STEREO CORRESPONDENCE

COMPSCI 773 S1 T

Vision Guided Control

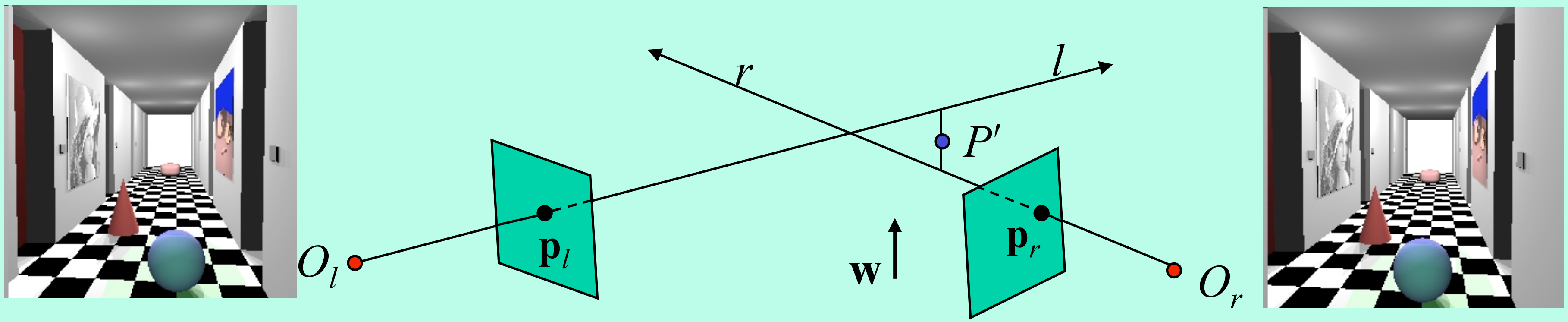
A/P Georgy Gimel'farb

Computer Stereo: Applications

Medical diagnostics, architecture, autonomous navigation, cartography, biometrics, material science, etc…

Triangulation from Projections

- Point P: projecting into the pair of corresponding points $\mathbf{p}_l$ and $\mathbf{p}_r$
- Ideally, P at the **intersection** of the two optical rays: respectively, from O_l through $\mathbf{p}_l$ and from O_r through $\mathbf{p}_r$
  - Due to approximate cameras' parameters and imprecise image locations: the actual two rays may not actually intersect in space
  - **Estimated intersection** : the point on minimum distance from both rays



Three Basic Cases

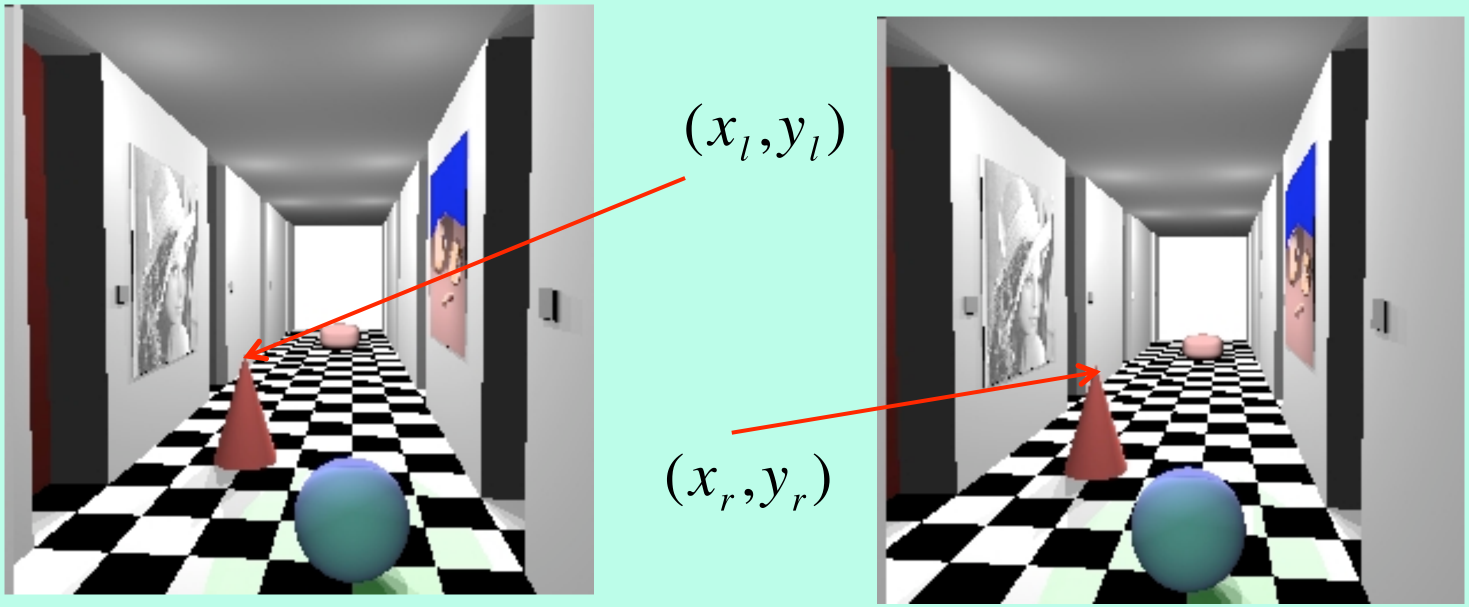
Depending on the amount of a priori cameras' knowledge:

1. **Both intrinsic and extrinsic parameters**: the unique reconstruction of a 3-D scene by ***triangulation***
2. **Only the intrinsic parameters**: a 3-D scene is still reconstructed and also the extrinsic parameters are estimated, but **up to an unknown scaling factor**
3. **Only pixel correspondences**: a 3-D scene is still reconstructed, but **up to an unknown, global projective transformation**

Matching Fundamentals

Computational stereo: 3-D coordinates of visible points from 2-D coordinates of the corresponding image pixels

– Difference of these pixel coordinates: **disparity** (or parallax)



Matching Fundamentals

- Generally, horizontal (x-) and vertical (y-) disparities
  - x-disparity: difference $d_{x,y} = x_l - x_r$ between the corresponding x-coordinates
  - y-disparity: difference $\delta_{x,y} = y_l - y_r$ between the corresponding y-coordinates
  - Usually, a horizontal stereo baseline and small y-disparities
  - *Canonical* stereo geometry (an *epipolar pair*): no y-disparities

Stereo image matching: to find disparities for all visible 3D points in a stereo pair

Disparity Map

Continuous vector-valued function $\mathbf{d}(x,y) = [d_{x,y}\,,\ \delta_{x,y}]$

- Mapping the coordinates of binocularly visible points in one image to the corresponding coordinates in the other image:

$$(x,y) \leftrightarrow (x-d_{x,y}\,,\ y-\delta_{x,y}) \Leftrightarrow \quad g_{l:x,y} \leftrightarrow g_{r:x-d_{x,y},y-\delta_{x,y}}$$

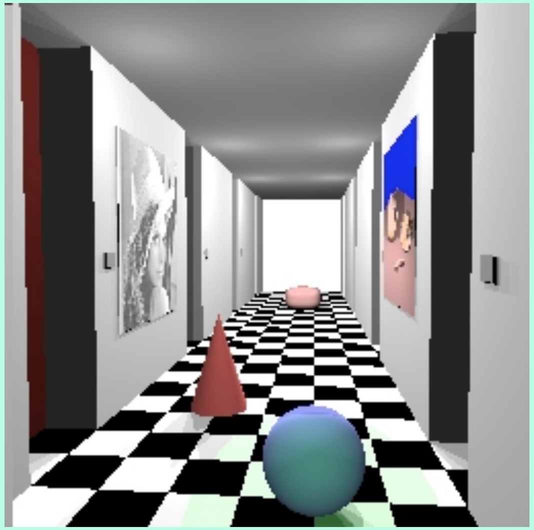
- Mapping is undefined for partially occluded points having no stereo correspondence

