ROBUST FOCUSED IMAGE ESTIMATION FROM MULTIPLE IMAGES IN VIDEO SEQUENCES

*Junlan Yang, Dan Schonfeld**

Multimedia Communications Lab
ECE Dept., University of Illinois
Chicago, IL, 60607
{jyang24,dans}@uic.edu

Magdi Mohamed

Physical Realization Research Center of Excellence
Motorola Labs,
Schaumburg, IL, 60196
Magdi. Mohamed@motorola.com

ABSTRACT

In this paper, we propose a novel technique for estimating focused image sequences captured by an out-of-focus camera. The basic concept used in the proposed algorithm employs multiple images with different blurs from the video sequence to estimate the blur model and construct a more robust method for focused image estimation. This algorithm can be used to correct focusing errors in the video capture process. It could potentially be used to replace the expensive apparatus required for auto-focus adjustments by miniature engines in many camera devices. We demonstrate the successful performance of our approach to focused image estimation through computer simulation experiments using various blur models.

Index Terms— Focusing, image restoration.

1. INTRODUCTION

Focusing is an important issue in digital camera design. Current auto-focus solutions widely applied in industry are mostly based on different focus measures. They search for the best focused images while moving the lens and can be tuned to perform fast. The shortage is that they require a focal-length changing lens and an accurate engine that can move the lens with a particular step size. From a different point of view however, image processing solutions model the out-of-focus phenomenon as focused images passing through a linear system. With the estimation of the point spread function (PSF) of the lens system, the focused images can be recovered through a deconvolution process. Discussion in this paper will lie mainly within this class and concentrate on PSF estimation.

The overall philosophy of estimating PSF and its Fourier Transform, also known as optical transfer function (OTF), is based on a fundamental observation, that is the blur characteristic relates only to the object depth and the camera settings. Despite the fact that the relationship is usually approximated by first order optics, this observation verifies itself through the success of depth-from-defocus (DFD) algorithms. For example, designer in [1] utilizes two settings of camera parameters for acquiring two differently blurred images. Assuming the PSF to be a Gaussian function, a close form solution of the blur parameter is given. More generally, the authors of [2] approximate the underlying OTF by a parametric polynomial and estimate the coefficients using a least-square criteria. In a more recent work [3], the technique of DFD is combined with stereo pairs and the estimation is performed with the tool of Markov random fields to improve the accuracy.

DFD for estimating the PSF has a solid and elegant theoretical foundation however, it poses a high requirement on the hardware. Due to the fact that changing camera settings such as camera aperture and focal length cannot be done without sophisticated experimental device, it limits the applications in practice. The algorithm proposed in this paper, on the other hand, is designed for a 'rigid' camera whose physical parameters are all fixed. Therefore it can be applied to simple digital cameras especially mobile-phone cameras. The other novelty of our algorithm is to exploit multiple images taken by a moving camera. Multiple images taken in variable positions not only provide differently blurred images but also reveal additional resources for improving estimation.

The rest of this paper is organized as follows. In Section 2, we begin with the problem definition and model formulation. In Section 3, we explain the main idea of blur estimation through three examples of PSF. We then discuss in Section 4 the idea of multiple-image estimation and in Section 5 noise analysis for the system. Section 6 provides the simulation results and Section 7 draws the conclusion.

2. CAMERA AND IMAGING MODEL

Assume a moving camera is looking at a static object and taking a video of it. The camera is a rigid camera, meaning that it has a fixed lens aperture, focal length and image plane-to-lens distance. One point in the object projects onto different image coordinates when the camera moves. In time t and time t', the camera takes two images, frame k and frame k'. The pixel lo-

*The authors would like to thank the Physical Realization Research Center of Excellence, Motorola Labs, for supporting this work.

cations (x_0, y_0) in image frame k and (x_1, y_1) in frame $k^{'}$ are related by a 2D affine transform [4]:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \frac{z_0 - f}{z_1 - f} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (1)$$

where $r_{11}, r_{12}, r_{21}, r_{22}$ are rotation parameters and t_x, t_y are translation parameters of the camera motion. z_0 and z_1 are distances between the object and the camera lens, in time t and time $t^{'}$ respectively, which are commonly referred as depths of the object. f is the focal length.

