Reconstruction of three dimensional scenes

P. Strumiłło, P. Pełczyński

Problem formulation

Three dimensional automatic scene sensing means capturing shape, appearance and spatial coordinates of real objects.

Shape *- geometry of 3D object*

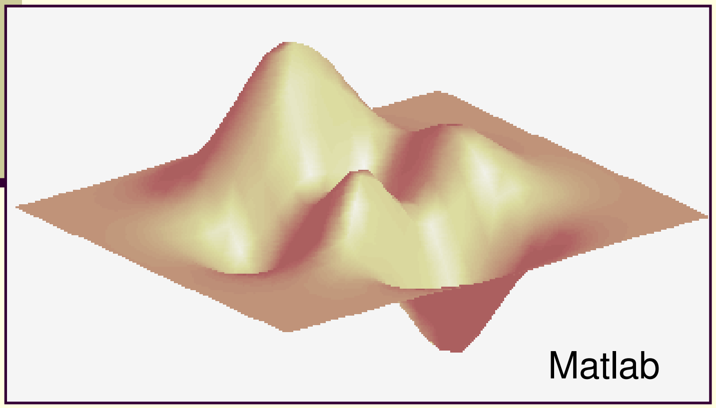
Appearance *- surface attributes: colour, texture, reflectance*

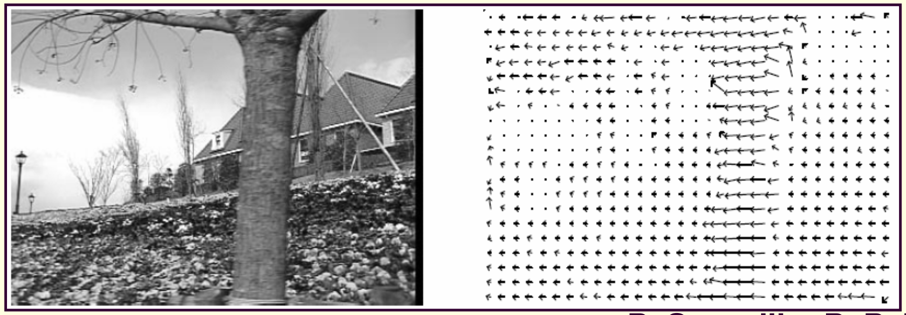
Spatial coordinates *– (X,Y, Z) coordinates in 3D*

There are methods of 3D scene reconstruction from a set of 2D images

3D scene reconstruction methods

- *Stereo image pair matching*
- *Structure from motion camera image sequence*
- *Shape reconstruction from shading*
- *Projection of structured light*
- *Laser scanning*





P. Strumiłło, P. Pełczyński

Applications

- *Robotics*
- *Medicine*
- *Electronic travel aid for the blind*
- *Product design*
- *Virtual reality*
- *Computer games*

Robotics - robot vision

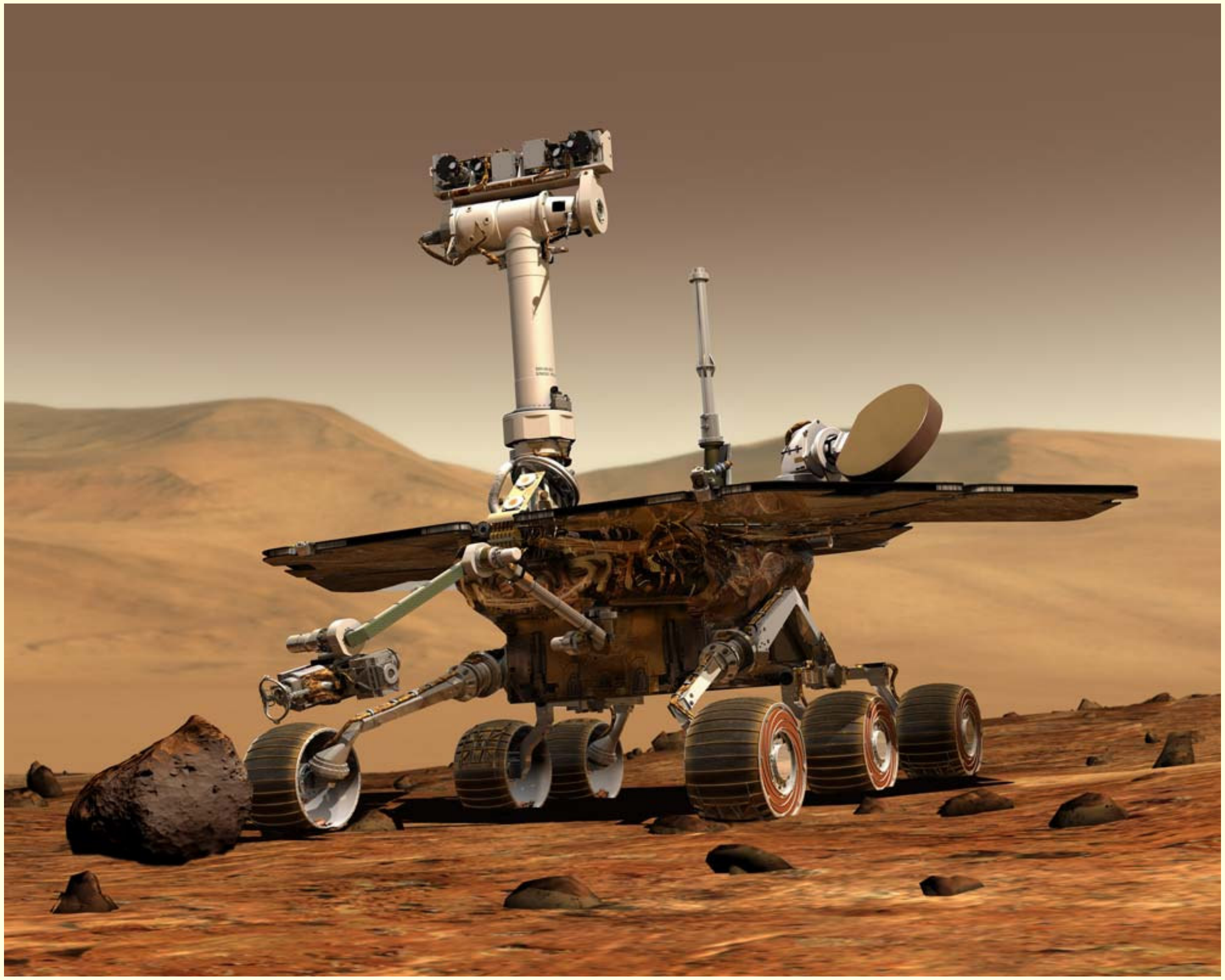
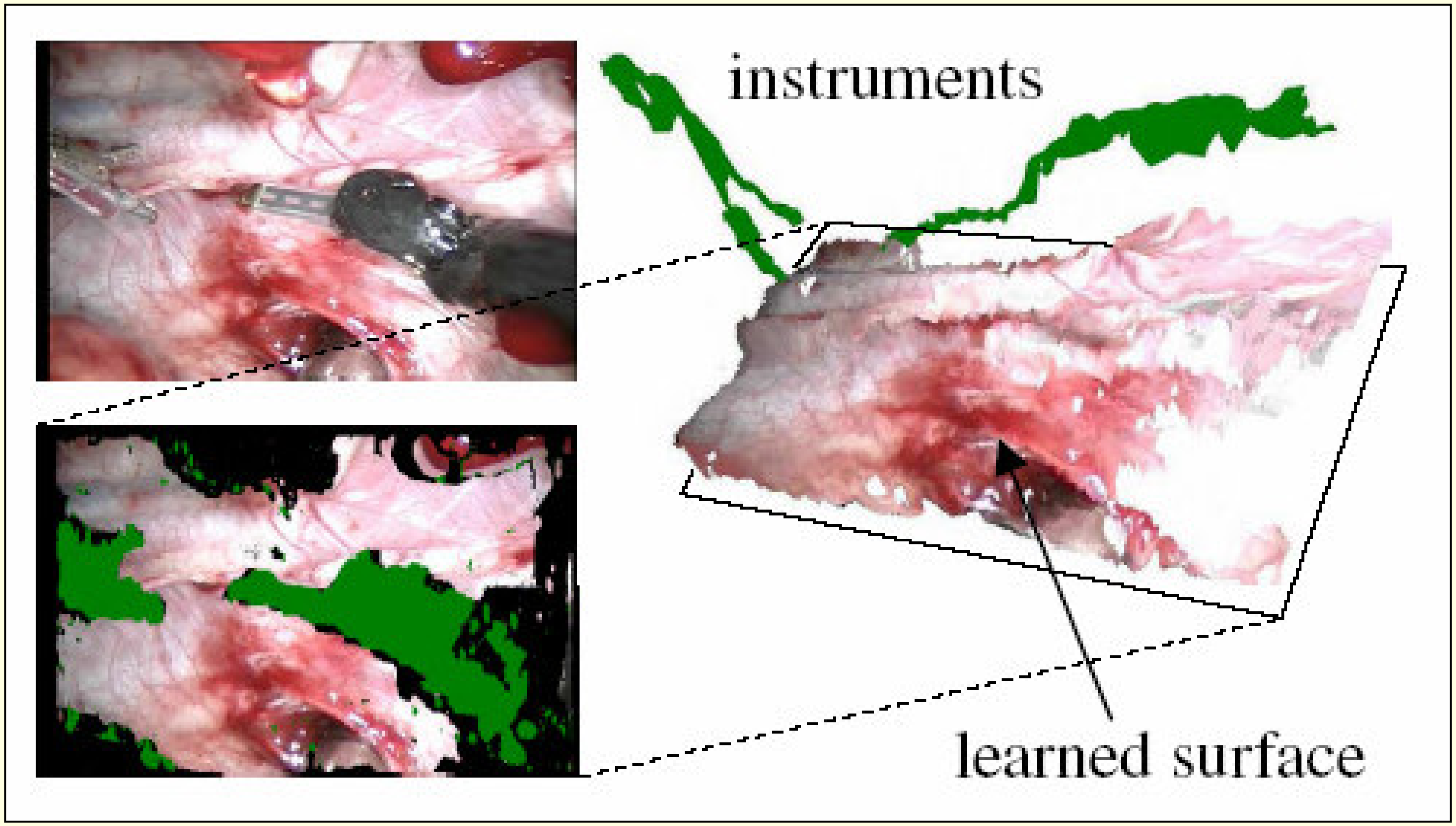


