Image mosaicing:
Create High Quality Panoramic Multi-spectral Image

by

Akifumi Ikari

A Senior Thesis

Submitted to
the Department of Information Science
Faculty of Science
the University of Tokyo
on February 12, 2002
in partial fulfillment of the requirements
for the Degree of Bachelor of Science

Thesis Supervisor: Katsushi Ikeuchi
Professor of Institute of Industrial Science

Abstract

In computer graphics and virtual reality, multi-spectral and panoramic images are needed. To get a panoramic multi-spectral images, the approach to use a filter transmitting spectral bands are spatially varying is proposed. A filter is attached to an 8-bit black/white camera. Then take images (there is no parallax).Each scene points are taken many times in different spectral band. After that mosaic images. Here, this image mosaicing have problems which different from traditional RGB-image mosaicing. First, illumination and spectral response greatly effect images. Second, because of using spatially varying filter, value of the scene point of each images are different depend on the position on the each images. Last, if put one upon another in a little shifted, the multi-spectral-data are shifted greatly. So exact mosaicing are needed.

This research propose the way of making rotational panoramas using images these are taken by camera which the filter is attached, and the way of mosaicing which is faster and more accurate.

Acknowledgement

I am more deeply than the heart thankful to Professor Katsushi Ikeuchi, Ko Nishino, and Robby T. Tan, who had many precious opinions enjoyed in creating this paper.

Moreover, I say gratitude also to you of the Ikeuchi laboratory who had you cooperate in order to advance research deeply at this place.

Contents

1. Introduction

Mosaicing is connecting two or more images and making a new and wide area image. In the case of a part of the needed scene can be taken at once for restriction of the resolution of a camera, a photography angle, etc, by taking the scene many times so that a part of image should be overlapped, and mosaicing the images, the scene can obtained. Thereby, 360-degree panorama picture etc. can be created. At this time, in mosaicing, it is the biggest problem how the position relation between two or more images is drawn. A highly precise integrated image can be obtained by drawing this relation precisely.

The various method of Mosaicing until now is proposed. The images which performs mosaicing does not include parallax fundamentally. And there are two kinds of images in which does not include parallax. One is the images which took a scene that is plane or approximated as a plane, and another is which were taken by rotate a camera at the main point.

In the case of the former mosaicing, it is known that the correspondence relation between pictures is expressed by eight parameters. Also in it, relation of the images which the distance of a camera from the plane is fixed and a camera is perpendicular to the plane can be expressed by three parameters of parallel / rotation movement.
And the determination of these eight or three parameters becomes very important in mosaicing. One of the general methods of taking out these parameters is to get the answer by make small the second power sum of the difference between two images by changing a parameter little by little. However, this method requires time very much. So the method of take out these parameters gradually by reducing the number of pixels of images, or by change these parameters only in the direction in which the second power sum become smaller. As other methods to take out these parameters, Chiba et al. [3] performed the feature extraction of an images, determined eight parameters by asking for movement of the 4 same points between pictures. Moreover, Chiba et al. [3] performed mosaicing also to the images in which include parallax by dividing a domain into a triangular patch which contain the same plane

Next, in the case of the latter, the images is taken by a camera is moved only pan rotation. Here, pan rotation is the rotation to the horizontal direction of a camera. And the images is mapped for a ball or a pillar. And the relation of the images expressed by three parameters. Then it is required that the focal length of a camera is known. And if the focal length is not accurate, it will become impossible to make an accurate image. Therefore, some techniques of rectifying the mistake of a focal length are proposed.

Next, there is another method of using a fish-eye lens etc. to obtain a panorama image.

A highly precise panorama picture can be obtained panoramic images from one image by easy conversion without mosaicing by using fish-eye lens. However, when a fish-eye lens is used, so the data of all directions are in one image, the resolution of panoramic image became low.

Multispectral is having not only the data of RGB but also having the data in much more wavelength. So multispectral images have many color information, that make possible it to process in the images such as create a new image that the light source is changed, etc. And there are the method of using a camera which can take multispectral a image, and the method of mosaic images, like Yoav at al. [5], which taken by a camera which attached spatially varying filter, etc, as method of obtain a multispectral image.

In this paper, we obtain highly precise plane and panoramic multispectral images by mosaicing carrying out edge detection of the image which taken by a camera which attached spatially varying filter. Section 2 explain about spatially varying filter. Section 3 explain about distortion compensation of a lens, edge detection and mosaicing. Section 4 explain making a RGB image. Section 5 explain how to create new images that the light source were changed.

2. Spatially Varying Filter

In this research, spatially varying filter is used. So, at first, explain about spatially varying filter. The wavelength of penetration light crosses filter is change across the filter in visible light from 700nm to 400nm. And this filter is attached to an 8-bit monochrome camera. If it says simply, in a image which taken with a camera attached this filter, the object of red, green, and blue will be took on the left of a picture, middle, and the right, respectively.

Then, if make a camera regularity move and take images and mosaic these, a multispectral image can be obtained. However, only the central part of the mosaiced image had data in full-visible wavelength, and the portion of the data has broken off around the central part. However, there is no broken off if creating a 360-degree panorama picture.

The advantages of obtaining a multispectral image by mosaic images which taken by a camera which spatially varying filter attached is only need to have monochromic camera and this filter, and being able to obtain a image which have information of the arbitrary numbers of wavelength.