Define $s \equiv \frac{z_0 - f}{z_1 - f}$. Denote the Fourier transform of frame k and frame $k^{'}$ as $F_0(u, v)$ and $F_1(u, v)$. According to the affine theorem for 2D Fourier transform [5], the amplitude of $F_0(u, v)$ and $F_1(u, v)$ has the following relationship:

$$|F_1(u, v)| = \frac{1}{|\Delta|} |F_0(\frac{sr_{22}u - sr_{21}v}{\Delta}, \frac{-sr_{12}u + sr_{11}v}{\Delta})|. \quad (2)$$

where $\Delta \equiv s^2(r_{11}r_{22} - r_{12}r_{21})$. Using the motion estimation and stabilization technique in [4], we can compensate for rotation before we further process the images. Therefore we only discuss here when the rotation matrix is identity. Then (2) can be simplified as

$$|F_1(u, v)| = \frac{1}{s^2} |F_0(\frac{u}{s}, \frac{v}{s})|. \quad (3)$$

When a camera is out of focus, the resulting image is blurred by a specific PSF, whose parameters are uniquely determined by the blur radius R. In frequency domain, the spectrum of blurred image $Y(u, v)$ will be the original spectrum times the OTF $H(u, v, R)$.

$$Y_i(u, v) = F_i(u, v) H(u, v, R_i), \qquad i = 0, 1. \quad (4)$$

Here we consider the PSF being a symmetric (even) function whose Fourier transform is real. With (3) and (4), we have

$$s^2 |Y_1(u, v)| = |Y_0(\frac{u}{s}, \frac{v}{s})| \frac{H(u, v, R_1)}{H(u/s, v/s, R_0)}. \quad (5)$$

To proceed, we will incorporate the knowledge from optic geometry. As illustrated in [1], the blur radiuses are given by

$$R_i = vL(\frac{1}{f} - \frac{1}{z_i} - \frac{1}{v}), \qquad i = 0, 1. \quad (6)$$

where v is the image plane-to-lens distance, L is the radius of lens aperture, and z is the depth of the object. It can be seen that the blur radius is affected only by the depth of the object for one particular camera. From the definition of s we can continue to arrive at

$$s = \frac{z_0 - f}{z_1 - f} = \frac{R_0 + L}{R_1 + L} \times \frac{R_1 + L - vL/f}{R_0 + L - vL/f}. \quad (7)$$

To estimate the focused images from the blurred images, we need to estimate the OTF, which equals identifying the blur radiuses. With v, L, f being known camera parameters, it is able to solve for s, R_0, R_1, thus $H(u, v, R_0)$ and $H(u, v, R_1)$, based on (5) and (7).

3. BLUR PARAMETER ESTIMATION AND IMAGE RECONSTRUCTION

In this section, we will discuss our algorithm for three types of PSF. In all the cases, we begin with assuming the energy conservation constraint, which means $H(0, 0, R) = 1$. Thus, s can be solved by noticing the DC components in (5) yields

$$s = \sqrt{Y_0(0, 0)/Y_1(0, 0)}. \quad (8)$$

3.1. Gaussian Blur Model

When PSF takes the form of a Gaussian function, we have:

$$H(u, v, R) = exp\{-\frac{1}{2}(u^2 + v^2)\sigma^2\}, \quad (9)$$

where $\sigma \approx R/\sqrt{2}$. Therefore, (5) becomes

$$s^2 \frac{|Y_1(u, v)|}{|Y_0(\frac{u}{s}, \frac{v}{s})|} = exp\{-\frac{1}{4}(u^2 + v^2)(R_1^2 - R_0^2/s^2)\}. \quad (10)$$

