International Conference on Computer Systems and Technologies
CompSysTech'11
16-17 June 2011, Vienna University of Technology, Vienna, Austria

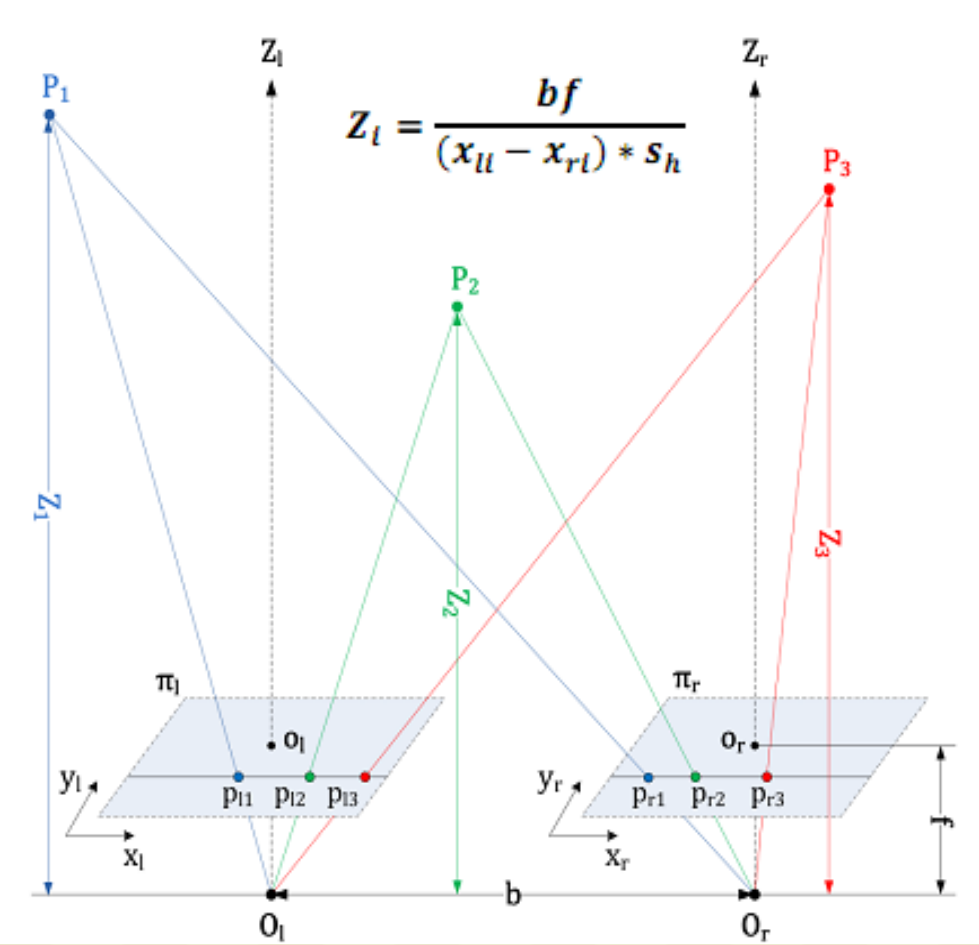
Depth Estimation using Shifted Digital Still Camera

Iva Nikolova, Assist. Prof., TU-Sofia, Bulgaria
Atanas Nikolov, Ph.D. student, BAS, Bulgaria
Georgi Zapryanov, Assist. Prof., TU-Sofia, Bulgaria

Depth Estimation using Shifted Digital Still Camera

- The present work's objective is to investigate the possibilities of a simple method for acquiring the depth using the principles of a canonical stereo vision system.
- The aim is to prove by physical experiments, using conventional digital still camera, when a real stereovision system is not available, is possible to effectively determine the z coordinates to particular object points in a given static scene.
- The main requirement is that the camera should have precise horizontal movement, high resolution and possibilities for adjusting the parameters of its optical system.

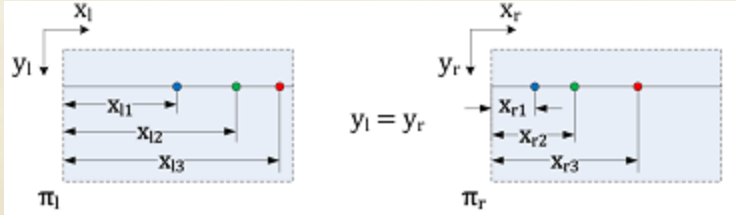
Geometrical Model of a Canonical Stereo Configuration



Geometrical model of a canonical stereo configuration

Parameters of a CSC:

(1) O_l and O_r - optical centers of the cameras; ***(2) Z_l and Z_r*** - parallel principal axes that pass through points O_l and O_r which are perpendicular to the image planes π_l and π_r; ***(3) base length, b*** - the distance between the points O_l and O_r; ***(4) focal length, f*** - the distance from the image planes to the central points O_l and O_r; ***(5) p_{li} and p_{ri}*** - the projections of an arbitrary point P_i within the scene on the image planes π_l and π_r; ***(6) x_{li} and x_{ri}*** - distances, measured from the top left corner of the images to the corresponding projections of a given point P_i (p_{li} and p_{ri}); ***(7) Z_i*** - the distance between the line connecting the optical centers of the cameras and the scene point P_i.



