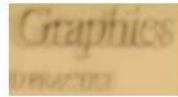# Synthetic Shutter Speed Imaging

J. Telleen , A. Sullivan , J. Yee , O. Wang , P. Gunawardane , I. Collins , J. Davis
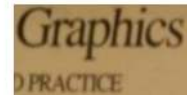
University of California, Santa Cruz

| No Stabilization | Optical Stabilization Lens | Synthetic Shutter Speed Image |

Abstract

Hand held long exposures often result in blurred photographs due to camera shake. Long exposures are desirable both for artistic effect and in low-light situations. We propose a novel method for digitally reducing imaging artifacts, which does not require additional hardware such as tripods or optical image stabilization lenses. A series of photographs is acquired with short shutter times, stabilized using image alignment, and then composited. Our method is capable of reducing noise and blurring due to camera shake, while simultaneously preserving the desirable effects of motion blur. The resulting images are very similar to those obtained using a tripod and a true extended exposure.

1. Introduction

Digital cameras offer a new opportunity for camera manufacturers to solve problems that plague photographers. Many solutions that were intractable with film cameras are possible in the digital realm. Unfortunately most current cameras treat digital sensors as mere replacements for analog film. By treating these sensors as flexible measurement devices new opportunities arise. This paper addresses one such challenge, hand held long exposures, by making use of rapid sequential images and embedded computation that would not have been possible on a film camera.

Photographers often need to open the camera shutter for extended periods of time in order to capture the desired image. Unfortunately it is difficult to hand hold a camera for long periods without introducing undesirable blurring due to camera shake. One motivation for an extended exposure comes when part of the scene is moving. A long exposure allows the motion to be captured as a localized blur, which conveys a sense of presence and action. Another motivation is noise reduction in low light situations. When the illumination level is low, the random arrival of photons on the camera's sensor results in a grainy image, which can only be compensated by a prolonged exposure. This unfortunately leaves the camera vulnerable to motion blur due to shake. To reduce camera shake, the photographer may carry a tripod or utilize an optical image stabilizer, but these are not always viable options.

This paper proposes a solution, synthetic shutter speed imaging(SSSI), which relies on trading a traditionally analog process for one in which parts of the process are replaced by computation. A camera's sensor integrates all of the incoming light for the entire time the shutter is open. If the camera moves while the shutter is open, then light from several scene points will be integrated together and a blurry image will result. Although this integration traditionally occurred chemically in film, and now occurs electrically on a chip, it is possible to perform this integration computationally as well. A similar blurry image could be computed by taking many photographs with short exposures in rapid succession, and summing (integrating) the sequence. Taking advantage of the computational realm and replacing the simple summation with a more sophisticated procedure allows an improved result, which does not exhibit blur.

Blur in the final result can be removed by adjusting each source photograph so that pixels, which sample light from the same scene point, are aligned. An automated image registration step can be used to compute this alignment. After alignment the images can be "correctly" integrated using a simple summation. This simple method is sufficient to remove blur. However, artifacts will remain due to the sparse temporal sampling available on most digital still cameras.

Even when taking photos in rapid succession, digital still cameras sample only discrete time intervals, and do not record photons in the intervening period. If the scene contains moving objects double images or ghosting may result when images are merged. This is not a problem when we use a single long exposure, and would not be a problem if we had an ideal camera, which had greater temporal sampling. For example, low-resolution video camcorders often have this ideal behavior, with very little "gap" between neighboring frames.

We address this challenge of double images by creating synthetic motion blur to fill in the temporal gaps between our samples. Optical flow is used to track the motion of different scene regions. This estimate of flow allows the creation of virtual photographs at a higher temporal sampling rate. By combining all images, both the original source sequence and the new virtual photos, results can be created that exhibit the expected smooth motion blur, rather than visually unappealing double images.

The primary contribution of this paper is a novel method for producing hand held pictures under long exposures without artifacts due to camera shake. The method reduces blur better than optical image stabilization lenses, and is cheaper to implement since it does not require special hardware. Furthermore it preserves the desirable motion blur caused by moving scene elements. Lastly, the method is extremely simple and one contribution of our work is showing that such a simple method works surprisingly well.

2. Related Work

The most common "anti-shake" technology in existing digital cameras is an inertial sensor embedded in the lens, which controls a compensating optical element. Canon calls this "Image Stabilization", while Nikon calls it "Vibration Reduction." Konica Minolta has a variation in which the image sensor, rather than lens elements, is adjusted in response to motion. Unfortunately these sensors drift over time and are useful only for relatively short exposures.

