Complex-Subband Transform for Subband-Based Motion Compensation[1]

Chung-Lin Huang and Jhy-Gau Chen
Institute of Electrical Engineering
National Tsing-Hua University
Hsin-Chu, Taiwan, ROC
e-mail: clhuang@ee.nthu.edu.tw

Abstract

A new method called Complex Subband Transform(CST) is introduced for subband signals decomposition, the motion estimation, and the subband image sequence coding. There are two subband-based methods proposed which combine the process of motion compensation and the process of subband decomposition, namely out-band and in-band compensation. In this paper, we modify the M-band Perfect Reconstruction Modulation Filter (PRMF) and propose a so-called Complex Subband Transform (CST). We also derive a method using CST-based phase correlation functions to calculate motion vectors for overlapped windowed regions. In the experiments, we show that the CST-based motion estimation generates a more accurate and coherent motion field than the conventional block matching methods.

1. Introduction

Recently, there are two subband-based motion compensation methods that combine the processes of motion compensation and subband decomposition, namely out-band and in-band motion compensation. In the first case, the high resolution signal is temporally predicted prior to the subband decomposition. In the second case, the subbands themselves are temporally predicted, i.e., the prediction and compensation processes are performed on subband basis.

A limitation of block matching[1] is that it generates a significant proportion of motion vectors that do not present the accurate motion field in the scene. Irregularities in the motion field due to incorrect motion vectors may make the motion field expensive to transmit, and for applications involving the motion compensation interpolation of skipped frame, that can result in disturbing artifacts in the reconstructed images. Overlapped window motion estimation and compensation has been proposed to reduce the displaced frame difference energy compared to classical block-based motion compensation and to obtain smoother prediction images[4].

The transform-based algorithms estimate motion vector using phase correlation. Phase correlation method [5] is similar to the block matching technique in the sense that it looks for the best matching location. However, DFT-based methods perform circular rather than linear correlation, resulting in corruption of the correlation function due to the edge effect. This may be overcome by applying zero padding to the data sequences, as in fast algorithms for performing correlation using the DFT [6], but the computational load is increased.

Young and Kingsbury [7] first defined the complex lapped transform (CLT) by extending the Lapped orthogonal transform (LOT)[8] to have complex basis functions. The CLT basis functions decay smoothly that are used for calculating vectors for overlapped windowed regions of data. In this paper, we apply the CLT concept on PRMF (perfect reconstruction modulated filterbank) filter [2] and make output subband signals be complex values. It is called *complex subband transform* (CST). Different from the CLT, the CST can be applied not only for in-band/out-band motion estimation, but also for subband signal decomposition. The overlapped, windowed nature of CST leads to smaller errors due to edge effects when the transform is used to estimate cross-correlation functions. Thus it can be used to estimate local motion vectors, without the need for zero-padding of the data. In the experiments, we illustrates that using CST for motion estimation generate a more coherent motion vector field.

2. Phase Correlation for Motion Estimation

For each individual band, we can not have the same motion field because of the aliasing effect of each decomposed subband. An alternative to spatial-domain block-matching methods is to estimate cross-correlation functions via the complex subband domain. It is common to normalize the frequency spectrum of the signal to give a phase correlation. Phase correlation is used in the complex subband domain to estimate motion vector.

2.1 Complex Subband Transform (CST)

Here, we modify the PRMF filter bank [2] and develop a complex subband decomposition. We can directly convert the complex subband to the original PRMF subband. We modify PRMF by a spatial shifting of M-1. The mth analysis filter becomes a non-causal filter as

$$f_m[n] = f_{1,m}[n+M-1] = \frac{\sqrt{2}}{M}\cos\left(\frac{n\pi}{2M}\right)\exp\left(\frac{-j(2m+1)(2n-M)\pi}{2M}\right) \quad (1)$$

[1] This work is finacially supported by NSC 84-2213-E-007-047.

