# A FAST MATCHING PURSUITS ALGORITHM USING SUB-BAND DECOMPOSITION OF VIDEO SIGNALS

*Kousuke Imamura*, Yoshimasa Koba†, Hideo Hashimoto‡*

Kanazawa University
Division of Electrical Engineering and Computer Science
Graduate School of Natural Science and Technology
Kakuma-machi, Kanazawa, Ishikawa 920-1192 Japan
Tel: +81-76-234-4894, Fax: +81-76-234-4900
Email: {*imamura,‡hasimoto}@is.t.kanazawa-u.ac.jp, †coba@gin.ec.t.kanazawa-u.ac.jp

## ABSTRACT

A high-efficiency video coding method using matching pursuits, which is a waveform coding technique, has recently been proposed. In this method, the motion compensated prediction error image is encoded by applying matching pursuits. In the present paper, we propose a matching pursuits coding method that encodes the sub-band images derived from motion compensated prediction error images. The complexity of the proposed method is reduced compared to the full-band matching pursuits because of lighter inner product computation due to reduction of both the resolution of the sub-band image and the basis function length of the dictionary. We evaluated the coding performance and computational complexity of the proposed method via computer simulations.

## 1. INTRODUCTION

Current standard video coding methods are based on a hybrid system of motion compensation (MC) and discrete cosine transform (DCT). In these methods, a prediction error image is generated by MC, and the prediction error image is encoded by DCT. Recently, another high-efficiency video coding method using a waveform coding technique called matching pursuits (MP) in place of DCT has been suggested[1]-[6].

The matching pursuits algorithm approximates a signal using a dictionary, and therefore its performance depends heavily on the dictionary. More specifically, a dictonary that is well-matched to the image characteristic is capable of improving the coding performance.

In the present paper, we propose a video coding method in which a motion compensated prediction error image is decomposed into sub-bands and each sub-band image is encoded by matching pursuits. A dictionary that has a suitable sub-band frequency characteristic is constructed, and the subsampled and half-size reduced images are encoded by these dictionaries. Hence, the complexity of the proposed method is reduced compared to the full-band matching pursuits because of lighter inner product computation due to reduction of both the resolution of the sub-band image and the basis function length of the dictionary. We attempt to improve the coding performance and computational complexity of matching pursuits for video coding.

## 2. MATCHING PURSUITS

### 2.1. Basic Principle

Matching pursuits (MP) reconstructs a signal $\hat{f}(t)$ (an approximated version of $f(t)$) using a linear combination of basis function $g_\gamma(t)$ included in overcomplete dictionary $\mathcal{D}$, as follows:

$$\hat{f}(t) = \sum_{k=1}^{m} p_k \cdot g_{\gamma_k}(t - \tau_k) \tag{1}$$

where $\tau_k$, $\gamma_k$ and $p_k$ are parameters of the basis function and denote its position, type and scale, respectively. A waveform obtained from the basis function and this parameter set is referred to as an atom.

In matching pursuits, atoms are selected by orthogonally projecting the signal $f(t)$ on basis function $g_\gamma(t)$ sequentially. First, the signal $f(t)$ is projected on the basis function $g_{\gamma_1}(t-\tau_1)$, and the signal $f(t)$ is denoted

$$f(t) = p_1 \cdot g_{\gamma_1}(t - \tau_1) + Rf(t), \tag{2}$$

$$p_1 =< f(t), g_{\gamma_1}(t - \tau_1) > . \tag{3}$$

Here, $p_1$ is the inner product of $f(t)$ and the basis function $g_{\gamma_1}(t-\tau_1)$. $Rf(t)$ is a residual signal that is orthogonal to $g_{\gamma_1}(t-\tau_1)$ and is projected on the basis function $g_{\gamma_2}(t-\tau_2)$.

