PRO-CAM DISPLAY SYSTEMS

IAS SRFP Project Presentation

SHWETHA RAM
B.TECH 3RD YEAR, NITK SURATHKAL.

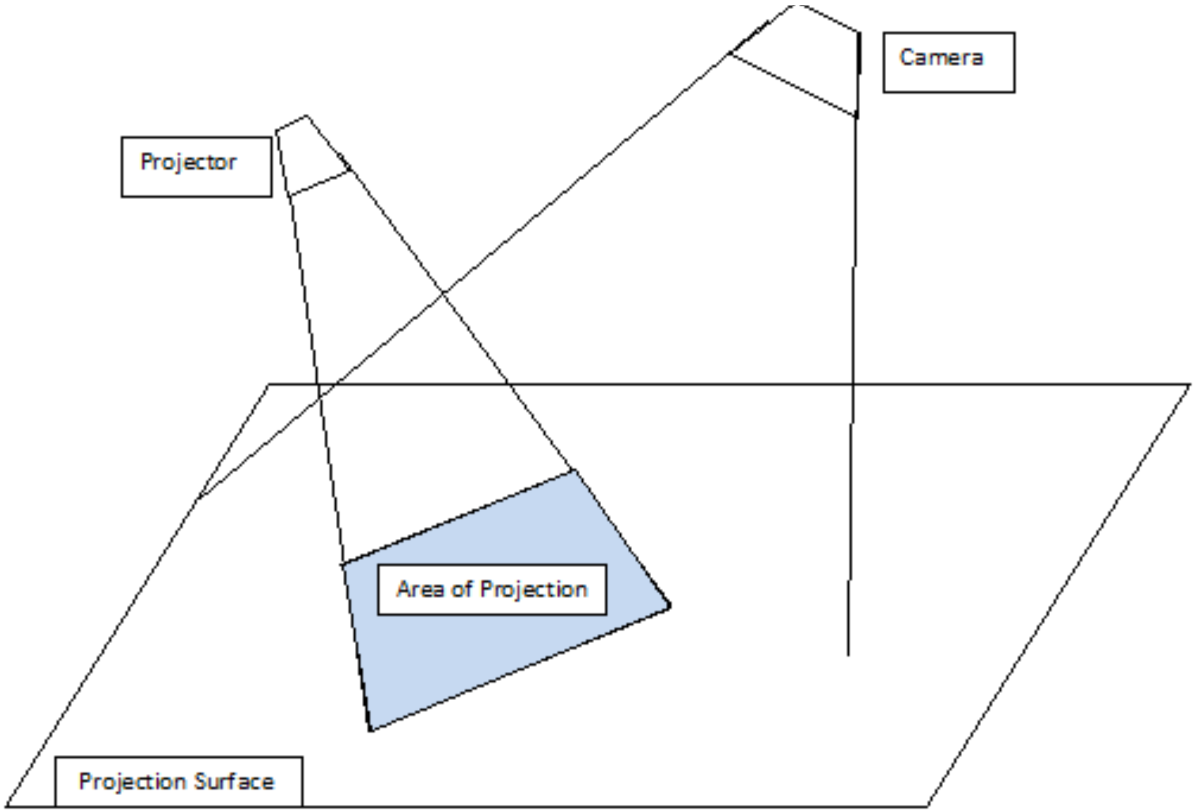
PROJECT GUIDE: DR.K.R.RAMAKRISHNAN
CVAI LAB, EE DEPT, IISC BANGALORE.

OUTLINE

- Introduction
- Problem Statement
- Related Work
- Static Display Systems
- Movable Hand-held Display Systems
- Using Depth Information
- Conclusion
- Future Work: Projection on Irregular Surfaces
- References

Introduction

The image projected by the projector is compensated to suit the view-point of the camera.



