Optimization of The Bases to Use and The Blocks to Apply in DCT-ICA Hybrid Coding Methods

Yudai Nakada
Graduate School of Software and Information Science
Iwate Prefectural University
152-52, Sugo, Takizawa, Iwate, Japan
g231t027@s.iwate-pu.ac.jp

Masashi Kameda
Graduate School of Software and Information Science
Iwate Prefectural University
152-52, Sugo, Takizawa, Iwate, Japan
kameda@iwate-pu.ac.jp

Abstract— The Discrete Cosine Transform (DCT) can extract and preserve the statistical features that an image needs with a generic basis, but it can cause visual distortion at lower bit rates. Therefore, we propose the hybrid DCT-ICA coding method using independent component analysis (ICA), which can extract and preserve features handled by human vision. However, optimizing the classification of the spatial region to which DCT or ICA is applied and the appropriate combining ICA bases to use DCT and ICA together is required. In this paper, we determine the optimal combination of ICA bases used for a given bit rate and the optimal segmentation to apply the DCT and ICA bases. Our proposed method improves coding performance over DCT only at all bit rates for some natural images.

Keywords—image compression, lossy compression, discrete cosine transform, independent component analysis.

I. INTRODUCTION

JPEG and MPEG is one of the effective methods based on orthogonal transforms with energy compression [1] in the coding methods based on the discrete cosine transform (DCT). In these methods, the DCT basis is applied to any input image, and high-frequency information is removed by the Q table. Therefore, DCT is effective in preserving regions of the image that have common statistical features. However, distortions such as mosquito noise and block noise are well known to be problematic at low bit rates. This is considered that DCT removes even visually necessary information.

Sparse coding, which models the human early visual system, focuses on the sparsity of the coefficients and can preserve features in a small region in the image by only a few bases [2,3]. It has been shown that independent component analysis (ICA), a form of sparse coding, can be applied to images to extract features handled by human vision [4,5]. We have proposed the DCT-ICA hybrid coding methods to preserve the information needed in both image data and human vision [6~8]. This method improves coding performance by classifying a given image in terms of image quality degradation and applying DCT or ICA. First, the input image is divided into blocks of 8 × 8 pixels, and each block is classified by comparing its bit rate when DCT or ICA is applied respectively. The bit rate of each block is controlled by the Q table for DCT or by the choice of basis used for ICA. If an appropriate ICA basis is not chosen, the bit rate will increase, and therefore, the coding performance will decrease. Noting that a unique ICA base is used when the bit rate is extremely low, the conventional method has realized on improved coding performance over DCT only by choosing the bases based on the similarity between the signal of a base and a block [6,7]. However, for multiple ICA bases are combined in most cases to preserve information in a region as the bit rate increases [8], and then this paper proposes a new optimal combination of bases to improve the coding performance in DCT-ICA hybrid coding method. Since the basis set obtained by ICA are different for each input image, these are shared between the encoder and the decoder. Therefore, it needs to limit the number of ICA bases used for the encoding process to reduce the information shared. We found the optimal ICA basis that can maximize the image quality of each block, and only some of them were used in [7]. However, since there are several ICA bases that can improve image quality from DCT except for one that can maximize block quality, the proposed method determines the truly necessary ICA basis for the entire image.

This study aims to develop the DCT-ICA hybrid coding method for practical bit rates. The results of applying the proposed method to some natural images show that the coding performance can be improved from that of DCT only for all bit rates.

II. IMAGE CODING METHOD USING ICA

A. Independent Component Analysis (ICA)

Independent component analysis (ICA) is a method of transforming observed multi-dimensional random vectors into original signals that are as independent as statistically possible. When an input signal $X = (x_1, x_2, \cdots, x_m)^T$ is represented by a linear combination of independent bases $S = (s_1, s_2, \cdots, s_n)^T$ can be written as

$$X = AS. \tag{1}$$

Note that A is the coupling coefficient represented as an $(m \times n)$ matrix, and element a_{ij} represents the contribution of the independent basis s_i to the input signal x_j. Since ICA does not have the information of the bases and coefficients, it must recover the basis S from the input signal X only. By denoting the inverse of A and the approximation of S as W and $Y = (y_1, y_2, \cdots, y_n)^T$ respectively, (1) can be transformed as

