3D VISION

Stereo Correspondence (Matching)

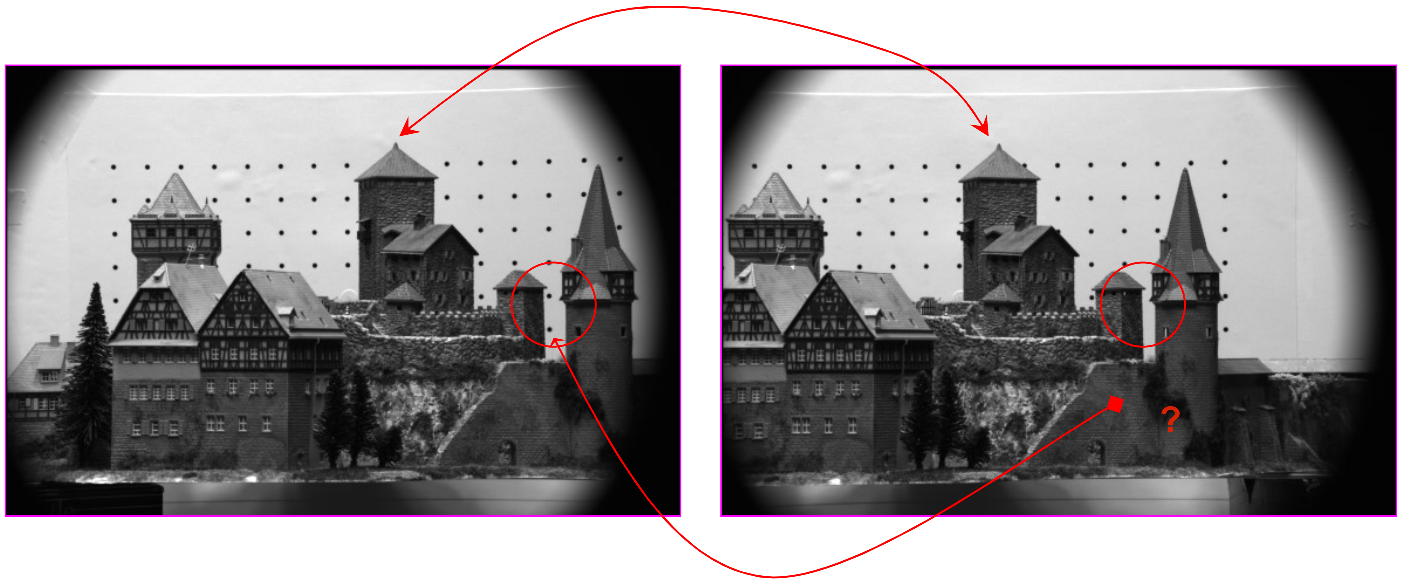
Systems Engineering and Automation

Miguel Hernández University

Table of Contents

Correspondence Problem

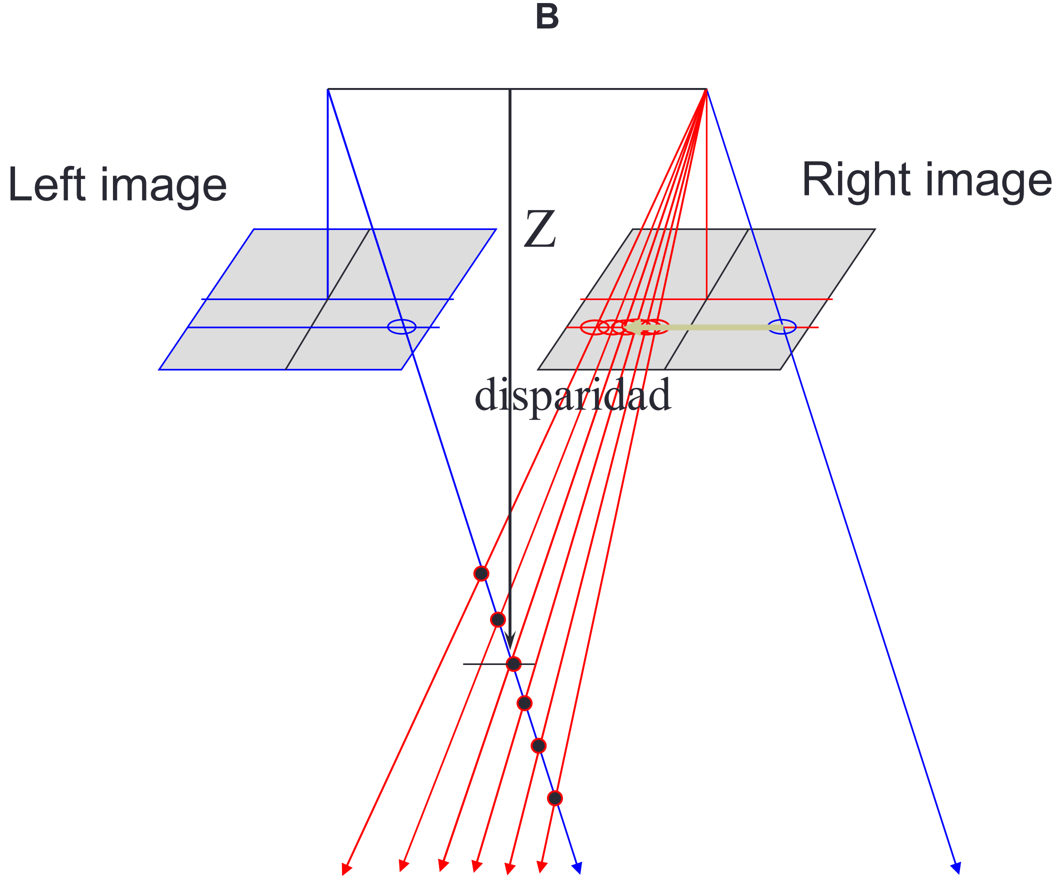
- Given a *feature* in an image (point, edge,...), projection of a *primitive* in a 3D space (whose location is unknown), obtain what is the projected *feature*, of the same *primitive*, in another image taken from a different point of view



Correspondence Problem

- Disparity
  - The correspondence problem can be formulated as the search for a disparity function:

$$D: F(x,y) \rightarrow (d_x, d_y) \perp$$
$$(x', y')^T_F = (x, y)^T_F + (d_x, d_y)^T$$



Correspondence Problem

- This is an ill-conditioned problem:
  - The left and right images captured by a stereoscopic vision system are obtained from different positions and angles.
  - The lighting conditions can be very different.
  - In addition, the cameras are intrinsically different (different calibration data).
  - Geometrically, there can be multiple solutions or no solution: one point can be found in one of the images and not in the other (occlusions, or outside the FoV of the camera).
  - You can also find a false correspondence of the feature (e.g. points of very similar appearance).

Correspondence Problem

- 3D reconstruction is linked to solving the problems of Correspondence and Calibration
- Solving the correspondence problem requires imposing restrictions on the geometric model of cameras, the photometric model of scene objects, and the used primitives
- Stereo correspondence is generally limited to a range of distances (depth).
    - In the human case this range is located between 40 cm and about 100m (from 10m depth perception is very limited). (Pannum fusional limit)
- Its solution generally involves a high computational cost.

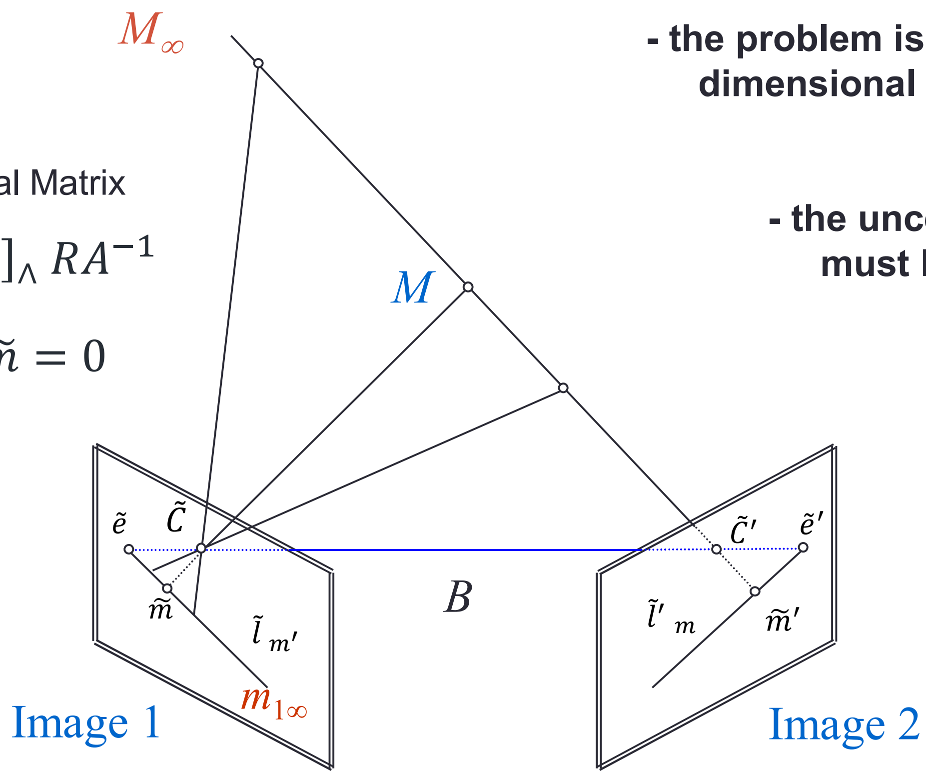
Table of Contents

Restrictions on Correspondence

- Geometric restrictions imposed by the imaging system:
  - Epipolar restriction
  - Disparity limit
- Geometric constraints from the viewed objects:
  - Positional order
  - Disparity gradient limit
  - Uniqueness
  - Surface continuity
  - Figural continuity
- Photometric restrictions
- Local primitive restrictions

Epipolar Restriction

- Given a point m' in the image 2, its corresponding point must be on the epipolar line $l_{m'}$ (intersection of the epipolar plane Π{m',C,C'} and the image plane 1).



Fundamental Matrix

$$F \cong A'^{-T}[t]_\wedge RA^{-1}$$

$$\tilde{m}'^T \tilde{F} \tilde{m} = 0$$

- the problem is reduced to a one-dimensional search

- the uncertainty in C's position and C' must be taken in account

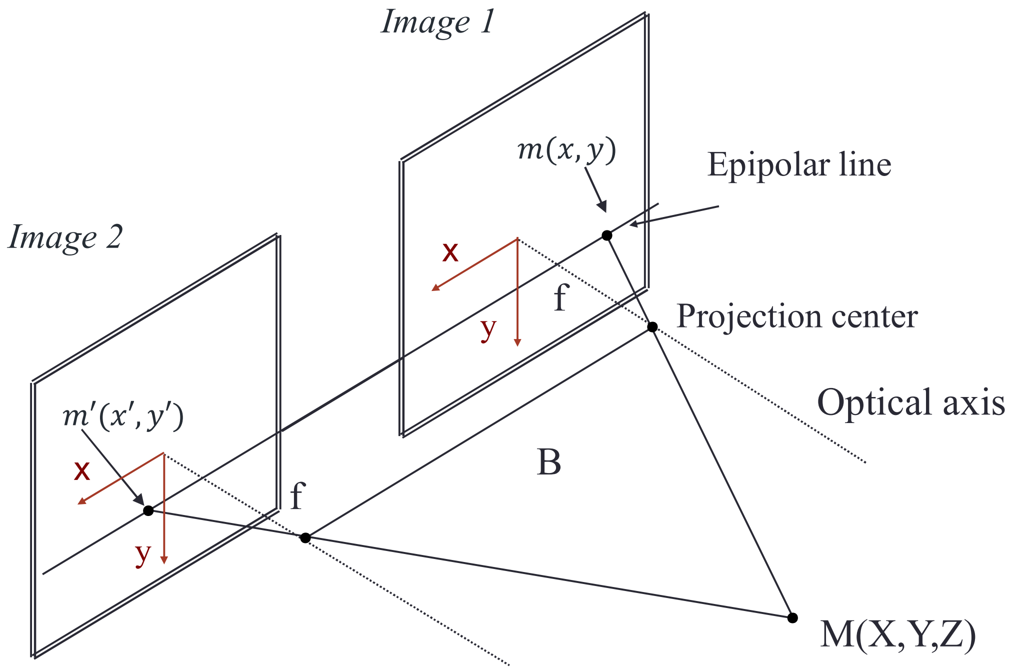
$$M_\infty \cong R^T A'^{-1} \tilde{m}'$$

$$\tilde{m}_\infty \cong A \cdot R^T A'^{-1} \tilde{m}'$$

$$\tilde{e} \cong -A R^T t$$

Epipolar Restriction

- Aligned axes systems



Epipolar restriction:

$$y = y'$$

Restrictions on Correspondence

- Geometric restrictions imposed by the imaging system:
  - Epipolar restriction
  - Disparity limit
- Geometric constraints from the viewed objects:
  - Order
  - Disparity gradient limit
  - Uniqueness
  - Surface continuity
  - Figural continuity
- Photometric restrictions
- Local primitive restrictions

