Optik 126 (2015) 426–431

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Image matching of junction-type feature points based on the third-order moment

Jian Li, Bingxuan Guo*, Zhe Peng

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing of Wuhan University, Wuhan, China

ARTICLE INFO

Article history:
Received 24 October 2013
Accepted 1 May 2014

Keywords:
Image matching
Feature-point detector
Third-order moment
Orientation vector
Junction-type feature points

ABSTRACT

Image matching of junction-type feature points in different views is a difficult process because of the complex characteristics of such corners, including significant translation, rotation, shear and scaling differences. A novel method is proposed where, initially, feature points are extracted based on Harris algorithm. Next, the third-order moment of each candidate point is used to compute its orientation vectors. Finally, an affine-invariant representation is determined by mapping the point and its orientation vectors to an orthogonal frame, and a characteristic scale is selected using the extreme of the LOG (Laplacian of Gaussians). Evaluations of the proposed method on different types of image pairs prove its efficacy with repeatability and number of correct matches (NCMs) compared to Harris-based methods.

1. Introduction

Image matching is of fundamental importance in photogrammetry, remote sensing and computer vision. It has been used in 3D reconstruction, target tracking, image stitching and other applications. Image matching aims to identify the correspondence between two different images of the same scene or objects in different poses, illumination conditions and environments. In this paper, we focus on local feature-based image matching. The challenges of this work lie in the stable and invariant extraction of features in various situations and in the robust matching.

Generally speaking, the framework of a feature-based image-matching method consists of three steps: feature detection, feature description and feature matching [1–3]. The result of image matching largely depends on the stability and invariance of the feature detection [4–6]. A feature point is often used as the local feature because of its stable performance in detection and description. A feature point is usually derived from a circle or window with a certain location and radius; this method is more effective and efficient than using other types of features, such as edges and contours. Therefore, feature points are extensively used in real applications.

Most existing feature-point detectors can be categorized into two types: (i) corner detectors and (ii) region detectors [7]. Corner detectors use changes in the intensity of the gray level of the images and have been developed by a number of authors, including Moravec [8,9], Förstner [10–12], Harris [13] and Hession [14]. Harris proposed the use of a structure tensor, a matrix of partial derivatives, to represent the gradient information in the fields of image processing and computer-vision research. This method is referred to as the Harris corner detector, and the eigenvalues of the matrix determine whether a point is a corner. The Förstner operator also utilizes a corner measure defined by the structure-tensor matrix. It was developed with the intent of creating a fast operator for the detection and precise location of distinct points, including corners and centers of circular image features. The threshold is determined by local statistics.

Region detectors include the scale-invariant feature transform (SIFT) algorithm [15,16], the maximally stable external region (MSER) algorithm [17], the Harris-affine detector and the Hession-affine detector [18]. SIFT uses a Gaussian weighting function to assign a weight to the gradient magnitude of each sample point when computing the descriptor. It places less emphasis on gradients that are far from the center of the descriptor. As a result, the problems caused by applying SIFT without affine invariance can be partially offset. The concept of a terrain watershed was proposed for the extraction of MSERs. A gray-level image is divided into a series of binary images using various thresholds to extract the MSER. The Harris-affine and Hession-affine methods detect interest points using the multi-scale Harris corner measure and the Hession corner measure, respectively, and iteratively select the characteristic scale to determine the affine-invariant region associated with an interest point using the second-moment matrix.

A junction-type feature point is a type of corner, such as a roof corner, a window corner or the corner of a field boundary. In

* Corresponding author.
E-mail address: bingxguo@gmail.com (B. Guo).

http://dx.doi.org/10.1016/j.ijleo.2014.05.054
0030-4026/

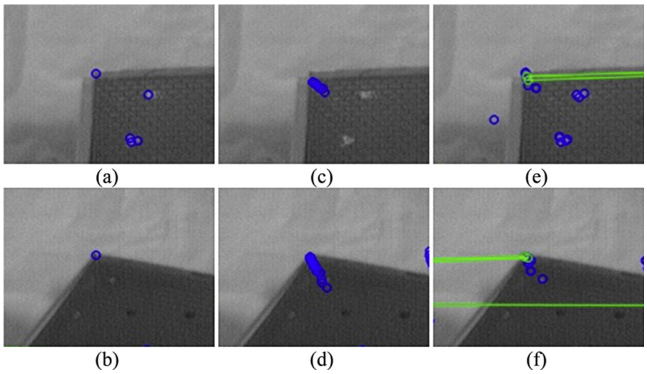

Fig. 1. The extraction of features from a challenging pair of images of a roof corner from our test data using three algorithms (left: Harris, center: Harris-affine, right: our algorithm). The upper row of images was recorded in a horizontal direction of 0°, whereas the lower row of images was recorded at 60°. The left-hand column shows the results of matching using the Harris detector, the middle column shows the Harris-affine results, and the right-hand column shows the results of our algorithm. All three algorithms used SIFT descriptors. In each image, small circles represent extracted features, whereas linked lines represent relations between matched features.

photogrammetry and remote sensing, junction-type feature points are widely used as image features, primarily because of their invariance with respect to imaging geometry and because they are easily perceived by a human observer. The image matching of junction-type feature points, as illustrated in Fig. 1, is a vital process in many tasks, such as image calibration, localization, change detection and 3D reconstruction.