$$Y = WX. \tag{2}$$

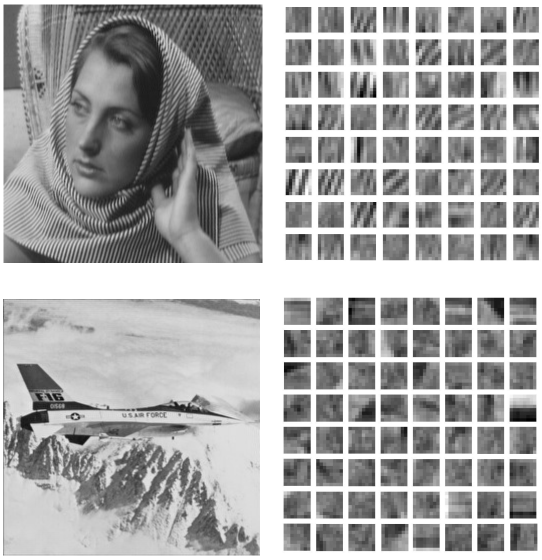


Fig. 1. The input image and obtained ICA basis sets

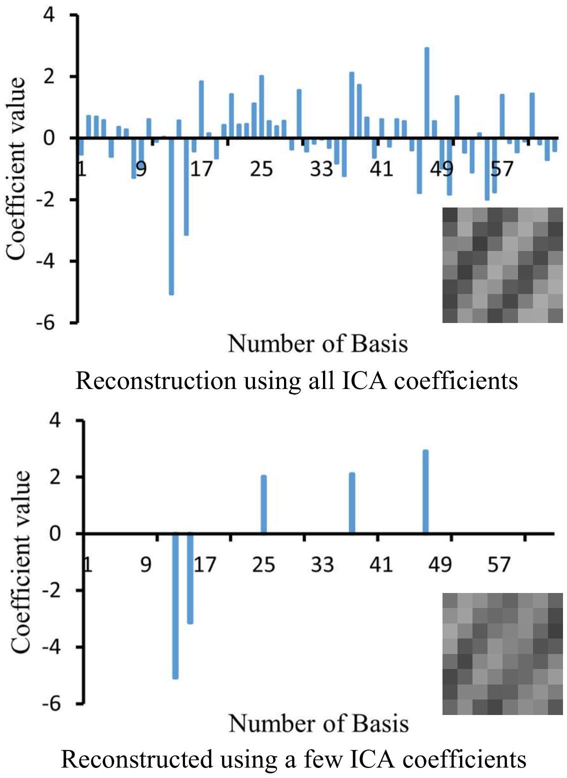


Fig. 2. Sparseness of the ICA coefficients and quality of reconstructed block

In ICA, the objective is to find the ICA coefficients W that makes each component of the ICA bases Y independent. The Kullback-Leibler information content [9] is often used as the assession criterion for independence, and by applying the method based on the steepest descent method proposed by Bell et al., we can obtain an update rule for Y that minimizes the mutual information content,

$$W^{k+1} = W^k + \mu[I - \varphi(Y)Y^T]W^T, \tag{3}$$

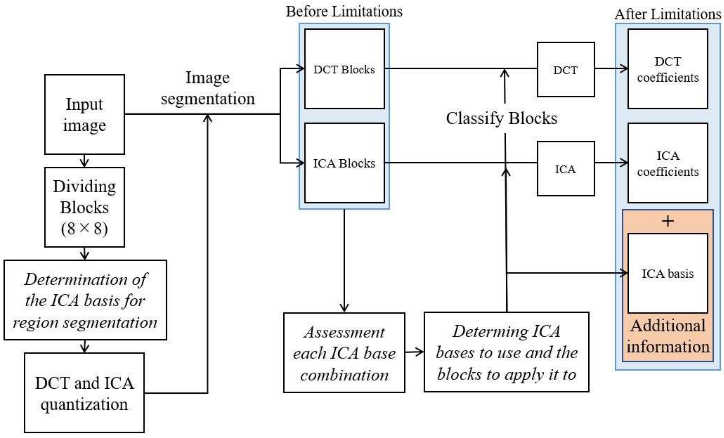


Fig. 3. Configuration of the DCT-ICA Hybrid coding method

where μ is the learning coefficient, I is the unit matrix, and φ is an approximation of the probability density function of Y. Typically, the Sigmoid function is used as φ.

