# A High-Resolution Panoramic Camera

**Hong Hua[1] and Narendra Ahuja[1, 2]**
*Beckman Institute[1], Department of Electrical and Computer Engineering[2]*
*University of Illinois at Urbana-Champaign, Urbana, IL, 61801*
*Email: Honghua|n-ahuja@uiuc.edu*

**Abstract**

*Wide field of view (FOV) and high resolution are two desirable properties in many vision-based applications such as tele-conferencing, surveillance, and robot navigation. In some applications such as 3D reconstruction and rendering, it is also desired that all viewing directions share a single viewpoint, the entire FOV be imaged simultaneously, in real-time, and the depth of field be large. In this paper, we review such a panoramic camera proposed by Nalwa in 1996 that uses reflections off planar mirrors to achieve the first four of the aforementioned capabilities. He uses a single mirror pyramid (SMP) and a number of cameras that point to the individual pyramid faces. Together the cameras yield a visual field having a width of 360 degrees and a height same as that of the individual cameras. We propose a double mirror-pyramid (DMP) design that still achieves a 360-degree FOV horizontally but doubles the vertical FOV. It retains the other three capabilities namely high resolution, a single apparent viewpoint across the entire FOV, and real-time panoramic capture. We specify the visual field mapping from the scene to the sensor realized by the proposed camera. Finally, an implementation of the proposed DMP design is described and examples of preliminary panoramic images obtained are included.*

## 1. Introduction

A sensor with a wide field of view (FOV) and high resolution is highly desirable in many applications such as tele-conferencing, surveillance, and robot navigation [1]. In addition, a single viewpoint for all viewing directions, a large depth-of-field (omni-focus), and real-time acquisition are desired in some imaging applications (e.g. 3D reconstruction and rendering) [2, 16, 24]. The FOV of a conventional digital camera is limited by the size of the sensor and the focal length of the lens. For example, a typical 16mm lens with 2/3" CCD sensor has a $30^\circ \times 23^\circ$ FOV. The number of pixels on the sensor ($640 \times 480$ for NTSC camera) determines the resolution. The depth-of-field is limited and is determined by various imaging parameters such as aperture, focal length, and the scene location of the object [12]. Many efforts have been made, which have succeeded in achieving various subsets of these properties (wide FOV, high resolution, large depth-of-field, a single viewpoint, and real-time acquisition). A summary of these methods is presented in the next paragraph. One of these efforts [16], to which the present work is most closely related, uses a right mirror-pyramid, and as many cameras as the number of pyramid faces, each located and oriented to capture the part of the scene reflected off one of the faces (See Fig. 1). Images from the individual cameras are concatenated to yield a 360-degree wide panoramic FOV whose height is the same as that of the FOV of the individual cameras. This paper presents a design that has a 360-degree wide FOV as in [16], but uses two right mirror-pyramids to double the vertical FOV. The taller FOV comes at the cost of twice as many pyramids and cameras, additional complexity in camera geometry, and processing of individual camera images to obtain the single, larger, panoramic image.

Before we describe our design, we will summarize some of the past work on panoramic and omni-directional image acquisition. These methods fall into two categories: dioptric methods, where only refractive elements (lenses) are employed, and catadioptric methods, where a combination of reflective and refractive components is used. Typical dioptric systems include: the camera cluster method where each camera points in a different direction and together the cameras cover all different directions [1]; the fisheye method where a single camera acquires a wide FOV image through a fisheye lens [15, 23, 27]; and the rotating camera method where a conventional camera [21] pans to generate mosaics, or a camera with a non-frontal, tilted sensor [12, 13, 14] pans around its viewpoint to acquire panoramic omni-focused images. The catadioptric methods include: sensors in which a single camera captures the scene as reflected off a single non-planar mirror [4, 5, 7, 8, 9, 18, 20, 24, 26], or sensors in which multiple cameras image the scene as reflected off the faces of a planar right mirror-pyramid [16, 11]. The dioptric camera clusters achieve good resolution across a wide FOV at video rate. However, typically the cameras in the cluster do not share a unique

0-7695-1272-0/01 $10.00 © 2001 IEEE

viewpoint so that there may be uncovered space between adjacent cameras, and therefore, it is difficult or even impossible to seamlessly combine individual images to form a panoramic view without image blending. The sensors with fisheye lens are able to deliver large FOV images at video rate, but suffer from low resolution, irreversible distortion for close-by objects, and non-unique viewpoints for different portions of the FOV. The rotating cameras deliver high-resolution wide FOV via panning, as well as omni-focus when used in conjunction with non-frontal imaging, but they have limited vertical FOV. Furthermore, because they sequentially capture different parts of the FOV, moving objects may be imaged incorrectly. The sensors that use a parabolic- or a hyperbolic-mirror to map an omni-directional view onto a single sensor are able to achieve a single viewpoint at video rate, but the resolution of the acquired image is limited to that of the sensor used, and further, it is greatly reduced in the peripheral fields. Analogous to the dioptric case, this resolution problem can be resolved by replacing the simultaneous imaging of the entire FOV with panning and sequential imaging of its parts, followed by mosaicing the images [6, 19, 22], but at the expense of video rate.