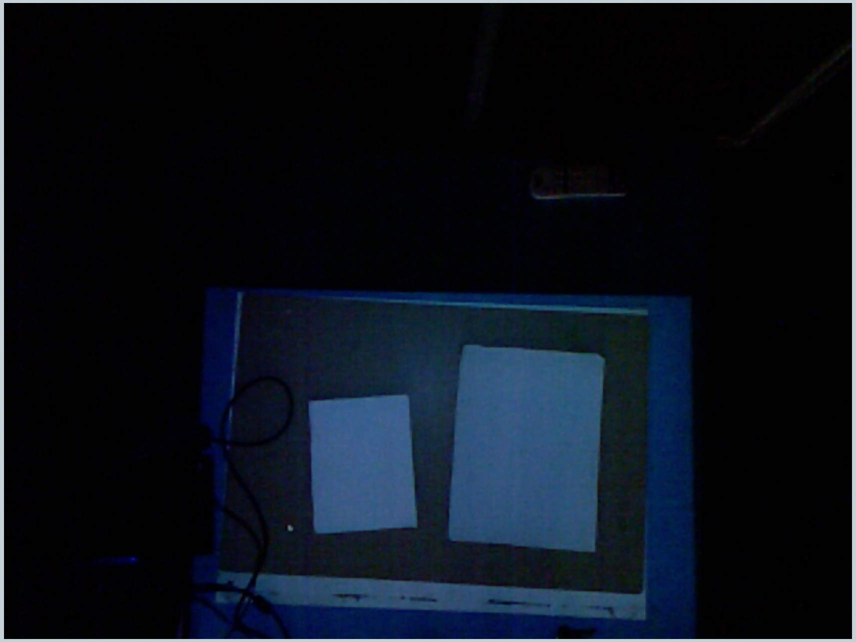
Problem Statement

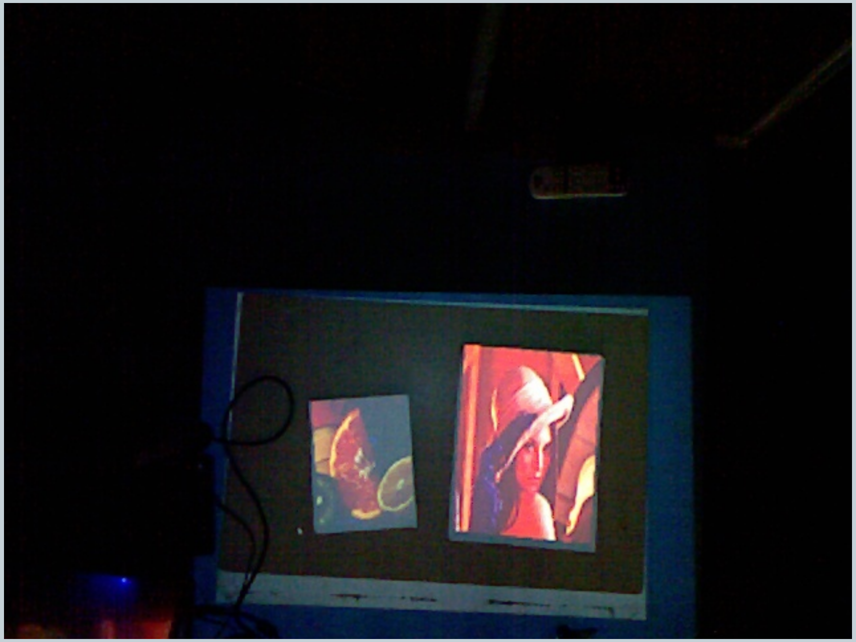
- Projecting display content on multiple display units using a single pro-cam system.
- Display units may be static/dynamic.
- No constraint on the number of display units or the size or movement of a display unit provided a display unit is always within the area of projection of the projector.

Related Work

- Sukthankar et al, J.C.Lee et al – Sensor based approaches.
- S. Gupta and C. Jaynes – Universal Media Book
- ‘A projector-based Movable Hand-held Display System’, Leung et al

Static Display System

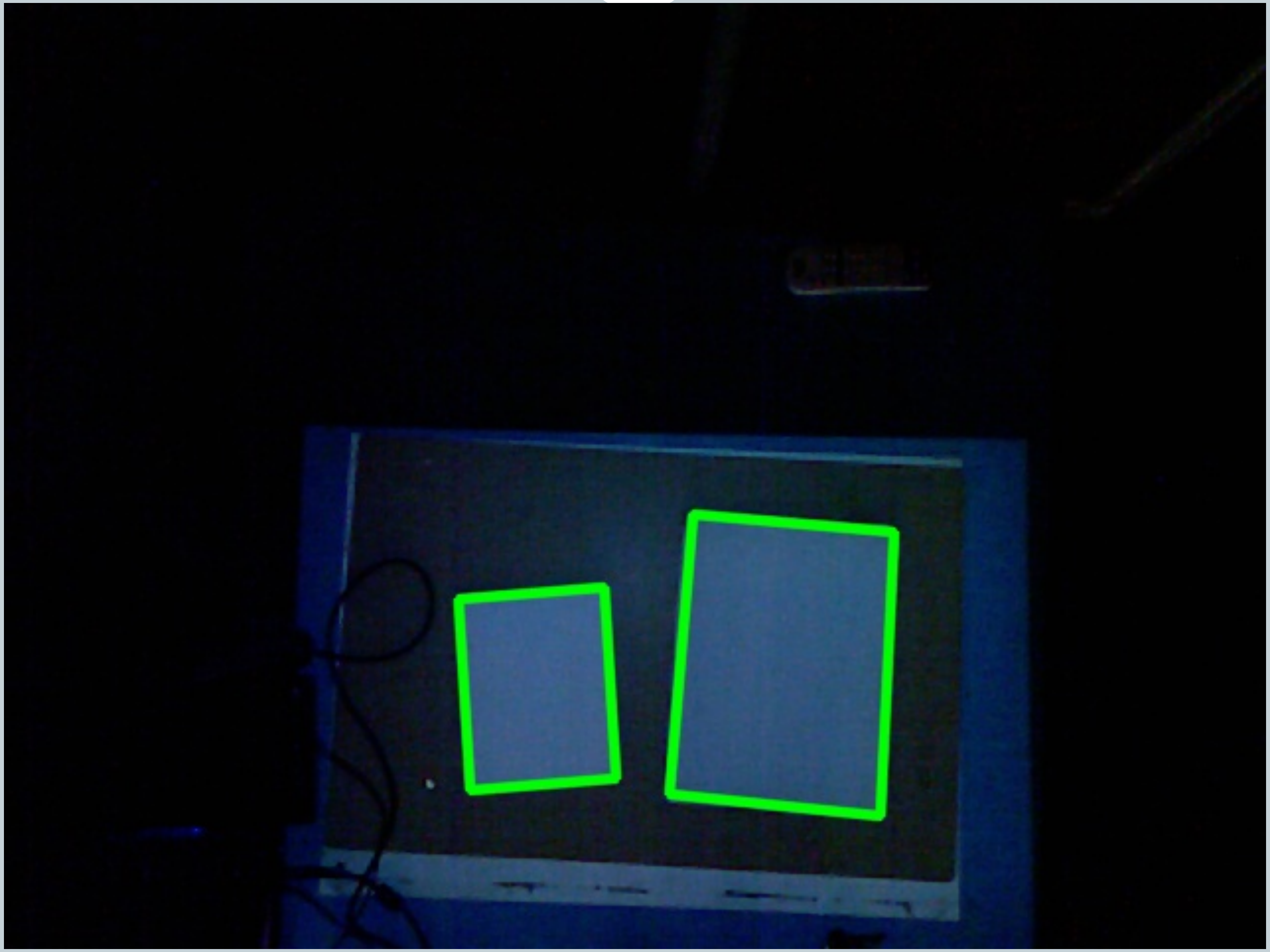




1. Detecting the Display Units

- Detect the outermost contours in the image (contours of hierarchy 1)
- A detected contour is considered to be representing a quadrangle if it satisfies the following conditions.
  - The contour should be convex.
  - The contour should have four vertices after approximation and a relatively large area (this is imposed in order to eliminate very small contours that can create noise).
  - The angles should lie in the specified range (close to 90 degrees).

Detected Quadrangles



Projecting Display Content

- Two projective Transformations need to be considered.
  - **Transformation from the source image to the detected quadrangle**
  - **Transformation from the camera image plane to the projector image plane**

Transformation from the Source Image to the Detected Quandrangle

- (x,y) are the transformed image coordinates and (X,Y) are the coordinates of source image.