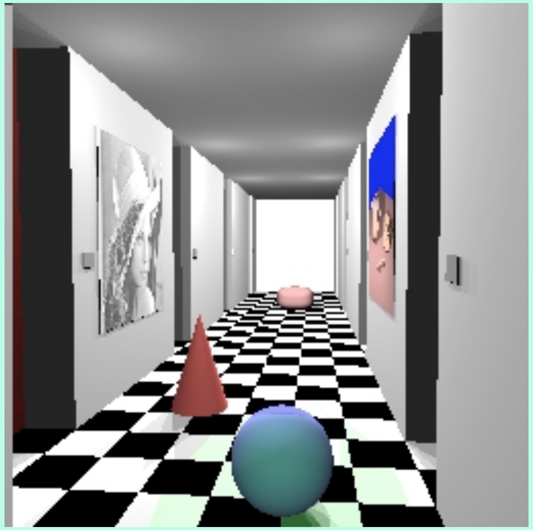
Search region: for a position (x,y) in the left image $\leftrightarrow$ a set of candidats to be explored in the right image:

$$\{(x',y') : x-d_{\max} \leq x' \leq x-d_{\min}; y-\delta_{\max} \leq y' \leq y-\delta_{\min})$$

“Corridor”: Disparity Map

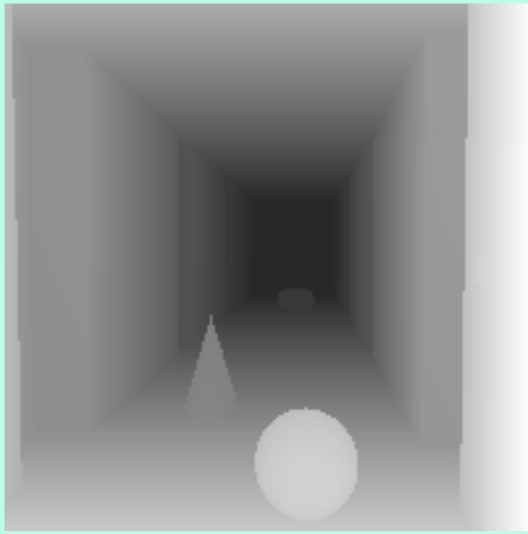
Left image of a stereo pair

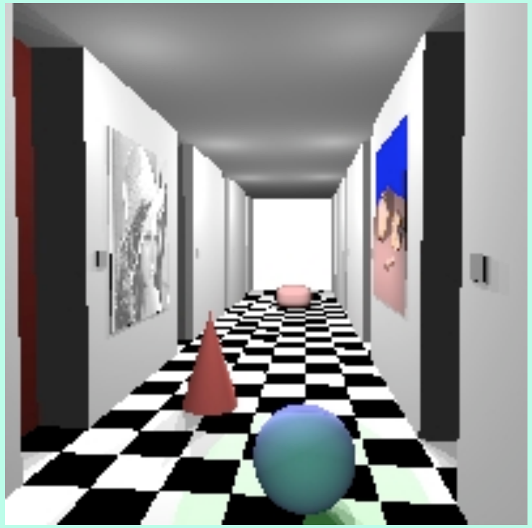




Right image of a stereo pair

Grey-coded disparity map





Cyclopean image of the 3-D scene

“Artificial Rock”: Disparity Map

Left image of a stereo pair

Right image of a stereo pair

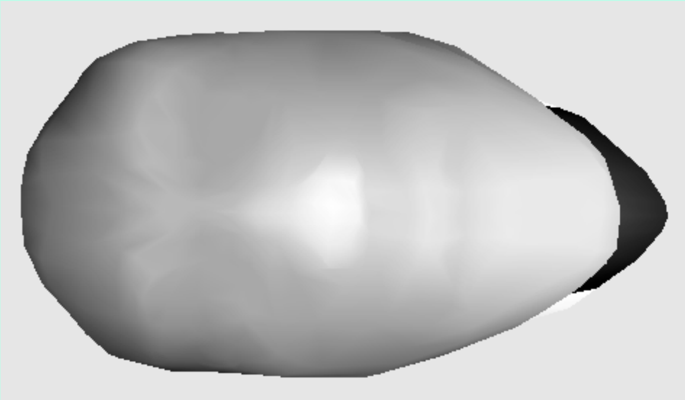
Grey-coded disparity map

Cyclopean image of the 3-D scene

"Alex": Disparity Map

Upper image of a vertical stereo pair

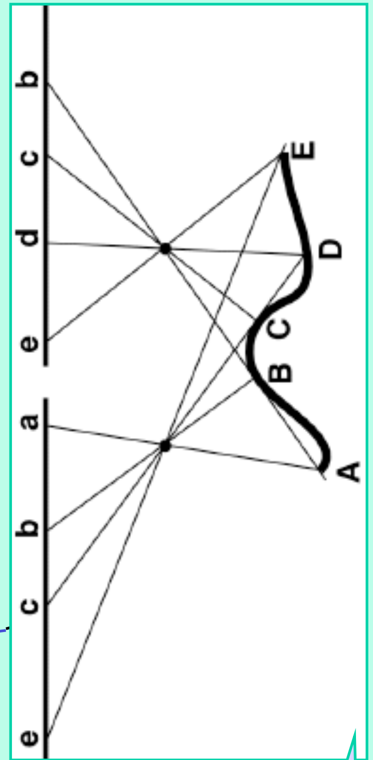
Bottom image of a vertical stereo pair



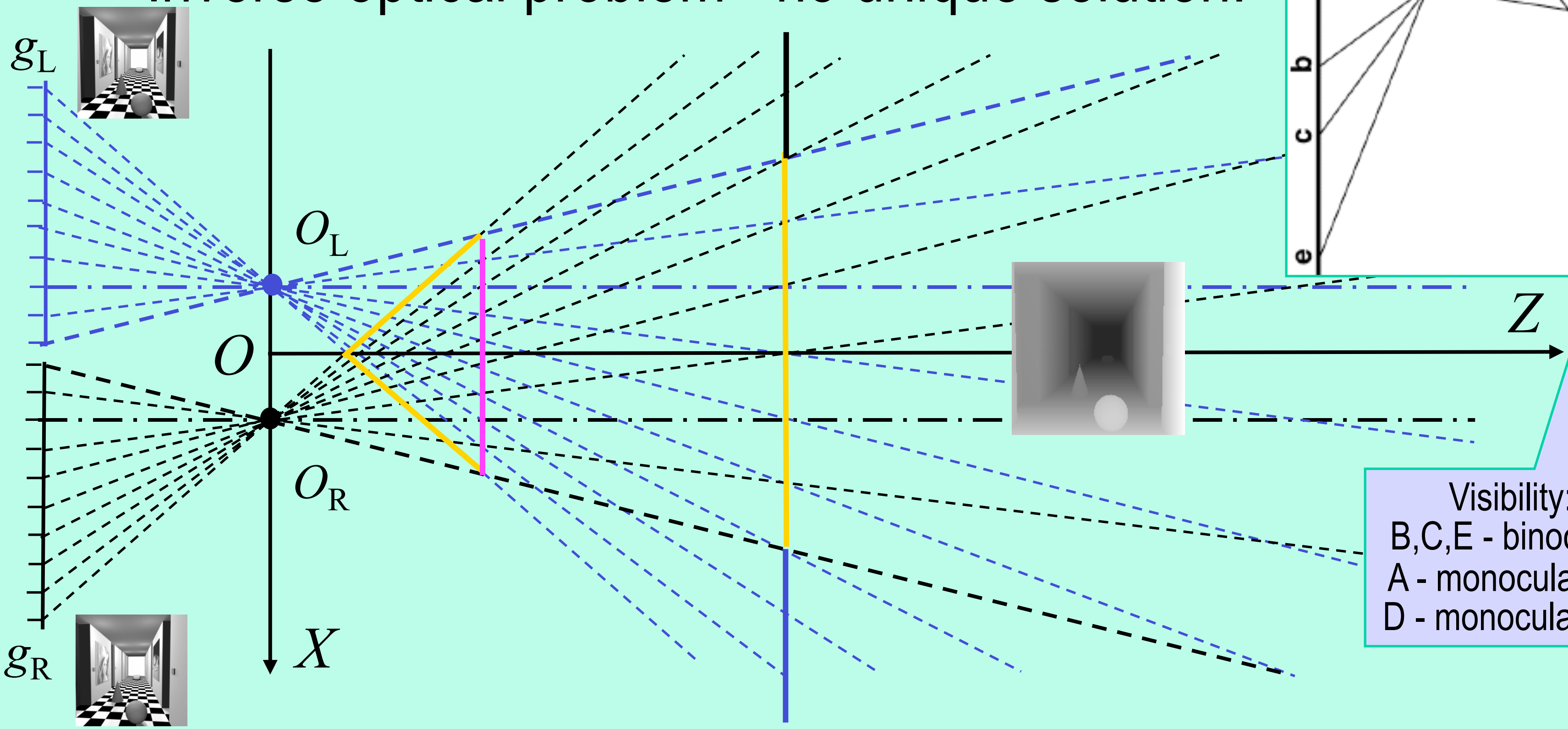
Grey-coded disparity map

Cyclopean image of the 3-D face

