DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

Incorporating Gradient Magnitude in Computation of Edge Oriented Histogram Descriptor

Liangpeng Xu, Yong Li, Chunxiao Fan, Hongbin Jin and Xiang shi; Beijing Key Laboratory of Work Safety and Intelligent Monitoring, School of Electronic Engineering, Beijing University of Posts and Telecommunications; Beijing, 100876, China

Abstract

This paper proposes an approach to employing the gradient magnitude in computing EOH descriptors. EOH has a better matching performance than SIFT (scale invariant feature transform) on multispectral images but does not utilize the gradient magnitude. In EOH, every edge pixel has the same contribution to the orientation histogram, which suppresses the usage of gradient magnitude. Observing this, we propose utilizing gradient magnitude with a logistic sigmoid function. The gradient magnitude of a pixel serves as the input to a sigmoid function, and the output is used as the weight of the pixel. Experimental results show that the proposed approach performs more robustly than the original EOH on multispectral images.

Introduction

Image registration has been widely applied in computer vision. The applications of registration include stereo vision, 3D scene reconstruction, and human activity recognition, etc. In medical field, with medical image registration doctors can find spatial mapping of corresponding anatomical or function of structures among images. In military field, radar image registration technique can help detect sensitive changes in battlefield.

Image registration methods are classified as two kinds: area-based and feature-based. Area-based methods choose a similarity metric and maximize it to find parameters of transformation functions. Feature-based methods often include three steps: keypoint detection, keypoint description, and match with discriptors. The matching ability of descriptors is measured with the repeatability and distinctiveness, and in practice a trade-off is often made between them.

On single-spectral images, as a feature-based method, SIFT [1] and its variants with post processing techniques (e.g., RANSAC) have witnessed many successful applications. The keypoint is defined to be the extrema of the difference of Gaussians (DOG). The local gradient pattern around a keypoint with respect to an assigned main orientation is computed as its descriptor. Bay *et al.* [2] proposed SURF (Speeded-Up Robust Features). SURF has the same repeatability and distinctiveness as SIFT but is computed faster than SIFT.

Recently, multispectral image registration became an attractive research topic since they provide a rich representation of scene richer [3]. From different spectral images, we can acquire different information. But on multispectral images, SIFT descriptors generate few correct mappings, since on multispectral image, intensities are not linearly relate, so we can not use the algorithms based on intensity gradient magnitude directly.

Saleem and Sablatnig [4] proposed DE-SIFT that computes descriptors using differential excitation gradient. DE-SIFT outperforms SIFT on image pairs of a visible image and a near infrared image. Chen *et al.* [5] proposed LC-SIFT that use local contrast gradient magnitude to improve the performance of SIFT in multisensor images registration. Dellinger *et al.* [6] proposed SAR-SIFT for SAR images. SAR-SIFT uses a new gradient computation method, gradient by ratio (GR), which is robust to speckle noise so that it performs better on SAR images than SIFT. Since on mutlispectral images, gradient orientation can reversed in some cases, Yi *et al.* [7] proposed GOM-SIFT that limits the gradient orientation from $[0, 2\pi)$ to $[0, \pi)$ to solve the problem.

When computing descriptors, SIFT and other SIFT-like algorithms use all the pixel surrounding keypoints. Recently, edge oriented histogram (EOH) [8] was proposed that utilizes only edge pixels and 5 bins for computing descriptors. Compared with SIFT, edge pixels is more stable than the other pixels on multispectral images, so EOH has a better matching performance on multispectral images than SIFT. But EOH does not consider the gradient magnitude of keypoints although the magnitude is not the same in different domain. In this article, we will assign the gradient magnitude to the EOH descriptor to improve its accuracy.

