Image Motion Deblurring

Daniel Cunningham, s0198594

daniel.cunningham@gmail.com

Introduction

When a photograph is taken in low light conditions or of a fast moving object, motion blur can cause significant degradation of the image (see figure 1). This is caused by the movement of the object relative to the sensor in the camera during the time the shutter is open. Both the object moving and camera shake contribute to this blurring. The problem is particularly apparent in low light conditions when the exposure time can often be in the region of several seconds.

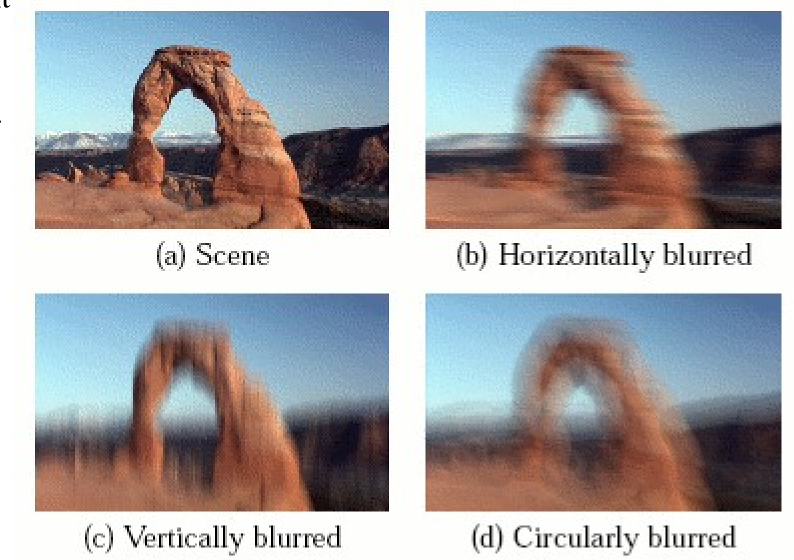

Figure 1: Motion Blur (image from Ezra and Nayar [1])

There are several techniques for either preventing image motion blurring at the time of image capture or post-processing images to remove motion blur. As well as in every day photography, the problem is particularly important to applications such as video surveillance - where low-quality cameras are used to capture sequences of photographs of moving objects (usually people).

Current techniques can be split roughly into the following categories:

- Hardware in the optical system of the camera to stabilise the image
- Post-processing of the image to remove motion blur by estimating the camera's motion
  - From a single photograph (blind deconvolution)
  - From a sequence of photographs
- A hybrid approach that measures the camera's motion during photograph capture

Image Stabilisation

Optically Stabilised lenses are often used in video cameras and more expensive still cameras to reduce the effects of small amounts of camera shake. These use a system of gyroscopes and inertial sensors to keep the optical systems of the camera steady during image capture. This is only really effective for removing a small amount of camera shake at relatively short exposures (less than 1/15th second) Since this technique doesn't really involve any computer vision I won't discuss it in any more detail. More details can be found on Canon's web site. [2][3]

Another hardware approach is to use customised CMOS image detectors that selectively stop image integration more quickly in areas where movement is detected [4][5]. However, this does not really solve the problem of camera shake when long exposures are necessary because of low light conditions.

Motion estimation concepts

Fundamental resolution trade off

Images are formed in a camera by light energy falling on an array of detectors. A certain minimum level of light energy must fall on each pixel for light to be detected. In order to reach this minimum level, the exposure time required for a pixel to detect light is inversely proportional to the area of the pixel. With bigger pixels (low resolution), exposure time can be short. With smaller pixels (higher resolution), exposure time has to be longer. This competition between spatial resolution (number of pixels) and temporal resolution (number of photographs per second) is the fundamental resolution trade off [1].

Modeling the image formation process

In order to process images to remove motion blur, a model of the image formation process is useful as a conceptual starting point [6]. Imagine an "ideal" image to be one that captures a moment in time instantaneously and therefore with no motion blur. This is the image we are trying to estimate in the process of motion deblurring. When a real photograph is taken, we can think of the deficiencies of the photography process as a series of transformations applied to this ideal image:

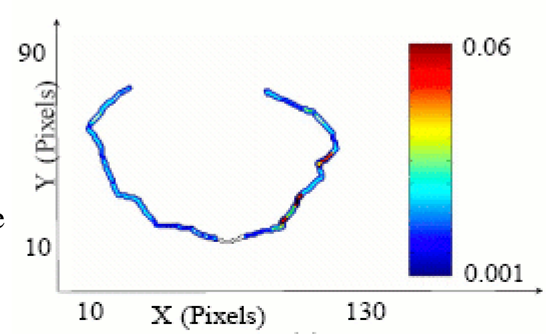

Figure 2: Example of Point Spread Function (PSF) (taken from Ezra and Nayar [1])

1. The image is integrated over time during exposure, resulting in motion blur. Light incident from each point of the scene contributes to several different pixels in the real image. This stage can be thought of as a transformation mapping each pixel of the ideal image to several pixels in the real image according to a point spread function (PSF) (see figure 2). This can really be split into two transformations – one corresponding to the point spread function caused by camera shake, the other corresponding to the movement of objects in the world relative to the photographer.

