

Full Length Research Paper

Statistical digital image stabilization

M. H. Shakoor* and M. Moattari

Fars Science and Research Branch, Islamic Azad University, Shiraz, Iran.

Accepted 17 February, 2011

In this paper, an efficient algorithm for the statistical digital image stabilization (SDIS) is proposed to reduce the computational cost of block matching algorithm, for local motion estimation. This method is based on calculation of statistical functions, mean and variance of pixels in each block; four blocks are used in each frame. By using the statistical functions, the best block is selected then the local motion vector (LMV) of selected block is estimated by using full search algorithm. LMV of selected block is used as global motion vector (GMV) and it is used to stabilize current frame. Full search algorithm is used for motion estimation, but instead of searching points in a block for full search, it is used in the partial distortion elimination method to terminate the improper candidate blocks and reduce computation for block matching.

Key words: Digital image stabilization, motion estimation, block matching algorithm, global motion vector, statistical functions, partial distortion elimination.

INTRODUCTION

The goal of image stabilization systems is to compensate for the position offset, caused by external variations of the camera, while excluding the effects of the undesired movement vectors by irregular conditions in the image sequence.

There are many types of image stabilization systems; one method for this purpose is digital image stabilization (DIS). The process of the DIS mainly consists of two steps: the motion estimation and the motion compensation. The task of motion estimation is used to estimate the global camera movement vector from the acquired video sequence and select the general global motion vector from several local motion vectors. The current video frame can be compensated by estimated global motion vector. The motion estimation plays an important role in the whole stabilization process. Many algorithms have been proposed to obtain fast or precise local motion vectors, such as, block matching (Vella et al., 2002), edge pattern matching (Parik et al., 1992), bit-plane matching (BPM) (Jeon et al., 1999), representative point matching (RPM) (Hsu et al., 2005), etc. Among these algorithms, block matching is precise and reliable, but it is a costly computation. The RPM can greatly reduce the complexity of computation in comparison with the other methods. However, it is sensitive to irregular

conditions such as internal moving objects, low-contrast block, and texture pattern. Therefore, the reliability evaluation is necessary to screen the undesired motion vectors for the RPM method (Hsu et al., 2005). For the purpose of increasing the speed and precision of local motion estimation vectors, a fast method for digital image stabilization based on block matching is proposed. This method uses statistical properties to select one of four blocks. It is used as statistical properties (mean and variance of pixels in each block) to select the best block from four candidates. The best block has two properties; it should not have a uniform area and it should belong to the background which does not have any motion objects in it. After the selection of the block, a method of block matching must be used to calculate LMV. When the best block is selected, LMV is estimated using block matching algorithm.

There are many methods for block matching, such as three-step search (3SS) (Koga et al., 1981), new three-step search (3SS) (Li et al., 1994) and four-step search (4SS) (Po and Ma, 1996), etc. These methods have errors because of some irregular conditions such as moving objects, intentional panning, noise, and cause error. The most accurate BMA is the full search algorithm (FS) that compares every candidate blocks in the search window with the current block. Although FS predicts the most similar block, it is computationally intensive. But instead of searching points in a block for full search, it is used in the partial distortion elimination (Park et al., 2008;

*Corresponding author. E-mail: mhshakoor@gmail.com.

Montrucchio and Quaglia, 2005) method to terminate the improper candidate blocks and reduce computation for block matching. So, instead of matching all the pixels in a block, the partial matching error can give the opportunity to make the fast, full search, possible. The partial distortion elimination (PDE) is the most popular example of the fast matching methods. PDE can easily be combined with the fast searching approach. Generally, PDE algorithm prunes the improper blocks by comparing the current minimum distortion with the partial distortion within the block. After that, LMV is calculated by full search (only in selected block). This LMV is used as GMV to apply to frame and motion compensation.

This paper is organized as follows: Local motion vector estimation, conventional fast algorithm for calculation of LMV and partial distortion elimination (PDE) are presented subsequently, after which the proposed statistical algorithm is presented for selecting the best block. Furthermore, local motion estimation is presented, before global motion estimation and motion compensation are introduced. Thus, simulation results are given before the concluding remarks are made.

