

## Research Article

# A Robust Subpixel Motion Estimation Algorithm Using HOS in the Parametric Domain

**E. M. Ismaili Aalaoui,<sup>1,2</sup> E. Ibn-Elhaj,<sup>2</sup> and E. H. Bouyakhf<sup>1</sup>**

<sup>1</sup> *Faculté des Sciences de Rabat, Université Mohammed V Agdal, 4 avenue Ibn Battouta, B.P. 1014 RP, 10000 Rabat, Morocco*

<sup>2</sup> *Department of Telecommunication, Institut National des Postes et Télécommunication, avenue Allal Al Fassi, Madinat El Irifane, 10000 Rabat, Morocco*

Correspondence should be addressed to E. M. Ismaili Aalaoui, ismailimehdi@gmail.com

Received 28 April 2008; Accepted 24 October 2008

Recommended by Simon Lucey

Motion estimation techniques are widely used in today's video processing systems. The most frequently used techniques are the optical flow method and phase correlation method. The vast majority of these algorithms consider noise-free data. Thus, in the case of the image sequences are severely corrupted by additive Gaussian (perhaps non-Gaussian) noises of unknown covariance, the classical techniques will fail to work because they will also estimate the noise spatial correlation. In this paper, we have studied this topic from a viewpoint different from the above to explore the fundamental limits in image motion estimation. Our scheme is based on subpixel motion estimation algorithm using bispectrum in the parametric domain. The motion vector of a moving object is estimated by solving linear equations involving third-order hologram and the matrix containing Dirac delta function. Simulation results are presented and compared to the optical flow and phase correlation algorithms; this approach provides more reliable displacement estimates particularly for complex noisy image sequences. In our simulation, we used the database freely available on the web.

Copyright © 2009 E. M. Ismaili Aalaoui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

The importance of image sequence processing is constantly growing with the ever increasing use of television and video systems in consumer, commercial, medical, and scientific applications. Image sequences can be acquired by film-based motion picture cameras or electronic video cameras. In either case, there are several factors related to imaging sensor limitations that contribute to the graininess (noise) of resulting images. Electronic sensor noise and film grain are among these factors [1]. In many cases, graininess may result in visually disturbing degradation of the image quality, or it may mask important image information. Even if the noise may not be perceived at full-speed video due to the temporal masking effect of the eye, it often leads to unacceptable single-frame hardcopies and to poor-quality freeze-frames that adversely affect the performance of subsequent image analysis [2].

The motion estimation process must be able to track objects within a noisy source. In a noisy source, objects appear to change from frame to frame because of the noise, not necessarily as the result of object motion [3]. Tracking objects within a noisy environment is difficult, especially if the image frames are severely corrupted by additive Gaussian noises of unknown covariance; second-order statistics methods do not work well.

Higher-order statistics (HOS) in general and the bispectrum (order 3) in particular have recently been widely used as an important tool for signal processing. The classical methods based on the power spectrum are now being effectively superseded by the bispectral ones due to some definite disadvantages of the former. These include the inability to identify systems fed by non-Gaussian noise (NGN) inputs and nonminimum phase (NMP) systems and identification of system nonlinearity [4]. In these cases, the autocorrelation-based methods offer no answer. Out of all

these, the identifiability of NMP systems has received the maximum attention from researchers.

HOS-based methods have been proposed to estimate motion between image frames [5–9]. In, the motion estimation is based on the bispectrum method for sub-pixel resolution of noisy image sequences. In [7], the displacement vector is obtained by maximizing a third-order statistics criterion. In [8], the global motion parameters are obtained by a new region recursive algorithm. In [6], several algorithms are developed based on a parametric cumulant method, a cumulant-matching method, and a mean kurtosis error criterion. The latter is an extension of the quadratic pixel-recursive method by Netravali and Robbins [10]. In [11], it is shown that such statistical parameters are insensitive to additive Gaussian noises. In particular, bispectrum parameters are insensitive to any symmetrically distributed noise and also exhibit the capability of better characterizing NGN and identifying NMP linear systems as well as nonlinear systems. Therefore, transformation to a higher-order domain reduces the effect of noise significantly. In this correspondence, a novel algorithm for the detection of motion vectors in video sequences is proposed. The algorithm uses bispectrum model-based subpixel motion estimation in the parametric domain for noisy image sequences to obtain a measure of content similarity for temporally adjacent frames and responds very well to scene motion vectors. The algorithm is insensitive to the presence of symmetrically distributed noise.

The outline of this paper is as follows. First, the problem formulation is introduced in Section 2. In Section 3, we first present briefly the definitions and properties of the bispectrum and cross-bispectrum. Next, we describe the motion estimation in the parametric domain. High-accuracy sub-pixel motion estimation is discussed in Section 4. Section 5 presents an evaluation of the computational complexity of our algorithm. The results of the experimental evaluation of the proposed method are shown in Section 6 and compared to existing methods while Section 7 concludes the paper.

## 2. Problem Formulation

The problem of motion estimation can be stated as follows: “Given an image sequence, compute a representation of the motion field that best aligns pixels in one frame of the sequence with those in the next” [9]. This is formulated as follows:

$$g_{k-1}(x, y) = f_{k-1}(x, y) + n_{k-1}(x, y), \quad (1)$$

$$\begin{aligned} g_k(x, y) &= f_k(x, y) + n_k(x, y) \\ &= f_{k-1}(x - d_x, y - d_y) + n_k(x, y), \end{aligned} \quad (2)$$

where  $(x, y)$  denotes spatial image position of a point;  $g_k(x, y)$  and  $g_{k-1}(x, y)$  are observed image intensities at instants  $k$  and  $k - 1$ , respectively;  $f_k(x, y)$  and  $f_{k-1}(x, y)$  are noise-free frames;  $n_k(x, y)$  and  $n_{k-1}(x, y)$  are assumed to be spatially and temporally stationary, zero-mean image Gaussian (or non-Gaussian) noise sequences with unknown

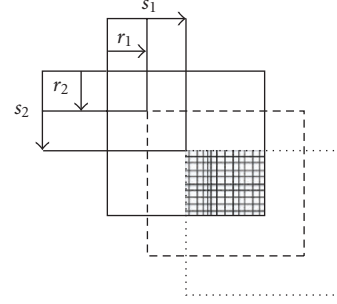


FIGURE 1: Physical representation of cumulants [4].

covariance; and  $(d_x, d_y)$  is the displacement vector of the object during the time interval  $[k - 1, k]$ .

The goal is to estimate  $(d_x, d_y)$  from  $g_k(x, y)$  and  $g_{k-1}(x, y)$ .

## 3. Bispectrum-Based Image Motion Estimation

**3.1. Definitions and Properties.** In this subsection, some HOS functions are defined and their properties are described in order to provide the necessary tools to understand the motion estimation methodology.

Also, the third-order autocumulant/moment sequence,  $C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2)$ , of  $g_k(x, y)$  is defined as follows [5]:

$$\begin{aligned} C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2) \\ = E[g_k(x, y)g_k(x + r_1, y + r_2)g_k(x + s_1, y + s_2)], \end{aligned} \quad (3)$$

where  $E\{\cdot\}$  denotes the expectation operation;  $(r_1, r_2)$  and  $(s_1, s_2)$  are two shifted versions of the  $g_k(x, y)$ .

