

Panoramic Stereo Videos with a Single Camera

Rajat Aggarwal* Amrisha Vohra* Anoop M. Namboodiri
 Kohli Center on Intelligent Systems,

International Institute of Information Technology- Hyderabad, India.

{rajat.aggarwal@research, amrisha.vohra@research, anoop@}.iiit.ac.in

Abstract

We present a practical solution for generating 360° stereo panoramic videos using a single camera. Current approaches either use a moving camera that captures 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 panoramic stereo 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 several properties of the design as well as present panoramic stereo and depth estimation results.

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. However, accurate omnidirectional stereo requires the capture of a four-dimensional light-field. Peleg *et al.* [17] showed that 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 [1].

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

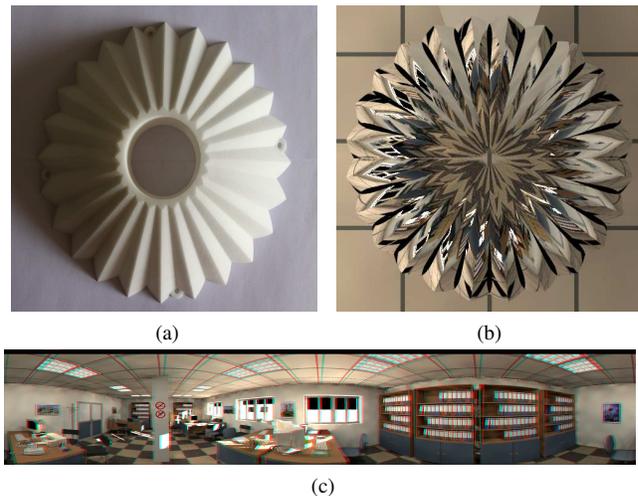


Figure 1: (a) 3D printed model of the proposed *coffee-filter mirror* (b) Image of the proposed mirror placed in a scene (POVRay) (c) Stereoscopic panorama (red-cyan anaglyph) recovered from the image in (b).

one to stand at any location in the world and look around as if they were present in the real environment. However, 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. 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. 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 in-

*Equal Contribution

volve either multiple cameras to capture various perspective directions [12, 1, 5], or a single camera that rotates along a horizontal circle to acquire the images [17, 19, 20]. 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 [13]. 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 [1] or the system proposed by Amini *et al.* [5], 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.
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 $360^\circ \times 270^\circ$ 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.

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. Gluckman *et al.* [12] 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. Kawanishi *et al.* [15] 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 resolution of the omnidirectional images, the disparity remains vertical as before. Lin *et al.* [16] created omnidirectional stereo using two cameras and a conical mirror, which limits the vertical FOV. Yi and Ahuja [23] 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. Peleg *et al.* [17] 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 misperceptions when viewed in 3D as explained by Held and Banks [13]. Couture *et al.* [8, 9] 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.

Richardt *et al.* [20] 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.

Peleg *et al.* [17, 19] 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.

Principle	Camera Name	Device setup	Type of scenes	Perceived Artifacts	Image Resolution
Single Camera based	Omnistereo [17]	Rotating	Static	Mis-alignments, stitching artifacts & Vertical parallax	Medium
	Megastereo [20]	Rotating	Static	No artifacts but extensive run time for SFM	High
	Coffee-filter (proposed)	Fixed	Dynamic	No stitching artifacts and horizontal disparity errors. Very less vertical disparity errors.	Medium
Multiple camera based, needs camera synchronization	Omnistereo, fresnel Lens solution[17]	2 Fixed cameras	Dynamic	Chromatic aberrations, difficult to manufacture and use , Self-occlusions	Low
	Panoramic Stereo Video Textures [8, 9]	2 Rotating Cameras	Dynamic textures	Visible seams and vertical parallax	High
	Google Jump [1]	16 Fixed Cameras	Dynamic	No stitching artifacts but setup is expensive and bulky	High
	Omnipolar [6]	6 Fixed Cameras	Dynamic	Self occlusions , setup is expensive and bulky	High
	Hexagonal Pyramidal Mirrors [15]	12 Fixed cameras	Dynamic	Vertical disparity errors	High

Table 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.

