A Thesis for the Degree of Ph.D. in Engineering

Pre/Post Blur Correction from a Single Photo Shooting

August 2011

Graduate School of Science and Technology

Keio University

Yuji Oyamada

Abstract

Recently, projector based display technologies have attracted much attention. In contrast to computer monitors, projectors have a big merit that projectors can fit their projection to arbitrary shape display. In addition to the development of computer power, miniaturization and price reduction of cameras and projectors encourages researches on such technologies. Such systems implicitly assume that we can display observation as high quality as we expect. However, this assumption is sometimes violated by motion/defocus blur occurred during imaging process.

This thesis focuses on image restoration problem aiming to remove blur effect. I propose single shoot based blur estimation method using constraints on target blur and scene.

To remove motion blur effect occurred during image acquisition process, it is required to know how the image is blurred. Proposed method analytically estimates this blur information by using target blur as constraints. The method takes a single blurred image captured by a normal camera and the blur on the captured image is assumed to be uniform on the entire image. Under such conditions, the cepstrum of the blurred image has partial information of target blur. Using the characteristics, I propose a PSF estimation method estimating a PSF from the cepstrum of a blurred image. Since the method uses constraints on target blur, the method is applicable to various types of images.

In-focus projection removes the projector defocus occurred during image display process. To realize it, we have to know the amount of defocus blur on the display. Furthermore, the amount of defocus varies across the display. Addition to an image captured by a camera, an image projected by a projector is also used as input. Proposed method estimates spatially varying defocus blur information

using the projector-camera image pair. Since the method does not require fiducial pattern projection, the method is available on on-line systems.

To validate the proposed methods, both synthesized images and real world images are used. Experimental results show that the proposed methods enable blur correction under the assumed conditions.

Acknowledgements

During my Ph.D. program, I have met many people those who influenced me in any sense. Without them, I could not have finished my Ph.D. work.

First of all, I would like to thank my thesis advisor, Prof. Hideo Saito, for the liberty he gave me and his encouragement and strong support on my way.

I express my gratitude to the thesis committee members: Prof. Ken-ichi Okada, Prof. Issei Fujishiro, and Prof. Masaaki Ikehara for accepting to evaluate this work. Their insightful comments and suggestions contributed to make this thesis quality much much better.

I would like to sincerely thank Dr. Yasuyuki Matsushita, my mentor during my internship at Microsoft Research Asia. Even though he told me a mentor and an intern are on equal relationship, he had collaborated with me as both a coworker and a supervisor. His comments, suggestions, and advice I had received and discussion and conversation with him took me higher than where I was before the internship. I also wish to thank Dr. John Wright for interesting and exciting discussions we had on our research. Special thanks go to my friends at Microsoft Research Asia, Peng He, Tomoaki Higo, Boxin Shi, Dirk Schnieders, and Takafumi Taketomi for giving me good suggestions and enjoying life in Beijing.

I don't know how to express my appreciation to three key persons in our laboratory. Dr. Vincent Nozick and Edmée Amstutz, who had already left our laboratory, opened my eyes to the world. Since my appearance is very identifiable at least in our laboratory and I knew so many things about the laboratory, they got used to speak to me. Even though my English skill was really awful, they were very very patient with me. Vincent told me how to enjoy research with his excellent sense of humor. Conversation with Edmée on both daily life and research work was good training of my English skill. Without them, I cannot get where

I am. I really love them. Dr. Hideaki Uchiyama, postdoctral fellow at INRIA Rennes, is my soul mate. Since we joined our laboratory, we have discussed so many things about research, teaching, *etc*. Working with a talented people like him keeps my motivation higher. I hope we will collaborate and publish a good work in the future.

I would also like to thank Prof. Kouhei Ohnishi and his laboratory members. Since I was a master course student, we have shared experimental rooms. When Prof. Ohnishi has time, he lets me join conversation with his students on various topics. His incredibly broad knowledge-base and his philosophy on both research and education helped me to find out how to survive as a researcher.

Not technically but mentally, Masayuki Takahashi, whose high-school basketball team I belonged to, has strengthened my character. Through daily practice and tournaments, he taught me how to set a direction to reach a critical goal, contribute to a team, conduct a team, and how further team chemistry takes us. Based on what I have learned under his direction, I have tough worked during my Ph.D. course.

Finally, my greatest gratitude goes to my family. Their patience, encouragement, and support have supported me in any situations. Without their support, I would never have achieved what I have ever done.

August 8, 2011
Yuji Oyamada

Contents

List of Figures

List of Tables

Chapter 1

Introduction

When we convey a thing to others, how do we do? In face to face communication, we rely on both oral and non-verbal communication. Oral communication, spoken verbal communication in other words, typically relies on words. In contrast, non-verbal communication meaning wordless communication relies on gesture and facial expression and *etc*.

When we want to do it across time and space, what should we do? The old-fashioned way follows the above type of communications. For example, folk stories and songs passed by word of mouth are categorized into this type. However, the reprise of such style varies the terms and the expression of the contents, thus we may lose the part of the contents as a result. Using telephone and video chat such as Skype ™, indirect communication can be achieved across space but not time. For communication across time, the better way is written communication. Ancient people used pictograph, resemblance of objects for the purpose. They painted images onto walls or incised into stones using mineral pigments. Figure 1.1 shows a cave painting of a horse drawn by Cro-Magnon peoples. Even though we have no idea what they wanted to tell by such pictograph, the graph can tell the information of the era. Printing technology further encouraged this type of communications, especially for text. A big leap in the technology occurred by printing-press technology intended in the 15th century. The printing-press devices enabled rapid and precise copy of text document. This is the reason why the invention and spread of the technology are regarded as the revolutionizing events in the

Figure 1.1: Written communication at ancient time: Cave painting of a horse at Lascaux drawn by Cro-Magnon peoples.

second millennium. With the printing technology, the contents can be preserved semi-permanently.

An idiom *Seeing is believing* means that physical or concrete evidence is convincing. This indicates that conveying a thing prefers showing the thing rather than telling the thing. Thus, it is natural that we have developed devices taking a photograph, an image of projecting lights of a scene. Before the first photographs, the principal of pinhole camera was mentioned by Mo Di, Chinese philosopher, and Aristotle, Greek mathematician, in the fifth and fourth centuries BC. *Camera obscura* consisting of a box or a room with a hole in one side is the concrete device of pinhole camera. Light from an external scene of the camera passes through the hole and then reaches a surface inside. The image of the scene can be projected onto the surface, and can then be manually traced to produce a photograph of the scene. Joseph Nicéphore Niépce, a French inventor, invented revolutionizing camera like printing-press technology for text [Gernsheim, 1977]. His key idea is to omit manual drawing from imaging process by relying on photochemical action so that we can automatically obtain a photograph. Before Niépce, photographs were not permanent, unable to permanently secure the images from fading. Gorman mentioned that his camera was designed based on heliograph [Gorman, 2002]. Figure 1.2 shows the earliest surviving photograph taken by Niépce. The big lim-

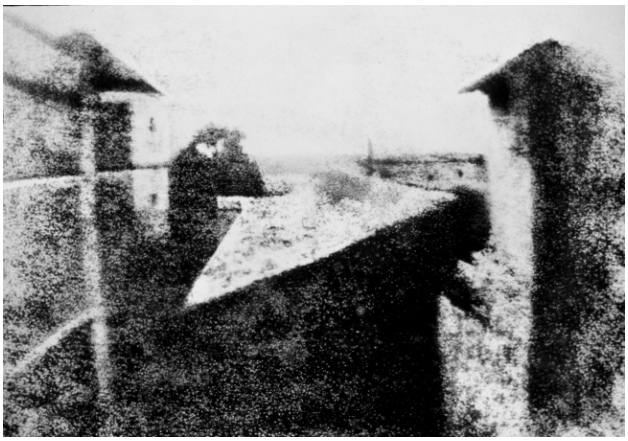


Figure 1.2: The earliest surviving photograph of a scene from nature taken with a camera obscura: *View from the Window at Le Gras*. [Joseph Nicéphore Niépce, 1826]

itation of Niépce's camera is its exposure time. It takes about eight hours for the camera to yield the photochemical action. Thus, his follower focused on achieving shorter exposure time. Through the 19th century, many advances in photographic glass plates and printing were made in. George Eastman replaced photographic plates to photographic film. This replacement was distributed through the late 19 century and results the technology of today's film camera. Nowadays, digital camera is one of the most popular devices for photo shooting. The difference between digital camera and film camera is their memory media. As memory medium, film camera uses photographic film while digital camera does memory devices such as memory card by converting the received light to digital data format via an electronic image sensor.

Digital camera takes two steps to provide the observation of the photograph as shown in Fig. 1.3 while Niépce's camera directly generated a photo of the scene via one process. First process is image acquisition process. The process receives lights from the scene and then converts the received light as the latent image. For this process, film camera uses photographic film or plate while digital camera uses imaging sensor, *e.g.*, a Charge-Coupled Device (CCD) image sensor or Complementary Metal-Oxide-Semiconductor (CMOS) sensor. Next is image display process. This step transforms the latent image into a visible image. For

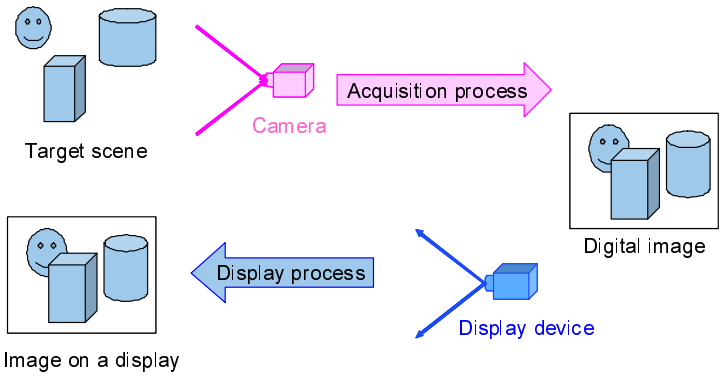


Figure 1.3: Imaging processes of digital camera: Image acquisition process converts the energy of lights coming from the target scene to measurable value. Image display process shows the digital image using a display device.

images saved on the film, we follow photographic processing, which is the chemical ways to produce a negative or positive image. On the other hand, digital image has various ways of displaying the photo. One may use printers to make the photo permanent while another may use display devices to see the photo temporarily. Typical display device is computer monitors including Cathode Ray Tubes (CRT) display and Liquid Crystal Display (LCD). Thanks to the recent development of display technologies, bigger and brighter display is available with cheaper cost. In contrast to such monitors, projectors only have light emitting devices. To form an image, projectors require display surface, onto which they emit the light.

1.1 State-of-the-art imaging technologies

Where can the imaging technologies do so far? Here, I briefly overview the state-of-the-art acquisition and display technologies independently.

State-of-the-art acquisition technology

As I mentioned in the previous section, camera has evolutionary been developed. Ancient cameras like Niépce's camera realized automatic photo shooting resulting permanent photo even though single shot takes longer than our senses.

Due to the direct imaging style, this type of cameras lacks reproducibility. Since film cameras split the imaging process into two sub processes, this type of cameras enabled reproduction of taken photo. With the development of high-speed Internet and high-power computer, digital cameras provided ease of reproducibility and editability.

Contrast to the change of form, its basic model has still been same as the camera obscura, which was originally mentioned more than two thousands year ago. What makes the revolution of image acquisition? I think computational photography has great potential for this question. The concept of computational photography is to redefine/reconstruct the acquisition process by considering the post-processing scheme as a part of acquisition process so that we can optimize the optics and processing scheme for the purpose. One famous example is a plenoptic camera for light field acquisition [Ng et al., 2005]. Light field describes light in every direction of the scene. Once we obtain the light field of a scene, it allows us to: view the scene from another point of view; refocus the image; and change perspective of the image. In conventional camera, a pixel receives the light passing through the lens but not consider the direction. To obtain light field with a conventional camera, we take photos from multiple positions, *e.g.*, by single moving camera [Levoy and Hanrahan, 1996] or multiple fixed cameras [Wilburn et al., 2005]. Ng et al. enabled a single fixed camera to capture a light field by integrating microlens array in the optical design [Ng et al., 2005]. For further information including theory and applications, readers can refer to the articles on this topic [Cohen and Szeliski, 2006; Debevec, 2006; Levoy, 2006; Nayar, 2006].

State-of-the-art display technology

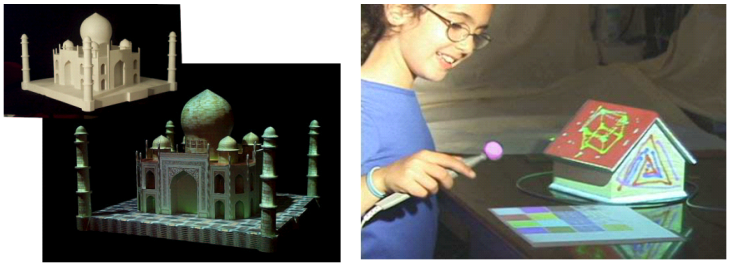
Different from acquisition technologies, display technologies are heading in different directions.

One direction considers display 3D information similar to the light field acquisition. One way is to show several images on a 2D display to the viewer so as to perceive the images as 3D. Very typical one is stereoscopic display. The display separately shows two offset images to the left and right eye of the viewer.The offset is combined in the brain and then be perceived as they have 3D depth. Another

type use a lenticular lens that emits different images to different angles [Roberts, 2003]. With the lens, a 2D display can show the different images to both eyes. The limitation of this type of displays is that the position of the viewer is restricted.

Another way uses volumetric displays rather than 2D displays. To form 3D image on volumetric object, we generally use projector(s). By changing the projecting image, we can change the appearance of the volumetric object so that the lit object looks like textured object. Paul Debevec and his colleague proposed an interactive 3D display [Jones et al., 2007]. The display consists of a high-speed video projector, a spinning mirror covered by a holographic diffuser, and FPGA circuitry to decode specially rendered DVI video signals. Since the spinning mirror reflects the projected 3D object to all angles, multiple viewers around the display can simultaneously see the 3D object. They further enhanced the display to work it on-line such as 3D video teleconference [Jones et al., 2009]. The system captures the user's 3D face data by real-time 3D scanning technique. Then, the captured data is transmitted to the remote location and showed the 3D face data using the 3D display.

As mentioned by Bimber and Raskar in their book [Bimber and Raskar, 2005], projectors allow daily objects, not only a spinning mirror, to be 3D display. We can control the appearance of the display object by somehow adjusting the projection on the display shape. Famous example is *Shader Lamps* originally proposed by Raskar et al. [Raskar et al., 2001]. Figure 1.4 shows Shader Lamps series. Original Shader Lamps uses multiple projectors to render a virtual texture onto the physical object of the same shape [Raskar et al., 2001]. As shown in Fig. 1.4 (a), the appearance of the non-textured wooden model becomes painted Taj Mahal. Dynamic Shader Lamps is an extension of the original Shader Lamps to movable objects [Bandyopadhyay et al., 2001]. The Dynamic Shader Lamps allows the display object movable and provides a 3D painting interface as interaction with the display as shown in Fig. 1.4 (b). Shader Lamps Avatars focuses on person-person communication [Lincoln et al., 2011]. Both the dynamic motion and the appearance of a real person is captured and projected onto a human-shaped display surface as shown in Fig. 1.4 (c). For more information about 3D display with projectors including Shader Lamps, readers can refer to a book on the topic [Bimber and Raskar, 2005].

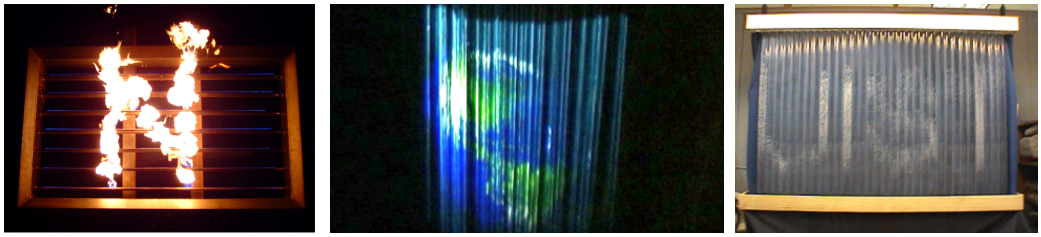


(a) Shader Lamps (b) Dynamic Shader Lamps

(c) Shader Lamps Avatars

Figure 1.4: Shader Lamps series: (a) Shader Lamps: (Left) Original appearance of wooden Taj Mahal model. (Right) The same model enhanced by adjusted projection. Courtesy of [Raskar et al., 2001]. (b) Dynamic Shader Lamps: A user is painting on movable object. Courtesy of [Bandyopadhyay et al., 2001]. (c) Shader Lamps Avatars: Implementation and diagram of the system. (Left) A user captured by a camera at the capture site. (Middle) An avatar projected by a projector at the display site. (Right) The un-illuminated avator. Used courtesy of Department of Computer Science, UNC Chapel Hill from [Lincoln et al., 2011].

Another direction of display technologies uses natural material as display object as shown in Fig. 1.5. The infernoptix Digital Pyrotechnic Matrix uses computer-controlled bursts of fire as a display [NAO Design]. Barunm et al. use water drops as display [Barnum et al., 2009]. The display is created by a projector that illuminates water drops falling from a drop generator. Detecting the 3D position of the drop based on the computer vision technique, the display projects an image to the drops. Display shape varies according to the set up of the drop generator. Heiner et al. use air bubbles rising up tubes of water [Heiner et al., 1999]. Accurate control of air release enables scrolling up display of image. This type of displays is well-reviewed by Moere [Moere, 2008].



(a) Fire display (b) Water drops display (c) Air bubble display

Figure 1.5: (a) Fire display showing a character ‘N’. Courtesy of [NAO Design]. (b) Curved water drop display showing the 3D model of a globe. Courtesy of [Barnum et al., 2009]. (c) Air bubble display showing text ‘UIST’. Courtesy of [Heiner et al., 1999].

1.2 For the next generation imaging technology

At this point, let us take a moment to imagine the configuration of imaging technology in the future. There exist several issues to be considered.

I think cheaper and more readily available device is preferred by consumers and such devices also provide the opportunity to researchers more. Famous example is iPhone [Apple Inc.] and Kinect [Xbox.com]. Before iPhone (or other current smartphone), Augmented Reality (AR) researches were mainly done with desktop/laptop PC and camera(s) connected to the PC. This set up limits where the application runs. Since smartphone consists of CPU, camera, and display, it is regarded as a mobile AR set up. As a result, the concept and potential benefit of AR have rapidly become public knowledge. Same as iPhone, Kinect opens the door to Human Computer Interaction (HCI) researchers. One way of HCI researches uses human action, *e.g.*, gesture and pose, as the interface. This way is very straightforward and intuitive strategy but requires human action recognition technique. With its SDK provided by Microsoft, human pose can be recognized in real time so that we can focus on interaction design more. These situations insist that cheaper mass product device is preferred for the configuration of next imaging process. Readers can check how much the devices mentioned here get attention in the world by searching keywords ‘iPhone AR’ and ‘Kinect’ on Internet.

Considering the cost and portability, camera and projector is a reasonable configuration candidate. State-of-the-art display technologies indicate that we are

going to use projector in dynamic scene. Currently, projected content is simple image and computer graphics objects. In a sense, such content is static, somehow generated beforehand. Let me show some related works of static contents projection. Shader Lamps [Raskar et al., 2001] mentioned above is an example of static contents projection in static scene. The system projects pre-rendered textured 3D model to fixed wooden model, the position of which is known. Yotsukura et al. proposed HyperMask [Yotsukura et al., 2002] that is an example of static contents projection in dynamic scene. HyperMask projects an animated face onto a physical mask worn by a moving actor. Using infrared LED embedded on the mask, the system adjusts the projection onto the moving display, the actor's mask. Audet et al. proposed a method for static content projection in dynamic scene [Audet et al., 2010]. Their method is based on gradient method based object tracking algorithm [Baker and Matthews, 2004]. Tracking the textured planar screen and, their method can adjust the projection to the moving screen in real time. Short conclusion of the literature is that

- projected contents are static such as pre-taken photo/movie and pre-rendered 3D model, and
- display object is moving plane or fixed volumetric objects.

What kind of configuration we should develop in the future? My tentative answer is dynamic contents projection in dynamic scene, *e.g.*, Shader Lamps Avatars [Lincoln et al., 2011]. Figure 1.6 illustrates the concept. We capture a scene of interest, moving people. Virtual information, hair, is overlaid on the captured contents and then projected on the volumetric and deformable objects, probably mannequin like object, in real time. The point of this dynamic contents projection in dynamic scene is that all materials in the scene can be dynamic. Once the guy moves, overlaid hair should be synchronized to the motion to fit on his head. Furthermore, display object also should be synchronized to his motion. This is my answer to the question.

To realize this dynamic projection in dynamic scene, we have to tackle several difficulties and one of them is image degradation. Figure 1.7 shows a projector based 3D human face display I have developed [Oyamada and Saito, 2006] and its fatal problem. In the system, a user holds a white mannequin and changes

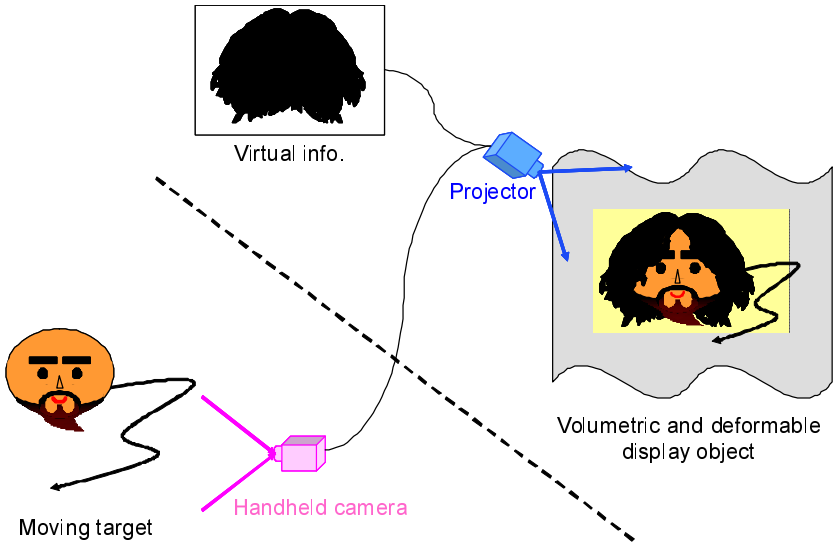


Figure 1.6: Configuration of imaging process in the future: Concept of dynamic projection in dynamic scene. Target object, moving people, is captured by a camera. Virtual information, hair, is overlaid on the captured contents and then a projector projects the contents on the volumetric and deformable surface.

its position and orientation. A camera tracks the mannequin and then a projector fits projected 3D face model by using tracking information. Since projector has narrow depth-of-field, the appearance on the display object may be blurred by projector defocus as shown in Fig. 1.7 (c). Such degradation is occurred during imaging processes. Addition to projector defocus, degradation occurred during the image acquisition process, *e.g.*, motion blur, may occur if the projected texture is captured/rendered in real time. Even with expensive devices and experienced users, satisfying visualization is difficult and challenging task. Figure 1.8 shows examples of image degradation. When camera/target object moves during the exposure time, obtained photograph is blurred by the motion. On an unoptimized display, displayed image forms on the display may contain defocus effect due to projector's narrow depth-of-field. Comparing with Fig. 1.3, the top figure is horizontally blurred and the bottom figure is defocus blurred by the projection.

To prevent such degradation, manufactures integrate solutions directly into devices. Anti-motion blur technologies use hardware support for the purpose. *Image Stabilizer* [Canon, 1995], *Vibration Reduction* [Nikon, 2000], and *Anti-*

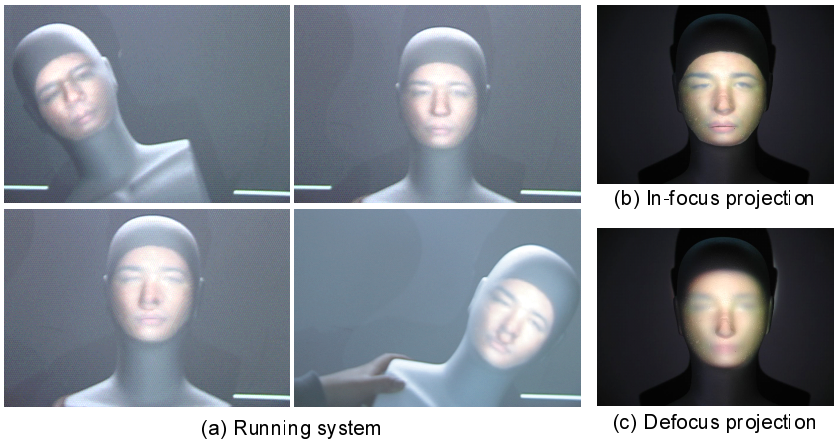


Figure 1.7: Projector based 3D human face display [Oyamada and Saito, 2006] and its fatal problem. (a) Figures of running system. (b) In-focus projection. (c) Defocus projection.

Shake [KonicaMinolta, 2003] detect camera motion by motion sensors in the camera, *e.g.*, gyro sensor, and shift lens [Canon, 1995; Nikon, 2000] or CCD array [KonicaMinolta, 2003]. Shifting lens or CCD arrays, geometric relation between the scene and the CCD array is stabilized, thus blur effect of camera motion can be cancelled on the observed image. Projector manufactures also provide image adjustment functions such as *Keystone correction* [NEC Display Solutions Europe, a] and *Wall color correction* [NEC Display Solutions Europe, b]. Keystone function fits projection onto slanted surface that is not perpendicular to the projector. There are two types of keystone functions, optical or digital. Optical keystone adjusts the image by physically modifying the light-path through the lens system while digital one does the image by shrinking the image before projection. Wall color correction function enables projection onto colored surface such as blackboard. When a user specifies the color of wall, the projector modifies the projecting color based on the pre-defined color lookup table.

We, computer vision researchers, have also provided solutions. Image restoration means recovering the degradation-free image given degraded image(s). For example, denoising targets noisy image damaged during acquisition process and deblurring restores blurred image degraded by motion during its exposure time. Stretching the notion of image restoration, projector image adjustment onto un-

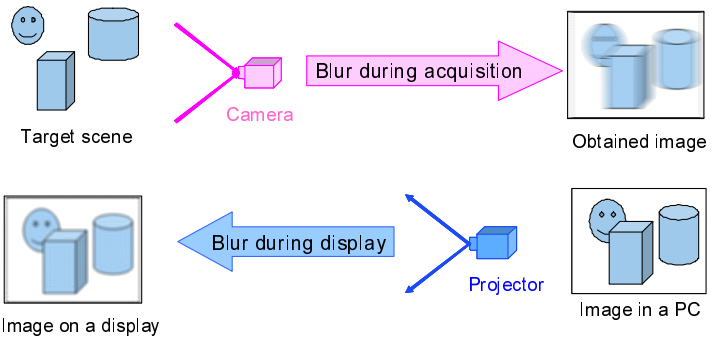


Figure 1.8: Image degradation during the imaging processes. Motion blur during the acquisition process and defocus blur during the display process.

optimized surface can be regarded as one of restoration works. When we use a projector under unoptimized situation, *e.g.*, off-axis projection and projection onto volumetric surface, normal projection results degraded display on the surface. To adjust the projection, we have to adjust the projection somehow. This adjustment can be considered as restoration operation.

For upcoming imaging technologies era, how can the computer vision technique contribute to enjoy dynamic projection in dynamic scene? From my perspective, image restoration technique is necessary. But how to do that? Two solutions are considered. First solution is for a simpler system just display the raw dynamic contents, no virtual contents. In such applications, what we should take care is final observation, intermediate observation, *e.g.*, the captured photo, is not important. Thus, a single restoration operation performed on the display process side is enough. On the other hand, the other solution for a system adding virtual contents requires several restoration operations. Since the system add the virtual information to the captured dynamic contents, the captured contents should be stabilized. Therefore, we have to take care of intermediate observation not only final one. From a resolvability point of view, the latter type is easier than the former one. Restoration of single process should solve several types of degradations merged on the final observation within a step while one of several processes does it step by step. Even though we have to separate the data flow somehow, sequential solution is generally easier than batch solution.

This thesis assumes that blur occurred during imaging processes is separable

(separating process is not discussed here) so that we can separately address image restoration technique upcoming imaging technologies era. Specifically, I focus on blur estimation of the following topics:

- **Motion deblurring**: removing motion blur effect occurred during acquisition process, and
- **In-focus projection**: removing projector defocus blur effect occurred during display process.

Figure 1.9 illustrates these two issues. Even though both topics seem to be in different research categories, they share the same purpose that removes the blur effect from observation. Motion deblurring removes the motion blur effect after the image is blurred while in-focus projection enhances the image component to cancel the defocus blur effect before the image is blurred. This means that both they are blur correction researches but the type of correction differs. Motion deblurring is a post blur correction while in-focus projection is a pre blur correction.

1.3 Summary of original contributions

This section outlines major contributions of this thesis. The proposed method should consider the conditions of upcoming imaging technologies, *i.e.*, dynamic projection in dynamic scene. To reduce the hardware constraint, simple set up is preferred. Thus, I put constraints on target blur and estimate blur information from a single photo shooting based on traditional signal processing ways. Here, the contributions are described for each of topics.

Cepstral Analysis based Non-Linear Motion PSF Estimation

Motion deblurring is one of the long existing problems in computer vision research field. For motion deblurring, there are two research issues to be solved. One is to estimate Point Spread Function (PSF) that represents degradation process, *e.g.*, motion path in the case. The other is to recover the unknown latent

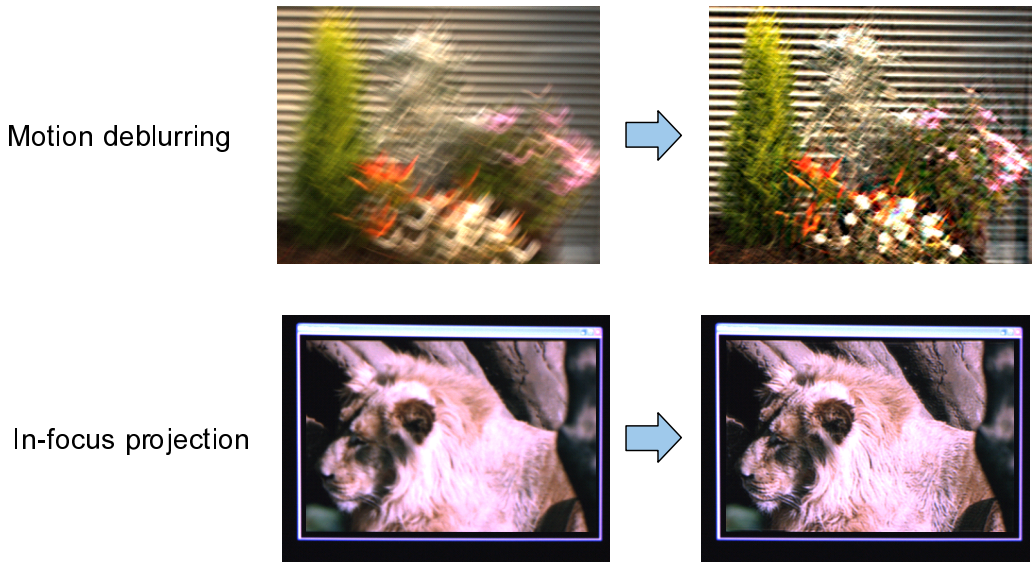


Figure 1.9: Two blur correction addressed in this thesis. (Top) Deblurring for motion blur occurred during acquisition process. (Bottom) In-focus projection for projector defocus blur occurred during display process.

image, namely removing the blur effect from the blur image. In this thesis, I focus on the former issue, namely PSF estimation of motion deblurring.

There exist several types of PSF estimation approaches, detail of the literature is mentioned in Ch. 2. Typically, the cepstral approach is used only for parametric PSFs, *e.g.*, linear motion. For more complicated PSF, *e.g.*, non-parametric motion, state-of-the-art works follow natural image statistics and estimate such PSF based on regularized minimization algorithms. The main contribution of this work is that I extend the cepstral analysis based PSF estimation to handle non-linear motion PSF. To introduce the cepstral approach for non-linear motion PSF estimation, I will analyze the cepstral behavior of non-linear motion. Then, I propose a non-linear motion PSF estimation method based on the analyzed cepstral behavior. The proposed method is validated with both synthesized images and real world images. The experimental results show that cepstral analysis based method can work for non-linear motion PSF estimation.

In-Focus Projection from a Single Projector-Camera Image Pair

Projector is well-used display devices with strong limitations. Since it was released to the world, we have wondered flexible usage of the device. Suppose that we can fully control the projection, the device can contribute to wide range of research fields and also our daily life as mentioned in Sec. 4.1. Projector image adjustment is to adjust a projector image onto unoptimized surface, *e.g.*, off-axis projection and projection onto volumetric surface. There exist three types of adjustment works; geometric adjustment, color adjustment, and focal adjustment. In this thesis, I focus on focal adjustment, so called in-focus projection. Similar to motion deblurring, there are also two research issues to be solved. One is to estimate PSF, *e.g.*, amount of projector defocus, here. The other is how to remove the defocus effect occurring after projection. In this thesis, I tackle both issues but mainly focus on PSF estimation.

Not only in-focus projection, almost the existing projector image adjustment works use fiducial patterns, *e.g.*, chess board like pattern, to estimate the information of image degradation. This is straightforward strategy for projector-camera systems that we project such fiducial patterns for estimation. In fact, there exist few works that passively estimate the information of degradation. However, all the existing in-focus projection works use fiducial patterns for estimation. To mount focal adjustment framework on upcoming imaging technologies, we require passive blur estimation method. The main contribution of this work is that I introduce PSF estimation strategy into this field so that the proposed method can passively estimate the information of projector defocus without fiducial patterns projection. The experimental results show that the proposed method achieves in-focus projection without using fiducial patterns.

1.4 Organization of the thesis

This thesis is divided into five chapters. This chapter stated global background of this thesis and described the main contribution of the latter parts of the thesis. In Ch. 2, the related works of blur correction are reviewed to clarify the position of this thesis. Chapter 3 describes motion deblurring work and Ch. 4 presents

in-focus projection work respectively. Finally, Ch. 5 concludes this thesis with a summary of the contributions and several directions of future works.

Chapter 2

Related Works on Blur Correction

This chapter reviews the previous and related works. First of all, Sec. 2.1 and Sec. 2.2 clarify the technical problems of blur correction. Next, the previous and related works are reviewed. For ease of learning, image restoration algorithms are first reviewed in Sec. 2.4 and then Sec. 2.5 reviews blur estimation algorithms.

2.1 Problem statement

A blurred image g is described by a convolution of a latent image f and a blur kernel k plus image noise n as

$$g(\mathbf{x}) = f(\mathbf{x}) \otimes k(\mathbf{x}) + n(\mathbf{x}), \tag{2.1}$$

where $\mathbf{x} = (x, y)^{\mathrm{T}}$ denotes a pixel position. Since blur kernel represents how the blur process spreads an ideal single point, blur kernel k is called Point Spread Function (PSF).

The goal of image deblurring, deconvolution in other word, is to recover the unknown latent image f. The difficulty of image deblurring is its ill-posedness. Figure 2.1 shows an example of the ill-posedness. The blurred image is generated by middle row of (f, k) pair, woman face degraded by zigzag motion. Top and bottom are wrong pairs. Top row pair is no-blur case and bottom one mistakes blur as horizontal motion. Even though our perception can easily judge which pair is the correct, all these pairs computationally well-explain the blurred image.

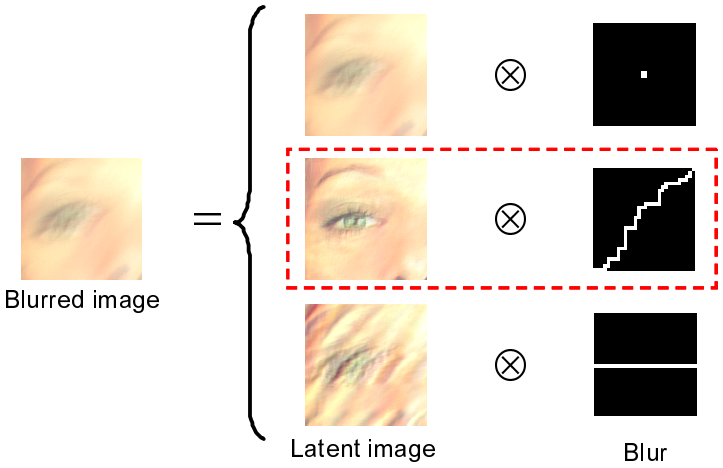


Figure 2.1: Example of Ill-posedness of image deblurring. Middle row of (f, k) pair is ground truth. Top and bottom pairs are wrong but theoretically well-explains the blurred image g.

