High Dynamic Range Image Reconstruction from Hand-held Cameras

Pei-Ying Lu   Tz-Huan Huang   Meng-Sung Wu   Yi-Ting Cheng   Yung-Yu Chuang

National Taiwan University

E-mail: {yummy,tzhuan,sula,zzz88213}@cmlab.csie.ntu.edu.tw, cyy@csie.ntu.edu.tw

Abstract

This paper presents a technique for reconstructing a high-quality high dynamic range (HDR) image from a set of differently exposed and possibly blurred images taken with a hand-held camera. Recovering an HDR image from differently exposed photographs has become very popular. However, it often requires a tripod to keep the camera still when taking photographs of different exposures. To ease the process, it is often preferred to use a hand-held camera. This, however, leads to two problems, misaligned photographs and blurred long-exposed photographs. To overcome these problems, this paper adapts an alignment method and proposes a method for HDR reconstruction from possibly blurred images. We use Bayesian framework to formulate the problem and apply a maximum-likelihood approach to iteratively perform blur kernel estimation, HDR image reconstruction and camera curve recovery. When convergence, we simultaneously obtain an HDR image with rich and clear structures, the camera response curve and blur kernels. To show the effectiveness of our method, we test our method on both synthetic and real photographs. The proposed method compares favorably to two other related methods in the experiments.

1. Introduction

High dynamic range (HDR) imaging has become more and more popular in recent years [8]. It attracts many professional and amateur photographers and finds utility in many applications such as visual effects production. Although a few special cameras can be used for directly capturing HDR images, they are still expensive and not prevalent. Most photographers still use standard cameras to take differently exposed photographs of the same scene and then composite HDR images from them by recovering the camera response curve [2, 6, 9]. One of the limitations for these approaches is that the camera must be kept still when taking photographs. A tripod or other specialized hardware is often necessary to enforce this requirement, but these devices are usually very bulky to bring with. Thus, most photographers would prefer to take photographs with conventional hand-held cameras. Unfortunately, images captured by hand-held cameras are likely to be blurry due to camera shake, especially for the ones with long exposures. Camera shake can be modeled as the convolution of the irradiance image with a blur kernel, describing the camera motion during exposure [4]. Since those differently exposed photographs taken by the hand-held camera always have different blur kernels, the directly recovered HDR image is often blurry even if they are somehow aligned.

A naïve solution is to apply image deblurring techniques [4, 10] to each of these blurred low dynamic range (LDR) images, and then to reconstruct the HDR image from the deblurred ones. However, there are problems with this approach. The quality of the reconstructed HDR image highly depends on the quality of the deblurred LDR images. However, the deblurred images may not be good enough, especially for the ones with long exposures. Individual deblurred images may have artifacts such as ringing. The problem is further aggravated in the combined HDR image because these artifacts are often not coherent spatially.

This paper proposes a new technique to reconstruct a sharp HDR image from a set of potentially blurry LDR images with different exposure times while estimating the blur kernels and the camera response curve at the same time. Taking the advantage of having multiple observations for the same scene, we formulate the problem based on a Bayesian framework to obtain all parameters at once. Furthermore, since we have multiple images, the dimension of the observations is actually larger than the one of the parameters to be estimated. Thus, a maximum-likelihood method is applied to iteratively optimize for the HDR irradiance image, the blur kernels and the camera response curve until converging to a good solution.

2. Background and related work

This section first introduces background for the problem. Next, we review the related work of two categories, HDR image reconstruction and image deblurring. For the former,
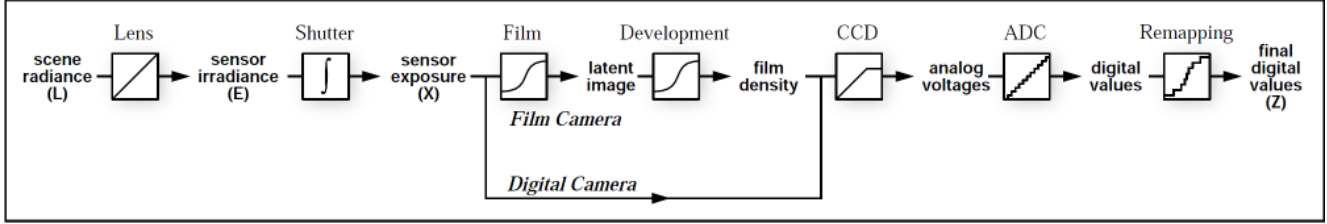
Figure 1. **Camera pipeline.** This pipeline shows how scene radiance maps to pixel values for both film and digital cameras. Unknown nonlinear mappings can occur during exposure, development, scanning, digitization, and remapping stages [2]. (Courtesy of Paul Debevec.)

we discuss related work on combining differently exposed images into a single HDR image. For the later, we discuss methods for recovering a sharp image from a single blurred image or multiple blurred images due to camera shake.