Couture and Roy [6] 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. Tanaka and Tachi [22] also proposed a method to capture omnistereo video sequences. Their rotating optics system consisted of prism sheets, circular or linear polarizing films, and a hyperboloidal mirror. The latest and most effective solution is a simple extension of the idea by Peleg *et al.* [17], 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’ [1]. 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 1 shows the more detailed comparison of our device with other omnistereo devices/approaches.

3. 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*

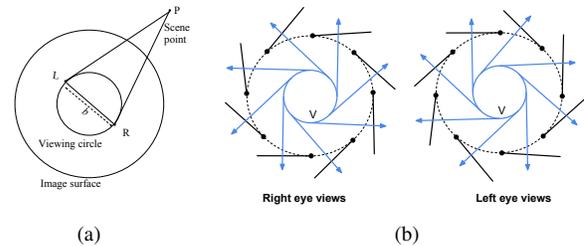


Figure 2: (a) The scene point P viewed by L and R eyes forming a viewing circle with diameter equal to baseline (b) Arrangement of mirrors for capturing right and left views. Tangential rays are captured by placing a mirror normal to the viewing circle.

circle, as shown in Fig. 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 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 Google cardboard [2].

Consider the arrangement of two flat mirrors P_1P_2 and P_2P_3 as shown in Fig. 3b, PB is the 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 ,

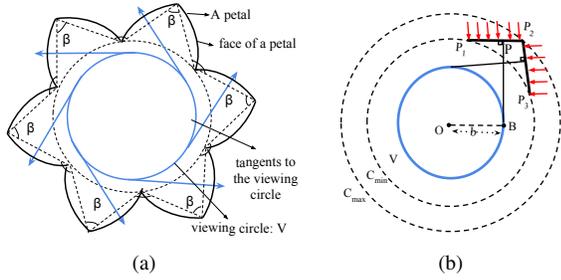


Figure 3: (a) A general horizontal cross-section of the proposed coffee filter mirror. (b) Combination of the two arrangements as shown in Fig. 2b to capture both eye views in a single design.

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 [17]. Each such mirror is referred to as a *face*. Multiple faces, when arranged in the configuration shown in Fig. 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. 3b.

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. 2b in a single design, as shown in Fig. 3b, 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. 3a 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.

4. 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.

4.1. 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. 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 flat mirrors are arranged in a circular manner, as shown in Fig. 4a. Hence, certain areas of the world will be left unseen by the camera and will be missing from the captured images.

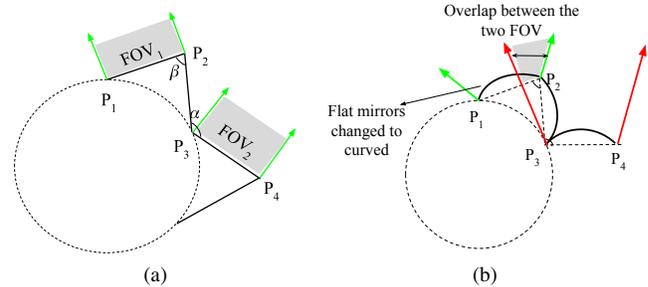


Figure 4: (a) Flat mirrors have limited FOV, causing blind spots (b) Curved mirrors increase the FOV and the overlap between consecutive faces of an eye.

Field of view of each face is increased by opting for a horizontal curvature for the face as shown in Fig. 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-reflections between the neighboring faces.

4.2. Inter reflections

Due to the curvature in the faces in the proposed design, the FOV of the adjacent faces aligned at an obtuse angle β , overlap. As shown in Fig. 5, a ray originating from the camera that strikes the face P_2P_3 , may get reflected multiple times because of the FOV overlap between faces P_2P_3 and P_3P_4 . These inter-reflections will cause mis-captured information at the camera sensor, thus decreasing the resolution of the final de-warped panoramas. Hence, an optimal amount of overlap is kept such that inter-reflections are kept to a minimum, while also solving the problem of missing regions as explained in Section 4.1.