2. Optical blur is caused by the lens not perfectly focusing the image onto the detector. This can be modeled as a Gaussian function.

3. The detector has a limited resolution so many pixels in the arbitrarily high resolution ideal image are averaged to produce a value for each pixel in the lower resolution real image.

To remove motion blur, we must form an inverse function of the transformation in step 1. The first step to do this is to determine the original motion function, or point spread function (PSF), as accurately as possible. The techniques discussed later tackle this problem in several different ways.

Removal of blur once motion function is known

Once we do know the PSF we can restore the original “ideal” image using image deconvolution to transform the blurred image to an estimate of the ideal image. Algorithms for carrying out this process have been well-studied [7][8]. So all we need to do is estimate this PSF. The PSF is an energy density function that describes the amount of time light from a single point in the scene exposes each (x, y) pixel position in the image detector. Techniques vary in the complexity of the PSF they estimate, ranging from simple linear translation functions to continuous functions in 2D or even 3D.

Motion estimation from a single image

It is possible to estimate the motion blur function from the examination of a single image but this requires strong assumptions about the motion function in order to constrain the parameters of potential PSFs. This method is called blind image deconvolution and often assume that the motion was at a constant velocity or was linear harmonic motion and was uniform across the entire scene. Techniques to estimate the function generally rely on the examination of the frequencies present in the image to determine the direction of motion [9][10][11]. These are very limited in their application as motion is usually far more complex. Estimating the PSF incorrectly results in far more serious artifacts than the original problem of motion blur.

Motion estimation from multiple images

Bascle, Blake and Zisserman [6] propose a technique that uses multiple motion-blurred images to determine the PSF function and deblur each image. The first step of this is to track the object in the photographs. This is done with a combination of area-based (texture correlation) and contour-based (contour matching) deformable models [12]. These give us an affine motion model estimating the 2D motion of the object between consecutive images.

If we let $M(b)_{t-T}^{T} = (A, B)$ be the affine motion transformation measured between the two images taken at time t-T and T, with the model parameterised by b, then the image I_t^{Mot} blurred by the motion (A, B) can be written as:

$$I_t^{Mot}(x, y) = \int_{r=t-T}^{r=t} I_{t-T}^{ideal}(u_r, v_r)\, dr$$

$$\text{with } \begin{pmatrix} x \\ y \end{pmatrix} = A(r) \begin{pmatrix} u_r \\ v_r \end{pmatrix} + B(r)$$

$$\text{and } A(r) = (A-I)\, r + I\ ,\ B(r) = Br$$

The transform between the ideal image and the motion blurred image can now be written in matrix form as $I^{Mot} = B\ I^{Ideal}$.

Once this motion has been estimated, the deblurred image is iteratively estimated by minimising the difference between the sequence of real images and the sequence images predicted by applying the model starting from the first image in the sequence. This is done by minimising the least-squares difference by conjugate gradient descent. Smoothness constraints are added to improve robustness to noise present in the image.

Motion measurement during capture (hybrid imaging)

Another approach, proposed by Ezra and Nayar [1] takes advantage of the fundamental trade off between spatial and temporal resolution to directly measure camera motion during image capture. The basic idea is to have two sensors: one sensor with high spatial resolution (number of pixels) but low temporal resolution (image captures per second) and one lower quality sensor with low spatial resolution but high temporal resolution. A sequence of images captured by the second sensor during the exposure time of the first sensor is used to determine the point spread function (PSF) that describes the equipment's movement during capture.

The motion between consecutive frames is limited to a very simple global rigid transformation model defined by the following translation and rotation model:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\Theta & \sin\Theta & \Delta x \\ -\sin\Theta & \cos\Theta & \Delta y \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where $(\Delta x, \Delta y)$ is the translation vector and Θ is the rotation angle about the optical axis. To compute the parameters of this model, the following optical-flow based error function is iteratively minimised:

$$arg\ \underset{(u,v)}{min} \sum \left(u \frac{\delta I}{\delta x} + v \frac{\delta I}{\delta y} + \frac{\delta I}{\delta t} \right)^2$$

where $\frac{\delta I}{\delta x}, \frac{\delta I}{\delta y}, \frac{\delta I}{\delta t}$ are the spatial and temporal partial derivatives of the image

and $(u, v)$is the instantanseous motion at time t

Once we have these discrete motion samples they must be converted into a continuous PSF. This is done according to several constraints: energy should be neither lost nor gained during the blurring and the function should be continuous and twice differentiable due to physical speed and acceleration constraints. Also, we can assume that scene radiance does not change during image integration.

