

# 基于模板迭代更新的图像配准算法

李慧敏 张之江

(上海大学通信与信息工程学院 上海 200072)

**摘要:** 在计算机视觉中,用变换矩阵对图像对进行配准一直是具有挑战性的,尤其是在变换矩阵难以获取且噪声很大的情况下。针对图像对之间重叠区域很小,以至于变换矩阵的获取精度不甚理想,首先用 SIFT 算子提取出图像的特征点,然后在对图像进行匹配时,提出一种基于模板迭代更新的图像匹配方法,推导出最小二乘约束下的变换矩阵求解方法。实验对比了一些实际场景图像序列的匹配结果,验证了本算法的可靠性和稳定性。

**关键词:** 图像配准; SIFT 算子; 模板迭代

中图分类号: TP751; TN911.73 文献标识码: A 国家标准学科分类代码: 510.4050

## Iterative updating affine alignment method for image group with little similarities

Li Huimin Zhang Zhijiang

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

**Abstract:** The alignment of an image pair by means of transformation matrices is a challenging problem in computer vision, especially when the transformation is unavailable to compute and free of noise. In this paper, we first use the SIFT algorithm to detect the feature points, and then during the image alignment process, we propose to iterate the reference images and solve the transformation matrix under the least squares constraint. Experimental results on aligning some real images with few or none overlap regions show the proposed alignment algorithm outperforms the state-of-the-art image alignment methods.

**Keywords:** image alignment; SIFT algorithm; iterative updating

## 0 引言

在图像领域中,通常需要将不同时间、不同成像设备或不同条件下获取的若干幅图像进行匹配叠加,这就叫做图像配准。图像配准算法一直是计算机视觉领域的一个重要问题,随着虚拟现实、模式识别等现代科技的飞速发展,其在目标识别、视频跟踪、全景融合、临床医学应用<sup>[1-2]</sup>等领域的应用越来越广泛。

近年来提出的针对图像匹配算法的研究主要可以归为两类:基于灰度的匹配和基于特征的配准<sup>[3-4]</sup>。前者是直接利用图像的灰度值来确定配准的空间变换,该法认为参考图像和待配准图像上的对应点及其周围区域具有相同或者相似的灰度<sup>[5]</sup>,并以此为基础,运用相似度函数,寻找一组使得相似度函数最大的几何变换参数,从而实现图像的配准<sup>[6-7]</sup>。该方法的主要特点是实现简单,但是对图像的灰度变化及目标旋转比较敏感<sup>[8-9]</sup>,在最优变换的过程中需要巨大的运算量。基于特征的图像匹配方法是近年来最热门的

图像匹配分支,这类方法的核心是通过在图像中提取显著的特征以寻找匹配点对,从而求解出图像变换模型。David 等人<sup>[10]</sup>通过比较各种局部描述符的仿射变换、规模变化、旋转、光照变化等特性,发现在基于特征的方法中 SIFT 算子是最优秀的<sup>[11-12]</sup>。SIFT 算子能够大大压缩图像信息量,在变换过程中运算量小,速度快且具有鲁棒性。

传统图像匹配方法主要是用来解决图像对之间的配准问题,但对于图像序列的配准问题,尚未有很好的解决方法。一般来说,对图像的匹配要求待匹配图像之间具有较多的重合部分,而对于若干张相似区域较少的图像而言,直接运用传统方法会导致极大的误匹配、漏匹配、错误的变换模型参数估计等现象,从而导致图像匹配扭曲变形,直接引起后续操作如图像拼接、图像融合等的差错。因此本文首先采用 SIFT 算子,提取出图像特征点,在保证特征点选取的精确度的前提下,针对图像序列的匹配问题,提出了一种基于模板迭代更新的图像匹配方法,推导出最小二乘约束下的变换矩阵求解方法,在每次匹配过程中不断迭代更新模板图