Proposed Method

A new EOH descriptor algorithm that makes use of keypoint gradient magnitude based on image quality was developed to further improve the accuracy of EOH. The original EOH algorithm uses the canny operator [9] to get the edge pixel and then compares the filters result at each edge pixel through the following five Sobel operators. These Sobel operators correspond to the 0°,45°, 90°, 135°, and non-direction as shown in Fig. 1. The filters shown in Fig. 1(a), Fig. 1(b), Fig. 1(c), and Fig. 1(d) are called direction Sobel operators, while the one shown in Fig. 1(e) is called non-direction Sobel operator.

$$\begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix}$$
(a) 0°

$$\begin{vmatrix} -1 & 2 & 2 \\ -1 & -1 & 2 \\ -1 & -1 & -1 \end{vmatrix}$$
(b) 45°

$$\begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{vmatrix}$$
(c) 90°

$$\begin{vmatrix} 2 & 2 & -1 \\ 2 & -1 & -1 \\ -1 & -1 & -1 \end{vmatrix}$$
(d) 135°

$$\begin{vmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{vmatrix}$$
(e) non-direction

Figure 1. *The five sobel operators used in [8, 10]. The filters compute directional derivatives in 0°, 45°, 90°, 135°, and non-direction. (a), (b), (c), and (d) are called direction filters, and (e) called non-direction filter.*

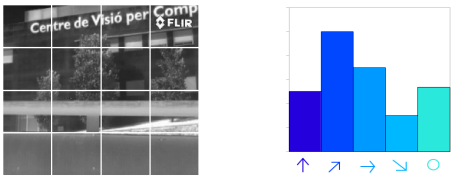
The five operators could give an idea of the strength of the

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

gradient in five particular directions by using equation (1),

$$G_i(x,y) = Sobel_i * I_c(x,y), \quad i \in [1,5]. \tag{1}$$

After comparing these five filter result, the maximal value will contribute 1 to the keypoint. For example, if the filter result of a edge pixel is $(8,10,11,13,11)$, the descriptor vector is $(0,0,0,1,0)$. As the surrounding area of a keypoint is designated as a small $4*4$ part, the descriptor vector of one keypoint has $4*4*5 = 80$ components. Fig. 2 shows 4×4 region divisions of a infrared image and the Edge orientation histogram for 5 orientations. The steps of the proposed algorithm is listed in Algorithm 1.



(a) 4×4 region divisions (b) Edge Orientation Histogram

Figure 2. *Edge orientation histogram of the infrared image. 2(a) the* 4×4 *region divisions of a infrared image and 2(b) the Edge orientation histogram for* 5 *orientations: vertical, horizontal, diagonals and non-directional [11].*

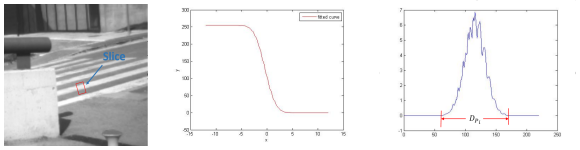
From the original computation method of EOH, we can learn that the gradient magnitude of pixels are treated as the same. It merely considers the main direction of gradient to decide which bin the pixel belong to. Keypoint gradient magnitude $G(x,y)$ can be calculated as equation (2),

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}, \tag{2}$$

it contains so much information that can improve the accuracy of keypoint matching, which have not been used by the original EOH. Simultaneously, gradient magnitude behaves unstable enough on multispectral image, which deserves great attention to. Therefore, we are supposed to operate regularization before using to avoid the deviation from the nonlinearity of gradient magnitude on multispectral image. Meanwhile, compared to high quality visible images, infrared images vary in their quality. For the low ones, noise will have great impact on gradient magnitude. As result, we make classification on image, construct EOH descriptors with adopting different stairs of gradient magnitude as well as operate enhancement on worse images.

Quality Classification on Image

Main factors which have impact on image quality are optical distortion, aliasing, motion blur and noise, among which first three ones influence on Point-Spread Function [12]. Hence the magnitude of PSF (Point-Spread Function) could measure quality of image. As known from [13], blurred edge width of image equals the diameter of Point-Spread Function in that direction. Thus the practice is detecting image's straight line edge with Canny operator and Hough transformation [14], then sampling several points $P_i, i = 1,2,\ldots,N$ evenly on the edge, get Slice S_{P_i} whose normal direction is perpendicular to edge's tangent and use bilinear interpolation algorithm to figure out the pixel value, as shown in 3(a). To eliminate interference from noise, we take average of pixel value corresponding in different segments of the same direction after normalizing them. Eventually, we count the distance D_{P_i} between the maximum and minimum for S_{P_i} and take the largest one as the Point-Spread Function diameter of image, as shown in 3(c).