4.3. 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. For better quality and perception

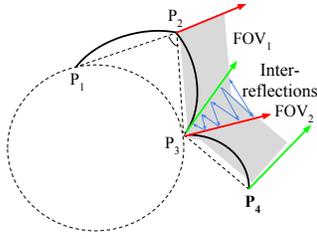


Figure 5: Inter-reflections cause wrong world points to be captured due to FOV overlap between each pair of left and right eye view mirrors

of the de-warped images, it is important that the resolution is uniform both horizontally and vertically. In a flat mirror, vertical FOV is equal to the height of the mirror and hence the resolution remains constant. Our proposed design is such that at each horizontal cross section the faces can be confined within a circle. The radii of these circles increase from 0 to R_{max} . As we need uniform resolution, we see that a parabolic curvature in the vertical direction is more desirable as compared to a hyperbola or straight line (see Fig. 6)¹. There have been several approaches proposed to attain uniform resolution on some planes or uniform angular resolutions [7, 11, 14]. However, their work focuses mainly on creating sensors with uniform resolution along the radial line, whereas our mirror is designed to capture stereoscopic views. Therefore, to simplify the set up and the dewarping process, at present, we seek optimality in terms of the size, complexity, and minimizing inter-reflections. Creating a device that can optimally capture stereoscopic views with uniform radial resolution is an aspect that needs to be explored in future.

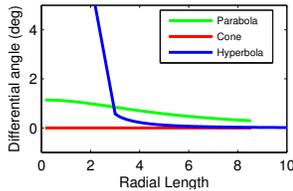


Figure 6: Uniformity of vertical resolution in terms of differential angle of the incident rays along the radial length for a cone, parabola and hyperbola.

5. The Mirror Surface

We now derive the equation of the mirror surface. Multiple factors can be varied to make the device adaptive to specific applications. We derive the expressions for only one petal APB as shown in the Fig. 7a. and the same expressions hold for all 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 β . Each petal subtends an angle θ at the center, where $\theta = \frac{2\pi}{n}$. Hence, we get n views each

¹Please see Supplementary Material for more mathematical details

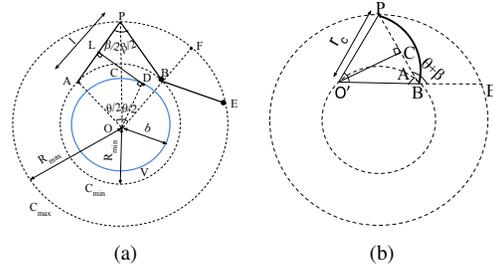


Figure 7: Geometry of the petal surface used to obtain optimal design parameters.

for left and right eye. The design of the mirror is symmetrical, and all the petals are of same size and dimensions. The distance between the two extreme points of a mirror surface i.e AP as shown in Fig. 7a is referred as *petal length*, denoted by l . 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. 7a, $OA = R_{min}$ and $OP = R_{max}$. From $\triangle OAP$ and $\triangle OBP$, by sine rule we get the relations as,

$$\frac{l}{\sin(\frac{\theta}{2})} = \frac{R_{max}}{\sin(\pi - \frac{(\theta + \beta)}{2})} = \frac{R_{min}}{\sin(\frac{\beta}{2})}$$

Note that $\angle APO = \angle BPO = \frac{\beta}{2}$ and $\angle AOP = \angle POB = \frac{\theta}{2}$, because each face is symmetrical and oriented at equal separation. $\therefore \angle OAP = \pi - \frac{\theta}{2} - \frac{\beta}{2}$. LD is the perpendicular bisector of the chord AP and is tangent to the viewing circle V . $OD = b$ is the radius of the viewing circle. In $\triangle OCD$ and $\triangle CLP$, we get $LP = CP \cos(\frac{\beta}{2})$, implies $CP = LP \sec(\frac{\beta}{2}) = \frac{l}{2} \sec(\frac{\beta}{2})$.

