Restoration of Temporal Image Sequence from a Single Image Captured by a Correlation Image Sensor

Kohei Kawade[1], Akihiro Wakita[1], Tastuya Yokota[1], Hidekata Hontani[1] and Shigeru Ando[2]
[1]*Computer Science and Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya-shi, Aichi 466-8555, Japan*
[2]*Richo Elemex Corporation, 3-69, Ida-cho, Okazaki-shi, Aichi, 444-8586, Japan*

Keywords: Optical Flow, Correlation Image Sensor.

Abstract: We propose a method that restores a temporal image sequence, which describes how a scene temporally changed during the exposure period, from a given still image captured by a correlation image sensor (CIS). The restored images have higher spatial resolutions than the original still image, and the restored temporal sequence would be useful for motion analysis in applications such as landmark tracking and video labeling. The CIS is different from conventional image sensors because each pixel of the CIS can directly measure the Fourier coefficients of the temporal change of the light intensity observed during the exposure period. Given a single image captured by the CIS, hence, one can restore the temporal image sequence by computing the Fourier series of the temporal change of the light strength at each pixel. Through this temporal sequence restoration, one can also reduce motion blur. The proposed method improves the performance of motion blur reduction by estimating the Fourier coefficients of the frequencies higher than the measured ones. In this work, we show that the Fourier coefficients of the higher frequencies can be estimated based on the optical flow constraint. Some experimental results with images captured by the CIS are demonstrated.

1 INTRODUCTION

Each pixel in a traditional image sensor measures the temporal integration of light strength over an exposure time. Because of the temporal integration, the measured pixel value does not contain information of the temporal change of the light strength generated during the exposure time. Moving objects or moving cameras generate the temporal change of the light strength, but a traditional camera fails to record this temporal change and hence restoring the temporal change from a given still image is difficult. In this article, we propose a method that restores the temporal image sequence, which describes how the image temporally changed during the exposure period, from a given still image captured by a correlation image sensor (CIS)(Ando et al., 1997)(Ando and Kimachi, 2003)(Wei et al., 2009)(Hontani et al., 2014). Several methods have been proposed for motion estimation or for temporal change restoration from a single image mainly because a single still image does not have enough information for restoring temporal changes.

As described above, the proposed method captures an image by using a CIS. Each pixel of a CIS measures not only a temporal integration of the light strength but also a temporal *correlation* between the light strength and a reference temporal signal supplied from the outside of the CIS to each pixel during an exposure period. By using sinusoidal functions as the references, one can measure a set of the Fourier coefficients of the light strength at each pixel. These measurements of the Fourier coefficients include the information on the temporal change of the light strength during the exposure period and it was reported that the CIS made a problem of the optical flow computation well-posed(Wei et al., 2009). The proposed method restores the temporal change of the image intensity at each pixel by using the measured Fourier coefficients: Just by computing the Fourier series of the temporal change at each pixel, one can approximately restore the temporal change of the image. The method improves the accuracy of the approximation by using other information obtained though optical flow computation. The temporal change approximately restored by the Fourier series is always periodic, but this is not always true. In other words, if one restores the temporal change of the light intensity only by using the Fourier series, the restored light intensity at the

Kawade K., Wakita A., Yokota T., Hontani H. and Ando S.
Restoration of Temporal Image Sequence from a Single Image Captured by a Correlation Image Sensor.
DOI: 10.5220/0006115301810191
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 181-191
ISBN: 978-989-758-227-1

beginning of the exposure period is always the same as the intensity at the end of the period. The proposed method thus improves the accuracy by explicitly estimating the difference between the light intensity at the beginning and that at the end of the exposure period through optical flow computation as will be described later. The proposed method can restore an instantaneous image that describes the spatial distribution of the light intensity on the image sensor corresponding to an arbitrary point of time during the exposure time. The restored instantaneous image should have motion blur less than the given original image. Our proposed method has strong relationships with methods used for motion blur reduction. The authors are interested in the comparison of the performance of the motion blur reduction between the proposed method and the other existing methods, but the comparison is out of the scope of this article. One of the strong points of the proposed method is that it restores a series of images with higher spatial resolution, which represents the temporal change of the images during the shutter is open. In the next section, we describe methods of temporal change restoration and of blind motion deblurring, which can be interpreted as the restoration of the instantaneous image that describes the light intensity at some specific point of time during the exposure period.

2 RELATED WORK

Several methods that can restore the temporal change of the light strength during the exposure time have been proposed. To the best of our knowledge, all those methods employ an *active sensing* strategy: A coded light, which is synchronized with the timing of the camera shutter, is used when a still image is captured and the temporal change of the light strength is decoded by a temporal high-frequency pattern of the shutter (Kadambi et al., 2013)(Velten et al., 2013)(Heide et al., 2013). Such a camera can measure the depth in realtime and can restore the temporal change of the light strength during the exposure time. In contrast to these methods, our method employs a *passive sensing* strategy: No encoded light that is specific to the camera used is required.

From the view point of the passive sensing strategy, the proposed method has strong relationships with methods for blind motion deblurring. One can divide the methods of the nonuniform blind deblurring into two classes: In one class, a given image is captured by a traditional image sensor. In the other class, a nontraditional sensor is used for capturing the image. The majority of the methods in the former class employ a Bayesian framework for removing motion blurs. These methods combine natural image priors(Levin, 2006)(Fergus et al., 2006)(Shan et al., 2008)(Cho and Lee, 2009)(Cai et al., 2012)(Xu et al., 2013)(Deshpande and Patnaik, 2014), blurring kernel priors(Jia, 2007)(Shan et al., 2008)(Cho and Lee, 2009)(Xu and Jia, 2010)(Cai et al., 2012)(Xu et al., 2013), or models of motions(Jia, 2007)(Shan et al., 2007)(Whyte et al., 2012), and optimization techniques to recover both the blurring kernel and the deblurred latent image simultaneously.

