quantizer is designed in the noise-free situation; that is, the chosen channel coder is able to correct the errors induced by the channel noise. In reality, the "true" noise-free environment is difficult to achieve. Particularly, in the case of a noisy channel with unknown characteristics, we can hardly correct all errors using only channel codes because the channel error rate can sometimes be very high. It is reported[2] that when the channel error probability is higher than the *error-correcting length/codeword length*, the channel code can induce a higher error rate than the original channel error rate. If this happens, the error-free condition is not valid anymore. On the other hand, if we use a very strong error control code to combat the worst possible case, we waste too many bits on the average.

When the separated coder is used in the non-error-free situation, there are two problems: quantizer mismatch and VLC error propagation. Because the quantizer in the separated coder is designed for the noise-free channel, the additional distortion due to channel noise may be significant.[3] Although the raw channel rate is low, the error propagation phenomenon in the VLC decoding process can magnify the output error rate. A single bit error may lead to the insertion of additional symbols or the erasion of the original symbols. We use Table 1 as an example of symbol insertion and erasure. Suppose the input sequence is $AABCB$, and thus the coded bit stream is 110001000. If the noisy channel causes errors at the 1st and 8th bits, the received bit stream is 010001010. The decoded symbol sequence becomes $CBAD$. In this case, a symbol erasure occurs when 5 input symbols is decoded into 4 symbols. If errors occur at the 1st, 2nd, and 5th bits, the received bit stream becomes 000011000. The decoded symbols sequence is thus $BBAAB$ with a redundant bit "0". Here, we retrieve 5 decoded symbols plus a left-over bit. This is a symbol insertion case. The symbol insertion and erasure not only leads to asynchronous decoding, but also produces burst errors. The performance of the separated coder, therefore, degrades rapidly in a noisy environment. Hence, a robust combined source and channel coding approach is able to offer a superior performance than the separated coding approach.

Part of the IS&T/SPIE Conference on Visual Communications
and Image Processing '99 • San Jose, California • January 1999
SPIE Vol. 3653 • 0277-786X/98/$10.00

Table 1. *Example of a variable wordlength code*

Symbol	Probability	Codeword
A	0.4	1
B	0.3	00
C	0.2	010
D	0.1	011

Table 2. *Threshold and Reconstruction levels of 4-level quantizers with 2-bit quantization index representation*

	Threshold levels			Reconstruction levels			
	x_1	x_2	x_3	y_1	y_2	y_3	y_4
Lloyd-Max (noise-free)	-0.9816	0.0000	0.9816	-1.5105	-0.4528	0.4528	1.5105
Noise rang=(0.0,0.01)	-0.9248	0.0000	0.9248	-1.4389	-0.4281	0.4281	1.4389
Noise range=(0.0,0.1)	-0.6326	0.0000	0.6326	-1.1263	-0.2724	0.2724	1.1263

2. OPTIMUM QUANTIZER DESIGN

In our proposed quantization process, a quantizer Q maps the input signal X into one of the M *reconstruction levels*, $y_1, y_2, \cdots, y_M$; that is, $X \in (x_{l-1}, x_l)$ is mapped to y_l and $\{x_l,\ l = 1, 2, \cdots, N\}$ are the *threshold levels.* Then, a binary codeword (index) is assigned to a reconstruction level. The resultant binary sequence is transmitted directly to a noisy channel without VLC coding and channel coding. At the decoder end, the received binary sequence is de-quantized into X' by an inverse quantizer Q^{-1}. Due to channel errors, the source index l may become index k at the receiving end. Let $P_{k/l}$ denote the channel transition probability,

$$P_{k/l} = P\{X' = y_k / Y = y_l\};\ k = 1, 2, \cdots, M,\ and\ l = 1, 2, \cdots, N. \tag{1}$$

Here, we focus only on the optimization of the quantizer parameters for a fixed index assignment. It has been observed that the reconstruction level indexing (binary representation) may affect the total distortion.[4] However, how to optimally select the index is a complicated issue and is not considered here.