The matching of junction-type feature points in various views is a difficult task, as Fig. 1 demonstrates, because right-angle junctions may be deformed into non-right-angle junctions. From different viewpoints, the junction may appear to undergo various geometric deformations, such as translation, scaling, rotation and shear deformation. For this reason, the two existing methods demonstrated here cannot effectively treat this problem. In (a) and (b), the Harris detector was used to extract the junction-type feature points; however, when the SIFT descriptor was applied in the local area of these feature points, it failed to find a match. The Harris-affine detector also failed to match the junction-type feature points, as shown in (c) and (d). However, the method proposed in this paper extracted the correct correspondences, as shown in (e) and (f). It can be concluded that the failure to match junction-type feature points in differing views is caused by geometrical deformation. It is also apparent that, unlike the Harris detector, the Harris-affine detector extracts an affine-invariant local area for blobs or regions and yields many redundant points at different scales. This issue motivated us to design a novel method of extracting corresponding junction-type feature points under view variation.

In this paper, a novel method, the TMAI (Third-order Moment Affine-Invariant) algorithm, is proposed. First, the Harris detector is used to extract candidate points. Second, the third-order moment of each candidate point is determined, and the orientation vectors of the feature points are extracted. Next, an affine-invariant representation is computed by mapping the region formed by the corner point and its orientation vectors to an orthogonal frame. Finally, the SIFT descriptor is used to identify correspondences between TMAI feature points. This approach can effectively identify correspondences between junction-type feature points and thus provides the required features for 3D reconstruction and other processes in photogrammetry and remote sensing.

The remainder of the article is organized as follows: in Section 2, we introduce the four main steps of the algorithm. In Section 3, an experiment and its analysis are discussed. Finally, in Section 4, we present our conclusions.

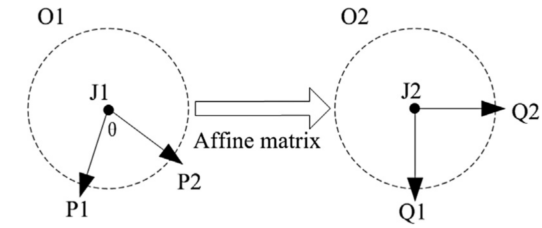

Fig. 2. Geometry of TMAI. Left: J1 is the feature point, and P1 and P2 are the two local orientation vectors, which are determined by the third-order moment. O1 is the local area of J1. Right: For the orientation, P1 and P2, with an enclosed angle of θ, are mapped to an orthogonal frame using the affine transform matrix. The transform maps the feature point J1 to the origin and P1 and P2 to the axes of the coordinate system in a local area, as shown in the right-hand image.

2. Proposed TMAI algorithm

The geometry of TMAI is illustrated in Fig. 2. In the following, we describe the 4 main steps of TMAI in greater detail.

2.1. Harris corner-point extraction

The Harris operator [13] is a well-known corner-point-detection algorithm, and it is used in this research to detect and extract feature points. The Harris detector is defined as the positive local extreme of the following operator:

$$R = \det(M) - k(trace(M))^2 \tag{1}$$

where $\det(M) = f_x^2 \cdot f_y^2$ and $trace(M) = f_x^2 + f_y^2$. $\det(M)$ and $trace(M)$ are the determinant and trace of the matrix M, and the scalar k is an empirical value. The matrix M, which is related to the auto-correlation function, is given by

$$M = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \otimes \begin{bmatrix} f_x^2 & f_xf_y \\ f_xf_y & f_y^2 \end{bmatrix} \tag{2}$$

where $\otimes$ is the convolution operator and f_x^2 and f_y^2 are the gradients of the gray-level intensity in image along the x and y axes, respectively.

2.2. Abstracting two orientation vectors of the feature point

The estimation of the local orientation of feature points in images can be conceived as a problem of finding the minimum gray-level variance axis in a local neighborhood [19]. To find two orientation vectors, let us now assume that within Ω, $f(x)$ is additively composed of two non-opaque oriented sub-images $f_1(x)$ and $f_2(x)$ as follows:

$$f(x) = f_1(x) + f_2(x), f_1, f_2 : R^2 \to R. \tag{3}$$

where f denotes a bivariate gray-level image. Assume that $K(\Phi)$ is the gray-level variance along orientation Φ. Generally, because of noise and the fact that the orientation may not be perfectly constant over Ω, it is impossible to find an orientation Φ such that $K(\Phi) = 0$. We therefore seek to minimize $K(\Phi)$.

$$\Phi = arg \min_{-\pi/2<\phi<\pi/2} K(\Phi). \tag{4}$$

$$K(a) = \int_{\Omega} [a^T df]^2 d\Omega = a^T Ta = 0. \tag{5}$$

where T is a third-order moment, a is a three-dimensional vector, and df is the second derivative of $f(x)$, which is given by $df = (f_{xx}, f_{xy}, f_{yy})^T$, or the square of the derivative, which is given by $df_0 = (f_x^2, f_{xy}, f_y^2)^T$. The 3×3 tensor T is the result of taking the outer