2014

ROBOMIRROR: A SIMULATED MIRROR DISPLAY WITH A ROBOTIC CAMERA

Yuqi Zhang
University of Kentucky, yuqizhang2009@hotmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Yuqi Zhang, Student

Dr. Sen-ching Samson Cheung, Major Professor

Dr. Cai-Cheng Lu, Director of Graduate Studies

RoboMirror: a Simulated Mirror Display with a Robotic Camera

THESIS

A thesis submitted in partial
fulfillment of the requirements for the
degree of Master of Science in
Electrical Engineering in the College
of Engineering at the University of
Kentucky

By
Yuqi Zhang
Lexington, Kentucky

Director: Sen-ching Samson Cheung, Ph.D., Professor of Electrical and Computer
Engineering
Lexington, Kentucky

2014

ABSTRACT OF THESIS

RoboMirror: a Simulated Mirror Display with a Robotic Camera

Simulated mirror displays have a promising prospect in applications, due to its capability for virtual visualization. In most existing mirror displays, cameras are placed on top of the displays and unable to capture the person in front of the display at the highest possible resolution. The lack of a direct frontal capture of the subject's face and the geometric error introduced by image warping techniques make realistic mirror image rendering a challenging problem. The objective of this thesis is to explore the use of a robotic camera in tracking the face of the subject in front of the display to obtain a high-quality image capture. Our system uses a Bislide system to control a camera for face capture, while using a separate color-depth camera for accurate face tracking. We construct an optical device in which a one-way mirror is used so that the robotic camera behind can capture the subject while the rendered images can be displayed by reflecting off the mirror from an overhead projector. A key challenge of the proposed system is the reduction of light due to the one-way mirror. The optimal 2D Wiener filter is selected to enhance the low contrast images captured by the camera.

KEYWORDS: Face Tracking, See-through Display, Image Reshaping, Contrast Enhancement, 2D Wiener Filter.

Author's signature: _____ Yuqi Zhang

Date: _____ June 25, 2014

RoboMirror: a Simulated Mirror Display with a Robotic Camera

By

Yuqi Zhang

Sen-ching Samson Cheung, Ph.D.
Director of Thesis

CaiCheng Lu, Ph.D.
Director of Graduate Studies

June 25, 2014
Date

ACKNOWLEDGEMENTS

Contents

List of Figures

Chapter 1

Introduction

Mirrors are very useful objects in our daily lives. The capability of simulating a mirror on a computer display, augmented with virtual scenes and objects, opens the door to many interesting and useful applications, such as fashion design, cosmetic surgery, and behavior modeling. Realistic simulation of a mirror is challenging as it requires accurate viewpoint tracking and rendering, wide-angle viewing of the environment, as well as real-time performance to provide immediate visual feedback. The goal of this proposal is to develop a small-scale mirror simulation system using a see-through display and a robotic camera.

1.1   Mirror Display with Robotic Camera

A mirror display is a joint camera-display system that renders images of the person in front of the system as if she/he is looking at a real mirror. A camera is used to capture the environment from the viewer's perspective and project the output on a display in front of the viewer. Different content should be rendered on the display depending on the viewer's perspective. To this regard, the naive setup of having a video camera on top of the monitor showing the output of the camera on the display, as shown in Fig. 1, is clearly insufficient - the viewpoint of the camera is fixed while the mirror image depends on the position of the viewer.

In this research, a mirror display with a robotic camera is proposed in which a camera is placed behind the screen, just as the ideal camera shown in Figure 1. The camera is endowed with a robotic moving system which enables the camera to move appropriately according to the user's perspective.

The main challenge of mirror display is to render the perspective dependent on the display. The camera-display system must be able to capture the 3D details of the world, track the time-varying viewpoint of the user, render the new view based on the position of the a mirror display, and possibly add new visual content that is compatible with the scene geometry. Moreover, it must be able to accomplish all these tasks in real-time and with extremely high fidelity.
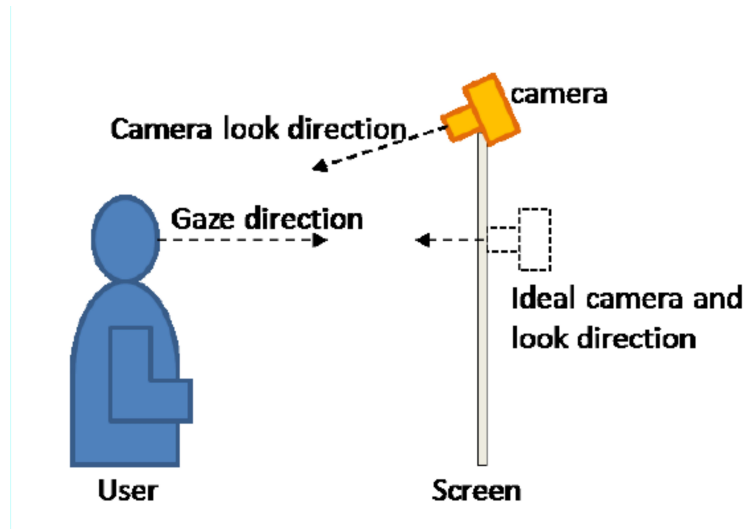
Figure 1.1: Naive setup of the mirror display system [38].

1.2 Mirror Display with Robotic Camera Applications

The mirror display is the foundation of a virtual mirror which depicts virtual scenes alongside with the mirror image to render a desired visual feedback to the user.

Mirror displays/virtual mirrors have been found applications in fashion. We can try on new clothes or shoes without making a trip to the department stores. A real time system is proposed for the visualization of customized sports shoes [12]. Similar systems have been developed to generate virtual mirror images for clothing [11], [31]. As the target objects in these system are already known, they usually have a pre-computed model from either 3D models or collections of image data. The viewpoint dependency is usually neglected in these systems due to its specific commercial intention.

The use of simulated mirror images go beyond fashion applications. One of the most well-known example is the use of the "mirror box" in helping patients suffering from chronic neurological disorders such as phantom pain, hemispheres from stroke, and other complex regional pain syndrome [26]. A mirror box is a simple top-open box with an opening on the side through which the patient can put his/her arm inside. There is mirror in the middle facing the inserted arm and the mirror image creates an illusion that of the presence of the opposite arm. It is this illusion that creates a visual feedback to the brain that alleviates the phantom pain caused by the amputated arm. While this form of mirror visual feedback is still not fully understood, it has already helped countless number of patients and has significant impact on the understanding of the elasticity of adult human brain in rewiring itself [25]. While a real mirror can only create the illusion of symmetric motion of the two hands, an augmented-reality mirror can synthesize asymmetric motion which might better reflect the state of mind of the patient.

Beyond a simple mirror, videos with self imageries and particular target behavior can be rendered and used by therapists to help patients in learning new behaviors or extinguishing bad ones. The approach of Video Self-Modeling (VSM) is a form of behavioral intervention where an individual learns from videos of the "self" engaged in target behavior [10]. The

feed-forward VSM focuses on teaching new skills by showing novel skills that have never been observed but are still within the reach of the learner.

The VSM has been found suitable for Autism treatment. Autistic children have significant difficulty in socially interacting with others. One theory to explain such a poor social interaction is the indifference of human faces. Recent fMRI studies have shown low activity in the brains of autistic children compared with typically-developed peers when viewing pictures of human faces [33]. The same study, however, shows that these children maintain high level of brain activity as viewing images of themselves. While a full neurological explanation of self recognition is still under investigation, such affinity to self images can be capitalized for VSM to help autistic children to learn from watching themselves.

Evidence shows that the feed-forward approach delivers a more dramatic learning affect than traditional therapist-driven behavioral intervention [7]. Besides Autism, the effectiveness of VSM has been demonstrated for many different types of disabilities and behavioral problems ranging from stuttering to inappropriate social behaviors [15].

Currently, effective VSM relies on tediously captured and heavily edited videos which lack interactivity and visual feedback. Using a mirror display augmented with virtual images, flexible visual contents and instant feedbacks to the patient can be provided which is far beyond the rigid video editing tools available today.

1.3   Contribution of the Thesis

In this thesis, we design a mirror display system consisting of a see-through screen and a robotic camera mounted on a motion system. By tracking the 3D coordinates of the viewer's perspective using Microsoft Kinect, the camera moves according to the obtained viewpoint such that the "mirror"-like image is rendered on the display. Due to the low contrast of the obtained image on the display, three image enhancement algorithms have been studied to improve the contrast of the image. Our contributions include the followings: 1) a mirror-display is designed capable of rendering the mirror-like image on the display by tracking the viewer's perspective in real-time; 2) two histogram equalization algorithms and an optimal 2D wiener filter have been applied to enhance the rendered image. The effectiveness of our designed mirror display system is demonstrated by our experimental results.

1.4   Organization of the Thesis

The rest of this thesis is organized as follows: in Chapter 2, we review existing techniques for mirror display system. Chapter 3 details the design of the proposed mirror display system. The equalization algorithms are presented, analyzed and compared with each other in Chapter 4. Experimental results are shown in Chapter 5, followed by conclusions and future work in Chapter 6.

Chapter 2

Related Work

In this chapter we discuss existing techniques for virtual mirror displays. Based on the applications, these techniques can be divided into two categories. The first category includes virtual mirror displays with no virtual objects rendered in the display. In this methodology, the main challenge is to adjust the perspective view, i.e. the gaze direction, light intensity and to avoid backscatter problems. They are also referred to as see-through screens which are discussed in the first section of this chapter. The second category includes advanced virtual mirror displays that can render virtual objects in the display, as discussed in the second section of this chapter.

2.1 See-through Screen

See-through screens are not a new concept in computer vision. They have attracted research studies for decades. A camera is normally rigidly mounted behind the translucent screen to capture the front view of the user. A projector is utilized to render images on the front side of the screen to the user. The screen has a mechanism to show images from the projector while allowing the camera to take images of the user's front view. Although the rendered images cannot be adjusted accordingly to the variations of the user's gaze direction due to the fixed camera, it has been extensively used in teleconferencing where eye contact and image quality are the essential concerns. In those applications, the light intensity of the rendered images and the backscatter problems are the main issues. There are various technologies, as discussed in this section, that have been proposed to tackle these issues.

Half silvered mirrors were intensively used in early studies, such as Teleprompter [23] and Gazecam [4]. The problem is that the half silvered mirror needs to be angled at approximately 45 degrees from the display. Such a design leads to the whole system being inconveniently large. Furthermore, the inherent stray reflections resulted from the mirror are also distracting. To improve the half silvered mirror implementations, a prototype of ClearBoard is implemented based on a clever "Drafter-Mirror" architecture [17]. The display surface was placed 45 degrees with respect to the ground. Polarization films are installed on the screen to avoid the displayed image being captured by the camera. However, polarization films cut down lights significantly, leading to poor display quality.