Ill-posed Stereo Problem

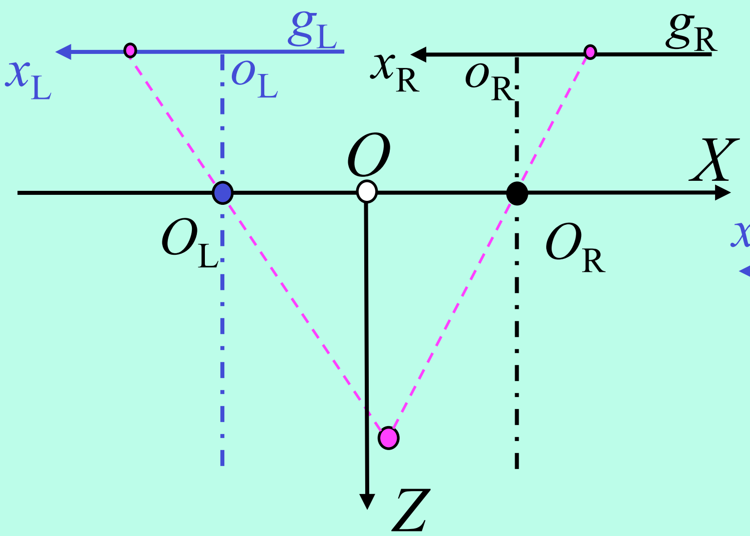


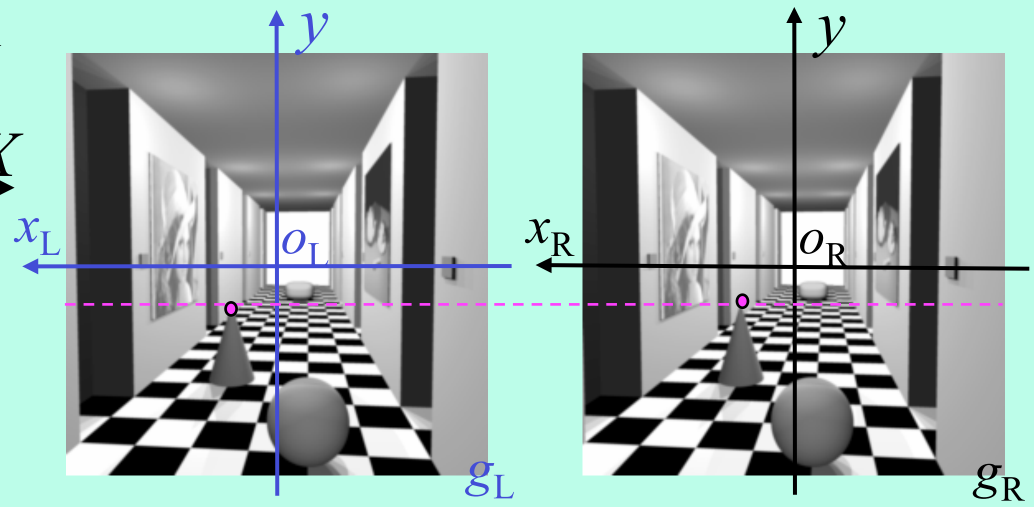
- Inverse optical problem - no unique solution!



Symmetric Canonical Geometry

- Disparity map: a set of epipolar profiles
  - Points of each profile and corresponding points along the conjugate scan-lines in images have the same y-coordinates
  - Only the signals along these scan-lines have to be matched





Symmetric Canonical Geometry

- Symmetric coordinates: $[X,y,Z]^{\mathrm{T}} \leftrightarrow [x_{\mathrm{L}},y], [x_{\mathrm{R}},y]$
  - Disparity: $d = x_{\mathrm{L}} - x_{\mathrm{R}} = bf/Z$
- Cyclopean image / disparity map: $(X,y,Z) \rightarrow (x,y,d)$

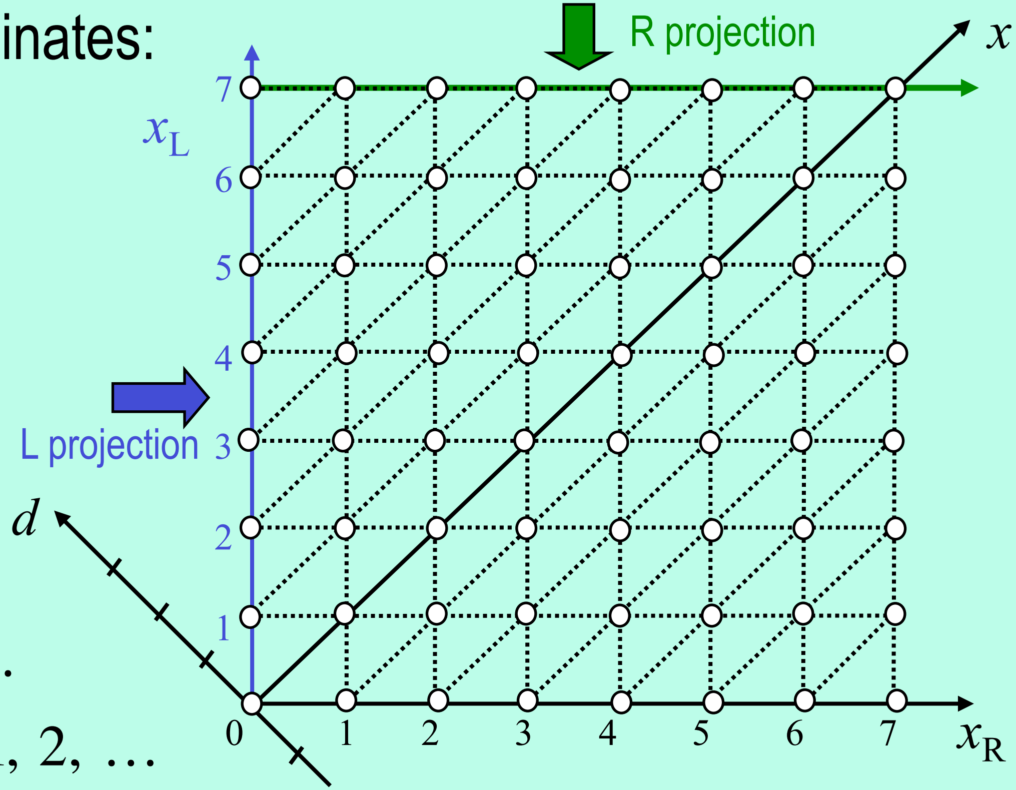
Symmetric (x,d)-coordinates:

$$\begin{pmatrix} x = (x_{\mathrm{L}} + x_{\mathrm{R}})/2 \\ d = x_{\mathrm{L}} - x_{\mathrm{R}} \end{pmatrix} \leftrightarrow \begin{pmatrix} x_{\mathrm{L}} = x + d/2 \\ x_{\mathrm{R}} = x - d/2 \end{pmatrix}$$

