

BIL 717 Image Processing-Spring 2016

Final Project

Analysis and Evaluation of Two Non-Uniform Motion Deblurring Studies

Mustafa Turan
Hacettepe University
mstftrn@gmail.com

Abstract

In this paper we evaluate two blind motion-deblurring methods in the recent literature using a non-traditional metric that was developed in order to create a more perceptually sound deblurring comparison criterion. In general terms, the first method, which was developed by Sun et al. (2015) [8], uses a convolutional neural network (CNN) to learn non-uniform motion blur removal. The second method, which was developed by Whyte et al. (2012) [10], proposes a geometric model to define the motion of the camera in 3D. Both methods tackle with the same blind non-uniform motion blur removal problem. In the upcoming sections we first define the blind motion-deblurring problem, then we discuss the related work focusing mostly on the recent literature, next we describe the two methods, subsequently we define the methodology and experimentation along with the no-reference metric, and we finish with conclusion.

1. Introduction

Motion blur is caused by moving the camera or objects within view when the camera shutter is open. This harms sharpness and causes losing the edges and thus objects in the scene seem intermingled. If the entire scene was blurred in the same way, this type of blur is called uniform blur. Non-uniform blur arises when the blur does not show the same type or magnitude of intermingling throughout the scene. A typical example of non-uniform blur emerges when the camera is rotated when the camera shutter is open and the recording of the information of the outside world is under way. When this happens the scene gets blurred in a rotational pattern (different parts of the scene blurred depending on the distance to a certain rotational center). Another example would be moving the camera towards or

away from the scene (transposition in the depth axis). In this case the blurring pattern looks more like a tunnel effect; namely blurring happens less in magnitude around a certain center in the scene (like a target), and more and more towards the image boundaries. More complex movements such as a combination of transposition and rotation of the camera complicate the issue even further.

Uniform deblurring typically have been defined as convolution of an image with a kernel and added noise. Therefore, the basic approach is to subtract the noise and deconvolve the image with the same kernel. The problem with this approach is that the kernel is generally unknown. When the blur kernel is unknown, the problem is called blind deblurring problem. Researchers have attacked this problem with a range of different approaches. Some of the approaches will be covered in the related work section. In this paper we analyze two of the recent papers that tackle with blind motion-deblurring problem.

1.1. Sun et al. (2015)

In Sun et al. (2015), the authors attack the problem of removing non-uniform motion blur from a single image using deep learning. The method they use depends on predicting motion kernels at a patch level. In order to do this they formulate the non-uniform blur as a field throughout the image. To find this field they divide the image into overlapping 30×30 patches, estimate the kernels at this level and then smooth the field depending on a smoothness of motion assumption. CNN is used to learn deblurring at a patch level. The general view of the CNN is seen on figure 1 [8].

Before training, the authors generated motion blur kernels with varying sizes from 1 pixel to 25 pixels with interval of two, and orientations ranging from 0 degree to 150 degree with intervals of 30, totaling 73 kernels. These kernels can be seen in figure 2. To train the CNN model, the authors used these 73 kernels to artificially generate 1.4 million blurred patch/kernel pairs using PASCAL VOC 2010