Above is true for all u, v so an averaged solution is suggested in [1] as well as the following:

$$c \equiv \frac{1}{A_1} \int \int_{I_1} \frac{-4}{u^2 + v^2} ln(s^2 \frac{|Y_1(u, v)|}{|Y_0(\frac{u}{s}, \frac{v}{s})|}) dudv, \quad (11)$$

$$R_1^2 - R_0^2/s^2 = c, \quad (12)$$

where I_1 is the region within which the integral is well-defined and A_1 is the area of I_1. With (7), (8) and (12), we can solve R_0 and R_1 uniquely. Here gives a solution with approximation based on the fact that $v \approx f$ and $L \gg R$. We can get the simplified version of (7) as $R_1 = sR_0$, so that $R_1^2 - R_0^2/s^2 = R_0^2(s^2 - 1/s^2) = c$. Hence, $R_0 = \sqrt{\frac{cs^2}{s^4 - 1}}$. This approximation avoids measuring v, L, f and is found to be accurate enough in experiments.

3.2. Geometric Blur Model

According to geometric optics, the first order approximation of the PSF takes the form of a cylindrical function in the case of a circular aperture. Therefore we have

$$H(u, v, R) = 2 \frac{J_1(R\sqrt{u^2 + v^2})}{R\sqrt{u^2 + v^2}}. \quad (13)$$

We adopt the polynomial expansion [6] of a bessel function,

$$J_1(x) = \frac{x}{2} - \frac{x^3}{2^2 \cdot 4} + \frac{x^5}{2^2 \cdot 4^2 \cdot 6} - \frac{x^7}{2^2 \cdot 4^2 \cdot 6^2 \cdot 8} + ..., \quad (14)$$

and (5) becomes

$$s^2 \frac{|Y_1(u, v)|}{|Y_0(\frac{u}{s}, \frac{v}{s})|} = 1 + a_1(u^2 + v^2) + a_2(u^2 + v^2)^2 + ...,$$

$$a_1 = -\frac{1}{8}(R_1^2 - R_0^2/s^2); a_2 = \frac{1}{192}(R_1^4 - R_0^4/s^4) + \frac{R_0^2}{8s^2} a_1; ...$$

If we can identify $a_n, n = 1...N$, we can solve for R_0 and R_1 with (7) and (8). N is the number of coefficients we plan to identify. Theoretically, identifying only a_1 is enough. However more coefficients are desired for a reliable solution. Thus the identification problem equals solving the matrix equation:

$$\begin{bmatrix} Z(u_0, v_0) \\ Z(u_0, v_1) \\ ... \end{bmatrix} = \begin{bmatrix} 1 & u_0^2 + v_0^2 & (u_0^2 + v_0^2)^2 & ... \\ 1 & u_0^2 + v_1^2 & (u_0^2 + v_1^2)^2 & ... \\ 1 & ... & ... & ... \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ ... \\ a_N \end{bmatrix} \quad (15)$$

where $Z(u,v) \equiv s^2 \frac{|Y_1(u,v)|}{|Y_0(\frac{u}{s}, \frac{v}{s})|}$. The LHS vector is $K \times 1$, K is the number of non-zero frequency components being used. The matrix in RHS is of size $K \times N$, and the vector in RHS is the unknown vector of size $N \times 1$. An least-square solution of $[a_1, ...a_N]$ will give us an overdetermined equation array for solving R_0 and R_1.