BLOCK MATCHING ALGORITHMS

Block matching algorithm (BMA) is the most popular motion estimation algorithm. BMA calculates motion vector for an entire block of pixels instead of individual pixels. The same motion vector is applicable to all the pixels in the block.

This reduces computational requirement and also results in a more accurate motion vector since the objects are typically a cluster of pixels.

BMA requires very heavy computational complexity. So, there are many methods to increase the speed of BMA calculations, such as histogram (Shakoor and Moattary, 2009) or PDE (Park et al., 2008; Montrucchio and Quaglia, 2005).

In block matching, each frame is segmented into blocks. Each block is compared against candidate blocks of the previous frame, and the best one is chosen upon the minimization of some matching criteria. The best match is chosen as a predictor for the next frame.

Conventional fast algorithms

By comparing the sum of block between current and candidate blocks, many improper candidate blocks can be eliminated with less computational costs. In other words, the calculation of the sum of absolute difference (SAD) could be reduced. SAD is a very time consuming part of each block matching algorithm. If the number of search points is decreased, the number of SAD calculations is decreased and algorithm works so fast (Shakoor and Moattai, 2009) (Figure 1).

$$SAD = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |f_{cur}(j, i) - f_{ref}(j + v_x, i + v_y)| \tag{1}$$

SAD is the typical matching criterion used in local motion

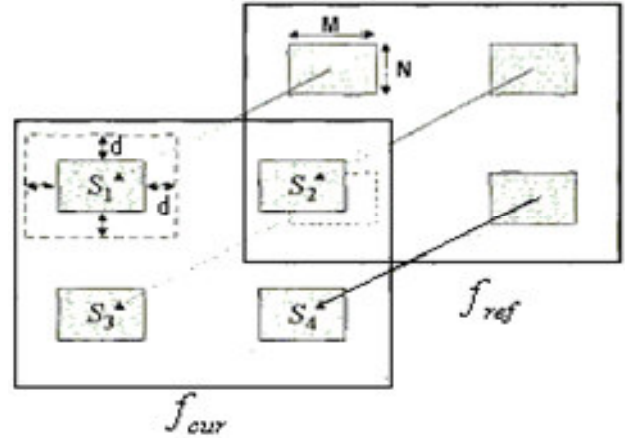


Figure 1. Four areas for the calculation of LMV.

estimation and it can be obtained through (1). N and M is the block height and width and $f_{cur}(x, y)$, $f_{ref}(x, y)$ are the pixel intensities at the current block and candidate block, respectively. i and j denote the index of horizontal and vertical directions within the block, respectively. v_x and v_y denote motion vectors of the candidate block in the reference frame. Minimum of SAD determines best block matching.

Partial distortion elimination (PDE)

One of the important methods that reduce the unnecessary computation during the SAD calculation within a block is PDE. General block matching algorithms compare the current minimum SAD (minSAD) with the final SAD, whereas PDE algorithm uses the partial sum of matching distortion to eliminate the improper candidate before finishing the matching distortion calculations. If an intermediate sum of matching distortion is larger than the minimum value of matching distortion at that time, the remaining computation for the matching distortion is unnecessary. In the PDE algorithm, the row based k-th partial SAD to check during the matching is as follows:

$$PartialSAD(k) = \sum_{i=0}^k \sum_{j=0}^{M-1} |f_{cur}(j, i) - f_{ref}(j + v_x, i + v_y)| \tag{2}$$

where $k=0, 1, \dots, N-1$

Here, in (2), M and k denote the block width and the number of rows already evaluated, respectively. If k is smaller than N and the partial SAD(k), calculated SAD until the k-th row, exceeds the current minSAD, then the remaining summation may be terminated. For example, if the Partial SAD(n) is larger than the minSAD at $k=n$, then $N-1-n$ SAD operations could be saved. This technique can be combined with the other block matching algorithms simultaneously and helps to reduce the computational load efficiently in FS algorithms.