In this paper, the number of ICA bases is determined to 64 in order to match the 8×8 pixels DCT bases which is used in the proposed method. Fig.1 shows the ICA bases derived by applying (3) when input images "Barbara" and "Airplane" are given. It is seen in Fig.1 that the ICA basis corresponds to the local features of each input image and the shape of bases is deferent from each input image. The ICA coefficients of an arbitrary block in the image "Barbara" are shown in Fig. 2. Fig. 2 compares the reconstructed block using all ICA coefficients with the block reconstructed using a few ICA coefficients with large coefficient values. It is seen in Fig. 2 that the ICA coefficients have sparsity [2~4] since only a few ICA bases can preserve the local features of the input image. Therefore, it is expected that the ICA basis can reduce the bit rates required to preserve the signal of blocks with local features to DCT.

B. DCT-ICA Hybrid Coding Methods

Fig. 3 shows the construction of the DCT-ICA hybrid coding method. As discussed in the previous section, the signal of a block with local features has sparser ICA coefficients, so applying ICA can be expected to reduce the bit rate. On the other hand, block with flat or regular features do not satisfy the sparsity of ICA coefficients, so it is better to have DCT applied to them to reduce the information. Here, the size of each block in the conventional and the proposed methods is 8×8 pixels. Therefore, by comparing the bit rates encoded by DCT and ICA respectively, some blocks with local features can be classified into ICA Blocks and the other blocks with flat or regular features into DCT Blocks. In ICA, the bit rate is controlled by the choice of the ICA basis used. In order to determine the ICA basis, it is optimal that the ICA basis with the lowest bit rate is chosen using the exhaustive search. However, since it is known that the number and type of bases used in each block is different as the bit rate changes, it is computationally impractical to choose the ICA basis to be used in all blocks by the exhaustive search. Therefore, the conventional method [6, 7] were proposed as the choice method of ICA bases based on the MP method [10]: the MP method chooses the basis with the highest similarity between

TABLE I. DEGRADATION OF IMAGE QUALITY FROM EXHAUSTIVE SEARCH DUE TO CHOICE OF COEFFICIENTS

Number of ICA basis used	Proposed Method	Conventional Method	Exhaustive Search
1	382.9	440.2	382.9
2	277.9	351.2	276.9
3	216.5	297.3	214.6

the signal of each block and the signal of the ICA base, in that order. By applying this choice method to all blocks in advance, and then determining the order in which the bases are chosen. As the result, it is possible to choose a basis with less computational complexity than the exhaustive search. However, as the bit rate increases, multiple ICA bases have to be combined to many encode blocks [8], moreover, it was confirmed that the conventional method based on the MP method chooses a different basis than the exhaustive search.

Since the basis set obtained by ICA are different for each input image, these are shared between the encoder and the decoder. Because the shared information increases in proportion to the number of ICA bases used, it is necessary to limit the number of ICA bases used in ICA-based coding. We have been chosen as the basis that maximizes the image quality of each block compared to DCT as the optimal basis that minimizes the bit rate of the entire image in [7]. However, since the existence of semi-optimal bases that can improve the image quality from DCT has been found in addition to the optimal basis, the addition of a semi-optimal basis that the conventional method did not focus on can further reduce the entire image bit rate, making the DCT-ICA hybrid coding method realizable even for practical bit rates.

Finally, regions that were initially classified as ICA Blocks based on all bases are reclassified because the bit rate may increase when using the semi-optimal basis that was finally determined. The process is completed by reclassifying only those blocks where the bit rate can be reduced by using chosen ICA basis as ICA Blocks and reclassifying the other blocks as DCT Blocks. As mentioned above, as the bit rate changes, the number, and type of bases used in each block also change, so the DCT-ICA hybrid coding method needs to separate the process for the given bit rate.