Figure 1:Spatially varying filter attached to a camera

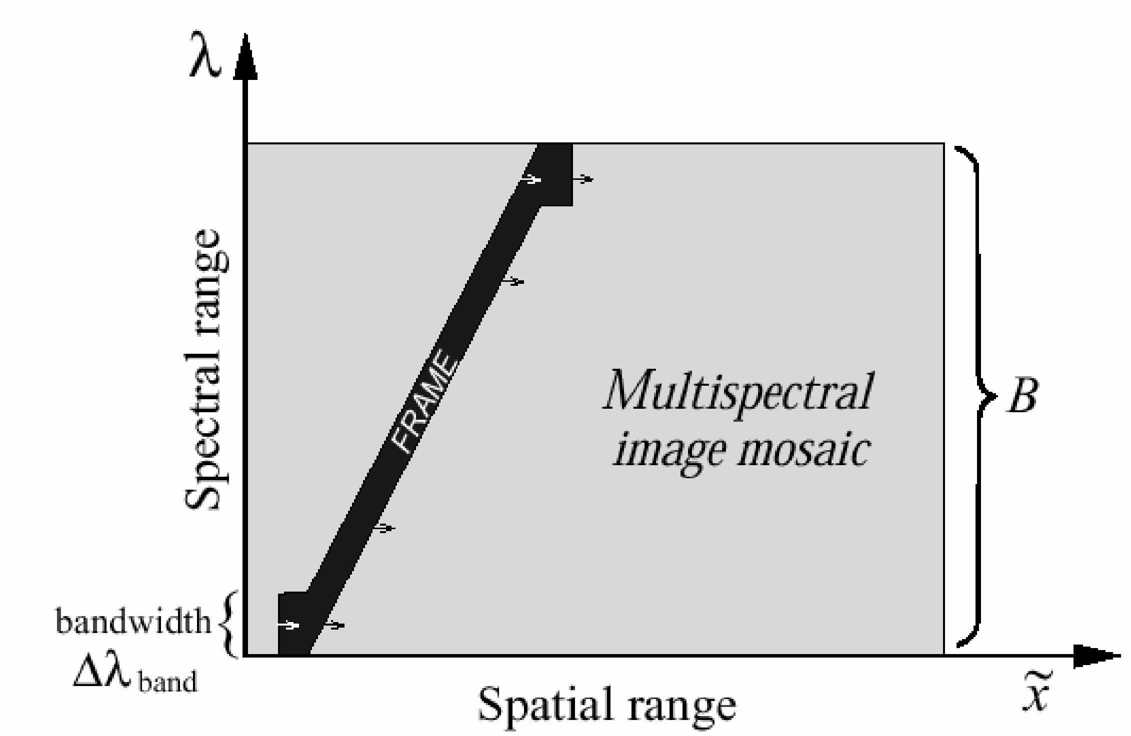


Figure 2

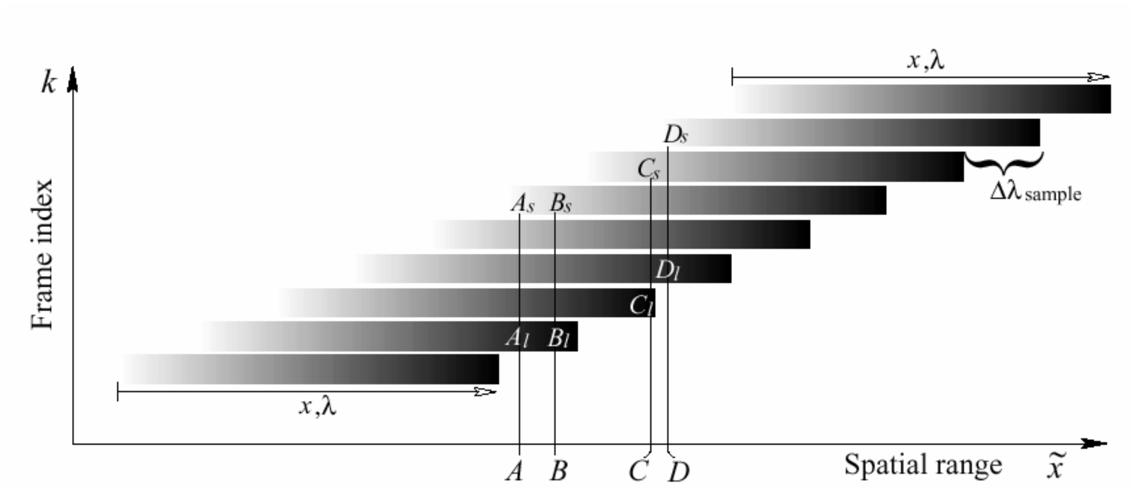


Figure 3

3. Mosaicing

3.1. Calibration

To obtain camera's internal parameters such as a focal length and an image center, and a distortion coefficient of a lens are called camera calibration. Because of the distortion of the lens and the gap of an image center, etc, the image which taken by an actual camera is different from the image which camera model should take. So, it is necessary to perform a camera calibration first, and get these parameters, and change an actual image to a model image by using these parameters.

In this experiment, get these parameters performing a camera calibration by using the method of Zhang [6]. And change images.

Look at Figure4 and 5. In figure 4, before changing, the image will have swollen round in the shape of a concentric circle from the center and the straight line will have bent. But, after changing that it is rectified and the straight line is straight.

A gap arises by next work and it will become impossible to generate a beautiful picture, if this work is not performed correctly.