To understand the theory of triple correlations physically for 2D data [4], the reader is referred to Figure 1. The Figure shows the spaces occupied by the original data (denoted by continuous box) and two shifted versions of the same data (denoted by dashed boxes). The shifts are made by the amounts  $(r_1, r_2)$  and  $(s_1, s_2)$ , respectively. It is now obvious that the product of the overlapping data positions (shown by the shaded portion) denotes the triple correlation function as defined by (3).

Also, the third-order autocumulant/moment sequence is defined as follows [12]:

$$\begin{aligned} C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2) &= C_3^{f_k f_k f_k}(r_1, r_2; s_1, s_2) \\ &\quad + C_3^{n_k n_k n_k}(r_1, r_2; s_1, s_2). \end{aligned} \quad (4)$$

Equation (4) states that the triple correlation of  $f_k(x, y)$  plus  $n_k(x, y)$  is equal to the triple correlation of the  $f_k(x, y)$  plus the triple correlation of the noise. For  $n_k(x, y)$  a zero-mean Gaussian noise, then its triple correlation is identically zero [7, 12]. This provides a theoretical basis for using the triple correlation (or the bispectrum) as a method of reducing the effects of additive noise. Then the term  $C_3^{n_k n_k n_k}(r_1, r_2; s_1, s_2)$  is negligible which renders the triple correlation very effective in detecting a signal embedded in noise. Therefore,

$$C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2) = C_3^{f_k f_k f_k}(r_1, r_2; s_1, s_2). \quad (5)$$

Also,  $n_k(x, y)$  can be non-Gaussian if it is independent and identically distributed (i.i.d.) and nonskewed (e.g., symmetrically distributed).

The bispectrum,  $B_3^{g_k g_k g_k}(u_1, u_2; v_1, v_2)$ , is defined as the 4D Fourier transform of the third-order autocumulant (or moment) [4]:

$$\begin{aligned} B_3^{g_k g_k g_k}(u_1, u_2; v_1, v_2) &= \mathcal{F}^4[C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2)] \\ &= \mathcal{F}^4[C_3^{f_k f_k f_k}(r_1, r_2; s_1, s_2)] \quad (6) \\ &= B_3^{f_k f_k f_k}(u_1, u_2; v_1, v_2), \end{aligned}$$

where  $\mathcal{F}^4$  denotes the 4D Fourier transform operation;  $(u_1, u_2)$  and  $(v_1, v_2)$  are the frequency coordinates for the 2D Fourier transform.

Let  $G_{g_k}(u)$  be the discrete Fourier transform (DFT) of the frame  $g_k(x, y)$ . Each component of the bispectrum is estimated by a triple product of Fourier coefficients as follows [12]:

$$\begin{aligned} B_3^{g_k g_k g_k}(u_1, u_2; v_1, v_2) \\ = G_{g_k}(u_1, u_2) G_{g_k}(v_1, v_2) G_{g_k}^*(u_1 + v_1, u_2 + v_2), \end{aligned} \quad (7)$$

where  $*$  indicates the complex conjugate.

The cross-bispectrum is obtained in a similar manner as the bispectrum. Thus,

$$\begin{aligned} B_3^{g_{k-1} g_k g_{k-1}}(u_1, u_2; v_1, v_2) \\ = B_3^{g_{k-1} g_{k-1} g_{k-1}}(u_1, u_2; v_1, v_2) e^{-j2\pi(u_1 d_x + u_2 d_y)}, \quad (8) \\ B_3^{g_{k-1} g_k g_{k-1}}(u_1, u_2; v_1, v_2) \\ = G_{g_{k-1}}(u_1, u_2) G_{g_k}(v_1, v_2) G_{g_{k-1}}^*(u_1 + v_1, u_2 + v_2). \end{aligned}$$

On close observation and after certain algebraic manipulations, (3) shows the third-order cumulant sequence to possess the following symmetry properties [4]:

$$\begin{aligned} C_3^{g_k g_k g_k}(r_1, r_2; s_1, s_2) &= C_3^{g_k g_k g_k}(s_1, s_2; r_1, r_2) \\ &= C_3^{g_k g_k g_k}(-s_1, -s_2; r_1 - s_1, r_2 - s_2) \\ &= C_3^{g_k g_k g_k}(s_1 - r_1, s_2 - r_2; -r_1, -r_2) \\ &= C_3^{g_k g_k g_k}(r_1 - s_1, r_2 - s_2; -s_1, -s_2) \\ &= C_3^{g_k g_k g_k}(-r_1, -r_2; s_1 - r_1, s_2 - r_2). \end{aligned} \quad (9)$$

This directly leads to the following symmetry properties of the bispectrum sequence [4]:

$$\begin{aligned} B_3^{g_k g_k g_k}(u_1, u_2; v_1, v_2) &= B_3^{g_k g_k g_k}(v_1, v_2; u_1, u_2) \\ &= B_3^{*g_k g_k g_k}(-v_1, -v_2; -u_1, -u_2) \\ &= B_3^{*g_k g_k g_k}(-u_1, -u_2; -v_1, -v_2) \\ &= B_3^{g_k g_k g_k}(-u_1 - v_1, -u_2 - v_2; v_1, v_2) \\ &= B_3^{g_k g_k g_k}(u_1, u_2; -u_1 - v_1, -u_2 - v_2) \\ &= B_3^{g_k g_k g_k}(-u_1 - v_1, -u_2 - v_2; u_1, u_2) \\ &= B_3^{g_k g_k g_k}(v_1, v_2; -u_1 - v_1, -u_2 - v_2). \end{aligned} \quad (10)$$

These symmetry properties reduce the computational burden while calculating the bispectrum.

**3.2. Parametric Model-Based Motion Estimation.** The problem of signal processing using bispectrum in the parametric domain has recently been widely addressed by researchers. Two primary schools of thought exist for this. The first line is headed by Raghuvver and Nikias [13], who have parametrized the bispectrum through solution of a cumulant matrix equation. The other school of thought headed by Giannakis [14] calculates the system impulse response coefficients directly using a linear combination of cumulant slices. Other papers have been subsequently published which give extensions and applications of these basic approaches (see [15] and references therein). Here, we propose the bispectrum model-based subpixel motion estimation in the parametric domain. Simulations demonstrate that this method requires large blocks of data and thus may be appropriate for estimating object displacement in background noise. Consequently, our approach will be derived in this context and hence,  $(d_x, d_y) = \text{const.}$  in (2). Substituting  $f_{k-1}$  from (1) into (2) we obtain

$$\begin{aligned} g_k(x, y) &= \sum_{i \in R} \sum_{j \in R} \alpha(i, j) g_{k-1}(x - i, y - j) \\ &\quad + n_k(x, y) + n_{k-1}(x - d_x, y - d_y), \end{aligned} \quad (11)$$

where, in theory,  $\alpha(i, j) = 0$  for all  $\{(i, j)\}$  except  $(i, j) = (d_x, d_y)$ , and  $\alpha(d_x, d_y) = 1$ . If the search region  $R$  contains the largest possible horizontal and vertical delays, and if we take  $g_k(x, y)$  for  $g_k(x + r_1, y + r_2)$ , we obtain

$$\begin{aligned} g_k(x + r_1, y + r_2) &= \sum_{i \in R} \sum_{j \in R} \alpha(i, j) g_{k-1}(x - i + r_1, y - j + r_2) \\ &\quad + n_k(x + r_1, y + r_2) \\ &\quad + n_{k-1}(x + r_1 - d_x, y + r_2 - d_y). \end{aligned} \quad (12)$$