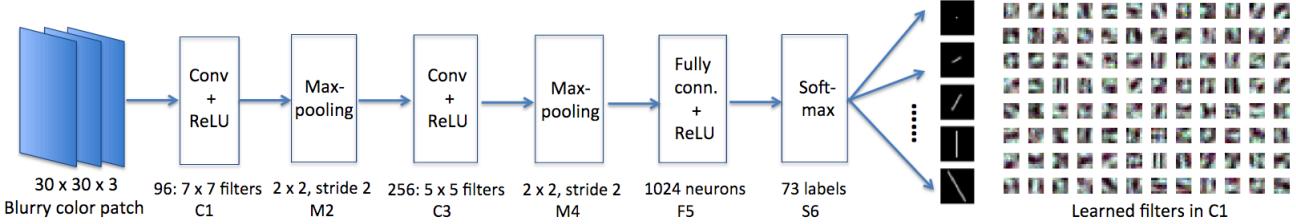


Figure 1: Convolutional neural network

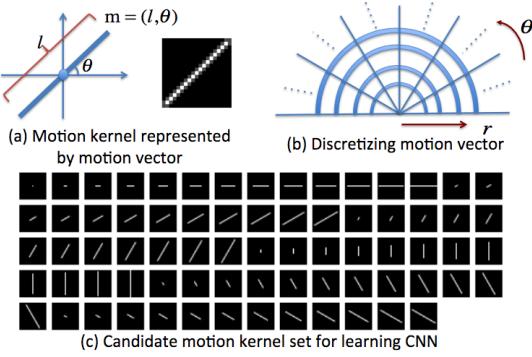


Figure 2: Convolutional neural network training kernels

database. They then used these patches as the input to the CNN during the training phase [8].

As can be seen in figure 1 the CNN finds a probability distribution at the softmax layer. The authors state that the kernel space, which consists of 73 kernels, is not dense enough to represent all types of motions. To overcome this problem, they extend the motion kernel set by rotating image patches in the range between 0 and 24 degrees and use the trained CNN on these patches. Note that at this point they do not retrain the CNN, but rather they use the trained CNN with rotated images. By doing this they overcome the trainability problem of a high number of motion kernels and get a good amount of motion orientation detail [8].

The next phase is using the Markov Random Field (MRF) to find a dense field of kernels on the image. The previous phases find a number of motion kernels at every patch location on the image. Here, the main assumption is that the camera moves in a smooth trajectory when the shutter is open and thus the change of kernels throughout the image must also be smooth. This implies that there must not be sudden changes when moving from one patch to another. This is made possible by using an MRF model and optimizing it. This enforces closeness of nearby kernels [8].

1.2. Whyte *et al.* (2012)

In Whyte *et al.* (2012), the authors emphasize that during the exposure, the camera sees a sequence of interrelated views and integrate them to form the blurry image. If we take one sensor pixel in the camera into consideration, it is subjected to photons coming from different points in space when the camera moves. It is also possible that nearby pixels are seeing the same points with passing time and recording the light coming from the same points. Therefore, they make an observation that the views of the camera are all related and they state that this relation can be explained using geometry [10].

The authors create a geometrical model using the geometry of a camera. They formulate how the translation and rotational movements of the camera can be expressed in terms of homographies, or in other words “projective transformations in 2D” [10]. They fix the origin of the coordinate axes as the camera’s optical center. They define the blurry image as an integral of the images during the exposure time. They then go on to express this integral in terms of angles instead of time [10].

In order to find these homographies, they need to find the so-called calibration matrix of the camera. This depends on the physical properties of the camera. The study uses EXIF tags of an image to populate these variables depending on the brand and model of the camera. When there is no EXIF tag available, the user might manually set these variables or these variables might be guessed depending on the other properties of the image [10].

There is not a single best algorithm the paper proposes but instead they plug the geometrical model into the well-established algorithms in the recent history. We use the Cho and Lee (2009) algorithm that depends on the image derivatives and Gaussian priors [1]. This is because, this code is readily available and expected to run faster than the alternatives.

2. Related Work

Non-uniform blur removal problem has been studied extensively. Many pieces of previous research has considered

the problem as a subset of uniform blur in a patch scale. In other words, these studies assumed that images that have non-uniform blur are much like uniformly blurred images in a smaller scale. Levin (2006) handled the problem of non-uniform blurring when some of the objects in the scene are moving independently. They divide the image into segments and find kernels using image statistics [5]. Cho *et al.* (2007) handles the problem of blind deblurring when the objects and the camera moving at the same time. They formulate the problem using an energy function and solve it [2].

Some other methods that did not handle blind deblurring as a uniform deblurring in a segment scale have been proposed. Shan *et al.* (2008) estimates blur kernel and deblurred image at the same time using a probabilistic model. They analyze the common artifacts of deblurring and they constrain some of the image features in order to get rid of these artifacts [7]. Tai *et al.* (2010) reduces the “spatially varying” blur using a hybrid camera system. This camera system has an extra camera that captures at a higher frame rate but at a lower resolution. They combine the two streams in order to reduce the non-uniform blur in video streams and in images [9]. Similarly, Joshi *et al.* (2010) uses inertial measurement sensors to measure six degrees of freedom motion to approximate the blur kernels [4]. Gupta *et al.* (2010) approximates the six degrees of freedom motion using “in-plane translation and rotation” similar to Whyte *et al.* (2012). They represent camera motion as a Motion Density Function. The algorithm in this paper first come up with a kernel and successively updates the image and the kernel in each step [3].