Depth Estimation Algorithm

The main problem in determining the distance to objects in a scene from a pair of stereo images, obtained by a canonical stereo configuration, is to find pairs of corresponding image points, which represent projection of one point from a 3D scene.

Step 1: Features detection in the stereo images by cornerness measure, proposed by Alison Noble, on the basis of the Harris corner detector.

Step 2: Searching for matches by the method of normalized cross correlation between the previously found feature points (corners).

Step 3: Depth estimation using the relationship between disparity, base length and focal length, obtained on the basis of the geometrical model of a canonical stereo configuration.

Corners Detection

Harris cornerness measure

$$C(x,y) = \det(M) - k\left(trace(M)\right)^2$$

where $k = 0.04 \div 0.06$

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix},\ \det(M) = AB - C^2,\ trace(M) = A + B,\ A = I_x^2 \otimes w,\ B = I_y^2 \otimes w,\ C = I_x I_y \otimes w$$

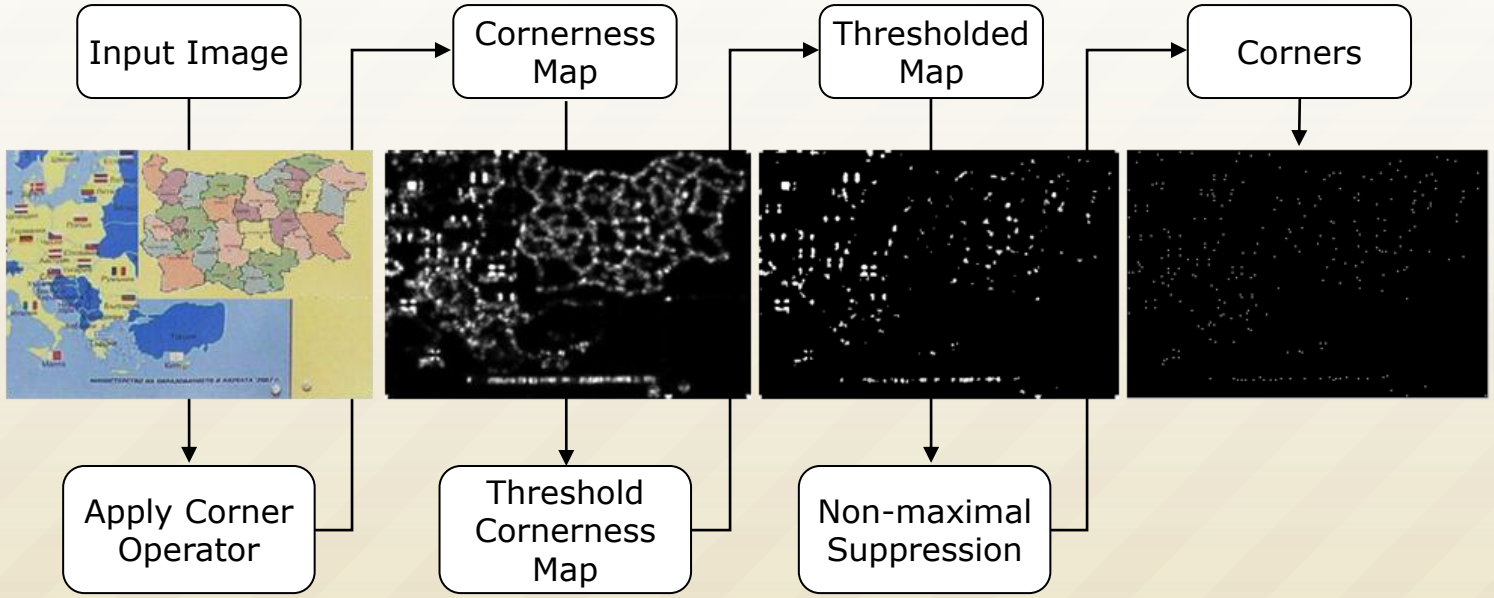
$\otimes$ *is the convolution opertor,*

w is the Gaussian window

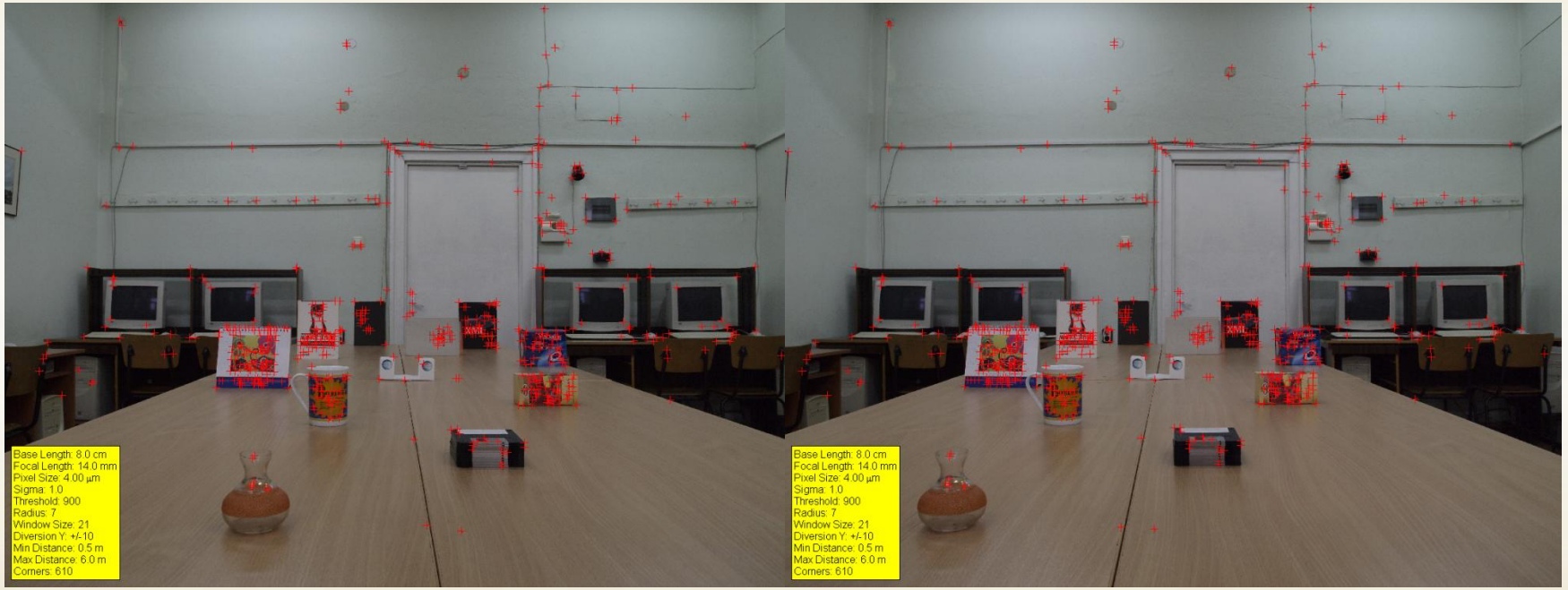
$$I_x = I \otimes (-1,0,1) \approx \frac{dI}{dx} \quad and \quad I_y = I \otimes (-1,0,1)^T \approx \frac{dI}{dy}$$

