MOTION ESTIMATION FROM MOTION SMEAR - A SYSTEM IDENTIFICATION APPROACH

Om Ji Omer , Sameer Kumar , Rajeev Bajpai , K.S.Venkatesh and Sumana Gupta

Indian Institute of Technology , Kanpur,
Department of Electrical Engineering.
e-mail: {omji , sameer , venkats , sumana}@iitk.ac.in

ABSTRACT

Motion smear, which arises because of fast motion relative to the shutter time of a camera, is generally considered as an artifact. Little work has been done to use motion smear as a visual cue for motion estimation or image restoration. Here, we present a new approach to estimate motion from two successive frames of smeared images. The blurring system is modeled as temporal integration of instantaneous images and has been estimated using *System Identification Theory*. Motion parameters have been extracted from the estimated system. As compared to earlier approaches having a similar objective, no edge detection or optical flow analysis is required. Our approach establishes a trade off between signal to noise ratio (SNR) and computational complexity. Highly accurate results have been observed with SNR as low as 12 dB. Experimental results with both simulated and real images are shown.

1. INTRODUCTION

The Human Visual System can perceive motion in motion-smeared images. This implies that motion-smear can be used in estimating the velocity of moving objects [1]. Applications like Water-Imaging, Astronomy have low illumination levels, and hence the shutter must be kept open for a longer duration to have significant SNR [2]. This results in motion-smear. Some standard motion filtering techniques have been illustrated in [3] to estimate motion from a sequence of images. To estimate motion from temporally integrated images, motion streaks have been used in [4]. In this paper, we present a simpler approach to estimate motion parameters using a System Identification Model [5].

2. LINEAR SYSTEM MODEL FOR MOTION SMEAR

2.1. Image Sequence Acquisition

In this paper, we model *motion smear* as an integration of all instantaneous linearly displaced images over a finite shutter time period. This type of model was first developed by Chen *et.al.* and is very well illustrated in [2]. Let $g(x, y, t_i, \Delta t_{i,i+1})$ be an integrated image of all instantaneous images $f(x, y; t)$ from t_i to t_{i+1}, taken by a camera with shutter time $\Delta t_{i,i+1} = t_{i+1} - t_i$. Mathematically, the motion smear is modeled as:
$g(x, y, t_i, \Delta t_{i,i+1}) =$

$$\frac{1}{\Delta t_{i,i+1}} \int_{t_i}^{t_{i+1}} f(x, y; t) l(t, t_i, \Delta t_{i,i+1}) dt \qquad (1)$$

where $l(t, t_i, \Delta t_{i,i+1})$ is the transmission coefficient(sensor gain) of the lens/filter system. We take $l(.)$ as unity over the shutter opening duration, which is true for common cameras. Smeared images $g(x, y, t_0, \Delta t_{01})$ and $g(x, y, t_1, \Delta t_{12})$ are obtained by opening the shutter from time t_0 to t_1 and t_1 to t_2 respectively, and summed up as :

$$s(x, y, t_0, \Delta t_{02}) = g(x, y, t_0, \Delta t_{01}) + g(x, y, t_1, \Delta t_{12}) \qquad (2)$$

2.2. Mathematical Formulation

In general, the *Point Spread Function* $h_{i,i+1}(x, y)$ is space-varying in nature over the entire image, but we may choose it to be locally space-invariant assuming motion to be uniform in that local region. For constant translational motion with velocity $\vec{v} = (v_x, v_y)$ during the interval $[t_i, t_{i+1}]$, $i = 0, 1$, we have

$$f(x, y; t_1) = f(x, y; t_0) \star \delta(x - L_x, y - L_y) \qquad (3)$$

where $L_x = v_x(\Delta t_{01})$ and $L_y = v_y(\Delta t_{01})$. Here, positive L_x and L_y correspond to vertically downward and horizontally rightward motion respectively. We can write shorthand notations for equation (2) as:

$$s_{02} = g_{01} + g_{12} \qquad (4)$$

$$= f_0 \star h_{01} + f_1 \star h_{12}$$

$$= f_0 \star h_{01} + f_0 \star h^{'} \star h_{12}$$

0-7803-8554-3/04/$20.00 ©2004 IEEE.

where $h^{'} = \delta(x - L_x, y - L_y)$.