The natural image priors include distribution models of spatial gradient magnitudes(Field, 1994)(Fergus et al., 2006)(Shan et al., 2008) and sparse representations based on local appearances(Deshpande and Patnaik, 2014) or on the gradients(Shan et al., 2008)(Xu and Jia, 2010)(Cai et al., 2012). The blurring kernel priors play important roles especially when the motion blurs result from camera shakes. For example, a motion density function(Gupta et al., 2010), spatially sparse local filter(Fergus et al., 2006), camera geometry model(Whyte et al., 2012), and framelet-based sparse representation(Cai et al., 2012) are introduced as the kernel priors. Object motion estimation is useful for removing motion blurs(Jia, 2007), and a parametric motion descriptor(Shan et al., 2007) is used for the estimation from a single given image. The proposed method also estimates optical flow from a single given image and uses it for removing motion blurs.

In the former class, nontraditional image sensors are used for obtaining the additional information useful for motion deblurring. Hybrid cameras(Nayar and Ben-Ezra, 2004)(Tai et al., 2010), which detect camera motion using data from a video camera attached to a still camera, are used for motion deblurring. Inertial measurement sensors are also used for the estimation of camera motion(Joshi et al., 2010). Among the methods in this class, the proposed method is most similar to the motion deblurring methods using coded exposure photography(Raskar et al., 2006)(Agrawal and Raskar, 2009)(McCloskey et al., 2012). Controlling the camera's shutter open and close during the exposure period with a binary pseudo-random sequence, one obtains a broad-band motion blurring kernel that preserves high-frequency spatial details. Combining the estimation of objects' motions, one can estimate the blurring kernel and compute the latent nonblurred image by deconvoluving the given motion blurred image with the estimated kernel.

In coded exposure photography, the camera encodes the temporal change of incident light strength at each pixel during the exposure period with the binary pseudo-random shutter pattern. A CIS used in

the proposed method, on the other hand, encodes the temporal change with analog sinusoidal reference signals. As will be described later, a CIS has three channels each of which encodes the temporal change with a different reference signal and makes the problem of optical flow computation well-posed. The proposed method first computes the optical flow of a given single image and then uses the optical flow for restoring higher frequency components of the temporal change of the light strength during the exposure period.

The contributions of this study are as follows: (1) A completely passive method is proposed for restoring the temporal change of the light strength during the exposure period from a single still image captured by a CIS, (2) To improve the quality of the restored images, a method is proposed for computing the temporal frequency components that are higher than those of the reference signals, and (3) It is shown that the restoration can reduce motion blur.

3 CORRELATION IMAGE SENSOR (CIS)

Let exposure time be denoted by T and let the time during which the shutter is open be denoted by t, where $0 \leq t \leq T$. Let the coordinates on an image sensor be denoted by $x = (x,y)^\intercal$ and let $f(t,x)$ denote the strength of the incident light that comes from the lens system to the pixel at location x at time t. The location variable, x, is often omitted as $f(t) = f(t,x)$ when the location is clearly known from the context.

The CIS used in this study has three channels and a pixel value of the s-th channel ($s \in \{1,2,3\}$) denotes the temporal correlation between the light strength, $f(t,x)$, and a reference signal, $r_s(t)$, which is supplied from the outside of the CIS to all pixels as

$$I_s(x) = \int_0^T f(t,x) r_s(t) dt.$$

Let $r_s(t) = \cos(n\omega_0 t + 2(s-1)\pi/3) + 1/3$ where $\omega_0 = 2\pi/T$ and $n \in \mathbb{Z}$. Then, we can measure the following three values at each pixel simultaneously(Ando et al., 1997)(Ando and Kimachi, 2003)(Wei et al., 2009) as follows:

$$g_0(x) = \frac{1}{T}\int_0^T f(t,x) dt, \tag{1}$$

$$g_{n,\mathcal{R}}(x) = \frac{1}{T}\int_0^T f(t,x)\cos(n\omega_0 t) dt, \tag{2}$$

and

$$g_{n,I}(x) = \frac{1}{T}\int_0^T f(t,x)\sin(n\omega_0 t) dt, \tag{3}$$

where $g_{n,\mathcal{R}}(x)$ and $g_{n,I}(x)$ are the real and imaginary parts of a complex Fourier coefficient g_n that corresponds to the frequency n/T. It should be noted that no image sensor can measure the light strength $f(t,x)$ directly and traditional image sensors can measure only $g_0(x)$. In this manuscript, we set $n = 1$ and a CIS sensor measures $g_0(x)$ and $g_1(x) = g_{1,\mathcal{R}}(x) + j g_{1,I}(x)$, where j is the unit imaginary number.

3.1 Optical Flow Computation with a CIS

We can compute an optical flow $v(x) = (v_x(x), v_y(x))^\intercal$ from a single image measured by the CIS. In the computation, we assume that the light strength arriving from each point of an object is invariant with respect to time and that the following equation is satisfied when the shutter is open, $t \in [0,T]$:

$$\frac{d}{dt} f(t,x) = \left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + \frac{\partial}{\partial t} \right) f(t,x) = 0. \tag{4}$$

Integrating (4) over the exposure time with a weight $e^{-jn\omega_0 t}/T$ where $n \in \mathbb{Z}$, we obtain the following equation:

$$\begin{aligned} 0 &= \frac{1}{T}\int_0^T \left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + \frac{\partial}{\partial t} \right) f(t,x) e^{-jn\omega_0 t} dt \\ &= \frac{1}{T}\int_0^T \left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} \right) f(t,x) e^{-jn\omega_0 t} + \frac{\partial}{\partial t} f(t,x) e^{jn\omega_0 t} dt \\ &= \left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} \right) g_n(x) \\ &\quad + \frac{1}{T}\left\{ [f(t,x) e^{jn\omega_0 t}]_0^T + jn\omega_0 \int_0^T f(t,x) e^{jn\omega_0 t} dt \right\} \\ &= \left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} \right) g_n(x) + F_0(x) + jn\omega_0 g_n(x), \end{aligned} \tag{5}$$

where $g_n(x)$ is an image of the Fourier coefficients of temporal signals $f(t,x)$ such that

$$g_n(x) = \frac{1}{T}\int_0^T f(t,x) e^{-jn\omega_0 t} dt \tag{6}$$

and F_0 denotes the difference of the boundary values such that $F_0(x) = [f(t,x)]_0^T/T = \{f(T,x) - f(0,x)\}/T$.