The use of the mirror pyramid, proposed by Nalwa [16] and mentioned earlier, achieves high resolution across a wide FOV at video-rate and with a single viewpoint for all directions across the FOV. In the next section, we first review Nalwa's panoramic camera system and then describe our design that doubles the vertical FOV in Sec. 3.

## 2. Panoramic imaging using a single mirror pyramid (SMP)

One of the major problems in designing an omni-directional sensor using a pinhole camera cluster in a straightforward manner [16] is to co-locate the multiple pinholes so that adjacent cameras cover contiguous FOV without obstructing the view of other cameras or itself. Nalwa [16] instead used planar mirrors to co-locate multiple pinholes [10] for panoramic imaging. Fig. 1a illustrates this through a pair of planar mirrors and two associated cameras, $C_1$ and $C_2$, positioned such that the mirror images of their projection centers coincide at a point $C$. Point $C$ becomes the re-located common viewpoint of the two cameras.

Nalwa proposed an n-sided right mirror-pyramid with a pinhole camera associated with each face such that the mirror image of every pinhole lies at the same location in space [16] as in Fig.1b. He reported an implementation using a 4-sided right pyramid and 4 pinhole-cameras. The pyramid stands on its horizontal base. Each triangular face forms a 45-degree angle with the base. The pinholes are positioned in the horizontal plane that contains the pyramid's vertex such that each pinhole is equidistant from the vertex and the mirror images of all the pinholes coincide at a single point on the axis of the pyramid. The cameras are pointed vertically downward at the pyramid faces, effectively viewing the world horizontally outward from the common virtual pinhole $C$ [16]. As seen in Fig. 1b, this common virtual pinhole is at a location on the pyramid axis whose vertical distance from the pyramid apex is determined by the horizontal distance of the actual camera pinholes from the pyramid apex. The virtual optical axes of the cameras are horizontal, all contained in a plane parallel to the pyramid base and intersecting at the common virtual pinhole. The mapping from the scene points to the image points is as usual, e.g., a rectangular planar object perpendicular to the pyramid base is imaged as a rectangle. Kawanishi et al. [11] used two such sensors to form a stereo pair. Each sensor uses a hexagonal pyramid and six cameras. The two pyramids, and therefore their common virtual pinholes, are separated vertically along their common axis by an amount equal to the desired stereo baseline. Nalwa suggested using the two pyramids back to back, with their bases coinciding, for such stereo viewing [17].

The vertical dimension of the panoramic FOV in each of the aforementioned cases is the same as that of each of the cameras used – only their horizontal FOV's are concatenated to obtain a wider, panoramic view. In the next section, we present an approach that doubles the vertical FOV while retaining the single-viewpoint and other characteristics of the panoramic image.

## 3. Proposed double mirror-pyramid (DMP) panoramic camera

In this section, we describe our proposed new design that uses a double mirror-pyramid, formed by joining two

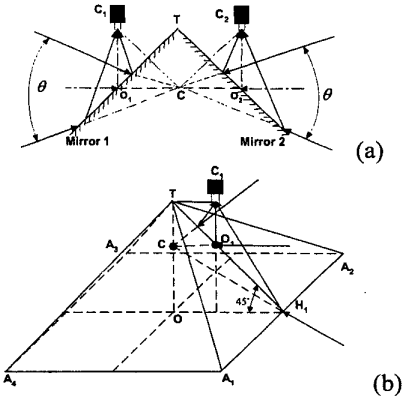



**Fig. 1: Using planar mirrors to co-locate viewpoints $C_1$ and $C_2$ of two different cameras at point $C$.**

mirror-pyramids such that their bases coincide (Fig. 2b), for panoramic viewing. The faces in each mirror-pyramid form an angle $\alpha$ with the base, which determines the maximum vertical FOV of a single mirror-pyramid. The base edge of each face subtends an angle $\gamma$ within the base plane, which determines the horizontal FOV covered by each camera, and is given by the number of pyramid faces. We show how the use of such a double mirror-pyramid facilitates adding another layer of cameras parallel to the single layer present in the single mirror-pyramid (SMP) system. With the second layer, the FOV of any arbitrary camera gets extended not only horizontally, by the two cameras associated with adjacent faces in the SMP system, but also vertically, by the camera associated with the adjacent face in the other pyramid. The resulting double mirror-pyramid (DMP) system thus doubles the vertical FOV while preserving the ability to acquire panoramic high-resolution images from an apparent single viewpoint at video rate.

To achieve the vertical contiguity in the panoramic image requires that both camera layers share a common virtual viewpoint. Clearly this is not possible by simply

