# A Robust Descriptor Based on Spatial and Frequency Structural Information for Visible and Thermal Infrared Image Matching

Zhitao Fu [a,b], Qianqing Qin [a], Chun Wu [a], Yunpeng Chang [a], Bin Luo [a], *

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China - (zhitaofu,cugwhu,c1571720262)@126.com, qqqin@lmars.whu.edu.cn, luob@whu.edu.cn
[b] Information and Network Centre, Yunnan Minzu University, Kunming, China

**Commission III, WG III/6**

**KEY WORDS:** Multimodal Descriptor, Visible and Thermal Infrared Image Matching, Spatial Structural Information, Frequency Structural Information, Multi-spatial Regions Division.

**ABSTRACT:**

Due to the differences of imaging principles, image matching between visible and thermal infrared images still exist new challenges and difficulties. Inspired by the complementary spatial and frequency information of geometric structural features, a robust descriptor is proposed for visible and thermal infrared images matching. We first divide two different spatial regions to the region around point of interest, using the histogram of oriented magnitudes, which corresponds to the 2-D structural shape information to describe the larger region and the edge oriented histogram to describe the spatial distribution for the smaller region. Then the two vectors are normalized and combined to a higher feature vector. Finally, our proposed descriptor is obtained by applying principal component analysis (PCA) to reduce the dimension of the combined high feature vector to make our descriptor more robust. Experimental results showed that our proposed method was provided with significant improvements in correct matching numbers and obvious advantages by complementing information within spatial and frequency structural information.

## 1. INTRODUCTION

Image matching is a very important process in image registration, image fusion, image mosaic and changing detection, just to mention a few. The points to be matched are described through their surrounding information providing rich representations. The goal of this process is to design good descriptor. The descriptor should have good distinguishability from other similar regions of the feature points. The descriptor should also be robust to different image transformations. In this paper, we focus on designing a robust descriptor which is robust against non-linear intensity variation between visible image and thermal infrared images, both of which belong to different modalities and different spectra.

Because of their robustness to geometric and illumination differences, most of the matching algorithms are based on local invariant features. Feature detection, feature description and feature matching are three main steps in these algorithms. The most familiar and popular matching method is SIFT (Scale Invariant Feature Transform) (Lowe, 2004), which uses the distribution of gradients for description. Although the SIFT algorithm can give good matching results for visible image pairs, the percentage of matches becomes worse for multi-spectral or multi-modal image pairs for their non-linear intensity differences, particularly in visible image and thermal infrared image pairs.

To cope with such non-linear pixel intensity differences, several modified methods to SIFT have been proposed. A gradient orientation modification SIFT (GOM-SIFT) descriptor was proposed to limit the gradient orientation within (0, π] (Yi, 2008). Orientation restricted SIFT (OR-SIFT) was proposed to combine the SIFT descriptor elements in the opposite orientation directions (Vural, 2009). A normalized gradients SIFT (NG-SIFT) descriptor was presented to compute descriptors using a normalized gradient (Saleem, 2013). These modified SIFT methods could work for image registration that the grey level intensities of the pixels in the same region between multispectral images are quite different, inverse better (Saleem, 2014).

Shape context (Belongie, 2000) and Edge Oriented Histogram Descriptor (EHD) (Aguilera, 2012 and Mouats, 2013) are very similar to SIFT descriptor, which computes a histogram based on the gradient distribution around the detected regions. However，both of them computed the histogram by describing the edge distribution around the detected region. Shape context was defined as a joint histogram of the angle and log distance of a point relative to other points. EHD only utilized the edge points which was combined with the spatial distribution of four directional edges and one non-directional edge (5 bins totally). The EHD descriptor was performed better than SIFT, GOM-SIFT and OR-SIFT for multispectral images matching (Aguilera, 2012). A Local Self-Similarity (LSS) descriptor was proposed to capture internal geometric layouts through accounting for small local affine deformations (Shechtman el at., 2007). The self-similarity of colour, edges, repetitive patterns and complex textures were captured in a single unified way. The modified LSS descriptors were applied to multispectral image matching (Kim, 2014 and Sedaghat, 2015) and image registration (Ye, 2014).

* Corresponding author