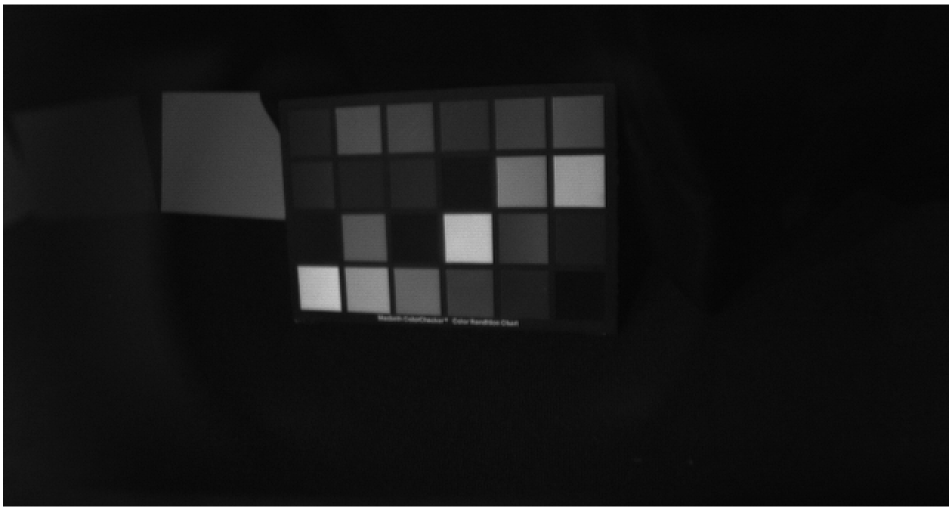


Figure 4:Before changing

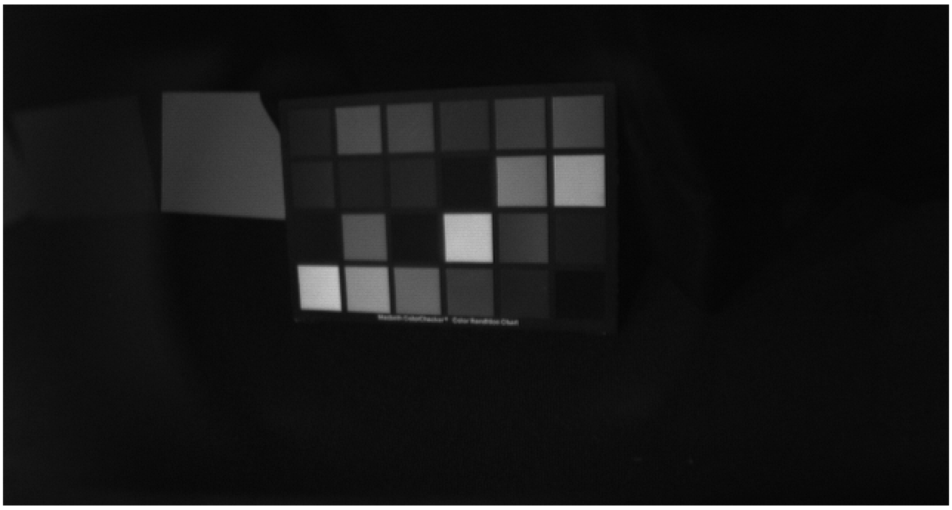


Figure 5:After changing

3.2. Assume Intensity

Looks like Figure 6, the images taken with a camera attached spatially varying filter is influenced greatly by the whole system of the camera containing a filter and by a light source. So it is difficult to mosaic these images. Since it needs to make those influences small.

First, consider the luminosity value of image I_k at (x, y) of the internal coordinates.

$$I_k(x, y) = L_{ill}(\lambda) l_{system}(\lambda) r(\tilde{x}, \tilde{y}, \lambda)$$

Here, L_{ill} is a light source, l_{system} is spectral response of a camera system and r is a spectral reflectance at $(\tilde{x}, \tilde{y})$ of the global coordinate.

Here,

$$x \sim \lambda$$

So,

$$I_k(x, y) = L'_{ill}(x) l'_{system}(x) r'(\tilde{x}, \tilde{y}, x)$$

And it is needed to make influences of L_{ill} and l_{system} as small as possible for mosaicing.

Then, an average of the luminosity value of the image in a horizontal direction is taken using all images

$$M(x) \propto \sum_k \sum_y I_k(x, y)$$

And, when the images taken by sufficient density,

$$r_{average}(\lambda) = \sum_{\tilde{x}} \sum_{\tilde{y}} r(\tilde{x}, \tilde{y}, \lambda)$$

From (4) and (5),

$$M(x) \propto L'_{ill}(x) l'_{system}(x) r_{average}(x)$$

is obtained.

If the scene is gray on average, $r_{average}(\lambda) = Const$, the average horizontal of luminosity value $M(x)$ is almost equal to the spectral of illumination and system response. And more, if $l_{system}(\lambda) = Const$, $M(x)$ is almost equal to the spectral of illumination.

Then, consider that $M(x)$ is the influence of a system and a light source. And make images by divide $I_k(x, y)$ by $M(x)$ in the direction of x-axis.

$$I'_k(x, y) = \frac{I_k(x, y)}{M(x)}$$

And edge detection and position determining are performed to these.

In order to perform this adjustment more strictly, it is necessary to take an image of what has a known spectral reflectance with the whole camera, and obtain the influence of a system response and a light source precisely.