For a specific binary channel with error rate ρ, the total mean square distortion is

$$D(\rho) = \sum_{k=1}^{M} \sum_{l=1}^{N} P_{k/l}(\rho) \int_{x_{l-1}}^{x_l} (x - y_k)^2 p_x(x) dx, \tag{2}$$

where

$$P_{k/l}(\rho) = (1 - \rho)^{n - d_H(k,l)} \rho^{d_H(k,l)}, \tag{3}$$

n is the codeword index length, and $d_H(k, l)$ denotes the Hamming distance between indices k and l. Suppose that the probability density function (*pdf*) of the channel error rates is $f_p(\rho)$ for $\rho \in (\rho_{min}, \rho_{max})$. The average distortion function D_t within the noise range (ρ_{min}, ρ_{max}) can be written as[3]

$$D_t = \int_{\rho_{min}}^{\rho_{max}} f_p(\rho) D(\rho) d\rho. \tag{4}$$

If the channel characteristics is known, its *pdf* can be used in Eq. (4). Otherwise, we assume the distortions induced by any two channel error rates within the noise range are equally important. Hence, the *pdf* function $f_p(p)$ used in designing a robust quantizer for an unknown channel is

$$f_p(p) = \begin{cases} \frac{1}{\rho_{max} - \rho_{min}} & for\ \rho_{min} \leq \rho \leq \rho_{max} \\ 0 & otherwise. \end{cases} \tag{5}$$

The optimum parameters x_l^* and y_k^* that minimizes the distortion function D_t can be obtained by taking the partial derivatives of (4) with respect to x_l and y_k, and set them to zero. We then obtain the optimal parameters x_l^* and y_k^* as follows[3]:

$$x_l^* = \frac{\int_{\rho_{min}}^{\rho_{max}} f_p(\rho) \sum_{k=1}^{M} {y^*}_k^2 (P_{k/l}(\rho) - P_{k/l+1}(\rho)) d\rho}{2\int_{\rho_{min}}^{\rho_{max}} f_p(\rho) \sum_{k=1}^{M} y_k^* (P_{k/l}(\rho) - P_{k/l+1}(\rho)) d\rho}, l = 1, 2, \cdots, N, \tag{6}$$

and

$$y_k^* = \frac{\int_{\rho_{min}}^{\rho_{max}} f_p(\rho) \sum_{l=1}^{N} P_{k/l}(\rho) \int_{x_{l-1}}^{x_l} x p_x(x) dx d\rho}{\int_{\rho_{min}}^{\rho_{max}} f_p(\rho) \sum_{l=1}^{N} P_{k/l}(\rho) \int_{x_{l-1}}^{x_l} p_x(x) dx d\rho}, k = 1, 2, \cdots, M. \tag{7}$$

For a fixed set of quantizer threshold levels $\{x_l\}$, it is easy to prove that updating $\{y_k\}$ according to equation (7) never increases the value of D_t. On the other hand, if we fix $\{y_k\}$ and update $\{x_l\}$ using Equation (6), it does not guarantee that D_t is not increased. To ensure that updating x does not increase the value of D_t, the second derivatives of D_t with respect to x_l must be non-negative. The convergence condition for searching for x_l parameters using the iterative method is derived in Farvardin and Vaishampayan,[3]

$$\sum_{k=1}^{N} \int_{\rho_{min}}^{\rho_{max}} f_p(\rho) P_{k/l+1} d\rho y_k \geq \sum_{k=1}^{N} \int_{\rho_{min}}^{\rho_{max}} f_p(\rho) P_{k/l} d\rho y_k, \tag{8}$$

where y_k is the optimal solution ((7)). We include the convergence condition (8) in the Farvardin's iterative algorithm[4] for searching for the optimum quantizer parameters. The iterative algorithm is outlined as follows.

Step 1. Set the iteration index $k = 0$. Choose an initial set of reconstruction levels $\mathbf{y}^{(0)} = \{y_1^{(0)}, y_2^{(0)}, \cdots, y_M^{(0)}\}$.
Step 2. Shuffle the reconstruction levels (and their indices) in such a way that the convergence condition (8) is satisfied. Then, determine the quantizer threshold levels $\mathbf{x}^{(k)} = \{x_1^{(k)}, x_2^{(k)}, \cdots, x_N^{(k)}\}$ using (6).
Step 3. Let k=k+1. Compute the reconstruction levels $\mathbf{y}^{(k)}$ using (7).
Step 4. Compute $D_t^{(k)}$. If $(D_t^{(k-1)} - D_t^{(k)})/D_t^{(k)} < \delta$, where δ is a pre-selected small positive value, stop; otherwise, go to Step 2.