If the norm of basis function $g_\gamma(t)$ is normalized to unit, then the following equation can be obtained:

$$||f(t)||^2 = p_1^2 + ||Rf(t)||^2. \tag{4}$$

**Table 1**. Parameters of the Basic Dictionary.

| $\vec{\alpha}\vec{\beta}$ | $s$ | $\xi$ | $\phi$ | $N$ | $\vec{\alpha}\vec{\beta}$ | $s$ | $\xi$ | $\phi$ | $N$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.0 | 0.0 | 0 | 1 | 10 | 5.0 | 1.0 | $\pi/2$ | 9 |
| 1 | 3.0 | 0.0 | 0 | 5 | 11 | 12.0 | 1.0 | $\pi/2$ | 21 |
| 2 | 5.0 | 0.0 | 0 | 9 | 12 | 16.0 | 1.0 | $\pi/2$ | 27 |
| 3 | 7.0 | 0.0 | 0 | 11 | 13 | 20.0 | 1.0 | $\pi/2$ | 35 |
| 4 | 9.0 | 0.0 | 0 | 15 | 14 | 4.0 | 2.0 | 0 | 7 |
| 5 | 12.0 | 0.0 | 0 | 21 | 15 | 4.0 | 3.0 | 0 | 7 |
| 6 | 14.0 | 0.0 | 0 | 23 | 16 | 8.0 | 3.0 | 0 | 13 |
| 7 | 17.0 | 0.0 | 0 | 29 | 17 | 4.0 | 4.0 | 0 | 7 |
| 8 | 20.0 | 0.0 | 0 | 35 | 18 | 4.0 | 2.0 | $\pi/4$ | 7 |
| 9 | 1.4 | 1.0 | $\pi/2$ | 3 | 19 | 4.0 | 4.0 | $\pi/4$ | 7 |

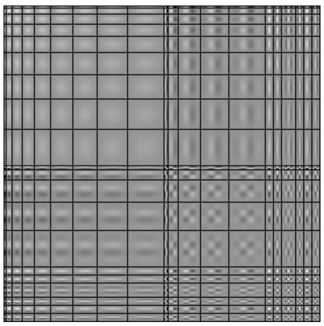

**Fig. 1**. Basic Dictionary for Matching Pursuits.

Namely, we can obtain the required parameters by maximizing the absolute value of the inner product, and the atom that minimizes the residual signal energy $\Delta e = ||f(t)||^2 - p_1^2$ is obtained. Since the residual signal $Rf(t)$ is also expanded

$$Rf(t) = p_2 \cdot g_{\gamma_2}(t - \tau_2) + R^2 f(t), \quad (5)$$

the signal $f(t)$ expanded by $m$ atoms is denoted as follows:

$$f(t) = \sum_{k=1}^{m} p_k \cdot g_{\gamma_k}(t - \tau_k) + R^m f(t). \quad (6)$$

Therefore, signal $f(t)$ can be approximately reconstructed by $m$ atoms.

### 2.2. Matching Pursuits Dictionary

The basis function of matching pursuits is constructed from the Gabor function shown in Eq. 7, and the two-dimensional separable dictionary is given as a Cartesian product of one-dimensional Gabor functions, as shown in Eq. 8:

$$g_{\vec{\gamma}}(n) = K_{\vec{\gamma}} \mathrm{g}\left(\frac{n - \frac{N}{2} + 1}{s}\right) cos\left(\frac{2\pi\xi(n - \frac{N}{2} + 1)}{N_b} + \phi\right) \quad (7)$$

$$G(i,j) = g_{\vec{\alpha}(i)} \otimes g_{\vec{\beta}(j)} \quad (8)$$
$$\otimes : Cartesian\ product$$

where $\mathrm{g}(\cdot)$ denotes a Gaussian window, $K_{\vec{\gamma}}$ denotes the normalization factor and $N$ denotes the basis function length.

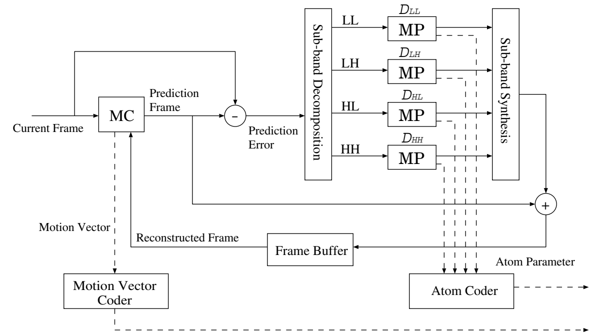


**Fig. 2**. Video Coding System using Sub-band Matching Pursuits.

These equations indicate that the Gabor function is decided by $\vec{\gamma} = (s, \xi, \phi)$. Here, $\vec{\alpha}$ and $\vec{\beta}$ denote parameter sets of horizontal and vertical components, respectively. In the present study, the parameter $N_b$ is fixed to 16.

