

# Visual Control of an Industrial Robot Manipulator: Accuracy Estimation

Gregor Papa\* - Drago Torkar  
Computer Systems Department, Jožef Stefan Institute, Slovenia

*To control robots by vision is a contemporary challenge in many industrial applications. The capabilities of a robot vision system are highly significant for planning any robotic task and must, therefore, be precisely established. In this paper we describe the procedure to estimate a static and dynamic accuracy of a robot stereo vision system consisted of two identical 1 Megapixel cameras. The accuracy was evaluated in 2D and 3D environment. We describe the methodology, the test setup, and the results of the evaluation.*

©2009 Journal of Mechanical Engineering. All rights reserved.

**Keywords:** robot vision, calibrated visual servoing

## 0 INTRODUCTION

In past decades, an increasing number of robots applications was used in industrial manufacturing accompanied by an increased demand for versatility, robustness and precision. The demand was mostly satisfied by increasing the mechanical capabilities of robot parts. For instance, to meet the micrometric positioning requirements, stiffness of the robot's arms was increased, high precision gears and low backlash joints introduced, which often led to difficult design compromises such as the request to reduce inertia and increase stiffness. This results in approaching the mechanical limits and increased the cost of robots decreasing the competitiveness of the robot systems on the market [1].

Lately, the robot producers have been putting many efforts in incorporating visual and other sensors into industrial robots, thus making a significant improvement in accuracy, flexibility and adaptability. The vision is still probably the most promising sensor [2] in real robotic 3D servoing issues [3]. It has been vastly investigated for the last two decades in the laboratories but it is only now that it has found its way in industrial implementation [4] in contrast to machine vision, which became a well established industry during the last years [5]. The vision systems used in robots must satisfy a few constraints that makes them different from machine vision measuring systems. Firstly, camera working distances are much larger, especially with larger robots that can reach a few meters. Measuring with high precision at such distances demands much higher resolution which consequently, increases the cost

of the vision system and the application very soon reaches its technological and cost limits. The dynamics of the industrial processes requires high frame rates which, in connection with real time processing, puts another difficult constraint to system integrators.

It is very important to be familiar with the precise capabilities of the vision system as a whole to be able to produce cost-effective robot vision applications. Accuracy is one of the most important parameters and vision system accuracy, especially when multiple vision sensors are involved can not be precisely estimated from the specifications of the components. In this paper we present a procedure for estimating the accuracy of a stereo vision system using an array of 10 infrared light emitting diodes (IR-LED) placed on a plate in a regular pattern.

## 1 METHODOLOGY

Four types of accuracy tests were performed: static and dynamic 2D test, and static and dynamic 3D test. With all the tests, the IR-LED array was used to establish its suitability for being used as a marker and a calibration pattern in robot visual servoing applications.

Within the static 2D test, we were moving the IR-LED array with the linear drive perpendicular to the camera optical axes and we measured the increments in the image. The purpose was to detect the smallest linear response in the image. The IR-LED centroids were determined in two ways: on binary images and on grey-level images as centers of mass. During the image grabbing the array did not move, thus

\*Corr. Author's Address: Computer Systems Department, Jožef Stefan Institute, Jamova c. 39, SI-1000 Ljubljana, Slovenia, gregor.papa@ijs.si

eliminating the dynamic effects. We averaged the movement of centroids of 10 IR-LEDs in a sequence of 16 images and calculated the standard deviation to get the idea of accuracy confidence intervals. With the dynamic 2D test, we investigated shape distortions in the images due to fast 2D movements of linear drive. We compared a few images of IR-LED array taken during movement to the statically obtained ones, which provided information of photocenter displacements and an estimation of dynamic error.

We performed the 3D accuracy evaluation with 2 fully calibrated cameras in a stereo setup. Using the linear drive again, the array of IR-LEDs was moved along the 3D line with different increments and the smallest movement producing a linear response in reconstructed 3D space was looked for. For the 3D dynamic test, we attached the IR-LED array to the wrist of an industrial robot, and dynamically guided it through some predefined points in space and simultaneously recorded the trajectory with fully calibrated stereo cameras. Then, we compared the reconstructed 3D points from images to these predefined points fed to robot controller.



