



Wide-angle camera distortions and non-uniform illumination in mobile robot tracking

Gregor Klančar*, Matej Kristan, Rihard Karba

Laboratory of Modelling, Simulation and Control, Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, SI-1000 Ljubljana, Slovenia

Received 15 January 2003; received in revised form 15 May 2003

Abstract

In this paper some fundamentals and solutions to accompanying problems in vision system design for mobile robot tracking are presented. The main topics are correction of camera lens distortion and compensation of non-uniform illumination. Both correction methods contribute to vision system performance if implemented in the appropriate manner. Their applicability is demonstrated by applying them to vision for robot soccer. The lens correction method successfully corrects the distortion caused by the camera lens, thus achieving a more accurate and precise estimation of object position. The illumination compensation improves robustness to irregular and non-uniform illumination that is nearly always present in real conditions. © 2003 Elsevier B.V. All rights reserved.

Keywords: Computer vision; Camera calibration; Non-uniform illumination correction; Mobile robots tracking

1. Introduction

Many possible ways exist of detecting moving objects using colour cameras. Sargent et al. [13] developed a fast real-time vision system with the aid of a special hardware accelerated system, which only makes sense when software optimisations or accelerations are not possible. More reliable vision tracking of moving objects can be achieved by using robust statistics and probability distributions. A good example of the latter is given in the colour-based face tracking implemented by Bradski [2]. Bruce et al. [3] suggested a fast vision system for mobile robots by means of efficient colour segmentation and a two-pass connected region determination algorithm.

Another important contribution to robot soccer vision design was introduced by Wyeth et al. [16], with special consideration given to the robustness of varying playground illumination conditions. Most of the approaches try to classify the pixels of an image into one of a predefined number of classes. The most common are: linear colour thresholding, K -nearest neighbour classification, neural net-based classifiers, classification trees and probabilistic methods [1,8,10].

In this paper an attempt to improve already-existing vision systems performance under poor operating conditions is presented. Two main factors which adversely affect performance are dealt with: non-uniform illumination and camera lens distortion. For the former, focus is placed on the problem area and its origins [6], with a solution for its compensation by means of the application of a multiplicative component defined by an illumination plane given. The latter consists of two steps. In the first, radial lens distortion

* Corresponding author. Tel.: +386-1 476 8763;
fax: +386-1 426 4631.
E-mail address: gregor.klancar@fe.uni-lj.si (G. Klančar).

fundamentals are discussed [9,14]. The suggested solution for its verification is realised by means of a geometric model of lens projection [11]. The second step covers perspective distortion originating from the tilt of the camera. For its correction, an efficient and robust method of vanishing point detection [4,12] is applied. The applicability of the presented approaches is confirmed on the robot soccer test bed. To improve the results of the robot soccer vision system [5], the proposed camera calibration and non-uniform illumination correction algorithm are implemented.

The paper is organised as follows. In Section 2, camera lens distortions are explained and a method for their correction is given. The method used for the correction of non-uniform illumination conditions is explained in Section 3. Section 4 focuses on the application of suggested algorithms to the popular robot soccer game. The paper ends with conclusions and some ideas for future work.

2. Camera distortion calibration

Colour (and also monochrome) cameras are often used for the position estimation or tracking of mobile robots. Without accurate position estimation capabilities (e.g. navigational capabilities), mobile robots cannot be assured of having effective movement and obstacle-avoidance behaviour. Whenever good position estimation based on acquired images is needed, the relationship between the pixel coordinates of the image and the real scene have to be known. This relationship also includes distortions caused by camera lenses. Several types of lens distortion exist; radial distortion is the most problematic [11], especially when inexpensive wide-angle lenses are used. The problem of camera calibration has therefore received wide attention in computer vision applications. The most frequently used method is the polynomial model for camera distortion [9,14]. Perš and Kovačič [11] suggested a mathematical model of radial distortion based on camera and lens projection geometry. Their idea is followed in the approach below.

