Pattern Recognition 47 (2014) 89–95

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Camera calibration in sport event scenarios

M. Alemán-Flores [a], L. Alvarez [a], L. Gomez [b], P. Henriquez [a,*], L. Mazorra [a]

[a] CTIM: Centro de I+D de Tecnologías de la Imagen, Departamento de Informática y Sistemas, Universidad de Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas, Spain

[b] CTIM: Centro de I+D de Tecnologías de la Imagen, Departamento de Ingeniería Electrónica y Automática, Universidad de Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas, Spain

ARTICLE INFO

Available online 22 May 2013

Keywords:
Camera calibration
Sport scenarios
Camera pose

ABSTRACT

The main goal of this paper is the design of a novel and robust methodology for calibrating cameras from a single image in sport scenarios, such as a soccer field, or a basketball or tennis court. In these sport scenarios, the only references we use to calibrate the camera are the lines and circles delimiting the different regions. The first problem we address is the extraction of image primitives, including the challenging problems of shaded regions and lens distortion. From these primitives, we automatically recognise the location of the sport court in the scene by estimating the homography which matches the actual court with its projection onto the image. This is achieved even when only a few primitives are available. Finally, from this homography, we recover the camera calibration parameters. In particular, we estimate the focal length as well as the position and orientation in the 3D space. We present some experiments on models and real courts which illustrate the accuracy of the proposed methodology.

1. Introduction

During the broadcast of sport events, video sequences are usually processed for a wide variety of purposes, such as mosaicing, change of the view point or insertion of virtual objects. Some of these tasks require a highly precise calibration of the cameras, that is, to know the *camera pose*: position and orientation of the camera and its intrinsic parameters (focal length, pixel size and the principal point). Moreover, the estimation of the lens distortion is likely to be needed for the case of highly precise camera calibration. In what follows, we will assume that all intrinsic parameters of the camera are known except for the focal length, which can be modified by the camera operator by changing the camera zoom. Such assumption comes from the fact that cameras at sports events are usually mounted on fixed tripods and are previously calibrated by means of standard off-line calibration methods to extract the intrinsic parameters. During the broadcast of the sport event, the cameras remain in the initial fixed position and it is even usual that they remain at the same position for different matches played at the same court.

In sport scenarios the main primitives we can use to calibrate the camera are the white lines and circles dividing the sport court. We assume that the court is a planar surface with known dimensions and a certain number of lines or circles (usually in white colour) dividing the different parts of the court. This is a quite common situation in sport scenarios, such as tennis, basketball or soccer.

As we will show below, when we capture a picture of a planar scenario, the positions of the primitives in the image are given by a homography. In Fig. 1, we illustrate this perspective transformation. We will show in this paper that if a minimum number of court primitives are visible in the scene, then we can recover the homography which transforms the image scene plane into the actual reference plane. Then, from this homography (and assuming that some internal camera parameters are known), we can recover the camera pose, that is, the position and orientation of the camera.

In the application of our methods, we focus on soccer images. These are one of the most challenging types, since they present natural light conditions (which are therefore quite variable), shaded regions, and heterogeneous background.

The main contribution of the paper is the introduction of a novel and fully automatic technique that allows estimating, from a single image, the camera calibration parameters in complex scenarios where we deal with a few visible primitives, poor light conditions, shaded areas or lens distortion. The proposed technique can be divided into the following steps:

1. Primitive extraction (white lines/circles) in the image. We also consider the challenging problem of images showing large shaded areas. In sport events broadcasted under natural light conditions (*daylight*), we can deal with large areas where the primitives are shaded, which makes the extraction of the primitives a more difficult task.

* Corresponding author. Tel.: +34 928 458708.
E-mail addresses: maleman@ctim.es (M. Alemán-Flores), lalvarez@ctim.es (L. Alvarez), lgomez@ctim.es (L. Gomez), phenriquez@ctim.es (P. Henriquez), lmazorra@ctim.es (L. Mazorra).

0031-3203/$ - see front matter
http://dx.doi.org/10.1016/j.patcog.2013.05.011

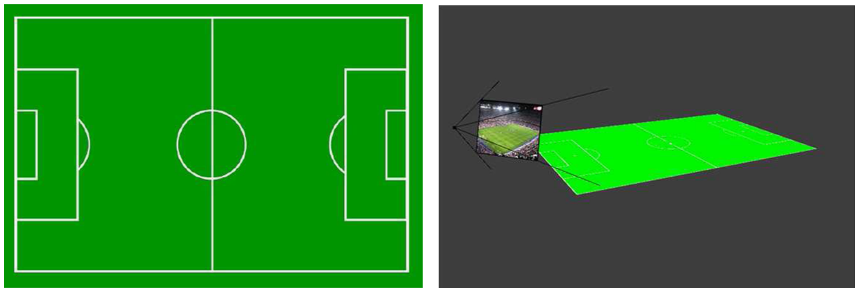

Fig. 1. Reference soccer field (left), and camera pose reconstruction showing the acquired image and the corresponding position and orientation of the camera with respect to the field (right).

