Chapters 1 – 10: Overview

- Chapter 1: Introduction
- Chapters 2 – 4: Data acquisition
- Chapters 5 – 10: Data manipulation
  - Chapter 5: Vertical imagery
  - Chapter 6: Image coordinate measurements and refinements
  - Chapters 7 – 10: Mathematical model, bundle block adjustment, integrated sensor orientation, and direct geo-referencing
- This chapter will cover the automated generation of 3D positional information from overlapping imagery: Digital Image Matching.

CE 59700: Chapter 11

Digital Image Matching

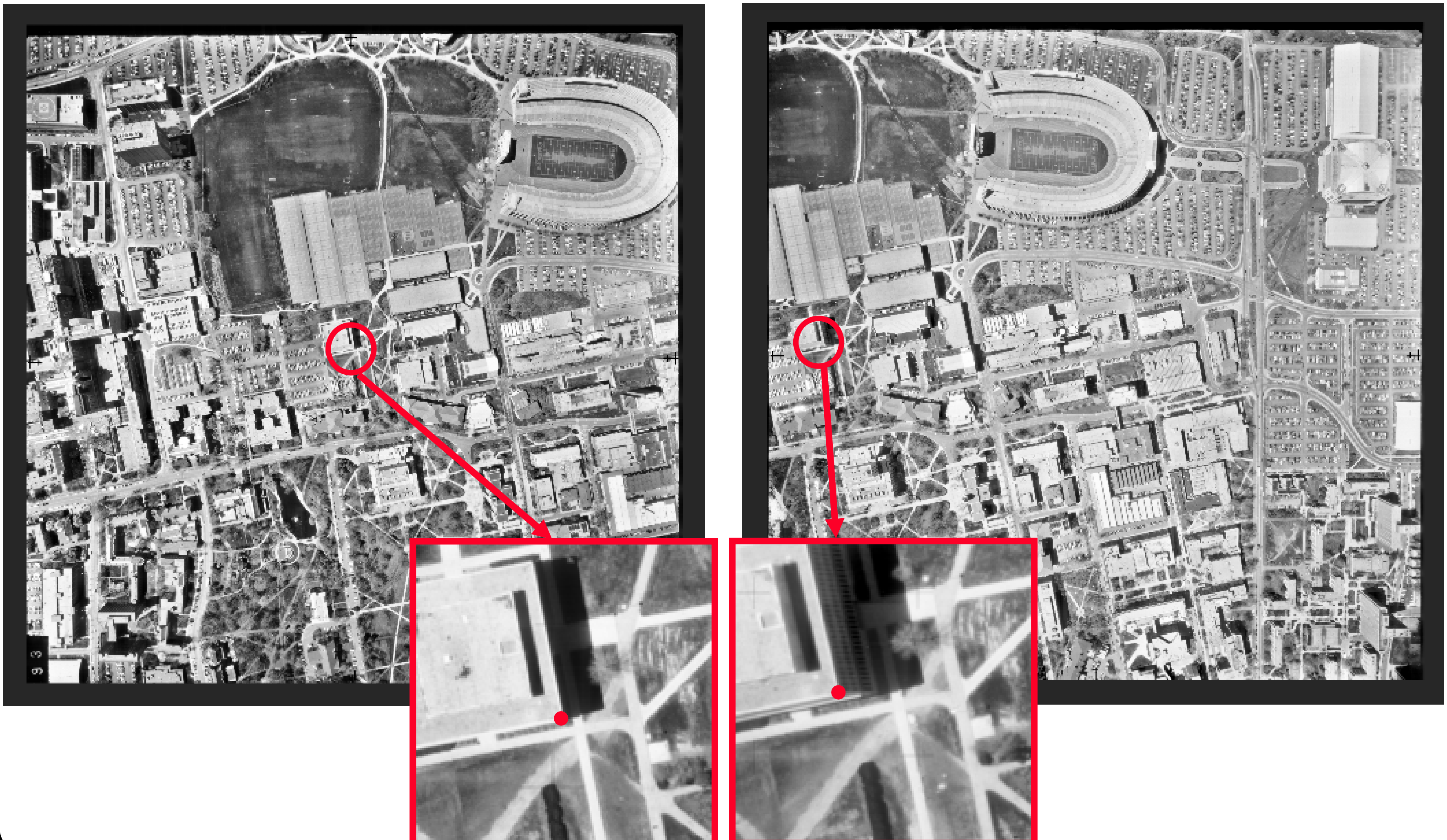
Digital Image Matching - Overview

- Objectives
- Terminology
- Problem statement and fundamental issues
- Area based matching
  - Correlation
  - Least squares
- Image Resampling according to epipolar geometry

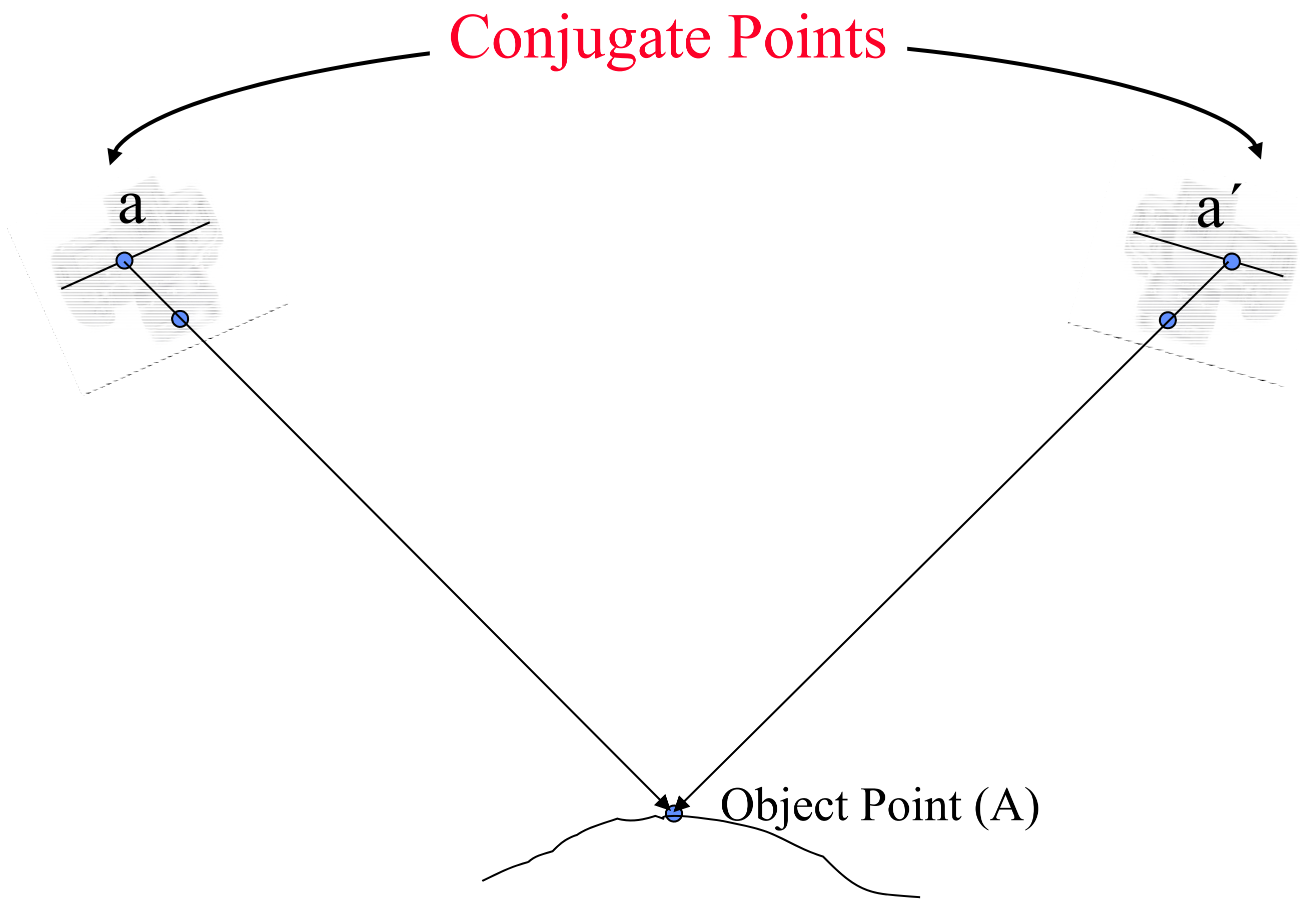
Digital Image Matching

- Objective:
  - Automatic matching of conjugate points and/or entities in overlapping images
- Applications include:
  - Automatic relative orientation
  - Automatic aerial triangulation
  - Automatic DEM generation
  - Automatic ortho-photo generation

Image Matching



3D Reconstruction



- The position and orientation of each camera station have to be known.

Terminology

- Conjugate entities:
  - Corresponding entities which are images of object space features such as points, lines, and areas
- Matching entity:
  - The primitive that is being compared with similar primitives in overlapping images to find conjugate entities
- Similarity measure:
  - A quantitative measure of how well the matching entities correspond to each other

Terminology

- Matching method:
  - Evaluate the similarity measure between the matching entities, such as:
    - Area-based matching,
    - Feature-based matching, and
    - Relational (symbolic) matching
- Matching strategy:
  - Refers to the concept or overall scheme of the solution to the matching problem. This entails the analysis of the matching environment, the selection of the matching method, and the quality control of the matching results

Problem Statement

- The problem of image matching and surface reconstruction can be summarized as follows:
  - Select a matching entity in one image,
  - Find the conjugate entity in the overlapping image,
  - Asses the quality of the match, and
  - Compute the three-dimensional location of the matched entity in the object space.
- The second task is the most difficult to solve and is the topic of our discussion.

Automatic Determination of Conjugate Points

Left Image

Right Image

Image Matching

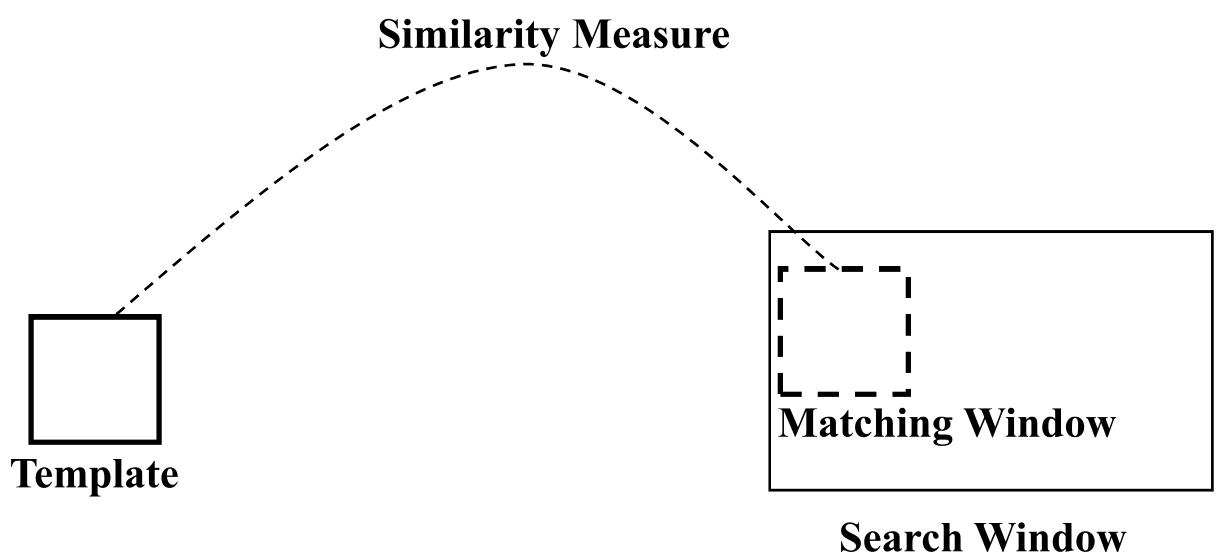
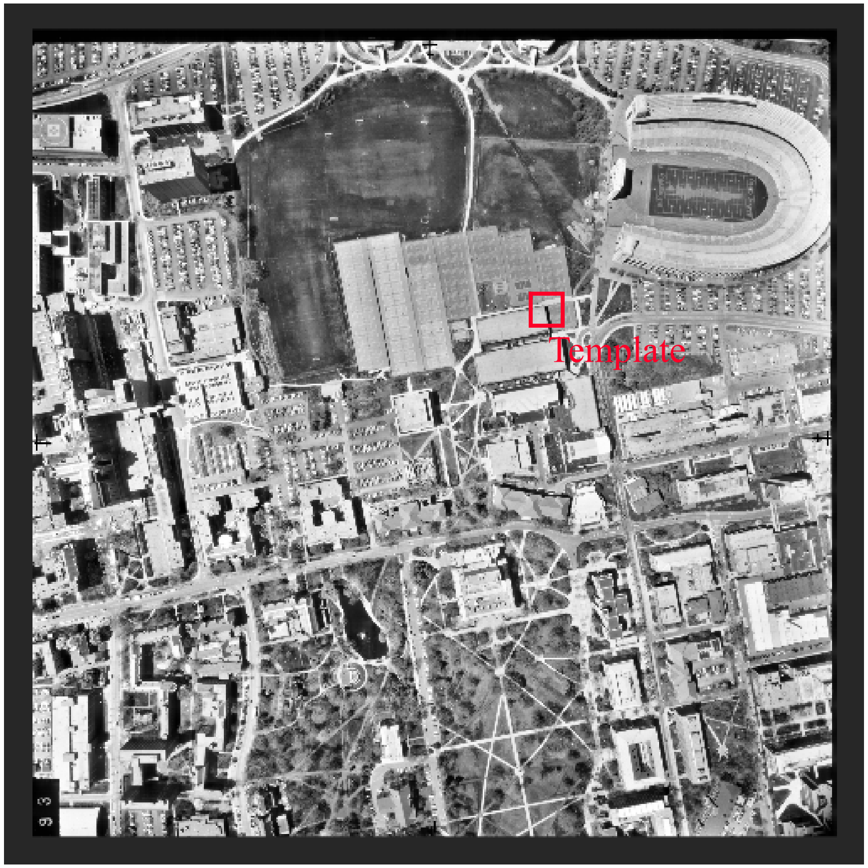
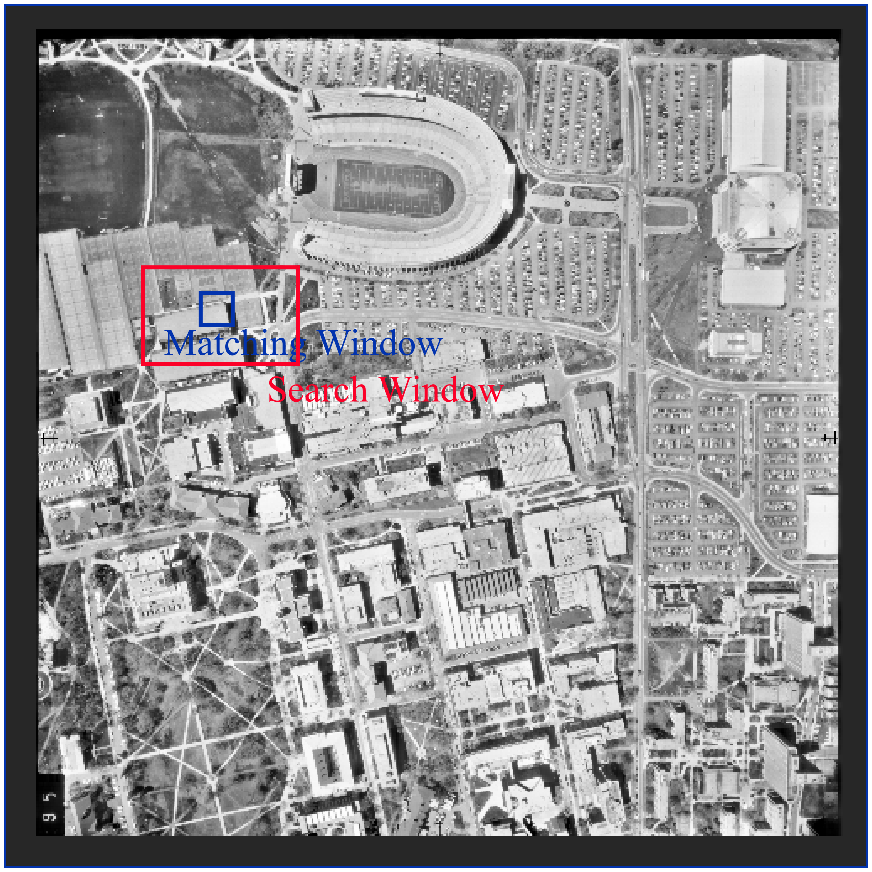


Image Matching





Fundamental Issues in Image Matching

- Search Space:
  - How can we determine the location (the center) and the size of the search space?
  - One should reduce the size of the search space to avoid combinatorial explosion which will be the case if the search space covers the whole image.
- The uniqueness of the Matching Entity:
  - The matching entity should be unique to avoid ambiguities in the matching process.
  - Interest points are used to avoid this problem.

Moravec Interest Operators

- This operator finds areas within the image with high variances.
- Variances along different directions, computed using all pixels in a window centered at a point, are good measures of the distinctness of the point.
- Variances are computed in windows with size ranging from 5X5 to 11X11 pixels.

Moravec Operator (Variance Computation)

$$I_1 = \sum_{(x,\,y)\in S} [f(x,\, y) - f(x,\, y+1)]^2$$

$$I_2 = \sum_{(x,\,y)\in S} [f(x,\, y) - f(x+1,\, y)]^2$$

$$I_3 = \sum_{(x,\,y)\in S} [f(x,\, y) - f(x+1,\, y+1)]^2$$

$$I_4 = \sum_{(x,\,y)\in S} [f(x,\, y) - f(x+1,\, y-1)]^2$$

Where S represent all the pixels in the window

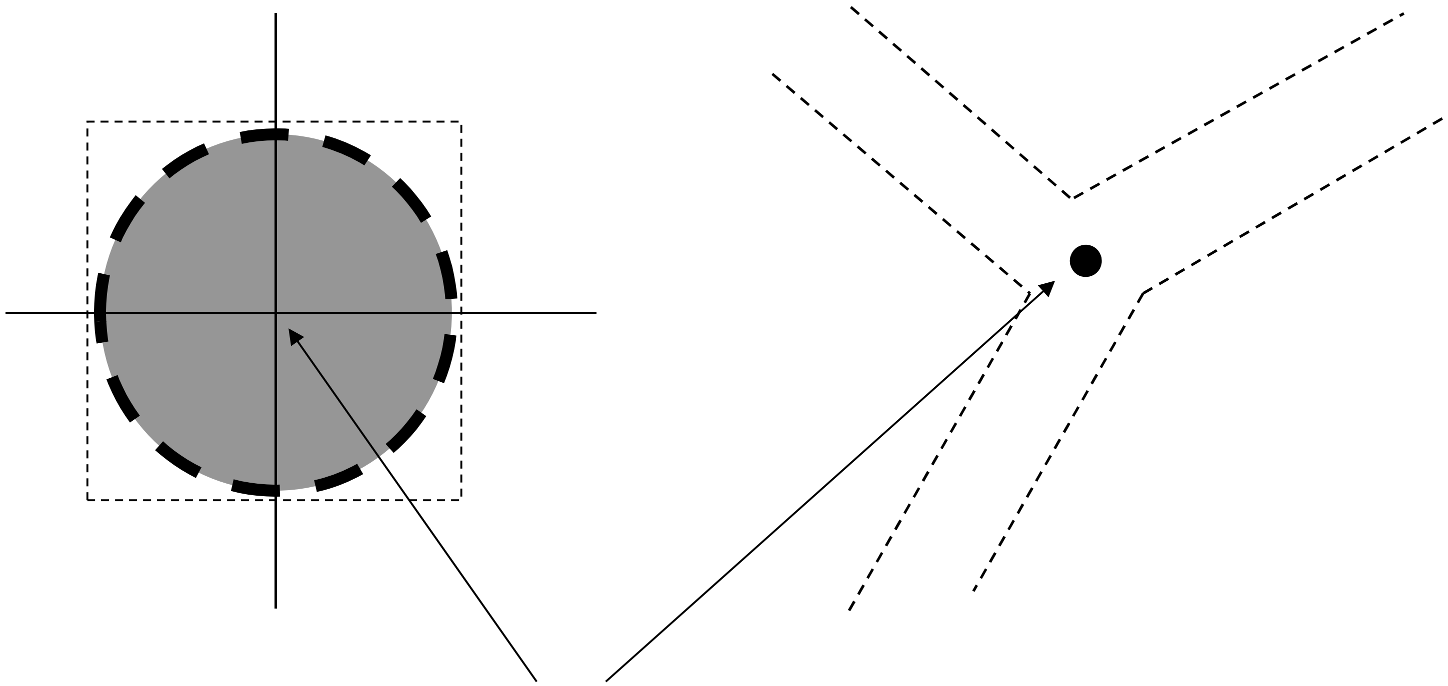
Moravec Interest Operators

- Edge pixels have no variance along the edge direction.
- The minimum value of the previous directional variances are taken as the interest value at the central pixel, (x_c, y_c):
  - $I(x_c, y_c) = \min(I_1, I_2, I_3, I_4)$
- Interest points if the interest value at their locations are local maxima and exceed a predefined threshold.

