Computational Imaging Techniques to Recover Omni3D Structure & Surface Properties

Thesis submitted in partial fulfillment
of the requirements for the degree of

Masters of Science
in
Computer Science and Engineering
by Research

by

Rajat Aggarwal
201201119
rajat.aggarwal@research.iiit.ac.in

International Institute of Information Technology
(Deemed to be University)
Hyderabad - 500 032, INDIA
November 2019

Copyright © Rajat Aggarwal, 2019
All Rights Reserved

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Computational Imaging Techniques to Recover Omni3D Structure & Surface Properties " by Rajat Aggarwal, has been carried out under my supervision and is not submitted elsewhere for a degree.

_____ _____

Date Adviser: Prof. Anoop M Namboodiri

To DreamVu.

Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. Anoop Namboodiri who not only brought me into the field of computational photography but also helped me in evolving as a person. His vision, creative thinking, mentorship and support to transform an idea into the reality is captured in this thesis. I can not be thankful enough to him to provide me this opportunity.

Every successful man has a woman behind him and it is true for me as well. I am blessed to have Sakshee in my life. She has supported me throughout this journey and played a pivotal role in making this journey incredible and unique. We are married now.

Nothing gets finished without the constant support and love from your friends and this acknowledgement is not enough to thank them. They have motivated me, helped me overcoming my personal problems, solved my issues throughout the research. I would also like to thank my friends Drushti, Krishnam, Urvashi and Divya. Special thanks to both my research partners Amrisha and Swetha, for teaching me so many things on my way.

Most importantly, my parents who have been the there always have told me to try different things and explore where one's passion lies. I can't thank them enough on the things they have taught me and always supported my decisions. I owe all my motivation and success to them.

Abstract

The process of imaging converts a 3D world into a 2D image. This process is inherently lossy, making the problem of understanding the world, ill-posed. During the imaging process, a light ray emanating from a source bounces of different surfaces, interacting with them according to the surface properties (albedo, color, specularity) before reaching the camera. The structure that we see in an image is primarily based on the final surface from which the light was reflected, except when that surface is highly specular (mirror-like). In this work, we explore imaging techniques the recover both the structure and surface properties of a scene.

For structure recovery, we extend the idea of stereo imaging and present a practical solution to capture a complete 360° panorama using a single camera. Current approaches either use a moving camera for capturing multiple images of a scene, which are then stitched together to form the final panorama, or use multiple cameras that are synchronized. A moving camera limits the solution to static scenes, while multi-camera solutions require dedicated calibrated setups. Our approach improves upon the existing solutions in two significant ways: It solves the problem using a single camera, thus minimizing the calibration problem and providing us the ability to convert any digital camera into a stereo panoramic capture device. It captures all the light rays required for stereo panoramas in a single frame using a compact custom designed mirror, thus making the design practical to manufacture and easier to use. We analyze the optimality of the design as well as present panoramic stereo and depth estimation results.

The methods for structure recovery, including stereo are often fooled when the surface is highly specular. To alleviate this, we propose an active-illumination based method to detect and segment mirror-like surfaces in a scene. In computer vision, many active illumination techniques employ Projector-Camera systems to extract useful information from the scenes. Known illumination patterns are projected onto the scene and their deformations in the captured images are then analyzed. We observe that the local frequencies in the captured pattern for the mirror-like surfaces is different from the projected pattern. This property allows us to design a custom Projector-Camera system to segment mirror-like surfaces by analyzing the local frequencies in the captured images. The system projects a sinusoidal pattern and capture the images from projector's point of view. We present segmentation results for the scenes including multiple reflections and inter-reflections from the mirror-like surfaces. The method can further be used in the separation of direct and global components for the mirror-like surfaces by illuminating the non-mirror-like objects separately. We show how our method is also useful for accurate estimation of shape of the non-mirror-like regions in the presence of mirror-like regions in a scene.

Contents

Chapter Page

1 Introduction 1

2 Related work 3

3 Camera Design to Capture Omni-Stereo Videos Using a Single CMOS Sensor 8
3.1 Design Overview 9
3.2 Design Details 11
3.2.1 Horizontal Field of View 11
3.2.2 Inter reflections 12
3.2.3 Vertical Field of View 13
3.2.4 Optimal surface for the coffee-filter shape 13
3.3 The Mirror Surface 15
3.3.1 Optimizing the design parameters 17
3.3.1.1 Optimal outer radius 17
3.3.1.2 Optimal angle between the two petals 18
3.3.1.3 Optimal angular curvature of a face 18
3.3.1.4 Optimal Radius of the curvature 19
3.3.2 Resultant Mirror Surface 19
3.3.3 Estimation of Surface normals 21
3.3.4 Epipolar Geometry and Stereo Depth Estimation 22
3.3.5 Derivation of Orthographic Projections 25
3.3.6 Calibration of the Proposed System 26
3.4 Results and Discussions 26

4 Camera-Projector system to Detect and Segment Mirror-like Surfaces Using Structured Illumination 32
4.1 Illuminating Mirror-like Surfaces 32
4.1.1 System Setup and Illumination Pattern 34
4.2 Segmentation by Illumination 35
4.2.1 Local Frequency Analysis 35
4.2.2 MRF for surface segmentation 36
4.3 Experimental Results 36
4.3.1 Separation of direct and indirect illumination components 38
4.3.2 3D Reconstruction 38

5 Conclusions 47

List of Figures

Figure Page

3.1 (a) 3D printed prototype of the proposed *coffee-filter mirror* (b) Image of the proposed mirror placed in a scene (POVRay) (c) Stereoscopic panorama (red-cyan anaglyph) recovered from the image in (b). 9
3.2 (a) The scene point P viewed by L and R eyes forming a viewing circle with diameter equal to baseline (b) Arrangement of mirrors for capturing right and left views. Tangential rays are captured by placing a mirror normal to the viewing circle. 11
3.3 Combination of the two arrangements as shown in Fig. 3.2b to capture both eye views in a single design. 12
3.4 (a) Flat mirrors have limited FOV, causing blind spots (b) Curved mirrors increase the FOV and the overlap between consecutive faces of an eye. 12
3.5 (a) Inter-reflections cause wrong world points to be captured due to FOV overlap between each pair of left and right eye view mirrors (b) A general horizontal cross-section of the proposed *coffee filter mirror*. 13
3.6 Uniformity of vertical resolution in terms of differential angle of the incident rays along the radial length for a cone, parabola and hyperbola. 14
3.7 Geometry of the petal surface used to obtain optimal design parameters. 17
3.8 Parameters of the mirror petal. 19
3.9 Geometry for deriving the normals direction at point **P**. 21
3.10 (a) Lines in black represents the locus of the intersection of consecutive rays for different radial lines lying on a single face (b) Red and blue curves represents the caustic curve for left and right faces respectively. 23
3.11 Geomtery for deriving the direction of incident ray direction. 23
3.12 Geometry showing the set of the points viewing the straight line **X** to derive the epipolar constraints. 24
3.13 Orthographic Projections using the proposed set up which includes the *coffee filter mirror* and a parabolic reflector 25
3.14 Red-Cyan anaglyph panorama obtained by using the proposed set up using POVRay dataset (1). 26
3.15 (a) Comparison of reconstructed depth as obtained using the proposed set up with the ground truth depth map 27
3.16 (a) A miniaturized model of the proposed design. (b) An example setup of the proposed device along with a consumer cellphone camera. 28
3.17 (a) Left-eye view (b) Right-eye view (c) Red-cyan stereo image of the Patio scene (19) 29
3.18 (a) Left-eye view (b) Right-eye view (c) Red-cyan stereo image of the Office scene (19) 30

3.19 (a) The red-cyan anaglyph of a scene (b) Depth map computed from the epipolar geometry of the proposed coffee-filter mirror design (c) Ground Truth depth of the scene . . 31

4.1 (a)Captured image with sinusoidal pattern projected onto the scene from projector's point of view. (b)Binary image representing the windows which are labelled as the specular regions. (c)Smooth segmentation obtained by applying MRF using Initial segmentation cues (d)Color Image of the specular regions present in the scene. 33