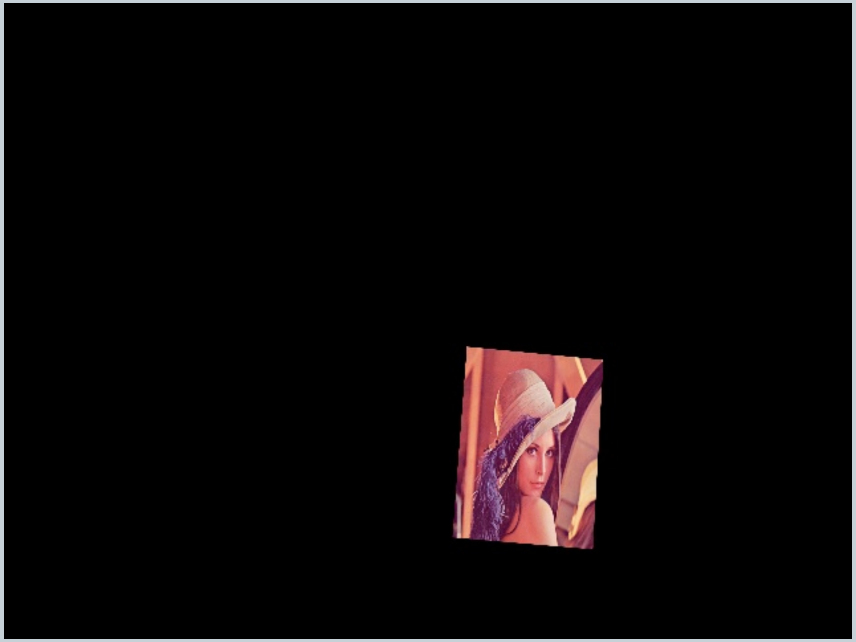
$$\mathbf{(x,y) = (M_{11}X+M_{12}Y+M_{13}/M_{31}X+M_{32}Y+M_{33},\ M_{21}X+M_{22}Y+M_{23}/\ M_{31}X+M_{32}Y+M_{33})}$$

- The four corners of the detected quadrangle are known. A four-point correspondence is known from which we solve for M.

Transformation from the Source Image to the Detected Quandrangle

Source Image S1

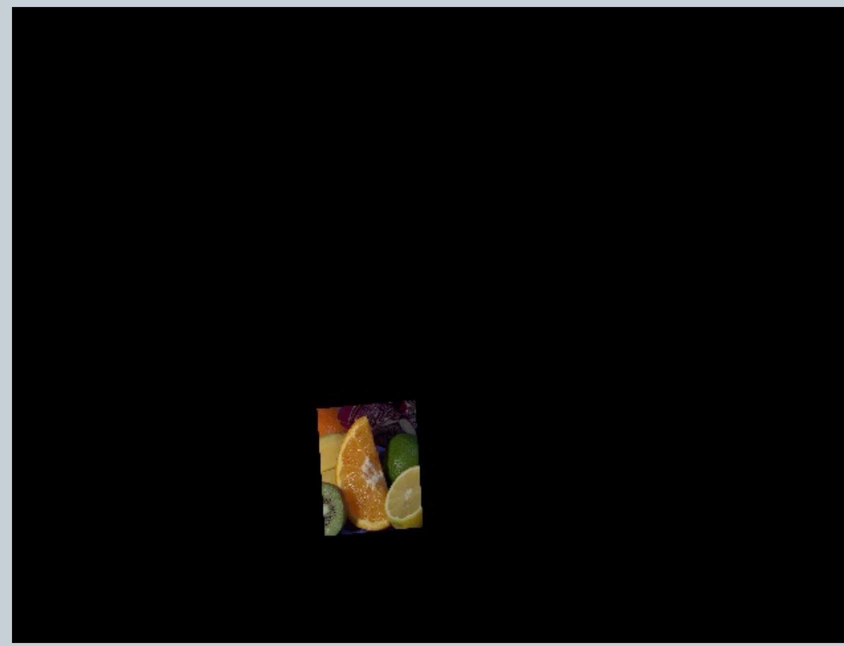
Transformed Image T1



Transformation from the Source Image to the Detected Quandrangle

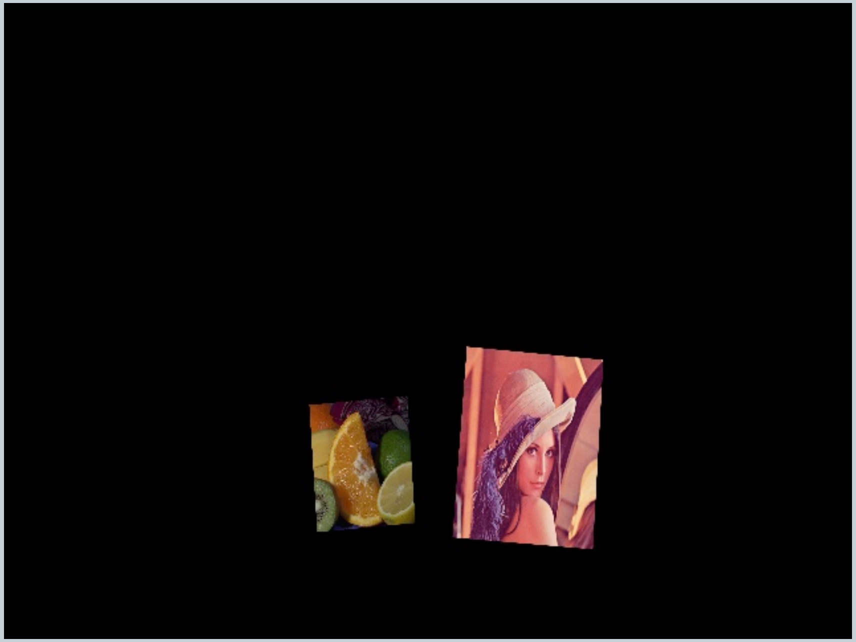
Source Image T2

Transformed Image T2



Final Transformed Image

- Let there be n display units.
- S1 to Sn are the source images to be displayed.
- T1 to Tn are the transformed images.
- Final image from camera's viewpoint T = T1+…+Tn.



Transformation from the Camera Image Plane to the Projector Image Plane

- When a source image is projected from the projector and this projection is viewed through the camera, the mapping between the source image and the camera image is given by a projective transformation.

