

# 10.4-3

## Low Complexity Global Motion Estimation Techniques for Image Stabilization

Felix Albu, Corneliu Florea, Adrian Zamfir, Alexandru Drimborean, *Members IEEE*  
FotoNation Romania, Bucharest, Romania

**Abstract—** Two new global motion estimation methods are proposed. The first one, called Sign Projection (SP) is obtained by modifying the integral projection estimation method using two thresholds for pixel values. The second method, called binary incrementation (BI), is obtained by using only one threshold for generating binary vectors from two images. It is shown that the proposed approaches provide similar motion estimation accuracy with the integral projection (IP) and phase correlation (PC) methods. Also, they have reduced numerical complexity and memory requirements, leading to shorter processing times as well as lower power consumption. The technique is particularly suitable for implementation in consumer devices such as digital video cameras.

### I. INTRODUCTION

The hand jitter is a biological phenomenon and the effect it has on the recorded video sequences is amplified by the small and lightweight modern cameras. To reduce the influence of the hand-jitter motion many image stabilization approaches have been proposed (the Optical Image Stabilization (OIS), the electronic image stabilization (EIS) and the digital image stabilization (DIS) techniques [1]. The proposed methods are dedicated DIS techniques for devices with limited computational power or memory (e.g. the mobile phone equipped with a digital camera).

The optimal procedure for global motion estimation is considered the full search (FS) method. However the computational complexity of the FS method is very high and alternative simpler and faster techniques have been explored. Two popular techniques used for motion estimation are: the phase correlation method (PC) [2] and the integral projection method [3]. The phase correlation method is based on the phase shift theorem [2]. In the integral projection method a pair of one-dimensional vectors is computed by summing the elements of each column or rows respectively. The projection vectors are correlated to find the best offset corresponding to the matched frames [3].

### II. THE LOW COMPLEXITY GLOBAL MOTION ESTIMATION METHODS

The proposed methods are derived from the integral projection method. They are inspired from standard methods used to determine the fundamental frequency of the speech signal [4]. For this purpose, the speech signal is “centrally clipped” in order to reduce the dynamic range of the signal and simplify the necessary hardware or software implementation. Further, the complexity is significantly

reduced if a level clipping or a limitation function is applied. Then the autocorrelation function or the average magnitude distance function (AMDF) is computed and the fundamental period of the speech signal is found [4].

#### A. The Sign Projection Method

A pair of one-dimensional vectors (horizontal and vertical) is computed from the pixel values of each frame or image plane crops. Instead of summing all elements and obtain a high final value (like in [3]), a lower value is obtained by incrementing or decrementing its value based on the result of comparison of the elements using two thresholds. We call this method the Sign Projection method.

For each row we add 1 if the value on the image plane is higher than a second threshold and add -1 if the value is lower than a first threshold. These thresholds are computed as a percentage of the dynamic range given by the maximum and minimum values of the image elements.

#### B. The Binary Incrementation Method

Further significant numerical complexity reduction can be obtained if only one threshold is used. Each pixel value of two consecutive frames is compared with only one threshold. For each line of the image, a counter counts the number of pixels above the threshold and a vector is obtained. A bit 1 is generated if each vector element is above its average (or median value), otherwise a bit 0 is generated. Two pairs of binary vectors (corresponding to horizontal and vertical computations) are obtained. For each pair of binary vectors, an exclusive OR (XOR) operation is performed between one vector and delayed versions of the other vector. The number of bits “1” shows the mismatch between the vectors for different displacement values and the minimum gives the displacement value. We call this global motion estimation method the Binary Incrementation (BI) method.

For all investigated methods, the following step involves low-pass temporal filtering. We are using the jitter improved calculation method proposed in [1]. Next, the image shift is performed using the maximum buffer area of the image sensor.

Figs. 1a and 1c show the computed horizontal projection vectors for the IP and SP methods. Figs. 1b and 1d show the corresponding SAD (sum of absolute difference) values for both methods. It can be seen from Fig. 1a and Fig. 1c that the dynamic range of the vectors of the SP method is reduced by about two orders of magnitude in comparison with those of the IP method. Fig. 1e shows the binary vectors and Fig. 1f shows the sum of XOR output values. The maximum value of the sum of the XOR output values depends only on the maximum allowed displacement value. All methods (BI, SP, IP, PC)

obtained the same simulated horizontal displacement value of 6.

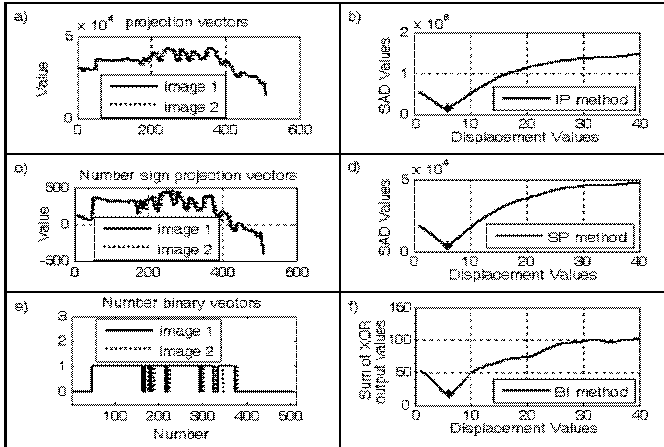


Fig. 1. a) The horizontal projection vectors for both image crops; b) The horizontal sign projection vectors for both image crops; c) The SAD values for the horizontal projection vectors; d) The SAD values for the horizontal sign projection vectors; e) The binary vectors for both image crops; f) The sum of XOR output values;

