光电技术及应用

DOI：10.16818/j.issn1001－5868.2024061405

基于二进制自相似描述符的多模态图像配准

张维刚，谢志华*，车　驰，付政权，徐本源
（重庆光电技术研究所，重庆 400060）

摘　要： 针对局部自相似性（Local Self-Similarity，LSS）描述符不适用于多模态图像配准，提出一种新的自相似描述符，该描述符能够有效应用于多模态图像配准。首先，对多模态图像采用相位一致性算法提取图像的最大矩；接着，对由最大矩信息得到的边缘图像进行 Harris 特征点提取；然后，基于边缘图像生成二值图像，并基于二值图像构造二值化自相似描述符；最后，对多模态图像进行描述符相似性计算以及特征点匹配。通过对比实验表明，所提出的二值化自相似描述符能够取代传统的 LSS 描述符，有效提高了自相似描述符对多模态图像的兼容性和运行效率。

关键词： 多模态；图像配准；二进制；自相关；描述符
中图分类号： TP391.41　**文章编号：** 1001－5868(2024)06－0966－05

Multimodal Image Registration Based on Binary Self-similarity Descriptors

ZHANG Weigang, XIE Zhihua, CHE Chi, FU Zhengquan, XU Benyuan
(Chongqing Optoelectronics Research Institute, Chongqing 400060, CHN)

Abstract: A notable limitation of the local self-similarity descriptor (LSS) is that it is considered unsuitable for multimodal image registration. In order to address this issue, a novel self-similarity descriptor is proposed and effectively applied for multi-modal image registration. First, the phase congruency algorithm is used to extract the maximum moment of the multimodal image. Second, Harris keypoints are extracted from the edge images obtained from the maximum moment information. Third, a binary image is produced based on the edge image, and a binary self-similar descriptor is constructed based on the binary image. Finally, descriptor similarity calculations and keypoint matching are performed for multimodal images. Comparative experiments demonstrate that the proposed binary self-similar descriptor serve as a replacement for the traditional self-similar descriptor, and it effectively improve the compatibility and efficiency of the self-similar descriptor for multimodal images.

Key words: multimodal; image registration; binary; self-similarity; descriptor

0　引言

随着光电成像技术和图像处理技术的发展，对多模态图像配准技术的要求越来越高。同源图像配准技术已相对成熟，但随着基于不同成像机理的成像设备的出现，需要对基于不同成像设备得到的图像进行有效配准处理，以满足实际工程需求。作为图像处理技术的基础，多模态图像配准工作尤为重要。由于多模态图像之间成像机理不同，不同类型的图像在同一特征上的灰度信息表达不一致性。因此，基于灰度特征的传统同源图像配准算法不适用于多模态图像。

尺度不变特征变换（Scale-Invariant Feature Transform，SIFT）算法[1]、加速强化特征（Speed Up Robust Features，SURF）算法[2]等传统经典算法广泛用于同源图像配准。然而这些传统算法不适用于