Föerstner Interest Operator

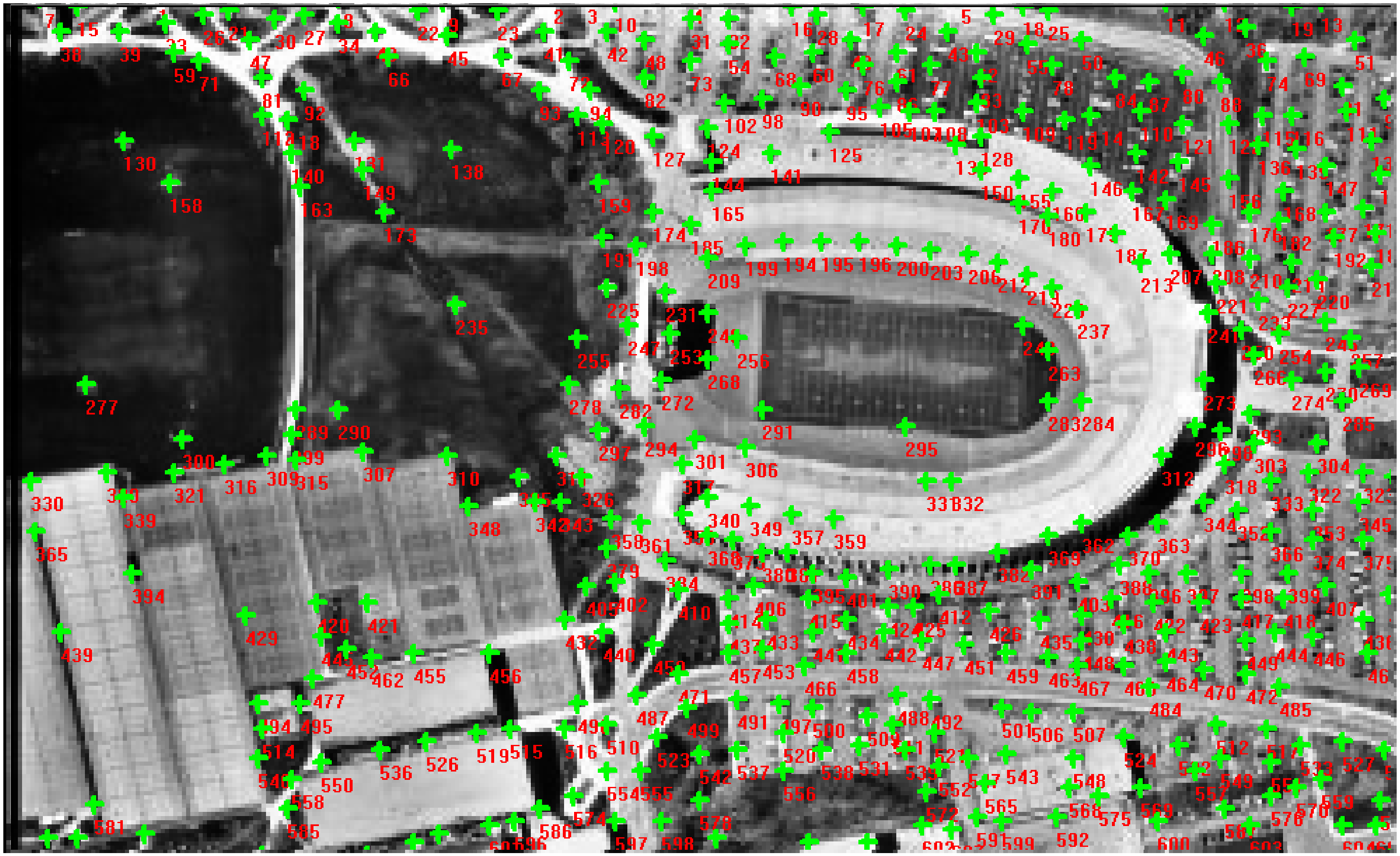
- Points of interest can be classified as corner and blob points. These points can be described as follows:
  - Corner is the closest point to all lines intersecting at that point.
  - Blob is the point that is closest to all lines that are normal to the tangents at the blob borders.
- Interest points are extracted after an estimation procedure (quantitative measures are included).

Föerstner Interest Operator



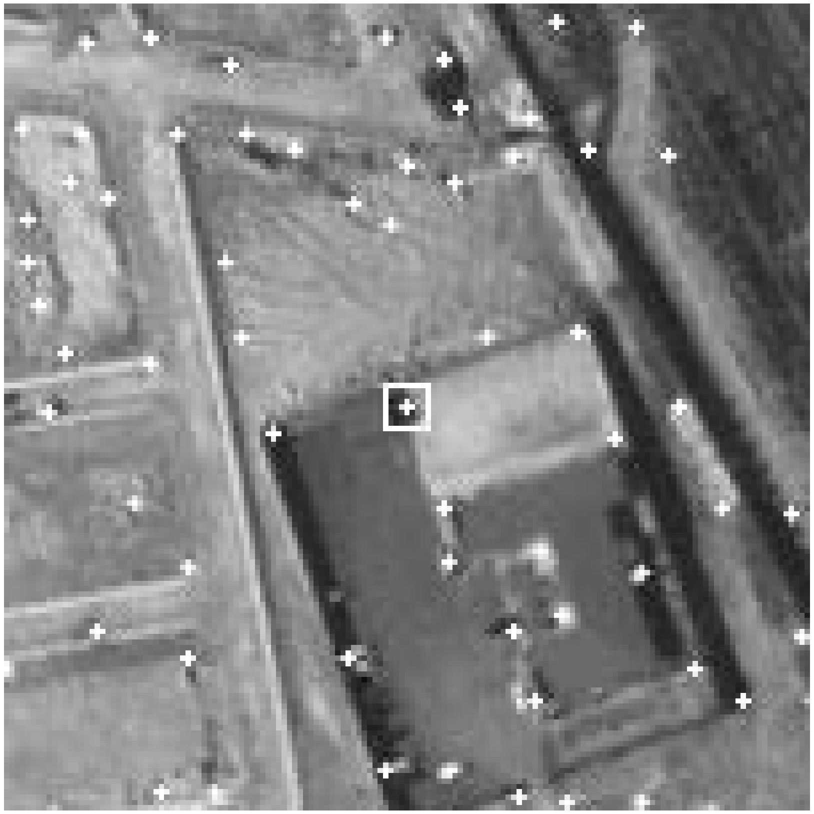
Representative Points

Interest Points

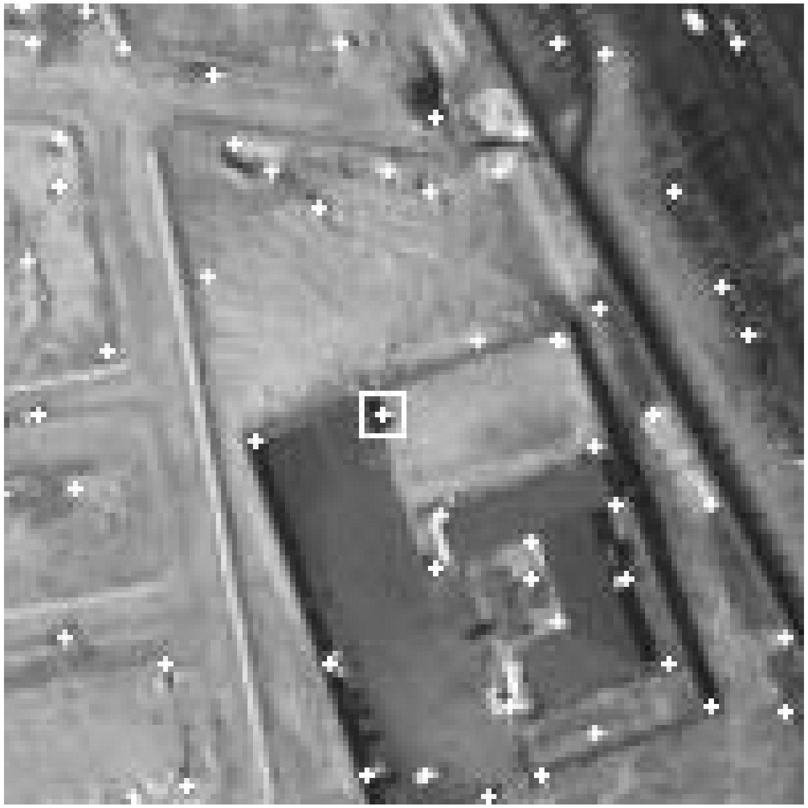


Automatic Determination of Conjugate Points

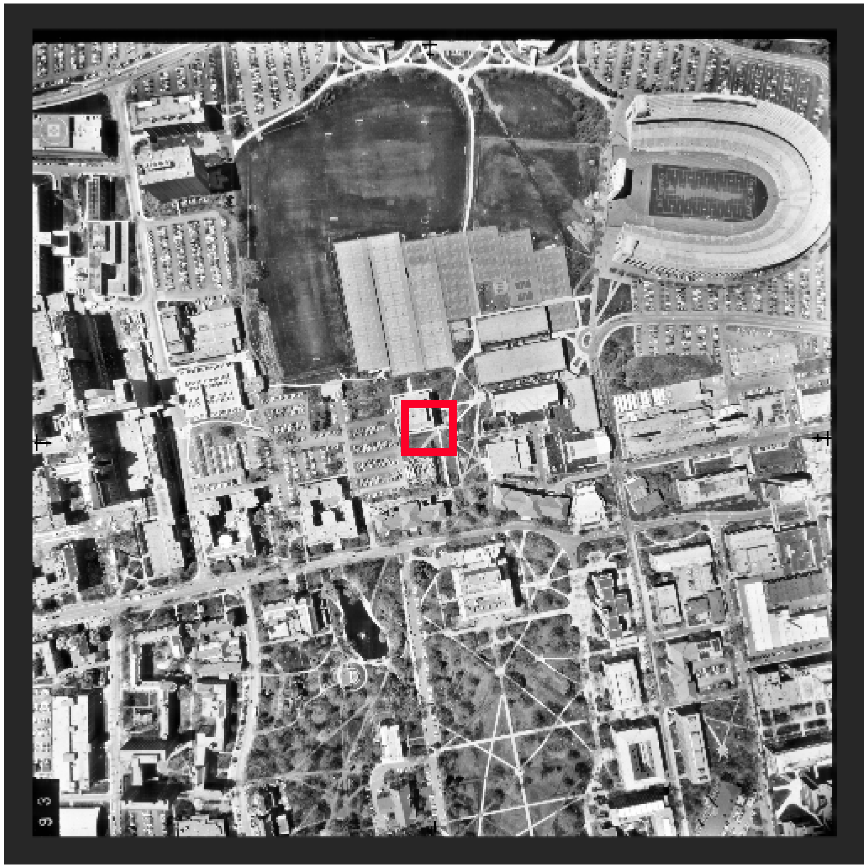
Left Image

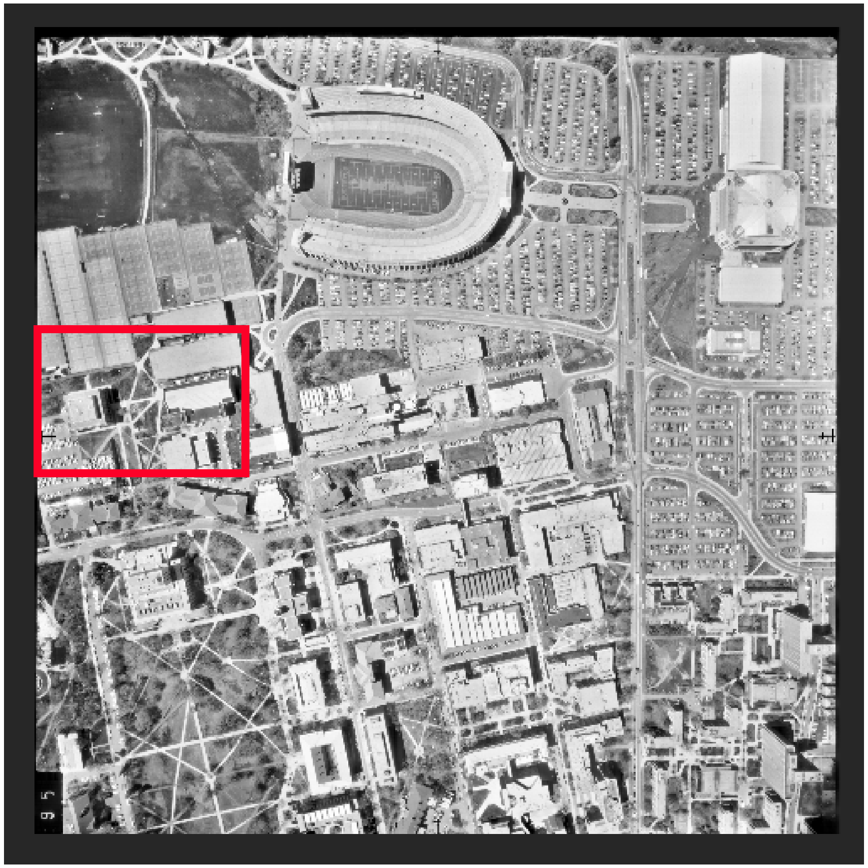


Right Image



Determination of the Search Space

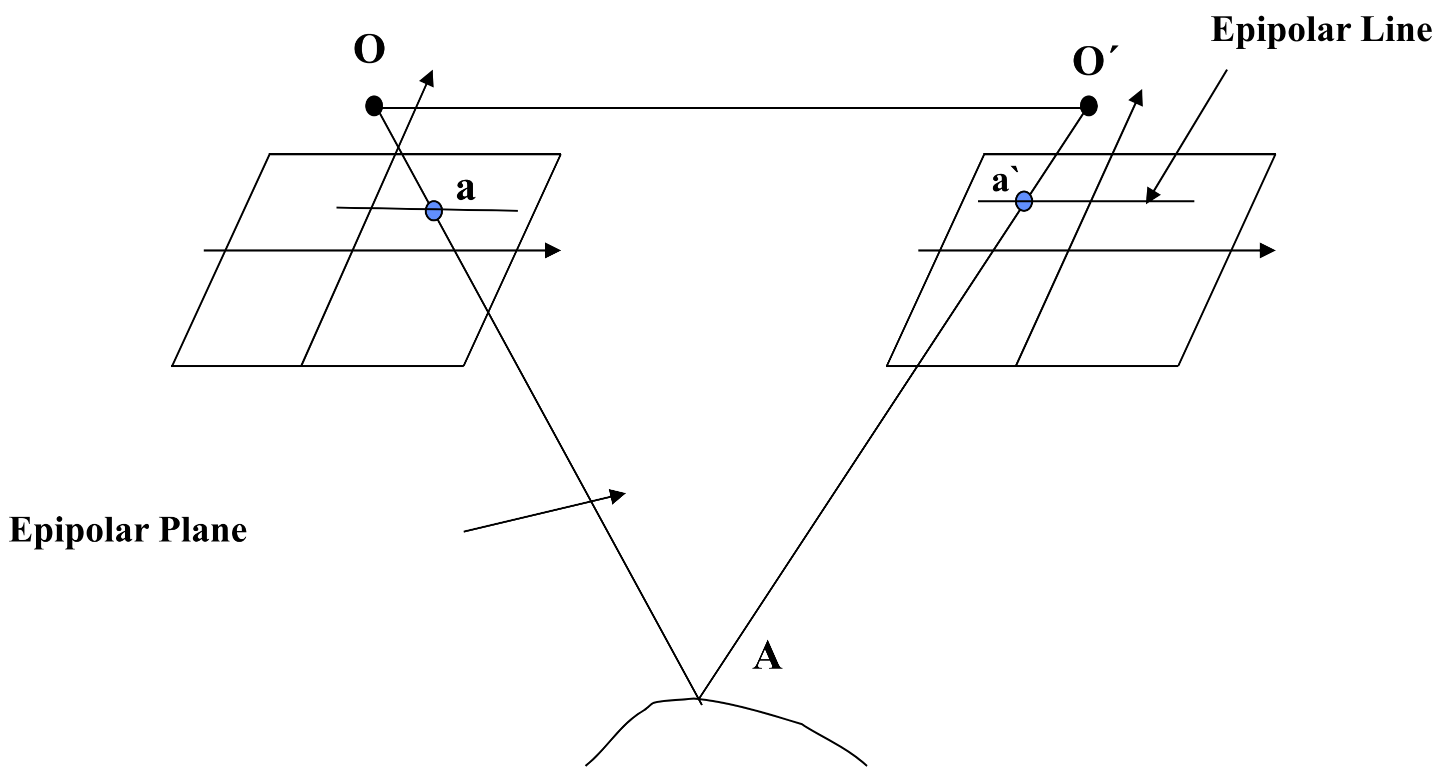




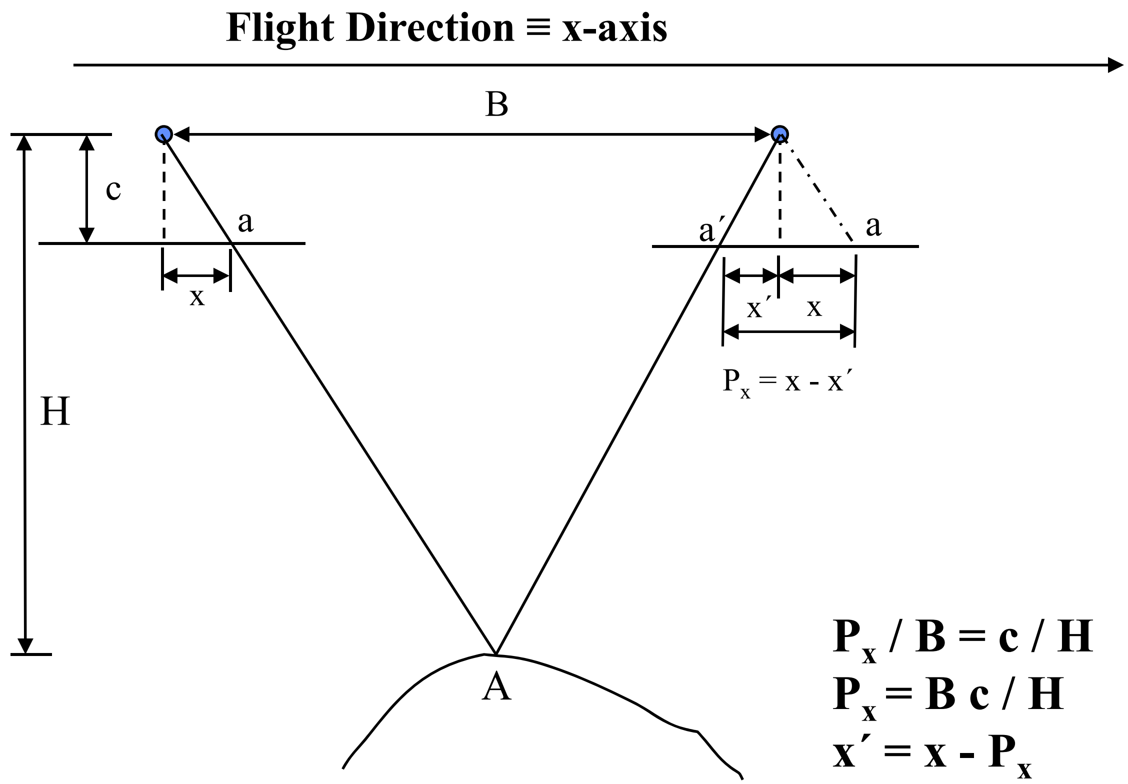
Determination of the Search Space

- The following parameters can be used to estimate the search space:
  - The average base (distance between successive exposures),
  - Average flying height above the terrain,
  - Approximate values for the exterior orientation parameters of the images of the stereo-pair under consideration, and
  - Approximate values for the interior orientation parameters of the used camera.