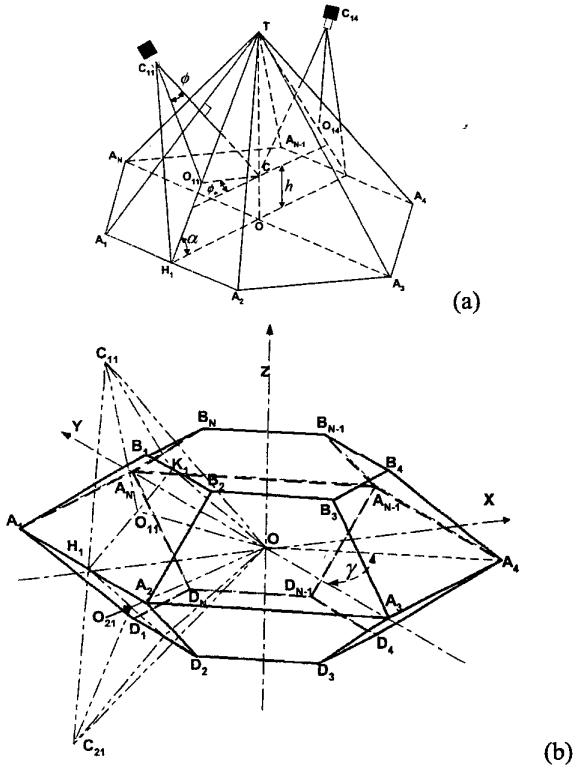



**Fig. 2: Comparison of mirror-pyramid panoramic cameras. (a) The SMP design consists of a single mirror-pyramid and a single horizontal layer of cameras; (b) The proposed DMP camera consists of a double truncated mirror-pyramid and two horizontal layers of cameras.**

replicating the SMP system vertically, because the common virtual viewpoint of each of the two SMP systems, although along the common axis of the two pyramids, is away from their common base. To achieve a common viewpoint, these two virtual viewpoints must be relocated to coincide, and the most direct way of achieving this is to relocate them to the center of the common base of the pyramids.

To see how this could be achieved, consider Fig. 2a that shows an arbitrary single mirror-pyramid and the associated cameras, with the camera locations and orientations being controllable to relocate the common virtual viewpoint. The right mirror-pyramid $TA_1A_2...A_N$ is comprised of $N$ identical triangular planar mirror surfaces $(TA_1A_2,\ TA_2A_3,\ ...,TA_NA_1)$. These mirror surfaces form an identical angle $\alpha$ with the base polygon $A_1A_2...A_N$. A camera cluster $C_1$ of $N$ cameras $(C_{11},C_{12},...C_{1N})$ is placed so that each camera points at a different face of the pyramid, and all cameras share the same virtual viewpoint $C$ on the axis $OT$ of the pyramid. The cameras appear as if they are all at point C but pointing in different directions. The offset of point $C$ from base $A_1A_2...A_N$ is $h$. Unlike the SMP system of [16], we make the cameras $(C_{11},C_{12},...C_{1N})$ point towards the locations $(O_{11},O_{12},...O_{1N})$, respectively, on the pyramid faces. The lines connecting points $(O_{11},O_{12},...O_{1N})$ with the viewpoint $(C_{11},C_{12},...C_{1N})$ are the optical axes of the corresponding cameras, and the lines connecting points $(O_{11},O_{12},...O_{1N})$ with the common virtual viewpoint $C$ are the corresponding virtual optical axes. The optical axis of each camera makes an angle $\phi$ with the normal of its corresponding pyramid face, or equivalently, each virtual optical axis makes an angle of $\phi_o$ with the base. The offset $h$ and tilt angle $\phi_o$ are the parameters whose values determine the relative locations of the two virtual viewpoints, and the angles formed by the two virtual viewing directions with the base.

There are three distinct configurations to consider. (1) $\phi_o = 0$ or $\phi = 90-\alpha$ and $h \neq 0$: This is the same configuration as used by Nalwa and Kawanishi for $\alpha = 45°$, provided the sensor and pyramid geometry, focal length, and the value of $h$ are such that the entire vertical dimension of the sensor is filled with the SMP FOV (Fig. 3a). The virtual optical axes are parallel to but are not contained in the base plane. (2) $\phi_o = 0$ or $\phi = 90-\alpha$, and $h = 0$: Here all virtual cameras face outwards from the common virtual pinhole, which is located at the center of the pyramid base. This is the same configuration as Nalwa's except that the virtual optical axes are contained in the base plane and bisect the

base edges. This is not a desirable solution since the lower half of the sensor does not receive any reflected light from its pyramid facet and therefore is unused (Fig. 3b). (3) $\phi_o \neq 0$ or $\phi < 90 - \alpha$ and $h = 0$ : Unlike cases (1) and (2), in this case the different optical axes are not coplanar. Consequently, vertical planar objects, although equidistant from the pyramid axis, are foreshortened in the vertical direction (Fig. 3c). However, this distortion, also called keystone distortion, can be corrected by re-projecting each acquired image onto a virtual vertical sensor plane.

Configuration (3), along with a double mirror-pyramid, leads to the proposed DMP design. A DMP is formed by stacking two truncated right mirror-pyramids, $A_1A_2...A_NB_1B_2...B_N$ and $A_1A_2...A_ND_1D_2...D_N$, back to back so their bases coincide (Fig. 2(b)). Each of the two truncated pyramids has the same geometry as the single pyramid operating under case (3) just described. The tilt angle $\theta_0$ is chosen to be $\theta_0 = \theta_V / 2$, where $\theta_V$ is the vertical FOV of each camera. This design achieves a 360-degree FOV horizontally and double the individual camera FOV vertically, and also preserves the SMP's ability for acquiring the entire panoramic image in high resolution from a single apparent viewpoint.

In the next section, we obtain the relationships between the parameters of the acquired panoramic image and those of the individual cameras and the imaging geometry.