PROPOSED ALGORITHM FOR SELECTING BEST BLOCK

Proposed algorithm has two states so it is proposed in two ways

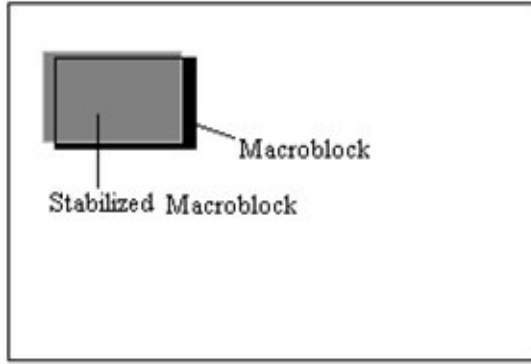


Figure 2. Block and stabilized block in an area.

and named SDIS1 and SDIS2. In these proposed method, two statistical functions are used to select the best block of four blocks. Then BMA is used only for the best block.

The first proposed algorithm for selecting best block (SDIS1)

In this method, we select four blocks of image, then calculate mean and variance (standard deviation) of pixels in each block, then, by selecting the best block and find the LMV only for this block, we can use LMV as GMV for current frame. The proposed method has flowing steps:

Step 1: Calculate mean and standard deviation of pixels in each area $[m1, m2, m3, m4]$ and $[s1, s2, s3, s4]$ for current frame (t-th frame).

Step 2: Calculate difference of mean for 2 previous frames (ms_i^{t-1} is the mean of t-1th (stabilized) frame in area i and m_i^{t-2} is the means of t-2nd frame of i-th area) for each area i: $mDiff_i = |m_i^{t-2} - ms_i^{t-1}|$. Sort the values of $[mDiff1, mDiff2, mDiff3, mDiff4]$ (in increasing order).

Step 3: Select the first mDiff (the smallest remain mean) (i-th area)

if $s^i > T1$ then select i-th macroback as best block and algorithm reach to end. Else repeat Step3 for the next area.

Step 4: If there are not any areas to satisfy condition of Step3 then select the area that has the minimum of $mDiff_i / (1 + s_i)$. T1 is a threshold value. For busy and noisy images this threshold should be large.

Hint: in Step1 Mean and Standard Deviation of the stabilized block is calculated. It does not mean current block must be shifted. According to Figure 2, it means we use some other pixels that belong to th block when it is stabilized.

If the global motion vector (GMV) of the previous frame is (v_x, v_y) :

$$m = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f^t(j - v_x, i - v_y) \quad (3)$$

$$s = \left(\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |m - f^t(j - v_x, i - v_y)| \right) / (N * M) \quad (4)$$

Best block is selected by 3 and 4 properties:

$$(1) s^i > T1 \quad (\text{in SDIS1})$$

It means, this block should not have uniform area, because, motion estimation in uniform area usually has many errors. T1 is a threshold value which should be large for busy and noisy image.

$$(2) mDiff_i < T2 \quad (\text{in SDIS2})$$

mDiff must be lower than a threshold value T2. It means, this block should not have motion object or it should belong to the background of frame. This condition is used in SDIS2.

The second proposed algorithm for selecting the best block (SDIS2)

If in SDIS1, step 2 and step 3 are changed as here, a new method named SDIS2 is proposed:

Step 2: Calculate the difference of mean of the two previous frames. For each area $mDiff_i = |m_i^{t-2} - ms_i^{t-1}|$ then sort the values of $[s1, s2, s3, s4]$ (in decreasing order).

Step 3: Select the first s (the biggest remain standard deviation) (in i-th area) if $mDiff_i < T2$ then select i-th block as the best block and algorithm reach to the end. Else repeat Step3 for the next area.

Step 4: If there is no area to satisfy the condition of Step 3 then select the area that has a max of $s_i / (1 + mDiff_i)$. T2 is a threshold value, for busy and noisy image, it should be much greater than zero, otherwise it should be near zero.