By multiplying both sides of (12) by  $g_{k-1}(x, y) g_{k-1}(x + s_1, y + s_2)$  and taking expectations, we obtain

$$\begin{aligned} E[g_{k-1}(x, y) g_k(x + r_1, y + r_2) g_{k-1}(x + s_1, y + s_2)] \\ = \sum_{i \in R} \sum_{j \in R} \alpha(i, j) E[g_{k-1}(x, y) g_{k-1}(x - i + r_1, y - j + r_2) \\ \quad \times g_{k-1}(x + s_1, y + s_2)] \\ + E[g_{k-1}(x, y) n_k(x + r_1, y + r_2) \\ \quad \times g_{k-1}(x + s_1, y + s_2)] \\ + E[g_{k-1}(x, y) g_{k-1}(x + s_1, y + s_2) \\ \quad \times n_{k-1}(x + r_1 - d_x, y + r_2 - d_y)]. \end{aligned} \quad (13)$$



Therefore,

$$\begin{aligned}
 C_3^{g_{k-1}g_k g_{k-1}}(r_1, r_2; s_1, s_2) \\
 = \sum_{i \in R} \sum_{j \in R} \alpha(i, j) C_3^{g_{k-1}g_k g_{k-1}}(r_1 - i, r_2 - j; s_1, s_2) \\
 + C_3^{g_{k-1}n_k g_{k-1}}(r_1, r_2; s_1, s_2) \\
 + C_3^{g_{k-1}n_{k-1} g_{k-1}}(r_1 - d_x, r_2 - d_y; s_1, s_2).
 \end{aligned} \quad (14)$$

However,  $C_3^{g_{k-1}n_k g_{k-1}}(r_1, r_2; s_1, s_2)$  and  $C_3^{g_{k-1}n_{k-1} g_{k-1}}(r_1 - d_x, r_2 - d_y; s_1, s_2)$  are identically zero for all  $(r_1, r_2; s_1, s_2)$  due to the fact that the signal and noise are zero-mean and independent. Consequently, (14) becomes

$$\begin{aligned}
 C_3^{g_{k-1}g_k g_{k-1}}(r_1, r_2; s_1, s_2) \\
 = \sum_{i \in R} \sum_{j \in R} \alpha(i, j) C_3^{g_{k-1}g_k g_{k-1}}(r_1 - i, r_2 - j; s_1, s_2).
 \end{aligned} \quad (15)$$

Taking the 4D Fourier transform of (15) and rearranging, the following pulse transfer function for the system is obtained:

$$\begin{aligned}
 B_3^{g_{k-1}g_k g_{k-1}}(u_1, u_2; v_1, v_2) \\
 = B_3^{g_{k-1}g_k g_{k-1}}(u_1, u_2; v_1, v_2) \sum_{i \in R} \sum_{j \in R} \alpha(i, j) e^{-j2\pi(u_1 i + u_2 j)}.
 \end{aligned} \quad (16)$$

The third-order hologram,  $h(r_1, r_2)$ , is then defined by

$$\begin{aligned}
 h(r_1, r_2) &= \mathcal{F}^{-4} \left[ \frac{B_3^{g_{k-1}g_k g_{k-1}}(u_1, u_2; v_1, v_2)}{B_3^{g_{k-1}g_k g_{k-1}}(u_1, u_2; v_1, v_2)} \right] \\
 &= \sum_{i \in R} \sum_{j \in R} \alpha(i, j) \delta(r_1 - i, r_2 - j),
 \end{aligned} \quad (17)$$

where  $\mathcal{F}^{-4}$  denotes the 4D inverse Fourier transform operation.

Selecting various integers for  $r_1$  and  $r_2$ , we form an overdetermined system of equations. For example, if the chosen search region  $R$  is a rectangular that varies from  $-P_x$  to  $P_x$  in the horizontal direction and  $-P_y$  to  $P_y$  in the vertical direction, and if  $r_1$  ranges from  $-P_x$  to  $P_x$ , and  $r_2$  ranges from  $-P_y$  to  $P_y$ , a set of linear equations can be produced as follows:

$$\hat{h} = \alpha \hat{\delta}, \quad (18)$$

where

$$\hat{h} = [h_l(-P_x, -P_y), h_l(-P_x + 1, -P_y + 1), \dots, h_l(P_x, P_y)]^T,$$

$$\alpha = [\alpha(-P_x, -P_y), \alpha(-P_x + 1, -P_y + 1), \dots, \alpha(P_x, P_y)]^T,$$

$$\hat{\delta} = \begin{bmatrix} \delta(0, 0) & \cdots & \delta(-2P_x, -2P_y) \\ \delta(1, 1) & \cdots & \delta(-2P_x + 1, -2P_y + 1) \\ \vdots & & \vdots \\ \delta(2P_x, 2P_y) & \cdots & \delta(0, 0) \end{bmatrix}. \quad (19)$$

