



Indian Academy of Sciences, Bengaluru  
Indian National Academy of Sciences, New Delhi  
The National Academy of Sciences India, Allahabad

**SUMMER RESEARCH FELLOWSHIPS — 2016**

# Catadioptric Scene Analysis

Report By

Arjun R Rao  
Information Science and Engineering  
M S Ramaiah Institute of Technology, Bangalore

Under the guidance of

Prof. Santanu Chaudhury  
Director  
CSIR-CEERI, Pilani

## **ACKNOWLEDGEMENT**

This research report is an outcome of the two month Summer Research Fellowship Programme conducted by the science academies IAS-NASI-INSA. I extend my heartfelt gratitude to the academies for providing students with an opportunity to work with leading researchers of our country through this fellowship. I would like to thank my guide Prof. Santanu Chaudhury, Director, CSIR-CEERI, for his continuous support and guidance throughout the fellowship program. His support and encouragement has been invaluable.

I would like to thank Dr. Sanjay Singh, and Mr. Sumeet Saurav in the Computer Vision Lab, CSIR-CEERI, Pilani for their insight and expertise that greatly assisted me in performing this research work. Their support was pivotal to making my experience at CEERI a valuable one.

I would also like to thank CSIR-CEERI, Pilani for providing accommodation and infrastructure, hence making my stay comfortable.

**Arjun R Rao**

## Table of Contents

<b>ACKNOWLEDGEMENT .....</b>	<b>2</b>
<b>ABSTRACT .....</b>	<b>4</b>
<b>1. Introduction.....</b>	<b>5</b>
<b>2. Theoretical Review of Catadioptric Image Formation .....</b>	<b>6</b>
<b>3. Approach to Unwrapping Omnidirectional Images .....</b>	<b>9</b>
Panoramic Image Unwrapping Method by Deriving a Mapping Relationship.....	9
Obtaining Cylindrical Panorama from Captured Omnidirectional Image .....	10
Obtaining Perspective Images from Panoramic Image.....	11
Automatic Circle Centre Detection .....	13
A note on the algorithm's Performance .....	13
Results of applying unwrapping algorithm on omnidirectional image .....	14
<b>4. Scene Analysis and Pedestrian Detection.....</b>	<b>15</b>
Integral Channel Features: A Theoretical Overview .....	15
Experiments and Results in MATLAB.....	18
<b>5. Conclusions.....</b>	<b>20</b>
A note on accuracy of detections .....	20
<b>REFERENCES.....</b>	<b>21</b>

## ABSTRACT

Many applications using computational vision techniques require that a large field of view be imaged. Examples include surveillance, teleconferencing, and model acquisition for virtual reality. A number of other applications, such as ego-motion estimation and tracking, would also benefit from enhanced fields of view. Unfortunately, conventional imaging systems are severely limited in their fields of view. Both researchers and practitioners have therefore had to resort to using either multiple or rotating cameras in order to image the entire scene. To overcome this limitation, a combination of lenses and mirrors are used to capture a 360 field of view. Such systems are called catadioptric-imaging systems. This report summarizes methods that are a feasible for working with catadioptric imaging systems, for scene analysis. In particular this research study is aimed at identifying methods for performing pedestrian detection in video streams captured using catadioptric cameras. Pedestrian detection as a problem, has applications in surveillance systems, and intruder detection systems, which are put in place to monitor any critical infrastructure. The findings of literature surveys and experimentation of algorithms conducted as part of the investigation of using catadioptric camera systems for scene analysis are presented.

## 1. Introduction

Many applications in computational vision require that a large field of view is imaged. Examples include surveillance, teleconferencing, and model acquisition for virtual reality. A number of other applications, such as ego-motion estimation and tracking, would also benefit from enhanced fields of view. Unfortunately, conventional imaging systems are severely limited in their fields of view. Both researchers and practitioners have therefore had to resort to using either multiple or rotating cameras in order to image the entire scene. To overcome this limitation, researchers have come up with a methodology to use mirrors in conjunction with lenses [1].

Such a system that uses lenses and mirrors in order to capture a large field of view, is termed in literature as '*Catadioptric Imaging*' or '*Catadioptric Image formation*'. As noted in [1] it is desirable to have catadioptric systems with a single view point as it permits the generation of geometrically correct perspective images from the images captured by the catadioptric cameras. The generation of perspective images allows for the application of existing image processing algorithms on the generated images thus allowing the use of catadioptric cameras in place of complicated multi-camera setups that are used to obtain a wider field of view. Obtaining perspective/panoramic images from omnidirectional 360° view, is termed as unwrapping of omnidirectional images and techniques for unwrapping are described by many researchers in literature, which are summarized systematically in this report. Once image unwrapping is performed, the perspective images obtained are used in conjunction with object tracking algorithms, to perform scene analysis. In this report, the focus is on application of pedestrian detection algorithms on unwrapped catadioptric images.

Before the discussion on unwrapping method used on omnidirectional images, a brief overview of catadioptric image formation and the single view point constraint is presented as this is fundamental to working with catadioptric imaging systems.

## 2. Theoretical Review of Catadioptric Image Formation

A catadioptric realization of omnidirectional vision combines reflective surfaces and lenses. Catadioptric imaging systems that preserve the uniqueness of the projection viewpoint are called central catadioptric systems [2]. Central catadioptric imaging can be highly advantageous for many applications because it combines two important features: a single projection centre and a wide field of view. The following figure shows the formation of images by a central catadioptric vision system.