Fig. 1. 1 Megapixel camera with IR filter

## 2 SETUP

The test environment consisted of:

- PhotonFocus MV-D1024-80-CL-8 CMOS camera; framerate: 75 fps at full resolution (1024x1024 pixels) (Fig. 1),

- Active Silicon Phoenix-DIG48 PCI frame grabber,
- Moving IR-LED array (mounted on linear guide in Fig. 2 and mounted on robot TCP in Fig. 3). The standard deviation of IR-LEDs accuracy is below 0.007 pixel [6],
- Festo linear guide (DGE-25-550-SP) with repetition accuracy of  $\pm 0.02$  mm (Fig. 2),
- ABB industrial robot IRB 140 with  $\pm 0.03$  mm repeatability, and linear path accuracy  $\pm 1.0$  mm (Fig. 3).

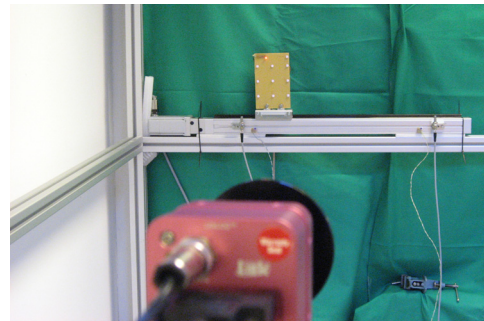


Fig. 2. Linear drive with IR-LED array

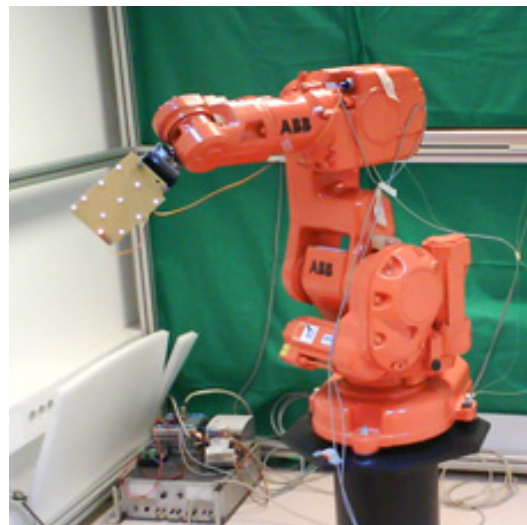


Fig. 3. ABB industrial robot IRB 140

For the static 2D test: the distance from the camera to a moving object (in the middle position) that moves perpendicularly to the optical axis was 195 cm; the camera field-of-view was  $220 \times 220$  cm, which gives the pixel size of 2.148 mm; Schneider-Kreuznach lens CINEGON 10 mm/1.9 F with IR filter; exposure time was

10.73 ms, while frame time was 24.04 ms. For the dynamic 2D test the conditions were the same as with the static test, except the linear guide was moving the IR-LED array with the speed of 460 mm/s.

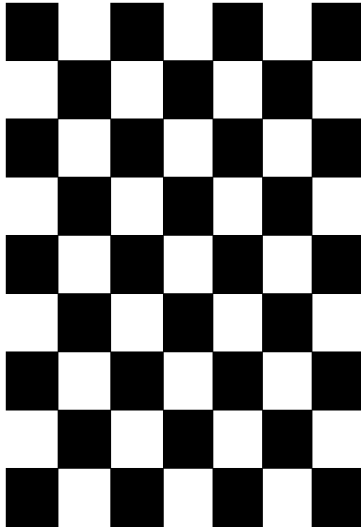


Fig. 4. Camera calibration pattern

For the 3D reconstruction test: left camera distance to IR-LED array and right camera distance to IR-LED array were about 205 cm;

baseline distance was 123 cm; Schneider-Kreuznach lens CINEGON 10 mm/1.9 F with IR filter; Calibration region-of-interest (ROI): 342 pixel  $\times$  333 pixel; Calibration pattern: 7  $\times$  9 black/white squares (Fig. 4); Calibration method [7]; Reconstruction method [8]. The reconstruction was done off-line and the stereo correspondence problem was considered solved due to the simple geometry of the IR-LED array and is thus, not addressed here. For 3D dynamic test, an ABB industrial robot IRB 140 was used with the standalone fully calibrated stereo vision setup placed about 2 m away from its base and calibrated the same way as before (Fig. 5). The robot wrist was moving through the corners of an imaginary triangle with the side length of approximately 12 cm. The images were taken dynamically when TCP was passing the corner points with the approximate speed of 500 mm/s and reconstructed in 3D. The relative length of such triangle sides were compared to the sides of a statically-obtained and reconstructed triangle.

### 3 RESULTS

#### 3.1 2D accuracy tests

Below are the results of the evaluation. Tests include the binary and grey-level centroids.