Image copied from : http://marsrovers.jpl.nasa.gov

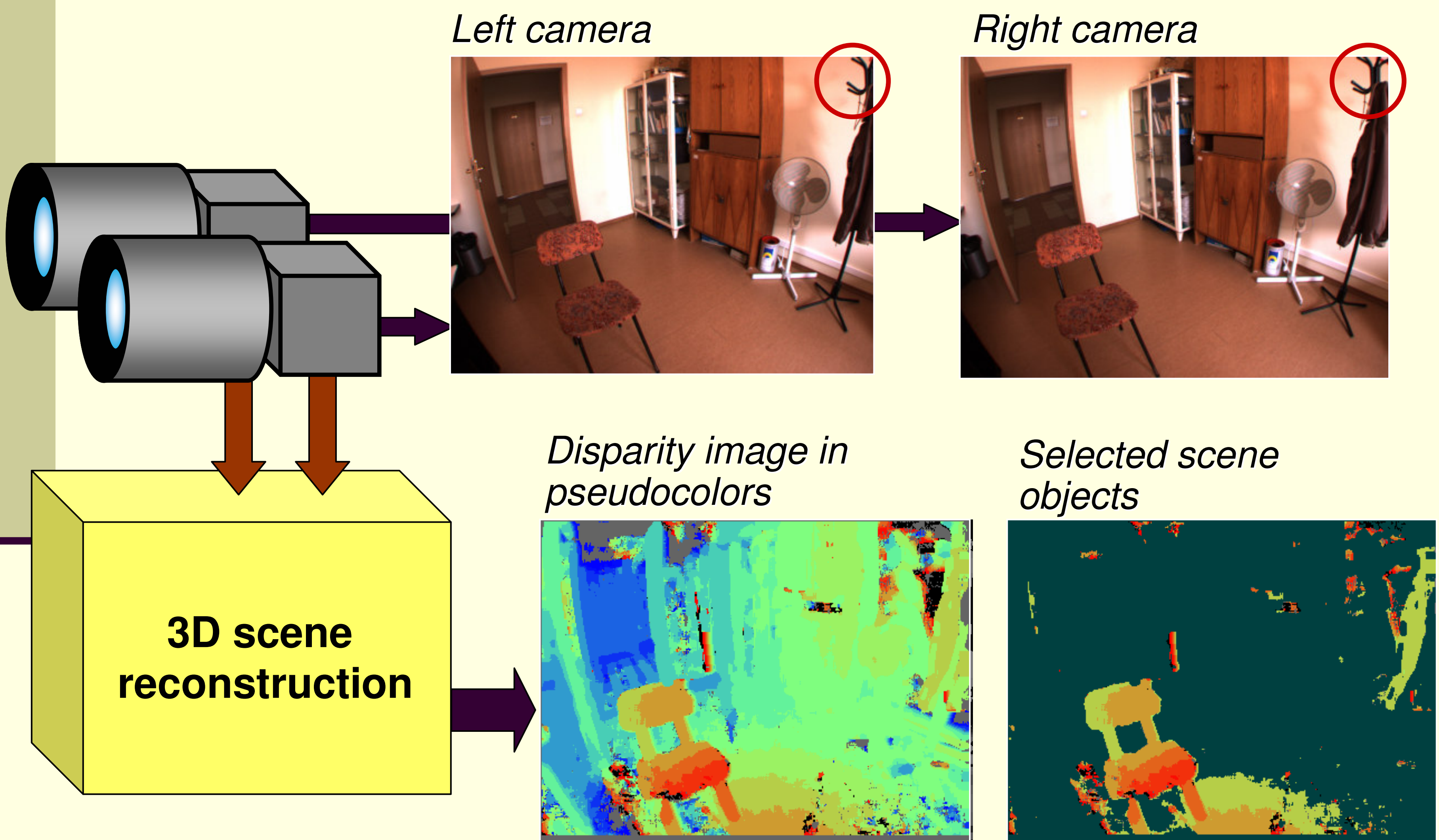
P. Strumiłło, P. Pełczyński

Medical imaging: 3D endoscopic surgery



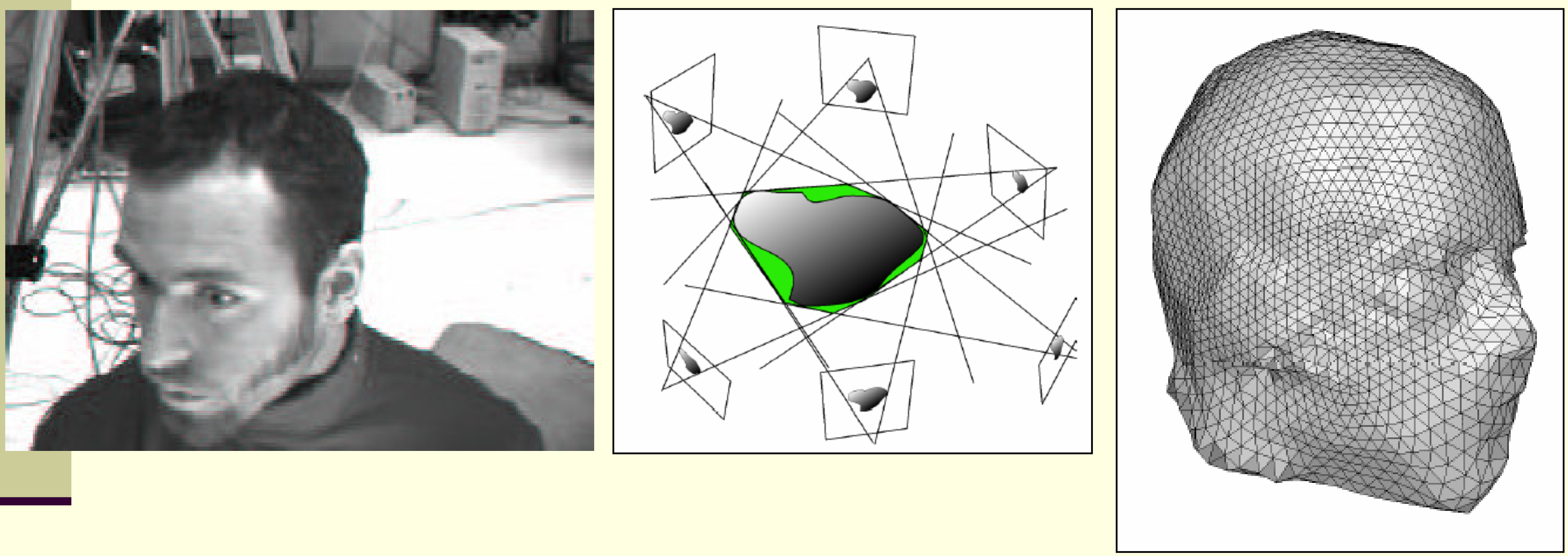
From: *"3D Reconstruction of the operating field for image overlay in 3D-endoscopic surgery" by F. Mourgues, F. Devernay and E. Coste-Maniere*