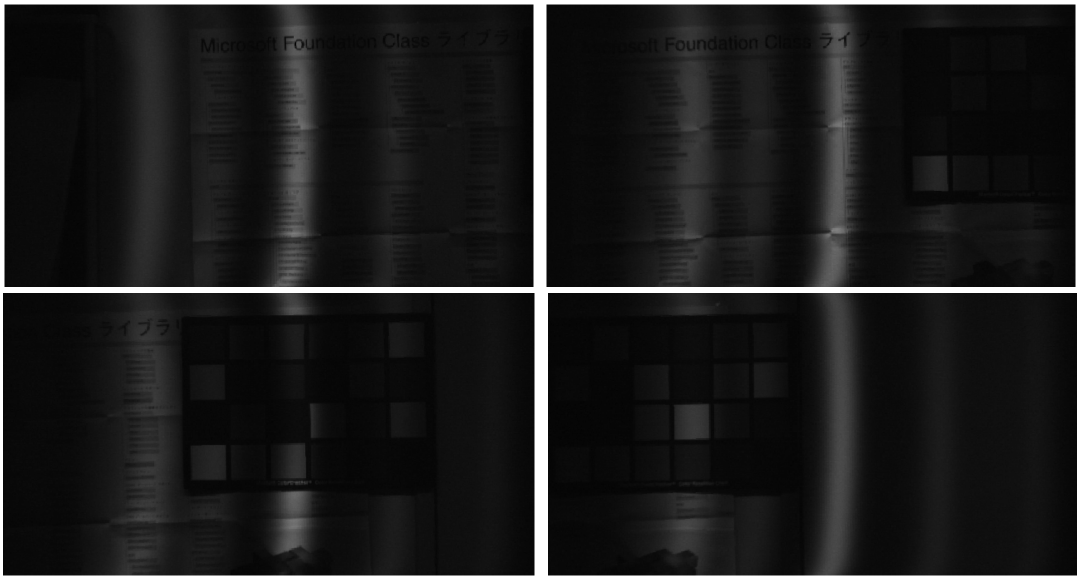


Figure 6: Before rectifies

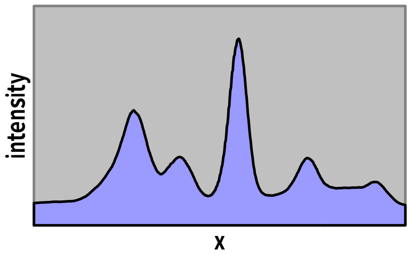


Figure 7: Assumed intensity

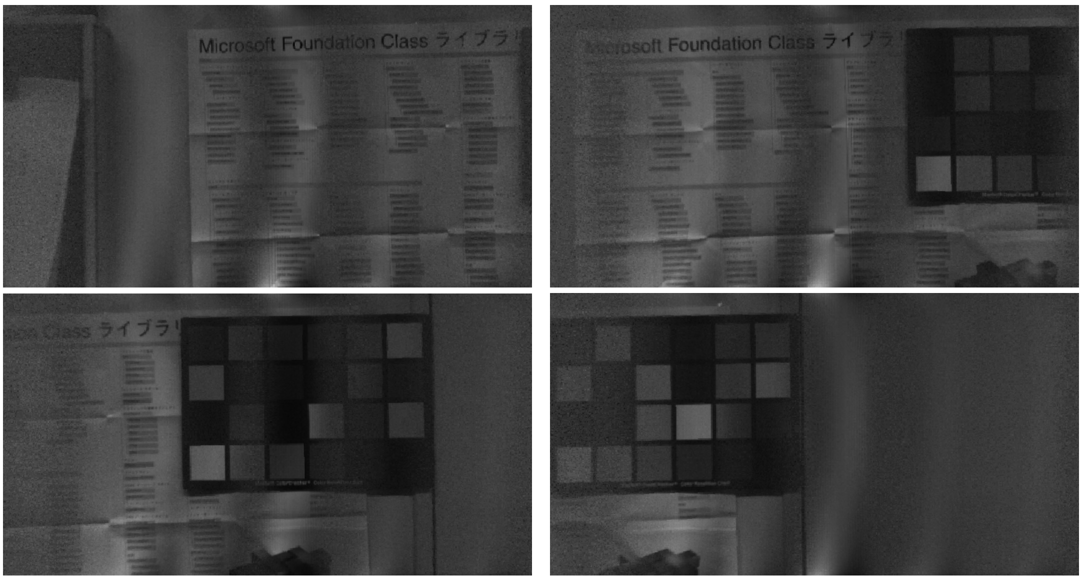


Figure 8: After rectifies

3.3. Edge detection

Even if the rectifies images which took with the same object, luminosity changes greatly with positions on a film, because of the spectral reflectance differ for every wavelength. For example, although a red object appears clearly on the left side of an image, but on the right side, it is hardly appeared. Therefore, in order to lessen this influence as much as possible, it is necessary to carry out the position relation by the images on which took an object on near wavelength, or on near position of an image. Next, different from general mosaicing that take an overlapped area as data overlapped too, our mosaicing takes overlapped area as data plotted by overlapped times. And in general mosaicing, miss positioning take influence to the overlapped area, but overlapped area is very small. Otherwise, in our mosaicing, miss positioning have a big influence to the overlapped area, but overlapped area is very large, because of necessary of plotting data in many wavelength. So it is necessary to carry out the position relation strictly as much as possible.

As mentioned above, to lessen the influence of the position on the image, use the edge detection.

Here, the technique of the edge detection used is explained. Edge detection consists of two processes, flat-and-smooth-izing and detection.

Before edge detecting, consider the error of each pixels. Because of the after rectifies images obtained by (7), the error of each pixels differ greatly. Consider the error of the before rectifies image as g, the error of the after rectifies images expressed as follows.

$$G(x) = \frac{g}{M(x)}$$

If a camera is 8-bit monochromic camera, it is thought that g is 0.5 or more because 8-bit expressed from 1 to 255.

Look at figure7. If under the fluorescent light, the difference of $G(x)$ will become five or more times. So, it appeared as a very strong noise according to an illumination.

So, it is necessary to reduce the noises.

Then, explain the method of noise removal.