One method of removing undesirable blur due to camera shake is through deconvolution. Since the underlying motion of the camera which caused the blur is not typically known, this is often formulated as a blind deconvolution task in which the point spread function (PSF) and the resulting image are estimated jointly. Obtaining the PSF prior to deconvolution can significantly improve results. For example, Ben-Ezra and Nayar built a special rig with a separate low-resolution high frame rate sensor used to estimate the PSF [BEN04]. Fergus et. al. use statistics in the gradient domain to form a prior to find the blur kernel and latent image with highest probability before image deconvolution [FSH*06]. Raskar et. al. use a coded exposure by fluttering the shutter during an exposure to preserve high frequency details to constrain deconvolution [RAT06]. All three methods produce amazing results that contain significantly more high frequency detail than does the original blurred image, however because the problem is under constrained, distracting visual artifacts remain. Furthermore, none of these deconvolution methods adequately deal with the need to preserve blur due to moving objects.

Whenever the need for an extended exposure is due to low light, an alternate approach is to capture a short exposure and then eliminate the resulting noise. A wide variety of single frame image denoising methods have been proposed, and a useful survey exists [BKE*95]. However these methods often have the side effect of removing non-noise high frequency components of the image, and cannot produce adequate results for our application.

Exposure correction has also been addressed in the context of low light photography as a multi-frame problem. For example, it is possible to combine a flash and non-flash image pair to produce a low noise image with pleasing lighting [PSA*04, ED04]. Similarly Kang et. al. captured video with alternating short an long exposures and used this to create high dynamic range video [KUWS03]. An alternate approach was introduced by Jia et. al. who corrects an underexposed image taken with a short shutter time by combining it with the color histogram taken from a blurry image acquired with a longer shutter [JSTS04]. Jia et. al. did not address the problem of image noise in their underexposed images, and none of the low light methods address preservation of desirable motion blur due to moving scene elements during a long exposure.

Several of the component technologies used in our work

have been explored in alternate domains. For example, it is common to stabilize a sequence of frames obtained from a video camera using frame-to-frame registration [HLL05]. Similarly, Bennett et. al. explores noise reduction in low light video sequences by aligning frames and then applying sophisticated temporal integration across frames [BM05]. Gamal et. al. use multiple exposures to simultaneously remove blur and generate high dynamic range images [LG01]. Finally, Brostow et. al. show that motion blur can be added into stop motion animation sequences using an optical flow calculation [BE01]. None of this prior work has addressed the need to reduce blur caused by camera shake while simultaneously retaining the artistically desirable blur caused by moving scene elements.

3. Method

3.1. Compositing

In traditional imaging, for each image sensor pixel, light is integrated over an interval of time (shutter time) starting from the moment T when the shutter opens. (Equation 1, Figure 2)

$$image(\vec{X}) = \int_{T}^{T+\Delta T} light(t, \vec{X}) dt \qquad (1)$$

ΔT is the shutter time, T the time at which the shutter opens, and $light(t, \vec{X})$ is the amount of light arriving at the sensor at time t at position $\vec{X}$.

Images taken with a short exposure time have a low signal to noise ratio(SNR) due to photon noise, which is related to the statisical nature of photons arriving at an image sensor. This is preferable to blurring due to camera shake, which is more difficult to eliminate in post processing. Noise is still undesirable and needs to be removed to synthesize a high quality image. Our process sums several images of the same scene. Since each pixel is sampled over a longer time interval there is an increase in the SNR of the final image.

We use an approximation of the imaging equation using discrete time intervals of length ΔT starting at times T_0 to T_n spanning a long exposure time instead of a continuous integration over the long exposure time (Equation 2).

$$image(\vec{X}) = \sum_{k=0}^{n} \int_{T_k}^{T_k+\Delta T} light(t, \vec{X}) dt \qquad (2)$$
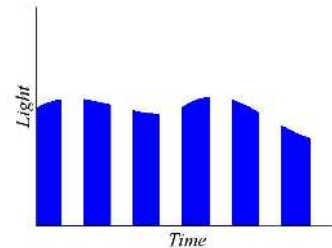
In an ideal scenario the delay between capturing a pair of images would be minimal (Figure 2), or non-existent ($T_k + \Delta T \longrightarrow T_{k+1}$). This gives a sampling which when combined equates to the traditional method.
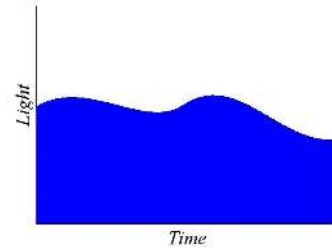
Traditional Long Exposure
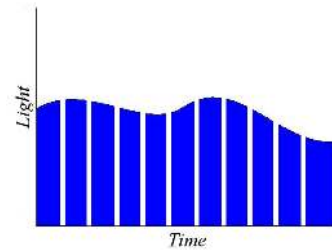
Ideal Camera Capture

