Motion Deblurring for Optical Character Recognition

Xing Yu Qi, Li Zhang, Chew Lim Tan
School of Computing, National University of Singapore
3 Science Drive 2, Singapore 117543
{qixingyu, zhangli, tancl}@comp.nus.edu.sg

Abstract

This paper investigates the problem of blurring caused by motion during image capture of text documents. Motion blurring prevents proper optical character recognition of the document text contents. One area of such applications is to deblur name card images obtained from handheld cameras. In this paper, a complete motion deblurring procedure for document images has been proposed. The method handles both uniform linear motion blur and uniform acceleration motion blur. Experiments on synthetic and real-life blurred images prove the feasibility and reliability of this algorithm provided that the motion is not too irregular. The restoration procedure consumes only small amount of computation time.

1. Introduction

When we take photos, there are three major types of image degradation: blurring, point wise nonlinearities and additive noise. The definition of blur is a cause of imperfect vision or more formally blur is a form of bandwidth reduction in the image formation process. Specifically motion blur is the consequence of relative motion between the camera and the object during the capture process. The motion is modeled by a constant velocity or a uniform acceleration in this paper. The problem of estimating the motion blur has received much attention because of its many applications. A particular application that we are interested in is document images. Motion blur will severely affect the performance of optical character recognition (OCR) results on blurred document images. In this paper, we take various name card images from handheld cameras and improve their OCR results with motion deblurring.

2. Literature

2.1 Modeling

We can model motion blurring as a spatially linear invariant system under certain conditions. If we assume the object translates at a constant velocity V during the exposure time T with angle α from the horizon, the distortion is one-dimensional. We use $d = VT$ and define the point spread function (PSF) as

$$h(x,y) = \begin{cases} 1/d, & 0 \le |x| \le d * \cos(\alpha), y = \sin(\alpha) * d \\ 0, & otherwise \end{cases} \quad (1)$$

If the object translates at a uniform accelerated velocity, S.C.Som [1] derived the spread function as

$$h(s) = \begin{cases} \dfrac{1}{T(V_0^2 + 2as)^{1/2}} & 0 \le s \le d \\ 0 & otherwise \end{cases} \quad (2)$$

where a is the uniform acceleration, V_0 is the initial velocity and s is the displacement from origin. We make the assumption that a and V_0 is in the same direction, and a is always positive. As orientation and magnitude in the uniform linear case, the ratio

$$R = V_0^2 / a \quad (3)$$

is an important parameter in the analysis of uniform acceleration case. When R is very large, i.e. $V_0 >> a^{1/2}$, the spread function will be close to uniform linear. On the other hand, when R is small, i.e. $V_0 << a^{1/2}$, the difference is obvious. We can normalize (2) by using R and $h(0)$, get

$$h(s) = [1 + (2s/R)]^{-1/2} \quad 0 \le s \le d \quad (4)$$

The PSF of motion blur gives the number of original scene points that affect a specific pixel in the blurred image. Now we can define motion blurring mathematically as the result of a linear filter

$$g(x,y) = h(x,y) * f(x,y) + n(x,y) \quad (5)$$

where $g(x, y)$ denotes the blurred image, $f(x, y)$ denotes the original image and $n(x, y)$ denotes additive noise. Note (*) here is used to denote 2-D convolution.

2.2 Methods

Motion deblurring algorithms usually can be divided into two categories. The first group identifies blur parameters then apply well known restoration algorithm, e.g. Wiener Algorithm to deblur the image, while the second group incorporates the identification procedure into the restoration algorithm. Cannon [2] first proposed the method to identify the blur parameters using frequency and cepstral domain techniques. Yitzhaky et al. [3]

proposed another method by making the observation that image characteristics along the direction of motion are different from the characteristic in other directions. Rekleitis [4] proposed a new method to estimate the optical flow map of a blurred image using only information from the motion blur. His algorithm consists of two parts. The first part estimates the orientation of the blur by using 2nd derivative of 2D Gaussian functions. The second part uses the estimated orientation as the input and estimate the extent of the blur by applying cepstral domain techniques. Recent developments in blur identification relate the identification process with the restoration process. ARMA parameter estimation methods estimate the true image and PSF by identification of the ARMA parameters [5]. However, these methods have assumed some *a prior* knowledge of original images.