Travel aid system for blind



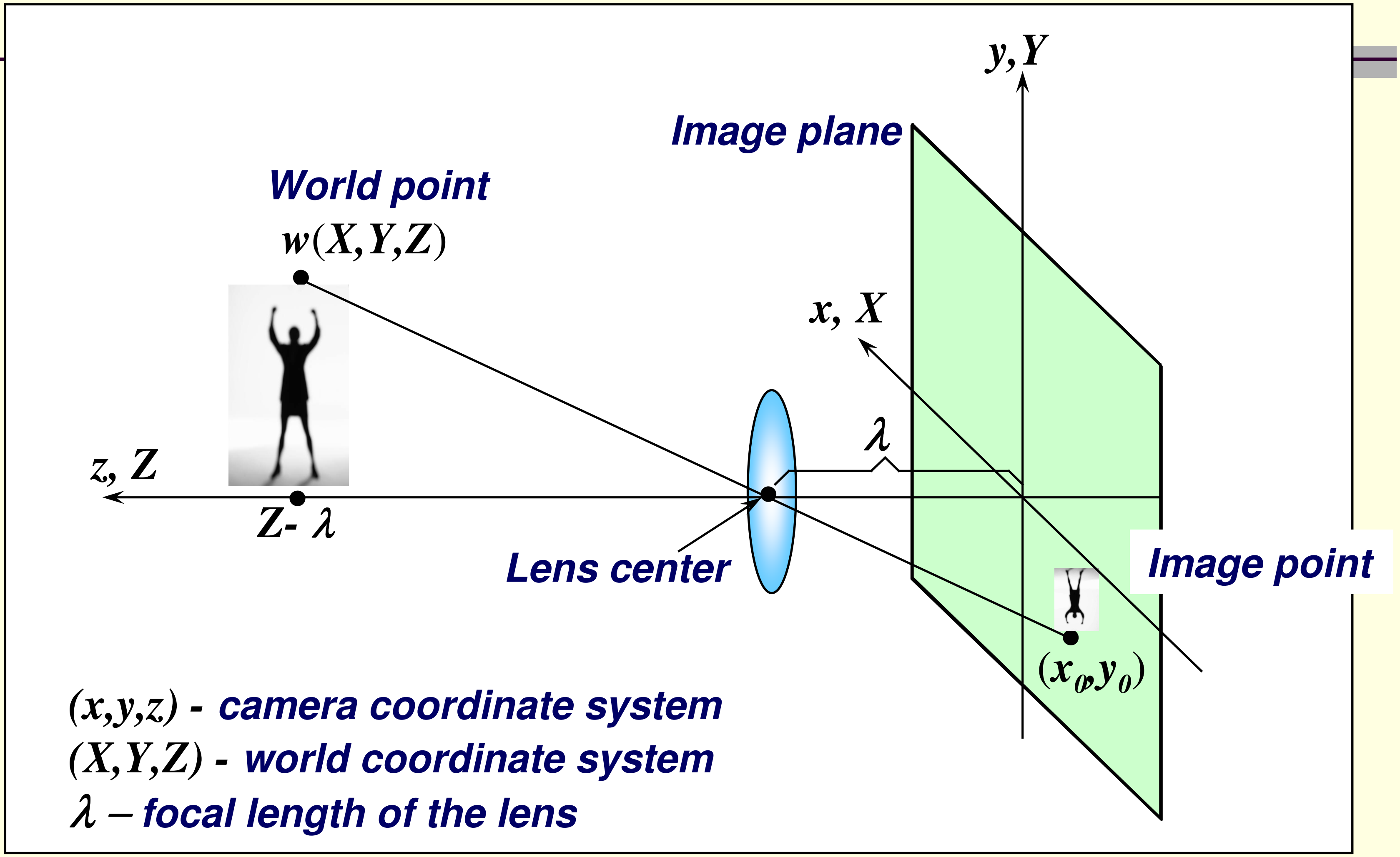
P. Strumiłło, P. Pełczyński

3D object modelling



Copied from the article*: “Spatio-Temporal Stereo using Multi-Resolution Subdivision Surfaces” by J. Neumann and Y. Aloimonos*

Perspective Transformations



"pin-hole" camera model

P. Strumiłło, P. Pełczyński

Perspective Transformation

Projection of the world point onto the image plane

Assumptions:

- world and camera coordinate systems are identically aligned

- $Z>\lambda$

$$\frac{x_0}{\lambda} = -\frac{X}{Z-\lambda} = \frac{X}{\lambda - Z}$$

$$\frac{y_0}{\lambda} = -\frac{Y}{Z-\lambda} = \frac{Y}{\lambda - Z}$$

Inverted images

$$x_0 = \frac{\lambda X}{\lambda - Z}$$

$$y_0 = \frac{\lambda Y}{\lambda - Z}$$

Nonlinear versus Z

P. Strumiłło, P. Pełczyński

Inverse Perspective Transformation

Projection of the image point onto the world point

$$X = \frac{x_0}{\lambda}(\lambda - Z)$$

$$Y = \frac{y_0}{\lambda}(\lambda - Z)$$

Two equations three unknowns

Conclusions

Mapping of 3D scene onto the image plane is a ***many-to-one transformation****: image point corresponds to a set of collinear 3D points*