Digital Camera Capture

Figure 2: *The blue regions represent the time when incoming light was sampled. A traditional exposure samples the incoming light over a continuous time interval. The ideal camera capture has small time interval between samples to minimize sampling artifacts. The current generation of digital still cameras have large intervals that can produce significant sampling artifacts.*

3.2. Alignment

Image sensors are not perfectly stable, over time the pixel being sampled would drift (as a result of 'shake' or 'jitter') and the light arriving at the sensor would no longer be coming from the same scene point resulting in blur. To enforce the constraint that the images contain the same spatial information, they must be aligned to form a correspondence between pixels. This alignment is impossible in a traditional camera since the integration of different scene elements occurs optoelectronically in the camera pixel.

There are a number of different methods, which can be

used for the initial alignment of images, and there has been a lot of research done in this area [LK81, BM04]. We used a publicly available Matlab registration toolkit to align images [Hee96]. It is an implementation of a gradient-based approach to align two images. It uses a multi scale, coarse to fine, refinement of the velocity estimations. These velocities provide information about the camera motion in order to generate a transformation matrix. The alignement step performs a pair wise alignment to generate a seperate transformation matrix for each image. Our process use the image that comes first temporally as the base image. All images are aligned to the same base image to avoid error accumulation in the alignment process.

When we capture a set of short exposure images in succession, the light is collected over n discrete sampling intervals of length ΔT_s starting at times T_0 to T_n. Each image is aligned according to a transformation matrix F_k obtained using an image registration algorithm. The imaging equation is approximated by summing over a set of discrete intervals spanning the desired imaging interval (Equation 3).

$$image(\vec{X}) = \sum_{k=0}^{n} \int_{T_k}^{T_k+\Delta T} light(t, F_k \vec{X}) dt \qquad (3)$$

3.3. Adding Motion Blur

Most digital capture devices have a significant amount of delay between the capturing of two adjacent frames. In this case, the sampling of light arriving at a single pixel becomes a poor approximation to the true continous function.

Temporally spaced discrete sampling intervals introduce a disparity between the actual time the light was sampled, and the duration of the image. For example if we take two images with a shutter speed of 1/50 second, with a delay of 1/25 seconds between the pair, by adding the two images the scene has effectively been sampled over 1/25 second. However, the actual duration of the composite image spans 2/25 seconds. This would generate artifacts if there were moving scene elements. Of course if the scene was static, this would not matter, but that is an unrealistic assumption. So what we need is a way to imitate the motion blur generated by moving parts of the scene over the total imaging duration, and we achieve this by applying our synthetic motion blur technique.

As mentioned before, at the same time while eliminating the blur due to camera shake, it is desirable to preserve the motion blur due to moving scene elements. If we used simple summation that is spatially uniform we observe a stepping artifact where the demarcation between the images is clearly visible. Therefore, we use an optical flow based method to achieve a more continuous motion path across the set of images. Due to the discrete nature of our temporal sampling, a scene point in motion would appear at pixel co-ordinates

$(x_1, y_1), (x_2, y_2), ..(x_n, y_n)$ across the set of images (where n is the number of images) where (x_i, y_i) and (x_{i+1}, y_{i+1}) would likely be not adjacent pixels.

We use the input image sequence as key frames and interpolate to get in between frames using the motion vectors generated by computing optical flow on temporally adjacent pairs of images. A coarse to fine hierarchical optical flow algorithm is used to compute the local flow per pixel, this could be replaced by any well regarded local optical flow algorithm.

The flow field we obtain is used to create background template for the scene, This is used to fill in any holes present in inbetween frames. The background template is created by selecting pixels from each of the frames that are considered to be backgorund. A pixel is considered to be background if it motion is below a threshold, 2 pixels motion was determined to achieve adequate results. In the cases were multiple frames have a background pixel at the same location the pixel that comes first temporally is used.