3. Methodology/Approach

One of the reasons why we decided on to analyze these two studies was the availability of the codes and data related to the studies. We have got in contact with some other researchers about their papers and queried whether the codes and related material were available, but we have not been able to get the sufficient material to conduct the experiments.

Sun *et al.* (2015) code is available Dr. Sun’s website ¹. Whyte *et al.* (2012) code is available at the study’s website ².

In conducting the experiments, we used five blurry images that was provided with [10]. These blurry images are available on the study’s website ³. These pictures pose very challenging deblurring problems. The movements and the other features of the images show quite complex patterns and the amount of intermingling among nearby pixels is

¹<http://gr.xjtu.edu.cn/>

²<http://www.di.ens.fr/willow/research/deblurring/>

³<http://gr.xjtu.edu.cn/web/jiansun/codes>

quite high.

We simply ran the the two pieces of code that were provided with the two studies we study and measured the quality of deblurring using the no-reference metric we are going to cover and a more traditional metric, PSNR.

In this study, we use Liu *et al.* (2013)’s no-reference metric to measure the success of the two deblurring methods along with PSNR. We will describe the no-reference metric briefly in the next section.

3.1. No-reference metric of Liu *et al.* (2013)

The main goal this study tries to achieve is finding a good metric to measure specifically quality of deblurring. The authors argue that a metric specific to the problem of motion deblurring will yield better results than more general metrics that measure the quality of the solutions to other types of image processing problems such as denoising and so on. To this end, they measure the principal artifacts related to deblurring in general, i.e. ringing artifacts, noise, and residual blur and use these as features to learn a metric for blind deblurring [6].

The researchers used crowdsourcing to perceptually assign deblurring quality to deblurred images. In doing this, the users were shown image pairs and they decided which has a better deblurring quality. The researchers used this quality information to rank different images that were deblurred using different algorithms and parameters. They decided on which artifacts are most important in the process of deciding on the quality of deblurring and which features are more important to use when assigning a score to a deblurring result [6].

4. Experimental Results

In this section we share some of the results of our experiments. The main limitation is the number of images that were used in the evaluation. We think that more images with various types of blur would give a better understanding when evaluating these studies. One of the main constraining factors was the time length of the experiments. Each deblurring task took around thousand seconds and this made the experimentation slower and harder. We could not conduct every single experiment we planned in the first place. In the upcoming sections we discuss the visual and quantitative results of the experiments.

4.1. Visual Results

In figure 3 blurry image and the deblurring results of Sun *et al.* (2015) and Whyte *et al.* (2012) can be seen. Obviously, Whyte *et al.* (2012) approach gives a better result although some ringing artifacts are present.

The blur kernels associated with these two deblurring results are shown in figure 4. Note that Sun *et al.* (2015)

NOREF	Sun <i>et al.</i> (2015)	Whyte <i>et al.</i> (2012)
istanbul	-11.5762	-11.3754
garden	-10.3892	-12.3826
living	-12.1000	-11.0210
notredame	-7.5634	-8.8334
book	-11.7937	-10.0433

Table 1: No-reference metric measurements of the deblurring results of the two studies

PSNR	Sun <i>et al.</i> (2015)	Whyte <i>et al.</i> (2012)
istanbul	27.5377	23.8787
garden	33.4010	14.7554
living	40.1380	29.4494
notredame	36.3057	23.9578
book	42.7525	33.2875

Table 2: PSNR metric measurements of the deblurring results of the two studies

approach could not find a good approximation to a set of kernels that depict smooth motion.

In figure 7 the ringing effects are much more obvious in Whyte result. On the contrary, Sun result does not obviously have ringing.

