Title: ICA-DCT ハイブリッド符号化における，選出基底と適用領域の最適化

Abstract:

The Discrete Cosine Transform (DCT) can extract and preserve the statistical features that an image needs with a generic basis, but at lower bit rates it can cause visual distortion. 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 blocks to which DCT or ICA is applied and combining ICA bases to use DCT and ICA together is necessary. In this paper, we determine the combination of ICA bases used for a given bit rate and the blocks to which they are applied and show that applying this method to multiple images improves coding performance over DCT only at all bit rates.

1. 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 the 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. On the other hand, distortions such as mosquito noise and block noise are known to be problematic at low bit rates. This is because 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 block 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]. Therefore, 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 input images in terms of image quality degradation and applying DCT or ICA. First, the image dividing into blocks of 8 × 8 pixels, and each block is classified by comparing its bit rate when DCT or ICA is applied. The bit rate of each block is controlled by the Q table for DCT or by the choice of ICA basis used for ICA. If an appropriate ICA basis is not chosen, the bit rate will increase, and conversely, coding performance will decrease. Noting that an ICA base is used when the bit rate is extremely low, [6,7] 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. However, for multiple ICA bases are combined in most cases to preserve information in a region as the bit rate increases [8], this paper proposes a new optimal combination of bases to improve the coding performance. Since the basis set obtained by ICAI 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 in encoding to reduce the information shared. [7] found the optimal ICA basis that can maximize the image quality of each block, and only some of them were used. However, since there are other ICA bases that can improve image quality from DCT than the one that can maximize block quality, the proposed method determines the truly necessary ICA basis for the entire image.

Based on the above, this study develops the DCT-ICA hybrid coding method for practical bit rates. The results of applying the proposed method to natural images show that the coding performance can be improved from that of DCT only for all bit rates.

2．

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 the input signal is represented by a linear combination of independent bases can be written as

Note that is the coupling coefficient represented as an matrix, and element  represents the contribution of the independent basis to the input signal. Since ICA does not have the information of the bases and coefficients, it must recover the basis from the input signal only. By denoting the inverse of and the approximation of as and respectively, (1) can be transformed as

In ICA, the objective is to find the ICA coefficients that makes each component of the ICA bases 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 that minimizes the mutual information content,

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

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 of 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.

図、1　DCT基底、ICA基底の比較

2　係数を減らしたやつ

2.2　DCT-ICAハイブリッド符号化手法

Fig. 3 shows the structure of the DCT-ICA hybrid coding method. As discussed in the previous section, blocks with local features have sparser ICA coefficients, so applying ICA can be expected to reduce the bit rate. Here, the size of each block in the conventional and proposed methods is 8 × 8 pixels. On the other hand, blocks 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. Therefore, by comparing the bit rates encoded by DCT and ICA, blocks with local features can be classified into ICA Blocks and blocks with flat or regular features into DCT Blocks. In ICA, the bit rate is controlled by the choice of the ICA basis used. For basis choice, it is optimal that the ICA basis with the lowest bit rate is chosen by the exhaustive search. However, since it is known that the number and type of bases used in each block changes 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, [6, 7] ware proposed methods of choice ICA bases based on the MP method [10]: the MP method chooses the basis with the highest similarity between the signal of a block and the signal of the ICA base, in that order. By applying this method to all blocks in advance and determining the order in which the bases are chosen, it is possible to choose a basis with less computational complexity than the exhaustive search. However, as the bit rate increases, multiple ICA bases are combined to encode a block [8], so 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 ICAI are different for each input image, these are shared between the encoder and the decoder. Since 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. [7] is chosen as the basis that maximizes the image quality of each block compared to when DCT is used as the optimal basis that minimizes the bit rate of the entire image. However, since the existence of semi-optimal bases that can improve image quality from DCT has been found in addition to the optimal basis, the addition of a semi-optimal basis that [7] 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 using 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 described 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 each bit rate.