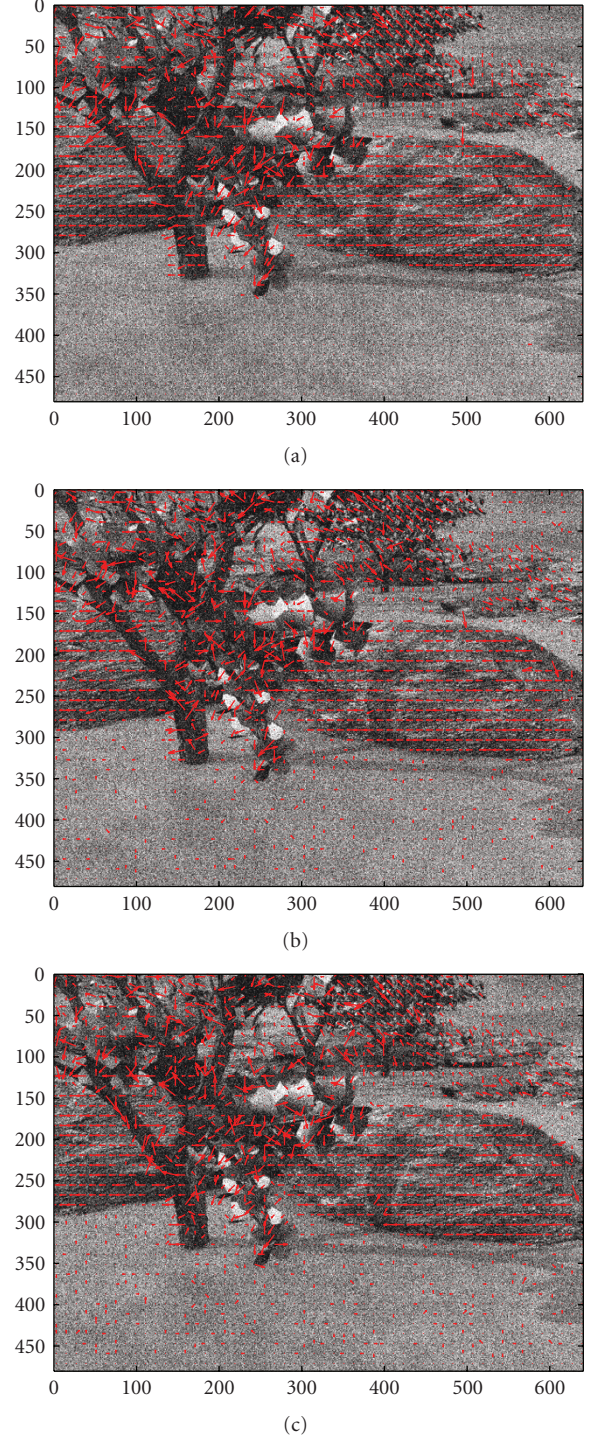


FIGURE 2: Motion vector fields of *Grove* sequence in the presence of noise using (a) our algorithm, (b) optical flow algorithm, (c) phase correlation algorithm.

The least-squares solution of (18) is given:

$$\alpha_{ls} = [\hat{\delta}^T \hat{\delta}]^{-1} \hat{\delta}^T \hat{h}. \quad (20)$$



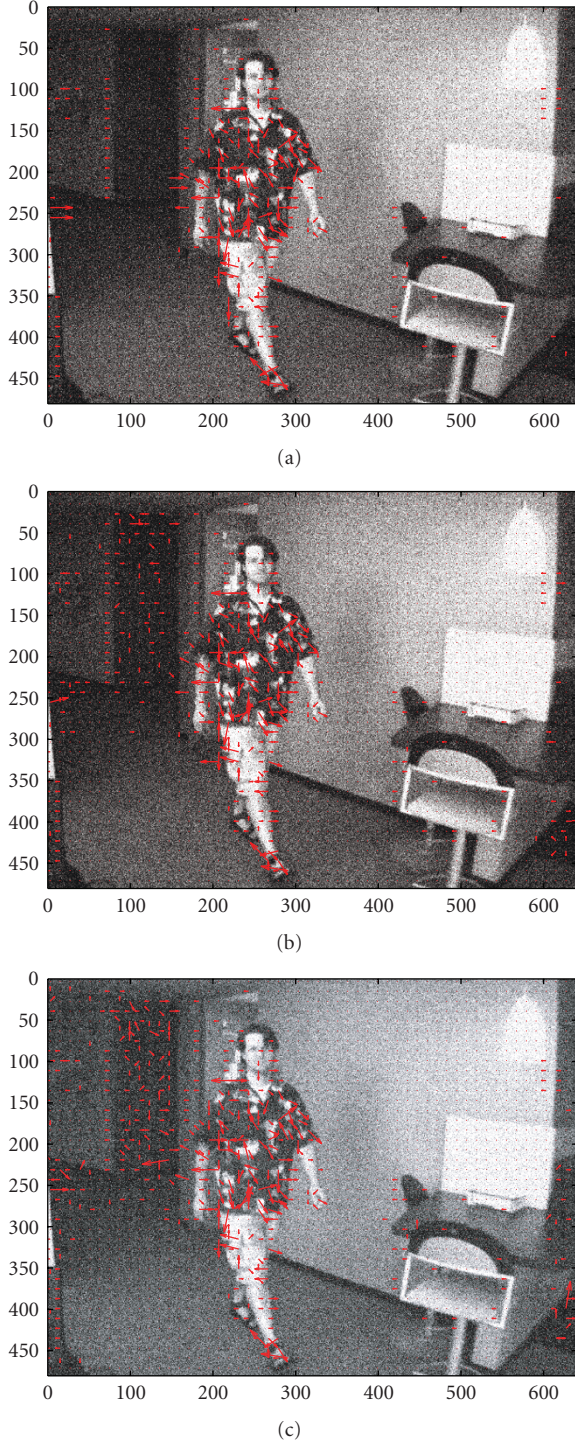


FIGURE 3: Motion vector fields of *Walking* sequence in the presence of noise using (a) our algorithm, (b) optical flow algorithm, (c) phase correlation algorithm.

The least-squares solution  $\hat{\alpha}_{ls}$  is obtained and its maximum  $\hat{\alpha}(\hat{d}_x, \hat{d}_y)$  is determined. The image motion estimate is then  $(\hat{d}_x, \hat{d}_y)$ .

Although we assume that the signals are non-Gaussian, it can be shown that for binary, deterministic signals and large

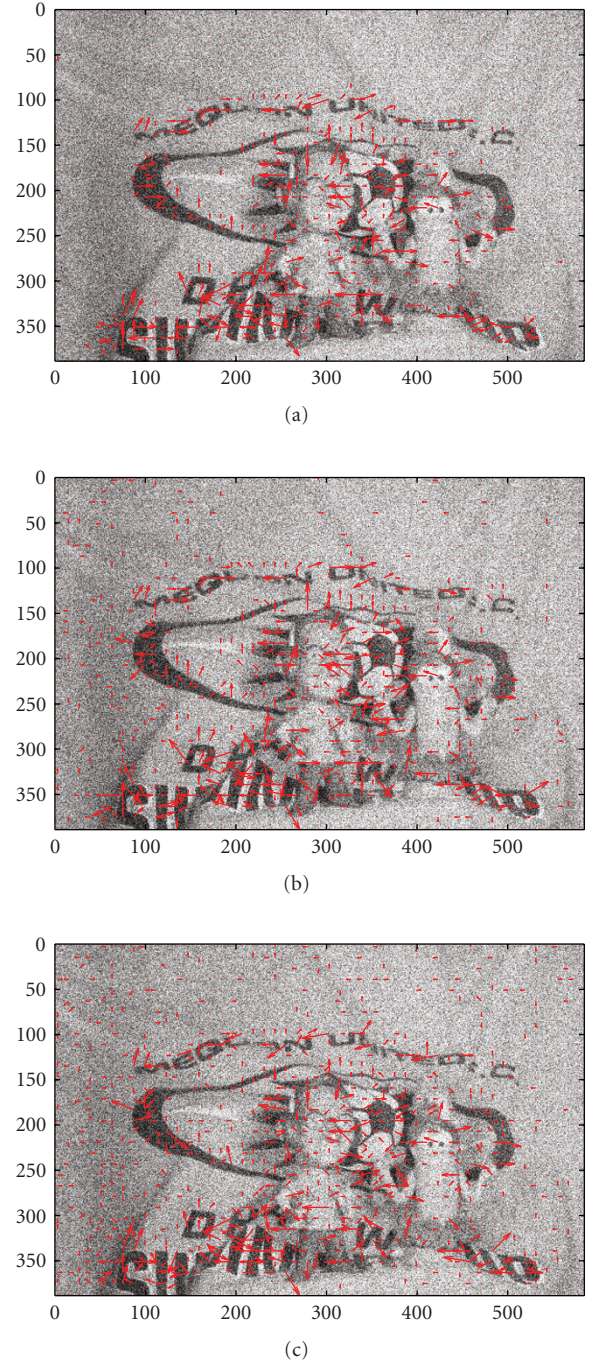


FIGURE 4: Motion vector fields of *Mequon* sequence in the presence of noise using (a) our algorithm, (b) optical flow algorithm, (c) phase correlation algorithm.