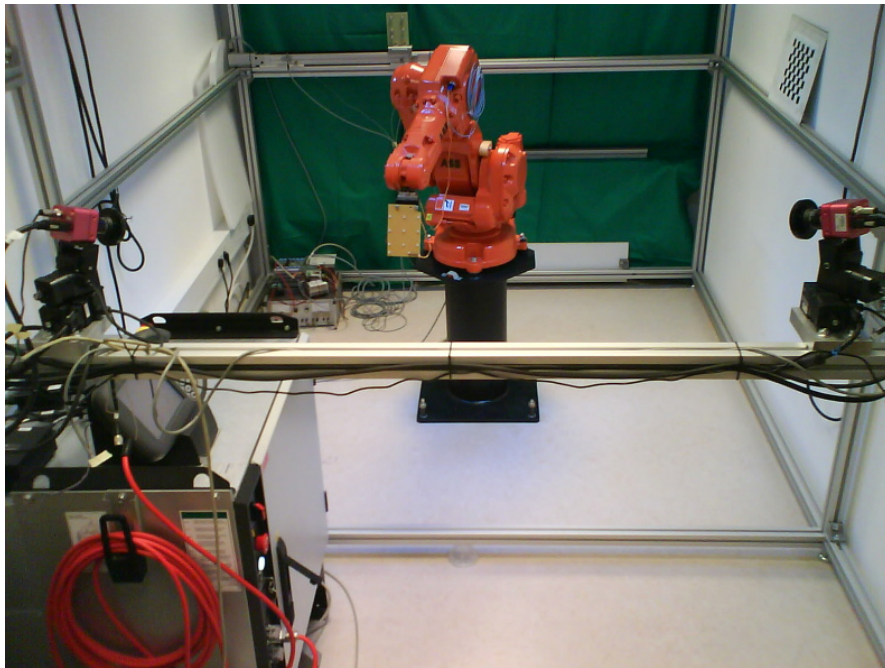


Fig. 5. Test setup for robot stereo vision 2D and 3D accuracy tests

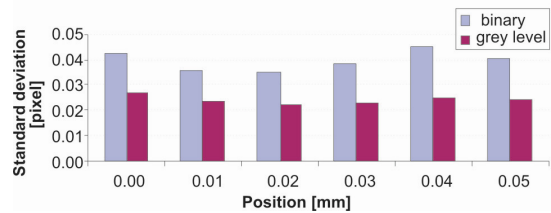
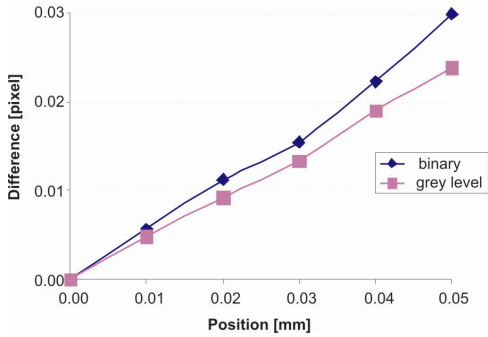
For each movement increments (0.01, 0.1, and 1 mm) the two graphs are presented (Figs. 6a, b, and c):

- Pixel difference between the starting image and the consecutive images – for each position the value is calculated as the average move of all 10 markers, while their position is calculated as the average position in the sequence of the 16 images grabbed at each position in static conditions.

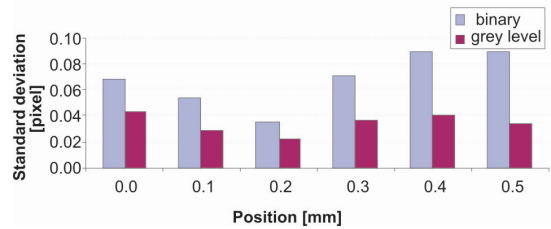
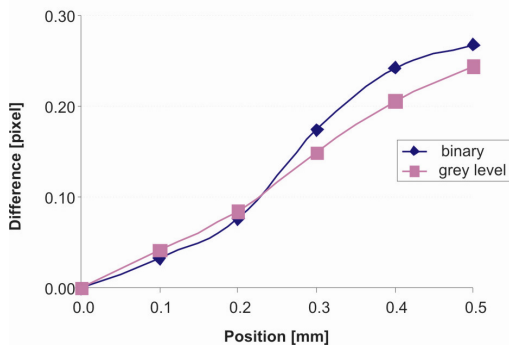
- Standard deviation of the centre positions of all markers regarding their move according to the first image.