The numerical complexity and memory requirements of the SP and BI methods are reduced. The integer additions are replaced with comparisons and simple counting is used in the first phase. The IP method needs about 2 integer additions per pixel, while the complexity of the PC method is much higher, because it involves three 2-D FFT's whose complexity depends on the particular implementation. Only slightly more than one comparison per pixel is required to generate the binary vectors for the BI method. There is no need to keep the full binary matrices in the memory like the method described in [5]. The proposed methods give robust displacement estimations on sufficiently large regions of both image crops depending on the picture content (at least few hundred samples for each direction). Bigger image crops should be chosen in case of pictures with little information.

### III. EXPERIMENTS AND RESULTS

Firstly, we tested the considered methods by computing the displacements between pictures taken with a camera on burst pictures mode operation. It can be seen from Fig. 2a that the global motion values indicated by all methods are close. Next, we tested the considered algorithms on a low quality video (320 x 240 pixels). As shown in Fig. 2b there are more differences between the displacement estimates of the considered methods for the low quality movie. The IP, SP, and BI displacement estimates are more sensitive in case of luminance variation and small intra-frame movements than PC estimates. Overall, the proposed stabilization methods reduced the hand jitter effect and improved the visual quality of a large variety of video sequences. Subjective test ranked the PC stabilized movies the first, followed by SP, IP and BI stabilized movies. Also, the performance of the algorithms was evaluated using the following error measure,  $E$ :

$$E = \frac{1}{N} \sqrt{\sum_{i=1}^N \left[ (h_i^{FS} - h_i)^2 + (v_i^{FS} - v_i)^2 \right]}, \quad (1)$$

where  $(h_i^{FS}, v_i^{FS})$  represents the global motion vector obtained by the FS method and  $(h_i, v_i)$  is the global motion vector obtained by aforementioned methods. Table I shows the error values obtained by the investigated global motion algorithms for a video sequence.

TABLE I  
ERROR VALUES FOR GLOBAL MOTION VECTORS COMPUTED BY DIFFERENT METHODS

Method	PC	IP	SP	BI
$E$ for burst pictures	0.1556	0.1860	0.1664	0.1951
$E$ for movie	0.1681	0.2661	0.2088	0.3243

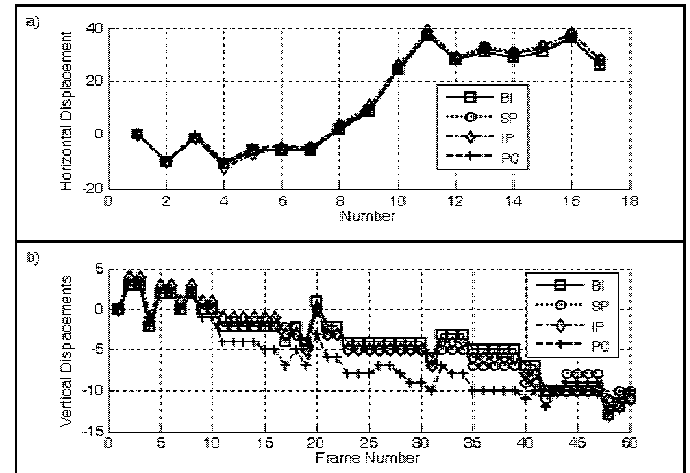


Fig. 2. a) The horizontal displacements estimated by the BI, SP, IP and PC methods for a burst of pictures. b) The vertical displacements estimated by the BI, SP, IP and PC methods for a low quality video sequence

### IV. CONCLUSIONS

The Sign Projection and Binary Incrementation methods are proposed for global motion estimation. Our experiments demonstrated that they provide similar motion estimation accuracy with other more complex global motion estimation methods. The proposed methods are obtained by central clipping and limiting the pixel values of consecutive frames or images. Due to their simplicity, these methods can have an efficient hardware or software implementation.

### REFERENCES

- [1] A. Deever, "In-camera, all-digital video stabilization", in proc. ICIS'06, Rochester, New York, May 2006, pp. 190-193.
- [2] Kuglin C D., Hines D C. The phase correlation image alignment method. In *Proc. Int. Conf. Cybernetics and Society*, IEEE, Bucharest, Romania, Sept. 1975, pp. 163-165.
- [3] K. Sauer and B. Schwartz, 1996, "Efficient Block Motion Estimation Using Integral Projections", *IEEE Trans. Circuits, Systems for video Tech.*, vol. 6, No. 5, October, pp. 513-518
- [4] J. Deller, J. Hansen, J. Proakis, *Discrete-Time Processing of Speech Signals*, 2nd. edition, Wiley-IEEE Press, 1999
- [5] B. Natarajan, V. Bhaskaran, and K. Konstantinides, "Low-complexity block-based motion estimation via one-bit transforms," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 7, no. 5, pp. 702-706, Aug. 1997.