2.1. Camera pipeline and HDR fusion

Figure 1 illustrates a typical camera pipeline [2]. In general, the recorded pixel values Z are proportional to the scene radiance E, transformed by some nonlinear mapping called the camera response function. The nonlinearity is induced by different parts of the imaging process. As the light rays carrying radiance L pass through the lens from different directions, they are attenuated and integrated by the lens and aperture to yield sensor irradiance E. The energy of light is accumulated on the image sensor during integration time Δt controlled by the shutter. The total recorded energy is called the sensor exposure X,

$$X = \int_{t=0}^{\Delta t} E \, dt = E \Delta t. \tag{1}$$

Typical sensors like CCD or CMOS are designed to produce electrical signals which are linearly proportional to the sensor exposure up to some saturation level. Above this level, the sensor is not able to distinguish between different exposure values. After exposure, the accumulated charge is converted to integer values using analog-digital converters. The process of digitization brings quantization noise to the recorded data. In addition, for most digital cameras, the digital values are further transformed nonlinearly to mimic film characteristics or scaled by a gamma function to properly display on LCD. Individual sources of nonlinearity in the imaging process are not important and the whole process can be represented by one nonlinear function — the camera response function f. The measured pixel values Z are proportional to the sensor exposure X remapped by the response function f, *i.e.* $Z = f(X)$.

2.2. HDR construction

The idea of fusing images with different exposures into an HDR image is not new. Its history can be traced back to the early sixties. To recover an HDR image, multiple im-

ages of the same scene differing only in exposures are captured. Each captured image records different ranges of light intensities. If there is enough overlap between these ranges, an HDR image can be reconstructed by properly combining these captured images. Debevec and Malik proposed one of the most popular methods based on the assumption of reciprocity [2]. This method can reconstruct HDR images up to a scale factor. Mitsunaga and Nayar proposed a method for recovering camera response curves assuming that they are polynomials of Z [6]. Robertson *et al.* proposed a more robust approach for HDR image reconstruction based on a probabilistic formulation [9].

2.3. Image deblurring

Blur caused by camera shake is a common artifact in digital photography. For the situations when there is not enough light, a long shutter speed must be used and the resultant photographs are often blurry and disappointing. Recovering an unblurred image from a single blurred image or multiple blurred images is an important topic for digital photography. If one assumes the blur kernel is shift-invariant, deblurring becomes a deconvolution problem:

$$B = I \otimes K + n, \tag{2}$$

where B, I, K and n denote the observed blurred image, the unknown latent sharp image, the blur kernel, and the noise respectively. In most cases, the deconvolution problem is underdetermined since the blur kernel, the latent image and the noise are all unknown. It is called blind deconvolution and good results usually can be obtained only for low-frequency blur kernels. Priors are often augmented to improve the results. Fergus *et al.* [4] used natural image statistics together with a sophisticated variational Bayes inference algorithm to estimate the kernel and the latent image. However, this approach is not robust enough and sometimes heavy human interactions must be involved [5]. Some techniques make the problem more tractable with additional inputs such as multiple images. Many deblurring methods using multiple images have been proposed [1, 7]. They seek to utilize the correlation among blurred images, based on the assumption that all blurred observations come from the same latent image. Promising results have been obtained

using multiple-image deblurring algorithms but they are often limited to blur due to simple directional motion. Yuan *et al.* [13] proposed an image deblurring approach using a pair of blurred/noisy images. It takes advantage of both images to produce a high-quality reconstructed image. By formulating the image deblurring problem using two images, they have developed an iterative deconvolution algorithm which can estimate a very good kernel and significantly reduce deconvolution artifacts.

3. Algorithm

The inputs to our algorithm is a set of LDR images Z_i with different exposure times Δt_i taken by a hand-held camera. Our goal is to recover the HDR scene irradiance image E from them. First, these LDR images are aligned by median threshold bitmap (MTB) algorithm [11]. However, here, we consider not only shifts but also rotations. Next, Debevec's algorithm [2] is used to combine the aligned LDR images to construct a blurred HDR image to be used as the initial guess to our algorithm. Our algorithm is based on a Bayesian framework and solves the problem by iteratively updating the HDR irradiance image, blur kernels and the camera response curve until convergence.

