

# AN OPTIMAL AND STATISTICALLY ROBUST CORRELATION TECHNIQUE FOR BLOCK BASED MOTION ESTIMATION

*F. Essannouni<sup>1,2</sup>, R. Oulad Haj Thami<sup>3</sup>, A. Salam<sup>1</sup> and D. Aboutajdine<sup>2</sup>*

<sup>1</sup> Laboratoire de Mathématiques Pures et Appliquées, Université du Littoral-Côte d'Opale.  
C.U. de la Mi-Voix, 50 rue F. Buisson, B.P. 699, 62228 Calais, Cedex, France.

Ahmed.Salam@lmpa.univ-littoral.fr

<sup>2</sup>GSCM, UFR IT

Faculté des Sciences, Université Mohamed V-Agdal, B.P. 1014 Rabat, Maroc  
aboutaj@fsr.ac.ma

<sup>3</sup> Laboratoire SI2M, Equipe WiM

ENSIAS, Université Mohamed V-Souissi, B.P. 713 Rabat, Maroc

oulad@ensias.ma

## ABSTRACT

In this paper we propose an optimal and robust correlation technique for the local motion estimation purposes. It is based on the maximization of a statistical robust matching function, which is computed in the frequency domain and therefore can be implemented by fast transformation algorithms. We show that our method achieves a significant speed up and robustness over the full search block-matching algorithm. We also present a comparative performance analysis, which shows that the proposed method greatly outperforms the state-of-the-art in correlation motion estimation.

## 1. INTRODUCTION

Block based motion estimation has many promising applications in the field of moving images analysis. Such as target detection and tracking, video compression, and template matching. The full search block matching algorithm (FS) is one of the most basic options. For every block of the current frame, FS searches all of the possible matching locations in the reference frame and finds a minimum point. The advantage of full search is that we can find the absolute optimal solution. However the FS requires a huge amount of computations. Many methods have been proposed to speed up the matching process where the simple sum absolute difference "SAD" or sum square difference "SSD" criterion is used [1–5]. In a cluttered environment, however, some outliers such as impulse noises or partial occlusions may occur during the motion estimation processes. In this situation, the SAD and SSD criteria are no longer suitable for block matching algorithms because they treat the outliers and inliers evenly when calculating the error measures.

Huber defines robust statistics as "insensitivity to small

deviations from the assumptions." [6]. Note that "small" may imply small deviations for all data (e.g., Gaussian noise) or large deviations for a small quantity of data (outliers). In relation to block based motion estimation, robustness implies correct matching in the presence of noise, occlusion, revealed regions, new objects, highlights in general, any effect that may cause deviation from a perfect match. There are several types of robust estimators. Among them, the M-estimator is one of the most popular estimator to solve the problem of robust parameter estimation. The basic idea of the M-estimator technique is to limit the influence of outliers in the matching error by replacing the SSD metric or the SAD metric by a less increasing function. One of the common M estimators is the Andrews wave cosine which satisfies the criteria for having an outlier process. In this paper we propose an optimal and robust correlation technique which is based on Andrews wave M estimator. Recently there has been a lot of interest in correlation motion estimation techniques operating in the frequency domain [8–11]. But in our knowledge no one of the existing correlation methods can guarantee the same level of performance as the full search block matching algorithm. In this paper however, we present a new correlation technique which can solve the dual problems of overwhelming complexity of the full search block matching algorithm and the sensitivity of the SAD and SSD metrics. The key idea is to combine the advantages of statistical robustness, offered by the Andrews' wave M-Estimator [6], with the speed and computational efficiency that typifies frequency domain processing, owing to the use of fast algorithms.

The rest of the paper is structured as follows. Sections 2-3 demonstrate the theoretical analysis behind our proposed correlation technique. In Section 4, the state-of-the-art in correlation motion estimation algorithms is briefly outlined. Section

5 presents some simulations results. Finally some concluding remarks are drawn in section 6.

## 2. ROBUST BLOCK MATCHING ALGORITHM

### 2.1. Problem formulation

The problem treated in block based motion estimation is to define for a given block  $g$  whose size is  $B \times B$  and a search area  $f$  whose size is  $w \times h$ , the position  $(mv_x, mv_y)$  with the minimum block distortion measure (BDM) among all possible search positions in the search area  $f$ .

$$BDM(x, y) = \sum_{l=0}^{B-1} \sum_{k=0}^{B-1} \rho(f(x+k, y+l) - g(k, l)), \quad (1)$$

where  $\rho$  is a symmetric, positive-definite function with a unique minimum at zero. The influence function  $\rho'(x)$  is defined as a derivative of the kernel  $\rho(x)$ . The influence function measures the influence of a datum on the value of the parameter estimate. For example, for the case of  $\rho(x) = x^2$ , the influence function is  $\rho'(x) = 2x$ , that is, the influence of a datum on the estimate increases linearly with the size of its error, which confirms the non-robustness of the sum square difference (SSD) to outliers.