By taking both the exposure durations Δt_{01} and Δt_{12} to be equal, *i.e.* $h_{01} = h_{12}$, we have

$$s_{02} = g_{01} \star (1 + h^{'}) \tag{5}$$

In frequency domain, we can write:

$$S_{02}(\Omega_x, \Omega_y) = H(\Omega_x, \Omega y) G_{01}(\Omega_x, \Omega_y) \tag{6}$$

$$H(\Omega_x, \Omega_y) = 1 + \exp(-j\Omega L) \tag{7}$$

where $\Omega L = \Omega_x L_x + \Omega_y L_y$.

Equation (6) indicates that s_{02} can be obtained by passing g_{01} through a linear system $h(x, y, \vec{v})$. Assuming local space-invariance, the motion estimation problem can be framed as a well known *System Identification Problem*.

$$H_{01}(\Omega_x, \Omega_y) = H_{12}(\Omega_x, \Omega_y) = \frac{\sin(\frac{\Omega L}{2})}{\frac{\Omega L}{2}} \exp(\frac{-j\Omega L}{2}) \tag{8}$$

Here, the zeros of H_{01} make image restoration error prone. This can be tackled by taking sensor gain to be triangular, but same, over the periods $[t_0, t_1]$ and $[t_1, t_2]$ [2].

2.3. Parametrization of the Transfer Function

In practice, only digitized images are available. Hence, we define $H(\omega_1, \omega_2), S_{02}(\omega_1, \omega_2)$ and $G_{01}(\omega_1, \omega_2)$ as the discrete time Fourier transforms of $h_{01}(k, l), s_{02}(k, l)$ and $g_{01}(k, l)$ respectively. These are digital representation of h_{01}, s_{02} and g_{01}, respectively. In order to estimate motion, we approximate the unknown transfer function $H(\omega_1, \omega_2)$, which is of polynomial form (equation(7)), with a non-causal polynomial transfer function having coefficients as b_{mn}'s, expressed as:

$$B(\omega_1, \omega_2) = \sum_{m=-M}^{M} \sum_{n=-M}^{M} b_{mn} \exp(-j\omega_1 m) \exp(-j\omega_2 n) \tag{9}$$

$B(\omega_1, \omega_2)$ is of Filter Order $2M$. We define M as the *System Order*.

3. THE ALGORITHM

3.1. Estimation of the Transfer Function Coefficients

We use the least mean squares error criterion to estimate the b_{mn}'s. Equation error algorithms for the minimization of the quadratic cost function are developed. Using equation (6), we reconstruct $S_{02}(\omega_1, \omega_2)$ as:

$$\hat{S}_{02}(\omega_1, \omega_2) = G_{01}(\omega_1, \omega_2) B(\omega_1, \omega_2) \tag{10}$$

We now define the error signal [6] as

$$e(k, l) = s_{02}(k, l) - \hat{s}_{02}(k, l) \tag{11}$$

Let

$\mathbf{a^T} = [0\ 0 \ldots\ldots 0\ 1\ 0 \ldots\ldots 0\ 0]$,

$\mathbf{b}^T = [b_{-M,-M} \ldots b_{M,-M} \ldots b_{0,0} \ldots b_{-M,M} \ldots b_{M,M}]$

$\mathbf{g_{01}}^T(k, l) = [g_{01}(k + M, l + M) \ldots g_{01}(k - M, l + M) \ldots g_{01}(k, l) \ldots g_{01}(k + M, l - M) \ldots g_{01}(k - M, l - M)]$

$\mathbf{s_{02}}^T(k, l) = [s_{02}(k + M, l + M) \ldots s_{02}(k - M, l + M) \ldots s_{02}(k, l) \ldots s_{02}(k + M, l - M) \ldots s_{02}(k - M, l - M)]$

Using equation (10) in spatial domain, we can write equation (11) as:

$$e(k, l) = \begin{bmatrix} -\mathbf{b}^T \mathbf{a}^T \end{bmatrix} \begin{bmatrix} \mathbf{g_{01}}(k, l) \\ \mathbf{s_{02}}(k, l) \end{bmatrix} \tag{12}$$

