Lecture 11: Homography Estimation using DLT - Short Notes

1. Homography and Point Correspondences

- Homography (H) is a 3x3 matrix that maps points from one image to another: x' ~ Hx.
- Uses homogeneous coordinates (e.g., (x, y, 1)).
- Each point pair gives 2 linear equations. Need at least 4 point pairs (8 equations).

2. Linear System and Solution

- Construct matrix A such that Ah = 0, where h is a 9x1 vector of H.
- Since we have 8 equations for 9 unknowns, the solution is unique up to scale.
- We solve using SVD (Singular Value Decomposition).

3. Degenerate Configurations

- A degenerate configuration is a point setup (e.g., 3 collinear points) that doesn't constrain H uniquely.
- Avoid collinear or clustered points. Prefer well-spread, non-collinear configurations.

4. DLT Algorithm

Steps:

1. Normalize points (centroid to origin, avg distance = sqrt(2)).
2. Build matrix A from point pairs.
3. Solve Ah = 0 using SVD.
4. Reshape h into H (3x3).
5. Denormalize: H = T'^-1 H T.

5. Verifying H

- Compute x'_hat = Hx and normalize it by dividing with 3rd coordinate.
- Check reprojection error: ||x'_hat - x'||. Small error = good mapping.

6. Points and Noise

- Use more than 4 point pairs to reduce error from noise.
- Overdetermined system (e.g., 6 or 8 pairs) is solved in least-squares sense using SVD.
- Normalization becomes critical when data is noisy or in large scale.

Flashcards (Q&A Review)

Q: How many linear equations does each point correspondence contribute?

A: 2 equations (1 for x', 1 for y').

Q: Why can we only recover H 'up to scale'?

A: Because in homogeneous coordinates, scaling the solution still satisfies the equations.

Lecture 11: Homography Estimation using DLT - Short Notes

Q: What happens if three of your four points are collinear?

A: The configuration is degenerate; it doesn't uniquely constrain H.

Q: Why do we normalize points before applying DLT?

A: To improve numerical stability by centering and scaling the data.

Q: What method is used to solve the matrix equation Ah = 0?

A: SVD (Singular Value Decomposition).

Q: What do we do at the end to get the final real-world H?

A: Denormalize using $H = T'?^{1} H T$.

Q: Why do we divide by the third coordinate after applying Hx?

A: To normalize the homogeneous result back to Cartesian coordinates.

Q: What does it mean if the reprojected point is far from the actual x'?

A: Your estimated H is inaccurate due to noise or poor configuration.

Q: Why is it better to use more than 4 points when noise is present?

A: To reduce risk of degenerate setups and get a more stable least-squares solution.

Q: What does normalization do to the coordinate values?

A: Centers them to mean 0 and scales them to avg distance sqrt(2).

Q: Which solution vector from SVD do we pick to solve Ah = 0?

A: The right singular vector corresponding to the smallest singular value.