A. Noble cornerness measure

$$C(x,y) = \det(M)/trace(M)$$



An Example of Found Corners in a Pair of Stereo Images

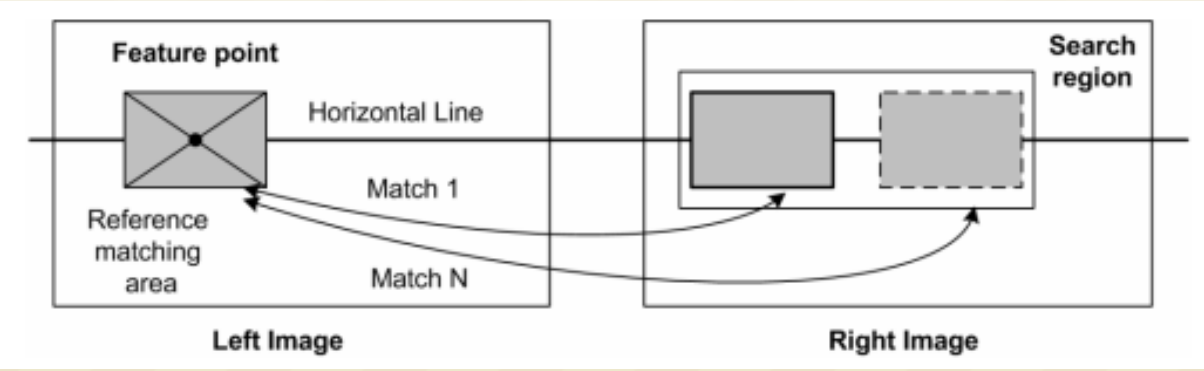


Searching for Matches

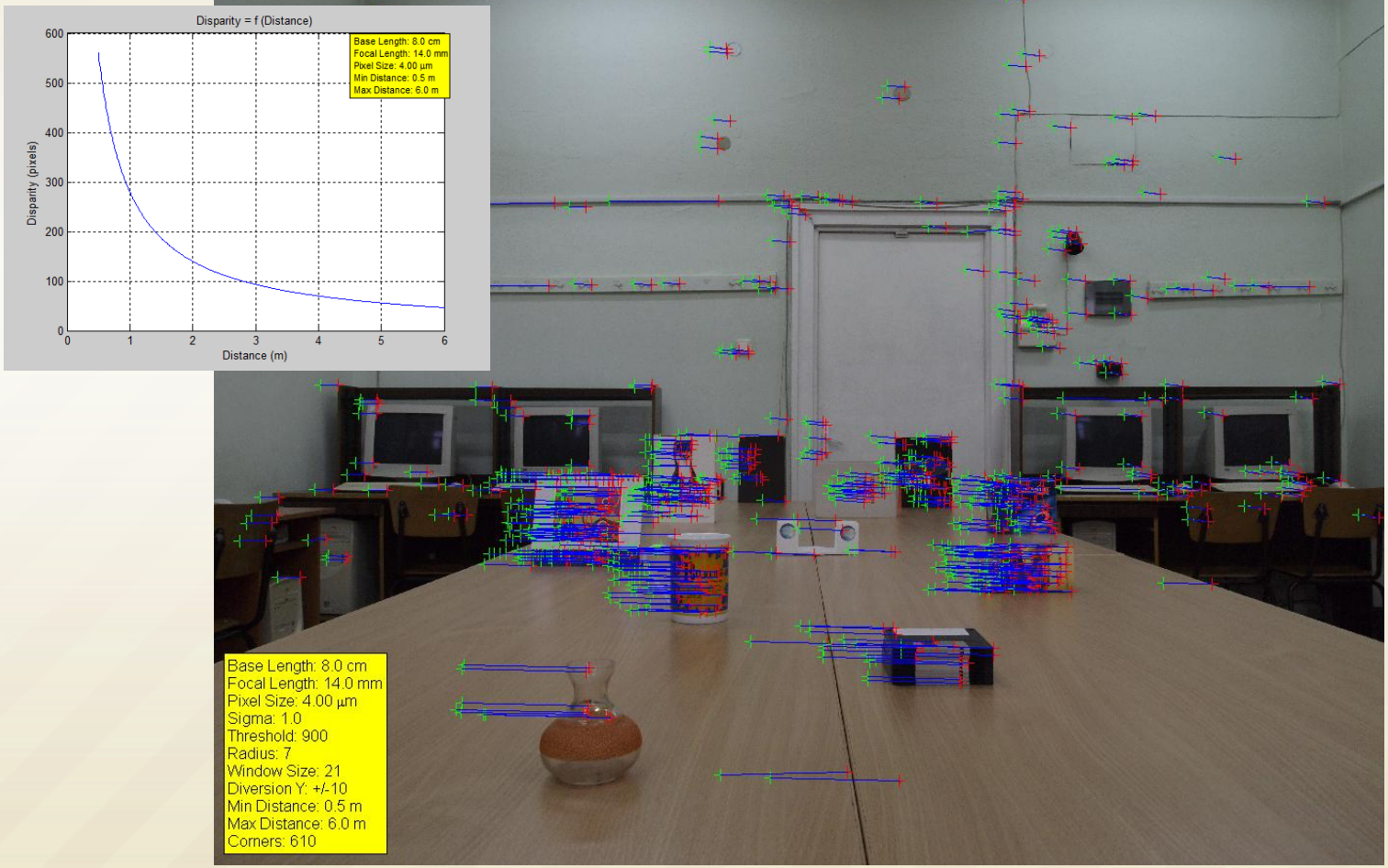
The metric used to determine which features in the stereo pair correspond to each other is based on Normalized Cross-Correlation:

$$NCC = \frac{\sum_{(i,j)\in U} I_1(x+i, y+j)\,.\; I_2\left(x+d_x+i, y+d_y+j\right)}{\sqrt{\sum_{(i,j)\in U} I_1(x+i, y+j)^2 .\; \sum_{(i,j)\in U} I_2\left(x+d_x+i, y+d_y+j\right)^2}}$$

where: $\boldsymbol{I_1}$ and $\boldsymbol{I_2}$ are two image regions being compared. The region $\boldsymbol{I_1}$ is built around a reference point $\boldsymbol{(x\,,\,y)}$, and the region $\boldsymbol{I_2}$ - around point $\boldsymbol{(x+d_x\,,\,y+d_y)}$, where with $\boldsymbol{d_x}$ and $\boldsymbol{d_y}$ are denoted the relative horizontal and vertical displacements of the two image blocks being compared. The matching regions are defined by a set $\boldsymbol{U}$ of offset values, measured from their reference points, i.e. $\boldsymbol{(x\,,\,y)}$ and $\boldsymbol{(x+d_x\,,\,y+d_y)}$, respectively.



An Example of Found Correspondences Between Previously Detected Corners

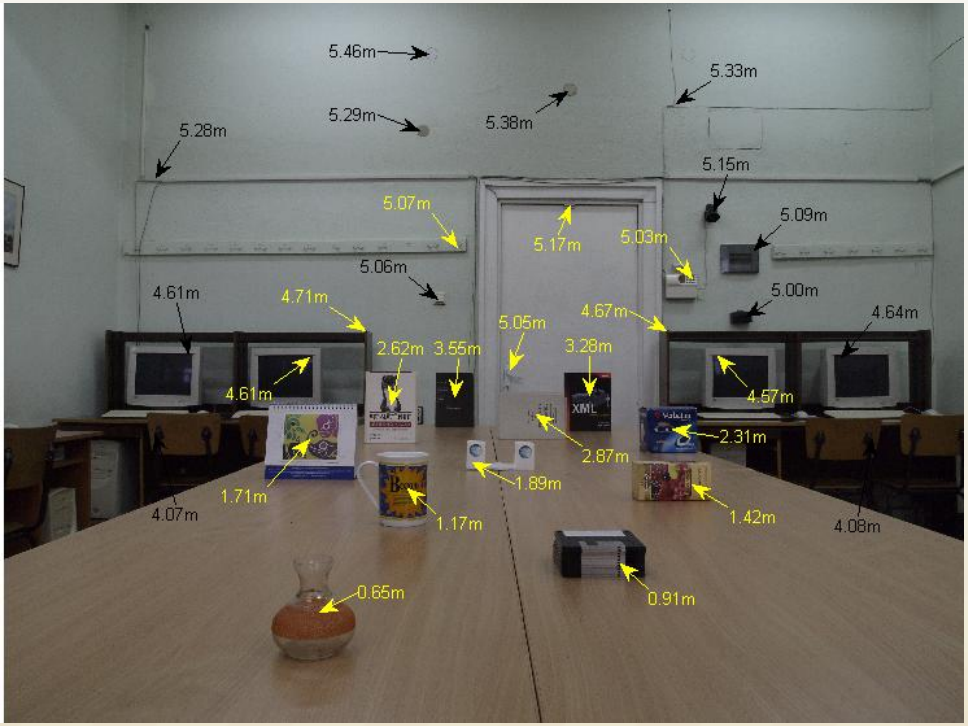


Experimental Results

Our experimental work has two goals:

1) to verify the applicability of the mathematical model, using a real camera system;

2) to test the accuracy of the estimated distances in a real scene.