The camera calibration routine consists of two steps originating from the demand for an easy-to-use and fully automated system, and from the fact that there are two main reasons for camera distortion. The first is radial distortion caused by the zoom (lenses); the

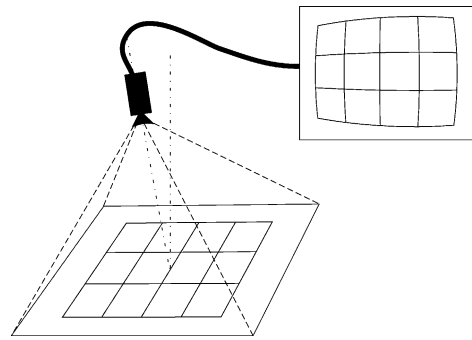


Fig. 1. Principle of camera lens distortion.

second is perspective distortion caused by the tilt of the camera. The effects of both are illustrated in Fig. 1.

2.1. Radial distortion rectification

To correct radial distortion, a corresponding model [11] based on camera and lens projection geometry is used:

$$R = f(r) = \frac{H}{2} \frac{1 - e^{-2r/H}}{e^{-r/H}} = H \sinh\left(\frac{r}{H}\right), \quad (1)$$

where R is the rectified radius, r is the radius from the distorted image (Eq. (2)) and H is the focal length. Let x and y be the coordinates of a pixel in the distorted image, and let X and Y be the coordinates of the same pixel in the rectified image. The origin of the transformation (1) is placed in the centre of the image. The relations between (x, y) and (X, Y) are as follows:

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, & \varphi &= \arctan 2\left(\frac{y}{x}\right), \\ X &= R \cos(\varphi), & Y &= R \sin(\varphi). \end{aligned} \quad (2)$$

The only unknown parameter is the focal length H , which can be set manually or automatically. Since the shape of the environment, its limitations and dimensions are often known (straight parallel and perpendicular lines in robot soccer), this knowledge can be used to estimate parameter H in the model (1). If straight parallel and/or perpendicular lines exist in the environment (e.g. boundaries, markings, etc.), the Hough transform should be used first to fit a line to the horizontal playground boundary (the longest one) on the distorted image obtained from the camera. Three

points should then be selected on the line: two at the ends and one in the middle. By moving in a normal direction from the line, the pixels on the thresholded image belonging to the boundary that are closest to the previously selected points are chosen. These three pixels are transformed by means of transformation (1) in a suitable optimisation procedure to change parameter H until the most satisfactory co-linearity of the transformed pixels is obtained.

2.2. Rectifying the effect of a tilted camera

The camera cannot always be placed over the centre of the scene under observation. Its symmetrical direction is therefore not in the normal direction of the scene under observation. With an assumption of perfect projection, e.g. with a pin-hole camera, a set of parallel lines in the scene is projected onto a set of lines in the image that meet at a common point. This point of intersection, perhaps at infinity, is called the vanishing point. The three orthogonal vanishing points form a triangle, and the intersection of the triangle's heights is called the principal point.

Many good and robust algorithms are available for the detection of perspective [12,15]. An efficient and robust method of vanishing point detection and transformation that relies only on the information gained from the playground is selected. Its basic idea was introduced by Fangi et al. [4]

$$\begin{aligned}(pX - 1)x + qYy &= -X, \\ pYx + (qY - 1)y &= -Y,\end{aligned}\quad (3)$$

where x and y are the coordinates of the rectified image, whereas X and Y are the coordinates of an image with rectified radial distortion only. The parameters p and q are estimated from the location of vanishing points, which are defined by the cross-sections of the lines fitted (by means of Hough transformation) to the playground boundaries of an image with rectified radial distortion. The principal point is placed in the centre of the image. The two lines running from the principal point to the calculated vanishing points are almost perpendicular. From the position of the vanishing points and the principal points, playground rotation is determined. If the detected rotation is not negligible, the rotation should also be rectified prior

to perspective transformation. Parameters p and q from Eq. (3) are obtained as

$$p = \frac{1}{d_{V1}}, \quad q = \frac{1}{d_{V2}}, \quad (4)$$

where d_{V1} is the distance from the principal point to the horizontal and d_{V2} is the distance to the vertical vanishing point.