Another popular scheme for the see-through screen is through switchable liquid crystal diffusers (SLCD), based on which both the system proposed by Shiwa [29] and the blue-c system [37] were developed. In these two systems, the screen is composed of a glass panel containing liquid crystal layers which can switch between the transparent state and a whitish opaque state based on the voltage applied on the layers. In the transparent state, the camera behind the screen can take images of the front view of the user. During the opaque state, the projector can render images of the desired user's views. The main challenge here is the synchronization of the two shutters for the projector and the camera. Yet, a known limitation of SLCD is the speed of the states switching that limits the update rate of the rendering, causing flickering images. In addition, the rendered images of SLCD tend to have poor contrast under ambient lighting conditions, making it unsuitable for many applications.

As a circumvention for the synchronization requirement of the projector to the camera, holographic projection screen (DNP HoloScreen) have received a lot attention [36] [21] [32]. The key feature of DNP HoloScreen is to diffuse only incident light from pre-specified angles, and allow the light to pass through otherwise. This screen does not require special synchronized cameras and projectors, thus offering greater freedom for designers. However, the disadvantage of this see-through screen is well-known, which is severe backscatters. Infrared cameras were used in the TouchLight system to circumvent this backscatter problem [36]. However, this solution is not suitable for mirror displays. HoloPort [21] proposed to synchronize the projector and the camera to counter the backscatter problem. Yet, it comes to a paradox of avoiding one problem (backscatter) by picking up another problem (synchronization). Although the backscatter has been successfully removed by using wavelength multiplexing in the ConnectBorad system [32], additional processing is required to pre-distort the projected image and color-correct the captured image in order to render qualified images.

Besides backscatters, HoloScreen only diffuses the incident light from a pre-specified angel, and a crisp image can also be seen on the ceiling or the floor of the room. Therefore, there are many angles for which the user is looking right into the beam, which can be quite painful. Furthermore, the display image quality of HoloScreen is similar to SLCD, which leaves room for improvement.

2.2   Virtual Mirror System

The applications of the virtual mirror differ in various fashions. Simple virtual mirror systems only alter the real image by augmenting some graphical expression in the user's face. Some provide virtual experience by rendering a virtual object, like a cloth or a pair of shoes, based on certain 3D models. While others create the whole background based on the user's perspective.

Several research groups have developed virtual mirror prototypes [5], [20], and [9]. Despite the minor distinctions in some aspects, all of them implement simple appearance modifications with limited viewpoints. Darrel *et al.* described a virtual mirror interface that applied different graphical effects on the face of the user in front [5]. However, this work does not consider the influence of the user's gaze direction on a larger virtual mirror.

A non-linear Kalman-based estimator is proposed for the 3D head pose and facial features locations based on the user's 3D face model [20]. Virtual facial modifications are performed by user-driven 3D-aware 2D warping of the image sequence. Although, the feature-based real-time 3D tracking algorithm is robust to moderate head movements, occlusion and facial animation, it cannot render the background of the image according to the user's viewpoint in a virtual mirror. The same issue also arises in which the same authors use a similar approach to implement a virtual mirror experience [19]. Due to the utilization of a monocular webcam, the user's face almost occupies the whole screen, and no significant changes of the background of the user/viewer can be conveniently captured. Thereby, for realistic rendering a virtual mirror, the viewer's gaze direction is limited.

A hand-held mirror simulation device is proposed by Francois and Kang [13]. The proposed system is built with a LCD screen manipulated by the user, a single camera fixed on the screen, and a tracking device. The continuous input video stream and tracker data is used to synthesize, in real-time, a continuous video stream displayed on the LCD screen. The synthesized video stream is a close approximation of what the user would see on the screen surface if it were a real mirror. Although they consider the impact of viewpoint change, their system uses a simplified model for rendering by assuming that the 3D world is parallel to the mirror. Such an assumption significantly limits the range of possible perspective change and is not suitable for rendering large mirrors.

A recent paper by Straka *et al.* described a system that allowed the user to view himself or herself anywhere within a 360° field of view [30]. Multiple RGB cameras were mounted around the user to capture the data for 3D reconstruction. However, in their proposed system only the person was rendered in the mirror without any background, which negated the essential characteristic of a mirror where the viewer could observe different scenes as the view point changed.

The capability of simulating a mirror on a computer display, augmented with virtual scenes and objects, opens the door to many interesting and useful applications from fashion design to medical interventions. Realistic simulation of a mirror is challenging as it requires accurate viewpoint tracking and rendering, wide-angle viewing of the environment, as well as real-time performance to provide immediate visual feedback. A virtual mirror rendering system is proposed in [28] using a network of commodity structured-light RGB-D cameras. The depth information provided by the RGB-D cameras is used to track the viewpoint and render the scene from different perspective. A novel depth denoising and completion algorithm is proposed in which the noise removal and interpolation procedures are guided by the foreground/background label at each pixel. The foreground/background label is estimated using a probabilistic graphical model that takes into account color, depth, background modeling, depth noise modeling, and spatial constraints. The wide viewing angle of the mirror system is realized by combining the dynamic scene, captured by the static camera network, with a 3D background model created off-line, using a color-depth sequence captured by a movable RGB-D camera. To ensure a real-time response, a scalable client-and-server architecture is used with the 3D point cloud processing, the viewpoint estimate, and the mirror image rendering are all done on the client side. The mirror image and the viewpoint estimate are then sent to the server for final mirror view synthesis and viewpoint refinement. Experimental results are presented to demonstrate the accuracy and effectiveness of each component and the entire system.

Virtual mirrors have also been used in some very specific applications. A real-time system

is proposed for the visualization of customized sports shoes [12]. A large display is used instead of a real mirror that shows the mirrored input of a camera capturing the legs and shoes of a person. 3D tracking of both feet and exchanging the real shoes by computer graphics models give the impression of actually wearing the virtual shoes. The 3D motion tracker proposed exploits mainly silhouetted information to achieve robust estimation for shoes from a single camera view. Hilsmann *et al.* proposed a dress-fitting system where users can virtually dress themselves in different clothes and see how they would look using the virtual mirror [14].

Similar systems have also been developed to generate virtual mirror images for fashion [11] [31] [34]. Since the target objects in these systems are already known, they usually have a pre-computed model from either an existing 3D model or a collection of large training data. Therefore, the on-line computation can focus on rendering the correct texture on the generated image. In addition, the view point change is usually neglected in these systems due to its particular commercial intention.

Chapter 3

Mirror Display System

Simulated mirror display is a camera-display system that renders a mirror image for an individual standing in front. The ability to capture and manipulate mirror image is important for many medical applications such as rehabilitation and behavioral therapy. The camera is expected to capture the three-dimension (3D) environment such that the view-dependent mirror image can be rendered. However, most existing simulated mirror displays fail to capture the front image of the subject because of their inherent limits. In this chapter, we present a design of a simulated mirror display system. A one-way mirror is used to reflect images from a projector in this research. Simultaneously, a robotic camera, which is placed in the back of the mirror, tracks the subject's frontal image. The movement of the camera is controlled by a two-axis T-slot table. The projector holding system is developed for the simulated mirror display. The proposed construction is expected to meet the requirements of a stable installation and be cost-effective. Moreover, because of the particular design of the simulated mirror display a person standing in front of the mirror would see a deformed image, a trapezoid image instead of a rectangular image. To this end, the perspective transformation is applied to reshape the projected image so that a rectified image can be presented on the mirror.

3.1 The Configuration of the Designed System and Its Principle

A key requirement for a virtual mirror system is to provide a high-quality rendering of self-image. This is particularly important to the facial area on which the subject is likely to focus. By using a robotics camera that follows the user's face, the robotic virtual mirror projector aims to create a virtual mirror view with a high-quality rendering of facial details.

In the current design of the virtual mirror display as shown in Fig. 3.1, a camera is set behind a one-way mirror which is positioned 45 degrees above the surface of the table. The camera is mounted on a Bislide system driven by two step motors which enables the camera to move along a two dimensional plane (right to left, up and down). The robotic camera aims to track the viewpoint of a viewer sitting or standing in front the mirror. To accomplish this tracking, the Microsoft Kinect system is mounted on top of the mirror box and used to provide the 3D coordinates of the viewpoint, which is the mid-point between
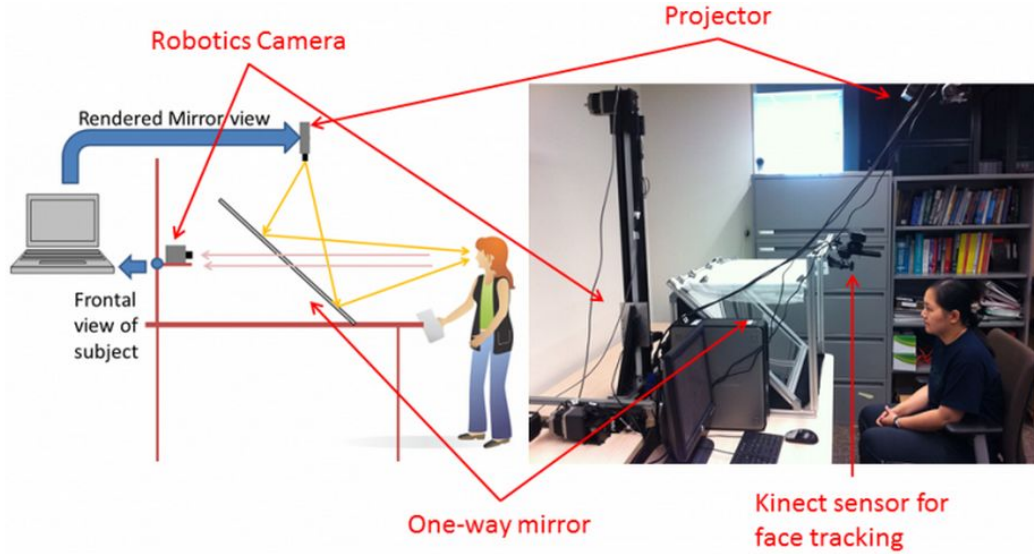
Figure 3.1: Illustration of the virtual mirror display with robotic camera [8]

the two eyes of the viewer. The viewpoints are forwarded to the computer, and the Bislide system is driven accordingly to enable the camera to track the viewpoint. The obtained scene from the camera is transmitted to the projector hanging above the mirror box. The rendered image is then projected on the screen on the top of the mirror box and imaged onto the mirror. In this way, the viewer in front of the mirror can see his/her images right on the mirror display, simulating a real mirror.

3.2 System Requirements and Mechanical Design

The movement of the camera in three-dimension (3D) is driven by the motors installed on the table. The movement accuracy of the motors is one of the factors to determine the movement accuracy of the camera. The motors using in most existing three-axis table can achieve the error of movement within 35 um [2]. Since the accuracy has met our requirement, the selection of the motor seems to be a minor issue for the movement accuracy of the camera.

Figure 3.2 shows one design of the three-axis table. The design only uses one track to support the movement in each direction. Such design would inevitably diminish the stability of movement of the camera even the system is properly and precisely installed.