LOCAL MOTION VECTOR ESTIMATION

When the best block is selected, LMV is calculated only for it. Full search is used for block matching and for the best estimation. The PDE technique is used to eliminate many computations in SAD and to reduce the search points. In this part, local motion vector of selected area is calculated by block matching full search method and this local motion vector is used as global motion vector for next step. The advantage of this method is calculation of only one LMV instead of four LMV (Shakoor and Moattai, 2009), so, it increases the speed of operation.

GLOBAL MOTION ESTIMATION AND COMPENSATION

In many algorithms, the global motion vector is obtained by :

$$GMV^t = \text{median}\{LMV1, LMV2, LMV3, LMV4\}^t$$

GMV^t is global motion vector of t-th frame and LMV1 to LMV4 are local motion vectors of each area. In the proposed method LMV that was estimated in previous parts is used as GMV of current frame:

$$GMV^t = LMV^t \quad (5)$$



Figure 3. Bike, person, road samples.

Also, we can use this formula to smooth GMV (Accumulated GMV):

$$\text{AccGMV}^t = k1 * \text{AccGMV}^{t-1} + k2 * \text{LMV}^t \quad (6)$$

The coefficient k1 and k2 are less than 1 and depend on current GMV^t and it is different from GMV^{t-1}. At the end of operations, AccGMV is applied to current frame to stabilize it.

EXPERIMENTAL RESULT

Computer simulations have been performed on, for example, bike, person and road, which consist of some different motions (Figure 3). Person has large global motion vectors, bike has some smooth motions, but there are many internal motions in all of four areas of each frame and road has smooth global motion and there is also a moving object (car) only in a small area of each frame. In simulation part, the original image of the previous frame was used as a reference frame to

generate a motion compensated prediction image. Image quality was measured by the peak signal-to-noise ratio (PSNR) as defined in Equation (8):

$$\text{MSE} = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} [I_o(i, j) - I_c(i, j)]^2 \quad (7)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right) \quad (8)$$

lo denotes the original current frame, and Ic, denotes the motion-compensated prediction frame. Block size = 16, search region = [-7, +7]. PSNR results are listed in Table 2.

In Table 1 the numbers of search points are shown. In all examples, the proposed method searches the points only in one area, but in other cases, sum of searched point's number in all four areas are used because, in proposed method, points of the best area are used not points of all

areas. In the last row, full search is used for calculation of LMV of each four areas. The number of search point for full search in each area is 225 but by using PDE some of these points are eliminated and the number of search point is less than 225. Table 1 compares the speed of proposed methods and other methods. The quality of proposed methods is compared to stabilize the frame, so, full search algorithm and proposed algorithms have approximately the same result (median method is used for calculation of GMV from LMVs for full search in the last row). In most cases, proposed methods have similar results as full search. In this simulation, T1=0.07*15*15 and T2 = 3 for all of the samples. T1 and T2 should be changed depending on the size of motion vectors and the amount of noise that exist in images. These values of T1 and T2 are good for road sample and they should be changed for other samples to produce better PSNR.

Figures 4 and 5 show accumulated LMV and

Table 1. Number of search points in each methods (all search points in each frame).

Method	Bike	Person	Road
SDIS1	215	219	208
SDIS2	215	219	208
TSS	100	100	100
NTSS	78.4	85.6	80.4
FSS	74.2	76.5	80.8
FS	900	900	900

Table 2. PSNR of each method (GMV estimated by full search method and other methods).

Method	Bike	Person	Road
SDIS1	18.93	17.59	31.72
SDIS2	19.82	17.77	31.65
TSS	18.88	16.48	29.50
NTSS	18.90	16.53	29.48
FSS	18.88	16.56	29.38
FS (median)	20.16	18.00	32.20