images sizes bispectrum is insensitive to Gaussian noise, and thus (20) is approximately true. Therefore,

$$\begin{aligned} \hat{h}(r_1, r_2) &= \hat{\alpha}(\hat{d}_x, \hat{d}_y) \hat{\delta}(r_1 - \hat{d}_x, r_2 - \hat{d}_y) \\ &= \hat{\delta}(r_1 - \hat{d}_x, r_2 - \hat{d}_y). \end{aligned} \quad (21)$$

TABLE 1: The comparison between three methods for the computation time.

Sequences	Methods	MECT (sec)
<i>Grove</i>	Our method	0.49
	Optical flow	0.25
	Phase correlation	0.38
<i>Beanbags</i>	Our method	0.50
	Optical flow	0.26
	Phase correlation	0.37
<i>Mequon</i>	Our method	0.37
	Optical flow	0.18
	Phase correlation	0.30
<i>Hydrangea</i>	Our method	0.37
	Optical flow	0.19
	Phase correlation	0.31
<i>Walking</i>	Our method	0.50
	Optical flow	0.25
	Phase correlation	0.38

#### 4. High-Accuracy Subpixel Motion Estimation

Subpixel performance is a critical element of the proposed algorithm. With reference to our previously published work [16, 17], we are introducing a number of important new features, which improve the accuracy of the motion estimates.

The coordinates  $(r_{1m}, r_{2m})$  of the maximum of the real-valued array  $\hat{h}(r_1, r_2)$  can be used as an estimate of the horizontal and vertical components of motion between  $g_k(x, y)$  and  $g_{k-1}(x, y)$  as follows:

$$(r_{1m}, r_{2m}) = \arg \max \text{Re}(\hat{h}(r_1, r_2)), \quad (22)$$

where  $\text{Re}\{\cdot\}$  denotes the real part of complex array  $\hat{h}(r_1, r_2)$ .

Subpixel accuracy of motion measurements is obtained by variable-separable fitting performed in the neighborhood of the maximum using one-dimensional quadratic function. Using the notation in (22), prototype functions are fitted to the triplets:

$$\{\hat{h}(r_{1m} - 1, r_{2m}), \hat{h}(r_{1m}, r_{2m}), \hat{h}(r_{1m} + 1, r_{2m})\}, \quad (23)$$

$$\{\hat{h}(r_{1m}, r_{2m} - 1), \hat{h}(r_{1m}, r_{2m}), \hat{h}(r_{1m}, r_{2m} + 1)\}, \quad (24)$$

that is, the maximum peak of the phase correlation surface and its two neighboring values on either side, vertically and horizontally.

The location of the maximum of the fitted function provides the required subpixel motion estimate  $(\hat{d}_k x, \hat{d}_k y)$ . Fitting a parabolic function horizontally to the data triplet (23) yields a closed-form solution for the horizontal component of the motion estimate  $\hat{d}_k x$  as follows:

$$\hat{d}_k x = \frac{\hat{h}(r_{1m} + 1, r_{2m}) - \hat{h}(r_{1m} - 1, r_{2m})}{2\hat{H}}, \quad (25)$$

where  $\hat{H} = [\hat{h}(r_{1m} + 1, r_{2m}) - 2\hat{h}(r_{1m}, r_{2m}) + \hat{h}(r_{1m} - 1, r_{2m})]$ .

The fractional part  $\hat{d}_k y$  of the vertical component can be obtained in a similar way using (24) instead of (23).

Finally the horizontal and vertical components of the subpixel accurate motion estimate are obtained by computing the location of the maxima of each of the above fitted quadratics.

In [18], it is shown that half-pixel accuracy motion vectors lead to a very significant improvement when compared to one pixel accuracy, whereas a higher precision results in negligible changes. Therefore, a half-pixel accuracy was chosen in our simulations.

#### 5. Computational Cost Comparison

The majority of the computational cost of the proposed bispectrum is due to the fast Fourier transform (FFT). Therefore, the fundamental computation required for bispectral estimates is given by (7), the triple product of the three individual Fourier transformations, while this computation is straightforward, limitations on computer time and statistical variance impose severe limitations on implementation of the definition of the bispectrum [19]. On the other hand, we take advantage of the symmetrical properties of the bispectrum to reduce the computational complexity and memory requirements of calculating third-order statistics. It can now be calculated in any one sector and mapped onto the others [20].

The phase correlation is estimated by multiplying each coefficient  $G_{gk}(u_1, u_2)$  by its complex conjugate, but each component of the bispectrum is estimated by a triple product of Fourier coefficients as demonstrated in (7). Thus, the number of operations required to compute the bispectrum is significantly increased relative to the phase correlation. There are  $N^2/8$  independent components of the bispectrum while there are only  $N/2$  independent components of the phase correlation for an  $N \times N$  image [21].

#### 6. Simulation Results

Our experiments have aimed at evaluating the performance of the proposed approach and comparing it with that of the optical flow and phase correlation techniques. For the optical flow method we used the implementation obtained from Bruhn method [22]. In our simulation we used the database freely available on the web at <http://vision.middlebury.edu/flow/>. We contribute three types of data to test different aspects of all techniques: real sequences of independent motion; realistic synthetic sequences; and high frame-rate video. These sequences have been chosen for their difficult motion and their different characteristics. Although the original sequences are in color, only the luminance component is used to estimate the motion vectors.

Figure 2 shows the estimated motion vector fields for the *Grove* sequence using the three aforementioned motion estimation methods. Note that for a fair comparison we used optical flow technique and phase correlation algorithm with