Figure 3.3 shows another design of the three-axis table. Also, this design only uses one track to support the movement in each direction. Without steady and firm support of the tracks installed with the motor, the stability of the movement of the motor is questionable. However, the compactness is one merit of this design.

Figure 3.4 (a) and (b) show the improved designs of the three-axis table. The movement in X direction is supported by two tracks which makes the movement more stable. Using a relative heavy base for the three-axis motors in Figure 3.4 (a) might make it more stable. In terms of movement of stability and dimension compactness, the design in Figure 3.4 (a)

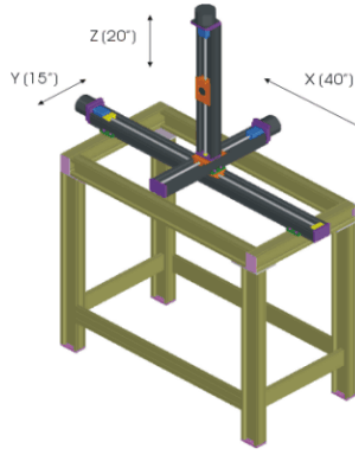Figure 3.2: Design No.1 [3]
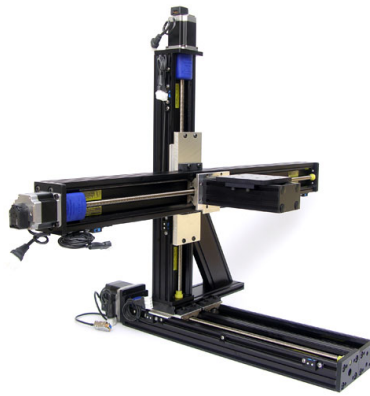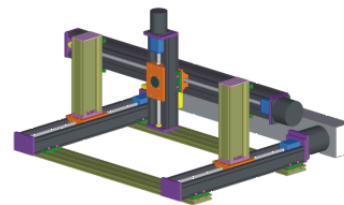

Figure 3.3: Design No.2 [1]

(a)   (b)

Figure 3.4: Design No.3 [3] [1]

is the most suitable selection of three-axis Bislide system for our application in simulated

mirror display.

In order to produce a relatively compact design, we decide to use a two-axis (x and y) Bislide system and simulate the z-direction movement based on image processing techniques. A two-axis (x and y) Bislide system is selected for our application. The reasons are the following: 1. Stability of the design; 2. Two-axis Bislide system uses space more effectively, because a two-axis system is much smaller than three-axis system; 3. Two-axis Bislide system is much cheaper than three-axis polyurethane belt system. The whole three-axis polyurethane belt system is about $19,000, while two-axis Bislide system is about $6000.

In the simulated mirror display system, the projector is required to be positioned downward pointing towards the mirror at a 45 degree angle (as shown in Figure 3.1). The lighter the projector, the more stable it is and the easier it is to install. In addition, the image generated by the projector should be bright enough to be used in environment with typical ambient light. According to the requirements for the projector mentioned above, Optoma's ML500 LED projector is selected. This projector is only 2.5 pounds, which satisfies the first requirement. It has 500 ANSI Lumens brightness, which satisfies the second requirement.

The projector holding system is shown in Figure 3.5. It can be divided into three parts. The first part is the horizontal rail and the ceiling beam connected by two clamps as shown in Figure 3.6.

Figure 3.5: Projector Holding System

The second part is the vertical rail and the horizontal rail installed by two brackets as shown in Figure 3.7.

The third part is a double flange bearing profile connecting the projector mount with the vertical rail as shown in Figure 3.8. The projector is fixed on the mount. The distance
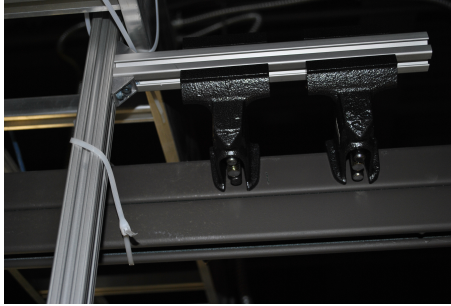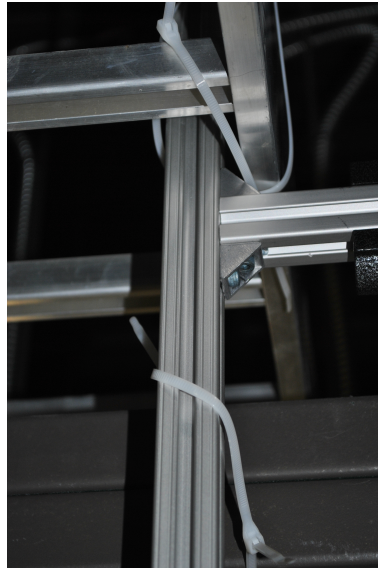
Figure 3.6: The First Part

Figure 3.7: The Second Part

between the projector and the mirror can be adjusted through the double flange bearing profile up and down along the vertical rail.

Figure 3.8: The Third Part

The entire projector system and the mirror are shown in Figure 3.9.

Figure 3.9: The Entire Projector System and The Mirror

Chapter 4

Mirror Image Enhancement

The camera in the proposed virtual mirror display system is mounted in a Bislide system behind the one-way mirror. The scene in front of the camera is captured, transmitted to the projector above the viewer, and eventually shown on the display in front of the viewer in real-time.

Although the mirror display system has been found to work satisfactorily, the quality of the images requires improvement. First, due to the misalignment of the projector and the projection's screen, there is a spatial distortion of the obtained image so that a rectangular image has been distorted to a trapezoid image. Second, the contrast of the image is comparatively low since the mirror display is a one-way mirror which passes though part of the light radiation, and the projector screen also blocks a portion of the light intensity.

In this chapter, we focus on image processing techniques to rectify these problems. In Section 4.1, we describe the transformation to reshape the distorted image. In Section 4.2, we apply contrast enhancement to improve the quality of the displayed image. In particular, three enhancement algorithms are introduced. The corresponding results are analyzed in Chapter 5.

4.1   Image Reshaping

The output of the camera is a "raw" digital image that consists of an array of digital count values with each value representing the color of a pixel in the image. The camera used in our display system is Point Grey Firefly MV (FMVU-03MTC-CS) . It provides 24 bit RGB color images with a capture rate at 30 frame per second (fps). In the simulated mirror display system, the projector is pointing downward towards the one way mirror. Since the projector cannot directly project image to human's eyes, it has to project image to a screen first, and then the image on the screen can be captured by human's eyes. In the simulated mirror display system, the screen is placed horizontally (as show in Figure 3.1). The rectangular image would be deformed to a narrow trapezoid if the projector directly projects images to the screen vertically. Thus, the projector should project images to the screen with an angle less than 90 degrees, so that the deformation of the projected image on the screen will be reduced. In order to get a rectangular shape on the screen, the original image is reshaped using a perspective transform to a trapezoid shape. The trapezoid shape

will be projected to the screen by the projector. A rectangle image, which is the same shape as the original image, will be displayed on the screen.

There are generally two geometric transforms used to reshape images. The first one is affine transform, which can convert rectangles to parallelograms. It can squash, rotate and/or scale the shape, but must keep the sides parallel. So an affine transform cannot convert a rectangle to a trapezoid. The other geometric transform is perspective transform, which offers more flexibility. A perspective transform can turn a rectangle into a trapezoid. Since parallelograms are also trapezoids, affine transformations are a subset of perspective transformation [6]. Figure 4.1 shows examples of various affine and perspective transformations. Since the rectangle shape will be deformed to a trapezoid shape when it is projected to the screen, perspective transforms will be adopted here to reshape images.
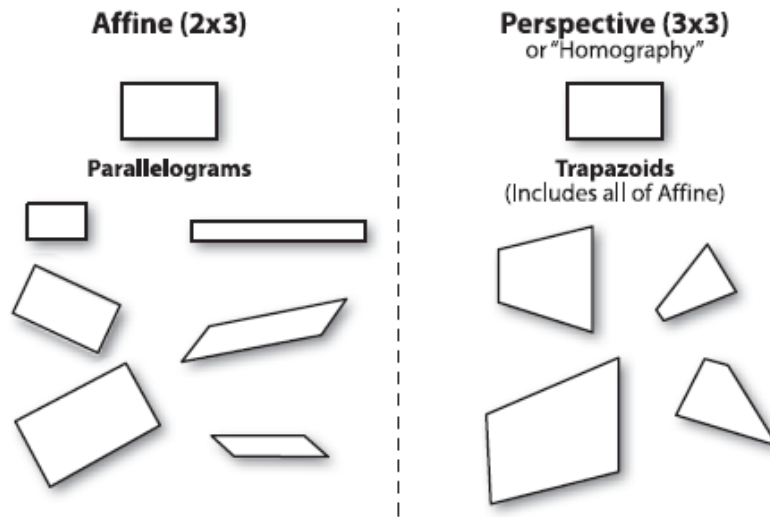

Figure 4.1: Affine and perspective transformations [6]

A perspective transformation is developed to reshape the projected image so that the image on the mirror has no distortion in shape. The transformation maps the points on original plane to the corresponding points on the second plane by a *plane to plane homography*, also known as a plane projective transformation. A homography is described by a 3x3 matrix H. Once this matrix is determined the back-projection of a point in second plane to a point on the original plane is straightforward.

Figure 4.2 shows the perspective transformation process. A point x on the second plane corresponds to point X in the original plane. Euclidean coordinates x-y and X-Y are used for the original and second planes, respectively.Under the projection corresponding points are related by :

$$X = Hx \qquad (4.1)$$

where H is a 3x3 homogeneous matrix, and "=" is equality up to scale [27] [22]. The points in the two planes are represented by homogeneous 3-vector as X=$(X, Y, 1)^T$ and x=$(x, y, 1)^T$.
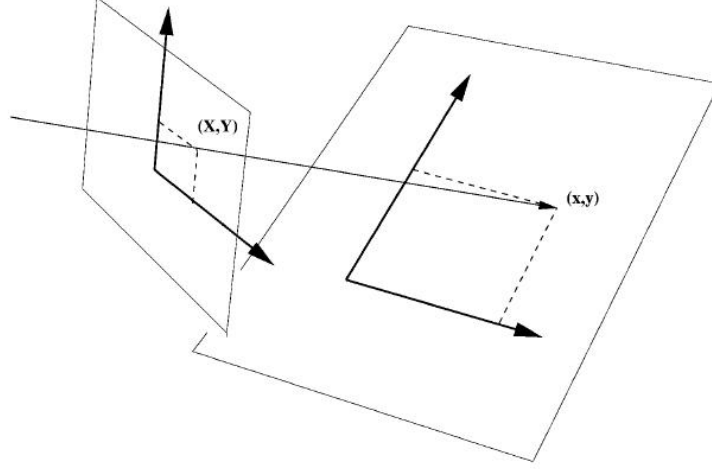
Figure 4.2: Diagram of perspective transformation

The scale of the matrix does not affect the equation, so only the eight degrees of freedom corresponding to the ratio of the matrix elements are significant. The corresponding relation can be detailed as:

$$\begin{bmatrix} XW \\ YW \\ W \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4.2}$$

then the distance coordinate W is:

$$W = gx + hy + 1 \tag{4.3}$$

divide W in both sides of equation 4.2, the X and Y can be rewritten:

$$\begin{aligned} X &= ax + by + c - 0d + 0e + 0f - Xxg - Xyh \\ Y &= 0a + 0b + 0c + xd + yd + f - Yxg - Yyh \end{aligned} \tag{4.4}$$

Given n pairs of corresponding points in the two planes, then the Equation 4.4 can be presented as:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -X_nx_n & -X_ny_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -Y_nx_n & -Y_ny_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ \vdots \\ X_n \\ Y_n \end{bmatrix} \tag{4.5}$$

According to Equation 4.5, 4 pairs of points are needed in order to obtain analytical solution of the transformation matrix.

