Infrared Physics & Technology 73 (2015) 1–7

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Detecting lines and building intersection correspondences by computing edge oriented histogram on multi-sensor images

Yong Li *, Xiang Shi, Hongbin Jin, Zhigang Wen

Beijing Key Laboratory of Work Safety and Intelligent Monitoring, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Article history:
Received 8 June 2015
Available online 24 August 2015

Keywords:
Line intersection
Correspondence
EOH
Multi-sensor

ABSTRACT

This paper proposes an approach to establishing correspondences for line intersections on multi-sensor images containing lines. Keypoints have been widely applied in a variety of computer vision fields. On multi-sensor images the number of keypoints may be much smaller than on single-sensor images due to the lack of texture. In addition, the multimodality often causes the incorrect assignment of main orientation to keypoints. Observing this, this paper proposes extracting line intersections as keypoints and utilizing one line as the main orientation to compute edge oriented histogram (EOH) descriptor. For EOH descriptor, gradient orientation is employed to compute the filter responses. Experimental results show that the proposed method can build more reliable keypoint matchings on challenging image pairs such as visible image and middle/long-wave infrared images.

1. Introduction

Image registration is a process of searching for feature correspondences between two images of the same scene taken from different viewpoints, and/or by different sensors [1,2]. Multispectral images are captured at specific spectral bands. Dividing the spectrum into multiple bands, we can employ instruments that are sensitive to particular wavelengths to collect the light of a spectral band. The imaging spectral bands include visible spectrum, near infrared (NIR), short-wave infrared (SWIR), middle-wave infrared (MWIR), long-wave infrared (LWIR). In a wide sense, multispectral images also include SAR (synthetic aperture radar), CT, MRI images, etc. Multispectral imaging provides richer information as different spectral bands capture different physical properties of objects/scenes [3,4]. For example, Band 1 of Landsat 8 excels at imaging shallow water and tracking fine particles like dust, while Band 5 measures the near infrared (NIR) which can help get indexes like NDVI for a precise measurement of plant health. Multi-sensor imaging technique has been applied in a wide range of fields, e.g., remote sensing surveillance, stereo and motion estimation, etc.

Effectively utilizing multispectral images necessitates image registration as the first step, and establishing reliable salient point matches is a method of accomplishing registration. However, it is still an open and challenging problem to build keypoint matches, especially on multi-sensor (multi-source) images. This paper aims to establish reliable keypoint matchings on multi-sensor images containing lines. We propose an approach to establishing correspondences for line intersections by computing descriptors for intersections as keypoints. Typical images include man-made scenes such as buildings [5] and remote sensing data containing geological structure. On such images, line intersections can serve as keypoints and be matched by their associated descriptors.

1.1. Related work

A variety of keypoints and descriptors have been proposed in the past. Detecting keypoints can date back to Harris corner [6] that detects corners to be points of large variations in neighborhood. Matas et al. [7] propose maximally stable extremal regions that are invariant under transformations. The scale invariant feature transform (SIFT) [8] has been proposed and applied to a wide range of images. SIFT detects keypoints invariant to scale and rotation that are defined to be the extrema of the difference of Gaussians (DOG). A main orientation is assigned to a keypoint, and the gradient pattern in a local window around a keypoint with respect to the main orientation is computed as its descriptor. Ke and Sukthankar [9] propose PCA-SIFT that computes PCA in the local window around a keypoint instead of gradient pattern.

Bay et al. [10] proposed SURF (Speeded-Up Robust Features). SURF has the same repeatability and distinctiveness as SIFT but is computed faster than SIFT. Rublee et al. [11] propose ORB, an alternative descriptor to SIFT and SURF. Alahi et al. [12] propose Fast Retina Keypoint (FREAK). FREAK is a cascade of binary strings

* Corresponding author.
E-mail address: yli@bupt.edu.cn (Y. Li).

http://dx.doi.org/10.1016/j.infrared.2015.08.013
1350-4495/

computed by comparing image intensities over a retinal sampling pattern. Ambai and Yoshida [13] propose compact and real-time descriptors (CARD). Compared with SIFT and SURF, CARD can be computed rapidly utilizing lookup tables to extract histograms of oriented gradients.

The above mentioned descriptors are initially designed for single-sensor images, and the matching ability of descriptors decreases on multi-sensor images since they are computed by utilizing the local gradient pattern. The nonlinear relationship between intensities of two multi-sensor images often yield different gradient pattern and as a result the above descriptors often fail to work [3]. Observing this, many algorithms have been proposed to adapt SIFT/SURF descriptors to multi-sensor images. The adaptation focuses mainly on two aspects: keypoint detection, and descriptor calculation.

Sedaghat et al. [14] propose uniform robust scale invariant feature matching (UR-SIFT). UR-SIFT selects SIFT features in the full distribution of location and scale where the feature qualities are quantized based on the stability and distinctiveness constraints. Park et al. [15] detect keypoints using higher-order scale space derivatives by observing that the underlying principle for the keypoint detection is to find local extrema in scale space. Specifically [15], extends the idea of utilizing $\frac{\partial L(x,y,\sigma)}{\partial \sigma} = 0$, and explores higher-order scale space derivatives $\frac{\partial L^2(x,y,\sigma)}{\partial \sigma^2} = 0$, $\frac{\partial L^3(x,y,\sigma)}{\partial \sigma^3} = 0$, $\frac{\partial L^4(x,y,\sigma)}{\partial \sigma^4} = 0$, etc. Accordingly, the number of scales in each octave is increased to $(s + 2 + i)$ for detecting the ith-order scale space derivative.