III. PROPOSED METHOD

Conventional methods were effective when the bit rate was extremely low, however, they had deterioration with the coding performance at practical bit rates. Therefore, we realize the DCT-ICA hybrid coding method at all bit rates by solving the problem of region segmentation and the problem of shared information in ICA by considering the features of ICA that are revealed at practical bit rates. The parts of the proposed method that correspond to the current proposal are indicated by shaded letters in Fig. 3.

A. Determination of the ICA basis for region segmentation

It is not realistic from a calculation cost point of view to perform the exhaustive searches in all blocks to choose a basis. In addition, at practical bit rates, multiple ICA bases are almost used to encode a block, and the bases used at each bit rate are different [8]. The proposed method determines the optimal basis for each bit rate as follows, by obtaining the order of basis selection for each region in advance, considering the fact that multiple bases are combined.

1) In a block, find the ICA base $B_i(i = 1, 2, \cdots, 64)$ *with the minimum MSE and set it as* P_1 *for that block.* P_1 *is the first base to be chosen when that block is encoded.*

2) Given the combined use of P_2*, let the second ICA base* $B_j(j = 1, 2, \cdots, 64, where\ j \neq i)$ *that can minimize the MSE be the* P_2 *for that block.*

3) Continue this process until P_{64} *is determined.*

4) Apply 1) ~ 3) to all blocks.

When a bit rate is specified, each block is chosen in order from P_1. For example, for a bit rate that uses three bases, $P_1 \sim P_3$ are chosen. This allows the basis to be determined for use in fewer calculations than the exhaustive search. Using the natural image "Airplane" as an example, Table I shows the MSE when using the bases determined by the proposed method, exhaustive search, and the conventional method (determination by the MP method). Table I shows that the MSE of the conventional method is significantly increase compared to the exhaustive search, while the MSE of the proposed method is almost the same as that of the exhaustive search. Note that the results in Table I are similar for other images. These results show that the proposed method is possible to solve the problem of the conventional method of reduced image quality when multiple bases are chosen.

B. Determination of optimal ICA basis and ICA Blocks

The DCT-ICA hybrid coding method reduces the information shared with the decoder by reducing the number of ICA bases used. The proposed method is defined to use up to three ICA bases for the entire image. In addition to the optimal basis that can maximize the image quality for each block, several semi-optimal bases that can improve the image quality from DCT has been confirmed to exist. Therefore, we propose a method to determine a few ICA basis combinations that minimize the bit rate for the same image quality by obtaining and comparing the image quality for all ICA basis combinations as follows,

1) The improvement in entire image quality when ICA bases a, b, c $(a, b, c = 1, 2, \cdots, 64, where\ a \neq b, a \neq c, b \neq c)$ *are used is*

$$Q_{a,b,c} = \sum_{All-Block} \left[\frac{1}{64} \sum_{j,k=0}^{8,8} \left\{ \left(org_{j,k} - DCT_{j,k} \right)^2 - \left(org_{j,k} - Hybrid_{j,k} \right)^2 \right\} \right], (4)$$