4.2 (a,b)Pattern projected on the specular surface is reflected from point C which is not necessarily in the scene. This changes the frequency of the pattern observed at a distance d from the projector's point of view P. 34

4.3 Derivation of the relation between the observed frequency and the projected frequency for the mirror-like surfaces. 35

4.4 Camera-projector alignment for capturing images from projector's point of view using One-way mirror. Rays pass through the One-way mirror and gets reflected back from point P in the direction of the camera. Camera is at distance d and captures the image from projector's point of view. 40

4.5 Segmentation Results for various kind of scenes. First column shows the actual image of the scene. Second column shows the image captured with sinusoidal pattern projected onto the scene from projector's point of view. Third column shows the Binary image representing the windows of size s which are labelled as the mirror-like regions. Fourth column shows the smooth segmentation obtained by applying MRF using Initial segmentation cues. Fifth column shows the colored segmentation result. 41

4.6 Segmentation result for a scene which includes multiple reflections from one mirror-like surface to another. It also includes the surface which is reflecting the pattern inside the scene. 42

4.7 Segmentation results where our approach fails. (a) Black case is dark in color, and hence frequency change can be seen. (b) The distance where the rays are reflected from the mirror-like object (red marked region) is very small. The local frequencies in this region is same as the projected frequencies. 42

4.8 (a-c) show the captured images with different frequency pattern. (d) R,G,B boundaries shows the segmentation result by a,b,c patterns, overlayed on the complete image. . . 42

4.9 (a-d)show the images captured at different camera distances. As the distance d, between the observation and projection point increases, observed frequency on the mirror-like regions decreases. 43

4.10 (a)shows the images of the scenes containing mirror-like surfaces. (b) shows the segmentation results from the proposed active illumination method. (c) shows the segmentation results by the method proposed in (18). 43

4.11 Average Recall/ Precision graph for (a)different window sizes used in STFT. (c) different frequency of the projected pattern. 44

4.12 (a)ROC curve illustrating the performance of the approach for segmentation of mirror-like objects in a scene. (b) Image of the experimental setup used in this work. 44

4.13 (a) Image obtained by complete illumination of the scene. (b) Direct Component of the illumination obtained from Nayar *et al.* (35). Inset images shows the reflections from the other objects in the scene in the direct component. (c) Image obtained after illuminating the non-specular regions only. (d) Direct component obtained after subtracting the non-specular illumination component. 45

4.14 (a) Sequence of images obtained at phases $-2\pi/3$, 0 and $2\pi/3$. (b) Segmented specular regions (b) Shape reconstruction using three phase shift method. The surface has errors in the regions where specular surface reflects light and causes change in the frequencies. (d) Correct shape reconstruction result obtained after projecting patterns on non-specular regions. In the inset, comparison between the before and after separation results are shown. 46

List of Tables

Table Page

2.1 Omnistereo device/approach comparison based on **Device Setup** complexity, **Type of Scenes** that the camera can handle, **Perceived Artifacts** in the omnistereo experience and **Resolution** of the omnistereo panoramas. 5

2.2 Comparison of different approaches for detection/segmentation of mirror-like surfaces, on the basis of the assumptions and Pros/Cons of each approach. 7

3.1 Omnistereo device/approach comparison based on **Device Setup** complexity, **Type of Scenes** that the camera can handle, **Perceived Artifacts** in the omnistereo experience and **Resolution** of the omnistereo panoramas. 10

Chapter 1

Introduction

The ability to capture and transmit high quality omnidirectional stereo videos enables several applications such as virtual tourism, remote navigation, and immersive entertainment. In recent years, display devices for stereo panoramas have become ubiquitous with virtual reality (VR) headsets that uses smart phone displays such as Google Cardboard and Samsung GearVR©. Stereo panoramic videos allows one to stand at any location in the world and look around as if they were present in the real environment. However, accurate omnidirectional stereo requires the capture of a four-dimensional light-field. If we assume the two eyes are restricted to move along a horizontal circle (panoramic stereo), one can create independent 2D panoramic images for each eye that closely reproduces the stereo views for a human. This result theoretically enabled the capture of omnidirectional stereo videos. However, practical systems for capturing stereo panoramic videos are just emerging (3).

The capture of stereo panoramas currently requires a camera that moves along a circle or a complex synchronized multi-camera setup. We aim to develop a simple solution that can capture 360° stereo panorama using an existing digital camera, thus making the creation process of immersive VR content accessible to a much larger population. The primary challenge here is to capture all the light rays corresponding two sets of cameras (left and right eye views) arranged along the same circle to a single sensor without causing blind spots or occlusions in the optical system.

The Omnistereo (Omni3D) cameras can also be widely used in applications such as machine vision, visual inspection, parts alignment, object recognition, estimating material properties, 3D scanners, space time stereos etc. However, capturing reliable information from a very complex scene including wide range of shapes and materials, is a difficult task. One of these complicated surfaces are mirror like surfaces. Mirror like surfaces are common objects in vision tasks and robot navigation. They are commonly found in indoor environments and its difficult to recognise for a camera whether it is an obstacle or a path. As service robots perform more and more tasks in indoor environments, the ability to recognize mirrors is necessary. In vision, these objects poses a problem for all techniques that are based on feature detection and matching, such as, binocular stereo and Structure from Motion. This is because, the appearance of mirror-like objects depends upon the environment and thus, it is difficult to determine whether an image feature corresponds to an actual scene point or it is the specular reflection

of another scene point. The representation of mirror-like surfaces depends not only on the surface properties but also on the properties of the surrounding scene and the viewpoint of the observer. One of the solutions to remove these kind of inaccuracies is to apply separate robust techniques for different surface materials. This makes segmentation of mirror-like surfaces an important problem. Detection of mirror-like surfaces in a scene, can simplify pre-processing steps in many practical tasks in robotic vision algorithms.

In this thesis, we present the computational imaging techniques and setups to recover the structure and the surface properties. Firstly, we present a camera system which captures Omni-Stereo videos in real time and secondly, we extend the camera system to capture the difficult surfaces like mirror surfaces by adding the projector system.

The second chapter in the thesis describes the detail design of the camera system to capture Omni-directional stereo videos. We show the mathematical design explanations, simulation results and also explain the practical use of such a camera system.

The third chapter presents an active illumination technique for segmentation of mirror-like surfaces from the scene. We propose a custom setup consisting of a camera, projector and a One-Way mirror which is used to capture the pattern projected onto the scene from the projector's point of view. We have developed a custom setup using which we can image the scene from the projector's point of view. We also show how detection the frequency change in the projected and the observed illumination pattern and use this difference as a cue for segmenting out the mirror like regions is useful. We show experiments on various scenes including mirrors and partially specular objects. We then demonstrate the usefulness of the segmentation to correctly separate out the direct and indirect illumination for the specular objects in a scene. Finally, we point out that this method can be employed to accurately reconstruct the surface if separate robust techniques are applied for different kind of surfaces.

Chapter 2

Related work

Last two decades have seen significant advances towards achieving omnistereo imaging. The first set of approaches generates a pair of panoramic images with a vertical offset. While such a geometry results in simple epipolar lines leading to fast disparity estimates, they are not suitable for human stereo perception as our brain expects a horizontal disparity. (21) proposed the use of two omnidirectional cameras, each consisting of a parabolic mirror, telecentric optics and a conventional camera, that were aligned vertically along the same axis. (25) proposed a more complex setup that replaces each omnidirectional camera in the above device with 6 cameras and a hexagonal mirror. While this increases the resoltuion of the omnidirectional images, the disparity remains vertical as before. (31) created omnidirectional stereo using two cameras and a conical mirror, which limits the vertical FOV. (51) improved on the above designs by capturing stereo omnidirectional panoramas with a single camera. Their system uses a mirror-lens combination to create two light paths that reflect off different positions in the mirror forming a stereo pair with a vertical disparity.