3.1. Image alignment

When taking a series of LDR images without a tripod, one takes a picture, adjusts the shutter speed, takes another one, and repeats these steps until a sufficient number of pictures have been taken. During picture taking and shutter speed adjustment, it is almost impossible to keep the camera still. Thus, there often exists shift or rotation movement between a pair of LDR images. Although, in principle, the shift movement can be modeled by the blur kernel convolution, it is beneficial to register the images first. This thereby reduces the kernel size, thus greatly speeding up the kernel estimation process. Without image registration, the kernel size must be set to be big enough to cover all shift movements between LDR images. Experiments indicates that the kernel needs to be more than 100 pixels wide. With alignment, kernel size can be reduced to 30 pixels.

We use the MTB alignment algorithm to register the differently exposed photographs [11]. This algorithm is not designed to register the blurred images originally, but it still works well in our experiments. In order to handle camera rotations, similar to estimating camera shift, we estimate the rotation angle by a coarse-to-fine multigrid search approach. Alternatively, Yuan *et al.*'s blurred/non-blurred image alignment algorithm could be used for such a task [12].

3.2. The Bayesian framework

We assume that each captured photograph Z_i comes from the same latent HDR irradiance image E. Each cap-

tured LDR image can be thought of as the output of the HDR irradiance going through a number of processes. The irradiance image is first convolved with the motion kernel and then scaled up by the exposure time. Finally, the blurred HDR irradiance image is transformed to the LDR one by applying the camera response function. The film reciprocity equation can be written as

$$Z_i = f\left((E \otimes K_i)\Delta t_i\right), \tag{3}$$

where Z_i, K_i and Δt_i denote the differently exposed LDR images, motion kernels and exposure times respectively, and $i = 1, 2, 3, \ldots, N$ ranging over all exposure durations. The response function of the camera is denoted by f and assumed to be monotonic. Thus, it is invertible and let g represent its inverse f^{-1}. We can then define B_i as follows.

$$B_i = g(Z_i)/\Delta t_i = E \otimes K_i. \tag{4}$$

The goal of our algorithm is to solve for the HDR irradiance image E, the motion kernels $K_1, K_2, \ldots, K_N$, and the camera response function f given the observed LDR images $Z_1, Z_2, \ldots, Z_N$. Our approach formulates the problem of computing these parameters in a well-defined Bayesian framework and solves it using the *maximum likelihood* (ML) technique.

We try to find the most likely estimates for E, f and K_i given the observation Z_i. We can express this as a maximization over a probability distribution P and then use Bayes's rule to express the result as the maximization over a product of probability distributions:

$$\arg \max_{E,f,K_i} P(E, K_1, K_2, \ldots, K_N, f | Z_1, Z_2, \ldots, Z_N)$$
$$= \arg \max_{E,f,K_i} P(Z_1, Z_2, \ldots, Z_N | E, K_1, K_2, \ldots, K_N, f)$$
$$P(E)P(K_1)P(K_2)\ldots P(K_N)P(f). \tag{5}$$

Since the dimension of the observations are often much larger than the dimension of parameters, the problem is essentially overdetermined. Thus, we do not need to include priors for E, f and K_i to constrain the solution. The problem is then turned into a maximum likelihood problem.

$$\arg \max_{E,f,K_i} L(Z_1, Z_2, \ldots, Z_N | E, K_1, K_2, \ldots, K_N, f)$$
$$= \arg \max_{E,f,K_i} - \sum_{i=1}^{N} \|E \otimes K_i - B_i\|^2. \tag{6}$$

Recall that $B_i = f^{-1}(Z_i)/\Delta t_i$ as defined in Equation 4. This leads to the minimization of the following objective function

$$O = \sum_{i=1}^{N} \|E \otimes K_i - B_i\|^2$$
$$= \sum_{i=1}^{N} \|\mathbf{E}\mathbf{K}_i - \mathbf{B}_i\|^2, \tag{7}$$

where $\mathbf{K}_i$ and $\mathbf{B}_i$ are the vector forms of K_i and B_i, and $\mathbf{E}$ is the matrix form of E [13]. In this energy function, the irradiance, the blur kernel of each observed image and the camera function are all unknown. We minimize the objective function with respect to a single variable at a time while keeping the other two fixed. In order to optimize the energy function O in Equation 7, we will iteratively estimate $\mathbf{K}_i$, $\mathbf{E}$, and g in $\mathbf{B}_i$ in turns.

Before starting the optimization, each parameter is initialized as the following.

Initialization for K_i***.*** We initialize $K_i = \delta$, the delta function, with all energy at the center of the kernel.