3.3. Arbitrary Polynomial Approximation

When we have no prior knowledge of PSF, we can approximate the OTF by an arbitrary Mth order polynomial. Consider:

$$H(u, v, R_0) = 1 + \sum_{n=1}^{M} b_n (u^2 + v^2)^n;$$
$$H(u, v, R_1) = 1 + \sum_{n=1}^{M} c_n (u^2 + v^2)^n.$$

According to [2], we have a constraint on b_1:

$$[\frac{\partial^2 H(u, v, R_0)}{\partial u^2} + \frac{\partial^2 H(u, v, R_0)}{\partial v^2}]|_{u=v=0} = 4b_1 = -\frac{R_0^2}{2}.$$

Similar constraint applies to c_1. One can verify that (13) satisfies this constraint. Therefore, we have

$$\frac{H(u, v, R_1)}{H(u/s, v/s, R_0)} = \frac{1 + \sum_{n=1}^{M} c_n (u^2 + v^2)^n}{1 + \sum_{n=1}^{M} b_n (u^2 + v^2)^n / s^n} = 1 + \sum_{n=1}^{N} a_n (u^2 + v^2)^n, \quad (16)$$

with $b_1 = -\frac{1}{8} R_0^2$ and $c_1 = -\frac{1}{8} R_1^2$. $a_n, n = 1, ..., N$ can be solved from (15). Here we define

$$\mathbf{b} \equiv [b_1/s, ...b_n/s^n, ...b_M/s^M]^T, \mathbf{c} \equiv [c_1, ...c_n, ...c_M]^T;$$
$$\mathbf{a}^{(1)} \equiv [a_1, ..., a_n, ..., a_M]^T, \mathbf{a}^{(2)} \equiv [a_{M+1}, ..., a_N]^T;$$

As long as $N \geq 2M$, a close form solution of **b** and **c** can be given by [7]:

$$\mathbf{c} = -\mathbf{A}^{-1}\mathbf{a}^{(2)}; \qquad \mathbf{b} = \mathbf{c} - \mathbf{K}\mathbf{a}^{(1)}, \quad (17)$$

where

$$\mathbf{A} \equiv \begin{bmatrix} a_M & ... & a_1 \\ a_{M+1} & ... & a_2 \\ ... & ... & ... \\ a_N & ... & a_M \end{bmatrix}; \mathbf{K} \equiv \begin{bmatrix} 1 & & & \\ c_1 & 1 & \mathbf{O} & \\ ... & ... & ... & ... \\ c_{M-1} & c_{M-2} & ... & 1 \end{bmatrix}.$$

Once we get the estimation of the transfer function, we can process the degraded image with an inverse filter or a Wiener filter to recover the focused image.

4. MULTIPLE IMAGE ESTIMATION

All above discussion is based on two frames. We can also use three or more frames for estimation, in order to improve the accuracy and robustness. For instance, in the case of the Gaussian PSF, if we have a third image with spectrum $Y_2(u, v)$, we can form another set of equations:

$$s'^2 \frac{|Y_2(u,v)|}{|Y_0(\frac{u}{s'}, \frac{v}{s'})|} = exp\{-\frac{1}{4}(u^2 + v^2)(R_2^2 - R_0^2/s'^2)\};$$
$$s' = \sqrt{Y_0(0,0)/Y_2(0,0)};$$
$$R_2 = s' R_0; \quad (18)$$

using the simplified version of (7). Along with (10), (8) and $R_1 = sR_0$, we have six equations for five unknowns. It is overdetermined, which enables us to use information from three frames to form one estimation. Define

$$w \equiv \frac{1}{A_2} \int\int_{I_2} \frac{-4}{u^2 + v^2} ln(s^2 \frac{|Y_1(u,v)|}{|Y_0(\frac{u}{s}, \frac{v}{s})|} s'^2 \frac{|Y_2(u,v)|}{|Y_0(\frac{u}{s'}, \frac{v}{s'})|}) dudv; \quad (19)$$

where the meanings of A_2 and I_2 are the same as in (11). The the estimation of R_0 is then given by

$$R_0 = ss' \sqrt{\frac{w}{s^2(s'^4 - 1) + s'^2(s^4 - 1)}}. \quad (20)$$

As we can see in the simulation results, the estimation based on three images improves the performance of the algorithm.