Although both radial distortion and perspective correction are very simple, running the algorithm on every incoming image from the camera is time-consuming. Pixel classification, image segmentation and other vision-based algorithms are therefore better performed on a distorted image and only the estimated objects coordinates are transformed (radial distortion and perspective rectification).

3. Non-uniform illumination compensation

Good illumination conditions contribute to the efficiency of every computer vision application. However, uniform illumination is usually difficult to obtain; vision algorithms must therefore somehow be adapted to such circumstances (especially in colour-based vision tracking). Non-uniform illumination caused by irregular illumination is the main reason for the bad tracking or even loss of position estimation of mobile robots. If a vision system is taught a specific colour when the carrier is somewhere in the centre of the observed scene, there is a strong possibility that this colour will not be recognised when its carrier is positioned in one of the corners. If non-uniform illumination and illumination conditions do not change drastically during operation, a static shade model could be built in the initialisation phase. The relation between the true (ideal uniform illumination) image $U(x, y)$ and acquired image $N(x, y)$ is usually described by means of a linear model of image formation:

$$N(x, y) = U(x, y)S_M(x, y) - S_A(x, y), \quad (5)$$

where $S_M(x, y)$ is the multiplicative and $S_A(x, y)$ the additive component. Retrospective and acquisition-based approaches are two types of possible non-uniform illumination correction method. Acquisition-based methods involve taking one or two reference images and interpreting them as multiplicative and additive components for the acquired image, while retrospective

methods rely solely on the information content of the acquired image. The most intuitive retrospective methods for correcting multiplicative and smooth intensity variations are homomorphic filtering, image blurring, smoothing, averaging, Fourier-domain filtering and homomorphic unsharp masking [7]. In [6], additive and multiplicative non-uniform illumination components were approximated by means of second-order polynomials. Our test showed that this type of model did not improve detection, and sometimes even made it worse. The reason for such results lies in the shape of the illumination plane. A typical illumination plane consists of several peaks and valleys which cannot be modelled sufficiently by means of second-order polynomials but need higher order approximations. It may be quite difficult to estimate the proper order of a polynomial in a model, hence a simpler method was chosen. The proposed method is acquisition-based; it employs a background image and the additive component is discarded. Non-uniform illumination is finally corrected by relation (6):

$$\hat{U}(x, y) = \frac{N(x, y)}{N_{\text{BCK}}(x, y)} C, \quad (6)$$

where $\hat{U}(x, y)$ is the corrected image, $N_{\text{BCK}}(x, y)$ is the background image and C is the normalisation constant needed to restore the desired grey level range.

To improve computational time efficiency, multiplication by component $C/N_{\text{BCK}}(x, y)$ in Eq. (6) should be realised by means of a look-up table.

3.1. Acquiring the multiplicative component

The camera's field of view should be covered with grey paper sheets (not necessary if the scene is uniformly coloured already). The acquired reference image $N_{\text{BCK}}(x, y)$ then represents the non-uniform illumination plane of a multiplicative component $C/N_{\text{BCK}}(x, y)$ in relation (6). The reference image is retouched with a slight Gaussian blurring to increase the smoothness of the non-uniform illumination plane. The illumination planes are then calculated for all three channels in RGB colour space (red, green, and blue). Fig. 2 shows a typical illumination plane (calculated from the robot soccer court in Fig. 4) for the red channel only, while the other two have a similar shape.