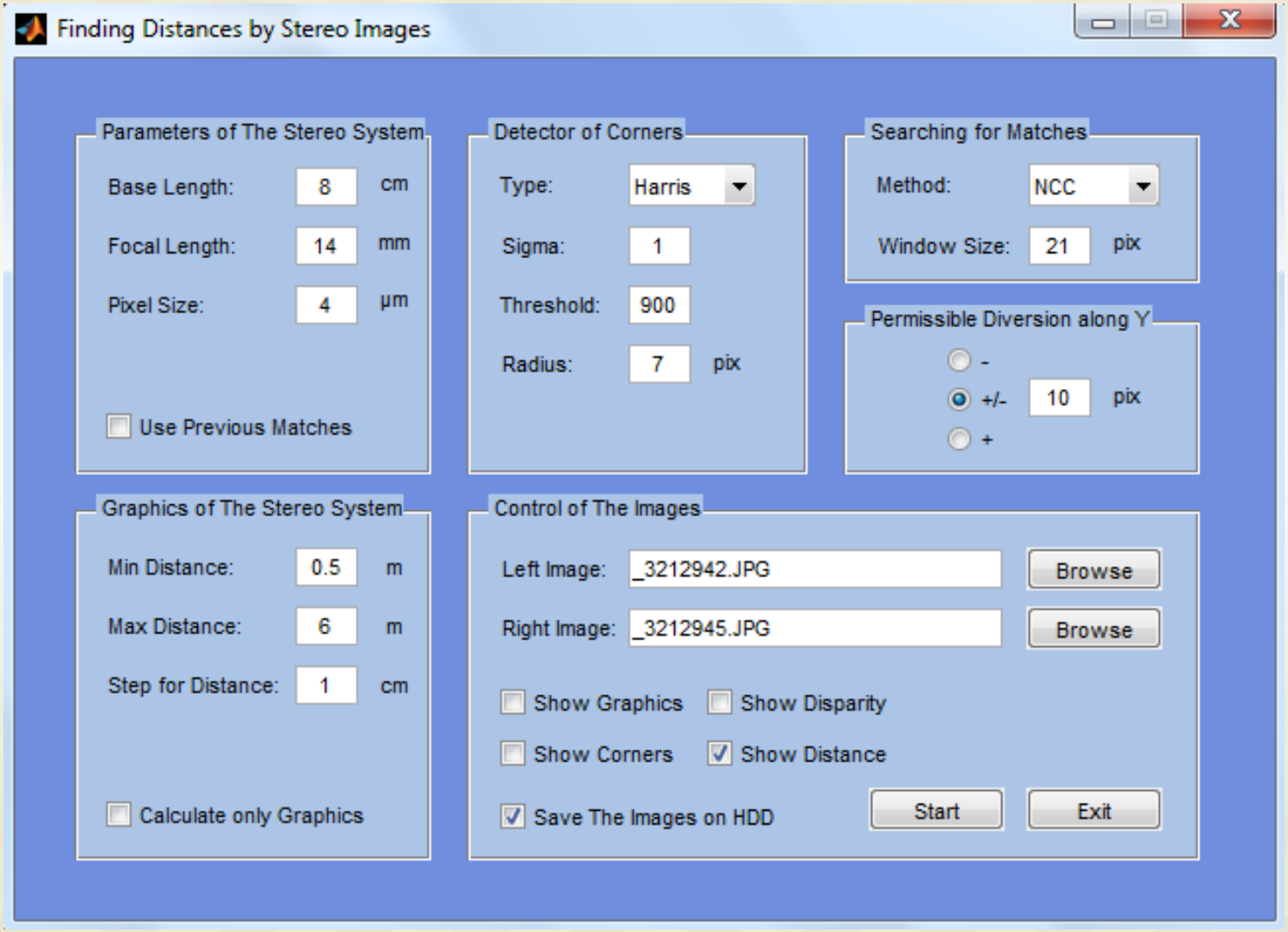


Real Distances to selected objects

Experimental Platform

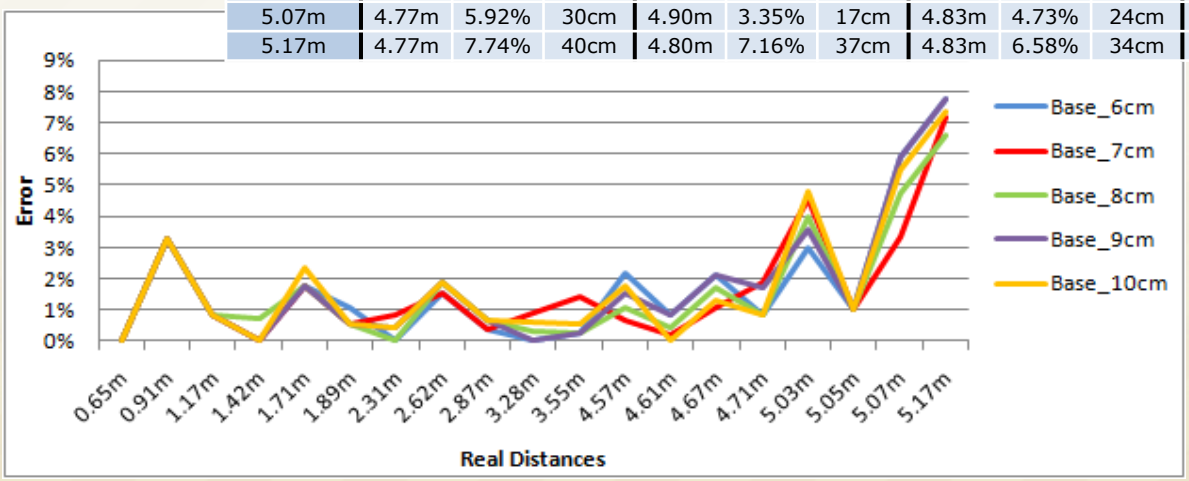
