

# 基于 Co-motion 的可见光-热红外图像序列自动配准算法

张秀伟<sup>1</sup> 张艳宁<sup>1</sup> 杨涛<sup>1</sup> 张新功<sup>1</sup> 邵大培<sup>1</sup>

**摘要** 提出了一种基于 Co-motion 的可见光-热红外图像序列自动配准方法, 引入 Co-motion 运动统计特征来解决异源图像序列配准问题, 从而避开了异源图像相似图像特征提取和精确运动检测的难题. 由于可见光和热红外成像机理不同, 在运用运动统计特征时, 会面临与同源配准不同的诸多问题, 如难以提取相似前景、大量外点干扰和易受大尺度变化的影响等. 本文通过对所面临难题的分析, 提出了适合于异源图像序列配准的同名点对选取方法和外点去除方法, 并精确迭代优化变换模型. 对俄亥俄州大学红外-可见光数据库和自建数据库共 8 组数据的实验分析结果表明, 该算法在大尺度、旋转、平移及视场角变化下均能精确配准.

**关键词** 图像配准, Co-motion, 可见光图像序列, 热红外图像序列

**DOI** 10.3724/SP.J.1004.2010.01220

## Automatic Visual-thermal Image Sequence Registration Based on Co-motion

ZHANG Xiu-Wei<sup>1</sup> ZHANG Yan-Ning<sup>1</sup> YANG Tao<sup>1</sup> ZHANG Xin-Gong<sup>1</sup> SHAO Da-Pei<sup>1</sup>

**Abstract** An automatic visual-thermal image sequence registration method based on co-motion was proposed. Different from other methods, co-motion (concurrent motions) statistics feature was adopted to regist heterogeneous image sequences. Compared with registration based on co-motion between homogeneous image sequences, our method faced many difficulties. By analyzing the difficulties, we proposed the corresponding point pairs extraction method and outliers removal method, which are suitable to visual-thermal image sequence. We demonstrated the performance of the method on eight groups of visual-thermal synchronous video sequences, and the results showed that the proposed algorithm carried out precise image registration under the change of image rotation, translation, scaling and viewing angle. Experiments demonstrated the accuracy and robustness of the proposed method.

**Key words** Image registration, concurrent motion (co-motion), visual image sequence, thermal image sequence

目前, 具有大范围、多视点、全天时等优点的多源图像传感器目标监控系统已成为视觉监控领域的研究热点<sup>[1-2]</sup>. 热红外、可见光传感器是该类监控系统常用的两类图像传感器, 它们之间的配准是多源图像信息融合和协同的前提, 是多源信息处理系统的关键问题之一. 图像配准是将不同传感器、不同时间、不同角度所获得的两幅或多幅图像变换到同一坐标系下的过程.

可见光-热红外图像配准除了面临同源图像配准中的旋转变换、尺度变化、仿射变形等问题外, 还

面临成像机理不同带来的如下诸多问题:

1) 灰度、纹理特征呈现巨大差异: 如图 1 中 4 组图片所示, 两类传感器在同一时刻对同一场景所成的像, 无论在灰度、灰度对比度还是纹理上差异都很大.

2) 难以提取相似几何特征: 角点、线、边缘等几何特征为同源图像配准所常用, 然而并非所有场景都拥有丰富的几何特征. 即使存在, 由于成像机理不同, 也难以保证在两类图像中都呈现出显著性, 如图 1(a) 所示, 图 1(a) 中可见光图像的显著线特征(用白色矩形线框标识) 在红外图像中并非显著. 加之红外图像边缘模糊(图 1(a) 红外图像中的汽车周围存在光晕现象, 轮廓不清晰), 使得几何特征很难成为可见光-热红外配准的鲁棒特征.

3) 两类不同阴影的影响: 可见光图像的阴影由光照引起, 而热红外图像的热阴影或冷阴影由目标与周围环境温度差异造成. 两类阴影的存在, 使得对场景目标的精确提取和自动识别变得更加困难.