像,该方法能够有效的保证图像配准中的结构完整性。

## 1 图像配准基本流程

图像配准就是根据图像在空间和灰度上的映射关系,求取空间几何变换模型的过程。假设  $I_1(x, y)$  和  $I_2(x, y)$  分别表示两幅待拼接图像在位置  $(x, y)$  上的灰度值,  $f$  表示二维空间的坐标变换,  $g$  表示一维的灰度变换, 图像配准关系可表示为:

$$I_2(x, y) = g[I_1(f(x, y))] \quad (1)$$

配准过程就是求解最优函数  $f$  和  $g$  的过程。无论图像源是何种类型,无论图像的种类是什么,其配准的基本流程都由以下 4 个步骤组成: 特征提取、特征匹配、参数估计和重采样。特征提取指的是通过算法将图像中最为显著的点或区域采集出来。特征匹配是指对上一步骤采集得到的显著点或者区域进行判断, 符合判断条件的点留下并进行匹配, 经过匹配后可以得到待配准图像间特征值的位置关系和对应情况。图像重采样是为防止图像经过变换以后可能出现非整数值的像素问题所进行的操作。为保证最后拼接的效果质量能够满足要求, 所提取的特征点需要有很好的稳定性, 如若提取的特征不够稳定, 则会大大增加拼接最后的失败概率。

## 2 基于模板迭代更新的图像匹配

对于从一个静态场景采集若干帧图片  $I_1, I_2, \dots, I_k, \dots, I_n$ , 由于图像获取设备移动会带来图像不对齐的问题, 干扰到后续对该帧序列图像的处理。本章提出的基于模板迭代更新的算法旨在解决该问题。对于第  $k$  帧图像  $I_k$ , 以  $\hat{I}_{k-1}$  为参考, 对  $I_k$  进行变换。利用 SIFT 特征点检测方法, 选取若干对应点  $V_k(u)$ , 得到两张图像之间的变换矩阵  $T_k$ 。将此变换矩阵反作用于此次变换的模板图像  $\hat{I}_{k-1}$ , 并与此次目标图像  $I_k$  进行融合, 记融合更新后的结果图像为  $\hat{I}_k$ , 将该图像作为下次配准的模板图像。迭代融合的过程, 也就是对每次配准的模板图像进行更新的过程。在图像匹配的过程中, 构建新的图像配准方法, 对模板图像进行实时更新矫正迭代, 以提高匹配的正确率。

### 2.1 基于 SIFT 的特征点检测方法

图像的特征点检测与匹配是基于特征的图像配准的基

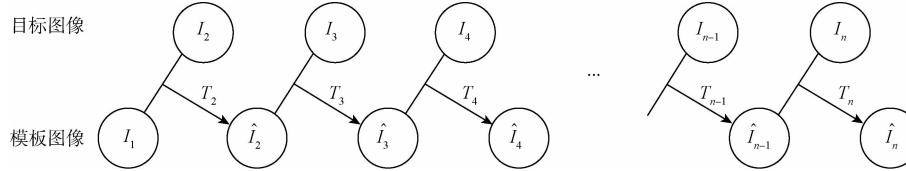


图 1 模板迭代更新匹配算法模型

下面以  $N$  幅待配准图像为例进行叙述。设从同一场景获取的图像序列为  $N$  张:  $I_1, I_2, \dots, I_k, \dots, I_n$ 。在 2.1 节利用 SIFT 算法对图像特征点进行检测后, 采取本小节

提出的模板迭代更新算法, 可以利用在图像中提取少量的特征点的匹配关系, 进而找到图像间的对应关系来进行图像的配准, 这样可以大大降低图像配准的复杂度和计算量。

本文局部特征点提取的目标图像由移动设备拍摄的, 该图像序列有如下特征: 1) 它们是由同一个图像采集设备, 取自同一时间、同一场景中; 2) 由于手持式图像设备的移动, 它们彼此间具有平移、缩放、旋转、仿射等变换关系; 3) 这些图像序列彼此间的重叠区域很小。一般来说, 多次多角度获取场景图像, 可以帮助用户更好的理解场景信息, 但在图像差异较大、重叠区域较少的情形下, 对图像序列的特征点检测方面的难度就比较大。为保证特征点选取的准确性, 本小节采用 SIFT 特征点<sup>[13]</sup> 提取算子进行特征点检测。

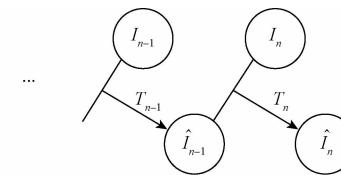
SIFT 算法主要有如下 4 步<sup>[14]</sup>:

- 1) 对尺度空间进行极值检测: 由原图像构建高斯金字塔, 使用高斯微分函数来搜索所有尺度上图像的位置, 找到特征点的候选集。
- 2) 对关键点进行定位: 通过构造插值函数, 剔除步骤 1) 的候选特征点中边缘位置的不稳定点, 并对关键点的位置和尺度进行定位。
- 3) 确定关键点的方向: 为每个关键点分配 1 个或多个主方向, 该方向信息由关键点邻域像素点的梯度大小和方法决定。
- 4) 对关键点进行描述: 计算每个关键点邻域内像素的梯度, 在选定的尺度商测量图像局部的梯度, 用这些梯度的大小和方向来描述关键点。

### 2.2 模板迭代更新机制

目前大多数图像匹配都是基于模板匹配的。即以 1 张图像为目标, 也就是模板图像, 匹配的目的是将另一张图像投射到该模板图像上。一般来说, 这张模板图像是不变的, 所做的就是将每帧图像映射到模板图像上, 求出一系列的变换矩阵, 从而对图像序列进行匹配。针对图像序列匹配的特殊性, 本章节提出了一种基于模板迭代更新的图像序列匹配方法, 即在图像匹配的过程中, 对模板图像进行实时更新迭代矫正的图像序列匹配方法。

基于模板迭代更新的图像序列匹配算法模型如图 1 所示。



提出的模板迭代更新算法, 先对前两张图像进行匹配, 以  $I_1$  为模板,  $I_2$  为待匹配图像, 通过前面求取到的变换模型进行映射, 用变换矩阵对  $I_2$  进行变换, 之后基于结构化的

图像拼接技术,对图像  $I_1$  和变换后的图像  $I_2$  进行融合得到图像  $I'_2$ ,将  $I'_2$  作为下一次匹配的模板图像,对图像  $I_3$  进行匹配操作。在以  $I_k$  为目标图像进行匹配时,模板图像为上一步融合得到的图像  $I'_{k-1}$ ,构造点对点能量代价函数:

$$E(T_k) = \sum_u \| (T_k I_{k-1}^m - I_k)^T \|_2 \quad (2)$$

设定阈值,得到相应的变换矩阵  $T_k$ 。设  $T_{map}$  为图像  $I_{k-1}^m$  和  $I_k$  之间的变换矩阵,用矩阵  $T_{map}$  来更新图像  $I_{k-1}^m$ ,由此可以将图像对的全局变换矩阵写为:

$$T_k \triangleq \prod_{i=0}^z T_{map}^i \quad (3)$$

同时将输出的融合图像记为  $\hat{I}_k^m$ 。在对图像变换矩阵求解过程中,将该变换矩阵看作为单应矩阵。一般来说,对于单一场景,单应矩阵是非常稳健的图像变换模型。单应矩阵模型是 1 个含有 8 个参数的图像变换模型,用坐标矩阵可以表示为:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \quad (4)$$

其中,矩阵的最后 1 个值可以认为是尺度参数,因此可认为此投影变换矩阵有 8 个自由度。设  $(x_0, y_0)$  和  $(x'_0, y'_0)$  分别为两幅待拼接图像中的 1 对匹配点的坐标,其具体变换关系如式(5)所示。

$$\begin{bmatrix} x'_0 \\ y'_0 \\ 1 \end{bmatrix} = H \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad (5)$$

在单应矩阵的求解过程中,倘若出现一些明显偏离的异常数据点,单应矩阵的估计精度就会大大降低。因此为了能够建立一个鲁棒的图像变换模型,需要进行鲁棒数据拟合,目前应用最广泛的是 Fischler 和 Bolles<sup>[17]</sup>提出的随机采样一致性算法 (random sample consensus, RANSAC)。在用 RANSAC 算法剔除错误匹配点对,得到最佳匹配点后,则可用最小二乘法来重新估计最终的图像变换模型。