Initialization for E***.*** We composite the HDR image directly from the registered LDR images using Debevec's method [2], and initialize the matrix $\mathbf{E}$ accordingly.

Initialization for g***.*** We initialize g to be a linear mapping from the pixel values to the irradiance values.

3.3. Optimizing K_i

In this step, we fix E and g to optimize for K_i. By taking $\mathbf{E}$ and $\mathbf{B}_i$ as constants, the optimal $\mathbf{K}_i$ can be obtained by finding the least-square solution to the linear system $\mathbf{E}\mathbf{K}_i = \mathbf{B}_i$. However, this naïve approach may lead to poor results because the problem is often ill-posed. In order to obtain a better result, we use Tikhonov regularization to stabilize the solution by solving

$$\min_{\mathbf{K}_i} \sum_{i=1}^{N} \|\mathbf{E}\mathbf{K}_i - \mathbf{B}_i\|^2 + \lambda_1^2 \|\mathbf{K}_i\|^2,$$

and the Landweber method is used to iteratively update the $\mathbf{K}_i$ as follows [3].

1. Initialize $\mathbf{K}_i^0 = \delta$ if it is the first time for optimizing K_i; otherwise, initialize $\mathbf{K}_i^0$ as the optimum from the previous K_i optimizing step.

2. Update $\mathbf{K}_i^{n+1} = \mathbf{K}_i^n + \beta_1 \left(\mathbf{E}^T\mathbf{E} + \lambda_1^2\mathbf{I}\right)^{-1} \left(\mathbf{E}^T\mathbf{B}_i - \left(\mathbf{E}^T\mathbf{E} + \lambda_1^2\mathbf{I}\right)\mathbf{K}_i^n\right)$.

3. set $\mathbf{K}_{ij}^{n+1} = 0$ if $\mathbf{K}_{ij}^{n+1} < 0$ and normalize $\mathbf{K}_{ij}^{n+1} = \mathbf{K}_{ij}^{n+1} / \sum_j \mathbf{K}_{ij}^{n+1}$, where $\mathbf{K}_{ij}^{n+1}$ is the j-th element of the vector $\mathbf{K}_i^{n+1}$.

The iteration stops when the change between two steps is sufficiently small. We typically run about 3 to 10 iterations by setting $\lambda_1 = 5$ and $\beta_1 = 0.8$. In order to reduce the cost of memory and reduce the running time, we applied the above method in the Fourier domain. Note that, after the first optimization for kernels, we can update the blur kernel of each input image. We will take these K_i as initialization in next iteration of optimizing K_i as described in the step 1 of the above procedure.

3.4. Optimizing E

In this step, K_i and g are fixed to optimize for E. The objective function in Equation 7 is transformed into

$$O_2 = \sum_{i=1}^{N} \|E \otimes K_i - B_i\|^2$$
$$= \sum_{i=1}^{N} \left\|\widetilde{\mathbf{K}}_i\widetilde{\mathbf{E}} - \mathbf{B}_i\right\|^2, \tag{8}$$

where $\widetilde{\mathbf{E}}$ and $\mathbf{B}_i$ are the vector forms of E and B_i, and $\widetilde{\mathbf{K}}_i$ is the matrix form of K_i. Comparing to the step of optimizing K_i, the roles of E and K_i are exchanged. Tikhonov regularization method is again used to stabilize the solution by solving

$$\min_{\widetilde{\mathbf{E}}} \left\|\widetilde{\mathbf{K}}\widetilde{\mathbf{E}} - \mathbf{B}\right\|^2 + \lambda_2^2 \left\|\widetilde{\mathbf{E}}\right\|^2,$$

where $\widetilde{\mathbf{K}}$ is the matrix formed by concatenating $\widetilde{\mathbf{K}}_1, \widetilde{\mathbf{K}}_2, \ldots, \widetilde{\mathbf{K}}_N$ and $\mathbf{B}$ is the vector formed by concatenating $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_N$.

We again adopt the Landweber method to find the optimum by iteratively updating $\widetilde{\mathbf{E}}$ as follows.

1. Initialize $\widetilde{\mathbf{E}}$ from the previous E optimizing step; If it is the first time, initialize E as the HDR image by directly compositing the input LDR images.

2. Update $\widetilde{\mathbf{E}}^{n+1} = \widetilde{\mathbf{E}}^n + \beta_2 \left(\widetilde{\mathbf{K}}^T\widetilde{\mathbf{K}} + \lambda_2^2\mathbf{I}\right)^{-1} \left(\widetilde{\mathbf{K}}^T\mathbf{B}_i - \left(\widetilde{\mathbf{K}}^T\widetilde{\mathbf{K}} + \lambda_2^2\mathbf{I}\right)\widetilde{\mathbf{E}}^n\right)$.