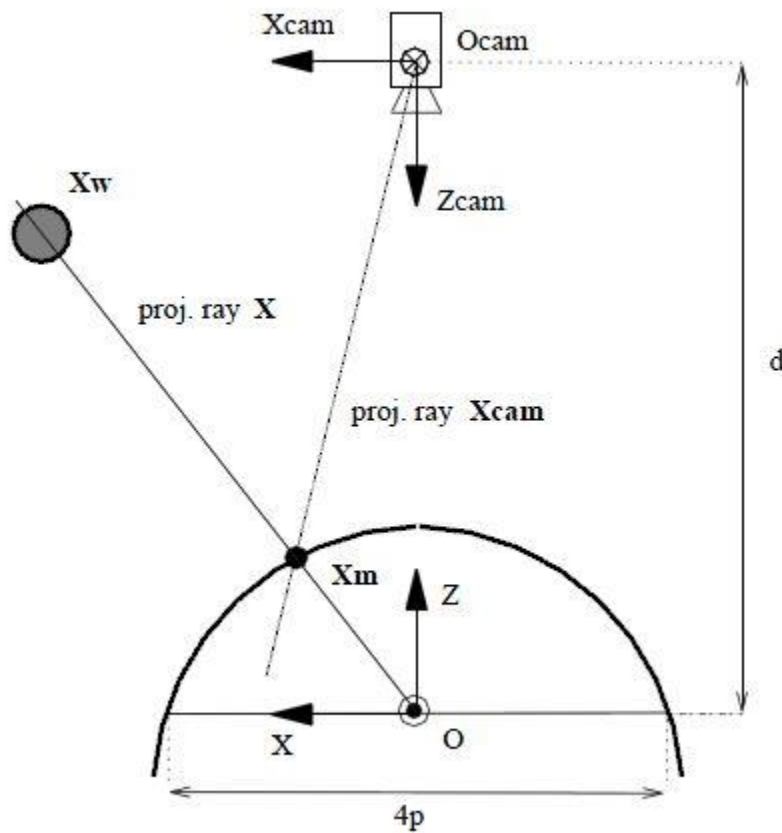


Figure 1: Central Catadioptric Vision System (Source: [2])

The system depicted combines a hyperbolic reflective surface with a perspective camera. The hyperbolic mirror is placed such that its axis is the  $Z$ -Axis, its foci are coincident with  $O$  and  $O_{cam}$ . Its latus rectum is  $4p$  and the distance between the foci is  $d$ . Light rays incident with  $O$  are reflected as rays incident with  $O_{cam}$ . The assumption is that the perspective camera is placed with its projection centre at  $O_{cam}$  and pointing towards the mirror surface. Thus the rays captured by the

camera go through the focus  $\mathbf{O}$  first, and thus the effective view point of the captured image is unique. A catadioptric system made up of a perspective camera steering a planar mirror also verifies the fixed view point constraint. The effective projection centre is behind the mirror in the perpendicular line passing through camera centre. Its distance to the camera centre is twice the distance between the planar mirror and the camera.

The fixed viewpoint constraint [1] is the requirement that a catadioptric sensor only measure the intensity of light passing through a single point in 3-D space. The fixed 3-D point at which a catadioptric sensor samples the plenoptic function is known as the *effective viewpoint*.

In the case that a single conventional camera is to be used as the only sensing element and a single mirror as the only reflecting surface. If the camera is an ideal perspective camera and defocus blur is ignored, it can be modeled by the point through which the perspective projection is performed; i.e. the effective pinhole. Then, the fixed viewpoint constraint requires that each ray of light passing through the effective pinhole of the camera (that was reflected by the mirror) would have passed through the effective viewpoint if it had not been reflected by the mirror.

The fixed view point constraint equation and its solution are derived in [1] using the following geometry:

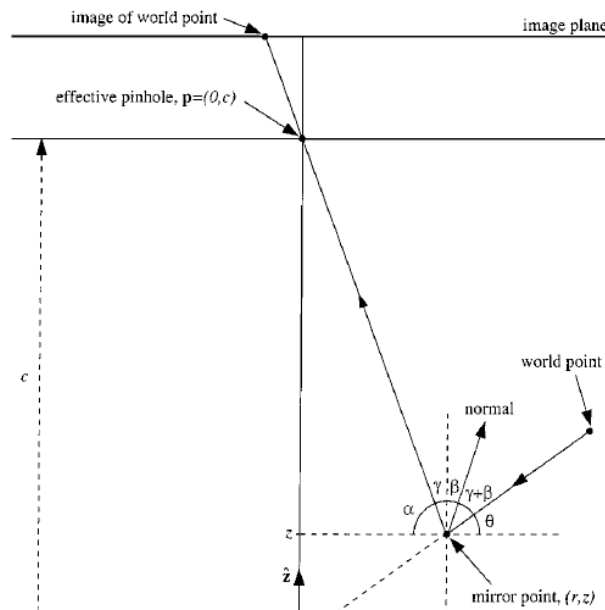


Figure 2: Geometry for derivation of fixed viewpoint constraint equation. (Source: [1])

