The Egyptian Journal of Remote Sensing and Space Sciences (2011) **14**, 63–72

National Authority for Remote Sensing and Space Sciences

The Egyptian Journal of Remote Sensing and Space Sciences

www.elsevier.com/locate/ejrs
www.sciencedirect.com

ORIGINAL ARTICLE

Multiscale registration of remote sensing image using robust SIFT features in Steerable-Domain

Xiangzeng Liu [a,*], Zheng Tian [a], Chunyan Chai [b], Huijing Fu [a]

[a] *Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an 710129, China*
[b] *Shanxi Provinal People's Hospital, Xi'an 710068, China*

Received 11 April 2011; accepted 23 August 2011
Available online 2 October 2011

KEYWORDS

Image registration;
Steerable pyramids transform;
SIFT;
Remote sensing image

Abstract This paper proposes a multiscale registration technique using robust Scale Invariant Feature Transform (SIFT) features in Steerable-Domain, which can deal with the large variations of scale, rotation and illumination between images. First, a new robust SIFT descriptor is presented, which is invariant under affine transformation. Then, an adaptive similarity measure is developed according to the robust SIFT descriptor and the adaptive normalized cross correlation of feature point's neighborhood. Finally, the corresponding feature points can be determined by the adaptive similarity measure in Steerable-Domain of the two input images, and the final refined transformation parameters determined by using gradual optimization are adopted to achieve the registration results. Quantitative comparisons of our algorithm with the related methods show a significant improvement in the presence of large scale, rotation changes, and illumination contrast. The effectiveness of the proposed method is demonstrated by the experimental results.

1. Introduction

Image registration is a process of determining the point-to-point correspondence between two images of the same scene, which are acquired by different sensors or by the same sensor at different times or with different parameters (Zitová and Flusser, 2003). Wavelet decomposition of the images was recommended for the pyramidal approach due to its inherent multi-resolution character. Le Moigne et al. (2002) utilized maxima of wavelet coefficients as the basic features and proposed a correlation-based automatic registration algorithm, which achieved higher computational speeds for comparable accuracies. Recently, several registration algorithms that combines wavelet-based pyramid with other similarity measures were proposed in Cole-Rhodes et al. (2003) and Zavorin and

* Corresponding author. Tel.: +86 13572133645.
E-mail address: lxzccy20062008@126.com (X. Liu).

Peer review under responsibility of National Authority for Remote Sensing and Space Sciences.
doi:10.1016/j.ejrs.2011.08.001

Production and hosting by Elsevier

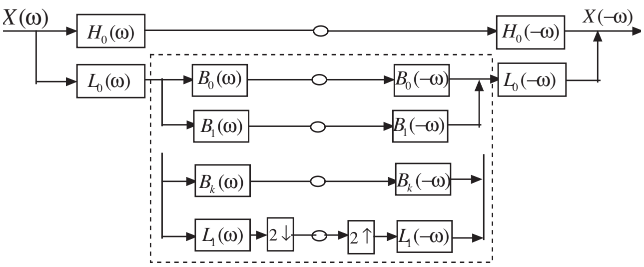

Figure 1 System diagram for a first derivative steerable pyramid.

Le Moigne (2003, 2005). However, orthogonal wavelet transforms are lack of translation and rotation invariance, which are considered as the source of mismatches when the amount of scene increases. Freeman and Adelson (1991) proposed a steerable pyramid transform, which is multi-scale, multi-orientation, and self-inverting image decomposition, and it has the advantage that the sub-bands are both translation and rotation invariant (Simoncelli and Freeman, 1992). These characteristics make it useful for registering remote sensing images, and therefore the registration with steerable pyramid transform will be more stable and robust under large rotation and scale changes (Liu et al., 2002).