where $-(M-1) \le t \le M-1$, and $0 \le m \le M-1$. We temporally ignore the phase shift in equation (1) (*i.e.*, $\exp(j(2m+1)\pi/4)$), and have equation (2). The phase shift is required only when the complex-subbands are converted to real-subbands (*i.e.*, original subbands from PRMF). In that case, we need to multiple complex-subband with this phase shift. In the following, we develop the modified mth complex analysis filter as

$$h_m[t] = \frac{\sqrt{2}}{M}\cos(\frac{t\pi}{2M})\exp(\frac{-j(2m+1)t\pi}{2M}), \quad (2)$$

where $-(M-1) \le t \le M-1$. So complex subband decomposition for the mth band $X_m[n]$ can be defined as

$$X_m(n) = \frac{\sqrt{2}}{M}\sum_{n=-(M-1)}^{(M-1)} x(n+t)\cos(\frac{t\pi}{2M})\exp(\frac{-j(2m+1)t\pi}{2M}) \quad (3)$$
$$= CST\{x[n+t]\},$$

where $m = 0,\ldots,M-1$, $n = 0,\ldots,N-1$, N indicates the size of signal. Similar to the complex lapped transform (CLT) [7], we do the complex subband transform via the modified PRMF concept and let the analysis filter become a non-causal filter. We may call this transformation a complex subband transform (CST). The modulated input signal can be obtained by inverse transform from M subbands signals $X_m[n]$ and $X_m[n+M]$ as

$$x[n+t]\cos(\frac{t\pi}{2M}) = \frac{1}{2\sqrt{2}}\mathrm{Re}\{\sum_{m=0}^{M-1} X_m[n]\exp(\frac{j(2m+1)t\pi}{2M})\} \quad (4)$$

where $m=0,\ldots M-1$. We shift n by M (*i.e.*, $n = n+M$) and then module t by M (*i.e.*, $t = t-M$).

$$x(n+t)\cos(\frac{(t-M)\pi}{2M}) = \frac{1}{2\sqrt{2}}\mathrm{Re}\{\sum_{m=0}^{M-1} X_m(n+M)\exp(\frac{j(2m+1)(t-M)\pi}{2M})\} \quad (5)$$

where $0 \le t \le M-1$. Equations (4) and (5) can be proved to have the orthogonal properties of the complex exponential term. We now derive an inverse transform to generate $x[n+t]$ from $X_m[n]$ and $X_m[n+M]$. We first decompose $x[n+t]$ by

$$x[n+t] = x[n+t]\cos^2(\frac{t\pi}{2M}) + x[n+t]\sin^2(\frac{t\pi}{2M})$$
$$= \left(x[n+t]\cos(\frac{t\pi}{2M})\right)\cos(\frac{t\pi}{2M}) + \left(x[n+t]\cos(\frac{t-M}{2M}\pi)\right)\cos(\frac{t-M}{2M}\pi) \quad (6)$$

where $0 \le t \le M-1$. Substituting equations (4) and (5) into equation (6) gives

$$x[n+t] = \frac{1}{2\sqrt{2}}\mathrm{Re}\{\sum_{m=0}^{M-1} X_m[n]\cos(\frac{t\pi}{2M})\exp(\frac{j(2m+1)t\pi}{2M}) + \sum_{m=0}^{M-1} X_m[n+M]\cos(\frac{(t-M)\pi}{2M})\exp(\frac{j(2m+1)(t-M)\pi}{2M})\} \quad (7)$$
$$= \frac{1}{2\sqrt{2}}\mathrm{Re}\left[ICST\{X_m[n]\} + ICST\{X_m[n+M]\}\right]$$

where $0 \le t \le M-1$. This method of inversion (i.e., Figure 1) is equivalent to the weighted overlapped add technique [3].