Laser distancemeter

Interface of the Software application



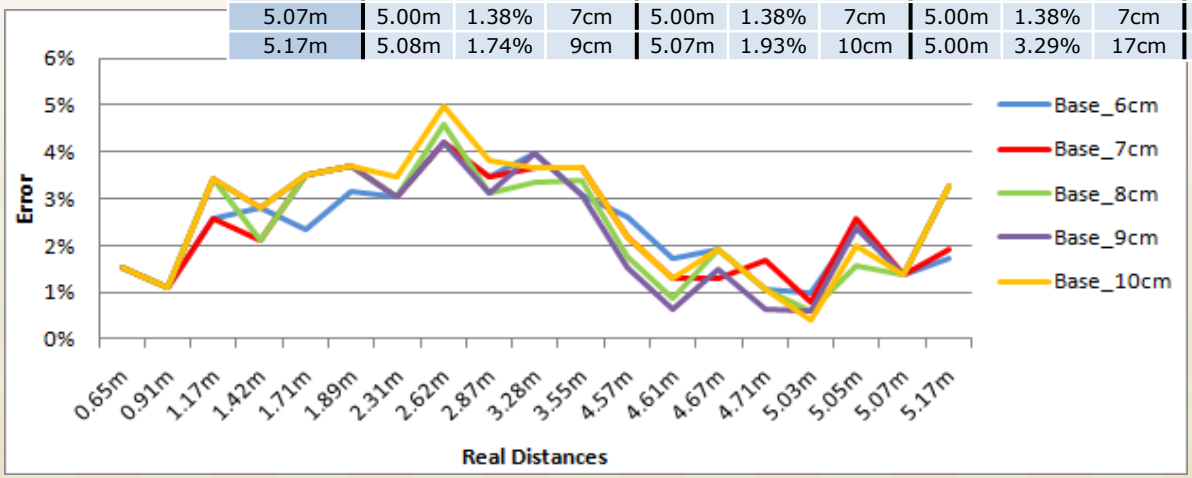
Experimental Results (f=14mm)