The derived constraint equation is given by the following quadratic first order ordinary differential equation:

$$r(c - 2z) \left( \frac{dz}{dr} \right)^2 - 2(r^2 + cz + z^2) \frac{dz}{dr} + r(2z - c) = 0$$

The general solution of the above differential equation and thus the specific solutions so generated, define the entire class of mirrors that satisfy the fixed viewpoint constraint, and hence can be used to define the shapes of the mirrors that can be used to produce omnidirectional images. This has been thoroughly described by Baker and Nayar in [1]. Their paper presents the solutions of the constraint equation mentioned above, that generate mirror systems that can be used in catadioptric sensors, namely, conical, spherical, ellipsoidal and hyperboloidal mirrors. Of these geometries obtained, conical mirrors and spherical mirrors cannot be used in actual sensors, as they turn out to be degenerate solutions of the fixed viewpoint constraint equation. They also show that paraboloidal mirrors can also be used conveniently in omnidirectional sensors, even though they aren't generated amongst the solutions of the fixed viewpoint constraint equation. The following figures illustrate the fixed viewpoint constraint for various mirror geometries discussed in [1].

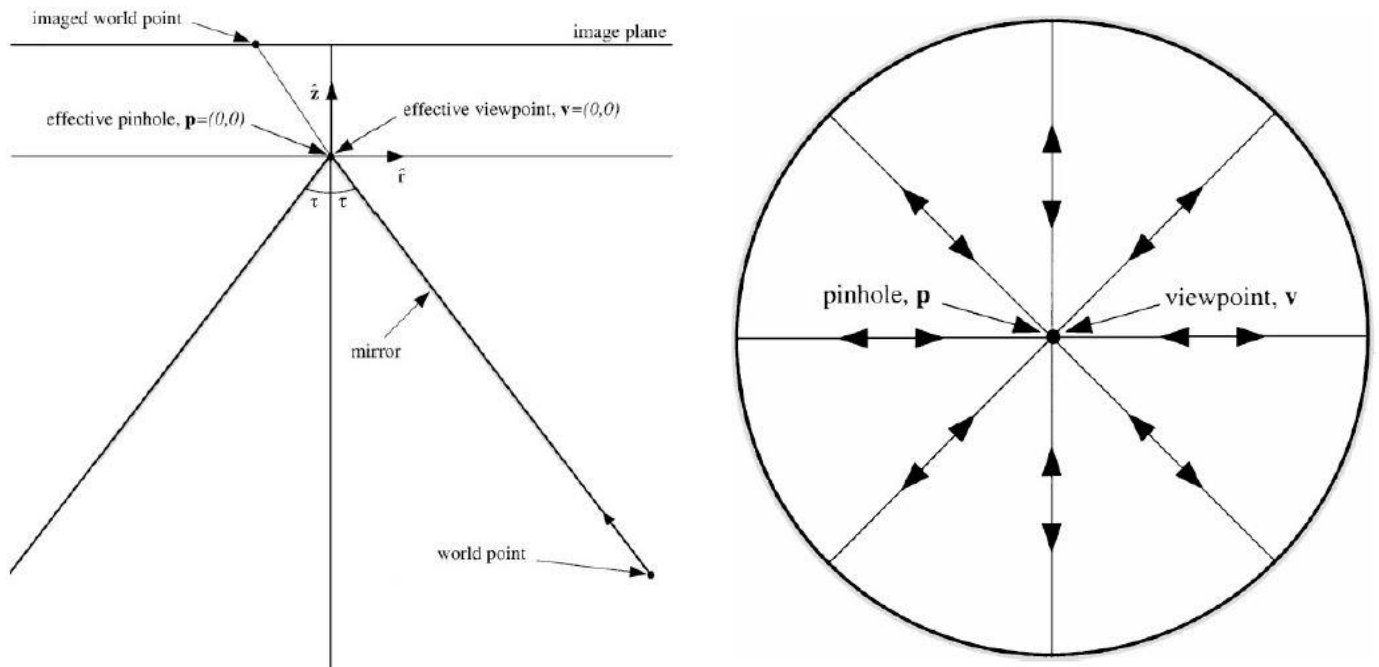


Figure 3: Conical (left) and Spherical Mirrors, degenerate solutions of the fixed viewpoint constraint equation.