(a) Slice (b) Curve Fitting (c) Diameter of blur kernel

Figure 3. *Estimate the diameter of the Point-Spread Function. 3(a) a slice extracted from the infrared image, 3(b) the curve fitting of slice and 3(c) the differential of 3(b). We can easily get the Diameter of blur kernel from 3(c).*

Algorithm 1: The steps of original EOH algorithm

Input: Image $I_c(x;y)$ to be registered.
Output: EOH descriptor for each keypoint.

Use the Canny detector to extract edges from the image;
Detect Sift feature points from image, denoted by KP_i, $i = 1,\ldots,N_s$;
for $i = 1 : N_s$ **do**
  Extract an area with point KP_i as center, 40 as edge length;
  Divided the region into 4×4 region divisions;
  for *each region* **do**
    Extract edge pixels from the result of Canny which denoted by EP_i, $i = 1,\ldots,N$;
    for $j = 1 : 5$ **do**
      Calculate the strength of the gradient by using Sobel operators: $G^j_{EP_i}(x,y) = Sobel_j * I_{EP_i}(x,y)$;
    $G_{EP_i}(x,y) = max(G^j_{EP_i}(x,y), j \in [1,5])$;
    $G_{EP_i}(x,y)$ contribute 1 to the 1×5 descriptor vector of KP_i;
  With combining 16 regions' 1×5 descriptors vector orderly, we construct a 1×80 descriptor vector for KP_i.
return the descriptor vectors of KP_i, $i = 1,\ldots,N_s$;

Finally, we classify images into 3 sets, among which low quality images D_{P_i} range from 10 to 15, the mid ones range from 5 to 10 and the high ones under 5.

Image enhancement

The low quality image enhancement, consists of the gray balance and Wiener filter, is operated for better detection of SIFT keypoints and straight line edge in infrared images [15]. As known to all that infrared image appears to be dark, gray histogram of it will collapse into a narrow section, which could influence on feature extraction. Thus gray balance operation which deallocates the pixel value of images is needed for a well-proportioned histogram. As for the gray balance, we assume

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

$I(x;y)$ is the observed image, it can be amended to get I_c according to equation (3),

$$I_c(x,y) = I_{black} + (I(x,y) - I_{min}) \times \frac{I_{white} - I_{black}}{I_{max} - I_{min}}. \quad (3)$$

Simultaneously, the image get exposed to noise deviation for dimness of infrared light when imaging. In that case, Wiener filter with Gaussian blur kernel, whose radiometer values 3 and σ is 0.7, is employed on low quality infrared images to reduce impact of noise and recover what the original appearance of images. Fig. 4 shows the result of image enhancement.

(a) original image (b) image enhancement

Figure 4. *Low quality image enhancement. 4(a) the original low quality image and 4(b) the result of image enhancement, which consists of the gray balance and Wiener filter.*

Adopt gradient magnitude to EOH descriptor

In fact, the gradient magnitude is not the same although EOH simply uses edge pixels. It will limits the effect of EOH. While at the same time, we could not ignored that gradient magnitude is not stable on multispectral images, as these images are generated by different sensors, which have varied principles, shown in Fig. 5. So before we employ the gradient magnitude, we must normalized them to reduce the influence caused by the variance of gradient magnitude. Firstly, we can perform sigmoid on each gradient magnitude G_{P_i}, and then quantify it according to several grades. In the surrounding of each salient point candidate, we normalized G_{P_i} as equation (4),

$$G_{norm_{P_i}} = \frac{G_{P_i} - G_{min}}{G_{max} - G_{min}}. \quad (4)$$

And it deserves great attention to the details of sigmoid utilization, as the below equation (5),

$$Sigmoid(G_{norm_{P_i}}) = \frac{1}{1 + e^{-G_{norm_{P_i}}}}. \quad (5)$$

As mentioned before, the images are classified into three classes, say the low, mid and high, by quality. Then, sigmoid calculate the first 20% gradient magnitude of the low quality images, the first 40% of the mid ones and the first 60% of the highs, as shown in equation (6),

$$G'_{P_i} = \begin{cases} Sigmoid(G_{norm_{P_i}}) & \text{if } G_{norm_{P_i}} \geq T, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