3. Motion Deblurring

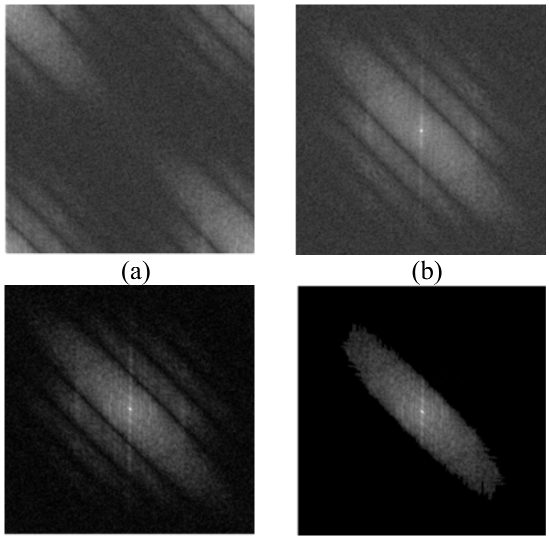
Rekleitis' method is the first algorithm that has been proven to work for a large variety of blur orientations and extents by our experimental results. However, the average estimation error of blur extent is as large as 5.7 pixels for synthetic motion blurred images. Such error totally fails OCR since OCR is more sensitive to errors in blur extent. Defined as the objective of our project, the motion estimation algorithm for OCR should report blur orientation in the range of 0°to 179° and blur extent in the range of 4 pixels to 25 pixels with average error less than 5 degree and 1 pixel respectively. Our deblurring method consists of two parts. The first part is adapted from Rekleitis' approach to find blur orientation and blur extent caused by uniform linear motion. The second part further estimates motion parameters due to uniform acceleration. A complete deblurring procedure using the parameters obtained is described at the end of this section.

3.1 Blur with Uniform Linear Motion

Differing from Rekleitis' method, we propose a new thresholding technique in blur orientation estimation and a new differential operation followed by radon transform in blur extent estimation.

To estimate blur orientation, we apply three steps on the blurred image. In the first step, Gaussian masking is used to obtain better results with the initial Fourier transform as shown in Figure 1(a). This step is optional in Rekleitis' algorithm but mandatory in ours. We reduce the boundary effect and keep the original blurred image maximally unchanged.

In the new second step, thresholding techniques are used to extract the smear lines from the Log spectrum of blurred images. We first use the MATLAB command *fftshift* to shift zero-frequency component to the center of spectrum in Figure 1(b). It is useful to visualize the smear effect of motion blurs. From the spectrum, we state that the orientation of blur is the direction perpendicular to the smear lines in the spectrum. It is because that the motion blur effectively performs a lowpass on the image in the direction of the blur, thus high frequency components diminish significantly in this direction. The smear lines show how the frequency component attenuates from center to sides. An intuitive way to distinguish the smear lines from the dark background is to use contrast stretching. This simple technique improves the contrast of the image by stretching the range of pixel intensities it contains. The smear lines are brighter compared to the dark area in Figure 1(c) after stretching.



(a) (b) (c) (d)

Figure 1. (a) Masked Fourier Spectrum, (b) Shifting zero components to center, (c) With Contrast stretching, (d) After Smear line extraction.

Our next task is to extract the smear lines from the spectrum. Normally, there will be numbers of parallel lines perpendicular to the blur orientation depends on the blur magnitude. Only the main smear line in the center is used for estimation in Figure 1(d). The intuitive idea is to start from the center pixel and gradually expands its surrounding pixels until any pixel falls below the threshold defined by the average pixel values of the spectrum. We use the simple mean value of the spectrum since our main objective is to remove unwanted information and it does yield satisfactory results. When the average neighbor pixel values of a specific pixel *(i, j)* defined by (6) is below the average of the spectrum, we just set *p(i, j)* = 0.

$$\begin{aligned}\hat{p} = 1/8\sum[\, & p(i-1, j-1) + p(i-1, j) \\ + \quad & p(i+1, j) + p(i, j-1) + p(i, j+1) \\ + \quad & p(i+1, j-1) + p(i+1, j) + p(i+1, j+1)]\end{aligned} \qquad (6)$$

The success of this method depends on how accurately we extract the smear line from the spectrum. The zigzag structures at the border of extraction may pose problems in the estimation. It is especially severe when the magnitude is small.