下面将从矩阵的角度对最小二乘的原理进行阐述。设  $A$  为一个  $m \times n$  的矩阵,  $b$  为一个  $m \times 1$  的矩阵, 最小二乘可看作是在欧几里得空间中,以 2-范数作为距离,求使得向量  $Ax$  与  $b$  之间距离最小的  $x$ :

$$\min_{x \in R^n} \|Ax - b\|_2 \quad (6)$$

也即求使得向量  $Ax$  与  $b$  之间距离平方最小的  $x$ :

$$\min_{x \in R^n} \|Ax - b\|_2^2 \quad (7)$$

将式(7)向量之间距离的平方展开,可得到:

$$\|Ax - b\|_2^2 = (Ax - b)^T \cdot (Ax - b) \quad (8)$$

为求得极值,将上式化简并对  $x$  求导:

$$\|Ax - b\|_2^2 = x^T A^T Ax - b^T Ax - x^T A^T b + b^T b \quad (9)$$

$$\frac{\partial \|Ax - b\|_2^2}{\partial x} = 2A^T Ax - 2A^T b \quad (10)$$

令式(10)等于 0,最小二乘解的矩阵形式可表示为式(11):

$$x = (A^T A)^{-1} A^T b \quad (11)$$

根据上述最小二乘原理,在图像变换矩阵的求解中,可以将图像变换模型矩阵  $H$  看作为最小二乘中要求解的  $x$ ,匹配点对  $x$  和  $x'$  看作最小二乘中对应的矩阵  $A$  和  $b$ ,则可求出借助最小二乘法求解图像变换矩阵的公式:

$$H = (x^T x)^{-1} x^T x' \quad (12)$$

其中  $H$  为 8 参数的单应矩阵,  $x$  为待配准图像中的特征点矩阵集,  $x'$  为模板图像中的特征点矩阵集。通过最小二乘法,求到图像变换模型的最优解,能够提高图像变换矩阵的精度,改善后续图像配准和全景融合拼接的效果。

### 3 实验结果与分析

本章节将提出的算法应用在多个图像数据集上,并将实验结果与文献[16-17]进行定性对比。本章主要从两个角度进行比较:1) 应用在 5 个不同的数据集上,横向比较不同算法之间配准结果的差异;2) 应用在同一个数据集上,改变图像配准中的模板图像,从而对匹配结果进行纵向比较。

#### 3.1 不同数据集图像对齐比较

本小节实验的输入图像是在同一场景下获取到的图像序列中的任意两张,这两张图像处于未配准状态,并且彼此之间的重叠区域较少。在原始方法中,这种重合区域极少的图像匹配是比较困难的。

图 2 表示了 5 组图像数据集在不同的配准算法下的实验结果。按照数据集的顺序从上往下,“基布尔学院”的屋顶、“小桥图”的船只、“故宫”的屋顶、“吉他图”的伞以及“房屋图”的楼阁可以显著表示图像配准后的结构变化。为了方便起见,用方框标记出了这些显著区域。可以看出使用本文提出的算法进行序列图像匹配,变换后的图像与模板图像具有相似度较高的结构性,对比实验 Bernard<sup>[15]</sup> 和 Ray<sup>[17]</sup> 的结构与模板图像相比有了较大变化。可以看出在结构化信息的不变性上,本文提出的算法有效性较高。

#### 3.2 同一数据集图像对齐比较

本小节实验的输入为同一图像数据集“房屋图”,在保证目标图像不变的基础上,任意选取 3 张图像作为模板,以此来检测在模板图像任意变化的情况下,本文算法匹配结果在结构上的鲁棒性。

图 3 是保持目标图像不变,以不同的图像为模板进行图像匹配的结果。在模板图像上用方框标记出其与目标图像的重叠区域,由上往下,可见随着模板图像与目标图像之间的重叠区域越来越小,从“阁楼”细节可以定性的表征出本文提出的匹配算法保持了相对稳定的结构不变性。

