Reconstruction of three dimensional scenes

Paweł Pełczyński, Paweł Strumiłło

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*

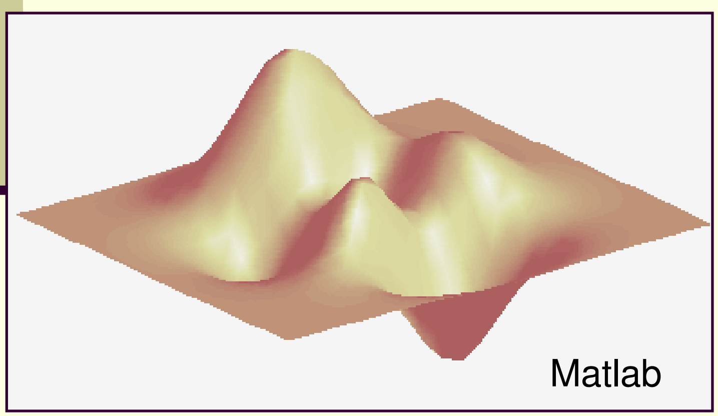
There are methods of 3D scene reconstruction from a set of 2D images

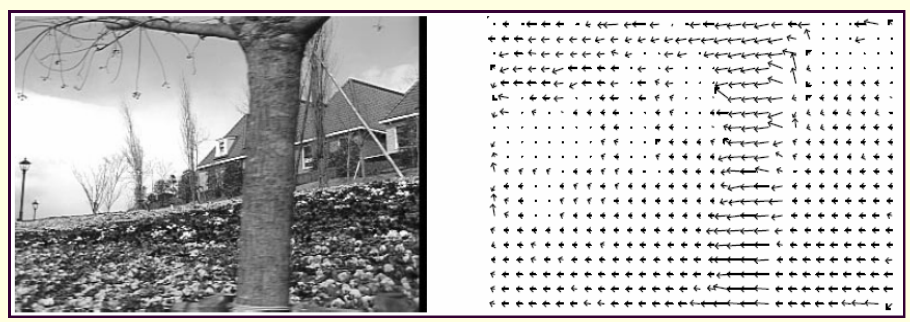
3D scene reconstruction methods

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


Gaël Neuez


©Tokyo-Boeki


Matlab



Applications

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

Robotics - robot vision

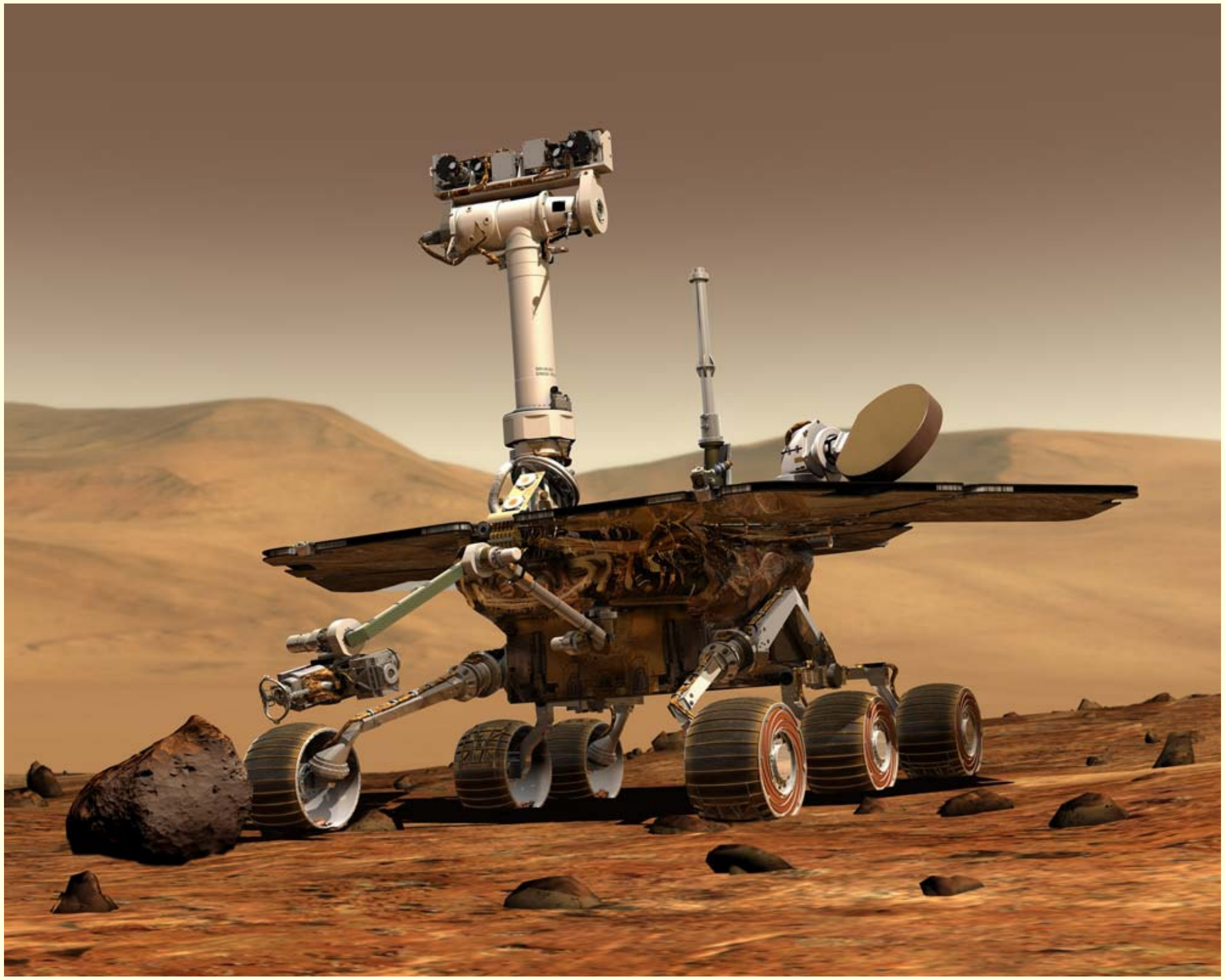
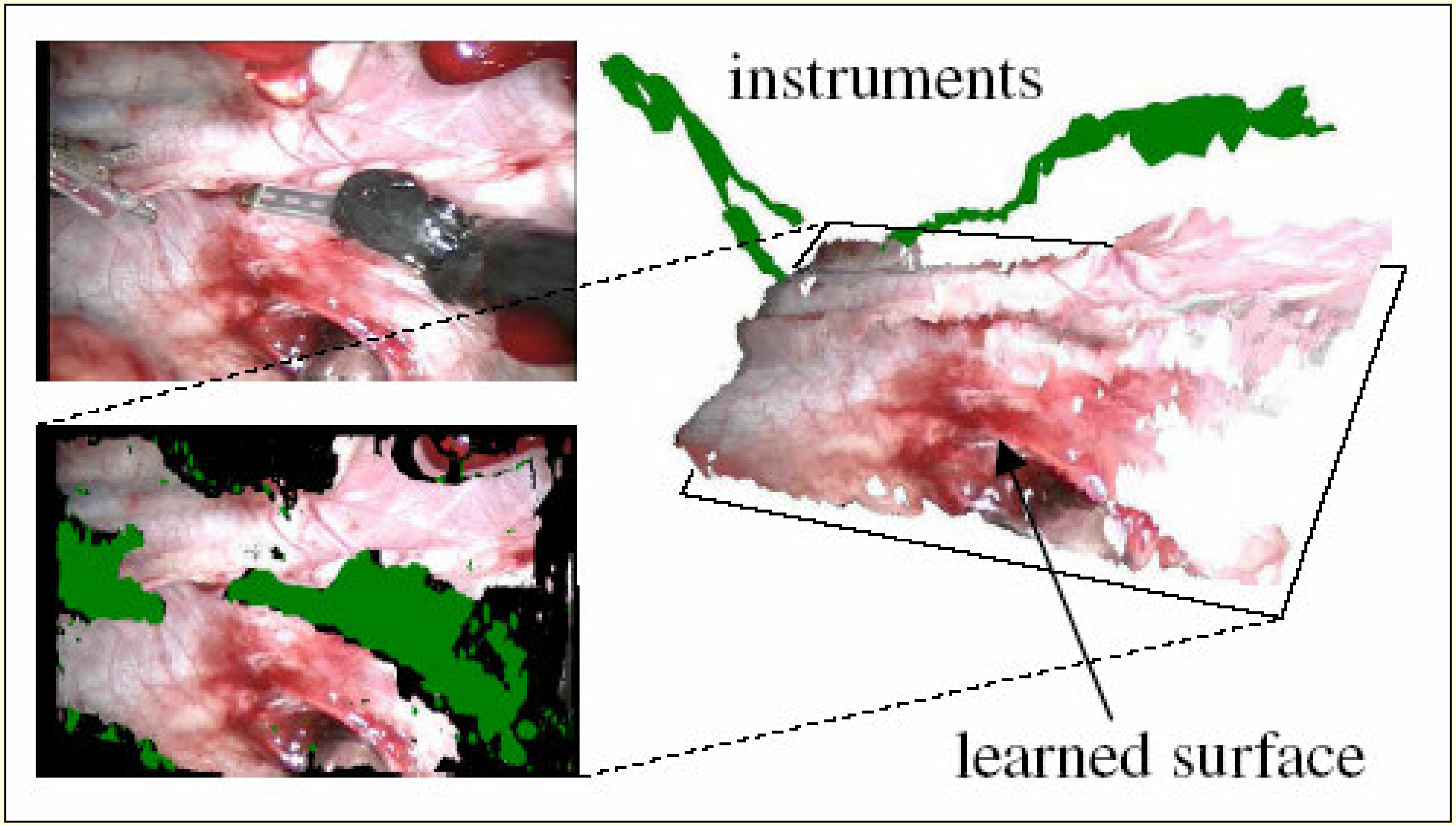