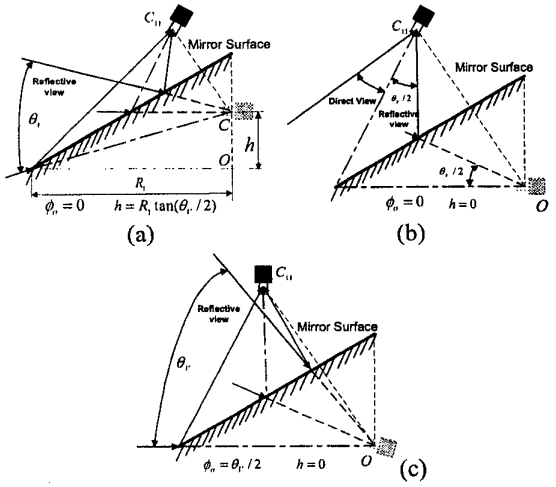



**Fig. 3: The position and orientation of a camera relative to the mirror facet affect the reflective FOV captured by the sensor: (a) when $\phi_o = 0$ and $h \geq R_1 \tan(\theta_V / 2)$, the entire vertical extent of the sensor captures the reflected scene and the image is free from keystone distortion; (b) when $\phi_o = 0$ and $h = 0$, only half of the vertical extent of the sensor captures the reflected scene and the image is free from keystone distortion; (c) when $\phi_o = \theta_V / 2$ and $h = 0$, the entire vertical extent of the sensor captures the reflected scene, but the image suffers from keystone distortion**

## 4. Parameters of the DMP panoramic camera

This section describes the parameters of the DMP imaging. We will refer to surface $A_1A_2B_2B_1$ and its associated camera $C_{11}$, whenever a mirror face and the associated camera are considered. This is without loss of generality because the faces are symmetrically located about the pyramid axis as well as the pyramid base. Furthermore, and again without loss of generality, we use virtual viewpoint $O$ and virtual optical axis $OO_{11}$ to represent viewpoint $C_{11}$ and optical axis $C_{11}O_{11}$ respectively. The derived equations and observations will directly apply to all face-camera pairs.

### 4.1. DMP parameters

The determination of the minimum acceptable number and dimensions of the pyramid faces is a key step in the DMP camera design. The FOV of each camera is computed in terms of the focal length $f$ of the camera lens and the sensor size $p(H)mm * q(V)mm$. The horizontal, vertical and diagonal FOVs of a specified lens and CCD sensor are given by $\theta_H = 2\arctan(p/2f)$, $\theta_V = 2\arctan(q/2f)$, and $\theta_D = 2\arctan(\sqrt{p^2+q^2}/2f)$, respectively. To avoid a visual field gap, the reflective visual field of each mirror face must be equal or smaller than the FOV of an individual camera.

For a given lens and a CCD sensor, the shape and size of a pyramid can be uniquely specified by the number of pyramid faces, $N$, the angle between the mirror faces (e.g. $A_1A_2B_2B_1$) and base polygon $A_1A_2...A_N$, $\alpha$, the radius of the polygon $A_1A_2...A_N$, $R_1$, and the height of

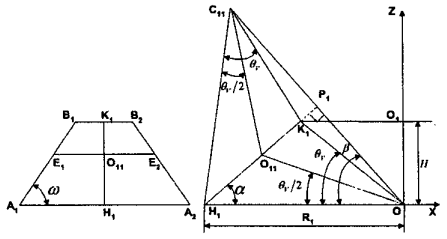


**Fig. 4: Parameter definitions in the DMP**

each truncated pyramid, $H$ (Fig. 4). Given the knowledge of geometrical optics [10] and the FOV constraint described above, these parameters should meet the criteria:

$$\begin{cases} \cos(\frac{180}{N}) \geq \frac{2f}{\sqrt{4f^2 + p^2 + q^2}\cos(\theta_v/2)} \\ \alpha \leq 90 - \theta_v \\ H \geq \frac{R_1 \sin\theta_v \tan\alpha}{\sin(\theta_v + \alpha)} \end{cases} \quad (1)$$

The number of pyramid faces, $N$, gives the angle $\gamma$ subtended by each base edge with the base center (Fig. 2b), i.e. $\gamma = \frac{360}{N}$.

## 4.2. Visual field mapping

We will now define the visual field mapping from the scene to the sensor achieved by the DMP camera. To do so, we first determine the reflective FOV covered by each camera. Then, the visual field mapping from the scene to the image sensor is given by the mapping from the mirror surface to the image sensor.

The ray passing through the center of the entrance pupil of an optical system is defined as the principal ray of a scene point [10]. With the assumption of a thin-lens model, the center of entrance pupil is the viewpoint, or the projection point. A scene point is uniquely defined by the angle that its principal ray makes with its optical axis, known as the field angle in optics.

In Fig. 2(b), a right-hand coordinate system $OXYZ$ is defined in which axis $OZ$ is perpendicular to polygonal plane $A_1A_2...A_N$ and points upward, axis $OX$ is perpendicular to $A_1A_2$ and points to the right, and axis $OY$ is defined according to the right-hand rule. As shown in Fig. 5, a scene point $Q$ on a plane $P$ within the FOV of camera $C_{11}$ maps onto point $T$ on the mirror surface $A_1A_2B_2B_1$. On the other hand, with the assumption that the reflective FOV of the mirror is equal to or narrower than the camera FOV, a line connecting point $T(x,y,z)$ on the mirror surface $A_1A_2B_2B_1$ with the virtual viewpoint $O$ (or real viewpoint $C_{11}$) uniquely defines a principal ray, and therefore, a scene point $Q$.