Acquiring a horizontal disparity panoramic stereo requires one to record a 3D light-field that is the equivalent of all cameras with centers along a horizontal circle. (40) showed that this may be compressed to two 360° panoramas, one for each eye, without affecting the visual perception. They proposed the use of a single camera that rotates in a circular trajectory to acquire the 3D light-field. Strips of images are then extracted from left and right ends of the frames that are stitched to create the right and left eye panoramas respectively. While this method works quite efficiently for static scenes, it is affected by artifacts like visible seams and vertical parallax for moving objects or uneven camera motions. These errors cause major mis-perceptions when viewed in 3D as explained by (22). (15; 16) proposed the use of a rotating stereo camera pair and stitching complete frames instead of small strips, which can generate panoramic stereo video textures. Use of rotating cameras limits the system to small repetitive motion and suffers from the perception artifacts like visible seams and vertical parallax.

(44) proposed a flow based blending approach that helps reducing these visual artifacts and works well for images captured using hand held cameras. However, the use of SFM makes this approach computationally expensive when applied to high resolution images.

(40; 42) also explored creating panoramic images using a perspective camera and a spiral mirror. They could capture the 360° panorama using three such setups, each acquiring 132°. They also proposed the use of a fresnel lens cylinder around an omnidirectional camera for each eye. Both systems were complex and hard to manufacture and use because of its large size and delicate structure.

To reduce visual artifacts and capture dynamic scenes, synchronized multi-camera setups were recently proposed. (14) have proposed a setup which uses 6 cameras with fisheye lenses. Large occlusions are visible in the produced omnistereo panoramas due to the huge curvature of the lenses. Also to reduce depth distortions, the number of cameras needs to be increased, which in turn makes the system bulkier. The latest and most effective solution is a simple extension of the idea by (40), where the rotating camera is replaced by 16 static cameras along a circle. This virtual reality camera was introduced at the Google I/O conference, called *'Google Jump'* (3). All the cameras are synchronized to take frame-aligned videos, which are then stitched to form a complete 360° video.

Although several systems have been proposed in the past most of them either suffer from limited FOV, or are limited by their size, calibration and alignment issues and cost of manufacturing. Our goal is to reduce this to a simple catadioptric system using a custom designed mirror and a single camera. We also provide a theoretical foundation of the design choices and demonstrate results for both static and dynamic scenes. Table 3.1 shows the more detailed comparison of our device with other omnistereo devices/approaches.

There has been a lot of work on removing specularity (highlights) from the image as shown in (10) but only a small number of researchers have looked into the issue of detecting specular surfaces and correcting the errors automatically. People have used different sensors to detect mirrors and windows , used the information from the surrounding to detect the feature points and the specular highlights to recognize specular surfaces.

Yang and Wang (49; 50) introduced a sensor fusion technique to detect potential mirror-like obstacles in a 2D occupancy grid map using sonar sensors and a laser scanner. However, their approach works in 2D only. Koch et. al. (27) presented a specular reflectance detection approach applicable to multi-echo laser scanners in order to identify and filter mirrored objects. Agha-mohammadi and Song (8) and Lu et.al. (32) proposed a technique of estimating a mirror transformation matrix and geometric constraints for corresponding real and virtual feature points in the image. However, these approaches assume the presence of feature points from the surroundings as well as their mirror images in the captured scene. Our technique assumes no prior environment information to detect the mirror-like surfaces. Kashammer and nuchter (24) captures a series of 3D laser scans. Laser completely gets reflected when strikes the mirror surface and do not appear in the point cloud at all. However, this is applicable only for framed rectangular mirrors whose dimensions are known.

Other approaches are based on the information from the environment to detect the distortions in a known shape. Oren and Nayar (37) analyze the characteristic and governing geometry of specular surfaces. However, their proposed method is limited to surfaces with high curvature and does not address detecting and modeling planar specular surfaces. Adato et. al. (5) proposes specular flow, but reconstruct

Principle	Camera Name	Device setup	Type of scenes	Perceived Artifacts	Image Resolution
Single Camera based	Omnistereo (40)	Rotating	Static	Mis-alignments, stitching artifacts & Vertical parallax	Medium
	Megastereo (44)	Rotating	Static	No artifacts but extensive run time for SFM	High
	Coffee-filter (*proposed*)	Fixed	Dynamic	No stitching artifacts and disparity errors	Medium
Multiple camera based, needs camera synchronization	Omnistereo, fresnel Lens solution(40)	2 Fixed cameras	Dynamic	Chromatic aberrations, difficult to manufacture and use , Self-occlusions	Low
	Panoramic Stereo Video Textures (15; 16)	2 Rotating Cameras	Dynamic textures	Visible seams and vertical parallax	High
	Google Jump (3)	16 Fixed Cameras	Dynamic	No stitching artifacts but setup is expensive and bulky	High
	Omnipolar (14)	6 Fixed Cameras	Dynamic	Self occlusions , setup is expensive and bulky	High
	Hexagonal Pyramidal Mirrors (25)	12 Fixed cameras	Dynamic	Vertical disparity errors	High

Table 2.1: Omnistereo device/approach comparison based on **Device Setup** complexity, **Type of Scenes** that the camera can handle, **Perceived Artifacts** in the omnistereo experience and **Resolution** of the omnistereo panoramas.

general surface shape under distant, unknown illumination by an environment map and a static observer. In their approach, the relative positions between the camera and the object must remain static, that is, only the environment map is allowed to move. Vasilyev et. al. (47) extends this approach and shows that one specular flow is needed to construct the shape of mirror-like surfaces. Savarese and DalPozo (18) proposed the features called as static specular flow (SSF), which are based on capturing the distortions of the surrounding scene due to the curvatures in the reflecting surface. However, such kind of specular features are only applicable to the curved reflective surfaces in proximity of either occluding contours of the object or regions where the difference of the two surface principal curvatures is high. In practice, objects need not always be curved and have SSF will not detect any distortion. Secondly, this approach is only applicable to the surrounding scenes with textures. Yilmaz *et al.* (52) proposed an algorithm that relies on scale and rotation invariant feature extraction techniques and uses motion cues to detect and localize specular surfaces. Geometry of the object requires to be known a priori in such approach. Osadchy *et al.* (38; 39; 36) uses specular highlights as the cue for segmentation. However, specular highlights are not always detectable and often even completely absent from the object. This is because

the perceived brightness becomes a strong function of the viewpoint due to highlights or reflections from the source. Hence they cannot be used as a main or unique cue for recognizing specular surfaces. To reduce these kind of ambiguities in the specular highlights, we propose to capture the specular reflections from the source's point of view. The main goal of our work is to provide an approach which only depends on the surfaces which are present in the scene.

Our goal is to come up with an active illumination technique which can be used to detect the specular surfaces without any prior information about the surrounding environment. Active illumination is also being used for image segmentation as shown in the work by Raskar et. al (43) who strategically positioned flashes to produce an edge depth map computed from a series of images, whereas koh et. al (28) uses multi-coloured lamps are used in place of the flashes which allows simultaneous acquisition of the images. Active lighting is also used in assisting the detection of specular surfaces. For example, Reiner and Donnaer (26) utilize stereo vision and a two dimensional array of light sources for constructing specular surfaces. Kutulakos and Steger (30) introduce a light-path triangulation method for constructing refractive and specular 3D objects, in which the light source must move along the light ray while the camera captures two consecutive images of the reflected light. Other active illumination methods are reconstruction methods (20; 48; 11) use a continuous area illumination or a single display to cast coded patterns onto the mirror-like object and use a multi-view approach to resolve the surface shape. Tin et. al. (46) proposes efficient ray coding scheme to reconstruct mirror type surfaces. Nayar and Gupta (34) proposes diffuse structured light by placing a diffuser between the light source and the scene. Their approach helps in reducing both specularities and shadows in structured light methods. These active illumination methods reconstructs the surface, when the scene contains specular objects and can not be used for discriminating between specular and non-specular regions in a scene. Whereas our method does not have any prior information about the presence of specular surface. Comparison of different approaches for detection/segmentation of mirror-like surfaces, on the basis of the assumptions and Pros/Cons of each approach is shown in Table 2.2.