The 2nd derivative of a Gaussian and its Hilbert transform is applied as bandpass filter [7] to the modified spectrum in the last step. Oriented filter is used in motion analysis and Gaussian derivatives fall into this category. Hilbert transform is equivalent to an interesting kind of filter, in which the amplitudes of the spectral components are left unchanged, but their phases are altered by 90, positively or negatively according to the sign of frequency. The orientation that maximizes (7)

$$E_2(\theta) = [G_2^\theta]^2 + [H_2^\theta]^2 \quad (7)$$

is returned as the local dominant orientation in the spectrum. Expanding the bandpass filter as a sum of basis filter outputs times interpolation functions, the equation simplifies to a Fourier series in angle in (8),

$$E_n(\theta) = C1 + C2\cos(2\theta) + C3\sin(2\theta) + [highorder\ldots] \quad (8)$$

with only even frequencies present. The lowest frequency term approximates the dominant direction. Clearly more frequency components hinder the thresholding process and obscure the finding of blur orientation.

To estimate the blur extent, we again apply three new steps on the blurred image. The blur orientation is required in the extent estimation. In the first step, differential filter is used in the direction perpendicular to the blur orientation to decorrelate the blurred image. Real-life document image is characterized by high spatial correlation. Pixels that are next to each other in an image are highly positively correlated, i.e. bright pixels tend to be next to other bright pixels and dark pixels tend to be next to other dark pixels. Pixels that are close to each other are still correlated, but not as much as next to each other. Name card images exhibit strong correlation properties in the background and the text area. Such property leads to difficulties in the cepstral domain analysis of the blurred image since the overlaying structure of the original image in the cepstrum significantly frustrates the identification of the PSF.

To suppress the correlation, we normally use differential operation as the decorrelating filter. The differential operation is simply to replace an original pixel with the difference between the pixel and its adjacent pixel. In the uniform linear blur, the pixels close to each other in the direction of the blur are also correlated because they are mostly imaging the same set of pixels in the original image. Such property shows the characteristics of the blur PSF. To maximally not affect this information, we use the differential filter in the direction perpendicular to the blur orientation. Since we deal with digitized images, the neighboring pixel to *p(i, j)* in the direction α where $0 \le \alpha \le 45°$ is approximated by intensities from both *p(i, j + 1)* and *p(i + 1,j + 1)*. Other directions are done in the similar way. The detailed formula for the differential operation is given in (9),

$$p(i,j) = \begin{cases} \tan\alpha \times p(i+1, j+1) + (1-\tan\alpha) \times p(i, j+1) \\ \quad -p(i,j), \quad 0 \le \alpha < 45° \\ \tan(90°-\alpha) \times p(i+1, j+1) + (1-\tan(90°-\alpha)) \times p(i+1, j) \\ \quad -p(i,j), \quad 45° \le \alpha < 90° \\ \tan(\alpha - 90°) \times p(i+1, j-1) + (1-\tan(\alpha-90°)) \times p(i+1, j) \\ \quad -p(i,j), \quad 90 \le \alpha < 135° \\ \tan(180°-\alpha) \times p(i+1, j-1) + (1-\tan(180°-\alpha)) \times p(i, j-1) \\ \quad -p(i,j), \quad 135° \le \alpha < 180° \end{cases} \quad (9)$$

where *p (i, j)* is the pixel of the blurred image and α is the blur orientation.

In the second step, Radon transform is applied to the Log Fourier spectrum of the blurred image in the blur orientation to get a collapsed 1D signal in Figure 2. Radon transformation can be used to project 2D objects into one line. The line whose normal vector is in θ direction and whose distance from the origin is s satisfying the following equation,

$$g(s,\theta) = \int \int_{-\infty}^{\infty} f(x,y)\sigma(x\cos\theta + y\sin\theta - s)dxdy \quad (10)$$

In the frequency domain, the Log spectrum = Log Original image + Log PSF. The differential operation has decorrelated the original image such that Log PSF is dominant in the Log spectrum. By performing Radon transform in blur orientation, we obtain signal approximates the blur PSF.