The variance of the error signal given by equation (12) is easily found by:

$$E[e^2(k, l)] = \begin{bmatrix} -\mathbf{b}^T \mathbf{a}^T \end{bmatrix} \begin{bmatrix} \mathbf{R_{gg}}(k, l) & \mathbf{R_{gs}}(k, l) \\ \mathbf{R_{gs}}^T(k, l) & \mathbf{R_{ss}}(k, l) \end{bmatrix} \begin{bmatrix} -\mathbf{b} \\ \mathbf{a} \end{bmatrix} \tag{13}$$

where

$$\begin{aligned} \mathbf{R_{gg}}(k, l) &= E[\mathbf{g_{01}}(k, l) \mathbf{g_{01}}^T(k, l)] \\ \mathbf{R_{gs}}(k, l) &= E[\mathbf{g_{01}}(k, l) \mathbf{s_{02}}^T(k, l)] \\ \mathbf{R_{ss}}(k, l) &= E[\mathbf{s_{02}}(k, l) \mathbf{s_{02}}^T(k, l)] \end{aligned} \tag{14}$$

Here $E(.)$ is the expectation operator and $\mathbf{R_{gg}}, \mathbf{R_{gs}}, \mathbf{R_{ss}}$ are the signal correlation matrices. Now minimizing equation (13) with respect to $\mathbf{b}$ gives

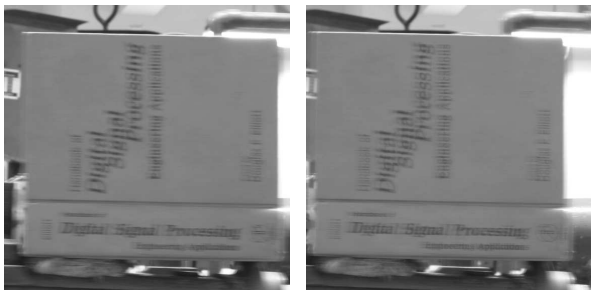
$$\mathbf{b} = \mathbf{R_{gg}^{-1}} \mathbf{R_{gs}} \mathbf{a} \tag{15}$$

3.2. System Order and Window Size

In order to estimate $\mathbf{b}$ uniquely, we assume that the input image over a local region around a given point is stationary, ergodic and persistently exciting of sufficient degree, such that $\mathbf{R_{gg}}$ is invertible. The image being ergodic and stationary ensures that the expected values in equation (14) at point (k, l) can be calculated by taking the average of each of $\mathbf{g_{01}} \mathbf{g_{01}}^T$, $\mathbf{g_{01}} \mathbf{s_{02}}^T$, and $\mathbf{s_{02}} \mathbf{s_{02}}^T$ separately in a region of size, say $(2M_1 + 1) \times (2M_1 + 1)$, around the point (k, l).

Here, it should be noted that for $\mathbf{g_{01}}$ and $\mathbf{s_{02}}$ at point (p, q), we require an image region $(p - M : p + M, q - M : q + M)$. Moreover, to compute the expected values at (k, l), we have to calculate $\mathbf{g_{01}}(p, q)$ and $\mathbf{s_{02}}(p, q)$ at each point of the region $(k - M_1 : k + M_1, l - M_1 : l + M_1)$. Hence the region of operation (Window Size, say $W \times W$, of the image) should be $(2M_1 + 2M + 1) \times (2M_1 + 2M + 1)$.

The matrix $\mathbf{g_{01}} \mathbf{g_{01}}^T$ is of rank one, and matrix $\mathbf{R_{gg}}$ (of dimension $(2M + 1)^2 \times (2M + 1)^2$) is the average value of $(2M_1 + 1)^2$ such matrices. So $(2M_1 + 1)^2$ should be greater than or equal to $(2M + 1)^2$ (*i.e.* $M_1 \geq M$) to get invertible $\mathbf{R_{gg}}$. Hence we should keep window size $W \geq (4M + 1)$.



(a) Image g_{01} (b) Image $s_{02} - g_{01}$

Fig. 1. Real Smeared Images: (a) First Smeared Image (b) Second Smeared Image.