2.2 Phase-correlation Function

Suppose we have two 1-D input sequences $x_1[n]$ and $x_2[n]$. We may define a cross-correlation function, $R_{x1,x2}[k]$ as

$$R_{x_1,x_2}[k] = \sum_{t=-(M-1)}^{M-1} x_1[n+t]\cos(\frac{t\pi}{2M})x_2[n+k+t]\cos(\frac{t\pi}{2M}) \quad (8)$$

where $0<k<M$. It can be rewritten as

$$R_{x_1,x_2}[k] = \frac{M}{4}\mathrm{Re}\left\{\sum_{m=0}^{M-1} X_{1,m}[n]X^*_{2,m}[n+k]\right\} = \frac{M}{4}\mathrm{Re}\left\{\sum_{m=0}^{M-1} X^*_{1,m}[n]X_{2,m}[n+k]\right\} \quad (9)$$

In DFT based methods for motion estimation, it is common to compute the phase correlation function[5,6]. The phase correlation is obtained by normalizing the magnitudes of the transform coefficients, prior to computing the correlation in the frequency domain. Similarly, the phase correlation between windowed data sequences can be computed approximately in the complex subbands by normalizing the complex subbands. The normalizing complex subbands are given by

$$X^p_m[n] = \frac{X_m[n]}{|X_m[n]| + \alpha} \quad (10)$$

where α is small constant. This avoids amplifying spectral components with a low signal to noise ratio by an excessively high gain, and tends to suppress any peaks in the phase correlation function.

2.3 Interpolation in the complex subband domain

From equation (9), we need to solve the problem of interpolating the decimated $X_{2,m}[n+k]$ between two complex subband signals $X_{2,m}[n]$ and $X_{2,m}[n+M]$. This can be achieved exactly if each interpolated coefficient is allowed to be a function of all the complex subbands coefficients. To simplify our computation, we only use coefficients from the same frequency subband and ignore the aliasing between the different frequency subbands. Although aliasing distortion does occur, the two properties of the complex subband (*i.e.*, the smoothly decaying half-cosine window function and the overlapped nature) are both beneficial for obtaining accurate interpolation. We can express $X_{2,m}[n+k]$ by

$$X_{2,m}[n+k] \cong \left(X_{2,m}[n]\cos(\frac{k\pi}{2M}) - j(-1)^m X_{2,m}[n+M]\sin(\frac{k\pi}{2M})\right)\exp(\frac{j(2m+1)k\pi}{2M}) \quad (11)$$

Substituting equation (11) into equation (9), we have equation (12) as

$$R_{x_1,x_2}[k] \cong \frac{M}{4}\mathrm{Re}\left\{\sum_{m=0}^{M-1} X^*_{1,m}[n]\, X_{2,m}[n]\cos(\frac{k\pi}{2M})\exp(\frac{j(2m+1)k\pi}{2M}) + \sum_{m=0}^{M-1} X^*_{1,m}[n]\, X_{2,m}[n+M]\cos(\frac{k-M}{2M}\pi)\exp(\frac{j(2m+1)(k-M)}{2M}\pi)\right\} \quad (12)$$

Similar to equation (13), $Rx_1,x_2[k]$ could be approximately by

$$R_{x_1,x_2}[k] \cong \frac{M}{\sqrt{2}}\mathrm{Re}\left(ICST\{X^*_{1,m}[n]X_{2,m}[n]\} + ICST\{X^*_{1,m}[n]X_{2,m}[n+M]\}\right) \quad (13)$$

where $0 \le k \le M$.

2.4. Motion Estimation Algorithm

The algorithm for performing 1-D phase correlation, may be easily extended to 2-D. The algorithm is illustrated in Figure 2 and the procedures are described as follows:

1. Calculate the complex subbands $\{A_{m1,m2}(x, y)\}$ and $\{B_{m1,m2}(x, y)\}$, where $m1=0,\ldots M-1$ and $m2=0,\ldots 2M-1$, of

the previous frames {a(x, y)} and current frame (b(x, y)} respectively.