	Method	Assumptions	Pros/Cons
(50; 27; 8)	Sensors or Laser Scan	Only Mirrors	Bulky and Expensive
(32; 24)	Geometric Transformations	Only Mirrors; known mirror dimensions	Both world points and its mirror image is captured
(18)	Specular Flow	Curved Surfaces; Textured Environment	mirror-like surfaces, single natural image
Ours	**Structured Illumination**	**known frequency value**	**mirror-like; single image**

Table 2.2: Comparison of different approaches for detection/segmentation of mirror-like surfaces, on the basis of the assumptions and Pros/Cons of each approach.

Chapter 3

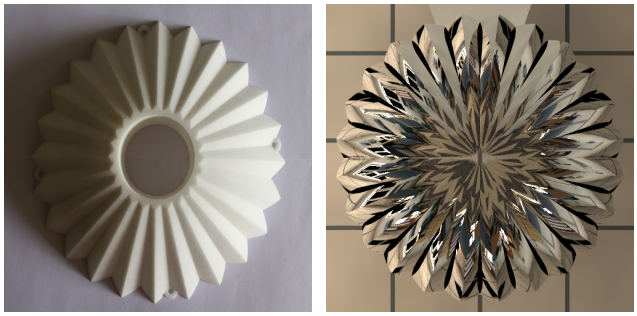
Camera Design to Capture Omni-Stereo Videos Using a Single CMOS Sensor

We show for the first time that such an optical system is possible and practical using a set of custom designed reflective surfaces that we refer to as the *Coffee Filter Mirror* (See Fig. 3.1). We derive the surface equations of the mirror petals and discuss the optimization of its parameters for maximizing the visual quality of the captured images.

As mentioned before, most of the existing approaches involve either multiple cameras to capture various perspective directions (21; 3; 9), or a single camera that rotates along a horizontal circle to acquire the images (40; 42; 44). The problem with moving or multiple cameras is that it makes the calibration and camera positioning difficult thus making the system bulky and delicate to use. They also give rise to visible artifacts like motion parallax, visible seams, synchronization errors and mis-alignments (22). Moving cameras also limit their use to static scenes and require extensive post processing to get the left and right panoramas. In comparison, our approach has the following advantages:

1. **Simplicity of Data Acquisition:** In the multi-camera systems like Google Jump (3) or the system proposed by (9), all the cameras need to be synchronized using an electronic system to ensure that the images are captured at the same time. The use of a single camera eliminates the need for synchronization and reduces the size of the device making it easy to be used and handled. Data is acquired in the form of a regular image or video and may be stored in standard formats.

2. **Ease of Calibration and Post Processing:** Our approach solves the problem without any moving parts, thus simplifying the calibration process. As we explain later, simple binary patterns can be used to calibrate the relative configuration of the camera and the mirror. This also simplifies the post-acquisition de-warping process to obtain the left and right panoramas. The whole process of data acquisition and post processing can easily be done on a smartphone, making it a panoramic stereo video capture and display device.



(a) (b)

(c)

Figure 3.1: (a) 3D printed prototype of the proposed *coffee-filter mirror* (b) Image of the proposed mirror placed in a scene (POVRay) (c) Stereoscopic panorama (red-cyan anaglyph) recovered from the image in (b).

3. **Adaptability to Various Applications:** Our custom designed mirror can be easily used as an attachment to any consumer camera to convert it into a stereo panoramic capture device. The size of the mirror can be adapted according to the application and the field of view can be controlled.

4. **Complete omni-directionality:** Our design can be extended using a concave lens to capture mono images of the top or bottom region. We can thus generate a complete 360° x 270° view of the world captured using a simple set up.

We describe in detail the design of the proposed catadioptric omni-stereo system. We also show that our system works efficiently for acquiring both 3D images and videos and for both static and dynamic scenes. We discuss the optimization of the design parameters and present the reconstructed stereo panoramas.

3.1 Design Overview

3D information can be extracted from two images taken by two cameras horizontally displaced by a baseline. For a complete 360° view, the two cameras, analogous to the two eyes, can move around the center of a circle called *viewing circle*, as shown in Fig. 3.2a. The diameter of the viewing circle is equal to the baseline. For each viewpoint, the set of tangential rays in the clockwise direction account for the left eye views, and the set of tangential rays in the anti-clockwise direction account for the right

Principle	Camera Name	Device setup	Type of scenes	Perceived Artifacts	Image Resolution
Single Camera based	Omnistereo (40)	Rotating	Static	Mis-alignments, stitching artifacts & Vertical parallax	Medium
	Megastereo (44)	Rotating	Static	No artifacts but extensive run time for SFM	High
	Coffee-filter (*proposed*)	Fixed	Dynamic	No stitching artifacts and disparity errors	Medium
Multiple camera based, needs camera synchronization	Omnistereo, fresnel Lens solution(40)	2 Fixed cameras	Dynamic	Chromatic aberrations, difficult to manufacture and use , Self-occlusions	Low
	Panoramic Stereo Video Textures (15; 16)	2 Rotating Cameras	Dynamic textures	Visible seams and vertical parallax	High
	Google Jump (3)	16 Fixed Cameras	Dynamic	No stitching artifacts but setup is expensive and bulky	High
	Omnipolar (14)	6 Fixed Cameras	Dynamic	Self occlusions , setup is expensive and bulky	High
	Hexagonal Pyramidal Mirrors (25)	12 Fixed cameras	Dynamic	Vertical disparity errors	High

Table 3.1: Omnistereo device/approach comparison based on **Device Setup** complexity, **Type of Scenes** that the camera can handle, **Perceived Artifacts** in the omnistereo experience and **Resolution** of the omnistereo panoramas.

eye views. To accurately capture stereo information, the camera should be able to capture all the rays tangential to the viewing circle. For this purpose, we propose a special mirror design for omnistereo viewing, called *coffee filter mirror*, owing to the similarity in shape. Using the image captured by this mirror we generate panoramas for left and right eye views with appropriate disparity. The disparity in these images is used to perceive depth when seen in 3D using a VR headset such as (2).

Fig. 3.3 shows a horizontal cross section of the proposed mirror surface, where PB is tangent to the viewing circle V and represents the direction of the rays that are required to be collected for the right eye view. Using a mirror P_1P_2, normal to the rays falling in the direction PB , the incident rays can be reflected to the camera placed at the center of the design. P_1P_2 provides a horizontal field of view to the eye and is analogous to the strip width in the image-based approaches as mentioned in (40). Each such mirror is referred to as a *face*. Multiple faces, when arranged in the configuration shown in Fig. 3.2b, captures all the tangential rays required for constructing the panorama for a single eye. These faces are arranged at equal angular separation such that P_i (even i) lies on the circle C_{max} and P_j (odd j) lies on

the circle C_{min}. Rays that are falling in the direction of PB correspond to the view of the world as seen from the right eye, as shown in Fig. 3.3.

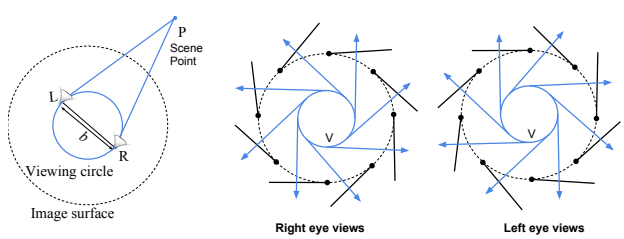
Our design is motivated from the idea to capture both eyes' views in a single device for omnistereo imaging. Between two consecutive faces that are catering to same eye view, we introduce a similar face that captures a second eye view.We combine the two arrangements of Fig. 3.2b in a single design, as shown in Fig. 3.3, such that rays for both eyes' views are reflected to a single camera. The combination of one left and one right face is referred to as a *petal* in rest of the paper. Different number of petals can be used depending upon the application requirements. Fig. 3.5b shows a general horizontal cross-section of the proposed device to explain the structure and the nomenclature that would be used further in the paper.

3.2 Design Details

In order to provide a better 3D experience to the user, it is important that the horizontal FOV of each face of the mirror is sufficient enough to avoid any stitching artifacts and mis-alignments. Also, vertical FOV should cover the appropriate height of the world and the resolution must be uniform across all the regions of the captured scene. In this section, we explain how these factors affect our choice of design parameters.

