Robotics and Autonomous Systems 62 (2014) 1271–1281

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

Spatial layout recovery from a single omnidirectional image and its matching-free sequential propagation

G. López-Nicolás *, J. Omedes, J.J. Guerrero

Instituto de Investigación en Ingeniería de Aragón. Universidad de Zaragoza, C/María de Luna 1, E-50018 Zaragoza, Spain

HIGHLIGHTS

- Method for scene layout recovery from an omnidirectional single image.
- Matching-free homography-based extension to sequences of images.
- All the hypotheses are propagated along the images using homographies.
- Experimental results with different image databases show good performance.

ARTICLE INFO

Article history:
Available online 5 April 2014

Keywords:
Omnidirectional vision
Vanishing points
Homography
Manhattan-world
Spatial layout

ABSTRACT

The goal of this work is to recover the spatial layout of indoor environments from omnidirectional images assuming a Manhattan world structure. We propose a new method for scene structure recovery from a single image. This method is based on the line extraction for omnidirectional images, line classification, and vanishing points estimation combined with a new hierarchical expansion procedure for detecting floor and wall boundaries. Each single omnidirectional image independently provides a useful hypothesis of the 3D scene structure. In order to enhance the robustness and accuracy of this single image-based hypothesis, we extend this estimation with a new homography-based procedure applied to the various hypotheses obtained along the sequence of consecutive images. A key point in this contribution is the use of geometrical constraints for computing the homographies from a single line of the floor. The homography parametrization proposed allows the design of a matching-free method for spatial layout propagation along a sequence of images. Experimental results show single image layout recovery performance and the improvement obtained with the propagation of the hypothesis through the image sequence.

1. Introduction

In this work, we address the problem of recovering the spatial layout of a scene from a set of lines extracted from indoor omnidirectional images. Although related works have been proposed in the literature, they are exclusively intended for conventional cameras. Omnidirectional cameras provide information in a wide field of view in one single image. Among these cameras we find the catadioptric systems, which are the combination of mirrors and cameras. A survey of the different classes of central catadioptric sensors with one mirror and lens are treated in [1]. For many applications, these camera systems are usually spotted in wheel-based robotic platforms with the mirror and the camera vertically aligned with respect to their symmetry axes. This consideration is useful since vertical lines of the scene become straight radial lines in the image and their vanishing point appears in the center of the image. When this assumption is not satisfied, all the lines of the scene become conic lines in the image and vanishing point detection is harder to execute given the additional geometrical complexity. Another good property in omnidirectional catadioptric systems as a consequence of their wide field of view is the observation of longer lines of the scene than in conventional camera systems. Additionally, vanishing points usually fall inside the omnidirectional image boundaries, being easier and more accurate to detect. Given these characteristics, the omnidirectional images allow a better way to deal with certain problems than conventional cameras, such as the problem of scene layout recovery from single images.

Spatial layout recovery of a scene requires the analysis of the surrounding structures to be able to recognize its global geometry, boundaries between floor and walls, relative orientation of

* Corresponding author. Tel.: +34 876555556.
E-mail addresses: gonlopez@unizar.es (G. López-Nicolás), jason.omedes@gmail.com (J. Omedes), josechu.guerrero@unizar.es (J.J. Guerrero).

http://dx.doi.org/10.1016/j.robot.2014.03.018
0921-8890/

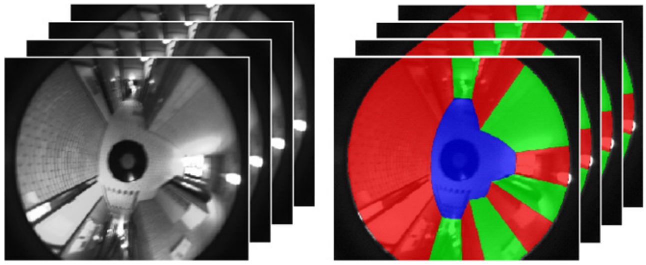

Fig. 1. Example illustrating the goal of the presented approach. An input image (left) and the result of the algorithm (right). Walls are in green or red, and floor is in blue. The single view procedure is also enhanced through a sequence of images avoiding image feature matching. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

surfaces, corners location, and relative distances between different elements of the scene. In general, a conservative spatial layout is enough for defining a navigability map of the environment, and this can be a powerful tool that provides very useful information for performing tasks such as navigation [2] or obstacle detection [3]. Therefore, rather than a precise and detailed map of the scene [4–6], we focus in this work on providing a conservative map in which the distribution of the different elements of the scene are classified as floor or walls. An example of the expected result for an input omnidirectional image after recovering its spatial layout is illustrated in Fig. 1.