Real Distance	Base length between cameras with focal length 14 mm														
	6cm			7cm			8cm			9cm			10cm		
	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]
0.65m	0.65m	0.00%	0cm	0.65m	0.00%	0cm	0.65m	0.00%	0cm	0.65m	0.00%	0cm	0.65m	0.00%	0cm
0.91m	0.88m	3.30%	3cm	0.88m	3.30%	3cm	0.88m	3.30%	3cm	0.88m	3.30%	3cm	0.88m	3.30%	3cm
1.17m	1.18m	0.85%	1cm	1.18m	0.85%	1cm	1.18m	0.85%	1cm	1.18m	0.85%	1cm	1.18m	0.85%	1cm
1.42m	1.42m	0.00%	0cm	1.42m	0.00%	0cm	1.41m	0.70%	1cm	1.42m	0.00%	0cm	1.42m	0.00%	0cm
1.71m	1.74m	1.75%	3cm	1.74m	1.75%	3cm	1.74m	1.75%	3cm	1.74m	1.75%	3cm	1.75m	2.34%	4cm
1.89m	1.91m	1.06%	2cm	1.90m	0.53%	1cm	1.90m	0.53%	1cm	1.90m	0.53%	1cm	1.90m	0.53%	1cm
2.31m	2.31m	0.00%	0cm	2.33m	0.87%	2cm	2.31m	0.00%	0cm	2.32m	0.43%	1cm	2.30m	0.43%	1cm
2.62m	2.66m	1.53%	4cm	2.66m	1.53%	4cm	2.67m	1.91%	5cm	2.67m	1.91%	5cm	2.67m	1.91%	5cm
2.87m	2.88m	0.35%	1cm	2.88m	0.35%	1cm	2.89m	0.70%	2cm	2.89m	0.70%	2cm	2.89m	0.70%	2cm
3.28m	3.28m	0.00%	0cm	3.31m	0.91%	3cm	3.29m	0.30%	1cm	3.28m	0.00%	0cm	3.30m	0.61%	2cm
3.55m	3.56m	0.28%	1cm	3.60m	1.41%	5cm	3.54m	0.28%	1cm	3.54m	0.28%	1cm	3.57m	0.56%	2cm
4.57m	4.47m	2.19%	10cm	4.54m	0.66%	3cm	4.52m	1.09%	5cm	4.50m	1.53%	7cm	4.49m	1.75%	8cm
4.61m	4.57m	0.87%	4cm	4.62m	0.22%	1cm	4.59m	0.43%	2cm	4.57m	0.87%	4cm	4.61m	0.00%	0cm
4.67m	4.57m	2.14%	10cm	4.62m	1.07%	5cm	4.59m	1.71%	8cm	4.57m	2.14%	10cm	4.61m	1.28%	6cm
4.71m	4.67m	0.85%	4cm	4.62m	1.91%	9cm	4.67m	0.85%	4cm	4.63m	1.70%	8cm	4.67m	0.85%	4cm
5.03m	4.88m	2.98%	15cm	4.80m	4.57%	23cm	4.83m	3.98%	20cm	4.85m	3.58%	18cm	4.79m	4.77%	24cm
5.05m	5.00m	0.99%	5cm	5.00m	0.99%	5cm	5.00m	0.99%	5cm	5.00m	0.99%	5cm	5.00m	0.99%	5cm
5.07m	4.77m	5.92%	30cm	4.90m	3.35%	17cm	4.83m	4.73%	24cm	4.77m	5.92%	30cm	4.79m	5.52%	28cm
5.17m	4.77m	7.74%	40cm	4.80m	7.16%	37cm	4.83m	6.58%	34cm	4.77m	7.74%	40cm	4.79m	7.35%	38cm

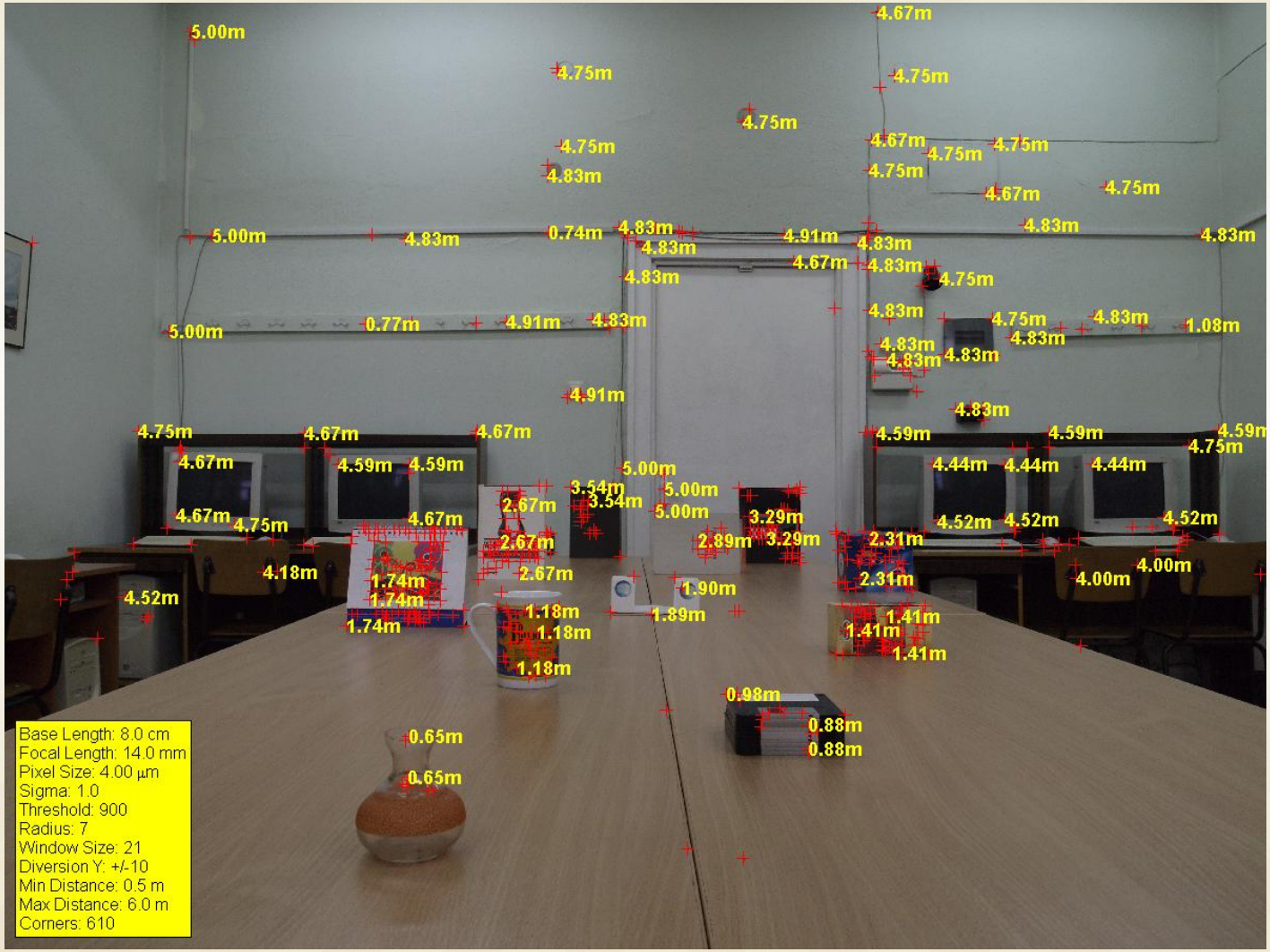


Experimental Results (f=20mm)