4.2 Image Contrast Enhancement

Contrast enhancements improve the perceptibility of objects in the scene by enhancing the brightness difference between objects and their backgrounds. Contrast enhancements are typically performed as a contrast stretch followed by a tonal enhancement, although these could both be performed in one step. A contrast stretch improves the brightness differences uniformly across the dynamic range of the image, whereas tonal enhancements improve the brightness differences in the shadow (dark), midtone (grays), or highlight (bright) regions at the expense of the brightness differences in the other regions.

4.2.1 Global Histogram Equalization (GHE)

Most contrast enhancement methods make use of the gray-level, created by counting the number of times each gray-level value occurs in the image, then dividing by the total number of pixels in the image to create a distribution of the percentage of each gray level in the image. The gray-level histogram describes the statistical distribution of the gray levels in the image but contains no spatial information about the image.

The image contrast enhancement is to find a monotonic transformation function τ to map the gray level p in the current dynamic range $[p_0, p_k]$ to a new gray level q in the new dynamic range $[q_0, q_k]$, such that the perceptibility of the objects in the image is enhanced. The transformation is shown in Eq. 4.6.

$$q = \tau(p) \tag{4.6}$$

The histogram equalization technique is intensively used for the image contrast enhancement, in which the gray level transformation function can be automatically found. The aim of histogram equalization is to render an image in which the distribution of each gray level is identical throughout the whole dynamic range (as shown in Fig. 4.3).

Figure 4.3: Illustration of a histogram equalization
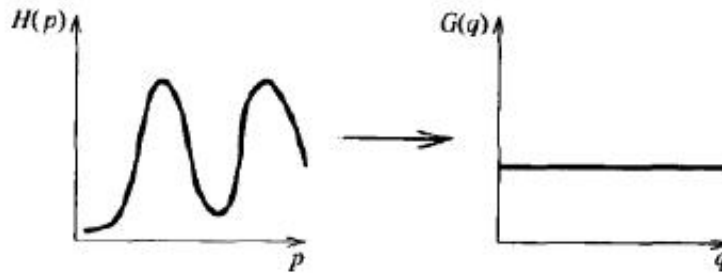
The input images histogram is denoted as $H(p)$, and the output images histogram is denoted as $G(q)$. The histogram can be considered as a discreet probability density function of the gray level in the whole image. Due the monotone feature of the transformation function means, the following equation can be held:

$$\sum_{i=0}^{k} G(q_i) = \sum_{i=0}^{k} H(p_i) \tag{4.7}$$

The accumulation of in Eq. 4.7 can be considered to calculate the discreet distribution function. Given the column and row number denoted as M and N, respectively, the equalized probability density function f corresponding to the histogram $G(q)$ is a constant, expressed in Eq. 4.8.

$$f = \frac{NM}{q_k - q_0} \tag{4.8}$$

Using Eq. 4.8, the accurate expression of the input histogram $H(p)$ can be thus obtained with an 'ideal' continuous probability density:

$$NM \int_{q_0}^{q} \frac{1}{q_k - q_0} ds = \frac{NM(q - q_0)}{q_k - q_0} = \int_{p_0}^{p} H(s)ds \tag{4.9}$$

Therefore, the desired transformation function for histogram equalization is obtained in Eq. 4.10, and its discreet approximation is shown in Eq. 4.11.

$$q = \tau(p) = \frac{q_k - q_0}{NM} \int_{i=p_0}^{p} H(s) + q_0 \tag{4.10}$$

$$q = \tau(p) = \frac{q_k - q_0}{NM} \sum_{i=p_0}^{p} H(s) + q_0 \tag{4.11}$$

The pseudo-code of global histogram equalization is summarized in Algorithm 4.2.1.

Algorithm 1 Global histogram equalization (GHE)

Require: A $M \ timesN$ image with G gray levels

  function HISTOGRAM EQUALIZATION(M, N, G, g_p)

    % g_p: the gray level for an arbitrarily pixel p%
    Build an array H with a length G, and initialize the array with 0.
    for each pixel p in the image **do**
      $H[g_p] = H[g_p] + 1$.
    end for

    % To build the histogram of this image, denoted as H_c%
    $H_c[0] = H[0]$
    for $i \leftarrow 1$ to $G - 1$ **do**
      $H_c[i] = H_c[i - 1] + H[i]$.
    end for
    % Find the gray level transformation function $T[i]$%
    $\tau[i] = round(\frac{G-1}{NM} H_c[i])$
    return A $N \times M$ image in which the gray level of each pixel $g_q = \tau[g_p]$
  end function

4.2.2 Adaptive Histogram Equalization (AHE)

The global histogram equalization algorithm applies the same transformation function $\tau(\cdot)$ derived from the image histogram to all the pixels. It works well if the distribution of the pixels' gray levels is comparatively homogeneous through the image. However, if the gray levels in some regions are orders of magnitude larger than those in the rest of the area in the image, the contrast in those local regions will not be sufficiently enhanced using GHE. Furthermore, GHE might not work well for images with a large range of contrast information, such as computed tomography (CT) image, and (MR) image.

An alternative method for contrast enhancement, adaptive histogram equalization, (AHE) has been developed by transforming each pixel with a transformation function derived from the local region ([18], [16]). In its simplest form, the gray level of each pixel is transformed based on the histogram of a rectangular surrounding the pixel, as shown in Fig. 4.4. Those rectangular are refereed as tiles.
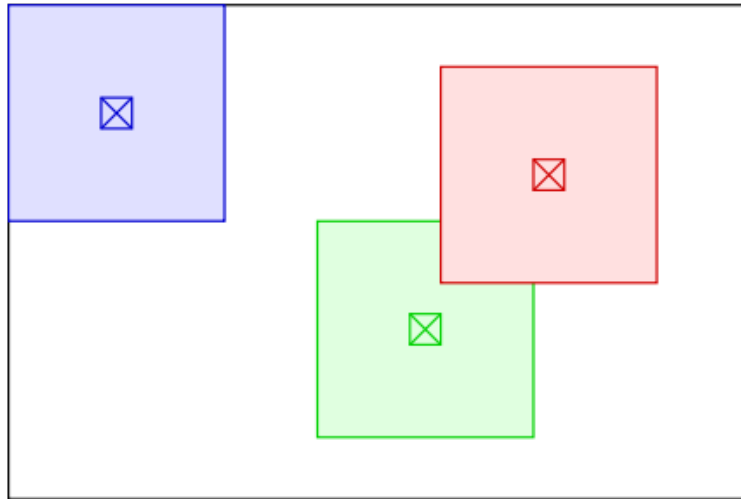

Figure 4.4: Tiles in a image

The procedure of the transformation function $\tau(\cdot)$ from the histogram can be the same as from the global histogram equalization in Algorithm 1, which is a flat histogram distribution. Besides, there are other options of the histogram distributions, such as Bell-shaped histogram and Curved histogram.

The AHE allows information in all intensity ranges of the image to be viewed simultaneously compared with the global histogram equalization. Studies have shown its promising usage in MRI and tomography [35].

However, there are two noticeable disadvantages associated with the AHE:

- It has a tendency to over-amplify noise in relatively homogeneous regions of an image. When an image region containing a pixel's neighborhood is fairly homogeneous, its histogram will be dramatically peaked, and the transformation function will map a narrow range of gray levels to the whole homogeneous regions of the resultant image [39]. This causes AHE to over-amplify small amount of noise in largely homogeneous region of the image.

- It is not computationally effective, since it requires the computation of a different neighborhood histogram and transformation function for each pixel in the image.

There are two AHE variants regarding the two disadvantages, respectively, referred to as Contrast limited AHE and Interpolated AHE [39].

Contrast limited AHE (CLAHE)

By limiting contrast enhancement, AHE's over amplification of the noise can be prevented. In basic AHE, contrast enhancement in the vicinity of a given gray level can be defined as the slope of the transformation function. Assuming the range of the input gray level is the same as the output image's gray level range, then a slope of 1 indicates no contrast enhancement, and higher slopes give increasingly higher enhancement. Thus, the limitation of the contrast enhancement can be achieved by restricting the slope of the transformation function. According to Eq. 4.11, the transformation function $\tau(\cdot)$ is proportional to the cumulative distribution function (CDF), $\sum_{i=p_0}^{p} H(s)$. Since the derivative of the cumulative distribution function is the histogram $H(s)$, the slope of $\tau(\cdot)$ at any input gray level, i.e. the contrast enhancement, is proportional to the height of the histogram at that gray level. We can express the slope of $\tau(\cdot)$ as

$$\frac{d\tau}{dp} = \frac{q_k - q_0}{NM} H(p) \tag{4.12}$$

Therefore, limiting contrast enhancement is equivalent to clipping the height of the histogram.

Based on the aforementioned mechanism, CLAHE limits the contrast enhancement by clipping the histogram at a predefined value before computing the CDF, as shown in Fig. 4.5. This limits the slope of the CDF and therefore limits the slope of the transformation function. The value at which the histogram is clipped, called the clip limit, depends on the normalization of the histogram and therefore on the size of the neighborhood region.

The part of the histogram that exceeds the clip limit is redistributed equally among all the histogram bins [24], as shown in Fig. 4.5. The redistribution will push some bins over the clip limit again (green region in the right figure), resulting in an effective clip limit larger than the prescribed limit, and the difference between the two limits depends on the image. If this is undesirable, the redistribution procedure can be repeated recursively until the excess is negligible.

Remarks:

1. The pixels near the image boundary have to be treated specially in CLAHE, since their tiles don't lie completely within the image. This can be solved by extending the image by mirroring pixel lines and columns with respect to the image boundary.