收稿日期：2024－06－14.
* 通信作者：谢志华

多模态图像配准。研究者基于这些传统算法提出适用于多模态图像配准的新算法，例如，Ma 等[3]基于位置、尺度和方向信息对 SIFT 进行改进并提出 PSO-SIFT 算法，成功将 SIFT 用于多模态图像配准；Xiang 等[4]在 POS-SIFT 基础上进一步优化，提高了多模态图像的配准精度和鲁棒性；Zhao 等[5]基于 SURF 算法进行改进，提出适用于多模态图像配准的描述符。然而，这些算法大都是基于图像梯度信息进行优化，对多模态图像之间的几何差异和辐射差异比较敏感。近年来，众多研究者们基于相位一致性算法[6]进行了相关研究，并发现相位一致性算法对多模态图像之间的几何和辐射差异不敏感，对边缘信息敏感程度高，更加适用于多模态图像配准。例如，Ye 等[7]基于相位一致性模型构建方向直方图，对图像中的结构信息进行有效表达，解决多模态图像之间的非线性辐射差异。但其构建的描述符不能兼容多模态图像之间存在的旋转差异，对多模态图像之间角度变化比较敏感。Li 等[8]采用改进后的相位一致性最大矩算法进行多模态图像配准，该算法能够兼顾图像之间的灰度和几何畸变等差异，且算法运行效率和配准精度均佳。但该算法采用梯度信息统计的方式计算主方向，而梯度信息对多模态图像中的非线性辐射差异兼容性较差，容易影响特征描述符的准确表达。Wang 等[9]同样基于相位一致性算法进行改进，提出基于相位一致性最大矩和最小矩结合的算法来提取特征和构建描述符。该算法有效提高了多模态图像的配准精度，然而算法的复杂度相对较高，运行效率较低。此外还有一些研究者提出了基于相位一致性的多模态图像配准算法，但如何较好地兼顾多模态图像之间的配准精度和配准效率始终是目前亟待解决的难题。由此可见，相较于传统的基于灰度和梯度的算法，基于相位一致性的配准算法更加适用于多模态图像。但基于相位一致性的配准算法运行效率较难满足实际应用，需要进一步的研究和优化，提出更加高效的多模态图像配准算法。

针对多模态图像配准技术展开研究，提出一种新的多模态图像配准算法，该算法解决了传统算法复杂度高、运行效率低的问题。首先，使用相位一致性算法提取多模态图像中的边缘信息，进而获取边缘图像。然后，基于边缘图像进行后续的特征点提取和二值图像的生成，并基于二值图像进行描述符的构建。最后，使用 Hamming 距离[10]和快速抽样一致性(Fast Sample Consensus，FSC)算法[11]完成描述符相似度计算和特征匹配。所提出的算法有效解决了传统图像配准算法中描述符计算过程繁琐的问题，并可适用于可见光图像、红外图像以及合成孔径雷达等多模态图像的配准。

1　配准算法

配准算法主要包括相位一致性最大矩计算、边缘图像生成、二值图像生成、特征提取、描述符构建等步骤，多模态图像配准算法的具体流程如图 1 所示。其中，特征点提取采用经典的 Harris 角点检测算法。充分考虑相位一致性算法对多模态图像之间几何和辐射差异的兼容性，利用相位一致性最大矩生成边缘图像，并基于边缘图像和阈值进行二值图像的生成。同时，基于二值化自相似描述符对多模态图像中的特征进行表达。所提出的算法直观易理解，且相较于传统的局部自相似性(Local Self-Similarity，LSS)描述符，运行效率得到了极大提高。

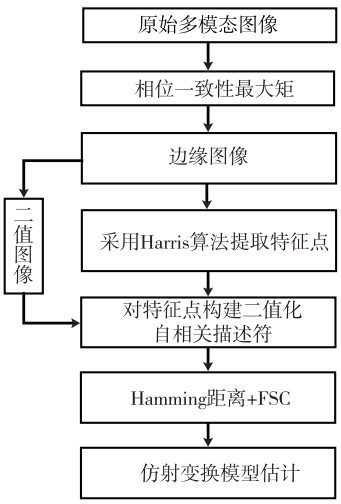


图 1　所提出的算法配准流程图

2　算法流程

2.1　相位一致性最大矩

相位一致性最大矩 M 是由相位一致性模型 PC 通过以下方式计算得到的，其计算式为：

$$M=\frac{1}{2}(c+a+\sqrt{b^2+(a-c)^2}) \tag{1}$$

式中，参数 a，b 和 c 的计算式分别为：

$$a=\sum_{\theta}(PC(\theta)\cos\theta)^2 \tag{2}$$

$$b=2\sum_{\theta}(PC(\theta)\cos\theta)(PC(\theta)\sin\theta) \tag{3}$$

$$c=\sum_{\theta}(PC(\theta)\sin\theta)^2 \tag{4}$$

2.2 边缘图像生成及特征点提取

首先，使用第一步中得到的相位一致性最大矩 $\boldsymbol{M}$ 按照以下公式生成边缘图像 $\boldsymbol{I}$：

$$\boldsymbol{I}=\boldsymbol{M}\times 255 \tag{5}$$

然后，在边缘图像 $\boldsymbol{I}$ 中采用经典的 Harris 角点检测算法提取特征点，即：

$$\boldsymbol{H}=\sum_{x,y}\boldsymbol{w}(x,y)\begin{bmatrix}\boldsymbol{I}_x^2 & \boldsymbol{I}_x\boldsymbol{I}_y\\ \boldsymbol{I}_x\boldsymbol{I}_y & \boldsymbol{I}_y^2\end{bmatrix} \tag{6}$$