There is an additional Fig. 7 to compare normalized movement increments: pixel differences of the single marker when working with binary images, and grey-level images.

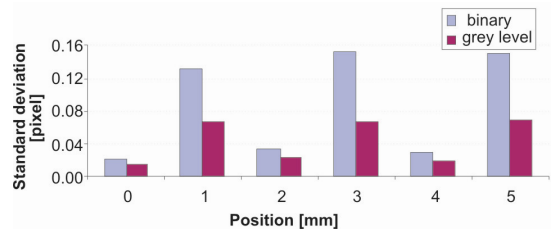
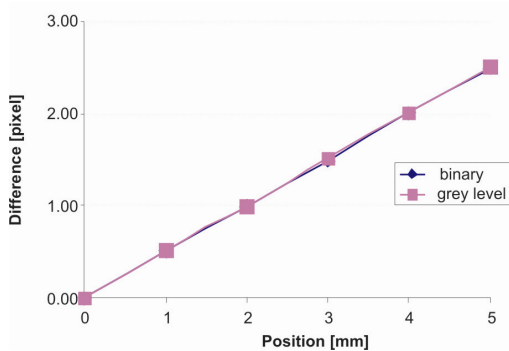
We applied a linear regression model to measured data, and we calculated the  $R^2$  values to asses the fitting quality.



a)



b)



c)

Fig. 6. Pixel difference (left) and standard deviation (right).  
 Increments: a) 0.01 mm, b) 0.1 mm, c) 1 mm

The results are presented in Table 1 for 2D tests and in Table 3 for 3D tests. The  $R^2$  value can be interpreted as the proportion of the variance in  $y$  attributable to the variance in  $x$  (see Eq. 1), where value 1 stands for perfect matching (fitting) and a lower value denotes some deviations.

$$R^2 = \left( \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \right)^2 \quad (1)$$

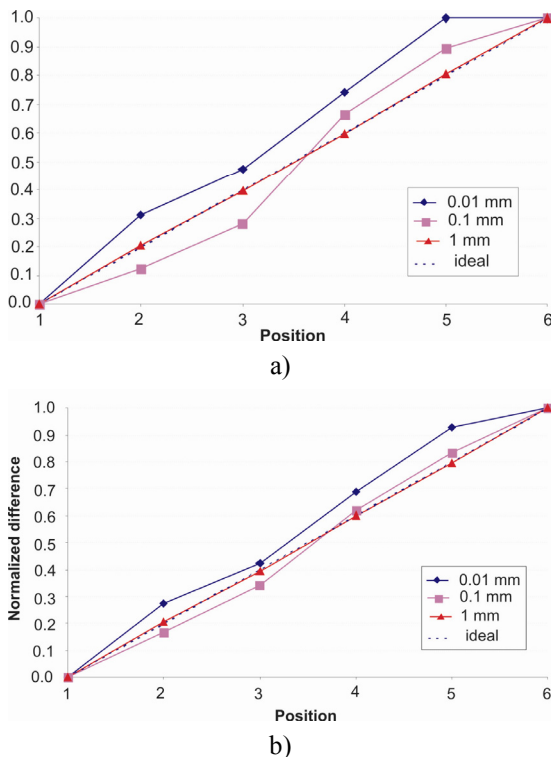


Fig. 7. Normalized differences of a) binary and b) grey-level images for each position comparing different increments

The dynamic 2D test showed that when comparing the centers of the markers of the IR-LED array and the pixel areas of each marker in statically and dynamically (linear guide moving at full speed) grabbed images, there is a difference in center positions. Furthermore, the areas of markers in dynamically grabbed images are slightly wider than those of statically grabbed images.

Table 2 presents the differences of the centers of the markers, and the difference in sizes of the markers of the statically and dynamically grabbed images. Fig. 8 compares the statically

Considering the  $R^2$  threshold of 0.994 we could detect the increments of the moving object in the range of 1/5 of a pixel. The threshold value was chosen in a way that gives a sufficient approximation of the linear regression model, to ensure the applicable results of the measurements. and dynamically grabbed image, and the merged image presents the pixel differences of the marker taken at the same position. Here, the outlined parts of the image represent the common pixels of both images, while the black part on the left side of each marker belongs to the dynamic image only.

