An Inexpensive, High Resolution Scan Camera

by

Shuzhen Wang

B.E., Tsinghua University, 2000

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES

(Department of Computer Science)

We accept this thesis as conforming
to the required standard

The University of British Columbia

December 2003

Library Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

_____Shuzhen   Wang_____   _____23 /12/ 2003_____
Name of Author *(please print)*   Date (dd/mm/yyyy)

Title of Thesis: _An Inexpensive, High Resolution Scan Camera_

_____

Degree: _Master of Science_   Year: _2003_

Department of   _Computer Science_
The University of British Columbia
Vancouver, BC   Canada

Abstract

The use of digital imaging devices has been growing very fast and having amazing influence over the last decade. Being easier to integrate with other digital media, digital imaging is taking the place of analog imaging in more and more fields. Although the resolution and color quality of digital cameras have reached those of 35mm films, there are still a number of applications which require better quality, such as museum catalogs, professional digital photography and research in image based modeling and rendering. These applications all benefit from high resolution digital imaging. Our work extends digital photography in this particular direction.

We present the design of a high-resolution scan camera using a flatbed scanner as the backend of a large format camera. The scan camera we built can take images with the resolution of up to 122 million pixels, while the camera itself can be built from off-the-shelf components for only 2,000 dollars. If we simply attach the two parts of the system together (the large format camera and the flatbed scanner) in their original setup, the system won't work properly because of mechanical and optical constraints. We dealt with these constraints by removing the light source and lenses from the scanner, and aligning the scanner with the imaging plane of the view camera. Due to the changed optics in the scanner, we can not directly use the commercial scanning software from the vendor. Instead, we get the raw image data from the scanner, then do denoising and calibration to acquire high quality images. A more advanced process is proposed to first detect artifact features, then remove them by image inpainting. Finally, some quantitative measurement of the light sensitivity and the optical resolution of the camera are obtained.

Contents

List of Tables

List of Figures

Acknowledgements

The University of British Columbia
December 2003

To My Grandpa.

Chapter 1

Introduction

1.1 Motivation

Digital cameras have become very popular in the photographic community since being first introduced in the mid of 90's. Compared to the traditional film camera, it has many advantages. First, it saves the cost of buying and developing films. Second, when the user is not satisfied with the photo he took, it is easy to erase and take it again. Today the image quality of consumer level digital cameras and 35mm analog cameras has been indistinguishable when both printed on regular photo papers. These advantages have made digital photography a popular alternative to conventional film photography.

Resolution is an important factor in digital imaging. The larger the resolution is, the more information the image acquires. The resolution of consumer level digital cameras has reached 5 million pixels, and the resolution of some latest semi-professional digital cameras has reached more than 8 million pixels. This resolution range is enough for scenarios like web-browsing and small to medium size photo printing.

However, there are some applications that require higher resolution. For example,

- Professional digital photography, such as advertising, commercial, and industrial, needs much higher resolution than consumer or semi-professional digital cameras. These applications require large prints of the photographs that have very fine details. The conventional approach is using medium/large format film camera, since this kind of photography is either in studio or on location.

- Museum catalog and art reproduction also require high quality images. Digitization of museum collections and art objects can overcome the physical constraints of traditional museum exhibition and archival. Some art items or historical artifacts are expensive and difficult to move around. If their 3D models and 2D images can be obtained in very fine detail, the "digital format" of the artifacts can be easily accessed by the people all around the world. Another reason to do a digital museum catalog is that the digital form of the artifacts lives longer than the physical one. The reproduction of the artifacts needs as much detail as possible.

- Research projects in image based modeling and rendering is another area which will benefit from high resolution images. One example is that when we try to reconstruct 3D models with complex surface property from images, such as hairy toys, obviously the resolution of the image is an important factor that influences the accuracy of the reconstruction.

Unfortunately, the resolutions required for these applications are only provided by digital backends for traditional large format camera. These backends are either one-shot backends with two-dimensional sensor arrays, or scan backends with

one-dimensional linear sensors. Although they can reach higher resolution, they are also much more expensive than consumer level digital cameras, because the manufacturers need to build their own mechanics according to the larger sensor array, and develop their customized control and calibration software.

1.2   System Overview

The resolution of a digital camera equals to the density of the sensor element times the size of the imaging plane. Commercial non-professional digital cameras keep the size of the imaging plane at 35mm or smaller in order to keep the cameras compact, and improve the resolution by making the sensor element smaller. But it is hard for the manufacturing process to make the sensor element small enough to get higher resolution. On the other hand, professional digital cameras are mostly used for photography in studio or on location, they need high resolution rather than high portability. They apply larger area of image sensors to achieve high quality.

Motivated by the design principle of professional digital camera, we use the image plane of large format camera to get higher resolution. To reduce the cost, we choose a consumer-level flatbed scanner as the digital backend. This is the main contribution of our scan camera.

In brief, we built a very high resolution scan camera by converting a consumer-level flatbed scanner to the scan backend of a 10" by 8.5" large format camera. Since the optical resolution of the flatbed scanner is 1,200 DPI (dot per inch), the resolution of the scan camera can reach 122 million pixels. This scan camera can take grayscale, color and near infrared images with appropriate color filters. The whole camera system, including the large format camera, the flatbed scanner, color filters and other accessories, costs less than 1,200 dollars. Figure 1.1 is a picture of our

camera system. The scan camera is connected to a computer running scanning soft-ware via a USB 1.1 port. Figure 1.2 shows an image of a fluffy dinosaur taken by our scan camera, and a zoom-in region.

The design of our camera system has two parts: the hardware construction of the scan camera and the software for the calibration of the images. On the hardware side, we dealt with both the mechanics and optics issues in the construction of the camera. On the other hand, because the optics of the scanner is changed, we also implemented our own calibration software. The calibration software takes the raw image data as input, and generates the resulting color image.
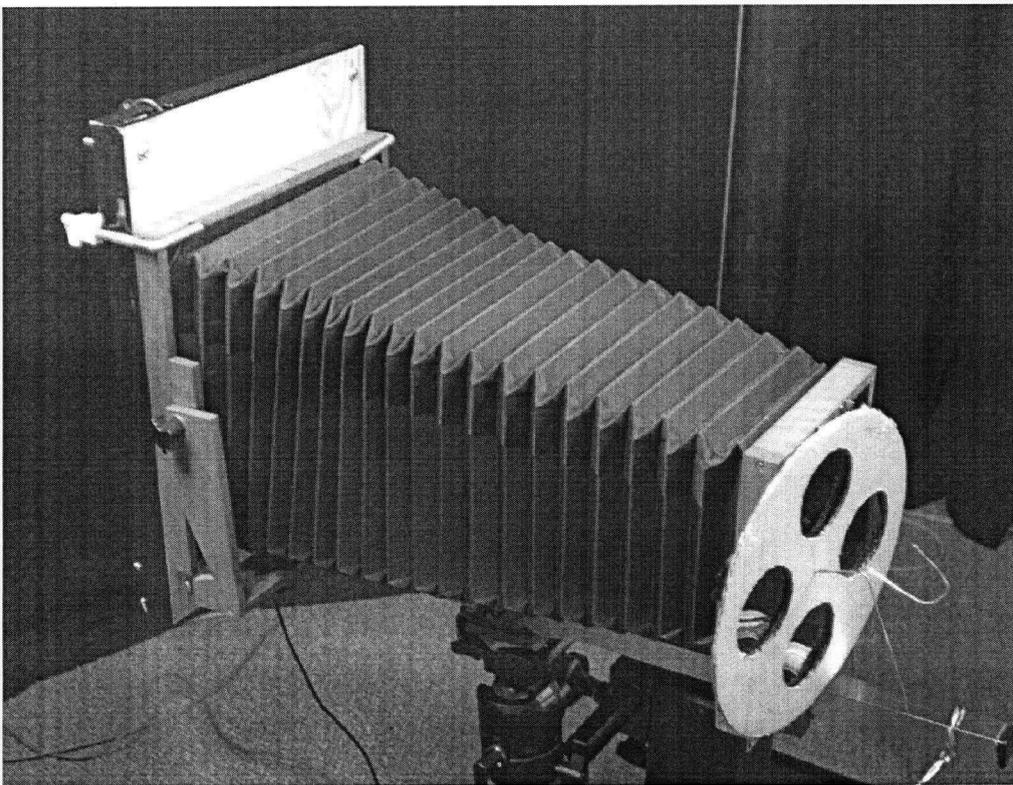
Figure 1.1: Picture of the scan camera.
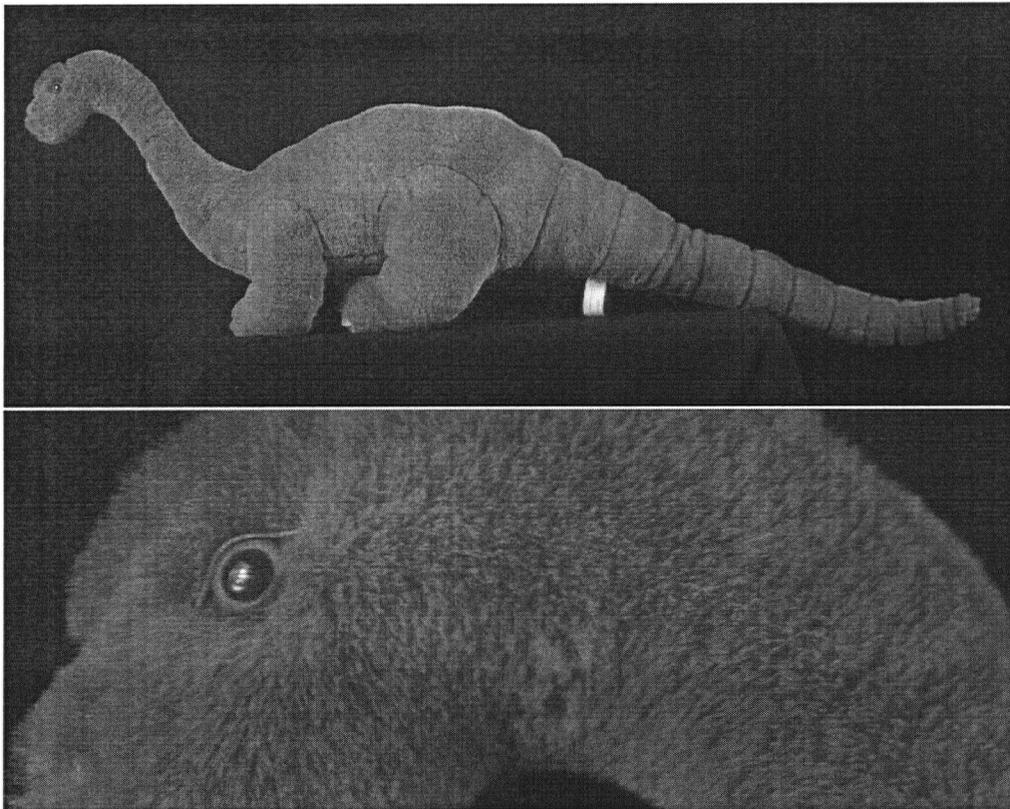
Figure 1.2: Pictures taken by the scan camera. The top is the whole image (12,390 by 4,834 pixels) scaled down to fit in the page, and the right is a region of the top image (1,650 by 644 pixels) printed in 300 DPI.

1.3 Thesis Outline

The rest of this thesis is organized as follows. In chapter 2, we discuss the related work to this project, including image sensor technologies, and high resolution digital imaging research in the commercial market and academics. Chapter 3 explains the hardware issues on the system, and how we worked around them. Chapter 4 describes the image processing we have done on the raw scanner output to get calibrated images. We discussed the results in chapter 5. Chapter 7 concludes the

thesis by the benefits, limitations and possible future work.

Chapter 2

Background

Digital imaging industry is always driven by the development of the image sensor technology. In this chapter we survey different image sensor technologies and briefly overview commercial high-end digital cameras. We also discuss some related work to our project.

2.1 Image Sensor Technology

The image sensor is the essential part of digital imaging devices. It consists of several individual elements (pixels) which can convert photons (light) into electrons (electrical charge). The brighter the light illuminates the pixel, the greater the electrical charge accumulate. When the shutter of the camera closes, the electrical charge of all the pixels are read out and saved in the memory of the camera, after processed by an analog-digital converter. There are different kinds of image sensors, which can be applied in different imaging applications.

2.1.1 CCD (Charge Coupled Device)

A CCD is capable of collecting, storing and transmitting electrical charge from one sensor element to another. There are one or more output amplifiers at the edge of the chip to collect the data from the CCD. Shift registers are used to read out pixels before the output amplifiers collect them. In CCD digital cameras, a separate circuit board is used to place all other functions: clock and timing drivers, analog-digital conversion and so on. There are three types of CCD architectures (Figure 2.1).

In Full Frame (FF) CCD, after the sensor array has been exposed for some time, series of sampling pulses are applied via parallel shift registers to transfer pixel signals to serial shift registers, one row at a time. Each row in the serial registers is then collected by the output amplifier. Because of the imaging process and readout process both occur on each pixel, the exposure is controlled by a mechanical shutter or strobe to guarantee the image integrity (i.e. all pixels have the same exposure time). The advantage of Full Frame CCD is that because it has such simple structure, it has highest resolution and highest density. The imaging time for one frame is the exposure time plus the readout time.