In $\triangle PLC$, $\angle LCP = \frac{\pi}{2} - \frac{\beta}{2}$. $\angle OCD = \angle LCP$, being vertically opposite angles. Hence, $\angle COD = \frac{\pi}{2} - \angle OCD = \frac{\beta}{2}$. Also, $OC + CP = R_{max}$, which gives, $OC = R_{max} - \frac{l}{2} \sec(\frac{\beta}{2})$

In $\triangle OCD$, $\frac{OD}{OC} = \cos \frac{\beta}{2}$, which gives $OC = b \sec(\frac{\beta}{2})$.

Using this we get, $R_{max} - \frac{l}{2} \sec(\frac{\beta}{2}) = b \sec(\frac{\beta}{2})$ and $R_{max} = (b + \frac{l}{2}) \sec(\frac{\beta}{2})$. Combining all these, we get:

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

$$R_{min} = R_{max} \frac{\sin(\frac{\beta}{2})}{\sin(\frac{\theta + \beta}{2})} \quad (2)$$

$$l = R_{max} \frac{\sin(\frac{\theta}{2})}{\sin(\frac{\theta + \beta}{2})} \quad (3)$$

5.1. Optimizing the design parameters

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

filter mirror i.e. R_{max} is dependent upon β . At petal angle, $\beta_{opt} = \frac{\pi-\theta}{2}$, R_{max} is minimum, and hence we get the minimum size of the device.

In Fig. 7a, Let $\angle PBE$ be α , the angle between two petals. In $\triangle OBP$, $\angle OPB = \frac{\beta}{2}$, $\angle POB = \frac{\theta}{2}$ and $\angle OBP = \pi - \frac{\theta+\beta}{2}$. Therefore, $\angle PBF = \pi - (\pi - \frac{\theta+\beta}{2}) = \frac{\theta+\beta}{2}$. $\angle PBE = 2\angle PBF$, which means $\alpha = \theta + \beta$. Consider Fig. 7b where O' is the center of curvature of the face PB . PO' and $O'B$ are the radii of curvature i.e r_c and $\angle PO'B = 2\gamma$ is the angle subtended by each face at the center of curvature. In $\triangle PO'B$, $\angle A = \pi - (\theta + \beta)$, which implies, $\gamma = \frac{\pi}{2} - \angle A = (\theta + \beta) - \frac{\pi}{2}$. In order to have small mirror size, $\gamma_{opt} = (\theta + \beta_{opt}) - \frac{\pi}{2}$. Therefore, the optimal horizontal angular field of view $\gamma_{opt} = \frac{\theta}{2}$ and is independent of the obtuse angle $\angle PBE$ between nearby faces.

$O'C$ is the perpendicular bisector of PB , $CB = \frac{l}{2}$. In $\triangle O'CB$, $\frac{l}{2r_c} = \sin \gamma$. Radius of curvature r_c can be optimized by using the optimal value of γ . Therefore, $r_c = \frac{l}{2\sin \frac{\theta}{2}}$, 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-reflections, as explained in Section 4.2, it is important to collect the maximum scene information in the captured image. Each face covers $\frac{2\theta}{n}$ angular FOV, thus a total of n such faces for each view covers complete 2π 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 $O'P$ 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 $\frac{\pi+\theta}{2}$. The amount of inter-reflections depends upon the angle between two consecutive petals, α , which depends upon the sampling angle of the mirror $\frac{2\pi}{n}$. Ideally, the amount of inter-reflections 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-reflections, we introduce a small angle δ such that the angle of curvature becomes $2\gamma + \delta$. This makes sure some overlap is there, so that some redundant information is captured, which can be used while dewarping. However, the value of δ is kept sufficiently low, such that inter-reflections are also reduced to a huge extent.

5.2. Resultant Mirror Surface

In this section, we obtain the surface equations of the proposed coffee filter mirror in terms of polar coordinates ϕ and r . As explained earlier, the surface of the *coffee filter mirror* is such that each face is paraboloidal vertically and at each horizontal cross section the faces can be confined within a circle. The radii of these circles increase from 0 to R_{max} . Let us consider the central axis of the mirror to be the z axis. Then the surface equation can be written as a

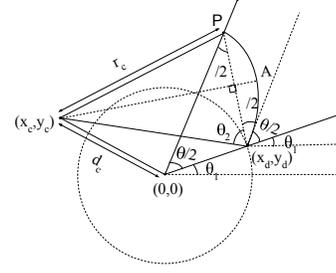


Figure 8: Parameters of the mirror petal.