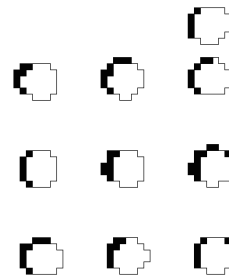


Fig. 8. Comparison of static and dynamic image. Common part outlined, while the black part belongs to dynamic image only

Following the results presented in Table 2, the accuracy of the position in direction  $x$  of dynamically grabbed images compared to the statically grabbed is in the range of 1/3 of a pixel, due to the gravity centre shift of the marker pixel area during the movement of the linear guide.

Table 1. Comparison of standard deviations and  $R^2$  values for different moving increments in 2D

increments [mm]	standard deviation [mm]		$R^2$	
	binary	grey-level	binary	grey-level
0.01	0.045	0.027	0.9624	0.9825
0.1	0.090	0.042	0.9702	0.9937
1	0.152	0.069	0.9998	0.9999

Table 2. Comparison of the images grabbed in static and dynamic mode

	X	Y	width	height	area
static	484.445	437.992	6	6	27
dynamic	484.724	437.640	7	6	32



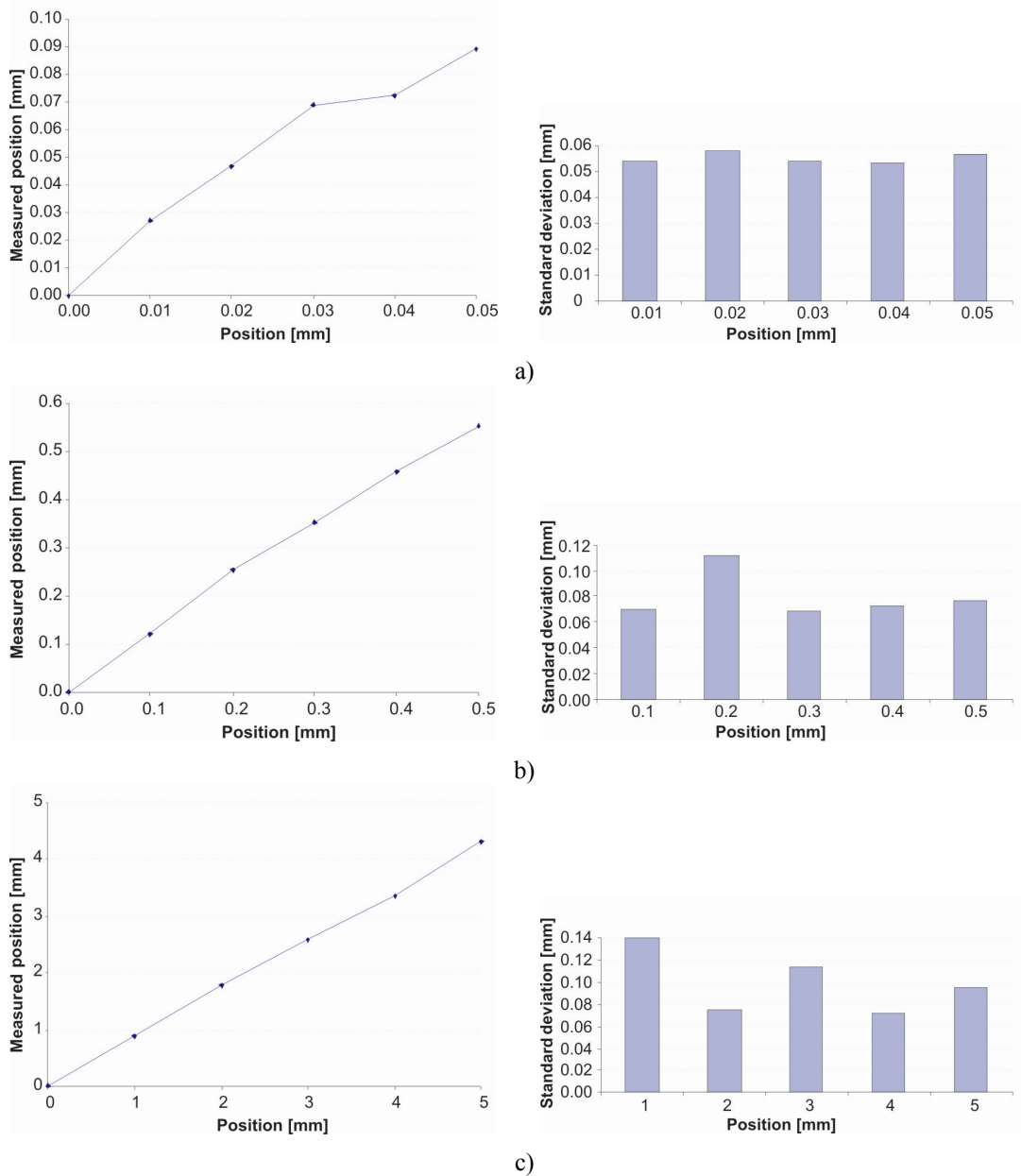


Fig. 9. Pixel difference and standard deviation. Increments: a) 0.01 mm, b) 0.1 mm, and c) 1 mm