2. The dimension of the tile is a parameter of the CLAHE. It constitutes a characteristic length scale: contrast at smaller scales is enhanced, while contrast at larger scales is reduced.
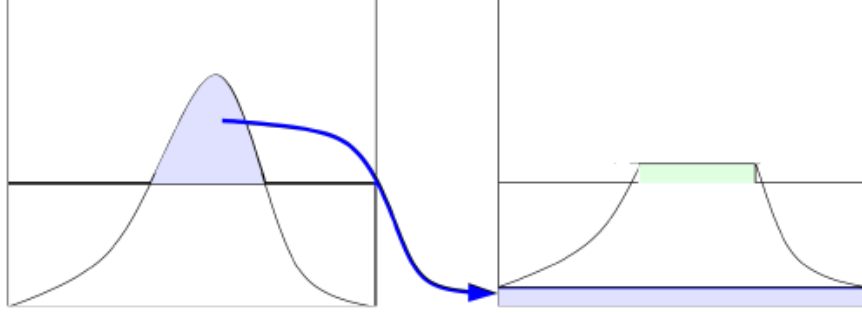
Figure 4.5: Distribution of clipped histogram

Interpolated AHE (IAHE)

The basic AHE require computation time $O(n^2(m + k))$ for an $n \times n$ image with a range of k gray level and an $m \times m$ tile region size. One of the simple alternative for the computationally inefficient AHE is to calculate the desired mapping only at the tile center pixel, and interpolate the gray levels of other pixels using neighborhood transformation functions. This can significantly reduce the computation complexity of the basic AHE.

Like the basic AHE, Interpolated AHE partitions input image into equally sized rectangular tiles as shown in the right part of Fig. 4.6.
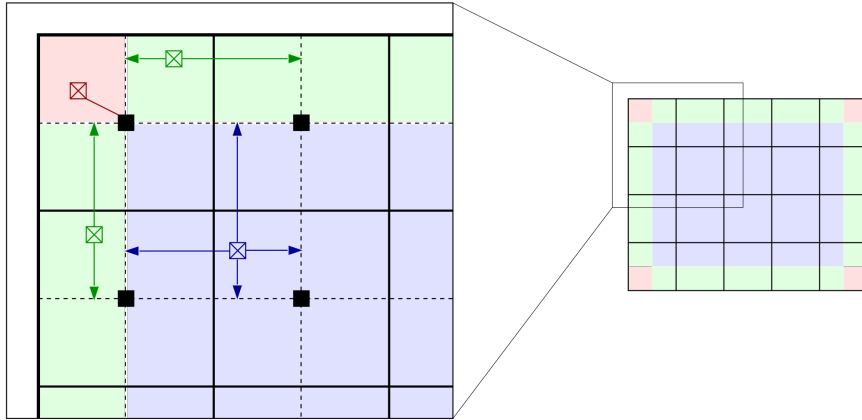

Figure 4.6: Interpolation of the basic AHE. The image divides into 64 tiles in 8 columns and 8 rows is a common choice.

A histogram, CDF and transformation function is then computed for each of the tiles. The transformation functions are appropriate for the tile center pixels, black squares in the left part of Fig. 4.6. All other pixels are transformed with up to four transformation functions of the tiles with center pixels closest to them, and are assigned interpolated values. Pixels in the bulk of the image (shaded blue) are bi-linearly interpolated. The interpolation coefficients reflect the location of pixels between the closest tile center pixels, so that the result is continuous as the pixel approaches a tile center.

Specifically, given an arbitrary pixel located at (x, y) inside the rectangular are cornered with the four tile center pixel in Fig. 4.6, its gray level is denoted as i. Denote τ_{ur} is the transformation function associated withe tile center pixel (x_u, y_r) to the upper right

of (x, y), ans similarly with subscripts lr, ll, and ul for the transformations functions and locations of the tile center pixels to the lower right, lower left, and upper left respectively of (x, y), then the interpolated AHE result at (x, y) is given by

$$\tau_{IAHE}(i) = a[b\tau_{ul}(i) + (1 - b)\tau_{ur}(i)] + (1 - a)[b\tau_{ll}(i) + (1 - b)\tau_{lr}(i)] \qquad (4.13)$$

where

$$a = \frac{y - y_l}{y_u - y_l} \qquad b = \frac{x - x_l}{x_r - x_l} \qquad (4.14)$$

Pixels close to the boundary (shaded green in Fig. 4.6) are linearly interpolated, which are given by

$$\tau_{IAHE}(i) = a\tau_{ul}(i) + (1 - a)\tau_{ll}(i) \qquad (4.15)$$

if the pixel are in the left boundary, or

$$\tau_{IAHE}(i) = b\tau_{ul}(i) + (1 - b)\tau_{ur}(i) \qquad (4.16)$$

if the pixel are in the upper boundary.

Pixels near corners (shaded red in Fig. 4.6) are transformed with the transformation function of the corner tile.

The interpolation allows a significant improvement in efficiency without compromising the quality of the result. This procedure reduces the number of transformation functions to be computed dramatically and only imposes the small additional cost of linear interpolation.

In this thesis, the CLAHE and IAHE are combined together, refereed as Contrast limited and interpolated AHE (CLIAHE), the pseudo-code of CLIAHE is summarized in Algorithm 4.2.2.

4.2.3 Optimal 2D Wiener Filter

The aforementioned two algorithms might not adaptively improve the contrast because they do not have the mechanism to evolve the performance of image processing. Hence an adaptive processing algorithm is proposed based a 2D Wiener filter.

Give pixel intensity (luminance only) $X(x, y)$ and $Y(x, y)$ as captured by camera with and without the one-way mirror, we model the relationship as follows:

$$Y(x, y) = f\left(\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C(m, n) \cdot X(x - m, y - n) \right) \qquad (4.17)$$

Physically, the convolution with the channel function $C(x, y)$ represents local light dispersion when passing through the glass slab of a certain thickness. The function $f(\cdot)$ represents the non-linear distortion on the signal brightness level.

Algorithm 2 Contrast limited and interpolated adaptive histogram equalization

---

Require: 1) gray levels of input image; 2) the tile dimension; 3) maximum and minimum gray level of input image; 4) clip limit value; 5) number of gray bins for histogram

   function CIAHE

1. Linearly map the gray level range of input image into range [1 number of gray bins]
   for each tile in the input image **do**
      2. Classify the gray levels present in the input image into a gray level histogram.
      3. Clip the histogram, count the number of excess pixels, and redistribute the excess pixels equally across the whole histogram.
      4. Calculate the equalized lookup table (mapping by cumulating the input histogram.
   end for

5. Calculate the new gray level assignments of pixels withing a sub-matrix of the image with a tile size, using a bilinear interpolation between four different mappings in order to eliminate boundary artifacts.
      return A $N \times M$ equalized image
   end function

---

(1) Inversion of $f(\cdot)$

By displaying a single gray-level image on a large display monitor as the light source for the camera, we first identify 256 levels that map to 256 images such that each image corresponds $X(x, y) = k$ for all x, y and $k = 0, 1, 2, ..., 255$. Since $X(x, y)$ is a constant image, eq 4.17 becomes

$$Y(x, y) = f(ck) \tag{4.18}$$

$c = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C(m, n)$. A d-degree polynomial fit can be used to compute the inverse mapping $g(\cdot)$:

$$g(Y(x, y)) \triangleq \frac{f^{-1}(Y(x, y))}{c} \tag{4.19}$$

(2) Channel Equalization

After estimating the function in (4.19), equation (4.17) can be rewritten as follows:

$$Y^{'}(x, y) \triangleq g(Y(x, y)) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C^{'}(m, n) \cdot X(x - m, y - n) \tag{4.20}$$

where $\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C^{'}(m, n) = 1$. To compensate for such effect, An approximation of the channel equalization function is performed using a $(2h+1) \times (2h+1)$ FIR filter $H^{*}(m, n)$

with unity DC gain, the approximation is formulated below:

$$H^*(m,n) \triangleq argmin_{H(m,n)} \sum_{xy}(X(x,y) - \sum_{m,n} H(m,n) \cdot g(Y(x-m,y-n)))^2 \quad (4.21)$$

subjected to

$$\sum_{m,n} H(m,n) = 1 \quad (4.22)$$

Taking the partial derivatives of the cost function with respect to $H(r,s)$ for $r = -h, ..., h$ and $s = -h, ..., h$ and setting them to zero, one can obtain:

$$\sum_{m,n} a(m,n,r,s)H^*(m,n) + \frac{\lambda}{2} = b(r,s) \quad (4.23)$$

where

$$a(m,n,r,s) \triangleq \sum_{x,y} g(Y(x-m,y-n)) \cdot g(Y(x-r,y-s)) \quad (4.24)$$

and

$$b(m,n) \triangleq -\sum_{x,y} X(x,h) \cdot g(Y(x-r,y-s)) \quad (4.25)$$

The linear system of equations composed by all the equations described in eq 4.22 and eq 4.23 uniquely determine the solution $H^*(m,n)$ and λ. To provide enough diversity to ensure non-singularity, we have used a number of different $X(x,y)$ and $Y(x,y)$ for scenes with different objects and illuminations in computing eq4.24 and eq4.25.

After the above off-line estimation process, the contrast enhancement process described below is applied in real-time during the mirror rendering:

$$X'(x,y) \triangleq \sum_{m,n} H^*(m,n) \cdot g(Y(x-m,y-n)) \quad (4.26)$$

Chapter 5

Experimental Results for the Mirror Display System

In this chapter, the experiment to validate the accuracy of the our virtual mirror display system are designed, which is detailed in Section 5.1. The experimental results are presented and discussed in Section 5.2.

5.1 Experiments Design

In order to validate the accuracy of the virtual mirror display system, the generated image on the display is compared with a real photo taken by a digital camera looking at a real mirror aligned with the display. The camera is in the same position as the asserted viewpoint that the mirror system uses for rendering. A 24×24 chessboard is set right beside the viewer such that both images can capture the chessboard. Since there is no virtual object or background generated in the generated image, the difference between the generated image and real photo comes from the inaccurate tracking of the viewpoint by the robotic camera. To measuring the rendering accuracy, 100 corner points in the chessboard are manually selected for analysis. To ensure that the exact pixel locations of the corners are applied, the normalized cross correlation (NCC) is computed over a 3×3 neighborhood around each corner point such that the corner position is refined with the maximum local NCC value.