4) 时相性变化的影响: 可见光序列的时相性表现为对光照变化的敏感性, 而热红外序列的时相性表现为对周围环境温度变化的敏感性. 如图 1(b)、1(c) 和 1(d) 分别为正午、傍晚和夜晚对同一

收稿日期 2009-03-17 录用日期 2010-03-08  
Manuscript received March 17, 2009; accepted March 8, 2010  
国家高技术研究发展计划(863 计划)(2009AA01Z315), 国家自然科学基金(60872145, 60903126), 教育部科技创新工程重大项目培育资金(708085), 中国博士后科学基金(20090451397) 资助  
Supported by National High Technology Research and Development Program of China (863 Program) (2009AA01Z315), National Natural Science Foundation of China (60872145, 60903126), Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (708085), and China Postdoctoral Science Foundation (20090451397)  
1. 西北工业大学计算机学院陕西省语音与图像信息处理重点实验室 西安 710129  
1. Shanxi Provincial Key Laboratory of Speech and Image Information Processing, College of Computer Science, Northwest Polytechnical University, Xi'an 710129



(a) 道路场景红外-可见光图像对

(a) Thermal-visual image pair of path scene



(b) 建筑物场景正午时的红外-可见光图像对

(b) Thermal-visual image pair of building scene at noon



(c) 建筑物场景傍晚时的红外-可见光图像对

(c) Thermal-visual image pair of building scene in the afternoon



(d) 建筑物场景夜晚时的红外-可见光图像对

(d) Thermal-visual image pair of building scene at night

图 1 热红外-可见光图像配准示例

Fig. 1 Examples of thermal-visual image pair

场景拍摄的三组图像, 它们间的灰度变化十分明显。

与单类型传感器相比, 从红外和可见光图像中提取一致性图像特征的困难更大. 因此基于静态图像特征的异源配准方法很难获得鲁棒的配准结果, 它主要包括基于像素灰度的配准<sup>[3-4]</sup>、基于图像几何特征的配准<sup>[5-7]</sup>和基于图像分类的配准<sup>[8]</sup>.

基于运动特征的配准是近年来出现的一类颇受关注的方法, 该类方法利用图像序列中的运动信息(运动目标的侧影、轮廓、轨迹等)进行配准. 在同源图像序列配准中该方法使用的较多, 它们大多需要精确的运动轨迹或运动检测结果<sup>[9-10]</sup>. Ju 等<sup>[11]</sup>也尝试利用运动信息来配准可见光-热红外图像序列, 在仅允许一个目标在室内场景中运动的情况下, 通过提取准确的运动目标轮廓进行可见光-热红外

图像序列的配准. 由可见光-热红外图像配准面临的问题可知, 在无任何约束(如背景和运动都很复杂的室外监控场景)情况下, 同时从可见光序列和热红外序列中获得精确的运动信息是十分困难的. 由 Szlavik 等提出的基于 Co-motion 的可见光序列配准<sup>[12-16]</sup>为运动特征配准注入了新鲜血液, 它是一类不依赖精确运动检测的配准方法, 通过对具有部分重叠视野的同步视频的运动检测, 统计两个序列中像素点同时发生运动的情况, 并依据该统计信息选择同名点对.

基于 Co-motion 的配准方法为可见光-热红外图像序列自动配准提供了可行的参考. 但是由于成像机理不同, 将 Co-motion 特征运用于可见光-热红外图像序列配准仍面临一些问题:

1) 要求观测场景中具有可视的、运动的热目标. 它是确保两序列在重叠区域中提取的前景具有交集的前提. 这一点较易满足, 因为在大多数监控场景中, 被观测的对象多为行人和汽车.