As described above, a CIS can measure both $g_0(x)$ and $g_1(x)$ at each pixel. Substituting $n = 0$ and $n = 1$ to (5), we obtain two different equations. Eliminating the variable F_0 from these two equations, we can derive a linear complex equation as follows:

$$\left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} \right) \{g_1(x) - g_0(x)\} + j\omega_0 g_1(x) = 0. \tag{7}$$

Equation (7) consists of two real coefficient equations and we can derive the following linear equation of v:

$$Av = d, \tag{8}$$

where $d = [\omega_0 g_{1,I}, -\omega_0 g_{1,\mathcal{R}}]^T$ and

$$A = \begin{bmatrix} \partial_x\{g_{1,\mathcal{R}} + g_0\} & \partial_y\{g_{1,\mathcal{R}} + g_0\} \\ \partial_x g_{1,I} & \partial_y g_{1,I} \end{bmatrix}. \quad (9)$$

One can estimate $v(x)$ by solving the linear equation shown in (8). It should be noted that the problem of optical flow computation is *well-posed* when one captures images with a CIS though the problem is ill-posed when one captures images with a traditional image sensor. The added measurements, $g_n(x)$, can introduce extra information that constrains the solution.

Substituting $n = 0$ and $n = 1$ to (5), we obtain the following two equations:

$$\left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y}\right) g_0(x) + F_0(x) = 0, \quad (10)$$

$$\left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y}\right) g_{1,\mathcal{R}}(x) + F_0(x) - \omega_0 g_{1,I} = 0. \quad (11)$$

Solving the system of linear equations (10) and (11), we can obtain the following equation.

$$F_0(x) = -\frac{1}{2}\left(v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y}\right)(g_{1,\mathcal{R}}(x) + g_0(x)) + \frac{\omega_0}{2} g_{1,I}(x). \quad (12)$$

Once we compute the flow $v(x)$, the difference of the boundary values $F_0(x)$ can be estimated by solving (12). The values, v_x, v_y, and F_0, computed at each pixel are used in the restoration process described in the next section.

4 RESTORATION OF TEMPORAL CHANGE OF LIGHT STRENGTH DURING EXPOSURE PERIOD

The objective is to restore the temporal change of the strength of light strength, $f(t,x)$ $(0 \le t \le T)$, at each pixel from the measurements, $g_0(x)$ and $g_1(x)$, obtained by a CIS. It is not difficult to approximately restore $f(t,x)$ by using the N-th order Fourier series expansion of $f(t) = f(t,x)$ $(N \in \mathbb{N})$ such that

$$\tilde{f}_N(t) = \sum_{n=-N}^{N} g_n e^{jn\omega_0 t}, \quad (13)$$

where g_n $(n \in \mathbb{Z})$ are the Fourier coefficients as defined in (6). As $g_{-n} = g_n^*$ holds when $f(t)$ is a real function, the original signal $f(t)$ can be approximated by using the Fourier coefficients g_0 and g_1 measured by a CIS as follows:

$$\tilde{f}_1(t) = \sum_{n=-1}^{1} g_n e^{jn\omega_0 t} = g_0 + 2\mathrm{Re}[g_1 e^{j\omega_0 t}], \quad (14)$$

where $\mathrm{Re}[c]$ denotes the real part of a complex number c. The proposed method restores $f(t)$ by improving the approximation $\tilde{f}_1(t)$ by restoring (i) the difference of the boundary values $f(T) - f(0)$, and (ii) the higher frequency components g_n $(n > 1)$ that are not measured by the CIS.

4.1 Restoration of Boundary Values

$\tilde{f}_N(0)$ and $\tilde{f}_N(T)$ approximate $f(0)$ and $f(T)$ inaccurately in general. $\tilde{f}_N(0) = \tilde{f}_N(T)$ always holds because $\tilde{f}_N(t)$ shown in (13) is a linear combination of $e^{jn\omega_0 t}$ $(n = 0, 1, \ldots, N)$ that are all periodic functions with period T. It leads to $\tilde{F}_0(x) = \{\tilde{f}_N(T) - \tilde{f}_N(0)\}/T = 0$, which is not always consistent with the true value of F_0. The value of $\tilde{f}_N(t)$ changes extremely rapidly near the boundaries $t = 0$ and $t = T$. The left panel of Fig.1 shows examples of the signals restored by g_n $(n = 1, 2, \ldots)$. A true and unobservable signal $f(t)$, where $f(T) \neq f(0)$, is shown at the top of Fig.1. As N increases, the resultant $\tilde{f}_N(t)$ converges to the original one, but we can notice that $\tilde{f}_N(T) = \tilde{f}_N(0)$ is always satisfied and can see the rapid change near the measurement boundaries $t = 0$ and $t = T$.

The proposed method compensates for the rapid changes by adding a new function $s(t)$ to $\tilde{f}_N(t)$ so that the resultant function $\hat{f}_N(t) = \tilde{f}_N(t) + s(t)$ is consistent with the value of F_0 estimated by using (12) and with the Fourier coefficients $g_n(x)$ that are measured (or would be computed) from $f(t,x)$. The consistency with respect to the Fourier coefficients requires $s(t)$ to satisfy the following equation:

$$\int_0^T \tilde{f}_N(t) e^{-jn\omega_0 t} dt = \int_0^T \hat{f}_N(t) e^{-jn\omega_0 t} dt. \quad (15)$$