Disparity Limit

- The disparity is a function of the distance between the object and the camera:
  - Parallel axes:

$$\Re_D = [d_{\min}, d_{\max}] \Leftrightarrow \Re_Z = [Z_{\min}, Z_{\max}]$$

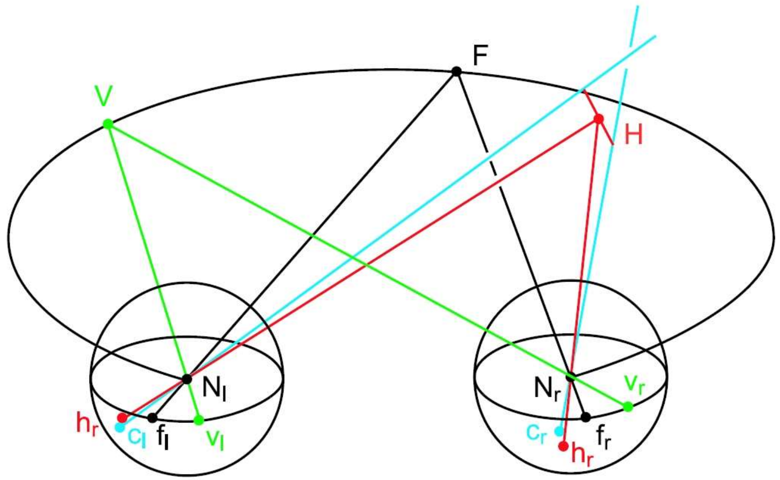
- This restriction imposes a range of depths on objects in the scene

$$\mathrm{Z} = \frac{\mathrm{f\ B}}{d} \quad d = x' - x$$

- Limits a segment of the epipolar line for matching search
- Allows you to reject correspondences whose disparity is outside that range
- Human stereoscopic vision also restricts maximum disparity (Panum's fusional limit)

Disparity Limit

- Non-aligned axes: $d = (d_1, d_2) = \left(x_f^l - x_f^r, y_f^l - y_f^r\right)$
  - There is difficulty defining it as a scalar. It is also not inversely proportional to the distance with the camera. Increases with distance to convergence point (near and far)
- Options:
  - Rectifying images: aligned axes
  - Relative distance along the epipolar line (difference in distance to the epipole)
  - Difference between the angle of the point of view
- Horopter: regions of the same disparity (same angle)



Restrictions on Correspondence

- Geometric restrictions imposed by the imaging system
- Geometric constraints from the viewed objects :
  - Order
  - Disparity gradient limit
  - Uniqueness
  - Surface continuity
  - Figural continuity
  - General position
- Photometric restrictions
- Local primitive restrictions

Positional Order

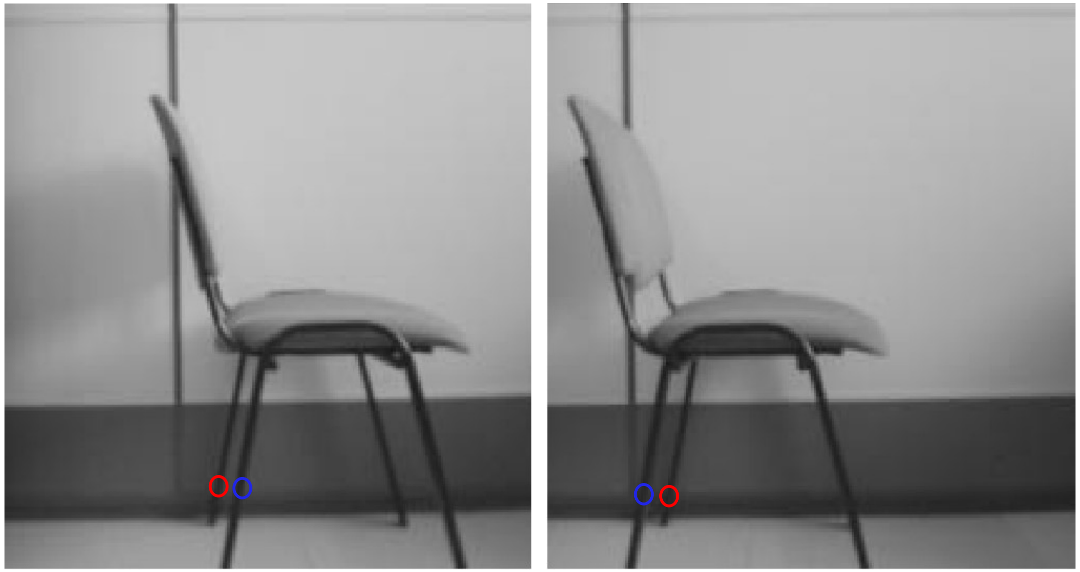
- Given two points on the epipolar line E1 their corresponding points should be located on the epipolar line E2 in the same order.

Left image

Right image

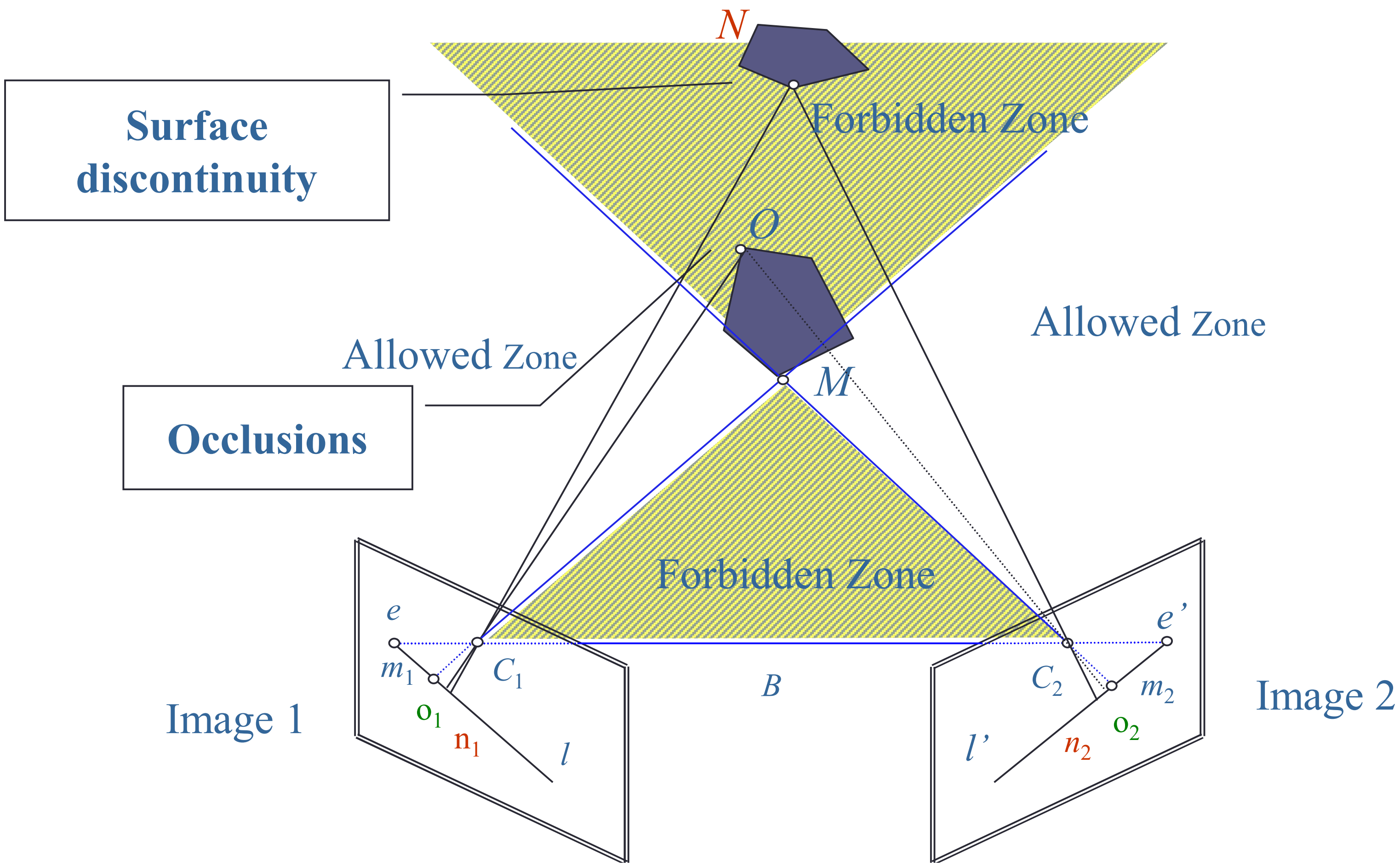
Positional Order

- When the scene has narrow objects located at clearly different distances, this condition may be violated. (No surface continuity)



Positional Order

- Geometric representation of the forbidden zone
  - Occlusions and objects at different distances



Disparity Gradient Limit

Disparity Gradient

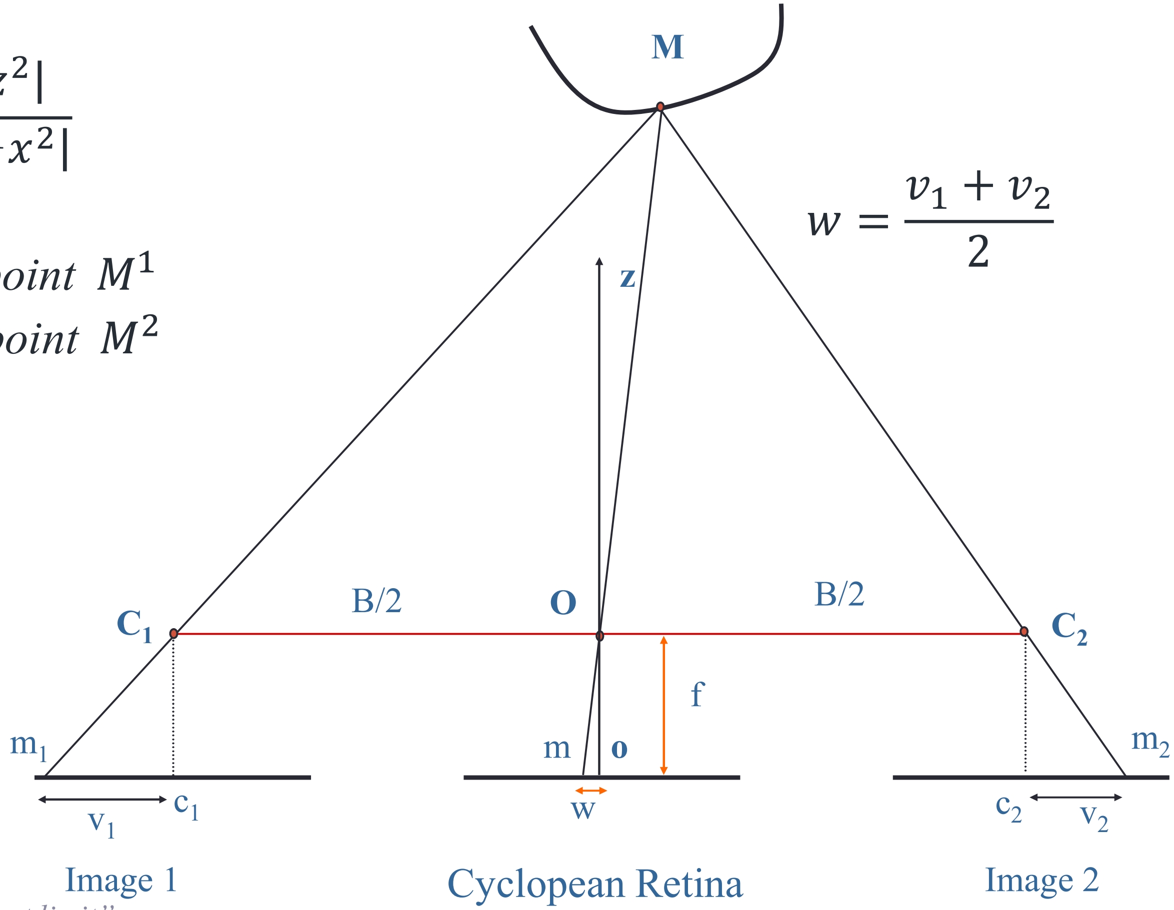
$$GD = \left|\frac{d^1 - d^2}{w^1 - w^2}\right| = \frac{B\ |z^1 - z^2|}{|z^2x^1 - z^1x^2|}$$

