Pattern Recognition 48 (2015) 2164–2184

Contents lists available at ScienceDirect

Pattern Recognition

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

Two-view line matching algorithm based on context and appearance in low-textured images

Juan López, Roi Santos*, Xosé R. Fdez-Vidal, Xosé M. Pardo

Centro de Investigación en Tecnoloxías da Información (CITIUS), Universidade de Santiago de Compostela, Campus Vida s/n, 15782 Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 25 February 2014
Received in revised form
19 September 2014
Accepted 27 November 2014
Available online 8 December 2014

Keywords:
Line detection and matching
Man-made environments
Differencial phase congruency
Adaptive incremental
Context-based matching

ABSTRACT

A novel approach for line detection and matching is proposed, aimed at achieving good performance with low-textured scenes, under uncontrolled illumination conditions. Line detection is performed by means of phase-based edge detector over Gaussian scale-space, followed by a multi-scale fusion stage which has been proven to be profitable in minimizing the number of fragmented and overlapped segments. Line matching is performed by an iterative process that uses structural information collected through the use of different line neighborhoods, making the set of matched lines grow robustly at each iteration. Results show that this approach is suitable to deal with low-textured scenes, and also robust under a wide variety of image transformations.

1. Introduction

Metrology is the science of measurement and is essential to ensure the suitability of industrial parts, their calibration and quality control. Several high-accuracy measurement systems capable of 3D measurement have been developed, such as coordinate measuring machines, laser-trackers, and optical scanners. The problem is that these systems are either expensive or impracticable in many industrial environments, such as in shipbuilding or in the inplace repairing of windmill turbine blades, where very large pieces are handled [1] under uncontrolled illumination conditions. In order to deal with these issues, photogrammetry has been used frequently in the industry in recent decades. It is an accurate, inexpensive, and non-contact metrology technique which involves estimating the three-dimensional coordinates of points on an object from multiple overlapping photographs taken from different poses, and has emerged as a desirable alternative to conventional metrology.

One of the key steps to obtain three-dimensional coordinates of points from multiple images is to identify homologous features, such as points or lines, across the different views. To achieve a reliable feature matching, most of the industrial photogrammetric systems usually require to manually place a set of retro-reflective landmarks on the surface of the object and control the illumination conditions. Often, these retro-reflective targets are illuminated by a light source near the lens to produce images with a high contrast between the target and the background. Moreover, the combination of low aperture and high shutter speed virtually suppresses background details allowing a reliable segmentation of the targets [2]. The main drawback of the use of targets is the need to attach them to the object surface in such numbers that the presence of a target is assured wherever a measurement is required, task that becomes complicated and expensive in many industrial environments with difficult access, or where large pieces are handled. Besides, the lighting requirements are difficult to satisfy in many outdoor settings when large focal lengths are handled and where the light conditions cannot be controlled.

In recent years, low-level image features have been proven useful for achieving reliable correspondences among images. Several feature-based methods have been proposed such as Moment Invariants [3], Steerable Filters [4], Differential Invariants [5] or SIFT [6]. It was shown that SIFT-based descriptors, which match keypoints by using a scale invariant region detector and a descriptor based on the gradient distribution, outperform the rest [7]. However, the drawback of these SIFT-based methods is that relying on properties like texture or local structure, they become deficient with low-textured objects and homogeneous surfaces, which are typical in industrial environments. In these cases, it is essential to use line segments as matching features due to their abundance on man-made objects, and also because they bring us greater structural information.

* Corresponding author.
E-mail addresses: juan.lopez.gomez@usc.es (J. López), roi.santos@usc.es (R. Santos), xose.vidal@usc.es (X.R. Fdez-Vidal), xose.pardo@usc.es (X.M. Pardo).

http://dx.doi.org/10.1016/j.patcog.2014.11.018

Due to several inherent difficulties only a few methods for automatic line matching are reported in the literature. These difficulties include the inaccurate locations of line endpoints, object occlusions leading to missing line counterparts, the fragmentation of lines, often causing the loss of topological connections among line segments, and also the lack of a global geometric constraint such as the epipolar constraint in point matching. Existing approaches for line matching can be divided into three types: those that match individual line segments, those that match groups of line segments, and those that use some point correspondences or epipolar constraints of line endpoints, to perform line matching. The main drawbacks of the last kind of approaches, such as the ones of Schmid and Zisserman [8] and Fan et al. [9], are the requirement of knowing the epipolar geometry in advance, and the reliance on some point correspondences which makes them inappropriate for low-textured scenes, typical in industrial environments.