The visual angle $\theta(T)$ subtended by point $T(x,y,z)$ with $O_{11}C_{11}$ is equivalent to the angle between vector $\overrightarrow{OT}$ and virtual optical axis $\overrightarrow{OO_{11}}$, and is given by

$$\cos\theta(T) = \frac{x^* x_{O_{11}} + y^* y_{O_{11}} + z^* z_{O_{11}}}{|OO_{11}|\sqrt{x^2 + y^2 + z^2}} \quad (2)$$

where

$$|OO_{11}| = \frac{R_1 \sin\alpha}{\sin(\alpha + \theta_v/2)}, \begin{cases} x_{O_{11}} = \frac{-R_1 \sin\alpha \cos(\theta_v/2)}{\sin(\alpha+\theta_v/2)} \\ y_{O_{11}} = 0 \\ z_{O_{11}} = \frac{R_1 \sin\alpha \sin(\theta_v/2)}{\sin(\alpha+\theta_v/2)} \end{cases}, \text{ and}$$

$T(x,y,z)$ is constrained by the plane equation:

$$\begin{cases} \frac{x}{R_1} - \frac{z}{R_1 \tan\alpha} = 1 \\ -R_1 \leq x \leq -R_1 \sin\alpha\cos\theta_v / \sin(\alpha+\theta_v) \cdot \\ -R_1 \tan(\gamma/2) \leq y \leq R_1 \tan(\gamma/2) \\ 0 \leq z \leq R_1 \sin\alpha \sin\theta_v / \sin(\alpha+\theta_v) \end{cases}$$

We are particularly interested in the mapping of boundary lines $A_1A_2$, $B_2B_1$, $A_1B_1$, and $A_2B_2$. These boundary mappings determine the trapezoidal shape of the rectangular object on the image sensor. For any point $T(x,y,z)$ on line $A_1A_2$ ($B_2B_1$, $A_1B_1$, or $A_2B_2$), its corresponding visual angle is given by

$$\begin{cases} \cos\theta(T_{A_1A_2}) = \frac{R_1\cos(\theta_v/2)}{\sqrt{R_1^2 + y_{A_1A_2}^2}} \\ \cos\theta(T_{B_1B_2}) = \frac{R_1 \sin\alpha\cos(\theta_v/2)}{\sqrt{R_1^2\sin^2\alpha + y_{B_1B_2}^2 \sin^2(\alpha+\theta_v)}} \\ \cos\theta(T_{A_1B_1}) = \frac{-x_{A_1B_1}\cos(\theta_v/2) + (x_{A_1B_1} + R_1)\tan\alpha}{\sqrt{\frac{x_{A_1B_1}^2}{\cos^2(\gamma/2)} + (x_{A_1B_1} + R_1)^2 \tan^2\alpha}} \\ \cos\theta(T_{A_2B_2}) = \frac{-x_{A_2B_2}\cos(\theta_v/2) + (x_{A_2B_2} + R_1)\tan\alpha}{\sqrt{\frac{x_{A_2B_2}^2}{\cos^2(\gamma/2)} + (x_{A_2B_2} + R_1)^2 \tan^2\alpha}} \end{cases}$$

$$\begin{cases} -R_1\tan(\gamma/2) \leq y_{A_1A_2} \leq R_1\tan(\gamma/2) \\ \frac{-R_1\sin\alpha\cos\theta_v\tan(\gamma/2)}{\sin(\alpha+\theta_v)} \leq y_{B_1B_2} \leq \frac{R_1\sin\alpha\cos\theta_v\tan(\gamma/2)}{\sin(\alpha+\theta_v)} \\ -R_1 \leq x_{A_1B_1} \leq -\frac{R_1\sin\alpha\cos\theta_v}{\sin(\alpha+\theta_v)} \\ -R_1 \leq x_{A_2B_2} \leq -\frac{R_1\sin\alpha\cos\theta_v}{\sin(\alpha+\theta_v)} \end{cases}$$

Specifically, for points $A_1$, $A_2$, $B_1$, $B_2$, and $O_{11}$, the corresponding visual angles are:

$$\begin{cases} \theta(O_{11}) = 0 \\ \cos\theta(A_1) = \cos(\theta_v/2)\cos(\gamma/2) \\ \cos\theta(A_2) = \cos(\theta_v/2)\cos(\gamma/2) \\ \cos\theta(B_1) = \frac{\cos(\theta_v/2)\cos(\gamma/2)}{\sqrt{1-\sin^2\theta_v\sin^2(\gamma/2)}} \\ \cos\theta(B_2) = \frac{\cos(\theta_v/2)\cos(\gamma/2)}{\sqrt{1-\sin^2\theta_v\sin^2(\gamma/2)}} \end{cases} \quad (3)$$

We assume that the CCD sensor is normal to the optical axis of the camera lens and is aligned so that the central pixel $O_{11}'$ of the sensor is the intersection of the optical axis $OO_{11}$ with the sensor plane, and the long axis of the sensor is parallel to $A_1A_2$, as shown in Fig. 5. We further assume that the camera lens is free from aberrations and the pinhole model applies. Then the