$$d^1 = {v^1}_2 - {v^1}_1 \quad \rightarrow \textit{disparity of point } M^1$$
$$d^2 = {v^2}_2 - {v^2}_1 \quad \rightarrow \textit{disparity of point } M^2$$

$$d = B\,\frac{f}{Z} \qquad x = \frac{B}{2d}(v_1 + v_2)$$

$$\boxed{GD < K}$$

Cyclopean Retina



"PMF: a stereo correspondence algorithm using a disparity gradient limit"
Pollard SB, Mayhew JE, Frisby JP Perception. 14(4) (1985)

Disparity Gradient Limit

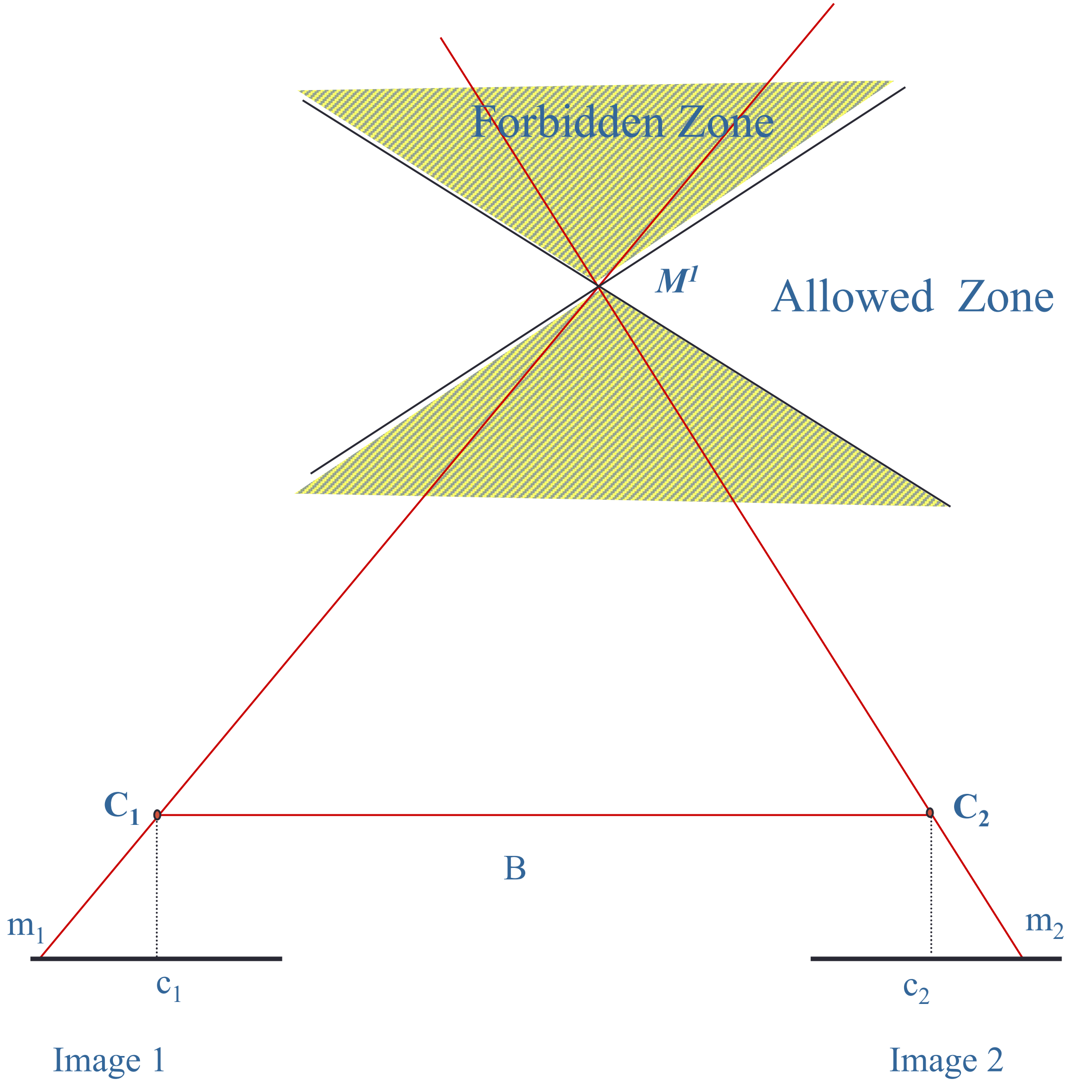
$$\boxed{GD < K}$$

$$M^1 = [x^1, z^1]^T$$

$$M^2 = [x^2, z^2]^T$$

$$z^1 - z^2 = \pm \frac{K}{B} (z^2 x^1 - z^1 x^2)$$

$$K \begin{cases} < 2 \rightarrow forbidden\ zone > Order\ Rest. \\ > 2 \rightarrow forbidden\ zone < Order\ Rest. \end{cases}$$



Disparity Gradient Limit

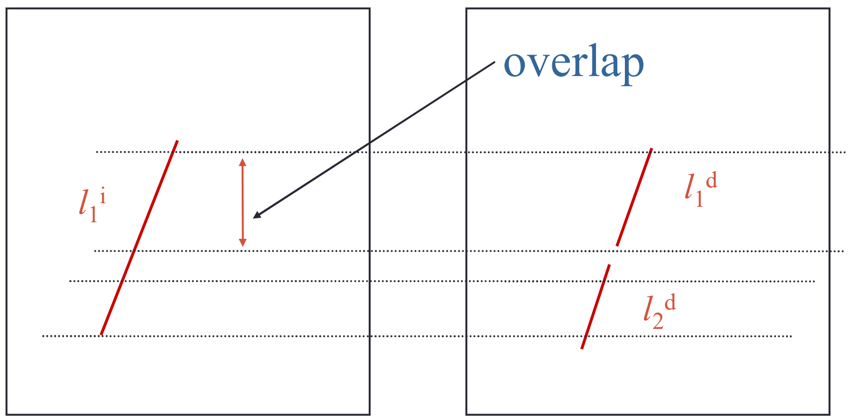
- Geometric interpretation:
  - This restriction imposes a limit on the tangents of the object's surfaces.
  - Variations should be small enough for surface points to lay outside the prohibited zone.
  - It is therefore a restriction on the shape of objects that can be reconstructed by a stereoscopic system.
  - Indirectly it also imposes the restriction of uniqueness.

Uniqueness

- For every point in the image 1 there is only one corresponding point in the image 2
  - The uniqueness condition is met as long as there are no semi-transparent objects in the scene (multiple points on the same projection line generate the same image point on the sensor).
  - The application is carried out by selecting the best of the possible correspondences obtained from the imposition of the other restrictions.
  - It is associated in some algorithms with the Relaxation process (iterative mechanism with inhibitory and excitatory components)

Uniqueness

- This restriction can be extended to higher-level primitives such as contours o edge segments. In this case the segmentation stage can produce different groupings in both images.
  - The uniqueness restriction must be reformulated in such a way that it incorporates not only the selected primitive, but also the three-dimensional geometric relationship with its neighbors.
  - It can be expressed based on admissible overlap.

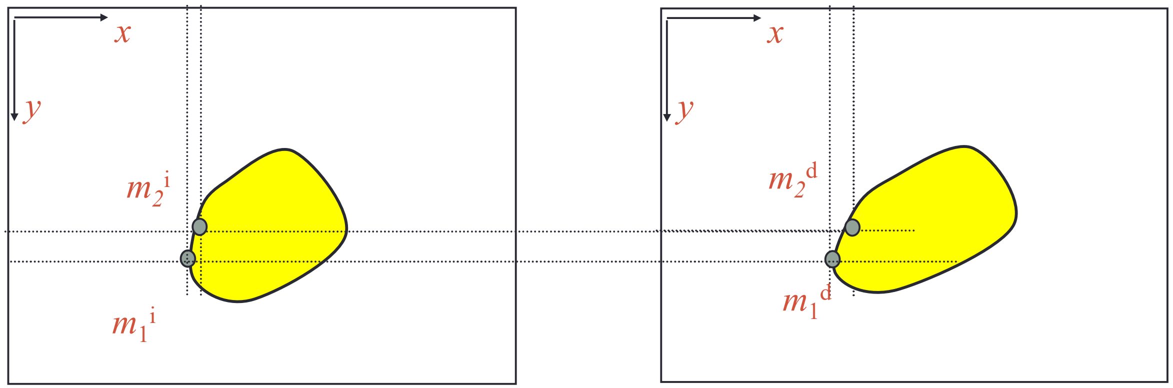


Surface Continuity

- The points projected onto the image belong to the surfaces of the objects in the scene. These surfaces are assumed to be continuous with discontinuities only in the boundaries of the objects (principle of matter cohesion).
  - The surface continuity results in a continuity in the depth map.
  - As we saw above depth and disparity are closely related, so this restriction is imposed as a continuity of disparities at the scene
  - In some algorithms it is linked to the relaxation process.

Figural Continuity

- Proposed by Mayhew y Frisby, establishes the continuity of the surfaces formulated as the continuity of the disparity along the contours of the objects (figures), and not through them.
- This formulation avoids problems arising from discontinuity at surface boundaries.



$$d_1 = x_1{}^i - x_1{}^d$$
$$d_2 = x_2{}^i - x_2{}^d$$

$$m_1, m_2 \in C_{ob} \quad |\overline{m}_1 - \overline{m}_2| < \Delta_a \Rightarrow |d_1 - d_2| < \Delta_b$$

"PMF: a stereo correspondence algorithm using a disparity gradient limit"
Pollard SB, Mayhew JE, Frisby JP Perception 14(4) (1985)

Restrictions on Correspondence

- Geometric restrictions imposed by the imaging system
- Geometric constraints from the viewed objects
- Photometric restrictions:
  - Surface reflectance restriction
  - Photometric compatibility restriction
  - Differential photometric compatibility restriction
- Local primitive restrictions

Photometric Restrictions

- These are restrictions based on models of interaction between objects and the illumination
  - Surface reflectance restriction:
    - Assumes a Lambertian reflectance model for surfaces. The intensity of the projection of a 3D point does not depend on the point of view
  - Photometric compatibility restriction:
    - The distribution of intensities between peer points should be similar

$$\left| I^l({x^l}_1, {y^l}_1) - I^r({x^r}_1, {y^r}_1) \right| < \Delta_a$$

  - Differential photometric compatibility restriction:
    - Given two points in an image nearby (surface continuity), the difference in intensities between the two points should be similar to the difference in intensities of their peers.

$$\left| (I^l({x^l}_1, {y^l}_1) - I^l({x^l}_2, {y^l}_2)) - (I^r({x^r}_1, {y^r}_1) - I^r({x^r}_2, {y^r}_2)) \right| < \Delta_b$$

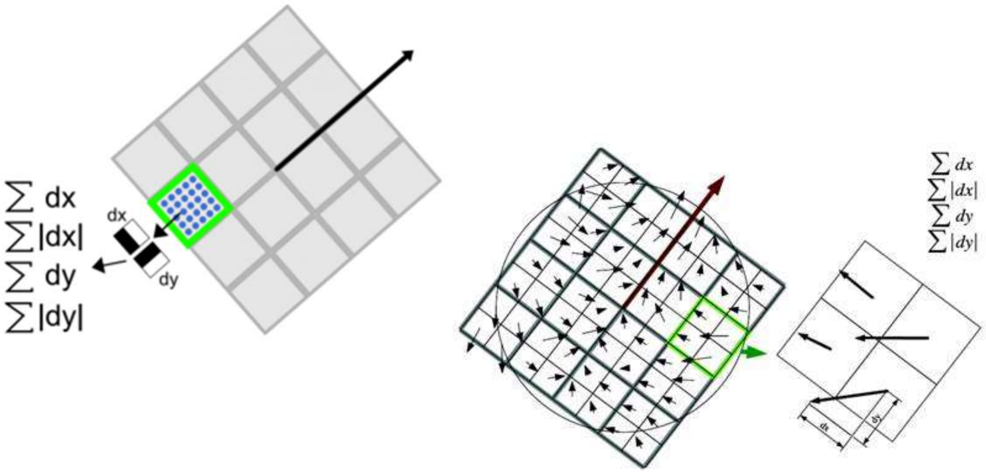
- The above restrictions should apply to the point neighborhood (regions) because intensity values in a pixel are subject to noise.