2) 两同源图像序列在重叠区域内的前景-背景对比度具有相似性, 因此两图像序列在重叠区域内的运动检测结果较为一致, 基于检测结果的运动统计特征也呈现出较强的相似性. 然而可见光和热红外具有不同的成像特性, 在重叠区域内两者的前景-背景对比度差异较大, 如何保证两图像序列在重叠区域内的运动前景具有尽量大的交集是影响运动统计特征使用的关键问题.

3) 文献 [12-16] 没有给出大尺度变化下配准结果, 也没有针对该问题进行分析和说明. 相对于同源系统, 异源系统更易面临传感器间尺度变化较大的问题. 这是由于热红外传感器较为昂贵, 异源系统更倾向于用热红外进行大范围观测和定位, 用可见光传感器获得精细观测信息.

4) 相对于同源序列配准, 不同的成像特性加之红外成像对比度低、边缘模糊, 使得不同源配准面临大量的外点干扰, 文献 [12-16] 等提出的外点去除方法并不适宜.

本文通过引入 Co-motion 特征来解决异类传感器图像序列配准问题. 由于同源配准和异源配准存在很大差异, 本文针对在异源配准应用 Co-motion 所面临的上述难题, 提出了适宜于异源序列的特征提取方法和外点剔除方法. 最后使用 LM (Levenberg-Marquardt) 算法<sup>[17-18]</sup>迭代优化模型参数. 实验结果表明, 本文算法在大尺度、旋转及平移变化下均能精确配准.

## 1 系统框架

图 2 给出了可见光-热红外图像序列配准系统的流程图, 它以可见光-热红外同步视频为输入, 主

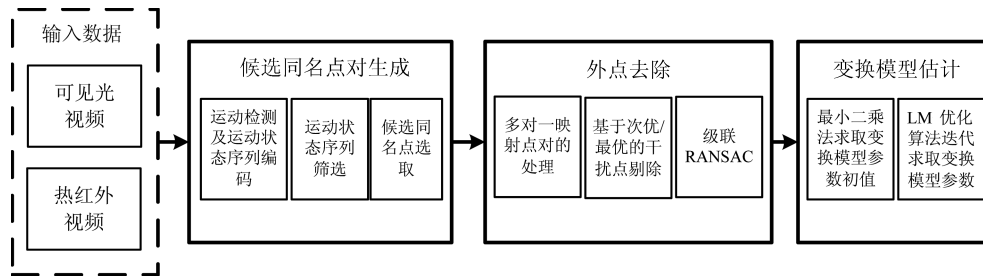


图2 基于 Co-motion 的多源图像序列自动配准算法流程图

Fig.2 Flow chart of the proposed method

要包括同名点对生成, 外点去除和变换模型估计三个模块. 其中, 可见光-热红外同步视频是指可见光图像序列和热红外图像序列中的各对应帧为同一时刻采集. 各模块概要说明如下:

### 1) 候选同名点对生成

针对可见光-热红外图像序列的特点, 分别对输入的可见光和热红外视频进行运动区域检测, 使得二者的运动前景尽量相似. 在运动检测结果的基础上, 为图像上的每个像素坐标生成运动状态统计特征描述. 并在基准图像序列和待配准图像序列间选择运动状态最相似的点对作为候选同名点对. 由于大量点的运动状态相似性度量是比较费时的, 因此在生成候选同名点对前, 需要对参与度量的点进行筛选.

### 2) 外点去除

首先分析了可见光-热红外图像序列所产生干扰点的特点, 特别是“多对一映射”干扰点对在配准中的影响, 然后使用次优/最优结合多层级联的随机抽样一致性算法 (Random sample consensus, RANSAC)<sup>[19]</sup> 进行干扰点去除.

### 3) 变换模型估计

将2)生成的同名点对作为输入, 使用最小二乘法估计变换参数初值, 并使用LM算法迭代优化两图像序列间的变换参数.