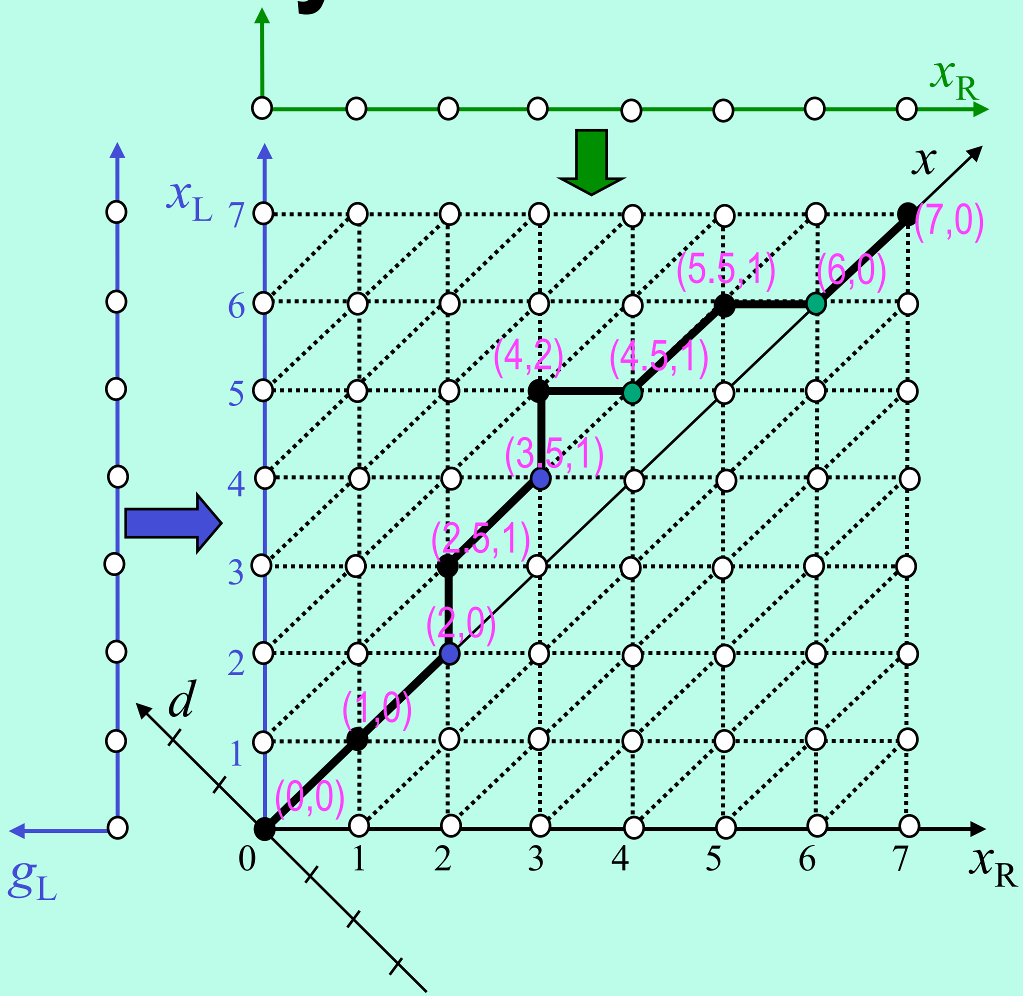
$$x_{\mathrm{L}}, x_{\mathrm{R}} = 0,\ 1,\ 2,\ \ldots$$

$$x = 0,\ \tfrac{1}{2},\ 1,\ 1\tfrac{1}{2},\ 2,\ \ldots$$

$$d = \ldots,\ -2,\ -1,\ 0,\ 1,\ 2,\ \ldots$$



Symmetric Canonical Geometry



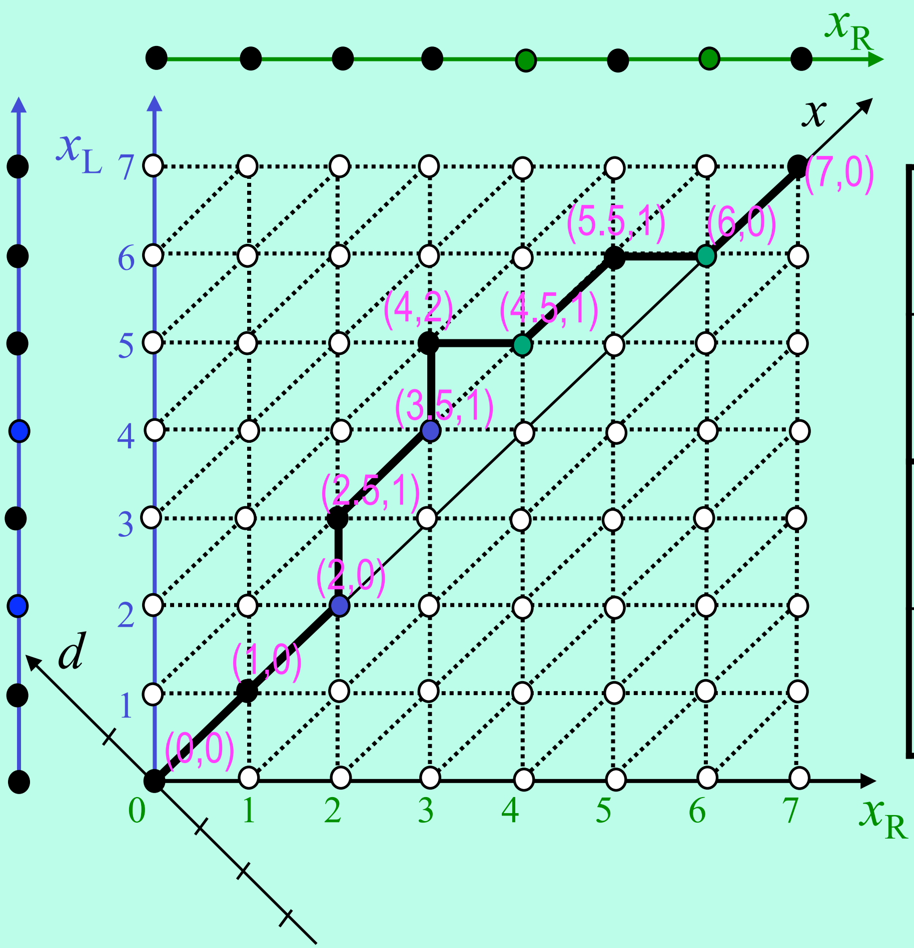
- B - binocular visibility
- ML - monocular visibility (left image only)
- MR - monocular visibility (right image only)