There is a lot of method to remove noises. One of these is the method of use Gauss function, as follows.

$$I'(x,y) = \frac{\sum_{i=-N}^{N}\sum_{j=-N}^{N} I(x+i, y+j)G(i,j)}{\sum_{i=-N}^{N}\sum_{j=-N}^{N} G(i,j)}$$

$$G(i,j) = (2\pi\sigma^2)^{-1} \exp[-\frac{i^2+j^2}{2\sigma^2}]$$

Here, σ is the range of flat-and-smooth-izing, N is the range of the close.

Under the fluorescent light, to get strong of flat-and-smooth-izing, σ is set to 3.

Then, using this noises removed image, perform edge detection.

There is a lot of method of edge detection too. And we use the simplest method. It expressed as follows.

$$I_x[x,y] = \frac{1}{6}(I[x+1,y+1] - I[x-1,y+1] + I[x+1,y] - I[x-1,y] + I[x+1,y-1] - I[x-1,y-1])$$

$$I_y[x,y] = \frac{1}{6}(I[x+1,y+1] - I[x+1,y-1] + I[x,y+1] - I[x,y-1] + I[x-1,y+1] - I[x-1,y-1])$$

$$I'[x,y] = \begin{cases} 1 & if\ I_x^2[x,y] + I_y^2 > k \\ 0 & else \end{cases}$$

Here, I is the image of before edge detection, I' is the image of after it, and k is the threshold.

And this two process collected into one process to improve the speed.

3.4. Mosaicing

3.4.1. Perspective panoramas

When mosaic images with which took same plane, it is known that the position relation is expressed as follow.

$$x \sim Mx = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

When parallel translation of the fixed distance from a camera to a plane and took image toward a camera vertical to the plane, (14) can changed as follow.

$$x \sim Mx = \begin{bmatrix} 1 & 0 & m_2 \\ 0 & 1 & m_5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

So, it is necessary to carry out the parameter m_2, m_5.

In this research, m_2 and m_5 are changed within fixed limits, and carrying out that make most small the second power sum of the difference between two images.

But, this method needs a lot of time. So a method of fast this method is needed. For the sake, do this registration hierarchically. Hierarchically is carrying out parameters at from coarse to fine images. When doing this, it becomes very fast.

As evaluation function, use as follows.

$$Error_{total} = \min_{i,j} \sum_{overlap} | E_1(x, y) - E_2(x + i, y + j) |$$

And search i, j, when $Error$ is the smallest.

3.4.2. Rotational Panoramas

When a camera is rotated on the main point and took an image, the position relation of images can be expressed with three parameters, θ, ϕ, and ϕ_{rot}, mapped the images in a ball. And, when mapped in a cylinder, it expressed with three parameters, θ, Y, and ϕ_{rot}. And when performing position determining of the images that has taken a camera only have pan rotation, in the cylinder system, only θ is need for a relation between images. But, because of a pitch of a camera system, Y is need. And when obtained all position relation, it is able to make multispectral panoramic image.

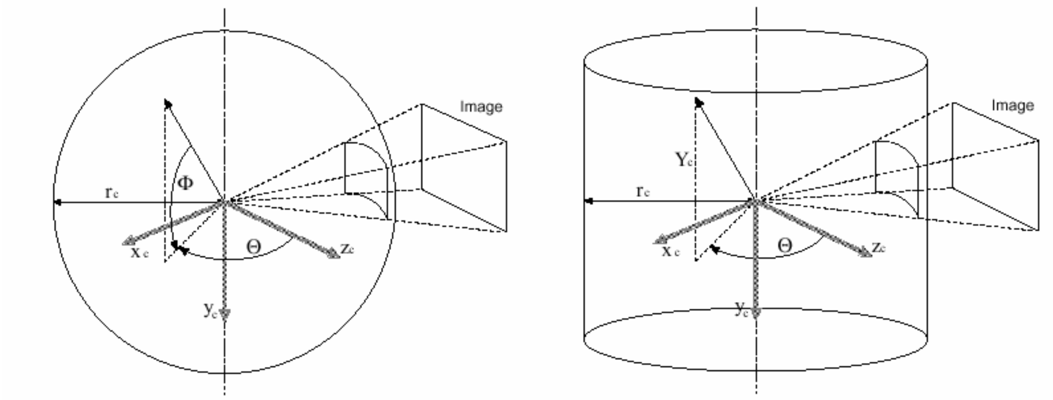


Figure 9: Mapped ball and cylinder

Now, when obtained θ and Y, there are some parameters with the necessity of understanding. The first is a center of images, (u,v), the second is a focal length, f, and the third is the size of a pixel element, k. And center of images is obtained when calibration.

The determination of these parameters is performed, like the time of perspective panoramas, minimize the difference between edge detected images, and expressed as following formula.

$$Error = \min_{\theta',y'} \sum_{overlap\,\theta,y} | E_1(M_1(\theta)+u, M_2(\theta,y)+v) - E_2(M_1(\theta-\theta')+u, M_2(\theta+\theta', y-y'))+v |$$

$$M_1(\theta) = \frac{f}{k}\tan\theta$$

$$M_2(\theta,y) = \frac{y}{k|\cos\theta|}$$

then, consider as follow,

$$y = kn$$
$$y' = kn'$$

then, from (14), (15), it can expressed as follow.

$$Error = \min_{\theta',n'} \sum_{overlap\,\theta,n} | E_1(M_1(\theta)+u, M_2(\theta,n)+v) - E_2(M_1(\theta-\theta')+u, M_2(\theta-\theta', n+n')) + v |$$

$$M_1(\theta) = \frac{f}{k}\tan\theta$$

$$M_2(\theta,n) = \frac{1}{|\cos\theta|}n$$

So, before determine θ and Y, it is necessary to be obtained f/k precisely. So, measure k and f, and gives as known. But, this value is not precisely. So it is necessary to rectifies.

