MODIFIED CLT-DOMAIN MOTION ESTIMATION

L.Y.BI,K.P.CHAN and Y.L.YU

DEPARTMENT OF ELECTRONIC ENGINEERING, SOUTH CHINA UNIVERSITY OF TECHNOLOGY,GUANGZHOU, PRC.
DEPARTMENT OF COMPUTERS SCIENCE AND INFORMATION SYSTEMS, UNIVERSITY OF HONG KONG

ABSTRACT

In this paper, we have proposed a modified complex lapped transform domain motion estimation (MCLT-ME) algorithm to estimate motion of video sequences. The proposed algorithm is based on phase correlation of current window region and the corresponding window region, and provides a better performance in motion estimation than CLT-ME, full search and three-step search algorithm.

1. INTRODUCTION

Motion estimation is important in video coding such as MPEG-1, MPEG-2, MPEG-4 and H.263. Block matching methods [1] is the common motion estimation method, which is based on searching the best candidate block of current block in search window of reference frame. However, it needs more sequential steps and produces sub-optimal estimate and irregularities in the motion field due to incorrect vector. An alternative method is to use phase correlation method in the transform domain. It can reduce the computational complexity, because it directly searches the motion vector from the phase correlation map. And phase correlation [2] is defined as:

$$\frac{F_1(\xi,\eta).F_2^*(\xi,\eta)}{\left|F_1(\xi,\eta).F_2^*(\xi,\eta)\right|} = e^{j2\pi(\xi x_0+\eta y_0)} \quad (1)$$

where $F_1(\xi,\eta)$ and $F_2(\xi,\eta)$ is frequency transform of two images $f_1(x,y)$ and $f_2(x,y)$ respectively, and the asterisk denotes the complex conjugate operation. The inverse frequency domain of Eqn.1 results in a phase correlation surface, and the location of peak is the displacement of consecutive images. Because $F_1(\xi,\eta)F_2^*(\xi,\eta)$ is normalized with phase correlation magnitude, the method is relatively insensitive to changes in illumination.

Phase correlation was initially used only with Fourier transform [1][2], and was late extended to the complex orthogonal transform domain motion estimation (CLT-ME) [3][4][5]. It can produce smooth motion field, however, because of approximation, it can't give correct motion vector and have a heavy computational load, as will be discussed in the next section. In this paper, we propose a modified complex orthogonal transform with a low computational load. Experiments have showed the proposed method has better performance than CLT-ME and block matching methods.

2. CLT-ME algorithm

In the complex orthogonal transform domain, CLT-ME algorithm [3][4] is based on the correlation between the $2N \times 2N$ current window region and window region of a displacement value k in the reference frame. Because the transform is separable, 2-D correlation can be easily obtained from the 1-D spatial domain. A modified correlation algorithm in 1-D spatial domain is given as follows:

$$R_{x_1x_2}(k) = \sum_{n=-(N-\frac{1}{2})}^{N-\frac{1}{2}} (x_1(n)\cos\frac{n\pi}{2N})(x_2(n+k)\cos\frac{n\pi}{2N}) \quad (2)$$

The modified correlation is a little different from the original correlation. It considers not only the correlation of current block, but also the correlation of neighboring block. Because the product of the current block is more important to the correlation than the neighboring block, so the cosine function is used as weight function.

And the standard CLT is:

$$X(m) = \frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})}^{(N-\frac{1}{2})} x(n) \exp\frac{-j(2m+1)n\pi}{2N} \cos\frac{n\pi}{2N}$$

$$m = 0,\ldots, N-1$$

The extended CLT can be obtained from the standard CLT by expanding m from $0,\ldots N-1$ to $-N,\ldots N-1$. This is necessary because the coefficients generated by the standard CLT are complex, and so do not exhibit spectral symmetry [3] . The extended CLT is defined as:

$$X(m) = \frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})}^{(N-\frac{1}{2})} x(n) \exp \frac{-j(2m+1)n\pi}{2N} \cos \frac{n\pi}{2N}$$

$m = -N, \ldots, N-1$

CLT-ME algorithm is based on the half-cosine windowed block, so it can alleviate the errors due to edge effects and produce a smoother motion field. However it has two disadvantages. First, approximation of replacing

$$\frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})+k}^{(N-\frac{1}{2})+k} x(n) \exp \frac{-j(2m+1)n\pi}{2N} \cos \frac{n\pi}{2N}$$

and

$$\frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})-(N-k)}^{(N-\frac{1}{2})-(N-k)} x(n+N) \exp \frac{-j(2m+1)n\pi}{2N} \cos \frac{n\pi}{2N}$$

by

$$\frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})}^{(N-\frac{1}{2})} x(n) \exp \frac{-j(2m+1)n\pi}{2N} \cos \frac{n\pi}{2N}$$

and

$$\frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})}^{(N-\frac{1}{2})} x(n+N) \exp \frac{-j(2m+1)n\pi}{2N} \cos \frac{n\pi}{2N}$$

, can not keep being accurate as the displacement k increases. Hence, correct motion vector cannot be produced when displacement k is close to the value of $\pm(N-1/2)$. Second, to compute the motion vector of one $N \times N$ block, a $2N \times 2N$ window region in the current frame and the nine nearest $2N \times 2N$ neighbor overlap window region in reference frame, which are centered at each $N \times N$ block, has to be used, so the computation load is high.