Frame Transfer (FT) CCD achieves a continuous imaging process without the shutter or synchronized strobe. The idea is that a storage array is used to cache the readout of the imaging array. When the exposure process is finished, the pixel values on the whole image array are first read to this storage array, which is not light-sensitive. Then the whole storage array is read out by the same process as in the Full Frame CCD. At the same time, the imaging array can start to acquire the next frame. This parallel architecture improves the frame rate of the sensor. The disadvantage is that the data transfer process between the imaging array to storage array may introduce "smear" effect. Because the sensor elements are exposed and

read out at the same time, different pixels have a little different exposure time. Meanwhile, the resolution and density of Frame Transfer CCD is less than Full Frame CCD because extra silicon circuit is need.

Interline (IL) CCD is a design which deals with the image "smear" problem in Frame Transfer CCD. Each pixel of the sensor is separated as a photodiode and a shift register. The charge of the photodiode is transferred to the shift register instantly after the exposing process. It has a faster frame rate because the readout process is faster than Frame Transfer CCD. The disadvantage is similar to Frame Transfer CCD. the resolution and light efficiency is sacrificed. This property of Interline CCD makes it work better in real-time imaging and motion picture.

2.1.2   CMOS (Complementary Metal-Oxide Semiconductor)

CMOS sensors convert light into electric charge and then process it to electronic signals like CCD sensors, but the design is different. In CMOS sensors, each element has an electron-to-voltage conversion amplifier, and most functions such as timing generation and signal processing are integrated into the chip. Compared to CCD sensors, CMOS sensors consume less power, and can be manufactured on any standard silicon production line. Nevertheless, in a CMOS sensor, since each pixel has its own electron-to-voltage conversion, the conversion at each pixel may be not uniform. This makes the image quality from CMOS is not as good as from CCD. A sample architecture of CMOS sensor is illustrated in Figure 2.2.

2.1.3   Linear CCD Image sensors

Linear CCD image sensors have only one row of sensor elements. Such sensors also have shift registers and output amplifiers for the data readout. Linear image

9
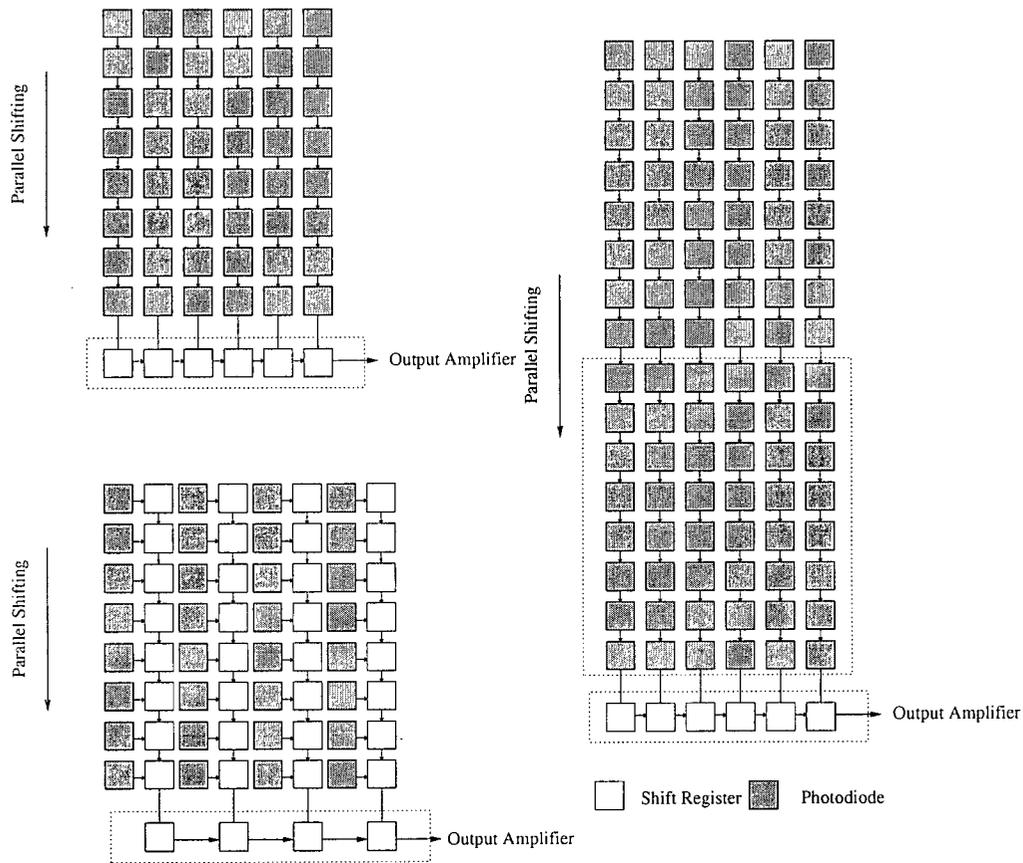
Figure 2.1: Architecture of three different CCD sensors. Left-top is a Full Frame (FF) CCD sensor. At the right is a Frame Transfer (FT) CCD. And at the left-bottom is an Interline (IL) CCD.

sensors are widely used in flatbed scanners and document copiers. They are also used in scan backs for professional digital photography and satellite imaging. Single-row linear sensors are monochrome. Color linear sensors are mostly made of three rows of sensors, one primary color for each row. Kodak manufactures an ultra-high resolution 14,400 pixel trilinear color image sensor for high-end color scanning system. [25]

To summarize, different image sensors have different advantages and disad-
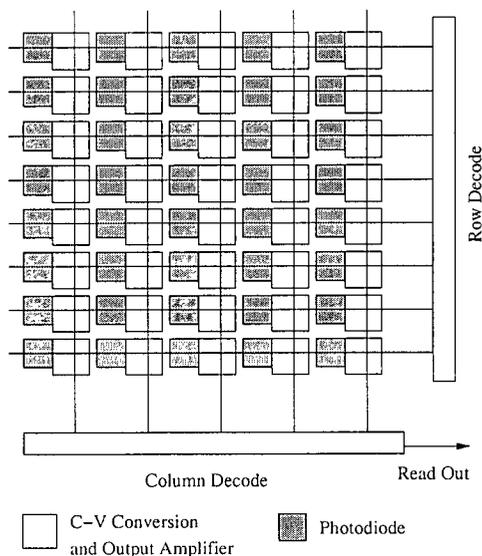
Figure 2.2: Architecture of CMOS image sensors.

vantages. In digital imaging, choosing which kind of sensors really depends on the actual applications, and the outcome versus the expense. To evaluate today's image sensors, important criteria include sensitivity, data rates and noise level. Exposure control and anti-blooming are also important factors.

2.1.4 Color for Sensors

Because the sensor element itself can not see color, it can only give the total amount of light striking its surface. To get full color image, in most design, each sensor element looks at the light through filters in its three primary colors. Once all three colors are recorded, they are integrated together to create the full color spectrum the human eye is used to see. There are several methods to record three color channels in digital camera.

- Some highest quality digital cameras use three seperate color sensors, each with a color filter on the top. A beam splitter divides the incoming light equally and redirects them to the three color filters. The advantage of this method is that the camera records each of three colors at each pixel, but the cameras using this technique tend to be expensive and bulky. Since three color sensors for a single pixel are at different positions, different color channels may not align well.

- The most common method uses only one color filter at each sensor pixel. The color filters are interleaved by alternating rows of red and green filters with rows of green and blue filters [3]. However, this means that only one color channel is measured at every pixel, and the remaining channels are interpolated from the neighbouring pixels. The problem is that since only one-third of the color information is actually measured, the other two-thirds are "guessed" by interpolation.

- In a scan camera, three rows of color sensor, in three color channels, are often used to measure full color of each pixel as the sensor sweeps across the image plane.

- Recently there has been a new sensor technique which captures three color channels by different layers of one single sensor. It is based on the fact that red, green and blue light penetrates silicon to different depths. [14]

2.2 Commercial Digital Cameras

There have always been efforts in pursuing higher quality digital photography. Consumer and semi-professional digital cameras build up to 6 million imaging ele-

ments [18] on an object frame smaller than a 35mm film. More increase of resolution in the design is restricted by manufacturing technology. But there are cameras aiming at the top professional photography in advertising and artistic market. Art reproduction and archival applications also need images with very fine details.

One end of this line of cameras are newest 35mm SLR digital cameras. Fujifilm S2 Pro, Canon EOS 10D and Nikon D100 are some examples of such cameras. They all have over 6 million pixels, with the advantage of portability of consumer-level digital cameras. However, the resolution they have is still not enough for certain applications we mentioned in Chapter 1.

The other end is digital backends for medium format camera. The development of the image sensor industry has made digital backends a strong competitor in high-end digital photography. Some digital backends are Kodak DCS Pro Back Plus [20] using Kodak 16 megapixel full frame CCD, FujiFilm Luna II with an 11 megapixel CCD. The vendors of professional backends develop their products from the new image sensors. More recently, Kodak-Sinar and Dalsa both introduced their cutting-edge 22 megapixel full frame CCD image sensor used for digital-back of large/median format cameras. Creo Leaf Valeo 22 and Sinar Back 54 are two large format camera back which use these CCD sensors. Since median/large format cameras have larger image area than normal 35mm cameras, theoretically, the improvement ratio on the resolution is the same as the enlargement of the image plane, assuming the imaging element is the same size.

Scan backends are another type of digital backend. Instead of using a two dimensional dense CCD array, scan backend uses a line of sensors to scan across the image plane, in a way similar to a flatbed scanner. Current scan backends on the commercial market can acquire higher resolution. For example, Better Light

[6] Super8K-2 Scan Back uses a Kodak trilinear color CCD for a 4" by 5" median format camera. The highest resolution for this scan back is 12,000 by 15,990 pixels. The weakness of this scanback technique is that it can only deal with static scenes and can only be used in the studio or on location.

Unfortunately, both 2D digital camera backs and scan backends are expensive to buy. In Table 2.1, we compare the resolution and price of different technologies applied in digital imaging, and compare it to our system. Note that the price of single shot camera backs and scan backs exclude the price of the the lens and the body of the median/large format camera.

Cameras	Resolution	Price (Canadian dollars)
Consumer	<4M pixels	$300–2,000
Semi-Professional (SLR)	<6M pixels	$2,500–5,000
Professional (single shot)	<22M pixels	$25,000–45,000
Professional (scan)	<100M pixels	$15,000–35,000
Our scan camera	122M pixels	$2,000

Table 2.1: Comparison of resolution and price for different digital camera technologies, and the comparison between them and our scan camera.

2.3   Related Work in Academia

Besides the commercial cameras for high resolution digital photography, there also have been some related work in the academics. We first review some work on the mosaicking technique used to acquire or view high resolution images and geometries, then we discuss some other scan camera projects.

2.3.1   Mosaicing

Medical images such as pathology images usually have very high resolution, which are obtained by stitching together images of microscope slides. This method needs

very accurate positioning of the camera or slides, or an optimization algorithm is required to stitch the images seamless. Ferreira et al. [13] designed a system to provide efficient manipulation of high resolution digital images of histopathology slides.

VASARI [22] is an art reproduction project to do high resolution digital imaging. It developed a colormetric scanner system to create more accurate records of museum paintings. The images are taken at the resolution of 10 to 20 pixels per millimeter by a CCD digital camera, then multiple images are mosaicked together, which in turn is calibrated to CIE XYZ color space by using color charts. The computer controlled digital camera can move back and forth to get automatic focusing, and the light source is mounted with it to get same light distribution for each image. The color information is obtained by combining images from 7 bandpass filters covering the whole visible spectrum. This scanning system produces promising results, however, it can only digitize planar objects, and it has a complex calibration process including lighting, color spectrum, mosaicking and so forth.

The digital Michelangelo project [21] aims at the high resolution digitization of geometry and surface characteristics of art objects. The Michelangelo Statues are reproduced by employing digital still cameras, laser rangefinders and a suite of software tools. The resolution of their system is very high. It can capture chisel marks smaller than a millimeter on the 23 feet tall David Statue. The way they get very high resolution is essentially a 3D mosaicing/stitching process. After the scanner and camera acquires the shape and color information of different parts of the statue, a post processing pipeline aligns and merge them together.

15

2.3.2 Scan Cameras

DavidHazy [8, 9] exposed his experience in obtaining a line-scan digital camera by his own. He turned a Kodak snapshot scanner into the digital back of a traditional 35mm camera. The sensor array is taken from the scanner and put on the focusing plane of the 35mm camera. In the snapshot scanner he used, the scanning is conducted by moving the document constantly; the sensor itself is fixed. For this reason, this setup can take panoramic photography by fixing the sensor array in the middle of the film gate and rotating the whole camera around the tripod's axis in constant speed. When the camera is fixed and the subject rotate in front of the camera, a peripheral or "rollout" photograph is acquired. However, this camera lacks mechanism for scanning across the focal plane, which makes it very difficult to do normal 2D photography as in commercial scanning camera backs. There isn't any calibration process in this work: the images come directly from the vendor supplied software. DavidHazy mentioned the infrared issues very likely met when scanning real work scenes using a flatbed scanner.