Among a number of available colour spaces, the RGB colour space is usually output in most colour cameras and frame-grabbers (easy to implement in computer systems, television, video, etc.) while in image processing the HSI and YUV colour spaces seem more appropriate when there are different illumination conditions. They code the information about chrominance in two dimensions (H and S, or U and V), with

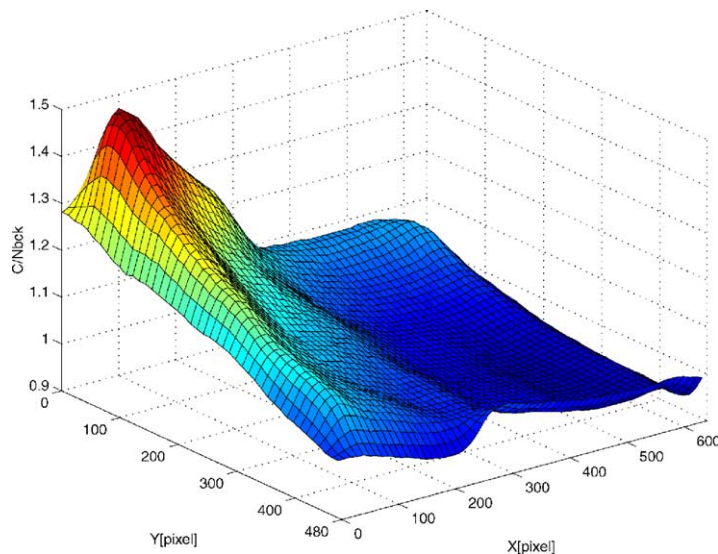


Fig. 2. Calculated multiplicative component (C/N_{BCK}) for the red channel.

only one dimension including information about intensity (I or Y).

In our experiments, both RGB and YUV colour spaces were tested on incoming camera images. The classification results obtained were similar. However, to obtain the YUV colour representation, transformation from the original RGB space had to be performed. This transformation was time-consuming, although optimisation by means of look-up tables was used. Including optimisation, it requires some 30 ms, while the rest of the program takes only 8 ms to identify objects from the image. RGB colour space is therefore used in the experiments presented. The illumination plane from Fig. 2 is then applied to incoming camera images (in accordance with Eq. (6)).

4. Application of proposed corrections to the soccer robot game

Suggested approaches for the rectification of camera distortions and non-uniform illumination are demonstrated on a robot soccer set-up. The set-up (Fig. 3) consists of 10 MiroSot category robots (forming two teams) of 7.5 cm × 7.5 cm × 7.5 cm, a rectangular playground of 2.2 m × 1.8 m, a colour NTSC camera, a Matrox Meteor 2 frame-grabber, and a Pentium 4 PC. The vision part of the program processes the incoming images at a resolution of 640 × 480 pixels in order to identify the positions and orientations of the robots and the position of the ball. Each robot has two

square-shaped colour patches. One is the team colour (blue or yellow) and the other (any colour but blue, yellow or orange) is the identification colour patch. The vision algorithm [5] finds objects on the playground by taking their colour and shape into consideration. If an incoming pixel colour falls inside one of the pre-defined boxes (defined by thresholds), it is classified as belonging to the class associated with this box. The thresholds are presented as boxes in three-dimensional colour spaces. The pixels belonging to one class (a connected region) are then distinctively labelled. The logic part and a simple optimisation method are employed to select the proper regions from the previously generated ones. Finally, the control part of the program calculates the linear and angular speeds, v and ω , the robots should have in the next sample time according to the current situation on the playground. These reference speeds are sent to the robots by means of a wireless connection, and they start moving according to received commands. The above-mentioned cycle repeats itself 30 times per second.

To improve the results of the vision system, the above-mentioned camera calibration and non-uniform illumination correction algorithm are implemented as seen in Fig. 3.