Epipolar Geometry



Expected P_x Parallax



$$P_x / B = c / H$$
$$P_x = B\,c / H$$
$$x' = x - P_x$$

The Location of the Search Window

- Assuming:
  - Vertical photography,
  - x-axis of the image coordinate system coincides with the flight direction, and
  - No change in the flying height
- The center of the search window (x_c, y_c) that corresponds to an image point (in the left photo) whose coordinates are (x_l, y_l) can be estimated as follows:
  - $x_c = x_l - P_x$
  - $y_c = y_l$

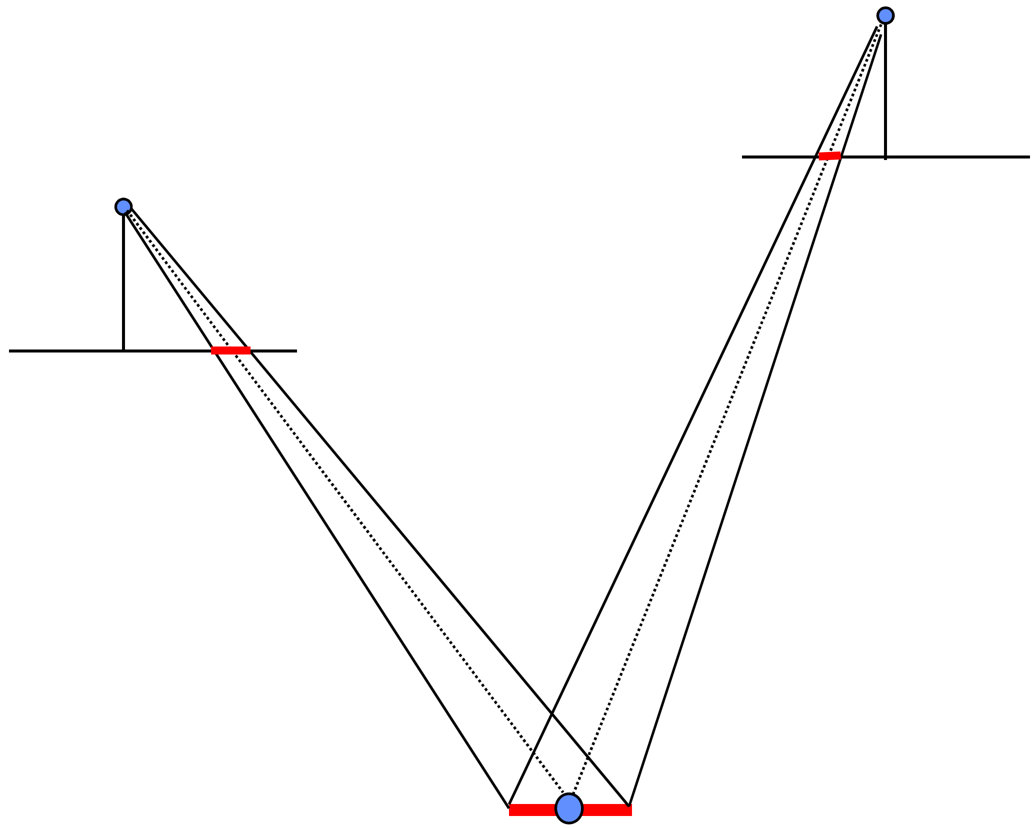
The Size of the Search Window

- The size of the search window depends on our confidence in:
  - The validity of the above mentioned assumptions,
  - The flying height above the object point under consideration,
  - The base line between successive exposures, and
  - The principal distance (focal length) of the camera.

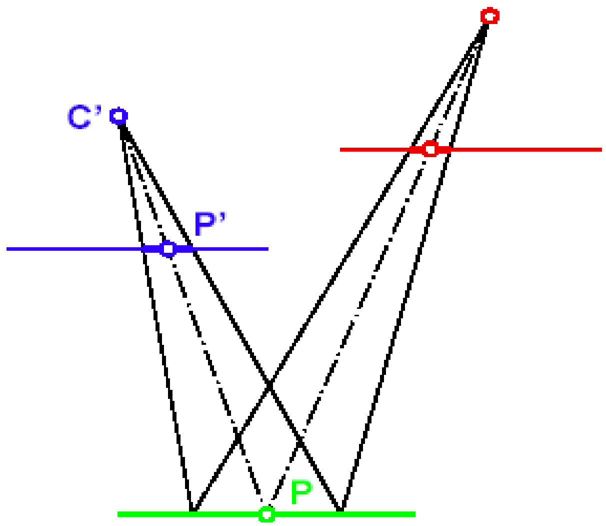
Fundamental Problems

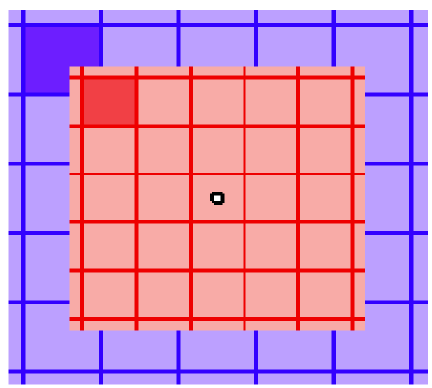
- Some problems that might complicate the matching problem include:
  - Scale differences between the two images,
  - Different rotation angles between the two images,
  - Tilted surfaces (foreshortening problem),
  - Occlusions,
  - Relief displacement (different background),
  - Different illumination conditions between the two images (different gray values), and
  - Repetitive patterns.

Scale Differences

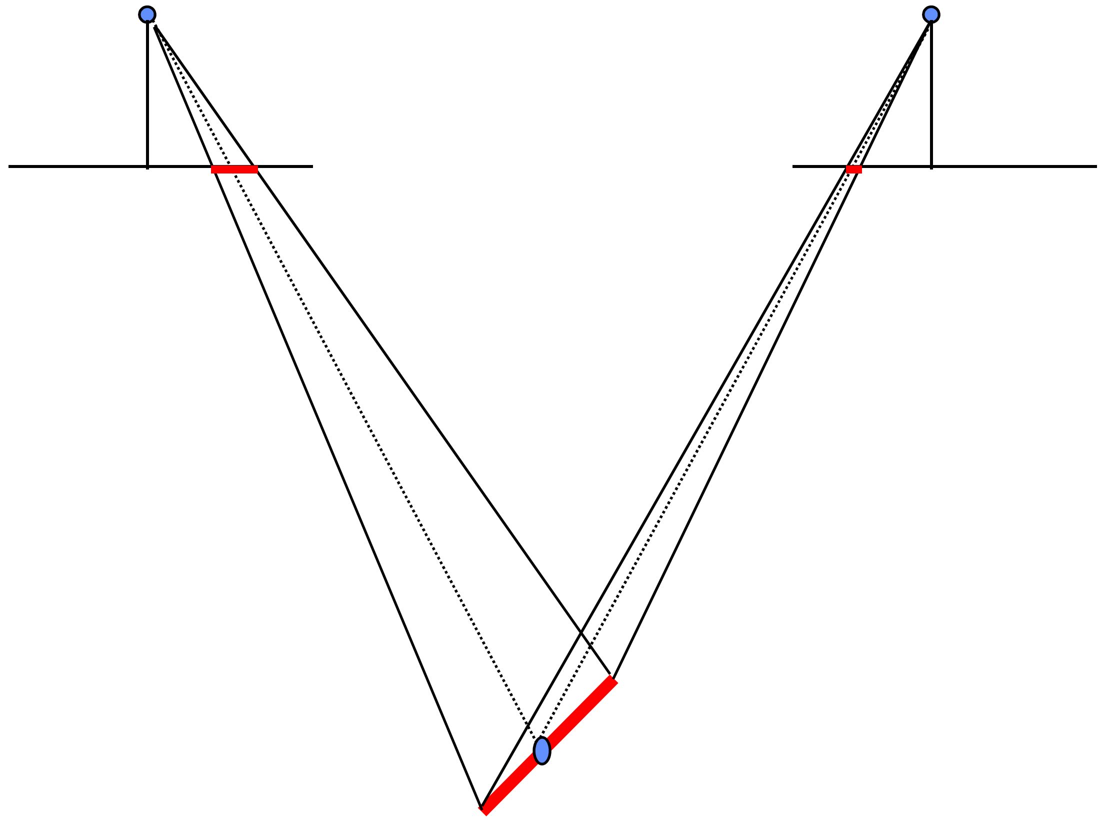


Scale Differences

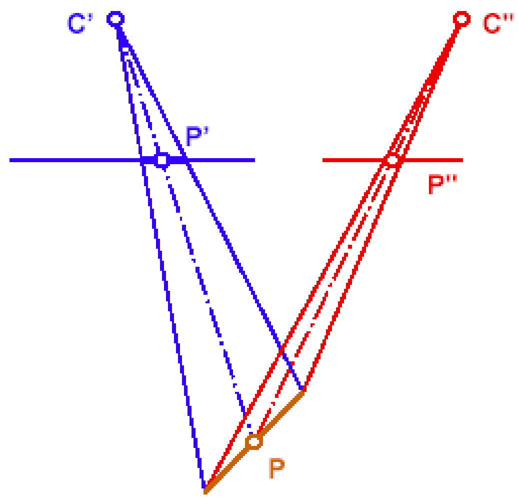


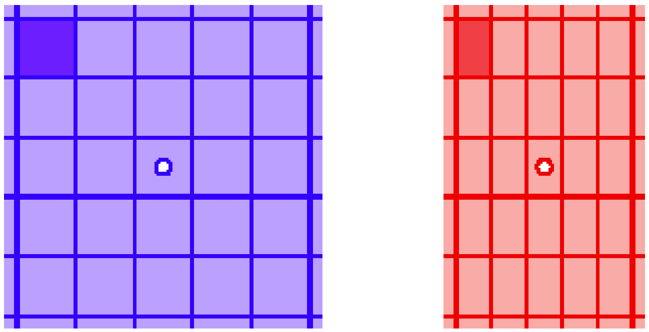


Foreshortening Problem

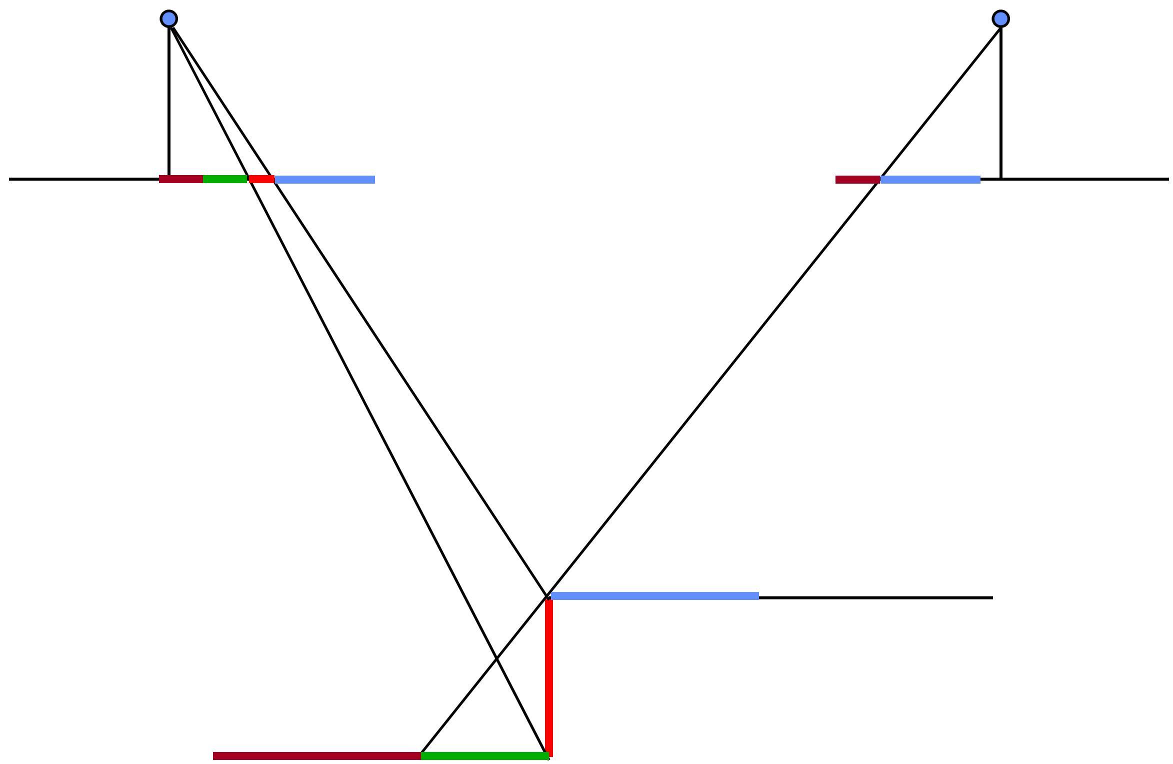


Foreshortening Problem



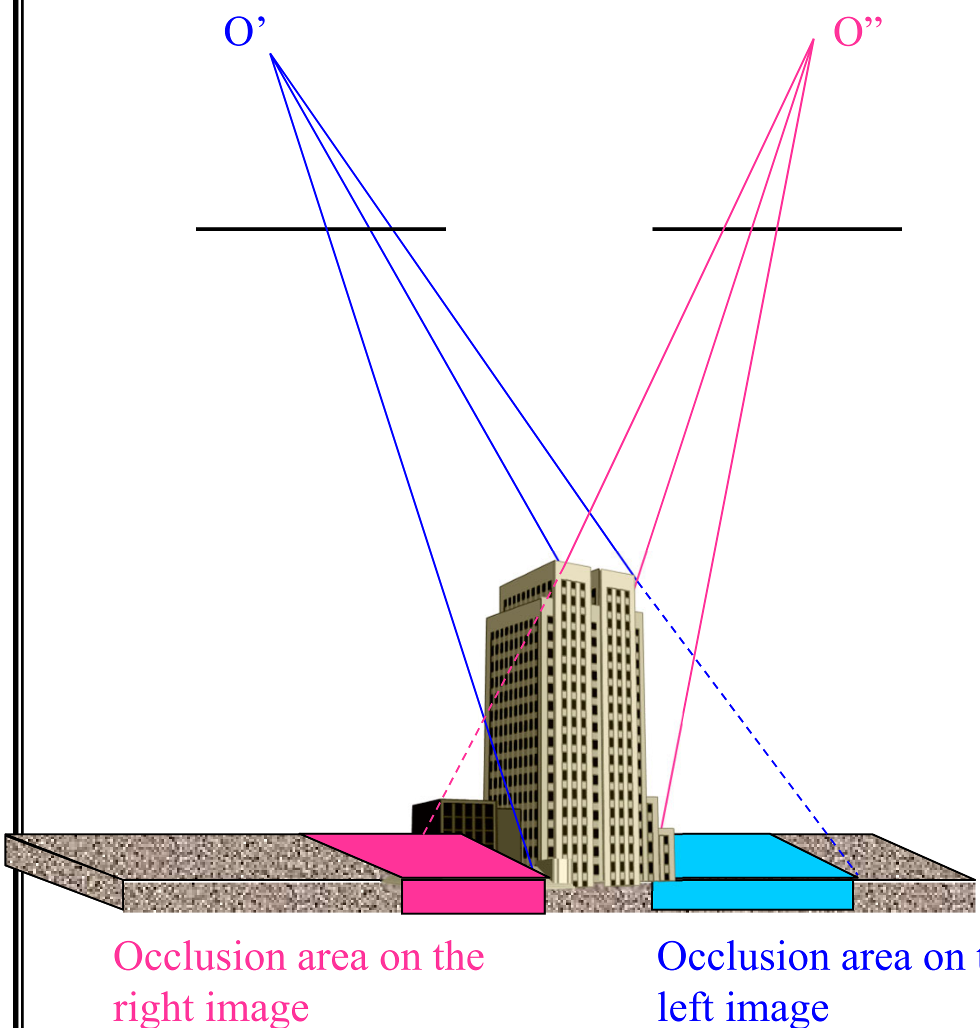


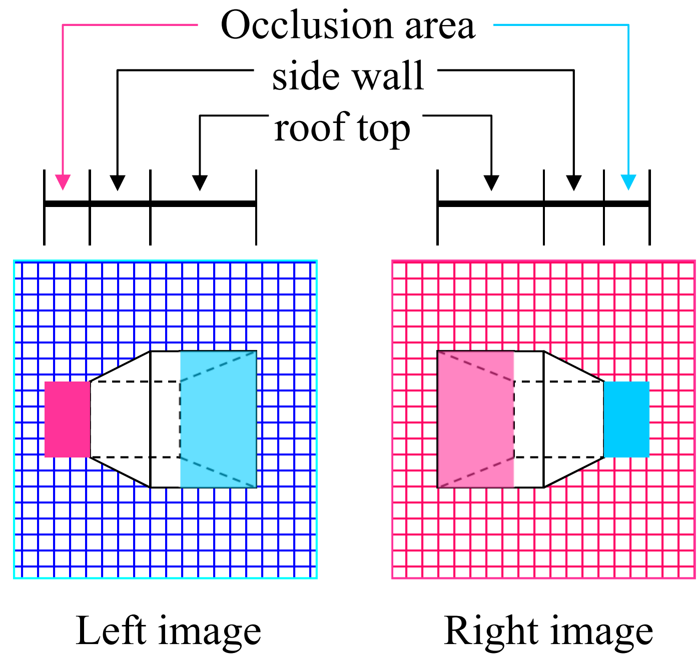
Occlusions



Occluded Area (right image)

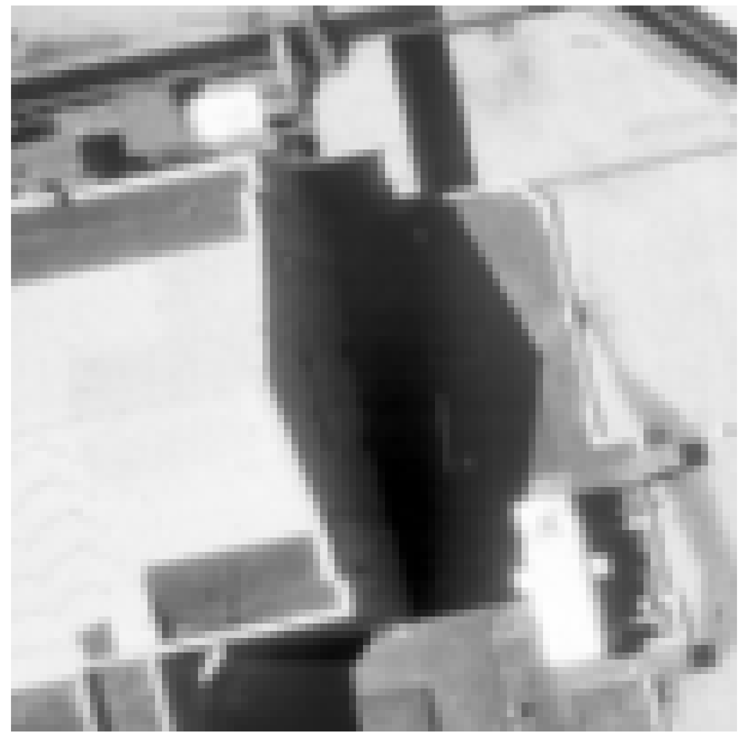
Occlusions



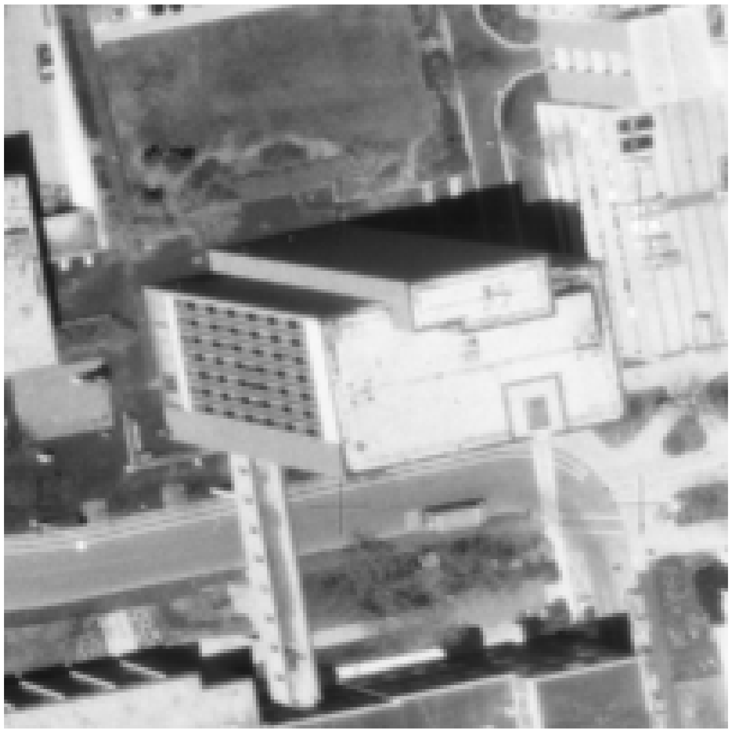


Matching problems take place at the area occluded by roof top or side walls.