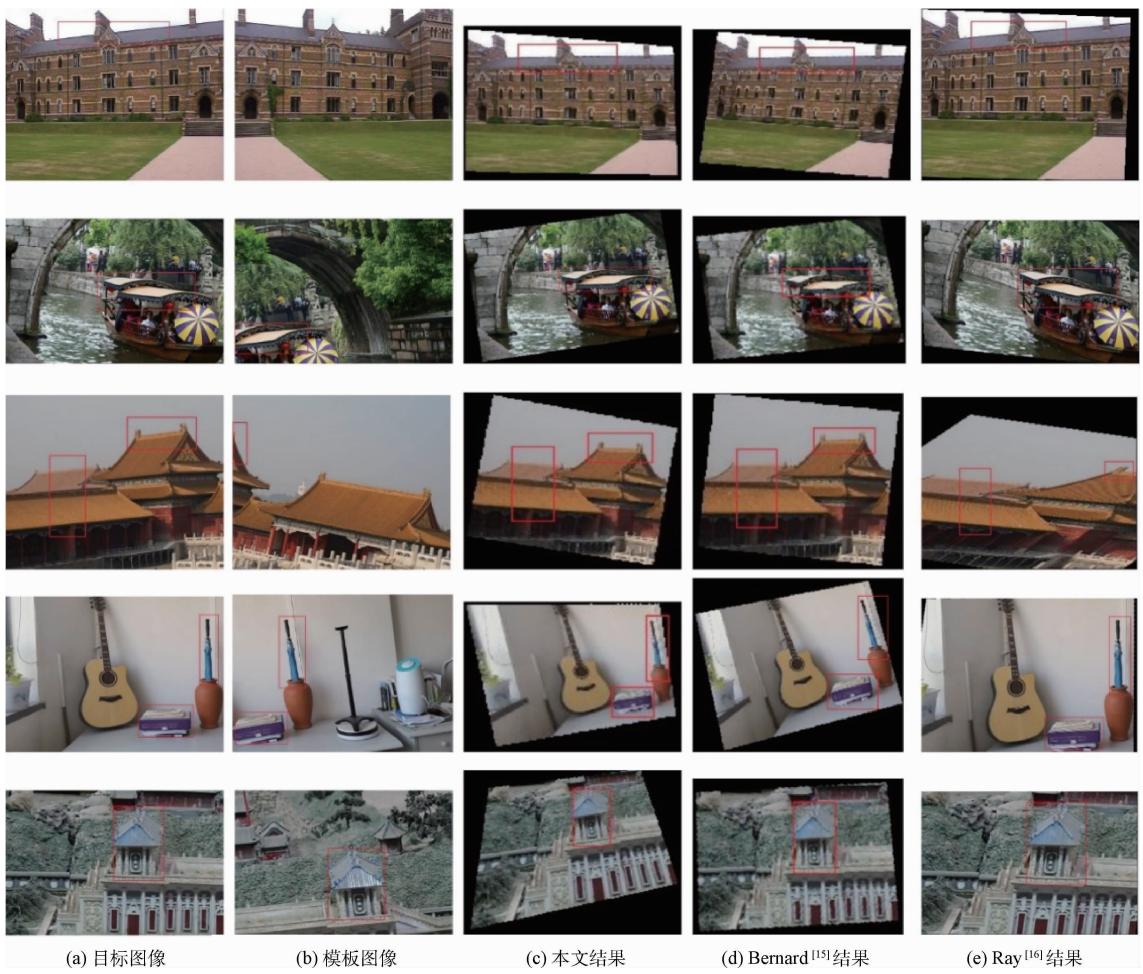


图 2 配准结果

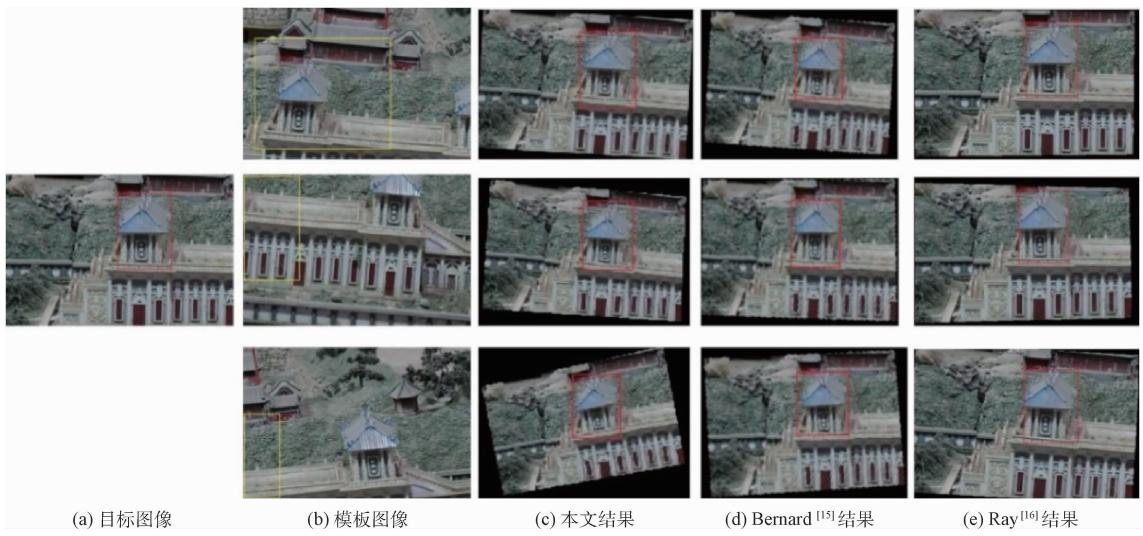


图 3 图像匹配结果

## 4 结 论

本文首先在 SIFT 特征点检测方法的基础上, 提取出

图像序列的特征点, 然后在图像序列的匹配问题上, 提出了一种基于模板迭代更新的图像匹配方法, 推导出最小二乘约束下的变换矩阵求解方法, 在每次匹配过程中不断迭

代更新模板图像以达到对图像序列进行连续匹配的目的。由实验结果可知,该法对于序列图像的匹配具有较好的效果,适用于该情形下的图像匹配。

## 参考文献

- [1] 朱奇光,张朋珍,李昊立,等. 基于全局和局部特征融合的图像匹配算法研究[J]. 仪器仪表学报, 2016, 37(1):170-176.
- [2] SU H R, LAI S H. Non-rigid registration of images with geometric and photometric deformation by using local affine fourier-monent matching[J]. Computer Vision and Pattern Recognition, 2015: 861-866.
- [3] CAO S. An edge-based scale-and affine-invariant algorithm for remote sensing image registration[J]. International Journal of Remote Sensing, 2013, 34(7):2301-2326.
- [4] 任克强,胡梦云. 基于改进 SURF 算子的彩色图像配准算法[J]. 电子测量与仪器学报, 2016, 30(5):748-756.
- [5] SOLERIBALTA A, SERRATOSA F. Graduated assignment algorithm for multiple graph matching based on a common labeling [C]. International Conference on Graph-Based Representations in Pattern Recognition, Springer-Verlag, 2013: 132-141.
- [6] 刘松. 基于改进的 SIFT 图像拼接及其并行化研究[D]. 合肥: 安徽大学, 2014, 4.
- [7] 彭坤杨, 董兰芳. 一种基于图像平均灰度值的快速图像修复算法[J]. 中国图象图形学报, 2010, 15(1):50-55.
- [8] CHEN X. Intensity-based image registration by nonparametric local smoothing[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(10): 2081-2092.