Then, insert the same images at both the beginning and the end of the images. And can consider the actual focal length f' as follows.

$$f' = \frac{2\pi + \theta}{2\pi} f$$

here, θ is the gap angle of the same two images.

By using this we can obtain more precise focal length.

4. Create Color Images

Make a RGB image use the images that the position relation is obtained. The reason of making RGB image is that it is general for use.

First, obtain spectral at each global pixel.

From (1), and l_{system} is known, expressed as follows

$$I'_k(x,y) = \frac{I_k(x,y)}{l_{system}(\lambda)} = L_{ill}(\lambda) r(\tilde{x},\tilde{y},\lambda)$$

Then by mosaic the spectral data in the position relation, every global pixel obtained data of spectral from 400nm to 700nm. Then express it as follows.

$$\tilde{I}(\tilde{x},\tilde{y},\lambda) \approx L_{ill}(\lambda) r(\tilde{x},\tilde{y},\lambda)$$

Next, make RGB image so that it may be easy to use this if needed.

The method expressed as follows, by the amount of reactions for every wavelength to each coloring matter, RGB, of an eye is D_r, D_g, D_b.

$$\tilde{I}_r(\tilde{x},\tilde{y}) = \int_\lambda D_r(\lambda)\tilde{I}(\tilde{x},\tilde{y},\lambda)d\lambda$$
$$\tilde{I}_g(\tilde{x},\tilde{y}) = \int_\lambda D_g(\lambda)\tilde{I}(\tilde{x},\tilde{y},\lambda)d\lambda$$
$$\tilde{I}_b(\tilde{x},\tilde{y}) = \int_\lambda D_b(\lambda)\tilde{I}(\tilde{x},\tilde{y},\lambda)d\lambda$$

From this, it is able to get RGB image.

5. Create New Image under Different Illuminant

Then, explain how to make the image that the light source is changed.

First, from (1), if take image of the same scene under the different light source it expressed as follows.、

$$I'_k(x, y) = L'_{ill}(\lambda) l_{system}(\lambda) r(\tilde{x}, \tilde{y}, \lambda)$$

Then, from (1) and (21)

$$L_{ill}(\lambda) = L'_{ill}(\lambda) \frac{I_k(x, y)}{I'_k(x, y)}$$

is obtained.

So, from (6) and (19)

$$\tilde{I}'(\tilde{x}, \tilde{y}, \lambda) = \frac{M'(x)}{M(x)} \tilde{I}(\tilde{x}, \tilde{y}, \lambda)$$

is obtained.

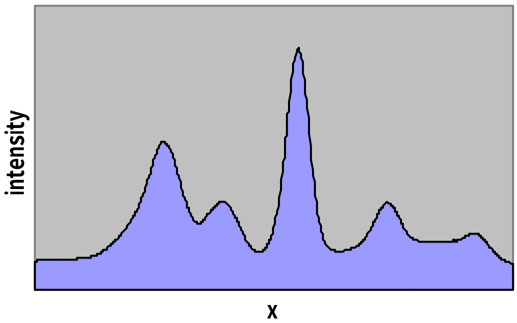
Moreover, the image under a new light source can be created by giving arbitrary light sources, without taking an image under the different light source, if is able to be estimated well. Namely if consider L_{ill} mostly equal to $M(x)$, it expressed as follows.

$$I''(x, y) = \frac{Lill(\lambda)}{M(\lambda)} NewLight(\lambda) r(x, y, \lambda) l_{system}(\lambda)$$