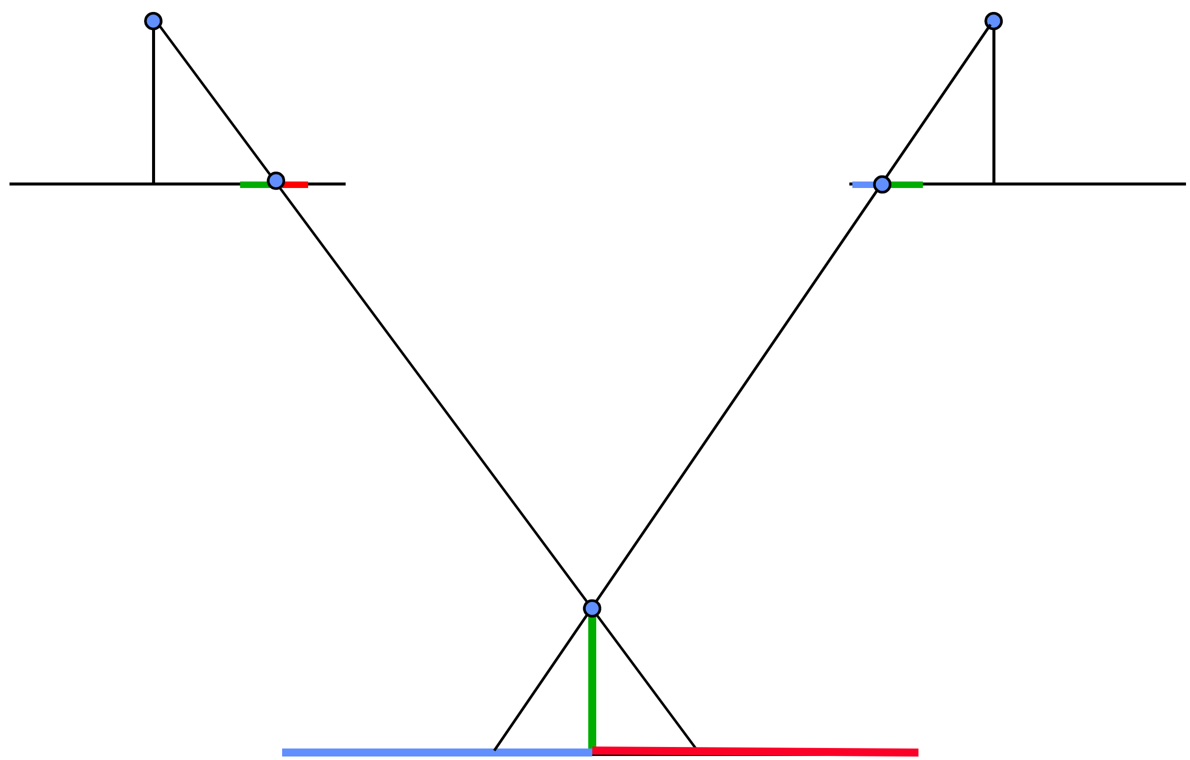
Occlusions



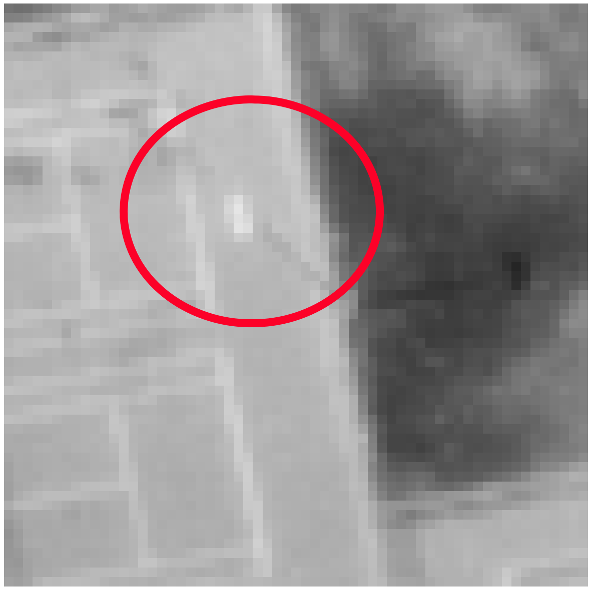
Occlusions & Foreshortening



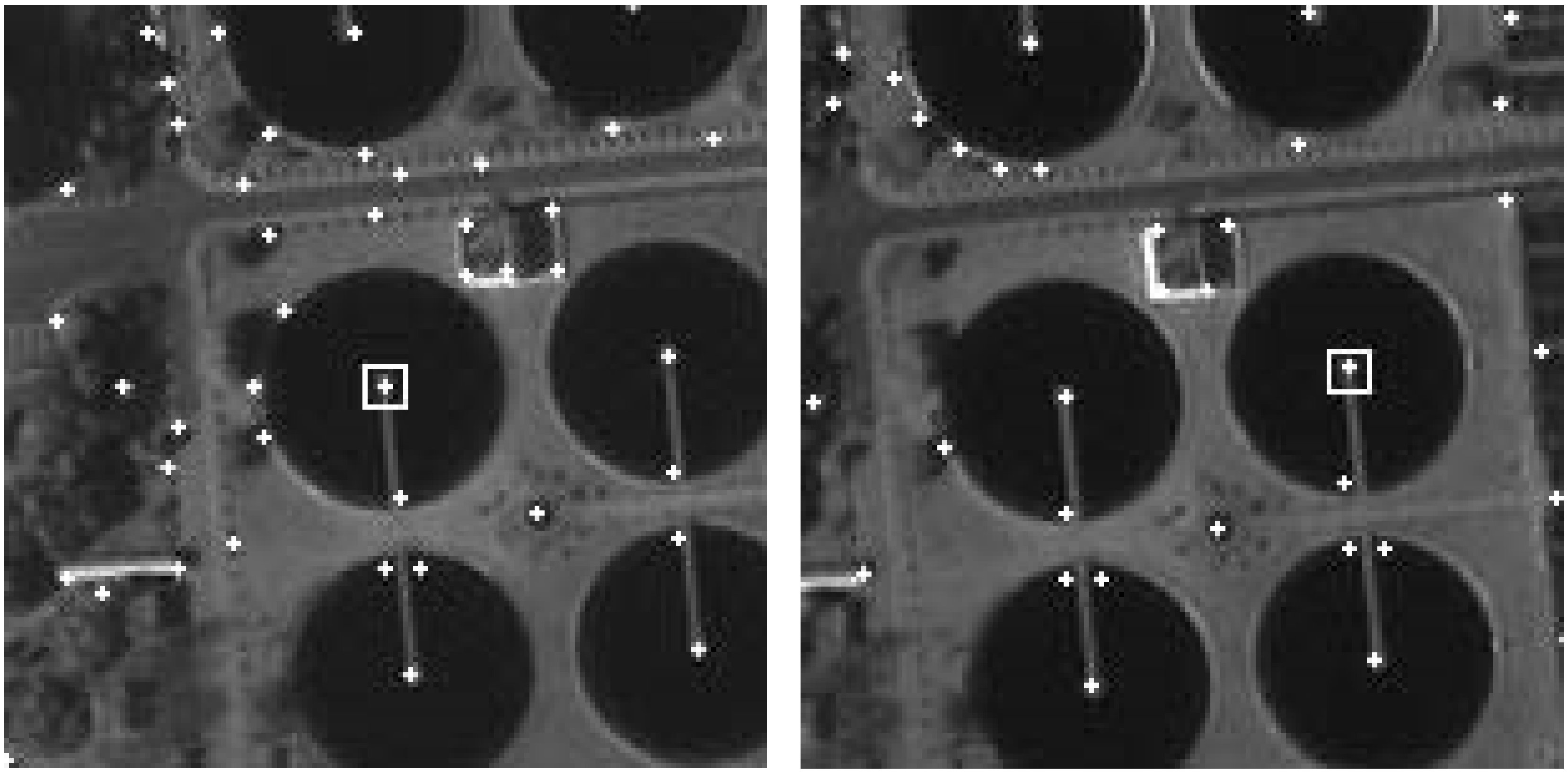
Relief Displacement (Different Background)



Relief Displacement (Different Background)



Repetitive Pattern

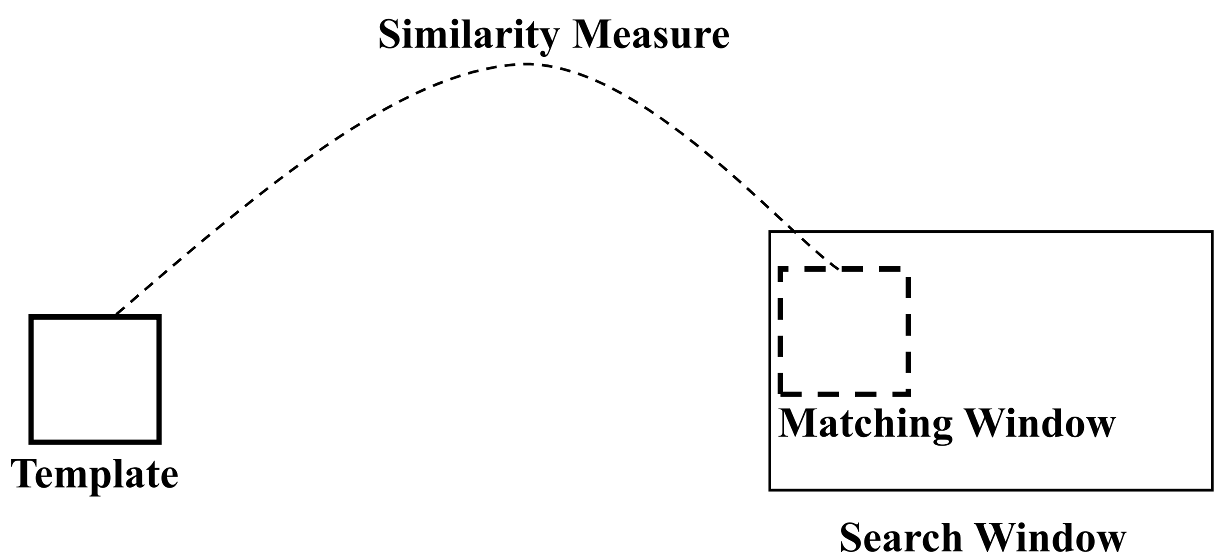


Area Based Matching

Area Based Matching

- Gray level distributions in small areas (image patches) in the two images of a stereo-pair are matched.
- Similarity measure between the image patches can be computed using:
  - Correlation coefficient, or
  - Least squares matching.
- Area based matching techniques are quite popular in photogrammetry.

Image Matching



Area-Based Matching

Correlation Matching

Correlation Coefficient

- Assuming that:
  - $g_l(x, y)$ is the gray value function within the template in the left image.
  - $g_r(x, y)$ is the gray value function within a matching window inside the search window in the right image.
  - (nxm) is the size of the template and matching windows.
- Then, the cross correlation coefficient (similarity measure) can be computed as follows:

Cross Correlation Coefficient

$$\overline{g_l} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} g_l(x_i, y_j)}{n\ m}$$

$$\overline{g_r} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} g_r(x_i, y_j)}{n\ m}$$

$$\sigma_l = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} [g_l(x_i, y_j) - \overline{g_l}]^2}{n\ m - 1}}$$

$$\sigma_r = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} [g_r(x_i, y_j) - \overline{g_r}]^2}{n\ m - 1}}$$

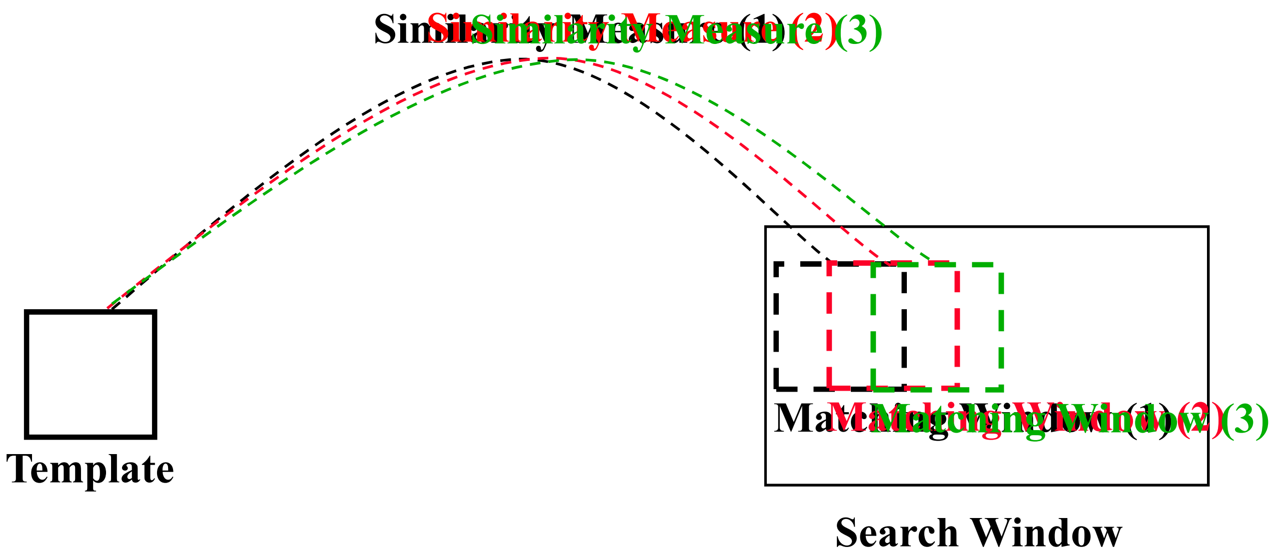
$$\sigma_{lr} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} [\{g_l(x_i, y_j) - \overline{g_l}\}\ \{g_r(x_i, y_j) - \overline{g_r}\}]}{n\ m - 1}$$

$$\rho = \frac{\sigma_{lr}}{\sigma_l\ \sigma_r}$$

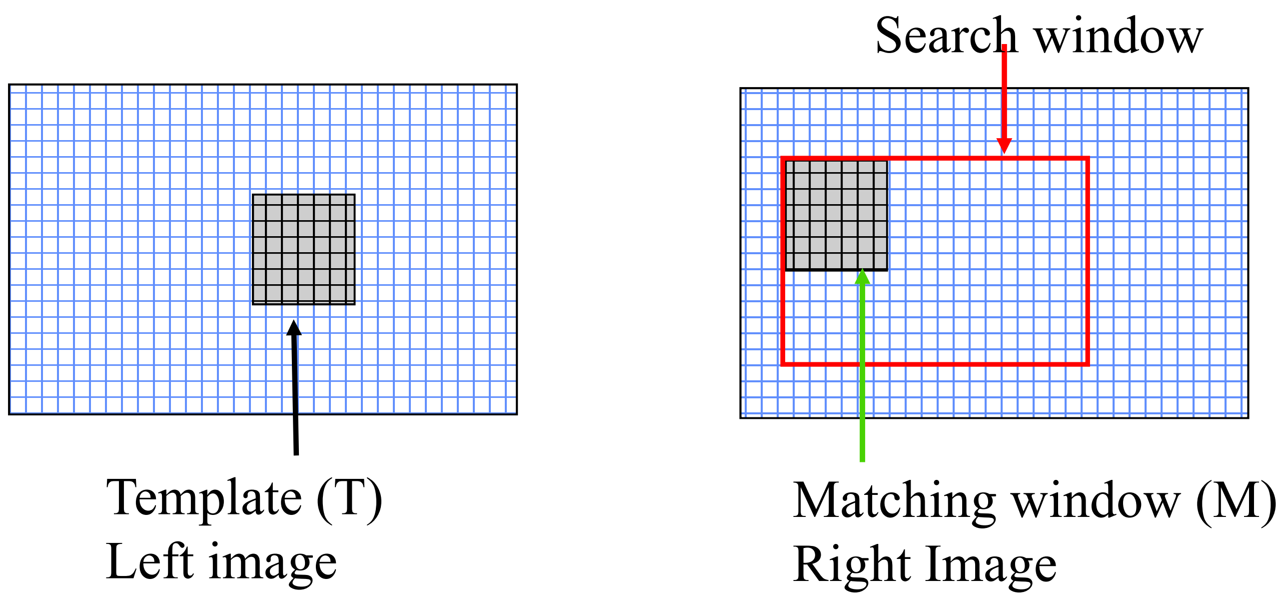
Cross Correlation Coefficient

- The correlation coefficient factor might take values that range from -1 to +1.
- $\rho = 0$ indicates no similarity at all.
- $\rho = -1$ indicates an inverse similarity (e.g., similarity between the dia-positive and the negative of the same image).
- $\rho = 1$ indicates a perfect match (the highest similarity possible).

Correlation-Based Matching



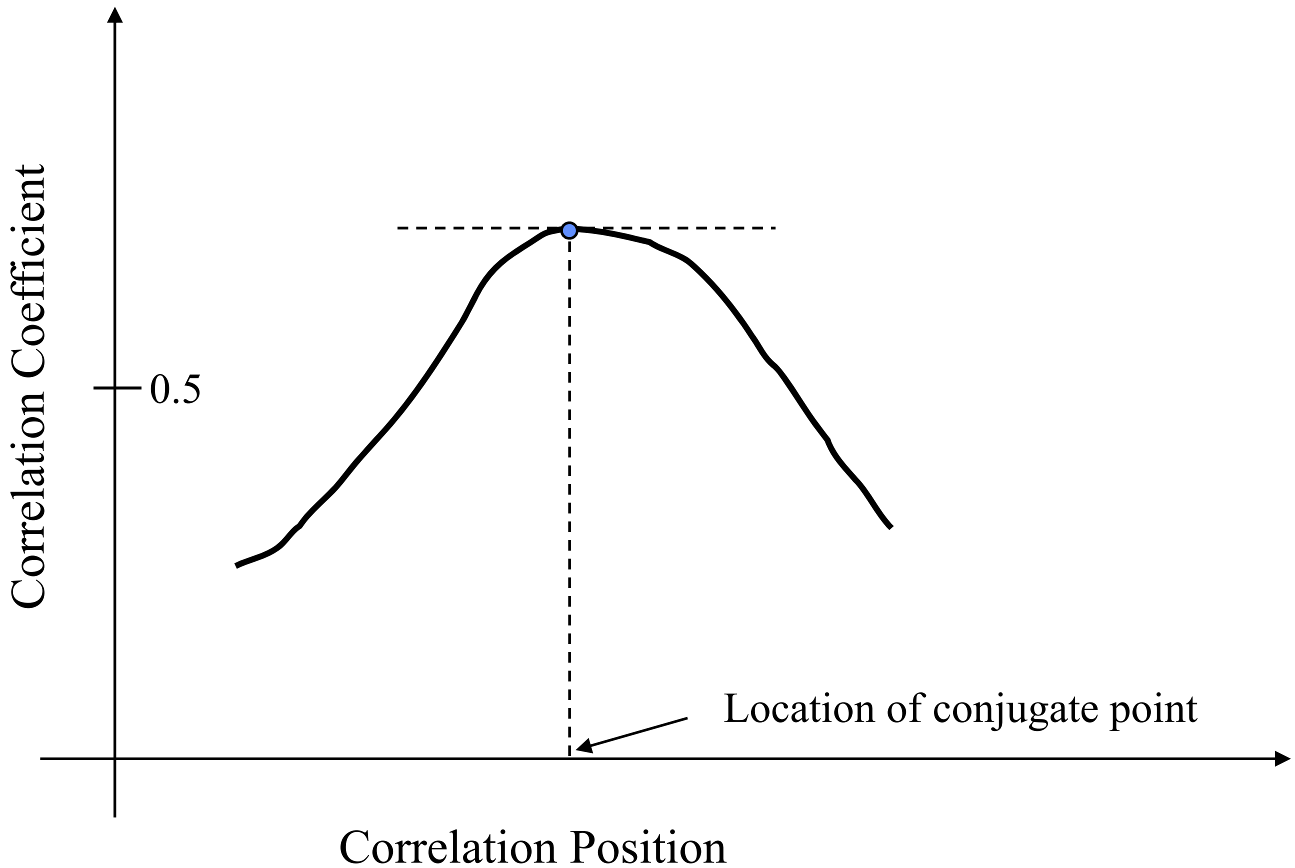
Correlation-Based Matching



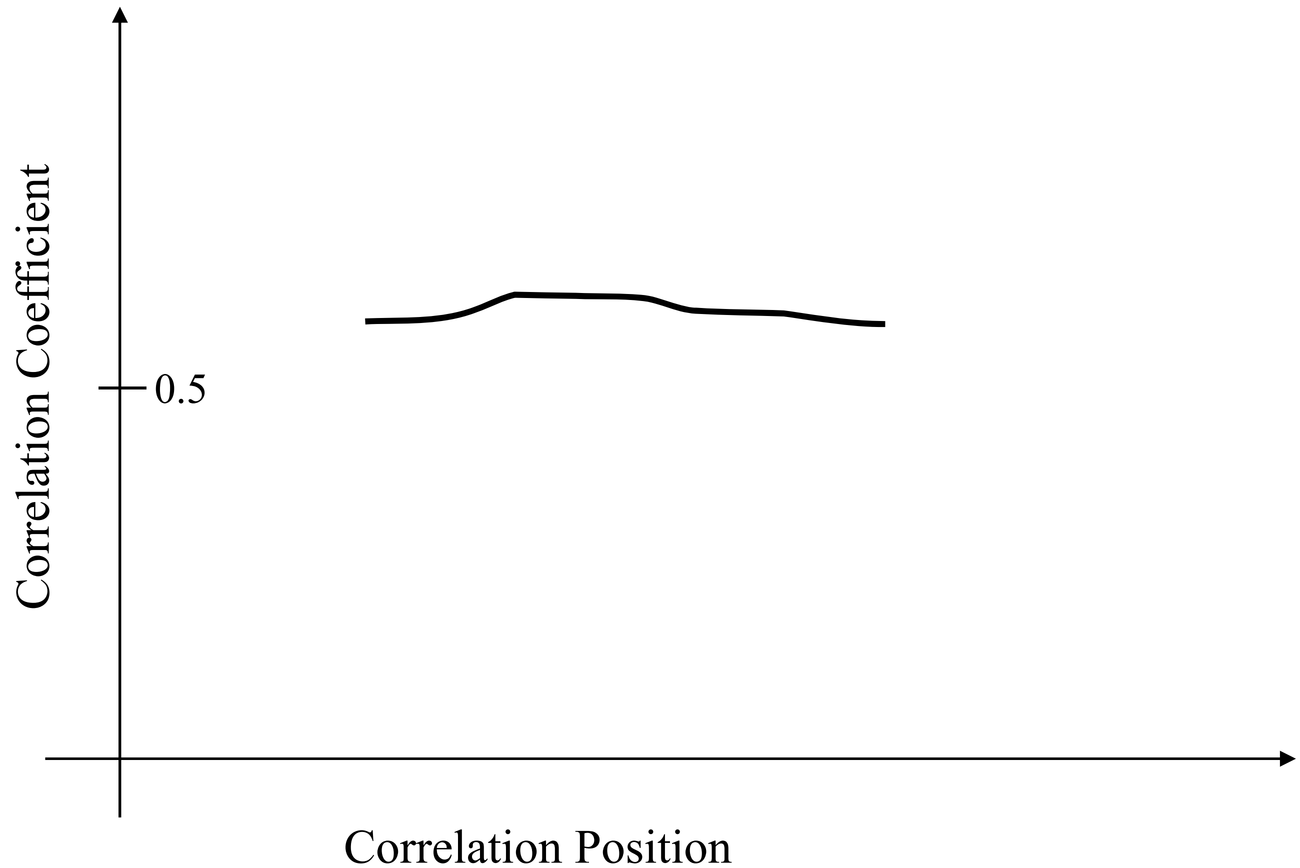
Correlation-Based Matching

- The correlation coefficient is computed for every possible position of the matching window within the search window.
- The position of the conjugate point is determined as the location with the maximum correlation coefficient.
- We will only accept correlation coefficients that are above a predetermined threshold (e.g., 0.5).