2. For each complex subband pel $B_{m1,m2}(x, y)$ in the current frame, calculate the element by element product of the conjugate of its 2MxM complex subbands with the pixel in the corresponding position and the eight nearest neighbor pixels $A_{m1,m2}(x+i, y+j)$, i , j = -1, 0, 1.
3. From the nine 2MxM product matrices $M_{i,j}$= {[$A_{m1,m2}$(x-i, y-j)$B_{m1,m2}$(x,y)] | m1=0,..,M-1, m2 = 0,...,2M-1}, where i,j=-1,0,1 which form four overlapping 4Mx2M matrices, calculate the inverse transform of each matrices to give four MxM matrices, {[$R_{A,B}$(k1,k2)] | k1=-M, ...M, and k2=-M,...M}. These matrices represent the four extending ±M pels horizontally and vertically around the current pel.
4. Find the position of the peak in the spatial correlation function which provides the motion vector for the pixel b(x,y).
5. If many local peaks are found, then adjustment is required, We only consider peaks with amplitude greater than 0.8 of the largest peak, then measure the mean square error (MSE) in the spatial domain over cosine-window of size (2M-1)x(2M-1). The vector giving the lowest error is selected for the motion vector for the pel b(x,y).

The MSE is measured over 2M x 2M windowed data block, and a cosine window is applied which is defined as

$$w(n_1,n_2) = \cos^2\left(\frac{n_1\pi}{2M}\right)\cos^2\left(\frac{n_2\pi}{2M}\right) \qquad (14)$$

where n_1, n_2 = -(M-1),...,(M-1). For most correlation function blocks, there is a single global peak, so adjustment usually is not required.

2.5 Fast algorithm for computing phase correlation function

In the following, we derive a fast method for equation (3) to speed up motion estimation algorithm based on FFT as

$$X_m[n] = \frac{\sqrt{2}}{M} FFT\{y_1[t_1 - M/2]\}\exp(\frac{jm(M-1)\pi}{M}) + \frac{\sqrt{2}}{M} FFT\{y_2[t_2 - M/2]\}\exp(jm\pi) \qquad (15)$$

Next, we can also derive a fast algorithm for ICST. Here, we also consider the 1-D case only. First, the ICST can be considered by dividing the sequence (equation (4) with t=-(M-1),.. ,(M-1)) into two subsequences z_1[n+t1] and z_2[n+t2] as

$$z_1[n+t1] = x[n+2t1-M]\cos((2t1-M)\pi/2M),\ t1 = 0, \dots, (M-1), \qquad (16)$$

$$z_2[n+t2]=x[n+2t2+1-M]\cos((2t2+1-M)\pi/2M),\ t2 = 0, \dots, (M-1). \qquad (17)$$

We let t=2t_1 -M and have z1[n+t1] be implement by IFFT as follows:

$$z_1[n+t_1] = \frac{1}{2\sqrt{2}} \mathrm{Re}\left\{ IFFT[X_m[n]\exp(\frac{-j(2m+1)\pi}{2})]\exp(\frac{jt_1\pi}{M})\right\} \qquad (18)$$

Similarly, z2[n+t2] can also be derived as

$$z_2[n+t_2] = \frac{1}{2\sqrt{2}} \mathrm{Re}\left\{ IFFT\left[X_m[n]\exp(\frac{j(2m+1)(1-M)\pi}{2M})\right]\exp(\frac{jt_2\pi}{M})\right\} \qquad (19)$$

z1[n+t] and z2[n+t] can be implement by M-point IFFT. Figure 3 illustrates the Fast ICST.

3. Experimental Results

The CST phase-correlation motion estimation has a better result than block-matching motion estimation. Irregularities in the motion field due to incorrect vectors may make the motion field expansive to transmit. The motion field estimated by CST phase-correlation method requires less bits to transmit than the motion field obtained by block-matching method.