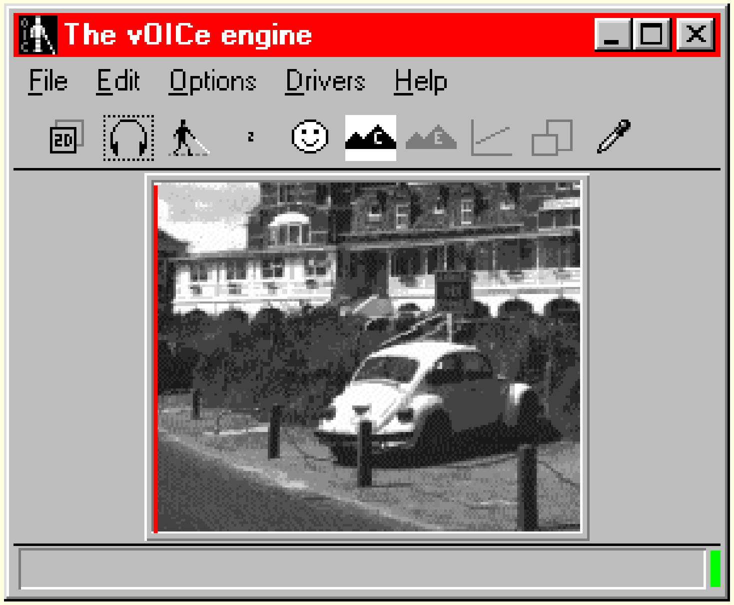
Image copied from : http://marsrovers.jpl.nasa.gov

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*

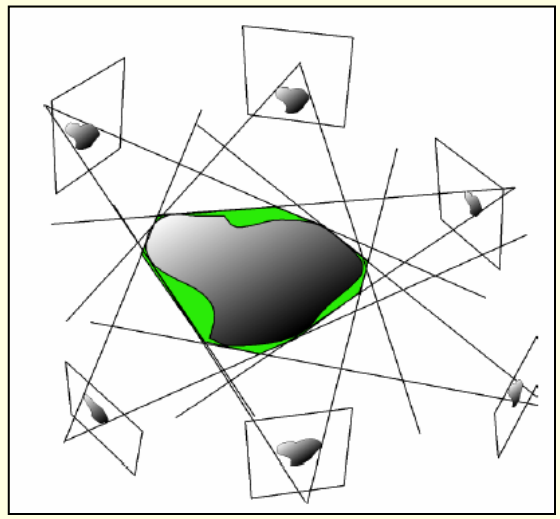
The vOICe
travel aid system for blind people

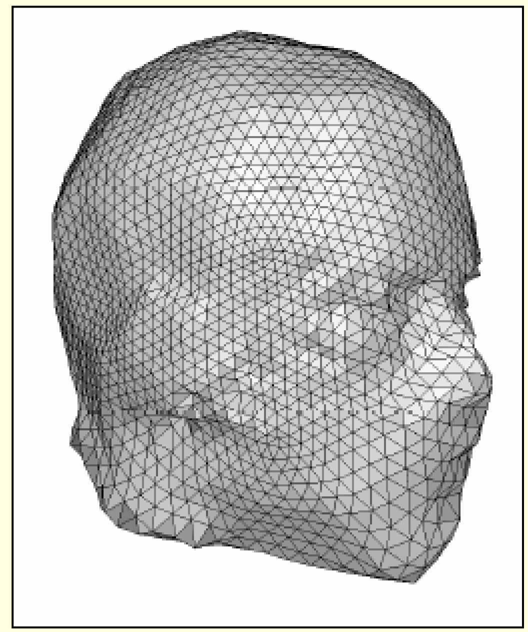


Acoustic code:

- *pitch* – *object location*
- *loudness* – *object brightness*

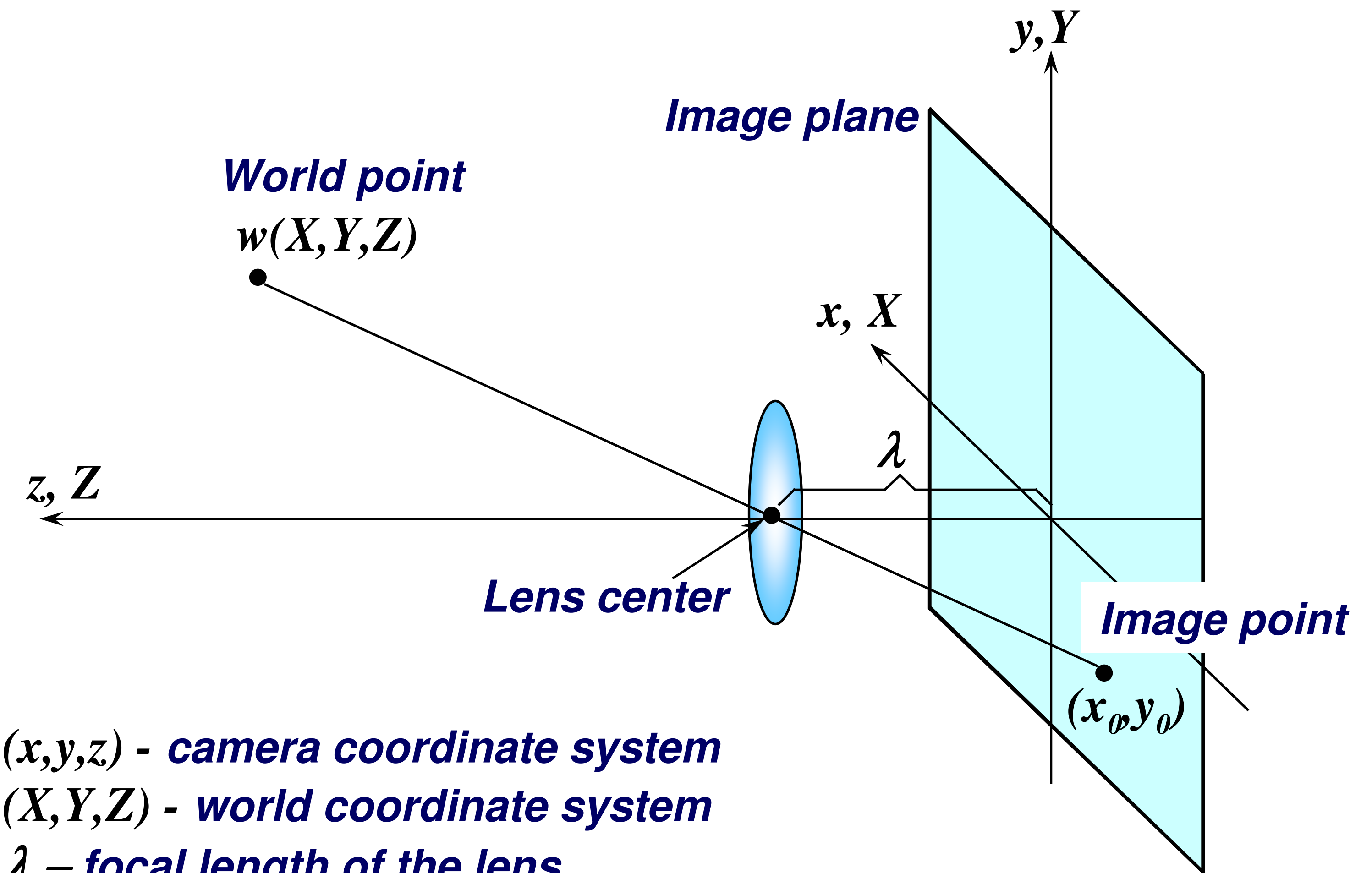
3D object modelling





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

Perspective Transformations



(x,y,z) - *camera coordinate system*

(X,Y,Z) - *world coordinate system*

λ – *focal length of the lens*

Basic model of the image formation process assumed "pin-hole" camera model

Perspective Transformation

Projection of the world point onto the image plane

Assumptions:

- world and camera coordinate systems are identically aligned

- $Z > \lambda$

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

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

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

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

Nonlinear versus Z

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

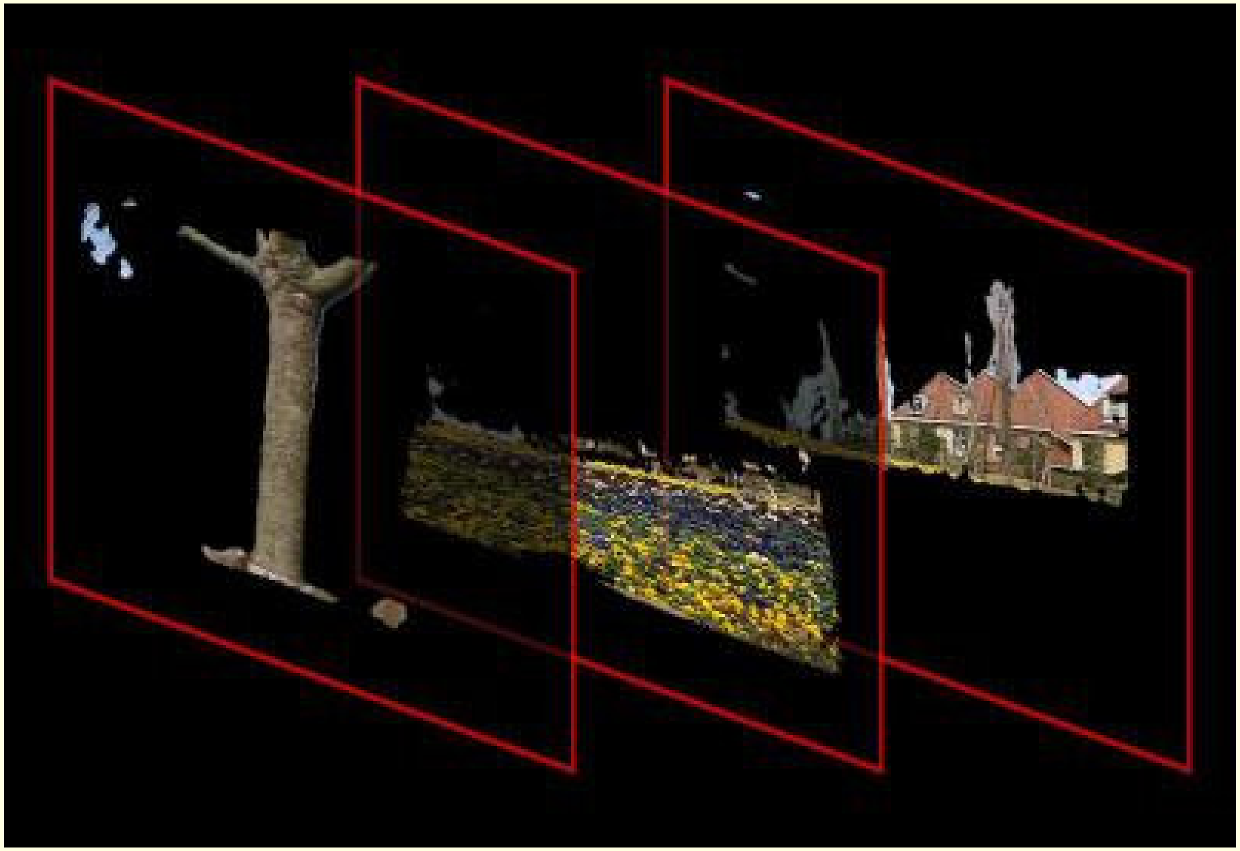
Mapping of 3D scene onto the image plane is a ***many-to-one transformation****: image point corresponds to a set of collinear 3D points*