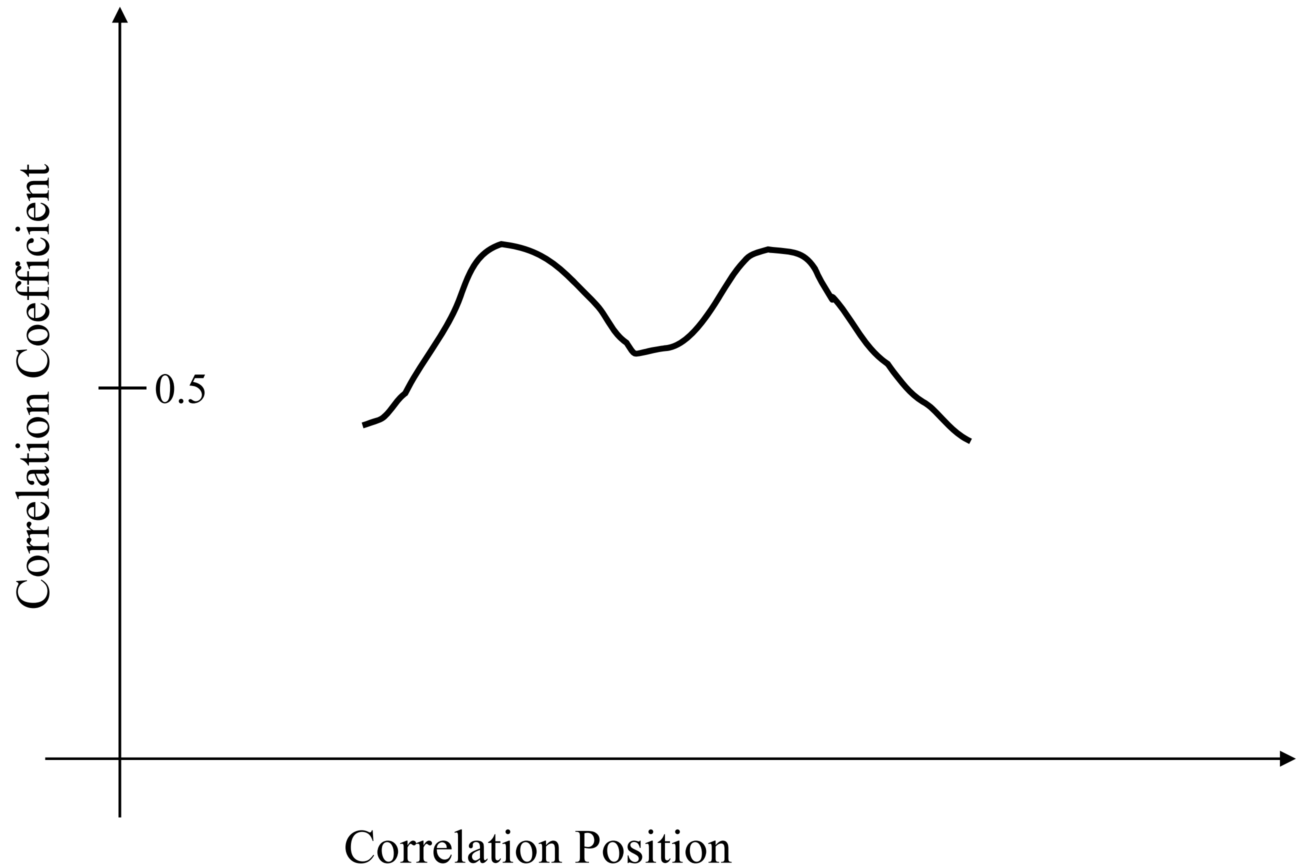
Optimum Correlation Coefficient Function



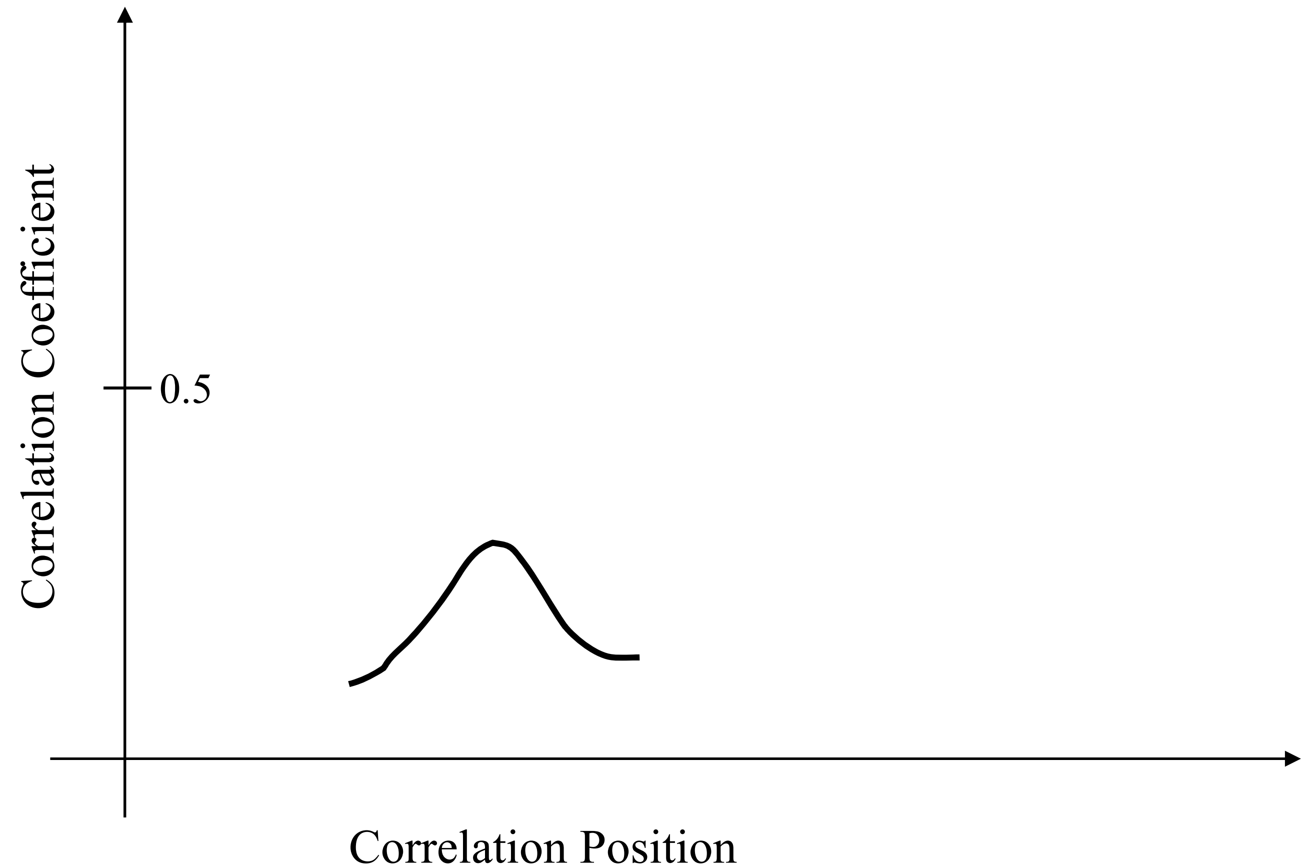
Typical Problems (1): Ill defined peak



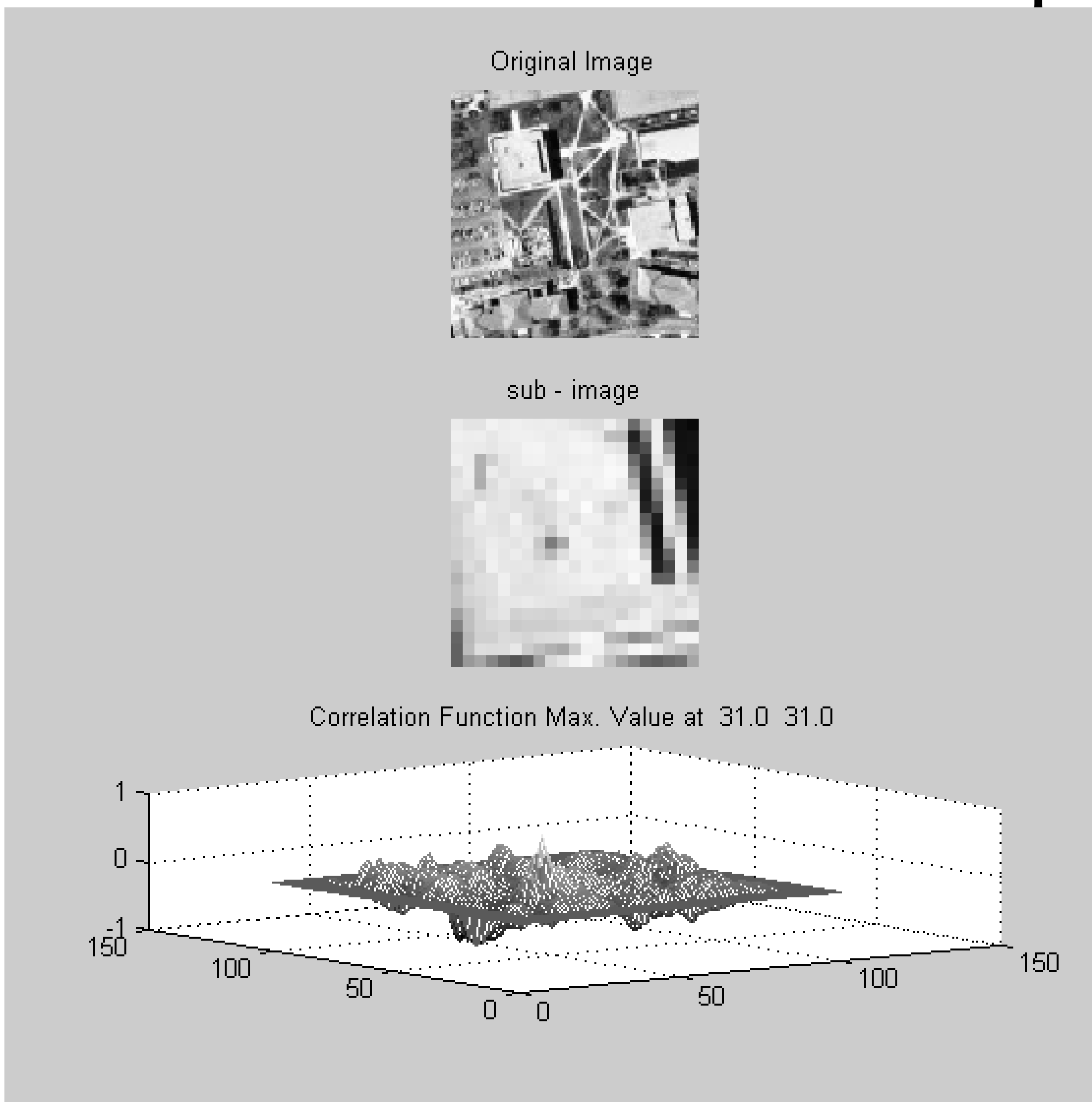
Typical Problems (2): Repetitive Pattern



Typical Problem (3): Low Maximum Value



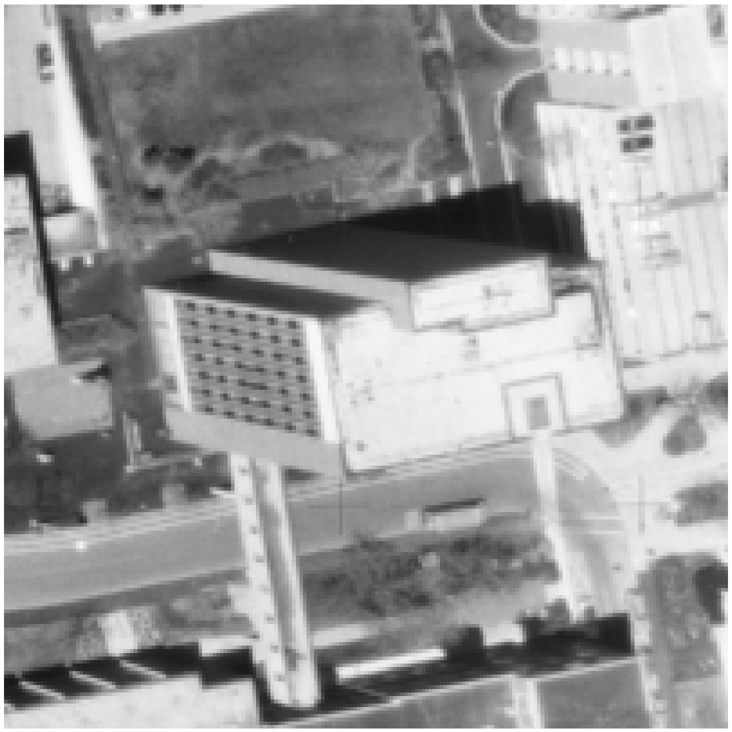
Correlation Function: Example



Correlation-Based Matching

- Main disadvantage:
  - We do not compensate for any geometric or radiometric differences between the template and matching windows.
- Geometric differences will happen due to different scale and rotation parameters between the two images, foreshortening, etc.
- Radiometric differences will happen due to different illumination conditions.

Geometric Differences



Radiometric Differences

Left Image

Right Image

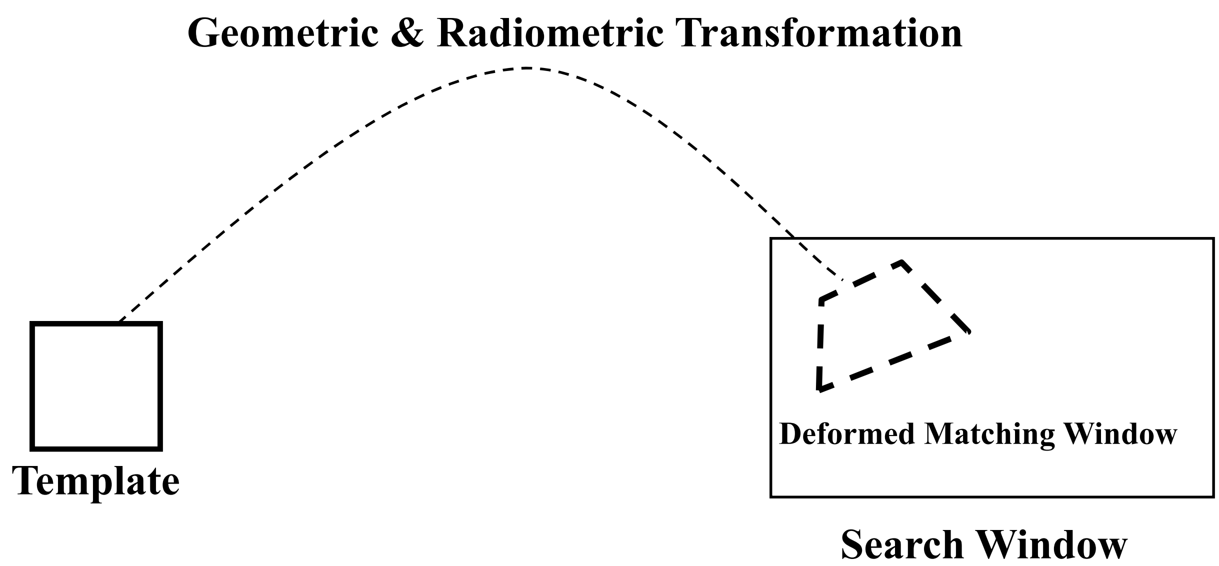
Area-Based Matching

Least Squares Matching

Least Squares Matching

- Concept:
  – Minimize the gray value differences between the template and the matching windows whereby the parameters defining the position and the shape of the matching window are determined using an adjustment procedure
- In other words:
  – The position and the shape of the matching window are changed until the gray level differences between the deformed matching window and the template reach a minimum.

Least Squares Matching



Geometric Transformation

- Geometric transformation between the template and the matching window can be done by either:
  - Projective transformation (8-parameters),
  - Affine transformation (6-parameters),
  - Two dimensional similarity transformation (4 parameters), or
  - Two shifts.

Affine Transformation

- $x_m = t_x + x_t a_1 + y_t a_2$
- $y_m = t_y + x_t b_1 + y_t b_2$
- Where:
  - (x_t, y_t) are the coordinates in the template,
  - (x_m, y_m) are the corresponding coordinates in the matching window, and
  - $(t_x, t_y, a_1, a_2, b_1, b_2)$ are the parameters of the affine transformation.

Radiometric Transformation

- One way of applying radiometric transformation that compensates for brightness and contrast differences between the template and matching window can be done as follows:
  - $g_t = t_o + t_1 \, g_m$
    - Where:
      - g_t is the modified gray value within the template,
      - g_m is the gray value within the matching window,
      - t_o brightness modification factor, and
      - t_1 contrast modification factor.