Table 1 shows the parameter sets, and Fig. 1 shows the dictionary given by Eqs. 7 and 8 [2][4][5]. We herein refer to this dictionary as the basic dictionary. A two-dimensional separable dictionary is often used in matching pursuits because of the ease with which its inner product can be computed.

## 3. VIDEO CODING USING SUB-BAND MATCHING PURSUITS

A video coding system using sub-band matching pursuits is presented in Fig. 2. In this system, the motion compensated prediction error image is decomposed into four sub-bands, and matching pursuits is applied to each sub-band. The dictionary used for each sub-band is suitably constructed for the frequency characteristic of each sub-band.

### 3.1. Dictionary for Sub-band Matching Pursuits

The dictionary for each sub-band is generated from the basic dictionary in Fig. 1. The basis function length for a sub-band image is half of that for the corresponding original image, because sub-band images are sub-sampled along both the horizontal and vertical directions. However, in the present paper, the basis function length remains odd. The parameter $s$ is the scaling factor of the Gaussian window in Eq. 7, which is defined as follows:

$$\mathrm{g}(t) = \sqrt[4]{2} e^{-\pi t^2}. \quad (9)$$

As a result, the parameter $s$ is reduced to half of the original value for the purpose of matching the size of the decimated Gaussian window. In contrast, the other parameters, $\xi$ and $\phi$, are used at their full-band values. In addition, $N_b$ is set to 8 in order to maintain the spatial frequency resolution.

**Table 2**. Parameters of the Sub-band Dictionary.

(a) Low Frequency

| $\vec{\alpha}_L \vec{\beta}_L$ | $s$ | $\xi$ | $\phi$ | $N$ | $\vec{\alpha}_L \vec{\beta}_L$ | $s$ | $\xi$ | $\phi$ | $N$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.0 | 0.0 | 0 | 1 | 8 | 7.0 | 1.0 | $\pi/2$ | 11 |
| 1 | 1.4 | 0.0 | 0 | 3 | 9 | 8.0 | 1.0 | $\pi/2$ | 13 |
| 2 | 3.0 | 0.0 | 0 | 5 | 10 | 10.0 | 1.0 | $\pi/2$ | 17 |
| 3 | 4.0 | 0.0 | 0 | 7 | 11 | 1.4 | 3.0 | 0 | 3 |
| 4 | 7.0 | 0.0 | 0 | 11 | 12 | 4.0 | 3.0 | 0 | 7 |
| 5 | 9.0 | 0.0 | 0 | 15 | 13 | 1.4 | 4.0 | 0 | 3 |
| 6 | 10.0 | 0.0 | 0 | 17 | 14 | 1.4 | 2.0 | $\pi/4$ | 3 |
| 7 | 3.0 | 1.0 | $\pi/2$ | 5 | | | | | |

(b) High Frequency

| $\vec{\alpha}_H \vec{\beta}_H$ | $s$ | $\xi$ | $\phi$ | $N$ |
|---|---|---|---|---|
| 0 | 1.0 | 0.0 | 0 | 1 |
| 1 | 1.4 | 0.0 | 0 | 3 |
| 2 | 3.0 | 1.0 | $\pi/2$ | 5 |
| 3 | 1.4 | 3.0 | 0 | 3 |
| 4 | 4.0 | 3.0 | 0 | 7 |
| 5 | 1.4 | 4.0 | 0 | 3 |
| 6 | 1.4 | 2.0 | $\pi/4$ | 3 |

Table 2 shows the parameter sets for the low- and high-frequency sub-bands. These parameters were obtained experimentally for each sub-band based on the adaptation probability of atoms.