3.2.1 Horizontal Field of View

Horizontal field of view refers to the amount of the scene captured by a face in the horizontal direction. In our design, this is directly dependent on the number of petals, say n. To cover the complete 360° view, each face should cover at-least $\frac{2\pi}{n}$ FOV. As shown in Fig. 3.4a, face P_1P_2, which is a flat mirror, views only the region parallel to its length. The best arrangement of the next face for the same eye view, P_3P_4 would be such that the FOV of the two mirrors cover consecutive areas of the scene. However, this is only possible if the mirrors are arranged linearly. FOVs can not intersect at any point if



(a) (b)

Figure 3.2: (a) The scene point P viewed by L and R eyes forming a viewing circle with diameter equal to baseline (b) Arrangement of mirrors for capturing right and left views. Tangential rays are captured by placing a mirror normal to the viewing circle.

Figure 3.3: Combination of the two arrangements as shown in Fig. 3.2b to capture both eye views in a single design.

at mirrors are arranged in a circular manner, as shown in Fig. 3.4a. Hence, certain areas of the world will be left unseen by the camera and will be missing from the captured images.

(a) (b)

Figure 3.4: (a) Flat mirrors have limited FOV, causing blind spots (b) Curved mirrors increase the FOV and the overlap between consecutive faces of an eye.

Field of view of each face is increased by opting for a horizontal curvature for the face as shown in Fig. 3.4b. Center of each curved face lies on the tangent to the viewing circle. Increasing the curvature of the faces, increases the overlap between the FOVs of two consecutive same eye view's faces. This overlap is advantageous while de-warping the captured image into the panoramas, as it resolves the problem of missing regions. However, the increase in curvature also increases the inter-re ections between the neighboring faces.

3.2.2 Inter re ections

Due to the curvature in the faces in the proposed design, the FOV of the adjacent faces aligned at an obtuse angle, overlap. As shown in Fig. 3.5a, a ray originating from the camera that strikes the face P_2P_3, may get re ected and fall on the face P_3P_4. Such inter-re ections will cause wrong world points to be captured at the camera sensor. The repetition of such wrongly captured pixels decreases the resolution of the nal de-warped panoramas. Optimal amount of overlap is kept such that inter-re ections can be kept to a minimum, while also solving the problem of missing regions as explained in Section 4.1.

3.2.3 Vertical Field of View

In the past, several catadioptric systems using various combinations of flat and curved mirrors have been used to increase the amount of the scene captured by the camera in the vertical direction. Our proposed design is circular at each horizontal cross section. The radii of these circles increase from 0 to R_{max}. For better quality of the de-warped images, the resolution needs to be uniform. In a flat mirror, vertical FOV is equal to the height of the mirror and the resolution remains constant. As we need a uniform resolution, we see that a parabolic curvature in the vertical direction is more desirable compared to a hyperbola or straight line(see Fig.[1] 3.6)

3.2.4 Optimal surface for the coffee- filter shape

In this section, we explain the reason for choosing paraboloid surface for the design of coffee- filter mirror. Let us take a general quadric surface equation $a_1x^2 + a_2y^2 + a_3z^2 + a_4x + a_5y + a_6z + a_7 = 0$. For a flat surface, $a_1; a_2; a_3$ are zero, for a paraboloid surface, $a_3; a_4; a_5$ are zero and for hyperboloid surfaces $a_4; a_5; a_6$ are zero. To calculate the variation in the resolution of the image captured along a radial line, we find the difference between the direction of the consecutive incident rays. In the case of orthographic projections, direction of reflected light rays is same and the direction of incident light rays is directly proportional to normals of the mirror's surface. We find the variation of normals n_v which is given by,

$$n_v = \begin{bmatrix} \cos & \sin & \frac{dz}{dr} \end{bmatrix} \tag{3.1}$$

for a constant ϕ and varying r. We find $\frac{dn_v}{dr}$ by,

$$\frac{dn_v}{dr} = \begin{bmatrix} 0 & 0 & \frac{d^2z}{dr^2} \end{bmatrix}$$

Using $x = r\cos\phi$ and $y = r\sin\phi$, we get a general quadratic function given by $f(x; y; z; \phi)$.

[1] Please see Supplementary Material for more mathematical details

(a) (b)

Figure 3.5: (a) Inter-reflections cause wrong world points to be captured due to FOV overlap between each pair of left and right eye view mirrors (b) A general horizontal cross-section of the proposed coffee filter mirror.

Figure 3.6: Uniformity of vertical resolution in terms of differential angle of the incident rays along the radial length for a cone, parabola and hyperbola.

$$f(r; z; \quad) = a_1 r^2 \cos^2 \quad + a_2 r^2 \sin^2 \quad + a_3 z^2 + a_4 r \cos$$
$$+ a_5 r \sin \quad + a_6 z + a_7$$

Double differentiating $f(r; z; \quad) = 0$, we get,

$$\frac{d^2 f}{dr^2} = 2 a_1 \cos^2 \quad + 2 a_2 \sin^2 \quad + 2 a_3 \frac{dz}{dr}^2$$
$$+ 2 a_3 z \frac{d^2 z}{dr^2} + a_6 \frac{d^2 z}{dr^2}$$

Consider the case of a at mirror, where $a_1 = 0; a_2 = 0; a_3 = 0$, we get $\frac{d^2 z}{dr^2} = 0$, which means that the at surface has uniform eld of view and does not change with This means that the eld of view is totally dependent upon the height of the mirror.

For the case of a paraboloidal surface, where $a_6 = 0$, we get,

$$\frac{d^2 z}{dr^2} = \frac{2(a_1 \cos^2 \quad + a_2 \sin^2 \quad)}{a_6} \tag{3.2}$$

which means that the resolution for the paraboloidal surface is uniformly increasing with respect to As r increases, resolution increases. Also the FOV captured in paraboloidal shape is more than that of the at mirror.

Considering hyperboloidal surface, where $a_6 = 0$, we get,

$$\frac{d^2 z}{dr^2} = \frac{2(a_1 \cos^2 \quad + a_2 \sin^2 \quad + a_3 \frac{dz}{dr}^2)}{a_6 + 2 a_3 z} \tag{3.3}$$

Although, FOV is more than that captured by the at mirror, the resolution increases non-uniformly with r. The resolution difference between the pixels at the upper part of the design and the lower part of the design is drastically high, which makes the choice of hyperbolic design irrelevant.

3.3 The Mirror Surface

In this section we derive the equation of the mirror surface. Multiple factors can be varied to make the device adaptive to speci c applications. We derive the expressions for only one petal APB as shown in the Fig. 3.7a. and the same expressions hold for n petals rotated by $2\pi/n$. Circular surfaces AP and PB are used to capture the right and left eye view respectively. Let us consider the angle between the chords of these two faces as α such that $\angle APB = \alpha$. Each petal subtends an angle θ at the center, where $\theta = \frac{2\pi}{n}$ such that $\angle AOB = \theta$. Hence, we get n views each for left and right eye. The design of the mirror is symmetrical, and all the petals are of same size and dimensions. The length of each face, referred as petal length, denoted by l, as shown in Fig. 3.7a. Each petal, say P_i, where $i = 1$ to n is bounded by a circle C_{max} with radius R_{max}, and inside by a circle C_{min} with radius R_{min}. V is the viewing circle with radius equal to b. From Fig. 3.7a, $OA = R_{min}$ and $OP = R_{max}$. From $\triangle OAP$ and $\triangle OBP$, by sine rule we get the relations as,

$$\frac{l}{\sin(\frac{\alpha}{2})} = \frac{R_{max}}{\sin(\frac{(\pi+\alpha)}{2})} = \frac{R_{min}}{\sin(\frac{\alpha}{2})} \tag{3.4}$$