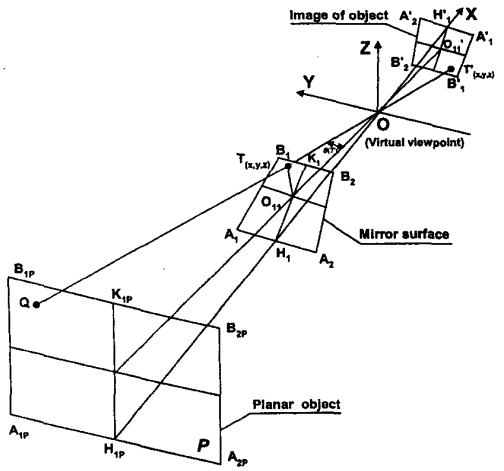



**Fig. 5: Mapping a point on plane *P* onto image sensor. O is the virtual viewpoint of camera $C_{11}$, $OO_{11}$ is the virtual optical axis. A rectangular object $A_{1P}A_{2P}B_{2P}B_{1P}$ is mapped onto a trapezoidal shape $A'_{1P}A'_{2P}B'_{2P}B'_{1P}$.**

mapping of point $O_{11}$ (i.e. the scene center) on image sensor would be the central pixel $O_{11}'$, and the mapping point $Q'$ where scene point $Q$ is mapped is given by $|O'Q'| = f\tan\theta(T)$. The mapping is illustrated in Fig. 6, which shows that a rectangular object is foreshortened and imaged as a trapezoid due to the camera tilt. This phenomenon is referred to as keystone distortion and must be corrected by dewarping to recover the rectangular image.

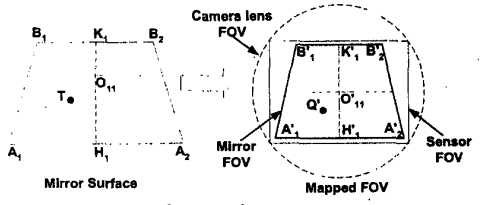



**Fig. 6: Visual field mapping on sensor.**

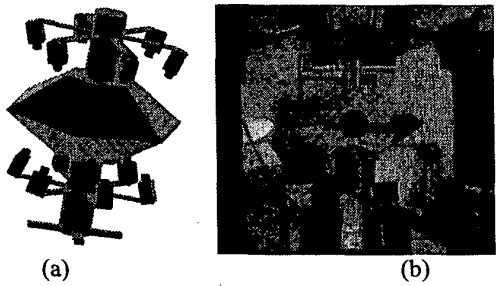
## 5. Implementation and experimental results

Using the equations derived in section 3, we designed a DMP panoramic camera with two right-hexagonal ($N = 6$) truncated pyramids. The small-circle radius of the right hexagon is $R_1 = 86.6mm$. The angle $\alpha$ between each mirror surface and the base of the hexagon is 40 degrees, and the tilt angle of the optical axes is 20 degrees ($\theta_0 = 20^\circ$ and $\theta_v = 40^\circ$). A 3D CAD model of the system and a prototype of the mirror system are shown in Fig. 7. The cameras used are Pulnix with 2/3", 640×480 black/white CCD sensors and 6.5mm lenses. Each of the cameras effectively covers 60 degrees FOV horizontally and 40 degrees FOV vertically. The total FOV of the DMP is $360^\circ(H)*80^\circ(V)$. The field angles corresponding to corners $A_1, A_2, B_1, B_2, H_1, K_1$ are $\theta(A_1) = 35.53^\circ$, $\theta(A_2) = 35.53^\circ$, $\theta(B_1) = 30.75^\circ$, $\theta(B_2) = 30.75^\circ$, $\theta(H_1) = 20^\circ$, and $\theta(K_1) = 20^\circ$, respectively.

In order to capture a panoramic image from a single viewpoint using the DMP camera, the camera clusters must be placed properly with respect to the mirror-pyramids. Otherwise, there may be gaps between the visual fields of the different cameras. For lack of space, we will not present the details of the calibration process in this paper.

Since the FOV of each camera is more than 60 degrees horizontally, pincushion or barrel distortion is almost inevitable and needs to be compensated for. For this purpose, each camera is calibrated using Zhang's calibration method [25], which images a planar pattern at different orientations to estimate the intrinsic and extrinsic parameters, and therefore radial distortion. Again, for lack of space, we will omit the details of this step. Fig. 8a shows an original image and Fig. 8b shows the image after compensating for the radial distortion.

Furthermore, each image, after compensating for the radial distortion, needs to be re-projected onto a virtual sensor that is perpendicular to the pyramid base. This step removes the keystone distortion. An alternative is to avoid the distortion in the first place by tilting the CCD sensor of the camera by $-\theta_0$ degrees with respect to its optical axis, which maps a vertical rectangular object in the scene as a rectangle on the tilted sensor. This avoids the computation required for digital re-projection, and eliminates the loss of quality due to interpolation that accompanies digital re-projection, but is more demanding because it requires a nontrivial manipulation of the camera hardware to adjust sensor orientation. In our



(a) (b)

**Fig. 7: A 6-faced DMP camera: (a) A CAD model; (b) Our implementation.**

current implementation, we have used the first approach, that of digital re-projection. We will omit the details of re-projection here for brevity. Fig. 8c shows the resulting image after compensating for both radial and keystone distortions.