## 2 基于多分辨率 Co-motion 的可见光-热红外图像序列配准算法

### 2.1 基于 Co-motion 的候选同名点对生成

#### 1) 运动区域检测及运动状态序列编码

由第1节中分析可知, 在同步序列公共区域内能否检测出可见光和热红外同时可见的运动目标是影响 Co-motion 特征在异源序列配准中应用的关键问题. 考虑到改进的混合高斯背景差算法<sup>[20]</sup> 具有较好背景变化适应能力和一定的阴影去除能力, 本文采用其进行可见光图像序列的运动检测. 但红外图像序列对比度低、时向性变化快等特点使得背景差

方法易出现大量虚警和漏检, 因此本文采用灰度分层算法<sup>[21]</sup> 检测运动的热目标. 并使用形态学滤波对两序列的检测结果进行膨胀, 以使得同步序列公共区域内的检测结果尽量相似.

这里使用时间序列来描述像素点的运动状态. 将像素点的二值运动检测结果按照时间顺序排列, 形成一个二值序列  $\text{seq}(i, j)$ , 如式(1)所示(见下页上方), 称之为像素点  $(i, j)$  的运动状态序列, 这与文献[16]类似. 其中,  $A$  为基准图像序列,  $B$  为待配准图像序列,  $N_A$ 、 $M_A$  为基准图像  $A$  的宽和高,  $N_B$ 、 $M_B$  为待配准图像  $B$  的宽和高,  $T$  为参与配准的同步视频的长度.

#### 2) 运动状态序列筛选

文献[14-15]所使用的基于熵的运动状态序列筛选方法和基于非极大值抑制的筛选方法都不适合于可见光-热红外序列配准. 这是由于基于熵的筛选方法需要背景-前景差值, 公共区域内同源序列的背景-前景差值具有相似性, 而异源的差异较大. 对于非极大值抑制筛选, 由于可见光-热红外序列提取的前景存在很大差异, 通过该方法筛选出的序列, 在待配准序列中并不能够找到精确的对应点, 从而降低了配准的精度.

通过对可见光-热红外序列运动检测结果的分析发现, 在基准和待配准图像序列中一些像素点几乎没有运动, 而另外一些像素点则频繁运动, 这两类像素点多由噪声引起. 我们将这类像素点的运动状态序列去除, 使其不参与后续候选同名点对选取计算, 一则通过去除噪声提高了同名点对的精确度, 二则通过减少参与计算的运动状态序列个数降低了运算量. 本文采用双阈值法对运动状态序列进行筛选, 其计算公式如式(2)(见下页上方)所示, 当  $\text{validateSeq}(L, i, j)$  取值为 true 时, 表示集合  $L$  中点  $(i, j)$  的运动状态序列有效; 否则为无效的运动状态序列. 其中,  $v_{\max}$  为两图像序列中像素点发生运动次数的上限,  $v_{\min}$  为两图像序列中像素点发生运动次数的下限,  $low\_thd$ 、 $hgh\_thd$  为低阈值和高阈值.



$$\begin{aligned} \text{seq}_L(i, j) &= [\text{seq}_L(i, j, 1), \dots, \text{seq}_L(i, j, t), \dots, \text{seq}_L(i, j, T)] \\ \text{seq}_L(i, j, t) &= \begin{cases} 0, & (i, j) \text{ detected as background at } t\text{-frame} \\ 1, & (i, j) \text{ detected as foreground at } t\text{-frame} \end{cases} \end{aligned} \quad (1)$$

$$1 \leq i \leq N_L, 1 \leq j \leq M_L, L \in \{A, B\}$$

$$\begin{aligned} \text{validateSeq}(L, i, j) &= \left( \sum_{t=1}^T \text{seq}_L(i, j, t) > v_{\min} \right) \wedge \left( \sum_{t=1}^T \text{seq}_L(i, j, t) < v_{\max} \right), L \in [A, B] \\ v_{\max} &= \max_{1 < i < N_L, 1 < j < M_L, L \in [A, B]} \left( \sum_{t=1}^T \text{seq}_L(i, j, t) \right) \end{aligned} \quad (2)$$

$$v_{\min} = v_{\max} \times \text{low\_thd}$$

$$v_{\max} = v_{\max} \times \text{hgh\_thd}$$

$$\text{mapA}(i, j) = \min_{(i, j) \in V} \left( \frac{\sum_{t=1}^T |\text{seq}_A(i, j, t) - \text{seq}_B(p, q, t)|}{\sum_{t=1}^T \text{seq}_A(i, j, t)} \right) \quad (3)$$