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

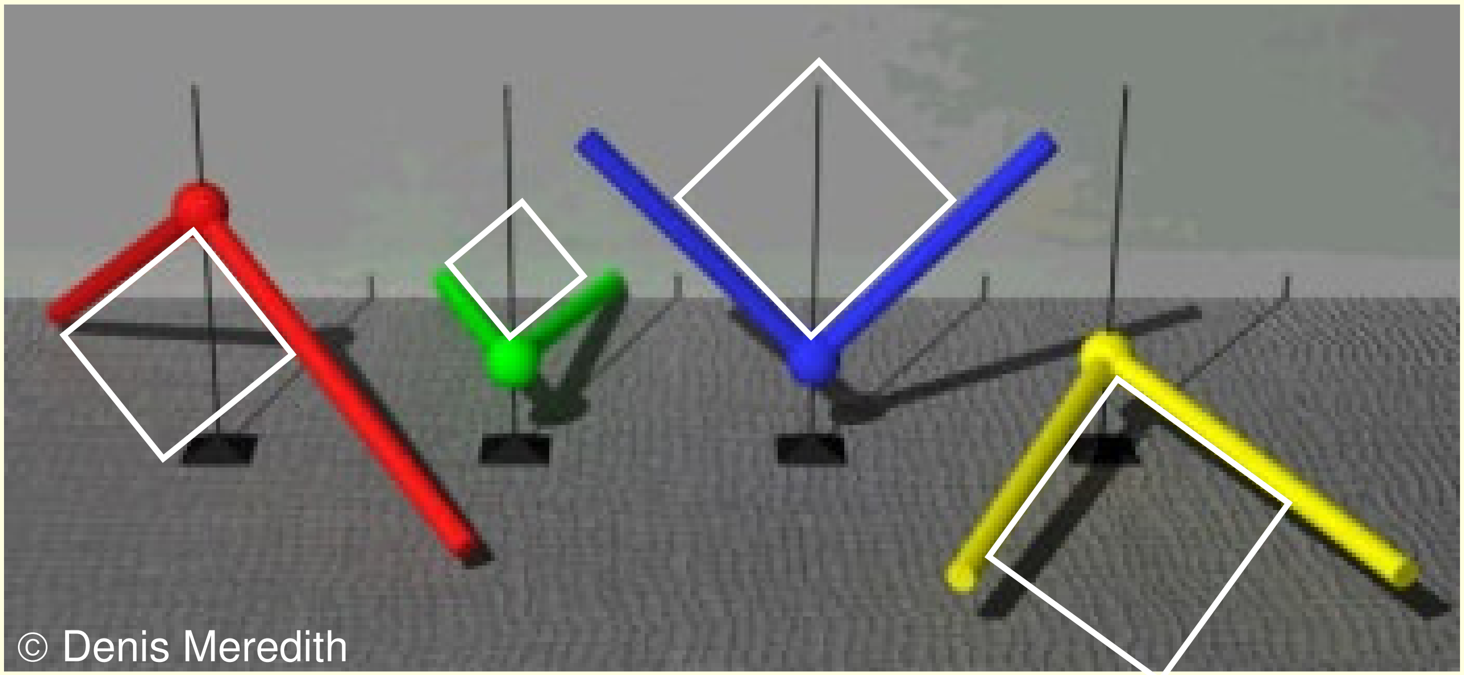
Other cues about distance from a monocular view