3. set $\widetilde{\mathbf{E}}_j^{n+1} = 0$ if $\widetilde{\mathbf{E}}_j^{n+1} < 0$, where $\widetilde{\mathbf{E}}_j^{n+1}$ is the j-th element of the vector $\widetilde{\mathbf{E}}^{n+1}$.

In our experiments, setting λ_2 as 3 or 4 often leads to good results. We set β_2 as 0.8 and run about 6 to 15 iterations. The optimization is also performed in the Fourier domain to reduce the memory and computation cost. After optimization, E can be set according to the optimum $\widetilde{\mathbf{E}}$.

3.5. Optimizing g

After optimizing K_i and E, we fix K_i and E to optimize the inverse camera function g. We rewrite Equation 7 and add a smoothness term to ensure the function g is smooth:

$$O_3 = \sum_{i=1}^{N} \sum_{j=1}^{P} \|(E_j \otimes K_i)\Delta t_i - g(Z_{ij})\|^2$$
$$+ \lambda_3 \sum_{z=Z_{min}+1}^{Z_{max}-1} g''(z)^2, \tag{9}$$

where P is the number of pixels in an image; Z_{ij} is the pixel value of the j-th pixel in the i-th image; $Z_{max} = 255$ and $Z_{min} = 0$. Recovering g only requires finding the finite number of values that $g(z)$ can take. Since z represents a pixel brightness value, it can only take the 256 discrete values within [0..255]. By setting $g''(z) = g(z-1) - 2g(z) + g(z+1)$, Equation 9 leads to a quadratic function that can be optimized as a linear least square problem [2]. The scalar λ_3 weights the smoothness term relative to the data fitting term, and should be chosen appropriately for compensating the amount of noise in the Z_{ij} measurements. As suggested by Debevec [2], we use the following hat weighting function to downweight the unreliable observations:

$$w(z) = \begin{cases} z - Z_{min}, & \text{for } z \leq \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z, & \text{for } z > \frac{1}{2}(Z_{min} + Z_{max}) \end{cases} \quad (10)$$

Instead of using all pixels, we only need to use enough pixels to optimize Equation 9 as long as the resulting linear system is sufficiently overdetermined. We prefer to select pixels from flat regions to avoid problems with misalignment. We take the mid-exposed photograph as the reference image and apply edge detection. We randomly pick samples from the regions which are not on edges. After the process of inverse camera curve optimization, we can update the inverse camera function g and thereby B_i.

These three steps complete one iteration, and the process repeats by updating kernel K_i of each image (Section 3.3), the irradiance image E (Section 3.4) and inverse camera function g (Section 3.5) until convergence. We typically run 10 to 15 iterations.

3.6. Implementation details

Irradiance scaling. Because of the regularization term in optimizing $\|\mathbf{E}\|^2$, the optimal irradiance values E are always scaled down by a factor. Thus, after optimizing E in each iteration, we scale up the reconstructed irradiance E to keep its values in a similar scale as the initialized irradiance $E_{initial}$. The scaling factor is determined by the ratio of their mean values, $mean(E_{initial})/mean(E)$. We determine the scale factor for each color channel separately.

Color handling. We make the assumption that all three color channels have the same blur kernels. Thus, in our implementation for optimizing K_i, we only consider the green channel. The same set of kernels K_i are used for optimizing E and g for all three color channels separately. However, optimizing E and g separately for three color channels sometimes causes a scale difference between color channels. In order to produce a color-balanced high dynamic range image, we need to relate color channels after the algorithm converges. We assume that the camera response functions of three channels are very similar, and find the scale factors for red channel and blue channel by taking the inverse camera function g of green channel as reference.

4. Results and comparisons

To demonstrate the effectiveness of our algorithm, we tested our method on both synthetic examples and real photographs. For comparisons, since there is no other algorithm solving the same problem this paper addresses, we compared our algorithm to the naïve applications of two image deblurring algorithms, Fergus' algorithm [4] and Yuan's algorithm [13]. For Fergus' algorithm, each LDR image is deblurred individually, and then all deblurred images are composited together to reconstruct the HDR image using Debevec's algorithm. Yuan's algorithm needs a noisy/blurred image pair for deblurring. We use the LDR image with the shortest exposure time as the noisy one, and others as the blurred images to obtain the deblurred images. After deblurring, the HDR image is also reconstructed using Debevec's algorithm. The implementation of Fergus' algorithm is directly taken from the authors, and Yuan's algorithm is implemented by ourselves. We also tried to use Shan *et al.*'s single-image deblurring algorithm [10], but found it frequently failed when dealing with images of our resolution, $2,000 \times 1,500$.

