Infrared Physics & Technology 77 (2016) 219–225

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

A novel coarse-to-fine method for registration of multispectral images

Hongbin Jin, Chunxiao Fan *, Yong Li, Liangpeng Xu

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

HIGHLIGHTS

- We use projective transformation to register multispectral images of large depth.
- Global information is exploited with local descriptors to build keypoint mappings.
- The edge overlap information is utilized as similarity metric.
- The outliers of keypoint mappings are determined by multiple steps.

ARTICLE INFO

Article history:
Received 21 November 2015
Revised 8 May 2016
Accepted 9 May 2016
Available online 3 June 2016

Keywords:
Multispectral
Keypoint mappings
Edge
Descriptor

ABSTRACT

Due to non-linear intensity changes between multispectral images, the existed descriptors often yield low matching performance. In order to build reliable keypoint mappings on multispectral images, a novel coarse-to-fine method is designed using projective transformation and the information of edge overlap. The method consists of a coarse process and a fine-tuning process. In the coarse process, initial keypoint mappings are built with the descriptors associated with keypoints and the relative distance constraints are employed on them to remove outliers. In the fine-tuning process, the edge overlap information is utilized as similarity metric and an iterative framework is applied to search correct keypoint mappings. The performance of the proposed is investigated with keypoints extracted by speeded-up robust features. The experiment results show that the proposed method can build more reliable keypoint mappings on multispectral images than existed methods.

1. Introduction

The fusion of multispectral images is widely used in a variety of fields including civilian surveillance, intelligent navigation, and automatic target recognition, etc. [1]. For example, an object that is difficult to detect and/or identify in visible light may be easier to detect in an infrared image. So a fusion image of the visible light and infrared light images contains more information than one band image [2]. The high quality image fusion relies on the reliable registration technique, which aims to search a geometrical transformation that can align two images. Currently, one class of commonly used registration methods is to build keypoint mappings between images and then use them to determine the transformation. Such methods are divided into three steps: detecting keypoints, calculating descriptors, and building keypoint mappings with the descriptors associated with the detected keypoints. Over the past decade, a variety of keypoints and descriptors have been designed. Scale-invariant feature transform (SIFT) [3] and speeded-up robust features (SURF) [4] have been designed to build keypoint mappings between two images on the assumption that corresponding keypoints have similar gradient pattern around them as shown in Fig. 1. Alahi et al. [5] propose a fast retina keypoint which is faster to compute and more robust than SIFT and SURF. Morel and Yu [6] propose ASIFT that is fully affine invariant. Cai et al. [7] then further propose a perspective scale invariant feature transform (PSIFT) using homographic transformation to simulate perspective distortion.

Due to non-linear intensity changes between multispectral images, on which these descriptors have low matching performance [8,9]. Many research have been done to improve the performance of the descriptors designed for single spectrum images (e.g., SIFT) on multispectral images. Observing that the orientation of the corresponding pixels may be same or inverse, Chen et al. propose gradient symmetric-SIFT (GS-SIFT) [10] and partial intensity invariant feature descriptor (PIIFD) [11] descriptor adopting gradient orientation instead of direction. GS-SIFT and PIIFD impose restrictions on the main orientation of keypoints to a number range $[0, \pi)$, which makes the same contribution to orientation histogram for corresponding pixels. Similar to PIIFD, Hossain et al. [12] propose improved symmetric-SIFT (ISS) to address the gradient reversal, region reversal, and the descriptor merging problem.

* Corresponding author.
E-mail address: cxfan@bupt.edu.cn (C. Fan).

http://dx.doi.org/10.1016/j.infrared.2016.05.025
1350-4495/

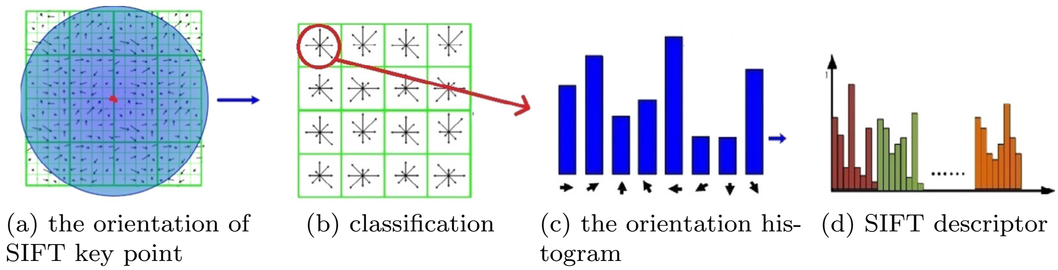

Fig. 1. The process of generating SIFT descriptor. (a) The orientation of pixel's gradient around a SIFT keypoint. (b) Classification of the pixel's gradients according to orientation. (c) A set of orientation histogram created on $4*4$ pixels neighborhoods with 8 bins each. (d) The dimension of SIFT descriptor.