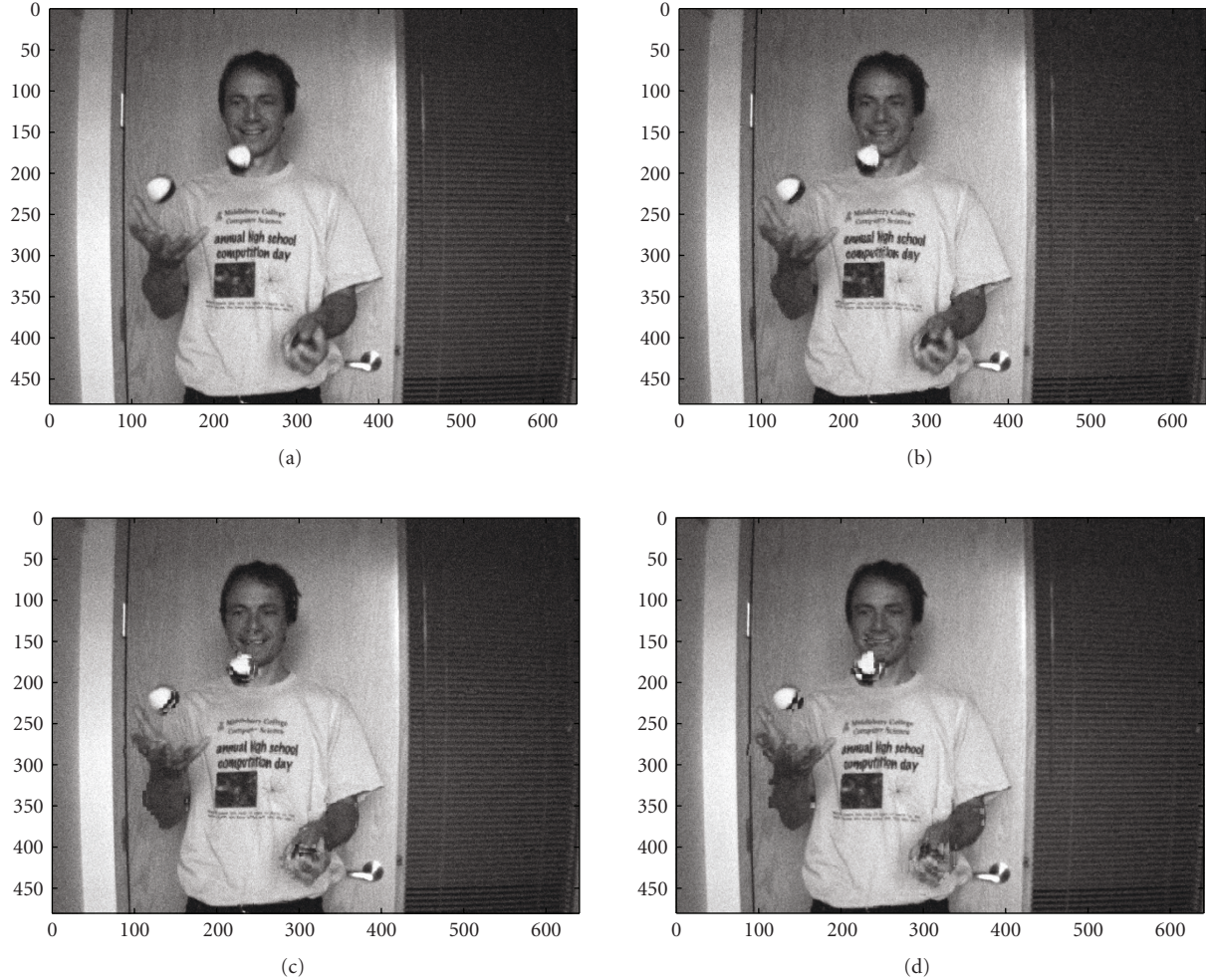


FIGURE 5: Prediction for frame 5 of the *Beanbags* sequence in the presence of noise using (b) our algorithm, (c) optical flow algorithm, (d) phase correlation algorithm, (a) is original image.

half-pixel accuracy. The motion vectors estimated between frames 6 and 7 are shown for the *Grove* sequence. For this particular sequence, our scheme provides the most consistent and reliable motion vector field. Both optical flow and phase correlation algorithms fail to detect the true motion vector. Similar results are shown in Figures 3 and 4 for the motion vectors estimated between frames 2 and 3, and between frames 5 and 6 in the *Walking* and *Mequon* sequences, respectively. Both optical flow and phase correlation algorithms produce abrupt motion vector fields. Although these abrupt motion vectors may lead to lower numerical mean squared errors (MSEs), they are incorrect motion vectors. Because of the noise resistant property of the parametric bispectrum method, it produces more reliable estimates. Therefore, our approach motion estimation results globally in motion fields more representative of the true motion in the scene.

To see more clearly the correctness of motion estimation, we use *Beanbags* sequence as an example. The motion compensated pictures using three methods are shown in Figure 5.

Portions of these three pictures are enlarged in Figure 6 to show the differences. We observe better compensated images by the proposed method. We also observe that the motion compensated images for our scheme are much closer to the original images. Thus, the scheme is able to measure the motion vector more accurately and is more robust in general. Overall, parametric bispectrum scheme typically offers better visual quality images than the other methods.

The detection of motion vectors relies on successive phase correlation operations applied to pairs of consecutive block partitioned frames of a video sequence. The heights of the dominant peaks are monitored, and when a sudden magnitude change is detected, then this is interpreted as a displacement vector. Figure 7 shows sample phase correlation surface between two blocks  $b_{k-1}(x, y)$  and  $b_k(x, y)$ , related to frames 3 and 4 of the *Hydrangea* sequence, respectively. The bispectrum retains both amplitude and phase information from the Fourier transform of a signal, unlike the other methods. The phase of the Fourier transform contains important shape information. Therefore,