As the figure shows, there is no unique answer. For such problems, we require additional queue(s) to resolve the ambiguity.

2.2 Additional queues for disambiguation

What kind of information can be the helpful for deconvolution? The conditions of target scene is considered as constraints *e.g.*, blur type and available data.

2.2.1 Blur type

Blur type of PSF can be a constraint of deconvolution. There exist various types of PSF as shown in Fig. 2.2. When target PSF can be described by a parametric form, it drastically decreases the ambiguity of potential PSF. Popular parametric models are a linear motion (LM) and defocus blur. An LM blur is parameterized by two parameters, motion direction and motion length, as

$$k^{\mathrm{L}}(\mathbf{x}) = \begin{cases} \dfrac{1}{L} & \text{if } ||\mathbf{x}|| \leq \dfrac{L}{2} \text{ and } \dfrac{y}{x} = \tan\theta \\ 0 & \text{elsewhere} \end{cases}, \tag{2.2}$$

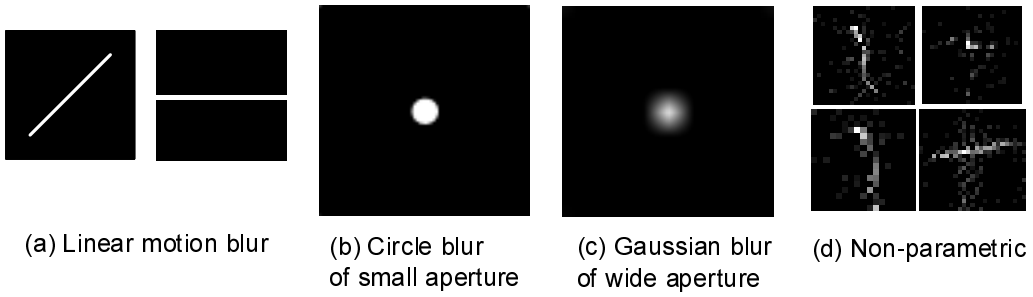


Figure 2.2: Various types of PSF. (a) linear motion PSF. (b) Circular disk PSF. (c) 2D Gaussian PSF. (d) non-parametric motion PSF.

where θ and L denote motion direction and length respectively. Defocus blur is parameterized as a circular disk function or an isotropic Gaussian. A circular disk PSF represents defocus blur of smaller aperture lens, *e.g.*, camera defocus, as

$$k^{\text{circle}}(\mathbf{x}) = \begin{cases} \dfrac{1}{\pi r^2} & \text{if } ||\mathbf{x}|| \leq r \\ 0 & \text{elsewhere} \end{cases}, \tag{2.3}$$

where r denotes the radius of the circle. On the other hand, an isotropic Gaussian PSF represents defocus blur of larger aperture lens, *e.g.*, projector defocus, as

$$k^{\text{Gaussian}}(\mathbf{x}; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \tag{2.4}$$

where σ denotes standard deviation of the Gaussian. When more complex motion occurs as shown in the right figure of Fig. 2.2, we give up the style of representing such motion with parametric form. Such motion is called non-parametric PSF.

Another classification of blur type is based on its uniformity. Classical approaches assume blur uniform on an image while some of recent works try to estimate non-uniform blur on an image. Motion blur parallel to its image plane and defocus blur of same depth scene are described by a uniform PSF. Non-uniform case is camera rotation, differently moving objects and defocus blur of different depth scene.

2.2.2 Available data

Available data of scene can also be a constraint of deconvolution. Additional image(s) of same target scene helps disambiguation. When we take two photos of a scene and they are differently blurred, how should we get the latent image? Rav-Acha and Peleg mention that using two differently blurred images are better than a single blurred image [Rav-Acha and Peleg, 2005]. When the camera exposure time varies, different types of degraded images are available. Yuan et al. use one blurred image with longer exposure time and one noisy image with shorter exposure time [Yuan et al., 2007]. The concept is to utilize color of the blurred image and sharpness of the noisy image. Combining information extracted from both blurred and noisy images, they recover the latent image rather than a recovered image from either image. An image set, each image of which is captured with different focal plane, of same target scene is useful to know defocus information. Since defocus blur is proportional to the distance between the object and camera's depth-of-field, depth information can be computed from such image set. By comparing such images, we can classify in-focus and defocus regions of images. Once the in-focus region of each image is extracted, all-in-focus image can be generated by merging the regions [Antunes et al., 2005].

If additional devices are available, what kind of devices can help deconvolution? PSF estimation of camera motion blur is equivalent to camera motion estimation. Thus, motion sensor is useful for camera motion blur estimation. Nikon provides lens containing motion sensors in its body [Nikon, 2000] while Joshi et al. put motion sensors inside camera body [Joshi et al., 2010]. Contrast to motion blur, depth/range sensors are useful for defocus blur estimation. The amount of defocus is proportional to the distance between object and the depth-of-field of the device. Therefore, NEC Display Solutions, Ltd. puts depth sensor inside projector body [NEC Display Solutions, Ltd.].

2.3 Deconvolution

Roughly speaking, deconvolution has two types, non-blind and blind deconvolution. Classical approaches assume that we somehow know the PSF, thus the

task is to estimate f given g and k as

$$\hat{f} = \text{Non-blind deconvolution}(g, k). \tag{2.5}$$

This type of deconvolution is called non-blind deconvolution. It is important to note that even non-blind deconvolution is ill-posed problem. On the other hand, blind deconvolution estimates f and k given g as

$$\left(\hat{f}, \hat{k}\right) = \text{Blind deconvolution}(g). \tag{2.6}$$

Note that correct information of k is necessary even though the goal is to estimate f. In the following, non-blind deconvolution works are first reviewed and then blind ones are reviewed.

2.4 Non-blind deconvolution

Non-blind deconvolution estimates a latent image f given a blurred image g with a known PSF k and possibly known noise n. Even though k is known, non-blind deconvolution is still ill-posed problem. Based on methods' strategies, previous works are categorized into two types of approaches, analytical solutions and numerical solutions.

2.4.1 Analytical solutions

Analytical solutions try to perform the inverse operation of convolution. Basically, they are based on the convolution theorem that the Fourier transform of a convolution is the pointwise multiplication of the Fourier transforms as

$$\mathcal{F}(f \otimes k) = \mathcal{F}(f)\mathcal{F}(k). \tag{2.7}$$

The simplest deconvolution method is inverse filtering in the frequency domain. Neglecting the noise term in Eq. (2.1), the latent image can be recovered by

inverse filtering in the frequency domain as

$$\hat{f} = \mathcal{F}^{-1}\left(\frac{F_g(\mathbf{u})}{F_k(\mathbf{u})}\right), \tag{2.8}$$

where F denotes the Fourier transform of the subscript and $\mathbf{u} = (u, v)^{\mathrm{T}}$ denotes a spectrum frequency. If the image g is perfectly noise free and the spectrum of the kernel $\mathcal{F}_k$ has no zero values at any frequencies, the inverse filtering should derive the perfect latent image.

Unfortunately, everyday photography usually violates the above assumption. They may contain some noise and F_k has negligible values at some frequencies. One straight forward solution is introducing denoising algorithm. Chesneau et al. combine image denoising algorithm with the inverse filtering [Chesneau et al., 2009]. They first apply denoising algorithm to blurred image and then apply the inverse filtering to the denoised image. Therefore, the restored image should have less noise artifacts than the inverse filtered image.

Wiener filtering considers the deconvolved noise of frequencies that have poor signal-to-noise ratio [Wiener, 1949]. The Wiener filter is formulated in the frequency domain as

$$F_{\text{Wiener}}(\mathbf{u}) = \frac{F_k^*(\mathbf{u})}{|F_k(\mathbf{u})|^2 + F_n(\mathbf{u})/F_f(\mathbf{u})} \tag{2.9}$$

where F^* is the complex conjugate of F, and F_f and F_n denote the power spectra of the ideal image f and the noise n, respectively. In the case of no noise, the Wiener filter becomes simple inverse filtering. Wiener filtering has been used for a lot of applications, however it's still have limitations. Wiener filtering assumes known signal-to-noise ratio for every frequency. Even though we assume uniform signal-to-noise ratio, it is typically unknown.

Another solution is to make the problem well-posed. Raskar et al. develop a coded-exposure camera that flutter the camera's shutter open and close during the exposure time with a binary pseudo-random sequence [Raskar et al., 2006]. Due to its randomness, coded-exposure camera prevents zero and negligible values and flattens its spectrum. Thus, the inverse filtering (Eq. (2.8)) works on the blurred images.

2.4.2 Numerical solutions

Contrast to the analytical solutions, numerical solutions estimate most likely f by numerical computation. Given g, k, and some prior knowledge on f, recovered image is obtained as

$$\hat{f} = \underset{f}{\operatorname{argmin}} |Q(g - f \otimes k) + \lambda R(f)|, \tag{2.10}$$

where Q denotes the data fidelity term that measures the distance between the observation g and the estimate f, R denotes the regularization term that is derived from our prior knowledge on f, and λ is a parameter that balances the trade-off between the fidelity term and the regularization term. In the literature, there exist two types of methods, non-regularized minimization and regularized minimization.

2.4.3 Non-regularized minimization

Non-regularized minimization methods consider only the fidelity term of Eq. (2.10) as

$$\hat{f} = \underset{f}{\operatorname{argmin}} |Q(g - f \otimes k)|. \tag{2.11}$$

The most well-known and well-used solution is Richardson-Lucy (RL) algorithm [Lucy, 1974; Richardson, 1972] that assumes Poisson distribution on image noise n. Based on the Bayes' theorem, RL algorithm iteratively finds the latent image f as

$$f_{t+1}(\mathbf{x}) = f_t(\mathbf{x}) \left(k(\mathbf{x}) * \frac{g(\mathbf{x})}{g_t(\mathbf{x})} \right), \tag{2.12}$$

where $*$ denotes the correlation operator and $g_t(\mathbf{x}) = f_t(\mathbf{x}) \otimes k(\mathbf{x})$. Considering that g_t is the prediction of a blurred image according to the current estimate f_t, the fraction $\frac{g}{g_t}$ can be regarded as the residual error between the real blurred image and the predicted blurred image.

As previously mentioned, RL algorithm assumes Poisson noise, means that the algorithm is suitable for very low light conditions such as astronomical images and microscopy images. However, noise in everyday photograph is usually modeled

by Gaussian noise. As Tai et al. mention in their paper [Tai et al., 2011], RL algorithm considering Gaussian noise is formulated as

$$f_{t+1}(\mathbf{x}) = f_t(\mathbf{x}) + (g(\mathbf{x}) - g_t(\mathbf{x})) * k(\mathbf{x}). \tag{2.13}$$

The derivation of RL algorithm for both Poisson noise and Gaussian noise are provided in Appx. A.

The RL algorithm assumes that residual $g - f \otimes k$ follows a parametric distribution such as Poisson distribution. However, such assumption is sensitive to outliers, or deviations, from the assumed model. Angelino et al. relax the parametric assumption on noise statistical model [Angelino et al., 2008]. Instead of parametric model, they use Parzen window estimation for describing the distribution of the residual. Their algorithm minimizes differential entropy of the residual reducing the dispersion of the residual.

2.4.4 Regularized minimization

Even though RL algorithm has been well-known and well-used, it has very important drawback that iteration amplifies the noise. This sensitivity to the noise can be reduced by introducing regularization term. Regularized minimization methods consider both the data fidelity term and the regularization term same as Eq. (2.10).

Total Variation (TV) norm assumes smooth intensity change in the latent image. TV regularization term is formulated as

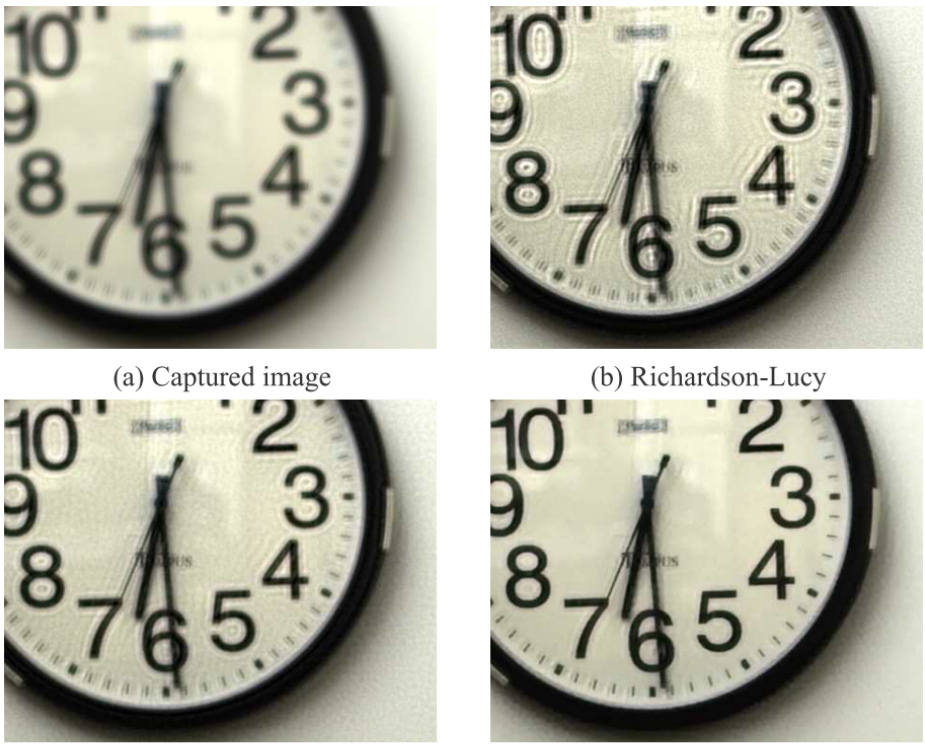
$$R_{\text{TV}}(f) = \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sqrt{|f(\mathbf{x}) - f(\mathbf{x}')|^2}, \tag{2.14}$$

where $\mathbf{x}'$ denotes nearest neighbors of $\mathbf{x}$. Minimization of TV regularization term prefers f has locally uniform color. Due to the form of regularization, it is impossible to derive analytical expression. Thus, one may use Bayesian inference [Ayasso and Mohammad-Djafari, 2009; Babacan et al., 2007, 2010], *e.g.*, variational Bayes and Markov Chain Monte Carlo (MCMC), while another may use Iterative Re-weighted Least Squares (IRLS) [Rodriguez and Wohlberg, 2009].

Recent works rely on the statistical property of natural images. Very famous and widely used property is that the distribution of the gradients of natural images is zero-peaked and heavy-tailed. In other words, the gradients of natural images can be represented by sparse vectors in appropriate domains. Thus, the sparsity regularization term is formulated as

$$R_{\text{sparse}}(f) = \sum_{\mathbf{x}} ||\partial^* f(\mathbf{x})||^{\alpha}, \tag{2.15}$$

where ∂^* returns derivatives of f, *e.g.*, 1st and 2nd order derivatives, and $||\cdot||^{\alpha}$ denotes L_{α} norm. Minimization of sparsity regularization term depends on α as Levin et al. mention [Levin et al., 2007]. When $\alpha = 2$, the objective function Eq. (2.10) become a convex function so that we can derive a closed form solution for minimization. This $\alpha = 2$ case is called Gaussian prior due to its form. The advantage of Gaussian prior is that we can analytically solve the problem. However, as some papers mention, the distribution of the gradient of natural images is more sharp than the sharpness of $\alpha = 2$. Thus, hyper-Laplacian prior, $\alpha < 1$, is used [Hou et al., 2010; Krishnan and Fergus, 2009; Levin et al., 2007]. Even though hyper-Laplacian prior represents the statistical property more than Gaussian prior, the object function is no longer convex, thus cannot be optimized in closed form. For minimization, IRLS is usually used [Levin et al., 2007]. Krishnan and Fergus decompose the minimization problem into two sub problems to fasten the computation [Krishnan and Fergus, 2009]. Using either lookup tables or analytical formulae, their approach provides comparable quality to IRLS in less than 3 seconds for a one mega pixel image. Hou et al. state the distribution of the gradient varies according to the order of derivative [Hou et al., 2010]. Using anisotropic derivatives as derivative function ∂^*, their algorithm can reduce the ringing artifaces caused by image noise. Figure 2.3 compares Richardson-Lucy algorithm (Eq. (2.12)), least-squares deconvolution with a Gaussian prior (Eq. (2.15), $\alpha = 1$), and one with a hyper-Laplacian prior (Eq. (2.15), $\alpha < 1$). Richardson-Lucy algorithm is damaged by ringing artifacts. Gaussian prior has less ringing artifacts but is smoothed. Hyper-Laplacian prior recovers sharper edges than others and has less noise and ringing artifacts. For further description on natural image statistics, please refer to Appx. B.



(a) Captured image (b) Richardson-Lucy

(c) Gaussian prior (d) Sparsity prior

Figure 2.3: Comparison of non-blind deconvolution algorithms. (a) Captured image. (b) Richardson-Lucy algorithm is damaged by ringing artifacts. (c) Gaussian prior is smoothed but less damaged by ringing artifacts. (d) Hyper-Laplacian prior recovers sharper edges and has less noise and ringing artifacts. Courtesy of [Levin et al., 2007].

Addition to the sparsity prior, Joshi et al. use two-color regularization term [Joshi et al., 2009]. The regularization term focus on the statistics that images can locally be described as a mixture of as few as two colors. The two-color model represents a pixel intensity as a linear combination of two colors. Thus, the two-color regularization term is formulated as

$$R_{\text{two}}(f) = \sum_{\mathbf{x}} |f(\mathbf{x}) - (\alpha P + (1-\alpha)S)|^2, \tag{2.16}$$

where P and S are primary and secondary colors and α is the linear mixing parameter. Since local color statistics reduces over-smoothing around step edges and

high-frequency texture, two-color regularization terms favors sharp edges while the sparsity prior does smooth edges.

Wang et al. combine existing works [Shan et al., 2008; Weiss and Freeman, 2007] for both fidelity term and regularization term [Wang et al., 2009]. Contrast to the fidelity term of other methods, Shan et al. use several orders of image derivatives as fidelity term [Shan et al., 2008]. As Simon proves, higher order partial derivatives of image noise follow Gaussian distributions with different standard deviations if image noise itself follows [Simon, 2002]. As regularization term, Wang et al. consider two image statistics, one represents global property of images and the other does local property [Wang et al., 2009]. For representing global property of natural images, they adopt Gaussian Scale Mixture Field of Experts (GSM FOE) model proposed by Weiss and Freeman [Weiss and Freeman, 2007]. GSM FOE model characters the image using a set of high dimensional linear filters to grasp the long range pixel correlations. Thus, the regularization term smooths out the image noise. As local property of natural images, they constrain smoothness of local areas. The regularization term enforces the blurred image gradient to be similar to the latent image gradient in the smooth area. For minimization of these terms, they use IRLS algorithm.

2.5 Blind deconvolution

Blind deconvolution estimates a latent image f and an unknown PSF k given a blurred image g. The methods are categorized into two types based on the types of target PSFs, one for parametric PSFs and the other for non-parametric PSFs. Readers may refer [Kundur and Hatzinakos, 1996] and [Levin et al., 2009] for further discussion.

2.5.1 Parametric PSF estimation

Classical approaches put a constraint on PSFs that target PSF has parametric form as mentioned in Sec. 2.1. The point of the constraint is that the assumption severely reduces the potential PSFs. For example, linear motion PSF estimation is equivalent to blur direction and length (θ, L) estimation and Gaussian defocus

PSF estimation is equivalent to σ estimation. Thus, PSF estimation problem is redefined as PSF parameters estimation problem.

2.5.2 Spectral/cepstral patterns identification

It is well-known that parametric PSFs have clearer features in spectrum/cepstrum domain rather than ones in image domain. One of the simplest approaches for estimating PSF parameters is to find such features. However, the features in spectral/cepstral domain are sensitive to presence of noise. Differences between the related works are how they identify the features against the noise.

The spectra of parametric PSFs have periodic patterns. The spectrum of LM is a 2D sinc function that has periodic lines of spectral zero. These lines are orthogonal to the motion direction θ and the period is inversely proportional to motion length L. The spectrum of defocus blur is the Bessel function of the first degree that has periodic circles of spectral zero. The radii of the circles are inversely proportional to defocus radius r. Thus, Spectral periodic patterns of zeros are the features of parametric PSFs and are corresponding to the blur parameters.

Such spectral zeros are corresponding to negative spikes in the cepstrum domain. The cepstrum of an image is the spectrum of log of the power spectrum of the image as

$$\mathcal{C}(\cdot) = \mathcal{F}^{-1}(\log|\mathcal{F}(\cdot)|), \tag{2.17}$$

where $\mathcal{C}$ denotes the cepstrum transform. The cepstrum of an LM has periodic negative spikes along the motion direction θ with period L. The cepstrum of defocus blur has periodic circles of negative spikes. The radii of the circles are proportional to doubled defocus radius r. Same as spectral zero patterns, cepstral periodic patterns of negative spikes are the features of parametric PSFs and are corresponding to the blur parameters.

Figure 2.4 shows PSFs in each domain. Top line shows LM PSF and bottom line shows defocus blur PSF. From left to right, PSF in image, spectrum, and cepstrum domain are shown. An LM PSF has periodic black lines in spectrum domain, corresponding to spectral zeros, and periodic negative spikes in cepstrum domain. Defocus PSF has periodic black circles in spectrum domain and periodic negative circles in cepstrum domain. These spectral zeros and cepstral negative

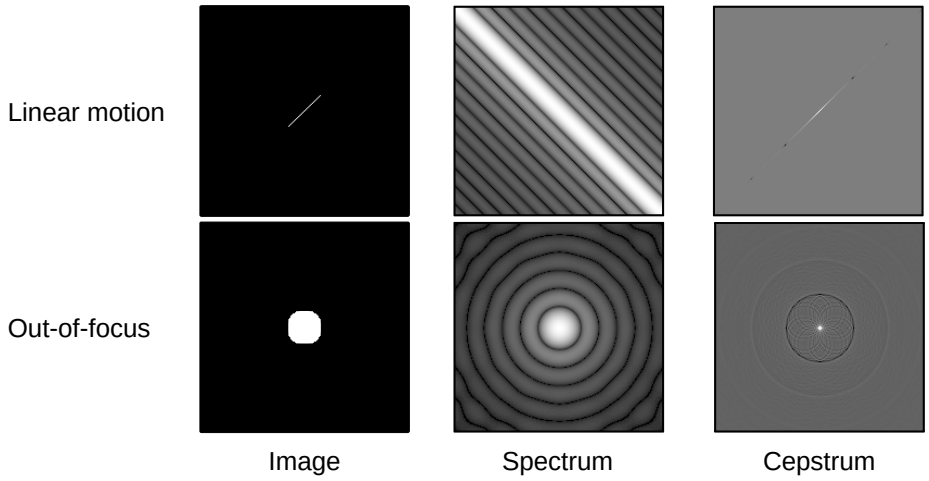


Figure 2.4: PSFs in each domain. Top line shows LM PSF and bottom line shows defocus blur PSF. From left to right, PSF in image, spectrum, and cepstrum domain are shown.

spikes are clear features of parametric PSFs.

Neglecting the noise term of imaging equation (Eq. (2.1)), the spectrum of a blurred image is rewritten as Eq. (2.7). The equation indicates that if F_k has zero value at frequency $\mathbf{u}$, F_g should also have zero value at the same frequency $\mathbf{u}$. Therefore, identifying periodic zeros of F_g is equivalent to PSF parameters estimation.

Regarding image deconvolution problem, the important cepstral property is that convolution of two images is corresponding to the sum of their cepstra as

$$\begin{aligned} C_g &= \mathcal{C}(g) \\ &= \mathcal{F}^{-1}\left(\log|\mathcal{F}(g)|\right) \\ &= \mathcal{F}^{-1}\left(\log|\mathcal{F}\left(f \otimes k\right)|\right) \\ &= C_f + C_k, \end{aligned} \tag{2.18}$$

where C denotes the cepstrum of the subscript. Note that C_k is relatively bigger than C_f at lower quefrencies. Since the distribution of spectrum of PSF is relatively smaller than that of a latent image, C_k converges at lower quefrencies while C_f is distributed from lower quefrencies to higher quefrencies. Thus, PSF

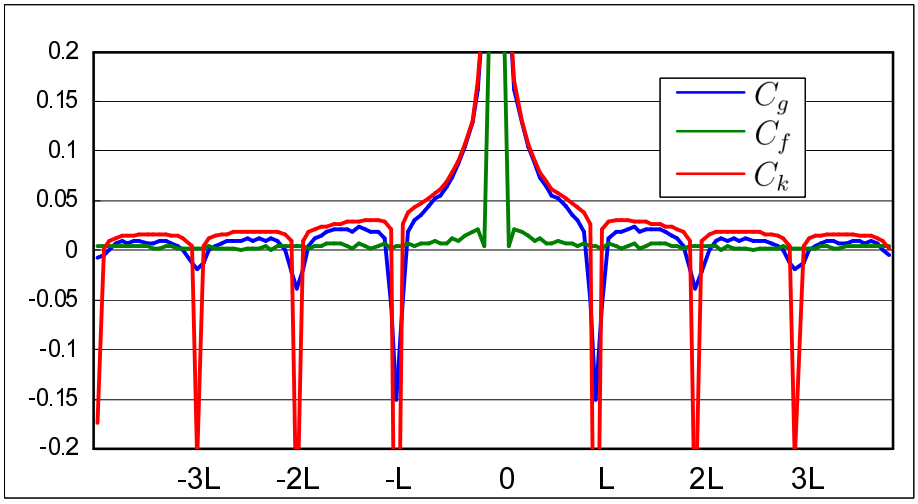


Figure 2.5: Comparison of cepstrum components of a motion blurred image. Blue, green, and red curves represent C_g, C_f, and C_k component along the motion direction.

component C_k is dominant in C_g at lower quefrencies. As an example of this behavior, Fig. 2.5 compares the cepstrum components of a motion blurred image. Blue, green, and red curves plot cepstrum of the blurred image C_g, latent image component C_f, and PSF component C_k extracted along the motion direction respectively. Comparison of the plots show that C_k is very closer to C_g while C_f has smaller values. Hence, we can say that C_k is dominant in the cepstrum of a blurred image at lower quefrencies. Moreover, periodic negative spikes of C_g is clear enough to identify. Therefore, identifying the negative spikes of C_g is approximately equivalent to parametric PSFs estimation problem.

Simplest approaches simultaneously estimate (θ, L) by just identifying the spectral zero values or cepstral negative spikes [Rom, 1975]. Such approaches are mathematically clear but are often frustrated due to noise effect and the overlying structure of unknown latent image component. They require blurred images to have high Signal-to-Noise ratio enough to recognize the patterns. Thus, the methods work only when the noise effect is enough small to recognize such patterns. Table 2.1 roughly classifies the related works according to how they treat such difficulties.

Table 2.1: Classification of the spectral/cepstral patterns identification works according to their key ideas.

Estimation	Key idea	Papers
(θ, L)	Raw cepstrum	[Rom, 1975]
	Take derivative	[Gennery, 1973; Ji and Liu, 2008]
	Spatial invariance of PSF	[Cannon, 1976; Kang et al., 2006]
	Natural image statistics	[Sun et al., 2009]
$\theta \rightarrow L$	Global behavior of OTF	[Mayntz et al., 1999]
	Spatial invariance of PSF	[Chang et al., 1991; Fabian and Malah, 1991]
	Global behavior of OTF	[Moghaddam and Jamzad, 2007; Oliveira et al., 2007; Wu et al., 2007]

One solution takes derivative of a degraded image to suppress the lower frequencies and to accentuate the high frequencies [Gennery, 1973]. Generally, the amplitudes of the lower frequencies are much greater than ones of the higher frequencies. Thus, taking derivative before the Fourier transform tends to flatten the spectrum so that the patterns in spectrum/cepstrum domain can be identified easier. Ji and Liu theoretically explain this property [Ji and Liu, 2008]. When a function $k(x)$ is differentiable, the Fourier transform of its derivative is given by $2\pi iuF_k(u)$. In the case of linear motion, $k(x)$ is a rect function and its Fourier transform $F_k(u)$ is a sinc function. The Fourier transform of the derivative of linear motion becomes

$$\begin{aligned} \mathcal{F}\left(\frac{d}{dx}k(x)\right) &= 2\pi iu\mathcal{F}\left(k(x)\right) \\ &= 2\pi iu\frac{\sin u}{u} \\ &= 2\pi i \sin u. \end{aligned} \tag{2.19}$$

Thus, taking derivative makes sinc function to sin function so that the spectral features are identical clearer than the Fourier transform of the original function.

Another type of approaches utilizes the spatial-invariance of PSFs [Cannon, 1976; Chang et al., 1991; Fabian and Malah, 1991; Kang et al., 2006]. As previously mentioned, most papers assume shift-invariant blur on a whole image while latent image and noise is globally shift-variant. By partitioning the blurred image into sub-images and then averaging the spectra of the sub-images, we can reduce the contribution from latent image component and noise while keeps the contribution from PSF component. Thus, spectral/cepstral patterns appear clearer.

Suppose we somehow extract the spectrum of PSF from the spectrum of a blurred image, PSF estimation becomes easier problem. Sun et al. achieved the above strategy by adding another constraint on latent images [Sun et al., 2009]. Their assumption is that the spectrum of latent images can be represented by monotonically decreasing isotropic polynomial function. Their method first estimates the global shape of the spectrum of unknown latent image from one of the blurred image. Then, the method extracts the spectrum of PSF by subtracting the estimated spectrum from one of the blurred image. As a result, we can obtain the modified spectrum of PSF. Once obtained, we can estimate θ and L simultaneously by autocorrelation.

Sequential estimation can also be robust to noise effect. Instead of direct identification, sequential estimation first estimates blur direction θ from the global shape of the spectrum of PSF and then estimates blur length L by identifying the patterns along the estimated motion direction [Chang et al., 1991; Fabian and Malah, 1991; Mayntz et al., 1999; Moghaddam and Jamzad, 2007; Oliveira et al., 2007; Wu et al., 2007]. For θ estimation, general approach utilizes the shape of Optical Transfer Function (OTF), the spectrum of a PSF. The idea of this method is to assess the anisotropy caused by linear motion blur in spectrum domain. The power spectrum of unblurred latent image is approximately isotropic, will discuss this features later. Since the spectrum of a blurred image is product of that of latent image and PSF, the spectrum of the blurred image becomes anisotropic. Motion direction θ estimation can be done by using this characteristic. OTF affects the coarse behavior of the spectrum of a blurred image along motion direction. Therefore, integral along a line on the spectrum can be useful for θ estimation. Oliveira et al. estimate a direction that has highest variance of Radon transformation as θ [Oliveira et al., 2007] while Moghaddam and Jamzad estimate paral-

lel lines of spectral zeros by Radon transform [Moghaddam and Jamzad, 2007]. Since the OTF is a 2D sinc function, θ estimation is equivalent to estimating the long and short axis of the sinc function. Thus, Mayntz et al. estimate θ from the inertia matrix, the eigenvectors of which are parallel and orthogonal to motion direction [Mayntz et al., 1999]. For L estimation, Fabian and Malah apply comb-like liftering [Fabian and Malah, 1991], filtering in cepstrum domain, so that negative spikes derived from noise component can be removed from L estimation. Bispectrum is known as insensitive to additive, symmetrically distributed noise [Chang et al., 1991; Moghaddam and Jamzad, 2007]. When the SNR is relatively high, both spectrum and bispectrum have observable patterns. On the other hand, when the SNR decreases, spectrum loses the patterns while bispectrum still keeps the patterns. Thus, L estimation can be done more reliably using the bispectrum for lower SNR images.

2.5.3 Parameter search

Another type of approaches estimates a PSF as a parameter search problem. They search blur parameters over the parameter space by evaluating a match metric computed for each parameter value as

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \, |Q(g - f \otimes k(\Theta))|, \tag{2.20}$$

where Θ denotes a set of parameters, *e.g.*, $\Theta = (\theta, L)$ for LM, and the data fidelity term Q is match metric evaluation function. When a parameter provides the best match according to the metric, the parameter is chosen as the optimum parameter $\hat{\Theta}$. The point of this type of approaches is what kind of metric the algorithm uses. Table 2.2 roughly classifies the related works according to their match metric.

Yitzhaky and Kopeika use the spatial property of PSF [Yitzhaky and Kopeika, 1997]. When we take a derivative along with the motion direction, sum of absolute difference of the gradient image become minimum. Based on this property, they first estimate the motion direction θ. Once the direction is estimated, motion length is estimated along the direction. They compute autocorrelation along with the estimated direction based on a property that only if the space shift is same as

Table 2.2: Classification of the parameter search works according to their match metric.

Match metric	Papers
Spatial behavior of PSF	[Yitzhaky and Kopeika, 1997]
Spectral behavior of restored image	[Tan et al., 1991]
Kurtosis of histogram of restored image	[Li et al., 2005]
Sharpness of restored image	[Rooms et al., 2004]
Restoration residual	[Levin et al., 2007; Savakis and Trussell, 1993]

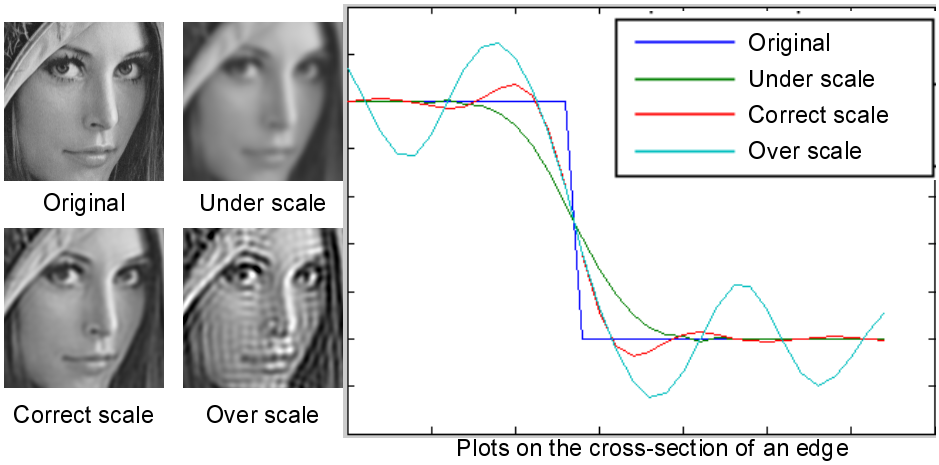


Figure 2.6: Physical versus numerical focus effect on the cross-section of an edge.

blur length, autocorrelation is minimum.

Restored image provides the information of amplitude of restoration. In theory, edges in a blurred image are fully recovered by a restoration algorithm with the optimum parameter. When the parameter is smaller than the optimum, the edges are less recovered while the bigger parameter results ringing artifacts. Figure 2.6 compares restored edge with varying amount of restoration. With correct scale, the edge gets closer to the original one while restored edge of under/over scales are less/more restored. As this figure shows, the behavior of edges in restored image can be a match metric. Tan et al. analyzed how the wrong PSF effect the restored edges [Tan et al., 1991]. Specifically, their method focuses on the

behavior of the Green's function defined as

$$\mathcal{F}^{-1}\left(\frac{F_k\hat{F}_k^*}{\hat{F}_k\hat{F}_k^*+\gamma}\right), \tag{2.21}$$

where F_k and $\hat{F}_k$ denote the true PSF and the estimate and γ denotes PSNR. The Green's function describes how a point in the latent image would appear in the restored image. For linear motion and accelerated motion, they analyzed how appropriate/inappropriate PSFs affect the recovered edges. Based on their observations, they classify the blur type, linear motion or accelerated motion, and estimate the blur parameter. Li et al. measure the sharpness by kurtosis of the histogram of a restored image [Li et al., 2005]. Their observation is that the larger the kurtosis is, the smoother the histogram is. Based on the observation, they compute the kurtosis of the restored images and then choose the restored image with the smallest kurtosis and the corresponding parameter is regarded as the optimum parameter. They had experiments on several types of blur, *e.g.*, atmospheric turbulence blur, defocus blur, etc. Rooms et al. use a wavelet basis as an edge detection filter [Rooms et al., 2004]. Natural images have a property that wavelet coefficients of the images are very sparse [Olshausen and Field, 1997]. To characterize the sparsity of wavelet coefficients, their metric for defocus blur is based on the kurtosis of the histogram of two sub-bands of wavelet coefficients. The optimum parameter is obtained by finding a parameter maximizing the sharpness of the kurtosis.