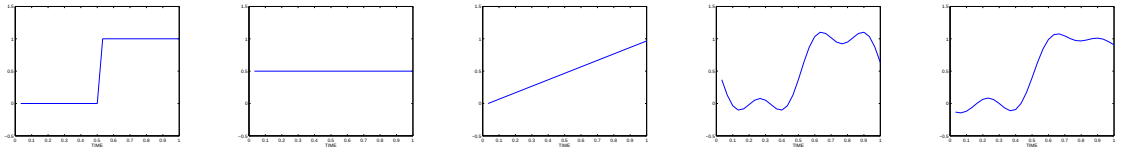
This leads to

$$\int_0^T s(t) e^{-jn\omega_0 t} dt = 0. \quad (16)$$

Substituting $n = 0$, we notice that $s(t)$ must be anti-symmetric with respect to the reflection at the center time of the exposure period $t = T/2$. Among the anti-symmetric functions, we select the following equation because it is consistent with the value of F_0:

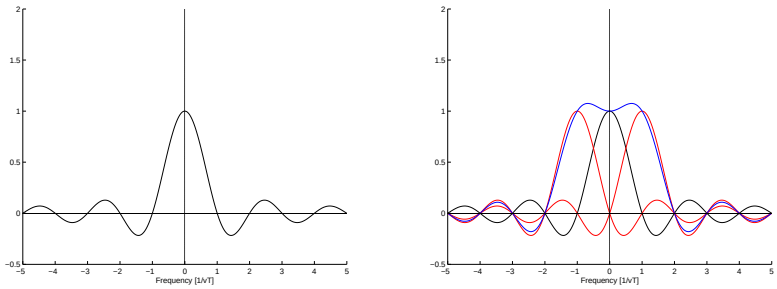
$$s(t) = s_N(t) = F_0 \frac{(t - T/2) + \sum_{n=1}^{N} a_n \sin(n\omega_0 t)}{T}, \quad (17)$$

which always satisfies $s_N(T) - s_N(0) = F_0$. The coefficients proposed in (17), a_n, are scalar and their values should be determined so that $\hat{f}_N(t)$ is consistent



$f(t)$ $(a)\tilde{f}_0(t)$ $(b)\hat{f}_0(t)$ $(c)\tilde{f}_3(t)$ $(d)\hat{f}_3(t)$

Figure 1: Examples of the approximations. An original signal $f(t)$, the Fourier approximations $\tilde{f}_N(t)$ shown in (13), and the modified approximation $\hat{f}_N(t)$ in (19) are indicated.



(A) (B)

Figure 2: (A) A graph of H_0 along the direction of v. (B) Graphs of H_0 (black), H_1 (red), and H_1^* (red). H_1 improves the bandwidth of the resultant filter (blue).

with g_n $(n \geq 1)$. Solving (16) with respect to a_n, we get

$$a_n = \frac{2}{n\omega_0}, \quad (n = 1, 2, \ldots, N). \tag{18}$$

As a result, the proposed method restores the latent image by using the following equation:

$$\begin{aligned} \hat{f}_N(t,x) &= g_0(x) + 2\sum_{n=1}^{N} \mathrm{Re}[g_n(x)e^{jn\omega_0 t}] \\ &+ \frac{F_0}{T}\left((t - T/2) + \sum_{n=1}^{N} \frac{2}{n\omega_0}\sin(n\omega_0 t)\right). \end{aligned} \tag{19}$$

(a) and (c) of Fig.1 show examples of $\hat{f}_N(t)$ that approximate $f(t)$, which is a step function in this example. Comparing (b) and (d) of Fig.1 respectively, we observe that the rapid changes near the measurement boundaries are suppressed.

4.2 Restoration of Higher Frequency Components

The number of channels of a CIS is limited to three and we use them for measuring the lower temporal frequency components of $f(t)$. Restoring the higher *temporal* frequency components of $f(t,x)$, one can restore higher *spatial* frequency components of $f(t,x)$ and can deblur the motion-blurred gray-scale image $g_0(x)$ more crisply. As will be described later, one can restore the higher temporal frequency components of $f(t,x)$ by using (5). Before the restoration algorithm is described, the relationships between the temporal frequencies of $f(t) = f(t,x)$ and the spatial frequencies of $g_n(x)$ are discussed.

4.2.1 Spatial Motion Blur in Fourier Coefficient Image

Let the velocity of a moving target in an image be denoted by $v = (v_x, v_y)^\intercal = v(\cos\theta, \sin\theta)^\intercal$, where v denotes the speed and θ denotes the angle between v and the x-axis. The motion blur generated on the gray-scale image $g_0(x)$ by the target motion can be represented by a spatial convolution with a one-dimensional box filter $h_0(x)$ that averages the intensities along a line segment of length vT, which is a trajectory of the moving target:

$$g_0(x) = f(x) * h_0(x), \tag{20}$$

where $*$ denotes a spatial convolution,

$$h_0(x) = \frac{1}{vT}\mathrm{rect}\left(\frac{x\cos\theta + y\sin\theta}{vT}\right)\delta(v_\perp^\intercal x), \tag{21}$$

$v_\perp$ denotes a unit vector perpendicular to v,

$$\mathrm{rect}(x) = \begin{cases} 1, & \text{if } |x| \leq 1/2, \\ 0, & \text{otherwise}, \end{cases} \tag{22}$$

and $f(x) = f(t = T/2, x)$. We fix the time to the center of the exposure time $t = T/2$ for avoiding the effects of the temporal boundaries of the measurements.