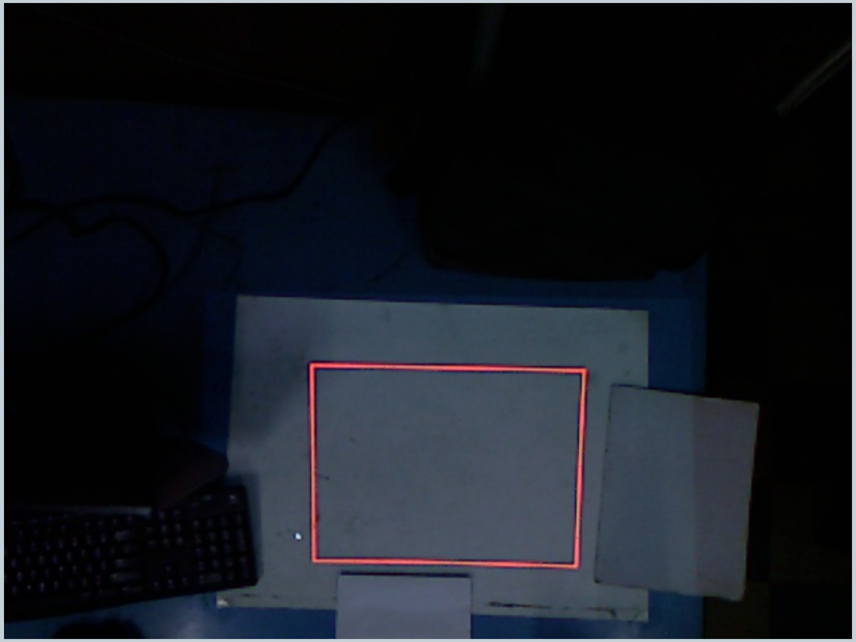
$$\mathbf{(x,y) = (H_{11}X+H_{12}Y+H_{13}/H_{31}X+H_{32}Y+H_{33},}$$
$$\mathbf{H_{21}X+H_{22}Y+H_{23}/\ H_{31}X+H_{32}Y+H_{33})}$$

- (x,y) are the coordinates of the camera plane image and (X,Y) are the coordinates of the projector plane image.