式中，$\boldsymbol{w}$ 为模板，$\boldsymbol{I}_x$ 和 $\boldsymbol{I}_y$ 分别表示 x 和 y 方向的梯度信息。然后，计算出二维矩阵 $\boldsymbol{H}$ 的特征值 λ_1 和 λ_2，通过特征值计算得到响应值 $\boldsymbol{R}$，即：

$$\boldsymbol{R}=\lambda_1\lambda_2-k(\lambda_1+\lambda_2)^2 \tag{7}$$

当响应值 $\boldsymbol{R}$ 计算完成之后，需要通过设置响应值的阈值得到相应的 Harris 角点，并进行非极大值抑制的方式得到所需特征点。

2.3 构造二值化自相关描述子

首先，计算边缘图像 $\boldsymbol{I}$ 的幅值和方向信息，考虑到多模态图像之间可能存在梯度反转，因此，将方向信息由[−90°～90°)限定到[0°-180°)之间，选取半径 91 pixels 的支撑区域，将方向信息划分到五个方向块 bin 中，即：

$$\begin{aligned}&bin=0,(0^\circ\leqslant\theta<22.5^\circ)\\&bin=1,(22.5^\circ\leqslant\theta<67.5^\circ)\\&bin=2,(67.5^\circ\leqslant\theta<112.5^\circ)\\&bin=3,(112.5^\circ\leqslant\theta<157.5^\circ)\\&bin=4,(157.5^\circ\leqslant\theta<180^\circ)\end{aligned} \tag{8}$$

接下来，基于边缘图像 $\boldsymbol{I}$ 通过以下公式获取由元素 0 和 1 组成的二值图像 b：

$$b(i,j)=\begin{cases}1,\boldsymbol{I}(i,j)\neq 0\\0,\boldsymbol{I}(i,j)=0\end{cases} \tag{9}$$

然后，分别在 0°，90°，180°和 270°四个方向上统计窗口内的灰度值，得到 θ 方向的二值图像 b_θ，根据经验将 N 设置为 5，阈值 T 设置为 13。

$$b_\theta(x,y)=\sum_{(i,j)\in N}b_\theta(i,j),((x,y)\in round((i,j)/N)) \tag{10}$$

$$\boldsymbol{L}(x,y)=\begin{cases}1,\prod_{\theta=1}^{4}b_\theta(x,y)\geqslant T\\0,\prod_{\theta=1}^{4}b_\theta(x,y)<T\end{cases} \tag{11}$$

采用四个方向上二值图像相乘的方式得到最终的二值图像 $\boldsymbol{L}$，即四个方向上均为元素 1 的位置灰度值定为 1，其余位置灰度值定为 0，以此得到的二值图像 $\boldsymbol{L}$ 对多模态图像具有较高的鲁棒性和兼容性。将支撑域划分到 49 个位置块 $\boldsymbol{A}_i$ 中，每个位置块大小为 26 pixel×26 pixel。其中，描述符的支撑区域划分方式如图 2 所示。

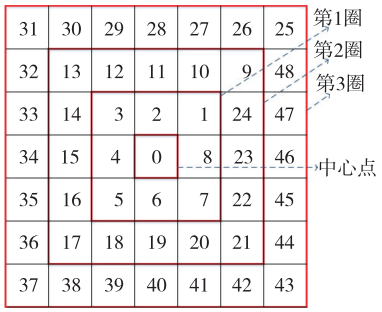


图 2 所提出的描述符支撑区域的划分方式

采用直方图统计的方式在二值图像 $\boldsymbol{L}$ 中获取每个位置块内五个方向块上的方向信息，得到 49 个 5 维的特征向量 $\boldsymbol{p}_i$。同时，将这 49 个特征向量按照主方向块 bin_0 进行排序得到每个位置块 $\boldsymbol{A}_i$ 的特征向量 $\boldsymbol{p}'_i$，排序方式如下：