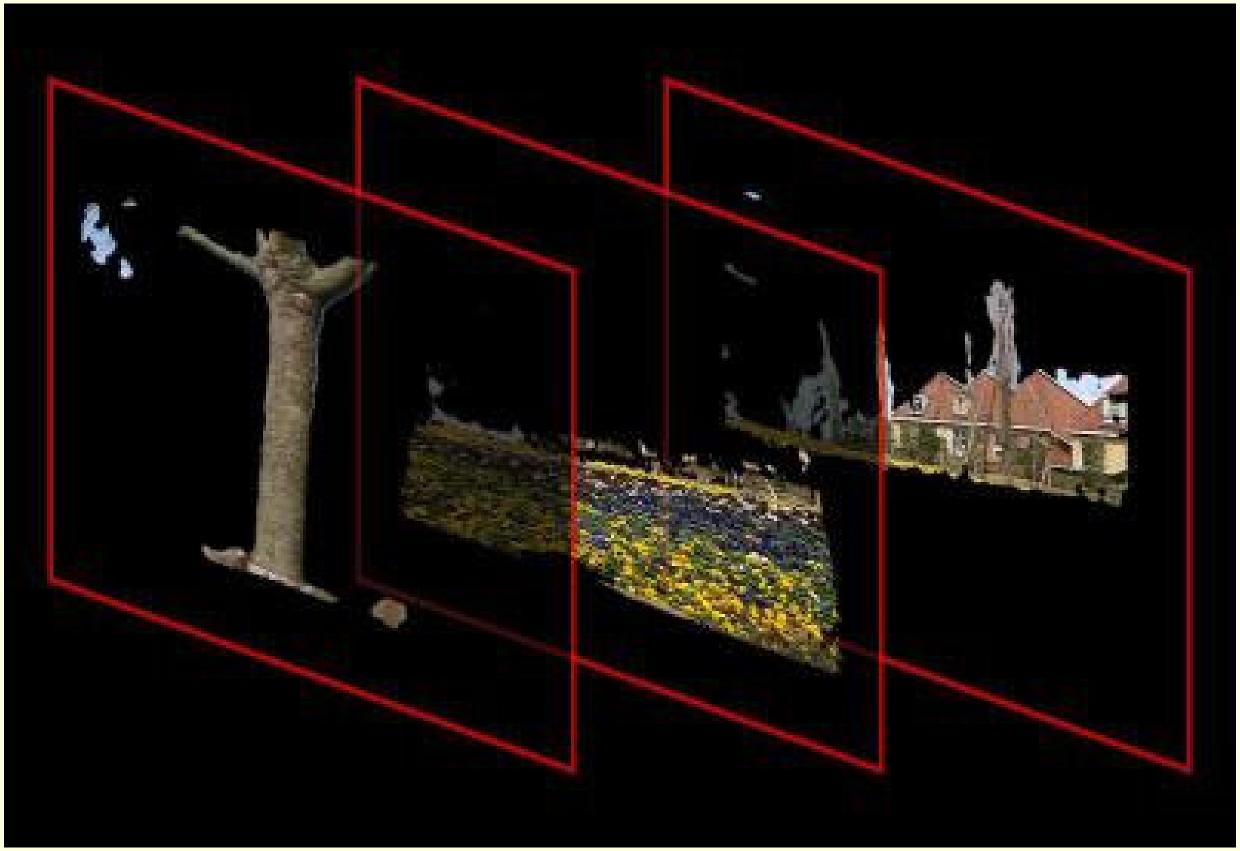
The inverse transformation cannot be performed on the basis of a single image