Restoration residual can be another match metric. Suppose we have a set of PSF candidates $k_n(\Theta_n)$, where $n = 1, \ldots, N$. The restoration residual of n th PSF candidate is computed as

$$r_n(\mathbf{x}) = g(\mathbf{x}) - f(\mathbf{x}) \otimes k_n(\mathbf{x}), \tag{2.22}$$

where r_n denotes the restoration residual corresponding to a PSF candidate k_n. The optimum parameter can be estimated by finding a parameter minimizing the

residual with the following Q function:

$$Q(r_n) = \sum_{\mathbf{x}} |r_n(\mathbf{x})|^2. \tag{2.23}$$

However, the latent image f is unknown, we alternatively use the restored image for the metric as

$$r_n(\mathbf{x}) = g(\mathbf{x}) - \hat{f}_n(\mathbf{x}) \otimes k_n(\mathbf{x}), \tag{2.24}$$

where $\hat{f}_n$ denotes the restored image obtained by the candidate k_n. The point of this type of approaches is how to estimate a latent image f_n for a PSF candidate k_n and how to compute the residual metric. Savakis and Trussell measure the residual in spectrum domain [Savakis and Trussell, 1993] as

$$F_{r_n}(\mathbf{u}) = F_g(\mathbf{u}) - F_{\hat{f}_n}(\mathbf{u}) F_{k_n}(\mathbf{u}). \tag{2.25}$$

As deconvolution algorithm, they evaluate three algorithms, the inverse filter, the linear maximum a posteriori probability (MAP) filter, and the Wiener filter. They conclude that the Wiener filter provides the most suitable analytical expression of the residual power spectrum. Furthermore, they evaluate three match metric, the Mean Square Error, the Chi Square test, the Kolmogorov-Smirnov test, and argued that the MSE test and Chi Square test performed the most consistently over the blur types and parameters. Levin et al. also measure the restoration residual in image domain [Levin et al., 2007]. In contrast to other works, they use their own deconvolution method, described in Sec. 2.4, and a camera with coded aperture. Their own deconvolution method provides sharper image with less ringing artifacts when the optimum parameter is used. Figure 2.7 illustrates the benefit of using a coded aperture. Those images are deblurred with the correct blur scale, larger scale, and smaller scale. With a coded aperture, deblurred images with both smaller and larger scale have ringing artifacts while with a conventional aperture results ringing artifacts only with larger scale. Thus, a coded aperture is better for evaluating the restoration residual.

ARMA parameter estimation methods parameterize both latent images and PSFs. The approaches model a latent image as a 2D Auto Regressive (AR) process and a PSF as a 2D Moving Average (MA) process. Based on the models, a blurred

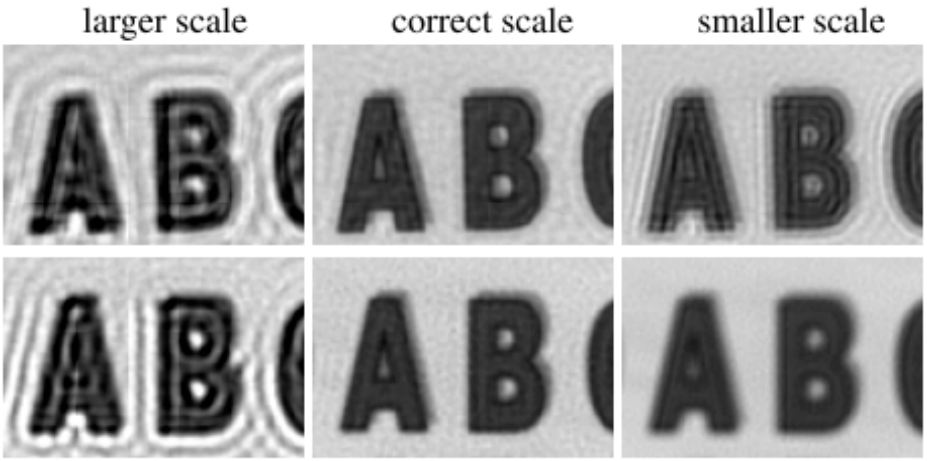


Figure 2.7: Deconvolution with varying blur scale. Top: Restored image captured by a coded aperture camera. Bottom: Restored image captured by a conventional aperture camera. Courtesy of [Levin et al., 2007].

image is represented as an Auto Regressive Moving Average (ARMA) process. The approaches estimate a PSF by identifying the ARMA parameters. A latent image is modeled as

$$f(\mathbf{x}) = a(\mathbf{x}) \otimes f(\mathbf{x}) + n_v(\mathbf{x}), \tag{2.26}$$

where n_v denotes modeling error. The problem of blind deconvolution is equivalent to estimating a and k. The Maximum-Likelihood (ML) methods attempt to derive restoration filters [Tekalp et al., 1986]. The Generalized Cross-Validation (GCV) methods, also known as leave-one-out, are based on second order statistics [Reeves and Mersereau, 1992].

2.5.4 Non-parametric PSF estimation

Even though the assumption on parameterization of PSF provides analytical solutions, real world images usually violates this assumption. Thus, it is natural to focus on non-parametric PSF estimation. Similar to numerical solutions of non-blind deconvolution, blind deconvolution for non-parametric PSF is formulated as

$$\left(\hat{f}, \hat{k}\right) = \underset{f,k}{\operatorname{argmin}} |Q(g - f \otimes k) + \lambda_f R_f(f) + \lambda_k R_k(k)|, \tag{2.27}$$

where R and λ denote the regularization term and the regularization parameter of the subscript.

There are two types of approaches for solving Eq. (2.27), sequential methods and iterative methods. Sequential methods [Chen et al., 2008; Dai and Wu, 2009; Harmeling et al., 2010; Rav-Acha and Peleg, 2005; Xu and Jia, 2010] separate the problem into two parts as

$$\begin{cases} \hat{k} = \underset{k}{\operatorname{argmin}} \, |Q(g - f \otimes k) + \lambda_k R_k(k)| \\ \hat{f} = \underset{f}{\operatorname{argmin}} \, |Q(g - f \otimes \hat{k}) + \lambda_f R_f(f)| \end{cases} . \tag{2.28}$$

The methods first estimate k from g, and then estimate f from g and $\hat{k}$. Due to the separation, they can directly apply existing non-blind deconvolution approaches for the latter problem. On the other hand, the other type of methods iteratively solve both k and f as

$$\begin{cases} \hat{k}_t = \underset{k}{\operatorname{argmin}} \, |Q(g - f_{t-1} \otimes k) + \lambda_k R_k(k)| \\ \hat{f}_t = \underset{f}{\operatorname{argmin}} \, |Q(g - f \otimes \hat{k}_t) + \lambda_f R_f(f)| \end{cases} . \tag{2.29}$$

As the equations indicate, the methods iterate k estimation and f recovery. For initial values $\hat{k}_0$ and $\hat{f}_0$, delta function and the blurred image g are usually chosen.

2.5.5 Unknown f for PSF estimation

For both types of methods, Eq. (2.28) and Eq. (2.29), PSF estimation requires the information of unknown f for the fidelity term evaluation. Since f is unknown, we should somehow prepare an alternative.

One solution is that we use another image or images as the alternative. Contrast to the appearance of natural images, their representation, *e.g.*, Fourier transform and wavelet transform, have similar distribution, so called natural image statistics [Torralba and Oliva, 2003] (see Appx. B for more detail). Following this property, we can use another image or images as the alternative of unknown latent image f for the fidelity term evaluation. Caron et al. assume a power-law distribution on the spectra of images [Caron et al., 2002]. The assumptions are that

the spectra of natural images have peak at a very narrow range of low frequencies and a flat tail at high frequencies and that Optical Transfer Function (OTF), which is the spectrum of a PSF, is slowly varying function. From the above assumptions, OTF can be represented by power-law scaled spectra of blurred image with a power-law scale factor. Thus, PSF estimation is equivalent to identifying the scale factor. Neglecting the noise effect and giving the spectrum of a reference image that shares similar frequency characteristics with the unknown latent image, their algorithm directly recovers OTF, and then apply inverse filtering in the frequency domain. Wan and Nowak assume the distributions of wavelet coefficients of natural images [Wan and Nowak, 1999]. The assumptions are that the distribution of wavelet coefficients of natural images has peaky and heavy-tailed symmetric shape and wavelet coefficients are statistically independent. Specifically, they model the distribution by mutually independent two Gaussian mixtures, one of lower variance represents sharp peak of the distribution while the other one of higher variance represents heavy-tail of the distribution. Fergus et al. and Shan et al. assume the distributions of image gradient [Fergus et al., 2006; Shan et al., 2008]. The assumptions is that the distribution of pixel intensities of image gradient of natural images has high-peak at zero and heavy-tailed symmetric shape. Fergus et al. represent the distribution of first order derivatives by Mixture of Gaussian distributions [Fergus et al., 2006]. Shan et al. mention that the distributions of several orders of derivatives still behaves following the 0-peak and heavy-tailed distribution [Shan et al., 2008]. For representation of the distribution, they introduce concatenating two piecewise continuous functions, one for 0-peak and the other for heavy-tailded representation. The representations differ by the papers, all the methods use the parameters of another image or images as the alternative of f in Q function.

The other solution estimates the alternative from the blurred image g or previous estimate f_{t-1}. Considering the numerical optimization, f in Q contributes to compute the direction to search the answer $\hat{k}$ and also $\hat{f}$. Ideal situation is that we have f (though this is not held for image deblurring) means that object function derives the steepest descent direction for each iteration. Even though we don't know true f, if we have approximation f_{apprx}, we can expect such approximation to provide the descent direction. Thus, straight forward solution is to use the

previous estimate of f_{t-1} directly [Almeida and Almeida, 2008, 2009; Cai et al., 2009a,b,c; Chen et al., 2008; Dai and Wu, 2009; Gupta et al., 2010; Harmeling et al., 2009, 2010; Hirsch et al., 2010; Hu et al., 2010; Huang et al., 2009]. However, there is a risk that the previous estimate f_{t-1} contains some deconvolution error such as ringing. Thus, some filtered image is used instead of f_{t-1} [Cho and Lee, 2009; Xu and Jia, 2010]. Well-used filters are shock filter [Osher and Rudin, 1990] to restore strong edges and bilateral filter [Tomasi and Manduchi, 1998] to suppress noise. Joshi et al. assume sharp edges on latent image f [Joshi et al., 2008]. Based on the assumption, they predict step edges on the latent image from gradients of g and use it as f of Q.

Some methods assume another image of same scene with shorter exposure time [Babacan et al., 2009; Šorel and Šroubek, 2009; Yuan et al., 2007]. Due to the exposure time, shorter exposure image should be blur free but noisy. Even though the noisy image is not perfect image, it can be a good reference for Q function evaluation.

2.5.6 Regularization term on PSF R_k

Similar to the regularization term on latent image R_f, one on PSF R_k varies the form depending on what kind of property R_k reflects. Without any loss of generality, we can put assumptions that PSF is non-negative and the energy of PSF is conserved as

$$k(\mathbf{x}) \geq 0, \tag{2.30}$$

$$\sum_{\mathbf{x}} k(\mathbf{x}) = 1. \tag{2.31}$$

Note that these constraints can reduce the potential PSFs not much.

One assumes smoothness on PSFs. Since a PSF represents motion, smooth PSF is plausible. Thus, TV regularization norm on k is used [Cho and Lee, 2009; Gupta et al., 2010; Harmeling et al., 2010] as smoothness constraint. The formulation of the regularization term is same as Eq. (2.14).

Addition to smoothness constraint, sparsity of PSFs is considered. The size of PSF is relatively smaller than one of images and PSF's values are almost zero.

Thus, Tikhonov regularization term on pixel intensities of k or on gradients of k are used [Cai et al., 2009a; Chen et al., 2008; Cho and Lee, 2009; Gupta et al., 2010; Harmeling et al., 2010; Hong and Park, 2010; Huang et al., 2009; Joshi et al., 2008; Shan et al., 2008; Xu and Jia, 2010; Yuan et al., 2007]. Smoothness of PSF is equivalent to sparsity of the distribution of PSF representation parameters of coefficients. Therefore, mixture of exponential is used to represent the distribution of PSF intensities [Babacan et al., 2009; Fergus et al., 2006] and curvelet transform is used to represent PSF [Cai et al., 2009b,c].

2.5.7 Other constraints

For a scene consists of two-layers, *e.g.*, foreground and background, we can derive an additional constraint on the scene. The constraint is that α matte of the latent image should be binary and motion smears the boundary between the objects. Thus, the regularization term on α matte is designed to favor binary α matte [Almeida and Almeida, 2008; Jia, 2007; Shan et al., 2007]. Dai and Wu also put constraint on α matte [Dai and Wu, 2009]. Their regularization term on the scene consists of two matrices. One minimizes the softcuts metric that helps obtaining smooth soft edges with transitions and the other one favors binary α matte.

Similar to α matte constraint, Huang et al. use bi-level region constraint [Huang et al., 2009]. Their method performs PSF estimation only on bi-level regions that seem to be binary in the latent image. Joshi et al. assume that observed edges in blurred image is originally ideal step edge [Joshi et al., 2008]. The method first finds the location and orientation of edges in the blurred image, and then predicts ideal sharp edges. PSF estimation is done by comparing edges of blurred image and the predicted ideal sharp edges.

2.5.8 Spatially varying PSF

The above discussion focuses on uniform blur on an image but the assumption of uniformity cannot cover the entire motion blur. Thus, some recent works try to solve spatially varying PSFs.

Piecewise homogeneous PSF

When PSF is locally homogeneous, *e.g.*, moving object captured by a fixed camera, we can assume that the scene consists of several regions each of which is blurred by homogeneous PSF. Simply speaking, applying uniform PSF estimation on each region is enough for such scene.

One uses multiple images of same scene so that geometric relation between the images helps PSF estimation. Cho et al. use several images of roughly the same scene [Cho et al., 2007]. Based on geometrical constraint, they iterate PSF refinement, image segmentation into regions of homogeneous motions, and the corresponding PSFs estimation.

When a scene consists of two-layers, *e.g.*, foreground and background, one may use a constraint on boundaries between the layers as mentioned in Sec. 2.5.7. Since motion smears boundaries between the objects, α matte of the scene is also blurred. Based on the assumption that α matte of the same scene but no motion is binary, we can estimate PSFs on each regions by comparing α matte of the blurred image and binarized one [Almeida and Almeida, 2008; Dai and Wu, 2009; Jia, 2007; Shan et al., 2007]. Note that this type of methods requires user input to segment the scene.

Instead of manual segmentation, we can use PSF estimated on every pixel to segment the scene. Levin use restoration error to estimate PSF of moving foreground [Levin, 2006]. Her method is based on the image statistics that linear motion blur changes the statistical property of an image. Based on the analysis on the statistical property, her method first estimates PSFs on several pixels. Restoration operation with estimated PSF recovers the image in the blurred areas but serious artifacts occurred in the background. Therefore, her method segments the image into blurred and unblurred layers by considering image smoothness and blur smoothness. Chakrabarti et al. point out the spectral behavior of local windowed image helps to identify blur length of linear motion [Chakrabarti et al., 2010]. Their method identifies PSF at every pixel based on the probabilistic model considering above characteristics. The obtained PSF map contains estimation error because the estimation considers only the probabilistic model. Thus, they refine the PSF map by considering object boundaries, blur smoothness, and color infor-

mation.

Piecewise non-homogeneous PSF

When PSF is smoothly varies along an image, *e.g.*, camera rotation, we have to estimate PSF on every pixel. In such case, we assume homogeneous PSF on neighboring pixels or regions.

One assumes 2D motion motions on local regions. Šorel and Šroubek estimate spatially varying PSFs by interpolating estimated PSFs on image grid of sub images [Šorel and Šroubek, 2009]. Using a pair of images of same scene, one blurred image and one underexposed noisy image, they estimate PSFs on grid. Since some of the PSFs contain estimation error, they replace such PSFs by the average of adjacent PSFs. Harmeling and his colleague introduce filter flow framework [Seitz and Baker, 2009] to represent smoothly varying PSFs [Harmeling et al., 2010; Hirsch et al., 2010]. Based on the filter flow model, we can efficiently describe spatially varying PSFs rather than PSFs on image grid.

Another focuses on camera motion what is well approximated by a few parameters and recover parameters of the motion rather than spatially varying PSFs. Gupta et al. represent 6D camera motion (3D translation and 3D rotation) by 3 degrees of motion (in-plane translation and rotation) [Gupta et al., 2010]. Whyte et al. focus on camera rotation about its optical center and describe a motion blur by sequential Homography instead of a convolution of the latent image and PSF [Whyte et al., 2010].

2.6 Blur estimation/correction suited to next generation imaging technologies

Here, let me consider what kind of blur estimation/correction approaches are suited to upcoming imaging technologies.

In the history of PSF estimation, we initially focused on *uniform & parametric* PSF. Then, *non-uniform & parametric* PSF was focused with an assumption that parametric PSF well-represents the blur in local region. In 1990s, some researchers started solving *uniform & non-parametric* PSF but they assumed sym-

metric form or simpler shape. Fergus et al.'s work [Fergus et al., 2006] provided big impact that we can estimate more complicated non-parametric PSF. Then, we get started estimating *uniform & non-parametric* PSF. Very recently, some works reported their works on *non-uniform & non-parametric* PSF estimation. Considering versatility to target scene, non-parametric PSF is better than parametric one and non-uniform PSF is better than uniform one. Some of recent works use additional data or devices, *e.g.*, multiple images with different camera parameters [Rav-Acha and Peleg, 2005] or data obtained by motion sensors [Joshi et al., 2010]. These works aim to relax the difficulty of the deconvolution problem.

What solution is suited to upcoming imaging technologies? To develop a system with cheaper cost, using additional devices is not suited. But how about additional images? In on-line system, sequential data, *e.g.*, previous image of video stream, is available in practice. Thus, using sequential data is better to relax the difficulty. However, approaches using sequential data implicitly assume that we get correct answer in the previous image. In other words, estimation/correction error propagates. In such sense, a single image based method is suited even though the difficulty still remains.

Another concern is its theoretical reliability. Even though *non-parametric* PSF estimation has more generality and has get attention, we cannot say whether this approach works or not because most of such works are numerical solutions. If we need guarantee that the approach works, analytical solutions, *e.g.*, spectral/cepstral analysis are suited.

Considering issues mentioned above, let me give a future perspective of 2the blur correction/estimation methods suited to upcoming imaging technologies. The method should be single image based approach without any additional data nor devices so that the approach can contribute to more people and be developed with cheaper cost. Furthermore, one with theoretical reliability is preferred. Thus, extension of analytical solution is suited to the imaging technologies in next era.

Chapter 3

Cepstral Analysis based Non-Linear Motion PSF Estimation

This chapter proposes a non-linear motion PSF estimation method from a single blurred image for motion deblurring. Based on the traditional signal processing theory, the proposed method estimates a PSF with two steps as shown in Fig. 3.1. First step (red frame in Fig. 3.1) is PSF candidates estimation. In this step, the method first estimates PSF candidates from the cepstrum of the blurred image based on the cepstral analysis. Second step (blue frame in Fig. 3.1) is PSF candidates evaluation. In the step, the method chooses the most likely PSF by evaluating the candidates based on the imaging equation.

3.1 Related works

There exist several works investigating camera motion. To know the real camera motion, they prepare known pattern consists of point light sources and shoot it. The light source appear in the image should represent the camera motion path. Xiao et al. investigate the 2D trajectory corresponding to camera motion in yaw and pitch axes [Xiao et al., 2006]. Figure 3.2 shows how the camera motion pattern changes according to exposure time change. With shorter exposure time, motion path looks point and line. The more exposure time is, the more complicated the PSF is. Even complicated PSF, the shape seems to be decomposable with line

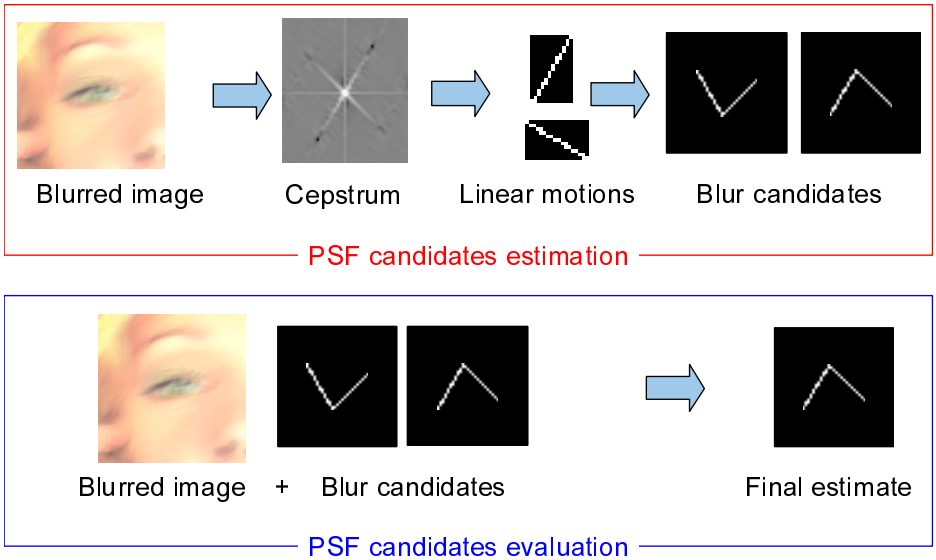


Figure 3.1: Overview of the proposed method. The proposed method takes a single blurred image as an input and estimates a PSF with two steps. The method first estimates PSF candidates from the cepstrum of the blurred image. Then, the most likely PSF is chosen by evaluating the candidates.

segments. Park et al. represent 3D trajectory by optical flow model [Park et al., 2004]. They mention that linear motion PSF can represent the basic camera motion with enough shorter exposure time. Nishi and Onda analyze the behavior of 3D camera motion [Nishi and Onda, 2010] for quantitative evaluation of camera manufacturers' image stabilizers. Figure 3.3 illustrates 3D camera motion with shorter and longer exposure time. 3D camera motion with longer exposure time appears on the image as shown in the left image. Taking light source by video camera with shorter exposure time, the sequence represents the motion segments of 3D camera motion. As shown in the right figure, each segment is represented by linear motion.

As mentioned in Sec. 2.5, the classic approaches target parametric motion while the recent approaches focus on non-parametric motion. Considering the capability for the various types of PSFs, non-parametric PSF estimation seems to be the best solution. The success of the approach is derived by the constraints on the latent image. Contrast to the approach, classic approaches target only parametric

Figure 3.2: How the camera motion changes with exposure time changes. From left to right and top to bottom, exposure time increases from 0.01 to 0.8 second. Courtesy of [Xiao et al., 2006].

PSF, *e.g.*, linear motion. The reason why the classic approaches focus on parametric PSF is because of our constraint on PSF not because of the methods' limitation. We assume parametric PSF in order to constraints on the spectral/cepstral behavior of the PSFs. Thus, it is natural that the approach handles only parametric PSF. Even though the target PSF is limited, the performance of the approaches is analytically guaranteed. On the other hand, the performance of the non-parametric PSF estimation is not guaranteed because the approach relies on the numerical minimization algorithms for computation. Therefore, some methods may require user's assist or have heavy computation cost.

3.2 Motivation

For upcoming imaging technologies, e.g., capturing moving contents by handled camera to show with volumetric display, what kind of deblurring technique is required? In other words, the questions is how complicated shape blur in video sequence is. Considering the works investigating camera motion [Nishi and Onda, 2010; Park et al., 2004; Xiao et al., 2006] mentioned above, linear motion repre-

Figure 3.3: 3D camera motion with shorter and longer exposure time. (Left) Actual camera motion shot with longer exposure time. (Middle) Test pattern consists of point light sources for acquiring 3D camera motion. (Right) Detected 3D camera motion shot with shorter exposure time. Courtesy of [Nishi and Onda, 2010].

sentation is enough for PSF of shorter exposure time and for PSF of segments of 3D camera motion with longer exposure time. As Xiao et al. mentioned [Xiao et al., 2006] (Fig. 3.2), blur with exposure time of video sequence, *e.g.*, 24 fps, is not simple linear motion but not too complicated, far from non-parametric motion PSF shown in Fig. 2.2. Even though such blur is not linear motion, we can assume locally linear motion as Nishi and Onda do [Nishi and Onda, 2010] (Fig. 3.3).

Considering such background, I focus on non-linear but locally linear motion for upcoming imaging technologies. When the motion is locally linear, it seems to be possible to apply classical parametric PSF estimation methods so that we can analytically solve the problem. Therefore, I set the goal of this work to find the possibilities of classical PSF estimation not to propose very complicated PSF estimation methods.

3.3 Overview of the proposed Method

I propose non-linear motion PSF estimation method based on traditional signal processing theory. Instead of putting constraints on latent image, the proposed method utilizes constraints on the cepstral behavior of target PSFs. As mentioned in Sec. 3.2, I focus on locally linear motion, *e.g.*, piecewise linear motion. How-

ever, I empirically find that the cepstrum helps PSF estimation when the motion is approximately linear motion but has small bounce. Thus, the proposed method focuses on two types of non-linear motion, Piecewise Linear Motion (PLM) and Noisy Linear Motion (NLM), both introduced in Sec. 3.4. Next, the detail of the proposed method is described in Sec. 3.5 and Sec. 3.6.

As an input, the proposed method takes a blurred image observed by a normal camera. For developing the analytical solutions, the proposed method assumes the cepstral behavior of target motions. However, the cepstral behavior can not directly solve PSF estimation problem, the proposed method separates the problem into two sub problems. First part estimates a set of PSF candidates from the cepsrtum of a blurred image. Next, the estimated candidates are evaluated and one of the candidates is chosen as the final estimate.

3.4 Target non-linear motions

Here, I describe the target non-linear motions, namely Piecewise Linear Motion (PLM) and Noisy Linear Motion (NLM). Figure 3.4 shows examples of target motions. One intuitive extension of linear motion representation is Piecewise Linear Motion that consists of several linear motions. The other one is Noisy Linear Motion (NLM). NLM is approximately linear motion but includes some small bounce.

3.4.1 Piecewise Linear Motion (PLM) PSF

One intuitive approximation of non-linear motion is to use a piecewise smoothness constraint like Ben-Ezra and Nayar do [Ben-Ezra and Nayar, 2004] and Nishi and Onda do [Nishi and Onda, 2010]. A Piecewise Linear Motion (PLM) is a non-linear motion but partially linear. The PSF of such motion can be represented by

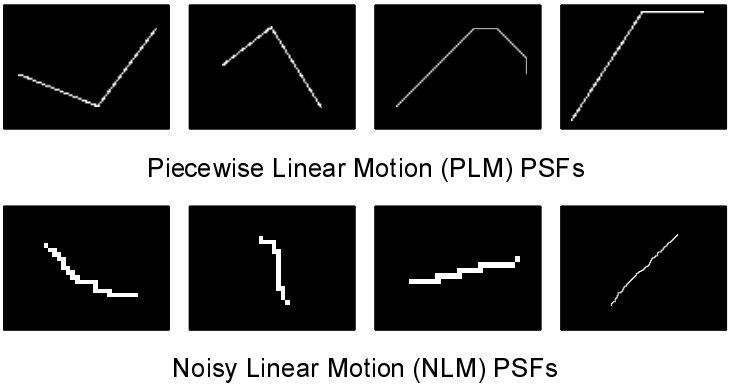


Figure 3.4: Target motion of the proposed method. (Top) Piecewise Linear Motion (PLM) PSFs. (Bottom) Noisy Linear Motion (NLM) PSFs.

sum of linear motion PSFs as

$$k^P(\mathbf{x}) = \sum_{n=1}^{N} \omega_n k_n^L(x - \Delta\mathbf{x}_n), \tag{3.1}$$

$$\omega_n = \frac{L_n}{\sum_{m=1}^{N} L_m},$$

where k^P and k^L denote PLM PSF and linear motion PSF, called component PSF, respectively, ω_n is scaling parameter forcing the constant speed motion, and $\Delta\mathbf{x}_n$ denotes connectivity of linear motions as

$$\Delta\mathbf{x}_n = \begin{pmatrix} \Delta x_n \\ \Delta y_n \end{pmatrix},$$

$$\Delta x_n = \begin{cases} 0 & n = 1 \\ \sum_{m=1}^{n-1} L_m \cos\theta_m & n \neq 1 \end{cases},$$

$$\Delta y_n = \begin{cases} 0 & n = 1 \\ \sum_{m=1}^{n-1} L_m \sin\theta_m & n \neq 1 \end{cases}.$$

With this parametric form, a PLM PSF is decomposed into two types of information. First one is blur information of each component PSF, *e.g.*, motion direction θ_n and motion length L_n. The other one is the order of the component PSFs. Thus, PLM PSF estimation is decomposed into two sub problems, component PSFs estimation and PSFs order estimation.

The cepstrum of a PLM PSF is obtained by applying the cepstrum transform to Eq. (3.1) as

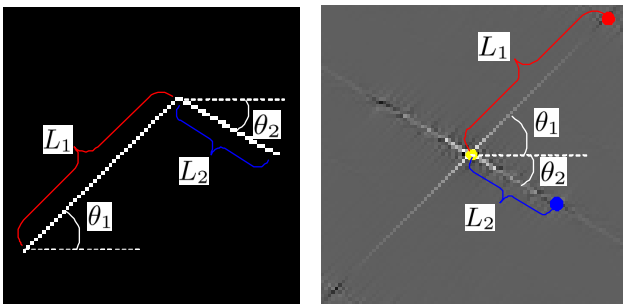
$$\mathcal{C}\left(k^P(\mathbf{x})\right) = \mathcal{F}^{-1}\left(\log\left|\mathcal{F}\left(\sum_{n=1}^{N}\omega_n k_n^L(\mathbf{x}-\Delta\mathbf{x}_n)\right)\right|\right). \tag{3.2}$$

Here, I limit the discussion to spectral/cepstral features. When all the blur directions differ, namely $\theta_n \neq \theta_m$ is held for all $n \neq m$, spectral zero values do not overlap. This means that each $\mathcal{F}(k_n^L)$ does not interfere, they are independent in other words. In such case, the Fourier transform of sum of component PSFs is equivalent to the sum of the Fourier transform of the component PSFs. Thus, Eq. (3.2) is approximated as

$$\begin{aligned}
\mathcal{C}\left(k^P(\mathbf{x})\right) &= \mathcal{F}^{-1}\left(\log\left|\mathcal{F}\left(\sum_{n=1}^{N}\omega_n k_n^L(\mathbf{x}-\Delta\mathbf{x}_n)\right)\right|\right) \\
&\approx \sum_{n=1}^{N}\mathcal{F}^{-1}\left(\log\left|\mathcal{F}\left(\omega_n k_n^L(\mathbf{x}-\Delta\mathbf{x}_n)\right)\right|\right) \\
&= \sum_{n=1}^{N}\mathcal{F}^{-1}\left(\log\left|\mathcal{F}\left(\omega_n k_n^L(\mathbf{x})\right)\right|\right) \\
&= \sum_{n=1}^{N}\mathcal{C}\left(\omega_n k^L(\mathbf{x})\right).
\end{aligned} \tag{3.3}$$

From line two to three, the Fourier transform property that shift in image space does not vary the power spectrum is used. Equation (3.3) indicates that the cepstra of a PLM PSF equals sum of the cepstrum of the component PSFs under the assumption. Therefore, we can expect that the cepstrum of a PLM PSF provides us blur information of all component PSFs.

Figure 3.5 shows the cepstrum of a PLM PSF. Figure 3.5 (a) and (b) show the



(a) PSF (b) Cepstrum

Figure 3.5: A PLM PSF and its cepstrum. Yellow point is the peak of the cepstrum, and red and blue points denote the negative spikes resulting from each component PSF.

PSF and the cepstrum, respectively. In the figure, yellow point is the peak of the cepstrum, and red and blue points denote the negative spikes resulting from each component PSF. For each motion direction θ_n, the spikes is located at L_n far from the peak.

3.4.2 Noisy Linear Motion (NLM) PSF

A Noisy Linear Motion (NLM) is a non-linear motion that is approximately linear motion but has small bounce. Since such small bounce is not easy to parameterize, we call this kind of motion path 'curve'. Suppose a curve C connects two points $\mathbf{x}_1$ and $\mathbf{x}_2$. The curve consists of N points $\mathrm{c} = (\mathrm{c}_1, \ldots, \mathrm{c}_N)$ and c_1 and c_N corresponds to $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. I put one-way constraint on noisy linear motion $\mathrm{c}_n = (x_{\mathrm{c}_n}, y_{\mathrm{c}_n})^{\mathrm{T}}$ as

$$x_{\mathrm{c}_n} \geq x_{\mathrm{c}_{n-1}}, \text{ and}$$
$$y_{\mathrm{c}_n} \geq y_{\mathrm{c}_{n-1}}.$$

Here, we analyze the behavior of the cepstrum of a NLM PSF. Empirically, I found that the cepstrum of a NLM PSF has strong values along the blur direction and that the cepstrum has distributed negative valleys along the motion direction. Figure 3.6 compares the cepstrum of a NLM PSF C_k and that of the blurred im-

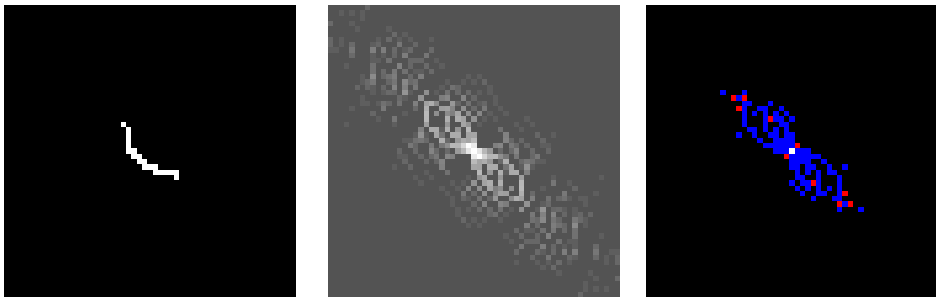


Figure 3.6: The cepstrum of a NLM PSF. From left to right, PSF, the cepstrum, and plotted values are shown. In the right figure, the cepstral peak of ceptrum, negative valleys, and strong values are plotted as white, red, and blue pixels, respectively.

age C_g. Red pixels in the right figures show that C_k has negative valleys along the motion direction but they are not periodic. Same as a linear motion PSF, C_g has additional negative valleys around the cepstral peak. Blue pixels in the right figures show that C_k has strong values along the motion direction. Different from linear motion PSFs, the strong values are distributed not located on an exact line. From this observation, I derive assumptions that

- the strong values of the cepstrum of a NLM forms approximate shape of its PSF, and
- such shape connects the cepstral peak and one of negative valleys.

3.5 PSF candidates estimation

This section describes PSF candidates estimation method using the cepstral behavior mentioned in Sec. 3.4. Since PLM and NLM have different cepstral behavior, I develop different PSF candidates estimation methods for each. For PLM, the method is based on the behavior that the cepstrum of a PLM PSF provides us blur information of all the component PSFs. For NLM, the method is based on the assumption that the strong values of the cepstrum of a NLM forms approximate shape of its PSF.

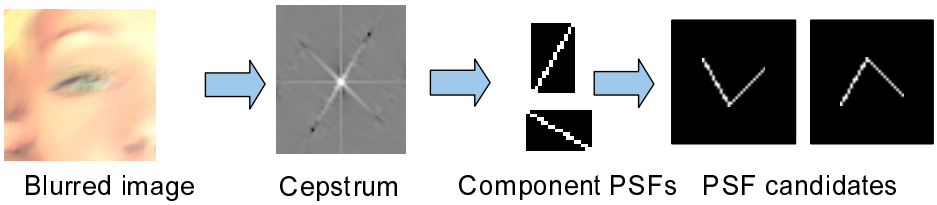


Figure 3.7: Overview of PLM PSF candidates estimation.