Least Squares Matching: Target function

- The radiometric and the geometric transformation parameters are determined in such a way to minimize the following target function:

$$\phi\,(t_o,\, t_1,\, t_x,\, t_y,\, a_1,\, a_2,\, b_1,\, b_2) =$$

$$\sum_{(x,y)\in\, Template\ Window} [g_t(x,\, y) - T_R\{g_m(T_g(x,\, y))\}]^2$$

$$= \min\,(t_o,\, t_1,\, t_x,\, t_y,\, a_1,\, a_2,\, b_1,\, b_2)$$

$NOTE:\ g_m(T_g(x,\, y))\ requires\ resampling$

Least Squares Matching

$$g_t(x_t, y_t) = t_o + t_1\, g_m(t_x + a_1 x_t + a_2 y_t, t_y + b_1 x_t + b_2 y_t)$$

$$g_t(x_t, y_t) - t_o^o - t_1^o\, g_m(t_x^o + a_1^o x_t + a_2^o y_t, t_y^o + b_1^o x_t + b_2^o y_t) =$$

$$g_t(x_t, y_t) - t_o^o - t_1^o\, g_m(x_m^o, y_m^o) = \Delta g =$$

$$dt_o + g_m(x_m^o, y_m^o)\, dt_1 + t_1^o \frac{\partial g_m}{\partial x} dt_x + t_1^o \frac{\partial g_m}{\partial x} x_t\, da_1 + t_1^o \frac{\partial g_m}{\partial x} y_t\, da_2$$

$$+ t_1^o \frac{\partial g_m}{\partial y} dt_y + t_1^o \frac{\partial g_m}{\partial y} x_t\, db_1 + t_1^o \frac{\partial g_m}{\partial y} y_t\, db_2$$

Least Squares Matching (Observation Equations – 3x3 Template)

$$
\begin{bmatrix} \Delta g^1 \\ \Delta g^2 \\ \Delta g^3 \\ . \\ . \\ . \\ . \\ . \\ \Delta g^9 \end{bmatrix}
=
\begin{bmatrix}
t_1^o g_x^1 & t_1^o x g_x^1 & t_1^o y g_x^1 & t_1^o g_y^1 & t_1^o x g_y^1 & t_1^o y g_y^1 & 1 & g^1 \\
t_1^o g_x^2 & t_1^o x g_x^2 & t_1^o y g_x^2 & t_1^o g_y^2 & t_1^o x g_y^2 & t_1^o y g_y^2 & 1 & g^2 \\
t_1^o g_x^3 & t_1^o x g_x^3 & t_1^o y g_x^3 & t_1^o g_y^3 & t_1^o x g_y^3 & t_1^o y g_y^3 & 1 & g^3 \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
t_1^o g_x^9 & t_1^o x g_x^9 & t_1^o y g_x^9 & t_1^o g_y^9 & t_1^o x g_y^9 & t_1^o y g_y^9 & 1 & g^9
\end{bmatrix}
\bullet
\begin{bmatrix} t_x \\ da_1 \\ da_2 \\ t_y \\ db_1 \\ db_2 \\ dt_o \\ dt_1 \end{bmatrix}
$$

Least Squares Matching (Final Remarks)

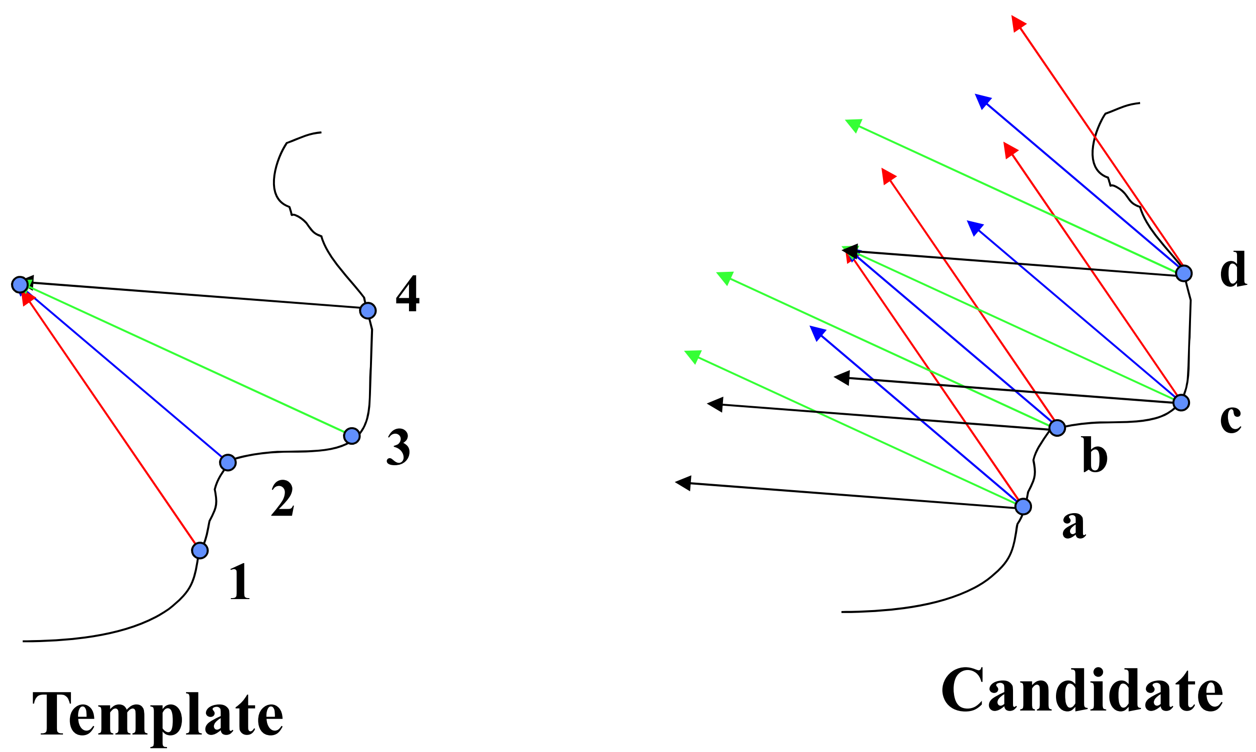
- Compensates for radiometric and geometric differences between the template and matching windows.
- Non linear procedure → approximations are necessary.
- It is very sensitive to the quality of the approximate values. It might diverge if we are more than four pixels away.

Feature Based Matching

Feature Based Matching

- Predominantly used in computer vision applications
- Edges or other features derived from the original images are compared to determine conjugate features.
- Similarity measures might include shape, magnitude and the orientation of the edge under consideration.
- An example of matching strategy is the generalized Hough transform.

Generalized Hough Transform

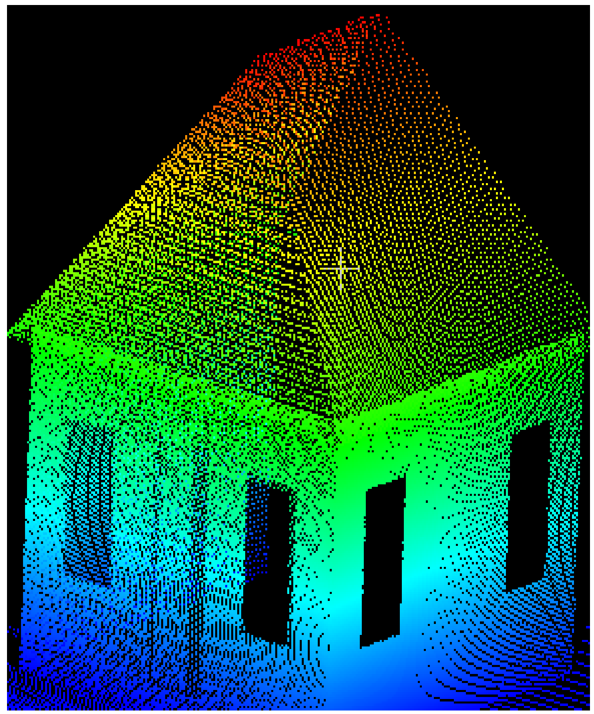


Symbolic (Relational) Matching

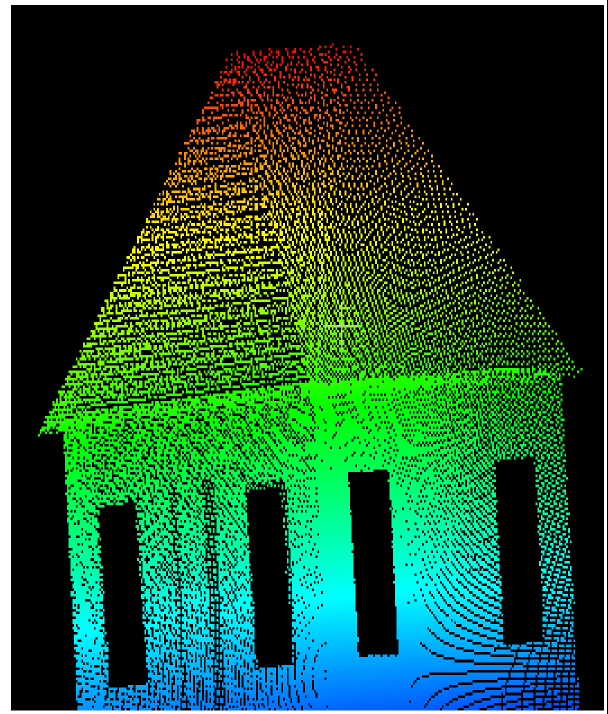
- Previous matching techniques are based on matching single primitives (i.e., one entity at a time).
- Relational matching provides a mechanism to consider relationships among the matching entities (i.e., more than one entity are matched simultaneously).
- Algorithms for relational matching include:
  - tree matching, and
  - Association matrices.

Symbolic (Relational) Matching: 3D Data

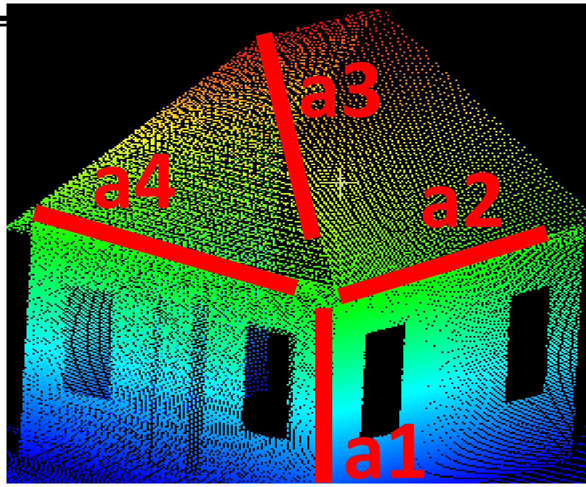
Building Model

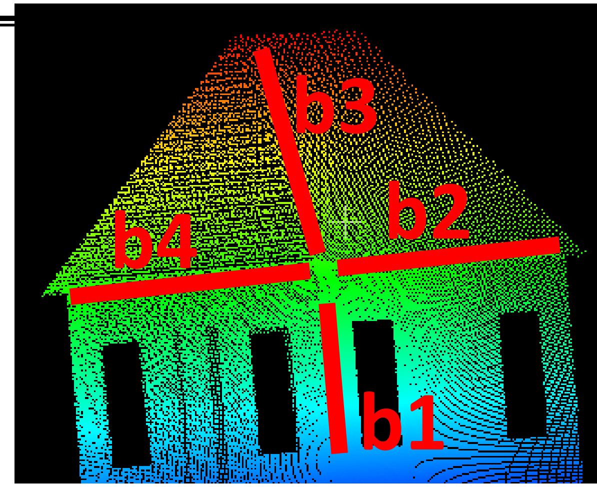


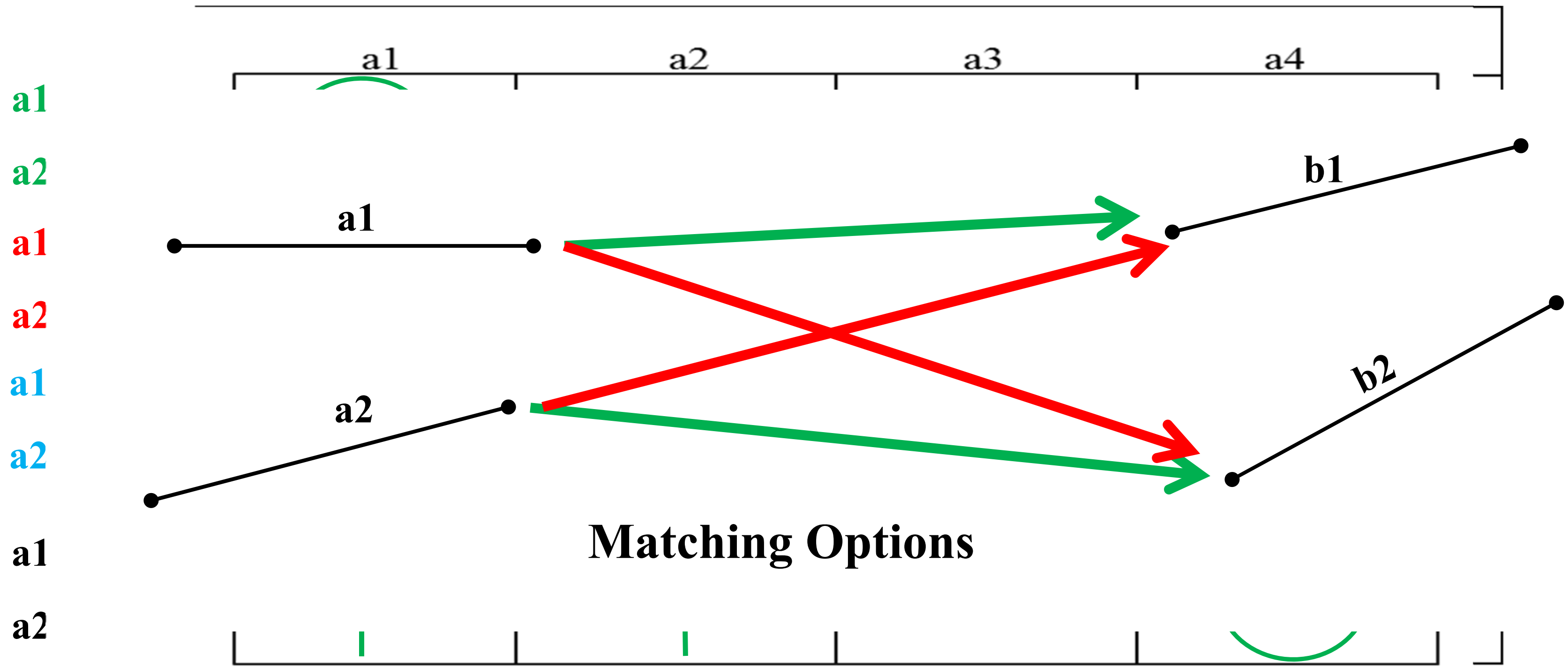
Scan 1



Scan 2





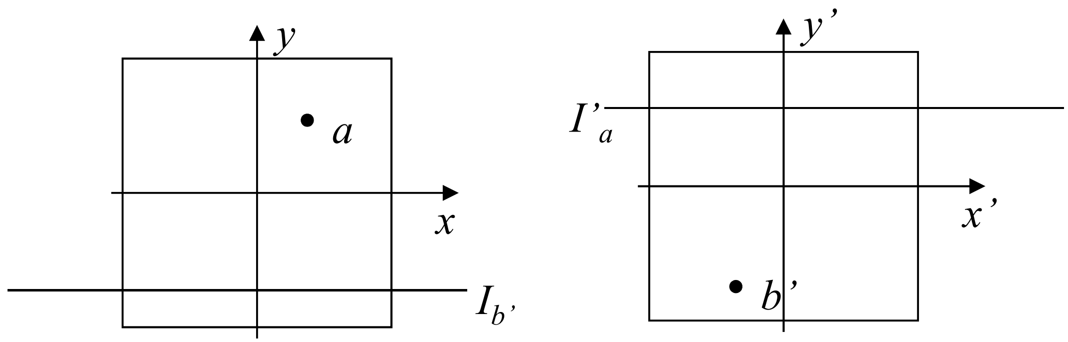


Association matrix of lines from Scans 1 and 2

Normalized Image Generation

Normalized Image Generation

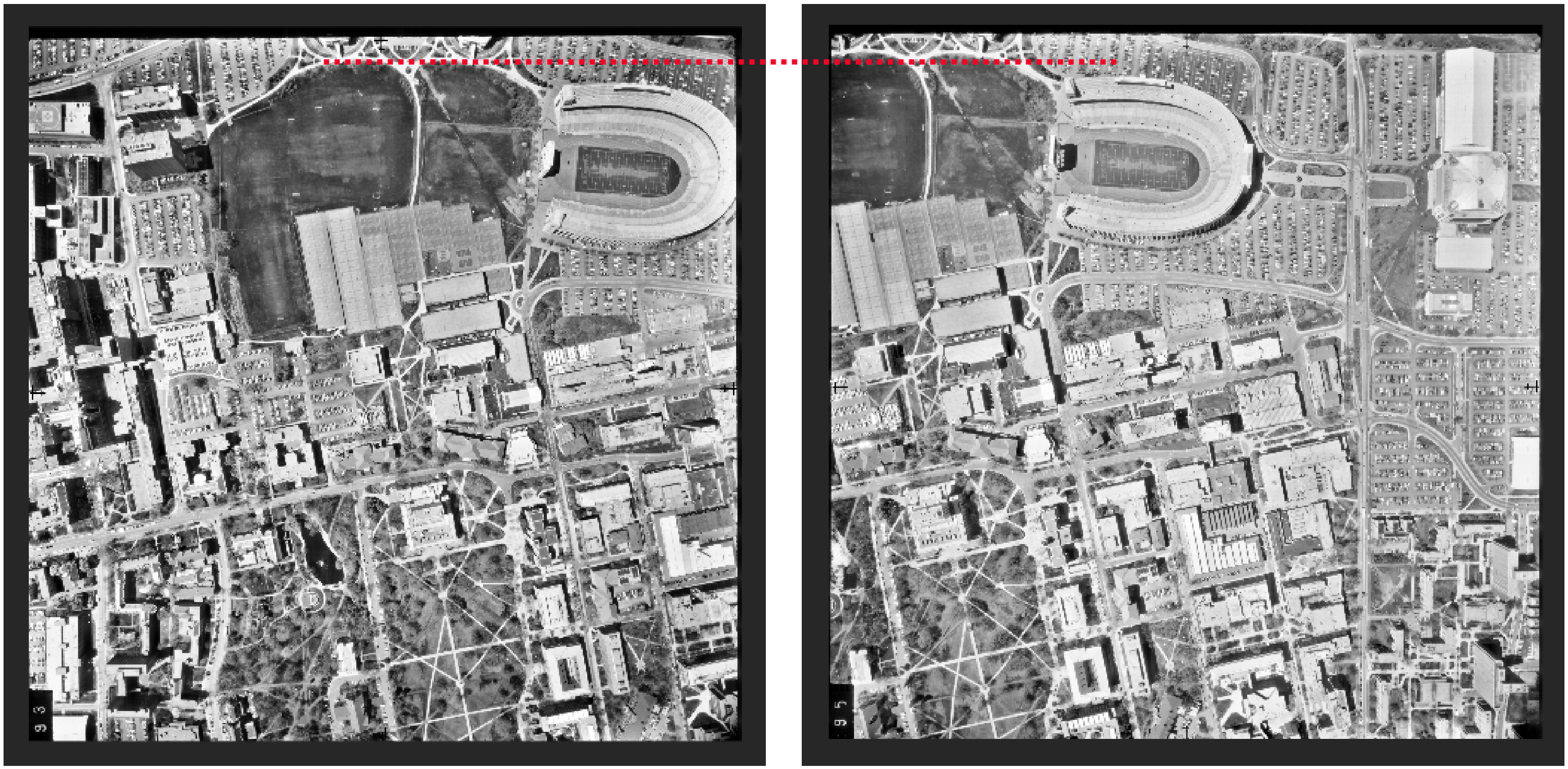
- Ensure that conjugate points in overlapping images have the same *y*-coordinates.
  - Reducing matching ambiguities
  - Reducing search space/time



- Important Applications:
  - Automatic matching
  - Automatic aerial triangulation
  - Orthophoto generation
  - Automatic relative orientation
  - Automatic DEM generation
  - Stereo viewing

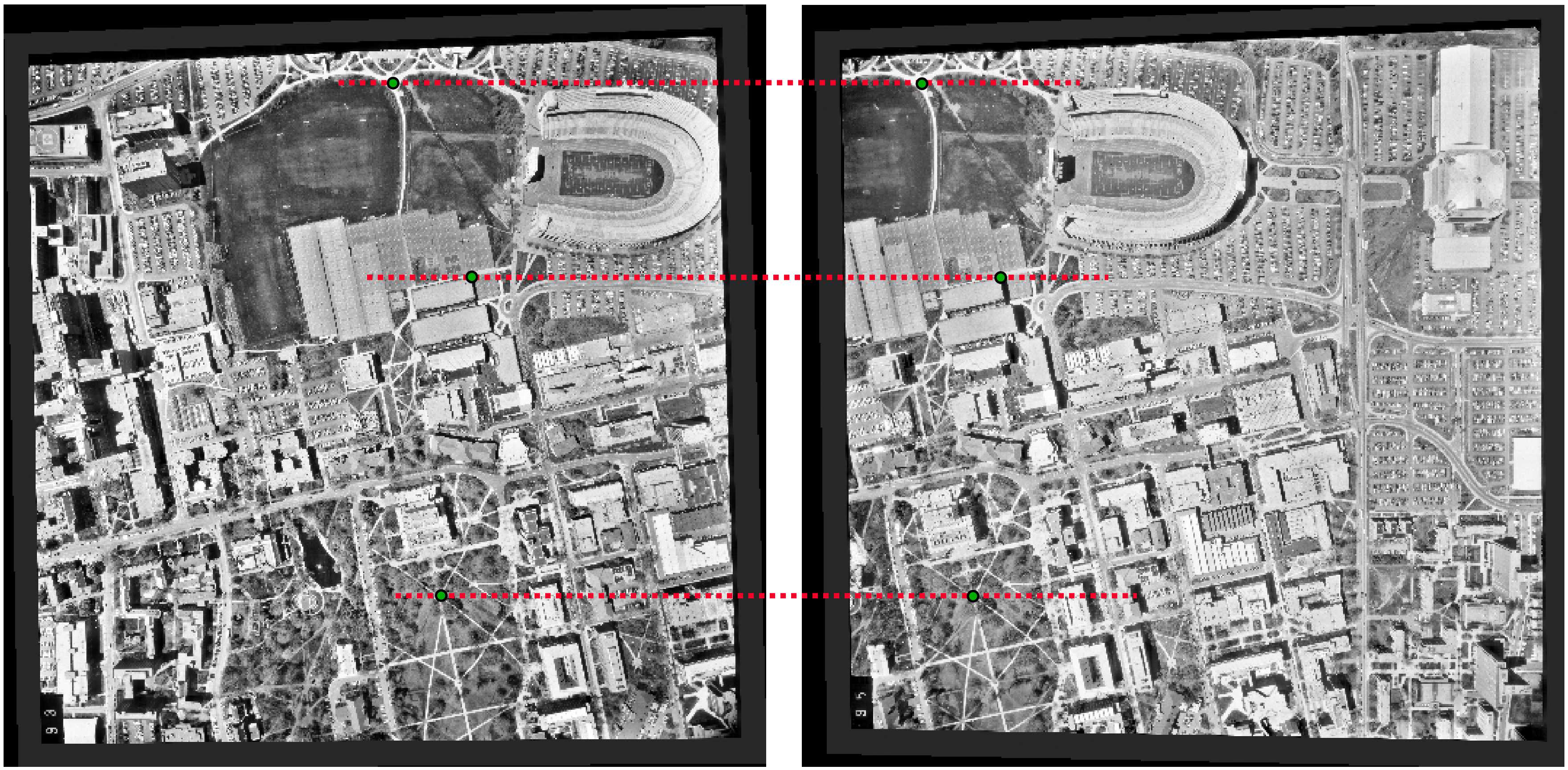
Original Images

Conjugate points do not have the same y-coordinates.