4.1. Camera distortion correction

In a robot soccer game, each team has its own camera. Each of them is placed slightly to the left or right of the centre of the playground. In order to cover as much of the acquired image of the playground as

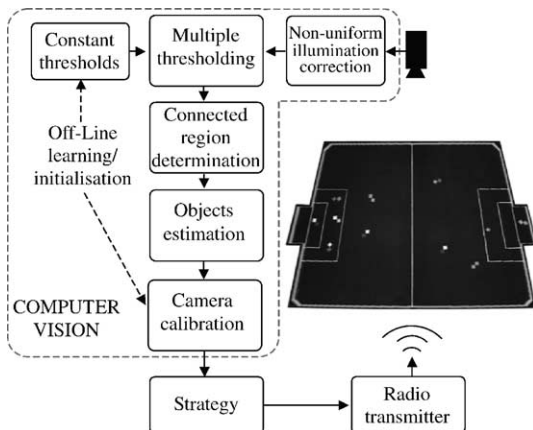


Fig. 3. Robot soccer overview.

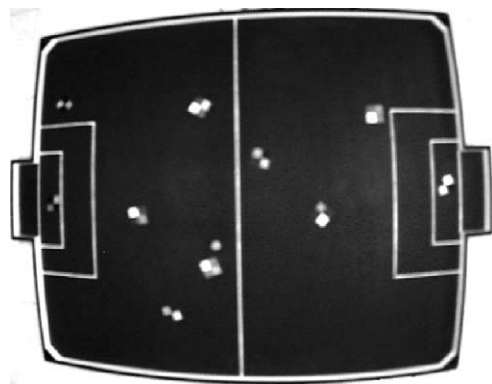


Fig. 4. Distorted image (image from the camera with radial and perspective distortion).

possible, the camera has to be tilted, which causes perspective distortion. The camera must also see the whole playground and must not be higher than 2 m above the playground (game rules). Wide-angle lenses are therefore used; these cause radial distortion.

An example of the complete camera calibrating procedure is shown in Figs. 4 and 5. The distorted image obtained from the camera is shown in Fig. 4, while the rectified image can be seen in Fig. 5.

4.2. Non-uniform illumination compensation

Although the playground is illuminated by a number of different light sources (reflectors or neon tubes), it is practically impossible to obtain uniform illumination. If lights are arranged carefully, compensation may not be needed, but this is usually not the case.

To test the efficiency of the non-uniform illumination correction algorithm, a number of colour patches of the same colour are arranged over the playground (see Fig. 6). From the colour patch in the centre of Fig. 6, the representative colour thresholds are determined. The pixel classification and segmentation algorithm is first represented on the incoming camera image (Fig. 7), and then on the same image with compensated non-uniform illumination (Fig. 8).

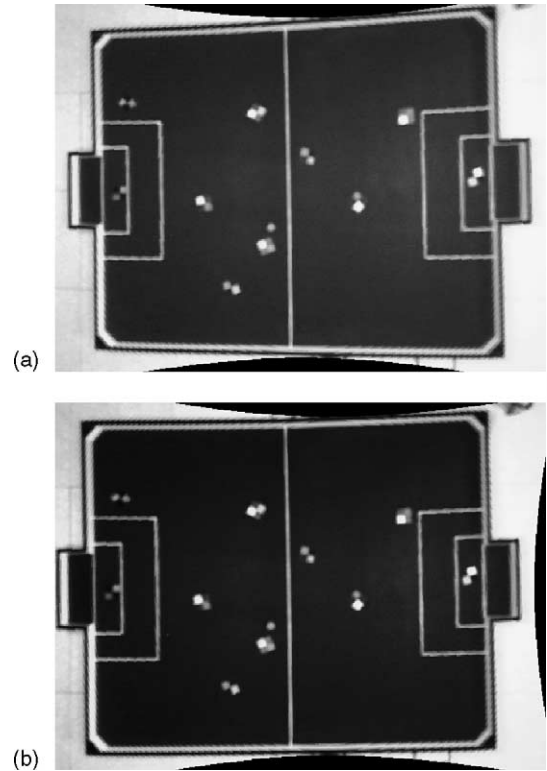


Fig. 5. Rectified image: (a) corrected radial distortion only; (b) corrected radial and perspective distortion.