$$\boldsymbol{p}'_i=\begin{cases}[\boldsymbol{p}_i((1\times bin_0+1):8);\boldsymbol{p}_i(1:1\times bin_0)],\\(\boldsymbol{A}_1\sim\boldsymbol{A}_8)\\[\boldsymbol{p}_i((2\times bin_0+1):16);\boldsymbol{p}_i(1:2\times bin_0)],\\(\boldsymbol{A}_9\sim\boldsymbol{A}_{24})\\[\boldsymbol{p}_i((4\times bin_0+1):24);\boldsymbol{p}_i(1:4\times bin_0)],\\(\boldsymbol{A}_{25}\sim\boldsymbol{A}_{48})\end{cases} \tag{12}$$

采用每个位置块 $\boldsymbol{A}_i$ 内的向量 $\boldsymbol{p}_i$ 与中心位置块 $\boldsymbol{A}_0$ 内的向量 $\boldsymbol{p}_0$ 作差的形式获取特征点所在支撑域内的自相关特征 $\boldsymbol{p}_{i,\mathrm{LSS}}$，并将 $\boldsymbol{p}_{i,\mathrm{LSS}}$ 中元素的正负特性进行二值化，得到二值化特征向量 $\boldsymbol{d}_i$，由此将 $\boldsymbol{d}_i$ 串联起来得到 240 维的二值描述符 $\boldsymbol{D}$。

本文通过提出的二值化自相关描述符提高多模态图像配准速度，并简化了多模态图像配准过程中构建描述符的复杂度。同时解决传统自相关描述符 LSS 不适用于多模态图像配准的问题。

2.4 描述符相似性计算以及特征点匹配

采用 Hamming 距离进行描述符之间相似性的度量，通过 FSC 算法进行误匹配点对剔除，最后使用仿射变换模型完成多模态图像匹配。

3 实验分析

为对比分析多模态图像配准性能，使用适用于

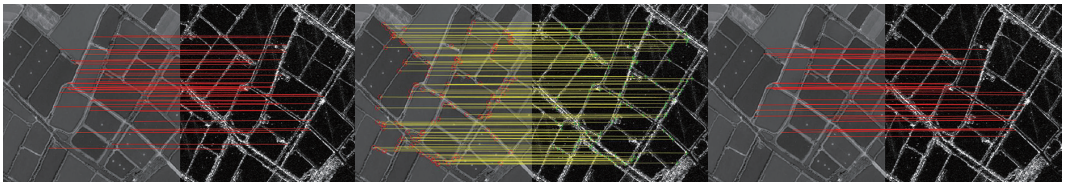
多模态图像配准的 RIFT 算法、BLSS 算法[12]与所提出的算法进行对比分析。两种对比算法和本文算法都是基于相位一致性进行优化，且在多模态图像配准的精度和运行效率等方面均表现良好。以上三种算法均基于 Matlab R2020b 软件平台实现。

3.1　参数设置与数据来源

本文算法中特征点提取阈值范围为 0.003～0.001，根据其数量进行微调，描述符构建过程中特征点所在支撑域范围为 183 pixel×183 pixel。对比算法按照作者指示进行参数设置；另外，本文为获取到相似数量的特征点，还对相关参数进行了微调，第一组和第二组多模态图像为可见光与合成孔径雷达图像[13]，大小均为 512 pixel×512 pixel；第三组和第四组多模态图像为可见光与红外图像[8,14]，大小分别为 600 pixel×600 pixel 和 512 pixel×512 pixel。