### 2.2. The proposed matching criteria

The M-estimators try to reduce the effect of outliers by replacing the SSD metric with a less rapidly increasing loss function of the data value. Andrews' Wave M-Estimator is a common robust estimator that satisfies the criteria for having an outlier process. Andrews proposed an influence function (derivative of kernel) as:

$$\rho'(x) = \begin{cases} \frac{\sin(\pi x)}{\pi} & \text{for } |x| \leq 1 \\ 0 & \text{elsewhere} \end{cases}, \quad (2)$$

where  $x$  is the quantity to be minimized. Thus Andrews kernel function is of the form:

$$\rho(x) = \begin{cases} \frac{1}{\pi^2} (1 - \cos(\pi x)) & \text{for } |x| \leq 1 \\ \frac{1}{\pi^2} & \text{elsewhere} \end{cases}. \quad (3)$$

Since the motion estimation process are generally taken on luminance channel. Then the values taken by  $f$  and  $g$  are in general in the range of  $[0..255]$ .

So if we denote:

$$d_{x,y}(k, l) = \pi \left( \frac{f(k+x, l+y) - g(k, l)}{255} \right). \quad (4)$$

Then we can define a new matching criteria "SCD" which is based on Andrews wave cosine as:

$$SCD(x, y) = \sum_{l=0}^{B-1} \sum_{k=0}^{B-1} \cos(d_{x,y}(k, l)). \quad (5)$$

The position of the motion vector  $(mv_x, mv_y)$  can be deduced from the **maximization** of the  $SCD$  metric. Since the cosine function  $\cos(x)$  can be approximated to  $1 - \frac{x^2}{2}$  for the values of  $x$  close to 0. Then we can expect that the  $SCD$  metric returns about the same result as the SSD metric in the absence of noise and a better result than it in the presence of outliers.

## 3. THE PROPOSED CORRELATION TECHNIQUE

The  $SCD$  function can be viewed as a cross correlation operation. Indeed, if we note:

$$g_c(x, y) = \exp\left(i \frac{\pi}{255} g(x, y)\right), \quad (6)$$

$$f_c(x, y) = \exp\left(i \frac{\pi}{255} f(x, y)\right), \quad (7)$$

where  $i$  is the square root of  $-1$  and  $\exp(ix)$  is the complex exponential. Using the Euler's identity ( $\cos(x) = \Re(\exp(ix))$ ), the  $SCD$  can be written as:

$$SCD(x, y) = \Re\left(\sum_{l=0}^{B-1} \sum_{k=0}^{B-1} g_c^*(k, l) f_c(k+x, l+y)\right). \quad (8)$$

From the Fourier correlation theorem, the  $SCD$  surface (8) can be computed using FFT algorithm as follows:

$$\Re\{IFFT(F_c(u, v)G_c^*(u, v))\}, \quad (9)$$

where  $G_c$  is the  $FFT$  of  $g_c$  and  $F_c$  is the  $FFT$  of  $f_c$ .  $\Re$  denotes the real part of a complex number and the asterisk denotes complex conjugation.

Note that the  $f_c$  and  $g_c$  are correlated with FFTs by zero padding the size of  $g_c$  to the size of  $f_c$  prior to taking the forward FFTs. There are two different reasons for zero padding. Firstly, for element-wise multiplication to occur, they must be of the same size, and secondly, to avoid cyclic correlation. For this reason, the last  $B-1$  rows and  $B-1$  columns of result will contain wrap-around data that should be discarded. Finally, the resulting motion vector  $(mv_x, mv_y)$  of the block  $g$  in the search area  $f$  is measured from the position of the maximum in (9).

Most computational cost is with the Fourier transform. The proposed algorithm requires two forward and one backward complex FFTs of size  $(wh)$ . Then the complexity of the proposed algorithm is about  $O((wh) \log_2(wh))$  per block. For basic FFT algorithms, the length of the sequence is usually chosen to be a power of two. However, modern FFTs, such as [7], works on sequences of any length. By utilizing the FFT, we are then able to make use of the existing body of work aimed at speeding up the discrete Fourier transform.