図　3　DCT-ICAハイブリッド符号化手法のシステム図

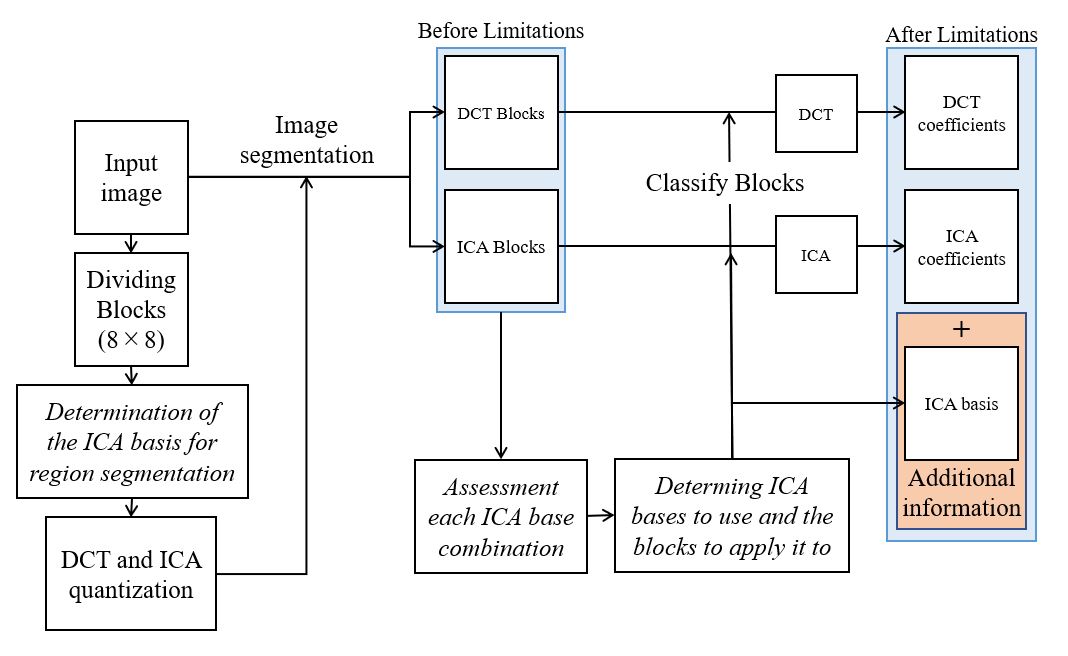


Fig. 3. Configuration of the DCT-ICA hybrid coding method

3提案手法

Conventional methods were effective when the bit rate was extremely low but had problems with coding performance at practical bit rates. Therefore, we realize the DCT-ICA hybrid coding method at all bit rates by solving the problem of applied blocks 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.

3.1領域分割のための基底の決定法

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 always 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 shown below, 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 Bi(𝑖 = 1,2, ⋯ ,64) with the minimum MSE and set it as P1 for that region. P1 is the first base to be chosen when encoding that block.
2. Given the combined use of P1, let the second ICA base Bj (j = 1,2, ⋯ ,64, where j ≠ 𝑖) that can minimize the MSE be the P2 for that block.
3. Continue this process until P64 is determined.
4. Apply 1)~3) to all blocks.

When a bit rate is specified, each block is chosen in order from P1. For example, for a bit rate that uses three bases, P1~P3 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 1 shows the MSE when using the bases determined by the proposed method, exhaustive search, and the conventional method (determination by the MP method). Table 1 shows that the image quality of the conventional method is significantly reduced compared to the exhaustive search, while the image quality of the proposed method is almost the same as that of the exhaustive search. That the results in Table 1 are similar for other images. These results show that the proposed method can solve the problem of the conventional method of reduced image quality when multiple bases are chosen.

表、１　優先度の画質比較

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 |
|  | 277.9 | 351.2 | 276.9 |
|  | 216.5 | 297.3 | 214.6 |

3.2　付加情報の削減のための基底の評価法

The DCT-ICA hybrid coding method reduces the information shared with the decoder by reducing the number of ICA bases used. The proposed method uses up to three ICA bases for the entire image. In addition to the optimal basis that can maximize the image quality for each block, semi-optimal basis 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. The specific process is as follows.