5. ERROR ANALYSIS

The performance of our estimation algorithm can be evaluated by introducing an additive noise in the model:

$$Y_i(u, v) = H(u, v, R_i) F_i(u, v) + N_i(u, v); \quad i = 0, 1. \quad (21)$$

Using the Gaussian blur model and simplified version of (7) as in Section 3.1, we can give the estimation of OTF for first image with presence of noise as:

$$\hat{H}(u, v, R_0) = [s^2 \frac{Y_1(u,v) - N_1(u,v)}{Y_0(u/s, v/s) - N_0(u/s, v/s)}]^{\frac{s^2}{s^4 - 1}}.$$

The estimation of the focused image is given by $\hat{F}_0(u,v) = Y_0(u,v)/\hat{H}(u,v,R_0)$, while the noise free estimation is $F_0(u,v) = Y_0(u,v)/H(u,v,R_0)$. This makes us arrive at the noisy version of focused image estimate as:

$$\hat{F}_0(u,v) = F_0(u,v)[\frac{Y_1(u,v) - N_1(u,v)}{Y_0(\frac{u}{s}, \frac{v}{s}) - N_0(\frac{u}{s}, \frac{v}{s})} \cdot \frac{Y_0(\frac{u}{s}, \frac{v}{s})}{Y_1(u,v)}]^{\frac{-s^2}{s^4 - 1}}. \quad (22)$$

Notice that the original additive noise becomes multiplicative noise in the final estimation. The statistical characteristic of the noise also changes. The random variable inside the square bracket is the ratio of two normal random variables with non-zero mean and its distribution has been studied in [8]. Based on that, we can also give the distribution and expectation of the noise.

6. SIMULATION RESULTS

We test the effectiveness of our algorithm in images and in video sequences captured by a digital camcorder. Here we report some of the experimental results.

Fig.1 shows (a) the original test image, (b) the image blurred by a Gaussian PSF and (c) the reconstructed image. (The second set of images is not shown here due to space limit). It can be seen that the blur effects are removed in the reconstructed image. The true blur parameter σ is 0.0354, and our estimation is 0.0361, which is a close estimation.

Fig.2 (a) shows the frame 1, frame 30, frame 60 and frame 90 of an indoor scene sequence. In this sequence, the camera is heading forward along the optical axis. Fig.2 (b) shows the resulting degraded frames blurred with the geometric model. Fig.2 (c) is the estimation of the focused frames. It can be noticed that the estimation is close to the original one.

Fig.3 (a) shows the frame 20, 50, 80 and 110 of same indoor sequence as in Fig.2, and Fig.3 (b) shows the sequence blurred by a Gaussian PSF instead. Fig.3 (c) is the estimation of the focused frames using two frames while Fig.3 (d) is the estimation using three frames. As we can see, three-frame estimation performs steadier than the two-frame one, especially in the first two frames where the latter makes mistakes.

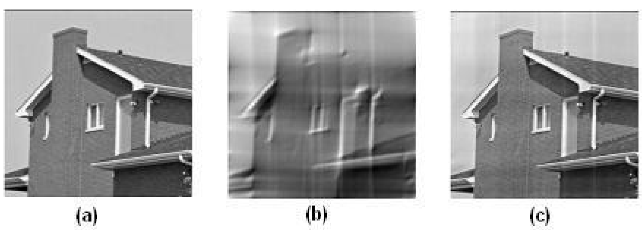

Fig. 1. (a) Original test image, (b) blurred image and (c) reconstructed image.