Since each face is symmetrical and oriented at equal separation, OP is the angle bisector of the $\angle APB$, such that

$$\angle APO = \angle BPO = \frac{\alpha}{2}$$

$$\angle AOP = \angle POB = \frac{\theta}{2}$$

Therefore, in $\triangle OAP$ we get

$$\angle OAP = \pi - \angle APO - \angle AOP$$

$$\angle OAP = \pi - \frac{\alpha}{2} - \frac{\theta}{2}$$

Since, LD is the perpendicular bisector of the chord AP and is tangent to the viewing circle V, $\angle DLP$ and $\angle CDO$ are the right angles and $LP = l/2$. $OD = b$ is the radius of the viewing circle. In $\triangle OCD$ and $\triangle CLP$, we get

$$LP = CP\cos(\frac{\alpha}{2})$$

$$CP = LP\sec(\frac{\alpha}{2})$$

$$= \frac{l}{2}\sec(\frac{\alpha}{2}):$$

We know, $OC + CP = R_{max}$, which gives

$$OC = R_{max} - \frac{l}{2}\sec(\frac{\alpha}{2}) \tag{3.5}$$

In 4 PLC,

$$\angle LCP = \quad \angle CLP \quad \angle CPL$$

$$= \frac{}{2} \quad \frac{}{2}$$

$\angle OCD = \angle LCP$ being vertically opposite angles. Thus we get,

$$\angle COD = \frac{}{2} \quad \angle OCD$$

$$= \frac{}{2} \quad (\frac{}{2} \quad \frac{}{2})$$

$$= \frac{}{2}$$

In 4 OCD,

$$\frac{OD}{OC} = \cos\frac{}{2}$$

$$OC = b\sec(\frac{}{2}) \tag{3.6}$$

Comparing Eqn 3.5 and Eqn 3.6 we get

$$R_{max} \quad \frac{l}{2}\sec(\frac{}{2}) = b\sec(\frac{}{2})$$

$$R_{max} = (b + \frac{l}{2})\sec(\frac{}{2}): \tag{3.7}$$

Combining Eqn 3.4 and Eqn 3.7, we get:

$$2R_{max}\cos\frac{}{2} = 2b + l$$

$$2R_{max}\cos\frac{}{2} = 2b + R_{max}\frac{\sin\frac{}{2}}{\sin\frac{+}{2}}$$

$$R_{max} \quad \frac{(2\cos\frac{}{2}\sin\frac{+}{2}) \quad \sin\frac{}{2}}{\sin\frac{+}{2}} = 2b$$

$$R_{max} \quad \frac{(2\cos\frac{}{2}(\sin\frac{}{2}\cos\frac{}{2} + \cos\frac{}{2}\sin\frac{}{2})) \quad \sin\frac{}{2}}{\sin\frac{+}{2}} = 2b$$

$$R_{max} \quad \frac{(2\cos^2\frac{}{2} \quad 1)\sin\frac{}{2} + \cos\frac{}{2}(2\sin\frac{}{2}\cos\frac{}{2})}{\sin\frac{+}{2}} = 2b$$

$$R_{max} \quad \frac{\cos \quad \sin\frac{}{2} + \cos\frac{}{2}\sin}{\sin\frac{+}{2}} = 2b$$

$$R_{max} \quad \frac{\sin\frac{+2}{2}}{\sin\frac{+}{2}}^{!} = 2b \tag{3.8}$$

Combining Eqn 3.4 and Eqn 3.8, we get

$$R_{max} = \frac{2b\sin(\frac{+}{2})}{\sin(\frac{+2}{2})} \tag{3.9}$$

$$R_{min} = R_{max}\frac{\sin(\frac{}{2})}{\sin(\frac{+}{2})} = \frac{2b\sin(\frac{}{2})}{\sin(\frac{+2}{2})} \tag{3.10}$$

$$l = R_{max}\frac{\sin(\frac{}{2})}{\sin(\frac{+}{2})} = \frac{2b\sin(\frac{}{2})}{\sin(\frac{+2}{2})} \tag{3.11}$$

3.3.1 Optimizing the design parameters

In our proposed design, disparity and device size can be altered depending upon the application requirement. Size of the device is proportional to R_{max}. In order to have a compact design of the mirror that generates human perceivable stereo panoramas, the design parameters need to be optimized.

3.3.1.1 Optimal outer radius

The value of the outer radius of the coffee lter mirror is R_{max} is dependent upon. We minimize the parameter R_{max} as given by Eqn 3.9 and get an optimal petal angle,

$$\boxed{_{opt} = \frac{}{2}} \tag{3.12}$$

at which R_{max} is minimum, and hence we get the minimum size of the device.

(a) (b)

Figure 3.7: Geometry of the petal surface used to obtain optimal design parameters.

3.3.1.2 Optimal angle between the two petals

We now find the optimal angle between the two petals. In Fig. 3.7a, let $\angle PBE$ be ϕ, the angle between two petals. Consider $\triangle OBP$ where

$$\angle OPB = \frac{\pi}{2}; \angle POB = \frac{\pi}{2}$$
$$\angle OBP = \frac{\phi + \theta}{2}$$

Therefore

$$\angle PBF = \angle OBP$$
$$= \left(\frac{\phi + \theta}{2}\right)$$
$$= \frac{\phi + \theta}{2}$$

Since OP is the angle bisector at equal separation, $\angle PBE = 2\angle PBF$, which means $\phi = \theta + \theta$, which means $\phi_{opt} = \theta + \theta_{opt}$. Hence, from Eqn 3.12 we get

$$\boxed{\phi_{opt} = \theta + \frac{\pi}{2} = \frac{\phi + \pi}{2}} \tag{3.13}$$

3.3.1.3 Optimal angular curvature of a face

Next, we find the optimal value of the angular curvature of a face. Consider Fig. 3.7b, where O^0 is the center of curvature of the face PB. PO^0 and O^0B are the radii of curvature i.e. r_c and $\angle PO^0B = 2\theta$ is the angle subtended by each face at the center of curvature. In $\triangle PO^0B$, $\angle A = \frac{\pi}{2}(\theta + \theta)$, which implies,

$$= \frac{\pi}{2} \quad \angle A = (\theta + \theta)\frac{\pi}{2}$$

In order to have a smaller device size,

$$\theta_{opt} = (\theta + \theta_{opt})\frac{\pi}{2}$$

Therefore, the optimal horizontal angular field of view is given by:

$$\boxed{\theta_{opt} = \frac{\pi}{2}} \tag{3.14}$$

and is independent of the obtuse angle $\angle PBE$ between two adjacent petals.

3.3.1.4 Optimal Radius of the curvature

From Fig ??, O^0C is the perpendicular bisector of CB, $CB = \frac{l}{2}$. In $\triangle O^0CB$, $\frac{l/2}{r_c} = \sin$. Radius of curvature r_c can be optimized by using the optimal value of . Therefore,

$$\boxed{r_{c_{opt}} = \frac{l}{2\sin\frac{}{2}}} \tag{3.15}$$

is the optimal radius of curvature. It is to be noted that these centers of curvature lie on a circle.

To avoid wastage of pixels due to inter-re ections, as explained in Section 4.2 in the main paper, it is important to collect the maximum scene information in the captured image. Each face $\frac{2}{n}$ covers angular FOV, thus a total of n such faces for each view covers complete 2 FOV. For no missing regions, FOVs of two faces for the same eye views should be covering consecutive areas of the scene. This is achieved by aligning one face in the direction of O^0P and the next face for the same eye view, in the direction BE. Hence the obtuse angle between the two faces PB and BE is $\frac{+}{2}$. The amount of inter-re ections depends upon the angle between two consecutive petals, which depends upon the sampling angle of the device $\frac{2}{n}$. Ideally, the amount of inter-re ections reduces down to zero, when the FOV of two consecutive faces do not intersect at all. However, this way, some of the scene regions will be left uncovered in the FOV of some faces and hence not imaged at all. In order to account for these inter-re ections, we introduce a small angle such that the angle of curvature becomes $2 +$. This makes sure some overlap is there, so that some redundant information is captured, which can be used while dewarping. However, the value of is kept suf ciently low, such that inter-re ections are also reduced to a huge extent.

3.3.2 Resultant Mirror Surface

Figure 3.8: Parameters of the mirror petal.