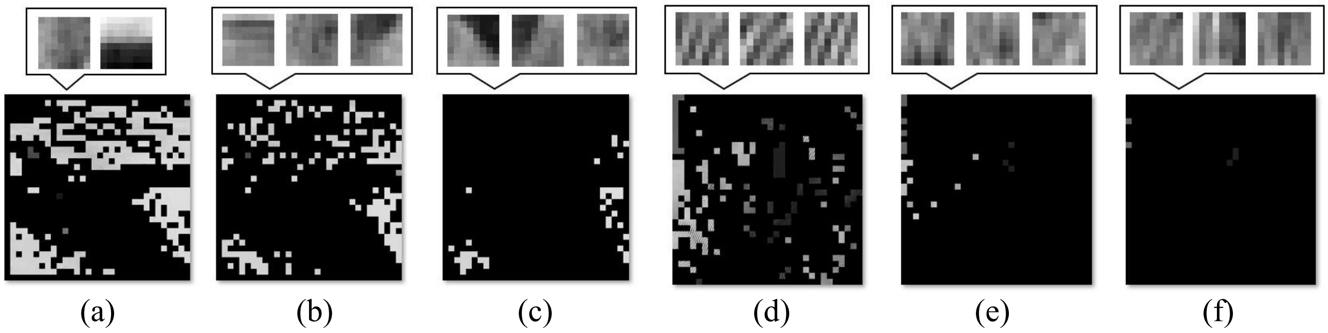


Fig. 4. Optimal ICA basis and ICA Blocks in each bit rate: (a) 28[dB], (b) 32 [dB], (c) 36 [dB] of "Airplane". (d) 27[dB], (e) 32 [dB], (f) 36 [dB] of "Barbara".

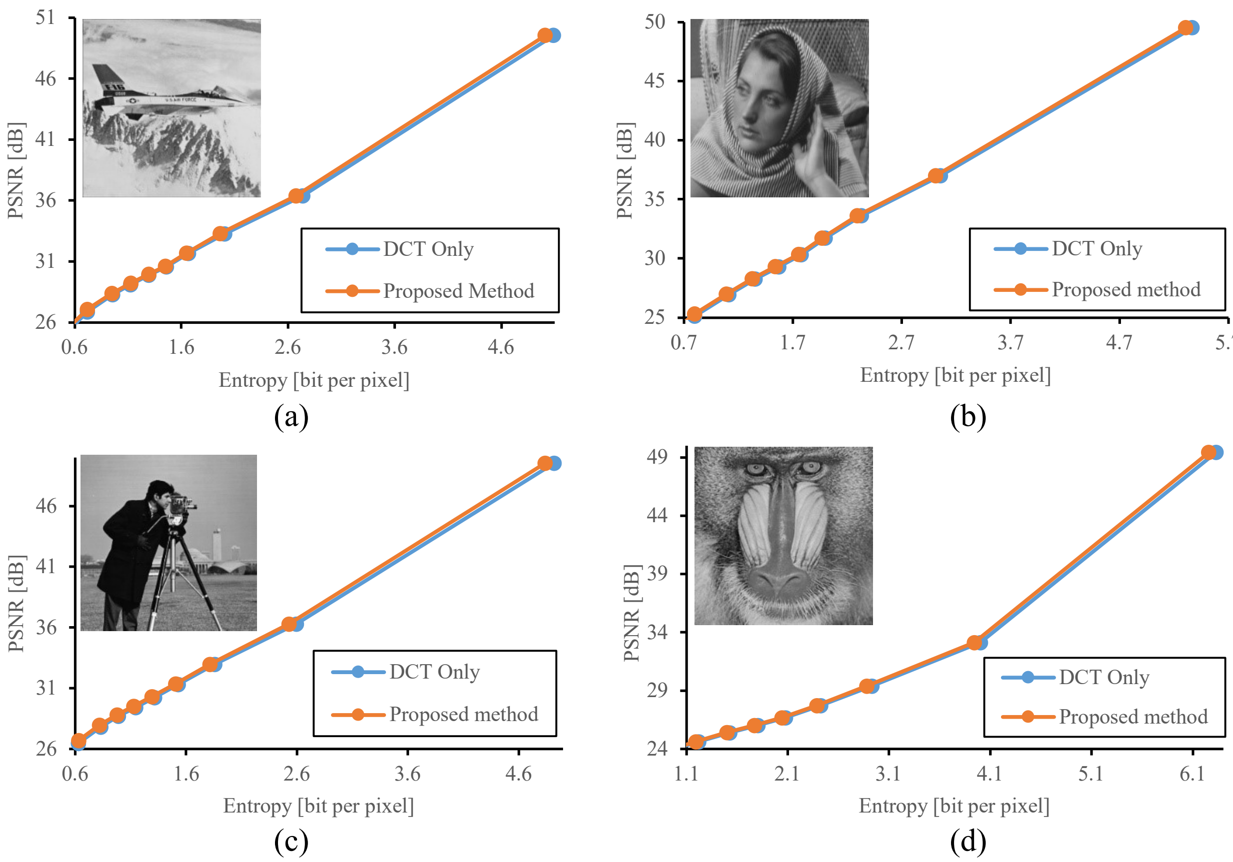


Fig. 5. Coding performances of the proposed method: (a) Airplane, (b) Barbara, (c) Cameraman, (d) Mandrill