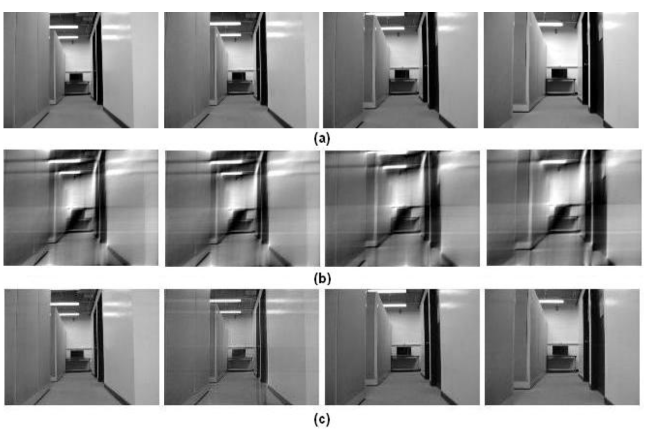

Fig. 2. (a) Original test sequence, (b) blurred sequence and (c) reconstructed sequence.

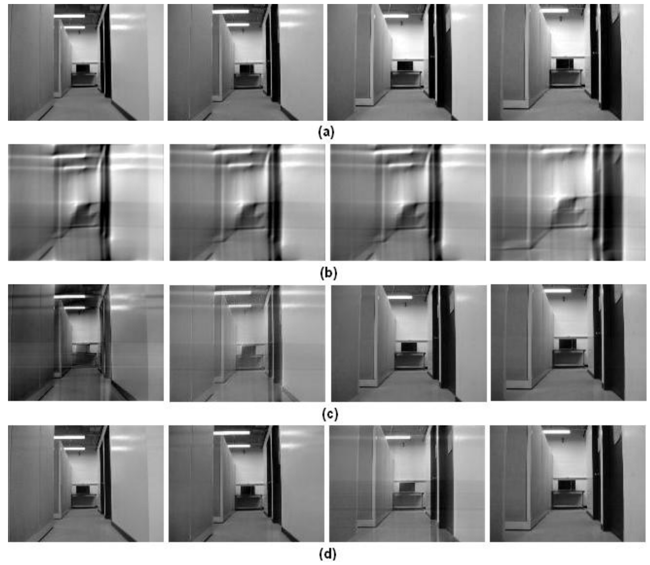

Fig. 3. (a) Original test sequence, (b) blurred sequence, (c) reconstructed sequence based on two frames and (d) reconstructed sequence based on three frames.

7. CONCLUSIONS

In this paper, we introduced a novel method for focused image estimation from defocused video sequences. The proposed algorithm relies on the differences in blur characteristic of multiple images resulting from camera motion in video sequences. This notion is exploited to estimate the blur model and obtain a more robust estimation of the focused sequence. The proposed algorithm can be used to correct out-of-focus video sequences as well as replace the expensive auto-focus apparatus in modern cameras. Noise performance is analyzed and computer simulations prove the success of our approach.

8. REFERENCES

[1] M. Subbarao, ”Efficient depth recovery through inverse optics”, *Machine Vision for Inspection and Measurement (H. Freeman, Ed.), New York: Academic*, 101-126, 1989.

[2] J. Rayala, S. Gupta, and S. Mullick, ”Estimation of depth from defocus as polynomial system identification”, *IEE Proceedings*, pp.356-362, 2001.

[3] A. Rajagopalan, S. Chaudhuri, and U. Mudenagudi, ”Depth estimation and image restoration using defocused stereo pairs”, *IEEE Trans. on PAMI*, vol.26, no.11, pp.1521-1525, Nov. 2004.

[4] J. Yang, D. Schonfeld, C. Chen, and M. Mohamed, ”Online video stabilization based on particle filters”, *IEEE ICIP*, Atlanta, GA, Nov. 2006.

[5] R. Bracewell, K. Chang, A. Jha, and Y. Wang, ”Affine theorem for two dimensional fourier transform”, *Electronics Letters*, Feb. 1993.

[6] F. Bowman, ”Introduction To Bessel Functions”, *Dover Publication Inc., New York*, 1958.

[7] D. Graupe, ”Time Series Analysis, Identification and Adaptive Filtering”, *R.E. Kreiger, Publ. Co.*, 1989.

[8] D. Hinckley, ”On the Ratio of Two Correlated Normal Random Variables”, *Biometrika*, 1969.