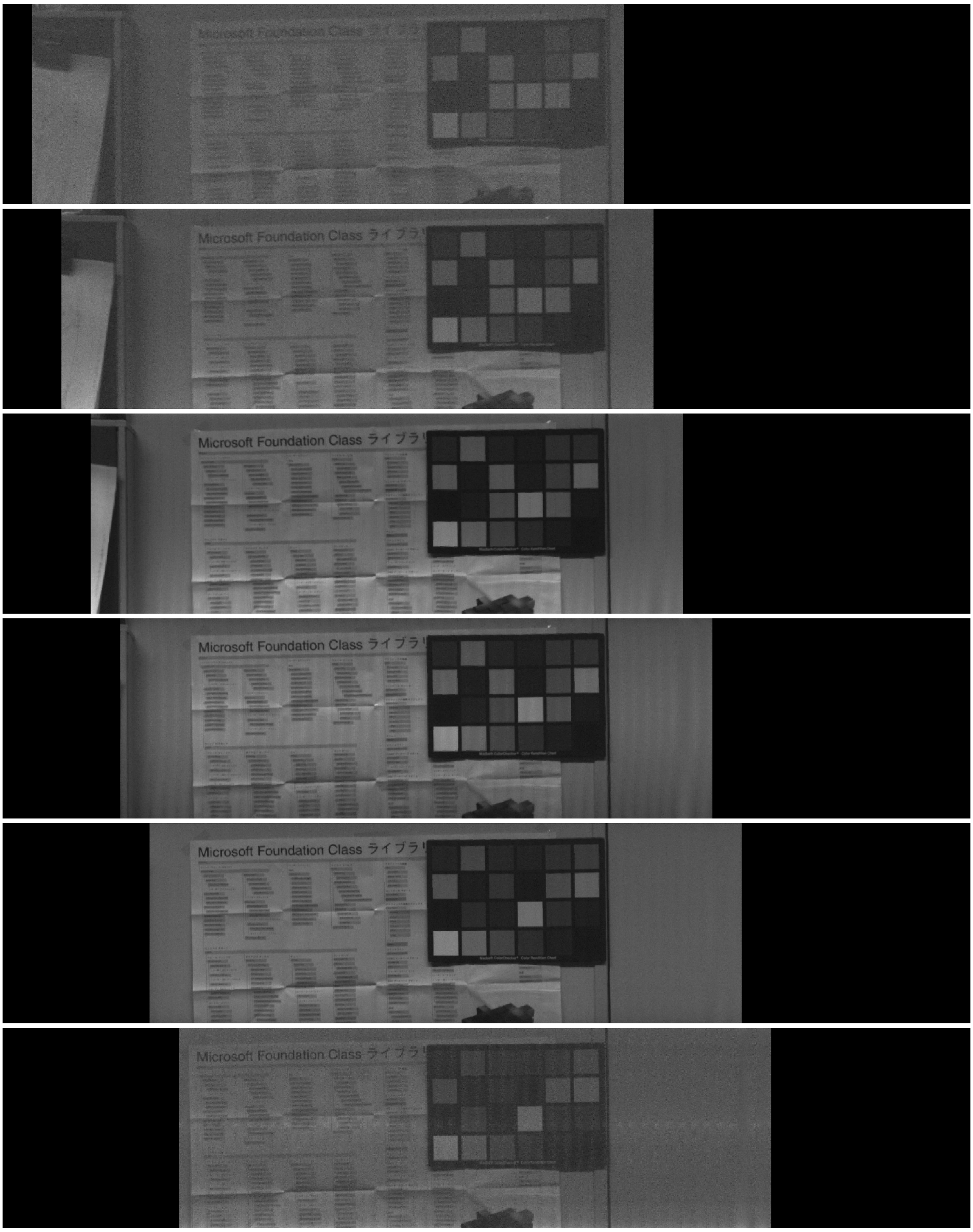
6. Experiments

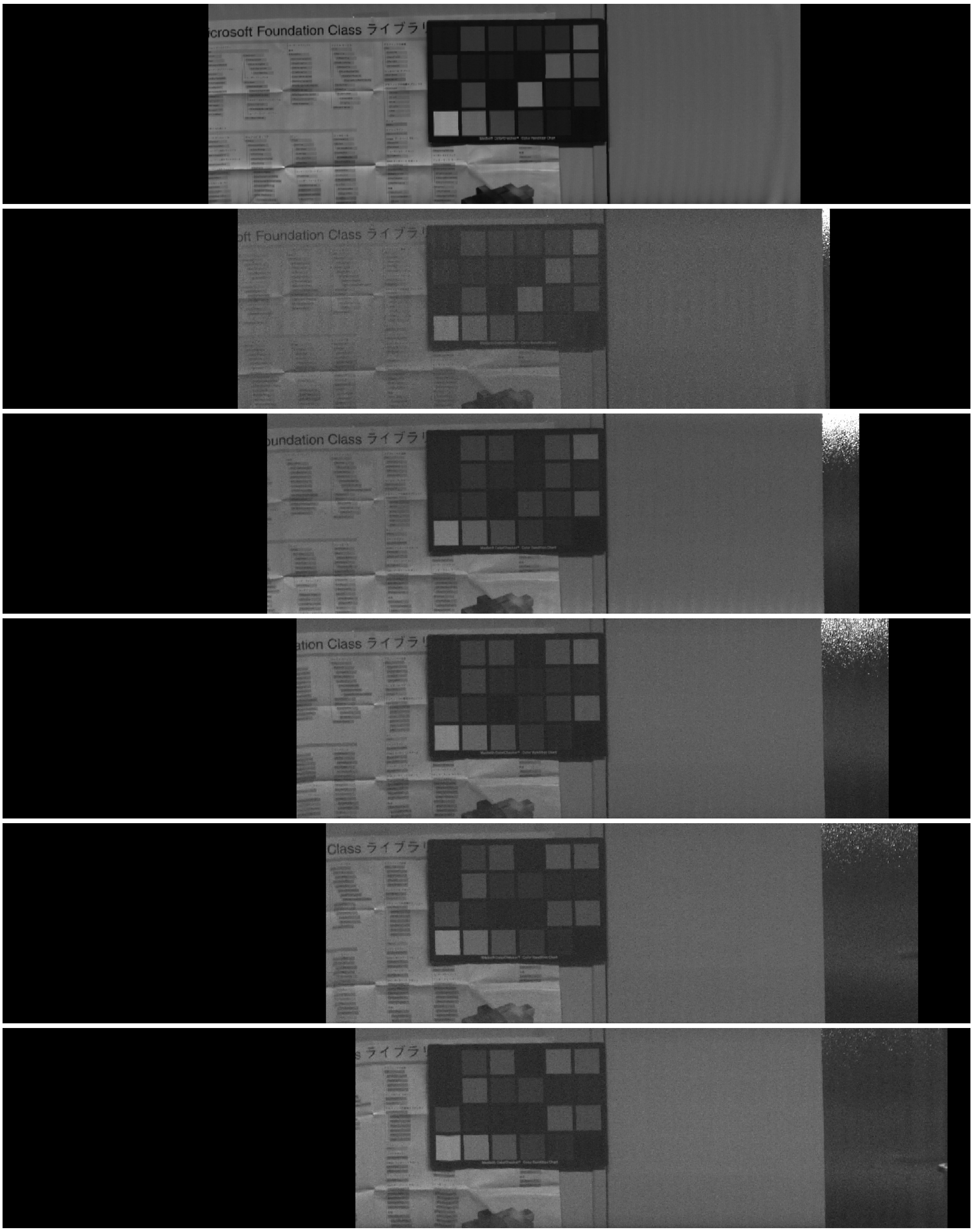
まず、カメラに spatially varying filter を装着して壁にマクベスチャートを貼り付け、壁からの距離が一定であるようにカメラを少しずつ動かしながら撮影して mosaicing を行った。$l_{system}(\lambda)$ が一定であると仮定して RGB 画像を作成し、M(x)を光源のスペクトルと考える。

mosaicing により作成された画像(蛍光灯の下で撮影)