Motion

Other cues about distance from a monocular view

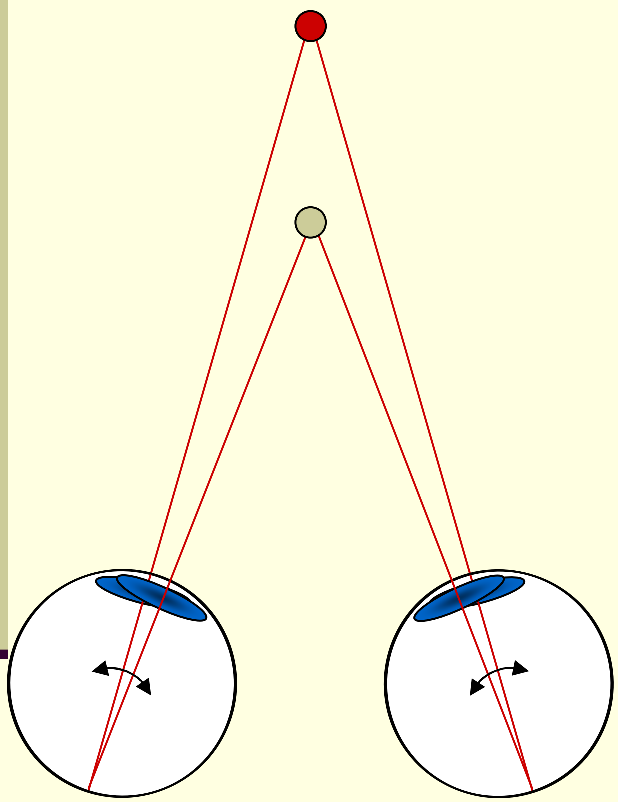


Other cues about distance from a monocular view

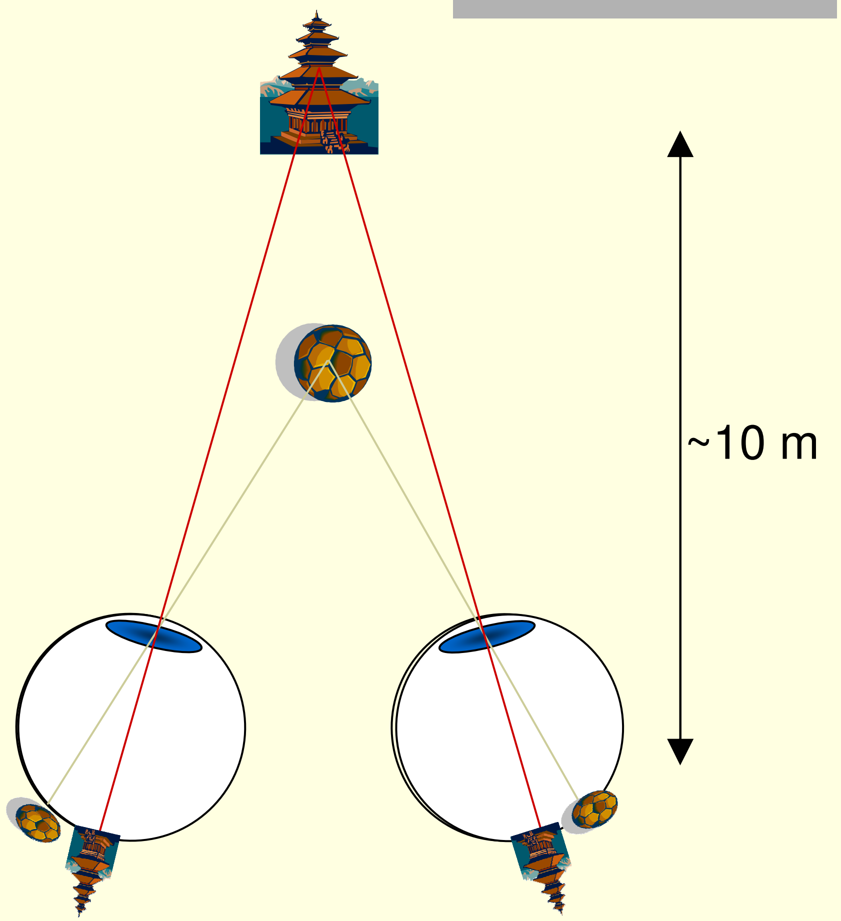


„Precoded" perception of perspective

Binocular vision

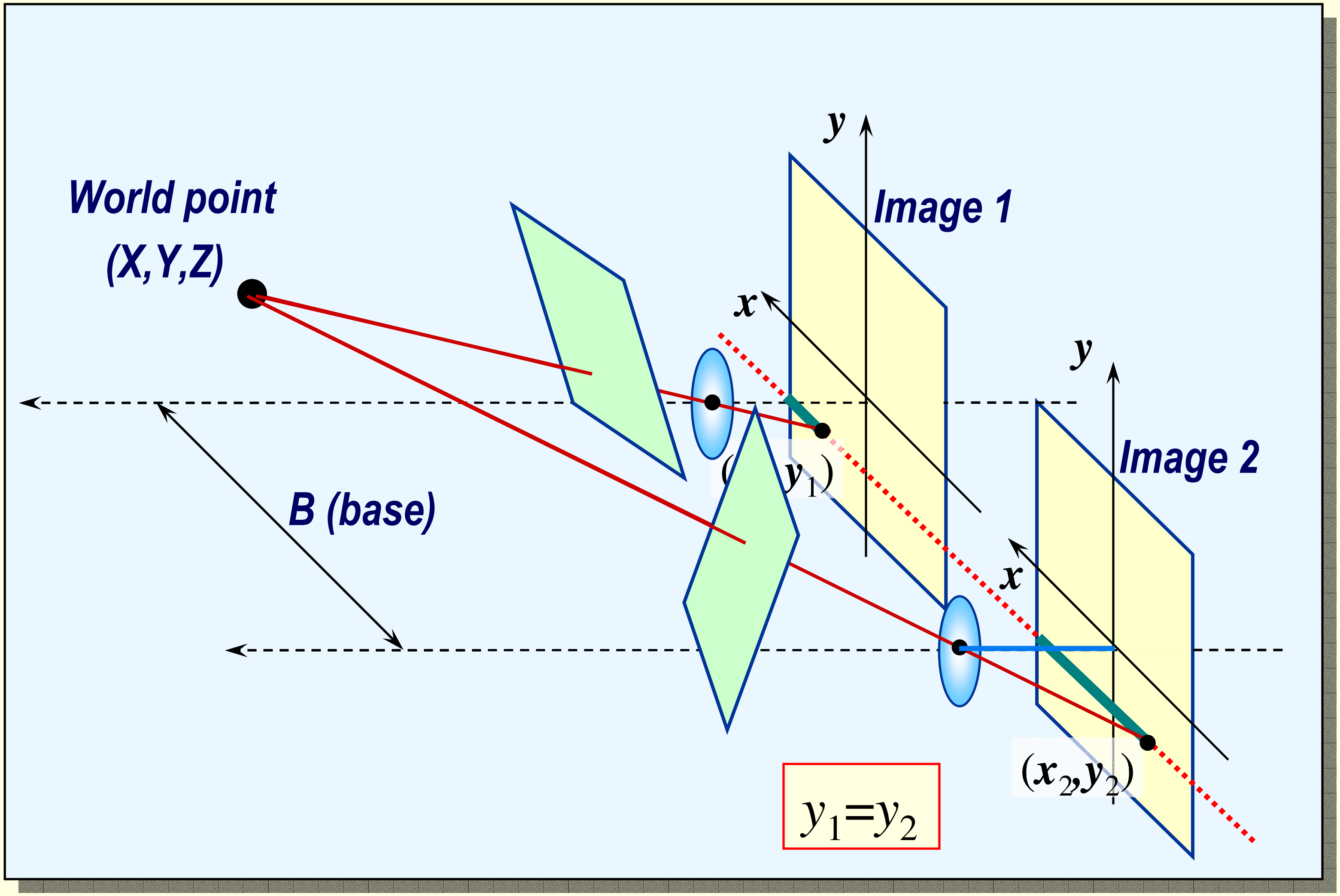


Convergence of eyes optical axis

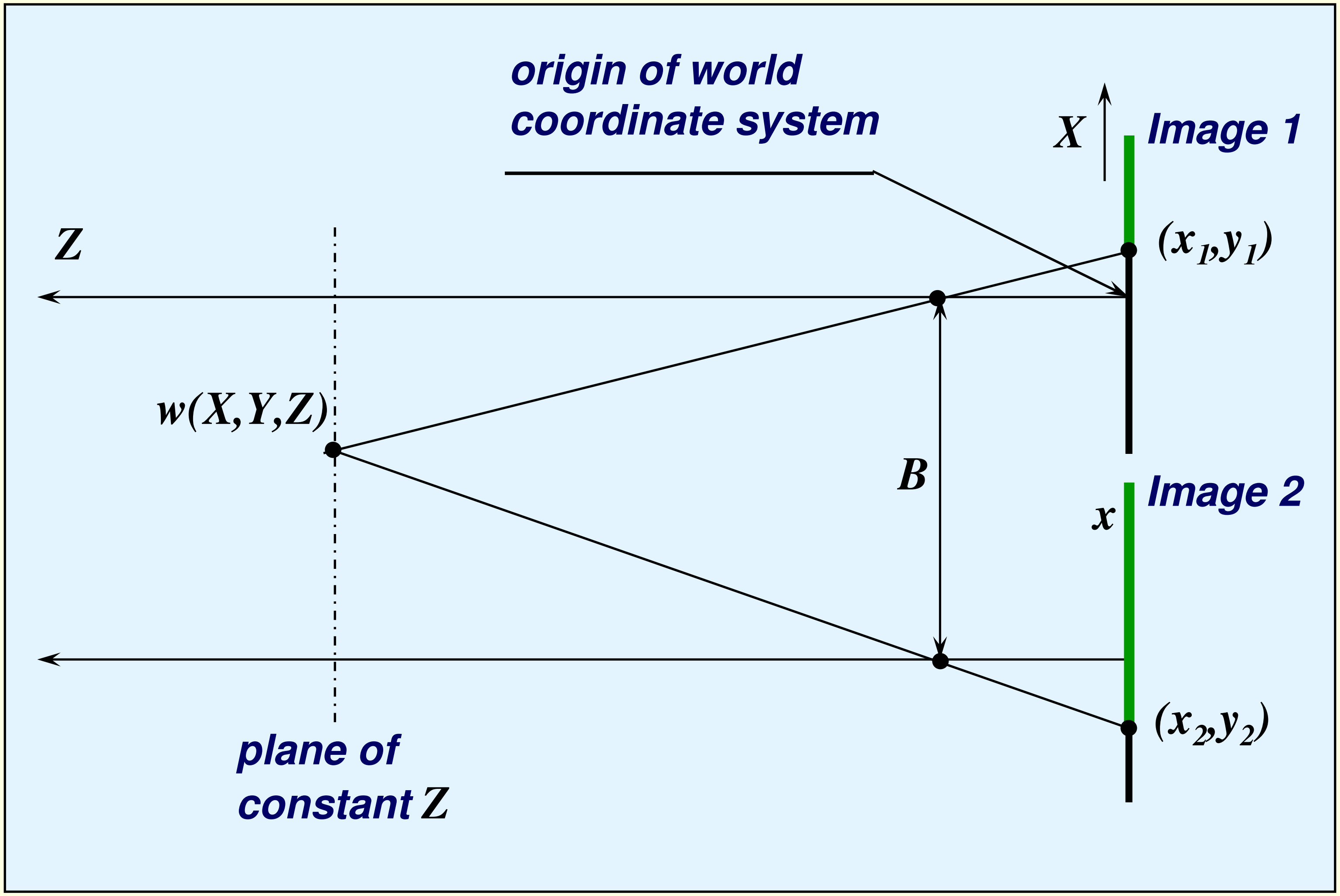


Binocular disparity

Stereo image acquisition: (epipolar case)



