🔑 Prior Terms You Must Know

Term	Meaning
Homography	A 3×3 matrix that maps points from one image plane to another in projective geometry.
Projective Space ($\mathbb{P}^2$)	A space where 2D points are represented in homogeneous coordinates $[x, y, w]$.
Correspondence	Matching the same point between two different images.
Over-constrained system	More equations than unknowns — no exact solution, so we estimate.
Least Squares	A method to find the best fit solution by minimizing the square of the error.
SVD (Singular Value Decomposition)	A method to solve linear systems like $Ah = 0$.
DLT (Direct Linear Transform)	An algorithm to estimate homography using linear equations.
RANSAC	A robust method to remove outliers when estimating transformations.
Bundle Adjustment	Joint optimization of points and camera parameters to minimize error.

✅ Q1. Why do we need at least 4 point correspondences to compute a homography?

- A homography matrix H has 9 elements, but is defined up to scale → **8 degrees of freedom.**
- Each correspondence gives **2 independent linear equations.**
- Therefore, $4 \text{ correspondences} \times 2 = 8$ equations → enough to solve for 8 DoFs.
- Fewer than 4 correspondences → underdetermined (no unique solution).
- More than 4 → overdetermined → solved using **least squares** (best-fit).

✅ Q2. Why is Singular Value Decomposition (SVD) used in DLT?

- We solve $Ah = 0$, a **homogeneous system.**
- **SVD** decomposes A into UDV^T, making it easy to find the solution.
- The solution h is the **right singular vector** of A corresponding to the **smallest singular value** in D.
- This gives the **best approximate solution** in the **least-squares sense.**
- It handles **noisy** and **overdetermined** systems well.

✅ Q3. What is the geometric meaning of $\vec{x}' = kH\vec{x}$?

- It represents a **projective transformation** between two images.
- Points are in **homogeneous coordinates**, so equality is **up to scale.**
- The vectors $\vec{x}'$ and $H\vec{x}$ are **collinear.**
- This means they point in the **same direction**, even if scaled.
- We often normalize by dividing by the last coordinate (e.g., $w = 1$).

🔍 1. Why estimation?

- In real-world images, measurements (points, features) are **not perfect**.
- There's **noise** due to lens, lighting, sensor, human error, etc.
- So, we **estimate** a transformation (like a homography) that "best fits" the data.

📐 2. When geometry meets statistics = Computer Vision

- Geometry handles shapes, transformations, mappings.
- Statistics helps us deal with **uncertainty**, **error**, **noise**.
- In vision, we combine both to **solve for transformation matrices** under uncertainty.

⚙️ 3. What is Homography?

- A **2D projective transformation.**
- A 3×3 matrix H maps 2D points in one image to another.

🔬 Four Major Estimation Problems in CV

Problem	What is given?	What to estimate?
1. 2D Homography	2D points $x_i \leftrightarrow x'_i$	Find H such that $x'_i = Hx_i$
2. 3D to 2D Projection	3D points X_i, image points x_i	Find camera projection matrix
3. Fundamental Matrix	Points in two images x_i, x'_i	Find matrix F linking epipolar geometry
4. Trifocal Tensor	Points in 3 images x_i, x'_i, x''_i	Find tensor T_{ijk} relating all views

✅ How many correspondences are needed?

- Homography H has **9 unknowns** (entries of a 3×3 matrix).
- But it's defined **up to scale**, so **only 8 degrees of freedom**.
- Each point correspondence gives **2 equations** (1 for x, 1 for y).
- So we need **at least 4 point correspondences** (4 × 2 = 8 equations).

🧮 If Points Are Perfect

- With 4 perfect correspondences, we can compute H exactly.
- The system $Ah = 0$ (where h is the vector form of H) has a **unique null space solution.**

🚫 What Goes Wrong in Real Life?

- **Measurement errors:** Even 1-2 pixel mismatch due to:
  - Human error
  - Lens distortion
  - Rounding/discretization
  - Imperfect feature detection
- These tiny mismatches can cause **big errors in** H.
- The transformation might warp or break the structure.