T equal to 0.4 for high quality images, 0.6 for mid ones and 0.8 for low ones.

After the interpolation and quantization, pixel values for computing filter response are obtained. The filter responses mostly is defined to be the direction at this pixel and contributes to EOH descriptors.

EOH descriptor, added with the gradient magnitude of keypoint neighborhood, will remarkably improve keypoint matching accuracy by its great information.

Matching Keypoints with Descriptors

After the construction of descriptors, we will match keypoints with those descriptors. The matching function of descriptors is to compare the Euclidean distance of the descriptor vectors [16]. In EOH, the size of a keypoint descriptor vector is 1*80. In proposed method, we will use common method to match descriptors like SIFT and other traditional algorithms. A reference keypoint $K_r^{j_0}$ is defined to be matched to a test keypoint $K_t^{i_0}$ if

$$D(f_t^{i_0}, f_r^{j_0}) < 0.8 \cdot D(f_t^{i_0}, f_r^{j_1}), \quad (7)$$

where $D(\cdot,\cdot)$ is the Euclidean distance and $f_r^{j_1}$ is the second-closest neighbor to $f_t^{i_0}$. The '0.8' in (7) can be changed to 0.6, which means a tighter matching criterion giving fewer matched keypoints.

During these matches, there must be some error ones, we can use techniques include RANSAC [17] and its variant fast sample concensus (FSC) [18], etc to remove these errors. However, when error matches occupy most of the matches, these technique like RANSAC have little effect to improve the algorithm, so the resulting improvement ought to be excluded for comparing descriptors. The most important part is to build a robust descriptor to describe the keypoint.

Experimental Results

This section presents experimental results. Visual matching results are provided firstly, followed by the quantitative analysis on matching results. The proposed method is compared with the original EOH. Two datasets EOIR and VS-LWIR are used for investigating the matching performance. EOIR includes 87 image pairs captured by ourselves, 12 Landsat image pairs from NASA. The 87 image pairs include outdoor and indoor scenes with one image taken with the visible light and the other taken with middle-wave infrared (MWIR) light. The 12 Landsat image pairs are downloaded from http://landsat.usgs.gov/ with one taken with visible band, e.g., Landsat 8 Band 3 Visible (0.53 - 0.59 μm), and the other taken with middle-wave light or Thermal Infrared Sensor (TIRS), e.g., Landsat 8 Band 10 TIRS 1 (10.6 - 11.19 μm). Dataset VS-LWIR is from [8] containing 100 image pairs, one image taken with visible bandwidth (0.4 - 0.7 μm) and the other taken with long-wave infrared bandwidth (LWIR, 8 - 14 μm).

Fig. 6 gives the keypoint matchings built with the original EOH without gradient magnitude normalization and the proposed method. The visible image serves as the reference image and the infrared image is used as the test image. Fig. 6(a), Fig. 6(c) from dataset VS-LWIR and Fig. 8(a) from EOIR show the matching result of original EOH between the reference and the test image. As a comparison, the proposed method provides sufficiently more correct matches on Fig. 6(b), Fig. 6(d), and Fig. 8(e).