Finding H : Four-point Correspondence

Camera Image Plane

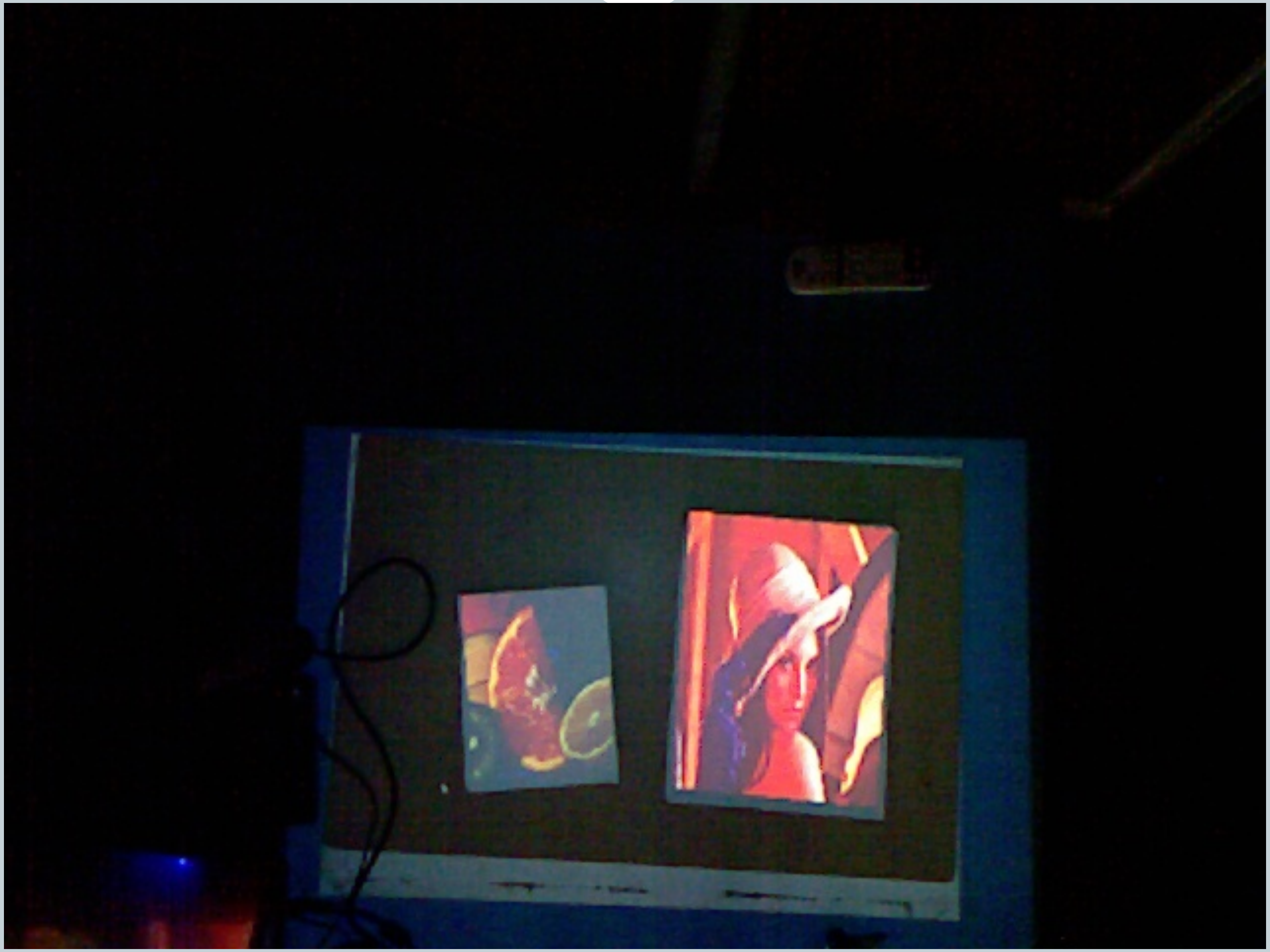
Projector Image Plane



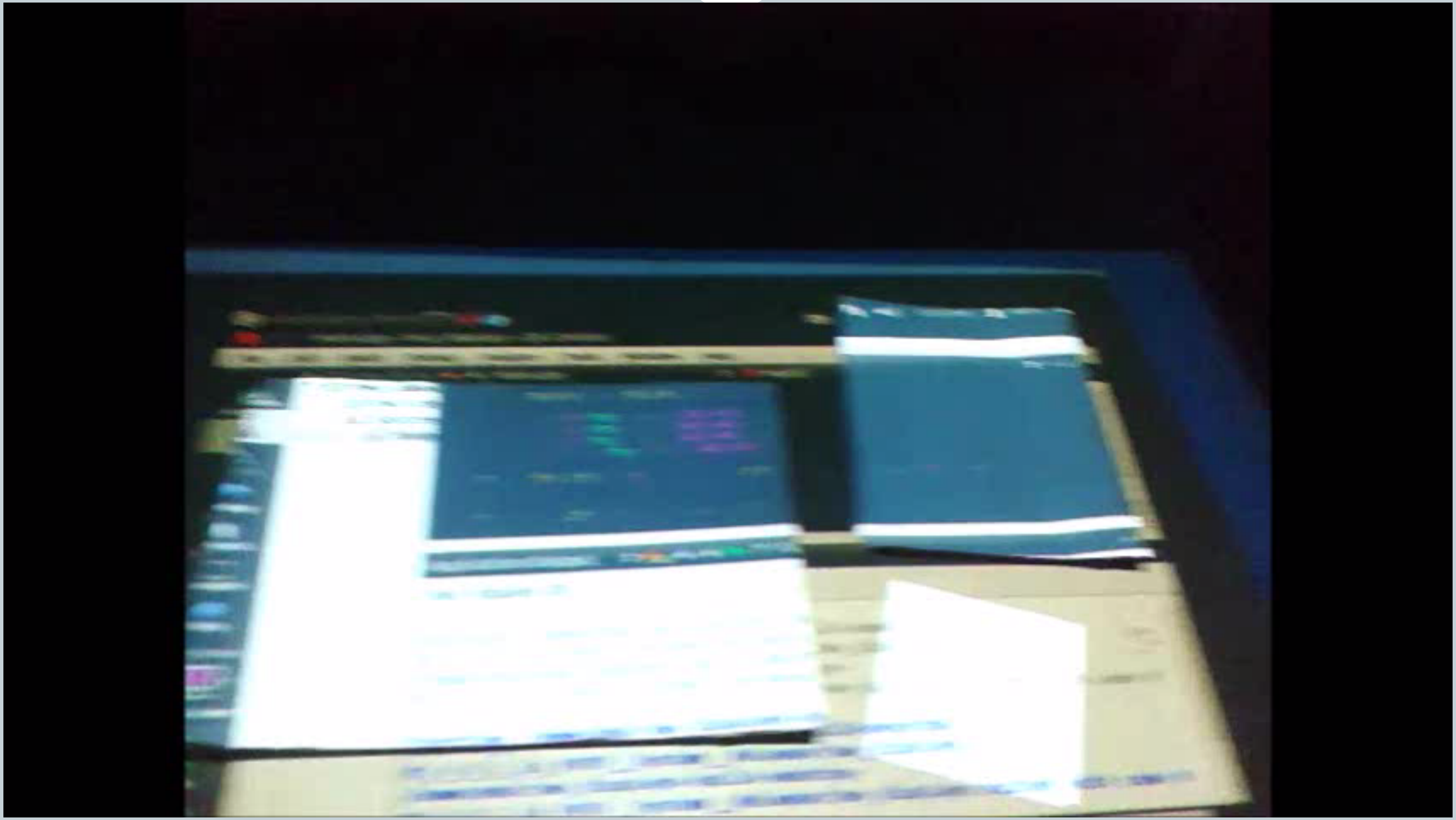
Final Image to be projected is $H^{-1}(T)$

The transformation H remains the same as long as there is no relative motion between the projector and the camera. Thus, it needs to be computed only once, when the system is initialized.

Result : Camera’s Viewpoint



Result



Hand-held Movable Display System

- As the display units are no longer constrained to be static, we have to track them in every frame.
- The transformation matrices Ms need to be computed in every frame.
- The Transformation matrix H needs to be computed only once, assuming there is no relative motion between the projector and the camera.

Tracking the Display Units : Kalman Filter

- Each corner of the display units is tracked separately.
- There are four trackers per display unit.
- One corner is essentially a particle moving in a plane with random perturbations in its trajectory.
- It is further assumed that the particle moves with constant velocity.

State Model

$$\begin{pmatrix} x1(t) \\ x2(t) \\ dx1(t) \\ dx2(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x1(t-1) \\ x2(t-1) \\ dx1(t-1) \\ dx2(t-1) \end{pmatrix} + \begin{pmatrix} wx1 \\ wx2 \\ wdx1 \\ wdx2 \end{pmatrix}$$

Here, x1 is the x-coordinate and x2 is the y-coordinate of the particle's position. dx1 and dx2 are the velocities along the x and y directions. W is the process noise matrix.

Observation Model

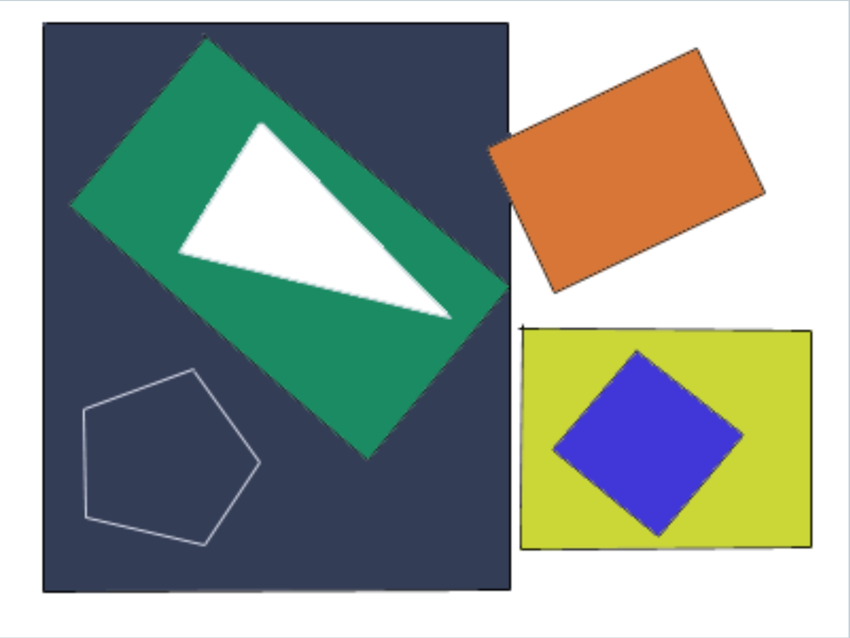
- Only the position of the particle is observed in the quadrangle detection result.

$$\begin{pmatrix} y1(t) \\ y2(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x1(t) \\ x2(t) \\ dx1(t) \\ dx2(t) \end{pmatrix} + \begin{pmatrix} vx1 \\ vx2 \end{pmatrix}$$