We first create a synthetic example so that we have the ground truth to compare with. We created five virtual LDR images synthesized from a HDR image E by setting $Z_i = Q(f((E \otimes K_i)\Delta t_i))$, where Q is a quantization function that quantizes values to 256 discrete levels. In this experiment, Δt_i's are set to $\frac{1}{512}, \frac{1}{256}, \frac{1}{128}, \frac{1}{64}$ and $\frac{1}{32}$ to generate five differently exposed LDR images. The camera function f is adopted from a real camera's (Canon G9) response curve, and K_i's are taken from previous deblurring work.

Figure 2 shows the results for the experiment with synthetic data. Figure 2(a) is the tone-mapped image of the ground truth HDR image taken from Debevec's website. Figures 2(b-f) show the LDR images synthesized using the above configuration and their kernels. Figure 2(g) is the initial guess for the irradiance image generated by directly applying Debevec's algorithm. Figures 2(h) and (i) are respectively the estimated blur kernels and the camera response curve outputted by our algorithm. In Figure 2(i), the sky blue curve stands for the ground truth camera response curve. The red, green and blue curves are the recovered response curves of red, green and blue channels, respectively. Figure 2(j), (k) and (l) are results of our algorithm, Fergus' algorithm and Yuan's algorithm. Our result clearly exhibits sharper image details than other two algorithms and more resembles the ground truth.

Note that those two algorithms are not designed for HDR reconstruction. Thus, it is not surprising that both have worse results than ours since our algorithm takes advantage of solving HDR images from all observations together. However, since our algorithm is the first of its kind, these two algorithms are probably the closest ones we can compare to. The comparison at least shows that the naïve ex-