Restrictions on Correspondence

- Geometric restrictions imposed by the imaging system
- Geometric constraints from the viewed objects
- Photometric restrictions
- Local primitive restrictions

Local Primitive Restrictions

- Compatibility between geometric characteristics of pairs of possible correspondences
  - Edge points:
    - Gradient orientation similarity (Sign)
    - Compatibility between different channels
  - Featured points:
    - Similarity in the HOG-based descriptor (SIFT Descriptor)
    - Similarity in the wavelet-based Haar descriptor (SURF Descriptor)



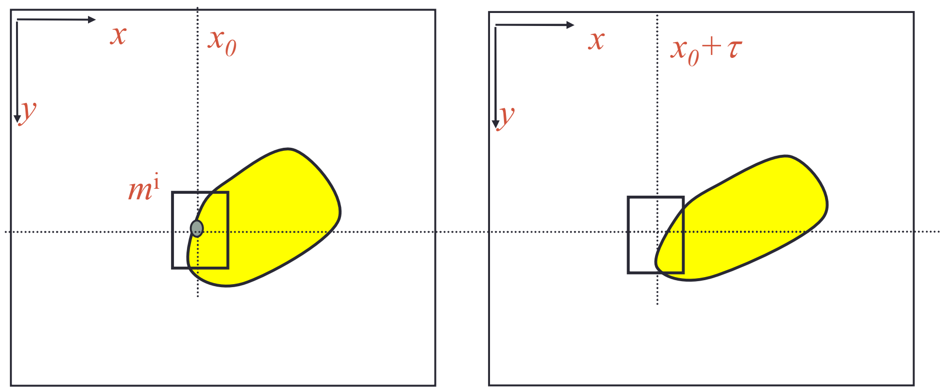
"SURF: Speeded Up Robust Features" Herbert Bay, T. Tuytelaars, L. Van Gool European Conference on Computer Vision, 2006

"Distinctive Image Features from Scale-Invariant Keypoints" David G. Lowe International Journal of Computer Vision, 60, 2 (2004)

Table of Contents

Area Correlation

- These methods consider the two images as a two-dimensional signal translated.
- They try to obtain, for each point of the image, such translation by minimizing or maximizing a certain criteria (Correlation)
- For each pixel of an image, the correlation between the intensity distribution of a window centered on that pixel and a window of the same size centered on the pixel to be analyzed from the other image is calculated.



Area Correlation

- Restrictions applied:
  - Lambertian restriction:
    - Surfaces must be Lambertian. The projected intensity of a 3D point does not depend on the point of view
  - Front-parallel restriction:
    - They assume that the disparity is constant locally (correlation window). Surfaces must be parallel to camera image planes (small slope)
  - Continuity restriction:
    - Surfaces are, at least locally, continuous.

Area Correlation

- Correlation criteria (maximizes the functional):
  - Sum of Square Difference:

$$SSD(m^l(x,y), m^r(x',y')) = -\frac{1}{N} \sum_{\substack{\forall i,j \perp \\ m(x+i,y+j) \in W(m)}} \left(I^l(m^l(x+i, y+j)) - I^r(m^r(x'+i, y'+j))\right)^2$$

$$\Updownarrow$$

$$SSD(m^l, m^r) = -\frac{1}{N} \sum_{\forall m_i \in W(m)} \left(I^l({m^l}_i) - I^r({m^r}_i)\right)^2$$

It's very sensitive to lighting differences

  - Average centered Sum of Square Differences:

$$ZSSD(m^l, m^r) = -\frac{1}{N} \sum_{\forall m_i \in W(m)} \left(\left(I^l({m^l}_i) - \overline{I^l(m^l)}\right) - \left(I^r({m^r}_i) - \overline{I^r(m^r)}\right)\right)^2$$

  - Average centered Normalized Sum of Square Differences

$$ZNSSD(m^l, m^r) = -\frac{\sum_{\forall m_i \in W(m)} \left(\left(I^l({m^l}_i) - \overline{I^l(m^l)}\right) - \left(I^r({m^r}_i) - \overline{I^r(m^r)}\right)\right)^2}{N\, \sigma^l(m^l)\, \sigma^r(m^r)}$$

Consider variations in the distribution of intensities within the window

Area Correlation

- Cross-correlation:

$$CC(m^l, m^r) = \frac{1}{N} \sum_{\forall m_i \in W(m)} I^l({m^l}_i) \cdot I^r({m^r}_i)$$

  - Promotes areas of high intensity

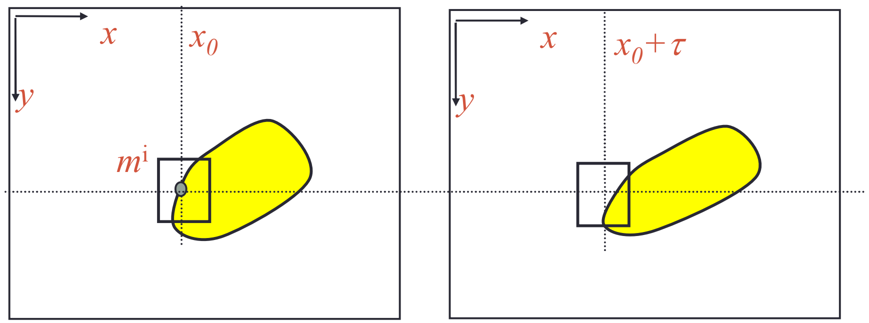
- Average centered Normalized Cross-Correlation:

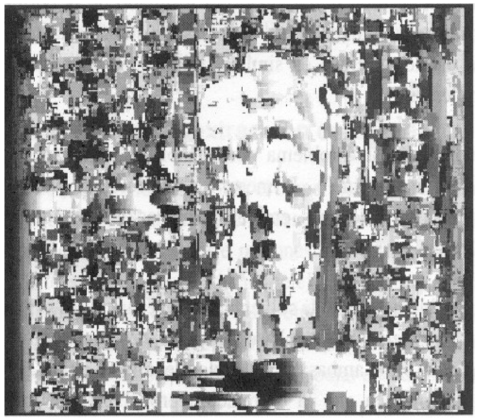
$$CC(m^l, m^r) = \frac{\sum_{\forall m_i \in W(m)} \left(I^l({m^l}_i) - \overline{I^l(m^l)}\right) \cdot \left(I^r({m^r}_i) - \overline{I^r(m^r)}\right)}{N\, \sigma^l(m^l)\, \sigma^r(m^r)}$$

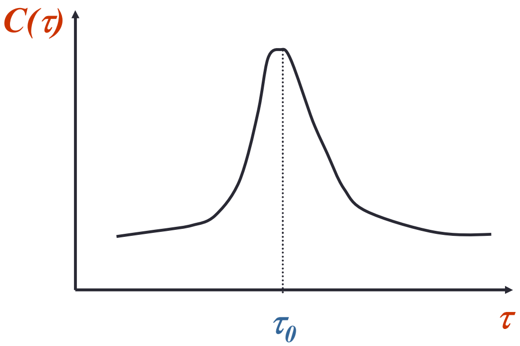
- Improvements:
  - Using adaptive windows

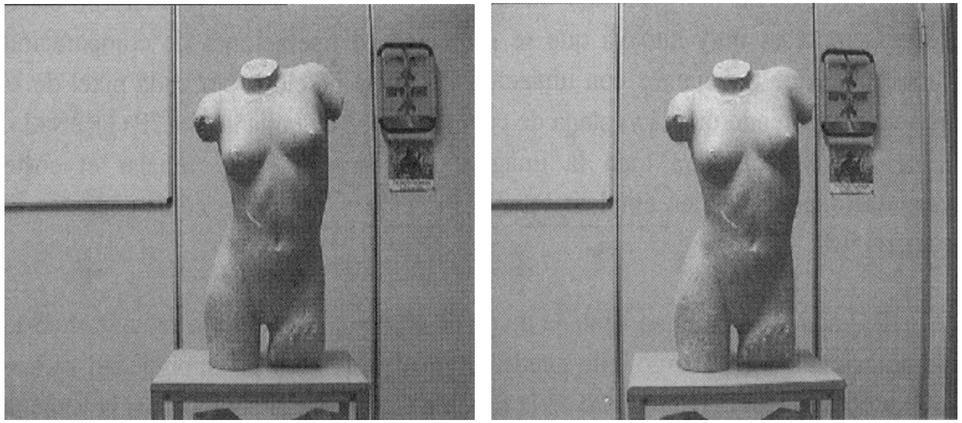
Area Correlation

- Example:









Area Correlation

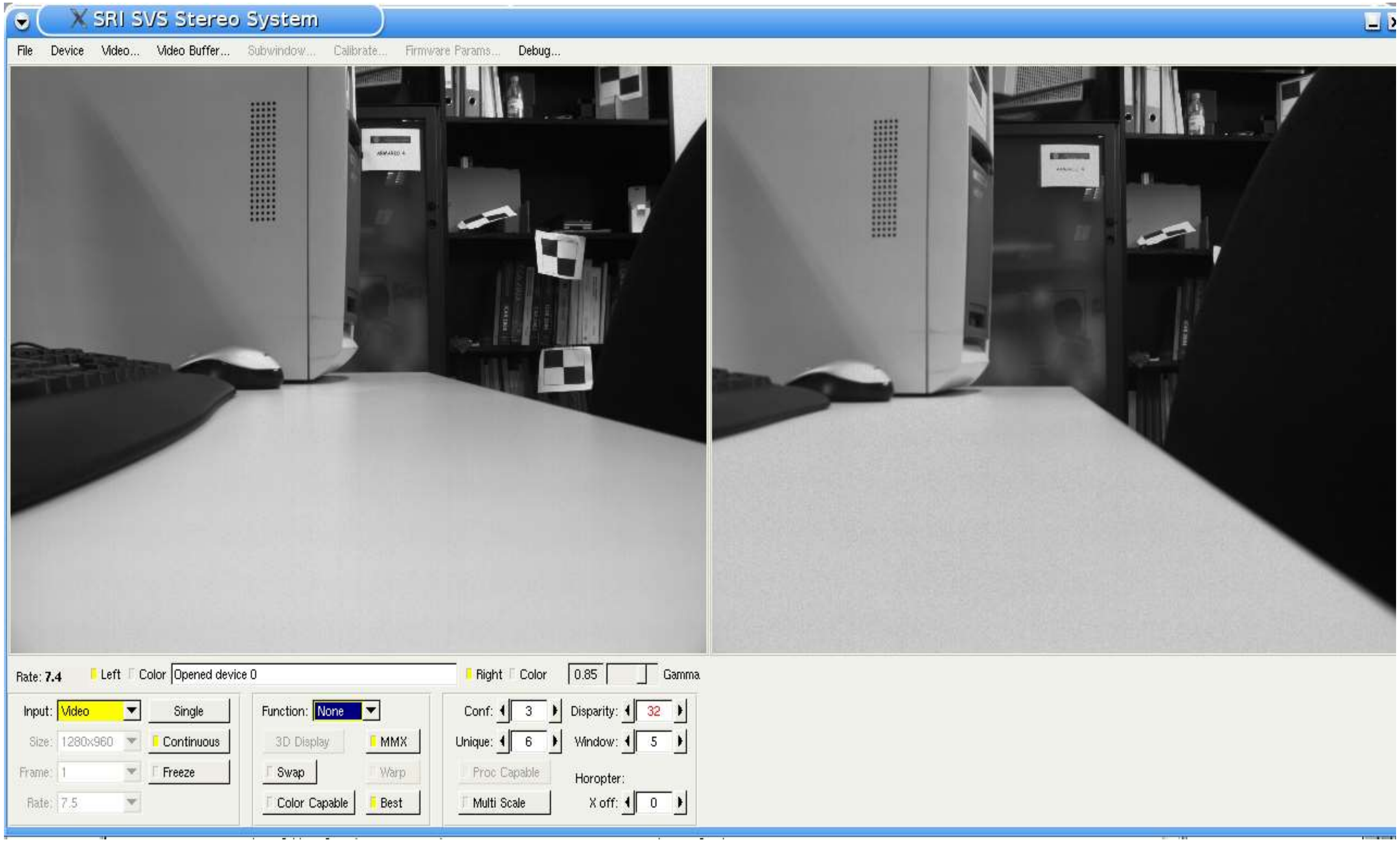
- Advantages
  - Give good results on images with important texture
  - Create dense maps of disparity
  - They're easy to parallelize
- Disadvantages
  - They have problems with images with high surface discontinuities
  - They are very sensitive to photomeric variations due to shadows
  - They require a subsequent process of filering false correspondences
  - Problems with occlusions