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

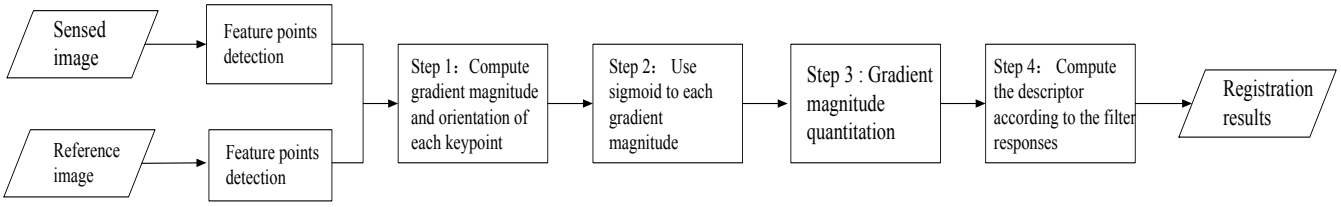
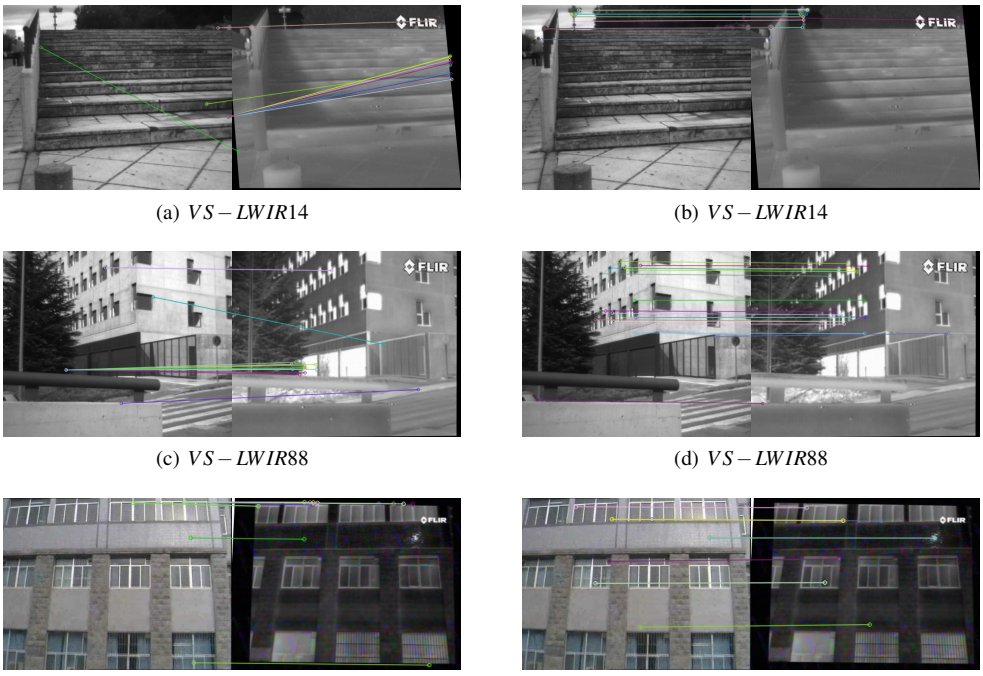


Figure 5. *Basic workflows of the computation of EOH considering gradient magnitude. The new EOH descriptor processing algorithm that make use of keypoint gradient magnitude was developed to further improve the accuracy of EOH.*



(a) $VS-LWIR14$ (b) $VS-LWIR14$

(c) $VS-LWIR88$ (d) $VS-LWIR88$

(e) $EOIR54$ (f) $EOIR54$

Figure 6. *The matching performance under EOH and the proposed method. The left column is the result of original EOH. The right column is the result of EOH utilizing gradient magnitude. 7(b) and 7(d) shows the result of image enhancement algorithms.*

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

Fig. 7 gives the keypoint matchings built for the low quality images with the original EOH without gradient magnitude normalization and proposed method. As shown in the Fig. 7, feature point matching pairs of infrared images after enhancement are much more than before. Meanwhile, wrong matching pairs are reduced a lot this time, which proves the obvious effect of our image enhancement algorithms.

Original EOH provides 11 keypoint matches in Fig. 8(a) and 3 are visually correct. The SIFT, DE-SIFT, and LC-SIFT descriptor do not give many correct matches shown in Fig. 8(b) ,Fig. 8(c) and Fig. 8(d), while the proposed method gives 6 keypoint matches in Fig. 8(e) and all the matches are visually correct.