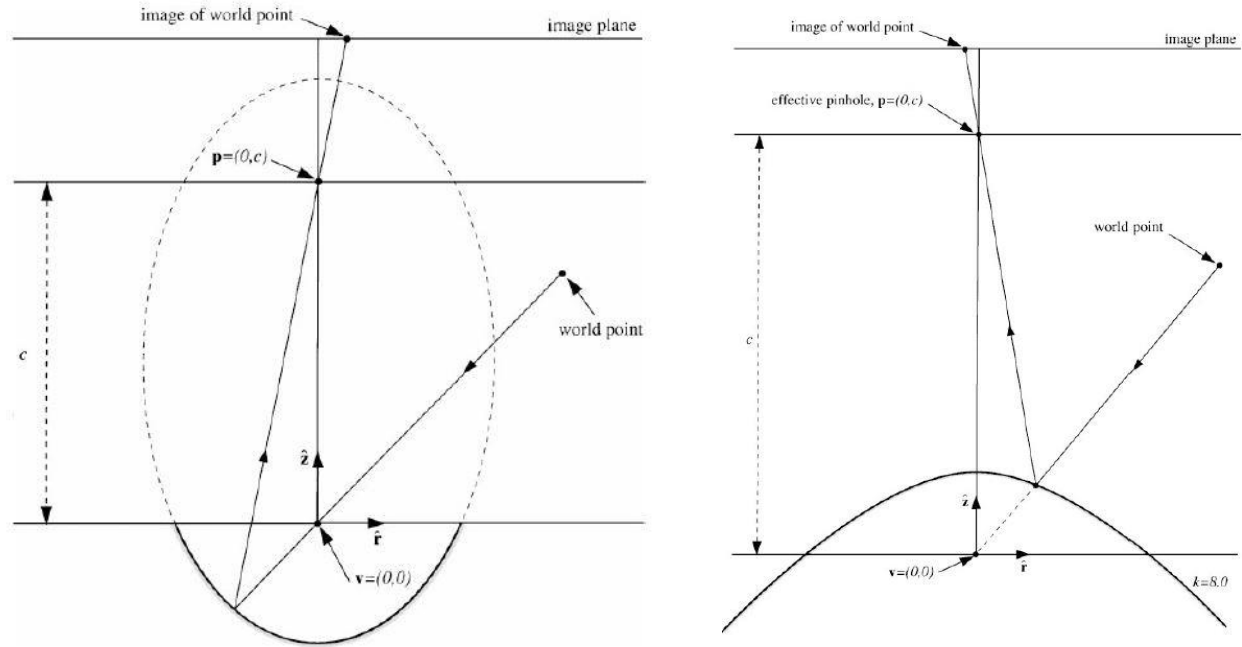


Figure 4: Ellipsoidal (left) and Hyperboloidal Mirrors, which are feasible solutions to the fixed viewpoint constraint equation

### 3. Approach to Unwrapping Omnidirectional Images

In this section, an approach to unwrapping an omnidirectional image to obtain its panoramic view is presented. The techniques used for unwrapping omnidirectional images, are varied due to the nature of the method used to detect the geometry of the mirror system, as well as the technique implemented to obtain a correlation between pixels of the omnidirectional image and pixels in its perspective/panoramic projection. Complexities are further introduced due to the difference in resolution of omnidirectional and panoramic images. The omnidirectional image being lower in resolution requires the unwrapping method to correct for the same using various techniques of interpolation. After a review of various methods used in literature (see references for other methods referred), the following approach has been selected for experimentation.

#### Panoramic Image Unwrapping Method by Deriving a Mapping Relationship

In this method the omnidirectional image is considered as the reflected cylindrical image around the mirror as shown in the following figure. The figure implies that in each cutting surface, there exists a one-to-one mapping between the cylinder edge and the Omni-image's radius, which can be used to recover the unwrapped image. This technique exploits this relationship to derive an

equation that allows calculation of coordinates of pixels on the panoramic image that correspond to coordinates of pixels in the omnidirectional image.