P. Strumiłło, P. Pełczyński

Other cues about distance from a monocular view

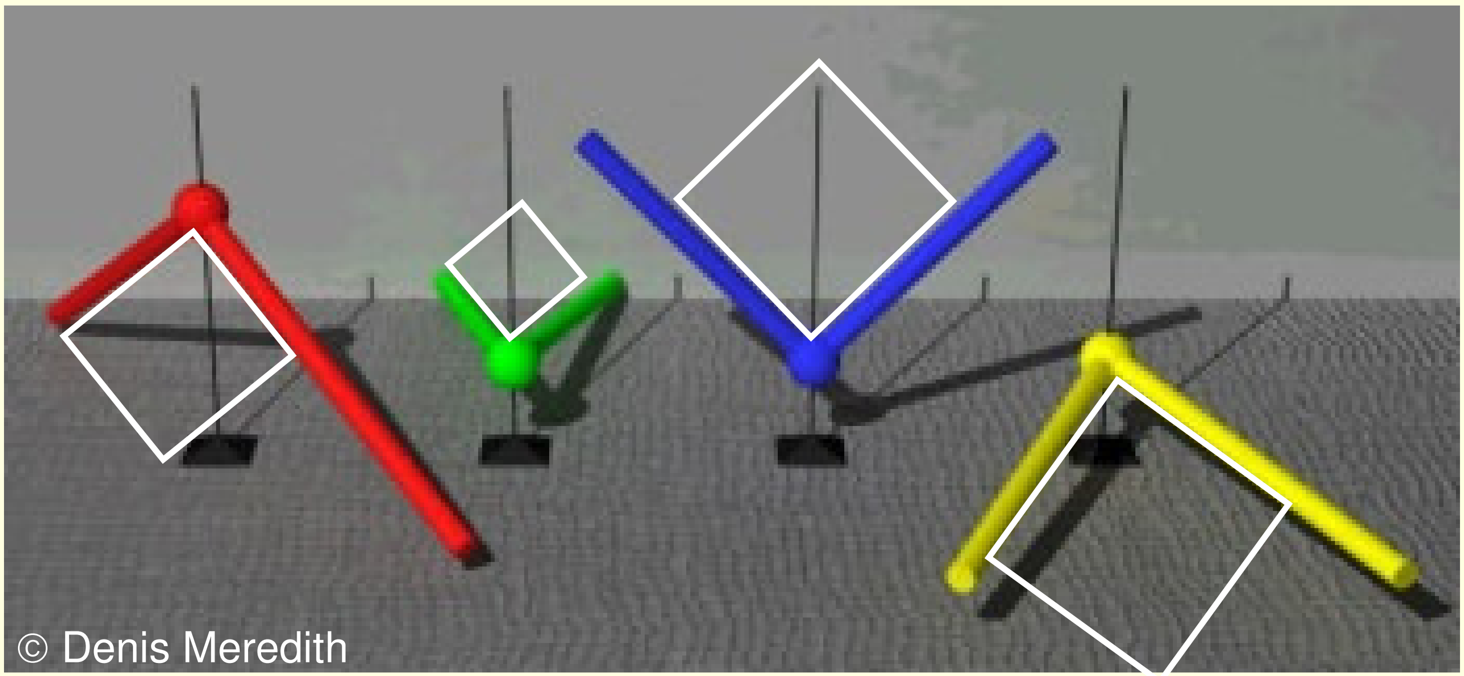
Motion

Other cues about distance from a monocular view



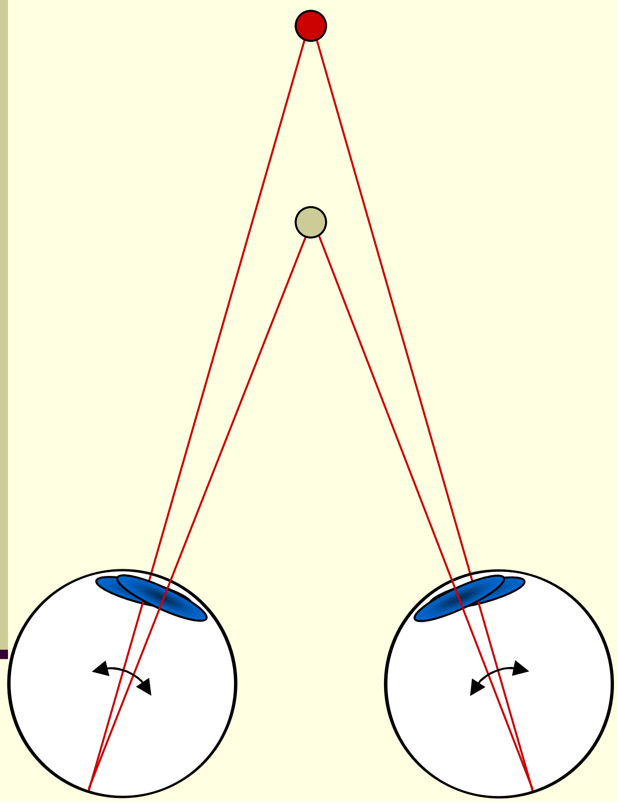
P. Strumiłło, P. Pełczyński

Other cues about distance from a monocular view

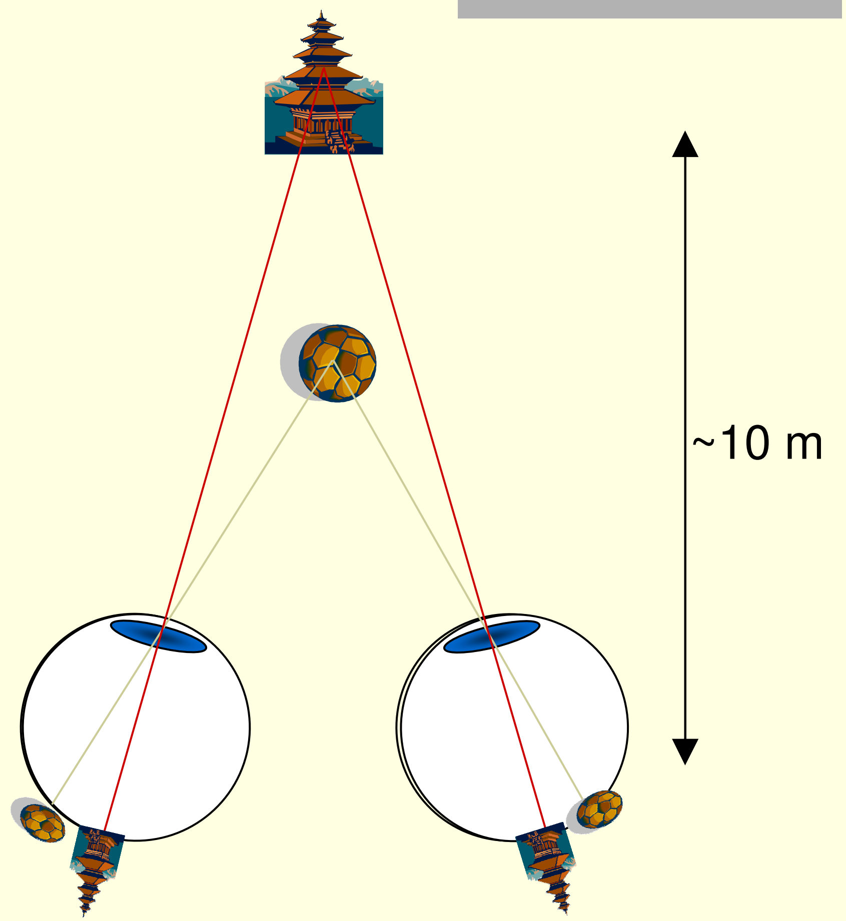


„Precoded" perception of perspective

Binocular vision



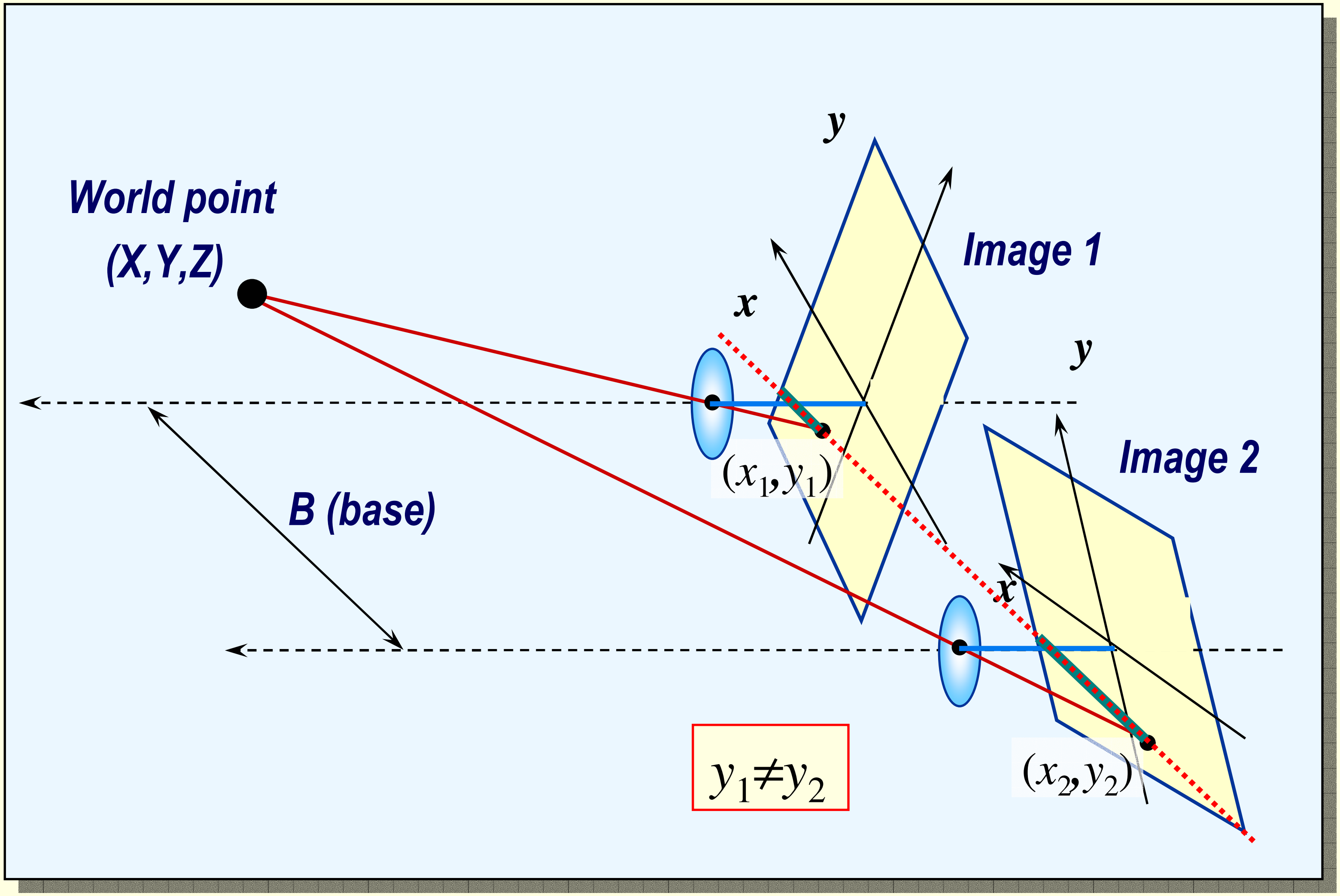
Convergence of eyes optical axis



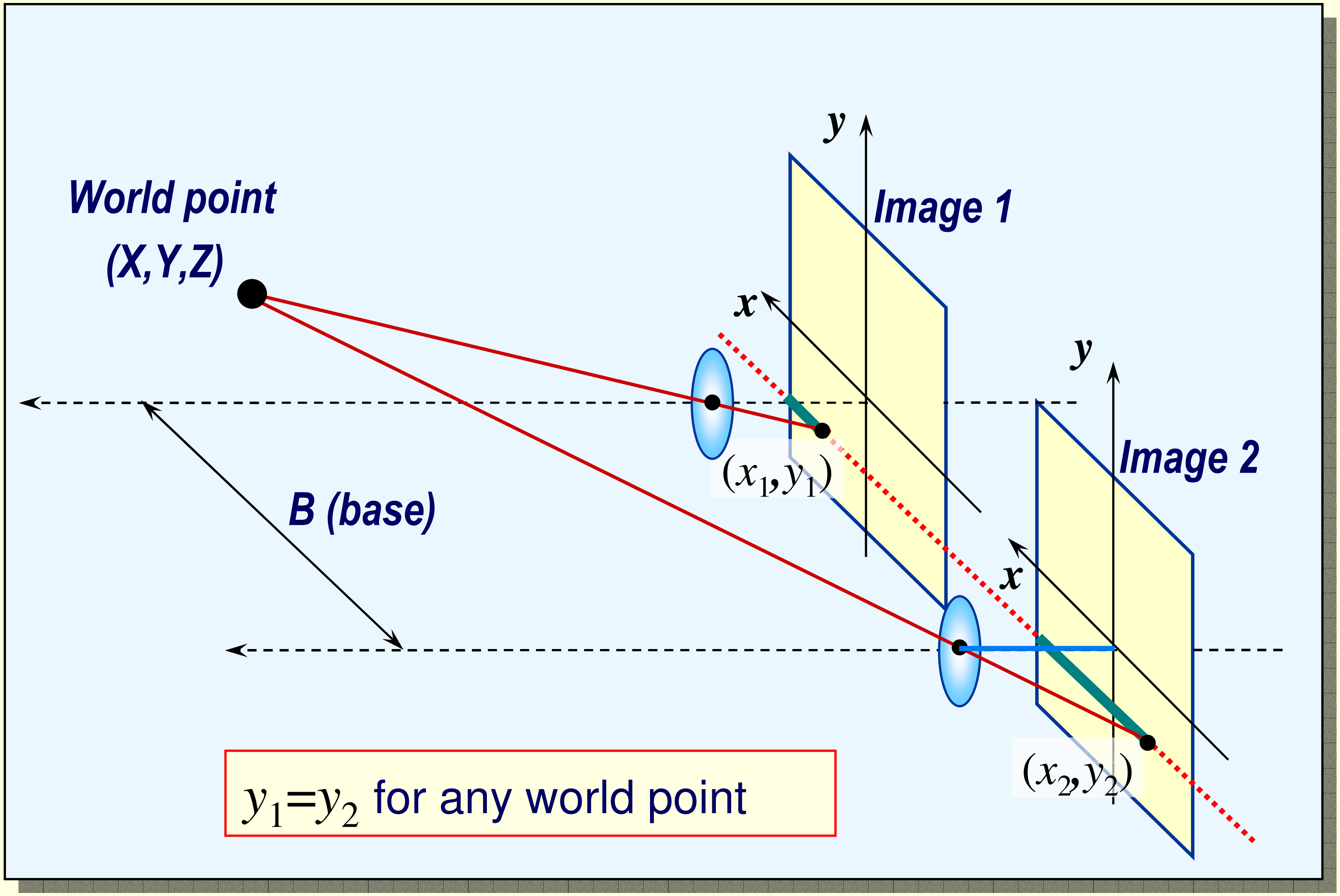