In the next section, we will propose a modified CLT-ME method that does not require the approximation as above, and with much lower computational requirement.

3. PROPOSED MCLT-ME ALGORITHM

The proposed MCLT-ME algorithm is based on a correlation between $2N \times 2N$ current window region and the corresponding window region in the reference frame. The 1-D correlation is obtained by modifying Eqn.1 to follows,

$$R_{x_1x_2}(k) = \sum_{n=-(N-\frac{1}{2})}^{N-\frac{1}{2}} \left(x_1(n) \cos \frac{n\pi}{2N}\right)\left(x_2(n+k) \cos \frac{(n+k)\pi}{2N}\right) \quad (3)$$

From the standard complex lapped transform [3] ,

$$x_1(n) \cos \frac{n\pi}{2N} = \frac{1}{\sqrt{N}} \mathrm{Re}\left\{\sum_{m=0}^{N-1} X_1(m) \exp\left(\frac{j(2m+1)n\pi}{2N}\right)\right\}$$

$n = -(N - \frac{1}{2}), \ldots, (N - \frac{1}{2})$ (4)

By substituting (4) into (3),

$$R_{x_1x_2}(k) = \mathrm{Re}\left\{\sum_{m=0}^{N-1} X_1(m) \left[\frac{1}{\sqrt{N}} \sum_{n=-(N-\frac{1}{2})}^{(N-\frac{1}{2})} x_2(n+k) \exp \frac{j(2m+1)(n+k)\pi}{2N} \cos \frac{(n+k)\pi}{2N}\right] \exp \frac{-j(2m+1)k\pi}{2N}\right\}$$

So,

$$R_{x_1x_2}(k) = \mathrm{Re}\left\{\sum_{m=0}^{N-1} X_1(m) X_2^*(m) \exp \frac{-j(2m+1)k\pi}{2N}\right\}$$

$k = -N, \ldots, N-1$ (5)

The phase correlation [2] between the two consecutive images differing by a simple shift is an impulse at displacement value, which can give better displacement estimation than correlation. Hence, we also use the phase correlation, rather than correlation in CLT domain. In other words, $X_1(m) X_2^*(m)$ can be replaced by $\frac{X_1(m) X_2^*(m)}{|X_1(m) X_2^*(m)| + \beta}$ in the Eqn. 5, where β is a small constant.

Similar result can be applied to the extended CLT transform, where only the Re operation is not needed. So, the 2-D phase correlation that MCLT-ME algorithm based on is defined as:

$$R_{x_1x_2}(k,l) = \mathrm{Re}\left\{\sum_{m_1=0}^{N-1} \sum_{m_2=-N)}^{N-1} \frac{X_1(m_1,m_2) X_2^*(m_1,m_2)}{|X_1(m_1,m_2) X_2^*(m_1,m_2)| + \beta} \exp \frac{-j(2m_1+1)k\pi}{2N} \exp \frac{-j(2m_2+1)l\pi}{2N}\right\}$$

$k = -N, \ldots, N-1, l = -N, \ldots, N-1$ (6)

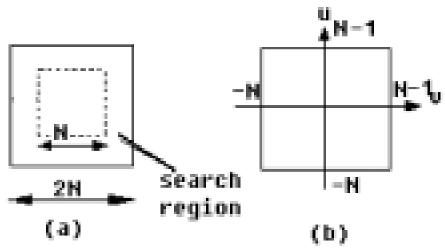


Figure 1 (a) $2N \times 2N$ search region (b) vector range.

So, the proposed MCLT -ME algorithm is described as below:

First, to every $N \times N$ block in the current frame, we choose a window region with $2N \times 2N$ size, which is centered at the block in each current frame and corresponding reference frame as Fig.1 (a).

Second, We use the 2-D CLT forward transform to get two $2N \times N$ complex coefficients of two consecutive frames, then get the corresponding conjugate coefficients of the reference frame, and find the product of the two coefficients and normalize the result.

Finally, According to the 2-D CLT phase correlation in Eqn.6, we can get the correlation surface in spatial domain. The displacement of the block can be obtained by locating the peak of the correlation surface. The displacement range is $[-N, N-1]$ as Fig.1 (b).