Area Correlation

- Choosing a window size:
  - Small window (5x5):
    - Computational cost is reduced, but can cause ambiguity issues.
    - Limit case: 1 pixel window size. it only considers the grey value of the pixel.
  - Large window (20x20):
    - Higher computational cost (more operations required for each comparison).
    - Pixels of different disparity are likely to exist within the same window, violating the fundamental assumptions of this technique.

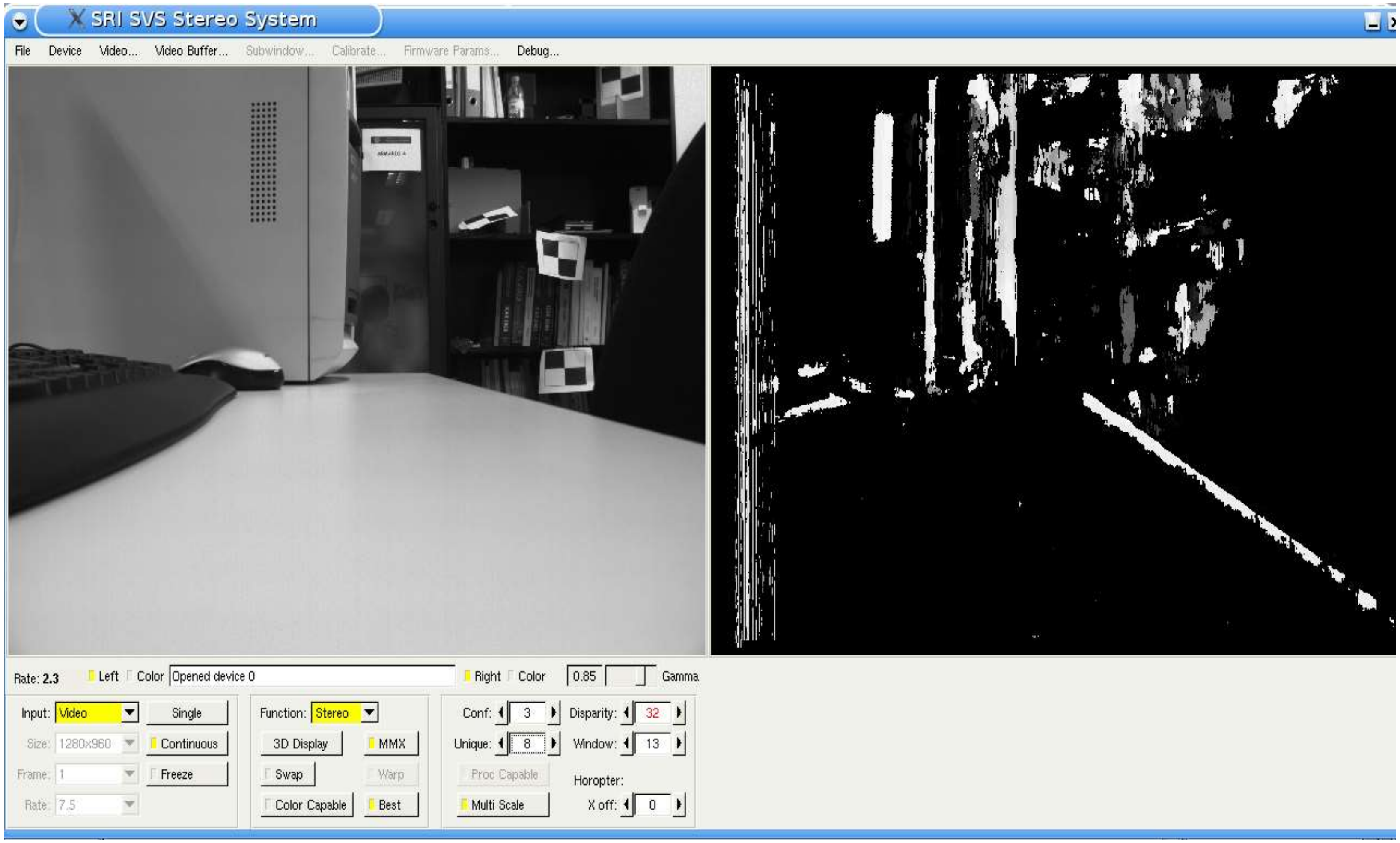
- One strategy is to correlate using different window sizes. The information obtained on a rude scale is used to guide and limit the search for correspondences on a finer scale.

Area Correlation : Typical Results



SRI SVS Stereo System

Area Correlation : Typical Results



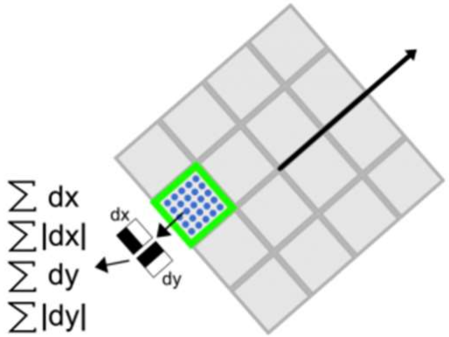
SRI SVS Stereo System

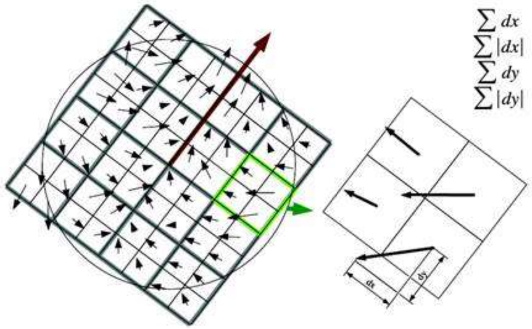
Feature-Based Techniques

- Feature-based techniques establish correspondence between primitives previously extracted from images.
- Detect high level features
    - Edge points
    - Edge segments
    - Regions
    - Corners: Harris
    - High gradient points: SIFT, SURF
- These elements are more stable in the face of changes in:
    - Illumination.
    - Point of view.
- Advantages
    - They allow comparisons between feature properties.
    - They do not have a strong dependence on epipolar restriction.
    - They focus on places in the scene where the information content is maximum (areas with more texture).
    - They're more robust.
- Drawbacks
    - Provide disperse information
    - Measurement accuracy depends on the type of feature chosen (e.g. point, edge).

Feature-Based Techniques

- Local Primitive Restrictions:
  - Properties of the primitives:
    - Edge points:
      - Gradient orientation similarity (sign)
      - Compatibility between different channels
    - Featured points:
      - Similarity in HOG-based Descriptor (SIFT Descriptor)
      - Similarity in Wavelet-Based Haar Descriptor (SURF Descriptor)





"SURF: Speeded Up Robust Features" Herbert Bay, T. Tuytelaars, L. Van Gool European Conference on Computer Vision, 2006

"Distinctive Image Features from Scale-Invariant Keypoints" David G. Lowe International Journal of Computer Vision, 60, 2 (2004)

Feature-Based Techniques

- The distance between two feature points is calculated from the value of their descriptor vectors.
- Metric Descriptors: (SIFT, SURF)
  - <u>Euclidea Distance</u>:
    - It needs that the descriptor vector is normalized
    - Otherwise, attributes with higher absolute value always prevail over others

$$d_e(\mathrm{x}^{(i)}, \mathrm{y}^{(j)}) = \sqrt{\left(x_1^{(i)} - y_1^{(j)}\right)^2 + \cdots + \left(x_n^{(i)} - y_n^{(j)}\right)^2}$$

$$\mathrm{x}^{(i)} = \left[x_1^{(i)}, \cdots, x_n^{(i)}\right]^T$$

$$\mathrm{y}^{(j)} = \left[y_1^{(j)}, \cdots, y_n^{(j)}\right]^T$$

Feature-Based Techniques

- The distance between two feature points is calculated from the value of their descriptor vectors.
- Metric Descriptors: (SIFT, SURF)
  - Mahalanobis Distance: normalized distances according to variance

$$d_m(\mathrm{x}^{(i)},\mathrm{y}^{(j)}) = \sqrt{\left(\hat{x}_1^{(i)} - \hat{y}_1^{(j)}\right)^2 + \cdots + \left(\hat{x}_n^{(i)} - \hat{y}_n^{(j)}\right)^2}$$

$$\hat{x}_k^{(i)} = \frac{x_k^{(i)} - \mu_k}{\sigma_k}$$

$$d_m(\mathrm{x}^{(i)},\mathrm{y}^{(j)}) = \sqrt{\left(\frac{x_1^{(i)} - y_1^{(j)}}{\sigma_1}\right)^2 + \cdots + \left(\frac{x_n^{(i)} - y_n^{(j)}}{\sigma_n}\right)^2}$$

Same mean and variance for both images

  - Cosine Distance: angle between the descriptor vectors

$$d_{cos}\left(\mathrm{x}^{(i)},\mathrm{y}^{(j)}\right) = 1 - \cos\left(\mathrm{x}^{(i)},\mathrm{y}^{(j)}\right) = 1 - \frac{\mathrm{x}^{(i)^T} \cdot \mathrm{y}^{(j)}}{|\mathrm{x}^{(i)}| \cdot |\mathrm{y}^{(j)}|}$$

Feature-Based Techniques

- The distance between two feature points is calculated from the value of their descriptor vectors.
- Binary Descriptors: (ORB, BRIEF)
  - Hamming Distance:
    - Number of different binary digits

$$d_{ham}(\mathbf{x}^{(i)}, \mathbf{y}^{(j)}) = \mathcal{W}(\mathbf{x}^{(i)} \oplus \mathbf{y}^{(j)})$$

$\mathbf{x}^{(i)} = [0,0,1,0,1, \ldots, 1]^T$

$\mathbf{y}^{(j)} = [1,0,0,0,1, \ldots, 1]^T$

Feature-Based Techniques

- Search for Correspondence:
  - Brute Force: lineal
    - Calculates the distance to all the features in other image
    - Distances: Euclidean, Mahalanobis, Cosine, Hamming (binary descriptors)
    - Type of search
      - **NN**: returns the nearest neighbor
      - **K-NN**: returns the ***K*** nearest neighbours (allows other restrictions to be imposed)
      - **FR-NN**: returns the neighbors at a distance ***R***
  - FLANN: "*Fast Approximate Nearest Neighbor Search*"
    - Accelerates search in high-size spaces
    - Requires a metric description vector (non-binary)
    - Use pseudo-optimal search techniques by indexing the search space using a tree (K-d-tree, K-means-tree)
      - *K-d tree*: multiple random trees are generated and a side-by-side search is carried out
      - *K-means tree*: a hierarchical tree is generated by grouping the data into clusters in each branch