In this section, we obtain the surface equations of the proposed coffee lter mirror in terms of polar coordinates and r. As explained earlier, the surface of the coffee lter mirror is paraboloidal vertically

and circular in each horizontal cross section. Let us consider the central axis of the mirror to be the z axis. Then the surface equation can be written as a function of x and y axis:

$$z = f(x; y) = m\ (x^2 + y^2); \tag{3.16}$$

where m is the slope of the parabola for a given. Let $x^2 + y^2 = r^2$, where r is the radial distance in the XY plane and is the angle of the radial line, then:

$$z = m\ r^2 \tag{3.17}$$

Eqn 3.17 represents the petal surface of our custom designed mirror centered around origin. Consider the uppermost and widest cross section of the mirror $r = r_1$ at z_{max}, such that

$$z_{max} = m\ r_1^2$$
$$m = \frac{z_{max}}{r_1^2}$$

Let $(x_c; y_c)$ be the center of the circle of curvature of a face of a petal (x_d, y_d) be the point which lie on the curvature $r_1^2 = k^2 r^2$ such that $x_d = kx$ and $y_d = ky$. r_c be the radius of the circle of curvature for a face. Combining this with Eqn 3.17, we get

$$m = \frac{z_{max}}{k^2 r^2}$$

which implies

$$z = \frac{z_{max}}{k^2}$$

Distance between $(0; 0)$ and $(x_c; y_c)$ is d_c such that $x_c^2 + y_c^2 = d_c^2$, Calculating distance from center of the curvature and the point on the curvature we have:

$$(x_d - x_c)^2 + (y_d - y_c)^2 = r_c^2$$
$$(kx - x_c)^2 + (ky - y_c)^2 = r_c^2$$

$$\Rightarrow \quad k = \frac{(xx_c + yy_c) + \sqrt{(xx_c + yy_c)^2 - r^2(d_c^2 - r_c^2)}}{r^2}$$

Since, $m = z_{max} = k^2 r^2$,

$$m = z_{max} \left(\frac{r}{(xx_c + yy_c) + \sqrt{(xx_c + yy_c)^2 - r^2(d_c^2 - r_c^2)}} \right)^2$$

Also, from the Fig. 3.8, it is to be noted that $(x_c; y_c)$ forms angle $_1 + \frac{}{2} + \frac{}{2} + \ _2$ from the horizontal. Hence,

$$x_c = x_d + r_c(\cos(\ _1 + \frac{}{2} + \frac{}{2} + \ _2)$$

$$y_c = y_d + r_c(\sin(\ _1 + \frac{}{2} + \frac{}{2} + \ _2))$$

where $_2 = \tan\ ^1(\frac{2r_c}{l})$.

From this and Eqn 3.17 we get,

$$z = z_{max} \quad \frac{}{(xx_c + yy_c) + p\frac{r^2}{(xx_c + yy_c)^2 \quad r^2(d_c^2 \quad r_c^2)}}^{!\ 2} \tag{3.18}$$

Therefore, Eqn 3.18 gives the equation of the paraboloidal surface of the mirror. Note that the slope m at every point is a function of r.

3.3.3 Estimation of Surface normals

In this section, we derive the equation of the normal vector of a point on coffee- lter mirror $\hat{n}$. Let be the direction of the normal vector of point $P(r; \ )$. We nd $\hat{n}$ by nding the normal vector of the tangent plane at point $P(r; \ )$ which consists of tangent vectors in horizontal and vertical plane such that $\hat{n} = PA \quad PB$. PA and PB are the tangent vectors at point P in the horizontal and vertical direction respectively.

(a) (b)

Figure 3.9: Geometry for deriving the normals direction at point P

Let us consider the horizontal plane containing point P as shown in Fig. 3.9a where P lies on a circular curvature with center O such that $P = (r\cos\ ; r\ \sin\ ; z_0)$, $O = (x_c; y_c; z_0)$ and

$$OP = P \quad O = \begin{bmatrix} r\cos \quad x_c \\ r\sin \quad y_c \\ 0 \end{bmatrix}^T$$

And the vector PA which is orthogonal to OP is this given by:

$$PA = n_h = \begin{bmatrix} y_c & r\sin & r\cos & x_c & 0 \end{bmatrix}^T$$

Similarly we calculate PB in the vertical direction for a $x$ed , where $P(x; y; z) = (r\cos; r\sin; m\ r^2)$ such that PB is given by

$$PB = n_v = \begin{bmatrix} \frac{dx}{dr} & \frac{dy}{dr} & \frac{dz}{dr} \end{bmatrix}^T$$
$$= \begin{bmatrix} \cos & \sin & 2m\ r \end{bmatrix}^T$$

In the end, we calculate n as $n = \hat{PA} \quad \hat{PB} = \hat{n}_h \quad \hat{n}_v$.

3.3.4 Epipolar Geometry and Stereo Depth Estimation

Consider the case of a general pinhole camera, the incident rays intersect at the camera center. Similarly, we derive the center of the proposed coffee- lter mirror. We nd the trajectory of the viewpoints by nding the intersection of the re ected rays for every pair of adjacent points. Let a point lying on a radial line with single $_j$ is denoted by $p_{i;\ j}$. Then the re ected rays for each pair of adjacent points $p_{i;\ j}$ and $p_{i+1;\ j}$ intersect at the $c_{i;\ j}$. The set of all such points form a line as shown in Fig. 3.10a. $L_j = f\, c_{i;\ j}\, g8i$. We take the average of these intersection points c^0_j, to nd the average viewing circle for complete 360 view. Locus of c^0_j for for a single face is a straight line parallel to central axis of the camera. Let c^{00}_k be the average center for a single face. Locus of c^{00}_k for all n faces is viewing circle, as shown in Fig. 3.10b. Thus, the locus of the viewpoints in the proposed design is the viewing circle. Thus, in order to have a device with diameter of the viewing circle equal to the baseline, diameter of the uppermost cross-section must be twice the human baseline. With the change of diameter of the cross-section from top to bottom, both linear and angular disparities change. However, for panorama to be perceivable, linear disparities are more important than the angular disparities. The change in linear disparities are overcame by interpolation during dewarping. The combination of mirrors and a conventional camera behave as a non-central catadioptric camera. For details on such classi cation, readers are directed to (45).

In this section we explain the derivation of the epipolar geometry in detail. In general, epipolar geometry is used for reducing the search space of matching points in the two stereo images. In our device setup, the arrangement of the mirrors capturing both left and right eye views, enables us to calculate the epipolar geometry by nding the direction of incident rays which are captured using the parabolic re ector. As described in the Section 5.3 in the main paper, the conventional camera used in the system captures the light rays which are parallel to the central axis of the camera using a parabolic re ector. Hence, all the rays which are incident on the coffee- lter mirror are re ected in the parallel direction. From the surface normals derived in the previous section, we nd out the direction of the set of the incident rays captured using the proposed system.

(a) (b)

Figure 3.10: (a) Lines in black represents the locus of the intersection of consecutive rays for different radial lines lying on a single face (b) Red and blue curves represents the caustic curve for left and right faces respectively.

Figure 3.11: Geomtery for deriving the direction of incident ray direction.

Consider Fig. 3.11 where $\hat{l}$ is the Incident ray vector on any point P, $\hat{n}$ is the normal vector and $\hat{r}$ is the re ected ray vector. Since the direction of re ected ray and the normal vector is already known, we calculate the direction of the incident ray to calculate the epipolar lines. From $\triangle PP_2P_3$, P_2P_3 is the projection of PP_3 on PP_2. Hence,

$$
\begin{aligned}
P_3P_2 &= PP_2 \quad \hat{l} \\
&= (PP_3 \quad P\hat{P}_2)P\hat{P}_2 \quad \hat{l} \\
&= (l \quad \hat{n})\hat{n} \quad \hat{l}
\end{aligned}
$$

Applying laws of re ection, the triangles $\triangle PP_3P_2$ and $\triangle PP_1P_2$ are congruent. This means,

$$
\begin{aligned}
P_3P_1 &= 2P_3P_2 \\
&= 2((l \quad \hat{n})\hat{n} \quad \hat{l}):
\end{aligned}
$$

Applying vector triangle law in $\triangle P_1P_2P_3$ we get,

$$\begin{aligned} PP_1 &= PP_3 + P_3P_1 \\ &= \hat{I} + 2((I \cdot \hat{n})\hat{n} - \hat{I}) \\ &= 2(I \cdot \hat{n})\hat{n} - \hat{I} \\ \hat{R} &= 2(I \cdot \hat{n})\hat{n} - \hat{I} \end{aligned}$$

Using principle of reversibility of light, one can simply derive

$$\hat{I} = 2(R \cdot \hat{n})\hat{n} - \hat{R}$$

With our setup the $\hat{R}$ is known, $\hat{I}$ becomes the function of r, θ for each point on the surface.