3. TRANSFORM IMAGE CODER

Figure 1 shows a typical image transform coding scheme. The input image data is first partitioned into blocks and then each block is processed by the DCT operator. In the quantization process — the gray box showed in Figure 1, each DCT coefficient is quantized by a specific quantization step size. An ordinary transform coder employs a VLC to increase compression efficiency and an error control code (ECC) to reduce channel error effect. In our combined source/channel coder, VLC and ECC are not used. Since the quantization indices are sent directly through a noisy channel to the receiver, the robust quantizer can be viewed as a compound of a quantizer, a VLC coder, and a channel coder. At the receiving end, the received data are processed by the dequantizer and the inverse DCT unit. Similarly, the robust dequantizer has the functions of a dequantizer, a VLC decoder, and a channel decoder.

The quantizer used here is the one proposed in section 2 for noisy channels. Since the bits representing the reconstruction level index is sent directly to the noisy channel, the quantizer designed for a noisy channel can be viewed as a compound of the a quantizer, a VLC coder, and a channel coder. Hence, in this formulation the variable r in the distortion function $d(r)$ denotes the transmission bit (the quantizer output bit number) instead of the entropy of the quantized data. Based on the proposed quantizer design, we can obtain a set of quantizers with r-bit outputs as shown in Figure 2, where $1 \leq r \leq r_{max}$. Then, the bit allocation problem is formulated as follows. Given a constrained transmission bit rate R, we like to find a bit allocation vector $\mathbf{r}$ which minimizes

$$D(\mathbf{r}) = \sum_{i=1}^{L} d(r_i) \tag{9}$$

subject to

$$0 \le r_i \le r_{max},\ i = 1, 2, \cdots, L \tag{10}$$

and

$$\sum_{i=1}^{L} r_i \le R, \tag{11}$$

where L is the number of coefficients in an image block, and $d(r_i)$ represents the MSE of a quantizer with r_i-bit codeword representation. To determine the best bit allocation vector $\mathbf{r}$, we use an algorithm[5] based on the Lagrangian method.

Define $B(\mathbf{r})$ as $\sum_{i=1}^{L} r_i$. As been proved in Shoham and Gersho,[5] for any $\lambda \ge 0$, the solution $\mathbf{r}^*(\lambda)$ to the unconstrained problem

$$\min_{\mathbf{r}}\{D(\mathbf{r})\lambda B(\mathbf{r})\} \tag{12}$$

is also the solution to the constrained problem. This algorithm is restated briefly below. To simplify notation, $B^*(\lambda) = B(\mathbf{r}^*(\lambda))$.

Step 1: Guess an initial value λ.

Step 2: Solve the unconstrained problem using (12). If λ is not singular, there is only one solution and one constraint $B^*(\lambda)$. Otherwise, there are at least two different solutions. Find the two solutions $\{\mathbf{r}^*(\lambda)\}$ corresponding to the greatest and the smallest constrains denoted by $B_l^*(\lambda)$ and $B_h^*(\lambda)$, respectively.

Step 3: If the desired constraint R is such that

$$B_l^*(\lambda) \le R \le B_h^*(\lambda),$$

then obtain all solutions and find the one of which the constraint, denoted by $B_a^*(\lambda)$, is closest to R. Stop the search.

Step 4: If $R < B_l^*(\lambda)$ or $R > B_h^*(\lambda)$, find the next singular λ. Then, go to Step 2.

If the search terminates with only an approximate solution, which means that the allocated bits is not equal to R, then we add one more bit using the following steps.

Step 5: Find the index k' that solves

$$min_{\mathbf{r}}\{d(r_k^* + 1) - d(r_k^*)\}.$$

Then, $r_{k'}^* = r_{k'}^* + 1$ bits.

Step 6: Repeat Steps 5 and 6 until the sum of the allocated bits equals R.

Table 3 shows the bits assignment matrix for $R = 128$ bits/block and $r_{max} = 8$. Since the low frequency coefficients have stronger power than the high frequency coefficients, more bits are assigned to the low frequency coefficients.

4. SIMULATION RESULTS

Assume that the input source has the Gaussian distribution with zero mean and unit variance. In our simulations, the reconstruction level index is represented by the natural binary code. In the iterative procedure described in section 2, the parameters of the Lloyd-Max quantizer are used as the initial values $\mathbf{x}$ and $\mathbf{y}$. The simulation results show that some of the quantizer threshold levels vanish; that is, the boundaries x_l and x_{l-1} of the l^{th} interval are equal. This interval vanish phenomenon implies a resolution-reduction of the quantizer to decrease the channel error effect on the destination distortion. Table 2 shows the optimal threshold levels and the reconstruction levels of a 4-level quantizer with 2-bit index representation for various noisy channels (Note: $x_0 = -\infty$, $x_4 = \infty$).