The discrete motion samples are merged using spline interpolation to create a continuous PSF that satisfies the constraints. Any of the deconvolution algorithms [7][8] can then be applied to the high spatial resolution image using the PSF to retrieve the deblurred image. (see figure 5)

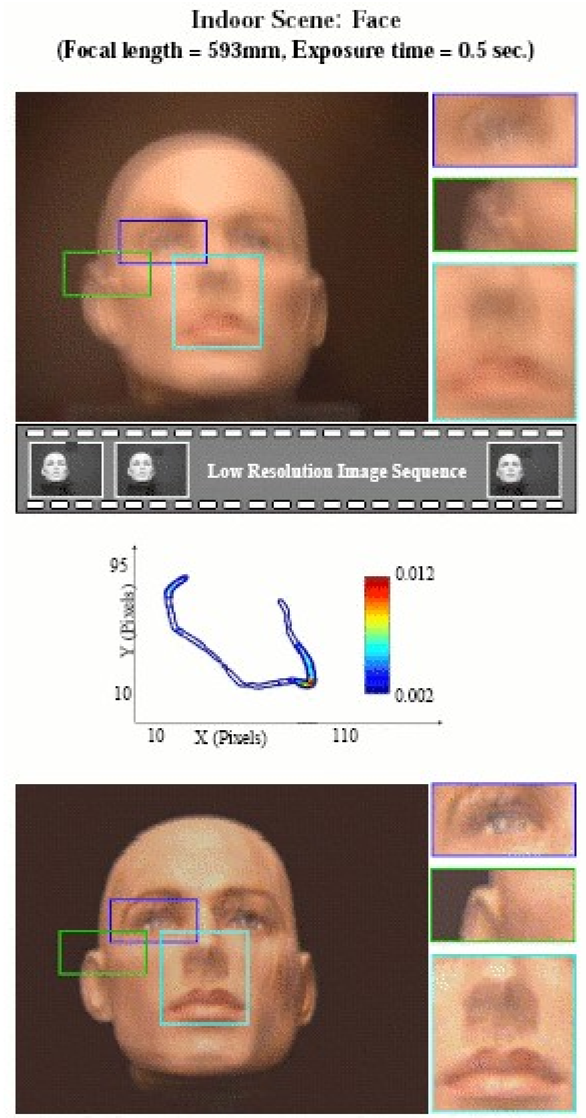

Figure 3: Motion Deblurring Using Hybrid Imaging: Top: The original high-resolution image, low-resolution image sequence; middle: the computed PSF; bottom: the resulting deblurred image (from Ezra and Nayar [1])

Conclusions

The hybrid imaging approach directly measures the motion of the camera to compute an accurate point spread function. This produces outstanding results allowing photographs to be taken in low light conditions on hand-held cameras with exposure times of up to 4 seconds.

References

[1] M. Ben-Ezra and S. K. Nayar. Motion Deblurring Using Hybrid Imaging. Proc. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2003.
[2] Shift-Method Optical Image Stabiliser. Canon, Inc. http://www.canon.com/technology/dv/02.html
[3] Vari-Angle Prism. Canon, Inc. http://www.canon.com/technology/dv/01.html
[4] T. Hamamoto and K. Aizawa. A Computational Image Sensor with Adaptive Pixel-Based Integration Time. 2001.
[5] X. Liu and A. El Gamal. Silmultaneous Image Formation and Motion Blur Restoration via Multiple Capture. Proc. 2001 IEEE International Conference Acoustics, Speech and Signal Processing. p1841. 2001.
[6] B. Bascle, A.Blake and A. Zisserman. Motion Deblurring and Super-resolution from an Image Sequence.
[7] P. A. Jansson. Deconvolution of Image and Spectra. Academic Press, second edition. 1997.
[8] D. P. MacAdam. Digital image restoration by constrained deconvolution. JOSA. 60 (12), p1617-1627. 1970.
[9] R. Fabian and D. Malah. Robust Identification of motion and out-of-focus blur parameters from blurred and noisy images. CVGIP: Graphical Models and Image Processing. Ch 55, p403. 1991
[10] Y. Yitzhaky, I. mor, A. Lantzman and N. S. Kopeika. Direct method for restoration of motion-blurred images. Journal of the Optical Society of America A (Optics, Image Science and Vision). Ch 15, p1512. 1998.
[11] Y. Yitzhaky, G. Boshusha, Y. Levy and N. S. Kopeika. Restoration of an image degraded by vibrations using only a single frame. Optical Engineering. Ch 39, p20832000
[12] B. Bascle and R. Deriche. Region tracking through image sequences. Proc. 5th International Conference on Computer Vision. 1995.