5.2 Results Discussion

5.2.1 Image Reshape Results and Discussion

It has been established that after deformation of the mirror, a rectangular image is reshaped to a trapezoid image. The diagram of the reshape is shown in Figure 5.1. B is the transformation matrix which is inevitable because of the geometry relationship between the projector and mirror in the installation. However the original image can be reshaped to a certain shape such that the output image in the mirror is identical to the original one. The image reshape diagram is shown in Figure 5.2. B' is the inverse transformation matrix of B.
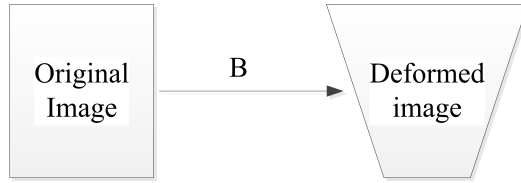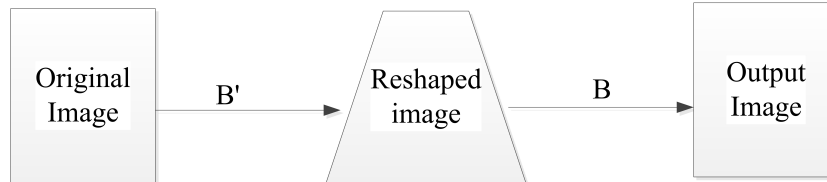
Figure 5.1: Deformation of the projector

Figure 5.2: Diagram of the reshape proposed in this report

In order to find transformation matrix B, the original and deformed images are required. Figure 5.3 shows the original image from the projector and Figure 5.4 is the deformed one shown on the screen.
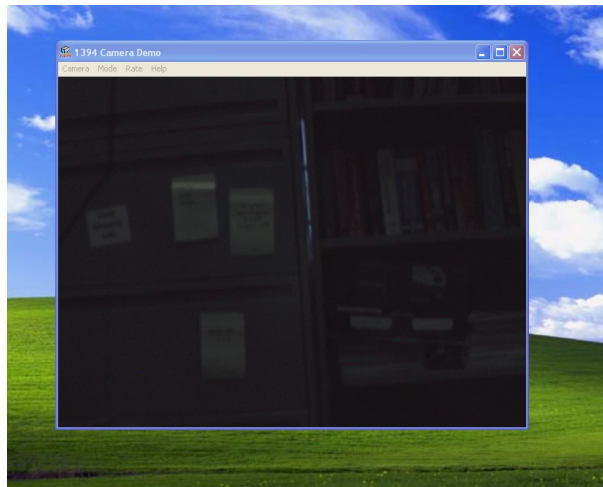

Figure 5.3: Original image

Finding four points in the original image and four corresponding points in the deformed image, the transformation matrix B is easily obtained. Using the program attached in the appendix, transformation matrix is

$$\mathbf{B} = \begin{bmatrix} 1.16664829 & 0.13163859 & -37.798832 \\ 0.031471398 & 1.2332906 & -10.524380 \\ 5.5596029e-6 & 0.00034266 & 1.0 \end{bmatrix} \tag{5.1}$$

With the warping using the matrix B, the resultant deformed image is shown in Figure 5.5. With the warping using the matrix B's inverse transformation, the resultant deformed image is shown in Figure 5.6. The result of output image after transformation matrix B
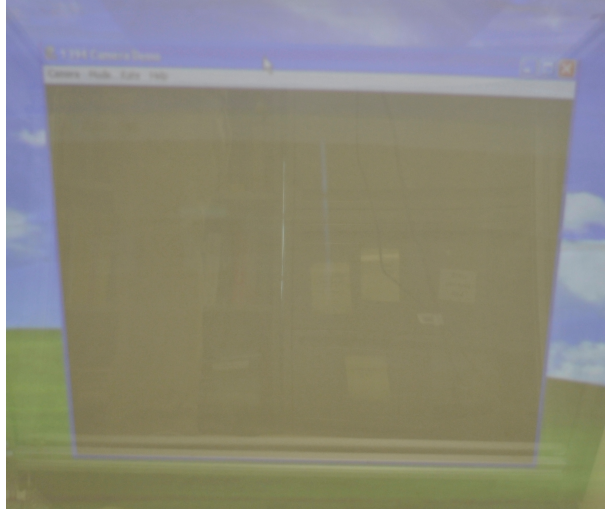
Figure 5.4: Deformed image

using Figure 5.6 as the reshaped image is shown Figure 5.7. Compared Figure 5.3 and Figure 5.7, the images are identical which verifies the effectiveness of the reshape process.
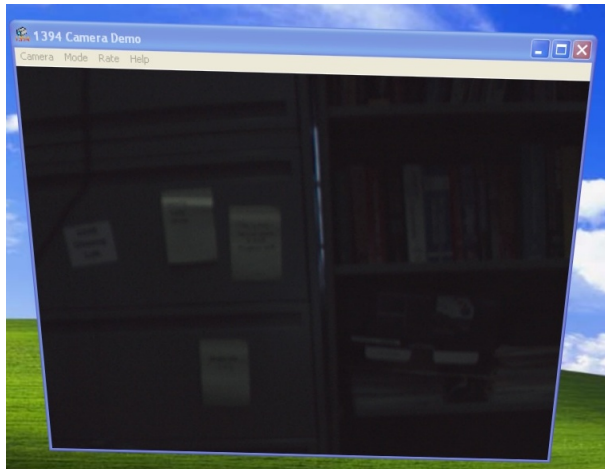

Figure 5.5: Resultant images of warping using matriax B

Figure 5.6 is the output image so that after the deformation of the mirror, image Figure 5.7 is obtained. There is no obvious distortion compared images Figure 5.7 and the original image Figure 5.3. The same procedure is applied to the Deformed image Figure 5.4, the resultant image is shown in Figure 5.8. Since the image shown in the mirror is dim, a color enhance algorithm is required to increase the contract of the color of the image. Using Laplace transfer to sharpen the image the resultant image is shown in Figure 5.9.

5.2.2   Image Contrast Enhancement Results and Discussion

Two images captured using the proposed virtual mirror display system, shown in Fig. 5.10a) and Fig. 5.11a) are used here to evaluate the performance of the two histogram equalization
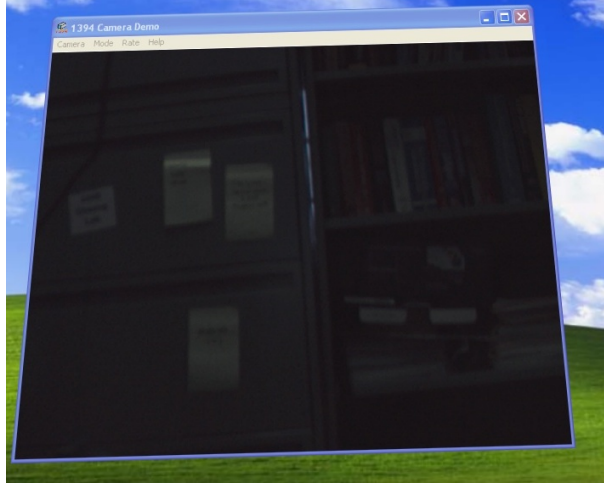
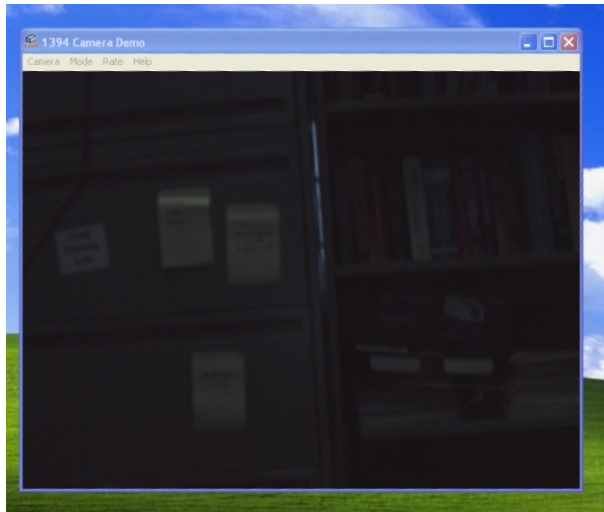Figure 5.6: Resultant images of warping using matriax B'

Figure 5.7: Resultant image using Figure 5.6 as the reshaped image

algorithms: GHE and CLIAHE. The GHE applied to the two images and the results are shown in Fig. 5.10b) and Fig. 5.11b). The corresponding resultant image accounting for the CLIAHE are shown in Fig. 5.10c) and Fig. 5.11c).

One can find from the two figures that the global histogram equalization does significantly increase the image contrast. However, it can also be found that the qualities of the two images are greatly degraded due the over-amplified noise associated with the GHE. On the other hand, using CLIAHE, the resultant images show the enhance contrast to some extent, i.e. compare with the raw input images, they are much brighter and clearer in details. The noise amplification is also appropriately reduced.

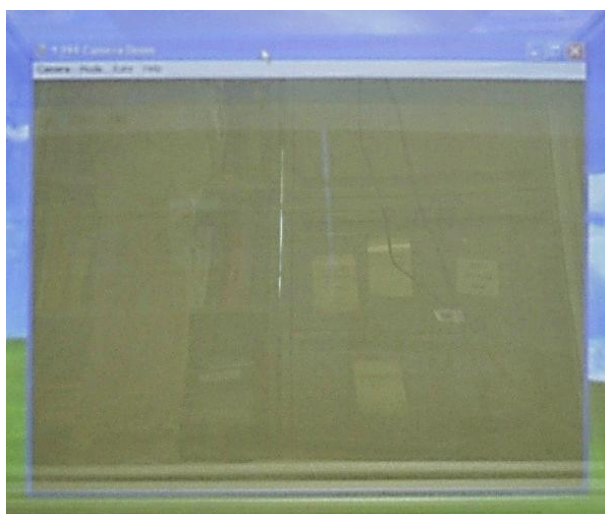Figure 5.8: Reshaped output image from the projector

Figure 5.9: Resultant image after sharpen process

Figure 5.10: Results of histogram equalization for image 1: a) the original image, b) resultant image using GHE, and c) resultant image using CLIAHE

Optimal 2D Wiener filter Enhancement Results

Figure 5.12 and Figure 5.13 show the images captured with and without the mirror respectively. The enhanced rendered image is shown in Figure 5.16. Figure. 5.14 and Figure.

Figure 5.11: Results of histogram equalization for image 2: a) the original image, b) resultant image using GHE, and c) resultant image using CLIAHE

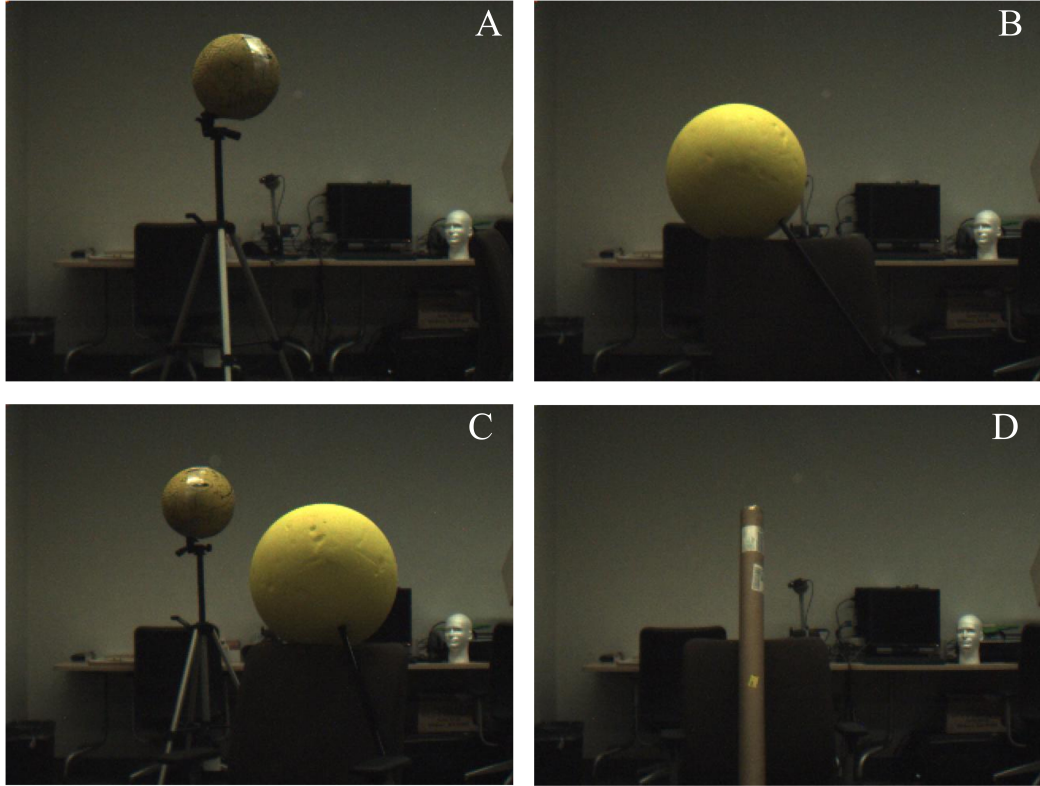5.15 are the results of GHE algorithm and CLIAHE algorithm, respectively.