- [9] KYBIC J, VNUCKO I. Approximate all nearest neighbor search for high dimensional entropy estimation for image registration [J]. Signal processing, 2012, 92(5): 1302-1316.
- [10] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2010, 60(2): 91-110.
- [11] 刘武. 地球同步卫星图像轮廓配准[J]. 电子测量技术, 2016, 39(7):87-91.
- [12] 戴金波, 赵宏伟, 刘君玲, 等. 一种针对于描述子的 SIFT 简化方法[J]. 仪器仪表学报, 2012, 33(10): 2255-2262.
- [13] 孙光民, 王晨阳. 一种基于改进 SIFT 的图像检索算法[J]. 国外电子测量技术, 2016, (8):32-37.
- [14] WANG S, LI W, WANG Y, et al. An improved difference of gaussian filter in face recognition[J]. Journal of Multimedia, 2012, 7(6).
- [15] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007, 74 (7): 59-73.
- [16] BERNARD F, THUNBERG J, GEMMAR P, et al. A solution for multi-alignment by transformation synchronization [J]. Computer Vision and Pattern Recognition, 2015, 31(42): 10226-38.
- [17] RAY N, MCARDLE S, LAY K, et al. Mistica: Minimum spanning tree-based coarse image alignment for microscopy image sequences[J]. IEEE Journal of Biomedical Health Informatics, 2015:1.

## 作者简介

李慧敏(通讯作者),1992 年出生,硕士研究生,主要研究方向为图像处理、计算机视觉。

E-mail: lihuimin0112@shu.edu.cn

张之江,1968 年出生,教授,主要研究方向为计算机视觉、数字全息等。

E-mail: zjzhang@staff.shu.edu.cn