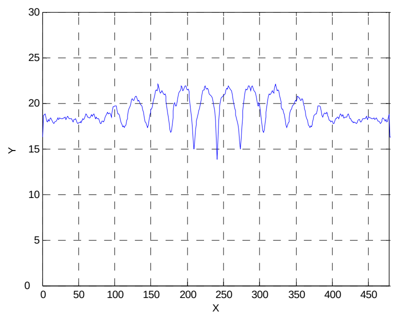


Figure 2. 1D Signal by Radon Transform

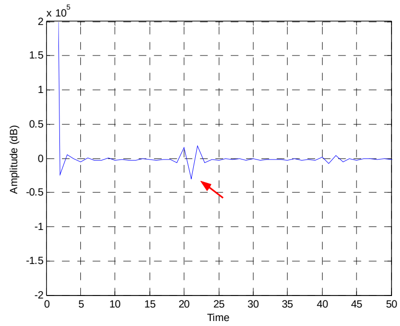


Figure 3. Cepstem of the Collapse Signal

Cepstral domain analysis is performed in the last step. We obtain the cepstrum by taking Fourier transform of the 1D signal in Figure 3. Cepstrum analysis is a nonlinear signal processing technique with a variety of applications in areas such as speech and image processing [6]. The independent variable in the Cepstral domain has the same dimension as the spatial domain variable. In the case of cepstrum of rectangle (PSF of uniform linear motion blur), big spikes indicate the width of the rectangle (blur extent). Since the cepstrum is computed with DFT, some of the smaller spikes are due to aliasing. The position corresponding to the local peak of the cepstrum is the blur magnitude.

3.2 Blur with Uniform Acceleration

We first look at properties of uniform acceleration blur. For the same amount of smear, the modulation transfer function (MTF) due to uniform acceleration is equal or greater than that due to uniform velocity in all R. It implies that blur due to uniform acceleration causes less distortion to images compared to that due to uniform velocity. But in the former case, phase of the transfer function (PTF) leads to additional degradations. We observe the MTF of uniform acceleration blur exhibits strong periodicity of the blur extent. Periodicity diminishes when R approaches 0, i.e. the acceleration is at infinity. However, in this case the image is virtually not blurred. In real world, acceleration is always bounded in a range, thus we define the possible values of R within [0.1, 50] here. Such assumption is reasonable as the PSF shape usually varies in a small margin when R approaches two ends of this range. The conclusion is that the methods for uniform linear motion blur also work on acceleration case, except that we need to estimate the possible range of R.

PSF of acceleration blur is usually asymmetric. For the same blur extent, we define two different acceleration PSFs. Since the spread function is not uniform over the extent of the smear, we have both forward and backward acceleration depending on how the image is blurred. To estimate the PSF curve, we make use of autocorrelation functions (ACF). The autocorrelation function $K(n)$ of an M pixel image line l is defined as,

$$K(n) = \sum_{i=-M}^{M} l(i+n)l(i) \quad n \in [-M, M] \tag{11}$$

where $l(i) = 0$ outside the image line range. Equation (11) describes how pairs of pixels at particular displacements from each other are correlated. It is high where they are well correlated and low where poorly correlated. For a normal image, the ACF will be some function of distance from the origin plus random noises. But for a motion blurred image, the ACF will decline much more slowly in the direction of the blur than in other directions. This is because pixels that are close together in the blurred direction are mostly representing the same real world image points. Obviously, when acceleration is higher namely R is larger, ACF will decline faster in the blur direction since the difference between neighboring blurred pixels increases.

We use Wiener filter to deblur after obtaining the estimated PSF. We reduce severity of the ringing by regularizing high-frequency regions less strongly. The resolution is enhanced at the same time.

To sum up, our estimation procedure on real-life blurred images is as follows:

1. Apply algorithm for uniform linear motion blur to estimate the blur orientation and extent.
2. Use the average autocorrelation function of the image lines in the blur orientation to adjust R.
3. Create both forward and backward acceleration PSF based on the estimated parameters in 1 and 2.
4. Use the created PSF in 3 to restore the blurred image. If the restoration from both forward and backward acceleration shows clear inverted scene imposed on the true scene, the image is blurred by a symmetric PSF, i.e. uniform linear motion blur; otherwise use the PSF that yields better OCR results.
5. (Optional) Adjust the parameter in Wiener filter.

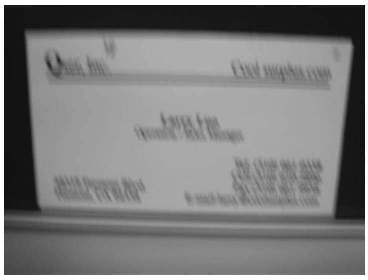


Figure 4. Blurred Name Card P1120541

Figure 5. Blurred Name Card P1120547

4. Experiments

We apply our proposed method to around 100 motion blurred name cards. Two examples are shown in Figures 4 and 5. The procedure of our experiment is as follows:

1. Use OCR software to recognize a blurred name card M_1 and record its precision p_1 and recall r_1.
2. Apply first part of our algorithm and obtain the blur orientation θ and magnitude d.