function of x and y axis:

$$z = f(x, y) = m_\phi(x^2 + y^2), \quad (4)$$

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

$$z = m_\phi r^2 \quad (5)$$

Eqn 5 represents the petal surface of our custom designed mirror centered around origin. Consider the uppermost and widest cross section of the mirror at $z = z_{max}$, such that $z_{max} = m_\phi r_1^2$, $m_\phi = \frac{z_{max}}{r_1^2}$. Let (x_c, y_c) be the center of the circle of curvature of a face of a petal and (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 5, we get $m_\phi = \frac{z_{max}}{k^2 r^2}$ which implies, $z = \frac{z_{max}}{k^2}$. It is to be noted 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:

$$(x_d - x_c)^2 + (y_d - y_c)^2 = (kx - x_c)^2 + (ky - y_c)^2 = r_c^2 \quad (6)$$

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

Since, $m_\phi = z_{max}/k^2 r^2$,

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

as $m_\phi = z_{max}/k^2 r^2$. Also, it can be derived from the Fig. 8, $(x_c, y_c) = (x_d, y_d) + r_c(\cos(\theta_1 + \frac{\theta}{2} + \frac{\beta}{2} + \theta_2), \sin(\theta_1 + \frac{\theta}{2} + \frac{\beta}{2} + \theta_2))$, where $\theta_2 = \tan^{-1}(\frac{2r_c}{l})$.

From this and Eqn 5 we get,

$$z = z_{max} \left(\frac{r^2}{(xx_c + yy_c) + \sqrt{(xx_c + yy_c)^2 - r^2(d_c^2 - r_c^2)}} \right)^2 \quad (7)$$

Therefore, Eqn 7 gives the equation of the paraboloidal surface of the mirror. Note that the slope m_ϕ at every point is a function of r . Any incident ray \mathbf{I} coming from the world point falls on the mirror surface at the point $P(r, \phi)$ and is reflected in the direction $\hat{\mathbf{R}} = 2(\hat{\mathbf{I}} \cdot \mathbf{n}(\hat{\mathbf{r}}, \phi))\mathbf{n}(\hat{\mathbf{r}}, \phi) - \hat{\mathbf{I}}$, where $\mathbf{n}(\hat{\mathbf{r}}, \phi)$ is the direction of normal vector at P . \mathbf{n} is the

normal vector to the tangent plane containing the tangents in both horizontal and the vertical direction. For a horizontal plane, the direction of normal vector \mathbf{n}_h at point (r, ϕ) is given by $[y_c - r \sin \phi \quad r \cos \phi - x_c \quad 0]^T$ where (x_c, y_c) are the center of curvature. Similarly, the normal vector \mathbf{n}_v in vertical direction is obtained by $[\frac{dx}{dr} \quad \frac{dy}{dr} \quad \frac{dz}{dr}]^T$ which is obtained as $[\cos \phi \quad \sin \phi \quad 2m_\phi r]^T$. Hence, the normal vector to the tangent plane at point (r, ϕ) is obtained by $\mathbf{n} = \hat{\mathbf{n}}_h \times \hat{\mathbf{n}}_v$.

5.3. Calibration and De-warping

Images captured by the camera using the coffee filter mirror are of the form as shown in Fig. 1b. Captured images varies with 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 use a generic non-parametric camera calibration method as proposed by Posdamer and Altschuler [18]. We project structured light binary patterns onto a display surface and project both normal and inverse binary sequence patterns. The obtained calibration images together will be used to compute a mapping from 3D world coordinates to 2D image coordinates which is used for de-warping the captured scene image into left and right eye panoramas.