✅ Q2: What kind of system is formed when using more than 4 point correspondences?

- More than 4 points = **more than 8 equations** for 8 unknowns (in H).
- This creates an **overdetermined system**.
- An overdetermined system usually has **no exact solution**, because the equations are inconsistent due to noise.
- So, we aim to find the **best possible approximation** — a **least-squares solution**.

✅ Q3: What technique helps find the best solution when exact solution doesn't exist?

- The most commonly used technique is **Least Squares Estimation**.
- Specifically, **SVD (Singular Value Decomposition)** is used to solve the system $Ah = 0$ robustly.
- **RANSAC** is also used when you suspect some correspondences might be **outliers** (wrong matches).
- These methods help compute a **best-fit homography matrix** that minimizes error across all points.

How to Solve an Over-Constrained System?

In computer vision, many estimation problems involve **more equations than unknowns** due to multiple data points — this is called an **over-constrained system**.

We can't get an exact solution, so we try to find the **best possible approximate solution**.

🔷 Key Concepts and Steps

• Use **Least Squares** to minimize total error between predicted values and actual measurements.

• Propose an **objective function** — a mathematical expression that defines what you want to minimize (e.g., squared error).

• A good estimate has a **small squared error** between actual and predicted measurements.

🔷 The Gold Standard Algorithm

• Introduced by **Hartley and Zisserman** in their book "Multiple View Geometry".

• It's considered the **best possible estimation algorithm** under a given statistical assumption.

• **Definition**: An estimation algorithm is optimal **if it minimizes the most suitable cost function** for the task.

Why Not Always Use the Gold Standard?

• It's not always practical. Even if we know the optimal algorithm, it may be:

- **Too slow (computationally expensive)**
- **Too complex to implement**
- Not needed if an approximate solution is **"good enough"**

• Example: In the **Travelling Salesman Problem**, the exact algorithm is too expensive — so we use approximations.

• Always ask: **"Why this method?" Is it optimal for our objective, or just practical?**

Minimizing Error in Homography Estimation

- ✅ Use **more than 4 point correspondences** for robustness.
- ✅ Apply **RANSAC** to detect and remove outliers.
- ✅ Use **least squares optimization** for best-fit homography.
- ✅ Perform **bundle adjustment** to jointly refine parameters and minimize global error.
- ✅ Preprocess images:
  - Correct for **lens distortion**
  - **Normalize coordinates** to improve numerical stability

Numerical vs Analytical Solutions

Type	Description	Example
Analytical	Solves equations **exactly**, usually symbolically	DLT (Direct Linear Transform)
Numerical	Uses **approximation methods**, especially for large data	Iterative optimization, SVD

✅ Why Avoid the Optimal Algorithm?

Even when we know the theoretically **best (optimal)** algorithm for solving an estimation problem, we often **choose not to use it** in practice. Here's why:

🔹 Key Reasons

• **Computational Cost**:
The optimal algorithm may be **too slow** or **computationally expensive**, especially on large datasets or in real-time applications.

• **Complexity**:
Implementing the optimal method might be **too complex**, requiring advanced mathematical tools or iterative solvers.

• **Efficiency Tradeoff**:
Sometimes we prefer **faster, analytical solutions** (like DLT) that are **"good enough"** even if they're not the best in theory.

• **Practical Constraints**:
In time-sensitive or resource-limited applications, **approximate solutions** can save time, memory, or power.

Feature	Geometric Function	Algebraic Function
Measures	Actual pixel distance (e.g., Euclidean)	How well equations like `Ah ≈ 0` are satisfied
Example	Reprojection error: `	
Accuracy	High (real-world relevance)	Lower (less visually meaningful)
Complexity	Non-linear, needs iterative methods	Linear, solvable via SVD
Usage	Used for final refinement (e.g., bundle adjustment)	Used for initial estimate (DLT)

📌 3. DLT and Error Functions

- DLT minimizes **algebraic error**, not the actual **geometric error**.
- Using **more correspondences** helps mitigate the effect of pixel imprecision.
- Though **algebraic error is computationally simple**, it lacks **geometric intuition**.