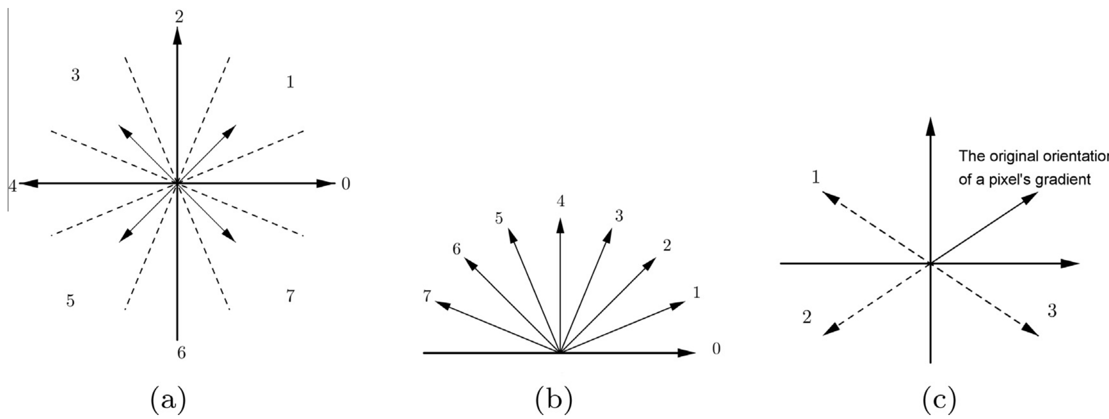

Fig. 2. The orientation of gradient (a), SIFT descriptor contains eight gradient orientations from $[0, 2\pi)$. (b) PIIFD deals with the orientation reversal by limiting the gradient orientation from $[0, \pi)$. (c) Assume a pixel's gradient orientation is in the first quadrant on an image of single spectrum (the solid line), the orientation of its corresponding pixel on another multispectral image may be reversed locating in third quadrant and also may be symmetrical and inverse locating in second quadrant or fourth quadrant (the dotted lines).

Saleem and Sablatnig [13] use normalized gradients as local image features for the description of keypoints in order to achieve robustness against non-linear intensity changes. Dellinger et al. [14] proposed SAR-SIFT for SAR images. Aguilera et al. [15] propose edge oriented histogram (EOH),which utilizes only edge points and five bins for computing descriptors. EOH has a better matching performance on multispectral images than SIFT, but does not assign a main orientation to keypoints.

All these improved methods perform better than those descriptors designed for single spectrum images. However, when the spectral difference of multispectral images increases, the common information between multispectral images will decrease, as described in Fig. 2. The gradient orientation of corresponding pixels may be reversed or may be symmetrical and inverse. This fact leads to corresponding pixels (keypoints) on multispectral images lack of linear relationship, making the existing descriptors difficult to build keypoint mappings on them. A common method is to use random sample consensus (RANSAC) [16] and its variants [17,18] to identify the correct mappings from the initial set built with the associated descriptors. However, because of the low ratio of correct keypoint mappings established on multispectral images, the rate of correct mappings being misidentified as incorrect (and the reverse) will be very high. Consequently, the set of the keypoint mappings preserved by RANSAC still contains a large percent of incorrect mappings, and also, some correct mappings are removed by mistake.

2. Proposed approach

To solve the aforementioned problem, we propose a coarse-to-fine framework to improve the correct rate of keypoint matchings on multispectral images. The framework of the proposed method includes two subprocedures, one is the coarse process and the other is the fine process. In the coarse process, we employ SURF with best-bin-first (BBF) rule [19] to establish an initial set of keypoint mappings and then use relative distance constraints to delete some wrong keypoint mappings. In the fine process, the similarity of edge information [20] on the images is introduced to assess the matching quality of keypoint mappings, incorporating global information in the evaluation of keypoint mappings. The possibility of overlapping edges is defined as the similarity metric. To calculate the similarity metric, our proposal adopts the edge orientation histogram which describes the shapes and contours from images. After that, an iterative algorithm is designed to search best keypoint mappings by iterating through all the combinations (quadruples) of four keypoint mappings. Once quadruples of keypoint mappings have been successfully traversed, the best keypoint mappings can be achieved by sorting through the value of the similarity metric. Note, this paper is focusing on describing the architecture of the algorithm. A variety of other keypoints have been used for building initial keyoint mappings to test the performance of proposed method.

The proposed method has three advantages. Firstly, we use projective transformation to calculate the misalignment, which can register multispectral images containing largely varying scene depth. Secondly, we exploit global information to build keypoint mappings in conjunction with local descriptors. The incorporation of global information ensures that the built keypoint mappings are consistent well with the content of entire images. Thirdly, the proposed method is composed of multiple steps to construct keypoint matchings. This makes up for the shortcoming that the idea of deciding on keypoint mappings in only a one-pass process. We