3.5.1 PLM PSF candidates estimation

Section 3.4.1 derives the cepstral behavior of PLM that the cepstra of a PLM PSF equals sum of the cepstrum of the component PSFs when all the blur directions differ, namely $\theta_n \neq \theta_m$ is held for all $n \neq m$. The proposed method estimates PSF candidates of a PLM PSF k^P from a single blurred image g based on the cepstral behavior. Figure 3.7 shows the overview of the method. First, a set of component PSFs is estimated based on the cepstral feature identification approaches from the cepstrum of the blurred image. Then, potential PLM PSF candidates are generated from the set by combining the components.

Cepstrum transform

The proposed method estimates component PSFs based on the cepstral behavior of linear motions. However, raw cepstrum obtained by Eq. (2.17) is noisy. For robust PSF estimation, the proposed make the cepstral feature by using traditional approaches.

One cepstral feature harming the proposed method is the effects of noise and the overlying structure of C_f. Even though C_k dominates the cepstrum around the cepstral peak, there exist the contribution of C_f and noise effect. To reduce the effect of such feature, we usually use spatially-invariance of PSF [Cannon, 1976; Chang et al., 1991; Kang et al., 2006; Maki and Sugimoto, 2007; Moghaddam and Jamzad, 2007]. Under the assumption, PSF is uniform on a blurred image while latent image component varies by region. Suppose we partition an input blurred image g into N sub images as $g_n(n = 1, \ldots, N)$. Taking an average of cepstra of

partitioned images, we have

$$\begin{aligned}\overline{C}_g &= \frac{1}{N}\sum_{n=1}^{N} C_{g_n} \\ &= \frac{1}{N}\sum_{n=1}^{N} \left(C_{f_n} + C_{k_n}\right),\end{aligned}$$

As previously mentioned, we assume uniform blur on the blurred image while f varies by region. In such case, C_{f_n} differs according to region while C_{k_n} is constant. Thus, the contribution of C_{f_n} decreases by taking average as

$$\begin{aligned}\overline{C}_g &= \frac{1}{N}\sum_{n=1}^{N} \left(C_{f_n} + C_{k_n}\right) \\ &= \frac{1}{N}\sum_{n=1}^{N} C_{k_n} \\ &= C_{k_n}. \end{aligned} \tag{3.4}$$

Figure 3.8 shows the effect of averaging cepstra. Comparing raw cepstrum and averaged cepstrum, averaged one has clearer lines than raw one.

Another feature is vertical and horizontal line due to discontinuities at the image boundaries. The cepstrum transform assumes periodic images but normal images have discontinuities at the image boundaries. Such discontinuities appear as vertical/horizontal lines going through the cepstral peak. Since the proposed method relies on the strong values on the cepstrum, those lines should be removed. Typical solution for this is windowing that tapers the image values at the boundaries. Since we have to take care of discontinuities at both image space and frequency space, we should apply windowing twice as

$$C_g^{\bigstar} = \mathcal{F}^{-1}\left(W\left(\log\left|\mathcal{F}\left(W\left(g\right)\right)\right|\right)\right), \tag{3.5}$$

where $C^{\bigstar}$ denotes cepstrum obtained with windowing and $W()$ denotes a windowing function. Specifically, Tukey window [Tukey, 1967] is used to reduce the discontinuities on image boundaries. Figure 3.9 shows the effect of windowing.

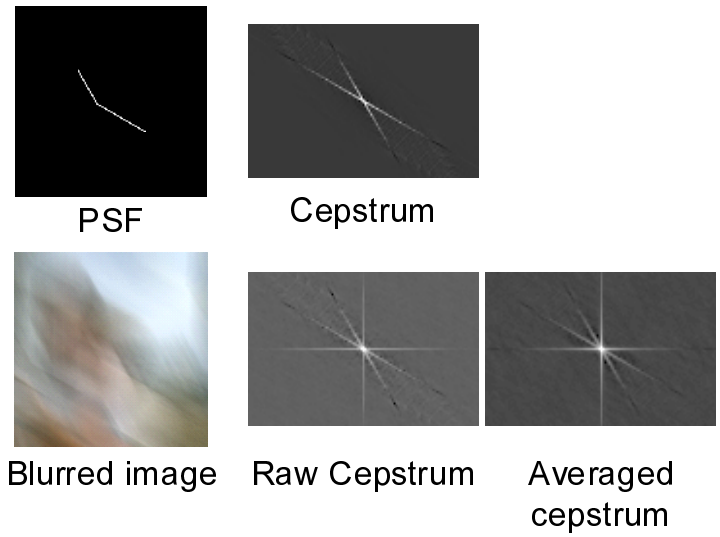


Figure 3.8: The effect of averaging cepstra. (Top) PLM PSF and its cepstrum. (Bottom) Blurred image, raw cepstrum, and averaged cepstrum.

The raw cepstrum has both vertical and horizontal lines derived from the discontinuities on image boundaries while windowed cepstrum does not have.

These two processes are concurrently applicable. As a result, the cepstrum transform of the proposed method averages the cepstra obtained with windowing of sub-images g_n as

$$\widehat{C}_g = \frac{1}{N}\sum_{n=1}^{N} C^{\bigstar}_{g_n}. \tag{3.6}$$

For ease of explanation, let C_g denote $\widehat{C}_g$ in the followings.

Component PSFs estimation

Next, the proposed method estimates a set of component PSFs by cepstral feature identification mentioned in Sec. 2.5.2.

First, the motion direction parameters θ_m are estimated from the cepstrum of a blurred image. As mentioned above, the cepstrum of a PLM PSF should keep the cepstral features of all the component PSFs when $\theta_n \neq \theta_m$ is held for all $n \neq m$. For θ_m estimation, the proposed method applies the Radon transform to extract the strong lines appearing along the motion directions. The Radon transform

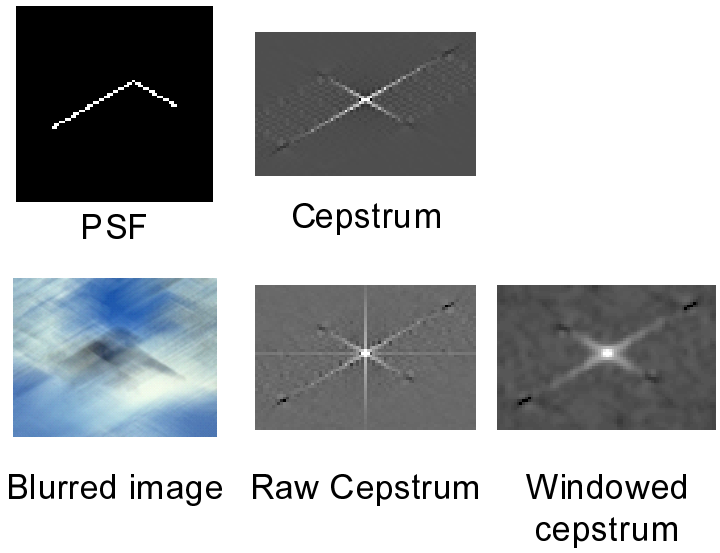


Figure 3.9: The effect of windowing. (Top) PLM PSF and its cepstrum. (Bottom) Blurred image, raw cepstrum, and windowed cepstrum.

$\mathcal{R}(C_g(\mathbf{p}), \rho, \psi)$ integrates the cepstrum $C_g(\mathbf{p})$ along a line $\rho = p\cos\psi + q\sin\psi$ as

$$\mathcal{R}(C_g(\mathbf{p})), \rho, \psi) = \iint C_g(\mathbf{p})\delta(\rho - p\cos\psi - q\sin\psi)dpdq, \qquad (3.7)$$

where δ denotes the Dirac delta function and $\mathbf{p} = (p, q)^{\mathrm{T}}$ denotes a quefrency. Since the strong lines pass through the cepstral peak, we have to apply the Radon transform for $\rho = 0$. To avoid the negative spikes works as negative bias, we use absolute value as input. Thus, motion directions estimation finds directions ψ maximizing $\mathcal{R}(|C_g(\mathbf{p})|, 0, \psi)$ as motion directions $\hat{\theta}$. Figure 3.10 shows the performance of Radon transform when $\theta_1 = 30$ and $\theta_2 = 120$. In the plot, blue and red curves plot Radon transforms of raw cepstrum and processed one. As the plot shows, the raw cepstrum has peak both correct directions $\psi = 30, 120$ and wrong directions $\psi = 0, 90$. On the other hand, the processed cepstrum has peak only correct directions $\psi = 30, 120$.

Next, the blur length parameters L_m are estimated along the estimated blur directions $\hat{\theta}$. The negative spikes, each of which is derived from one of the component PSFs, should appear along the strong lines and the distance between a

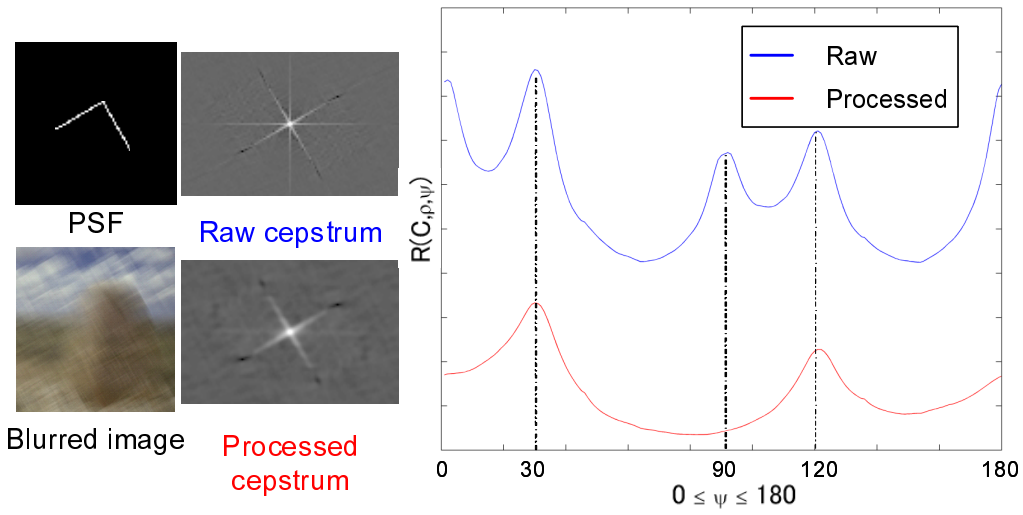


Figure 3.10: Motion direction estimation based on Radon transform when $\theta_1 = 30$ and $\theta_2 = 120$. (Left) PSF and blurred image degraded by the PSF. (Middle) Raw cepstrum (top) obtained by Eq. (2.17) and processed cepstrum (bottom) obtained by Eq. (3.6). (Right) Plot of Radon transform. Blue and red curves represent Radon transform of raw cepstrum and processed cepstrum, respectively.

spike and the peak is equivalent to blur length L. Thus, the method finds the location $\mathbf{p}$ of negative spike along each blur direction $\hat{\theta}_m$ and set $\hat{L}_m$ the distance between the spike and the peak.

PSF candidates generation

Now, we have a set of component PSFs with motion parameters $(\hat{\theta}_m, \hat{L}_m)$. The potential PLM PSFs should be one of a permutation of the estimated component PSFs. Thus, we generate all the permutation of the component PSFs. The generated PLM PSFs $\hat{k}_n^{\text{can}}$ $(n = 1, \ldots, N)$ are called PSF candidates.

3.5.2 NLM PSF candidates estimation

Section 3.4.2 derives assumptions that the strong values of the cepstrum of a NLM PSF forms approximate shape of its PSF, and shape connects the cepstral peak and one of negative valleys. If these assumptions are correct, NLM PSF can be estimated by finding a correct valley and then tracing the strong values between

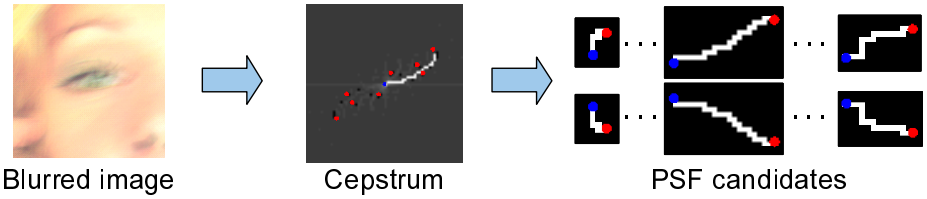


Figure 3.11: Overview of NLM PSF candidates estimation. Blue point represents the cepstral peak and red points represent negative valleys. The proposed method estimate PSF candidates each of which connects the peak and one of valleys.

the peak and the valley. Based on this assumption, the proposed method estimates PSF candidates of a NLM PSF k^N from a single blurred image g. Figure 3.11 shows the overview of the method. Instead of finding the correct valley, the proposed method estimates several PSFs each of which is corresponding to one of negative valleys as candidates. Since the cepstrum is symmetric around the peak, both the estimated candidates (top of the right figure) and their symmetric ones (bottom of the right figure) are regarded as PSF candidates.

LM PSF estimation from another aspect

Before describing the detail of NLM PSF candidates estimation, let me consider cepstral patterns identification method for LM PSF estimation [Chang et al., 1991; Fabian and Malah, 1991; Ji and Liu, 2008; Mayntz et al., 1999; Moghaddam and Jamzad, 2007; Oliveira et al., 2007; Wu et al., 2007] from another aspect. Figure 3.12 compares the cepstrum of a LM PSF and that of the blurred image. On the right figure, PSF shape (blue lines) is overlaid on the cepstrum. As you can see, both C_k and C_g have strong values along the motion direction and the appeared shape lies between the cepstral peak and the negative valley.

Let us consider LM PSF estimation using this cepstral features. Suppose we detect several negative valleys from the cepstrum of a blurred image C_g and check lines each of which connects the cepstral peak and one of the negative valleys. A detected valley on a line that has strong values should be equivalent to one of the periodic valleys. Thus, the LM PSF can be estimated by finding such line. This

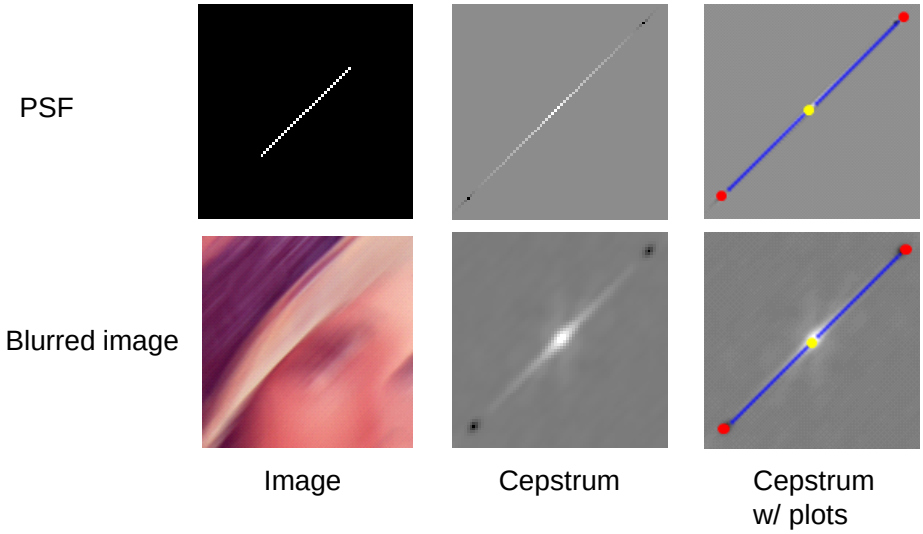


Figure 3.12: Cepstral strong values along the motion direction. A LM PSF of $(\theta, L) = (45, 30)$ (top) and blurred image degraded by the PSF (bottom). From left to right, image, cepstrum, and plotted values are shown. In the right figure, yellow point represents the cepstral peak, red points represent the negative valleys, and blue lines are overlaid PSF shape, respectively.

PSF estimation can be formulated by path integral equation as

$$\hat{\mathrm{L}} = \underset{\mathrm{L}}{\operatorname{argmax}} \int_{\mathrm{L}} C_g(\mathbf{p}) ds, \tag{3.8}$$

$$\hat{k}^L(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \hat{\mathrm{L}} \\ 0 & \text{otherwise} \end{cases},$$

where L denote a set of lines, each of which connects the cepstral peak and one of negative valleys and ds denotes an elementary arc length of lines. The estimated line $\hat{\mathrm{L}}$ has same shape as the PSF and also the position of the corresponding valley tells us the parameters of the PSF. This LM PSF estimation method can be regarded as a special case of LM PSF estimation based on Radon transform described in Sec. 3.5.1. Radon transform based method integrates the cepstrum over all directions and choose a direction having most strong value while this path integral based method integrates the cepstrum only for a few directions.

Then, same investigation is performed to NLM PSF. Figure 3.13 compares the

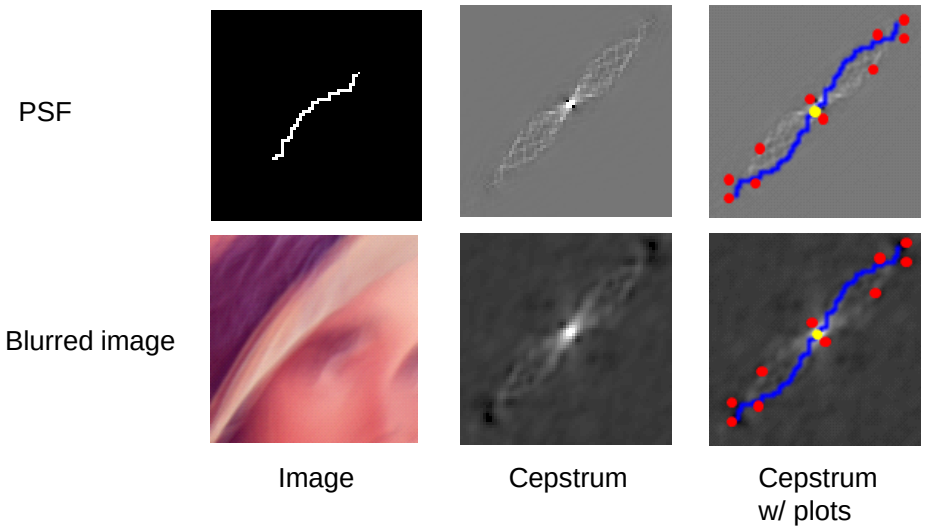


Figure 3.13: Cepstral strong values along the motion direction. A NLM PSF (top) and blurred image degraded by the PSF (bottom). From left to right, image, cepstrum, and plotted values are shown. In the right figure, yellow point represents the cepstral peak, red points represent the negative valleys, and blue curves are overlaid PSF shape, respectively.

cepstrum of a NLM PSF and that of the blurred image. On the right figure, PSF shape (blue curves) is overlaid on the cepstrum. On the cepstrum, strong values appear like Mobius strip. Comparing the cepstrum with the blue curves, strong values of the cepstrum of NLM seems to approximate PSF shape. Contrast to the cepstrum of LM PSF, the cepstrum of NLM PSF has distributed negative valleys.

NLM PSF candidates estimation based on path integral

From the discussion above, I derive an assumption that the cepstrum of a NLM PSF has approximate shape of the motion that lies between the cepstral peak and correct negative valley. If this assumption is correct, we can estimate NLM PSF by finding a curve, not a line, that maximizes the path integral between the cepstral peak and correct negative valley. However, there is no solution that finds correct negative valley from distributed negative valleys. Alternatively, the proposed method estimates a curve for each distributed negative valley as PSF candidate. Considering the cepstral symmetric property, symmetric PSFs of estimated curves

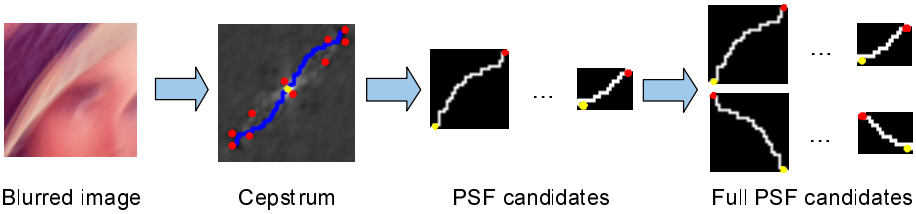


Figure 3.14: NLM PSF candidates estimation. Given a blurred image, the method first computes the cepstrum and detects several negative peaks. Then, the proposed method estimates a curve connecting the cepstral peak (yellow point) and one of negative valleys (red point) for each valley as PSF candidate.

are also added as PSF candidates. Figure 3.15 briefly explains the PSF candidates estimation method.

Here, I re-formulate the path integral based PSF estimation Eq. (3.8) for NLM PSF. Since the NLM PSF estimation method is based on path integral, C_f component and cepstral vertical/horizontal lines of image discontinuities may violate the integral. Same as PLM PSF estimation, the method takes averaging cepstra of sub images and apply Tukey window [Tukey, 1967]. Then, the method detects M negative valleys from the computed cepstrum. The number of negative valleys M is empirically decided 10.

The path integral equation Eq. (3.8) is re-formulated for NLM PSFs as

$$\hat{\mathrm{C}}_m = \underset{\mathrm{C}_m}{\operatorname{argmax}} \int_{\mathrm{C}_m} C_g(\mathbf{p}) ds, \tag{3.9}$$

$$\hat{k}_m^{\text{int}}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \hat{\mathrm{C}}_m \\ 0 & \text{otherwise} \end{cases},$$

where C_m denotes a curve connecting the cepstral peak and m the negative valley.

Now, the problem to be solved is finding a curve maximizing path integral given the cepstral peak $\mathbf{p}^{\text{peak}}$ and a negative valley $\mathbf{p}_m^{\text{valley}}$. Regarding this problem as a kind of shortest path searching problem, the proposed method solves this path searching problem by dynamic programming. Specifically, the Dijkstra's algorithm is performed [Dijkstra, 1959]. Figure 3.15 shows the strategy of the method. Given the cepstrum of a blurred image C_g, we assign the value of C_g

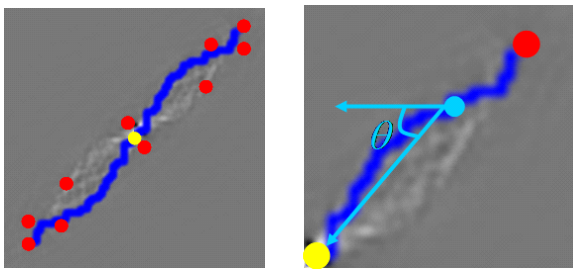


Figure 3.15: Path integral on the cepstrum of a NLM. In the figure, yellow point represents the cepstral peak, red points represent the negative valleys, and blue curves are overlaid PSF shape, respectively.

as a graph where each position $\mathbf{p} = (p, q)^{\mathrm{T}}$ corresponds to a node, and an edge is defined to connect three neighboring nodes $\mathbf{p}_j^{\text{nei}} \in \{(p, q+1), (p+1, q), (p+1, q+1)\}$. We define a weight over the nodes as

$$w\left(\mathbf{p}_j^{\text{nei}}\right) = \frac{\langle \mathbf{p}_j^{\text{vec}}, \mathbf{p}_m^{\text{vec}} \rangle}{\|\mathbf{p}_m^{\text{vec}}\| \cdot \|\mathbf{p}_j^{\text{vec}}\|}, \tag{3.10}$$

where $\mathbf{p}_m^{\text{vec}} = \overrightarrow{\mathbf{p}^{\text{peak}}\mathbf{p}_m^{\text{valley}}}$ is vector connecting the cepstral peak and the negative valley and $\mathbf{p}_j^{\text{vec}} = \overrightarrow{\mathbf{p}^{\text{peak}}\mathbf{p}_j^{\text{nei}}}$ is vector connecting the cepstral peak and a neighboring node $\mathbf{p}_j^{\text{nei}}$. The cost function represents the cosine of the angle formed by $\mathbf{p}_m^{\text{valley}}$, $\mathbf{p}^{\text{peak}}$, and $\mathbf{p}^{\text{nei}}$. Since the cost function enforces a path to connect the peak and the valley with shorter length, the estimated path tends to be a straight line rather than a zigzag line. After taking integral from $\mathbf{p}_m^{\text{valley}}$ to $\mathbf{p}^{\text{peak}}$, we find a path maximizing the path integral between the cepstral peak and the negative valley. This path estimation is performed for all detected negative valleys, thus we obtain M curves each of which corresponds to each negative valley.

Since a cepstrum is symmetric about the cepstral peak, the symmetric shape of an estimated path may be another candidate of the correct PSF. Thus, we regard both M estimated curves $\hat{k}^{\text{int}}$ and their symmetric curves $\hat{k}^{\text{sym}}$ as PSF candidates. Totally, we obtain $N = 2M$ PSF candidates $\hat{k}^{\text{can}} = \{\hat{k}_n^{\text{can}} \mid n = 1, \ldots, N\}$ from M negative valleys.

3.6 PSF candidates evaluation

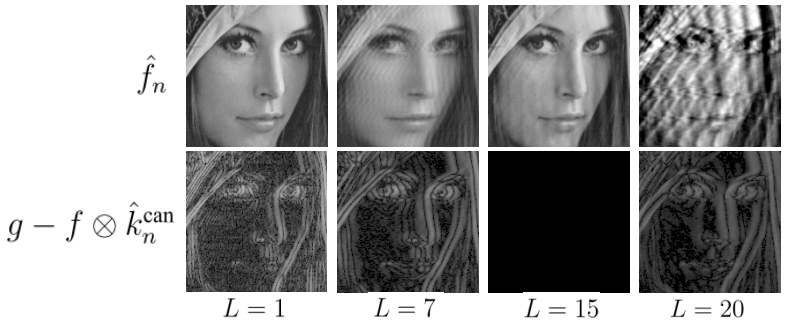
This section describes PSF candidates evaluation method based on the imaging equation. Regarding the estimated PSF candidates as PSFs with different parameters, the proposed method evaluates the candidates similar to the parameter searching methods, mentioned in Sec. 2.5.3, as

$$\hat{k} = \underset{\hat{k}_n^{\text{can}}}{\operatorname{argmin}} |Q(g - f \otimes \hat{k}_n^{\text{can}})|. \tag{3.11}$$

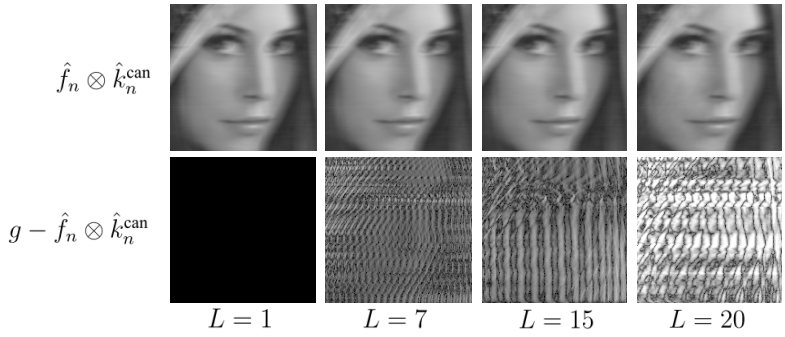
This evaluation should theoretically choose the best PSF among the candidates but in practice it does not work because Eq. (3.11) uses unknown latent image f. Alternatively, the proposed method uses a recovered image $\hat{f}_n$, which is the deconvolution result with a PSF candidate k_n^{can}, with a regularization term on PSF R_k as

$$\hat{k} = \underset{\hat{k}_n^{\text{can}}}{\operatorname{argmin}} |Q(g - \hat{f}_n \otimes \hat{k}_n^{\text{can}}) + \lambda_k R_k(\hat{k}_n^{\text{can}})|. \tag{3.12}$$

One may doubt that Eq. (3.12) can choose the correct PSF candidates because the fidelity term $Q(g - \hat{f}_n \otimes \hat{k}_n^{\text{can}})$ seems to be zero for any PSF candidate $\hat{k}_n^{\text{can}}$. However, it works. Since any deconvolution algorithm is pseudo inverse of convolution operation, recovered image has deconvolution error. Thus, the fidelity term $Q(g - \hat{f}_n \otimes \hat{k}_n^{\text{can}})$ is not zero. Figure 3.16 shows how the fidelity terms work on linear motion blur case. In this case, horizontal linear motion blur of $L = 15$ is applied to 'Lena' image. To the blurred image, Wiener filtering with different PSFs with different motion length parameters L are applied. Figure 3.16(a) shows restored images $\hat{f}_n$ and ideal fidelity term $g - f \otimes \hat{k}_n^{\text{can}}$ and Fig. 3.16(b) shows re-blurred images $\hat{f}_n \otimes \hat{k}_n^{\text{can}}$ and practical fidelity term $g - \hat{f}_n \otimes \hat{k}_n^{\text{can}}$. From left to right, PSF used for deconvolution differs its motion length L from 1, 7, 15, and 20. In ideal case, correct case $L = 15$ results zero error while other wrong cases contain some errors. On the other hand, practical case results zero error with $L = 1$ case but some errors with other cases. Even with correct PSF $L = 15$ case, residual of re-blurred image is not zero. Figure 3.17 plots the error value change w.r.t. motion length parameters. Blue plot validates that the ideal fidelity term can choose the correct PSF. On the other hand, pink plot shows interesting observa-



(a) Ideal case. (Top) Restored images $\hat{f}_n$. (Bottom) Ideal fidelity term $g - f \otimes \hat{k}_n^{\text{can}}$.



(b) Practical case. (Top) Re-blurred image $\hat{f}_n \otimes \hat{k}_n^{\text{can}}$. (Bottom) Practical fidelity term $g - \hat{f}_n \otimes \hat{k}_n^{\text{can}}$.

Figure 3.16: Example of PSF candidates evaluation with a case of horizontal linear motion blur of $L = 15$.

tions. Error value globally increases with the increase of L, especially the increase is accelerated over $L > 15$. However, there exists local minimum around $L = 15$. Considering these two observations, correct PSF seems to be estimated by finding such local minimum. Thus, the regularization term of Eq. (3.12) should be designed to find such local minimum against to the global behavior of the fidelity term.

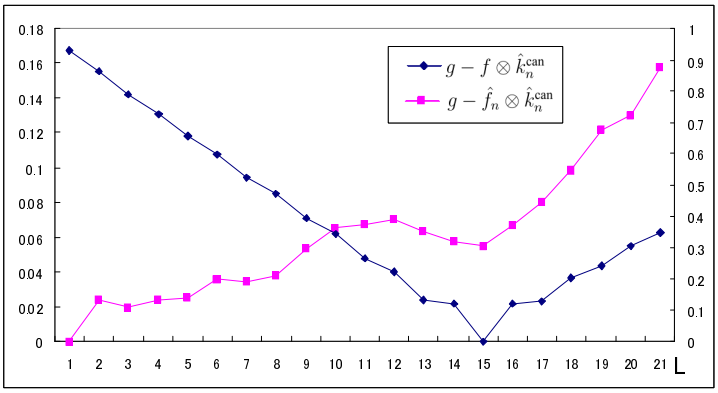


Figure 3.17: Plot of the fidelity term according to L. (Blue) Fidelity term values of ideal case. (Pink) Fidelity term values of practical case.

3.6.1 Data fidelity term

As mentioned in Sec. 2.5.3, the data fidelity term evaluates the mean square error as

$$Q(g - \hat{f}_n \otimes \hat{k}_n^{\text{can}}) = \sum_{\mathbf{x}} \sqrt{|g - \hat{f}_n \otimes \hat{k}_n^{\text{can}}|^2}. \tag{3.13}$$

Following Savakis and Trussell [Savakis and Trussell, 1993], Wiener filtering [Wiener, 1949] is used to compute recovered image $\hat{f}_n$.

3.6.2 Regularization term

As mentioned above, regularization term should be designed to find such local minimum against to the global behavior of the fidelity term. Thus, the regularization term considers the difference of neighboring motion direction parameters as

$$R_k\left(\hat{k}_n^{\text{can}}\right) = \frac{\text{Width}^2\left(\hat{k}_n^{\text{can}}\right) + \text{Height}^2\left(\hat{k}_n^{\text{can}}\right)}{\left|\left|\hat{k}_n^{\text{can}}\right|\right|^0} \tag{3.14}$$

where Width and Height functions return PSF's width and height, respectively, and L0 norm of a candidate $\left|\left|\hat{k}_n^{\text{can}}\right|\right|^0$ counts the number of non-zero components. Denominator works only for NLM to favor a straight line rather than a zigzag line.

3.7 Experimental results

This section validates the proposed method by using both synthesized images and real world images. Synthesized images are used to evaluate the performance of both each process and entire process of the proposed method. Then, the proposed method is applied to real world images. Furthermore, comparisons with other blind deconvolution methods are shown. All experiments are done Intel ® Core ™ i7 Quad CPU 3.20GHz and 6GB RAM.

3.7.1 Synthetic experiments

In these synthetic experiments, a set of 200 images from Berkeley Segmentation Dataset [Ber] is used as latent images and randomly generated PSFs is used to synthesize blurred images. This section first assesses sub processes of the proposed method, which are PSF candidate estimation process and PSF candidates evaluation process. Next, the entire method is evaluated with synthesized images. Since the synthesized images are noise-free, Wiener filtering [Wiener, 1949] is used as a deconvolution algorithm.

Synthetic experiments for PLM PSF

Here, the quantitative evaluation of NLM PSF estimation is shown. To generate random NLM PSFs, the experiments are done with the following conditions: number of component PSFs $N = \{2, 3, 4\}$; potential range of motion length of component PSF $L = \{10 \sim 30, 30 \sim 50, 50 \sim 70\}$ [pixel]; and minimum angle of two different motion direction $\Delta\theta = \{10, 20, 30\}$ [deg.]. For example, the case $(N, L, \Delta\theta) = (3, 10 \sim 30, 20)$ generates a PLM PSF consists of three component PSFs, motion length of each of them ranges $10 \sim 30$ pixels, and each motion direction parameter θ is distributed at least 20 degrees.

PLM PSF candidates estimation

First experiment is validation of Radon transform based motion directions estimation under the conditions mentioned above. Here, success case is defined to satisfy that the number of maximum of Radon transform is equivalent to N, that estimation error of θ is less than 5 degrees. Table 3.1 shows the number of success

Table 3.1: PLM PSF candidates estimation: The number of success cases.

$\Delta\theta$	10			20			30		
L \ N	2	3	4	2	3	4	2	3	4
$10 \sim 30$	110	50	12	114	58	20	180	145	82
$30 \sim 50$	159	92	75	161	117	87	199	194	189
$50 \sim 70$	173	123	84	164	115	82	200	200	193

Table 3.2: PLM PSF candidates estimation: RMSE of estimated motion directions [deg.]

$\Delta\theta$	10			20			30		
L \ N	2	3	4	2	3	4	2	3	4
$10 \sim 30$	0.5	0.5	0.4	0.4	0.5	0.6	0.3	0.4	0.4
$30 \sim 50$	0.1	0.1	0.1	0.0	0.0	0.1	0.0	0.0	0.0
$50 \sim 70$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

cases of 200 trials under different conditions. Almost failure cases are undetected cases. The table shows that both smaller $\Delta\theta$ and bigger N have more failure cases. This indicates that these conditions make Radon transform disable to localize two neighboring maxima. On the other hand, bigger L have more success cases. This indicates that bigger L clarifies the corresponding maximum of Radon transform. Next, Table 3.2 computes the Root Mean Square Error (RMSE) of estimated motion directions only for success cases. In all cases, RMSE is less than 1 degree. Especially, bigger L cases provide better results than smaller cases. Table 3.3 computes the RMSE of estimated motion length parameters only for success cases. Contrast to motion directions estimation, results differs according to the conditions. Especially, the condition $(N, L) = (4, 10 \sim 30)$ marks relatively worse results. So far, I have not found arrived at a clear conclusion for this results. However, I guess that each negative spike interferes each other, thus the results are relatively bad.

Table 3.3: PLM PSF candidates estimation: RMSE of estimated motion length parameters [pixel]

$\Delta\theta$	10			20			30		
L \ N	2	3	4	2	3	4	2	3	4
$10 \sim 30$	1.6	5.5	9.1	2.3	5.6	8.5	3.0	8.0	12.0
$30 \sim 50$	1.7	2.4	3.2	1.5	2.2	2.7	1.9	3.3	4.8
$50 \sim 70$	1.9	2.3	3.2	1.9	2.1	2.5	1.4	2.0	2.8

Table 3.4: PLM PSF candidates evaluation: The number of success cases.