Most of the approaches to match individual line segments are based on appearance similarities of the line segments, such as Bay et al. [10] where color histograms are used to obtain an initial set of candidate matches which grows iteratively, or Wang et al. [11] which defines MSLD line descriptor by using a SIFT-like strategy. However, the sole reliance on the appearance of lines also makes these methods inappropriate for industrial scenes, with textureless surfaces. Recently, Zhang et al. [12] have presented a line matching algorithm which considers not only the local appearance of segments but also the direction of lines. Direction histograms are used to estimate whether there is an approximate global rotation angle among image pairs, and if so, this angle is used to reduce the candidate matchings. Although it has been shown that this method achieves better results, the problem remains when there is no accepted rotation angle among images.

The main advantage of the methods that match groups of line segments is that more geometric information is available for disambiguation. In Wang et al. [13], line segments are grouped based on the spatial proximity and relative saliency, and then these groups are matched by using angles and length ratios to describe geometric configuration of their segments, and also average gradient magnitudes to describe appearance information. This strategy is shown to be useful to deal with large viewpoint changes and non-planar scenes. However, in order to improve the repeatability of line signatures among images, a multi-scale scheme for line extraction is applied to divide all the curves in many consistent ways, so that each curved connected-edge is polygonized at various scales and all possible segments are kept. This overestimation in the number of line segments makes this method computationally expensive. In Kim et al. [14], intersecting line pairs in 2D images that are coplanar in 3D are used as matching features by using Line Intersection Context Features (LICF). However, the method is restricted to coplanar lines and also due to the limited discriminative power of the similarity measure based on LICFs the method needs epipolar constraints known in advance or a refinement stage for estimating the camera geometry.

This paper presents a novel approach for line detection and matching aimed at achieving good performance with industrial image pairs, characterized by containing low-textured objects and homogeneous surfaces, and where illumination conditions may vary significantly. The main contributions are

- The detection of lines through a fusion process which combines all the multi-scale information obtained from a Gaussian scale-space to merge those line segments that correspond to fragments of the same line. This multi-scale fusion process appears to be essential in minimizing the number of fragmented and overlapped lines in the images, achieving a total number of extracted lines very close to the real number of perceived lines.
- An iterative method for line matching by means of their local appearance and their geometric properties that takes into account the geometrical arrangement among several kinds of neighboring segments. These different kinds of line neighborhoods are used to collect structural information around lines which is used to refine on-the-fly the similarity measure at each iteration, making the set of matched lines grow robustly.

The paper is organized as follows: Section 2 gives an overview of the approach; Sections 3 and 4 describes the method for line detection and line matching respectively; in Section 5 its performance is compared with two state-of-the-art line matching approaches, and finally, the main conclusions are summarized in Section 6.

2. Overview of the approach

An overview of our proposal for line detection and matching is shown in Fig. 1. The line detection algorithm is conducted independently for each one of the input images, and involves several steps. Firstly, a scale-space pyramid is generated by Gaussian blurring and downsampling the image in order to use multi-scale information to make detection more robust. Then, for each one of these scaled images, the edge features are detected by means of a phase-based feature detection, to achieve robustness to different illumination conditions. The next step in the process comprises the extraction of line segments based on continuity criterion, and finally, a fusion stage is performed along the scale-space to merge those segments that may correspond to fragments of the same line. This fusion stage takes into account all detected lines along scale-space with the aim of making a more meaningful merging, and hence, a more robust detection.

Once line segments are extracted from each image, they are introduced into the line matching algorithm, but in two different stages. These stages are implemented in an identical manner, differing only in their input set. Thus, in the first stage, the algorithm considers only those lines that have been detected in a robust way along scale-space, whereas in the second stage all unpaired lines are considered for matching. This is done because it is expected that most of the lines which are robustly detected in an image are also robustly detected on the other. Besides, this two-stage design makes the algorithm faster since it does not have to check all possible pairings among all lines at first time, but only those among robust lines, and also makes it more robust since the matching information of the first stage is used as a reference for the second stage, where there is a greater possibility of error due to the larger number of lines.

For each one of the stages, an iterative method is applied to achieve a higher number of matches and also to infer on-the-fly information about the transformation between the two images through the matched set of lines during each iteration. This information is then used in subsequent iterations as a feedback to ensure that the set of matched lines grows robustly. Each iteration is performed in four steps. In the first one, several kinds of neighborhoods are computed for each line, relying on different proximity relationships. The usage of various types of neighborhoods is intended to increase the robustness not only to different image transformations between views, but also to occlusions, line fragmentations and detection failures. After this neighborhood computation, the similarity measure is calculated for each pair of line segments, and then, the strongest correspondences are added to the global set of matched lines on the basis of a matching criteria. Once these steps are completed, a last one is performed to remove the weakest correspondences by checking several conditions over the whole set of matched lines. The loop ends when all the matching lines remain paired with the same partner for at least T iterations. This is done to ensure that the final matches are stable throughout the process. The choice of T is a trade-off between the stability of line matchings and time consumption. Details for each step are described in the following sections.