6. Epipolar Geometry and Stereo

To find the center of projections of the proposed coffee-filter mirror, we find the trajectory of the viewpoints by finding the intersection of the reflected rays for every pair of adjacent points. Let a point lying on a radial line with single ϕ_j is denoted by p_{i,ϕ_j} . Then the reflected rays for each pair of adjacent points p_{i,ϕ_j} and p_{i+1,ϕ_j} intersect at the c_{i,ϕ_j} . The set of all such points form a line as shown in Fig. 10a. $L_{\phi_j} = \{c_{i,\phi_j}\} \forall i$. We take the average of these intersection points c'_{ϕ_j} , to find the average viewing circle for complete 360° view. Locus of c'_{ϕ_j} for ϕ for a single face is a straight line parallel to central axis of the camera. Let c''_k be the average center for a single face. Locus of c''_k for all n faces is viewing circle, as shown in Fig. 10b. In order to have a mirror with diameter of the viewing circle equal to the baseline of human eye, 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 avoided by interpolation during de-warping. The combination of mirrors and a conventional camera behave as a non-central catadioptric camera. For details on such classification, readers are directed to [21].

Let us consider a point in 3D world defined by (X, Y, Z)

which is imaged by a mirror surface at point $P(r, \phi)$, then the ray coming from \mathbf{X} to P is viewed by some other mirror surface at location $P'(r', \phi')$. The set of such points form an epipolar curve for the point P . Each point is then transformed into the corresponding image coordinate using the calibration explained in previous section. Epipolar curve for a point in the left face is found by minimizing the distance between the reflected rays from a point in a left face P to every other point in its right face P' . Thus, for each ϕ in the mirror surface, we find the r_ϕ which intersects the reflected ray from point P such that,

$$\frac{|[\mathbf{PP}', \mathbf{I}_{r,\phi}, \mathbf{I}_{r',\phi'}]|}{|\mathbf{I}_{r,\phi} \times \mathbf{I}_{r',\phi'}|} = 0 \quad (8)$$

where $\mathbf{I}_{r,\phi}$ represents the direction of reflected ray from mirror surface. 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.

7. Results and Discussions

In order to test our design, we have modeled the coffee filter 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 [10, 3] to demonstrate how the proposed mirror is used to create stereo panoramas (red-cyan anaglyph) as shown in Fig. 1c and Fig. 9. Our mirror can also be used to capture dynamic scenes and create 3D stereo videos. As shown in the results, the proposed coffee filter mirror 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 5000×5000 and is de-warped into two left and right panoramas of resolution 512×8192 .

Fig. 11a shows a stereo depth map of the POVRay scene 'average office' [10] computed using the coffee filter mirror. The qualitative comparison of the depth map with the ground truth is shown in Fig. 11b. Fig. 12a and Fig. 12b shows left and right panoramas respectively created using 12 petals of the initial manufactured prototype of the *coffee-filter mirror*. Additional results including anaglyph images, videos of dynamic scenes and stereo depth maps may be found at the project website ².

We have created a physical prototype of the proposed mirror (Fig. 13(b)) with 24 petals, capturing 48 different left or right eye views. Each face subtends an angle $\theta = 15^\circ$ at the center of the coffee filter mirror. The total height of the mirror z_{max} is kept as 5 cm and the radius of the hole $r_{min} = 2.7$ cm. We have kept $b = 6.5$ cm which is equal to the average value for human baseline. This optimal petal angle β_{opt}

² <http://cvit.iiit.ac.in/research/projects/panoStereo/>

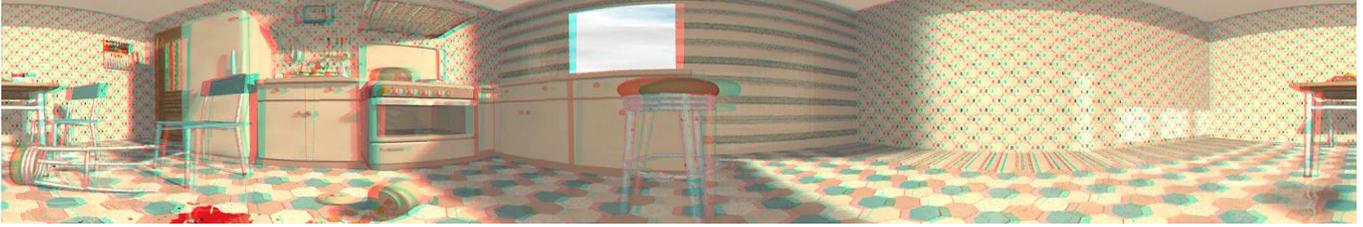


Figure 9: Red-Cyan anaglyph panorama obtained by using the proposed set up using POVray dataset [3].