4. EXPERIMENTS AND RESULTS

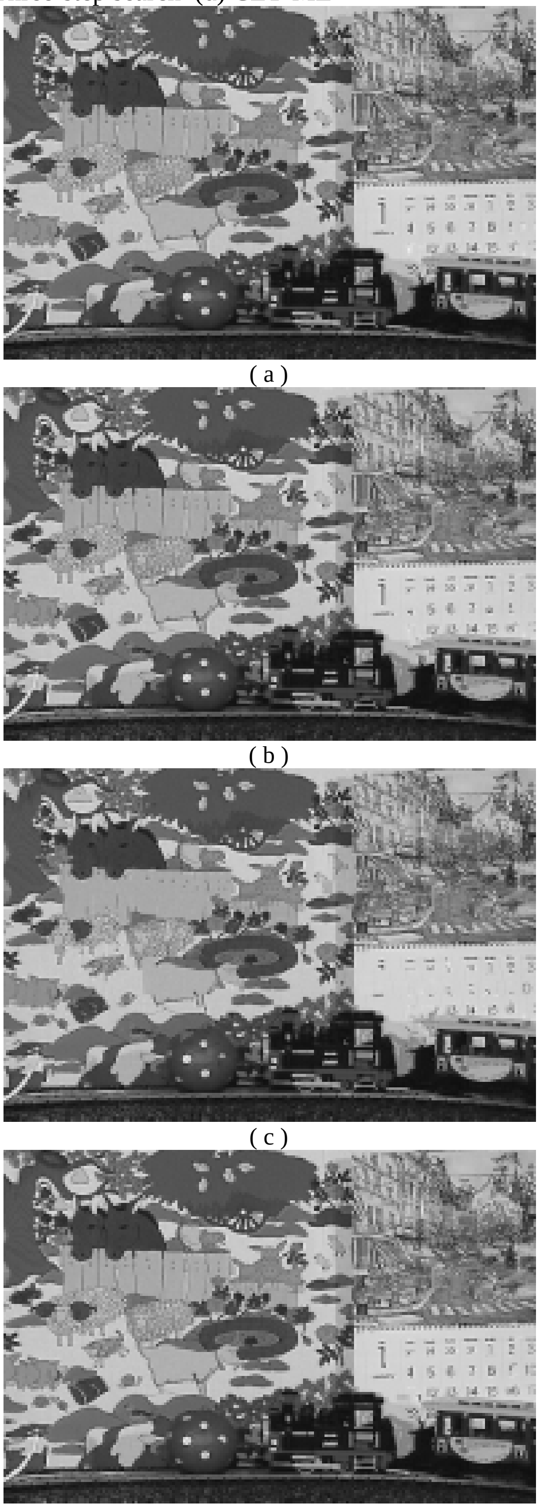
The MCLT-ME, CLT-ME, full search (FS) and three step search (TSS) [1] algorithms have been tested on the salesman, calendar and football sequence. Results obtained from these algorithms are shown in Table 1. The experiments were performed on a PC of 1.7GHZ Pentium IV cpu, 256M memory and matlab version 6.1.we use a block size of 8×8 and motion estimation is full pixel accurancy.

Table 1. A comparison of the PSNR, entropy of motion and computation between MCLT-ME, full search, three-step search and CLT-ME

	PSNR (dB)	Entropy of motion	Computational time(Second)
Salesman			
MCLT-ME	37.4687	1.7518	2.0470
FS	37.5058	2.1342	19.5160
TSS	37.2656	2.3656	3.7030
CLT-ME	37.0931	1.4715	5.8590
Calendar			
MCLT-ME	21.6967	1.8590	1.5940
FS	21.8228	2.4363	15.9070
TSS	19.9430	4.2917	3.0780
CLT-ME	20.9626	1.8385	4.6880
Football			
MCLT-ME	26.5169	3.5803	1.6720
FS	26.9085	3.8524	15.6560
TSS	25.6136	4.0090	3.0160
CLT-ME	25.5560	3.2932	4.8440