Figure 5.12: Images captured with the one-way mirror

Relative (Γ_1) and absolute (Γ_2) effectiveness coefficients are defined in eq. 5.2 and eq. 5.3. They quantify the effectiveness of the enhancement algorithm, where ξ denotes image pixels captured without the one-way mirror, and ξ' is the corresponding pixels rendered by proposed enhancement.

$$\Gamma_1 \triangleq \frac{\sum_{\zeta=r,g,b} \frac{\sum_{n=1}^{480} \sum_{m=1}^{640} \frac{|\xi_\zeta(m,n) - \xi'_\zeta(m,n)|}{\xi_\zeta(m,n)}}{640 \times 480}}{3} \quad (5.2)$$
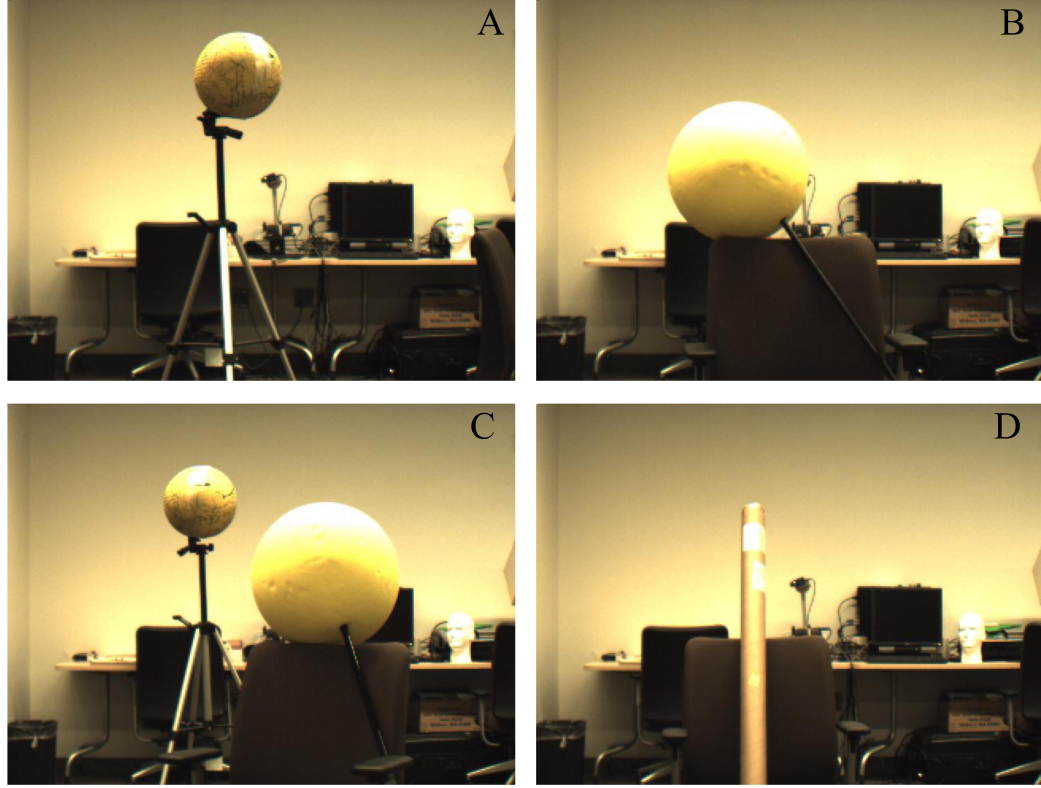
Figure 5.13: Images captured without the one-way mirror

$$\Gamma_2 \triangleq \frac{\sum_{\zeta=r,g,b} \frac{\sum_{n=1}^{480} \sum_{m=1}^{640} |\xi_\zeta(m,n) - \xi'_\zeta(m,n)|}{640 \times 480}}{3} \tag{5.3}$$

The effectiveness coefficient is a relative error measurement between the original image (without mirror) and rendered image. The images in Figure 5.13 and Figure 5.16 have a Γ_1 of 0.07, 0.07, 0.06, 0.05, respectively, giving an error of 6% of the original gray level. The results for Γ_2 are 6, 5, 5, 5, respectively, which indicate an absolute error of about 5% in gray level. The results are better than those using CLIAHE. Therefore, the optimal 2D wiener filter is selected in the proposed virtual mirror display based on its better performance.
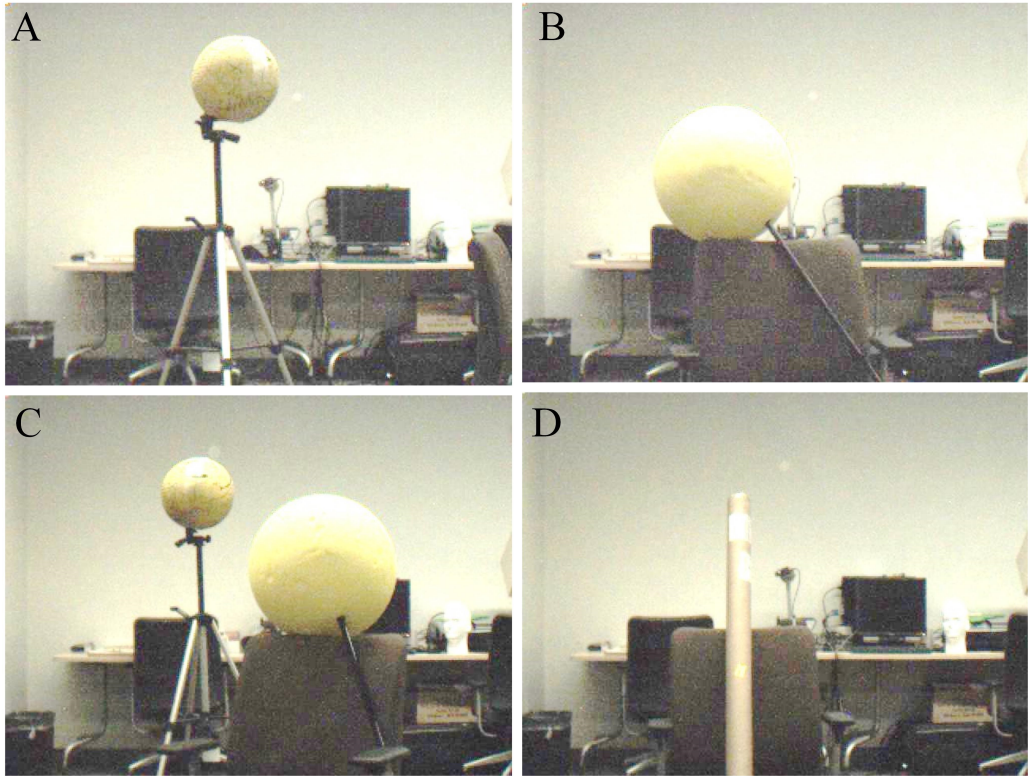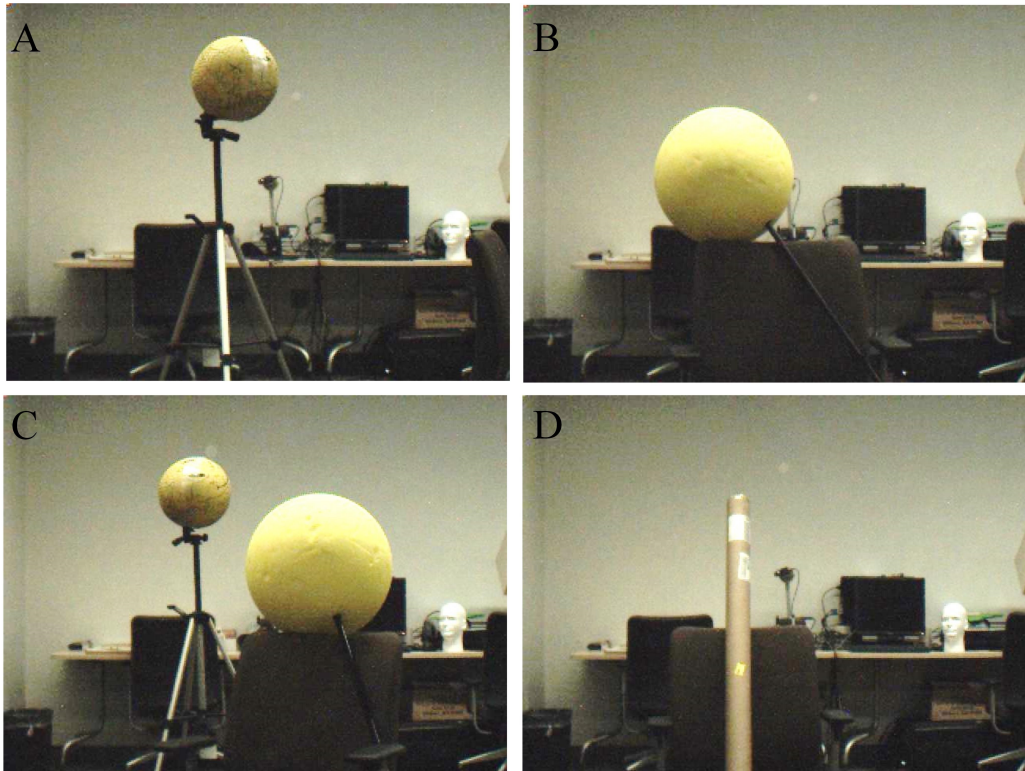
Figure 5.14: Rendered images using GHE algorithm