## 4. PREVIOUS CORRELATION METHODS

These type of methods usually fall in the category of so-called matching methods and offer well documented advantages in

terms of computational efficiency due to the employment of Fast Fourier transform algorithms [8–11]. Perhaps the most documented method in this class, the phase correlation, is based on the well known Fourier shift propriety [11]. The phase correlation formula between two images  $f_1$  and  $f_2$  is given in [11] by:

$$P(x, y) = IFFT \left\{ \frac{F_2(u, v) F_1^*(u, v)}{|F_2(u, v) F_1^*(u, v)|} \right\}, \quad (10)$$

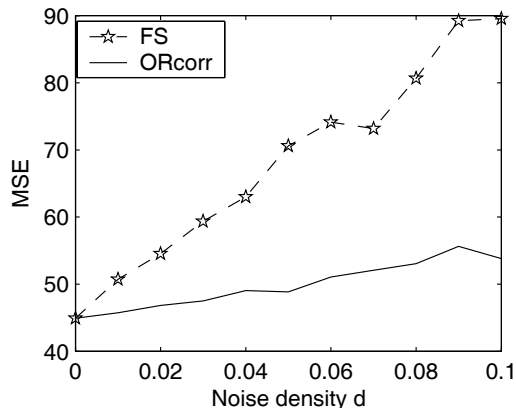
$F_1$  and  $F_2$  denote the Fourier Transforms of the two images  $f_1$  and  $f_2$ .

In block based motion estimation the phase correlation technique works by computing the phase correlation function (10) between two regions of image data, usually in the form of co-sited rectangular blocks in the current and the next frame of a video sequence [12–15]. However, in local motion estimation, the phase correlation method fails in several cases. Indeed the mathematical analysis of this technique assumes ideal translation which is not true between two compared blocks from real frames.

Recently, a new type of correlation technique "Fast Robust Correlation" has been proposed [10]. It works by expressing a robust matching surface as a series of correlations using cosine function. Speed is obtained by computing correlations in the frequency domain. An additional version of this cross correlation approach using angle gradient features has been proposed by the same authors in [9]. The "Orientation correlation" method operates by estimating the motion from orientation images. Each pixel in a orientation image is a complex number that represents the orientation of intensity. Authors have reported enhancements to the existing correlation methods. Despite the improvement, we can note through the experiments presented in [10] and [9], that both of these methods remain sub optimal when compared with the direct full search for block matching algorithm. Therefore, even though the existing correlation methods can present an attractive alternative for solving the overwhelming complexity of the full search block matching algorithm, these methods are not able to guarantee the same level of performance as the full search block matching algorithm

## 5. SIMULATION RESULTS

We compare here between our proposed correlation algorithm "ORcorr" (as detailed in section 3), the phase correlation method (PC) [15], the orientation correlation method (OC) [9] the fast robust correlation "FRcorr" as implemented in [10] ( video coding section) and the full search block matching algorithm under SSD metric (FS). The local motion estimation performance was assessed by applying a simple motion compensation scheme between the current frame and the previous frame using the estimated motion parameters and computing the relative mean square error "MSE". In these experiments, the maximum displacement is set to  $\pm 8$  pixels and the block size



**Fig. 1.** MSE between the reconstructed frame and the clean current frame from Carphone sequence using multiple values of the "salt and pepper" noise density  $d$ .

**Table 1.** Mean square prediction error "MSE"

	Carphone	Table	Foreman
ORcorr	26.42	82.92	24.39
PC	76.60	179.50	66.36
OC	47.68	143.59	43.29
FRcorr	46.29	138.38	64.60
FS	<b>26.41</b>	<b>82.92</b>	<b>24.22</b>

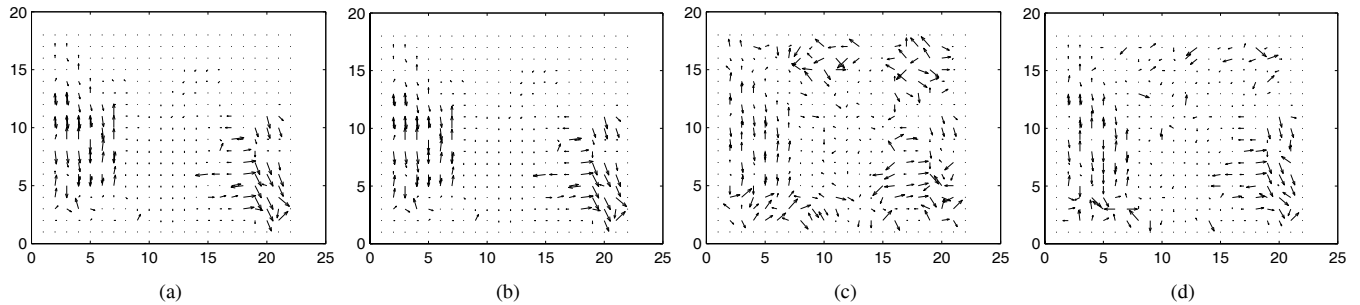
used is  $16 \times 16$  pixels.