"FLANN : Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration" Marius Muja, David G. Lowe
VISAPP International Conference on Computer Vision Theory and Applications , pp. 331-340 (2009)

Técnicas Basadas en Características

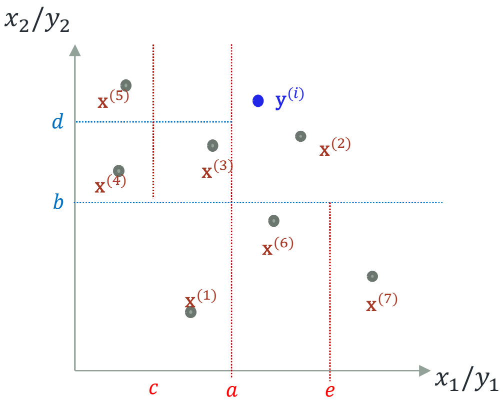
- FLANN: "*Fast Approximate Nearest Neighbor Search*"

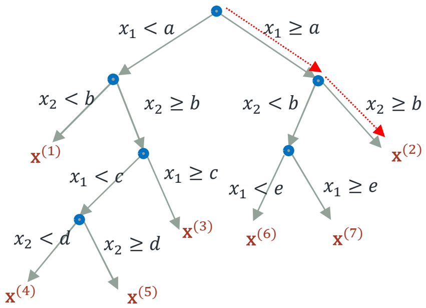
  - *K-d tree*:

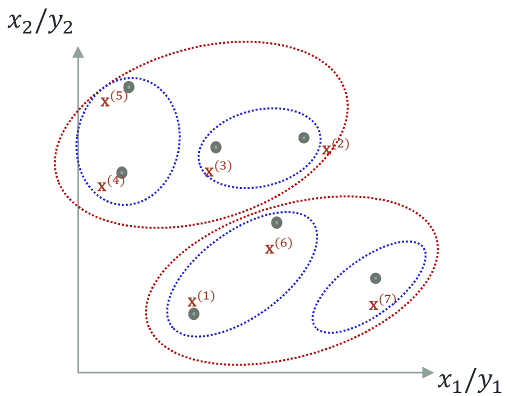
$$\mathbf{x}^{(i)} = \left[x_1^{(i)}, \cdots, x_n^{(i)}\right]^T$$

$$\mathbf{y}^{(j)} = \left[y_1^{(j)}, \cdots, y_n^{(j)}\right]^T$$

K-means tree:



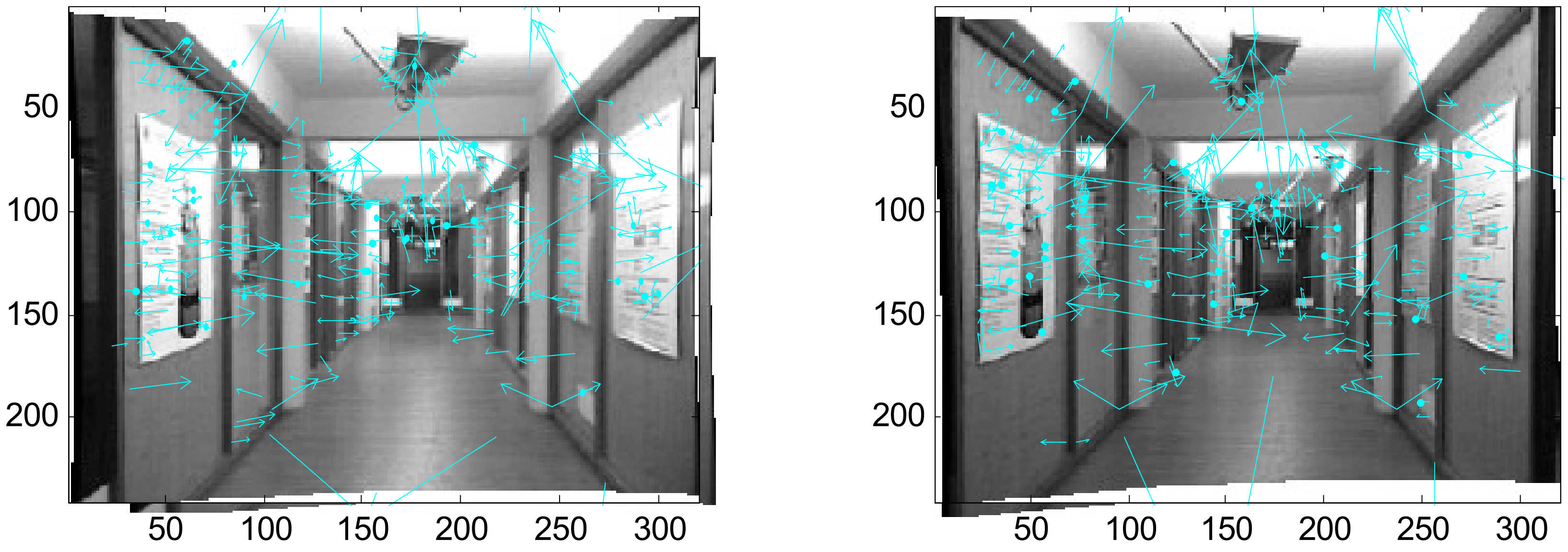




"FLANN : Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration" Marius Muja, David G. Lowe VISAPP International Conference on Computer Vision Theory and Applications , pp. 331-340 (2009)

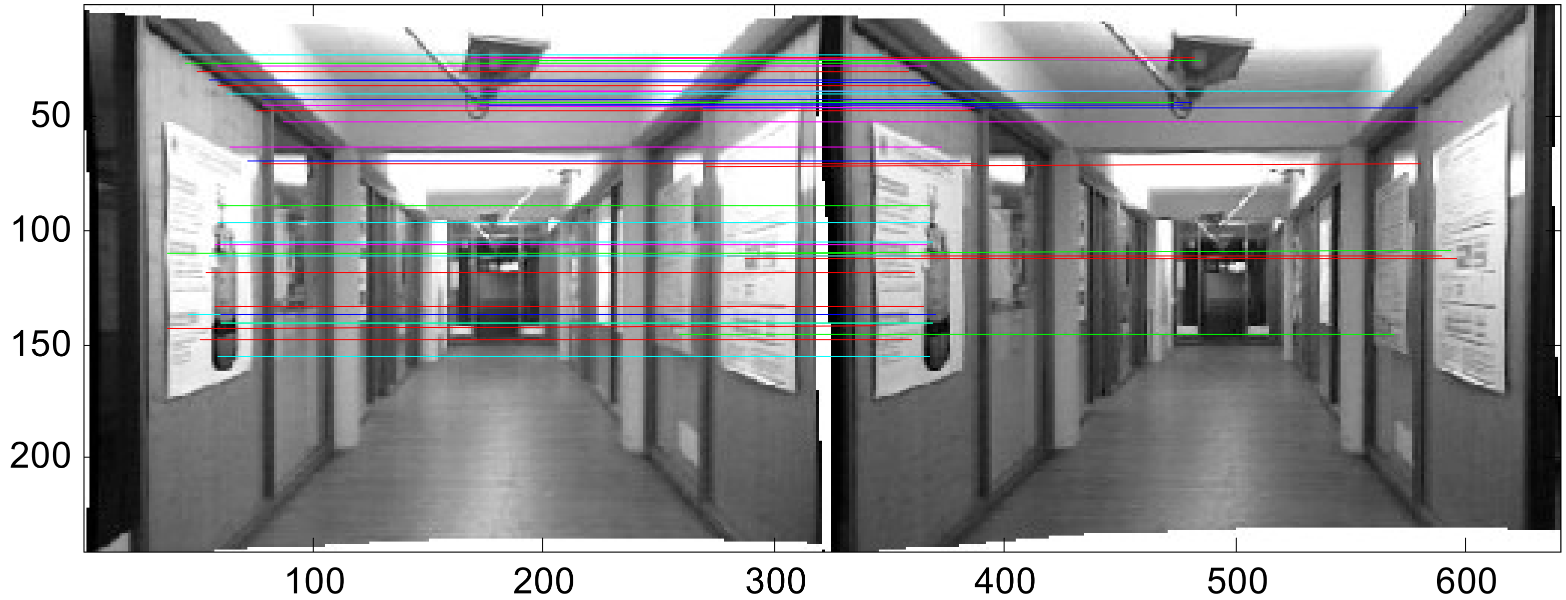
Example: Disperse Stereo Reconstruction

- Example of use of point characteristics
  - SIFT point extraction
  - Correspondence of those points ←→ Recognition



Example: Disperse Stereo Reconstruction

- Correspondences found. Restrictions:
  - Epipolar: +- 1 pixel
  - SIFT descriptor of each point



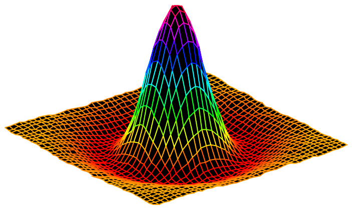
Marr-Poggio Computational Theory

- Marr [*Vision, 82*] presents one of the most important studies on the human visual process by establishing a complete computational theory.
- This theory includes the stereoscopic correspondence process
- Stages:
  1) Laplacian-Gaussian Filtering (different spatial scales), evaluating "**zero-crossings**"
  2) Correspondence between "**zero-crossings**" that have the ***same sign*** and about the same orientation.
  3) Correspondences obtained from larger masks control the convergence of the algorithm by eliminating ambiguities in the correspondence of smaller masks.
  4) A process of label ***Relaxation*** imposes the restrictions of ***surface continuity*** and ***uniqueness***
  5) The results of the map are stored in a buffer that calls the 2.5D sketch, representing the steps by zero next to the disparity between the two images.

"Vision: A Computational Investigation into the Human Representation and Processing of Visual Information" David Marr Freeman. (1982)

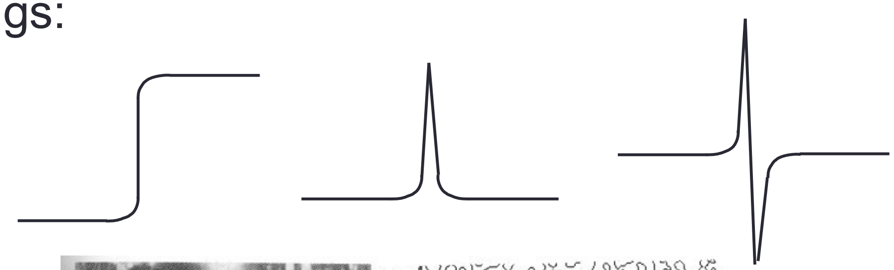
Marr-Poggio Computational Theory

- (1) Laplacian-Gaussian **Filtering** / zero-crossings:
  - Detects intensity discontinuities

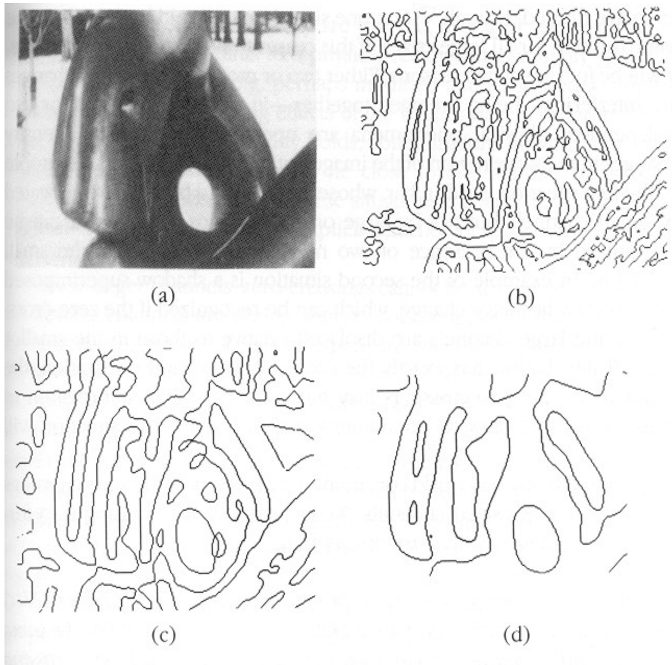


$$\omega_{2D} = 2\sqrt{2}\ \sigma$$

Filter width



$$\nabla^2 G(r,\vartheta) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{r^2}{\sigma^4}\right]\ e^{\left(\frac{-r^2}{2\sigma^2}\right)}$$