Table 3. *Bit assignment for the proposed quantizers at $R = 128 bits/block$, where the quantizers are designed for noise range*$=(0.0, 0.1)$

8	8	8	8	7	2	1	0
8	8	8	8	3	1	0	0
8	8	8	5	2	1	0	0
5	3	3	2	1	0	0	0
2	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

The block size in our transform coder is 8×8. Picture "*lena*" of size 512×512 is used as our test image. The JPEG default quantization matrix is used in our simulation. The transformed coefficients are assumed to have Gaussian distributions. Figures 3 and 4 show the performance of the proposed robust image coder and the JPEG coder with two BCH channel codes. The compressed bits produced by the JPEG coder are 168304. In Figure 3, since a BCH(255,99) code is used to protect the JPEG compressed data, the number of total transmission bits is $168304 \times 255/99 = 433510.3$; that is, the number of the transmission bits per image block equals to 433510.3/(4096 blocks in a picture)=105.8 bits/block. Therefore, we use 105 bits/block in designing our combined source/channel coder. The simulations show that the BCH(255, 99) code can achieve the nearly error-free situation when the error rate is smaller than 0.04. When the channel error rate is greater than 0.04, the channel decoder cannot decode the noise-corrupted received data correctly; hence, the compressed data are suffered badly from the channel noise. In the JPEG codec system shown in Figure 1, the quantization indices are first coded by a VLC coder (Huffman coder) and then further coded by a BCH coder. As discussed in section 1, a couple of bit errors may cause the VLC decoder to lose its synchronization and to produce successive errors. Figure 3 indicates that the performance of a JPEG codec decreases rapidly when the channel coder can not achieve the error-free situation (error rate > 0.04). When the error rate is greater than 0.06, the noise-induced errors may lead to illegal JPEG bit streams and thus the JPEG decoder halts its decoding operation.

Designed under the assumption that the channel error rate can be any value within $(0.0, 0.1)$, the proposed image coder performs robustly in this noise range. As shown in Figure 3, it has a higher distortion than the JPEG+BCH(255,99) coder for the channel error rate less than 0.04. However, it has a very graceful degradation for the channel error rate greater than 0.04, and it clearly outperforms the JPEG coder for $\rho > 0.045$. Similar results are shown in Figure 4, where a BCH(255, 131) channel code is used. Since the BCH(255, 131) code has a weaker error-correcting capability ($t = 18$), the performance of the JPEG coder with a BCH(255, 131) channel code degrades earlier than that of the JPEG coder with a BCH(255, 99) channel code. In the case of a BCH(255, 131) code, the number of transmission bits is $168304 \times 255/131 = 327614.7$; that is, the number of bits per block is 79.9. Therefore, the constrained bit rate in Figure 4 is 79 bits/block.

Figure 5 shows the reconstructed picture of "*lena*" processed by the JPEG codec with BCH(255, 131) code at channel noise= 0.01. Since the BCH(255, 131) code is able to correct nearly all the errors in this case, the overall distortion is made of only the source distortion due to the JPEG compression scheme, which is independent of channel. Figure 6 shows the reconstructed picture produced by the proposed coder. As described earlier, our proposed coder is designed for the robust operation in the noise range $(0.0, 0.1)$. Therefore, its performance is lower than that of the Shannon's separated coder operated in the error-free situation. On the other hand, our proposed coder outperforms the separated method operated in non-error-free situation. Figure 7 shows the reconstructed picture of the JPEG+BCH(255, 131) codec for $\rho = 0.04$. The asynchronous decoding of the VLC decoder induces disasters on the reconstructed images. Since no VLC coders are used in the proposed system, error propagation never occurs as shown in Figure 8.

5. CONCLUSIONS

In a noisy environment, a single bit error in the VLC-coded data may induce symbol insertion and erasure in the VLC decoding process. This symbol insertion and erasure not only leads to asynchronous decoding, but also produces burst errors. We propose a robust quantizer design that has a consistent performance for a wide channel noise range (ρ_{min}, ρ_{max}). Since the bits representing the reconstruction level are transmitted directly over a noisy channel, the proposed quantizer can be viewed as a compound of a quantizer, a VLC coder, and a channel coder. We then apply the proposed quantizer to an image transform coder. The simulation results show that our proposed transform coder has a graceful distortion within the designed noise range. On the other hand, the performance of the separated coder (JPEG coder and BCH code are used in this paper) degrades rapidly in a highly noisy environment when the channel coder can not achieve the error-free condition.