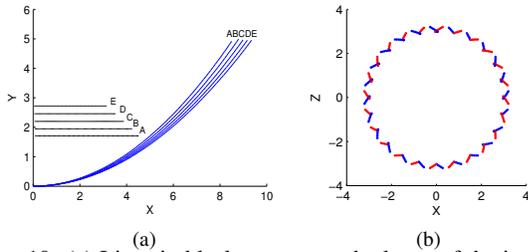


Figure 10: (a) Lines in black represents the locus of the intersection of consecutive rays for different radial lines lying on a single face (b) Red and blue curves represent the caustic curve for left and right faces respectively.

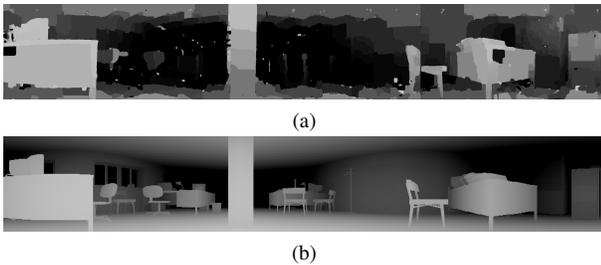


Figure 11: (a) Comparison of reconstructed depth as obtained using the proposed set up with the ground truth depth map.

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-reflections in the adjacent faces, as explained in Section 4.2, we introduced a small angular overlap between two adjacent petals as $\delta = 2^\circ$. This means each petal captures with a redundancy of $1/15^\circ$, since $\theta = 15^\circ$. The resulting panoramas are shown in Fig. 12. Degradations in quality of images are due to imperfections of the mirror surface that was created by SLA printing.

In our design, the horizontal shift in camera centers between faces is very small ($\approx 2mm$) reducing the blind spot to within 5 cm of the device. The shift decreases along the height of the mirror. To simplify the post-capture setup, we avoided depth computation step at dewarping. This retains a small disparity difference from the top of the image to its bottom. This difference is visible on close inspection and may be corrected with depth estimation. As our goal was human perceivable stereo and the artifacts are imperceivable while using stereo glasses or HMD, we avoided this step. There are radial and vertical shifts in camera cen-

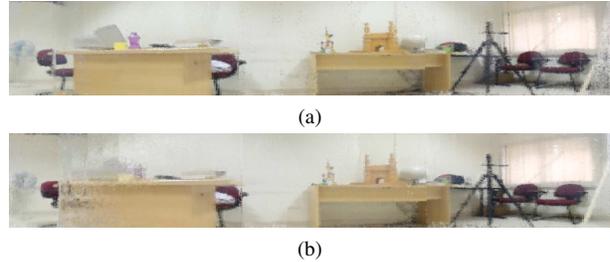


Figure 12: Left and right panoramas extracted from 12 petals of an initial prototype of the mirror shown in Fig 13(b).

ters between faces that causes the artifacts in Fig. 9. However, the ratio of the vertical shift to the horizontal shift is very small and does not create any visible artifacts beyond a depth of d . The value of d turns out to be 70 cm when the rendering cylinder is kept at 200 cm, for images captured at 5000×5000 resolution. These could also be corrected like other approaches by depth estimation at the dewarping stage.

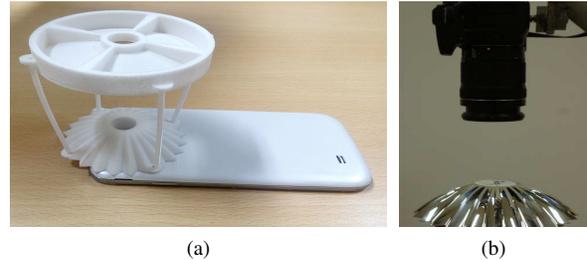


Figure 13: (a) Mockup of proposed device attached to a consumer cellphone camera (b) Set up created using initial manufactured prototype of *coffee filter mirror*.