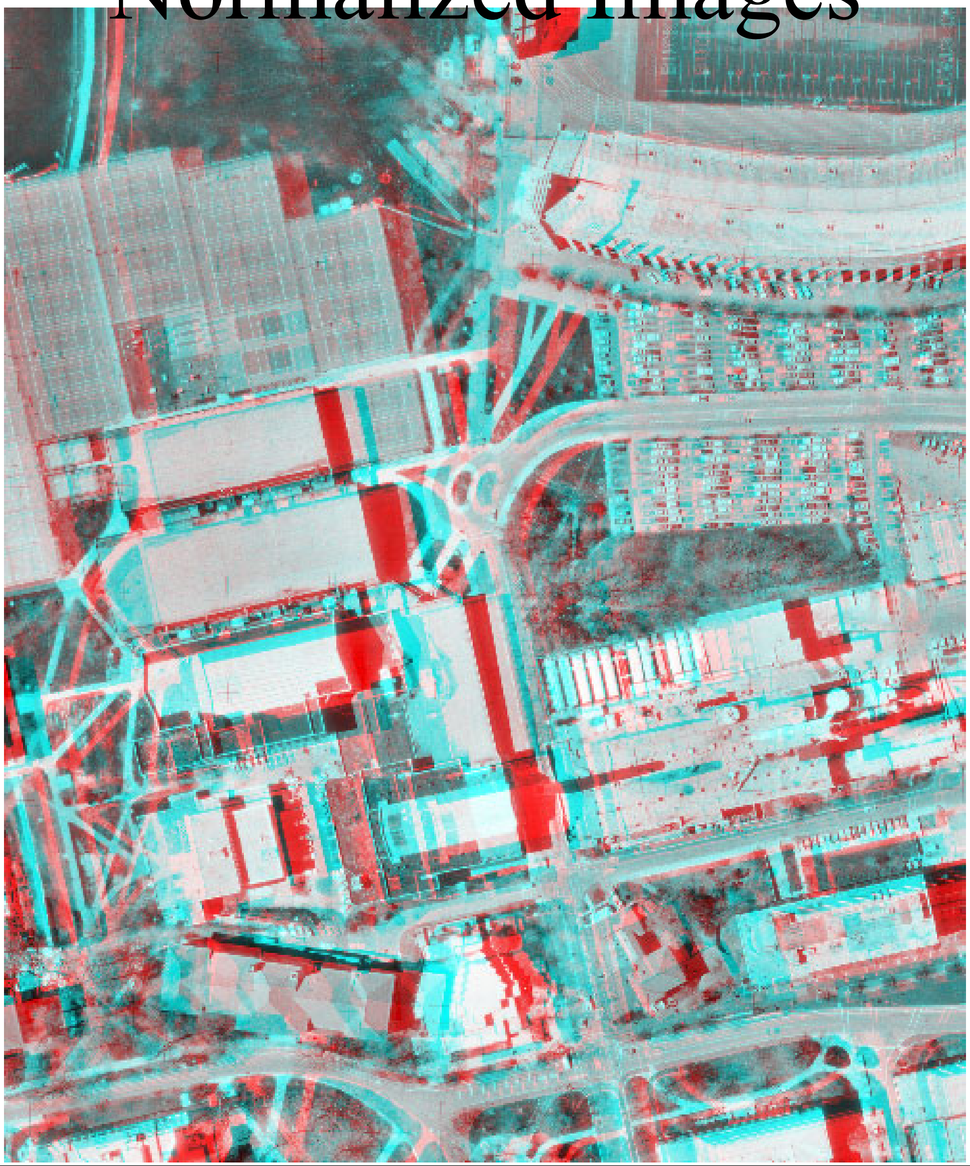


Normalized Images

Conjugate points have the same y-coordinates.



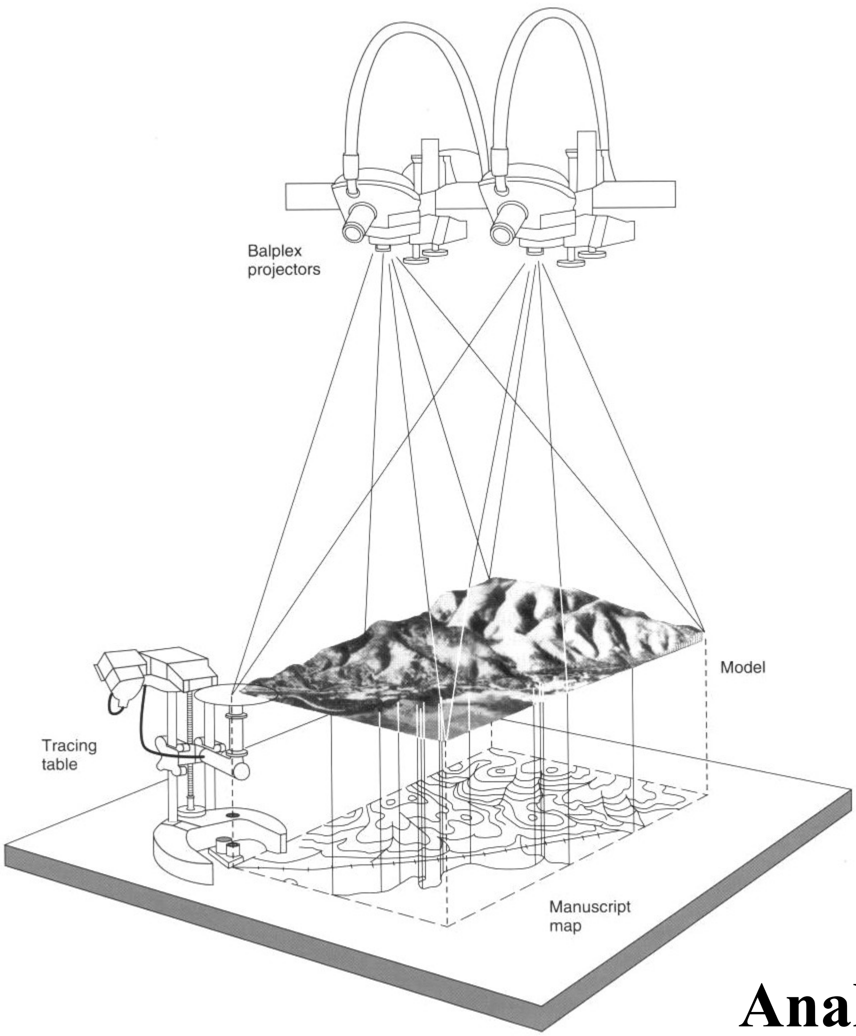
Normalized Images



Normalized Image Generation: Why?

- Normalized image generation is important for the following reasons:
  - 3-D viewing of 2-D images can be established without the need for photogrammetric plotters.
    - Photogrammetric plotters are expensive and need an experienced operator to establish the orientation procedure.
  - In automatic matching, the search space for conjugate points will be reduced from 2-D to 1-D.
    - Save time and avoid matching ambiguities whenever dealing with repetitive patterns.

3-D Viewing of 2-D imagery

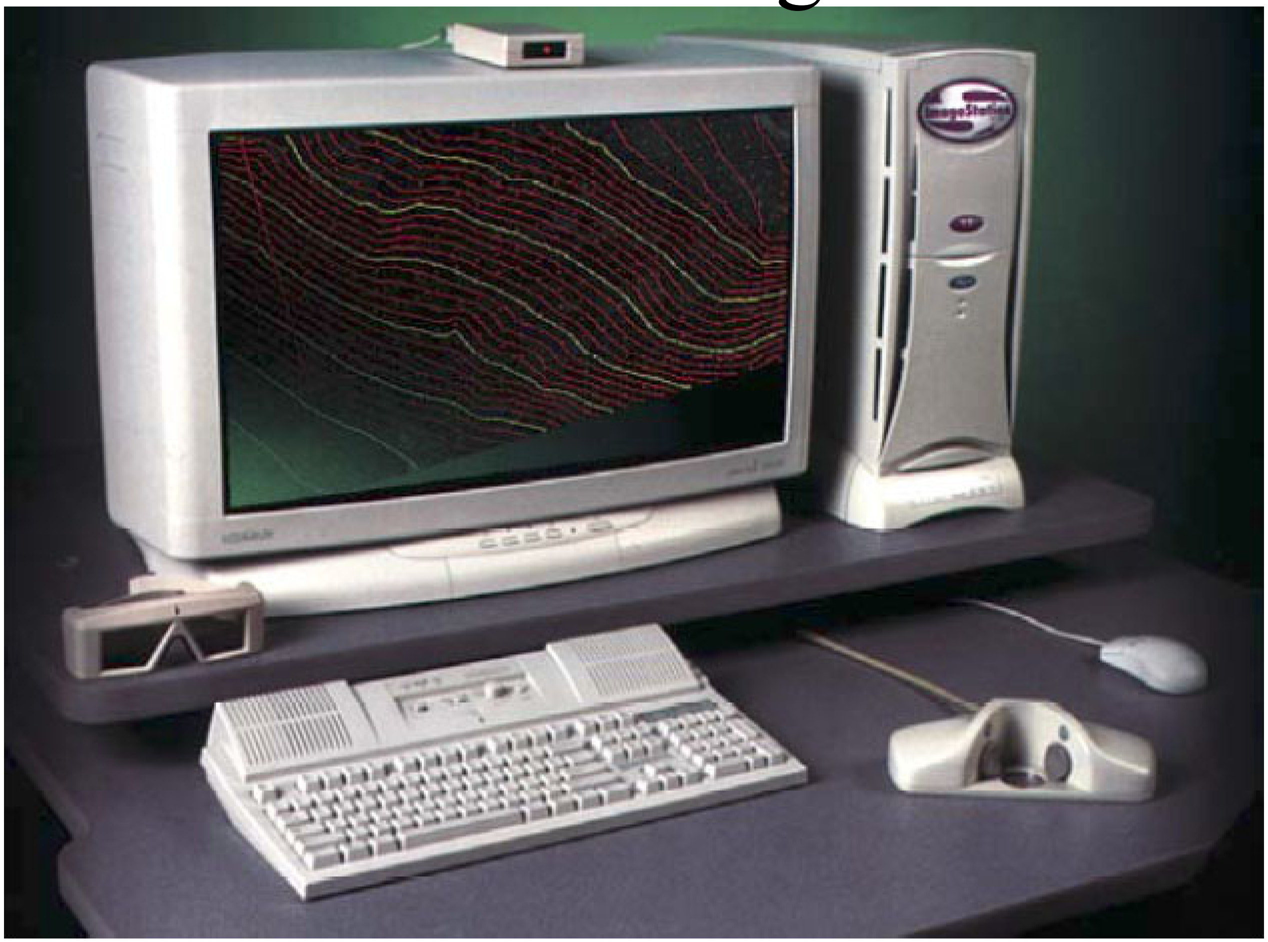


Analog Photogrammetric Plotter

3-D Viewing of 2-D imagery

Analytical Photogrammetric Plotter
Leica Geosystems: SD2000/3000

3-D Viewing of 2-D imagery

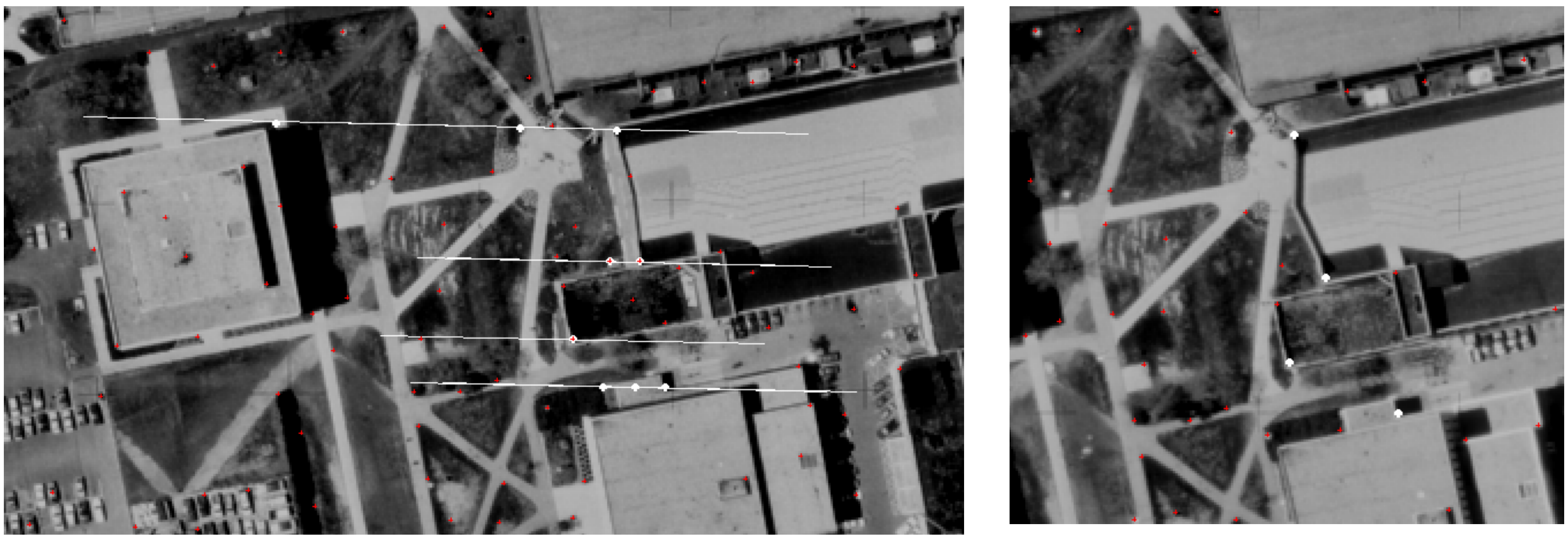


Digital Photogrammetric Plotter
Z/I Image Station

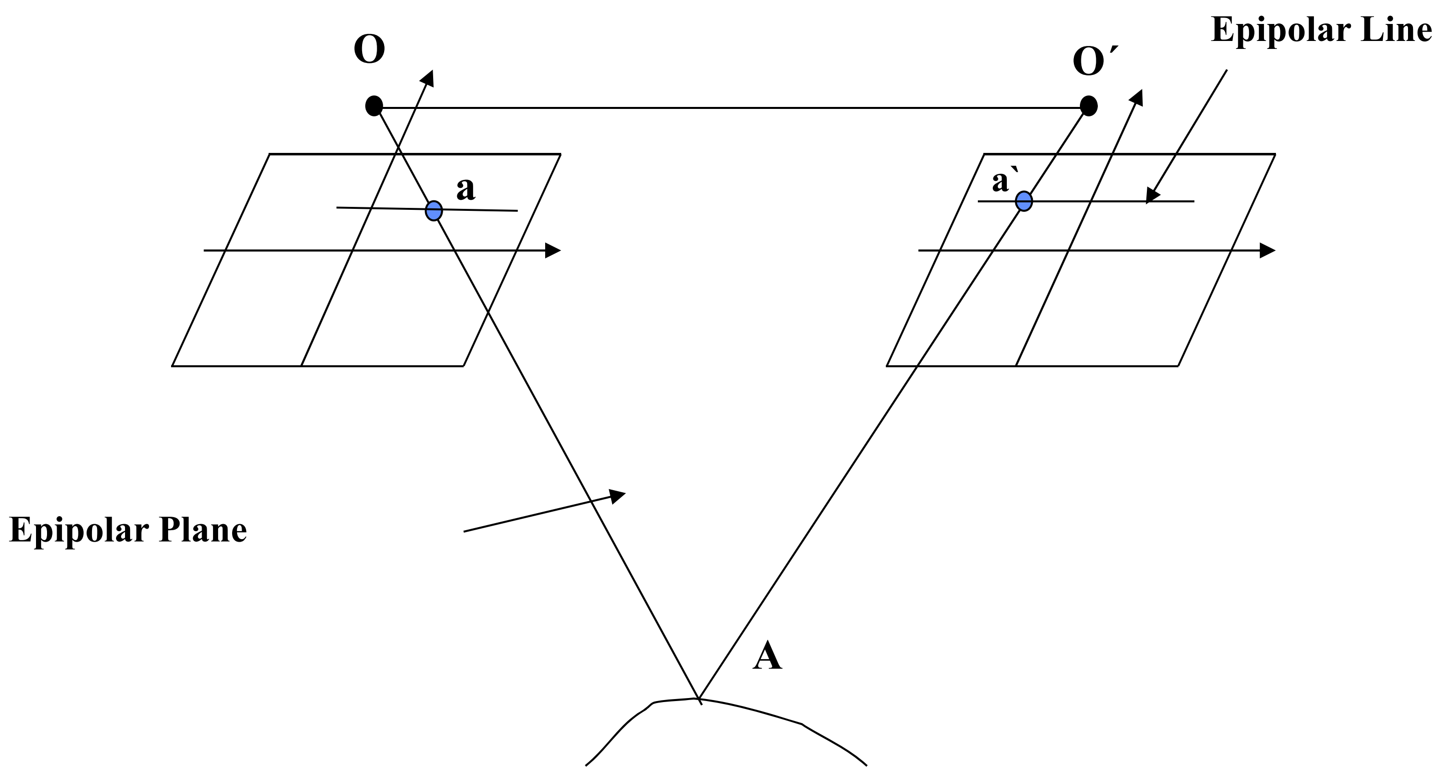
3-D Viewing of 2-D imagery

Digital Photogrammetric Plotter
Leica Geosystems

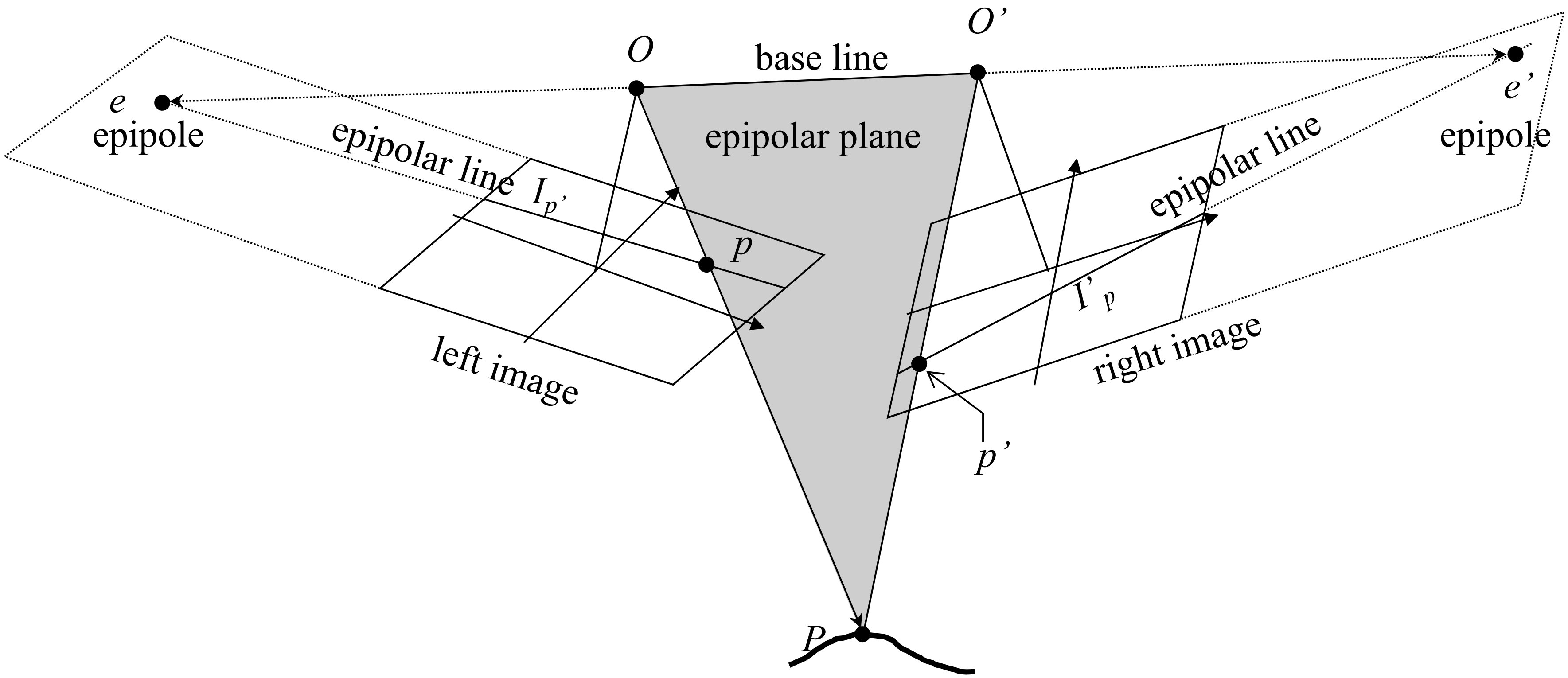
Epipolar Geometry



Epipolar Geometry



Epipolar Geometry: Definitions



Epipolar Geometry

- Epipolar plane for a given object point is the plane containing the perspective centers of a stereo-pair and the object point under consideration.
- Epipolar plane for a given image point is the plane containing the perspective centers of a stereo-pair and the image point under consideration.
- Epipolar line is the intersection of the epipolar plane with the image plane.