4.2. Quantitative Results

We measured the results of the experiments with two metrics. As we discussed earlier, the first metric we used is Liu *et al.* (2013)’s no-reference deblurring metric. The other metric is peak signal-to-noise ratio. In table 1 and table 2 we see a comparison of the two studies and their deblurring result scores for our five-image dataset. The correlation of the scores of two studies across no-reference metric and PSNR are positive and around 0.56 and 0.66, respectively. However, the correlation of the two metrics across studies are negative for Sun *et al.* (2015) (around -0.13), and positive for Whyte *et al.* (2012) (around 0.53). This is an interesting result and may be a sign that PSNR and no-reference metric are not behaving in the same way. We think that the PSNR metric might behave the opposite way because it is not meant to look at the same artifacts with the more deblurring-specific metric.

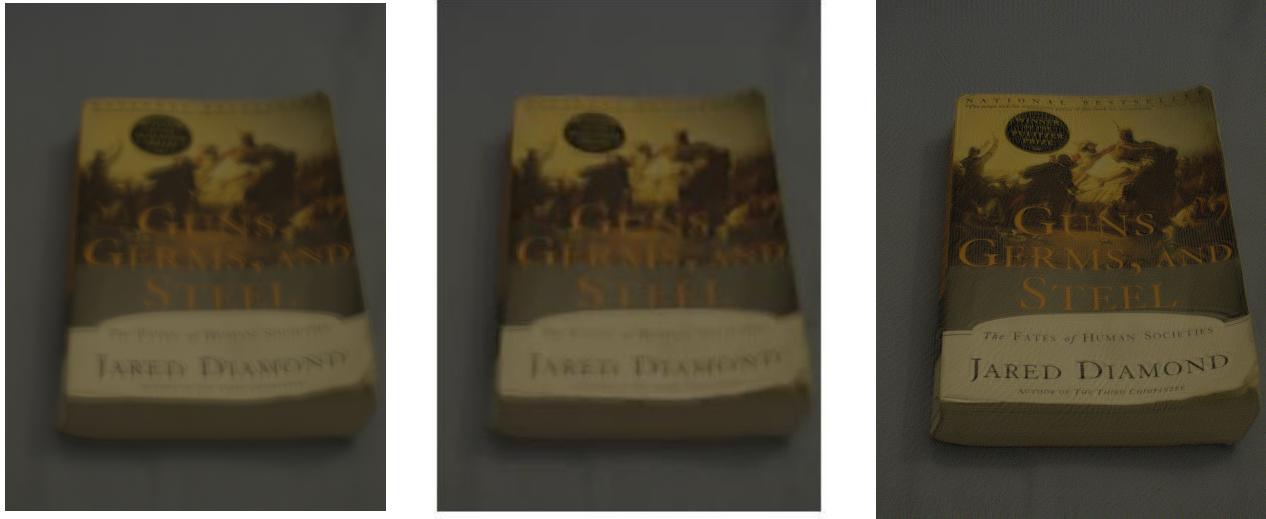
5. Conclusions

In this study, we analyzed and evaluated two of the deblurring studies in the current literature. We evaluated the results using a recently developed metric that was designed to measure motion deblurring specifically, along with more traditional PSNR. The results show that neither of the deblurring techniques is the panacea of the blind motion deblurring problem. Both techniques are good for certain

types of blurry pictures. Whyte *et al.* (2012) is no good for artifacts that arise after deblurring. Sun *et al.* (2015) cannot handle the problems with more complex motion types. Another interesting observation is that although visually more pleasing, some of the resulting images got low no-reference metric scores because of the ringing effects. Lastly there is no notable correlation between PSNR and no-reference metrics. The two metrics had a negative correlation for one metric and positive correlation across the two studies’ results. As a future work, parameters of the Whyte *et al.* (2012) can be changed and the algorithms can be evaluated more thoroughly.

References

- [1] S. Cho and S. Lee. Fast motion deblurring. In *ACM Transactions on Graphics (TOG)*, volume 28, page 145. ACM, 2009.
- [2] S. Cho, Y. Matsushita, and S. Lee. Removing non-uniform motion blur from images. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.
- [3] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In *Computer Vision–ECCV 2010*, pages 171–184. Springer, 2010.
- [4] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. In *ACM Transactions on Graphics (TOG)*, volume 29, page 30. ACM, 2010.
- [5] A. Levin. Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems*, pages 841–848, 2006.
- [6] Y. Liu, J. Wang, S. Cho, A. Finkelstein, and S. Rusinkiewicz. A no-reference metric for evaluating the quality of motion deblurring. *ACM Trans. Graph.*, 32(6):175–1, 2013.
- [7] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *ACM Transactions on Graphics (TOG)*, volume 27, page 73. ACM, 2008.
- [8] J. Sun, W. Cao, Z. Xu, and J. Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 769–777. IEEE, 2015.
- [9] Y.-W. Tai, H. Du, M. S. Brown, and S. Lin. 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, 2010.
- [10] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *International journal of computer vision*, 98(2):168–186, 2012.