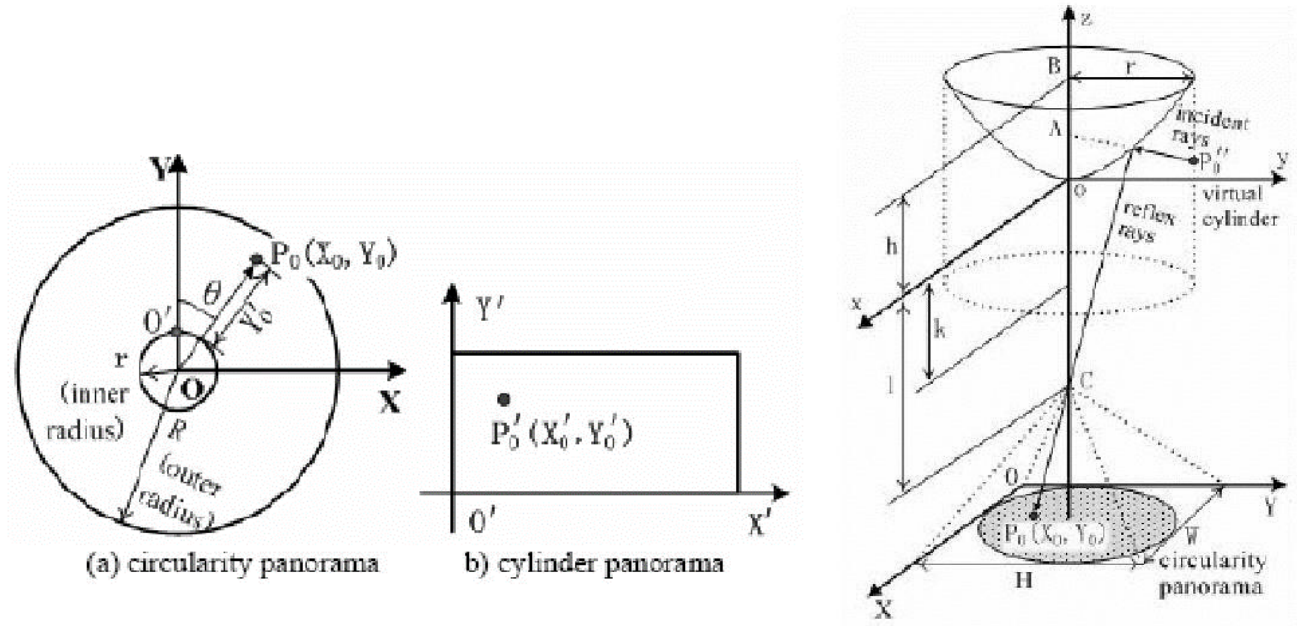


Figure 5- Relationship between cylindrical and omnidirectional image

### Obtaining Cylindrical Panorama from Captured Omnidirectional Image

As denoted by the figure, the omnidirectional image has an inner radius of  $r$  and an outer radius of  $R$ . The region between  $r$  and  $R$ , is the valid region of the image considered for unwrapping, this is due the reflection of the camera lens occupying the region inside the inner radius, which is irrelevant to the view being captured. Thus from the figure we can see that for a given pixel in the cylindrical panorama,  $P'_0(X'_0, Y'_0)$ , the polar coordinates of the corresponding pixel in the circular panorama  $P_0(X_0, Y_0)$  can be determined by the following formula:

$$h_0 = \frac{r_m}{y_m} Y'_0$$

$$\theta_0 = \frac{2\pi}{x_m} X'_0$$

Where  $(X_m, Y_m)$  the desired dimensions of the cylinder are image in pixels and  $r_m$  is the radius of the mirror as seen in the circular panorama in pixels, and  $\theta_0$  is in radians. Using these equations

the mapping between the polar coordinates and the rectangular coordinates in the circular panorama can be obtained as:

$$X_0 = X_c + h_0 \cos \theta_0 = X_c + \frac{r_m}{y_m} Y'_0 \cos \frac{2\pi}{x_m} X'_0$$

$$Y_0 = Y_c + h_0 \sin \theta_0 = Y_c + \frac{r_m}{y_m} Y'_0 \sin \frac{2\pi}{x_m} X'_0$$

Thus the above equations can be used to calculate the coordinates in the omnidirectional image, when given a point in the cylindrical panoramic image. It is to be noted that the coordinates  $(X_0, Y_0)$  need not be integers, hence the pixel values cannot directly be obtained from the image captured, and thus requires the use of some estimation technique such as interpolation methods to determine the pixel values.

The approach to ensure better quality of unwrapped images used in this method is through interpolation to approximate pixel values from neighbouring pixels. The methods described in [3] include nearest-neighbour interpolation, average interpolation, linear interpolation, bi-linear interpolation and cubic interpolation, while the experiments carried out in [3] ultimately use cubic interpolation techniques to better the resolution of unwrapped images. The choice is so made, as according to the discussion on cubic interpolation in [3], the cubic interpolation results in the best approximation of the *SINC* function which is required for upsampling for estimating arbitrary pixel values.

### **Obtaining Perspective Images from Panoramic Image**

The obtained panoramic image can further be used to generate perspective images from the panorama. This can further eliminate some amount of geometric distortion and improve the presentation of images captured to a viewer. The projection of panoramic to perspective views is done by deriving a mapping relationship between the panorama and the perspective image. The curved plane of the cylindrical panorama is used to reconstruct the perspective image by projecting it onto a virtual plane. The projections are done using a virtual view point, which is to be adjusted manually due to lack of a fixed focal point for the camera system used for obtaining omnidirectional images. The projection of the panoramic view to a plane is illustrated by the following figure.

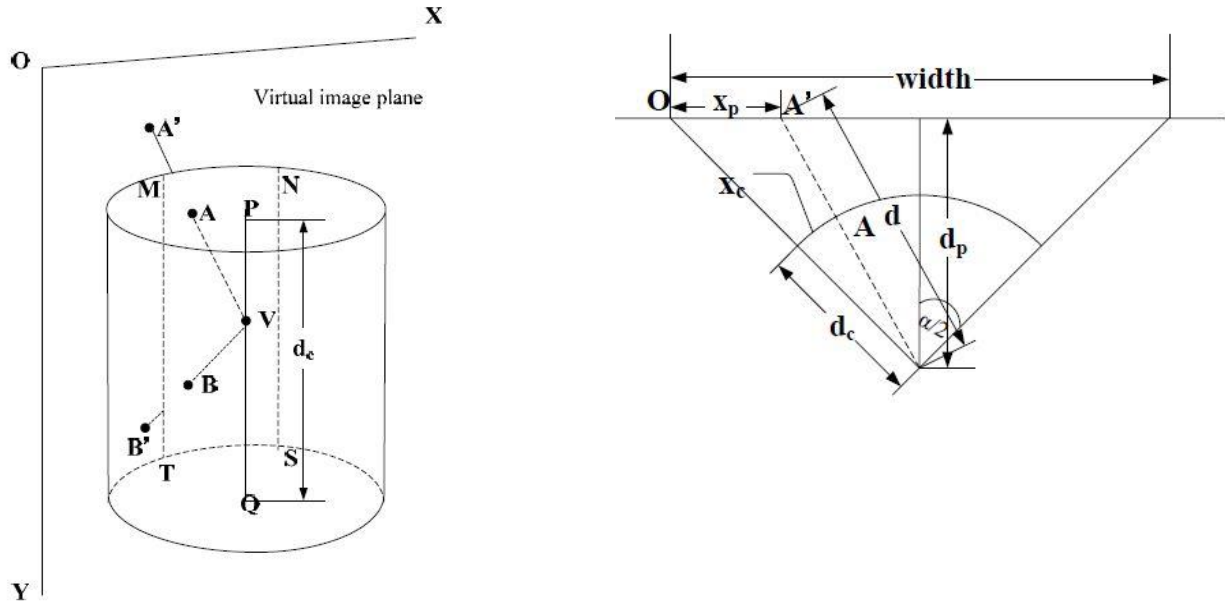


Figure 6: Projection from panorama to virtual plane

It can be seen that the lines VAA' and VBB' are projection lines that project the points A and B on the cylindrical panorama to the virtual plane. Thus using the geometrical representation above, a mapping relationship has been derived in [3] to obtain perspective images from the cylindrical panorama, which is given by:

$$x_c = d_c \left[ \frac{\alpha}{2} - \frac{1}{d_p} \tan^{-1} \left( \frac{width}{2} - x_p \right) \right]$$

$$y_c = y_v + \frac{d_c}{d_p} (y_p - y_v)$$

Using the above equations, a point on the perspective plane can be calculated using its corresponding mapping point on the panorama. In case the coordinates are non-integral values, interpolation techniques can be applied to estimate the pixel value.

### **Automatic Circle Centre Detection**

The techniques for unwrapping described here, rely on the position of the centre of the circle and the radius of the circular image obtained by the omnidirectional camera setup. The inner and outer radii along with the circle's centre determine the area of interest for the unwrapping process. Determination of this centre and radii is also addressed in [3], which allows for automatic circle centre detection along with determination of radii. The proposed approach is to apply Circle Hough Transform (CHT), which can accurately determine the circle centres and the radii. The centre-detection algorithm is based on the edge image (a binary image), which can efficiently reduce the computational complexity and improve the detection accuracy. The method used is to analyse the equation of a circle in X-Y plane given by  $(x - a_x)^2 + (y - a_y)^2 = R^2$  where  $x$ ,  $y$  are treated as parameters to determine the values of considered variables  $R$ , and  $(a_x, a_y)$  which denote the radius and position of the circle's centre respectively.

### **A note on the algorithm's Performance**

The primary assumption for the effective working of this method is that the camera system used must have a constant vertical resolution, which implies that objects at a fixed distance from the camera's optical axis will always be the same size in the image, independent of its vertical coordinates. Thus if the camera system used doesn't meet this requirement, the invariance of vertical resolution will not hold, and the obtained formulae for unwrapping can include transformation error, which will not result in satisfying results. The MATLAB implementation of the techniques described above, were executed on omnidirectional test images, to view the efficiency of the unwrapping approach. During MATLAB simulations, it is found that the CHT used to detect the position of the circle's centre and radii is time consuming due to the size of the search space. Thus the implementation of the circle detection algorithm needs to be optimized if this approach is to be used for real-time omnidirectional image/video processing. The algorithm can be made to run faster if human intervention is used to identify image centres and inner and outer radii limits of the omnidirectional image. This bypasses the need for automatic circle detection, which is a bottleneck in the otherwise convenient algorithm.

**Results of applying unwrapping algorithm on omnidirectional image**

The following images show one of the results of the unwrapping algorithm on various omnidirectional test images shot using the GoPano Plus catadioptric lens.



Figure 7a- Omnidirectional Image View (Source - [gopano.com](http://gopano.com) [14])



Figure 7b - Unwrapped Panoramic View obtained from MATLAB after running the unwrapping algorithm

#### 4. Scene Analysis and Pedestrian Detection

The following section and its subsections highlight the use of catadioptric images for scene analysis. In particular we focus on object detection and in specific, pedestrian detection and hence tracking. The discussion follows primarily based on work carried out by Dalal and Triggs on Histograms of Oriented Gradients for human detection [9] as well as that carried about by Piotr Dollár et.al on using Integral Channel Features for image classification [10]. Pedestrian Detection as a problem has key applications in surveillance, assistive robotics, etc. The problem of pedestrian detection with catadioptric cameras has a wide scope of application as the limitations of field of view are lifted due to the omnidirectional image. However the consequence of using a catadioptric camera is that direct application of algorithms as highlighted in [9], [10] and [11] are not applicable to the omnidirectional view, due to the nature of data and method used for classification. These require that the image on which feature computation is performed for detection, be a perspective/panoramic view. Thus in this research work, a method of unwrapping has been discussed in detail as the workflow used for pedestrian involves 3 stages as described below.

The image acquisition step being the first stage is performed using a catadioptric camera system, followed by application of the unwrapping algorithm on the catadioptric image. The resulting panoramic image is fed as an input image to the pedestrian detection algorithm, resulting in detection data that is overlaid on the image as bounding boxes and confidence scores. This workflow is illustrated through the outputs of images of each stage in the subsequent section on experimental results. Prior to presenting results of experiments carried out, a theoretical discussion of the method used for pedestrian detection using integral channel features as in [10] is presented.

##### **Integral Channel Features: A Theoretical Overview**

The work behind integral channel features is based on the use of the efficiency of computation of such features using Integral Images as highlighted in the landmark paper by Viola and Jones [12]. The method described in the section below is from the original work of Piotr Dollár, et al [10] where they highlight the use of Integral Channel Features for object detection in particular pedestrian detection. The classification method relies on use of a boosting framework, while the topic of feature selection is what is addressed in detail in [10]. Integral channel features is basically

based on the use of linear and nonlinear transformations of the input image to compute multiple registered image channels, followed by use of integral images to compute efficiently features such as histograms, local sums, and Haar-like features.