Wandel [27] built a similar digital camera from a flatbed scanner. The difference is that Wandel gave more elaborate mechanism change on the scanner. The whole flatbed scanner is disassembled and assembled together in a different mechanism. A 35mm screw mount SLR camera lens is mounted to the sensor, which are rotated by some gearing mechanism when the sensor is scanning. Wandel's system doesn't have any software processing either, but he discussed the blooming effects and color balance problem when scanning real world scenes.

Yang et al. [30, 29] proposed the design of a low-cost light field capture device, which is essentially a digital scan backend. In their system, a two-dimensional assembly of 88 lens in a 8 by 11 grid is mounted above the glass of a flatbed scanner.

16

The images formed by the lens are focused on the glass, then subsequently scanned by the flatbed scanner since the image sensors and lenses are also focused on the glass. For each scan of the whole glass, 88 images of the scene from different view points are acquired. These images can be used for light field rendering.

Yang et al. [30, 29] further discussed the capture and image processing in their light field camera. They analyzed the color problem when exposing the scanner sensor to lighting conditions other than fluorescent light source. First they remove infrared information by using an IR cutoff filter. Then they compared four possible color correction methods:

- "autolevels" color correction feature in Adobe Photoshop.

- white balancing using Matlab and its Image Processing Toolbox

- create a characteristic profile using a color calibration board

- adjust the response curve available in most image-editing software

They discussed the dependency of color profile on different illumination conditions. For the other three methods, they found that the white balancing keeps consistent histogram as the original image, but the "autolevels" and response curve methods do not. The reason is that the white balancing is a linear mapping between inputs and outputs, while the other two are nonlinear. In their system, they applied the "autolevels" in Adobe Photoshop because it's easy to use and fast. Finally, they characterized the radial distortion and perspective distortion by means of standard camera calibration approaches.

Unlike the previous two system, their system is a complete scan camera with both hardware and software functionality. However, their objective is a light field camera, instead of obtaining high resolution digital images.

Chapter 3

Hardware Setup

In our scan camera, we use a 10" by 8" image plane, which is over 90 times larger than 35mm films, and scan it with a flatbed scanner. The enlargement of the image plane gives us the corresponding scale-up on the resolution of the camera. However, when the large format camera and the flatbed scanner are attached together, they add mechanical and optical constraints to each other. In this section, we study the issues in the construction of the scan camera, and describe the way we dealt with them.

3.1   Large Format Camera

Large format cameras, also called view cameras, are widely used in conventional professional photography [1]. Generally, cameras with 6cm by 4.5cm films are called median format cameras and those with larger films, such as 4" by 5" or 8" by 10", are large format cameras. The body of such cameras is made of lens holder and back holder, connected with bellows. At the front, the lens is assembled on the lens holder. The back holder is a frame for focusing glass and film holder. In the process

of photography, the lens first focuses the subject on the focusing glass. This can be done by adjusting the bellows and camera position and looking at the image on the focusing glass. After the subject is properly focused, the focusing glass is replaced by the film to take an actual photo.

Large format cameras have advantages and disadvantages over 35mm film cameras. The main advantage is the large image size: the image quality of large format camera is hard to be matched by any enlargement of 35mm films. The image is sharper and grain-free. Another advantage is that there is more flexibility in the control of imaging process. The photographer gets more control not only on the focusing of the scene, but also on the perspective by adjusting the orientations of lens holder and back holder. However, large format cameras have some disadvantages. First, the camera is bulky and heavy. Second, although the camera gives more flexibility on the control, at the same time it requires more concerns from the photographer to get a nice photo: setting up camera, scene compositing, focusing, exposing and so on.

We bought the 8" by 10" view camera kit from Bender Photographic [4] and assembled it by ourselves. This wood material large format camera is light, cheap and supplies most of the functionalities we need. We use this view camera as the front end of our scan camera. Since we directly digitize the image plane, we removed the film and film holder and attached the flatbed scanner to the back holder.

The back holder of the camera has actually 10" by 10" imaging area. In film photography, a 10" by 8" film is slided into the film holder, which is attached to the back holder. The square back holder makes it possible to take landscape and portrait photos by putting the film holder in different directions. When we remove the film holder, the dimension of the flatbed scanner becomes the constraint. The effective

scanning area of the scanner is about 11" by 8.5". By taking the intersection of these two imaging dimensions, we get the actual imaging area for our scan camera: 10" by 8.5". The lens is a Nikkor M 300mm f/9 large format camera lens, which is an inexpensive large format lens which supplies high resolution and contrast.

3.2 Flatbed Scanner

3.2.1 How flatbed scanners work

The imaging process of most flatbed scanners is illustrated in Figure 3.1. A scan head scans across the document, one scanline at a time. On the scan head, a lamp illuminates the document, and a lens focuses the document on a linear CCD sensor array. In practice, most flatbed scanners use a sensor array shorter than the width of the scanning area, which means, the lens focuses the width of the scanning area to the whole range of sensor. The light striking each sensor element is essentially a cone shape beam, with the bottom at the lens. Since the focusing of the image needs a relatively long light distance, mirrors are used to bend the light path and keep the scan head compact. This design makes the optics consistent at every scanning position. For the two scanning positions in Figure 3.1, the lighting paths are exactly the same.

3.2.2 Problems

Now we want to adapt the flatbed scanner to the view camera. First of all, since we use environment lighting, we removed the light source inside the scanner. To make the lighting path clear, we removed all the materials which could possibly block the light between the camera lens and the sensor, including the glass. We also aligned
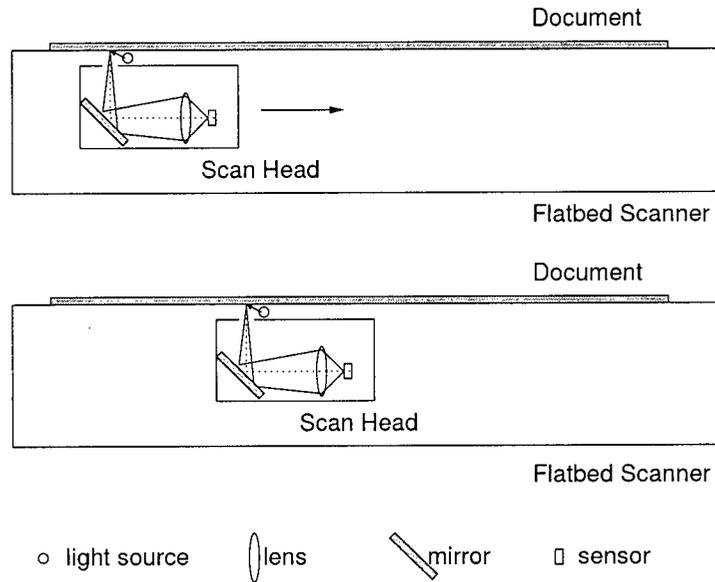
Figure 3.1: How most flatbed scanners work.

the scanner with the camera by facing the scanner towards the lens of the view camera. These changes are illustrated in Figure 3.2.

However, there is still a problem to realize our camera design: the inconsistent optics. In the two scanning positions depicted in Figure 3.2, the scanner scans different positions on the same focusing plane. In the flatbed scanner, both the orientation of the mirror and the direction of the reflection are constant over time. According to the law of reflection, the direction of the incoming light has to be constant. Obviously, this is not the case between the two positions in Figure 3.2. To make the optics work, the orientation of the mirror has to change over time. But it's very hard to make this change on the scanner.

An alternative is to remove the lens and mirror from the scanner, and directly face the sensor forward. We made such changes on a HP Scanjet CCD scanner. After we made those changes, it's hard to keep the optical system stable any more. The

21

o light source   () lens   \ mirror   □ sensor
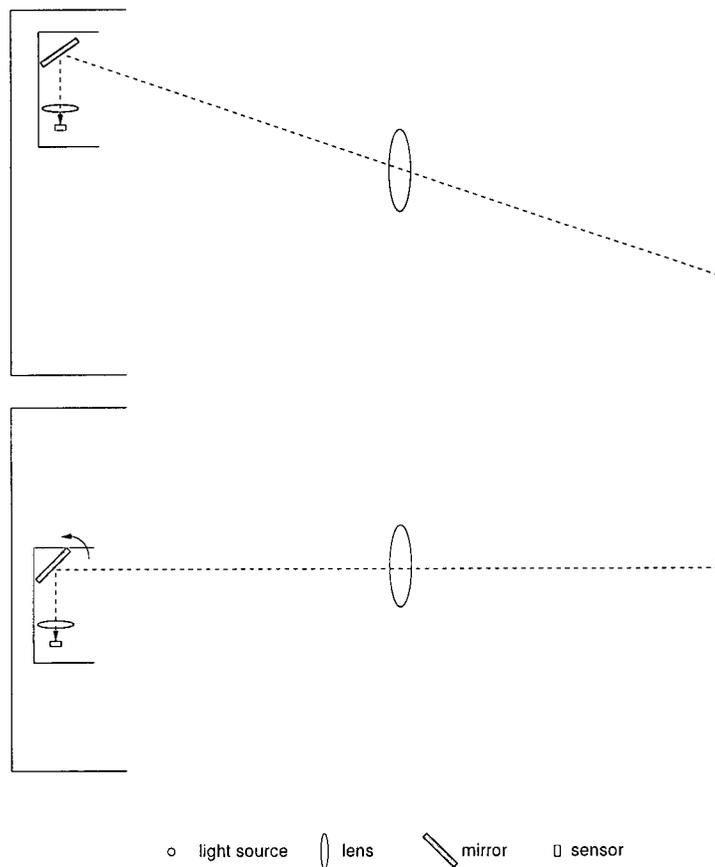
Figure 3.2: Optical problem of conventional flatbed scanners when they work with the large format camera.

other thing is that by using a CCD sensor shorter than the actual scanning area, we will lose some field of view.

3.2.3 Solution

The essential part of our system is a Canon LiDE 30 flatbed scanner. In contrast to conventional flatbed scanners, this model doesn't use any mirror combinations in the camera head. Instead, it has an image sensor covering the whole width of the scan area (about 8.5"), and the focusing is achieved through an array of rod

lenses directly on top of the sensor. The concept of this technique is depicted in Figure 3.3. A three-color (RGB) LED (light-emitting diode) illuminates the scan area uniformly through a light guide, one color at a time. The sensor itself, on the other hand, measures the the whole spectrum of the incoming light and doesn't have any color information. Since the light has only one color channel, the sensor thus obtains the measurement of the same color as the light. After all three color channels are measured, the full color for all pixels are read out and the ultra-micro motors drive the scan head to the next position.
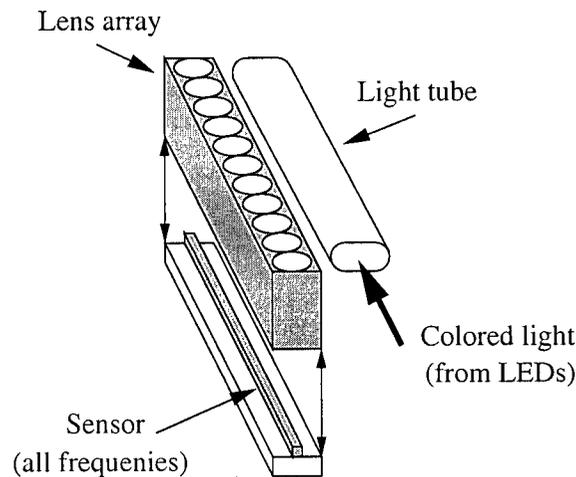
Figure 3.3: LED in Direct Exposure(LiDE) technology in the design of Canon scanner.