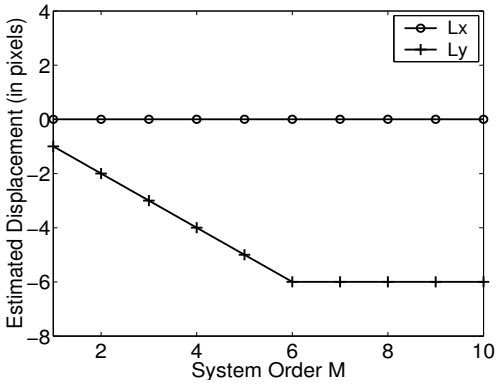


Fig. 2. Estimated Displacements at different System Orders for Down-Sampled Real Images.

From equations (7) and (9), we can deduce that the System Order M should be equal to or larger than the displacement in pixels of the moving object. But the computational complexity increases (order $\geq O(N^2)$) with the System Order M. Moreover, an increase in the System Order M results in an increase in the Window Size W, which conflicts with the assumption of space invariance over a local region. Hence, an optimal System Order should be chosen to efficiently estimate the motion.

3.3. Estimating the Displacements

Due to presence of unity term in $H(\omega_1, \omega_2)$ (equation (7)), we modified the coefficient b_{00} by subtracting one from it. Now we can deduce from (7) and (9) that, of all the b_{mn}'s, only $b_{L_x L_y}$ will be non-zero. Although, in case of real images, this may not hold due to the presence of noise, still $b_{L_x L_y}$ will be maximum over all other b_{mn}'s, above a certain signal to noise ratio (SNR). Hence we estimate (L_x, L_y) as:

$$(\hat{L}_x, \hat{L}_y) = (m, n) : b_{mn} > b_{ij}, \forall (i, j) \neq (m, n)$$

4. RESULTS: REAL IMAGES

A book was kept on a platform moving horizontally leftward with a constant velocity. Successive images of the cover-page of the book were taken with a camera having a fixed exposure time.

Two successive smeared images (566×566) were taken for estimation as shown in Fig.1. Since the observed images had a large displacement, they were down-sampled by a factor of four, hence reducing computational time by choosing smaller M. A window size of 61×61 was chosen to estimate displacements L_x and L_y for different System Orders M. Referring to Fig.2, stable values are attained for $M \geq 6$, and we take these values as our estimate for L_x and L_y. The coefficients b_{mn}'s, obtained for System Order M=8, are shown in Fig.3. The coefficient of highest intensity is at $(0, -6)$ which corresponds to $(\hat{L}_x, \hat{L}_y)$.

A finer estimation is done over original images (without down-sampling) after shifting $g_{01}(k, l)$ by $(0, -24)$, estimated from down-sampled images. The final estimated displacements are shown in Fig.4.

Sometimes, it is difficult to take two consecutive shots without any delay. In such cases, where the duration between the successive shots, δt (*i.e.* off time), is not negligible as compared to the exposure time of the camera, Δt, we should modify the estimated displacement $\vec{L}$ by $\vec{L'}$ as:

$$\vec{L'} = \vec{L}(\frac{\Delta t}{\Delta t + \delta t})$$

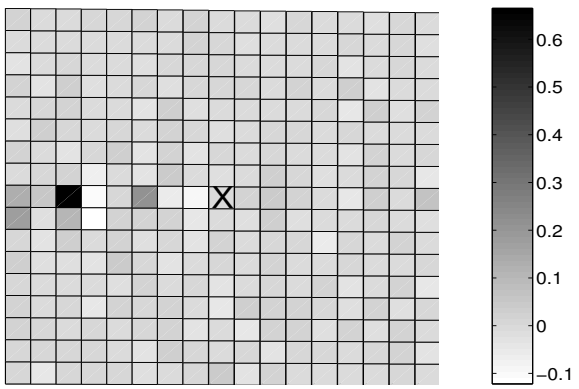


Fig. 3. Modified Transfer Function Coefficients b_{mn}'s for System Order M=8 using Down-Sampled images. "$\times$" in the center represents b_{00}.