Decomposed sub-images have different frequency characteristics, for example, the LH band includes the horizontal edge component and the HL band includes the vertical edge component. We generate the sub-band dictionaries by Eq. 10 using the parameters of Table 2.

$$\begin{aligned} D_{LL}(i,j) &= g_{\vec{\alpha}_L(i)} \otimes g_{\vec{\beta}_L(j)}, \\ D_{LH}(i,j) &= g_{\vec{\alpha}_L(i)} \otimes g_{\vec{\beta}_H(j)}, \\ D_{HL}(i,j) &= g_{\vec{\alpha}_H(i)} \otimes g_{\vec{\beta}_L(j)}, \\ D_{HH}(i,j) &= g_{\vec{\alpha}_H(i)} \otimes g_{\vec{\beta}_H(j)}. \end{aligned} \tag{10}$$

## 4. SIMULATION AND RESULTS

The sub-band matching pursuits video coding method was examined by computer simulation. The “akiyo”, “carphone”, “foreman” and “mother and daughter” (mother) sequences (QCIF, 10 fps, 50 frames, grayscale) were used as test sequences. An eight-tap Quadrature Mirror Filter (QMF) was used to decompose images into four sub-bands. Simulations were carried out on a Pentium IV 2.2 GHz PC.

We first verified the effectiveness of the dictionary for sub-band images in matching pursuits.

Figure 3 shows the decrease in signal energy for each sub-band prediction error image with the progress of matching pursuits at the 24th frame of the “carphone” sequence, where signal energy is defined as the squared sum of the signal over the sub-band image. The band including the low-frequency components has higher energy, which decreases rapidly with coding.

**Table 3**. Ratios of Atoms Assigned to Each Sub-band.

| Number of Atoms | Atom Ratio [LL:LH:HL:HH](%) akiyo | carphone | foreman | mother |
|---|---|---|---|---|
| 100 | 62:17:20:1 | 77:11:10:2 | 91:7:1:1 | 84:7:9:0 |
| 200 | 54:19:23:4 | 68:15:14:3 | 82:13:3:2 | 71:11:16:2 |
| 300 | 51:20:24:5 | 63:17:15:5 | 76:16:5:3 | 64:14:19:3 |
| 400 | 49:20:25:6 | 59:19:16:6 | 71:18:7:4 | 59:15:21:5 |

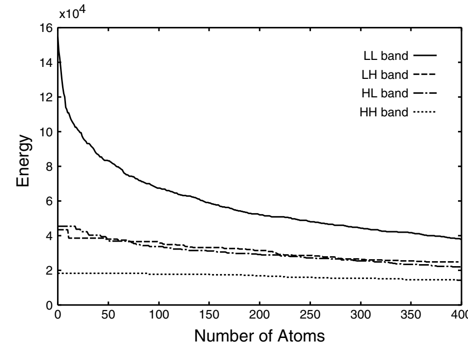


**Fig. 3**. Decrease in Signal Energy of Sub-bands (carphone, 24th frame).

Table 3 shows the ratios of atoms assigned to each sub-band. This table indicates that more than half of the atoms are assigned to the LL band and the atom ratio for high-frequency bands increases with the number of atoms. Moreover, more atoms are assigned to the LL band for images that include large motion, because a major part of the signal energy is concentrated on low-frequency bands for large-motion images. The atom assignment for sub-bands is well controlled.

Next, We compared the proposed sub-band matching pursuits with the existing full-band matching pursuits.

The searching process by the inner product of an approximated signal block and the applied atom requires large computation costs in matching pursuits. An approximated signal block generally exists in high-energy areas, because the decrease in the energy of the residual signal is maximized. Therefore, in matching pursuits, a block with the predetermined size (16×16 pixels) having the maximum signal energy is found over the full-band image in advance, and the approximated signal block is searched around this block with the predetermined size for the purpose of reduction of the computational complexity. In the proposed method, the approximated signal block is searched around the block with the predetermined size that has the maximum signal energy over all sub-band images.

Table 4 shows the average number of searching positions and calculated pixels for each atom in the searching process by inner product. The PSNR values listed in the table are average values for each sequence. The number of atoms was set to 200, and 50 frames were processed for both methods.

