Kivy - Bucătăria modernă

1 Introduction

In interactive systems where a camera observes hand positions and a projector displays targets, it is often necessary to map coordinates from the **camera space** (where the user's finger is detected) to the **projector space** (where the interface is projected). This mapping can be achieved using a **homography matrix**, assuming both the projector plane and camera-captured hand landmarks lie approximately on the same physical plane (e.g., a table surface).

2 Homography Overview

A **homography** is a transformation that relates two planes in projective space. It is represented as a 3×3 matrix H that maps homogeneous coordinates from one image (or plane) to another:

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow \left(\frac{x'}{w'}, \frac{y'}{w'} \right)$$

where:

- (x, y) are the input coordinates (e.g., from the camera).
- (x', y') are the output coordinates (e.g., on the projector).
- H is the homography matrix.

3 Homography Matrix Structure

The homography matrix has 8 degrees of freedom (up to scale) and is typically written as:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}, \quad \text{with } h_{33} = 1$$

4 Constructing the System of Equations

Given 4 pairs of corresponding points $(x_i, y_i) \leftrightarrow (x'_i, y'_i)$, we can derive the constraints needed to solve for H.

From the homography relation:

$$x'_i = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}}$$
$$y'_i = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}}$$

Multiply both sides by the denominator to eliminate the fraction:

$$x'_i(h_{31}x_i + h_{32}y_i + h_{33}) = h_{11}x_i + h_{12}y_i + h_{13}$$
$$y'_i(h_{31}x_i + h_{32}y_i + h_{33}) = h_{21}x_i + h_{22}y_i + h_{23}$$

Rearranging, we obtain two linear equations per point:

$$-h_{11}x_i - h_{12}y_i - h_{13} + h_{31}x_ix'_i + h_{32}y_ix'_i + h_{33}x'_i = 0$$
$$-h_{21}x_i - h_{22}y_i - h_{23} + h_{31}x_iy'_i + h_{32}y_iy'_i + h_{33}y'_i = 0$$

5 Matrix Formulation

Stacking the equations for all 4 point pairs yields a linear system $A\mathbf{h} = 0$, where $\mathbf{h}$ is a 9-element vector:

$$\mathbf{h} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{21} & h_{22} & h_{23} & h_{31} & h_{32} & h_{33} \end{bmatrix}^T$$

Each point pair contributes two rows to matrix A:

$$A_i = \begin{bmatrix} -x_i & -y_i & -1 & 0 & 0 & 0 & x_ix'_i & y_ix'_i & x'_i \\ 0 & 0 & 0 & -x_i & -y_i & -1 & x_iy'_i & y_iy'_i & y'_i \end{bmatrix}$$

6 Solving the System

The system $A\mathbf{h} = 0$ is homogeneous. To find a nontrivial solution, we compute the **Singular Value Decomposition (SVD)** of A:

$$A = U\Sigma V^T$$

The solution $\mathbf{h}$ is the last column of V (corresponding to the smallest singular value). Reshaping $\mathbf{h}$ into a 3×3 matrix gives the homography matrix H.

7 Application: Hand-to-Projector Mapping

In the calibration process:

1. The user touches 4 projected dots on a flat surface.
2. The camera captures the index fingertip positions (x_i, y_i) in the image.
3. The system knows the corresponding projected points (x'_i, y'_i).
4. A homography H is computed to map camera points to projector points.
5. Later, any new finger position (x, y) from the camera can be mapped to the projector using:

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow \left(\frac{x'}{w'}, \frac{y'}{w'} \right)$$

8 Conclusion

The homography provides a mathematically sound and computationally efficient way to map points between the camera and projector planes. With just four well-chosen calibration points, we can enable precise interaction in spatial augmented reality systems using only a webcam and a projector.