The reason we chose this scanner model is that we don't need complicated optical and mechanical changes on the scanner. In our system, for the purpose of taking pictures of real objects, we remove the LED and light guide from the scanner, so that only the light from the lens arrives at the sensor. To obtain color images, we take three images with different color filters (arranged in a filter wheel). The color

wheel is put in front of the camera lens. This method is suitable for both single-shot and scan cameras, although in this way it can no longer be used for moving targets. While this methods add some complexity to the image acquisition process, it allows for additional benefits by using non-standard filters in some specific ranges of light spectrum.
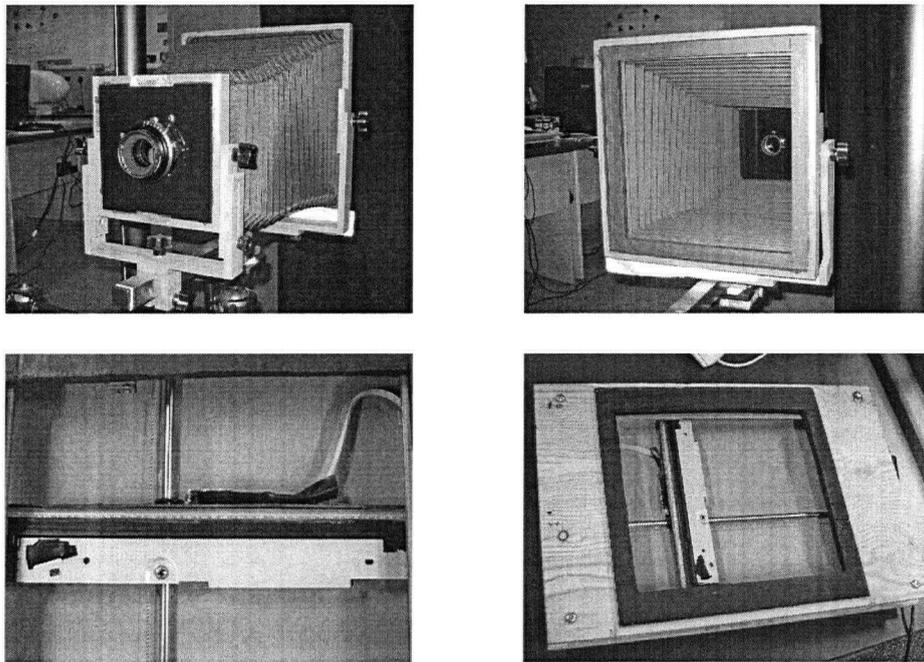
Figure 3.4: Photos of the front end (top) and backend (bottom) of our scan camera. At the top are two different views of our large format camera. The bottom-left is the sensor of the scanner, and the bottom-right is the actual backend.

The rod lens array is originally used to focus the illuminated document on the sensor. Now that we already have a lens on the large format camera, we don't need this lens array for focusing any more. The very small angle of view of these rod lens also prevent us from using them in our system. We removed this lens array and aligned the sensor on the image plane of the large format camera. Figure 3.4

shows the pictures of the front end and back end of the scan camera.

3.3   Imaging Issues

The integration of the view camera and the flatbed scanner brings several issues in the work process of our scan camera. Since we substituted the large format film with the flatbed scanner, the imaging process changed from one-shot to scanning. Because of this and some other changes in the imaging process, we encountered some focusing and lighting issues.

3.3.1   Focusing

Analog large format cameras use a frosted glass as the focusing aid. Covered by a piece of black cloth, the photographer can check the image on the glass and adjust the camera. When the image is properly focused, the focusing aid is replaced by the film.

However, since the original focusing aid and the flatbed scanner have different dimensions, their image planes do not align well. For our scan camera, we cannot do focusing by using the original focusing aid. We can resolve this problem by building our own focusing aid. We found that we could use the preview and scan function of the scanner to focus the image. We first used the preview to composite the scene and got approximate focusing. Then we further refined the focusing by scanning a small region in high resolution. In this way, we solved the alignment problem because the scanner is never detached from the view camera.

3.3.2 Lighting

Because our scan camera has an extended imaging process, the illumination has to be consistent over time. However, light sources like fluorescent lights have flickering effects, which results in different illumination levels at different scanning positions. Using such light sources, there may be some regular color patterns in the acquired images. There are some flicker-free light sources used in television and film industries. For our scan camera, we used a pair of K5600 [19] HMI (metal halide) lamps. The lamps have bluish color and are very similar to daylight. Note that we need to do calibration based on these light sources.

3.4 Design Summary

Our objective on this camera is high resolution and cheap price. The resolution we get is 122,200 pixels, a 10" by 8.5" image plane digitized at 1,200 DPI. The large format camera kit costs about $600, and the lens is about $900, which is the most expensive part in our system. In addition, we built a color wheel with cardboard and photographic filters made from polyester. We use Lee Tricolor filters model 25, 58, 47B for red, green, blue, which have peak transmission at 630, 540 and 440nm respectively. The fourth slot of the color wheel either remains empty for grayscale photography or is loaded with an infrared filter for near infrared imaging. The price of the filters is $20 each. When taking color images, we use a UV/IR cutoff filter($60) screwed on the lens. The total cost of our camera is less than $2,000. During the construction of the camera, we found that the difficult part is how to make the view camera and the flatbed scanner work together.

Chapter 4

Software Setup

Since we removed the LED and rod lens array, the calibration process embedded in the original scanning software won't work in our system. Instead, we get the uncalibrated images from the scanner, then apply our own calibration algorithms on them. The calibration algorithms take raw image data from the scanner as input, and output the calibrated grayscale or color images.

The essential functionality of our software system is to remove the variations introduced by our camera, and to compensate for different lighting conditions. In this chapter, we first describe the calibration pipeline for grayscale and near infrared imaging, then discuss the color calibration algorithm we realized on color images. At the end, we present an automatic approach to further improve the image quality by detecting and removing image defects.

4.1   Grayscale Imaging

The scan camera can acquire grayscale images without a color filter. We use the shareware program Vuescan [16] to get raw images. The whole $10'' \times 8.5''$ visible

area is scanned at 1,200 DPI in 16-bit grayscale mode and transfered to the host computer. Figure 4.1 is the basic calibration process for grayscale images. Firstly, the dark-current noise and flat field response are calibrated with a single linear mapping. Then the effects of scratch or dust are interpolated. Finally, a gamma correction step yields the end result.

Figure 4.1: Processing pipeline for Grayscale images.

4.1.1 Dark Current and Flat Field Response

Like all other image sensors, the sensor elements in the scanner are not perfectly identical on their characteristics. The most important sources of variations for both CCD and CMOS sensors are *dark current noise* and *flat field response* [18].

- Dark current means that a pixel may exhibit non-zero measurement even when there is no incoming photon. This phenomenon comes from the thermal energy within the silicon lattice comprising the sensor. Dark current effectively adds a positive constant to the response of each particular sensor element. The longer the exposure time is, the more dark current noise is accumulated.

- The relationship between the sensor response and the amount of incoming photons is very close to a linear mapping. However, the slope of this mapping varies slightly from element to element in the sensor array. This is called flat field response.

28

To correct these two effects, we take one "black" image (i.e., at zero aperture) and one "white" image (i.e., take an image of a piece of white paper). When taking the "white" image, the white paper is put out of focus to get even illumination. In order to make use of as much dynamic range as possible, we expose the paper just below saturation.

Intuitively, we only need to characterize the response of every sensor element. We averaged the pixels on the same scanline for both black and white images, then these average values defined a linear mapping that deals with both dark current and flat field response. However, this didn't work well for our scan camera. Although the sensor array is one dimensional, the whole imaging area has an extra dimension: the scanning direction. For a specific sensor element, the lighting conditions at different scanning positions are different. Exhibited on the raw image of the white paper, the pixels on the same scanline have different intensities. From Figure 4.2, not only can we see the different responses for different sensors (vertical), but also along the scan direction (horizontal).

Because of this observation, we calibrated the scan camera not only on each sensor element, but also on different positions of the same sensor. Assuming we have the "white" image I_w, the "black" image I_b and the raw image of the real scene I, the linear mapping calibration is:

$$I'(i,j) = 65535 \times \frac{I(i,j) - I_b(i,j)}{I_w(i,j) - I_b(i,j)}$$

4.1.2   Scratch/Dust Removal

The previous step removes the linear variations introduced by the scanner. In this section, we discuss the nonlinear variations in the imaging process and the way we dealt with them.
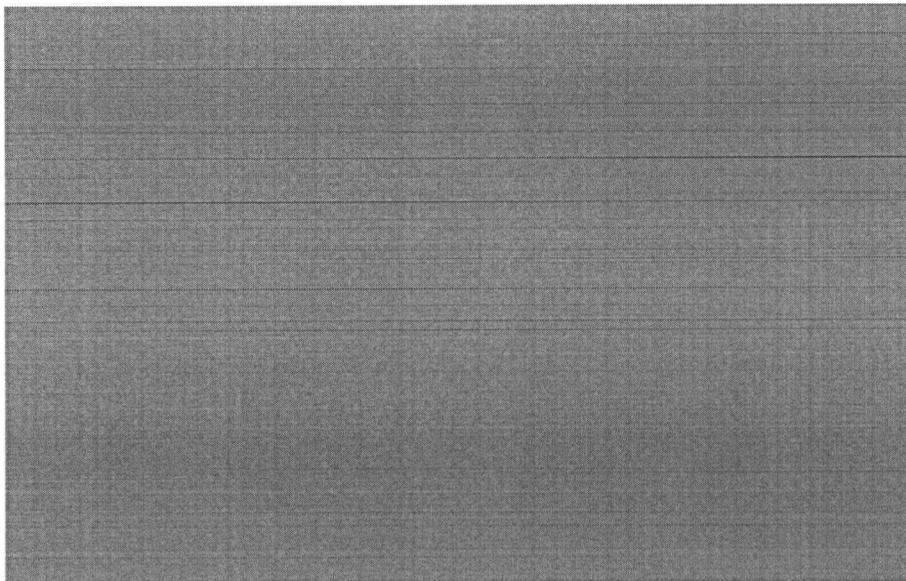
Figure 4.2: The raw image of a piece of paper with even illumination.

For flatbed scanner, it's important to keep the optical parts clean to obtain good image quality. Commercial flatbed scanners are assembled in dust free assembly line, and the light path has good optical quality. However, since we exposed the sensor in the air by removing the LED and rod lens array, scratches, particulates and other surface contaminates may degrade the quality of the images. After we applied the linear calibration on the images, some of the horizontal/vertical artifacts still existed. Our guess was that the effects of these scratches or dusts were not linear, so they were not easy to be calibrated out by the linear calibration.

Before installing the scanner and the view camera together, we cleaned the surface of the sensor with a soft cloth soaked with ethyl alcohol, and a pressed gas blower [11]. The scratches are hard to remove, and there may be dusts falling on the sensor when we put the scanner into the wood frame. We dealt with these defects by masking the corresponding scanlines and interpolating the color from the

30

neighbouring pixels. Since there are only a small number of such defects compared to the total resolution, and each defect only influence 1-2 scanlines, our simple approach works fairly well.

4.1.3 Gamma Correction

The final step is gamma correction depending on the characteristics of the specific display device and a subsequent reduction of bit depth to 8 bits.

Human eyes are sensitive to the ratios of intensity rather than the intervals, so in image displaying devices, the intensity levels should be placed logarithmically instead of linearly. The actual output intensity I for the computer display is:

$$I = kN^\gamma$$

where N is the number of electrons and k and γ are some constants. For most CRT monitors, the γ is roughly $2.0 - 2.5$. When our images are too dark, we can make them brighter by changing the gamma value. The Canon LiDE 30 flatbed scanner uses National LM9833 sensor array [23], which has 16 bits per channel. We reduced the color depth to 8 bits in favor of most image processing softwares. The results of the calibration process are shown in Figure 4.3.

4.2 Near Infrared Imaging

Image sensors are not only sensitive to visible range of light spectrum (400nm - 700nm), but also to light in near infrared spectrum (700nm - 1,200nm). This gives us the flexibility of taking infrared images by means of an IR filter. We loaded a Lee infrared (No. 87) filter in the 4th slot of our color wheel to stop all visible light. By
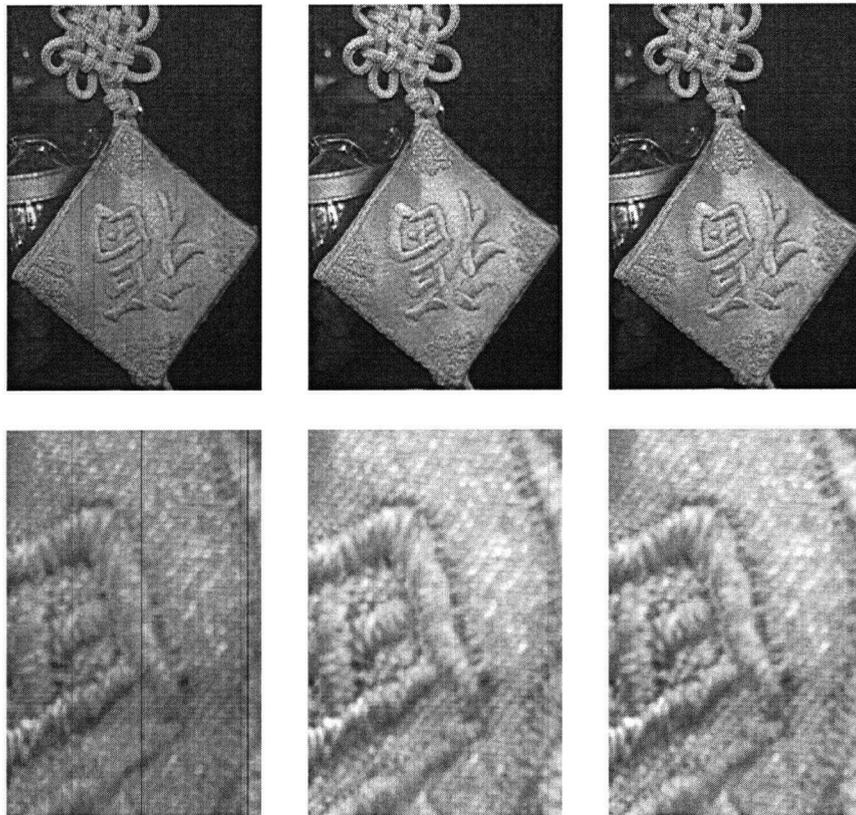
Figure 4.3: Effect of the black&white calibration step. Left: raw sensor data. Center: after dark current subtraction and flat-fielding. Right: calibrated image after interpolating faulty columns.

using a color filter with the camera, we can acquire an image at arbitrary spectrum range as long as we have the proper filter.