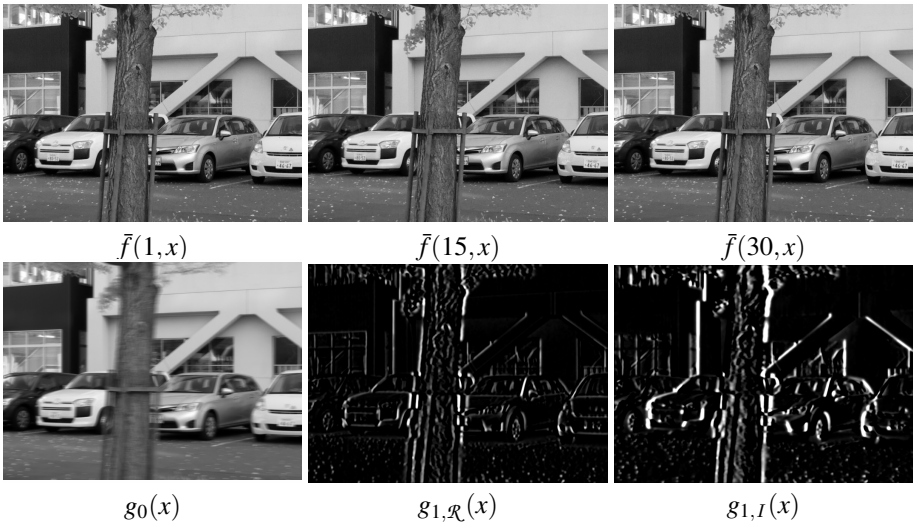


Figure 3: Top: Examples of the still images $\bar{f}(k,x)$. Bottom: Generated images, $g_0(x)$, $g_{1,\mathcal{R}}(x)$, and $g_{1,I}(x)$.

A motion blur generated on the complex Fourier coefficient image $g_n(x)$ by the target motion can be represented by a spatial convolution with a different one-dimensional filter $h_n(x)$. Let $x = vt$ denote the location of the moving target in the image, where the origin of the coordinate system is temporarily set to the target's location at $t = 0$. Then, multiplying $v^\intercal$ from the left, we obtains the equation $t = v^\intercal x/v^2 = (v_x \cos\theta + v_y \sin\theta)/v$. Substituting this equation and $\omega_0 = 2\pi/T$ into the representation of the reference temporal signal, $e^{jn\omega_0 t}$, we obtain a spatial filter that corresponds to the reference signal $e^{2\pi jn(v_x\cos\theta+v_y\sin\theta)/vT}$. The profile of this spatial filter along the trajectory of the moving target is a complex sinusoidal curve with frequency n/vT. Multiplying this spatial filter with h_0, we get $h_n(x)$, where

$$h_n(x) = h_0(x)e^{2\pi jn(x\cos\theta+y\sin\theta)/vT}, \qquad (23)$$

and this spatial filter generates the Fourier coefficient image, $g_n(x)$, as given by the following equation:

$$g_n(x) = f(x) * h_n(x). \qquad (24)$$

Let the spatial Fourier transformation of $f(x)$ be denoted by $F(u)$, where $u = (u_x, u_y)^\intercal$ denotes the two-dimensional spatial frequency and let the Fourier transformation of $\tilde{f}_1(x)$ in (14) be denoted by $\tilde{F}_1(u)$. We can now derive the following equation from (14)

$$\tilde{F}_1(u) = F(u)\{H_0(u) + H_1(u) + H_1^*(-u)\}, \qquad (25)$$

where $H_n(u)$ denotes the Fourier transformation of $h_n(x)$ and is given as

$$H_n(u) = H_0(u) * \delta(u_x\cos\theta + u_y\sin\theta + n/vT), \quad (26)$$

where $*$ is now a convolution with respect to the frequencies. The Fourier transformation of the one-dimensional box filter, $h_0(x)$, is given as a sinc function such that

$$H_0(u) = \frac{\sin(\pi(Tv_xu_x + Tv_yu_y))}{\pi(Tv_xu_x + Tv_yu_y)}. \qquad (27)$$

Finally, we obtain

$$H_n(u) = \frac{\sin(\pi(Tv_xu_x + Tv_yu_y + n))}{\pi(Tv_xu_x + Tv_yu_y + n)}. \qquad (28)$$

Figure 2 shows the profiles of H_0 and of H_1 along a line parallel to the motion direction. As shown by the profile of H_0, h_0 is a low-pass filter and generates the spatial motion blur of a moving target in $g_0(x)$. The graph of H_n is obtained by shifting that of H_0 by n toward the motion direction and h_n is a band-pass filter of which the spatial center frequency is n/vT. The approximation $\hat{f}_1(x)$ shown in (19) is computed with g_0 and g_1 and the latter measurement g_1 improves the spatial bandwidth by adding the two band-pass filters H_1 and H_1^* to the low-pass filter H_0 as shown in (25).

It should be noted that one can increase the bandwidth of the restoration and can obtain crisper images if one can estimate the higher frequency components $g_n(x)$ $(n > 1)$ and approximates as follows:

$$\tilde{f}_N(t,x) = \sum_{n=-N}^{N} g_n(x)e^{jn\omega_0 t}, \qquad (29)$$

where $N \geq 1$. The bandwidth of $\tilde{f}_N(t)$ is wider than that of $\tilde{f}_1(t)$ when $N \geq 1$ and the restoration accuracy is improved by the addition of g_n $(n \geq 1)$. The method for the estimation of $g_n(x)$ is described next.

4.2.2 Estimation with Sparsity Regularization

The proposed method estimates $g_n(x)$ $(n > 1)$ by using (5). As mentioned above, the values of $v_x(x)$,