Integer cyclopean coordinates x
- even disparities d

Half-int cyclopean coordinates x
- odd disparities d

Basic restriction: a continuous single surface in cyclopean space

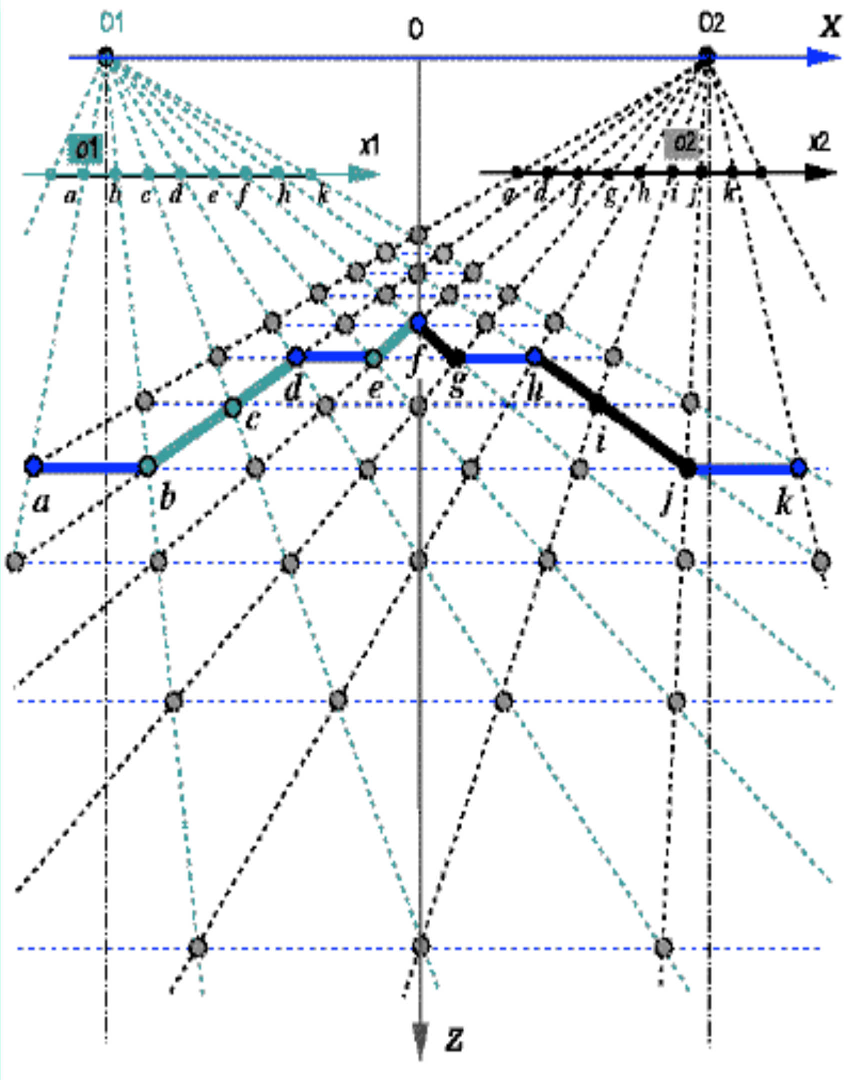
Symmetric Canonical Geometry

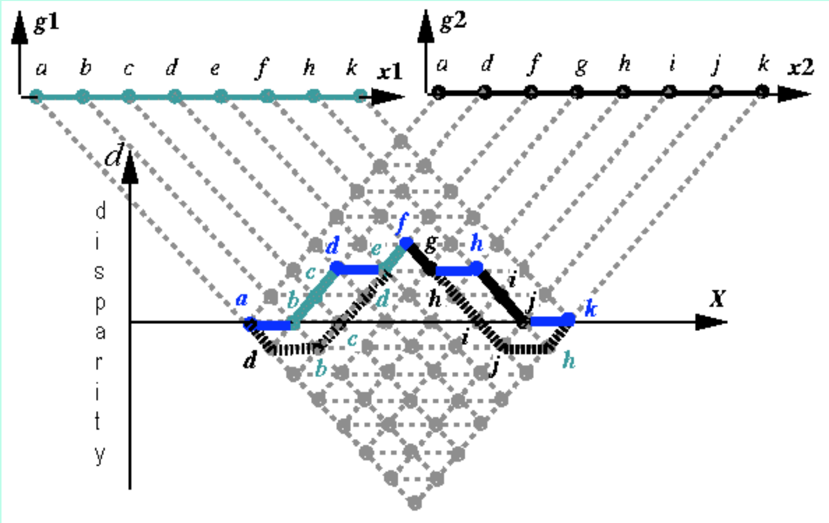


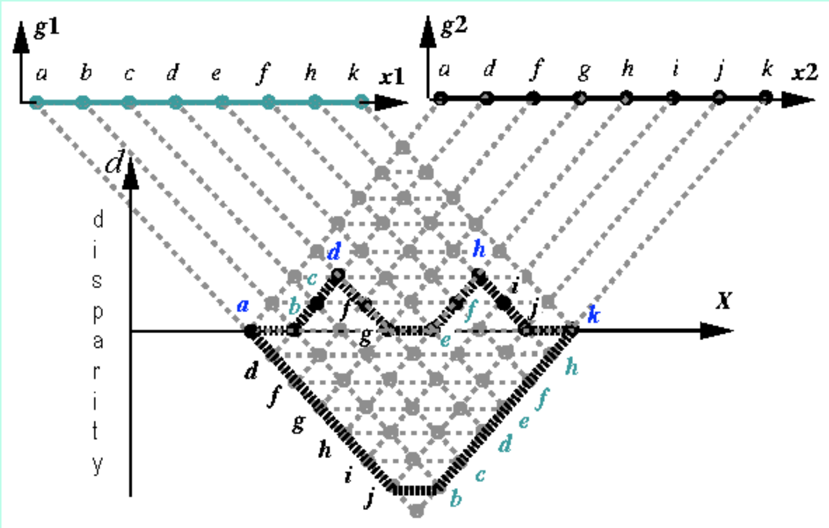
x	0	1	2	2.5	3.5	4	4.5	5.5	6	7
d	0	0	0	1	1	2	1	1	0	0
x_L	0	1	2	3	4	5	5	6	6	7
x_R	0	1	2	2	3	3	4	5	6	7