Stereo image acquisition - top view



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

Three equations - three unknowns

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*

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*

Digital image matching approaches

Local correspondence methods:

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

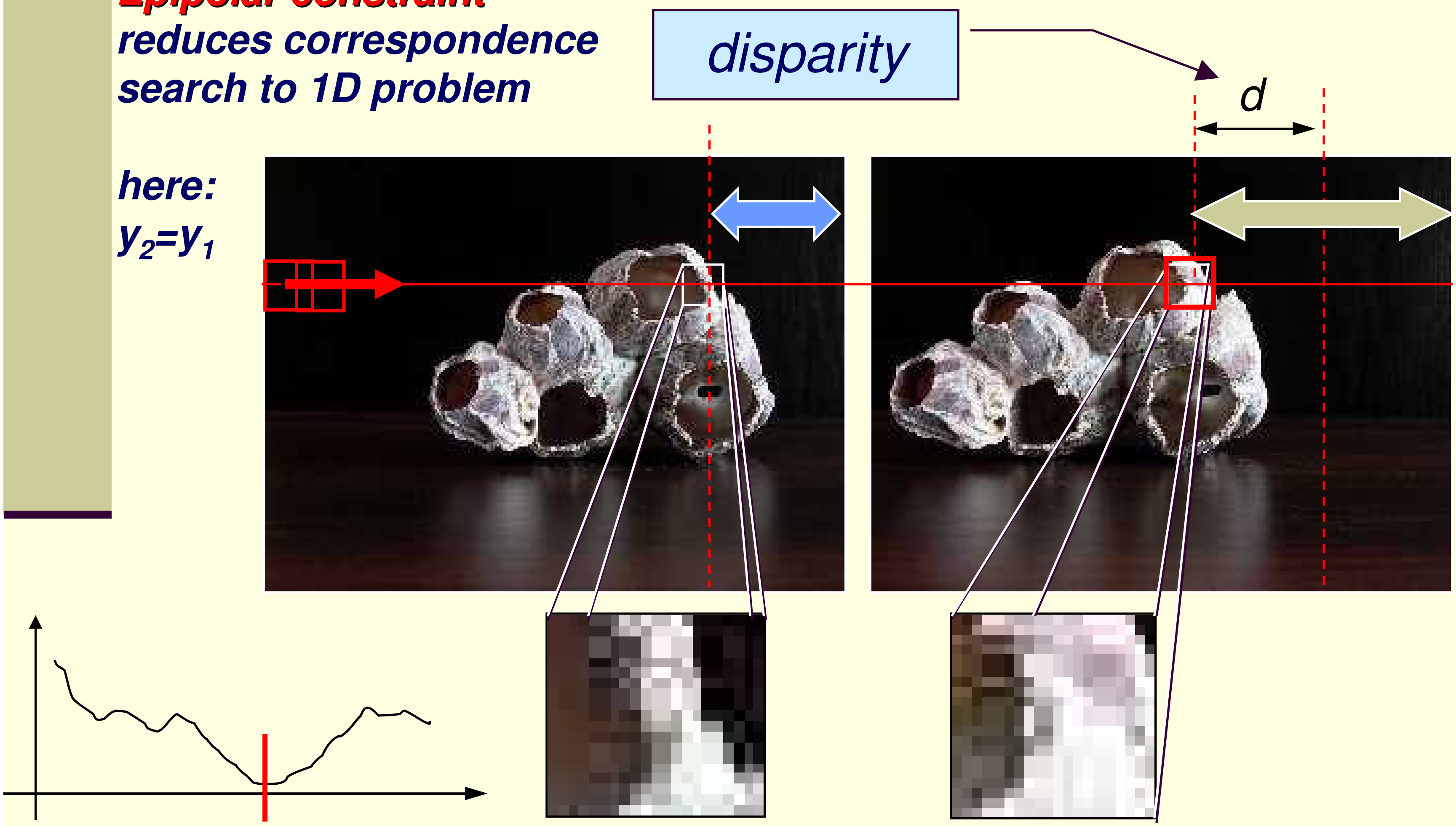
Global correspondence methods:

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

Digital image matching
- concept of disparity

Epipolar constraint *reduces correspondence search to 1D problem*

here:
$y_2 = y_1$



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

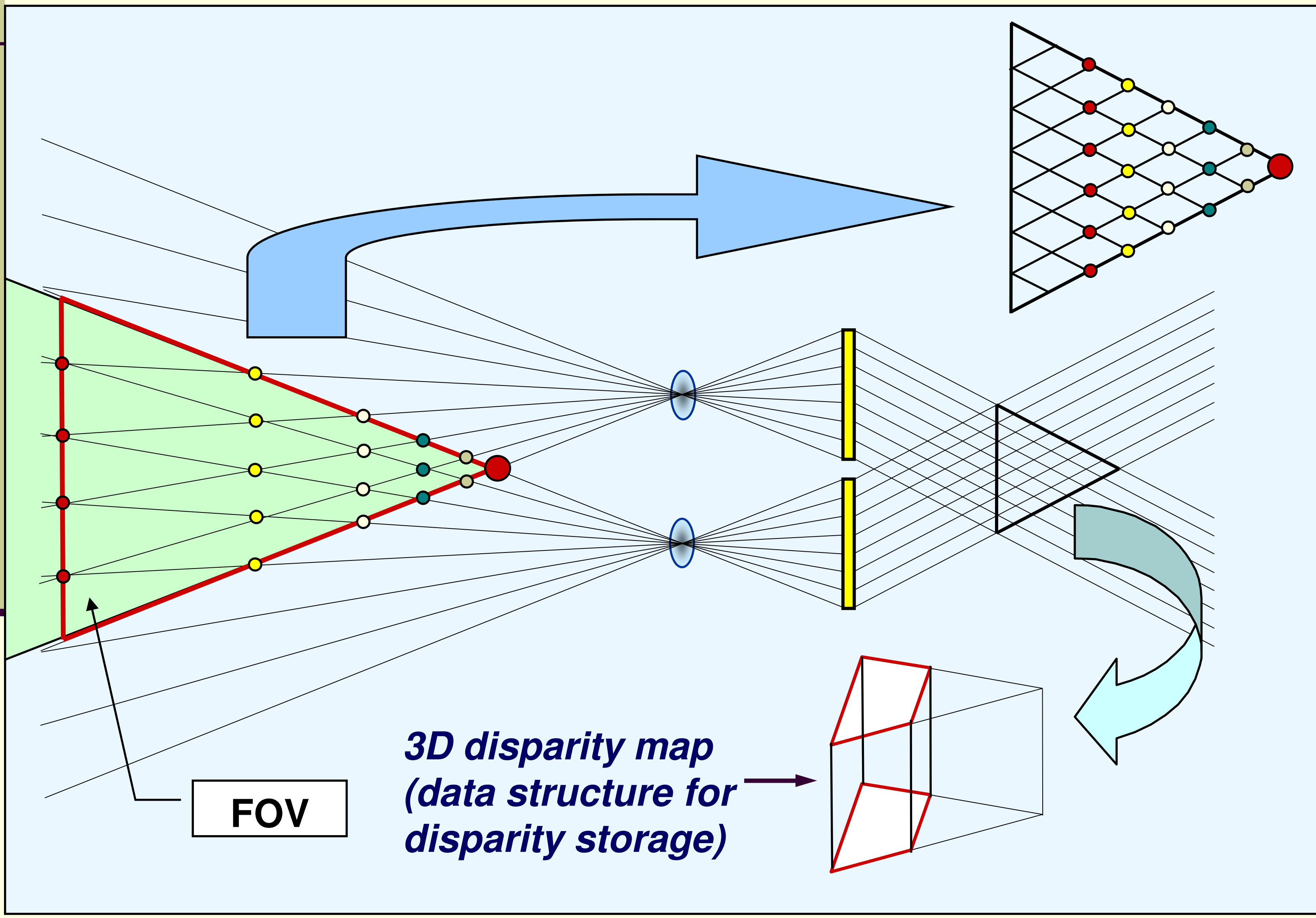
Rank

89	63	72
67	55	64
58	51	49

⇨ 00000110

Census

Digital image matching
- 3D disparity map building



Digital image matching - intersection of disparity map

Outex – University of Oulu Texture Database

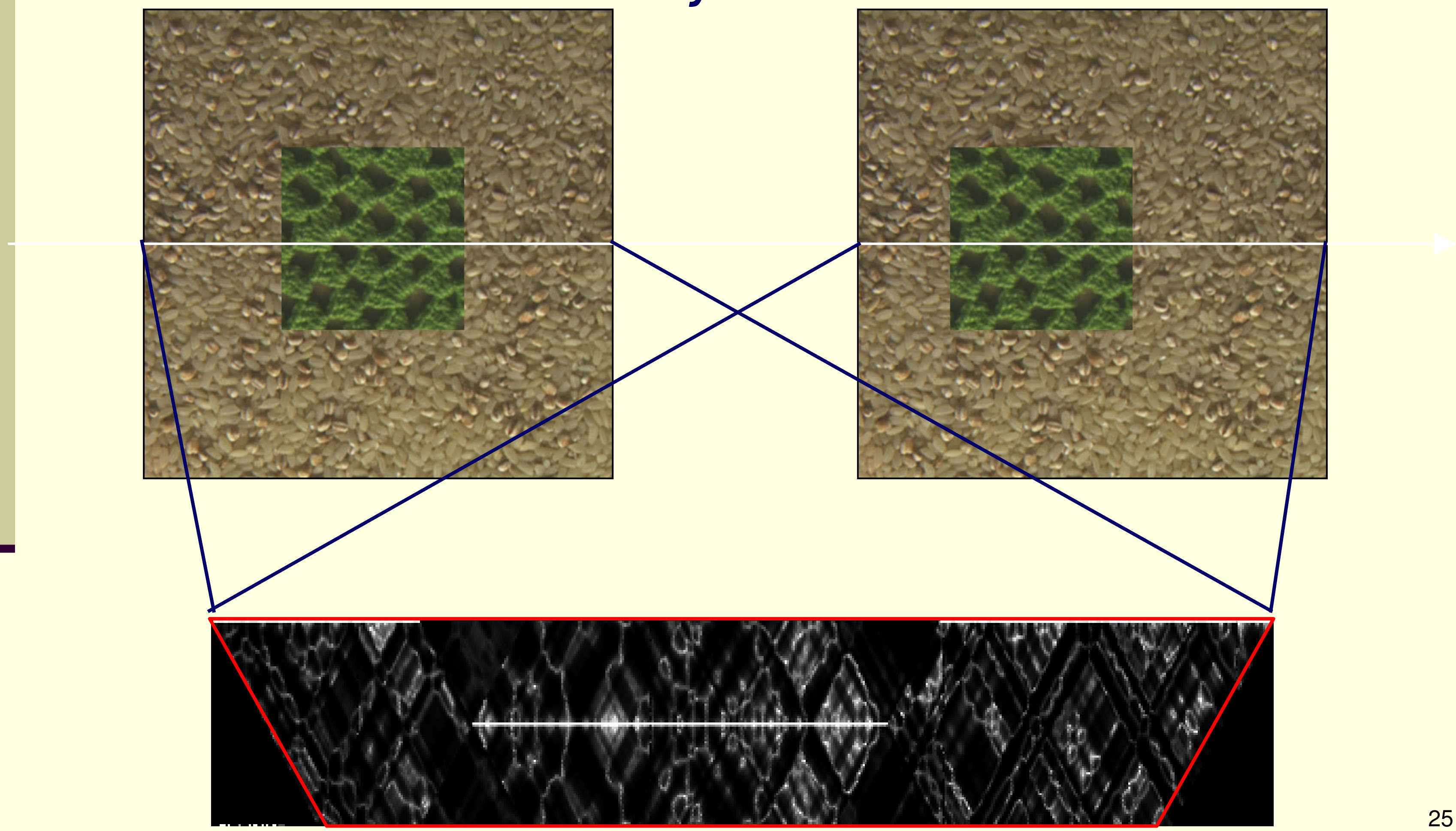
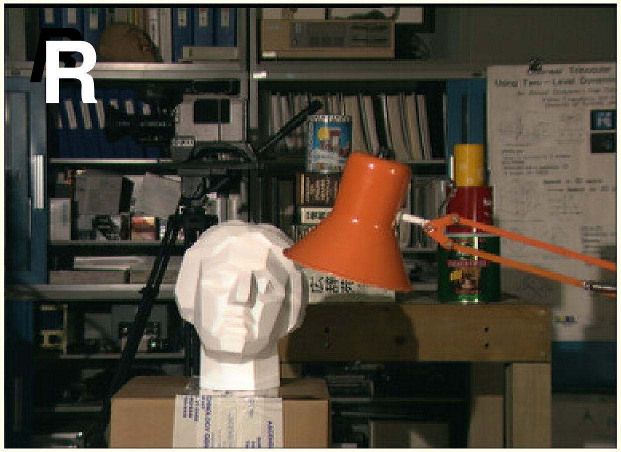
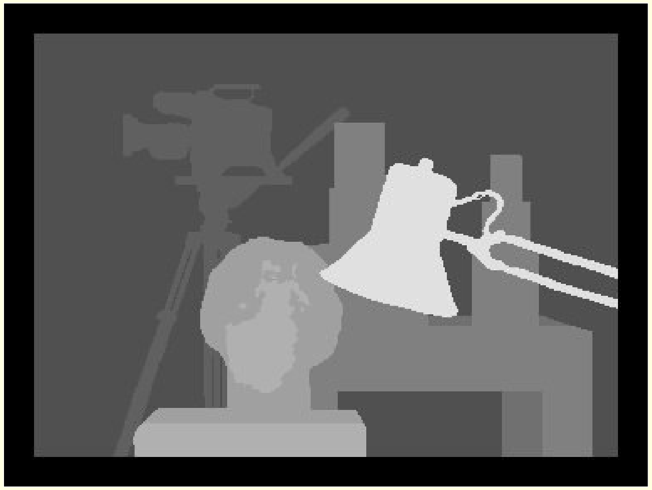