Descriptor adaptation considers two aspects, weighting of gradient magnitude and utilization of gradient direction. Saleem and Sablatnig [16] propose NG-SIFT utilizing normalized gradients. Normalized pixel gradients are of magnitude either equal to 1 for non-zero gradients or 0 for vanishing gradients. This means every pixel of non-zero gradient magnitude contributes as much to the orientation histogram for subblocks. Note, the normalization step does not change gradient directions, so a pixel of normalized gradient will contribute to the same orientation bin as the pixel of non-normalized gradient. A similar work, edge oriented histogram (EOH), is proposed in [17]. In EOH, every edge pixel contributes as much to the orientation histogram. Chen et al. [18] propose partial intensity invariant feature descriptor (PIIFD) in which the weighting of magnitude can be thought of as a digital version of logarithmic quantization to achieve intensity invariance. Formally, the top 20% strongest gradient magnitudes are normalized to 1, second 20% to 0.75, and the last 20% to 0. In a more general sense, the logarithmic weighting can be represented by a sigmoid function of "S" shape, e.g., $S(t) = \frac{1}{1+e^{-t}}$.

Besides the adaptation of magnitude weighting, direction adaptation has also been proposed. Dellinger et al. [19] proposed SAR-SIFT for SAR images. SAR-SIFT utilizes the ratio of average (ROA) for computing gradients. Define the ratio of exponentially weighted averages (ROEWA), and horizontal and vertical gradients are computed based on ROEWA. A fine derivation shows that defining gradients as such involves two procedures. The first is to apply a logarithmic function to average intensity, and the second is to compute gradients with the log-scaled intensities with Sobel operator. NG-SIFT [16] utilizes the same 8 orientation bins as SIFT. The orientation ranges from 0 to 2π with each bin covering $\pi/4$. PIIFD [18] still utilizes 8 orientations but the bin partitioning is $[0, \pi/8], [\pi/8, \pi/4], [\pi/4, 3\pi/8], [3\pi/8, \pi/2], [\pi/2, 5\pi/8], [5\pi/8, 3\pi/4], [3\pi/8, 7\pi/8], [7\pi/8, \pi]$. A gradient direction α and $\alpha + \pi$ go to the same bin in PIIFD under this bin partitioning scheme. EOH [17] utilizes 4 direction bins and a non-direction bin [20]. The 4 bins are exactly the first four bins used in SIFT. Similar to PIIFD, EOH quantizes α and $\alpha + \pi$ to the same bin.

The adapted descriptors including PIIFD still cannot provide a sufficiently high rate of correct matches on multispectral images. Then, a natural problem is how to find/identify correct matches initially built with them.

1.2. Proposed approach

The performance of matching keypoints relies mainly on two aspects, the repeatability of keypoints, and the matching ability of associated descriptors. It is assumed that in this work two images contain corresponding lines, which accordingly means corresponding line intersections exist. The goal is to find intersection correspondences. To this end, edges and lines are extracted from images and their intersections are calculated. The intersections are employed as keypoints, and then EOH descriptor is computed for every intersection. When images contain line features, the repeatability of line intersections is not inferior to commonly used salient points such as the SIFT keypoint or Harris corner. Since EOH does not assign a main orientation to keypoints, it is limited to image pairs containing only translation misalignment. In order to address the misalignment consisting of rotation element, which is typically the case in real applications, we assign a main orientation to an intersection with one of the two lines forming the intersection. When computing EOH, gradient orientation is utilized to substitute for the direction filters.

The rest of the paper is arranged as follows. Section 2 discusses extracting lines from edges, computing line intersections, and assigning main orientation to intersections, Section 3 presents matching intersections with descriptors computed with respect to main orientation, Section 4 investigates the performance for different methods and presents results, and Section 5 concludes this paper.

2. Compute descriptors for line intersections

This section discusses detecting lines, calculating line intersections, and then computing descriptors for intersections. One of the two lines forming an intersection is used as the main orientation for the intersection. Then, EOH descriptor relative to the main orientation is calculated using gradient orientation.

2.1. Extract lines and intersections

Let $I_t(x,y)$ denote the test image and $I_r(x,y)$ denote the reference image. This section discusses detecting lines (line segments) from $I_t(x,y)$ and $I_r(x,y)$. Lines are obtained with the following steps, (1) detecting edges from images with the Canny operator [21]; (2) extracting one-pixel-wide curves from edges [22]; (3) partitioning curves into segments with junctions [23]; (4) fitting lines to segments and preserving the segments of a sufficiently small fitting error. These preserved segments are called line segments; and (5) merging line segments to reduce their number and hence reduce the number of intersections. In addition, empirically longer line segments are more likely generated from real objects and hence are more reliable to be matched than shorter ones. Fig. 1 illustrates an image, detected lines and line intersections from it.

Let $A_1x + B_1y + C_1 = 0$ and $A_2x + B_2y + C_2 = 0$ denote two extracted lines. Their intersections are simply given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix}^{-1} \begin{pmatrix} -C_1 \\ -C_2 \end{pmatrix}. \quad (1)$$

When $\det \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix} = 0, A_1x + B_1y + C_1 = 0$ and $A_2x + B_2y + C_2 = 0$ are parallel and do not form an intersection. With (1), intersections of any two non-parallel lines can be easily obtained. Some intersections formed by extracted lines do not fall within the image domain.