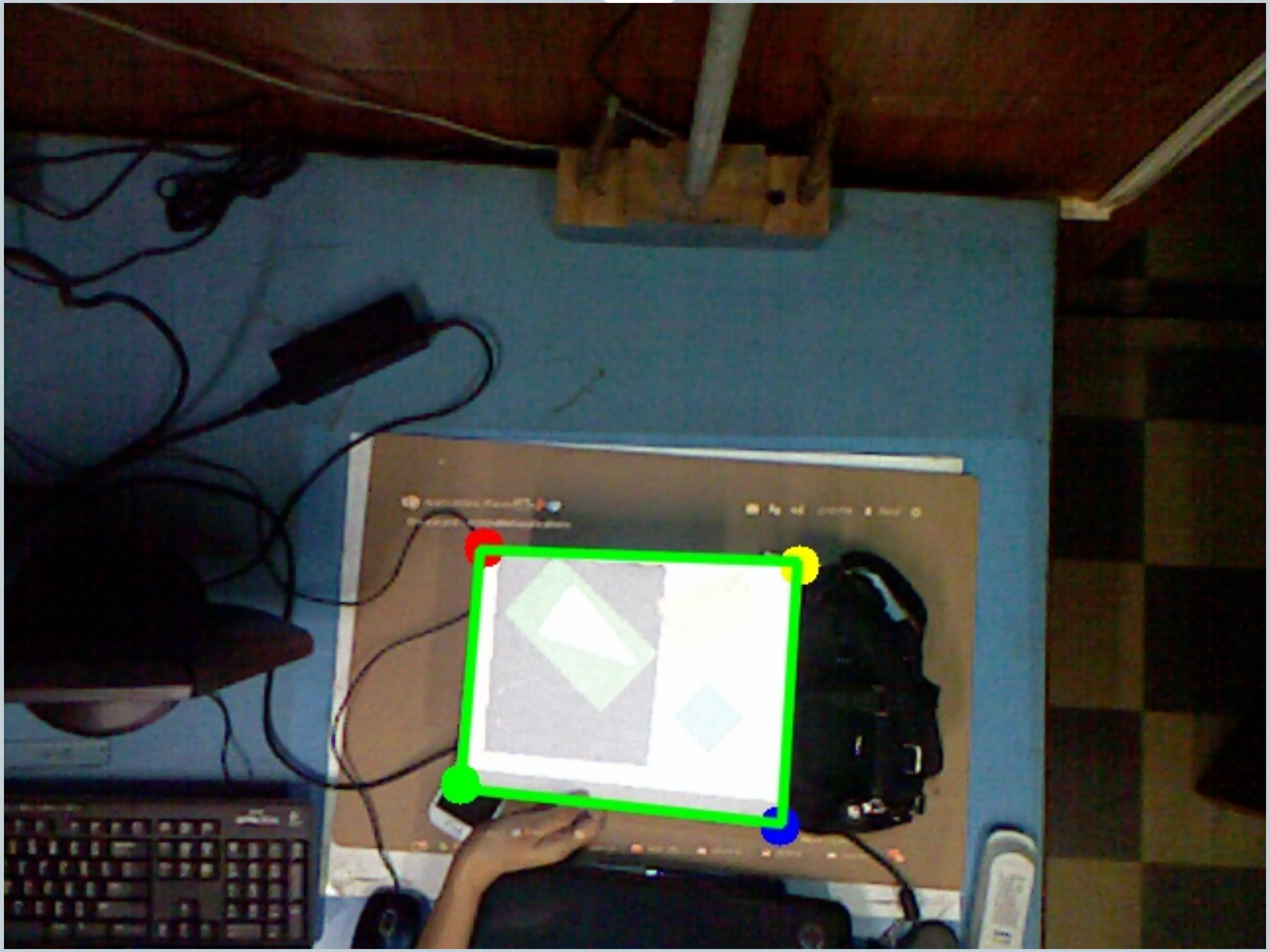
- The Observations y1 and y2 are the observed x and y coordinates of the particle's position. V is the measurement noise matrix.

Result

Input Image | **Source Image**

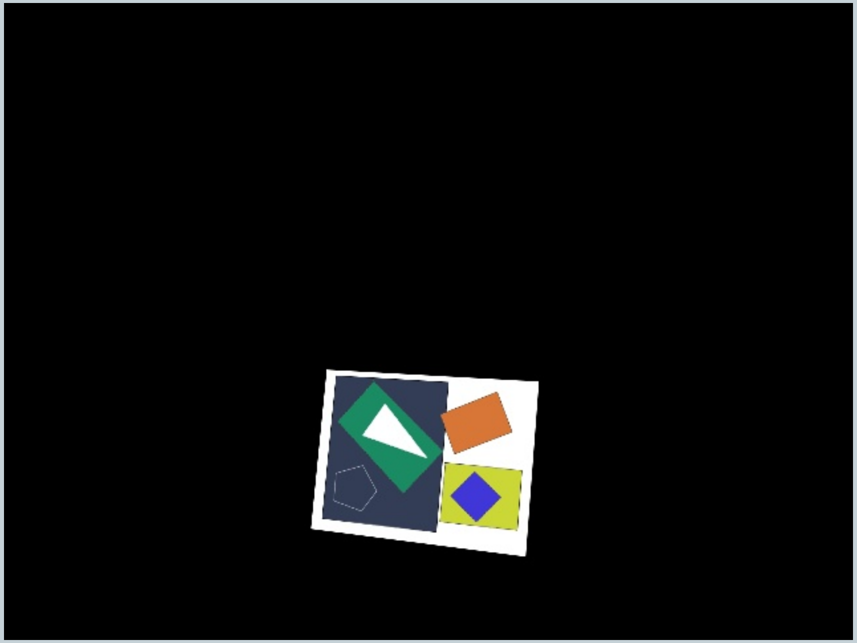


Result of Detection and tracking

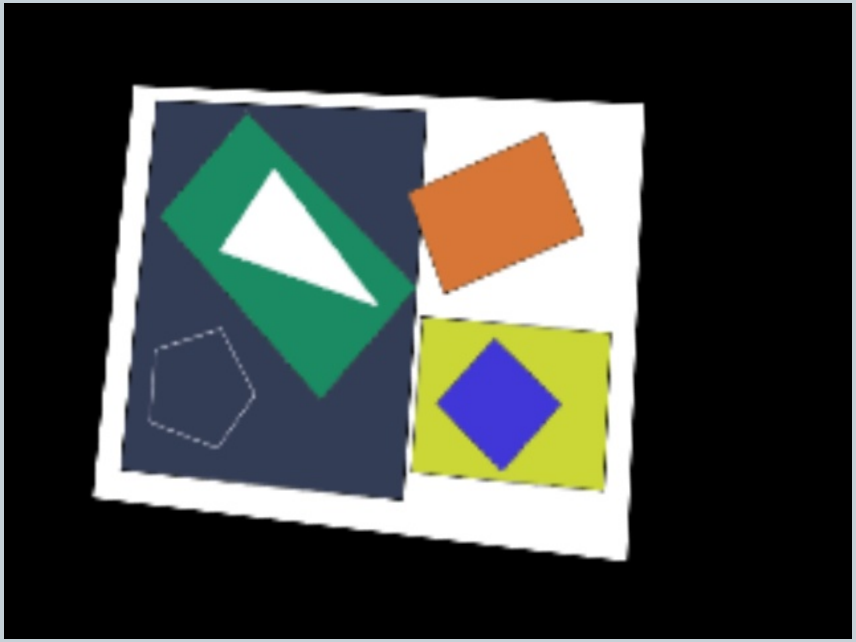


Result

Output-Camera image Plane



Output-Projector Image Plane



Results

- Video: single_display_result.avi
- Video: sigle_display_squares.avi
- Video: two_displays_rgb.avi