Figure 3.12: Geometry showing the set of the points viewing the straight line X to derive the epipolar constraints.

Let us consider a point in 3D world as shown in Fig. 3.12 defined by (X, Y, Z) which is imaged by a mirror surface at point $P(r, \theta)$, then the incident ray direction $\hat{I}$ that P is viewed by some other mirror surface at location $P^0(r^0, \theta^0)$. The set of such points form an epipolar curve for the point P. Epipolar curve for a point in the left face is found by minimizing the distance between the incident rays from a point in a left face P to every other point in it's right face P^0. Thus, for each θ in the mirror surface, we find the r which intersects the reflected ray from point P such that the triple vector product is zero which means,

$$\frac{|[PP^0, I_{r,\theta}, I_{r^0,\theta^0}]|}{|I_{r,\theta} \times I_{r^0,\theta^0}|} = 0 \tag{3.19}$$

where $I_{r,\theta}$ represents the direction of reflected ray from mirror surface.

Each point is then transformed into the corresponding image coordinate using the dewarping method explained in previous section. Since, the design behaves as a non-central camera, every point has different epipolar constraints. We calculate stereo disparity between the left and right views by finding the correspondences along these epipolar curves using sum of squared differences (SSD). For this, we formulate this problem as energy minimization problem and find solution using (12).

3.3.5 Derivation of Orthographic Projections

We observed that for most practical applications, important information perceived by humans is in the lower part of the world whereas upper part is mostly the sky. In our design, owing to the structure, upper part of the surface has better resolution than the lower regions. Therefore, we keep the coffee lter design inverted as shown in Fig. 3.13. A parabolic re ector is kept above it, both aligned along the same central axis. This parabolic mirror captures the orthographic rays and re ects them at the camera kept at its focus C.

Figure 3.13: Orthographic Projections using the proposed set up which includes the coffee lter mirror and a parabolic re ector

A hole of radius r_{min} is kept in the mirror so that the camera can capture the parallel rays re ected from the parabolic re ector. Also, $OB = R_{max}$ so that it captures all the rays coming from the entire radius of the mirror surface. Let a be the curvature of the parabolic re ector such that the equation of the parabolic re ector can be represented as

$$z = ar^2 \tag{3.20}$$

then $p_0 = aR_{max}^2$. Also, from Fig. 3.13:

$$\tan \quad = \frac{p_1}{R_{max} \quad r_{min}} = \frac{p_2}{r_{min}} \tag{3.21}$$

Since, $\frac{1}{4a}$ is the focal length of the parabolic re ector,

$$p_0 + p_1 + p_2 = \frac{1}{4a} \tag{3.22}$$

Solving Eqns 3.21 and 3.22, we get:

$$aR_{max}^2 + p_1 + p_2 = \frac{1}{4a}$$

$$4a^2R_{max}^2 + 4a(p_1 + p_2) \quad 1 = 0$$

$$(4R_{max}^2)a^2 + 4(p_1 + p_2)a \quad 1 = 0$$

Solving for a, we get,

$$a = \frac{\sqrt{(p_1 + p_2)^2 + R_{max}^2} \quad (p_1 + p_2)}{2R_{max}^2} \tag{3.23}$$

which can be used to nd the surface equations of the parabolic re ector.

3.3.6 Calibration of the Proposed System

In this section, we explain the calibration and dewarping process in detail. As explained earlier that the surface of the mirror is paraboloidal, the resolution is different at different points along each radial line. Also, each captured image depends upon the orientation and viewing angle of the camera. However, for stereo vision to be perceivable, camera's viewing axis must be aligned with the central axis of the device. To calibrate our device, we project structured light binary patterns onto a display surface. These patterns are used to compute a mapping from world coordinates to image coordinates which is used for de-warping the panoramas. We use the approach proposed in (41) and project both normal and inverse binary sequence patterns. These calibration images together will be used to de-warp the captured scene image into left and right eye panoramas as explained in the following steps.

1. Decoding the calibration images: At each pixel in the captured image, we nd the row and column it corresponds to in the de-warped panoramas, by decoding the observed binary sequence from the calibration images at that particular pixel.

2. Finding the correct eye views: For each pixel in the captured scene image we nd out the angle of the radial line it lies on from the center of the image. Each petal subtends an angle at the center. So pixels at angles 0 to $\frac{\pi}{2}$ belong to the left eye views, and those on angles between $\frac{\pi}{2}$ to belong to right eye views.

3. Creating the left and right panoramas: The captured scene image is de-warped into left and right panoramas using the spatial information obtained from step 1 and 2. The upper part of the image is of lesser resolution than the lower one. As a result, some portion of de-warped panorama has holes which can be easily lled by interpolating those regions.

Figure 3.14: Red-Cyan anaglyph panorama obtained by using the proposed set up using POVRay dataset (1).

3.4 Results and Discussions

We have created a prototype using the proposed coffee- lter mirror along with a parabolic re ector as shown in Fig. 3.16. The prototype has 24 petals in total catering to 48 different left and right eye

views. Each face subtends an angle $= 15$ at the center of the device. The total height of the device z_{max} is kept as 5 cm and the radius of the hole $r_{min} = 2.7$cm. We have kept $d = 6.5$ cm which is equal to the average value for human baseline. This optimal petal angle α_{opt} comes out to be 82.5, the optimal values of $R_{max} = 9.77$ cm, $R_{min} = 8.571$ cm and $l = 1.696$ cm. To mitigate the effects of inter-re ections in the adjacent faces, as explained in Section 4.2, we have kept a small angular overlap between two adjacent petals as $= 2$. This means each petal captures with a redundancy $= 15$, since $= 15$.

In order to test our design, we have modeled the coffee lter mirror using POVRay (4), a freely available ray tracing software tool that accurately simulates imaging by tracing rays through a given scene. We have used two 3D scene datasets (19; 1) to demonstrate how the proposed mirror is used to create stereo panoramas (red-cyan anaglyph) as shown in Fig. 3.1c and Fig. 3.14. Our device can also be used to capture dynamic scenes and create 3D stereo videos. As shown in the results, the proposed device is able to capture 103 FOV in vertical direction and 360 in the horizontal direction. The image of the simulated scene is captured using the simulated proposed setup and a virtual camera at a resolution of 5000 x 5000 and is dewarped into two left and right panoramas of resolution 512 x 8192.

Fig. 3.15a shows a stereo depth map of the POVRay scene `average of ce' (19) computed using the coffee lter mirror. The qualitative comparison of the depth map with the ground truth is shown in Fig. 3.15b. More results including anaglyph images, videos of dynamic scenes and stereo depth maps are given in the supplementary material. For a complete 3D immersive experience, we suggest the readers use Google Cardboard (2) for the videos provided in the supplementary material.

The dewarped left and right eye panoramas of two scenes are given in Figures 3.17 and 3.18 along with red-cyan anaglyph stereo.

Fig 3.18 below shows the anaglyph of an indoor scene along with the depth values that are recovered from our camera. Even through the design process is optimized for human consumption, the recovered stereo is highly consistent with the ground truth (See Figures 3.19b and 3.19c).

(a)

(b)

Figure 3.15: (a) Comparison of reconstructed depth as obtained using the proposed set up with the ground truth depth map

(a) (b)

Figure 3.16: (a) A miniaturized model of the proposed design. (b) An example setup of the proposed device along with a consumer cellphone camera.

Results included in the supplementary video for the dynamic scene are adapted from PovRay dataset `Domino' (17). The static scene results are shown both in anaglyphic and binocular formats.

(a)

(b)

(c)

Figure 3.17: (a) Left-eye view (b) Right-eye view (c) Red-cyan stereo image of the Patio scene (19)

.