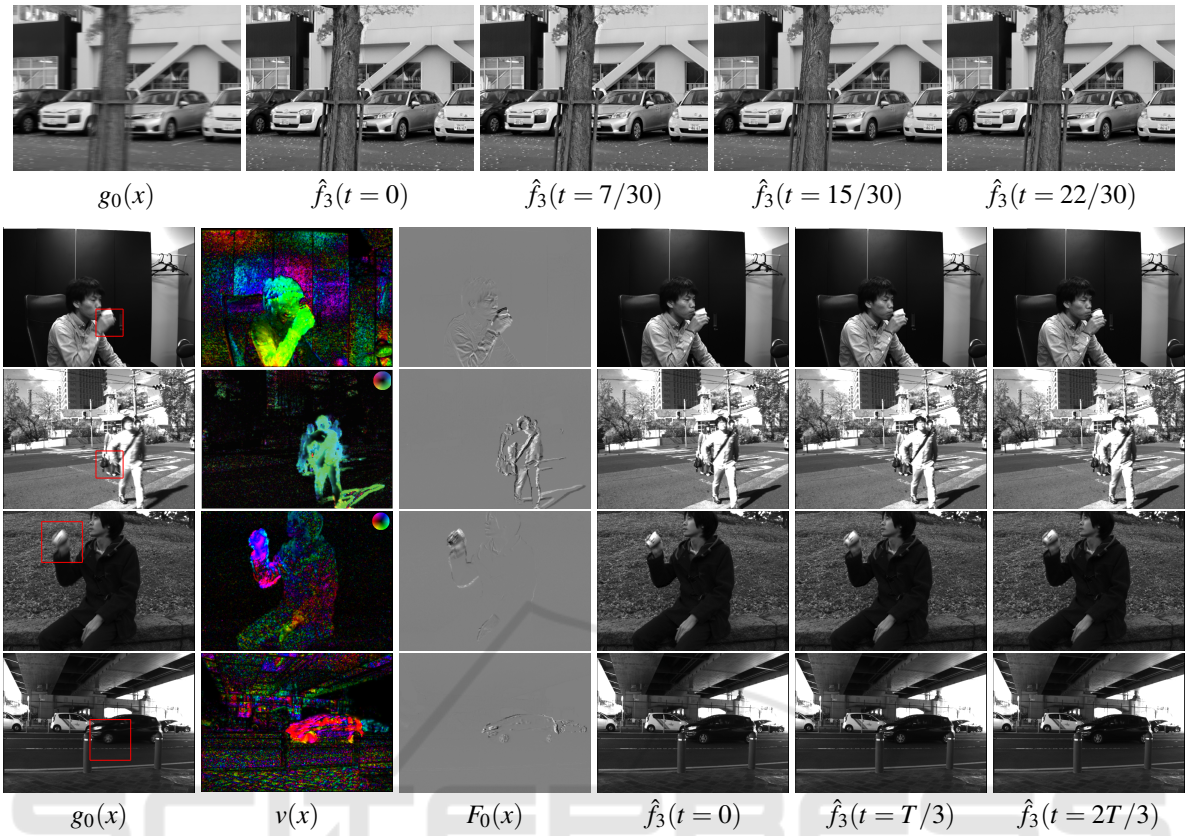


Figure 4: Top row: $g_0(x)$ artificially generated from $M=30$ still images shown in Fig.3 and $\hat{f}_3(t,x)$ for each t. The Second to Bottom rows: $g_0(x)$ measured by a CIS and the results, $v(x)$, $F_0(x)$, and $\hat{f}_3(t,x)$, computed from the images captured by the CIS.

$v_y(x)$, and $F_0(x)$ can be estimated by using the measured values of $g_0(x)$ and $g_1(x)$. No other measurements are needed for the estimation. Once these values are estimated, $g_n(x)$ is the only unknown variable in the linear complex equation shown in (5) and one can estimate its value by solving the equation.

These estimated values, though, can be inaccurate especially when the differential equation (5) does not hold as some irregular events like occlusions or specular reflections occur. We need a robust estimation method that automatically detects and excludes data that do not obey an employed model. Assuming that the regions in which the equation (5) does not hold are sparse in a given image, the proposed method uses a regularization technique proposed in (Ayvaci et al., 2012) for making the estimation robust against such irregulars.

Let $e(x)$ denote a residual of the right hand side of (5) defined as

$$e(x) \doteq \left(v_x\frac{\partial}{\partial x}+v_y\frac{\partial}{\partial y}\right)g_n(x)+F_0(x)+jn\omega_0 g_n(x). \tag{30}$$

Let D denote the entire image domain and let Ω denote subregions in D in which (5) does not hold. We assume that the residual $e(x)$ obeys a normal distribution with zero mean and small variance, if $x \in D\setminus\Omega$, where (5) is satisfied. If $x\in\Omega$, on the other hand, (5) is not satisfied and the residual, $e(x)$, can have an arbitrary value, $\rho(x)$.

Let $e(x)=e_1(x)+e_2(x)$ such that

$$e_1(x)=\begin{cases}\rho(x), & x\in\Omega,\\ 0, & x\in D\setminus\Omega,\end{cases} \tag{31}$$

and

$$e_2(x)=\begin{cases}0, & x\in\Omega,\\ \mathcal{N}(x), & x\in D\setminus\Omega,\end{cases} \tag{32}$$

where $\mathcal{N}$ denotes a variable that obeys the normal distribution. Then, e_1 is large and sparse, and e_2 is small and dense. Based on the discussion above, the proposed method minimizes the following cost function:

$$\begin{aligned} J_n(g_n,e_1) &= \|e_2\|_{\mathbb{L}^2(D)}+\alpha\|e_1\|_{\mathbb{L}^1}+\beta\|g_n(x)\|_{\mathbb{L}^2},\\ &= \|(v^\intercal\nabla)g_n+F_0+jn\omega_0 g_n-e_1\|_{\mathbb{L}^2(D)}\\ &+ \alpha\|e_1\|_{\mathbb{L}^1}+\beta\|g_n\|_{\mathbb{L}^2} \end{aligned} \tag{33}$$