$\Delta\theta$	10			20			30		
L \ N	2	3	4	2	3	4	2	3	4
$10 \sim 30$	189	170	149	189	166	146	191	170	156
$30 \sim 50$	173	115	83	163	128	95	176	127	79
$50 \sim 70$	134	70	38	139	94	41	143	87	75

PLM PSF candidates evaluation

Next, candidates evaluation process is validated. In this experiment, correct component PSFs are used to validate the evaluation function chooses the correct PLM PSF. Thus, Table 3.4 just counts the number of success cases of 200 trials for each condition. The result indicates that larger N cases and larger L cases seem to fail. Since larger size PSF causes more ringing artifacts even though the PSF is correct. Therefore, above conditions have more failure cases.

PLM PSF estimation

Finally, entire method is validated. Considering the above experiments, $\Delta\theta$ is set to 30 degree. Here, the success case is defined to satisfy that blur direction error is less than five degree. Table 3.5 shows the number of success cases, RMSE of motion directions, and RMSE of motion length parameters. The number of success cases differs according to the conditions more than previous experiments. In the previous experiments, larger PSF size is better for candidates estimation while smaller PSF size is better for candidates evaluation. In this experiment, the case of

Table 3.5: PLM PSF estimation: The number of success cases, RMSE of motion directions [deg.], and RMSE of motion length parameters [pixel].

	#success			RMSE of $\hat{\theta}$ [deg.]			RMSE of $\hat{L}$ [pixel]		
L \ N	2	3	4	2	3	4	2	3	4
$10 \sim 30$	150	63	12	0.3	0.6	1.5	1.1	1.5	1.1
$30 \sim 50$	157	92	71	0.0	0.0	0.0	0.9	1.8	2.1
$50 \sim 70$	136	76	49	0.0	0.0	0.0	1.5	1.8	2.7

$L = 30 \sim 50$ results most success cases. Calculated RMSE of success cases are that motion directions $\hat{\theta}$ is less than 1.5 degree and motion length parameters $\hat{L}$ is less than 3 pixels. For an image of 320×240 resolution, the proposed method takes 0.1 sec for PSF candidates estimation and $0.15N!$ sec for PSF candidates evaluation, *e.g.*, $0.1 + 0.15 \times 3! = 1.0$ sec for a PLM consisting of three linear motions.

Synthetic experiments for NLM PSF

Since NLM PSF does not parametric form, quantitative evaluation is not easy. To evaluate the accuracy of estimated PSFs, we compare a restored image using estimated PSF with one using the ground truth PSF not with the latent image because deconvolution algorithms cannot perfectly recover the latent image even with the ground truth PSF.

PSF shape estimation process

I first evaluate the PSF shape estimation process. The assumption is that the cepstrum of a NLM PSF has unclear PSF shape that lies between the positive peak and one of negative valleys. To validate the assumption, I input a correct negative valley, corresponding to a ground truth PSF, to Eq. (3.9) so that the process can ideally estimate the correct PSF. Since it is not easy to evaluate an estimated PSF by its shape, we compare the restored image obtained by the estimated PSF $\hat{f}$ with that by ground truth $\hat{f}_{\text{best}}$. Figure 3.18 shows the histogram of NCC of the restored images $\hat{f}$ and $\hat{f}_{\text{best}}$ w.r.t. PSF size. It is natural that the process recovers higher NCC images for smaller size of PSFs. Empirically, NCCs below 0.9 are visually

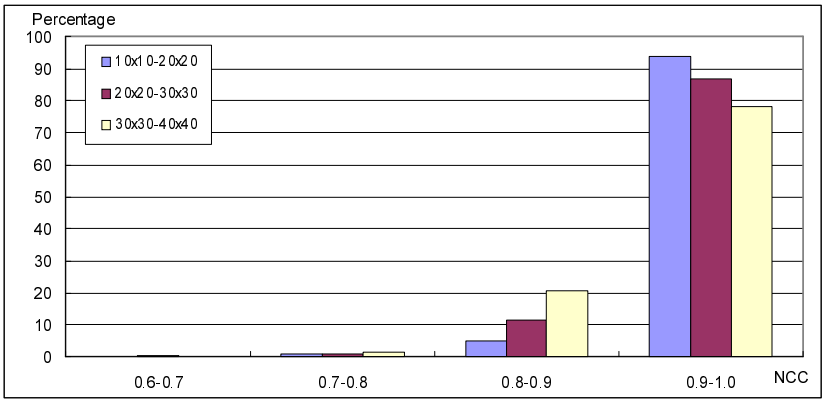


Figure 3.18: Experimental result: NCC histogram of the restored images $\hat{f}$ and $\hat{f}_{\text{best}}$ w.r.t. PSF size.

unacceptable. With this threshold, more than 70 percent of the trials are successful for all the sizes of PSFs. Figure 3.19 shows some of the restored images. From left to right, the size of ground truth PSF is increasing. Red framed figures in $\hat{f}$ show the estimated PSFs while that in $\hat{f}_{\text{best}}$ show the ground true PSFs. In both PENGUIN and FISH cases, PSF is well-estimated, thus the restored images by estimated PSF recover the detail of the latent images, *e.g.* penguin's fur skin and fish skin. The Goat case shows that the restored image is slightly damaged by even the ground truth PSF because of the bigger PSF size. In such case, the estimated PSF is not perfectly same as the ground truth. As a result, the ringing artifacts in $\hat{f}$ is more obvious than that in $\hat{f}_{\text{best}}$.

PSF candidates evaluation

Next, I evaluate the performance of the PSF candidates evaluation process. I synthesize 200 blurred images and prepare 10 PSF candidates, containing the ground truth and others are wrong. For validation, estimate PSF is evaluated using the objective function Eq. (3.12). Table 3.6 shows the number of success of 200 blurred images w.r.t. varying λ and varying PSF size. The cases $\lambda = 0$ denote that the objective function evaluates the PSF candidates with only error term. The results show that the objective function could successfully choose the ground truth PSF more than 80 percent of the trials without the regularization term. All the failure cases of $\lambda = 0$ chose the smaller size PSF than the ground truth. This result

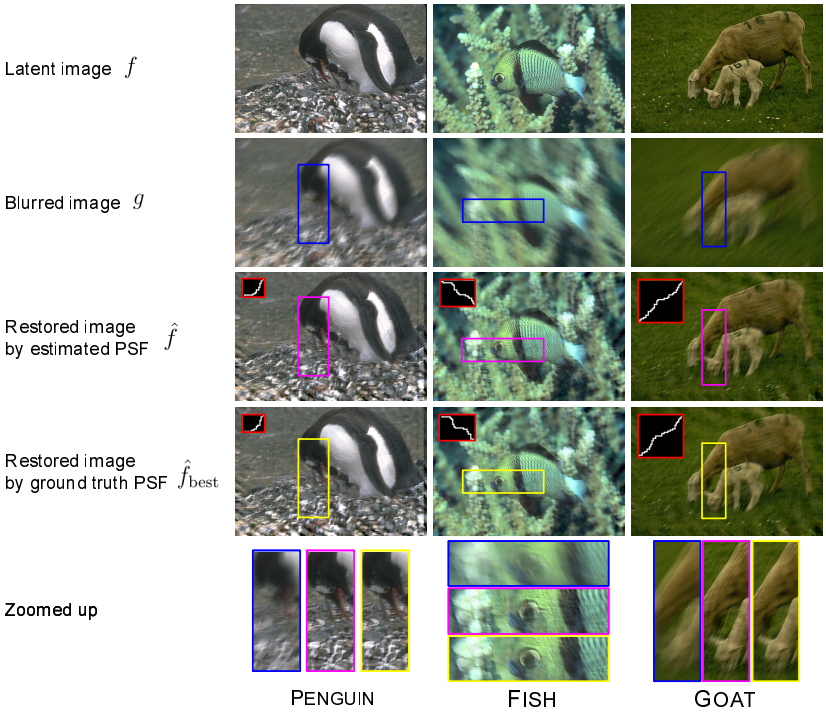


Figure 3.19: Examples of the PSF shape estimation experiment. From top to bottom, latent images, blurred images, restored images by estimated PSFs, restored images by ground truth PSFs, and zoomed up of the restored images are shown. Red framed figures in restored images are PSFs used for deconvolution (for better visualization, we enlarge the PSFs 3 times the normal size). From left to right, PSF size is increasing.

indicates that the error term can discriminate the ground truth PSF, however, the ringing artifacts caused by PSF size degrades the performance of the error term. With $\lambda = 5 \sim 15 \times 10^{-5}$, the objective function provides better results than that of $\lambda = 0$. However, the cases $\lambda > 15 \times 10^{-5}$ provide worse results according to λ. This result indicates that the regularization term assists the error term for the evaluation but relatively bigger λ makes the regularization term dominant in the evaluation function. In this experiment, the result has less correlation with PSF size. Thus, we expect that the objective function works invariant to PSF size with optimum λ value. In the latter experiments, λ is set to $5 \sim 20 \times 10^{-5}$.

Table 3.6: Experimental results of PSF candidates evaluation w.r.t. varying PSF size and λ of Eq. (3.12).

λ ($\times 10^{-5}$)	0	5	10	15	20	25	30	35	40	45
10×10∼20×20	176	192	191	186	179	172	160	152	142	131
20×20∼30×30	181	188	186	179	173	166	157	148	142	125
30×30∼40×40	174	179	178	176	169	165	152	147	137	127

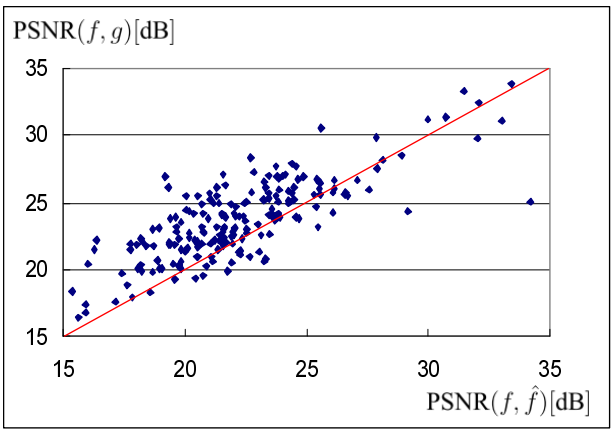


Figure 3.20: Experimental results: The plots of PSNR ratio. Red line denotes PSNR ratio PSNR(f,$\hat{f}$)/PSNR(f,g) equals 1. The ratio greater than 1 indicates that the restored image $\hat{f}$ is closer to the original image f than the blurred image g.

Performance of entire method

Here, I validate the performance of the entire method by using 200 blurred images. PSF size is set to 10×10∼20×20 pixels. For each image, we compute the Peak Signal-to-Noise Ratio (PSNR) of the blurred image and the one of the restored image and compare them. The case that PSNR(f,$\hat{f}$) is greater than PSNR(f,g) represents that the restored image $\hat{f}$ is closer to the latent image f than the blurred image g. Figure 3.20 plots PSNR ratio PSNR(f,$\hat{f}$)/PSNR(f,g) of the results and a red line in the figure represents the ratio equals 1. In the experiment, 159 of 200 cases result PSNR ratio greater than 1. Figure 3.21 shows some of the success cases. Each caption of the blurred images and one of the restored images by estimated PSF represent PSNR(f,g) and PSNR(f,$\hat{f}$), respectively. The MOUNTAIN

case is an example of higher PSNR(f,g) cases, which denotes less blurred case. Zoomed up figures show that wood area is well-recovered. Middle column shows the result of the WOMAN case. The restored image by estimated PSF is slightly damaged by more ringing artifacts than that by ground truth PSF, however, the detail of the image, *i.e.*, hair and eye, are recovered. Lower PSNR(f,g) case, the SHIP case, is severely blurred than other two examples. The estimated PSF is not perfectly same as the ground truth, however, the text on the ship VIKING LINE gets much better than the blurred image. For an image of 320×240 resolution, NLM PSF estimation method takes 0.001 sec for PSF candidates estimation and $0.15N$ sec for PSF candidates evaluation, *e.g.*, $0.001 + 0.15 \times 10 \approx 1.5$ sec for a NLM.

3.7.2 Real-world experiments

In the real world experiment, I compare the proposed method with a maximum likelihood algorithm [MathWorks] (Matlab's deconvblind function) and a Bayesian approach [Fergus et al., 2006] based on variational Bayes estimation [Miskin and MacKay, 2000] to validate the proposed method. To deal with the noise effect on blurred images, I use a regularized minimization deconvolution method [Levin et al., 2007].

Real-world experiments for PLM PSF

For PLM PSF estimation, following scenes are used: TREE (natural image, smaller blur), FLOWER (natural image, larger blur), DOLL (artificial object, larger blur), TEXT (text image, larger blur). Figure 3.22 shows blurred images, estimated PSFs, and recovered image using Wiener filter [Wiener, 1949]. The number of component PSFs are 2 for TREE, FLOWER, and DOLL scenes and 4 for TEXT scene. In the restored images of TREE and FLOWER scenes, leaves of tree and flowers are well-recovered without ringing artifacts, thus we can say that estimated PSF is correct. On the other hand, DOLL and TEXT scenes have little ringing artifacts. This indicates that the estimated PSF is not perfect, however, text in both scenes recovered enough to read. Therefore, the proposed method can estimate PLM PSF from various types of scenes.

Next, the proposed method is compared with the conventional methods, [Fergus et al., 2006] and deconvblind. Figure 3.23 shows the results. Fergus' method recover TREE and FLOWER scenes with less artifacts but fail on DOLL and TEXT scenes. deconvblind estimates obscure PSFs and recovered images are visually unacceptable.

Real-world experiments for NLM PSF

Figure 3.24 shows the results of four scenes, DOLL (natural image), ORANGE (artificial object), SIGN BOARD (natural image), and TEXT scene. The red frame in a restored image shows the estimated PSF. The caption of the restored image denotes size of the estimated PSF. The scenes DOLL and ORANGE are selected as examples of natural images. For such scenes, both our method and Fergus' method recover clearer images, *i.e.* doll's eye and the text in ORANGE scene, while deconvblind provides the restored images damaged by heavy ringing artifacts. The other scenes are selected as examples of less-textured scenes: text pattern in natural scene SIGN BOARD and text-pattern only scene TEXT. For SIGN BOARD scene, both proposed method and Fergus' method can recover satisfying quality images. The reason why Fergus' method can recover the satisfying image is that the background area of text part in the blurred image obeys the statistics of natural images in the case of text-pattern in natural scene. Contrast to above scenes, TEXT scene has only text component in the image. For the scene, our method recovers readable text even with ringing artifacts, while deconvblind and Fergus' method cannot recover clearly readable images. These results indicate that our method can estimate PSFs for various scenes.

3.8 Conclusion

The motivation of this work is to search for the answer to the question that *is it possible that the classic approach estimates non-linear motion PSF*. For this question, I tackle non-linear motion PSF estimation issues with cepstral approach. To achieve the purpose, I analyzed the cepstral behavior of two types of constant speed non-linear motion PSF, *e.g.*, PLM PSF and NLM PSF. Based on the anal-

ysis, the proposed method estimates non-linear PSF from a single blurred image. Experimental results show that the proposed method can estimate non-linear PSF under the condition that our assumptions are held.

3.8.1 Future direction

There are several future directions. One is the consistency of the proposed method. Now, the proposed method does not classify PLM and NLM because the method uses different types of cepstral features of PSFs. Thus, all experiments were performed separately. This style lacks of practicability. To develop a unique solution handling both type of motion, one solution is to merge both types of PSF such as a Piecewise Noisy Linear Motion PSF that is partially NLM. The cepstrum should have the sum of the cepstra of all NLMs but has several negative valleys which may correspond to one of the cepstrum of component NLM. To estimate each component NLM, our NLM PSF estimation is performed for all negative valleys. Since some of estimated NLMs may correspond to wrong negative valleys, we have to omit them but it is difficult to specify which valley is wrong. Thus, we first generate PSF candidates by all the permutation considering such ambiguity. For example, N component NLMs generates $N! + (N-1)!_N P_{N-1} + \cdots$ PNLM PSF candidates. Then, PSF candidates evaluation is applied to the candidates to choose the best estimate. To realize this PNLM PSF estimation, theoretical foundation of the cepstral behavior of NLM PSF is required. In this thesis, NLM PSF estimation method is designed based on assumption derived from my observation but lacks theoretical foundation. Therefore, further analysis on the cepstral behavior of NLM PSF is required.

Another issue is PSF candidate evaluation process. Regularization term of the PSF evaluation (Eq. (3.14)) considers only PSF size. This means that our regularization term considers the likelihood of unknown PSF. Similar to regularized minimization of non-blind deconvolution, mentioned in Sec. 2.4.4, can provide some information. If we can somehow find better regularization term, PSF candidate evaluation (Eq. (3.12)) can be a powerful tool for any blind deconvolution works because they can evaluate PSF with different parameters of their methods.

Another future direction is to extend cepstral approach as to estimate spatially

varying PSFs from a single blurred image. First step is discrete PSFs estimation. Since we don’t have prior knowledge on the scene, we cannot utilize sub regions that are suitable for PSF estimation. Thus, the discrete PSFs should be computed on image grid. Since there is no guarantee that a sub region is well-textured, we have to consider the possibility of mis-estimation. For this problem, I came up with some idea. First idea is to introduce constraint on latent image. Similar to the statistical properties of image gradient, same category scene at same depth scale follows the same distributions of power spectra at sub regions. Regarding the averaged power spectra map as reference, PSF estimation can easily be done.

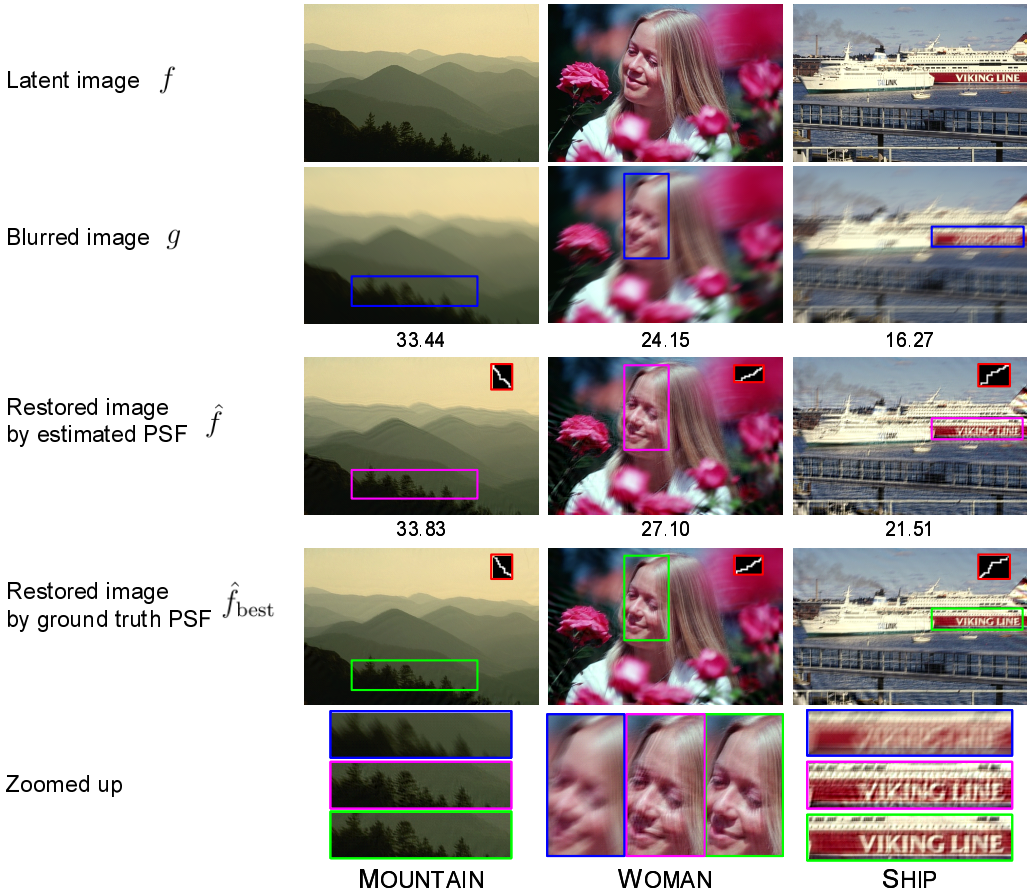


Figure 3.21: Experimental results: Restored images of the entire method. From left to right, MOUNTAIN, WOMAN, SHIP cases are shown with NCC values between the restored images. From top to bottom, original images, blurred images, restored images by estimated PSF, restored images by ground truth PSF, and zoomed up of the restored images are shown. Red framed figures in restored images are PSFs used for deconvolution (for better visualization, we enlarge the PSFs 3 times the normal size).

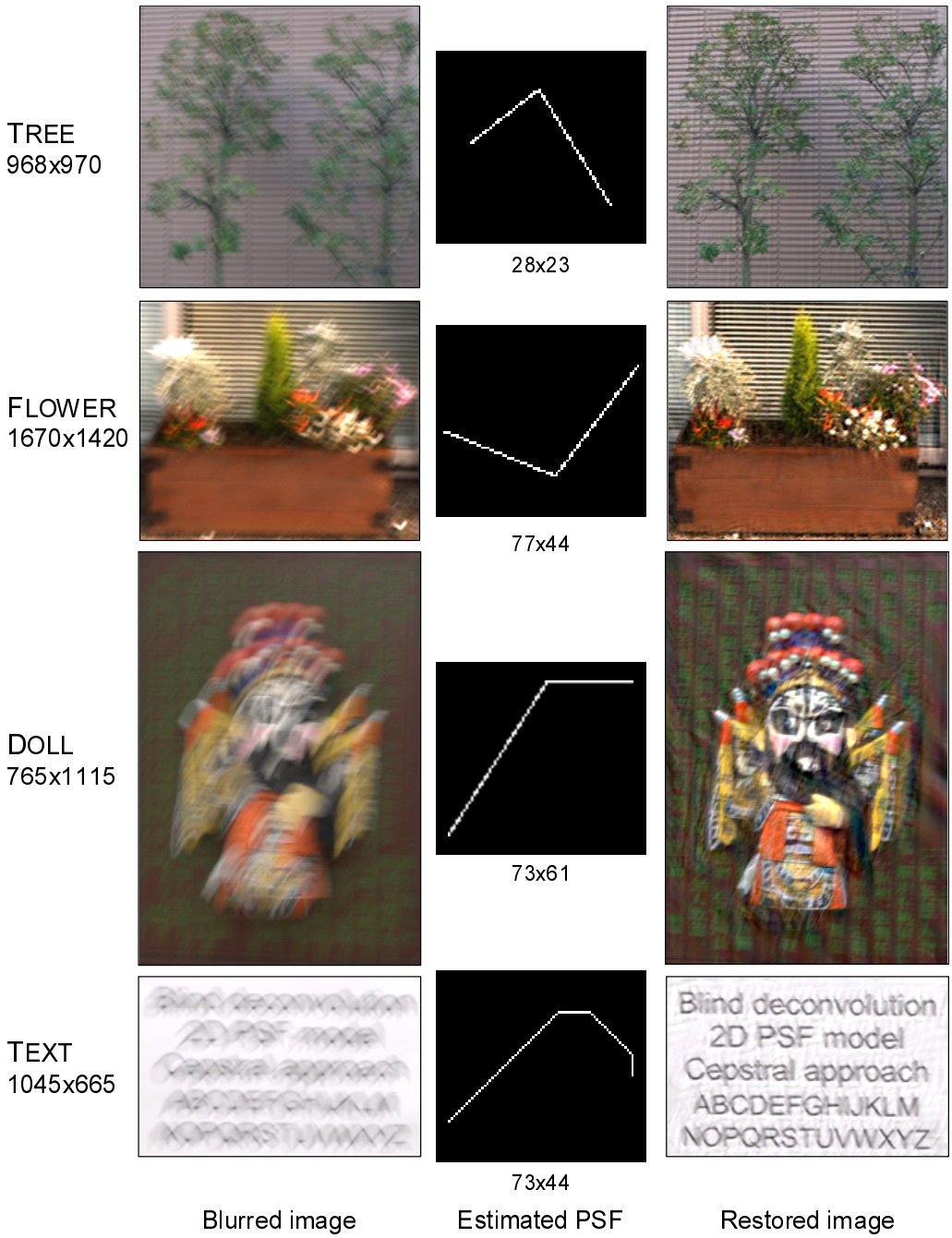


Figure 3.22: PLM PSF estimation for real world images. From top to bottom, TREE, FLOWER, DOLL, andTEXT scenes are shown. From left to right, blurred images, estimated PSFs, restored images. Captions of middle column images denotes estimated PSF size.

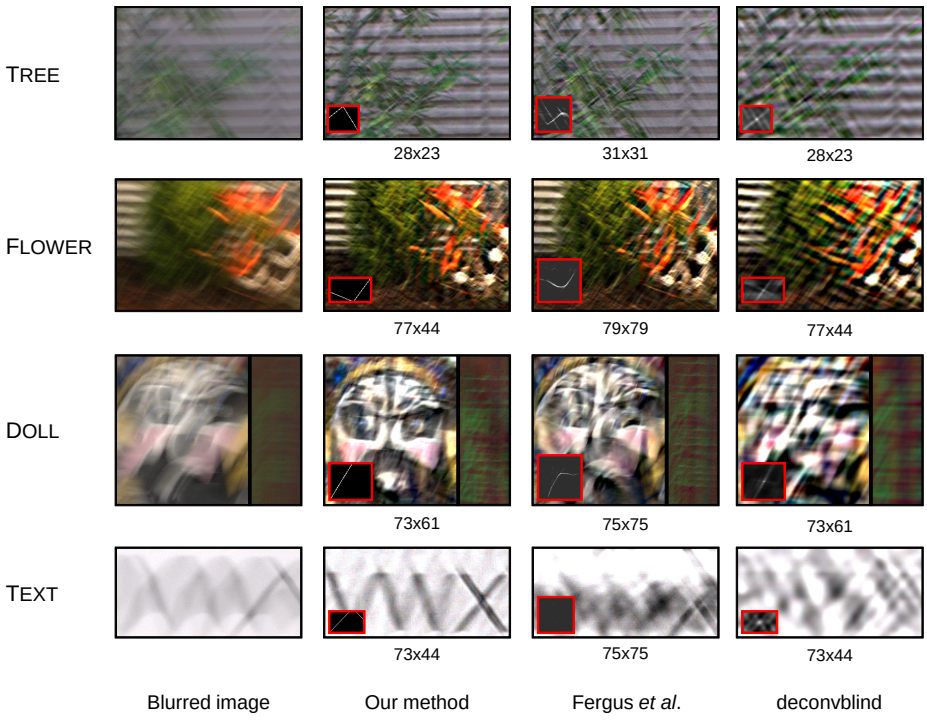


Figure 3.23: Comparison with traditional methods. From left to right, blurred images, our method, Fergus et al., and deconvblind, respectively. Red framed image in a restored image shows the estimated PSF.

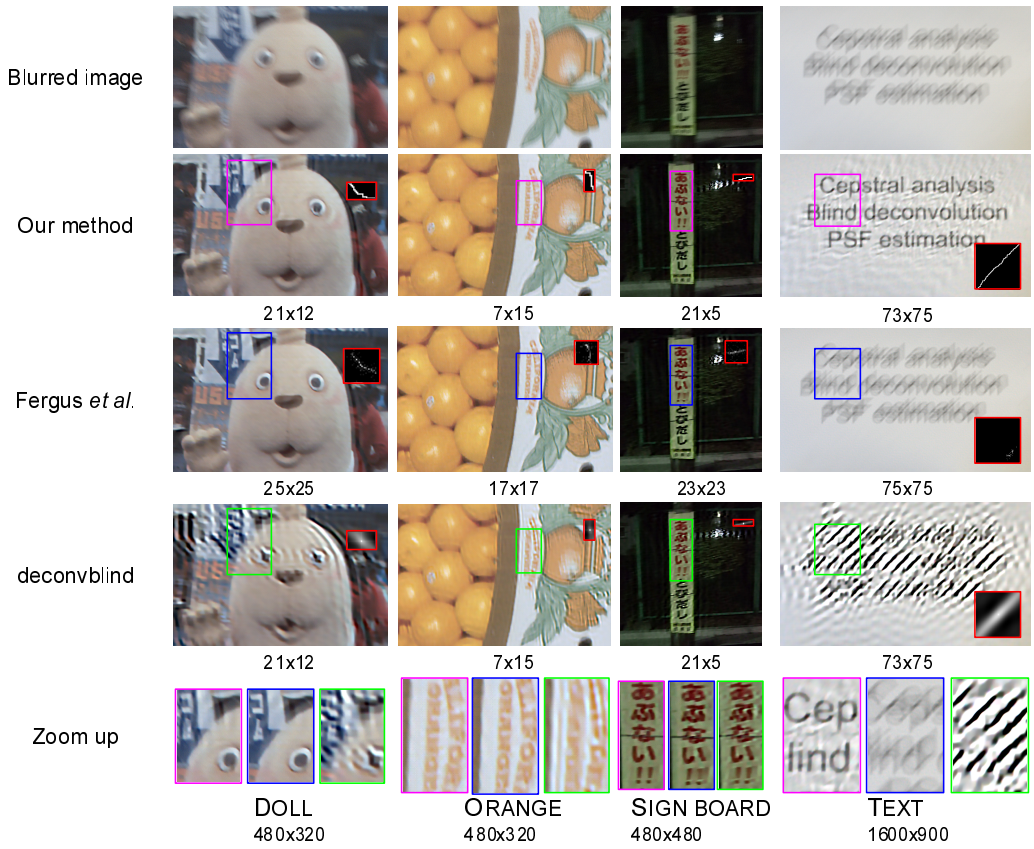


Figure 3.24: Experimental results: Restored images of the real world experiment. From left to right, DOLL, ORANGE, SIGN BOARD, and TEXT scenes are shown with the image resolution. From top to bottom, blurred images, restored images by our method, restored images by Fergus' method [Fergus et al., 2006], restored images by deconvblind [MathWorks], and Zoom up of restored images are shown. Red framed figures in restored images are PSFs used for deconvolution (for better visualization, we enlarge the PSFs 3 times the normal size) and each caption of restored images denotes the size of the estimated PSF. Other framed figures correspond to zoomed up regions of restored images.

Chapter 4

In-Focus Projection from a Single Projector-Camera Image Pair

In this chapter, I propose a method that displays an in-focus image onto off-axis surface based on a single projector-camera image pair matching algorithm as shown in Fig. 4.1. Making a pair of projector and camera, the proposed method first estimates defocus parameters on the off-axis surface. Then, generate a sharpened image that contains enhanced edge according to the estimated defocus parameters. Finally, the sharpened image is projected to the surface to cancel the defocus effect. The proposed method assumes that display surface is planar and Lambertian.

4.1 Projectors in computer vision and graphics

Thanks to the recent development of projectors, their capabilities, *e.g.*, brightness, resolution, contrast and throw-distance, have made projectors one of the popular display devices. The greatest merit of a projector is that a projector can project onto many screens of various sizes and scaled-up projection can show displayed image to many observers. Thus, projectors are known to be useful for both Computer Vision and Graphics researches.

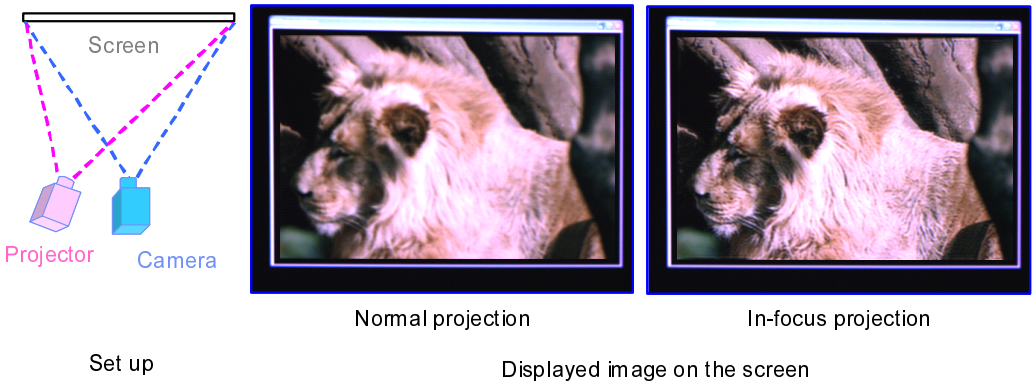


Figure 4.1: Overview of the proposed method. The propose method cancels the defocus effect appears on surface off-axis to a projector. By projecting a sharpened image, defocus effect is cancelled.

Real world measurement

With a projector, we can emit a light forming arbitrary shape. Such property is useful for real world measurement.

3D volume measurement

One famous application is active stereo that measures 3D volume of objects by the triangulation algorithm. Active stereo replaces a camera of stereo camera pair with a projector. Projecting known pattern, *e.g.*, points and marks, correspondence between the projector and the camera is easily done with the knowledge on the projected pattern.

One intuitive extension is to reduce the number of projected patterns. Kawasaki et al. use calibrated pro-cam set up for single-shot 3D reconstruction [Kawasaki et al., 2008]. For decode the projected pattern, they use grid pattern having coplanarity constraints. They further extend single-shot reconstruction system to work under uncalibrated pro-cam pair [Kawasaki et al., 2010]. To achieve single-shot system, their method does self-calibration using a projected De Bruijn grid pattern. Fernandez et al. do 3D dense reconstruction from a single-shot projection [Fernandez et al., 2010]. For uniquely decode the projected pattern, they design multiplexed color pattern. Thus, their pattern decode algorithm provides fast and reliable phase

map.

One of the big limitations of the above methods is sensitivity to objects' texture. Schmalz and Angelopoulou propose single-shot 3D reconstruction for textured objects [Schmalz and Angelopoulou, 2010]. To decode the pattern on textured surfaces, they introduce region adjacency graph. Watershed transform represents the input image with fewer color, thus the method becomes less sensitive to the noise caused by objects' texture.

Yamazaki and Xu measures 3D shape of glossy surface based on Dichromatic reflection model [Yamazaki and Xu, 2010]. The system consists of stereo camera, a projector located at closer position to the cameras, and a display located at opposite position to the cameras. Due to the positions, the cameras observe diffuse component of projector light and sum of specular and diffuse component display light. Thus, active binocular stereo measures diffuse component from projector light and shape-from-distortion measures specular component from display light. When the reflection on the surface is dominated by diffuse component, projector's projection is clearer while display's projection is blurred. On the contrary, when the surface reflection is mainly specular, display's projection gets clearer and projector's projection is blurred.

Cuypers et al. propose real-time 3D shape acquisition based on silhouette extraction, visual hull in other words [Cuypers et al., 2009]. Instead of multiple cameras, they use multiple colored point light sources. The setup consists of multiple colored point light sources, a diffuser and a digital camera. For each light source, a silhouette is extracted from the captured shadows of the scene. These silhouettes are used for visual hull reconstruction as well as image based collision detection.

Furukawa et al. measure entire 3D shape of an object using uncalibrated procam [Furukawa et al., 2009]. They estimate initial 3D shape using active stereo and compute initial estimate of extrinsic parameters. Next, feature points based rough registration is performed as the initial values of the motion parameters. Since the wrong correspondence may cause registration error, the motion parameters are refined by ICP algorithm. Finally, bundle adjustment is performed to optimize all the parameters.

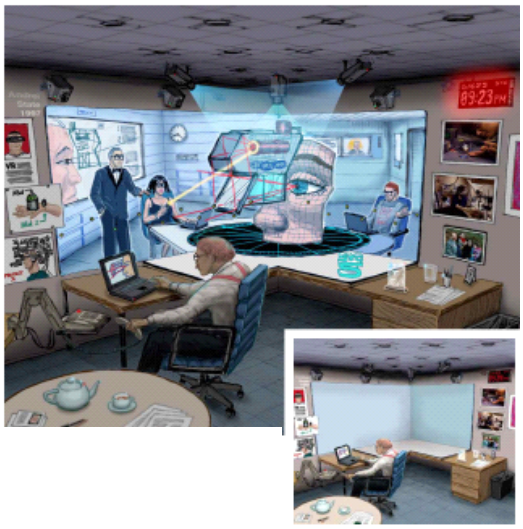


Figure 4.2: A conceptual sketch of the office of the future. Courtesy of [Raskar et al., 1998b].