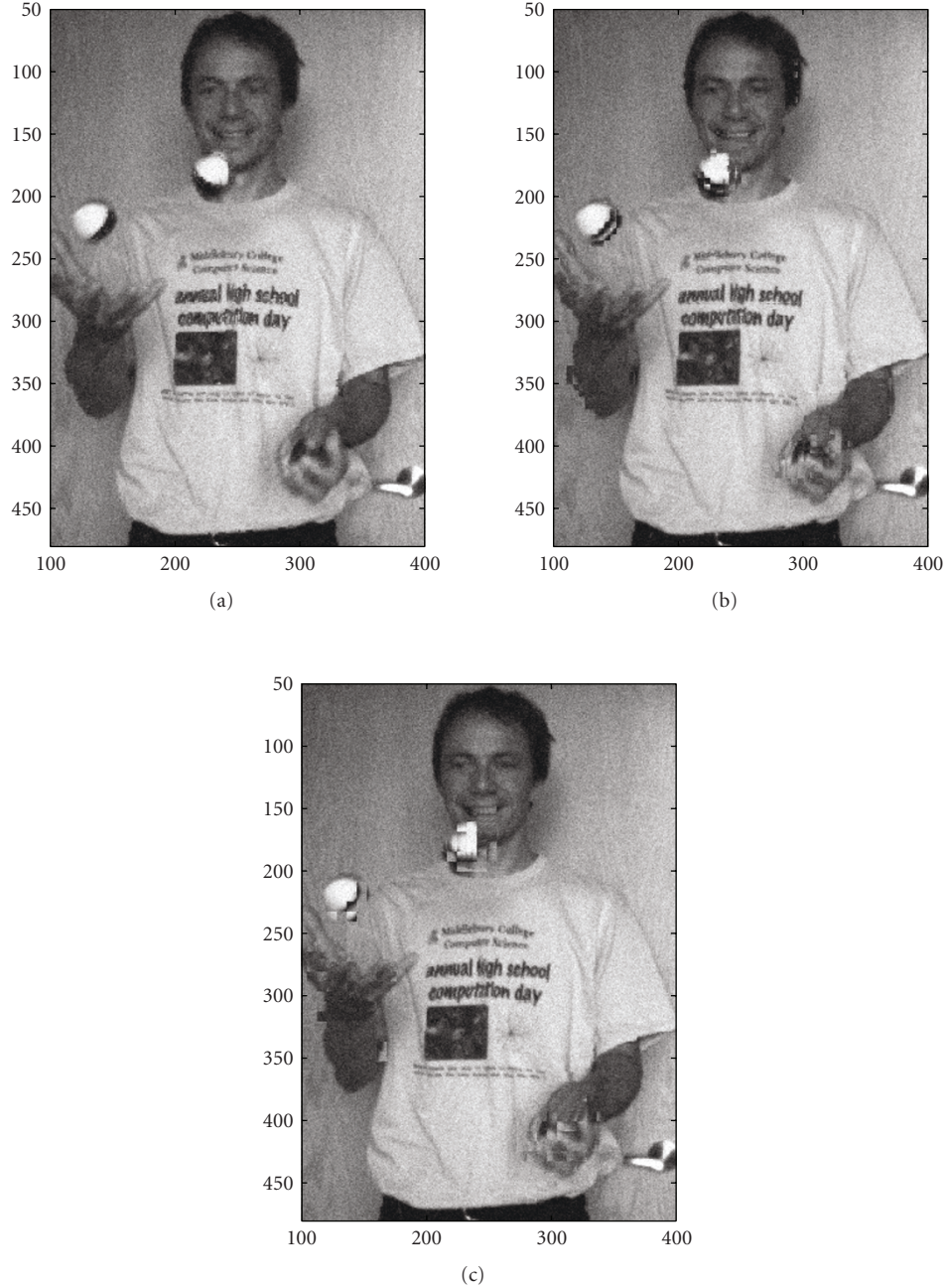


FIGURE 6: Enlarged portions of the motion compensated pictures of the *Beanbags* sequence using (a) our algorithm, (b) optical flow algorithm, (c) phase correlation algorithm.

the bispectrum minimizes the influence of the noise and simplifies the identification of the dominant peak on the correlation surface.

The PSNR of motion compensated is a popular performance measure for motion estimation, giving insight about the quality of the prediction. The PSNRs of the three motion estimation algorithms are shown in Figure 8. This result is obtained by using two real video sequences *Tempete* and *Stefan*. These sequences were run for 60 frames with a frame rate of 30 frame/sec. Both sequences are degraded with

additive zero-mean Gaussian noise to a signal-to-noise ratio (SNR) of 10 dB. Here we define

$$\text{SNR} = 10 \log_{10} \frac{\sigma_f^2}{\sigma_n^2}, \quad (26)$$

where  $\sigma_f^2$  is the variance of the frame,  $\sigma_n^2$  is the variance of the noise. From Figure 8, it is clear that the implemented optical flow technique is significantly less efficient than the parametric bispectrum technique. It is mainly due to the

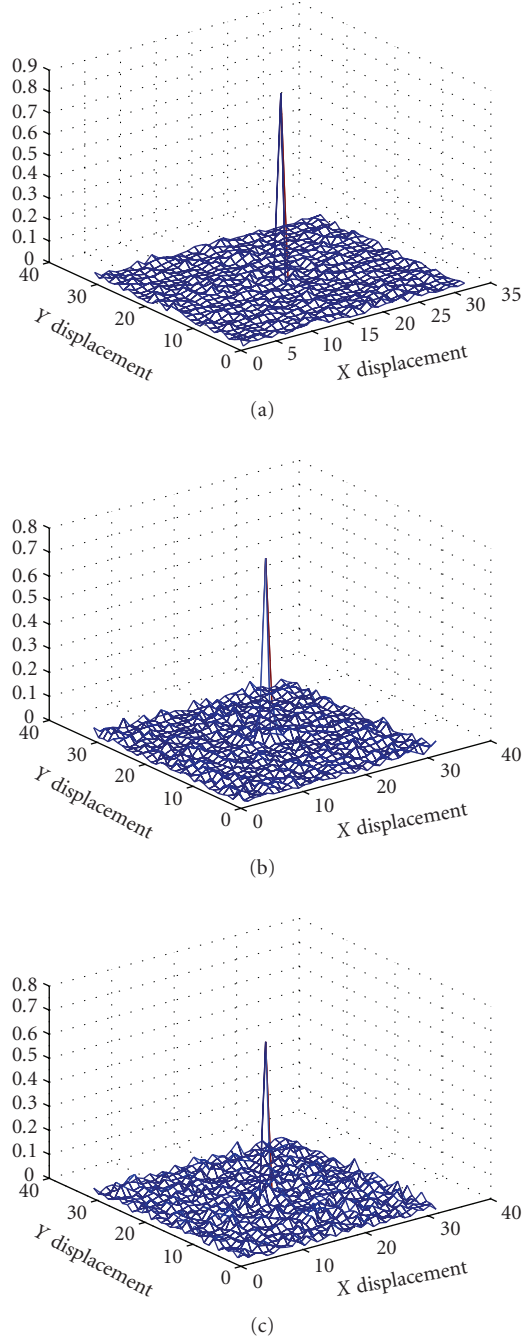


FIGURE 7: Phase correlation surfaces between two blocks using (a) our algorithm, (b) optical flow algorithm, (c) phase correlation algorithm.

difficulty of the optical flow technique to cope with large displacement and discontinuities in the motion field. On the other hand, the normalization (equalization) operation in the phase correlation technique enhances the noise power at high frequencies, and it produces incorrect displacement estimates on noisy image sequences. On the whole, the bispectrum retains both amplitude and phase information from the Fourier transform of a signal, unlike the other techniques. This confirms the motion that the proposed

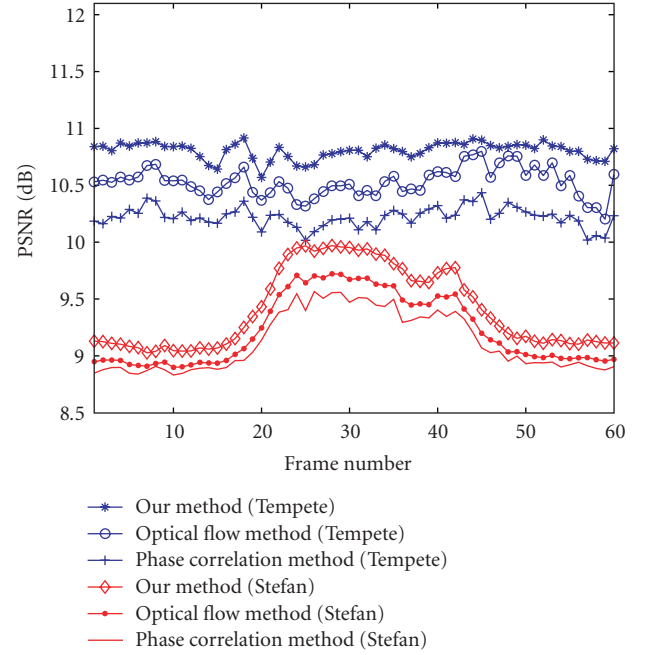


FIGURE 8: PSNR obtained for noisy sequences (SNR = 10 dB).

technique of an image is a superior feature selector utilizing the portions of the image spectrum most likely to contribute to reliable motion estimation.