It is clearly that the original EOH main orientation and the proposed method could give correct matches. The reason might be that although this pair of images are taken with a visible camera and an infrared camera they are very close to single-spectrum images, i.e., brighter (darker) areas in the visible image is also brighter (darker) in the infrared image the relationship between image intensities is not linear and the gradient orientation could reverse, which makes SIFT and SIFT-like descriptors do not perform very well.

On the one hand, as the gradient magnitude is taken into consideration, keypoints that could be matched will increase. On the other hand, EOH just uses edge pixels, if image texture information is not rich, the result of EOH or proposed method is not satisfied. This is also the next problem we will solve.

Conclusion

This paper proposed an approach to assigning gradient magnitude to EOH descriptor, as EOH does not consider the gradient magnitude. But the gradient magnitude is not stable on multispectrum images, we could not use it directly. In this article, we use sigmoid and quantization to normalize the gradient magnitude. Experimental results show that the proposed can improve the matching performance of EOH.

Acknowledgment

Thanks go to the authors of [8] for their kindly providing source code and test images.

This work was supported by the National Natural Science Foundation of China (Grants No., NSFC-61170176, NSFC-61471067), Fund for the Doctoral Program of Higher Education of China (Grants No., 20120005110002), Fund for Beijing University of Posts and Telecommunications (Grants No., 2013XD-04, 2015XD-02), Fund for National Great Science Specific Project (Grants No. 2014ZX03002002-004), Fund for Beijing Municipal Administration of Hospitals Clinical medicine Development of special funding support (code: XMLX201406).

References

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded up robust features (surf)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[3] F. Barrera, F. Lumbreras, and A. D. Sappa, "Multispectral piecewise planar stereo using manhattan-world assumption," *Pattern Recognition Letters*, vol. 34, no. 1, pp. 52–61, Jan. 2013.

[4] S. Saleem and R. Sablatnig, "A modified sift descriptor for image matching under spectral variations," *International Conference Image Analysis and Processing*, pp. 652–661, 2013.

[5] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1705–1720, Sep. 2010.

[6] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "Sar-sift: A sift-like algorithm for sar images," *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, vol. 53, no. 1, pp. 453–466, Jan. 2015.

[7] S. A. Chatzichristofis and Y. S. Boutalis, "Mutli-spectral remote image registration based on sift," *Electronic Letters*, vol. 44, no. 2, pp. 107–108, 2008.

[8] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, "Multispectral image feature points," *Sensors*, vol. 12, no. 9, pp. 12 661–12 672, Sep. 2012.

[9] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[10] B. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors."

[11] R. Ulloa, "Edge orientation histograms in global and local features," *http://roberto.blogs.cultureplex.ca/2012/01/26/edge-orientation-histograms-in-global-and-local-features/*, 2012.

[12] M. Delbracio, P. Muśe, and A. Almansa, "Non-parametric sub-pixel local point spread function estimation," *Image Processing On Line*, pp. 8–21, Mar. 2012.

[13] T. S. Cho and P. S. W.T, "Blur kernel estimation using the radon transform," *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 241–248, 2011.

[14] Y. Li, L. Xu, H. Jin, and J. Zou, "Localizing edges for estimating point spread function by removing outlier points," *Journal of Modern Optics*, vol. 63, no. 3, pp. 245–251, 2015.

[15] R. Schultz and R. Stevenson, "Video resolution enhancement," *Image and Video Proceeding III, Proceedings of the SPIE*, pp. 23–24, 1995.

[16] C.-L. Tsai, C.-Y. Li, G. Yang, and K.-S. Lin, "The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence," *IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 636–649, Mar. 2010.

[17] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[18] Y. Wu, W. Ma, and M. Gong, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 43–C47, 2015.