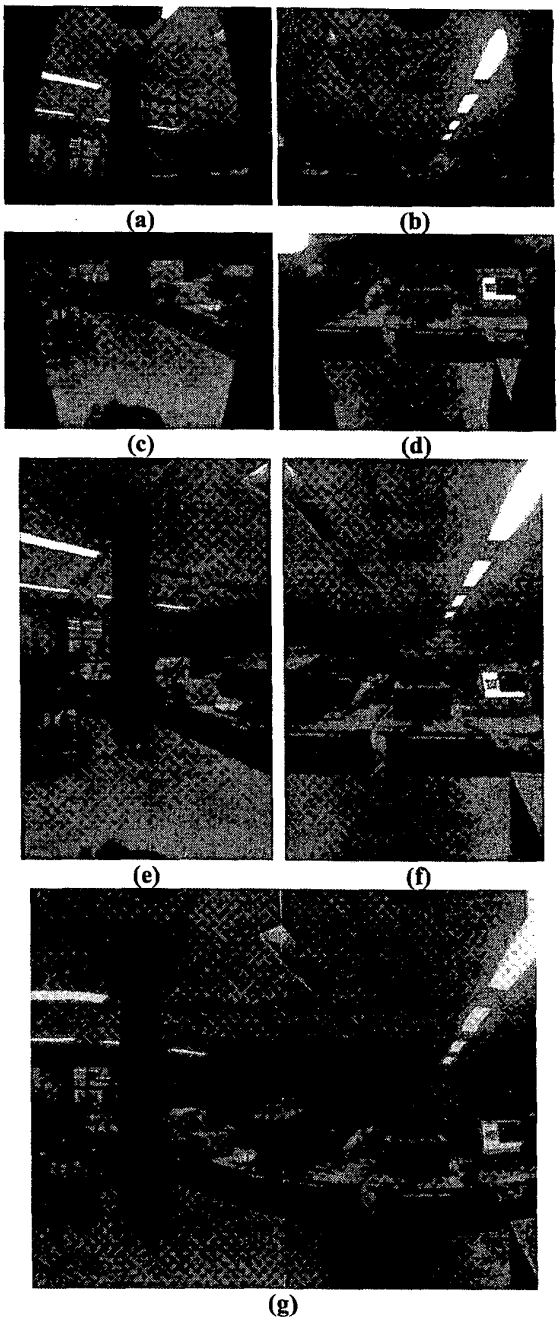
Figures 9 (a)-(d) show four images acquired by four adjacent cameras, two horizontally adjacent in the upper layer ($C_{11}$ and $C_{12}$) and the other two horizontally adjacent in the lower layer ($C_{21}$, $C_{22}$). $C_{11}$ and $C_{21}$ are vertically adjacent and so are $C_{12}$ and $C_{22}$. Figures 9(e) and (f) show the seamless mosaics of the images provided by the camera pairs $C_{11}$ and $C_{21}$, and $C_{12}$ and $C_{22}$, respectively, after post-processing for keystone and radial distortions. Fig. 9(g) shows the seamless cylindrical mosaic of Figs 9(e) and (f) [22]. Due to the mirror effect, the original images acquired by the cameras need to be flipped appropriately before they are mosaiced.

## 6. Discussion

In generating panoramic images using a mirror-pyramid, we have assumed that the cameras are pinholes, i.e., the aperture is very small, and the edges formed by adjacent mirrors are perfect, i.e. knife-shaped and free of rounding. However, these assumptions do not hold in practice. For example, a real camera has a non-pinhole aperture, and its size varies as the F-number setting is changed. Therefore, a bundle of rays from a scene point, instead of only the principal ray, go through the entrance pupil to form the image of the point. The effect of the non-pinhole aperture minimally affects the parts of the image formed by reflections off the interiors of the pyramid faces, but there is mixing of light arriving directly from the scene and that after reflection off the mirror. This leads to artifacts. Furthermore, since the two adjacent mirrors do not form a perfect knife-edge, the edge curvature further adds to the complexity of the image formed. For example, if two adjacent faces have a flat surface transition (i.e. an approximately planar patch connecting the two faces and having a normal close to the average of the two face normals), this may lead to a loss of light from the corresponding part of the scene, resulting in a dark band, which will occur periodically

(a) (b) (c)

**Fig. 8: Compensating for distortions. (a) Original DMP image; (b) Result after correcting radial distortion using Zhang's calibration method; (c) Result after compensating for both radial and keystone distortions.**



(a) (b) (c) (d) (e) (f) (g)

**Fig. 9: Sample images obtained by the DMP camera. (a)-(d) Original images of camera $C_{11}$, $C_{12}$, $C_{21}$ and $C_{22}$; (e) The mosaic of images (a) and (c) after post-processing of keystone and radial distortions; (f) The mosaic of images (b) and (d) after post-processing of keystone and radial distortions; (g) The cylindrical mosaic of the images (e)-(f).**

across the panoramic image along each boundary between sub-images acquired by adjacent cameras. Such artifacts will occur in all mirror-pyramid based systems. We plan to analyze these effects and explore possible ways of eliminating or softening them.