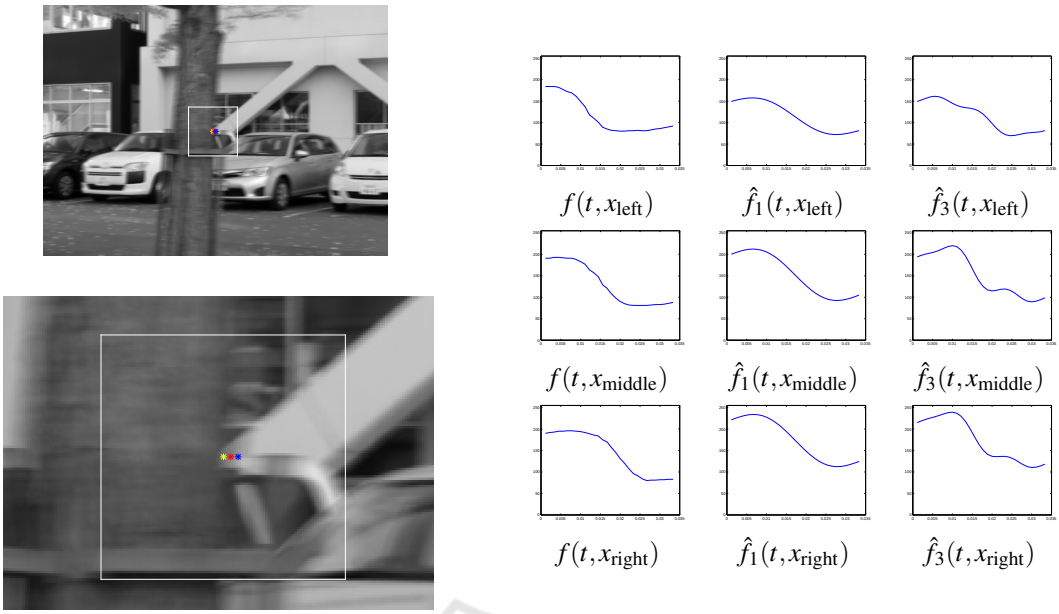


Figure 5: Left Panel: The location of three pixels indicated by three dots (yellow (x_{left}), red (x_{middle}). Right Panel: Examples of $f(t=(k-1)/M)$ (left column), $\hat{f}_{N=1}(t)$ (middle column), and $\hat{f}_{N=3}(t)$ (right column) at the three pixels.

where α and β are positive scalar coefficients for the regularization terms and their values are experimentally determined. $J_n(g_n,e_1)$ is convex with respect to g_n and e_1 when v and F_0 are fixed and then one can obtain the unique solution of $g_n(x)$ and $e_1(x)$.

5 EXPERIMENTAL RESULTS

The performance was evaluated with sets of simulated data and with images captured by a CIS.

5.1 Experiments with Simulated Data

Sets of M still images of a stationary scene, $\{\bar{f}(k,x)|k=1,2,\ldots,M\}$, were captured by a slowly translating traditional camera for simulating $f(t,x)$ $(0\le t<1)$ and artificially generated $g_0(x)$ and $g_1(x)$ from each of the sets as follows:

$$g_0(x) = \frac{1}{M}\sum_{k=1}^{M}\bar{f}(k,x), \quad (34)$$

$$g_{1,\mathcal{R}}(x) = \frac{1}{M}\sum_{k=1}^{M}\bar{f}(k,x)\cos(2\pi(k-1)/M), \quad (35)$$

$$g_{1,I}(x) = \frac{1}{M}\sum_{k=1}^{M}\bar{f}(k,x)\sin(2\pi(k-1)/M). \quad (36)$$

Figure 3 shows examples of $\bar{f}(k,x)$ $(M=30)$ and the corresponding $g_0(x)$ and $g_1(x)$. One can see the motion blurs in $g_0(x)$. The top row in Fig. 4 shows $g_0(x)$ again and some restored latent images, $\hat{f}_3((k-1)/M,x)$ $(k=0,7,15,22)$. Comparing with $g_0(x)$ shown in the left, we can see that the proposed method suppressed the motion blur in $g_0(x)$ and restored the temporal change of the light strength during the exposure time by computing $\hat{f}_3(t,x)$. Examples of the true signal $f(t=(k-1)/M,x)$ and the restored signals $\hat{f}_N(t,x)$ at three neighboring pixels are shown in Fig.5. The locations of the three pixels are shown at the left panel in the figure. They were located near the right boundary of a tree, which moved toward the right. The left column in the right panel shows the true profiles of $f(t,x)$ at the three points x_{left}, x_{middle}, and x_{right}. The value of f decreased rapidly when the tree, which moved from left to right in the image, reached to each pixel. The graphs in the middle column and in the right column show the restored temporal change of the light strength at each pixel with $N=1$ (middle column) and with $N=3$ (right column) as shown in (19), respectively. The latent true signals $f(t,x)$ in the left column were smooth enough and no higher frequency components were required for describing the true signal. Hence, $\hat{f}_1(t,x)$, which consists of only lower frequency components, approximated the latent signal more accurately than $\hat{f}_3(t,x)$, in which one can see some artifacts like aliasing. It is included in our future works to adaptively determine the appropriate value of N in (19) for each pixel.

$g_0(x)$ $\hat{f}_3(t=0,x)$ $\hat{f}_3(T/3,x)$ $\hat{f}_3(2T/3,x)$

Figure 6: Enlarged parts of the images shown in Fig.4

5.2 Experiments with Images Captured by a CIS

A set of images was captured by the CIS, and the image size was 512×704, with the exposure time set as $T = 1/30$ s. Examples of $g_0(x)$ are shown at the leftmost panels in the second to bottom rows of Fig.4. Firstly, solving the linear equation (8), we computed the flow v, and the difference of the boundary values $F_0(x)$ for each image. Examples of the obtained results are shown in Fig.4. As shown, $F_0(v)$ had nonzero values around the regions corresponding to the motion blurs. Then, minimizing $J(g_n, e_1)$ in (33), the method computed $g_n(x)$ $(n = 2,3)$ and e_1. Using the estimated values of g_n $(n = 2,3)$ with the measured values, g_0 and g_1, the method restored images $\hat{f}_3(t,x)$ as shown in the middle and the bottom rows in Fig.4. Enlarged parts of the images, $g_0(x)$ and $\hat{f}_3(t,x)$, are shown in Fig.6. Comparing with $g_0(x)$ at the leftmost panel, we can see that the restored images include less motion blurs.

6 SUMMARY AND FUTURE WORK

We proposed a CIS-based method that removes motion blurs from a single image and restores the latent temporal images, which represent the temporal change of the light strength during the exposure period. We believe that our proposed method would largely improve the stability and accuracy of motion analysis including landmark tracking or optical flow computation that are crucial in medical image analysis.