The calibration of near infrared images is the same as grayscale images. Figure 4.4 is a comparison between a grayscale image and an infrared image of a painted ceramic figurine taken with our camera. Note that in the infrared image, the pigment effects are almost completely removed. The ability to eliminate textures and pigmentation effects makes infrared imaging fit in the applications of industrial inspection and food processing [10]. The infrared image will also be a better input

to a shape-from-shading algorithm than a grayscale image.
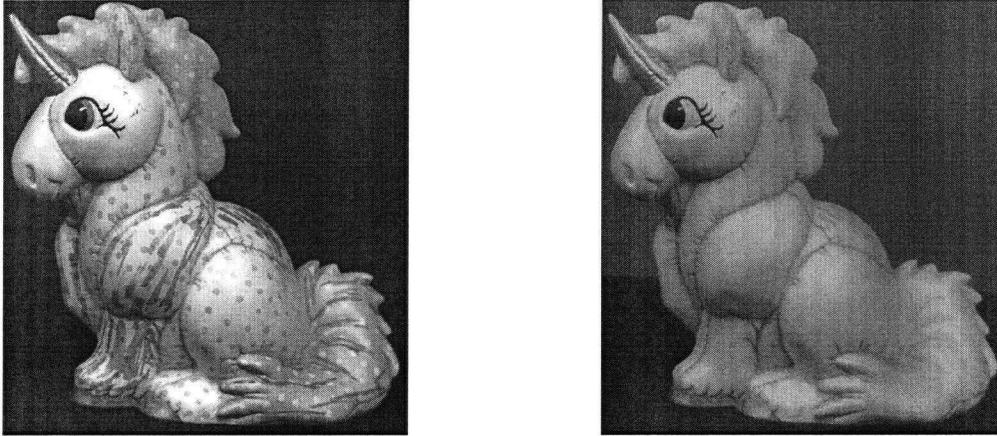
Figure 4.4: Comparison of a grayscale photograph (left) with an infrared photograph (right). Note that except for the black color of the eye, the effect of different paints is almost completely removed in the infrared image.

4.3 Color Imaging

In color imaging, we acquire one image for each of three primary color channels. For each channel, we do the same black&white calibration as in Figure 4.1 except the final gamma correction step. We then align and merge the three channels together to get a color image, which undergoes color calibration step to the final image. The gamma correction is actually absorbed in the final color calibration step. The whole calibration process is illustrated in Figure 4.5.

As we mentioned in the previous section, the scanner sensor is not only sensitive to the visible light, but also to infrared. On the other hand, human eyes are not sensitive to infrared. This means that infrared will make color look wrong in normal photography. Although most digital cameras use a hot mirror filter to block out the infrared, scanners do not need such a filter because there is very

little infrared component from the light source. However, there is plenty of infrared coming from real objects in the environment. We want to make sure that there is no infrared component in the acquired images.

The first approach we tried was to remove the infrared component from the red image. We took two images for the red channel: one with the red filter only, the other with the red filter and the IR filter. Note that by overlapping the red filter and the IR filter, only the intersection of the two filters' spectrum was captured. The subtraction of the two images should be the red channel image without any infrared component. However, this method didn't give satisfactory results because there was also infrared passing through the green and blue filters. To eliminate infrared from all the color filters, we need to do 6 scans instead of 3, which will increase the imaging time significantly. We avoided this problem by simply using a combined UV and IR cutoff filter (Schneider Optics B+W Filter no. 486). We measured the three color channels of a white card with and without the IR cutoff filter, from which we computed how much infrared passed each color filter. The result is depicted in Table 4.1, in which we can see the amount of infrared passing the green and blue filters without the IR cutoff filter.

Color Filter	Without IR Cutoff	With IR Cutoff	Infrared Amount	Percentage
Red	27502	9318	18184	66.1%
Green	15218	3779	11439	75.2%
Blue	10869	1756	9113	83.8%

Table 4.1: The measurements of the three color channels of a white card, with and without IR cutoff filter.
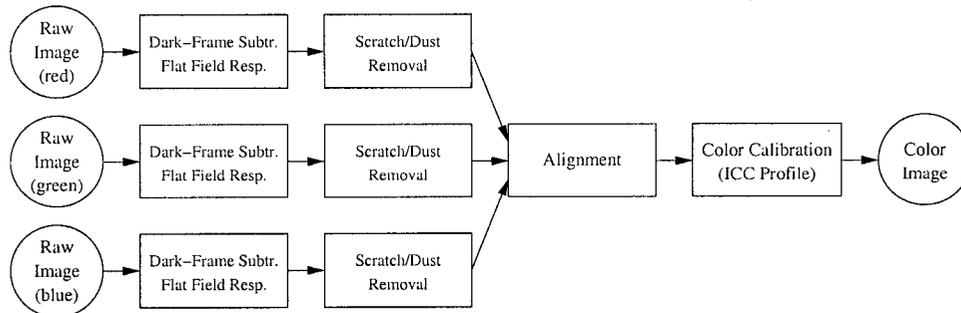
Figure 4.5: Processing pipeline for color imaging: the same calibration processes as in Figure 4.1 are made for each color channel, then the three calibrated images are aligned and merged together. At last is a color calibration process to get color fidelity.

4.3.1 Channel Plane Alignment

The images of different color channels align very well in the direction along the sensor array. On the scanning direction, especially when scanning in the full resolution, there are often several pixels offset between different channels. This results from the small differences in the positioning of the scan head when the scan begins. To compensate for this, we select a region of the image with high frequency details, and determine the one dimensional offset by manual alignment. The future work is to do alignment automatically by running an edge-detection algorithm on the region in three images, then searching in one direction for the best offset.

4.3.2 Color Calibration

ICC Profile

After converging the three channels, we did color calibration to get the result image. A standard method for color calibration is to use some professional color

management system and ICC profile. This method is used by many digital imaging devices [26].

ICC profile is a data file describing the color characteristics of a specific device. The main purpose of the profile is to be used by color management software to maintain color consistency across various devices. By means of the profile, the color management software first converts the color space of the source device to a device independent color space, which is called common profile connection space (PCS). CIE LUV and CIE LAB are two commonly used color PCS's, and can be converted to each other. An destination profile can in turn convert the image data from the PCS to the color space of the output device before it's displayed or printed. Gamma correction is taken care of as a transfer function in the profile of the display device. In our case, the source device is the scan camera, and the destination device is the computer monitor or printer. We use SCARSE [15], a free color calibration and management software package, and a Kodak Q60 color target to generate the profile of our camera.

The Kodak Q60 color target is used to do color calibration for flatbed scanners. When such a color target is obtained, its specific device independent color data can be downloaded from Kodak's website. The target is also scanned by the flatbed scanner. The color management software then reads in these two images and generates the color profile for the flatbed scanner. In our case, we took a photo of the color target, and used SCARSE to generate the profile of our scan camera.

Figure 4.7 are the images of the color target before and after the color calibration process. The leftmost one is the image in scan camera color space, which has more red component than green and blue when reproduced in monitor color space. We generated the profile from this image. At the same time, we generated

the profile of the computer monitor. The middle image was acquired by transferring the source image to the computer monitor color space using the two profiles.

Note that for different light sources, we need to generate different profiles. Even for the same light source, the different positions and orientations of the light source may still generate different profiles. Because of this problem, it's hard to acquire nice color calibration results on images of real world objects.
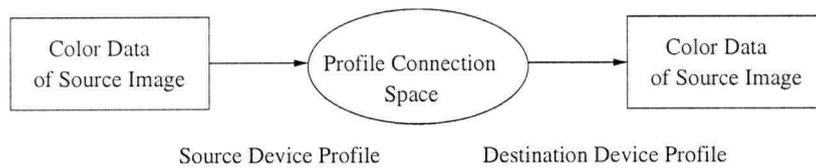

Figure 4.6: Color calibration with ICC profile: the image on the source device is transferred to the profile connection space(PCS), then to the color space of the destination device.
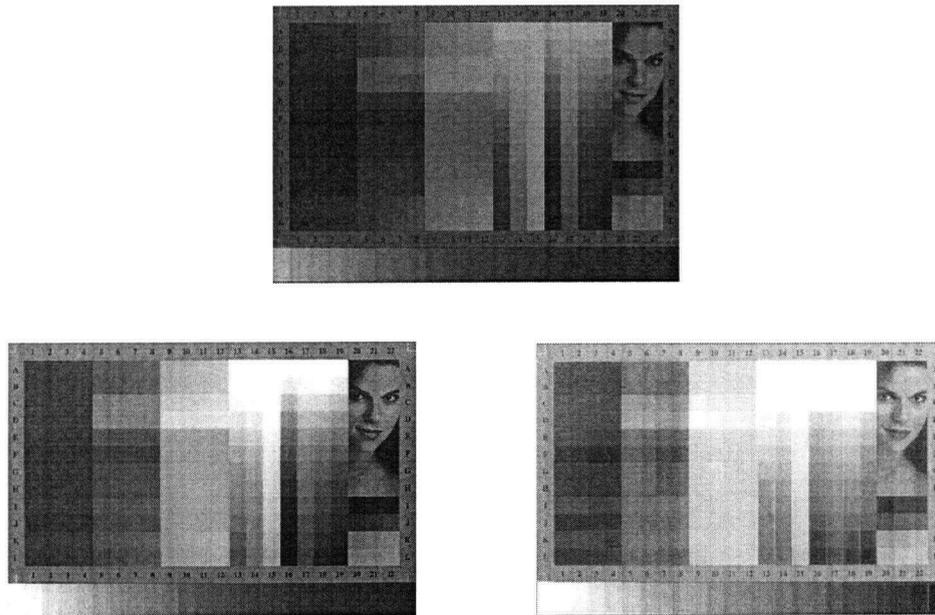

Figure 4.7: A picture of the color calibration target without any calibration (top), with white balance (bottom left), and with ICC profile color calibration (bottom right).

White Balancing

White balancing is a color adjustment method widely used in conventional photography. Human eyes compensate for the color tones in different lighting condition, although the precise method is not known. Digital cameras, on the other hand, need to know what the "white point" (the color of that point in current lighting condition is white) is to correct other colors cast by the light sources. Most digital cameras can calculate the white balancing automatically by evaluating the color of the whole image and calculating the best fit "white point". This automatic balancing does not always work, especially in the images with very few dominant colors. Middle to high end digital cameras often have the functionality of manual white balancing. The user selects a point which is supposed to be white, then the whole image is processed by white balancing algorithms.

We used this easier approach in our camera system by scaling the individual color channels. It is essentially the same as exposure compensation in astronomy tri-color imaging [24]. With the same aperture stop and the same exposure time, in the image of a pure white card illuminated by white light, the measurements of the red, green, and blue channels are not equal. The red component is larger than the green and blue because the image sensors are more sensitive to red light. In tri-color photography, this is compensated by different exposure time. In our system, we can't change the exposure time of the scanner. Instead, we scaled the sensor data accordingly.

Given the image of a white card, we calculate the average pixel value of each channel, e.g. R, G, and B. Assuming the image of the white card should be pure white in the current lighting, we calculates the scale factors for exposure compensation: $255/R$, $255/G$, and $255/B$. If the color image is I, the balanced

image is:

$$I'_R = \frac{255}{R} \times I_R$$

$$I'_G = \frac{255}{G} \times I_G$$

$$I'_B = \frac{255}{B} \times I_B$$

The rightmost image in Figure 4.7 is the image calibrated by this white balancing method. The color quality is close to the one calibrated by ICC profile. The advantage of white balancing is that the computation overhead is much less than ICC color calibration, and it's more robust to minor lighting changes.

4.4 Image Recovery

In the calibration process we discussed, we remove the scratch/dust effects by manual selection and linear interpolation. Although there are a small number of artifact scanlines compared to the resolution of the image, there are still many of those to be picked up when using high resolution. The selection process is inefficient and time-consuming. Linear interpolation removes the artifacts, but at the same time, it also blurs the image and loses some useful information.

In this section, we discuss an automatic process to do artifacts detection and removal. In the first step, a detection algorithm is applied to estimate the possibility of each scanline being artifact. Then we implemented the image inpainting algorithm [5] to recover as much information as possible at the damaged regions.