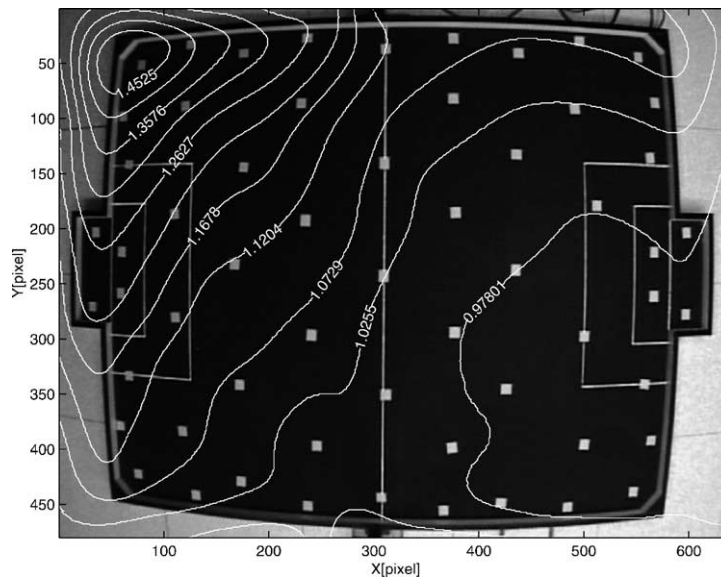


Fig. 6. Camera image with colour patches arranged over the playground and intensity illumination contours displayed.

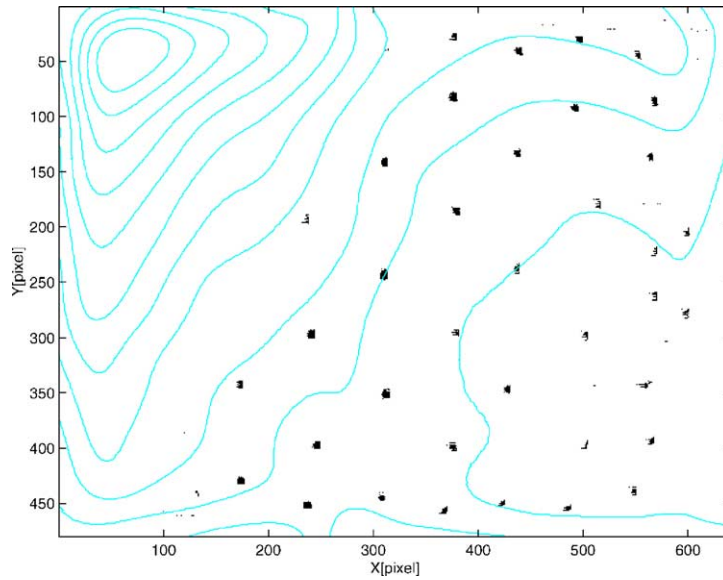


Fig. 7. Classified pixels from the camera image in Fig. 6.

In Figs. 7 and 8, the regions with black colour represent recognised pixels which belong to colour patches from Fig. 6. A pixel is classified as belonging to the colour patches if it fits the representative colour thresholds. Due to non-uniform illumination (see the intensity illumination contours in

Figs. 6–8 or the illumination plane in Fig. 2), some colour patches in Fig. 7 are purely detected or even undetected.