(a) Ground-truth

(b) Z_1 with $\Delta t = 1/512$

(c) Z_2 with $\Delta t = 1/256$

(d) Z_3 with $\Delta t = 1/128$

(e) Z_4 with $\Delta t = 1/64$

(f) Z_5 with $\Delta t = 1/32$

(g) $E_{initial}$

(h) Estimated blur kernels

(i) Response curves

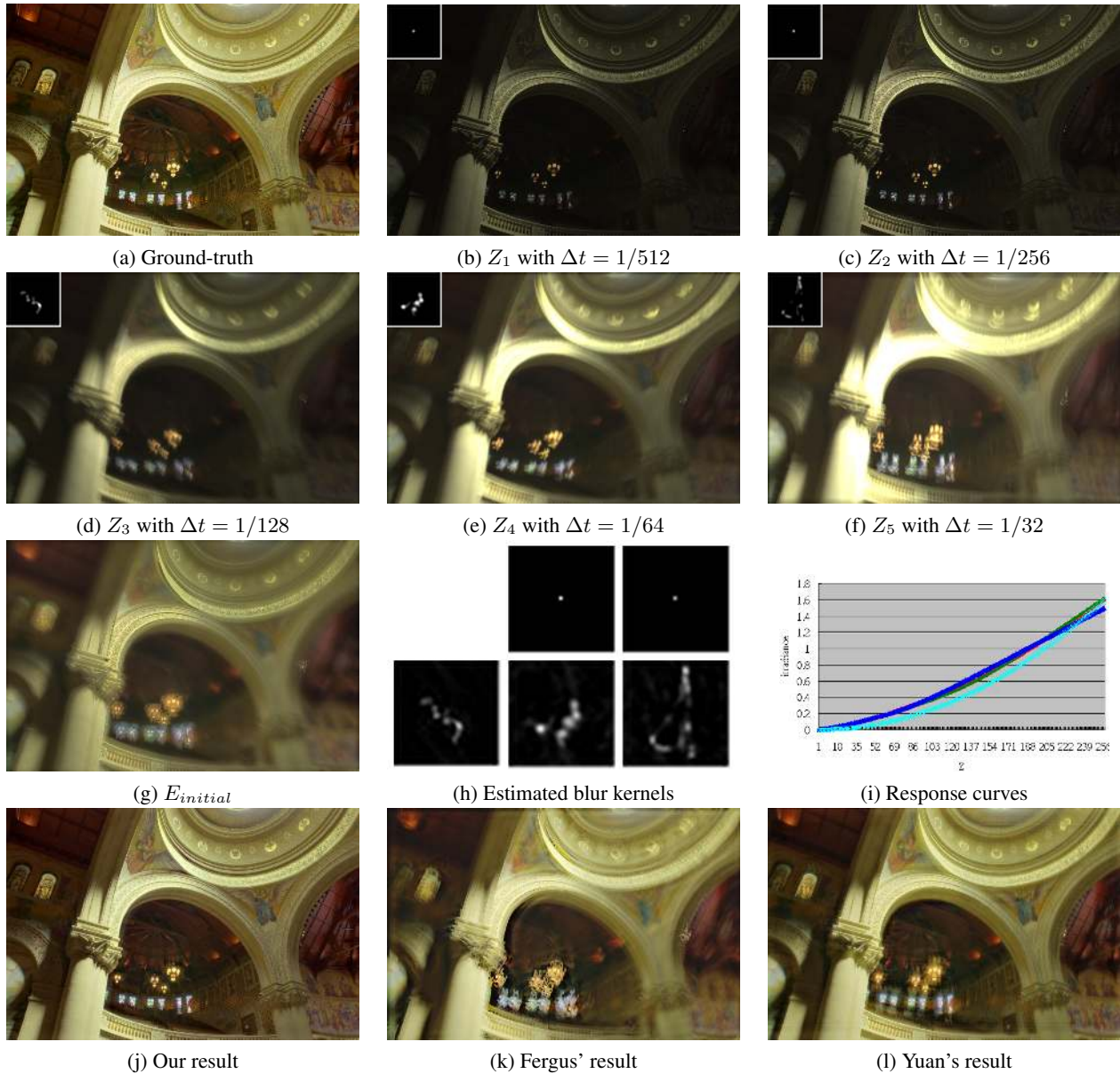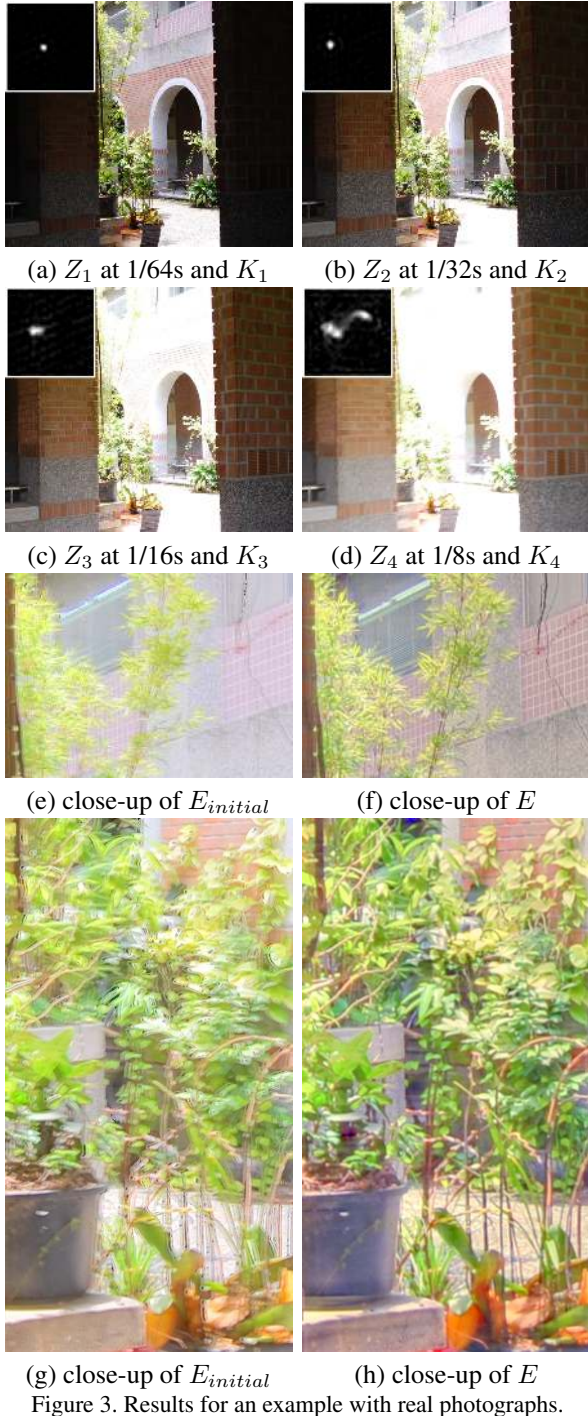(j) Our result

(k) Fergus' result

(l) Yuan's result

Figure 2. Results and comparisons for a synthetic example.

tensions of previous image deblurring algorithms will not lead to satisfactory results and solving the problem is not a trivial task.