Inbetween frames are generated by bilinearly interpolating along the motion vectors obtained by optical flow. The number of in between frames that are rendered is determined by the largest pixel motion across all frames (Equation 4), where S is a distance in pixels. This is done to ensures that all inbetween frames are spaced equally in time across the imaging interval. Any holes that are present are filled using the background template. The final set of images is then summed to get the final output image (Equation 5).

$$S = maxLength(\forall k : V_k(\overset{X}{})) \qquad (4)$$

$$image(\vec{X}) = \sum_{k=0}^{n} \sum_{i=1}^{S} \int_{T_k}^{T_k+\Delta T} light(t, F_k \vec{X} + \frac{i * V_k(F_k \vec{X})}{S}) dt \qquad (5)$$

The proccess creates several virtual snapshots that are used to fill in the time between the captured shots. The proccess eliminates unsightly stepping artifacts by summing the set of virtual snapshots along with the keyframes.

4. Results

The most common method of stabilizing long hand-held shots is with an optical stabalization lens. The Figure 1 shows a comparison of our method to traditional options. We used a Canon EOS 2D with an EF 75-300mm IS lens. The long exposures were taken with a shutter speed of $0.4sec$. at 800 ISO The short exposures were taken with a shutter speed of $0.008sec$. at 3200 ISO. Without stabilization, the text in the image is unreadable. The image that was captured with optical stabilization is more readable, however the image still contains noticeable blur due to camera shake. The
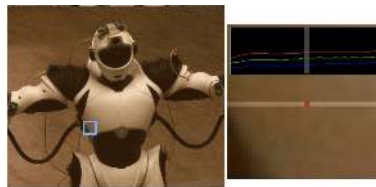
synthetic shutter speed image was created with 13 short exposure, the text is now clearly readable.
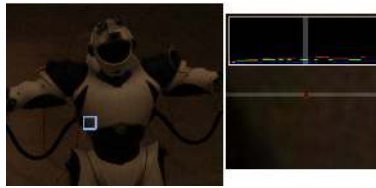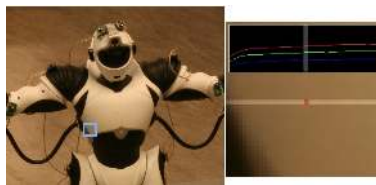

Short Exposure
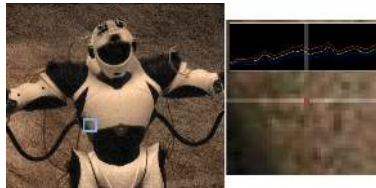
Short Exposure With Adaptive
Histogram Equalization

Synthetic Shutter Speed Image

Long Exposure on Tripod

Figure 3: *Much of the detail in the short exposure image is difficult to make out because it is too dark. The short exposure with adaptive histogram equalization reveals more detail but also has significant noise and color artifacts. The synthetic shutter speed image reveals detail in the image while reducing noise and providing better color. A long exposure on a tripod is present for comparison.*

Noise reduction is one important reason that a photographer might want take a long exposure. Figure 3 shows results that illustrate the noise reduction and exposure correction characteristics of synthetic shutter speed imaging compared with other exposure correction approaches. The short exposure is under exposed and has a low SNR making that image undesirable. One common approach for fixing an under exposed image is histogram equalization. The adaptively

histogram equalized image was created using the short exposure shown. The result is an image with better exposure but the color is muted and there is a noticeable amount of noise. The synthetic shutter speed image is the result from our method, which has a substantially reduced noise profile. A ground truth image taken with a tripod is provided for comparison. By comparing the enlarged regions from each image it is clear that adaptive histogram equalization makes the noise present in the under exposed image more apparent. Our method reduces the noise and looks closer to the long exposure taken with a tripod.

Long exposures are also used to intentionally record motion blur, giving photographs an enhanced sense of presence. Figure 4 provides an example of this usage. In the image a person is erasing a whiteboard, and the motion of the hand provides an important cue to understanding the action. The images were captured using a Point Grey Dragonfly Express high speed black and white camera. This camera acquires images with minimal time between exposures. The long exposures had a shutter time of $0.3secs$. The short exposures were taken with a shutter time of $0.03secs$. Notice in the long hand-held exposure that the entire image exhibits blur due to hand-shake. The blur is particularly evident in the blown up region of the top right of the white board, the equation written is not readable. A short exposure removes the blur due to hand shake, but also removes the desirable motion blur from the image. The synthetic shutter speed image was created using ten short exposures giving a total light integration time of 0.3. Notice that is preserves the motion blur but does not introduce blurring due to shake. A tripod mounted long exposure is presented for comparison.