In [10], a computed channel is a registered map of the original image, where the output pixels are corresponding patches of the input pixels. This preserves the overall image layout. For a greyscale image a trivial channel can be that which is same as the input image. While for a colour image, this can be derived by taking each colour channel as an individual channel image. Linear and non-linear transformations of the input image are then used to obtain other channels. This is illustrated in the figure below as shown in [10].

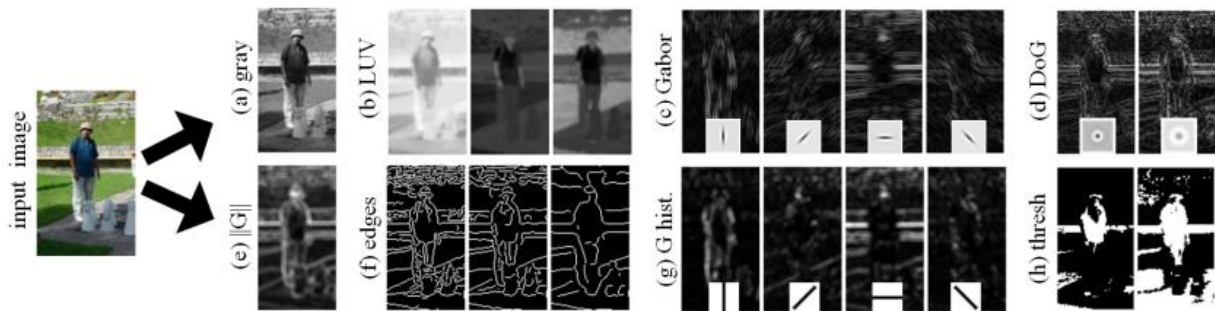


Figure 8- Multiple registered image channels computed from input image

From the channels obtained, feature extraction is performed. Using an integral image for each channel, computation of a first-order channel feature which is a sum of pixels in a fixed rectangular window in the channel is computed very efficiently. Using multiple first order channel features, higher order channel features are computed, for example using sums or differences of two simple features in the same channel.

A brief description of each of the channels used in [10] is presented below. The channels describe correspond to images (a) through (h) in figure 8.

The simplest channel used is a grayscale version of the image (fig. 8a). Colour channels used are shown in fig. 8b comprising of three CIE-LUV colour channels. As shown in fig. 8c, the input image is convolved with 4 oriented Gabor filters. Each channel contains edge information in the original image at different orientations. The results shown in fig. 8d are that obtained on convolution of the input using Difference of Gaussian (DoG) filters, which record the texture



variations at different scales. Edge information is captured using gradient magnitudes (fig. 8e) and canny edge operator (fig. 8f). Generally for colour images, 3 colour channels are used separately to compute gradients and the maximum response is used. Fig 8h shows the result of thresholding the input image with two different thresholds.

Apart from these methods, other methods such as using pointwise transformations, integral histograms, gradient histograms are also discussed in [10]. The gradient histograms can be used to approximate HOG features [9]. The channels obtained using these methods require very little code for implantation using standard image processing tools. Their efficient computation is one of the key strengths of using integral channel filters.

In the implantation of integral channel features, a combination of three types of channels i.e., gradient histograms, colour channels and gradient magnitudes are used as they capture diverse information about the input image. A large pool of candidate features are randomly generated. The first-order feature is a sum over a rectangular region in a given channel. The generation is performed by random selection of a channel and the rectangle having the constraint of minimal area being 25 pixels i.e. - 5x5 rectangle. Higher-order features are also generated randomly through weighted sums of first-order features, each spanning across multiple channels. Random feature sampling is used as it yields good results. The boosting framework as described in [12] is useful for learning a large number of candidate features and also serves as a basis of the Viola-Jones object detection framework. A variant of the approach known as soft-cascade is used in [10] for implementation. Evaluation of weak learners is performed by having a threshold instead of using multiple distinct cascade layers. A single boosted learner is trained on the entire data set and post training heuristic analysis is used to set thresholds this making the training procedure simple.

A sliding window over multiple scales is used during the detection process. Multiple nearby detections are overcome using Non-maximal suppression. In [10] a simplified non-maximal suppression method is used that suppresses the less confident of pairs of detection that overlap significantly.

## Experiments and Results in MATLAB

After detailed theoretical review of various methods of unwrapping omnidirectional images, and survey of object detection methods, practical implementations of the algorithms were carried out in MATLAB R2015b using both the Image Processing Toolbox as well as Computer Vision Toolbox by MathWorks®. In addition to the toolboxes supplied by MathWorks®, P.Dollár's Computer Vision Toolbox [13] was used for setting up the detection framework and performing pedestrian detection using Integral Channel Features method.