From Fig. 8 it can be seen that non-uniform illumination compensation improves the efficiency of the classification and segmentation algorithm. All

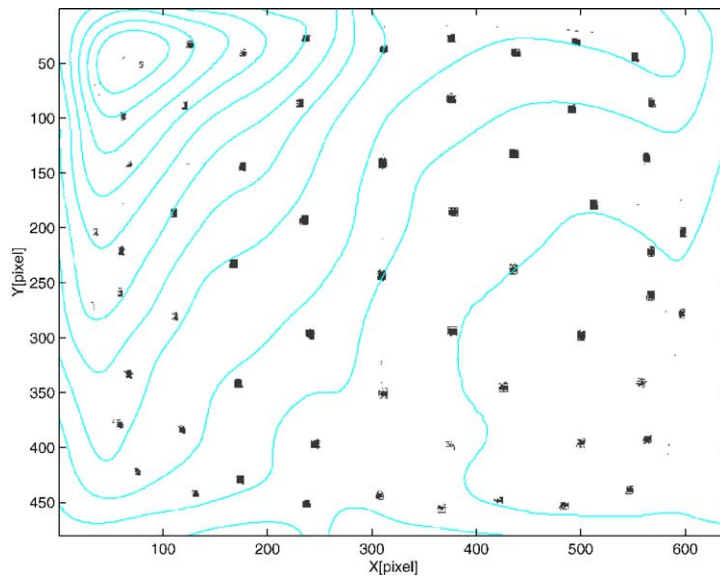


Fig. 8. Classified pixels from the camera image in Fig. 6 with compensated non-uniform illumination.

the patches from Fig. 6 can be reliably identified from the compensated image (Fig. 8); this could not be achieved by using the non-compensated image (Fig. 7).

4.3. Evaluation of results

The suggested approaches for the rectification of camera distortions and non-uniform illumination applied to a robot soccer set-up improve the game in the following ways.

Radial and perspective distortions are successfully corrected, which can be proved by obtaining straight parallel and perpendicular lines for playground boundaries (see Fig. 5) from the distorted image (Fig. 4) of the playground. More accurate and precise estimations of robot positions are therefore achieved, which is especially important for robot control when robots are close to boundaries. The improvement is difficult to estimate because this depends on where the robots are mostly moving during the game. If they are mostly moving in the corners of the playground, the precision of the robot's positions and the quality of the game (boundary avoidance, ball manipulation) is improved by up to 40%.

Non-uniform illumination compensation improves robustness to irregular illumination. Patches in certain areas on the playground which cannot be found from the distorted image are then reliably found (recognition rate of around 98%). Recognition rate is thus improved by up to approximately 30% compared to the rate for poor illumination conditions (statistics taken from our game).

5. Conclusions

An approach towards establishing a more robust and accurate vision system for mobile robot tracking under poor illumination and camera lens distortion conditions is presented. To improve the results of visual robot tracking, a camera calibration and non-uniform illumination correction algorithm is suggested. The former corrects the distortion caused by the camera lens, thus achieving a more accurate and precise estimation of object position, while the latter improves robustness to irregular illumination and non-uniform illumination conditions. The applicability of the sug-

gested solutions is demonstrated in a robot soccer game, where any incorrect or inaccurately estimated robot or ball position results in poor game-play (apart from perfection of the strategy control algorithm). The robustness of the vision system is therefore improved by means of camera calibration algorithms. The suggested procedure for shading correction proved useful when the illumination conditions remained more or less unchanged during the game. The procedure presented also assumes fixed camera view, as in central vision systems. In general mobile robotics, these conditions are not always met. If illumination conditions change during tracking, a more robust approach with an adaptation mechanism should be applied. This will be addressed in further research. The optimised algorithms presented enable the vision system to be used in real-time applications where robustness to irregular illumination and camera distortions are important.