Binocular disparity

P. Strumiłło, P. Pełczyński

Stereo image acquisition: (non-epipolar case)

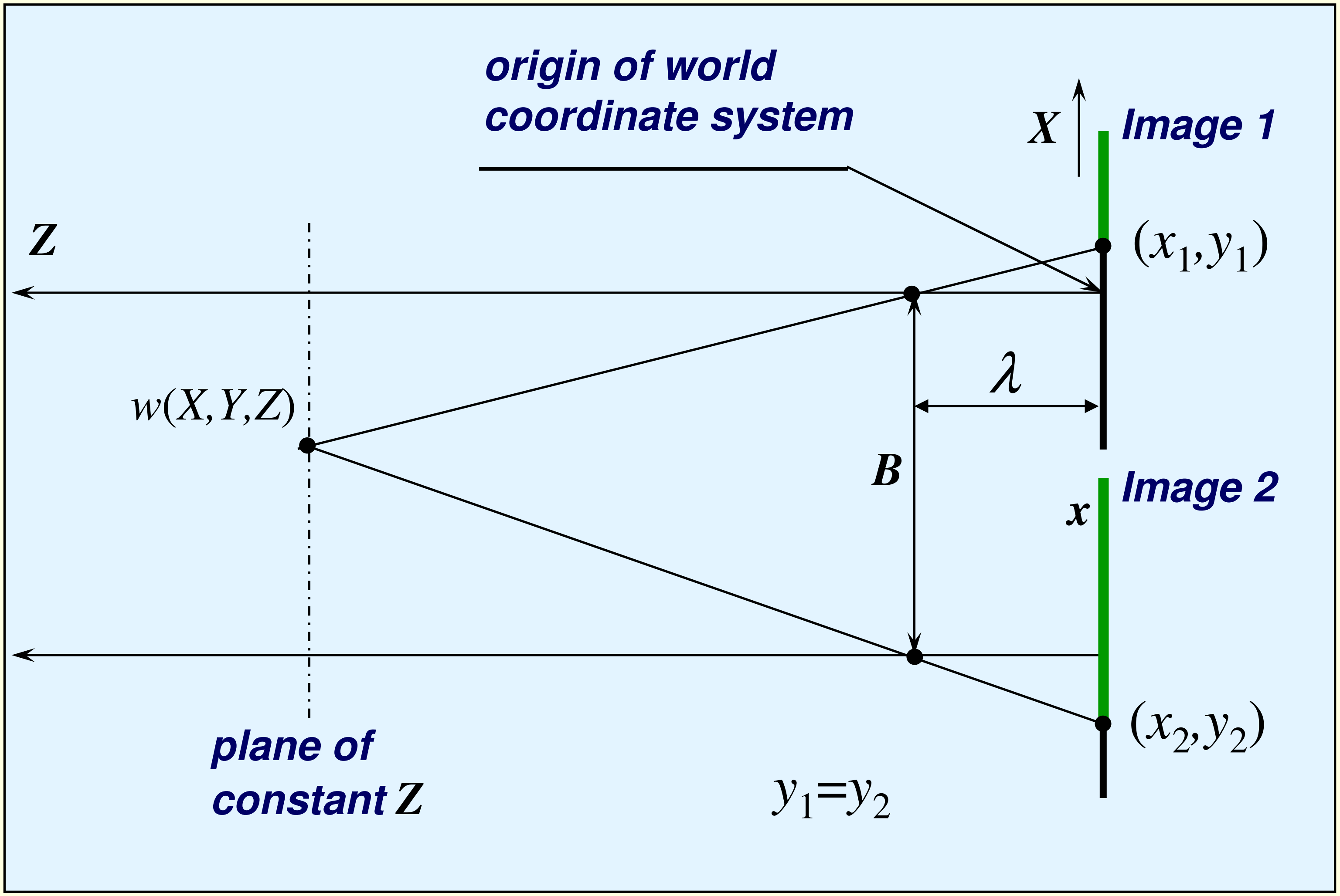


Stereo image acquisition: (epipolar case)



P. Strumiłło, P. Pełczyński

Stereo image acquisition - top view



P. Strumiłło, P. Pełczyński

3D point reconstruction in stereoscopy: Inverse Perspective Transformation

$$X = \frac{x_1}{\lambda}(\lambda - Z)$$

$$Y = \frac{y_1}{\lambda}(\lambda - Z)$$

$$Z = \lambda - \frac{\lambda B}{x_2 - x_1}$$

disparity ($x_2 - x_1$)