To provide a more careful comparison on a precisely repeatable motion we used a rotational stage controlled by a computer. Results are shown in Figure 5. The long exposures were captured with a shutter time of $1sec$, and 48 short exposures were captured with a shutter time of $0.02sec$. All of the images were taken with a Canon EOS 20D with an EFS 18-55mm lens. Since this camera is not capable of taking images in rapid succession, the sequence of short exposures spanned 5 seconds. We slowed the motion of the controllable motion stage so that the same angular motion was present in both the long exposures and the short exposure sequence. In the hand-held long exposure image you will notice that the entire image contains significant blur. Unlike the long hand-held exposure, the short exposure does not contain any noticeable blur. Synthetic shutter speed imaging preserves the desirable motion blur while removing camera shake, producing a result very similar to the ground truth image captured on a tripod.

Objects that move too quickly within a scene can have a negative effect on our composite image. If the motion is larger than 2 pixels ghosting can occur. In these cases we use optical flow to compute a new image with appropriate motion blur. Figure 6 provides an example. We took pictures
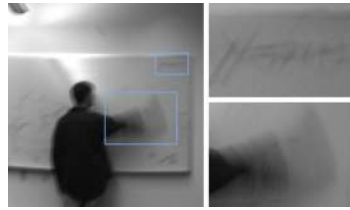
Long Exposure Hand-Held
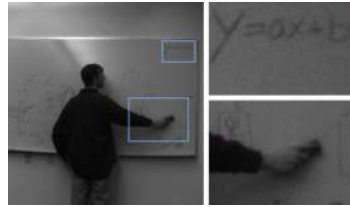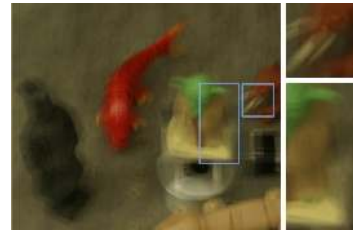

Short Exposure

Synthetic Shutter Speed Image

Long Exposure on Tripod

Figure 4: *The hand-held long exposure is blurry making the equation on the whiteboard unreadable. In the short exposure the equation is readable but there is no motion blur in the hand. The synthetic shutter speed image has motion blur in the hand and the equation is readable. A long exposure on a tripod is present for comparison..*

Long Exposure Hand-Held

Short Exposure

Synthetic Shutter Speed Image

Long Exposure on Tripod

Figure 5: *The hand-held long exposure contains blur everywhere in the image. In the short exposure there is no motion blur. Motion blur is apparent in zoomed in object on the rotational stage, while the rest of the scene remains unblurred. The long exposure on a tripod is presented for comparison.*

of a scene where the motion of an object was sufficiently rapid to cause ghosting artifacts if simple summation is used. All of the images were captured using a Canon EOS 2D with a EFS 18-55mm lens. The same sequence of three images was used to generate both photographs. The result generated without blur has noticable ghosting. Using optical flow to add motion blur removes the ghosting artifacts.

Fast non-rigid motions are a difficult case for our method. Current generations of digital cameras can not capture

frames quickly, and optical flow is insufficiently robust to track complex motions. Figure 7 is a scene containing a small waterfall with rapidly moving water. The hand-held long exposure contains blur across the whole image, and the sharp edges in the rocks are lost due to camera shake. The short exposure has been gain adjusted for visibility since it was originally very dark. Sharp edges are present in the short

Synthetic Shutter Speed Image No Blur

Synthetic Shutter Speed Image With Blur

Figure 6: *The synthetic shutter speed image without blur exhibits ghosting. The synthetic shutter speed image with our method for motion blur results in smooth motion blur.*

exposure that are blurred in the long exposure, however, the moving water regions do not exhibit any blur. The SSSI image is sharp, and contains a better appearance of flowing water than either of the traditional options. However, the region at the base of the waterfall shows a different kind of localized blur than is present in the long exposure taken on a tripod, and would represent an unacceptable substitute to most photographers. All of the images were captured with a Canon Rebel XT using an EFS 18-55mm lens. The long exposures were taken with a shutter time of 0.2*secs*. The short exposures were taken with a shutter time of 0.02*secs*. Our result was created from 10 hand held exposure with an effective shutter speed of 0.2*secs*.

5. Limitations

There are several places where our method can exhibit undesired behavior. This is largely related to instabilities in the underlying methods used in our image formation process such as image alignment and optical flow.