Due to assumed single continuous surface

Equivalent Surface Profiles

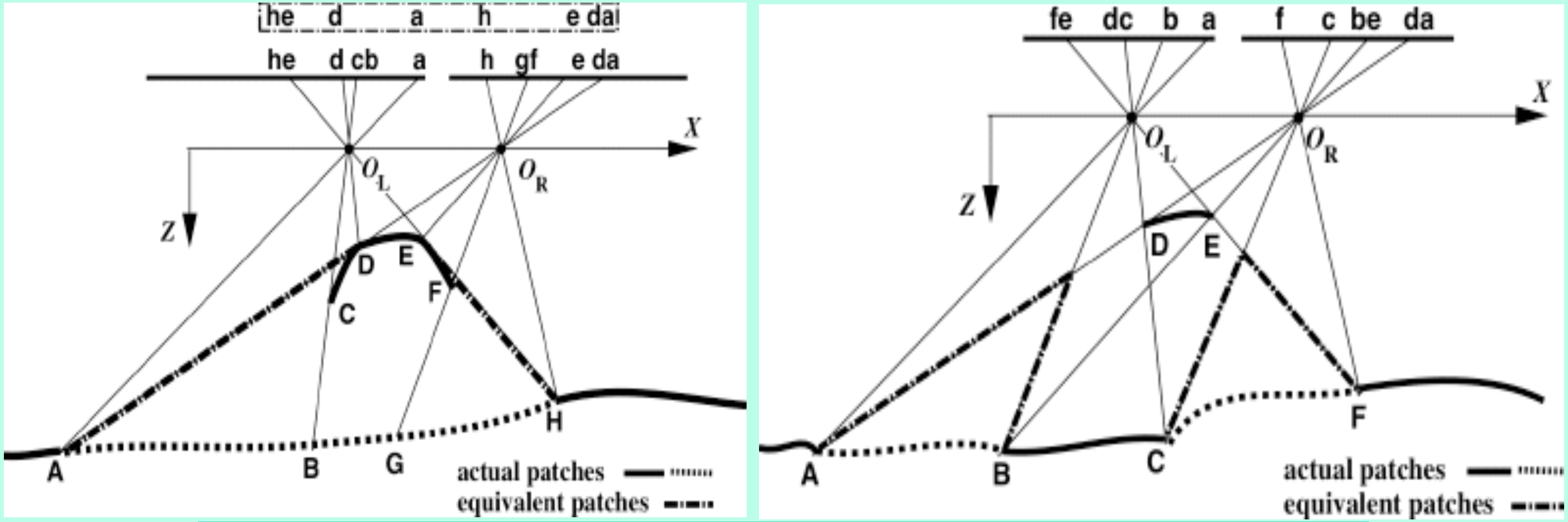






Sources of *Ill-Posedness*

- Multiple surfaces, partial occlusions, uniform texture…

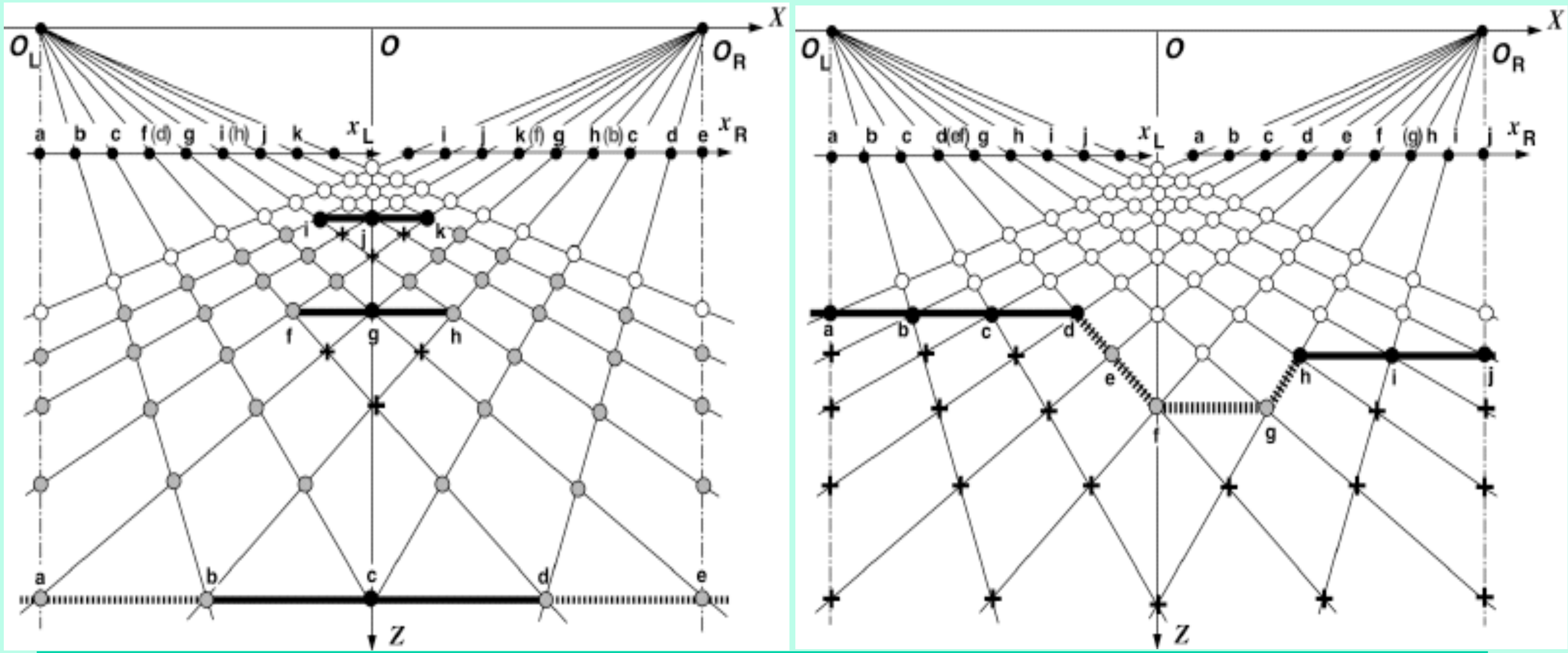


Actual disjoint and equivalent continuous surface profiles

Sources of *Ill-Posedness*

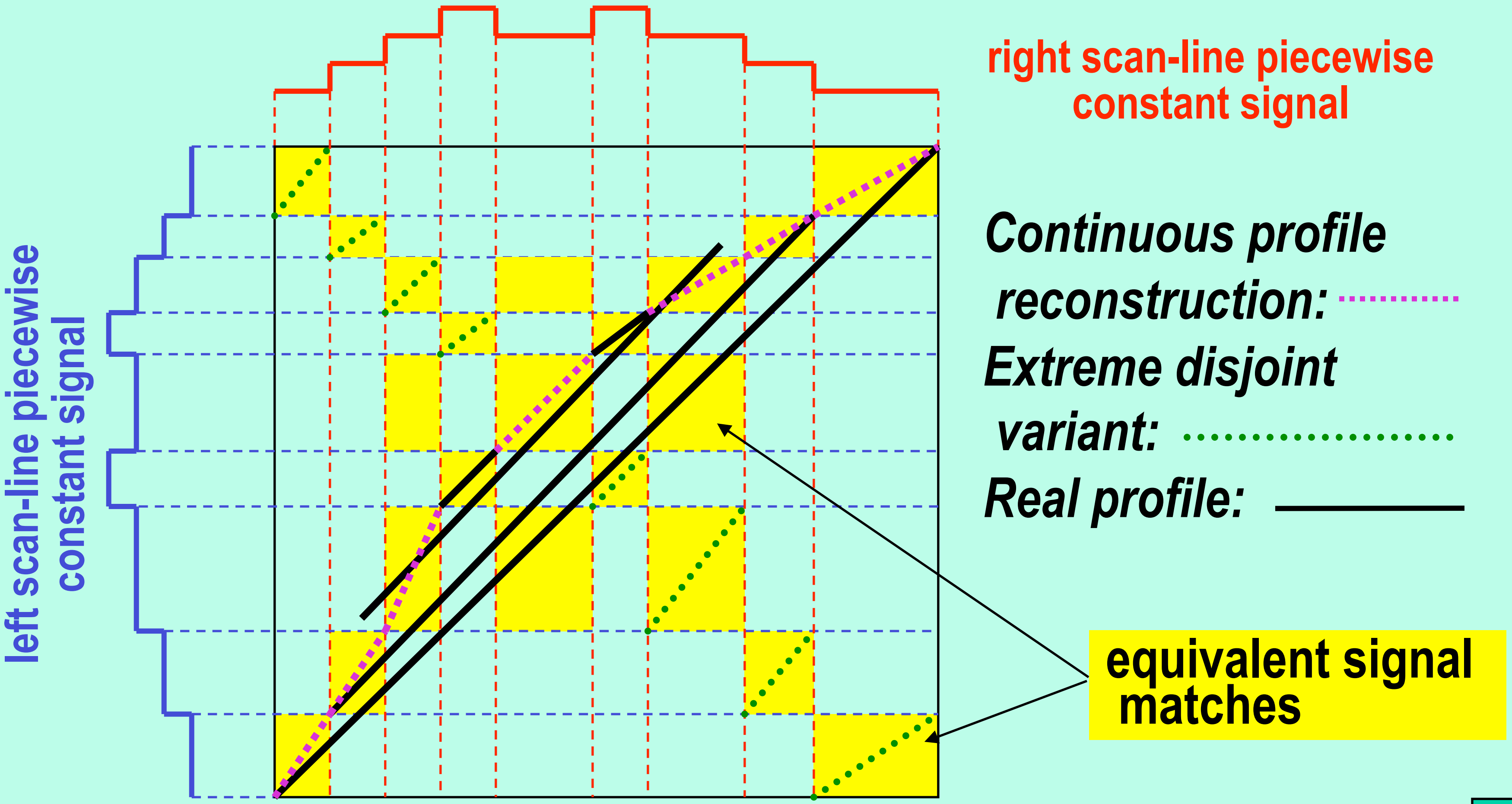
Most of 3D scenes do not conform to the single surface assumption and thus to the ordering constraint