4.4.1 Effects Detection

The scratch/dust effects are either horizontal or vertical lines, but they are always the same as the scanning direction. The horizontal or vertical lines do not necessarily

cover the whole scanlines. Based on this observation, we proposed the detection procedure as in Figure 4.8. Suppose the direction of the artifacts is horizontal, firstly a horizontal Sobel Operator filters the image to emphasize the horizontal edges. We then apply Hough Transform on the edge image to further characterize horizontal lines in the edge image. Hough transform is a global voting method for finding straight lines with different orientations in an image. In our case, we calculated the average of each scanline in the edge image. After Hough Transform, the scanlines with horizontal artifacts will have larger values. Finally, the result of the Hough Transform produces a histogram and is thresholded to estimate the artifacts positions.
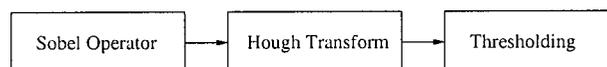
Figure 4.8: Process of scratch/dust effects detection.

Given the facts that the scratch/dust effects do not necessarily cover the whole scanline, we divided the image into tiles and applied the detection method on each tile. If the size of the tile is too small, some actual edges will be also detected as artifacts, which is not what we want. The tile size we chose is 1000 by 1000 pixels. Using tiles of this size, the artifacts can be detected fairly easily, without picking up actual edges. Figure 4.9 is an example of a specific detection, and it is a a 290 by 360 region in a 1000 by 1000 tile. After the detection on each tile, all the tiles are composed together as the mask for the whole image.

4.4.2   Image Inpainting

Image inpainting is an effective approach to restore damaged paintings or photographs. It can also remove selected objects from photographs. Given the source
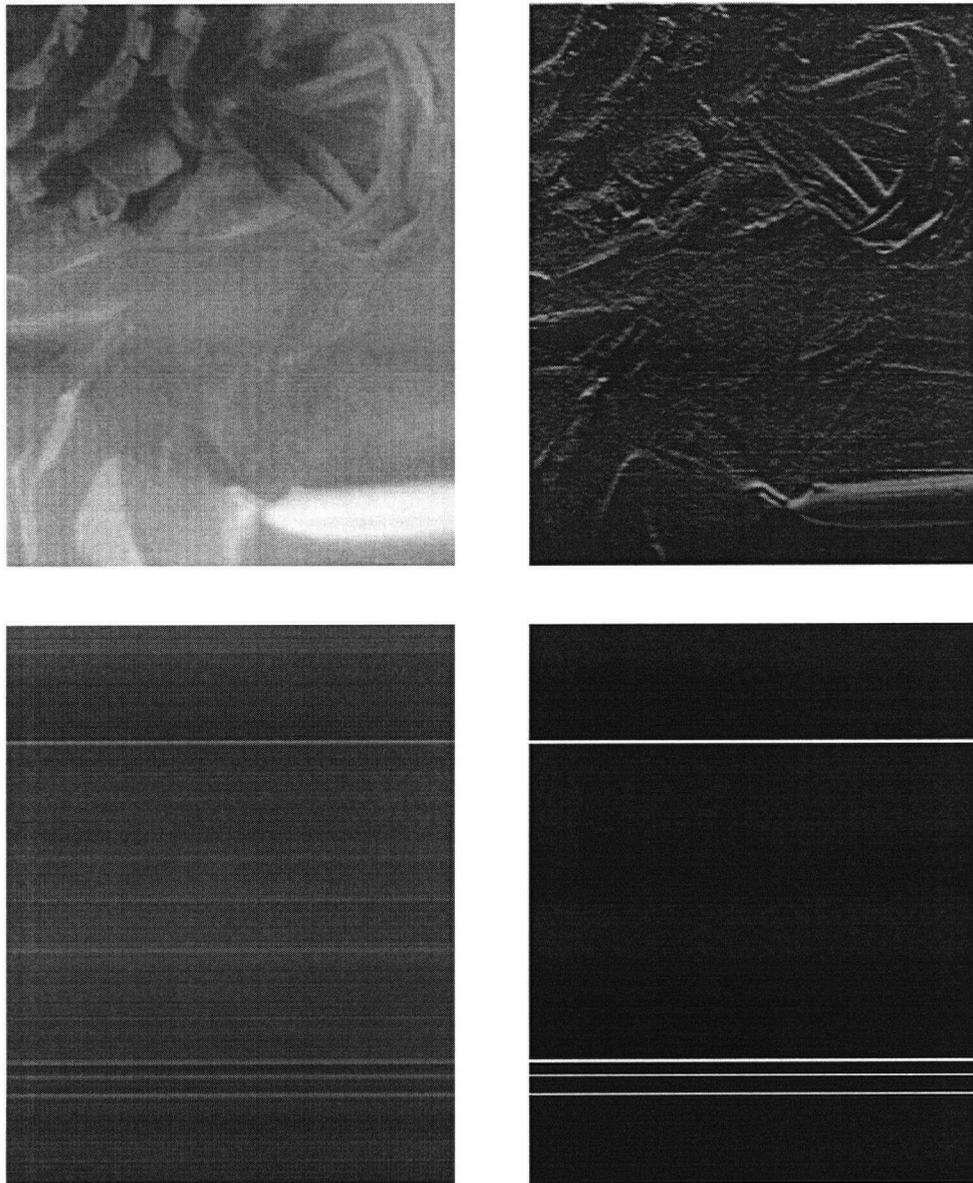
Figure 4.9: An example of the noise detection method. The source image(top left) is filtered by a horizontal Sobel Operator(top right), followed by Hough Transform (bottom left) and thresholding(bottom right).

image and the mask area to be recovered or removed, this approach fills the masked area by extending the isophote lines from the unmasked area [5].

The inpainting is an iterative algorithm, starting from the source image I^0. Let Ω stands for the region to be inpainted, and $\partial\Omega$ is the boundary of the region. In each iteration, an update image I_t^n times the step size $\triangle t$ is added to the image I^n. I^{n+1} is the "improved" image. The iteration stops when the update image I_t^n is small enough. Note that the update of the image only occurs in the inpainting region Ω. The following are the actual numerical equations of inpainting process taken from [5].

$$I^{n+1}(i,j) = I^n(i,j) + \triangle t I_t^n(i,j), \forall (i,j) \in \Omega \tag{4.1}$$

where

$$I_t^n(i,j) = \left(\overrightarrow{\delta L^n(i,j)} \cdot \frac{\overrightarrow{N(i,j,n)}}{\left| \overrightarrow{N(i,j,n)} \right|} \right) |\nabla I^n(i,j)| \tag{4.2}$$

$$\overrightarrow{\delta L^n(i,j)} := (L^n(i+1,j) - L^n(i-1,j), L^n(i,j+1) - L^n(i,j-1)) \tag{4.3}$$

$$L^n(i,j) = I^n{}_{xx}(i,j) + I^n{}_{yy}(i,j) \tag{4.4}$$

$$\frac{\overrightarrow{N(i,j,n)}}{\left| \overrightarrow{N(i,j,n)} \right|} := \frac{(-I_y{}^n(i,j), I_x{}^n(i,j))}{\sqrt{(I_x{}^n(i,j))^2 + (I_y{}^n(i,j))^2}} \tag{4.5}$$

$$\beta^n(i,j) = \overrightarrow{\delta L^n(i,j)} \cdot \frac{\overrightarrow{N(i,j,n)}}{\left| \overrightarrow{N(i,j,n)} \right|} \tag{4.6}$$

and

$$|\nabla I^n(i,j)| = \begin{cases} \sqrt{(I^n{}_{xbm})^2 + (I^n{}_{xfM})^2 + (I^n{}_{ybm})^2 + (I^n{}_{yfM})^2} & , \text{when} \quad \beta^n > 0 \\ \sqrt{(I^n{}_{xbM})^2 + (I^n{}_{xfm})^2 + (I^n{}_{ybM})^2 + (I^n{}_{yfm})^2} & , \text{when} \quad \beta^n < 0 \end{cases} \tag{4.7}$$

For each iteration, the update image is the image information propagated to the masked region along the isophote direction (4.2). The information to be

propagated is the 2D smoothness estimation of the image $L^n(i,j)$, calculated as in (4.4). The progagation of this information is realized by the 1st derivative or gradient of the smoothness estimation (4.3). The isophote direction is a vector perpendicular to the image gradient (4.5). Then the projection of $\overrightarrow{\delta L}$ onto the direction $\overrightarrow{N}$ is computed in (4.6). Since a central differences realization will turn the numerical algorithm unstable, the norm of a slop-limited gradient (4.7) is multiplied to β^n in the final step.

To ensure the correct evolution of the algorithm, an anisotropic diffusion process is interlaced with the image inpainting process described above. Every few inpainting iterations, a few diffusion steps are applied to curve the isophote lines so that they won't cross each other. The anisotropic diffusion has the advantage of not losing sharpness of the reconstructed region. We use the same anisotropic diffusion equation as in [5]:

$$I_t(i,j) = g_\epsilon(x,y)\kappa(i,j,t)\left|\nabla I(i,j,t)\right|, \forall(i,j) \in \Omega^\epsilon \qquad (4.8)$$

where Ω^ϵ is a dilation of the Ω with a ball of radius ϵ, κ is the Euclidean curvature of the isophotes of I and $g_\epsilon(x,y)$ is a smooth function in Ω^ϵ such that $g_\epsilon(x,y) = 0$ in $\delta\Omega^\epsilon$ and $g_\epsilon(x,y) = 1$ in Ω. In our case, we use the mean curvature algorithm [12] as the diffusion operator.

Figure 4.10 is the inpainting of the image in Figure 4.9. Note that there are several horizontal artifacts near the highlight area, and they are removed in the inpainted image.
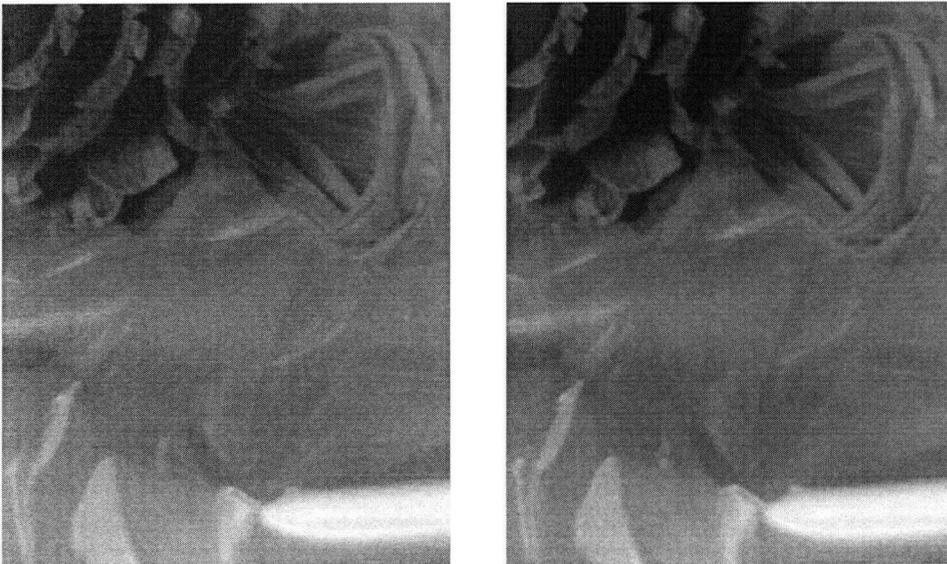
Figure 4.10: An example of image inpainting algorithm: The two images are before(left) and after(right) image inpainting respectively.

Chapter 5

Results

Our scan camera is connected to a Pentium 4 1.6GHz computer with 1 gigabytes memory, on which the scanning and calibration program run. The operating system is Redhat 2.4.18-27.7.x.

In order to obtain quantitative properties of our camera, we compared it against a Canon EOS D60 digital camera with a Canon EF $28 - 105$mm zoom lens. The camera is a 6 million pixels SLR camera. Note that the price of this camera is twice as expensive as our scan camera.

5.1 Light Sensitivity

Every camera has a certain light sensitivity. The light sensitivity of analog films in conventional film camera is also called ISO speed. The most common speeds are ISO 100, 200, 400 and 800. Each level is twice sensitive to light as the previous level. In photography, different light sensitivities are obtained by using films of different speeds. Digital cameras, however, only have one "film": the sensor array. Their light sensitivity is often expressed in the same way as analog films. In most digital

cameras, the speed can be changed in the setting of the camera.

We tried to get the light sensitivity of the scan camera. We first set the f-stops of both the Canon D60 and the scan camera to $f/9$, which was the largest aperture for the scan camera. After we took an photo of a particular scene with the scan camera, we changed the position and focal length of the Canon D60 to cover roughly the same field of view as the scan camera. Finally, we changed the exposure time to match the brightness of the image from digital camera to the one taken by the scan camera. In this way, we obtained the exposure time as 1/60 second. Since the speed of the digital camera is set at ISO 100, we knew that the light sensitivity of our scan camera is ISO 100 with exposure of 1/60 second, or ISO 300 with exposure of 1/180 second, and so on.

Although we can not directly change the exposure time of the scan camera, we can simulate changing the camera speed by scaling the raw image data from the camera (i.e. increasing or decreasing the slope of the flat field response). Note that by scaling up the raw image data, the noise is magnified as well. This is unavoidable even in commercial digital cameras supplying this feature.

5.2 Optical Resolution

Since the original optical system of the flatbed scanner is changed, the actual resolution of the scan camera is not necessarily equal to the nominal resolution of the scanner. In this section we discuss the resolution measurement approach we applied on the scan camera.