REFERENCES

1. C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, 1948.
2. G. C. Clark, Jr., and J. B. Cain, *Error-correction coding for digital communications,* New York, Plenum Press, 1982.
3. C. -H. Su, H. -M. Hang, and C. -H. Wei, "Optimum quantizer design for a slow fading frequency-nonselective channel," submitted for publication.
4. N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source-channel coding," *IEEE Trans. Inform. Theory,* vol. IT–33, pp. 827–838, Nov. 1987.
5. Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust. Speech. Signal Proc.,* vol. 36, pp. 1445–1453, Sept. 1988.

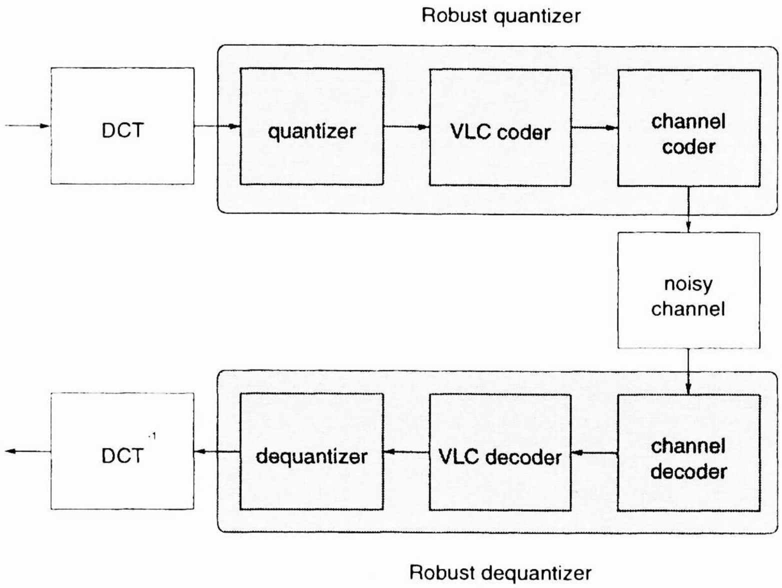

Figure 1. *A typical transform image coder.*

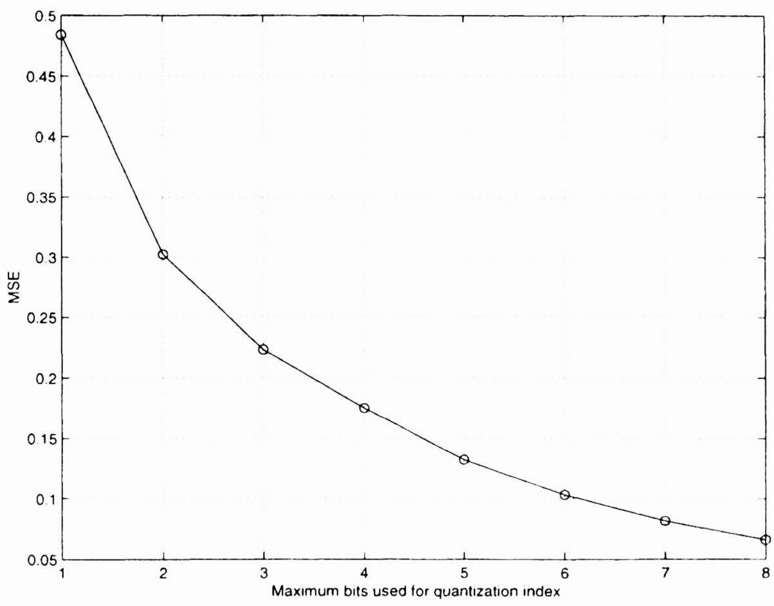

Figure 2. *The performance of the proposed quantizer.*