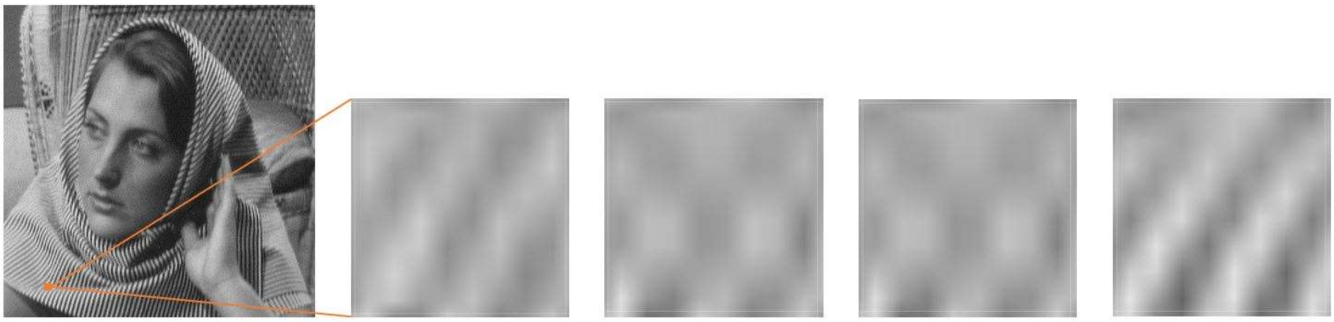
where $org_{j,k}$ *is the original image,* $DCT_{j,k}$ *by the coding image of DCT, and* $Hybrid_{j,k}$ *by the proposed method of the coding image. Blocks where applying basis a, b, c increases the bit rate from that of DCT only are excluded from* $Q_{a,b,c}$ *because they increase the bit rate of the entire image. In (4),* $Q_{a,b,c}$ *is the sum of all MSE that can be improved from DCT only by all combination of basis in* ${}_3P_1 + {}_3P_2 + {}_3P_3$ *when basis a, b, c are given.*

2) Since the number of bases chosen by basis a, b, c is 0 ~ 3, $Q_{a,b,c}$ *is obtained for 43745 patterns, which is the sum of 1 pattern when no base is used, 64 patterns when 1 base is used (*${}_{64}C_1$*), 2016 patterns when 2 bases are used (*${}_{64}C_2$*) and 41664 patterns when 3 bases are used (*${}_{64}C_3$*).*

3) All blocks that improve image quality compared to DCT only when using the combination of 2) are ICA Blocks, and all other blocks are DCT Blocks. The entropy of each DCT and ICA coefficient is combined with the entropy of the basis a, b, and c themselves to obtain the bit rate required for coding.

4) The ICA basis combinations with the maximum $Q_{a,b,c}$ *are compared in turn with the bit rate of the DCT, and the ICA basis a, b, and c that first reduces the bit rate from the DCT is determined to be the optimal ICA basis combination.*

Since $Q_{a,b,c}$ changes as the bit rate changes, the proposed method is applied to each bit rate. With the above process, the effects of ICA basis a, b, and c on the entire image can be



(a) (b) (c) (d)

Fig. 6. Comparison of the image quality in encoded image at 29[dB]: (a) Proposed method, (b) Conventional method, (c) DCT only, (d) Original image.

determined, and the bit rate with additional information can be compared, thus solving the problem of information to be shared.