Table 3. Comparison of standard deviations and  $R^2$  values for different moving increments in 3D

increments [mm]	standard deviation [mm]	$R^2$
0.01	0.058	0.7806
0.1	0.111	0.9315
1	0.140	0.9974

### 3.2 3D reconstruction tests

We tested the static relative accuracy of the 3D reconstruction of the IR-LED array movements by linear drive. The test setup consisted of the two calibrated Photonfocus cameras gazing at the IR-LED array attached to the linear drive, which exhibited precise

movements of 0.01, 0.1 and 1 mm (Fig. 9). The mass centre points of 10 LEDs were extracted in 3D after each movement and relative 3D paths were calculated and compared to the linear drive paths. As stated in Fig. 7 and in Table 1, only grey-level images were considered due to better results obtained in 2D tests.

Accuracy in 3D is lower than in 2D due to the calibration and reconstruction errors. According to the tests performed it is approximately 1/2 of a pixel.

Table 4 shows the results of the 3D dynamic tests where the triangle area and the side lengths  $a$ ,  $b$  and  $c$  reconstructed from dynamically-obtained images were compared to the static reconstruction of the same triangles. 10 triangles were compared, each formed by a diode in IR-LED array. The average lengths and the standard deviations are presented.

Table 4. Comparison of static and dynamic triangles (measurements are in mm)

	$\bar{a}$	$\sigma$	$\bar{b}$	$\sigma$	$\bar{c}$	$\sigma$
static	193.04	12.46	89.23	2.77	167.84	12.18
dynamic	193.51	12.43	89.03	2.77	167.52	12.03

We have observed a significant standard deviation (up to 7%) of triangle side lengths, which we ascribe to lens distortions since it is almost identical in the dynamic and in the static case. The images and the reconstruction in dynamic conditions vary only a little in comparison to the static ones.

#### 4 CONCLUSIONS

We performed the 2D and 3D accuracy evaluation of the 3D robot vision system consisting of two identical 1 Megapixel cameras. The measurements showed that the raw static 2D accuracy (without any subpixel processing approaches and lens distortion compensation) is as good as 1/5 of a pixel. However, this is reduced to 1/2 of a pixel when image positions are reconstructed in 3D due to reconstruction errors.

In the dynamic case, the comparison to static conditions showed that no significant error is introduced with moving markers in both, 2D and 3D environment. The accuracy of the dynamic 2D case was reduced to 1/3 of a pixel, while for the 3D case it remained 1/2 of a pixel.

#### 5 ACKNOWLEDGEMENTS

This work was supported by the European 6<sup>th</sup> FP project Adaptive Robots for Flexible Manufacturing Systems (ARFLEX, 2005 to 2008) and the Slovenian Research Agency program Computing structures and systems (2004 to 2008).

#### 6 REFERENCES

- [1] Arflex - European FP6 project. (2009) Adaptive Robots for Flexible Manufacturing Systems. Retrieved on 3. 12. 2009, from <http://www.arflexproject.eu>.
- [2] Ruf, A., Horaud, R. Visual servoing of robot manipulators, Part I: Projective kinematics, INRIA technical report, no. 3670, 1999.
- [3] Hutchinson, S., Hager, G., Corke, P. I. A tutorial on visual servo control, Yale University Technical Report, RR-1068, 1995.
- [4] Robson, D. Robots with eyes, *Imaging and machine vision-Europe*, 2006, vol. 17, p. 30-31.
- [5] Zuech, N. *Understanding and applying machine vision*, Marcel Dekker Inc, 2000, ISBN 0-8247-8929-6.
- [6] Papa, G., Torkar, D. Investigation of LEDs with good repeatability for robotic active marker systems, Jožef Stefan Institute technical report, no. 9368, 2006.
- [7] Zhang, Z. A flexible new Technique for Camera Calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, vol. 22, p. 1330-1334.
- [8] Faugeras, O. *Three-Dimensional Computer Vision: A geometric Viewpoint*, The MIT Press, 1993, ISBN 0-262-06158-9.