Pattern Recognition 47 (2014) 2280–2292

Contents lists available at ScienceDirect

Pattern Recognition

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

Automatic generation and detection of highly reliable fiducial markers under occlusion

S. Garrido-Jurado*, R. Muñoz-Salinas, F.J. Madrid-Cuevas, M.J. Marín-Jiménez

Computing and Numerical Analysis Department, Córdoba University, Edificio Einstein, 3ª planta, Campus de Rabanales, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 4 June 2013
Received in revised form
2 December 2013
Accepted 9 January 2014
Available online 21 January 2014

Keywords:
Augmented reality
Fiducial marker
Computer vision

ABSTRACT

This paper presents a fiducial marker system specially appropriated for camera pose estimation in applications such as augmented reality and robot localization. Three main contributions are presented. First, we propose an algorithm for generating configurable marker dictionaries (in size and number of bits) following a criterion to maximize the inter-marker distance and the number of bit transitions. In the process, we derive the maximum theoretical inter-marker distance that dictionaries of square binary markers can have. Second, a method for automatically detecting the markers and correcting possible errors is proposed. Third, a solution to the occlusion problem in augmented reality applications is shown. To that aim, multiple markers are combined with an occlusion mask calculated by color segmentation. The experiments conducted show that our proposal obtains dictionaries with higher inter-marker distances and lower false negative rates than state-of-the-art systems, and provides an effective solution to the occlusion problem.

1. Introduction

Camera pose estimation (Fig. 1(a,b)) is a common problem in many applications requiring a precise localization in the environment such as augmented and virtual reality applications, and robotics. [1–4]. Obtaining the camera pose from images requires to find the correspondences between known points in the environment and their camera projections. While some approaches seek natural features such as key points or textures [5–9], fiducial markers are still an attractive approach because they are easy to detect and allows us to achieve high speed and precision.

Among the several fiducial marker systems proposed in the literature, those based on square markers have gained popularity, especially in the augmented reality community [10–12]. The reason why is that they allow us to extract the camera pose from their four corners, given that the camera is properly calibrated. In most of the approaches, markers encode a unique identification by a binary code that may include error detection and correction bits. In general, each author has proposed its own predefined set of markers (*dictionary*). The problems of setting a predefined dictionary are twofold. First, in some cases, the number of markers required by the application might be higher than the dictionary size. Second, if the number of markers required is smaller, then it is preferable to use a smaller dictionary whose inter-marker distance is as high as possible, so as to reduce the inter-marker confusion rate.

Another common problem in augmented reality applications is related to the occlusion. The problem occurs when a real object appears occluding the virtual scene. In this case, the virtual objects are rendered on the real object, which should be visible (see Fig. 1 (c,d)). This is indeed a limitation to the augmented experience since the user cannot interact freely.

This paper presents a fiducial marker system based on square markers offering solutions to the above-mentioned problems. First, we propose a general method for generating configurable dictionaries (both in size and number of bits). Our algorithm creates dictionaries following a criterion to maximize the inter-marker distance and the number of bit transitions. In the process, we derive the maximum theoretical inter-marker distance that a dictionary of square binary markers can have. Then, a method for automatically detecting markers in images and correcting possible errors, based on our generated dictionaries, is presented. Third, we propose a solution to the occlusion problem based on combining multiple markers and an occlusion mask calculated using color information. While using multiple markers provides robustness against occlusion, color information is used to determine the occluded pixels avoiding rendering on them.

The rest of the paper is structured as follows. Section 2 presents the most relevant works related to ours. Section 3 explains the proposed method to generate marker dictionaries. Section 4 shows the process proposed for marker detection and error correction.

* Corresponding author. Tel.: +34 957212255.
E-mail addresses: i52gajus@uco.es (S. Garrido-Jurado), rmsalinas@uco.es (R. Muñoz-Salinas), fjmadrid@uco.es (F.J. Madrid-Cuevas), mjmarin@uco.es (M.J. Marín-Jiménez).