Table 1 illustrates the MSE results for the first 100 frames from Carphone(CIF  $352 \times 288$ ), Table (QCIF  $176 \times 144$ ) and Foreman (QCIF  $176 \times 144$ ) video test sequences.

As this table shows, our method greatly outperforms the existing correlation techniques in terms of minimizing MSE and gives about the same level of performance as the FS under SSD. We note also that, our proposed correlation method has been tested on many sequences using different values of the block size and the search area size, and it has given about the same MSE as the FS for all tests.

On the other hand, to demonstrate the robustness of our method against the FS under SSD. We consider the first and the second frame from Carphone sequence. We introduce an additive "salt and pepper" noise into the second frame and we apply each of the proposed correlation and the FS to find the motion field between the first frame and the second polluted frame. We compute after the MSE between the reconstructed frame and the **clean current frame**. Figure 1 illustrates the MSE computed for multiple values of the noise density. Figures 2.a and 2.b show the motion field found by the FS and the proposed correlation method in the absence of noise, meanwhile Figures 2.c and 2.d show the motion field in the presence of "salt and pepper" noise whose density  $d$  is 0.07.

From Figure 2, we can see that, the motion field found by the



**Fig. 2.** Motion field by a) the FS under SSD b) the proposed correlation in the absence of noise, and by c) the FS under SSD d) the proposed correlation in the presence of "salt and pepper" noise whose density  $d$  is 0.07.

FS under SSD has largely changed in the presence of "salt and pepper" noise unlike the one found by the proposed correlation method. Also Figure 1 highlights the fact that our proposed correlation is substantially more immune to outliers than the full search block matching algorithm under SSD metric.

## 6. CONCLUSION

In this paper we have presented an optimal correlation technique which although uses a method that is computationally simple, can provide a singular solution to the problems of the overwhelming complexity of the full search and the weakness of the SSD metric. Simulation results show that our algorithm offers a high degree of robustness in the presence of outliers when compared with the direct full search algorithm under SSD metric. Furthermore, it greatly outperforms the existing correlation techniques in terms of motion estimation accuracy while still keeping the same order of complexity.

## 7. REFERENCES

- [1] R. Li, B. Zeng, and M.L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. On Circuits and Systems for Video Technology*, vol. 4, no. 4, pp. 438–442, August 1994.
- [2] W. Li and E. Salari, "Successive elimination algorithm for motion estimation," *IEEE Transactions on Image Processing*, vol. 4, no. 1, pp. 105–107, January 1995.
- [3] Y.C. Lin and S.C. Tai, "Fast full-search block-matching algorithm for motion-compensated video compression," *IEEE Transactions on communications*, vol. 45, no. 5, pp. 527–531, May 1997.
- [4] Y.S. Chen, Y.P. Hung, and C.S. Fuh, "Fast block matching algorithm based on the winner-update strategy," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1212–1222, August 2001.
- [5] F. Essannouni, R. Oulad Haj Thami, A. Salam, and D. Aboutajdine, "A new fast full search block matching algorithm using frequency domain," in *IEEE ISSPA*, Sydney, Australia, 2005.
- [6] P. Huber, *Robust statistics*, John Wiley, New York, 1981.
- [7] Matteo Frigo and Steven G. Johnson, "FFTW: An adaptive Software Architecture for the FFT," *ICASSP*, vol. 3, pp. 1381, 1998.
- [8] V. Argyriou and T. Vlachos, "Using gradient correlation for sub-pixel motion estimation of video sequences," in *ICASSP04*, 2004, pp. 329–331.
- [9] A.J. Fitch, A. Kadyrov, W.J. Christmas, and J.V. Kittler, "Orientation correlation," in *BMVC02*, 2002, p. Matching/Recognition.
- [10] A.J. Fitch, A. Kadyrov, W.J. Christmas, and J.V. Kittler, "Fast robust correlation," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1063–1073, August 2005.
- [11] C.D. Kuglin and D.C. Hines, "The phase correlation image alignment method," in *IEEE 1975 International Conference on Systems, Man and Cybernetics*, September 1975, pp. 163–165.
- [12] G.A. Thomas, "Television motion measurement for datv and other applications," *Research Report*, 1987.
- [13] Y.M. Chou and H.M. Hang, "A new motion estimation method using frequency components," *JVCIR*, vol. 8, no. 1, pp. 83–96, March 1997.
- [14] T. Vlachos and G. Thomas, "Motion estimation for the correction of twin-lens telecine flicker," in *ICIP96*, 1996, p. 16A4.
- [15] Y. Liang, "Phase-correlation motion estimation," *EE392J Project report*, 2000.