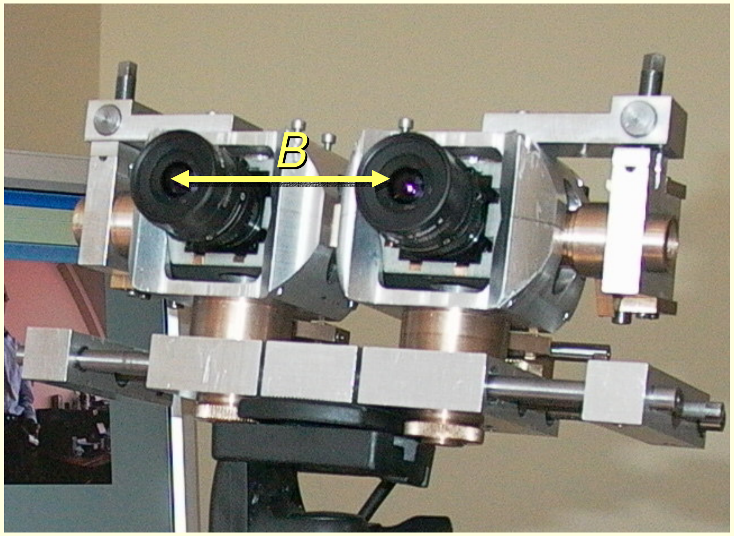
$$Z \cong \frac{\lambda B}{x_2 - x_1}$$

Three equations - three unknowns

P. Strumiłło, P. Pełczyński

3D point reconstruction in stereoscopy: Inverse Perspective Transformation

$$Z \cong \frac{\lambda B}{x_2 - x_1} = \frac{\lambda B}{\Delta_p n}$$



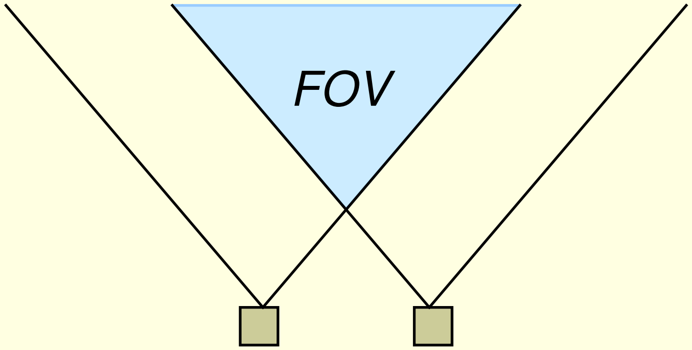
<u>Example:</u>

B =8cm – base of the stereovision system

λ =5mm – lens focal length

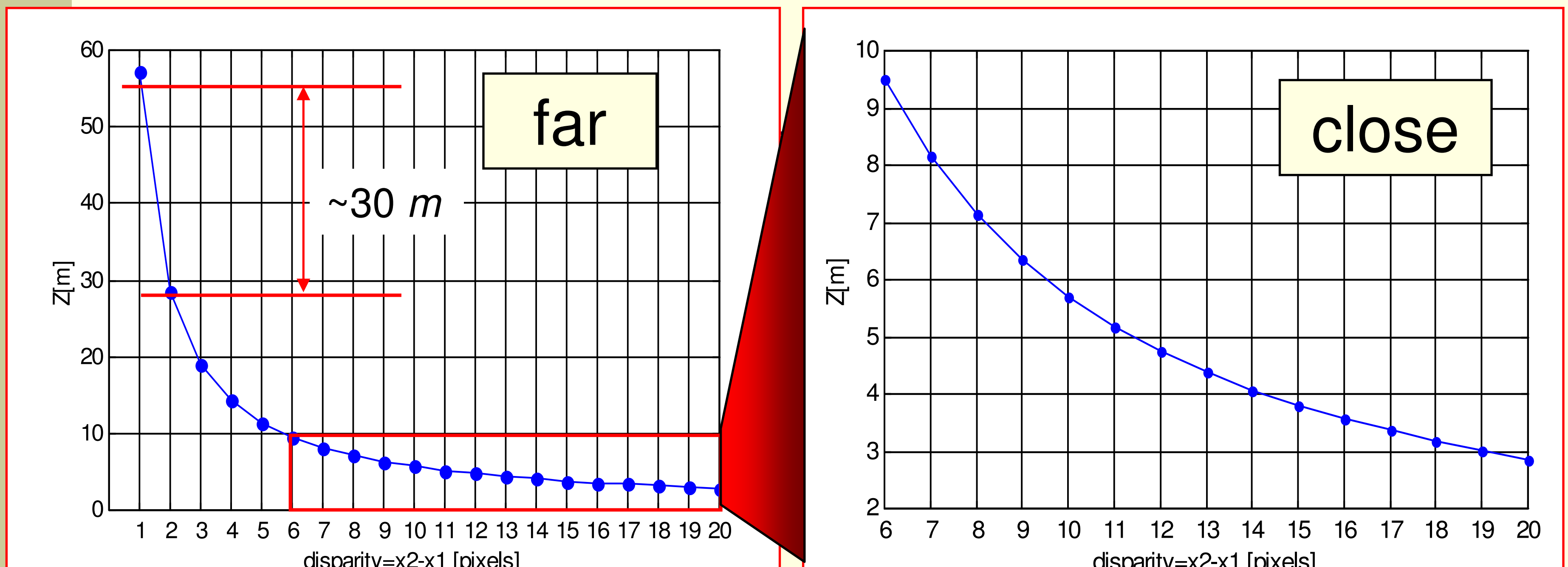
Δ_p=7μm – pixel size

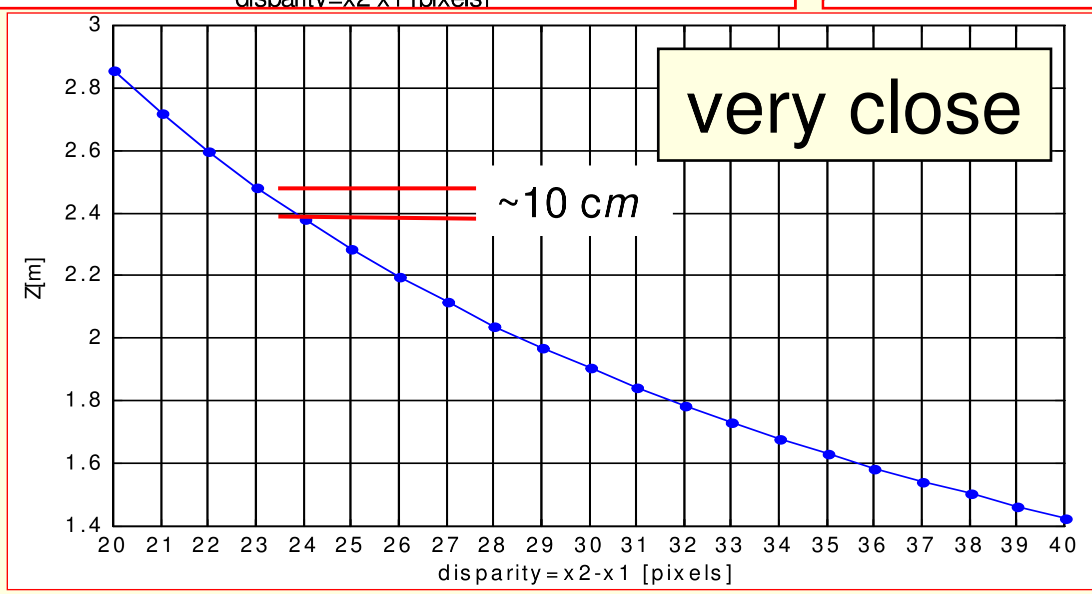
n – integer number $n \geq 1$



P. Strumiłło, P. Pełczyński

Depth resolution in stereoscopy

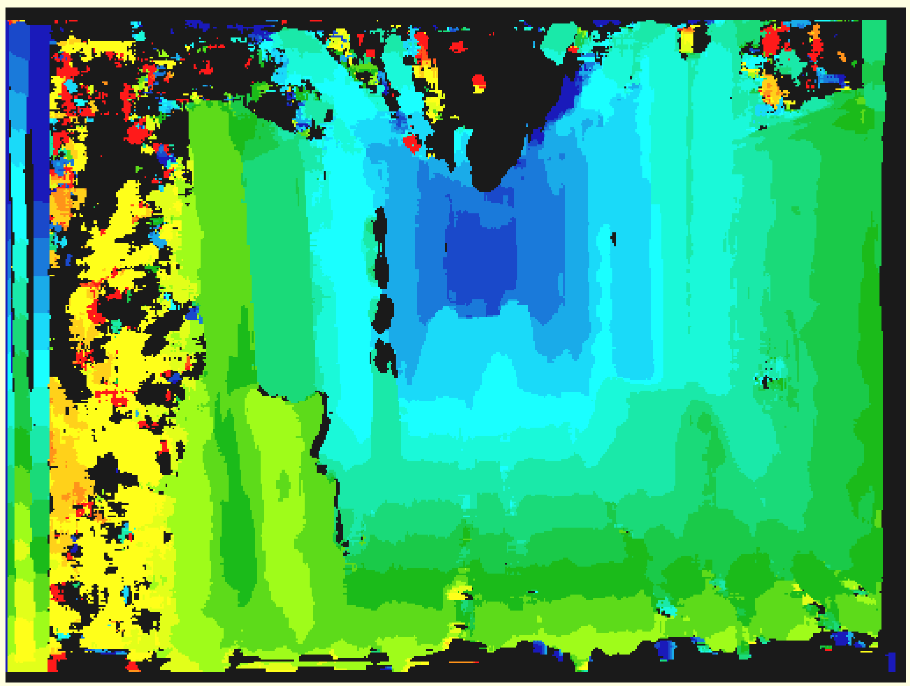




Depth error ∝ 1/B

Sub-pixel accuracy required for precise depth estimation!


How to match feature regions?