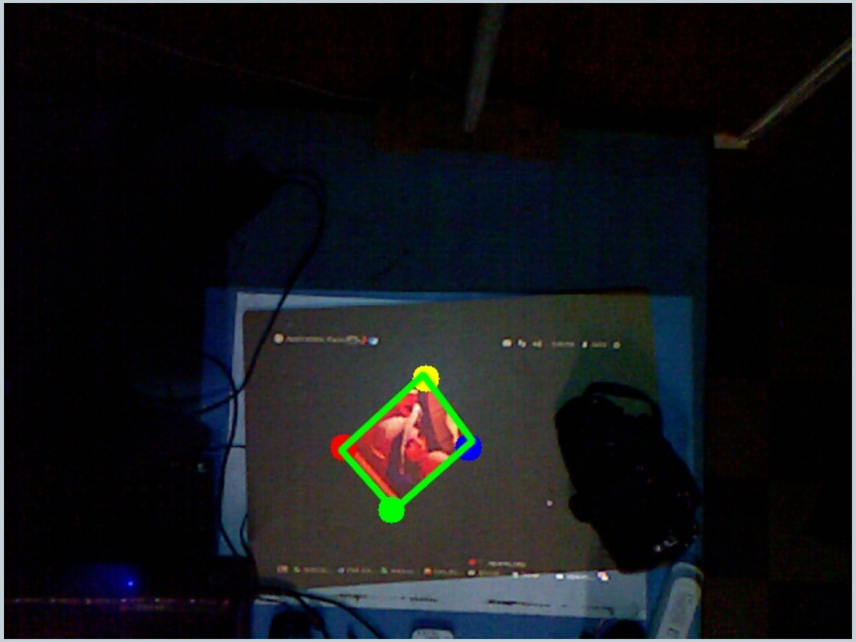
Comparison

- The case of a single hand-held movable display is considered for comparison with related works because this is the most implemented system.
- The system requires around 50ms for processing and projection which is comparable to the state of the art. This allows for real-time speed of 20 fps.
- The system requires no extensive calibration. The projective transform is established by a four-point correspondence, the determination of which takes less than half a second.
- There is no need for us to know the size of the cardboard in advance and there are no constraints on the movement of the cardboard.

Limitations of using only RGB image

- System does not function if display units and the background have the same colour.
- Detection of quadrangles from the projected display content.
- Projecting on other quadrangles in the frame which may not be Display Units.

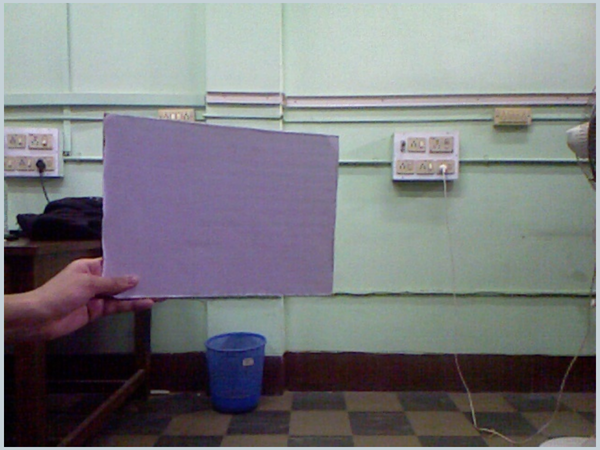
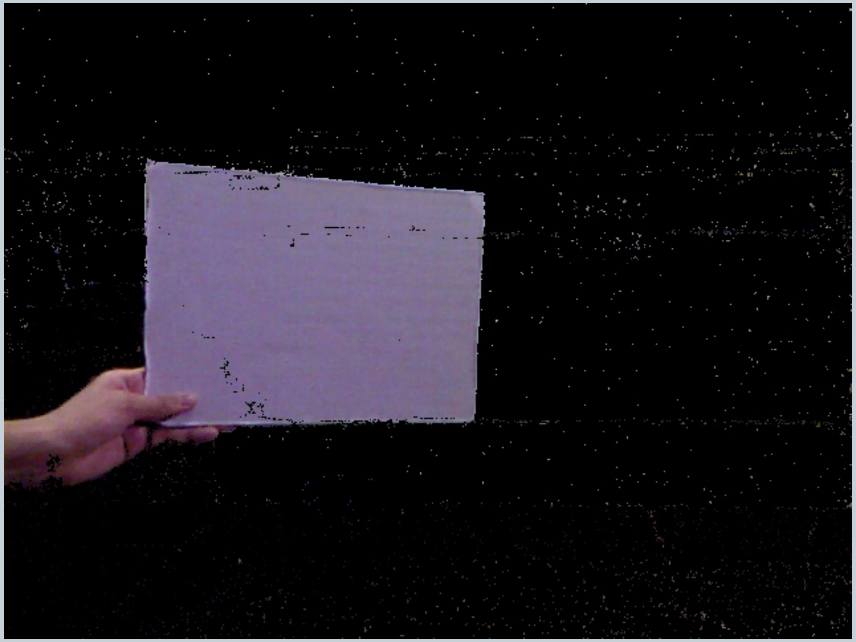
Using Depth Information

Background	Detection of quadrangle from projected display content
	

Project only on quadrangles which are moving

- Background Subtraction is performed on the depth image.
- Foreground mask obtained is applied to the RGB image.
- The foreground image now has the display unit and the hands holding it.
- Detect contours from this image and proceed as before.
- This system involves more processing. In order to make it real-time, a gpu implementation is required.