1. The improvement in entire image quality when ICA bases a, b, c are used is

Where is the original image, by the coding image of DCT, and by the proposed method of the coding image.

Regions where applying basis a,b,c increases the bit rate from that of DCT only are excluded from Qa,b,c because they increase the bit rate of the entire image.

In (4), Qa,b,c is the sum of all MSEs that can be improved from DCT only by the combination of 3P1+3P2+3P3 when basis a, b, c are given.

1. Since the number of bases chosen by basis a,b,c is 0~3, Qa,b,c is obtained for 43745 patterns, which is the sum of 1 pattern when no basis is used, 64 patterns when 1 basis is used (64C1), 2016 patterns when 2 bases are used (64C2) and 41664 patterns when 3 bases are used (64C3).
2. 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, c themselves to obtain the bit rate required for coding.
3. The ICA basis combinations with the maximum Qa,b,c are compared in turn with the bit rate of the DCT, and the ICA basis a,b,c that first reduces the bit rate from the DCT is determined to be the optimal ICA basis combination.

Since Qa,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, c on the entire image can be determined, and the bit rate with additional information can be compared, thus solving the problem of information to be shared.

4.実験結果

4.1　最適なICA基底の組み合わせとその適用ブロック

The optimal ICA basis combination and ICA Blocks when the proposed method is applied to the 256 x 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. Fig. 4 shows that the optimal basis combination is different for each bit rate, and the applied block 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.

4.2 符号化性能と主観評価

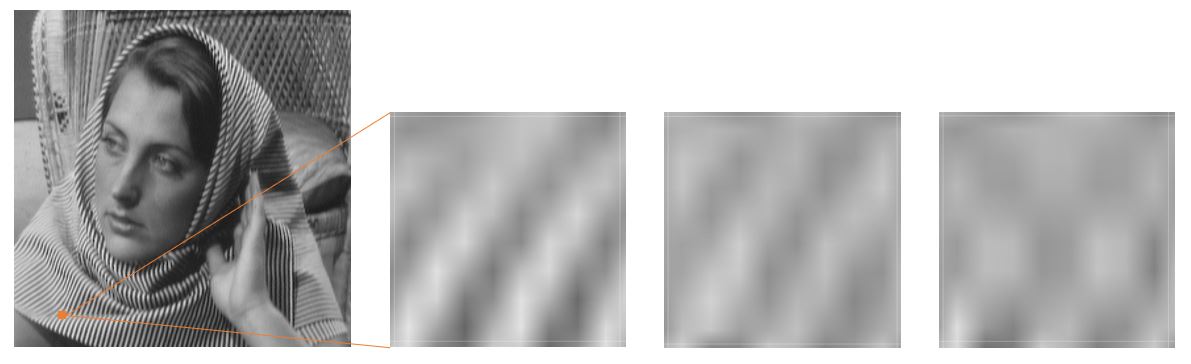
The results of applying the proposed method to the 256 x 256 pixels natural images "Airplane", "Barbara", "Cameraman", and "Mandrill" are shown in Fig. 5. The conventional method is applicable only to extremely low bit rates, and its performance is equivalent to that of the extremely low bit rates in 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 drastically improve image quality. This means that the proposed method reflects the ICA property of being able to preserve image quality with a few bases.

Fig. 6 shows the coding image when the proposed method is applied to the image "Barbara". Fig. 6 visually confirms that the proposed method preserves local features that could not be preserved by DCT by using ICA.

図、4　PSNR（Airplane、Barbara、Cameraman、Mandrill）

　　5　主観評価（バーバラ）

6　基底と適用ブロック（提案28dB,32dB,36dB）



原画像　　　　提案手法　　　　DCT

図5　29dBにおける符号化画像の比較

5.まとめ

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 used in Fig. 4 is different for each bit rate. However, redoing the basis selection each time a bit rate is specified would increase the processing cost. Therefore, reducing the processing cost of ICA basis selection is our future challenge.