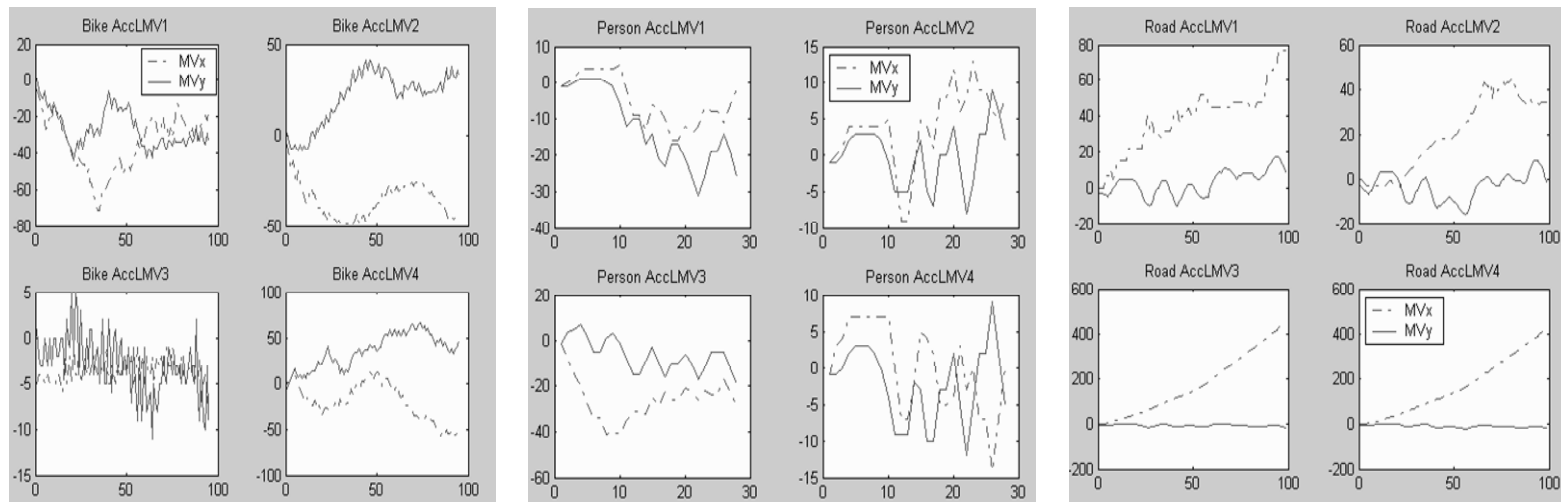


Figure 4. Result of AccLMV for SDIS2 in each area (bike, person and road samples).