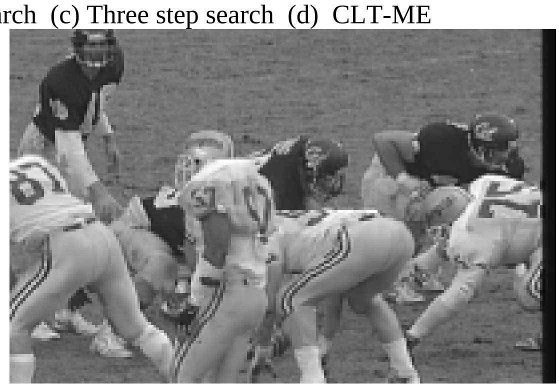
Through the comparison of PSNR, we can find the MCLT-ME give a comparable objective result with full search and a better result than other methods. Experiments on "football" sequence show that MCLT-ME gives significant improvement than the CLT-ME, because MCLT-ME can track faster motion vector than the CLT-ME. Through the comparison of entropy, we can find that MCLT-ME give a more homogeneous motion field and low entropy of motion so that it can reduce bit rates that are used to code motion field. The computational complexity of MCLT-ME is lower than other methods. As the result shown in the Table 1, the computational time of MCLT-ME is lower than other methods. From Fig.2 and Fig.3, we can find that the MCLT-ME has a good subject opinion on the image than other methods except full search method. In Fig.2, You can see the apparent line on barrier, clear contour of the sheep and more numbers on the calendar in the MCLT-ME, full search and CLT-ME methods. In Fig.3, you can find smooth line on the football player's trousers in MCLT-ME, full search and three step search methods, but there are some apparent block artifacts on the trousers in three-step search method.

Figure 2. The comparison of motion compensation image for the calendar sequence using different methods. Motion compensation image using (a) MCLT-ME (b) full search (c) Three-step search (d) CLT-ME

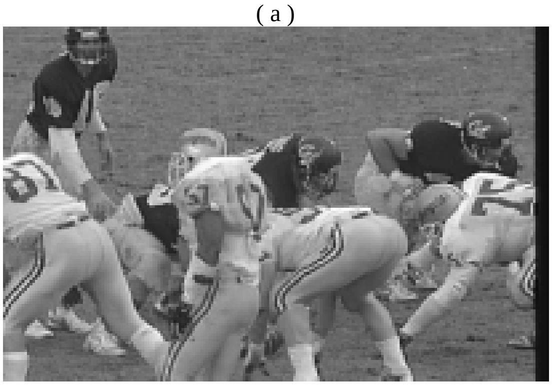


(a) (b) (c) (d)

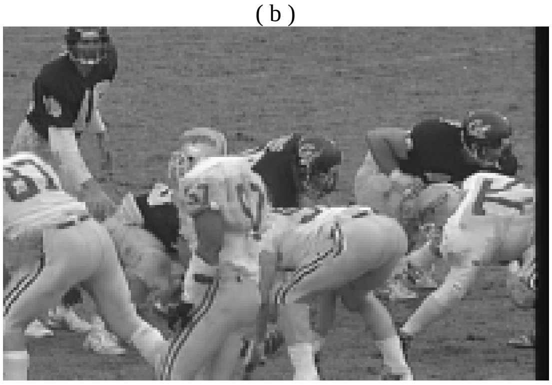
Figure 3. The comparison of motion compensation image for the football sequence using different methods. Motion compensation image using (a) MCLT-ME (b) full search (c) Three step search (d) CLT-ME



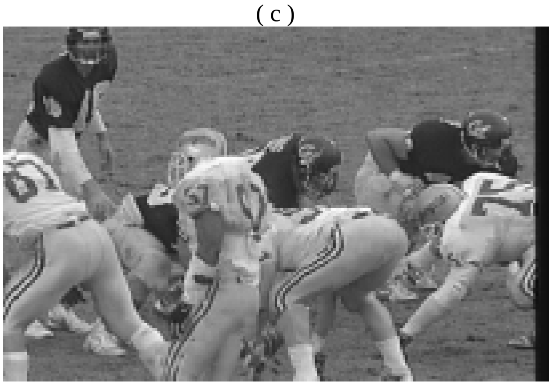
(a)



(b)



(c)



(d)

6. CONCLUSION

In this paper, we proposed a MCLT-ME algorithm. The MCLT-ME algorithm was based on a phase correlation between $2N \times 2N$ current window region and the corresponding window region in the reference frames. It is found that the MCLT-ME produced a higher PSNR and smoother vector than block matching methods, because of the use of the half-cosine window region, which cause the reducing of the block effect, and computational time is reduced compare to CLT-ME algorithm.

5. REFERENCES

[1] A.Murat Tekalp, *Digital Video Processin*g, Prentice Hall 1995.

[2] G.A.Thomas, “Television motion measurement for DATV and other application”, *BBC Research Department Report.* Volume 11,1987.

[3] R.W.Yong and N.G.Kingbury, “Frequency-domain motion estimation using a complex lapped transform”, *IEEE trans.Image Processing*, vol.2, pp2-17,Jan.1993.

[4] H.S.Malvar and D.H.Staelin, “The LOT: Transform Coding With Blocking Effects,” *IEEE trans.Acoustics,Speech,and Signal Processing*, vol37, April. 1989.

[5] Zixiang Xiong and H.S.Malvar, “ A nonuniform modulated complex lapped transform”, *IEEE Signal Processing Letters*, Volume: 8 Issue: 9 , Sept. 2001 Page(s): 257 –260.