Real Distance	Base length between cameras with focal length 20 mm														
	6cm			7cm			8cm			9cm			10cm		
	Est. [m]	Err [%]	Err. [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]	Est. [m]	Err [%]	Err [cm]
0.65m	0.66m	1.54%	1cm	0.66m	1.54%	1cm	0.66m	1.54%	1cm	0.66m	1.54%	1cm	0.66m	1.54%	1cm
0.91m	0.90m	1.10%	1cm	0.90m	1.10%	1cm	0.90m	1.10%	1cm	0.90m	1.10%	1cm	0.90m	1.10%	1cm
1.17m	1.20m	2.56%	3cm	1.20m	2.56%	3cm	1.21m	3.42%	4cm	1.21m	3.42%	4cm	1.21m	3.42%	4cm
1.42m	1.46m	2.82%	4cm	1.45m	2.11%	3cm	1.45m	2.11%	3cm	1.46m	2.82%	4cm	1.46m	2.82%	4cm
1.71m	1.75m	2.34%	4cm	1.77m	3.51%	6cm	1.77m	3.51%	6cm	1.77m	3.51%	6cm	1.77m	3.51%	6cm
1.89m	1.95m	3.17%	6cm	1.96m	3.70%	7cm	1.96m	3.70%	7cm	1.96m	3.70%	7cm	1.96m	3.70%	7cm
2.31m	2.38m	3.03%	7cm	2.38m	3.03%	7cm	2.38m	3.03%	7cm	2.38m	3.03%	7cm	2.39m	3.46%	8cm
2.62m	2.73m	4.20%	11cm	2.73m	4.20%	11cm	2.74m	4.58%	12cm	2.73m	4.20%	11cm	2.75m	4.96%	13cm
2.87m	2.97m	3.48%	10cm	2.97m	3.48%	10cm	2.96m	3.14%	9cm	2.96m	3.14%	9cm	2.98m	3.83%	11cm
3.28m	3.41m	3.96%	13cm	3.40m	3.66%	12cm	3.39m	3.35%	11cm	3.41m	3.96%	13cm	3.40m	3.66%	12cm
3.55m	3.66m	3.10%	11cm	3.68m	3.66%	13cm	3.67m	3.38%	12cm	3.66m	3.10%	11cm	3.68m	3.66%	13cm
4.57m	4.69m	2.63%	12cm	4.67m	2.19%	10cm	4.65m	1.75%	8cm	4.64m	1.53%	7cm	4.67m	2.19%	10cm
4.61m	4.69m	1.74%	8cm	4.67m	1.30%	6cm	4.65m	0.87%	4cm	4.64m	0.65%	3cm	4.67m	1.30%	6cm
4.67m	4.76m	1.93%	9cm	4.73m	1.28%	6cm	4.76m	1.93%	9cm	4.74m	1.50%	7cm	4.76m	1.93%	9cm
4.71m	4.76m	1.06%	5cm	4.79m	1.70%	8cm	4.76m	1.06%	5cm	4.74m	0.64%	3cm	4.76m	1.06%	5cm
5.03m	5.08m	0.99%	5cm	5.07m	0.80%	4cm	5.06m	0.60%	3cm	5.00m	0.60%	3cm	5.05m	0.40%	2cm
5.05m	5.17m	2.38%	12cm	5.18m	2.57%	13cm	5.13m	1.58%	8cm	5.17m	2.38%	12cm	5.15m	1.98%	10cm
5.07m	5.00m	1.38%	7cm	5.00m	1.38%	7cm	5.00m	1.38%	7cm	5.00m	1.38%	7cm	5.00m	1.38%	7cm
5.17m	5.08m	1.74%	9cm	5.07m	1.93%	10cm	5.00m	3.29%	17cm	5.00m	3.29%	17cm	5.00m	3.29%	17cm



Depth Estimation to Definite Points



Conclusion and Future Work

- The accuracy of the investigated method for distance determining to objects in a static scene by translation of one camera can be viewed from physical and algorithmic point:
  - For the physical accuracy improvement a subjective adjustments (like choice of optimal base and focal length, knowledge for optical distortions) need to be applied, in order to calibrate the stereovision system, according to the working distance range.
  - The algorithmic accuracy aspect depends on the software methods for determining the feature points (corners) and their correspondence.

- The future development of the investigated method can be focused on:
  - ✓ Each point from the stereoimages to be viewed as a characteristic point and its correspondence to be found in the other image;
  - ✓ Determining the 3D coordinates (X, Y, Z) of a random point;
  - ✓ The distance determining algorithm should not require the usage of the canonical stereovision system model;
  - ✓3D modeling and 3D object recognition.

THANK YOU FOR THE ATTENTION!

QUESTIONS?