The implementation workflow is a three-staged process as shown by the figure below:

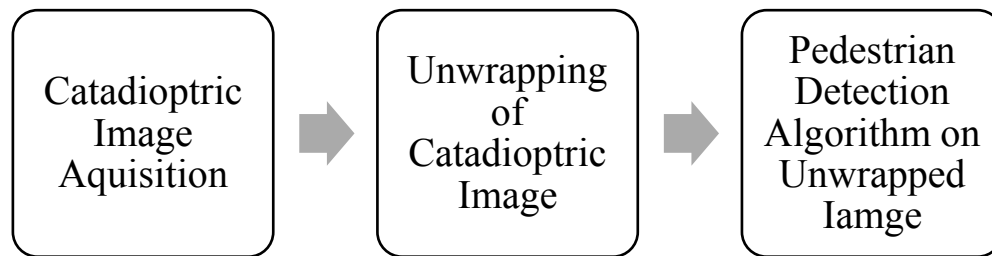


Figure 9- Three Staged Pedestrian Detection Workflow for Catadioptric Images

The image acquisition stage in this research work was done using existing video footage [14] obtained from a catadioptric camera system, EasyEye360 GoPano Plus. The GoPano Plus is a catadioptric lens system that can be mounted on compatible cameras for capturing 360° fields of view. The lens system is depicted in the following figure.



Figure 10- GoPano Plus Catadioptric Lens System

The obtained catadioptric videos were then fed as input the unwrapping algorithm implemented in MATLAB, in order to obtain the unwrapped panoramic views. The results of these stages are illustrated in Fig 7a and 7b respectively.

The unwrapped panoramic videos are given as input to the Integral Channel Features Detection Algorithm implemented in MATLAB using P.Dollár's toolbox and available functions from MATLAB's computer vision toolbox. The result of the algorithm is a set of bounding box coordinates and confidence scores of pedestrian detections in each frame of the image. This data is superimposed on the panoramic input videos to obtain an output video that contains tracked pedestrian detection across frames of the video. The results are shown for obtained detections in the following images.



Figure 11- Annotated results after people detection stage.



Figure 12 - Annotated results showing false detection

## 5. Conclusions

The above experimental results and the literature reviews of various techniques, algorithms and applications of catadioptric camera systems lead to an obvious conclusion that catadioptric cameras offer significant advantages for scene analysis. The ability to capture large fields of view using the catadioptric camera combined with the application of well-tested computer vision algorithms on unwrapped views offer significant gains in using them as opposed to traditional multi-camera setups. Albeit this approach increases complexity significantly due to the process of unwrapping as well as camera calibration, both of which are highly critical in order to obtain accurate results from algorithms applied. Techniques mentioned as references in this report, such as unwrapping methods using GPU accelerated computation of lookup tables [8] can greatly reduce the computational complexities of unwrapping algorithms and produce near-real time results of unwrapped panoramic/perspective images. Apart from [8], other algorithms for unwrapping were studied as part of the literature survey and are mentioned in references.

### **A note on accuracy of detections**

It can be observed in fig. 12, that there is a false detection as well as the person on the extreme right is not detected. This occurs due to the training set used for the classifier. Since majority of the training set used, contains frontal views of people, those people who not in full view of the camera i.e., who's side profiles are visible, might get missed out by the detection algorithm. However by using a more robust training set, and combining the algorithm with results of motion-detection algorithms, a highly accurate efficient tracking framework using catadioptric camera systems can be developed.

### References

1. Shree K. Nayar, Simon Baker, “*A Theory of Single-Viewpoint Catadioptric Image Formation*”, International Journal of Computer Vision 35(2), 175–196 (1999)
2. Helder Araujo, João P. Barreto, “*Issues on the Geometry of Central Catadioptric Image Formation*”
3. Kai Guo and Zhuang Li, “*Image Reconstruction From Omni-Directional Camera*”, Technical Report ECE-2007-06, Boston University
4. Omar EL Kadmiri, Lhoussaine Masmoudi, “*An omnidirectional image unwrapping approach*”, International conference on Multimedia Computing and Systems (ICMCS'11)
5. Zhipeng Chen, Bingwei He, “*A modified unwrapping method for omnidirectional images*”, International Conference on Electric Information and Control Engineering (ICEICE), 2011
6. Chong et al., “*A closed form unwrapping method for a spherical omnidirectional view sensor*”, EURASIP Journal on Image and Video Processing 2013,
7. S. W. Jeng and W. H. Tsai, “*A Unified Approach to Unwrapping Omni-Images into Panoramic and Perspective-View Images Using Pano-Mapping Tables*”, IEEE International Conference on Systems, Man, and Cybernetics 2006
8. Ngan Soon Chong, “*A parallel root-finding method for omnidirectional image unwrapping*”,
9. Dalal, N. and B. Triggs. “*Histograms of Oriented Gradients for Human Detection*”, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 2005, pp. 886-893.
10. Piotr Dollár, Zhuowen Tu, Pietro Perona and Serge Belongie. “*Integral Channel Features*”. Proceedings of the British Machine Vision Conference, pages 91.1-91.11. BMVA Press, September 2009.
11. P. Dollár, C. Wojek, B. Schiele and P. Perona, “*Pedestrian Detection: A Benchmark*”, Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2009.
12. Viola and Jones, “*Rapid object detection using a boosted cascade of simple features*”, Computer Vision and Pattern Recognition, 2001
13. P. Dollár’s Computer Vision Toolbox for MATLAB, <https://github.com/pdollar/toolbox>
14. Videos shot with GoPano Plus, Retrieved from <https://www.gopano.com/shotwith/plus> on 27th July 2016 at 11:00 AM