Image data for verification of matching algorithms

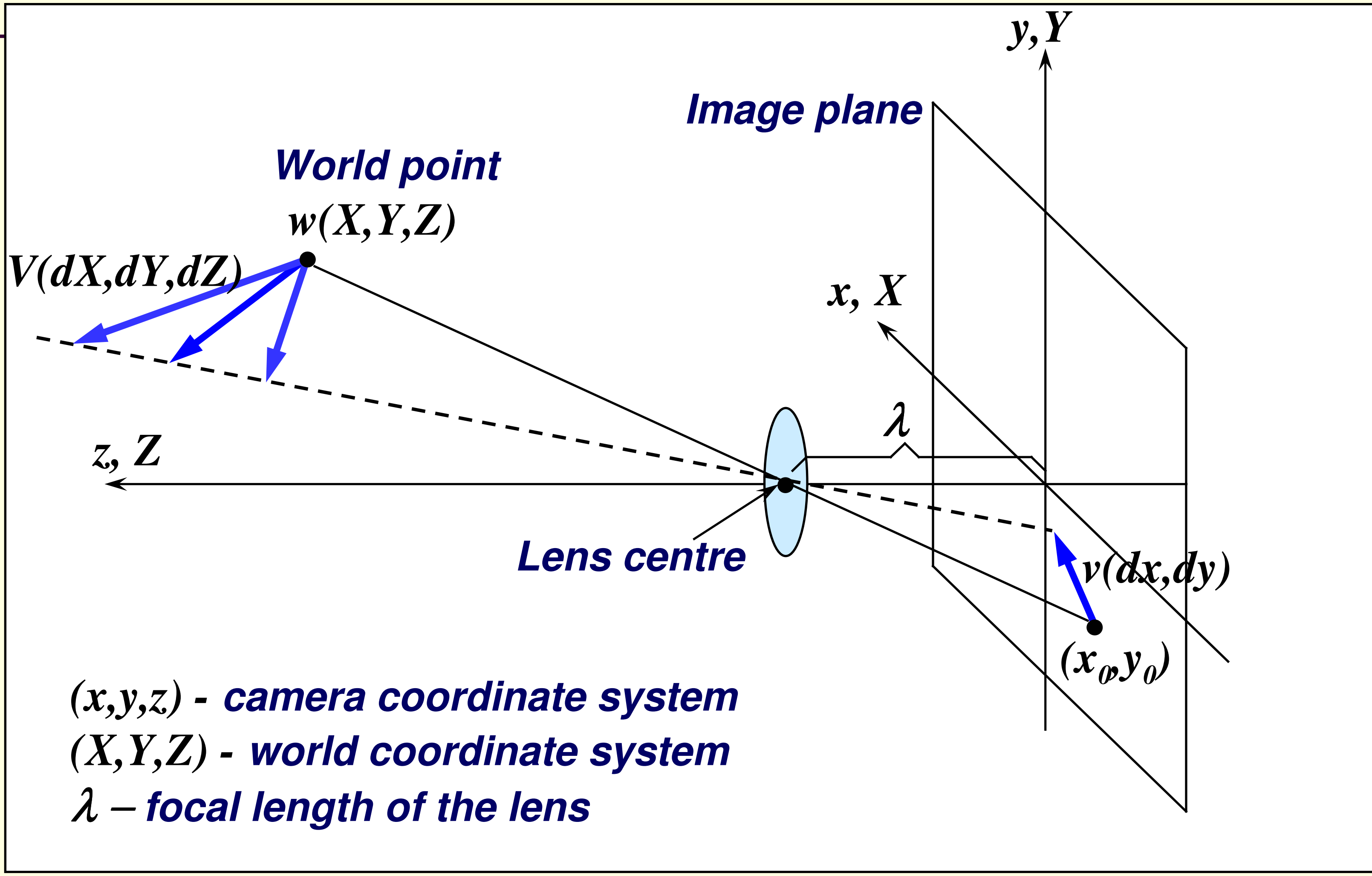




ground truth disparity image

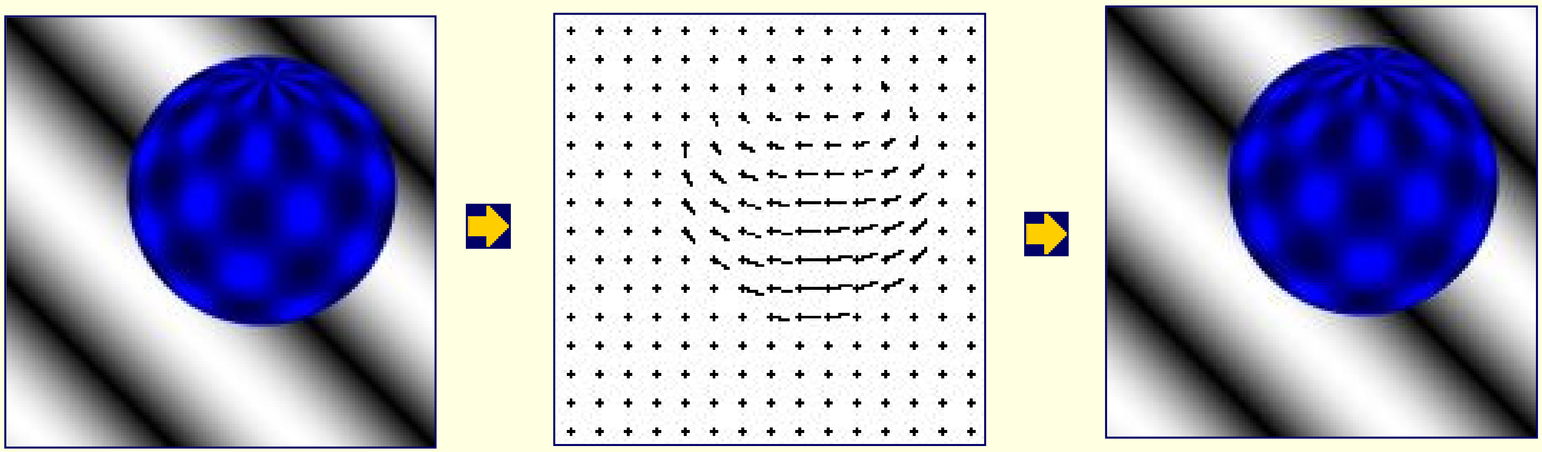
computed disparity image

Optical flow – projection of 3D movement vector



Perspective transformations apply

The concept of optical flow



Images from:
http://www.cs.otago.ac.nz/research/vision/Research/index.html

B.K.P. Horn, B.G. Schunk *„**Determining Optical Flow**”,* Artificial Intelligence, vol. 17, 1981, pp. 185-203