One advantage of the proposed method is found especially in the restoration of the latent temporal images. The restoration of the temporal change of the light strength during the exposure time from a single given image is very difficult when the image is cap-

tured with a traditional image sensor. On the other hand, a CIS modulates the temporal change of the light strength at each pixel with the sinusoidal reference signals and records the temporal change with its Fourier coefficients. Using these coefficients, one can compute optical flow, $v(x)$, and the difference of the boundary values, $F_0(x)$, from a single image captured by a CIS. Once one obtains $v(x)$ and $F_0(x)$, one can restore the higher temporal frequency components, $g_n(x)$ $(n \geq 2)$, based mainly on the optical flow constraint, which represents the temporal invariance of the strength of an incident light coming from an object point.

The biggest limitation of the proposed method is that the optical flow constraint (5) used in the proposed method assumes that a flow observed at each pixel, $v(x)$, is constant with respect to time during the exposure time. This is not true especially when the motion blurs are generated with a high-frequency motion such as a hand shake. Many blind motion deblurring methods can estimate spatial blurring kernels from a single blurred image by introducing the prior knowledge on natural images and/or on kernels. The future works would include the use of the strategies employed by those blind motion deblurring methods for estimating the spatial blurring kernels that are consistent not only with the blurred image $g_0(x)$ but also with the Fourier coefficient image, $g_1(x)$, so that one can restore more accurate and crisp images that represent the temporal change during the exposure time.

REFERENCES

Agrawal, A. and Raskar, R. (2009). Optimal single image capture for motion deblurring. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2560–2567. IEEE.

Ando, S. and Kimachi, A. (2003). Correlation image sensor: Two-dimensional matched detection of amplitude-modulated light. *Electron Devices, IEEE Transactions on*, 50(10):2059–2066.

Ando, S., Nakamura, T., and Sakaguchi, T. (1997). Ultrafast correlation image sensor: concept, design, and applications. In *Solid State Sensors and Actuators, 1997. TRANSDUCERS'97 Chicago., 1997 International Conference on*, volume 1, pages 307–310. IEEE.

Ayvaci, A., Raptis, M., and Soatto, S. (2012). Sparse occlusion detection with optical flow. *International Journal of Computer Vision*, 97(3):322–338.

Cai, J.-F., Ji, H., Liu, C., and Shen, Z. (2012). Framelet-based blind motion deblurring from a single image. *Image Processing, IEEE Transactions on*, 21(2):562–572.

Cho, S. and Lee, S. (2009). Fast motion deblurring. In *ACM Transactions on Graphics (TOG)*, volume 28, page 145. ACM.

Deshpande, A. M. and Patnaik, S. (2014). Uniform and non-uniform single image deblurring based on sparse representation and adaptive dictionary learning. *The International Journal of Multimedia & Its Applications (IJMA)*, 6(01):47–60.

Fergus, R., Singh, B., Hertzmann, A., Roweis, S. T., and Freeman, W. T. (2006). Removing camera shake from a single photograph. *ACM Transactions on Graphics (TOG)*, 25(3):787–794.

Field, D. J. (1994). What is the goal of sensory coding? *Neural computation*, 6(4):559–601.

Gupta, A., Joshi, N., Zitnick, C. L., Cohen, M., and Curless, B. (2010). Single image deblurring using motion density functions. In *Computer Vision–ECCV 2010*, pages 171–184. Springer.

Heide, F., Hullin, M. B., Gregson, J., and Heidrich, W. (2013). Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)*, 32(4):45.

Hontani, H., Oishi, G., and Kitagawa, T. (2014). Local estimation of high velocity optical flow with correlation image sensor. In *Computer Vision–ECCV 2014*, pages 235–249. Springer.

Jia, J. (2007). Single image motion deblurring using transparency. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE.

Joshi, N., Kang, S. B., Zitnick, C. L., and Szeliski, R. (2010). Image deblurring using inertial measurement sensors. In *ACM Transactions on Graphics (TOG)*, volume 29, page 30. ACM.

Kadambi, A., Whyte, R., Bhandari, A., Streeter, L., Barsi, C., Dorrington, A., and Raskar, R. (2013). Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (TOG)*, 32(6):167.

Levin, A. (2006). Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems*, pages 841–848.

McCloskey, S., Ding, Y., and Yu, J. (2012). Design and estimation of coded exposure point spread functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(10):2071–2077.

Nayar, S. K. and Ben-Ezra, M. (2004). Motion-based motion deblurring. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):689–698.

Raskar, R., Agrawal, A., and Tumblin, J. (2006). Coded exposure photography: motion deblurring using fluttered shutter. *ACM Transactions on Graphics (TOG)*, 25(3):795–804.

Shan, Q., Jia, J., and Agarwala, A. (2008). High-quality motion deblurring from a single image. In *ACM Transactions on Graphics (TOG)*, volume 27, page 73. ACM.

Shan, Q., Xiong, W., and Jia, J. (2007). Rotational motion deblurring of a rigid object from a single image. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE.

Tai, Y.-W., Du, H., Brown, M. S., and Lin, S. (2010). Correction of spatially varying image and video motion blur using a hybrid camera. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(6):1012–1028.

Velten, A., Wu, D., Jarabo, A., Masia, B., Barsi, C., Joshi, C., Lawson, E., Bawendi, M., Gutierrez, D., and Raskar, R. (2013). Femto-photography: capturing and visualizing the propagation of light. *ACM Transactions on Graphics (TOG)*, 32(4):44.

Wei, D., Masurel, P., Kurihara, T., and Ando, S. (2009). Optical flow determination with complex-sinusoidally modulated imaging. *relation*, 7(8):9.

Whyte, O., Sivic, J., Zisserman, A., and Ponce, J. (2012). Non-uniform deblurring for shaken images. *International journal of computer vision*, 98(2):168–186.

Xu, L. and Jia, J. (2010). Two-phase kernel estimation for robust motion deblurring. In *Computer Vision–ECCV 2010*, pages 157–170. Springer.

Xu, L., Zheng, S., and Jia, J. (2013). Unnatural l0 sparse representation for natural image deblurring. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1107–1114. IEEE.