### 3) 候选同名点对选取

运动状态序列是二值序列, 海明距离是度量这类序列相似性的一种有效方法. 考虑到运动序列间的距离还与运动状态序列 (1) 中的总数相关, 本文采用了如式 (3) 所示的相似性度量. 对基准集合  $A$  中的有效运动状态序列, 从待配准集合  $B$  的有效运动状态序列中, 选择与其距离最近的像素点作为匹配点, 共同形成候选同名点对. 函数  $\text{mapA}(i, j)$  求取的就是与基准  $A$  中某像素点运动状态最相似的  $B$  像素集合. 其中,  $V$  为图像  $A$  满足式 (2) 的像素点集合;  $U$  为图像  $B$  满足式 (2) 的像素点集合.

## 2.2 外点去除

### 1) 多对一映射的处理

基于 Co-motion 生成的候选匹配点对中存在两种多对一映射, 即一个基准点对应多个最优候选匹配点的映射和多个基准点对应同一个匹配点的映射. 在这些多对一映射中, 有些是由于尺度差异造成的, 属于正确的配准点对, 还有些是由于运动检测不准确或者非重叠区域的干扰点造成的, 属于错误的配准点对, 正确与错误的点对很难区分, 文献 [12-13, 16] 采取了直接剔除多对一映射的点对. 随着基准和待配准图像序列间尺度差异的增大, 由尺度差异引起的多对一映射的数目将急剧增加, “多对一映射干扰点的去除”会把大量正确的匹配点对剔除, 干扰点去除后的候选配准点对中错误点对比例急剧增大, 可致使错误模式成为最终的同名点对, 导致配准结果不准确, 因此文献 [12-13, 16] 并不适合于存在大尺度差异的图像序列配准. 图 3 给出了大

尺度变化下采用多对一映射去除干扰点的一个示例, 从图 3 (b) 中可以看出, 经多对一映射去除后, 剩下的同名点对大多为错误点对.

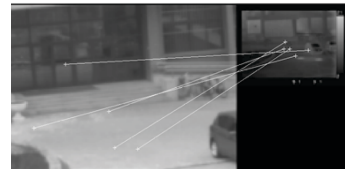


(a) 候选同名点对

(a) Original corresponding point pairs



(b) 去除多对一映射的候选同名点对  
(b) Corresponding point pairs after removing  
“N-1 mapping”



(c) RANSAC 算法产生的同名点对  
(c) Corresponding point pairs selected by RANSAC

图 3 大尺度变化下采用“多对一映射去除”的外点去除结果

Fig. 3 Removal results of “N-1 mapping” point pairs

为了保证可见光-热红外序列在大尺度变化下也能够精确配准, 本文不采用“多对一映射干扰点的去除”。由于在多对一映射中正确与错误的点对并存, 这一操作可能引入大量的外点, 后续的外点去除方法需针对这一问题进行处理。

### 2) 基于次优/最优的干扰点去除

对于正确的同名点对而言, 基准点除了与最优候选配准点的运动状态相似外, 也与最优候选配准点邻域内的一些点运动状态相似。因此最优距离(即基准点与最优候选配准点间的相似性度量值)与次优距离应当相距不远。本文基于此原理进行错误点对的去除, 通过对次优最优间差值的阈值化进行干扰点的初步筛选。