Different filters detect discontinuities at different spatial scales

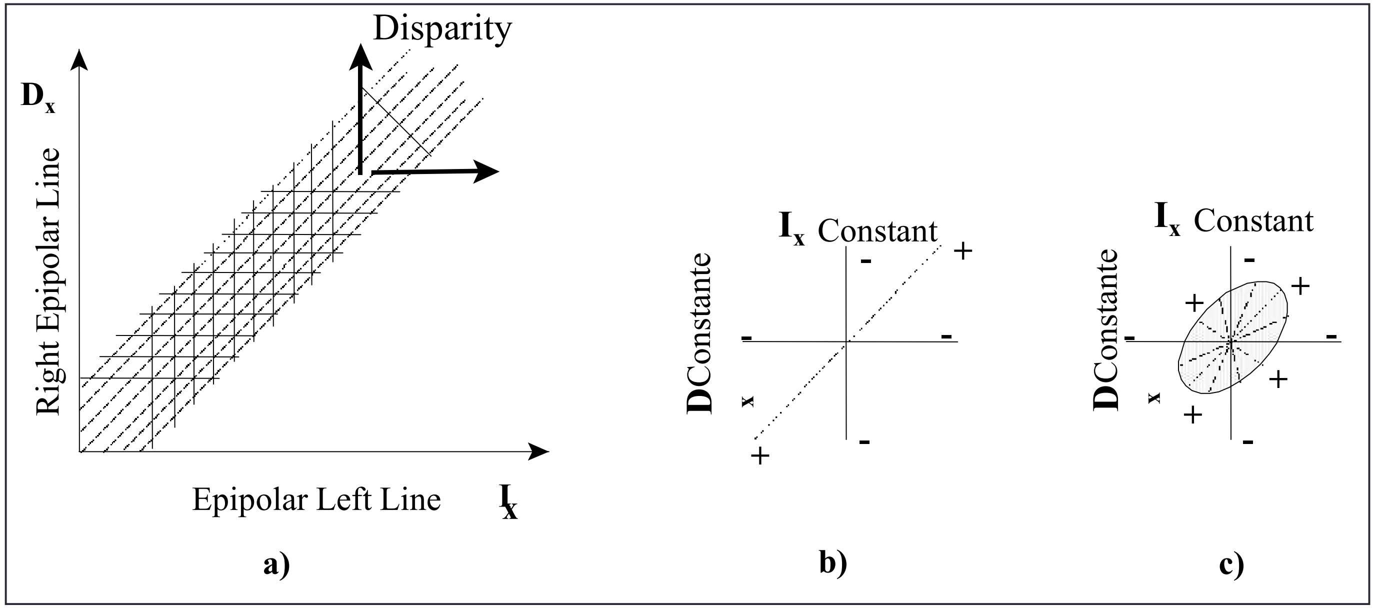
"Vision: A Computational Investigation into the Human Representation and Processing of Visual Information" David Marr Freeman. (1982)

Marr-Poggio Computational Theory

- (2) **Correspondence** between the zero-crossings:
  - For each scale the correspondence is evaluated along the epipolar of zero-crossings that have the same sign and approximately the same orientation.
  - **Límite fusional de Panum**:
    - The probability of finding a correct correspondence with a lower disparity than $\pm(\omega_{2D}/2)$ is the 95%
  - Ambiguities in local correspondence are resolved considering the disparity in the sign of neighboring non-ambiguous correspondences
  - Correspondences obtained from larger scales control algorithm convergence by eliminating ambiguities in the correspondence of smaller scales.

Marr-Poggio Computational Theory

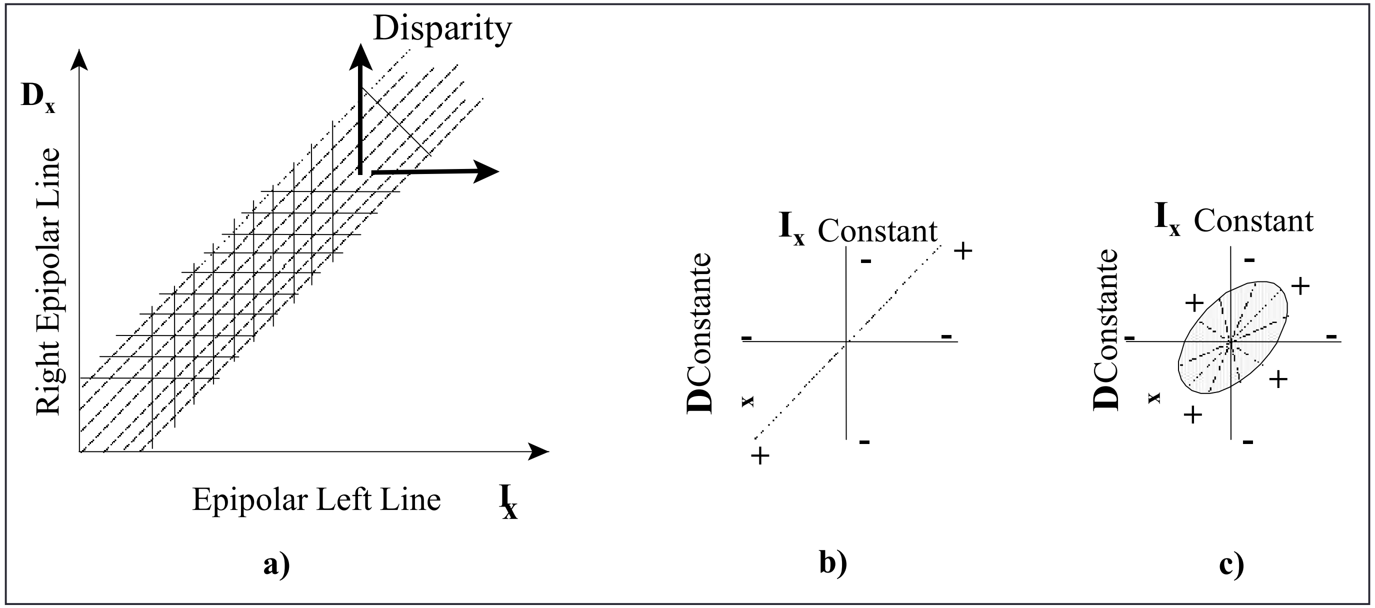
- (3) Relaxation process: Cooperative Algorithm



- **For each epipolar line of a stereo pair, a two-dimensional network of Nodes o Cells Interconnected (Figure a)**

Marr-Poggio Computational Theory

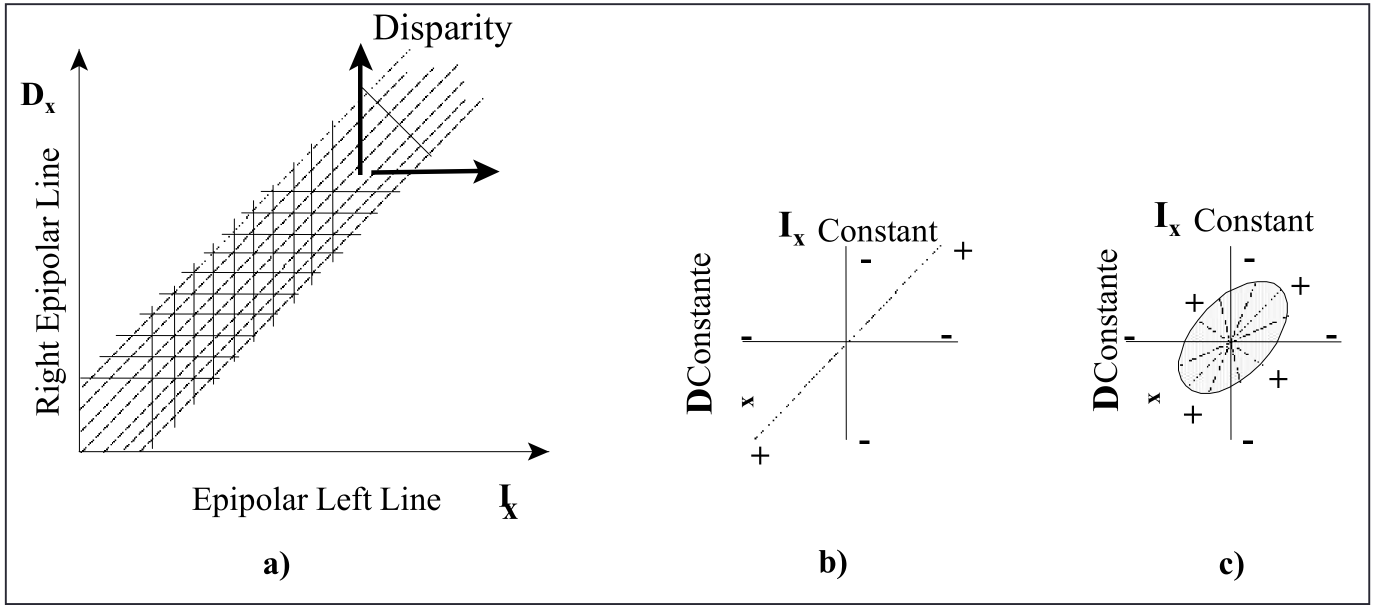
- (3) Relaxation process: Cooperative Algorithm



- **Horizontal and vertical connections establish an Inhibitory effect, so that nodes along a horizontal or vertical line inhibit each other, leaving at the end a single correspondence on each line (Restriction of Uniqueness) (Figure b)**

Marr-Poggio Computational Theory

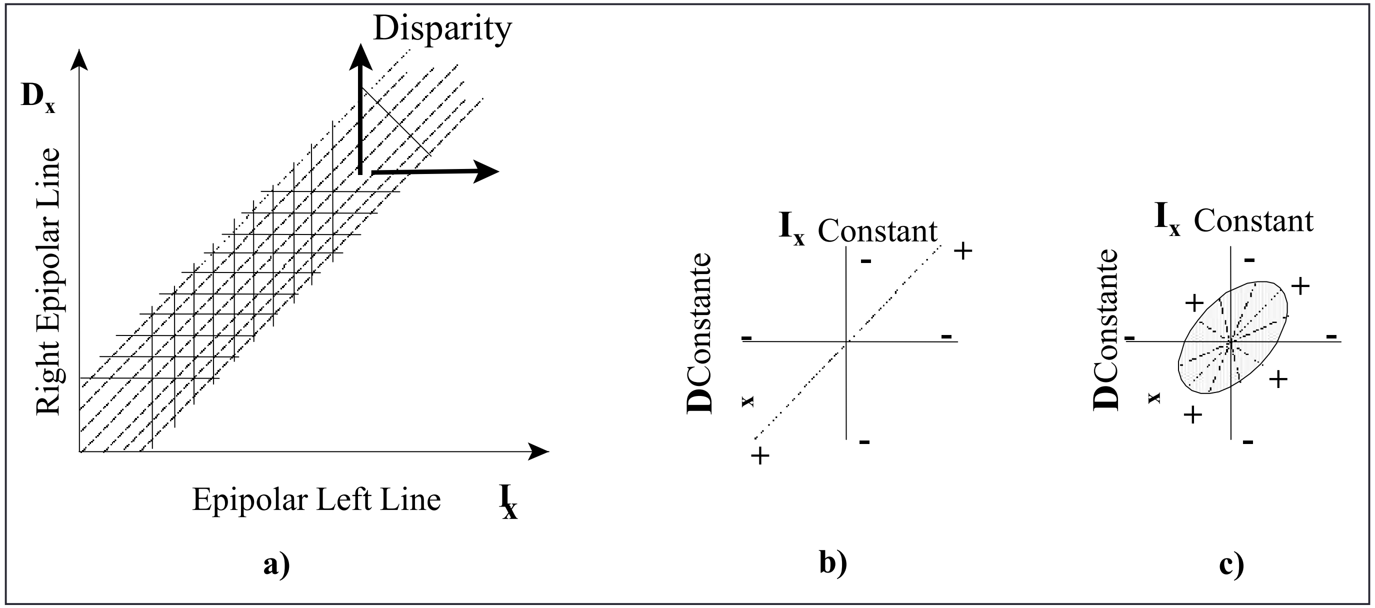
- (3) Relaxation process: Cooperative Algorithm