## Acknowledgements

This work was supported in part by National Science Foundation Grant IIS 00-83037 ITR. We thank Chunyu Gao for his assistance in the design, building and calibration of the camera.

## References

[1] P. I. Anderson, "From telepresence to true immersive imaging: into real-life video-now!", *Advanced Imaging*, 1995, Vol. 10(7), pp48- 50, 1995.

[2] S. Baker and S. K. Nayar, "A theory of single-viewpoint catadioptric image formation", *in International Journal of Computer Vision*, 1999, 35(2), pp.1-22.

[3] R. Benosman, E. Deforas, and J. Devars, "A new catadioptric sensor for the panoramic vision of mobile robots", in *Workshop on Omnidirectional Vision*, June 2000, pp. 112-116.

[4] J. S. Chahl and M. V. Srinivasan, "Reflective surfaces for panoramic imaging", *Applied Optics*, November 1997, 36(31):pp. 8275-8285.

[5] J. S. Chahl and M. V. Srinivasan, "A complete panoramic vision system, incorporating imaging, ranging, and three dimensional navigation", *in Workshop on Omnidirectional Vision*, June 2000, pp. 104-111.

[6] S. Coorg, N. Master, and S. Teller, "Acquisition of a large pose-mosaic dataset", in *Conference on Computer Vision and Pattern Recognition*, 1998, pp. 872-878.

[7] R. A. Hicks, "Reflective surfaces as computational sensors", *in Workshop on Perception for Mobile Agents*, 1999, pp. 82-86.

[8] R. A. Hicks and R. Bajcsy, "Catadioptric sensors that approximate wide-angle perspective projections", in *Workshop on Omnidirectional Vision*, June 2000, pp. 97-103.

[9] H. Ishiguro, M. Yamamoto, and S. Tsuji, "Omni-directional stereo", *IEEE Transactions On Pattern Analysis and Machine Intelligence*, February 1992, 14(2), pp257-262.

[10] B. R. Johnson, "Lenses", *Handbook of Optics: Device Measurement and Properties*, Michael Bass, Eric W. Van Stryland et. al. Editors, volume II. McGraw Hill, 2nd edition, 1995.

[11] Takahito Kawanishi, Kazumasa Yamazawa, et al. "Generation of high resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors," *14th International Conference on Pattern Recognition*, 16th-20th August, 1998, Brisbane, Australia, pp. 485-489, 1998.

[12] A. Krishnan and N. Ahuja. "Range estimation from focus using a non-frontal imaging camera," in *National Conference on Artificial Intelligence*, pages 830-835, Washington D.C., July 11-15 1993.

[13] A. Krishnan an N. Ahuja, Range Estimation from Focus using a Nonfrontal Imaging Camera, *Int. Journal of Computer Vision*, Vol. 20, No. 3, 1996, 169-185.

[14] A. Krishnan and N. Ahuja, "Panoramic image acquisition", in *Conference on Computer Vision and Pattern Recognition*, 1996, pp. 379-384.

[15] K. Miyamoto, "Fish eye lens", *Journal of Optical Society of America*, August 1964, 64:pp.1060-1061.

[16] V. Nalwa, "A true omnidirectional viewer", *Technical report, Bell Laboratories*, February 1996.

[17] V. Nalwa, "Stereo panoramic viewing system", US Patent 6141145.

[18] S. K. Nayar, "Catadioptric omnidirectional camera", in *Conference on Computer Vision and Pattern Recognition*, 1997, pp. 482-488.

[19] S. K. Nayar and A. Karmarkar, "360X360 mosaics", in *Conference on Computer Vision and Pattern Recognition*, June 2000, volume 2, pp. 388-393.

[20] S. K. Nayar and V. Peri, "Folded catadioptric cameras", in *Conference on Computer Vision and Pattern Recognition*, June 1999, volume 2, pp. 217-223.

[21] S. Peleg, "Panoramic mosaics by manifold projection", *in Conference on Computer Vision and Pattern Recognition*, June 1997, pp 338-343.

[22] H.-Y. Shum and R. Szeliski, "Panoramic image mosaics", *Technical Report MSR-TR-97-23*, Microsoft Research, 1997.

[23] Y. Xiong and K. Turkowski, "Creating image-based VR using a self-calibrating fisheye lens", in *Conference on Computer Vision and Pattern Recognition*, 1997, pp. 237-243.

[24] K. Yamazawa, Y. Yagi, and M. Yachida, "Omnidirectional imaging with hyperboloidal projection", in *International Conference on Intelligent Robots and Systems*, July 1993, pp 1029-1034.

[25] Zhengyou Zhang, "A flexible new technique for camera calibration", *Technical Report MSR-TR-98-71*, Microsoft Research, 1998.

[26] Z. Zhu, K. D. Rajasekar, E. M. Riseman, and A. R. Hanson, "Panoramic virtual stereo vision of cooperative mobile robots for localizing 3D moving objects", in *Workshop on Omnidirectional Vision*, 2000, pp 29-36.

[27] S. Zimmerman and D. Kuban, "A video pan/tilt/magnify/rotate system with no moving parts", in *IEEE/AIAA Digital Avionics Systems Conference*, 1992, pp. 523-531.