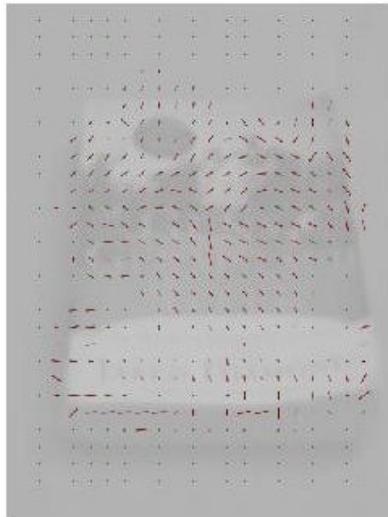


(a) Blurry book image

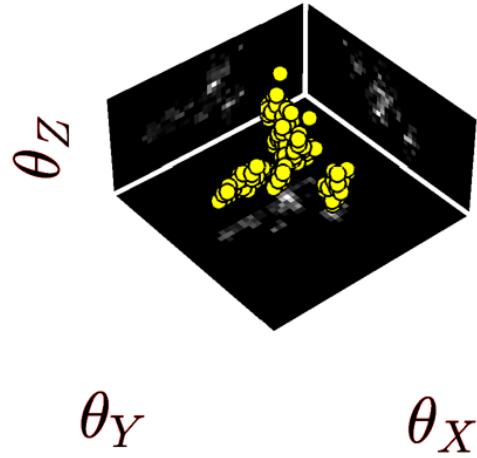
(b) Sun *et al.* (2015) result

(c) Whyte *et al.* (2012) result

Figure 3: Blurry book image and deblurring results of the two algorithms



(a) Sun *et al.* (2015) kernels



(b) Whyte *et al.* (2012) 3D kernel

Figure 4: Blur kernels associated with the two algorithms. Note that the Whyte *et al.* (2012) kernel is a 3D kernel and the Sun *et al.* (2015) set of kernels are 2D ones. This is the main difference Whyte *et al.* (2012) differs from the rest of the algorithms. Whyte *et al.* (2012)'s approach seems to be a better fit where the movement is more complex and cannot be expressed with simple movements as in Sun *et al.* (2015).



(a) Blurry image from Istanbul (b) Sun *et al.* (2015) result (c) Sun *et al.* (2015) kernels (d) Whyte *et al.* (2012) result

Figure 5: Blurry image from Istanbul and deblurring results of the two algorithms. Whyte *et al.* (2012) has a lot of ringing effects because of saturation. Ringing is not so obvious for Sun *et al.* (2015).