Modulation transfer function (MTF) [28] has long been a usual way to diagnose analog and digital imaging systems. It characterizes the frequency response caused by the lens and aperture for analog image capture. For digital cameras, this

measurement also takes into account the influence of sampling and image processing. If the reflected light from interleaved black and white bars is measured in luminance, the modulation is defined as:

$$modulation = (L_w - L_b)/(L_w + L_b)$$

where L_w is the luminance of the white color, and L_b is the luminance of the black. After the black and white pattern of a specific frequency f is measured by the camera, the modulation transfer function is the division of the output (image) modulation M_O and the input (real objects) modulation M_I.

$$MTF(f) = M_O/M_I$$

The slanted-edge edge gradient MTF is a commonly used technique in MTF measurements [17, 7, 28]: take an image of a slanted edge, then analyze the attenuations of different frequencies in the image. Since the imaging behaviors along the rows (different sensors) and columns (scan direction) are different, we analyzed these two dimensions separately [28]. We used the method by Williams and Burns to generate the MTF of our camera system (units of frequencies are cycles/pixel) as in Figure 5.1.

Based on the Rayleigh model, a modulation transfer value of 0.10 can be used as a criterion to evaluate the resolution [28]. On the other hand, from Nyquist theorem, the maximum useful frequency for any sampling devices is 0.5 cycles/pixel. If the modulation transfer value of the imaging device drops to 0.1 at frequency smaller than 0.5, the actual resolution is less than the nominal one, because the maximum frequency captured by the imaging system is smaller than the sampling frequency. On the other hand, if the modulation transfer value drops to 0.1 at frequency equal to or larger than 0.1, the imaging device has the claimed resolution.
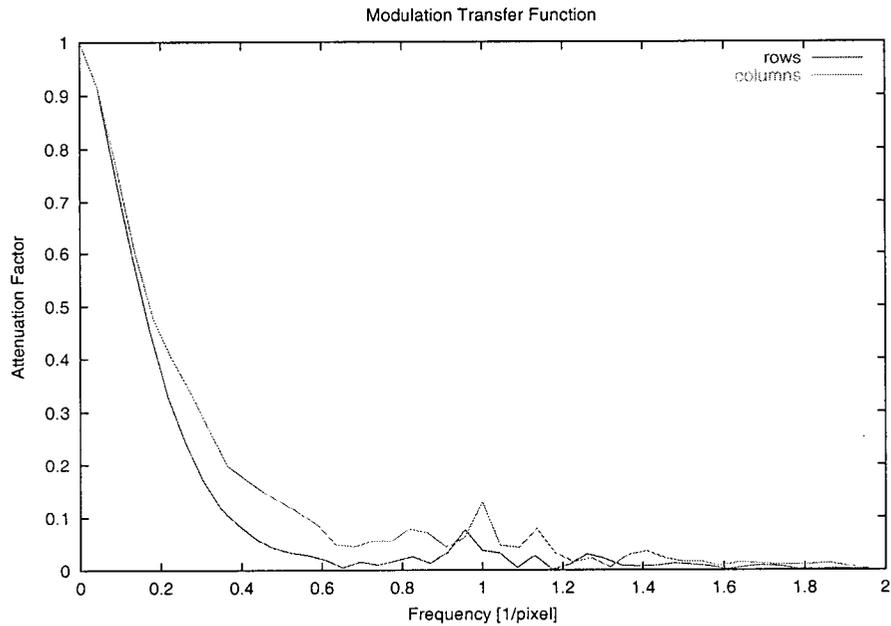
Figure 5.1: The modulation transfer function of our imaging system, in in directions of rows and columns, respectively. Courtesy of Michael Goesele.

If the cutoff frequency exceeds 0.5, there will be some aliasing. In the measurement of the scan camera system Figure 5.1, the 0.1 attenuation factor occurs at 0.36 cycles/pixel within a scanline, which means that there is a little bit of blurring in the scan direction. For the other dimension, the dropoff frequency is 0.54 cycles/pixel, which indicates a small amount of aliasing. From these measurements, we knew that the optical resolution of the scan camera was close to the nominal 1,200 DPI of the flatbed scanner.

5.3   Comparison

To compare the scan camera with the Canon EOS D60, we first set them to their respective highest resolution. The two cameras have different lenses: one is 300mm

48

large format lens, the other is $28mm - 105mm$ zoom lens. This difference is the main concern in the comparison, because different lenses also result in different depth of field and angle of view. We set up the two cameras to cover roughly the same field of view. Since now we have to adjust the lighting and aperture to compensate for different lenses, the illumination in the two images is not directly comparable.

We compared the full images of a plant in the top row of Figure 5.2. They were both slightly cropped and scaled down to fit in the page. We noticed that in the scan camera image, the three color channels at the top of leaves did not align well. This was because in the process of photography, the heat from the light sources dried the leaves. Note that the restriction to take photos of non-stationary objects applies to all scan cameras. In the bottom row of Figure 5.2 we compared a small region from the two images. We cropped a small region from the scan camera image and printed it at 300DPI. The right was the corresponding area taken from the Canon D60 image, scaled 4 times up to match with the size of the scan camera image. We can see significant quality difference in the zoom-in images.

Although the colors in these two images do not match perfectly, the differences are well within the range of color differences one can experience in commercial digital cameras. Different sensors and different color calibration process are both factors to make colors different, which is the case in our comparison. In comparing the two images, we found that green is slightly better in the scan camera image, while the red color is better in the Canon D60 image.

5.4   Full Resolution Imaging

Figure 5.3 and Figure 5.5 are prints of images taken by the scan camera. Both images were scaled down to fit into the pages in 300 DPI. Figure 5.3 is the image of

a Chinese jade ornament. The object itself is roughly 10" by 10" size (not including the knots). We took the image in the highest resolution, i.e. 12,000 by 10,200 pixels, which can be printed on a 34" by 40" poster in 300 DPI. Figure 5.5 is the image of a more colorful setup, and the image size is 9,244 by 5,784 pixels. From the glass grapes, we can see the reflection of the HMI light sources. Figure 5.4 and Figure 5.5 show regions from these two images in original resolution. In Figure 5.4, the uneven reflection from the metal surface is clearly visible. In the bottom image of Figure 5.5, we can even see the individual fur of the beard and the knitted thread in the dwarf's eyes.
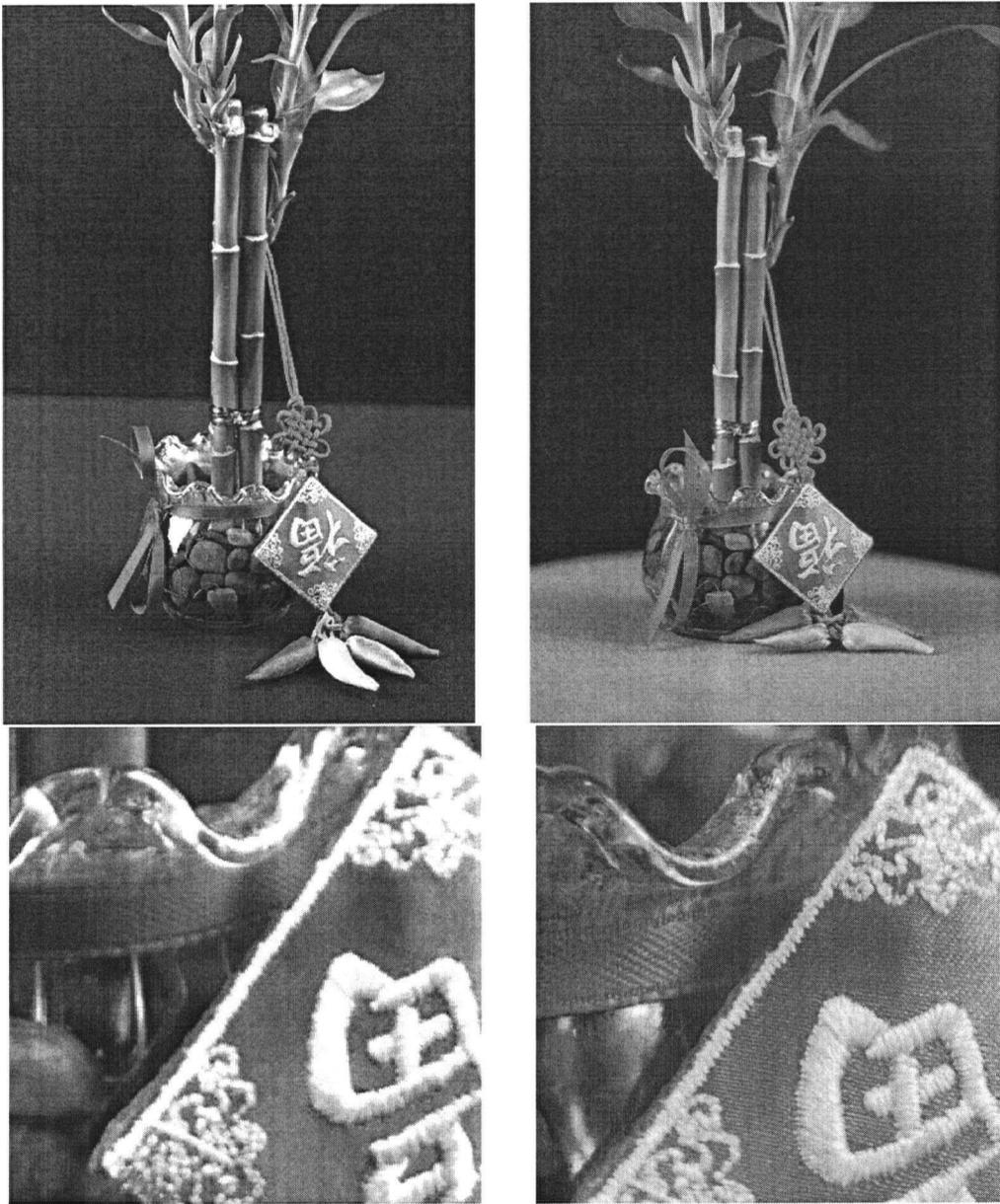
Figure 5.2: Comparison between the Canon EOS D60 (left) and our scan camera (right). The top row shows slightly cropped images. The bottom row shows magnification of one region to better compare the resolution.
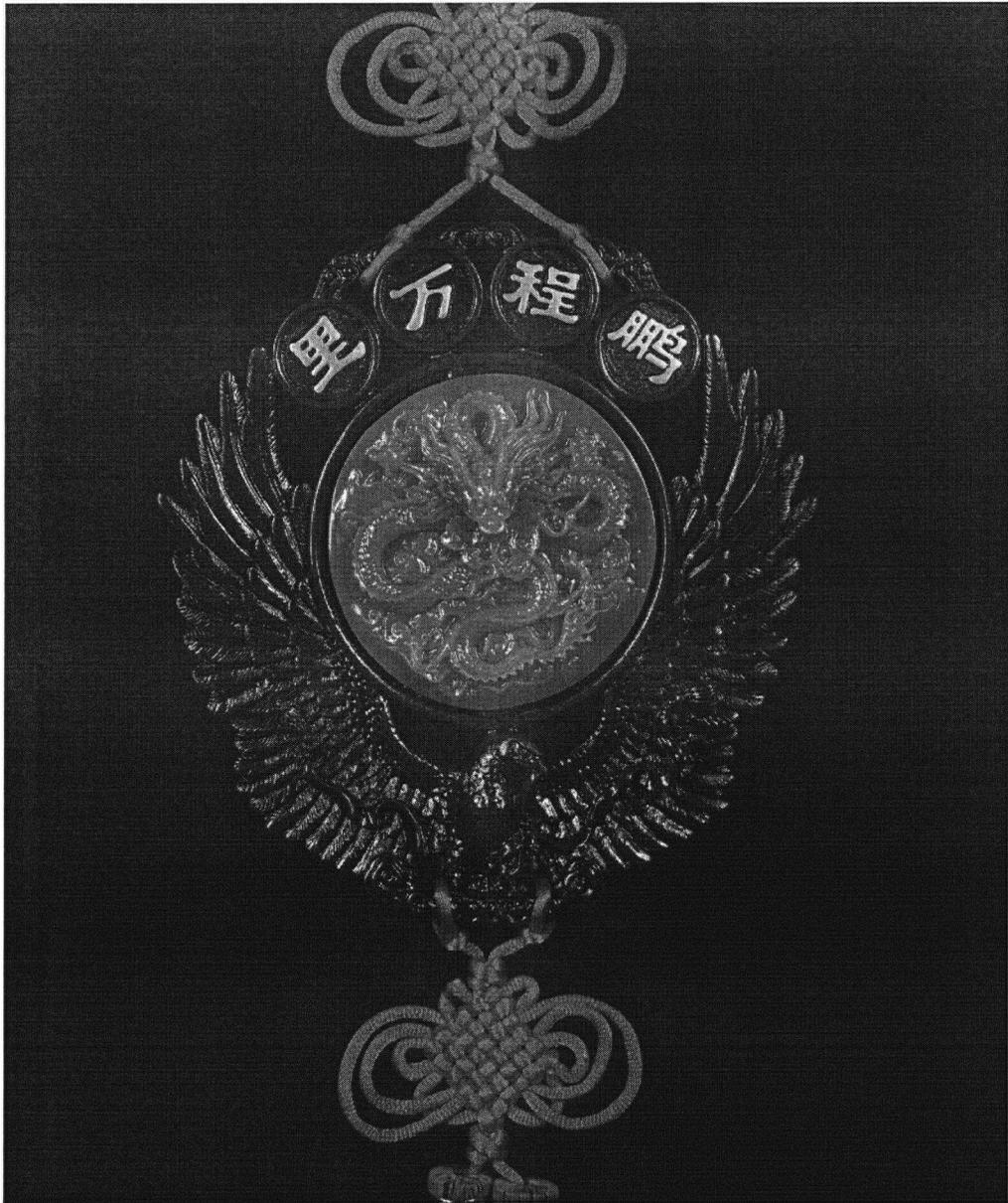
Figure 5.3: A scaled-down print of the image of a Chinese jade wall ornament. The image is taken by the scan camera at the highest 122 million pixels, and the full image can be printed on a 34" by 40" poster in 300 DPI. A region in full resolution is in Figure 5.4.