Table 1 shows the prediction error in 16 subbands of 23-24th frame of 288x352 salesman sequence. We can find that the in-band motion compensation with overlapped window is better than out-band motion compensation without overlapped window. Out-band motion compensation with overlapped window gives the best results. Table 1 shows the measured prediction mean square error in the subbands of the 23th frame when a 288x352 salesman sequence is predicted *(i.e.,* no coding). The in-band subbands still contain aliasing term which has to be canceled out by other subbands. We use equation (4) to interpolate those decimated subband pixels during downsampling, so the aliasing error can be reduced. It can be observed that the accuracy of the in-band window overlapped prediction is comparable with the out-band window overlapped prediction. By comparing the 4th column and the 5th column of Table 1, we can observe that window overlapped scheme can reduce the reconstruction error.

4. Conclusion

We develop a CST to estimate motion field using phase-correlation functions in the complex subband domain and estimating motion vectors for overlapping, windowed regions of image. This algorithm is invariant to illumination changes which is a common problem in estimating true motion vectors in image sequences. By comparing the performances of the phase-correlation algorithm and the block-matching algorithm for the salesman test sequence, the phase-correlation algorithm generate a smoother motion fields than block-matching. The CST can also be used for subband signal decomposition. We may apply CST for inband/outband motion compensation with/without overlapped windows.

References

[1] H. G. Musmann, P. Prisch, and H. Grallert, " Advances in picture coding ",Proc . IEEE , 73(4) , pp. 523 - 548, Apr. 1985 .

[2] J. Mau, " Perfect reconstruction modulation filter banks " , IEEE Conf. on ASSP , pp. 273 - 276 ,1992 .

[3] R. E. Crochiere,"A weighted overlapped-add method of short-time Fourier analysis/synthesis,"IIEEE trans. ASSP, vol. ASSP-28 , pp. 99-102 , Feb. ,1980.

[4] H. Watanabe and S. Singhal,"Windowed motion compensation"" Proc. of SPIE Visual Commun. Image Proc., vol. 1605 , pp. 582 - 589 ,1991.

[5] C. D. Kuglin and D. C. Hines," The phase correlation image alignment method ", Proc. of IEEE int. Conf. on Cybernetics and Society , pp. 163-165 , 1975 .

[6] G. A. Thomas, " Television motion measurement for DATV and another applications", BBC Res. Dep. Rep. 11 , 1987 .
[7] R. W. Young and N. G. Kingsbury, " Frequency-domain motion estimation using a complex lapped transform ", IEEE trans. on Image Proc., pp.2-17 , vol. 2 , No. 1 , Jan. 1993.
[8] H. S. Malver and D. H. Staelin," The LOT: Transform coding without Blocking Effect", IEEE Trans. on ASSP, vol. 37,no. 4. pp. 553-559, April 1989.

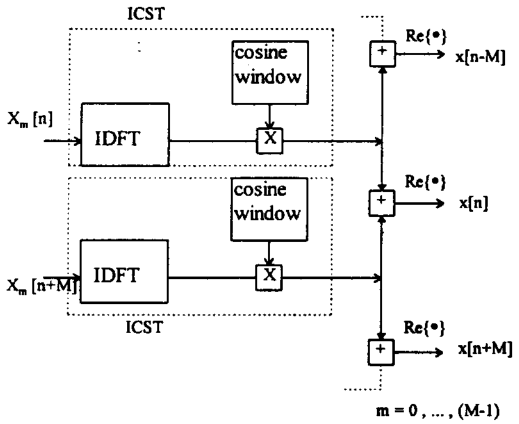

Figure 1. Complex Subband Reconstruction

Table 1 . Prediction error (mean square error)