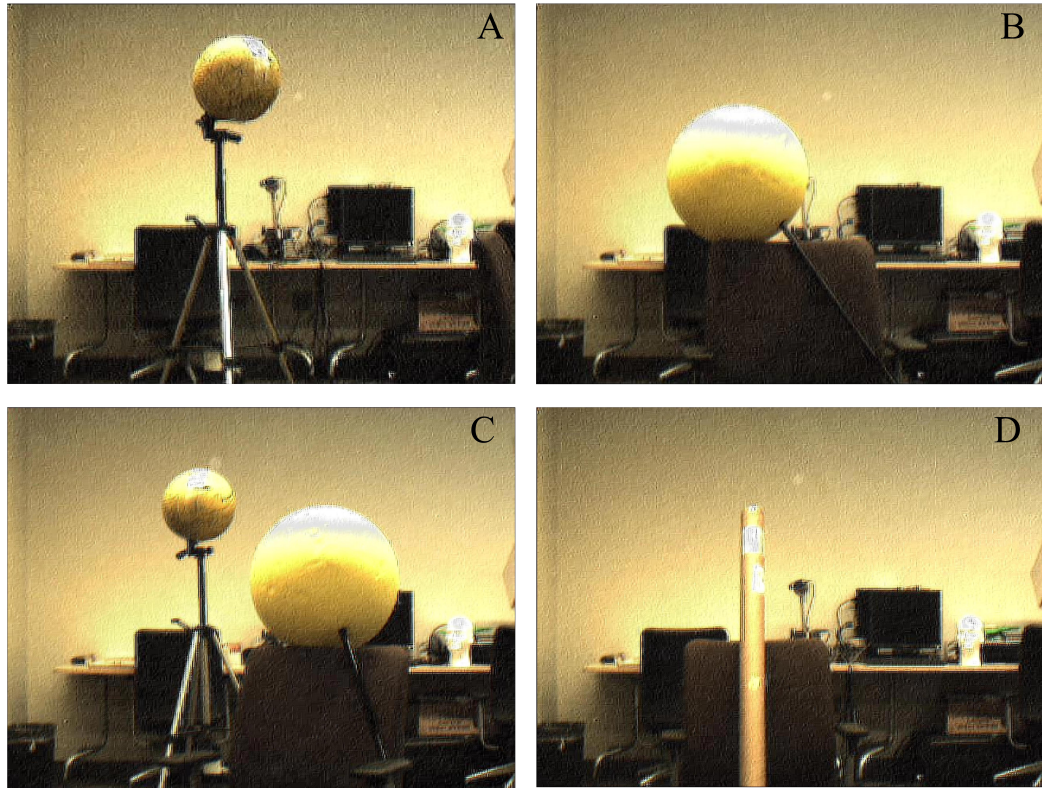Figure 5.15: Rendered images using CLIAHE algorithm

Figure 5.16: Rendered images using the proposed algorithm and image in Figure 5.12

Chapter 6

Conclusions and Future Work

In this research, a small-scale mirror simulation system using a see-through display and a robotic camera is developed. The aim of this system is to provide a mirror-like self-image. To accomplish this goal, a one-way mirror is placed on a table with an orientation of 45 degrees respect to the surface of the table. Robotics camera system, i.e. a high-quality camera mounted on a Bislide system, is set behind the mirror. It captures the front view of the user by tracking the user's viewpoint in real-time using a Microsoft Kinect. The obtained images are forwarded to a projector hanged above the mirror. They are then projected onto the screen on the top of the mirror and finally imaged into the mirror itself. The rendered image in the mirror has comparatively low contrast and deformed in shape, which significantly degrades the mirror-like experience. To this end, a reshape algorithm is introduced to correct the deformed image. Further, two histogram equalization algorithms and an optimal 2D wiener filter are analyzed in the thesis, and the optimal 2D wiener filter is selected for the contrast enhancement. Experimental statistics have validated the accuracy and effectiveness of the mirror display system.

In the future work, movements in the z direction in the rendered image need to be simulated. Virtual zooming, by depth-assisted foreground/background separation and independent scaling, is a promising direction to explore.

Bibliography

[1] *Examples of motorized linear stages & rotary indexing tables, http://www.velmex.com/motor examples.html.*

[2] *http://www.velmex.com.*

[3] *Multi axes bislide motor driven positioning system, http://www.velmex.com/motor bislide-multi.html.*

[4] S. Acker and S. Levitt. Designing videoconference facilities for improved eye contact. *Journal of Broadcasting and Electronic Media*, 31(2):181–191, 1987.

[5] M. Biehl, A. Ghosh, and B. Hammer. Learning vector quantization: The dynamics of winner-takes-all algorithms. *Neurocomputing*, 69:660–670, 2006.

[6] G. Bradski and A. Kaehler. *Learning OpenCV*. O'Reilly Media, Inc., 2008.

[7] T. Buggey. *Seeing Is Believing: Video Self-Modeling for People with Autism and Other Developmental Disabilities*. Woodbine House, 2009.

[8] S.-C. Cheung, Aug. 2013.

[9] C. Cullinan and S. Agamanolis. Reflexion: A responsive virtual mirror for interpersonal communication. In *In UIST 2002 posters and extended abstracts*. ACM Press, 2003.

[10] P. W. Dowrick. *Using video: Psychological and Social Applications*. New York, NY, USA: Wiley, 1983.

[11] J. Ehara and H. Saito. Texture overlay for virtual clothing based on pca of silhouettes. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 139–142, October 2006.

[12] P. Eisert, P. Fechteler, and J. Rurainsky. 3-d tracking of shoes for virtual mirror applications. In *Proceeding IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Anchorage, Alaska, June 2008.

[13] A. R. J. Francois and E. E. Kang. A handheld mirror simualtions. In *Proceedings. 2003 International Conference on Multimedia and Expo*, volume 2, pages II–745–8, July 2003.

[14] A. Hilsmann and P. Eisert. Realistic cloth augmentation in single view video. In *Proceeding of Visual, Modelling and Visualization Workshop*, 2009.

[15] C. H. Hitchock, P. W. Dowrick, and M. A. Prater. Video self-modeling intervention in school-based settings: A review. *Remedial and Special Education*, 24(1):36–45, 2003.

[16] R. Hummel. Image enhancement by histogram transformation. *Computer Graphics and Image Processing*, 6(2):184195, April 1977.

[17] H. Ishii and M. Kobayashi. Clearboard: a seamless medium for shared drawing and conversation with eye contact. In *CHI '92 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1992.

[18] D. J. Ketcham, R. W. Lowe, and J. W. Weber. Image enhancement techniques for cockpit displays. Technical report, HUGHES AIRCRAFT CO CULVER CITY CA DISPLAY SYSTEMS LAB, 1974.

[19] V. Kitanovski and E. Izquierdo. Augmented reality mirror for virtual facial alterations. In *18th IEEE International Conference on Image Processing (ICIP)*, 2011.

[20] V. Kitanvski and E. Izquierdo. 3-d tracking of facial features for augmented reality applications. In *Proceedings of Internaional Workshop for Image Multimedia Interaction Services*, 2011.

[21] M. Kuechler and A. Kunz. Holoport - a device for simultaneous video and data conferencing featuring gaze awareness. In *Proceedings of the IEEE conference on Virtual Reality*, pages 81–88, 2006.

[22] J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[23] J. Oppenheimer. Prompting apparatus, 1959.

[24] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz adn T. Greer, B. T. H. Romeny, J. B. Zimmerman, and K. Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39:355–368, 1987.

[25] V. S. Ramachandran and E. L. Altschuler. The use of visual feedback, in particular mirror visual feedback, in restoring brain functions. *Brain*, 132:1693–1710, 2009.

[26] V. S. Ramachandran, D. Rogers-Ramachandran, and S. Cobb. Touching the phantom limb. *Nature*, 377:489–490, 1995.

[27] J. Semple and G. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.

[28] J. Shen, S.-C. Cheung, and J. Zhao. Virtual mirror by fusing depth and color cameras. *IEEE Transactions on Image Processing*, vol. 22, issue 9:1–16, 2013.

[29] S. Shiwa. A large-screen visual telecommunications device using a liquid-crystal screen to provide eye contact. *Journal of the Society for Information Display*, 1(1):37–41, 1993.

[30] M. Straka, S. Hauswiesner, M. Ruther, and H. Bischof. A free-viewpoint virtual mirror with marker-less user interaction. In *Proceedings of the 17th Scandinavian conference on Image analysis*, pages 635–645, 2011.

[31] A. Taguchi, T. Aoki, and H. Yasuda. A study on real-time virtual clothing system based on two-dimensional plane model. In *6th Asia-Pacific Symposium on Information and Telecommunication Technologies*, pages 126–130, November 2005.

[32] Kar-Han Tan, I. Robinson, R. Samadani amd Lee Bowon, D. Gelb, A. Vorbau, B. Culbertson, , and J. Apostolopoulos. Connectboard: A remote collaboration system that supports gaze-aware interaction and sharing. In *IEEE International Workshop on Multimedia Signal Processing*, pages 1–6, 2009.

[33] L. Q. Uddin, M. S. Davies, A. Scott, E. Zaidel, S. Y. Bookheimer, M. Lacoboni, and M. Dapretto. Neural basis of self and other representation in autism: An fmri study of self-face recognition. *PLoS ONE*, 3(10), 2008.

[34] L. Wang, R. Villamil, S. Samarasekera, and R. Kumar. Magic mirror: A virtual handbag shopping system. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 19–24, 2012.

[35] Z. Wang. A fast implementation of adaptive histogram equalization. In *IEEE 8th International Conference on Signal Processing*, 2006.

[36] A. D. Wilson. Touchlight: An imaging touch screen and display for gesture-based interaction. In *ICMI04*, 2004.

[37] M. Grossand S. Wrmlin, M. Naef, E. Lamboray, C. Spagnoeth, A. Kunz, E. Koller-Meier, T. Svoboda, L. Van Gool, S. Lang, K. Strehlke, A. Vande Moere, and O. Staadt. Blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Transactions on Graphics (TOG)* -, 22(3):819–827, 2003.

[38] C. Zhang, R. G. Yang, T. Large, and Zhang Z. Y. A novel see-through screen based on weave fabrics. In *ICME '11 Proceedings of the 2011 IEEE International Conference on Multimedia and Expo*, 2011.

[39] K. Zuiderveld. *Graphics gems IV*, chapter Contrast limited adaptive histogram equalization, pages 474–485. Academic Press Professional, Inc, 1994.

Vita

YuQi Zhang was born on in YiChun, JiangXi, China.

Education

University of Kentucky, USA 06/2014

M.S. in Electrical Engineering

Sun Yat-sen University, China 10/2010

Master of Jurisprudence (M.J.)

Harbin Institute of Technology, China 07/2005

B.S. in Electrical Engineering

Professional Positions:

Teaching Assistant,

Electronics Laboratory, Circuit Design Laboratory, Matlab and Excel Laboratory,

University of Kentucky, 2012-2014

Research Assistant,

Multimedia Information Analysis (MIA) Laboratory,

University of Kentucky, 2013

Honors and Awards

Power and Energy Institute of Kentucky (PEIK) Graduate Certificate

Power and Energy Scholarship

International Student Tuition Scholarship

Microsoft Certified Professional

Microsoft Certified Professional Systems Engineer

Microsoft Certified Professional Database Administrator