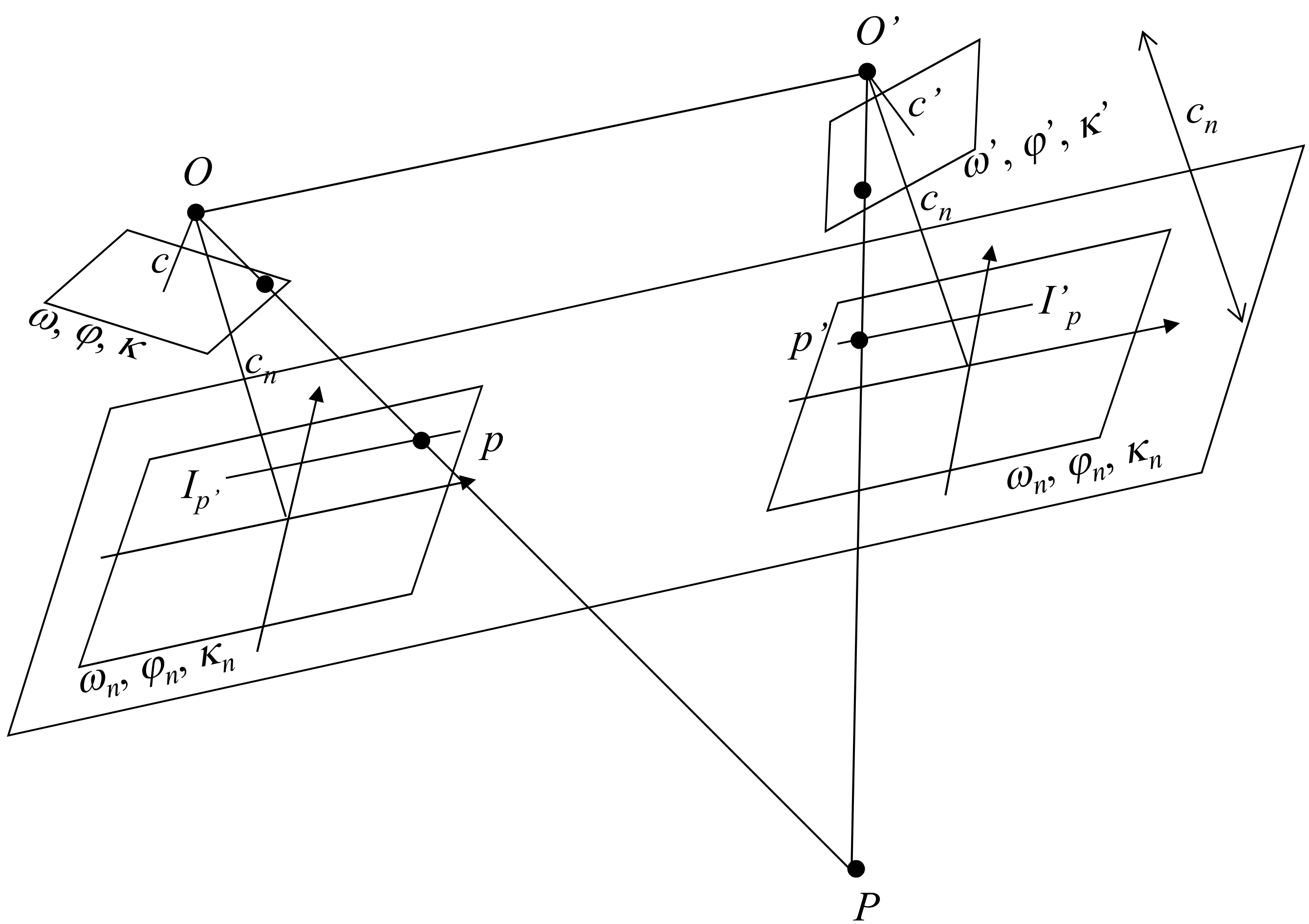
Epipolar Geometry

- The epipolar plane can be defined once we have:
  - Relative Orientation Parameters (ROP) relating the two images of a stereo-pair, and
  - Image coordinate measurements in either the left or right image.
- Epipole: The intersection of the base line with the image plane
  - All the epipolar lines pass through the epipole.
- Epipolar lines associated with different object points are not parallel except when:
  - The two images of a stereo-pair are parallel to the base line (the line connecting the two perspective centers).

Normalized Images

- Concept:
  - Create a new image at the same exposure station
  - This image is parallel to the base line.
  - The rows of the new image should be parallel to the base line.
  - The rows of the new image are the epipolar lines.
- EOP of the normalized image:
  - (X_o, Y_o, Z_o) are the same as the original image.
  - (ω_n, ϕ_n, κ_n) are chosen in such a way that:
    - The image plane is parallel to the base line (ω_n, ϕ_n), and
    - The rows are parallel to the base line (κ_n).

Normalized Image Generation



Original Image

$$\vec{v}_i = \lambda \ R^T(\omega, \phi, \kappa) \ \vec{V}_O$$

$$\begin{bmatrix} x_a \\ y_a \\ -c \end{bmatrix} = \lambda \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_A - X_O \\ Y_A - Y_O \\ Z_A - Z_O \end{bmatrix}$$

Where: λ is a scale factor

$$\vec{V}_O = 1/\lambda \ R(\omega, \phi, \kappa) \ \vec{v}_i$$

Normalized Image

$$\vec{v}_{i_n} = \lambda_n \; R_n^T(\omega_n, \phi_n, \kappa_n) \; \vec{V}_O$$

$$\begin{bmatrix} x_{a_n} \\ y_{a_n} \\ -c \end{bmatrix} = \lambda_n \begin{bmatrix} r_{11_n} & r_{21_n} & r_{31_n} \\ r_{12_n} & r_{22_n} & r_{32_n} \\ r_{13_n} & r_{23_n} & r_{33_n} \end{bmatrix} \begin{bmatrix} X_A - X_O \\ Y_A - Y_O \\ Z_A - Z_O \end{bmatrix}$$

Where: λ_n is a scale factor

$$\vec{V}_O = 1/\lambda_n \; R_n(\omega_n, \phi_n, \kappa_n) \; \vec{v}_{i_n}$$

Normalized Image

$$\vec{V}_O = 1/\lambda \; R(\omega,\phi,\kappa) \; \vec{v}_i$$

$$\vec{V}_O = 1/\lambda_n \; R_n(\omega_n,\phi_n,\kappa_n) \; \vec{v}_{i_n}$$

$$1/\lambda \; R(\omega,\phi,\kappa) \; \vec{v}_i = 1/\lambda_n \; R_n(\omega_n,\phi_n,\kappa_n) \; \vec{v}_{i_n}$$

$$\vec{v}_{i_n} = \lambda_n/\lambda \; R_n^T(\omega_n,\phi_n,\kappa_n) \; R(\omega,\phi,\kappa) \; \vec{v}_i$$

$$\vec{v}_{i_n} = \lambda_n/\lambda \; M(\omega_n,\phi_n,\kappa_n,\omega,\phi,\kappa) \; \vec{v}_i$$

Normalized Image

$$x_{a_n} = -c \; \frac{m_{11}\, x_a + m_{12}\, y_a - m_{13}\, c}{m_{31}\, x_a + m_{32}\, y_a - m_{33}\, c}$$

$$y_{a_n} = -c \; \frac{m_{21}\, x_a + m_{22}\, y_a - m_{23}\, c}{m_{31}\, x_a + m_{32}\, y_a - m_{33}\, c}$$

Normalized Image

$$\vec{v}_i = \lambda / \lambda_n \;\; M^T(\omega_n, \phi_n, \kappa_n, \omega, \phi, \kappa) \;\; \vec{v}_{i_n}$$

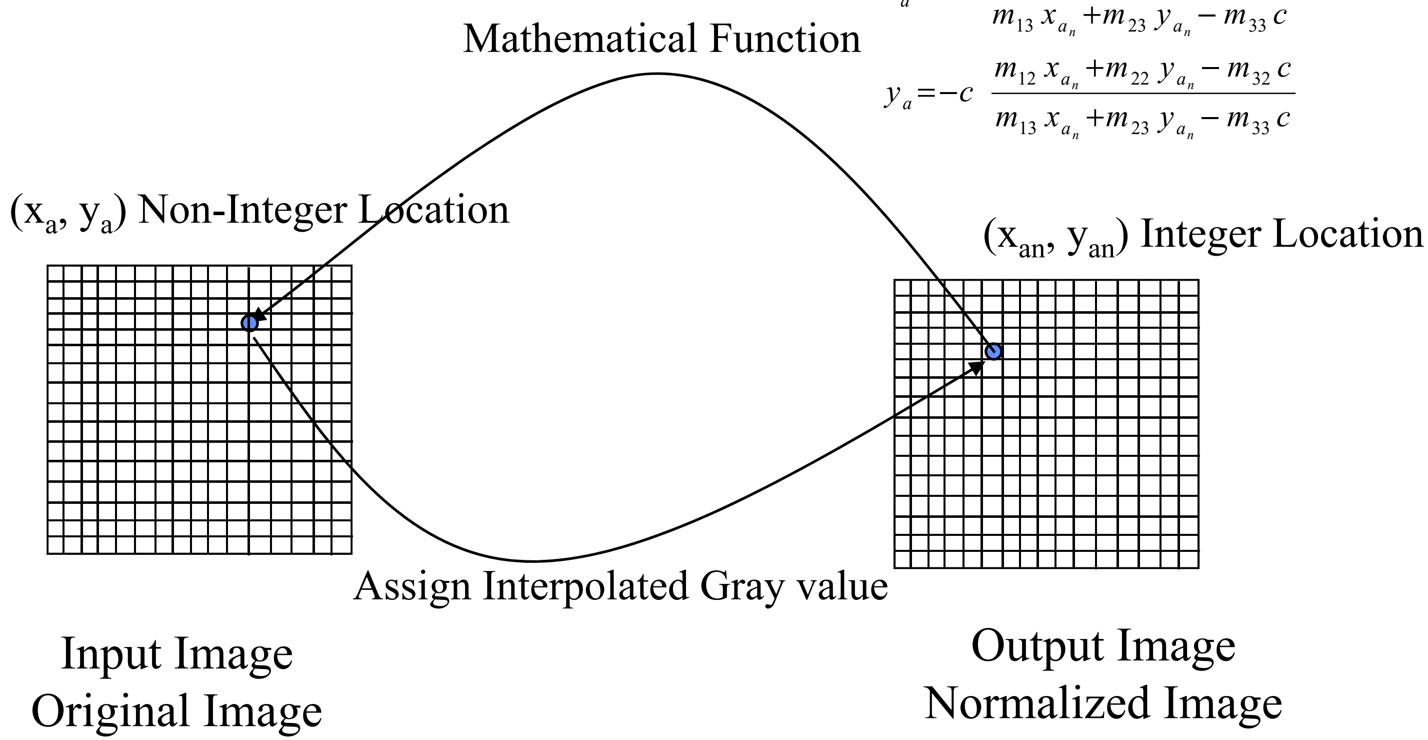
$$x_a = -c \;\; \frac{m_{11}\, x_{a_n} + m_{21}\, y_{a_n} - m_{31}\, c}{m_{13}\, x_{a_n} + m_{23}\, y_{a_n} - m_{33}\, c}$$

$$y_a = -c \;\; \frac{m_{12}\, x_{a_n} + m_{22}\, y_{a_n} - m_{32}\, c}{m_{13}\, x_{a_n} + m_{23}\, y_{a_n} - m_{33}\, c}$$

Normalized Image Generation

$$x_a = -c\,\frac{m_{11}\,x_{a_n} + m_{21}\,y_{a_n} - m_{31}\,c}{m_{13}\,x_{a_n} + m_{23}\,y_{a_n} - m_{33}\,c}$$

$$y_a = -c\,\frac{m_{12}\,x_{a_n} + m_{22}\,y_{a_n} - m_{32}\,c}{m_{13}\,x_{a_n} + m_{23}\,y_{a_n} - m_{33}\,c}$$

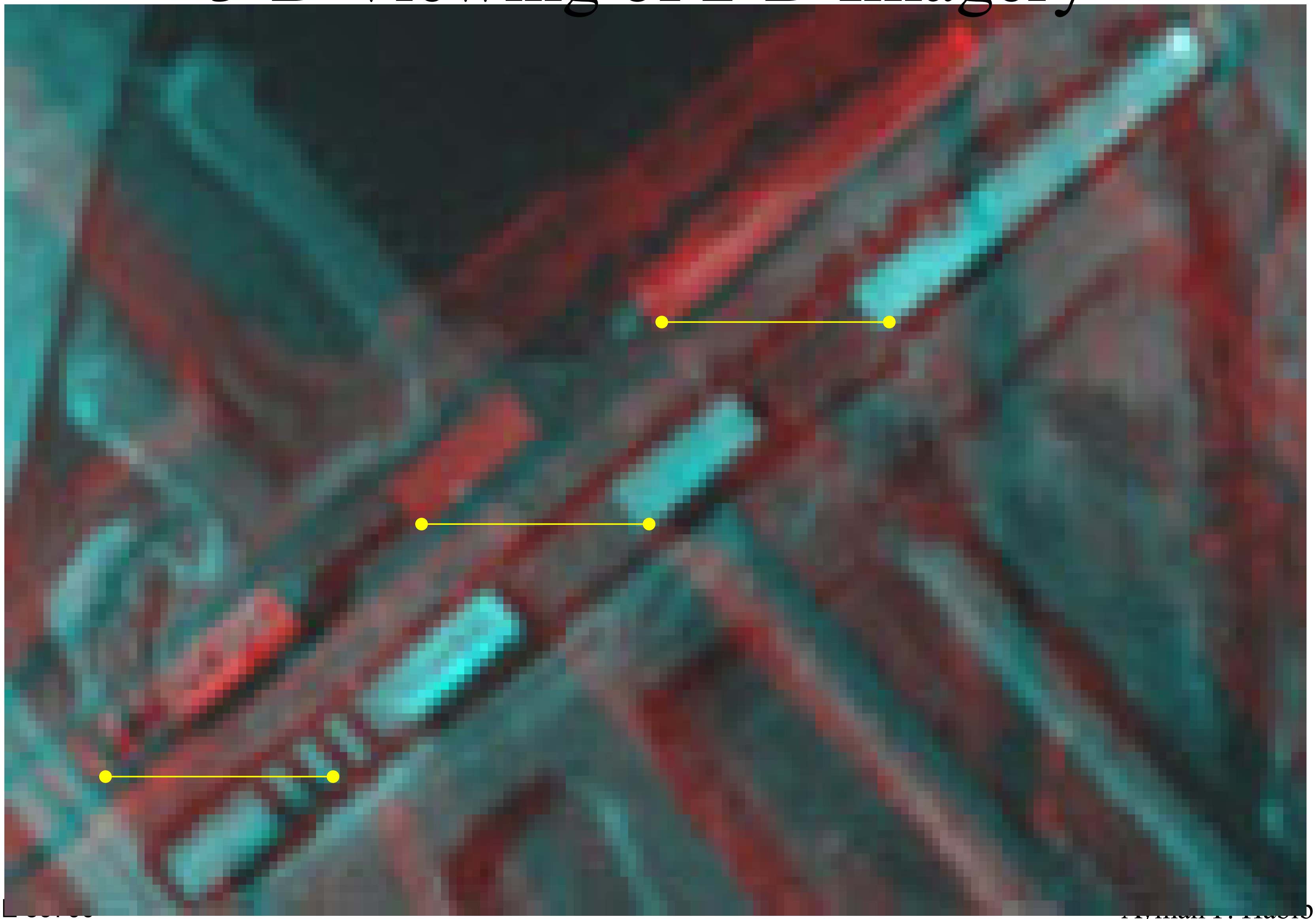


Interpolate for the gray value at (x_a, y_a)

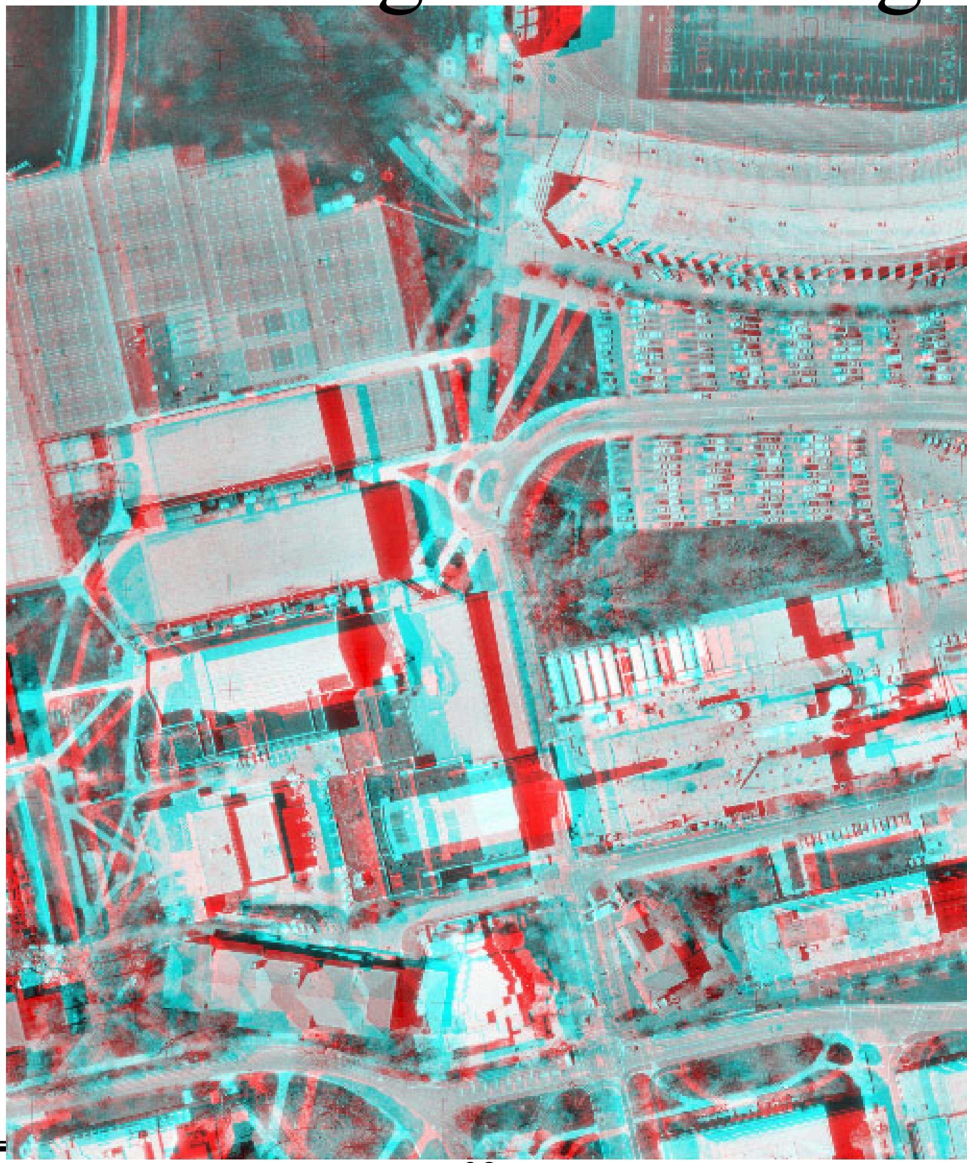
Normalized Image Generation

- Procedure:
  - Start from a certain pixel (x_{a_n}, y_{a_n}) in the normalized image,
  - Compute the corresponding (x_a, y_a) using the previous equations,
  - Compute g (x_a, y_a) using image resampling,
  - $g(x_{a_n}, y_{a_n}) = g(x_a, y_a)$, and
  - Repeat the above mentioned steps for every pixel in the normalized image.

3-D Viewing of 2-D Imagery



3-D Viewing of 2-D Imagery



3-D Viewing of 2-D Imagery

3-D Viewing of 2-D Imagery