IV. Expremental Results

A. Optimal ICA basis combination and ICA Blocks

The optimal ICA basis combination and the optimal ICA Blocks in each bit rate when the proposed method is applied to the 256 × 256 pixels natural image "airplane" are shown in Fig. 4. Fig. 4 shows the ICA Blocks in the original image and the DCT Blocks in black. It is seen in Fig. 4 that the optimal basis combination is different for each bit rate, and the blocks which are encoded by the obtained ICA basis is also different. These means that the proposed method was applied at each bit rate, resulting in the choice of an appropriate basis and the classification of suitable regions at each bit rate. In addition, three ICA bases are used at practical bit rates, suggesting that the entropy of the coefficients is drastically reduced by the proposed method.

B. Comparison coding performance and subjective evalution

The results of applying the proposed method to the 256 × 256 pixels natural images “Airplane”, “Barbara”, “Cameraman”, and “Mandrill” are shown in Fig. 5. The conventional method is applicable only to extremely low bit rates, and its performance is equivalent to that of the extremely low bit rates in the proposed method. Fig. 5 shows that the proposed method reduces the bit rate from DCT. In addition, the high bit rate can drastically reduce entropy, while the low bit rate can improve image quality. This means that the proposed method reflects the ICA property of being able to preserve image quality with a few bases. As the results, the DCT-ICA hybrid coding method is effective to any natural image, since similar results were obtained for all four images with different features.

Fig. 6 shows part of the coding image when the proposed method is applied to the image "Barbara". Fig. 6 compares the original image, the encoded image by the proposed method, the conventional method, and DCT (encoded at 29[dB]). Fig. 6 visually confirms that the proposed method preserves local features that could not be preserved by DCT and the conventional method. This result means that the proposed method can preserve local features of the input image with a fewer bit rate.

V. Conclusion

In this paper, we propose a method for determining the ICA basis considering the features of ICA at practical bit rates and extend the applicable bit rate range of the DCT-ICA hybrid coding method. In addition, the proposed method improved the coding performance from DCT only at all bit rates.

The optimal ICA basis combination shown in Fig. 4 is different for each bit rate. However, whenever the bit rate changes, re-determination of the optimal ICA basis would increase the processing cost. Therefore, reducing the processing cost of ICA basis determination is left for our future challenge.

References

[1] W.B. Pennebaker, J.L. Mitchell, JPEG still image data compression standard, Van Nostrand Reindold, 1993.

[2] H. Hino, N. Murata, “Mathematics of Sparse Representation and its Applications,” *IPSJ SIG Technical Report*, vol. 2009-CVIM-183, no.3, pp.1-10, 2012.

[3] Z. Zhang, Y. Xu, J. Yang, X. Li and D. Zhang, "A Survey of Sparse Representation: Algorithms and Applications," in *IEEE Access*, vol. 3, pp. 490-530, 2015.

[4] N. Murata, Introductory Independent Component Analysis, Tokyo Denki University Press, 2014.

[5] Y. Chen, Application of Independent Component Analysis (ICA) to Pattern Recognition and Image Processing and MATLAB Simulation, Triceps, 2007.

[6] M. Kameda, K. Kawamura, “A Still Image Coding Method Using Discrete Cosine Transform and Basis of Independent Component Analysis,” *The journal of the Institute of Image Electronics Engineers of Japan*, vol.45, no.2, pp.201-211, 2016.

[7] A. Tgashi, M. Kameda, “Selection of Important ICA Basis Based on Comparison of Image Quality Improvement in Still Image Coding Using Basis of Independent Component Analysis,” *IEICE technical report*, vol.118, no.501, pp.153-158, 2019.

[8] Y. Nakada, M. Kameda, “Important of Coding Performance by Optimal Basis Selection in Image Compression Method Using ICA Basis,” *Proc of The Eleventh International Workshop on Image Media Quality and its Applications*, 2022, pp.36-42.

[9] J. kauhunen, E. Oja, L. Wang, R. Vigario and J. Joutsensalo, “Aclass of neural networks for independent component analysis,” *IEEE Transactions. on Neural Networks*, vol.8, no.3, pp.486-504, 1997.

[10] S. Mallat, A. Zhang, ”Matching Pursuits with Time-Frequency Dictionaries, ” *IEEE Trans. on Signal Processing*, Vol.41, No.12, pp.3397-3415, 1993.