Figure 5.4: A zoom-in region (1940 by 1650 pixels) cropped from the image in Figure 5.3, printed in 300 DPI.
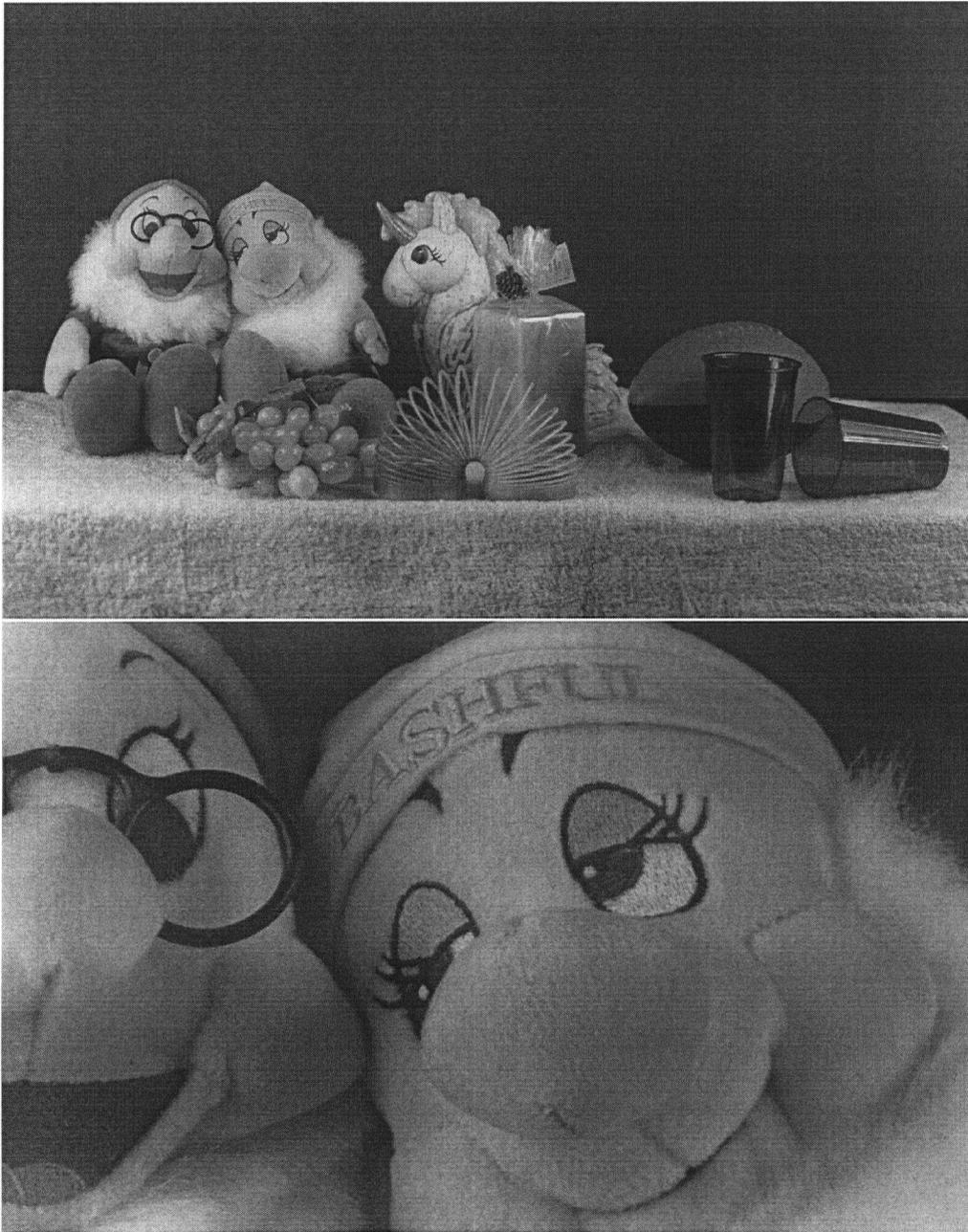
Figure 5.5: A scaled-down print of the image of colorful toys is in the top image. The image is cropped from the full field of view, and the size of the image is 53 million pixels. A region in full resolution is presented in the bottom.

Chapter 6

Discussion and Future Work

In summary, we described the building of a 122 million pixel scan camera. The main contribution is that the camera has very high resolution, and at the same time the cost of the system is reduced a lot by employing a consumer level flatbed scanner and a large format camera kit. We discussed the issues in both the hardware construction and the calibration software development. At the end, we acquired the measurement of the optical resolution, and made a comparison with a semi-professional SLR digital camera.

Next we give some discussion on the limitations and some possible improvements on our camera system. Note that the scan camera easily adapts to advances in the scanner technology. For example, the new Canon LiDE 80 scanner has 2,400 DPI resolution and uses USB 2.0 interface, which will give us the benefits of getting 488 million pixel resolution and higher imaging speed.

6.1 Exposure

One requirement for large format imaging is that more light is needed to illuminate the whole image plane. Because the exposure time of the scanner cannot be changed directly, we chose the HMI light sources to get enough light. If we use ordinary light sources, even when keeping the aperture widely open, the images are still not bright enough. Some commercial digital backends use longer exposure time to work around the need for expensive HMI lights. An alternative to change the exposure is to turn up or down the aperture, with the side effect of changing the depth of field.

6.2 Depth of field

The formula to calculate depth of field is [2]:

$$DOF = F^2 \cdot D \cdot c \cdot f_n \cdot \frac{D - F}{(F^2 + D \cdot c \cdot f_n) \cdot (F^2 - D \cdot c \cdot f_n)} \tag{6.1}$$

where F is the focal length, D is the subject distance, c is the circle of confusion and f_n is the f-stop of the lens. From this formula, we know that as long as the subject distance is not close to the macro range (where D is close to F) and is not close to the hyperfocal distance (where D is close to $\frac{F^2}{c \cdot f_n}$), the depth of field is approximately proportional to the f-stops and subject distance, but inversely proportional to square of the focal length.

The formula to calculate field of view is:

$$FOV = 2 \cdot \arctan \frac{l}{2 \cdot f} \tag{6.2}$$

where FOV is field of view, l is the size of the film, and f is the focal length. For the same angle of view, large format cameras need longer focal length compared to

35mm film camera. For example, we can get the same field of view by a 80mm lens with 35mm film, and by a 400mm lens with 10" by 8" film.

Unfortunately, the longer focal length introduces depth of field reduction as in (6.1). Although the depth of field can be increased by turning down the aperture or increasing the subject distance, these two alternates may easily conflict with other requirements: turning down the aperture will reduce the dynamic range of the captured image, and the change of subject distance will change the field of view as the same time.

The conclusion is that it's a restriction that the exposure time of the sensor can not be changed. If we directly make the exposure time longer, we will be able to increase the depth of field by turning down the aperture without sacrificing the dynamic range.

6.3   Scanning speed and image size

Since the scanning process is clearly the exposure of scanlines interlaced with the data readout, the scanning speed depends on the exposure time and the readout speed. When the scanner works in low resolution, the movement of the scan head is continuous. But in high resolution, it's easy to notice that the scan head stops for some time at each scanline. For a full resolution scan (10,200 pixels per scanline, 12,000 scanlines), it takes about 30 minutes. The scanning speed problem can be solved by new model of scanners with faster data transfer rate.

The raw image data is in 16-bit grayscale uncompressed tiff file format. We save the calibrated color image as 8-bit maximum compressed PNG file. The reason we chose PNG file format is that it supplies lossless compression. Although the PNG file of the full resolution image is 222 megabytes, the calibration program can

not load the three raw channels together into the main memory. In our calibration process, we only read in and process part of the whole image, and write it out before the next part is loaded in.

6.4 Fully automated calibration

Some future research can be done to achieve fully automated calibration. The mechanics and optics of the scan camera can be fine tunned to alleviate the complexity of the calibration process, and the calibration program can be improved for better performance too. The goal is to get an imaging software compared to the commercial ones.

Bibliography

[1] Ansel E. Adams and Robert Baker. *The Camera (Ansel Adams Photography, Book 1)*. Little Brown and Co, 1995.

[2] R. Atkins. Depth of field and the digital domain. http://www.photo.net/learn/optics/dofdigital/, 2003.

[3] B. E. Bayer. Color imaging array. U.S. Patent 3,971,065.

[4] Bender Photographic. http://www.benderphoto.com, 2002.

[5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proc. of ACM SIGGRAPH 2000*, pages 417–424, July 2000.

[6] Better Light. http://www.betterlight.com, 2003.

[7] P. Burns. Slanted-edge MTF for digital camera and scanner analysis. In *Image Processing, Image Quality, Image Capture Systems Conference (PICS)*, pages 135–138. Society for ImagingScience and Technology, April 2000.

[8] Andrew Davidhazy. Demonstration quality scanning digital camera. http://www.rit.edu/ andpph/text-demo-scanner-cam.html, 1999.

[9] Andrew Davidhazy. Very drafty - improvised scanning digital camera. http://www.rit.edu/ andpph/text-better-scanner-cam.html, 1999.

[10] DuncanTech. Multispectral imaging – expanding your vision horizons. http://www.duncantech.com/MS_overview.htm, 2002.

[11] Eastman Kodak Company. Application note: Cover glass cleaning for image sensors. http://www.kodak.com, 2001.

[12] Adel I. El-Fallah and Gary E. Ford. Mean curvature evolution and surface area scaling in image filtering. In *IEEE Transactions of Image Processing*, pages 750–753, 1997.

[13] R. Ferreira, B. Moon, J. Humphires, A. Sussman, J. Saltz, R. Miller, and A. Demarzo. The virtual microscope. In *Proceedings of the AMIA'97 Fall Symposium*, 1997.

[14] Foveon Inc. http://www.foveon.com/X3_tech.html, 2003.

[15] A. Frolov. SCARSE: Scanner CAlibration ReaSonably Easy. http://www.scarse.org/, 2002.

[16] Hamrick Software. http://www.hamrick.com/vsm.html, 2002.

[17] ISO. ISO 12233:2000 Photography -- Electronic still picture cameras -- Resolution measurements.

[18] J. Janesick. *Scientific Charge Coupled Devices*, volume PM83. SPIE PRESS, 2001.

[19] K5600 Lighting. http://www.k5600.com, 2002.

[20] Kodak. http://www.kodak.com, 2001.

[21] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, and Duane Fulk. The digital michelangelo project: 3D scanning of large statues. In Kurt Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, pages 131–144. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2000.

[22] K Martinez. High resolution digital imaging of paintings: The vasari project. In *Microcomputers for Information Management*, pages 277–283, 1991.

[23] National Semiconductor. Lm9833 48-bit color, 1200dpi usb image scanner. http://www.national.com, 2001.

[24] R. Provin. Tri-color with ccd imagers. http://voltaire.csun.edu/tc.html, 1996.

[25] Brent Kecskemety Thomas Carducci, Antonio Ciccarelli. Ultra-high resolution 14,400 pixel trilinear color image sensor. http://www.kodak.com, 2001.

[26] D. Wallner. Building ICC Profiles – the Mechanics and Engineering. available at http://www.color.org/iccprofiles.html, 2000.

[27] Matthias Wandel. Building a megapixel digital camera from a flatbed scanner. http://www.sentex.net/ mwandel/tech/scanner.html, 2000.

[28] D. Williams and P. Burns. Diagnostics for digital capture using MTF. In *Image Processing, Image Quality, Image Capture Systems Conference (PICS)*, pages 227–232. Society for Imaging Science and Technology, April 2001.

[29] J. Yang. A light field camera for image based rendering. Master's thesis, MIT, 2000.

[30] J. Yang, C. Lee, A. Isaksen, and L. McMillan. A low-cost, portable light field capture device. In *Siggraph Conference Abstracts and Applications*, page 224, 2000.