Result

Input Image	Back-ground Subtracted Image
	

- Video: output_using_depth.avi

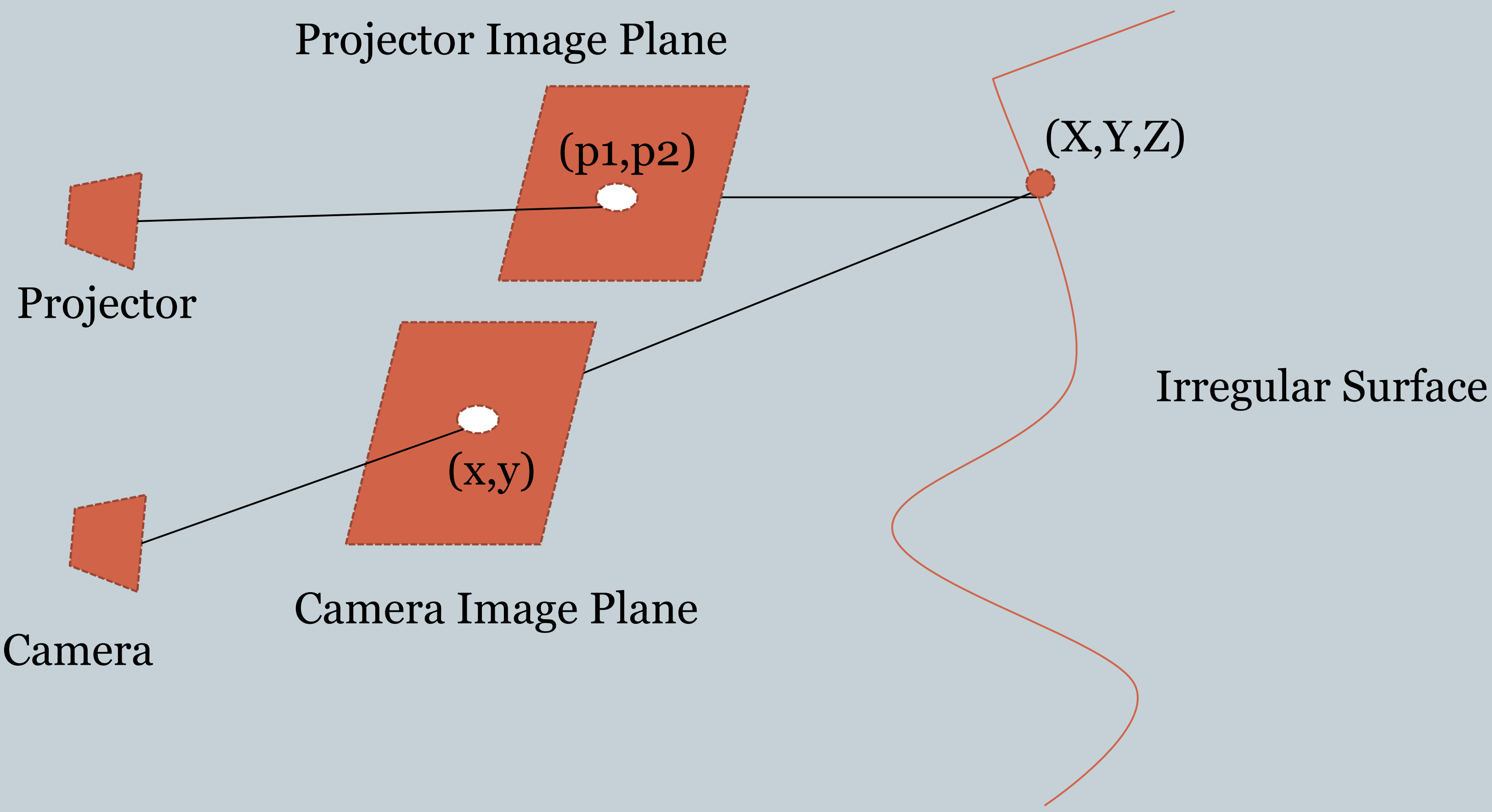
Conclusion

- The systems developed require no extensive calibration. By projecting a rectangle and taking the image of the projection, the projective transform is easily determined. The whole process takes less than half a second.
- The use of depth information in pro-cam display systems is introduced. This is conducive to developing a more robust application.

Conclusion

- Finding the Perspective Transformations (H and the Ms) involves solving a system of eight linear equations and can be done very quickly.
- The perspective warping needs to be done for every input frame and is computationally expensive. Its computational cost depends on the size of the source image ($O(mn)$ for an $m \times n$ image) and this puts an upper limit on the speed of the application.
- An efficient implementation would involve doing the warping for each source image parallely.

Future Work: Projecting on Irregular Surfaces



Future Work: Projecting on Irregular Surfaces

- Information projected at (p1,p2) in projector image plane is seen at (x,y) in camera image plane.
- So project the information desired at point (x,y) at point (p1,p2).
- Find mapping for all points.

$$(x,y) \rightarrow (X,Y,Z) \rightarrow (p1,p2)$$

- Intrinsic and extrinsic parameters of camera and projector need to be known.
- Greatly reduces the time and effort involved in calibration compared to Light Transport Matrix approach.
- When the surface of projection changes for given pro-cam system, re-calibration can be done very quickly.

References

- 'A Projector-based Movable Hand-held Display System', Leung et al, CVPR 2009.
- 'Automatic Keystone Correction for Camera-assisted Presentation Interfaces', Sukthankar et al, International Conference on Multimedia Interfaces, 2000.
- 'Stochastic Models, Estimation and Control- Volume 1', Peter S Maybeck.
- 'An Introduction to the Kalman Filter', Greg Welch and Gary Bishop.

References

- J. Summet and R. Sukthankar, "Tracking Locations of Moving Hand-Held Displays Using Projected Light", Proceedings of Pervasive 2005. Munich, Germany. Pp 37-46
- J. C. Lee *et al. "Moveable Interactive Projected Displays* Using Projector Based Tracking", Proceedings of the ACM Symposium on User Interface Software and Technology, October 2005, pp. 63-72.
- S. Gupta, C. Jaynes, "The Universal Media Book: Tracking and Augmenting Moving Surfaces with Projected Information". IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR), 2006, pp. 177-180
- 'Kalman Filter toolbox for MATLAB', Kevin Murphy.

References

- ‘Making One Object Look Like Another: Controlling Appearance Using a Projector-Camera System’, Grossberg et al, CVPR 2004.
- ‘The Visual Computing of Projector-Camera Systems’, Bimber et al, Eurographics 2007.
- ‘Fast High-Resolution Appearance Editing Using Superimposed Projections’, Aliaga et al,

Thank You