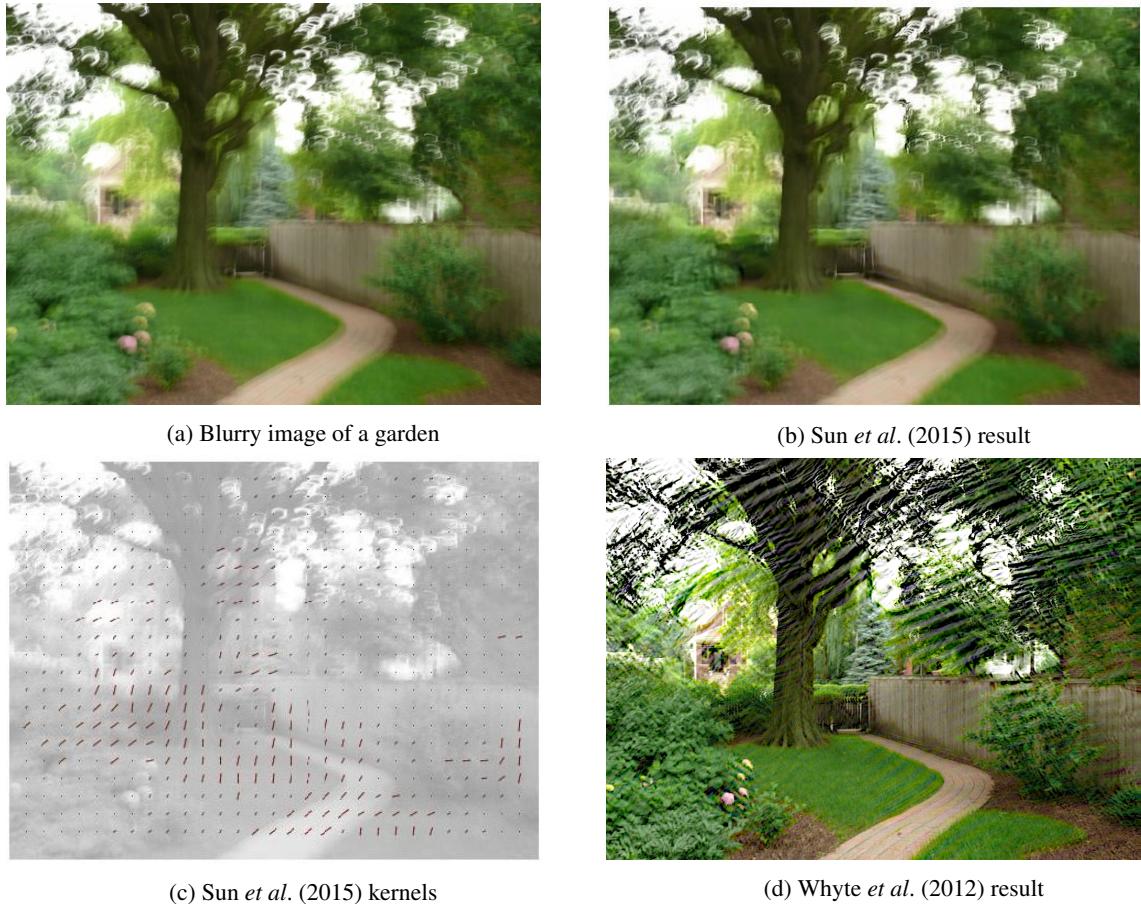
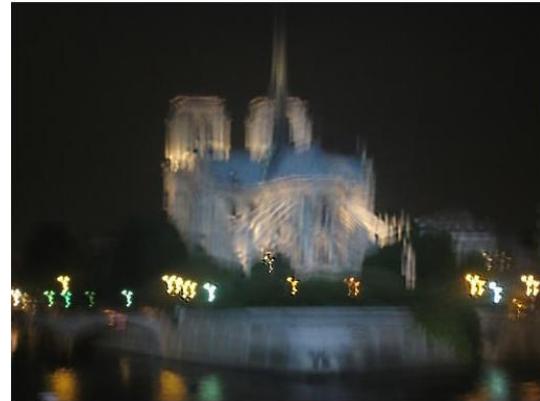


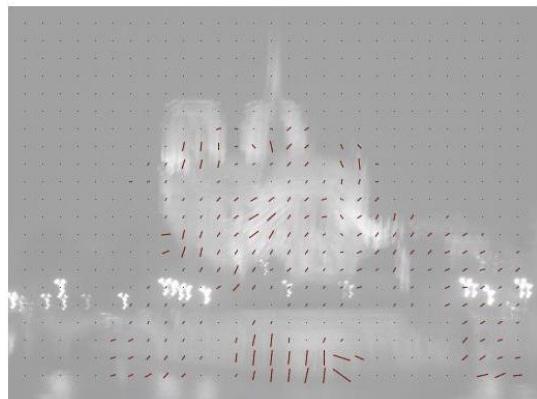
Figure 6: Blurry image of a garden and deblurring results of the two algorithms. Whyte seems to be much more successful but might get a lower score in no-reference metric because of the ringing effects.



(a) Blurry image from Notre Dame



(b) Sun *et al.* (2015) result



(c) Sun *et al.* (2015) kernels



(d) Whyte *et al.* (2012) result

Figure 7: Blurry image from Notre Dame and deblurring results of the two algorithms. Sun *et al.* (2015) finds some kernels but the smoothness is not what a normal camera motion is expected to produce.