8. Conclusions

We have proposed a simple practical solution to capturing 360° stereo panoramas using a single digital camera for immersive human experience. As the resolution of sensors increase, the quality of the panoramas also increase. We derived the optimal parameters of the design and experimental results show that we can avoid most visual artifacts in the panoramas. While designed with human consumption in mind, the stereo pairs could also be used for depth estimation.

References

- [1] Google Jump, <https://www.google.com/get/cardboard/jump/>. 1, 2, 3
- [2] Google Cardboard, <https://www.google.com/get/cardboard/>. 3
- [3] <http://www.ignorancia.org/en/index.php?page=Childhood>. 7, 8
- [4] POV-Ray, <http://www.povray.org/>. 7
- [5] A. S. Amini, M. Varshosaz, and M. Saadatseresht. Evaluating a new stereo panorama system based on stereo cameras. *International Journal of Scientific Research in Inventions and New Ideas*, 2, 2014. 2
- [6] V. Chapdelaine-Couture and S. Roy. The omnipolar camera: A new approach to stereo immersive capture. In *Computational Photography, IEEE International Conference on*, 2013. 3
- [7] T. L. Conroy and J. B. Moore. Resolution invariant surfaces for panoramic vision systems. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 392–397. IEEE, 1999. 5
- [8] V. Couture, M. S. Langer, and S. Roy. Panoramic stereo video textures. In *Computer Vision, IEEE International Conference on*, 2011. 2, 3
- [9] V. C. Couture, M. S. Langer, and S. Roy. Omnistereo video textures without ghosting. In *3D Vision, International Conference on*, 2013. 2, 3
- [10] F. Devernay and S. Pujades. Focus mismatch detection in stereoscopic content. In *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012. 7
- [11] S. Gächter, T. Pajdla, and B. Micusik. Mirror design for an omnidirectional camera with a space variant imager. In *IEEE Workshop on Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing*, pages 99–105, 2001. 5
- [12] J. Gluckman, S. K. Nayar, and K. J. Thoresz. Real-time omnidirectional and panoramic stereo. In *Proceedings of Image Understanding Workshop*, 1998. 2
- [13] R. T. Held and M. S. Banks. Misperceptions in stereoscopic displays: a vision science perspective. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, 2008. 2
- [14] R. A. Hicks and R. K. Perline. Equi-areal catadioptric sensors. In *null*, page 13. IEEE, 2002. 5
- [15] T. Kawanishi, K. Yamazawa, H. Iwasa, H. Takemura, and N. Yokoya. Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors. In *Pattern Recognition. Fourteenth International Conference on*, 1998. 2, 3
- [16] S. Lin and R. Bajcsy. High resolution catadioptric omnidirectional stereo sensor for robot vision. In *Robotics and Automation. IEEE International Conference on*, 2003. 2
- [17] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistereo: Panoramic stereo imaging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23:279–290, 2001. 1, 2, 3, 4
- [18] J. Posdamer and M. Altschuler. Surface measurement by space-encoded projected beam systems. *Computer graphics and image processing*, 18:1–17, 1982. 7
- [19] Y. Pritch, M. Ben-Ezra, and S. Peleg. Optics for omnistereo imaging. In *Foundations of Image Understanding*, pages 447–467. 2001. 2
- [20] C. Richardt, Y. Pritch, H. Zimmer, and A. Sorkine-Hornung. Megastereo: Constructing high-resolution stereo panoramas. In *Computer Vision and Pattern Recognition, IEEE Conference on*, 2013. 2, 3
- [21] P. Sturm, S. Ramalingam, and S. Lodha. On calibration, structure from motion and multi-view geometry for generic camera models. In *Imaging Beyond the Pinhole Camera*. 2006. 7
- [22] K. Tanaka and S. Tachi. Tornado: Omnistereo video imaging with rotating optics. *Visualization and Computer Graphics, IEEE Transactions on*, 11(6):614–625, 2005. 3
- [23] S. Yi and N. Ahuja. An omnidirectional stereo vision system using a single camera. In *Pattern Recognition. Eighteenth International Conference on*, 2006. 2