The PSNRs of the proposed sub-band matching pursuits are slightly superior to the current matching pursuits, indicat-

**Table 4**. Computational Complexity of Inner Product Calculation Per Atom.

| Method | Image | PSNR [dB] | Inner product calculation | |
|---|---|---|---|---|
| | | | number of positions | number of pixels [$\times 10^5$] |
| Full-band MP | akiyo | 37.7 | 217 | 186 |
| | carphone | 33.3 | 215 | 184 |
| | foreman | 30.6 | 195 | 167 |
| | mother | 35.7 | 217 | 186 |
| Sub-band MP | akiyo | 38.0 | 227 | 21 |
| | carphone | 33.6 | 224 | 24 |
| | foreman | 30.6 | 207 | 26 |
| | mother | 35.8 | 224 | 25 |

**Table 5**. Processing Times (200 atoms, 50 frames).

| Method | Image | Processing Time [sec] | | | |
|---|---|---|---|---|---|
| | | Total | MC | MP | Others |
| Full-band MP | akiyo | 155.13 | 23.65 | 130.77 | 0.71 |
| | carphone | 154.15 | 23.59 | 129.83 | 0.73 |
| | foreman | 143.30 | 23.69 | 118.96 | 0.65 |
| | mother | 155.43 | 23.72 | 130.96 | 0.75 |
| Sub-band MP | akiyo | 57.76 | 23.64 | 33.14 | 0.98 |
| | carphone | 60.37 | 23.60 | 35.78 | 0.99 |
| | foreman | 61.40 | 23.67 | 36.74 | 0.99 |
| | mother | 61.22 | 23.69 | 36.48 | 1.05 |

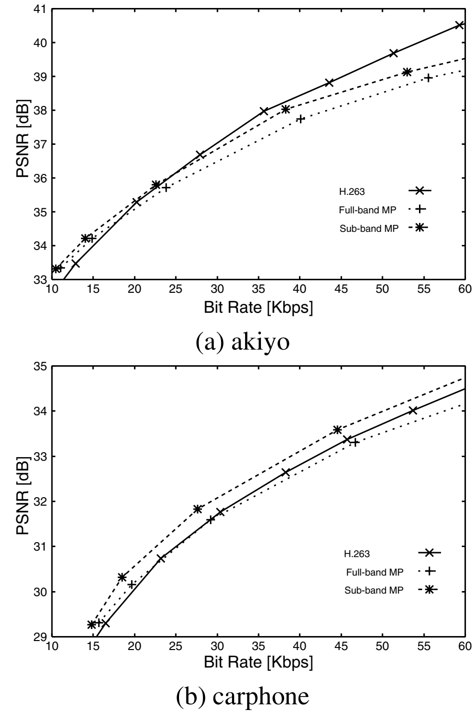


**Fig. 4**. Coding Performances of H.263, Full-band MP and Sub-band MP.

ing that the proposed method can utilize suitable dictionaries for each sub-band frequency characteristic. This table indicates that the average number of searching positions is approximately the same as that of the full-band matching pursuits, but the number of calculated pixels is decreased in the sub-band matching pursuits.

Table 5 shows the processing time of the current matching pursuits and the proposed sub-band matching pursuits. From this table, the processing time of sub-band matching pursuits is 25∼30% shorter than that of the current method.

Figure 4 shows the coding performance of the sub-band matching pursuits, the full-band matching pursuits and H.263 [7]. This figure indicates that the proposed method attained better performance than H.263 in all sequences at lower bit rates and, in some sequences, at higher bit rates.

## 5. CONCLUSION

We proposed a video coding method in which a motion compensated prediction error image is decomposed into four subbands and each sub-band image is encoded by matching pursuits. The simulation results showed that the proposed method improves both the coding performance and computational complexity for video coding.

## 6. REFERENCES

[1] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397-3415, Dec. 1993.

[2] R. Neff and A. Zakhor, "Very low bit-rate video coding based on matching pursuits," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 158-171, Feb. 1997.

[3] R. Banham and C. Brailean, "A selective update approach to matching pursuits video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 119-129, Feb. 1997.

[4] R. Neff and A. Zakhor, "Matching pursuits video coding. I. Dictionary approximation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 1, pp. 13-26, Jan. 2002.

[5] P. S.-Saugeon and A. Zakhor, "Dictionary design for matching pursuits and application to motion-comppensated video coding," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 14, no. 6, pp. 880-886, June 2004.

[6] K. Imamura, S. Kako and H. Hashimoto, "Matching Pursuits using Slant Patterns and Its Dictionary Design," *in Proc. IEEE International Symposium on Circuits and Systems (ISCAS'05)*, Kobe, Japan, May 2005, pp. 2080-2083.

[7] ITU-T Study Group 15, "Draft Recommendation H.263," Dec. 1995.