Appearance control

Initially, projector is designed to use with a planar that is white and Lambertian. However, adjusting the projection enables to control the appearance of real-world object. Such appearance control technique is useful for virtual reality, mixed reality, and augmented reality works.

Immersive display

Raskar and his colleague proposed office of the future [Raskar et al., 1998a,b] as shown in Fig. 4.2. Replacing the normal office lights with projectors, we can control all over the light in the office. In the back of the system, synchronized cameras capture the visible surface of the office so that we can control the appearance of images on the surfaces.

One of the biggest limitations of front projection systems is that occluder between the projector and the surface casts shadow on the surface and that projector illuminates undesirable projection onto occluder. Using multiple projectors, we can fill the shadow region [Audet and Cooperstock, 2007; Jaynes et al., 2004; Sugaya et al., 2010]. First step detects shadow region by comparing predicted

Figure 4.3: Example of appearance control results. (a) Original appearance. (b) Color saturation enhancement. (c) Color removal. (d) Color phase control. (e) Unique brightness. (f) Edge enhance and blur. Courtesy of [Amano and Kato, 2010].

view and actual observation. Once shadow region is detected, another supplemental projector projects occluded region. Sugaya et al. proposed shadow contrasting method for multiple projector display [Sugaya et al., 2010]. Their method recognizes which projector generates the shadow on the display from a single-shot. The method is available for removing the cast shadow by occluder and undesirable projection onto occluder.

Texture control

Contrast to the above immersive display, texture control of smaller size objects is also interesting. Amano and Kato implemented appearance control methods for the visually impaired [Jefferson and Harvey, 2007; Peli et al., 1994; Wolffsohn et al., 2007] on Projector-Camera systems [Amano and Kato, 2010]. Figure 4.3 shows the examples of appearance control results of color saturation enhancement, color removal, color phase control, brightness equalization, and edge enhancement and blur. Menk and Koch proposed appearance control under the influence of ambient light [Menk and Koch, 2010]. They use spectral data recording

projector, ambient light, and object separately. The data is used to decide projection pattern by physically-based simulation. The proposed radiometric model has the constraint that the pixels of the projector are treated independently of each other. Park and Kim use a mobile projector to support large display [Park and Kim, 2010]. Colored projection boundary enables the system to track the mobile projector. By extracting user's shadow generated by projection, the user can interactively access the content. As mentioned in Ch. 1, Raskar et al. presented the idea of shader lamps where projectors are used to project additional graphical content onto a neutral and diffuse real object [Raskar et al., 2001]. Bandyopadhyay et al. then extended [Raskar et al., 2001] to project onto movable objects, which additionally could be interactively colored with a tracked paint brush [Bandyopadhyay et al., 2001].

4.2 Related works

This section briefly overview the related works. The discussion is mainly related to projector image adjustment works but not limited to in-focus projection.

As mentioned above, computer vision and graphics research fields have expanded projector's potential probability during this decade. Projector's key property enables these works is that we can control projecting light so that environmental information can be extracted by making a pair of projector(s) and camera(s). For the upcoming imaging technologies, what kind of pro-cam applications are favored? Considering the number of intended people, projector image adjustment technique is the one favored. The release of books well-summarizing the related technologies indicates that my guess is not bad direction. Bimber and Raskar mention the projector based augmented reality technologies in [Bimber and Raskar, 2005] while Majumder and Brown mention the development of multi-projector display in [Majumder and Brown, 2007].

Generally speaking, we, projector-camera researchers, regard the view of camera is equivalent to user's view. The relationship between a projector image p and a camera image c is described as

$$c(\mathbf{x}) = D(p(\mathbf{x}), \Theta), \tag{4.1}$$

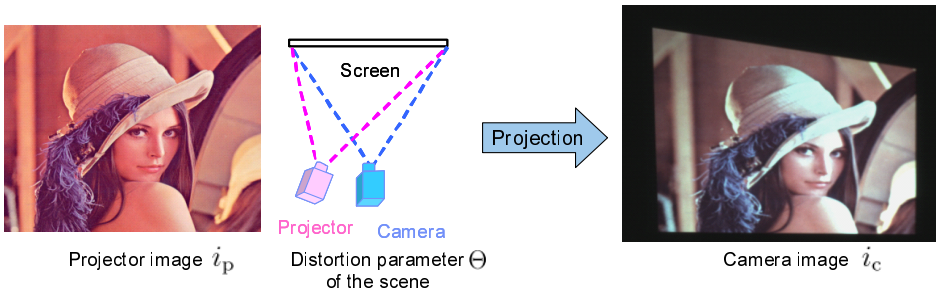


Figure 4.4: The relationship between a projector image and camera image.

where D represents any image distortion on the projector image and Θ is a set of distortion parameters. Figure 4.4 shows the relationship. The form of distortion function D varies with considered image distortion and parameter Θ depends on the scene. The goal of projector image adjustment is to fit the projected image onto any surface so that we can show what we want to show, namely distortion free image. For this purpose, two technical problems should be solved. One is how to synthesize such adjusted image and the other is how to estimate the distortion parameter Θ for image synthesis. In this section, I give the brief overview of the literature for each problem.

4.2.1 Adjusted image synthesis

First problem to be solved is adjusted image synthesis. Suppose we expect the camera to observe c. With known distortion parameter set Θ, adjusted image to be projected is synthesized as

$$p(\mathbf{x}) = D^{-1}(c(\mathbf{x}), \Theta). \tag{4.2}$$

Note that the equation implicitly assumes that the function D is invertible. Roughly speaking, there are two approaches for the synthesis, two pass rendering and pixel-by-pixel mapping.

Two pass rendering

Two pass rendering was first introduced by Raskar et al. [Raskar et al., 1998a]. The approach assumes known scene geometry and fully calibrated projector-camera pair. In the first pass, we render the goal image i_{g} from the perspective of the camera. Next, i_{g} is mapped onto the display surface. This mapped image is generated by projective texturing that projects the goal image onto the surface. Then, the second pass renders the mapped image from the perspective of the projector. This approach is used for projection onto non-textured object [Bandyopadhyay et al., 2001; Johnson and Fuchs, 2007; Raskar et al., 2003, 2001].

Pixel-by-pixel mapping

The assumption of two pass rendering is too strong to use in everyday applications because the geometry of the scene is generally unknown. For the case of unknown geometry, pixel-by-pixel mapping is suitable approach. The approach somehow finds the correspondence between c and p so that we can precisely control the projection without the scene information.

Light transport The light transport describes all global illumination effects between a light source (projector) and an imaging device (camera) [Sen et al., 2005]. Simply speaking, the light transport matrix is a lookup table describing all the pixel correspondence of the projector-camera pair. Thus, the transport matrix is obtained once, we can synthesize any image from the projector's perspective under camera illumination [Ding et al., 2009; Wetzstein and Bimber, 2007; Yamamoto et al., 2010]. The merit of this approach is that the light transport contains all the distortion in a matrix form. Thus, both geometric and color distortion can be cancelled by applying the inverse light transport matrix. Once the light transport of between the projector and the camera is obtained using structured patterns projection, we can synthesize image from the viewpoint of the projector, the scene illuminated by a synthetic light source, *etc*.

Approximate model With light transport, we can synthesize any image of the scene. However, acquisition takes so long to know all the pixel correspondence of the projector-camera image pair. One may prefer approximate models rather than pixel level mapping under a situation that display surface is locally

uniform. Such approximated model finds local region level mapping [Majumder and Brown, 2007]. Contrast to the light transport, this approach should separately consider the distortion. For geometric distortion, piecewise planar assumption is well-used. Interested readers may refer to [Brown et al., 2005] for further information.

For color adjustment, well-used distortion function is 3×3 color mixing matrix D_{color} that maps projector image to camera image as

$$\begin{bmatrix} c^R \\ c^G \\ c^B \end{bmatrix} = D_{\text{color}} \begin{bmatrix} p^R \\ p^G \\ p^B \end{bmatrix}, \tag{4.3}$$

where R, G, B denote red, green, and blue component of an image, *e.g.*, p^R denotes red color component of projector image. Using this expression, pixel level color mapping is realized with geometrically aligned projector-camera pair [Ashdown et al., 2006; Bimber et al., 2005; Fujii et al., 2005; Grossberg et al., 2004; Grundhöfer et al., 2007; Nayar et al., 2003; Wang et al., 2005].

Focal adjustment

The above approaches implicitly assume projector can always display in-focus image, however this assumption is not hold in practice. Due to its narrow depth-of-field, projector can make in-focus projection with strict environment. In order to increase the depth of field of conventional projectors, There are mainly two solutions, single or multiple projector based methods.

Multiple projectors for in-focus projection One intuitive solution is to use multiple overlapping projectors with different in-focus positions [Bimber and Emmerling, 2006]. Each projector projects an image onto a part of target surface that is located in the each projector's depth-of-field as shown in Fig. 4.5. This partial projection by multiple projectors minimizes a degradation caused by projector defocus on the screen surface. Thus, overlapped image forms in-focus image.

Single projector for in-focus projection When only a single projector is available, we should tackle the problem with inverse of distortion function Eq. (4.2). The inverse function against to defocus blur is edge sharpening. Thus, previous

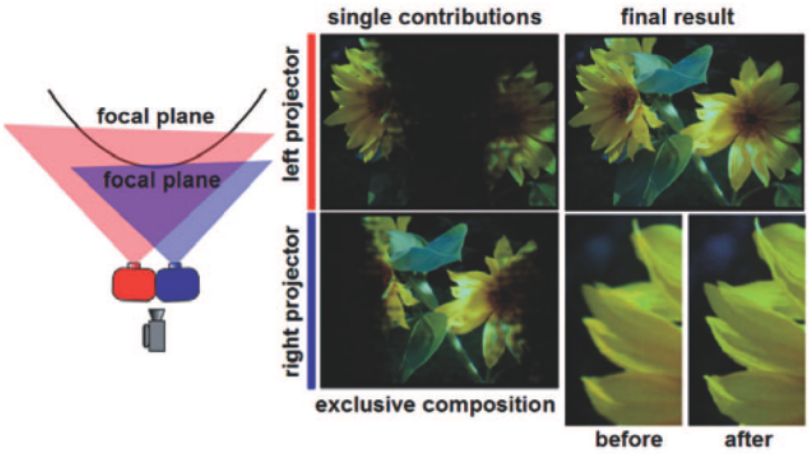


Figure 4.5: In-focus projection with multiple overlapping projectors that have different in-focus planes. Courtesy of [Bimber and Emmerling, 2006].

methods first sharpen an image and then project the sharpened image so that the sharpening effect cancels the defocus effect [Brown et al., 2006; Grosse and Bimber, 2008; Park et al., 2008; Zhang and Nayar, 2006]. Figure 4.6 shows an example of single projector based in-focus projection.

4.2.2 Distortion parameter estimation

The other problem to be solved is distortion parameter estimation. Any approach estimates Θ by pattern matching algorithm that compares the projection p and the observation c. Thus, the categorization considers what kind of patterns are used.

Structured light

The most popular approach is to use structured light (fiducial pattern). As listed by Salvi et al., coding strategy of such fiducial is classified according to the target [Salvi et al., 2004]. For geometry estimation, point is to make a set of corresponding points. Thus, point cloud, binary/gray code, chess board patterns, special markers, *e.g.*, ARTag [Fiala, 2004], are used. For color/radiometric parameters estimation, point is to make a correspondence of projector color and

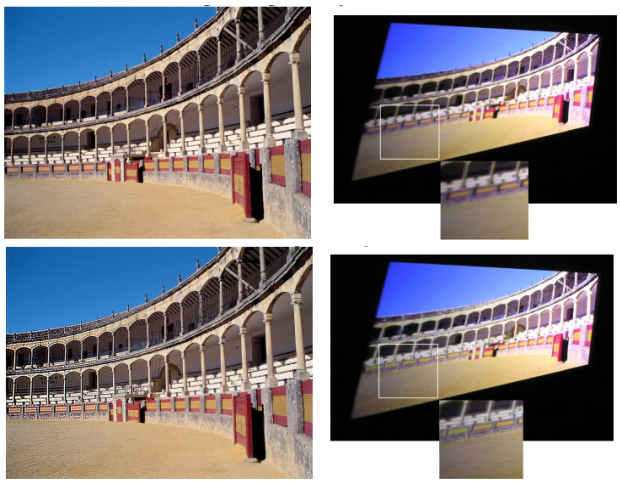


Figure 4.6: In-focus projection by sharpened image projection with single projector. Projector images (top) and their corresponding camera images (bottom). From left to right, fiducial patterns, original image, and sharpened image are shown. Courtesy of [Brown et al., 2006].

camera color. Therefore, gray color code and color code are used. For defocus parameter estimation, projected pattern should be sensitive to defocus blur. Thus, point/circle clouds are well-used. Table 4.1 classifies the related papers according to their target distortion and type of structured pattern.

Image matching based method

Thanks to the recent development of image description/matching researches, alternative approach relies on image features matching. For static scene, feature points matching, *e.g.*, SIFT feature points, are used [Takahashi et al., 2010, 2008; Yang and Welch, 2001]. For moving surface, Lukas-Kanade [Baker and Matthews, 2004] like gradient descent method [Audet et al., 2010] or stereo based tracking [Johnson and Fuchs, 2007] is suitable. Amano and his colleague directly modify the appearance [Amano and Kato, 2008, 2010]. They set the statistical property of goal image, *e.g.*, color profile and edge intensities of the camera image. Then, they can directly compute the parameter for image synthesis by comparing the statistical property of the goal image and the camera image. Another type of solutions uses some external devices for ease of track [Borkowski et al.,

Table 4.1: Classification of structured light.

Distortion	Pattern	Papers
Geometry	Point cloud	[Okatani and Deguchi, 2005, 2009]
	Binary codes	[Tardif et al., 2003]
	Chess board	[Draréni et al., 2009; Sun et al., 2008a,b]
	ARTag	[Audet and Okutomi, 2009; Fiala, 2005; Griesser and Gool, 2006]
Color	Color codes	[Fujii et al., 2005; Majumder et al., 2010, 2000, 2003; Nayar et al., 2003]
Focus	Cross dots	[Brown et al., 2006]
	Chess board	[Park et al., 2008]
	Circle dots	[Bimber and Emmerling, 2006; Grosse and Bimber, 2008; Grosse et al., 2009, 2010; Zhang and Nayar, 2006]

2003; Lee et al., 2004, 2005; Leung et al., 2009]. Contrast to the above method, it works faster and more robust.

Imperceptible structured light

As Raskar et al. mentioned [Raskar et al., 1998b], embedding imperceptible structured light [Cotting et al., 2004; Grundhöfer et al., 2007; Park et al., 2007; Zollmann and Bimber, 2007] is another solution. A temporal modulation of projector images, flickering projection in other words, allows us to embed structured light that is imperceptible for human visual system in the projector images. Zollmann et al. combine both normal structured light and imperceptible one [Zollmann et al., 2007]. The method embeds the imperceptible structured light in running phase and estimates the distortion parameters on-line. Once the on-line estimation fails, the method projects normal structured light for re-initialization.

4.3 Motivation

For upcoming imaging technologies, *e.g.*, dynamic projection in dynamic scene, what is required? In answer to this question, I develop an in-focus projection

method without implicit fiducial pattern projection. The reason of this answer comes from some issues described in the following.

Most of existing projector image adjustment works, *e.g.*, geometric adjustment and color adjustment, are constrained by hardware limitations of projector. Narrow depth-of-field results defocus blur with volumetric depth or moving screen object. When its resolution or dynamic range is not enough for the purpose, displayed image quality is significantly lost. For dynamic projection systems, one of the most important things is defocus blur. Thus, in-focus projection technique is a good topic for the future improvements of projector.

Thanks to the recent development of image feature descriptor, representation, and matching and tracking works, projector image adjustment researches are going to remove structured light projection from their method. This trend is persuasive because structured light projection is a tool not a purpose. Even though we can expect imperceptible structured light to provide as same quality result as normal structured light does, embedding the pattern itself is undesirable additional process. Therefore, developing image matching based approaches can be a suitable choice for a system under dynamic scene. Since most projector-camera researchers carefully use projector to make the display in-focus, focal adjustment was previously less required. Furthermore, in-focus projection technique is a fresh research field. Therefore, there does not exist an in-focus projection method based on image matching method.

Another concern is a way of in-focus projection. As mentioned above, there are mainly two solutions, multiple projectors or single projector to achieve in-focus projection. For developing a simple set up, I choose the latter approach.

4.4 Overview of the proposed Method

This section proposes in-focus projection method. The proposed method consists of two steps, distortion parameter estimation and adjusted image synthesis. Figure 4.7 shows overview of the proposed method. As input, the proposed method takes a projector image p and a camera image c. In the distortion parameter estimation step, the proposed method first estimates discrete PSFs at extracted regions from a pair of projector-camera image (blue dot framed).

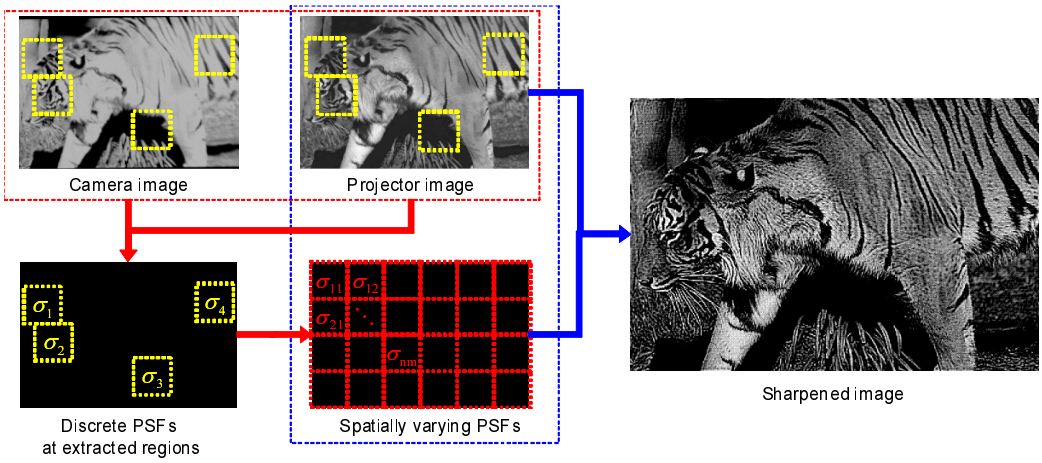


Figure 4.7: Overview of the proposed method. As pre-processing, rich textured regions (yellow dot framed) are extracted. From a pair of projector-camera image (blue dot framed), the proposed method first estimates discrete PSFs at extracted regions. Then, spatially varying PSFs covering the entire image is computed by interpolation/outerpolation. Once the spatially varying PSFs are computed, the original projector image is sharpened according to the computed PSFs.

Then, spatially varying PSFs covering the entire image is computed by interpolation/outerpolation. Then, the estimated PSFs are used to sharpen a projector image. Finally, the sharpened image is projected to the surface to remove the defocus effect. The proposed method assumes that display surface is planar and Lambertian.

4.5 Spatially varying PSFs estimation

The proposed method estimates spatially varying PSFs from a pair of projector image p and camera image c. When the light direction of the projector is not perpendicular to the display surface, projector defocus on the surface is not constant, spatially varying across the surface. In a strict sense, such PSFs should be estimated on pixel by pixel, however, it's applicable. Assuming the planar surface, we can regard the spatially varying PSFs locally uniform and smoothly varying. Thus, the proposed method estimates such PSFs from by interpolating PSFs estimated at several positions.

4.5.1 Discrete PSFs estimation

The projector defocus is described by a 2D Gaussian PSF due to its larger apertures as

$$k(\mathbf{x};\sigma) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{4.4}$$

where σ denotes standard deviation of the Gaussian. Thus, camera image c is described as

$$c(\mathbf{x}) = p(\mathbf{x}) \otimes k(\mathbf{x};\sigma). \tag{4.5}$$

Note that the amount of defocus σ is not constant on the surface.

Given projector-camera image pair, the proposed method estimate σ of the PSF as

$$\hat{\sigma} = \underset{\sigma}{\operatorname{argmin}}\, Q\left(c(\mathbf{x}) - p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right). \tag{4.6}$$

Considering projector light attenuation and noise effect, evaluation function Q computes NCC values of the argument as

$$\begin{aligned}
\hat{\sigma} &= \underset{\sigma}{\operatorname{argmin}}\, Q\left(c(\mathbf{x}) - p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right) \\
&= \underset{\sigma}{\operatorname{argmax}}\, \mathrm{NCC}\left(c(\mathbf{x}), p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right) \\
&= \underset{\sigma}{\operatorname{argmax}}\, \frac{\sum_{\mathbf{x}} \left(c(\mathbf{x}) - p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right)}{\sum_{\mathbf{x}} c(\mathbf{x}) \sum_{\mathbf{x}} p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)}.
\end{aligned} \tag{4.7}$$

Applying this σ estimation on sub images, we can estimate discrete PSFs at different positions.

Since we have both p and c, one may claim that straightforward strategy is spectral division as

$$\hat{k} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(c)}{\mathcal{F}(p)}\right), \tag{4.8}$$

however, the estimate is not correct because Eq. (4.5) ignores other projector image distortion. In addition to defocus blur, the camera image c actually contains other types of distortion, *e.g.*, geometric and color distortion. The assumption of geometrically registered projector-camera pair is theoretically cancel the geomet-

ric distortion but color distortion still remains on the image. Well-used form of the color distortion is formulated as

$$c(\mathbf{x}) = D_{\text{color}}\left(p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right). \tag{4.9}$$

The color distortion function D_{color} is single scalar for gray scale image and 3×3 matrix for color image. Loss of generality, let us consider the simple gray scale case. When color distortion function scales and biases input as

$$\begin{aligned} c(\mathbf{x}) &= D_{\text{color}}\left(p(\mathbf{x}) \otimes k(\mathbf{x};\sigma)\right) && (4.10)\\ &= C_{\text{scale}}(\mathbf{x})p(\mathbf{x}) \otimes k(\mathbf{x};\sigma) + C_{\text{bias}}(\mathbf{x}), && (4.11) \end{aligned}$$

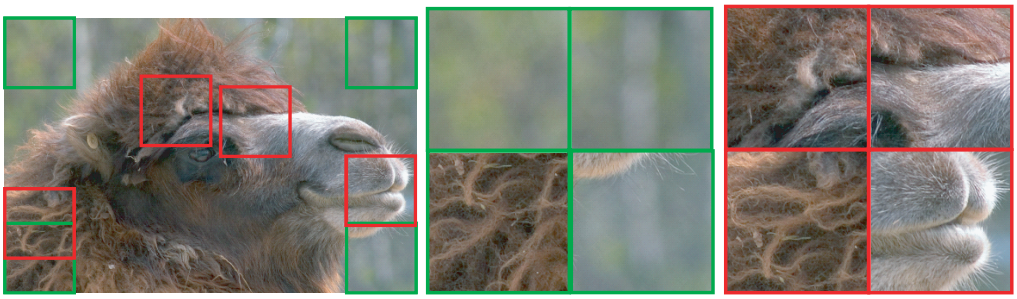
where C_{scale} and C_{bias} denote scale and bias factors. In supposed situation of the method, both C_{scale} and C_{bias} is not constant across the image. In such case, Eq. (4.8) is rewritten as

$$\begin{aligned} \mathcal{F}^{-1}\left(\frac{\mathcal{F}(c)}{\mathcal{F}(p)}\right) &= \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(C_{\text{scale}}(\mathbf{x})p(\mathbf{x}) \otimes k(\mathbf{x};\sigma) + C_{\text{bias}}(\mathbf{x})\right)}{\mathcal{F}(p)}\right)\\ &= \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(C_{\text{scale}}(\mathbf{x})\right) \otimes \mathcal{F}\left(p(\mathbf{x})\right)\mathcal{F}\left(k(\mathbf{x};\sigma)\right) + \mathcal{F}\left(C_{\text{bias}}(\mathbf{x})\right)}{\mathcal{F}(p)}\right)\\ &= \mathcal{F}^{-1}\left(\mathcal{F}\left(C_{\text{scale}}(\mathbf{x})\right) \otimes \mathcal{F}\left(k(\mathbf{x};\sigma)\right) + \frac{\mathcal{F}\left(C_{\text{bias}}(\mathbf{x})\right)}{\mathcal{F}(p)}\right). \quad (4.12) \end{aligned}$$

Therefore, spectral division cannot be applicable.

4.5.2 Spatially varying PSFs estimation

Then, the method interpolates the discrete PSFs for spatially varying PSFs estimation. Intuitively, we compute the discrete PSFs on image grid or on image corners. However, I empirically found that PSF estimation on such region sometimes fails. When a sub image p_{sub} has uniform color, Eq. (4.6) always returns $\hat{\sigma} = 0$ meaning no defocus blur. This indicates that the PSF estimation (Eq. (4.6)) implicitly assumes that projector image p_{sub} is well-textured. If we use PSFs estimated on less-textured regions, the σ estimation error propagates. Therefore, we should carefully extract rich textured sub images for stable σ estimation rather



(a) Original image (b) Image corners (c) Rich textured regions

Figure 4.8: Textures on image corners (green) and on rich textured regions (red).

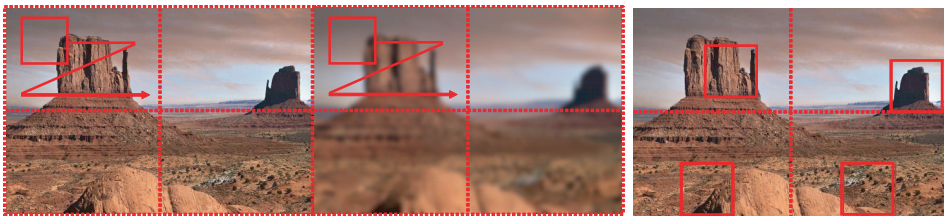
than less textured ones. Figure 4.8 compares image corners and well-textured regions. Contrast to the rich textured regions (red frames), image corners (green frames) are less textured. On such less textured regions, proposed discrete PSF estimation (Eq. (4.6)) fails.

For rich textured region extraction, the method should consider defocus effect. Since the proposed method is designed for defocus blur estimation, rich textured region should be sensitive to projector defocus. Thus, a rich textured region should satisfy the following equation as

$$\hat{p}_{\mathrm{sub}} = \underset{p_{\mathrm{sub}}}{\mathrm{argmax}} \sum_{\mathbf{x} \in \mathbf{x}_{\mathrm{sub}}} \left| p_{\mathrm{sub}}(\mathbf{x}) - p_{\mathrm{sub}}(\mathbf{x}) \otimes k(\mathbf{x}; \sigma) \right| , \tag{4.13}$$

where $\mathbf{x}_{\mathrm{sub}}$ denotes sub region of p_{sub}. Rasterizing the computation over entire image, the proposed method extract four rich textured regions and use them for the discrete PSFs estimation. Figure 4.9 depicts the extraction process.

Once four rich textured regions are extracted, the discrete PSFs are estimated on the regions. Then, spatially varying PSFs are estimated by simple linear interpolation and extrapolation. For the following process, the method spatially varying PSFs are interpolated/extrapolated on the image grid not on all the pixels.



(a) Original image p (left) and synthetically blurred image $p \otimes k$ (right).

(b) Extracted rich textured region.

Figure 4.9: Rich textured regions extraction.

4.6 Sharpened image synthesis

Next, the proposed method sharpens the projector image using the estimated PSFs. To cancel the spatially varying defocus, sharpening should also be spatially varying. Here, the method uses interpolation again for spatially varying sharpening.

Now, we have the projector image p and the estimated PSFs on the image grid σ_{g}. The pixel value of final sharpened image $\hat{p}$ at $\mathbf{x}$ is obtained by linear interpolation as

$$\begin{aligned}\hat{p}(\mathbf{x}) = D^{-1}(p(\mathbf{x}); \sigma(\mathbf{x})) = \quad & \alpha_{\mathrm{TL}} D^{-1}(p(\mathbf{x}); \sigma_{\mathrm{g}}(\mathbf{x}_{\mathrm{TL}})) \\ + \quad & \alpha_{\mathrm{TR}} D^{-1}(p(\mathbf{x}); \sigma_{\mathrm{g}}(\mathbf{x}_{\mathrm{TR}})) \\ + \quad & \alpha_{\mathrm{BL}} D^{-1}(p(\mathbf{x}); \sigma_{\mathrm{g}}(\mathbf{x}_{\mathrm{BL}})) \\ + \quad & \alpha_{\mathrm{BR}} D^{-1}(p(\mathbf{x}); \sigma_{\mathrm{g}}(\mathbf{x}_{\mathrm{BR}})), \end{aligned} \tag{4.14}$$

where D^{-1} is the sharpening function, α denotes weight parameter, σ_{g} denotes PSF at neighboring image grid, and the subscripts indicate (T)op, (B)ottom, (L)eft, and (R)ight respectively. Following this interpolation, spatially varying sharpening is done.

For sharpening function D^{-1}, the method uses the Wiener filtering [Wiener, 1949].

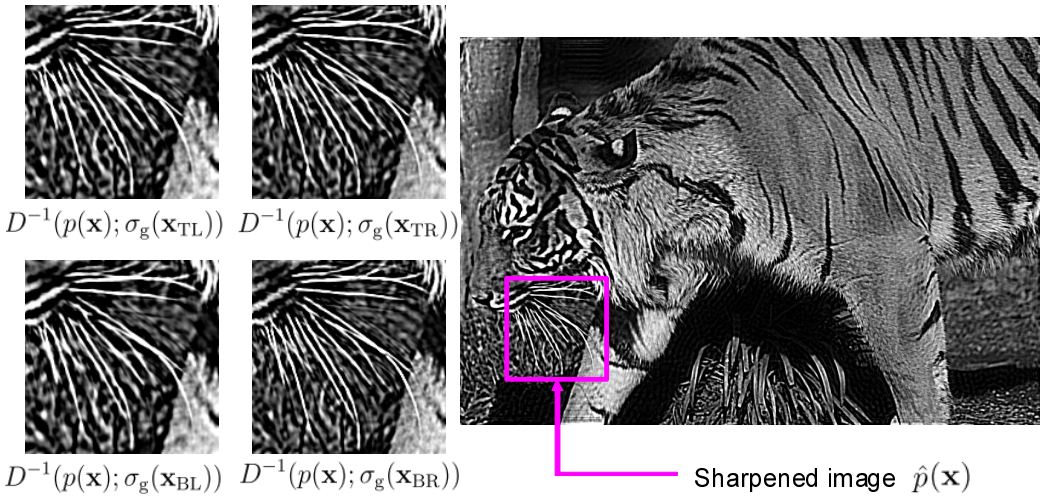


Figure 4.10: Sharpened image synthesis. As pre-processing, rich textured regions (yellow dot framed) are extracted. From a pair of projector-camera image (blue dot framed), the proposed method first estimates discrete PSFs at extracted regions. Then, spatially varying PSFs covering the entier image is computed by interpolation/outerpolation. Once the spatially varying PSFs are computed, the original projector image is sharpened according to the computed PSFs.

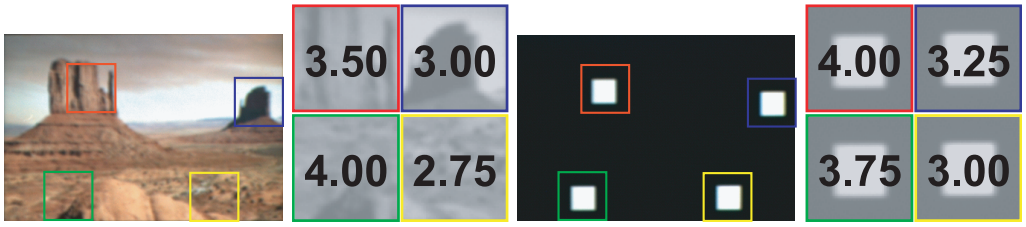
4.7 Experimental results

This section validates the proposed method with real world images. The experimental setup is as follows.

- Projector image: 960×640 resolution
- Camera image: 1024×768 resolution
- Sub image: 160×160 resolution
- γ of Wiener filter: 0.001

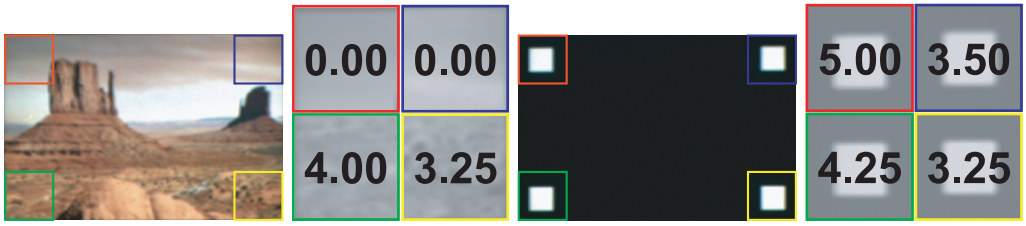
4.7.1 Spatially varying PSFs estimation

In this experiment, PSF estimation method is validated with on-axis case and off-axis case.



(a) Input image (b) Estimated σ (c) Fiducial pattern (d) Estimated σ

Figure 4.11: Rich textured region extraction and estimated PSFs. (Left) Input image and estimated PSFs on the extracted regions of input image. (Right) Fiducial pattern and estimated PSFs on the same regions.



(a) Input image (b) Estimated σ (c) Fiducial pattern (d) Estimated σ

Figure 4.12: Image corners and estimated PSFs. (Left) Input image and estimated PSFs on the corners. (Right) Fiducial pattern and estimated PSFs on the corners.

Rich textured region extraction

First experiment validates rich textured region extraction and PSF estimation on the regions. Figure 4.11(a) shows an input image and extracted regions on the image while Figure 4.11(c) does fiducial pattern that has fiducial markers on the corresponding regions. Here, estimated PSF on the fiducial markers (Fig. 4.11(d)) are regarded as ground truth. Comparing the estimated σ, the proposed method estimates similar values. Next, I use image corners for PSF estimation under same condition. Figure 4.12 shows the results. At the top corners, estimated σ using input image are zero while fiducial pattern estimates reasonable values. These results indicate that the proposed method extracts the rich textured regions and that PSF estimation performs as similar as fiducial pattern method does.

Table 4.2: PSF estimation on extracted regions. Both mean square error (MSE) and standard deviation (Std. dev.) of estimation error is shown. Small and big mean the relative amount of PSF. Partial and entire mean displayed image is partially blurred and entire image is blurred respectively.

	On-axis		Off-axis (small)		Off-axis (big)	
	small	big	partial	entire	partial	entire
MSE	1.09	0.61	0.55	0.88	0.87	0.97
Std. dev.	1.23	0.37	0.70	0.98	1.03	0.56

Discrete PSFs estimation

Same experiments are done with different conditions and different images. Table 4.2 summarizes the results. I compare the estimated PSFs on extracted regions with fiducial pattern approach. Both mean square error (MSE) and standard deviation (Std. dev.) of estimation error is shown on the table.

Spatially varying PSFs estimation

Next, spatially varying PSFs estimation method is validated. Here, I compare the proposed method with my previous method [Oyamada and Saito, 2007]. The previous method estimates spatially varying PSFs on image grids, thus it should mis-estimate PSF on less textured regions. Figure 4.13(a) tells that the amount of defocus blur is slightly increasing from top left to bottom right in the scene. Figure 4.13(b) shows that PSFs estimated on the image grid fails on less textured regions, top area of the image. On the other hand, proposed method estimates closer values and does not have outliers. Same as previous experiment, both methods are compared with fiducial pattern approach and estimation error is shown in table format. Table 4.3 and Table 4.4 show the results of previous method and ones of the proposed method respectively. These results indicate that the proposed method suppresses the mis-estimation and error propagation caused by interpolation.

(a) Fiducial pattern (b) [Oyamada and Saito, 2007] (c) Proposed method

Figure 4.13: Spatially varying PSFs estimation. From left to right, fiducial pattern approach (ground truth), previous method [Oyamada and Saito, 2007], and the proposed method are shown.

Table 4.3: Spatially varying PSFs estimation by previous method [Oyamada and Saito, 2007].