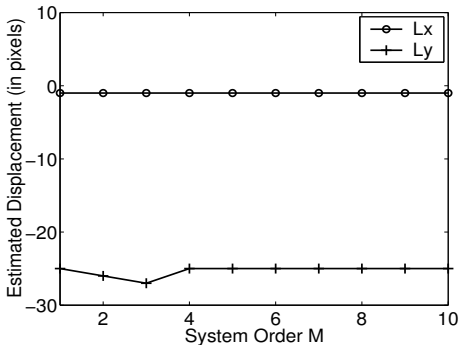


Fig. 4. Estimated Displacements at different System Orders for Original Images using Fine Estimation.

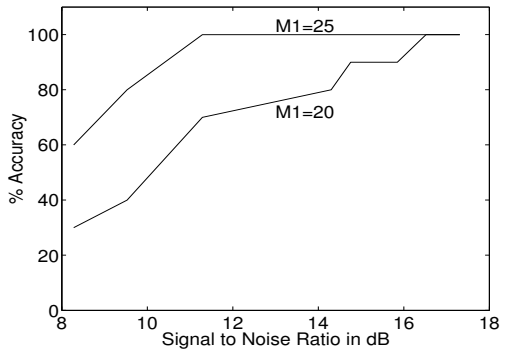


Fig. 6. % Accuracy obtained for System Order $M = 6$ with $M_1 = 20$ (W=53) and $M_1 = 25$ (W=63)

5. NOISE PERFORMANCE

The algorithm, presented in this paper, can also be applied to single frame (not smeared) images g_{01} and g_{12}. The only condition is that g_{01} and g_{12} should be linearly displaced with respect to each other. Hence, to illustrate the noise performance, we took a 512×512 Lena image as $g_{01}(k, l)$. This image was shifted by 4 pixels along the horizontal direction, and 3 pixels along the vertical direction, and the resulting image was added to $g_{01}(k, l)$ to obtain $s_{02}(k, l)$.

Additive White Gaussian noise was deliberately added to the images ten times corresponding to different SNR's. For this we used the random noise generator of MatLab Version 6.0.0.88. The noisy images are shown in Fig.5. %Accuracy is calculated as the ratio of number of correct estimations to the total number of simulations (=10) for each SNR.

(a) Image g_{01} (b) Image s_{02}

Fig. 5. AWGN noise added to the Lena Image (SNR = 14.3 dB): (a) first noisy image and (b) smeared noisy image.

As mentioned earlier (Section 3.2), $\mathbf{R_{gg}}$ and $\mathbf{R_{gs}}$ are calculated by averaging over the region of size $(2M_1 + 1) \times (2M_1 + 1)$. Hence the effect of noise can be reduced significantly by choosing larger M_1 (Fig.6).

6. CONCLUSION

This paper presents a new technique, based on *System Identification Theory*, to estimate motion parameters from motion smeared images of objects moving with a uniform linear velocity along the image plane. The extraction of motion parameters from transfer function coefficients can be improved by applying optimal statistics for better accuracy in presence of noise, even for smaller window size. In addition, this approach can also be applied to rotational motion, translational motion (perpendicular to image plane), and also defocus blur, *i.e.* estimating the 3-D motion.

7. REFERENCES

[1] David Burr, "Motion smear," *Nature*, vol. 284, pp. 164–165, March 1980.

[2] N . Nandhakumar W.G.Chen and W.Martin, "Image motion estimation from motion smear -a new computational model," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, April 1996.

[3] A. Murat Tekalp, "Digital video processing," *Prentice Hall, PTR*, 1995.

[4] Barry Sheppard Daniel Majchrzak, Sudeep Sarkar and Robin Murphy, "Motion detection from temporally integrated images," *ICPR*, vol. 3, pp. 3844–3847, 2000.

[5] J.D.Rayala, "Estimation of depth for monocular defocused images," *Ph.D Thesis , Department of Electrical Engineering ,IIT Kanpur*, May 1997.

[6] Rajeev Bajpai, "Estimation of motion and depth from smeared images," *M.Tech Thesis , Department of Electrical Engineering, IIT Kanpur*, 2001.