Digital image matching: correspondence problem

Digital image matching *automatically establishes the correspondence between primitives extracted from two or more digital images depicting at least part of the same scene*

Corner detection (Harris detector)

Image matching problems:

- *selection of primitives for matching*
- *choice of models for mapping of primitives*
- *measure of similarity of corresponding primitives*
- *matching algorithm*
- *matching strategy*

P. Strumiłło, P. Pełczyński

Digital image matching approaches

Local correspondence methods:

- *Block matching*
- *Gradient-based optimisation*
- *Feature matching*

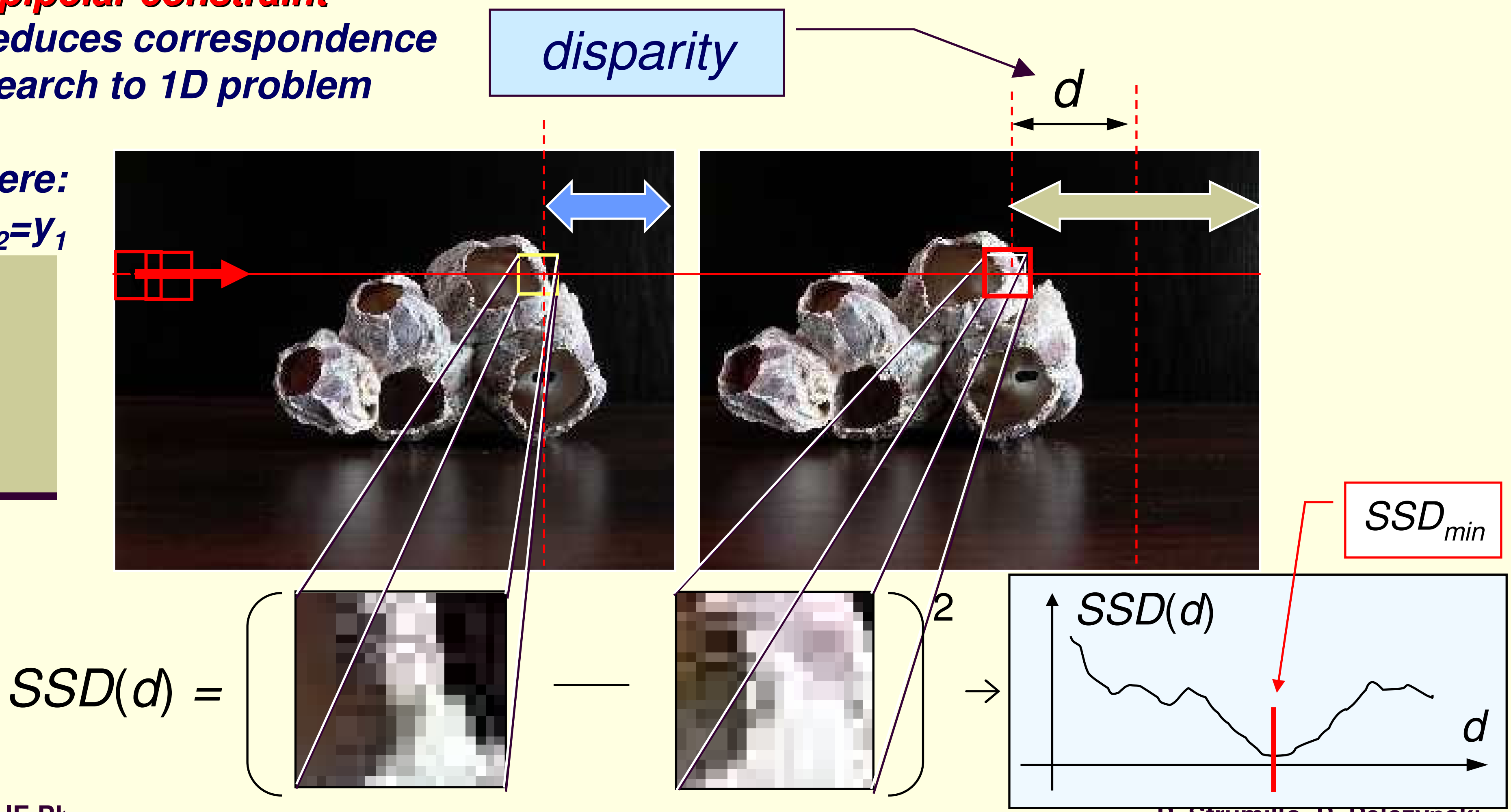
Global correspondence methods:

- *Dynamic programming*
- *Intrinsic curves*
- *Graph cuts*
- *Other methods*

Digital image matching - the concept of disparity

Epipolar constraint ***reduces correspondence search to 1D problem***

here: $y_2 = y_1$



P. Strumiłło, P. Pełczyński

Metrics in block image matching

- *Normalised cross-correlation*
- *Sum of squared differences SSD*
- *Normalised sum of squared differences*
- *Sum of absolute diffrences SAD*
- *Rank*
- *Census*

$$\sum_{x,y}\left(I_1(x,y)-I_2(x+d,y)\right)^2 \rightarrow \min$$

89	63	72
67	55	64
58	51	49

⇨ 2

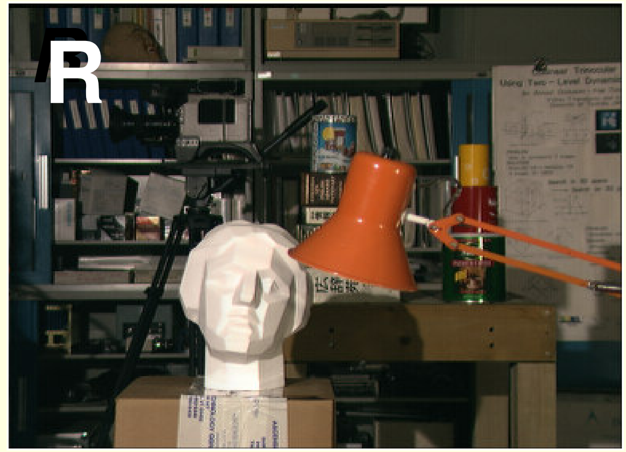
Rank

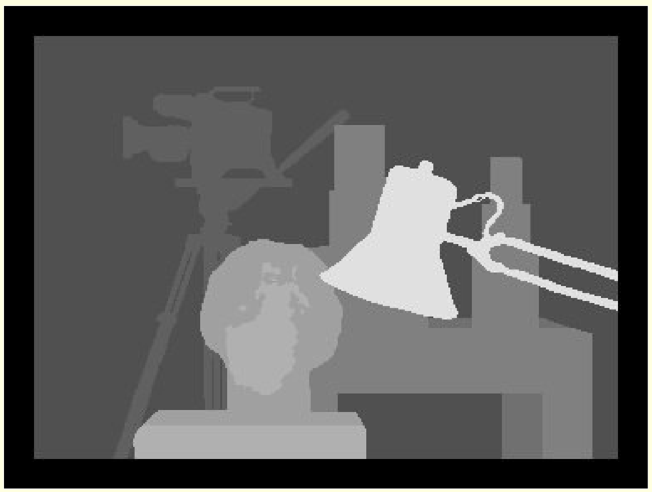
89	63	72
67	55	64
58	51	49

⇨ 00000110

Census

Image data for verification of matching algorithms





ground truth disparity image

computed disparity image

Summary – issues in 3D scene reconstruction using stereoscopy

Hardware:

1. *Good quality gen-locked cameras*
2. *Camera resolution vs. stereo-base tradeoff*
3. *Decide about FOV (viewing camera angles)*
4. *Rigid mounting of the cameras!*

Software:

1. *Geometric distortions need to be corrected (internal geometry)*
2. *Rectify stereo images to the canonical set-up (external geometry)*
3. ***Correspondence** (expect: occlusions, non-texture regions, non-Labertian surfaces)*
4. ***Reconstruction** (3D scene structure from disparity map, is dense disparity really required?)*
5. *Real-time operation?*

Calibration (items 1–2)