References

- [1] C.M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, UK, 1995.
- [2] G.R. Bradski, Computer vision face tracking for use in a perceptual user interface, *Intel. Technology Journal* 2 (2) (1998) 1–15.
- [3] J. Bruce, T. Balch, M. Veloso, Fast and inexpensive color image segmentation for interactive robots, in: *Proceedings of the IROS-2000*, Takamatsu, Japan, October 2000, pp. 2061–2066.
- [4] G. Fangi, G. Gagliardini, E.S. Malinverni, Photointerpretation and small scale stereoplotting with digitally rectified photographs with geometrical constraints, in: *Proceedings of the CIPA Congress 2001*, September 18–21, 2001. <http://cipa.icomos.org/papers/2001-htm>.
- [5] G. Klančar, O. Orqueda, D. Matko, R. Karba, Robust and efficient vision system for mobile robots control—application to soccer robots, *Electrochemical Review* 68 (5) (2001) 306–312.
- [6] B. Likar, B.A. Maintz, M.A. Viergever, F. Pernuš, Retrospective shading correction based on entropy minimization, *Journal of Microscopy* 197 (2000) 285–295.
- [7] B. Likar, M.A. Viergever, F. Pernuš, Retrospective correction of MR intensity inhomogeneity by information minimization, in: S.L. Delp, A.M. DiGioia, B. Jaramaz (Eds.), *Lecture Notes in Computer Science*, vol. 1935, *Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, 2000, pp. 375–384.
- [8] C.G. Looney, *Pattern Recognition Using Neural Networks, Theory and Algorithms for Engineers and Scientists*, Oxford University Press, UK, 1997.

- [9] T. Pajdla, T. Werner, V. Hlavač, Correcting radial lens distortion without knowledge of 3-D structure, Technical Report TR97-138, Faculty of Electrical Engineering, Czech Technical University, Praha, Czech Republic, June 1997.
- [10] N. Pavešič, Razpoznavanje vzorcev: Uvod v analizo in razumevanje vidnih in slušnih signalov (Pattern recognition: an introduction into the analysis and understanding of visual and audio signals), Faculty of Electrical Engineering, Ljubljana, Slovenia, 2000.
- [11] J. Perš, S. Kovačič, Nonparametric, Model-based radial lens distortion correction using tilted camera assumption, in: H. Wildenauer, W. Kropatsch (Eds.), Proceedings of the Computer Vision Winter Workshop 2002, Bad Aussee, Austria, February 2002, pp. 286–295.
- [12] C. Rother, A new approach for vanishing point detection in architectural environments, in: Proceedings of the 11th British Machine Vision Conference (BMVC'00), Bristol, UK, 2000, pp. 382–391.
- [13] R. Sargent, B. Bailey, C. Witty, A. Wright, The importance of fast vision in winning the First Micro-Robot World Cup Soccer Tournament, Robotics and Autonomous Systems 21 (1997) 139–147.
- [14] R.Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, IEEE Journal of Robotics and Automation 3 (4) (1987) 323–344.
- [15] T. Tuytelaars, M. Proesmans, L.V. Gool, The cascaded Hough transform, in: Proceedings of the International Conference on Image Processing (ICIP'97), vol. 2, 1997, pp. 736–739.
- [16] G.F. Wyeth, B. Brown, Robust adaptive vision for robot soccer, in: J. Billingsley (Ed.), Mechatronics and Machine Vision Baldock, Herts, Research Studies Press, UK, 2000, pp. 41–48.



Gregor Klančar received his B.Sc. degree in 1999 from the Faculty of Electrical Engineering of the University of Ljubljana, Slovenia, where he is currently employed as a member of the National Young Researcher Scheme. His research interests are in the area of fault diagnosis methods, multiple vehicle coordinated control and robot soccer.



Matej Kristan received his B.Sc. degree in electrical engineering from the University of Ljubljana, Faculty of Electrical Engineering, Slovenia, in 2003. His research interests are in the area of computer vision, image processing and machine vision applications.



Rihard Karba received his B.Sc. (1972), M.Sc. (1975) and Ph.D. (1981) degrees from the Faculty of Electrical Engineering of the University of Ljubljana, Slovenia, where he is currently a Professor. His research interests include modelling and simulation of dynamic systems, multivariable control systems analysis and design as well as modelling and simulation in biopharmacy and in biomedicine.