3. Apply Wiener filter with θ and d to deblur M_1 and the resulting image is M_2.
4. Use OCR software to recognize M_2 and record down its precision p_2 and recall r_2.

Both images are severely blurred. Nearly all the characters can not be recognized even by human eyes. We assume no noise is present and obtain the following OCR results in Table 1.

Table 1. OCR results for original real-life blurred name card images

BLURRED IMAGE	P1120541	P1120547
PRECISION 1	0%	0%
RECALL 1	0%	0%
EST. θ	114	90
EST. d	21	16
PRECISION 2	51.51%	46.67%
RECALL 2	34.46%	23.3%

After deblurring, we notice most texts including digits are readable. However, OCR results are not satisfactory since ringing effects occur around most characters. We apply the blur estimation procedure in the second part of our method and obtain the following results in Table 2.

Table 2. OCR results for real-life blurred name card images after complete deblurring

BLURRED IMAGE	P1120541	P1120547
PRECISION 2	51.51%	46.67%
RECALL 2	34.46%	23.3%
EST. R	24	18
EST. d	Backward	Forward
PRECISION 3	61.47%	62.61%
RECALL 3	48.23%	49.67%

The blur in the first image is estimated with lowly backward acceleration while the second image with lowly forward acceleration. The recall increases from 34.46%, 23.3% to 48.23%, and 49.67% respectively. The low acceleration means the actually blur is close to uniform linear motion blur. As we look at the restored image in Figure, the ringing effect is reduced by using an accelerated PSF. However, due to the asymmetry of the PSF, the ringing artifacts appear to exist on only one side of the texts. In horizontal blur, artifacts are on the right of the scene when using backward acceleration; left when using forward acceleration. The artifacts lessen when we use a proper form of the Wiener filter in Figures 6 and 7.

5. Conclusion

In this paper, a modified approach based on I.M.Rekleitis' method is formulated and evaluated experimentally. The orientation of the blur is first determined and then the blur extent in that direction is recovered. When we look into real-life blurred name cards, OCR fails or returns poor results even both parameters are correctly estimated. This leads us to think that more severe motion has occurred during the capture process. We assume uniform acceleration motion blur exists and new estimation procedure has been proposed. Experiments on blurred name cards show that OCR results improve in most cases.

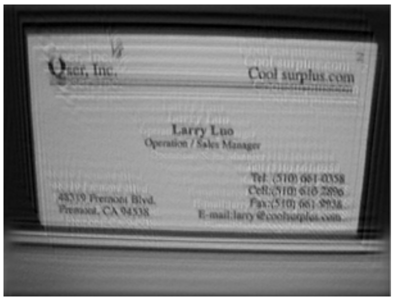


Figure 6. Restored P1120541

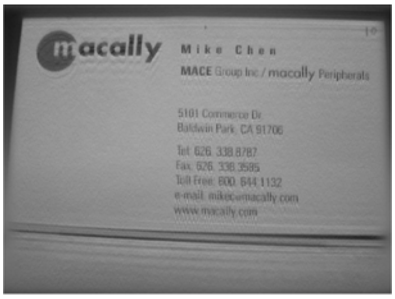


Figure 7. Restored P1120547

References

[1] S.C.Som, "Analysis of the effect of linear smear on photographic images," J.Opt.Soc.Am.Vol.61, 859-864, 1971.

[2] M.Cannon, "Blind deconvolution of spatially invariance image blurs with phase," IEEE Trans.Acoust.Speech Signal Process. 24,1976.

[3] Y.Yitzhaky, I.Mor, A.Lantzman, and N.S.Kopeika, "Direct method for restoration of motion-blurred images," J.Opt.Soc.Am.Vol.15, No.6, June, 1998.

[4] I.M.Rekleitis, "Visual motion estimation based on motion blur interpretation," MSc Thesis, School of Computer Science, McGill University, Montreal, Quebec, Canada, 1995.

[5] D.Kundur and D.Hatzinakos, "Blind image deconvolution," IEEE Signal Processing Magazine, May, 1996.

[6] R.Rom, "On the cepstrum of two-dimensional functions," IEEE Trans.Inform.Theory, Vol.IT-21, 214-217, Mar.1975.

[7] W.T.Freeman and E.H.Adelson. "The design and use of steerable filters," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.13, No.9, 891-906, September 1991.