Image alignment relies on finding some feature set that is

Long Exposure Hand-Held

Short Exposure

Synthetic Shutter Speed Image

Long Exposure on Tripod

Figure 7: *The hand-held exposure has noticeable blur across the entire image. The short exposure has sharp edges but lacks motion blur in the water regions. The synthetic shutter speed image maintains the sharp edges of the short exposure, but the blur effect that is created does not look like the blur in the long exposure. The long exposure is presented for comparison.*

common between images. Features are used to determine a transformation between image pairs. It is possible for alignment failures under certain conditions. One case were this might occur is a scene with little texture or potentially a region that is out of focus or blurred. It is also assumed that

the majority of a scene is static. If a significant portion of a the scene undergoes a non-rigid motion the alignment might also fail. In these cases it is not possible to properly align the image thus we cannot claim that the same scene coordinate maps to the same image space coordinate in all of the images.

Optical flow is another possible source of compositing artifacts. Scenes with little texture can pose a problem for optical flow computations. This can result pixel with mislabeled motion vectors. Scene elements that change appearance between frames will also cause problems with optical flow calculations. If optical flow does not provide a reasonable flow field appropriate blur cannot be computed. This can cause unexpected regions to be blurred and regions with motion to remain unblurred.

6. Conclusions and Future Work

Our methodology provides image stabilization while preserving localized motion blur within a scene. Its simplicity makes it an inexpensive solution to the image stabilization problem compared to expensive optical image stabilization lenses. It does not require cumbersome tripods. Our method also preserves motion blur in scenes with moving objects. These algorithms do not require any extra hardware.

In the future we'd like to see the method implemented in camera hardware. This would give access to individual frames at a sufficiently fast rate, and allow a seamless user experience of extremely long handheld exposures with no camera shake.

References

[BE01] BROSTOW G. J., ESSA I.: Image-based motion blur for stop motion animation. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2001), ACM Press, pp. 561–566.

[BEN04] BEN-EZRA M., NAYAR S. K.: Motion-based motion deblurring. *IEEE Trans. Pattern Anal. Mach. Intell. 26*, 6 (2004), 689–698.

[BKE*95] BRAILEAN J., KLEIHORST R., EFSTRATIADIS S., KATSAGGELOS A., LAGENDIJK R.: Noise reduction filters for dynamic image sequences: a review. In *Proceedings of the IEEE* (1995), vol. 83, IEEE Computer Society, pp. 1272–1292.

[BM04] BAKER S., MATTHEWS I.: Lucas-kanade 20 years on: A unifying framework. *Int. J. Comput. Vision 56*, 3 (2004), 221–255.

[BM05] BENNETT E. P., MCMILLAN L.: Video enhancement using per-pixel virtual exposures. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), ACM Press, pp. 845–852.

[ED04] EISEMANN E., DURAND F.: Flash photography enhancement via intrinsic relighting. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (New York, NY, USA, 2004), ACM Press, pp. 673–678.

[FSH*06] FERGUS R., SINGH B., HERTZMANN A., ROWEIS S. T., FREEMAN W. T.: Removing camera shake from a single photograph. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM Press, pp. 787–794.

[Hee96] HEEGER D.: Notes on motion estimation. *Stanford University, 1996.* (1996).

[HLL05] HSU S.-C., LIANG S.-F., LIN C.-T.: A robust digital image stabilization technique based on inverse triangle method and background detection. *IEEE Transactions on Consumer Electronics 51*, 2 (May 2005), 335–345.

[JSTS04] JIA J., SUN J., TANG C.-K., SHUM H.-Y.: Bayesian correction of image intensity with spatial consideration. *ECCV 2004* (2004), 342–354.

[KUWS03] KANG S. B., UYTTENDAELE M., WINDER S., SZELISKI R.: High dynamic range video. In *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers* (New York, NY, USA, 2003), ACM Press, pp. 319–325.

[LG01] LIU X., GAMAL A. E.: Simultaneous image formation and motion blur restoration via multiple capture. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on* (Salt Lake City, UT, USA, 2001), vol. 3, IEEE Computer Society, pp. 1841–1844.

[LK81] LUCAS, KANADE: An iterative image registration technique with an application to stereo vision. *Proceedings of the 7th International Joint Conference on Artificial Intelligence* (1981), 671–679.

[PSA*04] PETSCHNIGG G., SZELISKI R., AGRAWALA M., COHEN M., HOPPE H., TOYAMA K.: Digital photography with flash and no-flash image pairs. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (New York, NY, USA, 2004), ACM Press, pp. 664–672.

[RAT06] RASKAR R., AGRAWAL A., TUMBLIN J.: Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM Press, pp. 795–804.