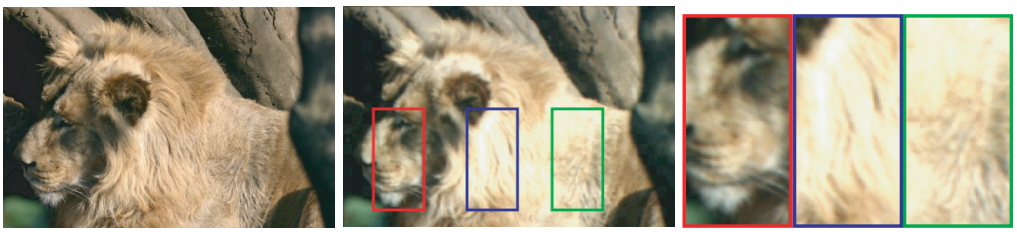
	On-axis		Off-axis (small)		Off-axis (big)	
	small	big	partial	entire	partial	entire
MSE	2.05	1.61	0.83	2.10	1.75	2.30
Std. dev.	1.26	2.35	1.08	3.46	1.35	4.23

4.7.2 Sharpened image projection

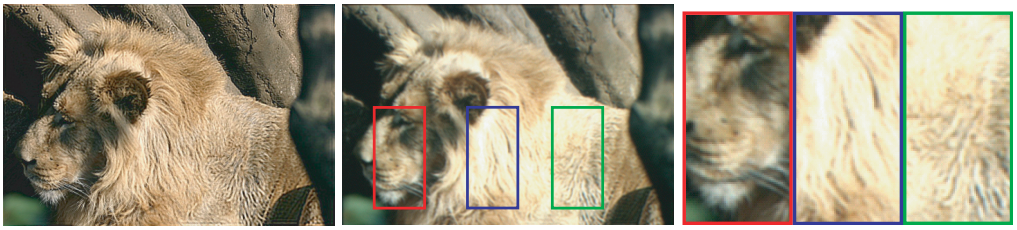
Next experiment is to validate that sharpening can reduce the defocus effect. Figure 4.14 shows the result of an off-axis case. Figure 4.14(a) and Fig. 4.14(d) show that edges in the sharpened image is emphasized. Comparing the displayed images, the sharpened image preserves lion's fur and beard while the original image loses. PSN ratio between the original image and displayed result of original

Table 4.4: Spatially varying PSFs estimation by the proposed method.

	On-axis		Off-axis (small)		Off-axis (big)	
	small	big	partial	entire	partial	entire
MSE	1.58	0.96	0.93	1.35	1.85	1.27
Std. dev.	1.19	0.60	0.86	1.06	1.20	1.43



(a) Original image (b) Displayed (a) on the surface (c) Zoom up of (b)



(d) Sharpened image (e) Displayed (d) on the surface (f) Zoom up of (e)

Figure 4.14: Sharpened image projection.

image and sharpened image are 21.3 and 21.7. Thus, the result indicates that sharpened image projection can reduce the projector defocus effect.

4.8 Conclusion

The motivation of this work is to achieve in-focus projection without explicit fiducial patterns projection. Since projector-camera image pair is available in projector-camera systems, the proposed method estimates spatially varying PSFs from the image pair. The method has two key points. First point is that the assumption that the projector image is available reduces the difficulty of PSF estimation. Using the assumption, discrete PSF on a sub image can be estimated by comparing two images. Second point is to rely on the discrete PSFs on only rich textured regions. Thus, we can prevent PSF estimation error propagation. Experimental results show that the proposed method realizes in-focus projection without explicit fiducial patterns projection.

One of the biggest contributions of this work to our research domain is to

introduce simple signal processing strategy to estimate projector defocus without fiducial patterns projection. In fact, the concept for geometric adjustment was first realized by Yang and Welch [Yang and Welch, 2001] but no prior work exists for focal adjustment. I first tried to realize in-focus projection without fiducial patterns projection. After I proposed the original idea of this approach in [Oyamada and Saito, 2007], Park et al. developed a prototype method of on-line in-focus projection [Park et al., 2008]. This indicates that this work contributes to research domain.

4.8.1 Future direction

To use the proposed method in dynamic scene, entire process should be completed in real time. Using a Pentium 4 processor of 3.2GHz and 1.0GB RAM, entire process takes about 8 sec, 1 sec for PSF estimation and 7 sec for image synthesis. For speed up, several approaches are conceivable. Simple but powerful tools for fast image synthesis is to develop the function with GPU. Another idea for dynamic scene is using sequential data even though I mentioned single shot based approach is better than sequential data based method in Sec. 2.6. Since the proposed method uses single projector-camera image pair, the discrete PSF estimation (Eq. (4.6)) searches all potential σ values. With sequential data, we can limit the parameter search range close in value to the previous estimate. Another approach is to compute the discrete PSFs at multiscale. Run PSF estimation at different scale, we obtain several PSF maps each of which corresponds to one scale as shown in Fig. 4.15. If a sub image contain enough texture for PSF estimation, the estimated PSF value, *e.g.*, σ value for defocus blur and probably shape for motion blur, should be similar to neighboring scale. When the estimated PSF value violates the above assumption, we stop estimation at the regions of finer scale.

Another potential extension of this work is to run the in-focus projection in dynamic environment. Spatially varying PSFs estimation itself is not computationally heavy, thus the method potentially runs in real time. Straightforward strategy is to merge object tracking algorithm [Baker and Matthews, 2004] in the process. At the initialization step, we extract the rich textured regions of projector image. On running phase, target surface tracking and PSF estimation are simultaneously

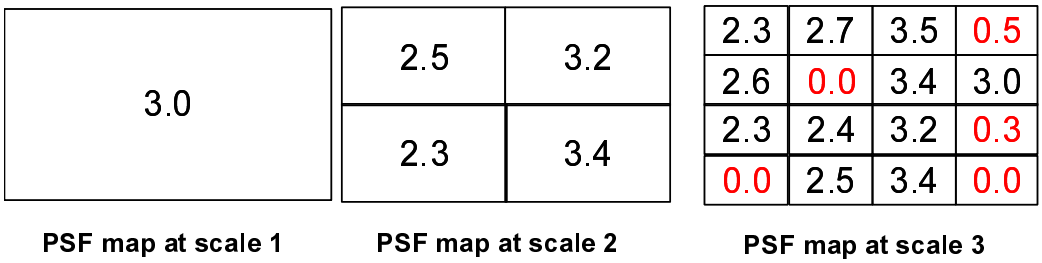


Figure 4.15: PSF map at different image scale. From left to right, sub image size gets smaller. Red text means outlier.

processed. One concern is that discrete PSF estimation takes time proportional to the amount of defocus. Since PSF estimation contains convolution operations, the more defocus blur is the bigger kernel size is. As a result, the method runs slowly for larger defocus scene. To solve this problem, PSF estimation can be run in cepstrum domain. Different from estimation in image domain, we don't need to compute convolution operations in cepstrum domain. Thus, adaptive PSF estimation according to defocus amount is possible solution.

Chapter 5

Conclusions

The history from ancient photograph to state-of-the-art image acquisition/display technology, I imagine the configuration of imaging technology in the future. That is dynamic contents projection in dynamic scene as shown in Fig. 1.6. For the upcoming technologies, we may have to prepare several things. To realize the future technologies, one issue to be solved is image degradation occurred during imaging process. To relax the difficulty of the issue, I have separately focused on image restoration works for both image acquisition and display process. Especially, key contribution is to introduce traditional signal processing theory to concrete problems, *e.g.*, non-linear motion PSF estimation for motion deblurring and spatially varying PSF estimation for in-focus projection. As the experimental results in Ch. 3 and Ch. 4, the proposed methods work in real cases.

Cepstral Analysis based Non-Linear Motion PSF Estimation

Chapter 3 focuses on non-linear motion PSF estimation from a single blurred image. To solve this ill-posed problem, I analyzed the cepstral behavior of non-linear motion, *e.g.*, PLM PSF and NLM PSF. To use the analyzed behavior for non-linear PSF estimation, the proposed method estimates a PSF with two steps, PSF candidates estimation and then PSF candidates evaluation. The cepstral behavior is used for the former process. Once the method estimates PSF candidates, they are evaluated by considering the imaging equation and the likelihood on PSF. The main contribution of this work is that the method extends the classic cepstral

analysis approaches for non-linear motion PSF estimation. With both synthesized images and real world images, the proposed method is validated. Experimental results showed that the proposed method can estimate non-linear PSF.

In-Focus Projection from a Single Projector-Camera Image Pair

Chapter 4 focuses on spatially varying PSFs estimation from a projector-camera image pair. For PSF estimation, we can put strong assumption on projector-camera systems that projector-camera image pair is available. With the assumption, image matching based algorithm can be applicable even for spatially varying PSFs. Another assumption is that projector image has rich textured region that is suitable for matching based PSF estimation. Experimental results showed that the proposed method realizes in-focus projection without explicit fiducial patterns projection. The main contribution of this work is that I introduce PSF estimation strategy into this field so that the proposed method can passively estimate the information of projector defocus without fiducial patterns projection. The experimental results show that the proposed method achieves in-focus projection without using fiducial patterns.

5.1 Future works

Let me introduce some potential future works in my mind.

Human perception based image restoration

The proposed method considers physical phenomena, *e.g.*, camera motion and projector defocus because target blur for the method is explained as physical phenomenon. If the goal is to find an answer that is physically correct, this type of approach is fine. However, the final purpose is to show satisfying photograph to the users. This indicates that we have to consider human perception as well. Not many but some projector color adjustment works consider human perception such as Ashdown et al.' work [Ashdown et al., 2006] and Amano and Kato's

work [Amano and Kato, 2010]. Blur correction considering human perception is one potential and interesting future works.

Blur information for realistic AR

Blur estimation method is useful for realistic AR. The key issue to develop a realistic AR system, consistency between real world and virtual world should be ensured. Thus, elemental technologies of AR have focused on finding geometric and photometric information from an observed image. In 2009, Park et al. proposed a method that estimates parametric motion blur information from an input image and then render virtual object under the blur [Park et al., 2009]. Since AR is getting a lot of attention as a dreamlike technology, blur estimation for realistic AR is also interesting topic. The point of the topic is computation time. Since most of AR system are required to be run in real time, some idea making the computation faster is necessary.

Enhanced display for weak eyesight

One application contains the component of the proposed method is image enhancement for weak eyesight people. When weak eyesight people go to an eyeglasses shop, he tries several types of eyeglasses. Unfortunately, eyeglasses set at shop front do not have lens. This means that he check how much the eyeglasses suits to him without lens. For such situation, the concept of in-focus projection can be helpful. Figure 5.1 shows the concept of the display. Top is normal eyesight people's view while bottom is weak eyesight people's view. By enhancing texture in the image, weak eyesight people can provably perceive the texture that is imperceptible without eyeglasses. To realize this application, we need to know the amount of weak eyesight and to enhance the view according to the weakness. Each of necessary part is corresponding to PSF estimation and to in-focus image synthesis respectively.

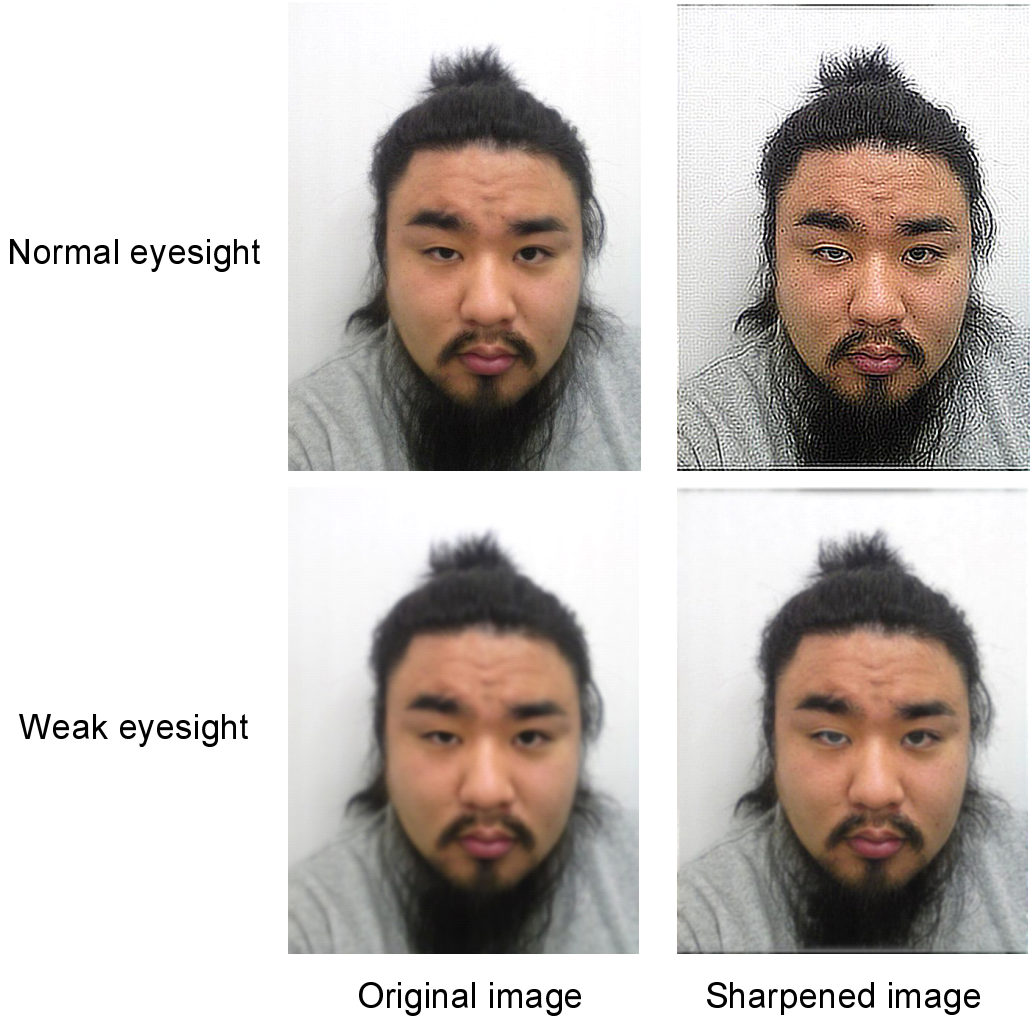


Figure 5.1: Concept of enhanced display.

Appendix A

Richardson-Lucy algorithm

Richardson-Lucy (RL) algorithm [Lucy, 1974; Richardson, 1972] is a non-blind deconvolution algorithm based on Bayes' theorem. Here, we derive original RL algorithm and an extended one.

A.1 Bayes' theorem

Following the Bayes' theorem, a posterior distribution is formulated as

$$P(f(\mathbf{x})|g(\mathbf{x})) = \frac{P(g(\mathbf{x})|f(\mathbf{x}))P(f(\mathbf{x}))}{P(g(\mathbf{x}))}, \tag{A.1}$$

where $P(g(\mathbf{x})|f(\mathbf{x}))$, $P(g(\mathbf{x}))$, and $P(f(\mathbf{x}))$ are the likelihood, the evidence, and the prior distribution.

A.2 Maximum Likelihood estimation

Bayesian estimation is to estimate f that maximizes the posterior distribution $P(f(\mathbf{x})|g(\mathbf{x}))$, thus called Maximum-A-Posteriori (MAP) estimation. MAP esti-

mation is formulated as

$$\begin{aligned}\hat{f} &= \underset{f}{\operatorname{argmax}} \prod_{\mathbf{x}} P(f(\mathbf{x})|g(\mathbf{x})) \\ &= \underset{f}{\operatorname{argmax}} \prod_{\mathbf{x}} \frac{P(g(\mathbf{x})|f(\mathbf{x}))P(f(\mathbf{x}))}{P(g(\mathbf{x}))}. \end{aligned} \tag{A.2}$$

Since we don't have any prior knowledge on images g and f, we assume the evidence term and the prior distribution as uniform distribution. Therefore, MAP estimation is converted to Maximum-Likelihood estimation as

$$\hat{f} \propto \underset{f}{\operatorname{argmax}} \prod_{\mathbf{x}} P(g(\mathbf{x})|f(\mathbf{x})). \tag{A.3}$$

Since the equation computes the products over $\mathbf{x}$, the resulting value may potentially be underflow. To reduce the risk, we usually minimize the negative log likelihood instead of maximizing the likelihood. Since $\log$ is monotonically increase, the conversion does not lose the contents. Thus, Eq. (A.3) is reformulated as

$$\hat{f} = \underset{f}{\operatorname{argmin}} \sum_{\mathbf{x}} L(\mathbf{x}), \tag{A.4}$$

where $L(\mathbf{x}) = -\log(P(g(\mathbf{x})|f(\mathbf{x})))$ is the negative log likelihood.

A.3 Poisson noise

When image noise n follows Poisson distribution, the likelihood is formulated as

$$P\left(g(\mathbf{x})|f(\mathbf{x})\right) = \prod_{\mathbf{x}} \frac{f(\mathbf{x}) \otimes k(\mathbf{x})^{g(\mathbf{x})} \exp\left(-f(\mathbf{x}) \otimes k(\mathbf{x})\right)}{g(\mathbf{x})!}. \tag{A.5}$$

In this case, the negative log likelihood is written as

$$\begin{aligned} L(\mathbf{x}) &= -\log\left(P\left(g(\mathbf{x})|f(\mathbf{x})\right)\right) \\ &= \sum_{\mathbf{x}} \log \frac{f(\mathbf{x}) \otimes k(\mathbf{x})^{g(\mathbf{x})} \exp\left(-f(\mathbf{x}) \otimes k(\mathbf{x})\right)}{g(\mathbf{x})!} \\ &= \sum_{\mathbf{x}} f(\mathbf{x}) \otimes k(\mathbf{x}) - g(\mathbf{x}) \log\left(f(\mathbf{x}) \otimes k(\mathbf{x})\right) + \log g(\mathbf{x})!. \end{aligned} \quad \text{(A.6)}$$

Consider a small perturbation $\Delta\mathbf{x}$. The negative log likelihood is

$$L(\mathbf{x}+\Delta\mathbf{x}) = \sum_{\mathbf{x}} f(\mathbf{x}+\Delta\mathbf{x}) \otimes k(\mathbf{x}) - g(\mathbf{x}) \log\left(f(\mathbf{x}+\Delta\mathbf{x}) \otimes k(\mathbf{x})\right) + \log g(\mathbf{x})!. \quad \text{(A.7)}$$

Removing constant term w.r.t. f, we obtain

$$L(\mathbf{x}+\Delta\mathbf{x}) = \sum_{\mathbf{x}} f(\mathbf{x}+\Delta\mathbf{x}) \otimes k(\mathbf{x}) - g(\mathbf{x}) \log\left(f(\mathbf{x}+\Delta\mathbf{x}) \otimes k(\mathbf{x})\right). \quad \text{(A.8)}$$

Assuming $f(\mathbf{x}+\Delta\mathbf{x}) = f(\mathbf{x}) + f(\Delta\mathbf{x})$, the equation is rewritten as

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= \sum_{\mathbf{x}} (f(\mathbf{x}) + f(\Delta\mathbf{x})) \otimes k(\mathbf{x}) - g(\mathbf{x}) \log\left((f(\mathbf{x}) + f(\Delta\mathbf{x})) \otimes k(\mathbf{x})\right) \\ &= \sum_{\mathbf{x}} f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left(f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x})\right). \end{aligned} \quad \text{(A.9)}$$

Here, let $f \otimes k(\cdot)$ represent $f(\mathbf{x}) \otimes k(\cdot)$. Then, the equation is expanded as

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= \sum_{\mathbf{x}} f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left(f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x})\right) \\ &= \sum_{\mathbf{x}} f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left[f \otimes k(\mathbf{x}) \left(1 + \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})}\right)\right] \\ &= \sum_{\mathbf{x}} f \otimes k(\mathbf{x}) + f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left(f \otimes k(\mathbf{x})\right) - g(\mathbf{x}) \log\left(1 + \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \\ &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left(1 + \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})}\right). \end{aligned} \quad \text{(A.10)}$$

Following the Taylor expansion that $\log(1+x) \approx x - \frac{x^2}{2}$, the equation is approximated as

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \log\left(1 + \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \\ &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})} + \frac{1}{2} g(\mathbf{x}) \left(\frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})}\right)^2. \end{aligned} \tag{A.11}$$

Omitting the last term, because it's too small,

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) - g(\mathbf{x}) \frac{f \otimes k(\Delta\mathbf{x})}{f \otimes k(\mathbf{x})} \\ &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) \left(1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})}\right). \end{aligned} \tag{A.12}$$

From the definition of convolution integral $\int ab \otimes cdx = \int ba \otimes \overline{k}dx$, where $\overline{k}k$ is the adjoint of k,

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \otimes k(\Delta\mathbf{x}) \left(1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \\ &= L(\mathbf{x}) + \sum_{\mathbf{x}} f \left(1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \otimes \overline{k}(\Delta\mathbf{x}). \end{aligned} \tag{A.13}$$

The partial derivative of $L(\mathbf{x})$ on $\mathbf{x}$ is derived as

$$\begin{aligned} \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} &= \frac{L(\mathbf{x}+\Delta\mathbf{x}) - L(\mathbf{x})}{\Delta\mathbf{x}} \\ &= \frac{1}{\Delta\mathbf{x}} \sum_{\mathbf{x}} f \left(1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \otimes \overline{k}(\mathbf{x}). \end{aligned} \tag{A.14}$$

Since the minimization of the negative log likelihood is obtained by finding $\mathbf{x}$ satisfying

$$\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = 0.$$

Thus, we have

$$\left(1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})}\right) \otimes \overline{k}(\mathbf{x}) = 0$$
$$1 - \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})} \otimes \overline{k}(\mathbf{x}) = 0 \tag{A.15}$$

Using the convergence condition $\frac{f_{t+1}(\mathbf{x})}{f_t(\mathbf{x})} = 1$, we obtain the update rule as

$$\frac{f_{t+1}(\mathbf{x})}{f_t(\mathbf{x})} = \frac{g(\mathbf{x})}{f \otimes k(\mathbf{x})} \otimes \overline{k}(\mathbf{x}). \tag{A.16}$$

Finally, we obtain the Richardson-Lucy deconvolution algorithm as

$$f_{t+1}(\mathbf{x}) = f_t(\mathbf{x}) \left(\frac{g(\mathbf{x})}{f_t \otimes k(\mathbf{x})} \otimes \overline{k}(\mathbf{x})\right). \tag{A.17}$$

A.4 Gaussian noise

When image noise follows Gaussian distribution, the likelihood is formulated as

$$\begin{aligned} P\left(g(\mathbf{x})|f(\mathbf{x})\right) &= \prod_{\mathbf{x}} \mathcal{N}(f \otimes k(\mathbf{x}), \sigma^2) \\ &= \prod_{\mathbf{x}} \exp\left(\frac{-|g(\mathbf{x}) - f \otimes k(\mathbf{x})|^2}{2\sigma^2}\right). \end{aligned} \tag{A.18}$$

The negative log likelihood $L(\mathbf{x})$ is

$$\begin{aligned} L(\mathbf{x}) &= -\log\left(P\left(g(\mathbf{x})|f(\mathbf{x})\right)\right) \\ &= \sum_x \left(g(\mathbf{x}) - f \otimes k(\mathbf{x})\right)^2. \end{aligned} \tag{A.19}$$

Consider a small perturbation Δx. The negative log likelihood $L((\mathbf{x}) + \Delta(\mathbf{x}))$

is

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= \sum_x g(\mathbf{x})^2 - 2g(\mathbf{x}) f \otimes k(\mathbf{x}+\Delta\mathbf{x}) + f \otimes k(\mathbf{x}+\Delta\mathbf{x})^2 \\ &= L(\mathbf{x}) + \sum_x -2g(\mathbf{x}) f \otimes \mathbf{x}(\Delta x) + 2f \otimes k(\mathbf{x}) f \otimes h(\Delta\mathbf{x}) + f \otimes k(\Delta\mathbf{x})^2. \end{aligned} \tag{A.20}$$

Omitting the last term, because it's too small,

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= L(\mathbf{x}) + \sum_x -2g(\mathbf{x}) f \otimes k(\Delta\mathbf{x}) + 2f \otimes k(\mathbf{x}) f \otimes k(\Delta\mathbf{x}) + f \otimes k(\Delta\mathbf{x})^2 \\ &= L(\mathbf{x}) + \sum_x -2g(\mathbf{x}) f \otimes k(\Delta x) + 2f \otimes k(\mathbf{x}) f \otimes k(\Delta\mathbf{x}) \\ &= L(\mathbf{x}) + 2\sum_x f \otimes k(\Delta\mathbf{x}) \left(f \otimes k(\mathbf{x}) - g(\mathbf{x})\right). \end{aligned} \tag{A.21}$$

From the definition of convolution integral $\int ab \otimes cdx = \int ba \otimes \overline{k}dx$, where $\overline{k}k$ is the adjoint of k,

$$\begin{aligned} L(\mathbf{x}+\Delta\mathbf{x}) &= L(\mathbf{x}) + 2\sum_x f \otimes k(\Delta\mathbf{x}) \left(f \otimes k(\mathbf{x}) - g(\mathbf{x})\right) \\ &= L(\mathbf{x}) + 2\sum_x f \left(f \otimes k(\mathbf{x}) - g(\mathbf{x})\right) \otimes \overline{k}(\Delta x). \end{aligned} \tag{A.22}$$

The partial derivative of $L(\mathbf{x})$ on $\mathbf{x}$ is derived as

$$\begin{aligned} \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} &= \frac{L(\mathbf{x}+\Delta\mathbf{x}) - L(\mathbf{x})}{\Delta\mathbf{x}} \\ &= \frac{2}{\Delta\mathbf{x}} \sum_{\mathbf{x}} f \left(f \otimes k(\mathbf{x}) - g(\mathbf{x})\right) \otimes \overline{k}(\mathbf{x}). \end{aligned} \tag{A.23}$$

Since the minimization of the negative log likelihood is obtained by finding $\mathbf{x}$ satisfying

$$\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = 0.$$

Thus, we have

$$f \left(f \otimes k(\mathbf{x}) - g(\mathbf{x})\right) \otimes \overline{k}\mathbf{x} = 0. \tag{A.24}$$

Using the convergence condition $f_{t+1}(\mathbf{x}) - f_t(\mathbf{x}) \approx 0$, we obtain the update rule as

$$f_{t+1}(\mathbf{x}) - f_t(\mathbf{x}) = (f \otimes k(\mathbf{x}) - g(\mathbf{x})) \otimes \overline{k}(\mathbf{x}). \tag{A.25}$$

Finally, we obtain the Richardson-Lucy deconvolution algorithm for Gaussian noise as

$$f_{t+1}(\mathbf{x}) = f_t(\mathbf{x}) + (g(\mathbf{x}) - f_t \otimes k(\mathbf{x})) \otimes \overline{k}(\mathbf{x}). \tag{A.26}$$

Appendix B

Natural image statistics

Recently, many computer vision researches have paid attention to the natural image statistics.

Natural image statistics represents the statistical properties of natural images. One may doubt that natural images have common statistical properties because there exist huge number of potential images, *e.g.*, a 100×100 8-bit gray scale image has more than 10^{24000} potential images. This doubt assumes all pixel intensity is generated following an i.i.d. (independent and identically distributed) uniform distribution, however natural images have some redundancy. Since redundancy helps reducing ambiguity of a problem, seeking the statistical property of natural images and its usage should take very important roles in computer vision researches. One example famous example using natural image statistics is JPEG image format. JPEG is an image format of lossy compression. JPEG compression is based on a statistical property of natural images that the energy of power spectra of natural images are concentrated mainly in lower frequencies. Thus, higher frequencies have less contribution to entire images.

Torralba and Oliva analyzed the statistical properties of the spectra of natural images for scene and object categorization tasks [Torralba and Oliva, 2003]. In their paper, they model the power spectra of images using polar coordinates as

$$E[|I(f,\theta)|^2] \simeq \frac{A_s(\theta)}{f^{\alpha_s(\theta)}}, \tag{B.1}$$

where f and θ denote frequency and orientation of frequency component respec-

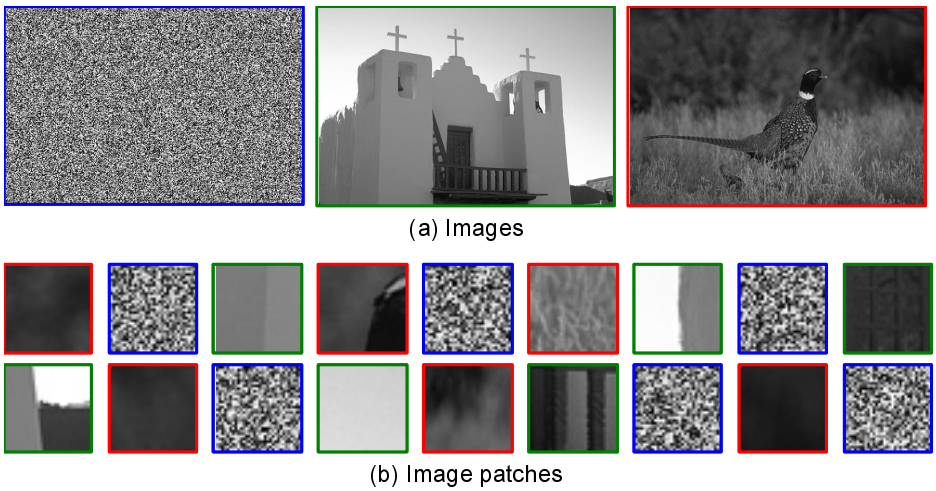


Figure B.1: Images and their patches. (a, left) An image randomly generated from an i.i.d. uniform distribution. (a, middle and right) Images of man-made object and natural object from Berkeley segmentation data set [Ber]. (b) Patches cropped from the images. Blue, green and red framed patches are the patches from random image, man-made object image, and natural object image, respectively.

tively, and A_s denotes an amplitude scaling factor for each orientation and α_s is the frequency exponent as a function of orientation. Note that f in Eq. (B.1) denotes a frequency while f in the other parts of this thesis denote unknown latent images. The equation implicitly tells that mean power spectrum of natural images is proportional to the inverse of frequency component. Though the paper focuses on image categorization not on image deblurring, the point of this paper is that the spectra of natural images and also man-made scenes can be modeled.

Recently, we pay attention to the statistical property of gradient of natural images. Figure B.1 compares a random image and natural images. From the patches cropped from the images, readers can easily classify which patches are from natural images. Since the random image is synthesized following i.i.d. uniform distribution, neighboring pixels in the random image have no relationship. Thus, patches of the random image look distributed. On the other hand, patches of natural images seem to consist of a uniform or a few colors. Even though we don't know exactly where a patch is coming from, our perception systems distinguish the difference between random image and natural images. Figure B.2 plots

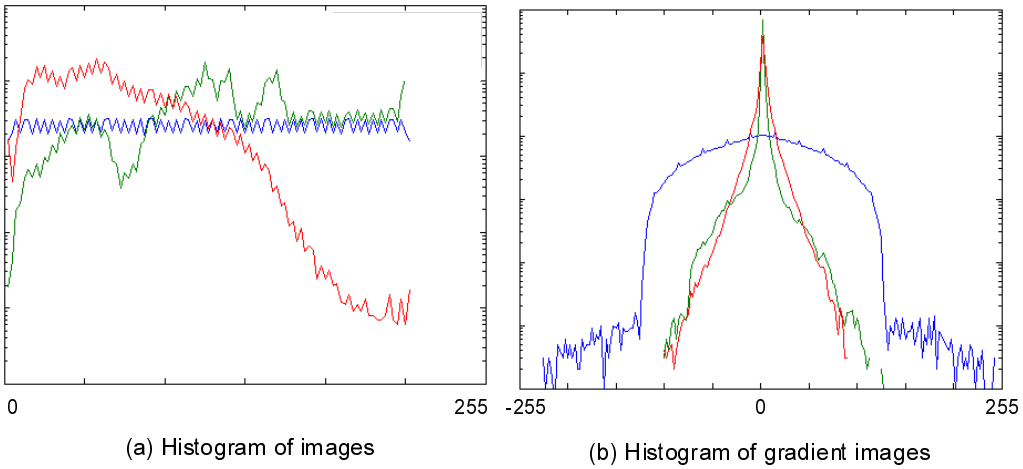


Figure B.2: Histograms of three images of Fig. B.1. Blue, green and red curves represent histograms of random image, man-made object image, and natural object image, respectively. (a) Histograms of images. (b) Histograms of gradient images.

histograms of the images and ones of gradient images. Histograms of images show that pixel intensities of random image are distributed while ones of natural images concentrate on. Histograms of gradient images show more interesting observation. Contrast to one of random image, the histograms of gradient of natural images are 0-peaked and heavy-tailed distributions. This observation means that natural images are locally uniform color or consist of a few colors. Thus, 0-peaked and heavy-tailed distributions is corresponding to the observation that patches cropped from natural images seem to consist of a few colors.

The above statistical property can be a prior knowledge of natural images as mentioned in Sec. 2.4.4 and Sec. 2.5.5. For further information, readers can refer to [Hyvärinen et al., 2009].