2. Sport court recognition. We automatically estimate the homography which transforms the actual sport court into its projection onto the image. To do that, we propose some novel strategies to match the primitives extracted from the image and the actual primitives in the reference plane.
3. Camera calibration. We recover the camera calibration parameters from the estimated homography. In Fig. 1, we illustrate the result of the calibration procedure.

The organisation of the paper is as follows: in Section 2, we present some related works. Section 3 is devoted to primitive extraction. In Section 4, we introduce the camera model which has been used. Section 5 presents the loss function which measures the accuracy of a homography. Section 6 is devoted to camera calibration. In Section 7, we present some experiments on simulated and real scenarios. Finally, in Section 8, we present some conclusions.

2. Related works

Camera pose determination is a major issue in computer vision. The problem was formulated by Fischler and Bolles [1] for the case of a camera with constant focal length, that is, no zooming was accounted, and the first attempts to solve it used 2D or 3D point correspondences [2,3]. Other works consider more features into the recognition phase, such as corners or man-made objects [4,5]. These methods exploit some knowledge within the scene (e.g. lines or corners) or even use standard calibration patterns (*chessboard*) and require multiple views. Then, a nonlinear optimisation problem is stated and solved by gradient-like optimisation methods to get a satisfactory local minimum. Radial distortion or the possibility of changing the camera zoom is not usually taken into account in order to simplify the problem, or is included in a post-processing stage. In this work, to deal with lens distortion we use the algebraic approach detailed in [6,7], where the authors use a least square approximation of those edges which should be a projection of 3D lines as a measure of distortion. Afterwards, they calculate the sum of the squares of the distances from the points to the lines to obtain the distortion error.

A method for self-calibration of rotating and zooming cameras is described in [8]. Other works present some methods for multi-camera calibration, such as [9,10]. Based on those techniques, some optimisations for the calibration of PTZ (Pan–Tilt–Zoom) cameras have been introduced [11], and there are also some methods that do not require predefined patterns for tracking, such as [12].

There has been a great deal of recent works in the pattern recognition research community tackling the specific case of the camera pose problem for planar scenarios. Video tracking for soccer matches, automatic analysis of soccer video sequences and summarisation of complete soccer matches require a precise camera calibration and camera pose recognition [13–17].

For the case of planar scenarios, there is a significant reduction of the complexity of the problem due to the fact that the camera is mounted on a fixed tripod and the correspondence to be established is also reduced to a 2D problem [18]. One of the first published works researching the particular problem of camera calibration under the geometrical 2D restriction was [19]. In this paper, the author shows that the camera pose for a moving projective camera can be estimated from a minimum number of five frames of a planar view under the constraint that the projections of orthogonal direction vectors (points at infinity) in the plane must be orthogonal in the calibrated camera frame of each image. In [20], an automatic camera calibration for short sequences of video soccer matches taken in the penalty area near one of the goals is introduced. The authors apply the well-known Tsai's camera calibration method [21] and a planar transformation (*homography*) is computed from some primitives extracted from the images to estimate the intrinsic and extrinsic camera parameters, including the zoom factor. In [22], by using images taken simultaneously by two PTZ cameras which are positioned at different locations and observe the same dynamic event taking place on a planar surface, a planar homography is calculated for each time instant using the image correspondences between both images and, from that homography, the camera pose is obtained.

Most line extraction methods applied to sport events commonly use a segmentation with dominant colour detection using HSV space [13] or Gaussian mixture models [23]. In this paper we apply the method described in [24] for extracting lines when dealing with interlaced images, HD definition images, or scenarios with significant contrast variations between the background and the lines.

Finally, to calculate the extrinsic parameters of the camera (rotation and focus) some methods have been proposed using lines/circles [25] or court models [26,27]. There are also some other methods which use the vanishing points of the lines, such as [28].

3. Primitive extraction

In order to calibrate the camera, we need to extract the primitives which will help us identify its parameters. These are the white lines and circles which characterise a soccer pitch. To this aim, we will make use of the following disk morphological operators:

Given an image $I:\Omega\rightarrow\Re$ and a disk D_s centred in $(0,0)$ and with radius s, we define:

$$\begin{aligned}
&\textit{Disk dilation}: && I\oplus D_s(x)=\sup_{y\in x+D_s} I(y)\\
&\textit{Disk erosion}: && I\ominus D_s(x)=\inf_{y\in x+D_s} I(y)\\
&\textit{Disk opening}: && I\bigcirc D_s(x)=(I\ominus D_s)\oplus D_s(x)\\
&\textit{Disk closing}: && I\bullet D_s(x)=(I\oplus D_s)\ominus D_s(x),
\end{aligned}$$