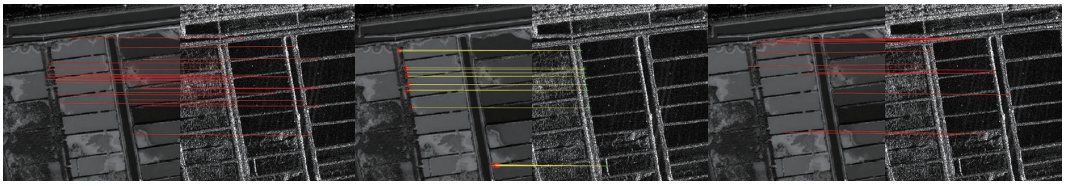
3.2　性能对比

首先，对配准性能进行定性和定量分析，如图 3～6 所示给出了三种算法的四组图像对配准效果图。由图可见，本文算法和 BLSS 算法均能成功对多模态图像对完成有效匹配，且视觉效果理想。由于第二组图像对中图像内的结构信息独特性较差，RIFT 算法对第二组图像配准效果不佳。



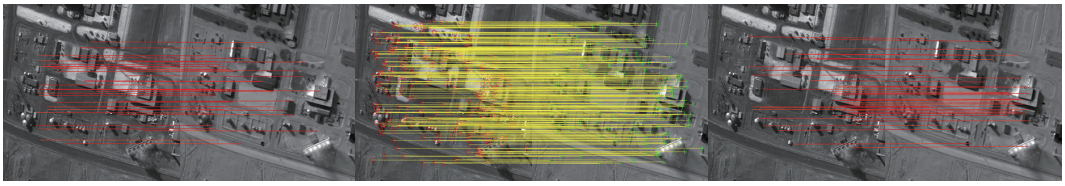
(a) BLSS 算法　(b) RIFT 算法　(c) 本文算法

图 3　第一组图像对配准结果



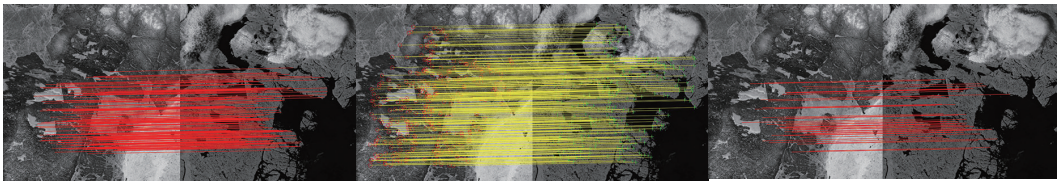
(a) BLSS 算法　(b) RIFT 算法　(c) 本文算法

图 4　第二组图像对配准结果



(a) BLSS 算法　(b) RIFT 算法　(c) 本文算法

图 5　第三组图像对配准结果



(a) BLSS 算法　(b) RIFT 算法　(c) 本文算法

图 6　第四组图像对配准结果

本文采用正确匹配点对数量(Number of Correct Matches，NCM)、均方根误差(Root Mean Square Error，RMSE)以及运行时间对所提算法进行定量分析。其中，RMSE 基于手工选取匹配特征点对的方式进行计算，计算公式为：

$$RMSE=\sqrt{\frac{1}{NCM}\sum_{i=1}^{NCM}\left[\left(X_i^{m'_1}-X_i^{m'_2}\right)^2+\left(Y_i^{m'_1}-Y_i^{m'_2}\right)^2\right]} \tag{13}$$

式中，$(X_i^{m_1},Y_i^{m_1})$和$(X_i^{m_2},Y_i^{m_2})$分别为采用随机抽样一致性算法进行误匹配点删除之后得到的待配准的多模态图像对中第 i 对特征点坐标值。上述公式中，坐标值$(X_i^{m'_1},Y_i^{m'_1})$可由$(X_i^{m_1},Y_i^{m_1})$基于变换矩阵 $\boldsymbol{H}$ 获得，采用 20 对手工标注的特征点对计算得到多模态图像对之间的变换矩阵 $\boldsymbol{H}$。