0031-3203/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2014.01.005

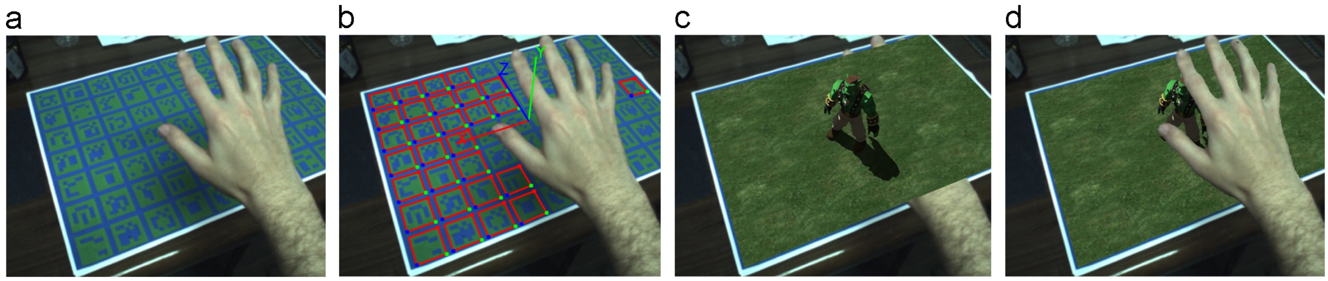

Fig. 1. Example of augmented reality scene. (a) Input image containing a set of fiducial markers. (b) Markers automatically detected and used for camera pose estimation. (c) Augmented scene without considering user's occlusion. (d) Augmented scene considering occlusion.

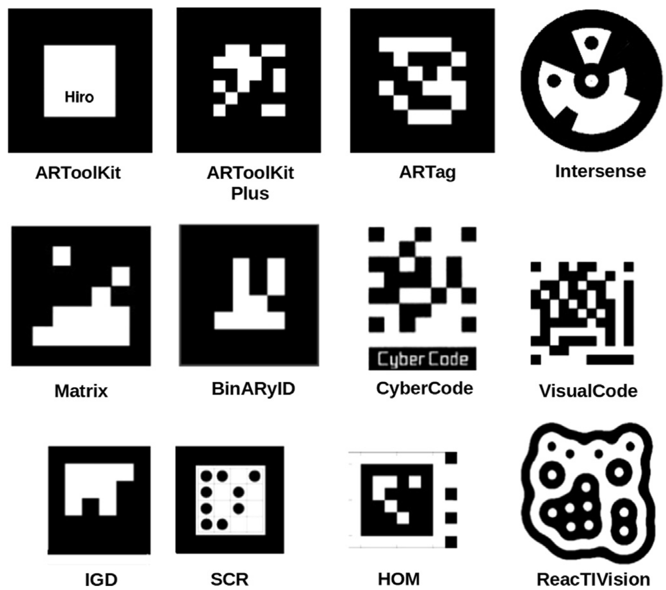

Fig. 2. Examples of fiducial markers proposed in previous works.

Section 5 presents our solution to the occlusion problem. Finally, Section 6 shows the experimentation carried out, and Section 7 draws some conclusions.

Finally, it must be indicated that our work has been implemented in the ArUco library which is freely available [13].

2. Related work

A fiducial marker system is composed by a set of valid markers and an algorithm which performs its detection, and possibly correction, in images. Several fiducial marker systems have been proposed in the literature as shown in Fig. 2.

The simplest proposals consist in using points as fiducial markers, such as LEDs, retroreflective spheres or planar dots [14,15], which can be segmented using basic techniques over controlled conditions. Their identification is usually obtained from the relative position of the markers and often involves a complex process.

Other approaches use planar circular markers where the identification is encoded in circular sectors or concentric rings [16,17]. However, circular markers usually provide just one correspondence point (the center), making necessary the detection of several of them for pose estimation.

Other types of fiducial markers are based on blob detection. Cybercode [18] or VisualCode [19] is derived from 2D-barcodes technology as MaxiCode or QR but can also accurately provide several correspondence points. Other popular fiducial markers are the ReacTIVision amoeba markers [20] which are also based on blob detection and its design was optimized by using genetic algorithms. Some authors have proposed the use of trained classifiers to improve detection in cases of bad illumination and blurring caused by fast camera movement [21].

An alternative to the previous approaches is the square-based fiducial markers systems. Their main advantage is that the presence of four prominent points can be employed to obtain the pose, while the inner region is used for identification (either using a binary code or an arbitrary pattern such as an image). In the arbitrary pattern category, one of the most popular systems is ARToolKit [10], an open source project which has been extensively used in the last decade, especially in the academic community. ARToolKit markers are composed by a wide black border with an inner image which is stored in a database of valid patterns. Despite its popularity, it has some drawbacks. First, it uses a template matching approach to identify markers, obtaining high false positive and inter-marker confusion rates [22]. Second, the system uses a fixed global threshold to detect squares, making it very sensitive to varying lighting conditions.

Most of the square-based fiducial systems use binary codes. Matrix [23] is one of the first and simplest proposals. It uses a binary code with redundant bits for error detection. The ARTag [11] system is based on the same principles but improves the robustness to lighting and partial occlusion by using an edge-based square detection method, instead of a fixed threshold. Additionally, it uses a binary coding scheme that includes checksum bits for error detection and correction. It also recommends using its dictionary markers in a specific order so as to maximize the inter-marker distances. Its main drawback is that the proposed