Local characteristics have shown to be well adapted to registration and recognition, as they allow robustness to large rotation and scale variations. The difficulty is to obtain invariance under arbitrary viewing conditions. Different solutions to this problem have been developed over the past few years. Lowe (1999, 2004) presented the SIFT method to extract distinctive invariant features from images. These features are invariant to image scale and rotation, and provide robust matching across a substantial range of affine distortion, addition of noise, and changes in illumination. Mikolajczyk and Schmid (2005) compared the performance of several descriptors for local interest regions, and concluded that the performance of the SIFT-based descriptors was best. Later, the SIFT method has been successfully applied in remote sensing image registration (Yu et al., 2008; Yi et al., 2008). Delponte et al. (2006) proposed SVD matching using SIFT features to cope with large scale variations. However, the SIFT descriptor is not invariant to affine transformation. Therefore, although decreasing the matching threshold in the SIFT algorithm can increase the matching features number, it increases mismatches simultaneously under the large scale variances, rotations and changes in viewpoint.

To achieve a robust registration for remote sensing images, we propose a multiscale registration method using robust SIFT features in Steerable-Domain. The main contribution of this paper can be divided into three aspects. First, a new robust SIFT descriptor is presented, which is invariant under affine transformation. The improved main orientation makes the feature descriptor robust to the difference in the pixel intensity and gradient intensity. Second, an adaptive similarity measure is developed according to the robust SIFT descriptor and the adaptive normalized cross correlation of feature point's neighborhood. The new measure not only considers the distance between feature points' descriptors but also considers the adaptive normalized cross correlation of the neighborhood of feature points. It enhances the robustness of the proposed algorithm to large changes in viewpoint. Finally, the images along certain orientations are used for initial matching at each layer, which are obtained by applying the steerable pyramid transform to the two input images. They make full use of the structural information of the two input images and enhance the robustness of the proposed algorithm to large rotations.

The rest of this paper is organized as follows. Section 2 gives an overview of the steerable pyramid transform and SIFT algorithm. In Section 3, we describe our algorithm for remote sensing images registration. Experimental results and conclusions are given in Sections 4 and 5, respectively.

2. Related work

In this section, we first review the details of the steerable pyramid transform and then introduce the SIFT algorithm for feature points matching.

2.1. Steerable pyramid transform

The steerable pyramid transform introduced by Freeman and Adelson (1991) is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications. It has been developed in 1990, in order to overcome the limitations of orthogonal separable wavelet decompositions that were popular for image processing. The ''steerable filter'' refers to a class of filters, in which a filter of arbitrary orientation can be synthesized as a linear combination of a set of ''basis filters''. For any function $f(x, y)$, $f^{\theta}(x, y)$ is $f(x, y)$ rotated through an angle θ about the origin. We call $f(x, y)$ is steerable if it satisfies the following equation:

$$f^{\theta}(x,y) = \sum_{j=1}^{M} k_j(\theta) f^{\theta_j}(x,y), \tag{1}$$

where $k_j(\theta)$ are the interpolation functions ($j = 1, \ldots, M$). The basis functions of the steerable pyramid are directional derivative operators that come in different sizes and orientations, and the number of orientations may be adjusted by changing the derivative order. The structure of the steerable pyramid in the frequency domain is shown in Fig. 1. The image is initially divided into high and low-pass sub-bands using filters $H_0(w)$ and $L_0(w)$. The low-pass branch is then further divided into lower-pass and oriented band-pass portions using filters $L_1(w)$ and $B_k(w)$. This lower-pass sub-band is sub-sampled by a factor of 2 in the X and Y directions. The recursive construction of a pyramid is achieved by inserting a copy of the shaded portion of the diagram at the location of the solid circle. The steerable pyramid performs a polar-separable decomposition in the frequency domain, thus allowing independent representation of scale and orientation. More importantly, the representation is translation-invariant and rotation-invariant. More details about steerable pyramid can be found in Freeman and Adelson (1991):

$$k_j(\theta) = \frac{1}{4}[2\cos(\theta - \theta_j) + 2\cos(3(\theta - \theta_j))], \quad j = 1,2,3,4, \tag{2}$$

where $\theta_j = j\pi/4$. According to Eq. (1), an image of arbitrary orientation can be synthesized as a linear combination of the four orientation band-pass components at each layer. Therefore, in a fixed orientation, the structural information of the reference and sensed images can be used for image registration.