In terms of complexity, this is measured by the computation time. All the computations are performed on Intel centrino duo machines (Toshiba Satellite A100-579 T5500, 2 GHz(2 CPUs)) with Windows XP. The three algorithms have been implemented using a prototype written in Matlab 6.5 R13. The comparison between three methods for the motion estimation computation time (MECT) is shown in Table 1.

We employ 60 frames of the video *Tempete* sequence. We perform the motion compensation procedure for each current frame  $k$  with respect to reference frames  $k - r$ , where  $r = 1, 2, 3$ , and 4. The average PSNR of the motion compensated images is given in Table 2, with *Tempete* sequence degraded with additive zero-mean Gaussian noise to an SNR of 10 dB.

The average PSNR,  $\text{PSNR}_{\text{avg}}$ , is given as follows:

$$\text{PSNR}_{\text{avg}} = \frac{1}{F} \sum_{i=1}^F \text{PSNR}_i, \quad (27)$$

where  $\text{PSNR}_i$  is the measured PSNR for frame  $i$  and  $F$  is the total number of frames. In Table 2, we observe that the  $\text{PSNR}_{\text{avg}}$  decreases with larger apparent disparity between the global motion of the background and the local motion of the foreground. For each value of  $r$ , we see that the  $\text{PSNR}_{\text{avg}}$  is higher for the proposed scheme than the other methods.

TABLE 2: Average PSNR of motion compensated images for the three motion estimation techniques (unit: dB) for *Tempete* sequence.

$r$	Our method	Optical flow	Phase correlation
1	10.80	10.50	10,21
2	10,23	9.83	9,59
3	9,93	9.59	9,25
4	9,68	9.24	8,81

## 7. Conclusion

In this paper, subpixel motion estimation algorithm using bispectrum in the parametric domain was presented. We have presented a collection of datasets for the evaluation of our method, available on the web at <http://vision.middlebury.edu/flow/>. In the case of the data is severely corrupted by additive Gaussian noises of unknown covariance, our method suppresses the effects of noise and simplifies the identification of the dominant peak on the correlation surface, unlike other techniques. At high noise levels SNR around 10 dB the optical flow and phase correlation techniques fail, yet even under these extreme conditions, the parametric bispectrum provides improvement in performance over the other algorithms. Overall, our scheme produces smoother displacement vector field with a more accurate measure of object motion in different SNR scenarios.

## References

- [1] N. Benmoussat, M. Faouzi Belbachir, and B. Benamar, "Motion estimation and compensation from noisy image sequences: a new filtering scheme," *Image and Vision Computing*, vol. 25, no. 5, pp. 686–694, 2007.
- [2] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsaggelos, and R. L. Lagendijk, "Noise reduction filters for dynamic image sequences: a review," *Proceedings of the IEEE*, vol. 83, no. 9, pp. 1272–1292, 1995.
- [3] R. M. Armitano, R. W. Schafer, F. L. Kitson, and V. Bhaskaran, "Robust block-matching motion-estimation technique for noisy sources," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, vol. 4, pp. 2685–2688, Munich, Germany, April 1997.
- [4] S. Bhattacharya, N. C. Ray, and S. Sinha, "2-D signal modelling and reconstruction using third-order cumulants," *Signal Processing*, vol. 62, no. 1, pp. 61–72, 1997.
- [5] E. M. Ismaili Aalaoui and E. Ibn-Elhaj, "Estimation of subpixel motion using bispectrum," *Research Letters in Signal Processing*, vol. 2008, Article ID 417915, 5 pages, 2008.
- [6] J. M. M. Anderson and G. B. Giannakis, "Image motion estimation algorithms using cumulants," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 346–357, 1995.
- [7] R. P. Kleihorst, R. L. Lagendijk, and J. Biemond, "Noise reduction of severely corrupted image sequences," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '93)*, vol. 5, pp. 293–296, Minneapolis, Minn, USA, April 1993.
- [8] E. Ibn-Elhaj, D. Aboutajdine, S. Pateux, and L. Morin, "HOS-based method of global motion estimation for noisy image sequences," *Electronics Letters*, vol. 35, no. 16, pp. 1320–1322, 1999.
- [9] E. Sayrol, A. Gasull, and J. R. Fonollosa, "Motion estimation using higher order statistics," *IEEE Transactions Image Processing*, vol. 5, no. 6, pp. 1077–1084, 1996.
- [10] A. N. Netravali and J. D. Robbins, "Motion-compensated television coding—part I," *Bell System Technical Journal*, vol. 58, no. 3, pp. 629–668, 1979.
- [11] V. Murino, C. Ottonello, and S. Pagnan, "Noisy texture classification: a higher-order statistics approach," *Pattern Recognition*, vol. 31, no. 4, pp. 383–393, 1998.
- [12] B. M. Sadler and G. B. Giannakis, "Shift- and rotation-invariant object reconstruction using the bispectrum," *Journal of the Optical Society of America A*, vol. 9, no. 1, pp. 57–69, 1992.
- [13] M. R. Raghuveer and C. L. Nikias, "Bispectrum estimation: a parametric approach," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 33, no. 5, pp. 1213–1230, 1985.
- [14] G. B. Giannakis, "On the identifiability of non-Gaussian ARMA models using cumulants," *IEEE Transactions on Automatic Control*, vol. 35, no. 1, pp. 18–26, 1990.
- [15] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, 1991.
- [16] E. M. Ismaili Aalaoui and E. Ibn-Elhaj, "Estimation of motion fields from noisy image sequences: using generalized cross-correlation methods," in *Proceedings of the IEEE International Conference on Signal Processing and Communications (ICSPC '07)*, Dubai, UAE, November 2007.
- [17] E. M. Ismaili Aalaoui and E. Ibn Elhaj, "Estimation of displacement vector field from noisy data using maximum likelihood estimator," in *Proceedings of the 14th IEEE International Conference on Electronics, Circuits, and Systems (ICECS '07)*, pp. 1380–1383, Marrakech, Morocco, December 2007.
- [18] G. Madec, "Half pixel accuracy in block matching," in *Proceedings on the Picture Coding Symposium (PCS '90)*, Cambridge, Mass, USA, March 1990.
- [19] K. S. Lii and K. N. Helland, "Cross-bispectrum computation and variance estimation," *ACM Transactions on Mathematical Software*, vol. 7, no. 3, pp. 284–294, 1981.
- [20] J.-M. Le Caillec and R. Garello, "Comparison of statistical indices using third order statistics for nonlinearity detection," *Signal Processing*, vol. 84, no. 3, pp. 499–525, 2004.
- [21] R. W. Means, B. Wallach, and D. Busby, "Bispectrum signal processing on HNC's SIMD numerical array processor (SNAP)," in *Proceedings of the ACM/IEEE Conference on Supercomputing (SC '93)*, pp. 535–537, Portland, Ore, USA, November 1993.
- [22] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.