表 1 给出了多模态图像配准算法的定量分析结

果。实验结果表明,本文算法和 BLSS 算法的配准精度相近,均能成功用于多模态图像配准。本文算法的配准精度由于第二组图像的各区域之间信息差异较小,不同特征点所在支撑域的结构信息相似度较高,容易干扰正确匹配点对,RIFT 算法匹配失败。对于其他几组图像,RIFT 算法的配准精度是最高的。除此之外,由于本文算法采用二值化的思想改进构建描述符的方式,缩短了构建描述符的时间。因此,本文算法在三种对比算法中配准效率最高,运行时间最短。由此可见,所提算法能够在配准精度可接受的前提下有效提高算法的运行效率,可用于对配准效率要求较高的场景。

表 1 三种算法的定量分析结果

图像对编号	算法	*NCM*	*RMSE*	描述符运行时间/s	总运行时间/s
1	BLSS	46	3.135 3	1.452	3.923
	RIFT	125	2.523 2	2.473	5.231
	本文算法	41	3.307 8	1.45	3.311
2	BLSS	20	3.795 4	1.803	4.985
	RIFT	—	—	2.074	4.496
	本文算法	13	3.907 5	1.677	3.573
3	BLSS	37	2.390 7	2.127	4.412
	RIFT	327	2.151 9	3.654	9.629
	本文算法	42	2.503 9	1.686	4.252
4	BLSS	172	1.686	2.06	4.914
	RIFT	322	1.552	2.527	4.908
	本文算法	41	1.98	2.018	4.4

4 结论

为提高基于相位一致性框架的多模态图像配准算法对多模态图像之间几何和辐射差异的兼容性,提出了基于二值化自相似描述符的多模态配准算法,有效提高了基于相位一致性配准算法的运行效率。首先,利用相位一致性算法中的最大矩信息生成边缘图像,并基于边缘图像提取角点特征。然后,对图像进行二值图像的生成。最后基于二值图像进行描述符构建,基于二值化的思想提高传统自相关描述符对多模态图像中特征表达的有效性和鲁棒性。实验结果表明,本文算法能够有效提高多模态图像配准效率,具有重要的实际应用价值。

参考文献:

[1] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[2] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features[C]// Proceedings of the 9th European Conference on Computer Vision (ECCV), 2006: 404-417.

[3] Ma W, Wen Z, Wu Y, et al. Remote sensing image registration with modified SIFT and enhanced feature matching [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 14 (1): 3-7.

[4] Xiang Y, Wang F, You H. OS-SIFT: A robust SIFT-Like algorithm for high-resolution optical-to-SAR image registration in suburban areas[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(6): 3078-3090.

[5] Zhao D, Yang Y, Ji Z, et al. Rapid multimodality registration based on MMSURF[J]. Neurocomputing, 2014, 131: 87-97.

[6] Kovesi P. Phase congruency: A low-level image invariant[J]. Psychological Research, 2000, 64(2): 136-148.

[7] Ye Y, Shan J, Bruzzone L, et al. Robust registration of multimodal remote sensing images based on structural similarity[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2941-2955.

[8] Li J, Hu Q, Ai M. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform[J]. IEEE Transactions on Image Processing, 2020, 29: 3296-3310.

[9] Wang L, Sun M, Liu J, et al. A robust algorithm based on phase congruency for optical and SAR image registration in suburban areas[J]. Remote Sensing, 2020, 12(20): 3339.

[10] Yong A, Hong Z. SIFT matching method based on *K* nearest neighbor support feature points[C]// 2016 IEEE International Conference on Signal and Image Processing (ICSIP), 2016: 64-68.

[11] Zhao M F, Chen H J, Song T, et al. Research on image matching based on improved RANSAC-SIFT algorithm[J]. Laser Journal, 2018, 39(1): 114-118.

[12] Xie Z, Liu J, Liu C, et al. Optical and SAR image registration using complexity analysis and binary descriptor in suburban areas[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.

[13] Xiang Y, Tao R, Wang F, et al. Automatic registration of optical and SAR images via improved phase congruency model [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 5847-5861.

[14] Ye Y, Bruzzone L, Shan J, et al. Fast and robust matching for multimodal remote sensing image registration[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57 (11): 9059-9070.

作者简介:

张维刚(1981—),男,主要从事 DSP,FPGA 应用技术的研究;

谢志华(1991—),女,主要从事图像处理算法、FPGA 应用技术、摄像机等方面的研究。