Works in the literature addressing the problem of spatial layout recovery have been proposed for images acquired by conventional cameras, and most of them work under the Manhattan-world assumption [7]. This assumption states that the scene is essentially composed of three main directions orthogonal to each other, which is usually accomplished for indoor environments. This is used for instance in [8] to create a cubic room model attempting to recognize surfaces in cluttered scenes. Straight lines are usually easy to find and extract in structured environments like indoor scenes of buildings. Many techniques start with line detection as main image feature and impose geometrical constraints in order to find corners or relevant characteristics such as parallelism or orthogonality between elements to generate plausible hypothesis of the scene structure [9]. All of these techniques were designed to be used with conventional images. In this context, the use of omnidirectional vision would be advantageous due to the wider field of view and the other characteristics commented above. However, omnidirectional vision presents challenges such as geometrical complexity that makes the previous proposed techniques for conventional images not plausible, and new algorithms are required to take into account this different geometry.

In this paper, we present an efficient method for spatial layout recovery from single omnidirectional images. Additionally, the method proceeds by propagating the individual hypothesis along a sequence composed by consecutive images of the scene. To our knowledge, this is the first spatial layout recovery algorithm that addresses the problem with omnidirectional vision. The first contribution is a new algorithm that provides the spatial layout recovery from a single image. In particular, the algorithm for estimating the structural layout hypothesis for a single image proceeds as follows. First, lines from the omnidirectional image are extracted and classified depending on their orientation by estimating the vanishing points (VPs). Examples of works using VPs are the 3D reconstruction from single standard images presented in [10] or the line clustering to determine the VPs from omnidirectional images [11]. Using the VPs classification, our proposed algorithm selects a set of lines producing different possible wall–floor boundaries. Imposing geometrical constraints an initial four walls-room hypothesis is generated followed by a hierarchical expansion process that accommodates the room hypothesis according to the image data.

This single image based algorithm shows good performance and it is robust to occlusions of the scene lines. However, depending on the complexity of the scene, misclassifications may occur. In order to improve the robustness of the approach, we use the spatial layout recovery algorithm for single images to extract the hypothesis for a whole sequence of consecutive images. The idea is to combine the information obtained from the set of hypothesis to get a better approximation of the real structure. Regarding the use of a sequence of images to estimate the spatial layout, a related work is [12], where motion cues are used to compute likelihoods of indoor structure hypotheses by comparing the predicted location of point features on the environment model to their actual tracked locations in the standard image stream. Scene understanding has been also considered by combining geometric and photometric cues learned across multiple standard views [13]. However, that approach requires solving the stereo problem whereas our method relies on the monocular information extended through a sequence of images without performing feature matching. In particular, our second contribution is a matching-free homography-based procedure that extends the single view approach to sequences of images improving the accuracy and robustness of the results. In our proposal, we avoid the necessity of classical methods of the prone to error feature matching process, especially harder in omnidirectional images, improving efficiency and robustness. In order to be able to compare and integrate several hypotheses, all of them have to be projected over the same frame. This projection is carried out by using homographies of the floor. The homography computation procedure presented is also novel since it uses geometrical constraints to reduce from four to one the minimal set of line correspondences necessary to compute each homography, which allows to skip computational expensive matching algorithms. Once the different hypotheses for some consecutive frames of the sequence are projected on an image, the algorithm looks for the one with the highest similarity rate with respect to the other hypotheses. This one is selected as a basic layout, and later on is averaged with the characteristics of the other hypothesis depending on how similar they are.

In our previous work [14], a preliminary version of our single-view spatial layout recovery approach was presented. Here we contribute with an improved version of that method, extend the experimental results with different datasets, and propose the matching-free method for sequential propagation of the layout. Experimental results showing the improvement of that propagation on image sequences are also included.

The paper is organized as follows. In Section 2 we present the method for obtaining the spatial layout recovery from a single image. The following Section 3 presents the procedure to propagate the hypotheses along a sequence of images avoiding feature matching. In Section 4, the proposed approach is tested with real images of three datasets acquired with different cameras. Finally, conclusions are given in Section 5.

2. Single view spatial layout recovery

In this section, we describe the proposed algorithm to come up with the spatial layout of the scene from a single omnidirectional image. We start extracting lines from the image, which are then classified according to their orientation in order to carry out the estimation of vanishing points (VPs). Combining this information with a set of geometrical constraints we generate hypothesis about the floor contour. From the classified lines, a set of points is selected. These points are used to fit conic lines which represent plausible wall–floor boundaries. Then, a conservative four walls-room hypothesis is generated by selecting the four most voted conic lines. Finally, the initial hypothesis is expanded, according to the image data distribution, to obtain a representative hypothesis.