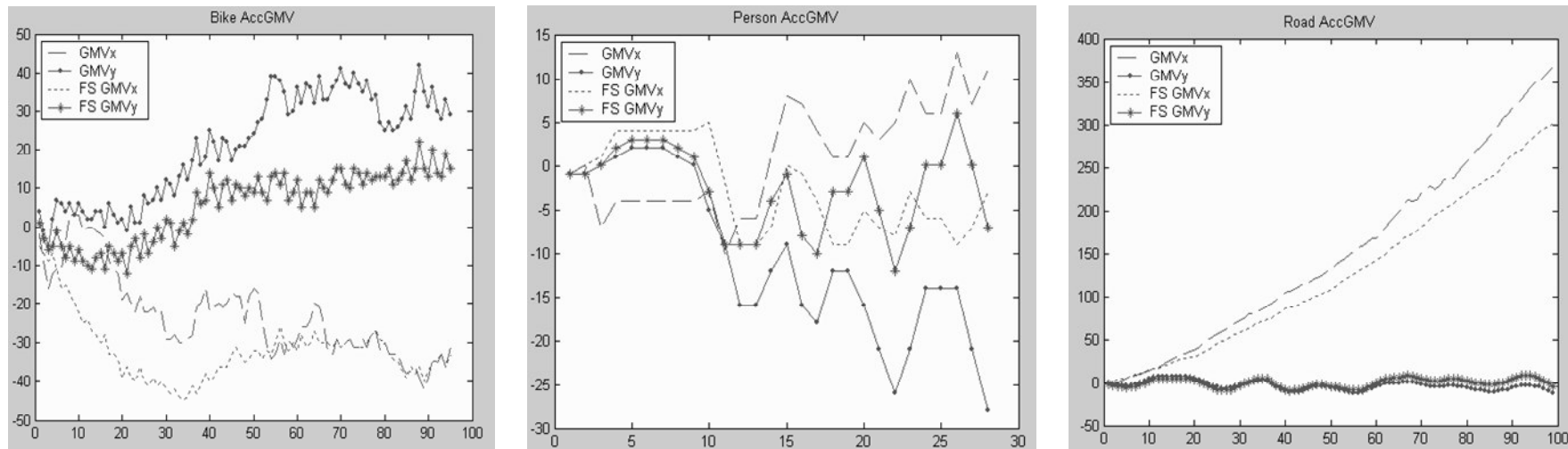


Figure 5. Result of AccGMV for SDIS2 and FS (bike, person and road samples).

Table 3. Specification of each samples.

Frame	Frames number	Size	Comment
Bike	96	256*256	Smooth motion
Person	28	320*240	Large motion
Road	99	336*224	Smooth and small motion

accumulated GMV for each sample that uses SDIS2, in these figures, accumulated vector is obtained by adding new vector to the previous accumulated vector. There are some differences between full search and the proposed plot because of the intentional panning. Some plots in Figure 4 show LMVs of each area and in many samples, they show that there are many differences between LMVs in each area. LMV of each area is calculated by FS method. But Figure 5 shows GMVs that are extracted from the best area in each frame. In some cases that intentional

panning occurred, there are some differences. In Table 3, specifications of each sample are listed

Conclusion

In this paper, a statistical digital image stabilization method is proposed that used fast block matching algorithm. A statistical method is used for selecting the best block for full search LMV. We have considered Partial SAD criteria for BMA, and also used PSNR criteria for comparing

results with traditional BMAs. The new algorithm can reduce most of the computational complexity compared to previous developed fast BMAs and has better quality than many other BMA algorithms.

ACKNOWLEDGEMENTS

We would like to thank Sang-Jun Park, Soonjong Jin and Jechang Jeong for using PDE in our paper. We also wish to thank the Islamic Azad

University, Fars Science and Research Branch for supporting this research.

REFERENCES

- Hsu SC (2005). Robust Digital Image Stabilization Technique Based on Inverse Triangle Method and Background Detection, *IEEE Trans. Consum. Electron.*, 51(2) : 335-345.
- Jeon SW (1999). Fast digital image stabilizer based on Gray-coded bit-plane matching, *IEEE Trans. Consum. Electron.*, 45(3): 598-603.
- Koga T, Iinuma K, Hirano A, Iijima Y, Ishiguro T (1981). "Motion compensated interframe coding for video conferencing," In Proc. Nat. Telecommun. Conf., New Orleans, LA, pp. G. 5.3.1-G.5.3.5.
- Li R, Zeng B, Liou ML (1994). "A new three step search algorithm for block motion estimation." *IEEE Trans. Circuits Syst. Video Technol.*, 4: 438-442.
- Messina G (2002). Digital image stabilization by adaptive block motion vectors filtering, *IEEE Trans. Consum. Electron.*, 48: 796-801.
- Montrucchio B, Quaglia D (2005). "New sorting-based lossless motion estimation algorithms and a partial distortion elimination performance analysis," *IEEE Trans. Circuit Syst. Video Technol.*, 15(2): 210-220.
- Park SJ, Jin S, Jeang J (2008). "Adaptive partial distortion search algorithm using histogram-based sorting," In Proc. ICME, pp. 833-836.
- Po LM, Ma WC (1996). "A novel four-step search algorithm for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, 6(3): 313-317.
- Shakoor MH, Moattary M (2009). "Digital Image Stabilization Using Histogram-Based Sorting," in Proc. KAM. 3: 266-268.
- Shakoor MH, Moattai M (2009). "A Fuzzy Method Based on Bit-Plane Images for Stabilizing Digital Images", *IEEE Conf. CONIELECOMP*.106-108.
- Vella F, Castorina A, Parik JK, Parrk YC, Kim DW (1992). An adaptive motion decision system for digital image stabilizer based on edge pattern matching, *IEEE Trans. Consum. Electron.*, 38(3): 607-616.