subband no.	out-band non-w	out-band with-w	in-band non-w	in-band with-w
0	2.161	2.188	4.152	2.132
1	1.194	1.134	2.791	1.130
2	0.676	0.520	1.234	0.450
3	0.468	0.337	0.540	0.470
4	0.936	0.911	1.231	1.011
5	0.571	0.548	0.789	0.431
6	0.308	0.241	0.532	0.237
7	0.168	0.140	0.190	0.158
8	0.637	0.522	0.345	0.598
9	0.374	0.309	0.467	0.356
10	0.168	0.115	0.157	0.123
11	0.099	0.076	0.115	0.085
12	0.487	0.417	0.414	0.456
13	0.190	0.161	0.180	0.173
14	0.095	0.071	0.109	0.089
15	0.098	0.078	0.034	0.076
total	8.685	7.806	13.280	8.212

non-w : motion compensation without overlapped window.
with-w : motion compensation with overlapped window.

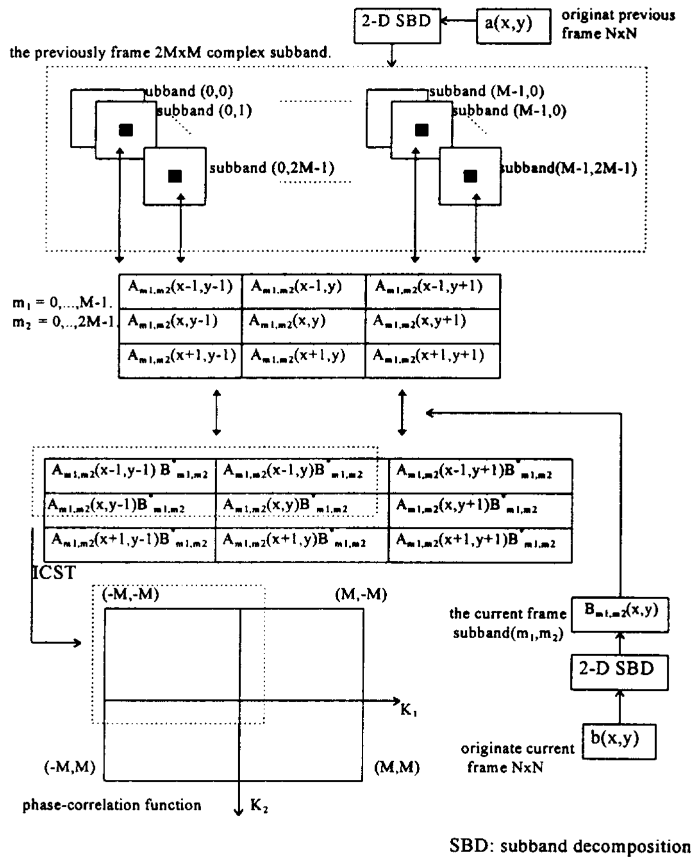

Figure 2. Schematic of the correlation function

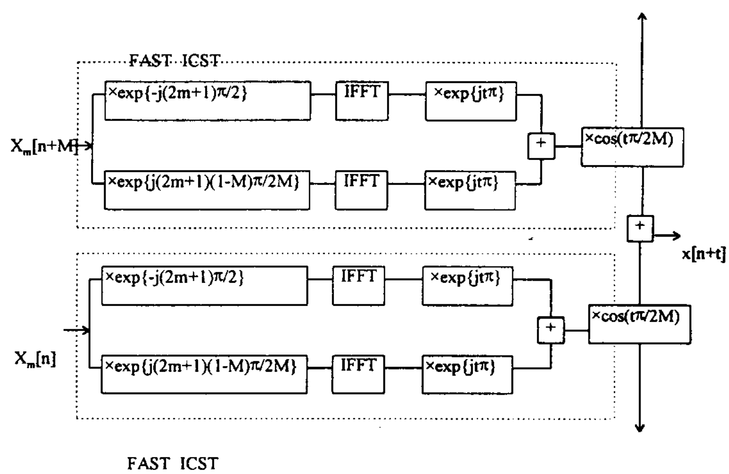

Figure 3. Fast Inverse Complex Subband Transform (ICST)