Figure 3 shows the application of our algorithm on a set of differently exposed real photographs taken by a hand-held camera. The image size is $2,000 \times 1,500$ and the kernel size is 31×31. Figure 3(a-d) shows four captured images and the blur kernels estimated by our algorithm. Note that, for this example, we did not perform image alignment. Thus, the kernels are not necessarily centered. Figures 3(e) and (g) show close-ups of the reconstructed HDR image using Debevec's algorithm. Figures 3(f) and (h) show close-ups of our results. Clearly, our result is shaper.

Figure 4 shows comparisons on two sets of real photographs. Figure 4(a) shows the real photos with shutter speeds of $\frac{1}{256}$, $\frac{1}{128}$, $\frac{1}{64}$, $\frac{1}{32}$ and $\frac{1}{16}$ seconds. Figure 4(b) shows the images taken with the longest exposure and their close-ups. Clearly, they are blurred. Figure 4(c), (d), (e) are the results and the close-ups of Fergus' algorithm and Yuan's algorithm and our algorithm respectively. Our algorithm has better results than the other two. Note that Yuan's method sometimes shows more high frequencies than our method, particularly for the leftmost brick wall in the close-up. However, at the same time, it also introduces high-frequency artifacts. For example, when looking closely at the brick wall in the middle, Yuan's method shows rip-

(a) Z_1 at 1/64s and K_1

(b) Z_2 at 1/32s and K_2

(c) Z_3 at 1/16s and K_3

(d) Z_4 at 1/8s and K_4

(e) close-up of $E_{initial}$

(f) close-up of E

(g) close-up of $E_{initial}$

(h) close-up of E

Figure 3. Results for an example with real photographs.

ples instead of vertical stripes that the wall should have. Such high-frequency artifacts also appear in the stone texture above the brick wall. It appears rougher than it should be. Stochasticity of the stone texture helps hide the artifacts. Such artifacts are more obvious in the black lamp cover and the left scene of Figure 4.

5. Conclusions and future work

This paper proposed a unified formulation for recovering radiance image, blur kernels and camera response curve. With this, we propose a technique for reconstructing a non-blurred HDR image from a set of differently exposed and blurred images taken with a hand-held camera. Bayesian framework is used to define the probabilistic model. A maximum-likelihood approach is applied to find the best solutions for the blur kernels, the HDR irradiance image and the camera response curve through an EM-like optimization algorithm. We found that the proposed technique successfully reconstructs HDR images from LDR images taken with hand-held cameras in most cases.

In the future, we plan to augment priors for kernel $p(K_i)$, for the reconstructed HDR image $p(E)$ and for the camera curve $p(f)$. By adding these priors, it could improve the results. In addition, our method could be more efficient and effective by taking account that short exposed images are less blurred. Finally, our approach shares the same limitation with most image deblurring techniques: the blur kernel must be shift-invariant. When the blurred images taken from hand-held cameras are affected by blur that is not shift-invariant, *e.g.* from slight camera rotation or non-uniform object motion, our approach will have problems. Image deblurring can be further improved if we replace the shift-variant assumption with a more general kernel assumption.

Acknowledgments

References

[1] B. Bascle, A. Blake, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *Proceedings of ECCV 1996*, pages 573–582, 1996.

[2] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proceedings of ACM SIGGRAPH 1997*, pages 369–378, 1997.

[3] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic, 2000.

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

[5] J. Jia. Single image motion deblurring using transparency. In *Proceedings of CVPR 2007*, 2007.

[6] T. Mitsunaga and S. Nayar. Radiometric self calibration. volume 1, pages 374–380, 1999.

[7] A. Rav-Acha and S. Peleg. Two motion-blurred images are better than one. *Pattern Recogn. Lett.*, 26(3):311–317, 2005.

[8] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Morgan Kaufmann, 2005.

Figure 4. Results for two sets of real photographs. (a) Input images. (b) The input images with the longest exposure and their close-ups. (c) Fergus' results. (d) Yuan's results. (e) Our results.

[9] M. A. Robertson, S. Borman, and R. L. Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12(2):219–228, Apr. 2003.

[10] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *SIGGRAPH*, 2008.

[11] G. Ward. Fast, robust image registration for compositing high dynamic range photographs from handheld exposures.

Journal of Graphics Tools, 8(2):17–30, 2003.

[12] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Blurred/non-blurred image alignment using sparseness prior. In *ICCV*, Rio de Janeiro, Brazil, 2007.

[13] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. 26(3):Article 1, 2007.