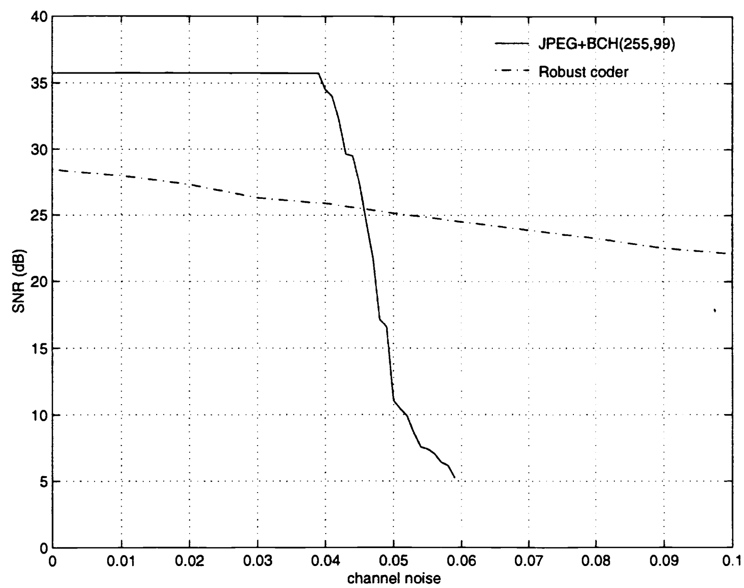

Figure 3. *The performances of the proposed robust codec scheme (105 bits) and the JPEG+BCH(255,99) code ("lena" picture).*

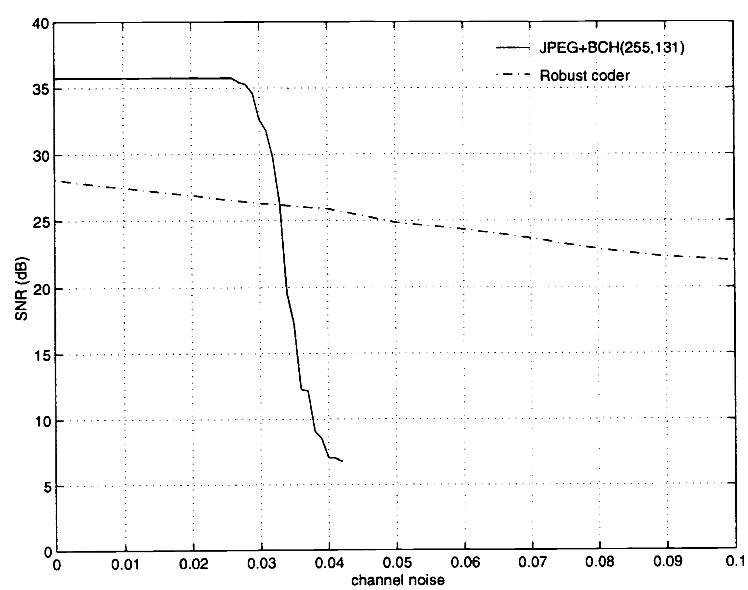

Figure 4. *The performances of the proposed robust codec scheme (79 bits) and the JPEG+BCH(255,131) code ("lena" picture).*

Figure 5. *The reconstructed picture by the JPEG+BCH(255,131) coder in the case of the channel noise=0.01.*

Figure 6. *The reconstructed picture by the robust scheme (79 bits) in the case of the channel noise=0.01.*

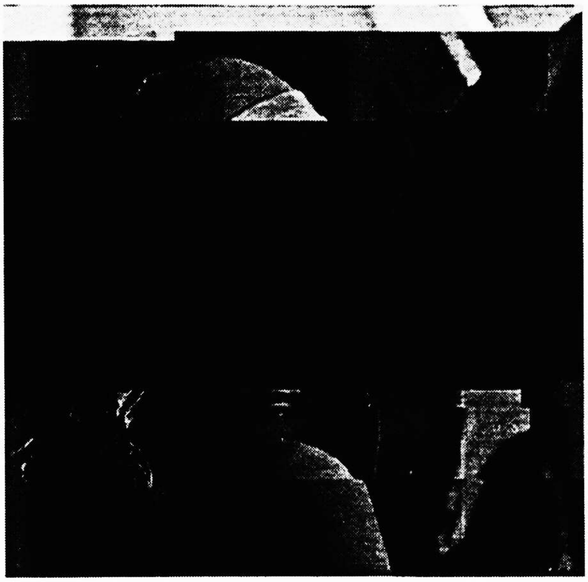

Figure 7. *The reconstructed picture by the JPEG+BCH(255,131) coder in the case of the channel noise=0.04.*

Figure 8. *The reconstructed picture by the robust scheme (79 bits) in the case of the channel noise=0.04.*