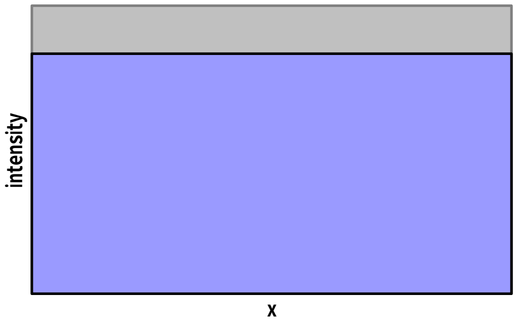
そしてこの場合、各波長におけるスペクトルの強さは次のようになっている、


Microsoft Foundation Class ライブラ


icrosoft Foundation Class ライブラ

次に、新しい光源を想定して、画像を作成する。

光源を一定であると仮定したもの



次式(Planck's equation)により導き出されたもの。

$$E(\lambda,T) = c_1\lambda^{-5}[\exp(\frac{c_2}{T\lambda}) - 1]^{-1}$$

K=2500

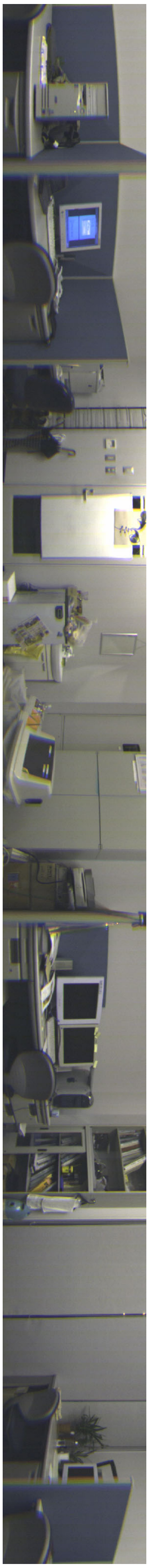
K=5500

K=10000

となる。

次に実際の光源で撮影されたものと比較してみる。

次にカメラをパン回転させて画像を撮影し mosaicing を行う。
カメラを 2.5°ずつ回転させ、144 枚の写真を撮影した。



7. Conclusion

Some problems came out by this experiment. The first, in figure 14 that is rotational panorama, a part of the image is faded. This is considered because the parallax to have arisen by the camera moves other than pan rotation. Therefore, in order to obtain a highly precise picture, it is required to make it rotate on the main point correctly. The second is that the difference has arisen between the image in which was taken with RGB camera, and the created image. It is thought that this is the problem of the reaction of RGB camera element and the problem of the reaction of the monochrome camera system. It is thought that the same image is created deducing this reaction correctly for every wavelength. The third is that a big error arises in the amount of reflection for every objective wavelength, when a photograph is taken under a fluorescent light. Therefore, for more exact presumption, it is considered good, having taken an image under the electric bulb. Or, as the method of lessening an error, the method to take an image of same scene two or more times, and to take an average, the method of extending the range of data by changing the shutter speed of a camera, etc. can be considered.

And, another problem is that assumed the light source and a reaction of a camera system. In order to obtain more exact data, the easy and more exact method of determining the amount of the light source and a reaction of a camera system is needed.

Next, the problem is that it needs a lot of time to determine the position relation between two images. By using hierarchical method, we attain to improve the speed. And since it is necessary to plot the data of wavelength two or more times for each point, only a fixed motion becomes possible to a camera substantially. Then, the range of a parameter can be narrowed down to some extent, and it can accelerate. However, in general mosaicing, you have to carry out this somehow or other.

And, it is necessity of taking a lot of images, so the way of gather efficiency, take the images by video camera as movie and mosaicing process enabled to the movie is considered.

And it is enabled by solve these things to obtain multispectral panorama, earlier, more exact, and more simply.

References

[1]Richard Szeliski and Hung-Yeung Shum. *Creating Full View Panoramic Image Mosaics and Environment Maps*. Microsoft Research. SIGGRAPH'97, pp.251-258, 1997.

[2]Yuu Nakatani. *Image Mosaicing Using Active Camera*. The Institute of Image Electronics Engineers of Japan, vol.29, no.5, 2000.

[3]Yalin Xiong and Ken Turkowski. *Registration, Calibration and Blending in Creating High Quality Panoramas.* Fourth IEEE Workshop on Applications of Computer Vision, Princeton, New Jersey, October 19-21, 1998, pp. 69-74.

[4]Naoki CHIBA, Hiroshi KANO, Michihiko MINOH, and Masashi YASUDA: *Feature-Based Image Mosaicing*. The Institute of Electronics, Information and Communication Engineers. D-Ⅱ, Vol.82-D-Ⅱ, pp.1581-1589, 1999.

[5]Yoav Y.Schechner and Shree K.Nayar: Generalized Mosaicing: *Wide Field of View Multispectral Imaging.* IEEE Trans. on Pattern Analysis and Machine Intelligence 2001.

[6]Zhengyou Zhang: *A Flexible New Technique for Camera Calibration.* Microsoft Research Technical Report MSR-TR-98-71. December 2, 1998.