Bibliography

The Berkeley segmentation dataset. http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/. xvii, 67, 119

Mariana S. C. Almeida and Luis B. Almeida. Blind deblurring of natural images. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1261–1264, 2008. 40, 41, 42

Mariana S. C. Almeida and Luis B. Almeida. Blind deblurring of foreground-background images. In *International Conference on Image Processing (ICIP)*, pages 1301–1304, 2009. 40

Toshiyuki Amano and Hirokazu Kato. Real world dynamic appearance enhancement with procam feedback. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2008. 92

Toshiyuki Amano and Hirokazu Kato. Appearance control by projector camera feedback for visually impaired. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 57–63. IEEE, 2010. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543478. xv, 86, 92, 109

C. V. Angelino, E. Debreuve, and M. Barlaud. A nonparametric minimum entropy image deblurring algorithm. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 925–928, 2008. 24

Mark Antunes, Michael Trachtenberg, Gabriel Thomas, and Tina Shoa. All-in-focus imaging using a series of images on different focal planes. In *Interna-*

tional Conference on Image Analysis and Recognition, pages 174–181, 2005. 20

Apple Inc. iPhone. http://www.apple.com/iphone/. 8

Mark Ashdown, Takahiro Okabe, Imari Sato, and Yoichi Sato. Robust content-dependent photometric projector compensation. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2006. 90, 108

Samuel Audet and Jeremy R. Cooperstock. Shadow removal in front projection environments using object tracking. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2007. 85

Samuel Audet and Masatoshi Okutomi. A user-friendly method to geometrically calibrate projector-camera systems. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 47–54, 2009. 93

Samuel Audet, Masatoshi Okutomi, and Masayuki Tanaka. Direct image alignment of projector-camera systems with planar surfaces. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 303–310. IEEE, 2010. URL http://www.computer.org/portal/web/csdl/doi/10.1109/CVPR.2010.5540199. 9, 92

Hacheme Ayasso and Ali Mohammad-Djafari. Joint image restoration and segmentation using gauss-markov-potts prior models and variational bayesian computation. In *International Conference on Image Processing (ICIP)*, pages 2265–2277, 2009. 24

S. Derin Babacan, Rafael Molina, and Aggelos K. Katsaggelos. Total variation image restoration and parameter estimation using variational posterior distribution approximation. In *International Conference on Image Processing (ICIP)*, pages I–97–I–100, 2007. 24

S. Derin Babacan, Rafael Molina, and Aggelos K. Katsaggelos. Sparse bayesian image restoration. In *International Conference on Image Processing (ICIP)*, pages 3577–3580, 2010. URL http://decsai.ugr.es/vip/files/conferences/0003577Icip2010restoration.pdf. 24

S. Derin Babacan, Jingnan Wang, Rafael Molina, and Aggelos K. Katsaggelos. Bayesian blind deconvolution from differently exposed image pairs. In *International Conference on Image Processing (ICIP)*, pages 133–136, 2009. 40, 41

Simon Baker and Iain Matthews. Lucas-kanade 20 years on: A unifying framework. *International Journal of Computer Vision (IJCV)*, 56(3):221–255, 2004. 9, 92, 105

Deepak Bandyopadhyay, Ramesh Raskar, and Henry Fuchs. Dynamic shader lamps : Painting on movable objects. In *International Symposium on Augmented Reality*, pages 207–216, 2001. x, 6, 7, 87, 89

Peter Barnum, Srinivasa Narasimhan, and Takeo Kanade. A projector-camera system for creating a display with water drops. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2009. x, 7, 8

Moshe Ben-Ezra and Shree K. Nayar. Motion-based motion deblurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):689–698, 2004. 49

Oliver Bimber and Andreas Emmerling. Multifocal projection: A multiprojector technique for increasing focal depth. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):658–667, 2006. xv, 90, 91, 93

Oliver Bimber, Andreas Emmerling, and Thomas Klemmer. Embedded entertainment with smart projectors. In *ACM SIGGRAPH Courses*, 2005. 90

Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A K Peters LTD, 2005. ISBN 1568812302. URL `http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/1568812302`. 6, 87

Stanislaw Borkowski, Olivier Riff, and James L. Crowley. Projecting rectified images in an augmented environment. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2003. 92

Michael Brown, Aditi Majumder, and Ruigang Yang. Camera-based calibration techniques for seamless multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):193–206, 2005. 90

Michael S. Brown, Peng Song, and Tat-Jen Cham. Image pre-conditioning for out-of-focus projector blur. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1956–1963, 2006. xv, 91, 92, 93

Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen. Blind motion deblurring from a single image using sparse approximation. In *IEEE Conference on Compuer Vision and Pattern Recognition (CVPR)*, pages 104–111, 2009a. 40, 41

Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen. Blind motion deblurring using multiple images. *Journal of Computational Physics*, 228(14):5057–5071, 2009b. 40, 41

Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen. High-quality curvelet-based motion deblurring from an image pair. In *IEEE Conference on Compuer Vision and Pattern Recognition (CVPR)*, pages 1566–1573, 2009c. 40, 41

Michael Cannon. Blind deconvolution of spatially invariant image blurs with phase. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24 (1):58–63, 1976. 31, 32, 54

Canon. Canon in-lens image stabilizers. `http://web.canon.jp/imaging/lens/index.html`, 1995. 10, 11

James N. Caron, Nader M. Namazi, and Chris J. Rollins. Noniterative blind data restoration by use of an extracted filter function. *Appl. Opt.*, 41(32):6884–6889, 2002. 38

Ayan Chakrabarti, Todd Zickler, and William T. Freeman. Analyzing spatially-varying blur. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2512–2519. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5539954`. 42

Michael M Chang, Murat A Tekalp, and Tanju A Erdem. Blur identification using the bispectrum. *IEEE Transactions on Signal Processing*, 39(10):2323–2325, 1991. 31, 32, 33, 54, 59

Jia Chen, Lu Yuan, Chi-Keung Tang, and Long Quan. Robust dual motion deblurring. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. 38, 40, 41

C. Chesneau, M. J. Fadili, and J. L. Starck. Image deconvolution by stein block thresholding. In *International Conference on Image Processing (ICIP)*, pages 1329–1332, 2009. 22

Sunghyun Cho and Seungyong Lee. Fast motion deblurring. *ACM Transactions on Graphics (SIGGRAPH ASIA)*, 28:5, 2009. 40, 41

Sunghyun Cho, Yasuyuki Matsushita, and Seungyong Lee. Removing non-uniform motion blur from images. In *International Conference on Computer Vision (ICCV)*, pages 1–8, 2007. 42

Michael F. Cohen and Richard Szeliski. The moment camera. *Computer*, 39(8): 40–45, 2006. 5

Daniel Cotting, Martin Naef, Markus Gross, and Henry Fuchs. Embedding imperceptible patterns into projected images for simultaneous acquisition and display. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 100–109, 2004. 93

Tom Cuypers, Yannick Francken, Johannes Taelman, and Philippe Bekaert. Shadow multiplexing for real-time silhouette extraction. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2009. 84

Shengyang Dai and Ying Wu. Removing partial blur in a single image. In *IEEE Conference on Compuer Vision and Pattern Recognition (CVPR)*, pages 2544–2551, 2009. 38, 40, 41, 42

Paul Debevec. Virtual cinematography: Relighting through computation. *Computer*, 39(8):57–65, 2006. 5

E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959. 62

Yuanyuan Ding, Jing Xiao, Kar-Han Tan, and Jingyi Yu. Catadioptric projectors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009. 89

Jamil Draréni, Sébastien Roy, and Peter Sturm. Geometric video projector auto-calibration. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2009. 93

R. Fabian and D. Malah. Robust identification of motion and out-of-focus blur parameters from blurred and noisy images. *Graphical Models and Image Processing (CVGIP)*, 53:403–412, 1991. 31, 32, 33, 59

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. *ACM Transactions on Graphics*, 25(3):787–794, 2006. xv, 39, 41, 44, 74, 75, 81

Sergio Fernandez, Joaquim Salvi, and Tomislav Pribanic. Absolute phase mapping for one-shot dense pattern projection. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 64–71. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543483`. 83

Mark Fiala. Vision guided control of multiple robots. In *Canadian Conference on Computer and Robot Vision*, pages 241–246. IEEE Computer Society, 2004. ISBN 0-7695-2127-4. 91

Mark Fiala. Automatic projector calibration using self-identifying patterns. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2005. 93

Kensaku Fujii, Michael D. Grossberg, and Shree K. Nayar. A projector-camera system with real-time photometric adaptation for dynamic environments. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 814–821, 2005. 90, 93

Ryo Furukawa, Kenji Inose, and Hiroshi Kawasaki. Multi-view reconstruction for projector camera systems based on bundle adjustment. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2009. 84

Donald B. Gennery. Determination of optical transfer function by inspection of frequency-domain plot. *Journal of the Optical Society of America*, 63(12): 1571–1577, 1973. 31

Helmut Gernsheim. The 150th anniversary of photography. *History of Photography*, 1(1):1:3–8, January 1977. 2

Jessica Gorman. Photography at a crossroads. *Science News*, 162(21):331–333, November 2002. URL `http://www.jstor.org/stable/4013861`. 2

Andreas Griesser and Luc Van Gool. Automatic interactive calibration of multi-projector-camera systems. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2006. 93

Michael D. Grossberg, Harish Peri, Shree K. Nayar, and Peter N. Belhumeur. Making one object look like another: Controlling appearance using a projector-camera system. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 452–459, 2004. 90

Max Grosse and Oliver Bimber. Coded aperture projection. In *International Workshop on Immersive Projection Technologies/Emerging Display Technologiges (IPT/EDT)*, pages 1–4, 2008. 91, 93

Max Grosse, Gordon Wetzstein, Anselm Grundhöfer, and Oliver Bimber. Adaptive coded aperture projection. In *ACM SIGGRAPH*, pages 68:1–68:1. ACM, 2009. URL `http://doi.acm.org/10.1145/1597990.1598058`. 93

Max Grosse, Gordon Wetzstein, Anselm Grundhöfer, and Oliver Bimber. Coded aperture projection. *ACM Transactions on Graphics*, 29:22:1–22:12, 2010. ISSN 0730-0301. URL `http://doi.acm.org/10.1145/1805964.1805966`. 93

Anselm Grundhöfer, Manja Seeger, Ferry Häntsch, and Oliver Bimber. Dynamic adaptation of projected imperceptible codes. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, 2007. 90, 93

Ankit Gupta, Neel Joshi, C. Lawrence Zitnick, Michael Cohen, and Brian Curless. Single image deblurring using motion density functions. In *European Conference on Computer Vision (ECCV)*, pages 171–184. Springer, 2010. URL `http://www.springerlink.com/index/Y07467510533568U.pdf`. 40, 41, 43

Stefan Harmeling, Hirsch Michael, and Bernhard Schölkopf. Online blind deconvolution for astronomical imaging. In *International Conference on Computational Photography (ICCP)*, pages 1–7, 2009. 40

Stefan Harmeling, Hirsch Michael, and Bernhard Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems (NIPS)*, pages 829–837, 2010. 38, 40, 41, 43

Jeremy M. Heiner, Scott E. Hudson, and Kenichiro Tanaka. The information percolator: Ambient information display in a decorative object. In *ACM Symposium on User Interface Software and Technology (UIST)*, pages 141–148. ACM Press, 1999. x, 7, 8

Michael Hirsch, Suvrit Sra, Bernhard Scholkopf, and Stefan Harmeling. Efficient filter flow for space-variant multiframe blind deconvolution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 607–614. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5540158`. 40, 43

Hanyu Hong and In Kyu Park. Single image motion deblurring using anisotropic regularization. In *International Conference on Image Processing (ICIP)*, pages 1149–1152. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5650775`. 41

Tingbo Hou, Sen Wang, Hong Qin, and Rodney L. Miller. Image deconvolution using multigrid natural image prior and its applications. In *International Conference on Image Processing (ICIP)*, pages 3569–3572. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651083`. 25

Zhe Hu, Jia-Bin Huang, and Ming-Hsuan Yang. Single image deblurring with adaptive dictionary learning. In *International Conference on Image Processing (ICIP)*, pages 1169–1172. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651892`. 40

Po-Hao Huang, Yu-Mo Lin, Hao-Liang Yang, and Shang-Hong Lai. Image deblurring by exploiting inherent bi-level regions. In *International Conference on Image Processing (ICIP)*, pages 1321–1324, 2009. 40, 41

Aapo Hyvärinen, Jarmo Hurri, and Patrik O. Hoyer. *Natural Image Statistics – A Probabilistic Approach to Early Computational Vision*. Springer-Verlag, 2009. 120

Christopher Jaynes, Stephen Webb, and R. Matt Steele. Camera-based detection and removal of shadows from interactive multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics*, 10(3):290–301, 2004. 85

Luke Jefferson and Richard Harvey. An interface to support color blind computer users. In *Conference on Human Factors in Computing Systems (SIGCHI)*, pages 1535–1538. ACM, 2007. ISBN 978-1-59593-593-9. URL `http://doi.acm.org/10.1145/1240624.1240855`. 86

Hui Ji and Chaoqiang Liu. Motion blur identification from image gradients. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. 31, 59

Jiaya Jia. Single image motion deblurring using transparency. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007. 41, 42

Tyler Johnson and Henry Fuchs. Real-time projector tracking on complex geometry using ordinary imagery. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2007. 89, 92

Andrew Jones, Magnus Lang, Graham Fyffe, Xueming Yu, Jay Busch, Ian McDowall, Mark Bolas, and Paul Debevec. Achieving eye contact in a one-to-many 3d video teleconferencing system. *ACM Transactions on Graphics (SIGGRAPH)*, 28:64:1–64:8, 2009. ISSN 0730-0301. URL `http://doi.acm.org/10.1145/1531326.1531370`. 6

Andrew Jones, Ian McDowall, Hideshi Yamada, Mark Bolas, and Paul Debevec. Rendering for an interactive 360° light field display. *ACM Transactions on Graphics (SIGGRAPH)*, 26:1–10, 2007. ISSN 0730-0301. URL `http://doi.acm.org/10.1145/1276377.1276427`. 6

Neel Joshi, Sing Bing Kang, Lawrence C. Zitnick, and Richard Szeliski. Image deblurring using inertial measurement sensors. *ACM Transactions on Graphics (SIGGRAPH)*, 29:30:1–30:9, 2010. ISSN 0730-0301. URL `http://doi.acm.org/10.1145/1778765.1778767`. 20, 44

Neel Joshi, Richard Szeliski, and David J. Kriegman. Psf estimation using sharp edge prediction. In *IEEE Conference on Compuer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. 40, 41

Neel Joshi, C. Lawrence Zitnick, Richard Szeliski, and David J. Kriegman. Image deblurring and denoising using color priors. In *IEEE Conference on Compuer Vision and Pattern Recognition (CVPR)*, pages 1550–1557, 2009. 26

Xiumei Kang, Qingjin Peng, Gabriel Thomas, and Chunsheng Yu. Blind image restoration using the cepstrum method. In *Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1952–1955, 2006. 31, 32, 54

Hiroshi Kawasaki, Inose Kenji, Toshihiro Kawasaki, Ryo Furukawa, Ryusuke Sagawa, and Yasushi Yagi. Projector camera system for realtime 3d scanning. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2008. 83

Hiroshi Kawasaki, Ryusuke Sagawa, Yasushi Yagi, Ryo Furukawa, Naoki Asada, and Peter Sturm. One-shot scanning method using an uncalibrated projector and camera system. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 104–111, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5544604`. 83

KonicaMinolta. Anti-Shake. `http://ca.konicaminolta.com/products/consumer/digital_camera/slr/dynax-7d/02.html`, 2003. 11

Dilip Krishnan and Rob Fergus. Fast image deconvolution using hyper-Laplacian priors. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1033–1041, 2009. 25

Deepa Kundur and Dimitrios Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, 1996. 27

Johnny C. Lee, Paul H. Dietz, Dan Maynes-Aminzade, Ramesh Raskar, and Scott E. Hudson. Automatic projector calibration with embedded light sensors. In *ACM Symposium on User Interface Software and Technology (UIST)*, 2004. 93

Johnny C. Lee, Scott E. Hudson, Jay W. Summet, and Paul H. Dietz. Moveable interactive projected displays using projector based tracking. In *ACM Symposium on User Interface Software and Technology (UIST)*, pages 63–72, 2005. 93

Man Chuen Leung, Kai Ki Lee, Kin Hong Wong, and Michael Ming Yuen Chang. A projector-based movable hand-held display system. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009. 93

Anat Levin. Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems (NIPS)*, pages 841–848, 2006. 42

Anat Levin, Rob Fergus, Fredo Durand, and William T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics*, 26:1–10, 2007. xi, xii, 25, 26, 34, 36, 37, 74

Anat Levin, Yair Weiss, Fredo Durand, and William T. Freeman. Understanding and evaluating blind deconvolution algorithms. Technical report, MIT, 2009. 27

Marc Levoy. Light fields and computational imaging. *Computer*, 39(8):46–55, 2006. 5

Marc Levoy and Pat Hanrahan. Light field rendering. In *SIGGRAPH*, pages 31–42, 1996. 5

Dalong Li, Russell M. Mersereau, and Steven Simske. Blur identification based on kurtosis minimization. In *International Conference on Image Processing (ICIP)*, pages 20–23, 2005. 34, 35

Peter Lincoln, Greg Welch, Andrew Nashel, Andrei State, Adrian Ilie, and Henry Fuchs. Animatronic shader lamps avatars. *Virtual Reality*, 15(2-3):225–238, 2011. x, 6, 7, 9

L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745–754, 1974. 23, 111

Aditi Majumder and Michael S. Brown. *Practical Multi-Projector Display Design*. A. K. Peters LTD., 2007. ISBN 1568813104. 87, 90

Aditi Majumder, Robert G. Brown, and Hussein S. El-Ghoroury. Display gamut reshaping for color emulation and balancing. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 17–24. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543467`. 93

Aditi Majumder, Zhu He, Herman Towles, and Greg Welch. Achieving color uniformity across multi-projector displays. In *International Conference on Visualization (VIS)*, pages 117–124, 2000. 93

Aditi Majumder, David Jones, Matthew McCrory, Michael E Papka, and Rick Stevens. Using a camera to capture and correct spatial photometric variation in multi-projector displays. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2003. 93

Wikky Fawwaz Al Maki and Sueo Sugimoto. Blind deconvolution algorithm for spatially-invariant motion blurred images based on inverse filtering and dst. *International Journal of Circuits, Systems and Signal Processing*, 1:92–100, 2007. 54

MathWorks. Matlab function deconvblind. `http://www.mathworks.com/help/toolbox/images/ref/deconvblind.html`. xv, 74, 81

C. Mayntz, T. Aach, and D. Kunz. Blur identification using a spectral inertia tensor and spectral zeros. In *International Conference on Image Processing (ICIP)*, pages 885–889, 1999. 31, 32, 33, 59

Christoffer Menk and Reinhard Koch. Physically-based augmentation of real objects with virtual content under the influence of ambient light. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 25–32. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543472`. 86

James Miskin and David J. C. MacKay. Ensemble learning for blind image separation and deconvolution. In *Advances in Independent Component Analysis*. Springer-Verlag, 2000. 74

Andrew Vande Moere. Beyond the tyranny of the pixel: Exploring the physicality of information visualization. In *International Conference on Information Visualisation*, pages 469–474. IEEE Computer Society, 2008. ISBN 978-0-7695-3268-4. URL `http://portal.acm.org/citation.cfm?id=1439280.1440187`. 7

Mohsen Ebrahimi Moghaddam and Mansour Jamzad. Motion blur identification in noisy images using mathematical models and statistical measures. *Pattern Recognition*, 40:1946–1957, 2007. 31, 32, 33, 54, 59

NAO Design. infernoptix - digital pyrotechnic matrix. `http://www.infernoptix.com/`. x, 7, 8

Shree K. Nayar. Computational cameras: Redefining the image. *Computer*, 39 (8):30–38, 2006. 5

Shree K. Nayar, Harish Peri, Michael D. Grossberg, and Peter N. Belhumeur. A projection system with radiometric compensation for screen imperfections. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2003. 90, 93

NEC Display Solutions Europe. Keystone correction. `http://www.nec-display-solutions.com/p/edu/en/technology/keystonecorrect.xhtml?e=e2s1`, a. 11

NEC Display Solutions Europe. Wall colour correction. `http://www.nec-display-solutions.com/p/edu/en/technology/wallcolorcorrect.xhtml?e=e3s1`, b. 11

NEC Display Solutions, Ltd. Autofocus. `http://www.nec-display.com/global/technology/con_pj_adjust.html`. 20

Ren Ng, Marc Levoy, Mathieu Bredif, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light field photography with a hand-held plenoptic camera. Technical report, Stanford University, 2005. 5

Nikon. Vibration Reduction. `http://imaging.nikon.com/products/imaging/technology/vr/index.htm`, 2000. 10, 11, 20

Kazuki Nishi and Tsubasa Onda. Evaluation system for camera shake and image stabilizers. In *International Conference on Multimedia and Expo (ICME)*, pages 926–931, 2010. xii, 46, 47, 48, 49

Takayuki Okatani and Koichiro Deguchi. Autocalibration of a projector-camera system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 (12):1845–1855, 2005. ISSN 0162-8828. 93

Takayuki Okatani and Koichiro Deguchi. Easy calibration of a multi-projector display system. *International Journal of Computer Vision*, 85:1–18, 2009. 93

João P. Oliveira, Mário A.T. Figueiredo, and José M. Bioucas-Dias. Blind estimation of motion blur parameters for image deconvolution. In *3rd Iberian conference on Pattern Recognition and Image Analysis, Part II (IbPRIA)*, pages 604–611, 2007. 31, 32, 59

Bruno A. Olshausen and David J. Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37(23):3311–3325, 1997. URL http://www.sciencedirect.com/science/article/B6T0W-494SR70-19/2/b4c138506b06df6f332ced73e8501a3e. 35

Stanley Osher and Leonid I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 27:919–940, 1990. ISSN 0036-1429. URL http://portal.acm.org/citation.cfm?id=80058.80065. 40

Yuji Oyamada and Hideo Saito. New type face display with augmented reality system (in japanese). In *IEICE General Conference*, page 257, 2006. xi, 9, 11

Yuji Oyamada and Hideo Saito. Focal pre-correction of projected image for deblurring screen image. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2007. xvi, xviii, 102, 103, 105

Hanhoon Park, Moon-Hyun Lee, Byung-Kuk Seo, Yoonjong Jin, and Jong-Il Park. Content adaptive embedding of complementary patterns for nonintrusive direct-projected augmented reality. In *International Conference on Virtual Reality (ICVR)*, pages 132–141. Springer-Verlag, 2007. ISBN 978-3-540-73334-8. URL http://portal.acm.org/citation.cfm?id=1770090.1770106. 93

Hanhoon Park, Byung-Kuk Seo, and Jong-Il Park. A nonintrusive method for generating all-focused projection. In *International Conference on Image Processing (ICIP)*, pages 529–532, 2008. 91, 93, 105

Jiyoung Park and Myoung-Hee Kim. Interactive display of image details using a camera-coupled mobile projector. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 9–16. IEEE, 2010. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543466. 87

Sang-Cheol Park, Hyoung-Suk Lee, and Seong-Whan Lee. Qualitative estimation of camera motion parameters from the linear composition of optical flow. *Pattern Recognition*, 37(4):767–779, 2004. 46, 47

Youngmin Park, Vincent Lepetit, and Woontack Woo. Esm-blur: Handling & rendering blur in 3d tracking and augmentation. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 163–166, 2009. ISBN 978-1-4244-5390-0. URL http://dx.doi.org/10.1109/ISMAR.2009.5336480. 109

Eli Peli, Estella Lee, Clement L. Trempe, and Sheldon Buzney. Image enhancement for the visually impaired: The effects of enhancement on face recognition. *Journals of Optical Society of America A*, 11(7):1929–1939, 1994. URL http://josaa.osa.org/abstract.cfm?URI=josaa-11-7-1929. 86

Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Transactions on Graphics*, 25 (3):795–804, 2006. 22

Ramesh Raskar, Matt Cutts, Greg Welch, and Wolfgang Stuerzlinger. Efficient image generation for multiprojector and multisurface displays. In *Eurographics Workshop on Rendering*, pages 139–144, 1998a. 85, 89

Ramesh Raskar, Jeroen van Baar, Paul Beardsley, Thomas Willwacher, Srinivas Rao, and Clifton Forlines. iLamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics*, 22:809–818, 2003. 89

Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *ACM SIGGRAPH*, pages 179–188, 1998b. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.1129. xv, 85, 93

Ramesh Raskar, Greg Welch, Kok lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Eurographics Workshop on Rendering*, pages 89–102, 2001. x, 6, 7, 9, 87, 89

Alex Rav-Acha and Shmuel Peleg. Two motion-blurred images are better than one. *Pattern Recognition Letters*, 26(3):311–317, 2005. 20, 38, 44

Stanley J. Reeves and Russell M. Mersereau. Blur identification by the method of generalized cross-validation. *IEEE Transactions on Image Processing*, 1(3): 301–311, 1992. 37

William Hadley Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972. 23, 111

David E. Roberts. History of lenticular and related autostereoscopic methods. Leap Technologies, 2003. URL `http://www.microlens.com/pdfs/history_of_lenticular.pdf`. 6

Paul Rodriguez and Brendt Wohlberg. A generalized vector-valued total variation algorithm. In *International Conference on Image Processing (ICIP)*, pages 1309–1312, 2009. 24

Raphael Rom. On the cepstrum of two-dimensional functions. *IEEE Transactions on Information Theory*, 21(2):214–217, 1975. 30, 31

Filip Rooms, Wilfried Philips, and Javier Portilla. Parametric PSF estimation via sparseness maximization in the wavelet domain. In *SPIE - Wavelet applications in industrial processing II*, pages 26–33, 2004. 34, 35

Joaquim Salvi, Jordi Pages, and Joan Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, 2004. 91

Andreas E. Savakis and H. Joel Trussell. Blur identification by residual spectral matching. *IEEE Transactions on Image Processing*, 2(2):141–151, 1993. 34, 36, 66

Christoph Schmalz and Elli Angelopoulou. A graph-based approach for robust single-shot structured light. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 80–87. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543492`. 84

Steven M. Seitz and Simon Baker. Filter flow. In *International Conference on Computer Vision (ICCV)*, pages 143–150, 2009. 43

Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. *ACM Transactions on Graphics*, 24:745–755, 2005. ISSN 0730-0301. URL `http://doi.acm.org/10.1145/1073204.1073257`. 89

Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics*, 27(3):1–10, 2008. 27, 39, 41

Qi Shan, Wei Xiong, and Jiaya Jia. Rotational motion deblurring of a rigid object from a single image. In *International Conference on Computer Vision (ICCV)*, pages 1–8, 2007. 41, 42

Marvin K. Simon. *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*. Springer-Verlag, 2002. ISBN 0387346570. 27

Yoshiko Sugaya, Isao Miyagawa, and Hideki Koike. Contrasting shadow for occluder light suppression from one-shot image. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 96–103. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5544603`. 85, 86

Hongwei Sun, Michel Desvignes, and Yunhui Yan. Motion blur adaptive identification from natural image model. In *International Conference on Image Processing (ICIP)*, pages 137–140, 2009. 31, 32

Wei Sun, Irwin Sobel, Bruce Culbertson, Dan Gelb, and Ian Robinson. Calibrating multi-projector cylindrically curved displays for ”wallpaper” projection. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2008a. 93

Weibin Sun, Xubo Yang, Shuangjiu Xiao, and Wencong Hu. Robust checkerboard recognition for efficient nonplanar geometry. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2008b. 93

Yu-Wing Tai, Ping Tan, and Michael S. Brown. Richardson-lucy deblurring for scenes under projective motion path. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, TBD:TBD, 2011. 24

Toru Takahashi, Tatsuya Kawano, Koichi Ito, Takafumi Aoki, and Satoshi Kondo. Multi-projector display using SIFT and phase-only correlation. In *International Conference on Image Processing (ICIP)*, pages 1189–1192, 2010. 92

Toru Takahashi, Norihito Numa, Takafumi Aoki, and Satoshi Kondo. A geometric correction method for projected images using SIFT feature points. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2008. 92

Kah-Chye Tan, Hock Lim, and B. T. G. Tan. Restoration of real-world motion-blurred images. *CVGIP: Graphical Models and Image Processing*, 53(3):291–299, 1991. URL `http://www.sciencedirect.com/science/article/B6WDC-4D7CT66-21/2/6523df54c8d1ebbb146d9192eb6b414b`. 34

Jean-Philippe Tardif, Sébastien Roy, and Martin Trudeau. Multi-projectors for arbitrary surfaces without explicit calibration nor reconstruction. In *International Conference on 3-D Digital Imaging and Modeling (3DIM)*, pages 217–224, Oct. 2003. 93

Murat A. Tekalp, Howard Kaufman, and John W. Woods. Identification of image and blur paramters for the restoration of noncausal blurs. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34(4):963–972, 1986. 37

C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *International Conference on Computer Vision (ICCV)*, pages 839–846, 1998. 40

Antonio Torralba and Aude Oliva. Statistics of natural image categories. *Network: Computation in Neural Systems*, 14:391–412, 2003. 38, 118

John W. Tukey. An introduction to the calculations of numerical spectrum analysis. In *Spectral Analysis of Time Series*, pages 25–46, 1967. 55, 62

Michal Šorel and Filip Šroubek. Space-variant deblurring using one blurred and one underexposed image. In *International Conference on Image Processing (ICIP)*, pages 157–160, 2009. 40, 43

Yi Wan and Robert Nowak. A bayesian multiscale approach to joint image restoration and edge detection. In *in Wavelet Applications in Signal and Image Processing VII, Proc. SPIE*, pages 73–84, 1999. 39

Chao Wang, LiFeng Sun, ZhuoYuan Chen, ShiQiang Yang, and JianWei Zhang. Robust inter-scale non-blind image motion deblurring. In *International Conference on Image Processing (ICIP)*, pages 149–152, 2009. 27

Dong Wang, Imari Sato, Takahiro Okabe, and Yoichi Sato. Radiometric compensation in a projector-camera system based on the properties of human vision system. In *International Workshop on Projector-Camera Systems (PROCAMS)*, 2005. 90

Yair Weiss and William T. Freeman. What makes a good model of natural images? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007. 27

Gordon Wetzstein and Oliver Bimber. Radiometric compensation through inverse light transport. In *Pacific Conference on Computer Graphics and Applications (PG)*, page 38, 2007. 89

Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 491–498. IEEE, 2010. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5540175`. 43

Norbert Wiener. Extrapolation, interpolation, and smoothing of stationary time series, 1949. 22, 66, 67, 74, 99

Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. High performance imaging using large camera arrays. In *SIGGRAPH*, pages 765–776, 2005. 5

James S. Wolffsohn, Ditipriya Mukhopadhyay, and Martin Rubinstein. Image enhancement of real-time television to benefit the visually impaired. *American Journal of Ophthalmology*, 144(3):436–440, 2007. ISSN

0002-9394. URL http://www.sciencedirect.com/science/article/B6VK5-4P6MC0N-3/2/263ae4cb68aa4c782b7243818a0ac930. 86

Shiqian Wu, Zhongkang Lu, Ee Ping Ong, and Weisi Lin. Blind image blur identification in cepstrum domain. In *16th International Conference on Computer Communications and Networks (ICCCN)*, pages 1166–1171, 2007. 31, 32, 59

Xbox.com. Kinect. http://www.xbox.com/en-US/Kinect. 8

Feng Xiao, Amnon Silverstein, and Joyce Farrell. Camera–motion and effective spatial resolution. In *International Congress of Imaging Science*, pages 33–36, 2006. xii, 45, 47, 48

Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *European Conference on Computer Vision (ECCV)*, pages 157–170. Springer, 2010. URL http://www.springerlink.com/index/Y24XK016678HT3V3.pdf. 38, 40, 41

Satoshi Yamamoto, Yasumasa Itakura, Masashi Sawabe, Gimpei Okada, Norimichi Tsumura, and Toshiya Nakaguchi. Precomputed ROMP for light transport acquisition. In *International Workshop on Projector-Camera Systems (PROCAMS)*, pages 49–56. IEEE, 2010. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5543460. 89

Masaki Yamazaki and Gang Xu. 3d reconstruction of glossy surfaces using stereo cameras and projector-display. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1213–1220. IEEE, 2010. URL http://www.computer.org/portal/web/csdl/doi/10.1109/CVPR.2010.5539830. 84

Ruigang Yang and Greg Welch. Automatic and continuous projector display surface calibration using every-day imagery. In *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2001. 92, 105

Y. Yitzhaky and N. S. Kopeika. Identification of blur parameters from motion blurred images. *Graphical Models and Image Processing (CVGIP)*, 59(5):310–320, 1997. 33, 34

Tatsuo Yotsukura, Shigeo Morishima, Frank Nielsen, Kim Binsted, and Claudio S. Pinhanez. Hypermask - projecting a talking head onto a real object. *The Visual Computer*, 18(2):111–120, April 2002. 9

Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics*, 26(3):1, 2007. 20, 40, 41

Li Zhang and Shree Nayar. Projection defocus analysis for scene capture and image display. *ACM Transactions on Graphics*, 25(3):907–915, 2006. 91, 93

Stefanie Zollmann and Oliver Bimber. Imperceptible calibration for radiometric compensation. In *Eurographics (short paper)*, pages 61–64, 2007. 93

Stefanie Zollmann, Tobias Langlotz, and Oliver Bimber. Passive-active geometric calibration for view-dependent projections onto arbitrary surfaces. *Journal of Virtual Reality and Broadcasting*, 4(6):1–11, 2007. 93

Publication list

Journal papers

Yuji Oyamada, Haruka Asai, and Hideo Saito. Blind deconvolution for a curved motion based on cepstral analysis. *IPSJ Transactions on Computer Vision and Applications (CVA)*, 3:32–43, 2011.

Peer-reviewed conference papers

Haruka Asai, **Yuji Oyamada**, Julien Pilet, and Hideo Saito. Cepstral analysis based blind deconvolution for motion blur. In *IEEE International Conference on Image Processing (ICIP)*, pages 1153–1156, 2010.

Yuji Oyamada and Hideo Saito. Focal pre-correction of projected image for deblurring screen image. In *IEEE International Workshop on Projector-Camera Systems (PROCAMS)*, pages 1–8, 2007a.

Yuji Oyamada and Hideo Saito. Pre-correction of projected images for deblurring on projector screen. In *Japan-Korea Joint Workshop on Frontiers of Computer Vision (FCV)*, 2007b.

Yuji Oyamada and Hideo Saito. Defocus blur correcting projector-camera system. In *Advanced Concepts for Intelligent Vision Systems (ACIVS)*, pages 453–464, 2008a.

Yuji Oyamada and Hideo Saito. Estimation of projector defocus blur by extracting texture rich regionin projection image. In *International*

Conference in Central Europe on Computer Graphics,Visualization and Computer Vision (WSCG), pages 153–160x, 2008b.

Yuji Oyamada and Hideo Saito. Blind deconvolution based projector defocus removing with uncalibrated projector-camera pair. In *IEEE International Workshop on Projector-Camera Systems (PROCAMS)*, 2009.

Yuji Oyamada, Hideo Saito, Koji Ootagaki, and Mitsuo Eguchi. Cepstrum based blind image deconvolution. In *International Workshop on Vision, Communications and Circuits (IWVCC)*, pages 197–200, 2008.

Others

Yoshimitsu Aoki, Yuko Uematsu, **Yuji Oyamada**, Ugur Tumerdem, Baris Yalcin, Wataru Yamanouchi, and Yuki Yokokura. Panel discussion: Multi-sensory telepresence in access space. International Workshop on Vision and Control for Access Space, 2009.

Yuji Oyamada, Haruka Asai, and Hideo Saito. Single image blind deconvolution: Psf estimation methods based on cepstral analysis. CVL-HVRL Joint Workshop, 2010.

Yuji Oyamada, Yuichiro Ikuma, Tomoya Miyanishi, Toyofumi Ishikawa, Rii Hirano nd Lin I-Te, Juan Camilo Corena, Haruki Nishimura, and Yuki Yokokura. Panel session: ”access space: This is my proposal”. International Symposium on Access Space, 2011.

Yuji Oyamada and Hideo Saito. Projector defocus deblurring. G-COE Winter camp, 2008.

Yuji Oyamada and Hideo Saito. Image based projector image enhancement. MSRA Keio-GCOE Workshop, 2010a.

Yuji Oyamada and Hideo Saito. Visual image enhancement by post/pre-image restoration approaches. G-COE Winter camp, 2010b.

Domestic journal papers

小山田 雄仁, 斎藤 英雄. プロジェクタスクリーン上に生じる焦点ボケ解消のための投影画像の事前補正法. バーチャルリアリティ学会論文誌, 12 巻 4 号, 479–486 頁, 2007.

Domestic conferences

小山田 雄仁, 斎藤 英雄. プロジェクタスクリーンにおけるボケ除去のための投影画像の事前補正法. 映像情報メディア学会 2006 年冬季大会, 2006a.

小山田 雄仁, 斎藤 英雄. 拡張現実を用いた顔の立体表示システム. 電子情報通信学会総合大会, 2006b.

小山田 雄仁, 斎藤 英雄. Projection based augmented reality のための投影画像の事前補正によるボケ除去法. 電子情報通信学会技術報告, 2007a.

小山田 雄仁, 斎藤 英雄. 投影面上に生じる焦点ボケ補正のための投影画像の事前補正法. 画像の認識・理解シンポジウム (MIRU), 2007b.

小山田 雄仁, 斎藤 英雄. 特徴領域抽出に基づくプロジェクタの焦点ボケ補正. 画像の認識・理解シンポジウム (MIRU), 2008a.

小山田 雄仁, 斎藤 英雄. 表示画像の特徴領域抽出に基づくプロジェクターカメラシステムの焦点ボケの自動補正法. 電子情報通信学会技術報告, 2008b.

小山田 雄仁, 斎藤 英雄. 補正画像の繰り返し投影によるプロジェクタの焦点ボケ推定. 電子情報通信学会総合大会, 2008c.

小山田 雄仁, 斎藤 英雄, 太田垣 康二, 江口 満男. 劣化画像のケプストラムを利用した手ブレ補正. 電子情報通信学会技術報告, 2008.

小山田 雄仁, 浅井 晴香, 斎藤 英雄, 太田垣 康二, 江口 満男. ケプストラムに基づいたブレ画像補正. ビジョン技術の実利用ワークショップ (ViEW), 2008.

浅井 晴香, 小山田 雄仁, 斎藤 英雄. ケプストラム解析を用いたブレ画像の blind deconvolution. 画像の認識・理解シンポジウム (MIRU), 2009a.

浅井 晴香, 小山田 雄仁, 斎藤 英雄. ケプストラム解析を用いたブレ画像の blind deconvolution. CVIM 研究会, 2009b.

浅井 晴香, 小山田 雄仁, 斎藤 英雄. Blind deconvolution のための劣化画像のケプストラム解析. CVIM 研究会, 2010.

中村 隆之, 小山田 雄仁, 斎藤 英雄. プロジェクターカメラシステムによる移動非剛体に対する動的テクスチャリングに向けた基礎検討. CVIM 研究会, 2010.

中島 悠, 野崎 貴裕, 小山田 雄仁, 大西 公平. カメラを用いた物体空間座標系におけるバイラテラル制御系. 平成 22 年度電気学会産業計測制御研究会, 2011.

Magazines

小山田 雄仁, 斎藤 英雄. プロジェクタ投影画像の焦点ボケ補正. 月刊画像ラボ 解説記事, 2007 年 8 月号.

浅井 晴香, 小山田 雄仁, 斎藤 英雄. ケプストラム解析に基づいたブレ画像の blind deconvolution. OplusE 解説記事, 2010 年 3 月号.