Visibility of 3-D Points

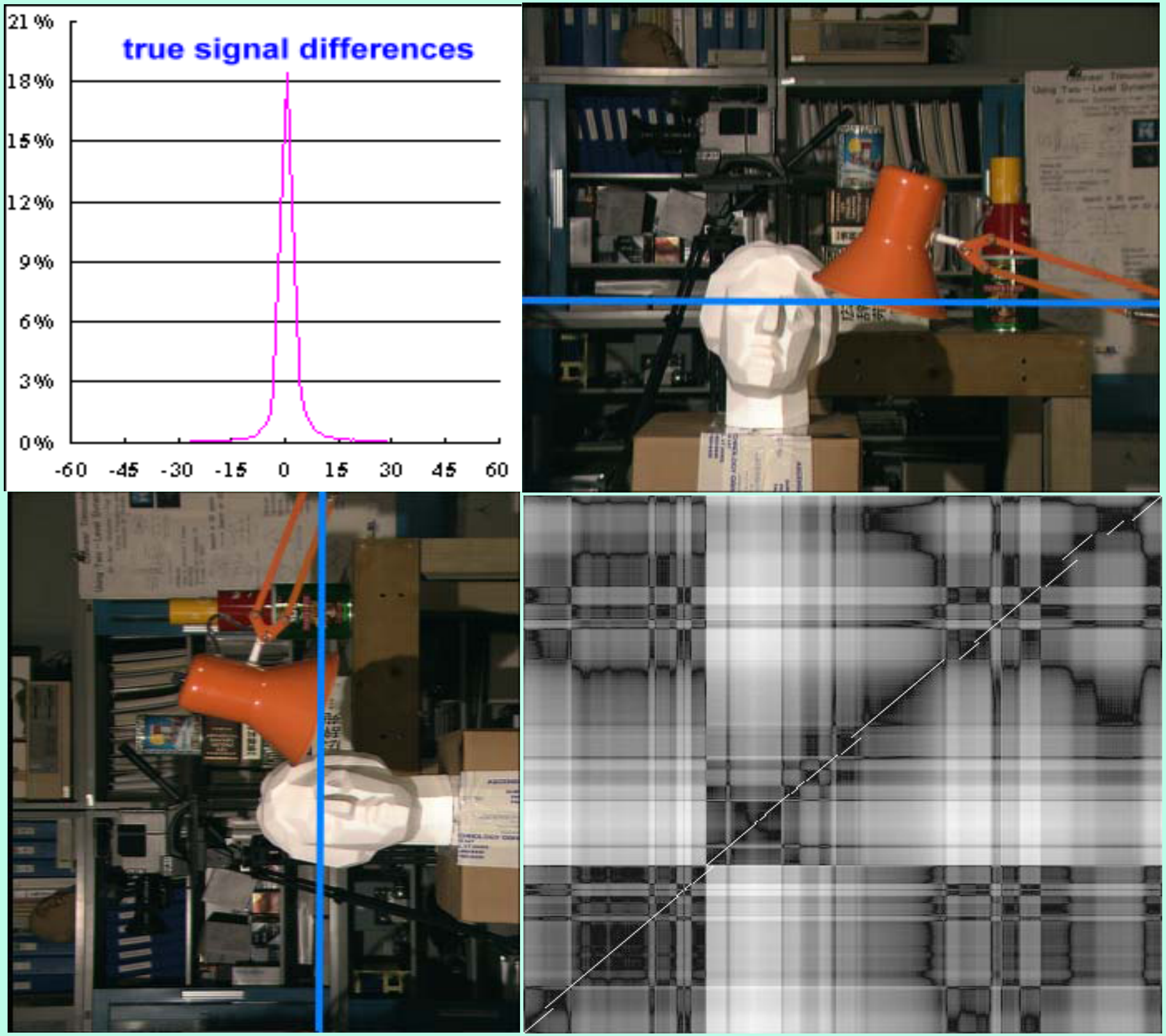


Multiple disjoint and continuous 3-D profiles: ● binocularly visible points; **+** : invisible (occluded) points; ◉ partially occluded points; ○ transparent points

Real Vs. Equivalent Profiles



Distribution of signal differences



Signal-to-signal matching along the same scan -line 173; stereo pair "Tsukuba"

Grey-coded signal differences

3-D Reconstruction

- Because of ill-posedness, it is impossible to reconstruct precisely the original 3D scene from a stereo pair
  - Goal of stereo matching is therefore more limited and more practical: to bring the reconstructed surfaces close enough to those perceived visually or with the photogrammetric tools
- Due to a multiplicity of visually observed scenes - only very general prior knowledge to constrain optical 3-D surfaces under reconstruction
  - E.g. expected smoothness, curvature, discontinuities, etc

General Matching Constraints

Reflecting intrinsic properties of stereo viewing and a 3-D scene

Epipolar constraint: 1-D search along the conjugate scan-lines

- Rectified stereo pairs reduced to the canonical stereo geometry
- Reduced search region; excluded false matches across the scan-lines

Uniqueness constraint: every pixel in one stereo image has at most one corresponding pixel in the other image

- Every visible 3-D point is observed either binocularly or only monocularly
- Monocular observation: *partial occlusion* (no stereo correspondence)

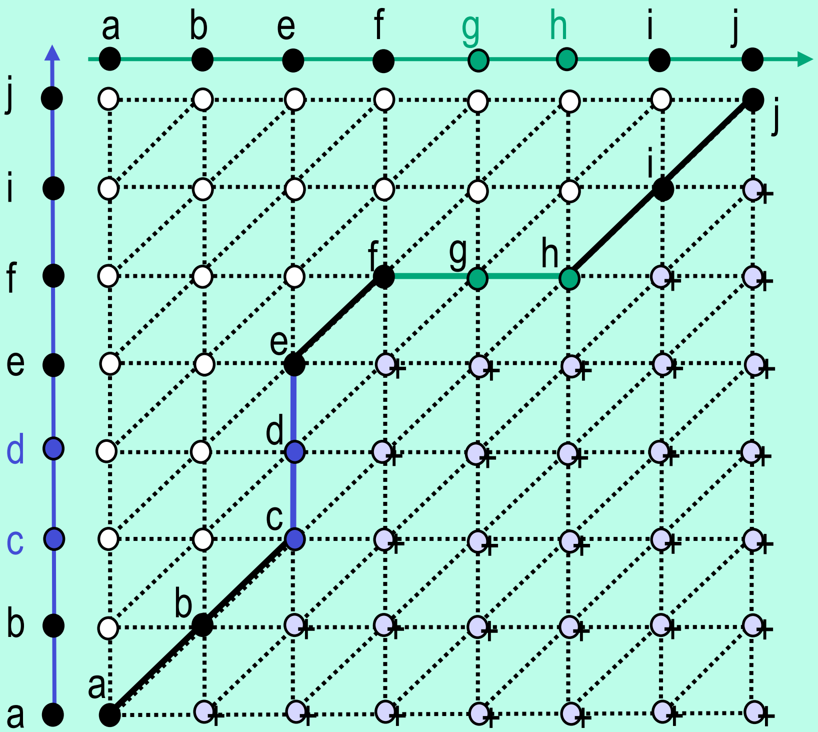
Disparity range $[d_{\min}, d_{\max}]$ is typically known for a 3-D scene