- **Diagonal connections (excitatory connections) favor the correspondence of adjacent elements on the diagonal with the same disparity (Disparity Continuity) (Figure b), considering therefore in this model that the surfaces are continuous and smooth.**

Marr-Poggio Computational Theory

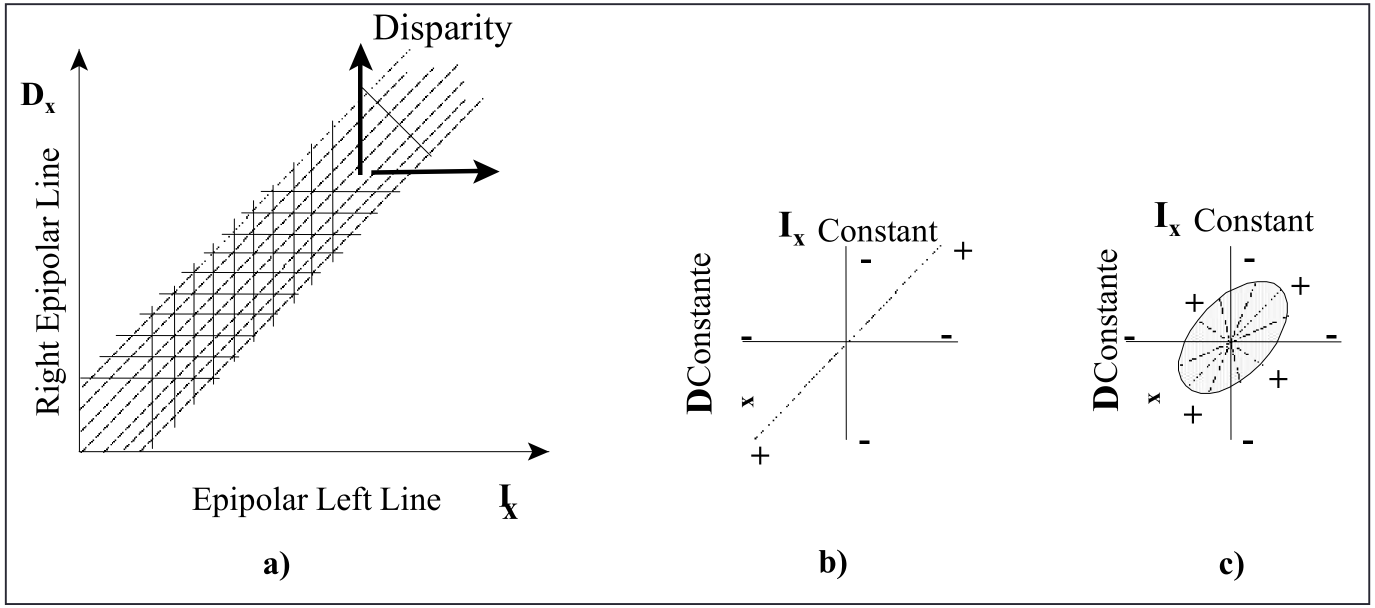
- (3) Relaxation process: Cooperative Algorithm



- **The above restriction, disparity continuity, is set only in the diagonal direction. To consider a two-dimensional constraint it can be used an excitatory disk-shaped neighborhood environment S(x,y,d) (Figure c).**

Marr-Poggio Computational Theory

- (3) Relaxation process: Cooperative Algorithm



$$C^{t+1}_{x,y\,;\,d} = \sigma \left\{ \sum_{x',y'\,;\,d' \in S(x,y,d)} C^{t}_{x',y'\,;\,d'} - \varepsilon \sum_{x',y'\,;\,d' \in O(x,y,d)} C^{t}_{x',y'\,;\,d'} + C^{0}_{x,y\,;\,d} \right\}$$

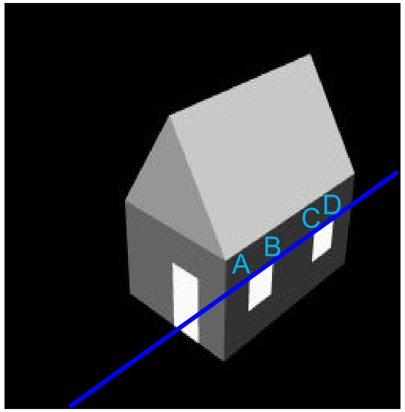
Excitatory Component

Inhibitory Component

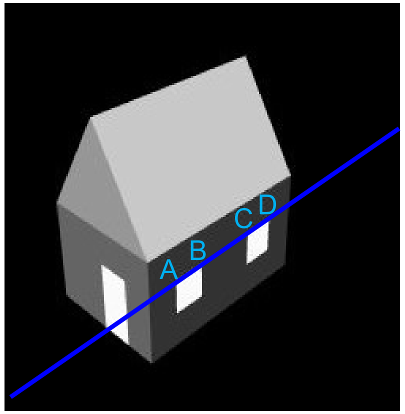
Initial Value

Marr-Poggio Computational Theory

Left Image



Right Iamge



Initial Probability

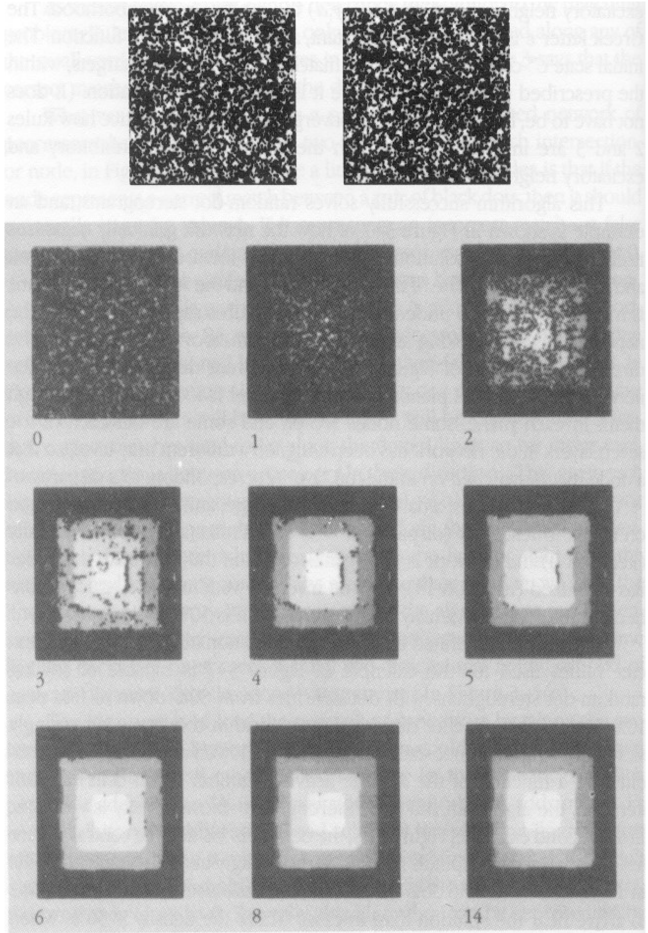
	A^i	B^i	C^i	D^i
D^d	0	1	0	1
C^d	1	0	1	0
B^d	0	1	0	1
A^d	1	0	1	0

Initial Probability

	A^i	B^i	C^i	D^i
D^d	0	0.1	0	0.9
C^d	0.1	0	0.9	0
B^d	0	0.9	0	0.1
A^d	0.9	0	0.1	0

"Vision: A Computational Investigation into the Human Representation and Processing of Visual Information" David Marr Freeman. (1982)

Marr-Poggio Computational Theory



Random point stereograms

Depth maps in different iterations of the relaxation process

"Vision: A Computational Investigation into the Human Representation and Processing of Visual Information" David Marr Freeman. (1982)

Advanced Matching Techniques: CNN

- Siamese Convolutional Networks: GC-Net / PSMNet

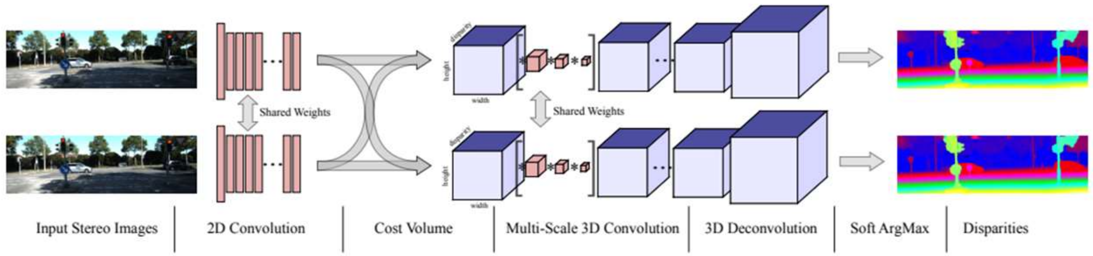


Figure 1: **Our end-to-end deep stereo regression architecture, GC-Net** (Geometry and Context Network).

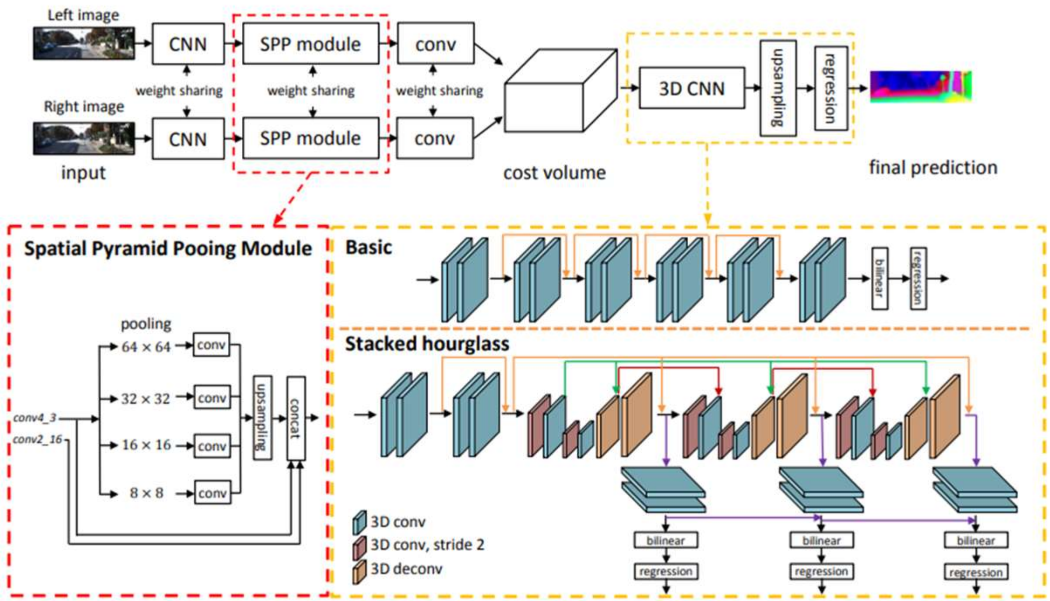


Figure 1. Architecture overview of proposed PSMNet. The left and right input stereo images are fed to two weight-sharing pipelines

"End-to-End Learning of Geometry and Context for Deep Stereo Regression" A. Kendall, H. Martirosyan, S. Dasgupta, P. Henry, R. Kennedy, A. Bachrach, A. Bry, (2017) IEEE International Conference on Computer Vision

"Pyramid Stereo Matching Network" Jia-Ren Chang Yong-Sheng Chen (2018) IEEE/CVF Conference on Computer Vision and Pattern Recognition

Table of Contents