1. What is 2D homography estimation, and in what kind of problems is it commonly used in computer vision?
2D homography estimation involves computing a projective transformation matrix that maps points from one image to corresponding points in another.
It is commonly used in image stitching, panorama creation, planar surface alignment, and perspective correction, where relationships between two views of the same scene are analyzed.

---

2. Why are at least 4 point correspondences needed to compute a 2D homography matrix using DLT? Explain with degrees of freedom.
The 2D homography matrix has 9 elements, but only 8 are independent due to the scale ambiguity (H is up to scale).
Each point correspondence provides 2 linear equations. Thus, a minimum of 4 point pairs gives 8 equations, which is just sufficient to solve for the 8 degrees of freedom in the homography matrix using the DLT method.

---

3. What are the key problems faced in 2D homography estimation in real-world scenarios?
Real-world point correspondences are rarely perfect due to noise, pixel quantization, lens distortion, or errors in feature matching.
Even a minor error (like 1-2 pixels) can lead to a significantly inaccurate homography matrix.
These inaccuracies make homography estimation sensitive and require robust approaches to handle imperfect data.

---

4. How does using more than 4 point correspondences improve the accuracy of homography estimation? What method is used to handle such over-constrained systems?
Using more than 4 correspondences introduces redundancy, allowing the averaging out of random measurement errors.
This over-constrained system is typically solved using **least squares** optimization, which finds the best-fit homography matrix that minimizes the overall error in a statistically optimal way.

---

5. Describe the DLT (Direct Linear Transformation) method and explain what kind of error it minimizes.
DLT is an analytical method that formulates a system of linear equations from point correspondences and solves it using techniques like SVD.
It minimizes **algebraic error**, which reflects how well the mathematical model fits the data but does not necessarily correspond to visual accuracy.
It is efficient and often used as an initial estimate before refinement.

---

6. What is the difference between algebraic and geometric error functions in homography estimation? Which is more accurate and why?
Algebraic error measures how well linear equations (like $Ah \approx 0$) are satisfied; it is easy to compute using linear algebra.

Geometric error, such as reprojection error, measures actual pixel distance between projected and true points, offering more visual accuracy.
Although geometric methods are more accurate, they require nonlinear optimization and are computationally more intensive.

---

7. What is the role of the Gold Standard algorithm in estimation problems, and why might it not always be used in practice?
The Gold Standard algorithm provides the most accurate estimate by minimizing the best cost function under certain assumptions, often involving geometric error.
However, it may not always be used due to its high computational cost, complexity, or the need for fast approximate solutions in real-time applications.
Sub-optimal but efficient methods like DLT are often used as practical alternatives.

---

8. Explain the concept of collinearity in homography estimation with reference to $xi' = kiHxi$. What does it imply geometrically?
In the equation $xi' = kiHxi$, the resulting vectors are scalar multiples of each other, meaning they lie in the same direction.
This implies **collinearity** between the original point and the transformed point in homogeneous coordinates.
Geometrically, their cross product is zero, indicating perfect alignment in projective space.

---

9. What is RANSAC and how does it contribute to improving the quality of homography estimation?
RANSAC (Random Sample Consensus) is a robust algorithm that helps in estimating parameters like homography by repeatedly selecting random subsets of data.
It identifies and eliminates **outliers** in point correspondences, retaining only the inliers to compute a more accurate transformation.
This leads to a much more reliable and noise-resistant homography matrix.

---

10. How is Singular Value Decomposition (SVD) used in the DLT algorithm to compute the homography matrix? What role does the smallest singular value play?
In DLT, the system of equations is represented as Ah = 0.
SVD decomposes the matrix A into UDV^T, and the solution **h** (the flattened homography matrix) is the **last column of V**, which corresponds to the smallest singular value.
This vector represents the direction in which the error is minimized in the null space of A.