Author Biography

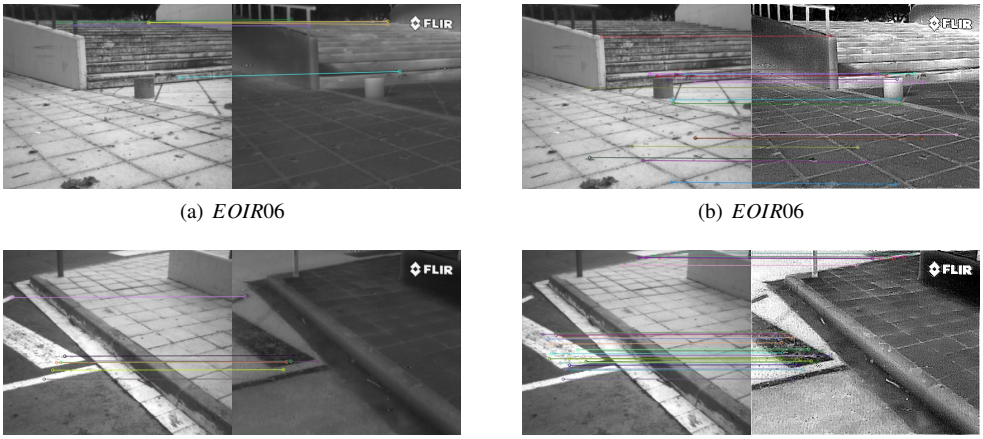
Liangpeng Xu is a graduate student of Beijing University of Posts and Telecommunciations. His research is focused on image super-resolution and deep learning.

Chunxiao Fan is a professor of Beijing University of Posts and Telecommunciations.

Hongbin Jin is a PHD student of Beijing University of Posts and Telecommunciations. His research is focused on computer vision and image processing.

Yong Li received his Master of Science in applied mathematics with Prof. Gerald Misiolek, and his PhD with Prof. Robert L. Stevenson, both from the University of Notre Dame. His research is focused on computer

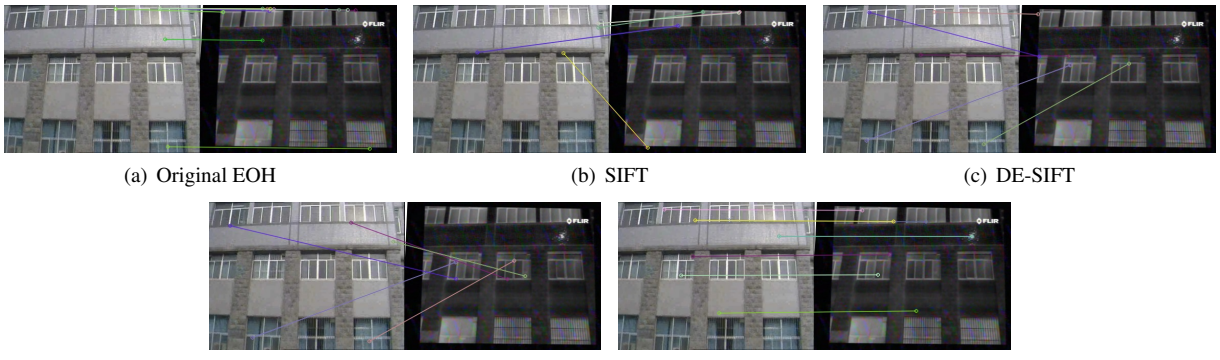
DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241



(a) $EOIR06$ (b) $EOIR06$

(c) $EOIR50$ (d) $EOIR50$

Figure 7. *The matching performance for low quality images under image enhancement. The left column is the result of EOH utilizing gradient magnitude without image enhancement. The right column is the result of EOH utilizing gradient magnitude with image enhancement.*



(a) Original EOH (b) SIFT (c) DE-SIFT

(d) LC-SIFT (e) Proposed method

Figure 8. *The matched keypoints built with descriptors. (a) the original EOH ,(b) SIFT, (c) DE-SIFT, (d) LC-SIFT, (e) the proposed method.*

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-241

vision, image processing, and differential geometry.

Xiang Shi is a graduate student of Beijing University of Posts and Telecommunciations. His research is focused on computer vision and image processing.