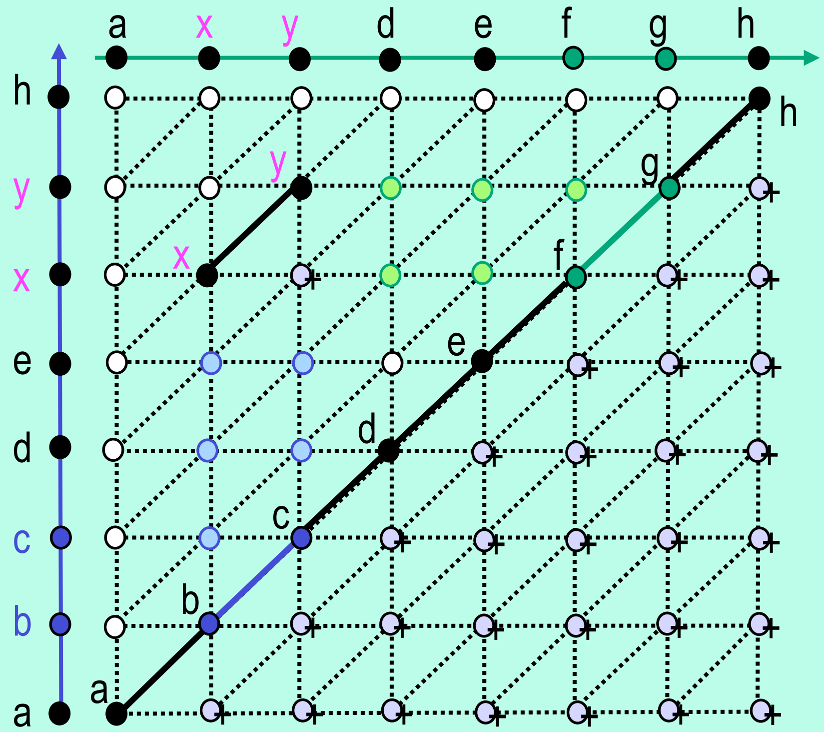
- Reduced search region; excluded false matches outside the range

Continuity constraint: smooth surfaces except for object boundary

General Matching Constraints

Ordering constraint: the same order of corresponding points along the conjugate scan-lines - but only for a ***continuous single*** profile!





Simplifying Constraints

- **Equal corresponding signals**
  - Lambertian (direction-independent) reflection of 3-D surfaces
  - Simple matching scores like $|g_{\mathrm{L}:x,y} - g_{\mathrm{R}:x,y}|$ or $(g_{\mathrm{L}:x,y} - g_{\mathrm{R}:x,y})^2$
- **Frontal parallel surfaces**
  - Area-based correlation matching: constant disparity and no occlusions over the matching windows
- **Similarity of features**
  - Feature-based matching: similar and mutually consistent groups of corresponding features in both images

Stereo Correspondence

- Similarity between (dissimilarity of) stereo images
  - Under their relative photometric and geometric distortions
  - Different projective views, camera noise, occlusions, …
- **Photometric distortions**
  - Non-uniform reflection of observed 3-D surface points in different directions
  - Non-uniform and noisy transfer factors over a field-of-view (FOV) of every stereo camera
    - Spatially variant contrast and offset deviations between corresponding signals in stereo images

Stereo Correspondence

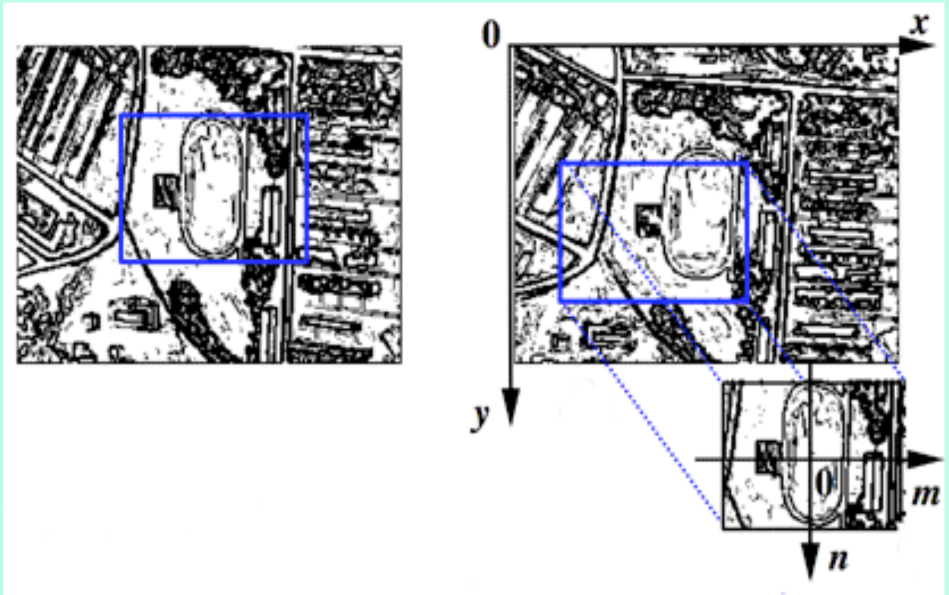
- **Geometric distortions** due to projection of a 3-D scene onto the two image planes
  - Spatially variant disparities of the corresponding points
    - Corresponding regions in stereo images may differ in positions, scales, and orientations
  - Partial occlusions (monocular visibility) of some 3-D points
    - Such regions have no stereo correspondence in principle
  - *A single continuous visible surface*: the images preserve the natural x- and y-order of binocularly visible points (BVP)

Stereo Correspondence

- Stereo matching techniques differ in:
  - Which image similarities (dissimilarities) are measured
  - Which relative image distortions are taken into account
  - Which constraints / regularising heuristics are involved
  - How a stereo pair is matched as a whole
- Image signals: grey values, colours (RGB, HSV, etc)
- Feature-based vs. intensity-based stereo matching
  - Dense and sparse disparity maps

Feature-based Matching

- Specific area, linear, and point features individually found in each image of a stereo pair
  - Edges, corners, T- junctions, isolated local shapes, etc
  - Only features are tested for similarity
  - Feature matching usually cannot produce dense disparity maps



Intensity-based Matching

- Similarity (or dissimilarity) between the images in terms of the initial signals for all binocularly visible 3-D points
  - Signals: grey levels, colours, or multi-band signatures
  - Math models to relate optical signals from the observed 3-D points to the image signals in the corresponding pixels
  - Similarity (dissimilarity $D(\mathbf{d} \mid \mathbf{g}_L, \mathbf{g}_R)$) between the corresponding image pixels or regions is derived from the model
  - The similarity score has to be invariant to relative geometric and photometric image distortions the model accounts for

Local and Global Optimisation

- 3-D reconstruction: search for the max similarity (or min dissimilarity) between the corresponding regions / pixels
  - Similarity measure accounts for admissible image distortions and includes regularising constraints:
    - E.g. to deal with partial occlusions or multiple equivalent optima
    - For a single continuous surface: visibility and ordering constraints
- Scenarios of reconstructing a 3-D scene: $\mathbf{d}^* = \max_{\mathbf{d}} \{ D(\mathbf{d} \mid \mathbf{g}_L, \mathbf{g}_R) \}$
  - Exhausting variants of visible surfaces by global optimisation
  - Independent selection of each 3-D point by local optimisation
  - Successive search for each next small surface patch by local optimisation in order to add it to the already found surface

Local Optimisation

- **Pros**: simple computations; easily takes account of both *x*- and *y*-disparities of the corresponding pixels
- **Cons** for independent selection of 3-D points: the found surfaces may violate visibility and continuity constraints
- **Cons** for guiding next search by the current surface: due to accumulation of local errors, the search regions after a few steps may become completely wrong
- **Cons** in both cases: it needs intensive on- or off-line editing of the reconstructed 3-D scene to fix errors

Global Optimisation

- **Pros**: less sensitive to local errors due to constraints on the conjugate scan-lines or the entire stereo images
- **Cons**: generally, it is an **NP-hard** problem (due to 2-D constraints on disparities in the neighbouring points)
  - It is feasible only in particular cases when direct exhaustion of all the variants (with the exponential complexity) is avoided
  - Known approximate solutions are still too complex and thus too slow for processing large-size images of practical interest
  - Profile-by-profile reconstruction by 1-D dynamic programming is fast but takes no account of constraints across the profiles

Global Optimisation

- Two popular tools for approximate global optimisation:
  - Graph min-cut iterative algorithms
    - Exact solution for a minimum cut / maximum flow problem
    - Approximation of stereo matching with a sequence of graph min-cuts
    - Disparity map: the constrained dissimilarity between two stereo images within a fixed factor from the global minimum
  - Loopy belief-propagation algorithms
    - Computing marginal posterior probabilities of constrained disparities for each surface point
    - Convergence on loop-less graphs (trees)
    - Under specific conditions, convergence on loopy graphs, too