### 3) 级联 RANSAC 算法

由于异源图像序列所面临的运动检测不一致性及两类多对一映射点对的保留, 异源序列相对于同源, 在提取的候选同名点对中将面临更多不符合实际模型的错误匹配点对(称之为外点 Outliers)。

由 Fischler 和 Bolles 提出的 RANSAC 算法<sup>[19]</sup>, 对于错误超过 50% 的数据仍然能够处理, 是一种鲁棒的外点去除方法。因此, 本文采用该算法来剔除外点。但在应用该算法时, 由于候选同名点对的外点比例过高, 使得算法在运行成千上万次后仍很难获得满足要求的同名点对。针对该问题本文设计了级联 RANSAC 算法, 通过逐次降低容忍度值来剔除外点, 从而保证了同名点对的精确度。级联 RANSAC 算法流程如图 4 所示, 最小容忍度选择为 0.01。

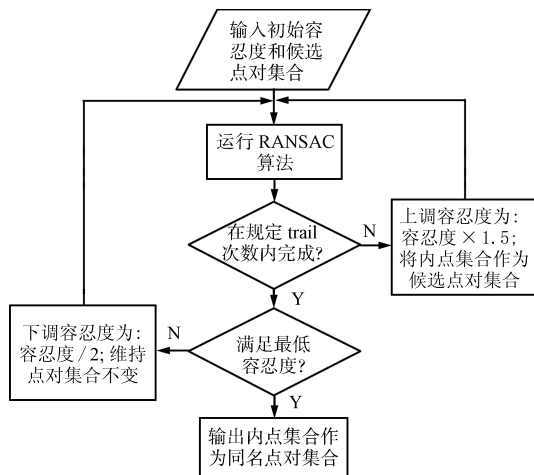


图 4 级联 RANSAC 算法的流程示意图

Fig. 4 Our flow chart of cascade RANSAC method

## 2.3 鲁棒的模型估计

在获得了同名点对后, 如何精确计算变换模型参数是模型估计的工作重点, 变换模型参数估计主要包含两个步骤, 变换模型选择和模型参数估计。

本文方法采用的是图像像素的运动统计特征, 同一像素上的运动都认为是空间中同一点被运动物体遮挡所引起, 较适合于中远距离场景。透视投影变换可满足上述的假设, 其模型变换函数  $H(\cdot, \cdot)$  及参数矩阵形式为

$$\begin{cases} x' = \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + 1} \\ y' = \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + 1} \end{cases}$$

$$H_{\text{perspective}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \quad (4)$$

其中,  $(x, y)$  为基准点,  $(x', y')$  为匹配点。

透视投影模型属非线性变换, 尽管可以将其线性化, 但仍然容易面临严重的病态, 甚至退化的问题。LM 优化算法<sup>[17-18]</sup> 是一种无条件约束的优化方法, 拥有梯度下降算法的全局特性和高斯牛顿法的局部收敛性。当优化参数接近误差目标时, 算法能以最快的速率收敛, 同时保持了主矩阵的正定性和空间解的稳定性。因此, 本文使用 LM 算法来精确估计透视投影变换模型的参数。

本文使用同名点对矩阵变换前后坐标值欧氏距离的平方和作为误差指标函数  $E(H)$ , 并通过使  $E(H)$  最小化来获得精确的模型参数。  $E(H)$  如下所示:

$$E(H) = \sum_{i=1}^N [(x''_i - x_i)^2 + (y''_i - y_i)^2] \quad (5)$$

$$(x'', y'') = H(x', y')$$

其中, 点  $(x_i, y_i)$  是第  $i$  组同名点对中的基准点坐标,  $(x'_i, y'_i)$  是第  $i$  组同名点对中匹配点坐标,  $(x''_i, y''_i)$  是  $(x'_i, y'_i)$  经模型变换函数  $H(\cdot, \cdot)$  变换获得的新坐标,  $N$  为同名点对的总数。求解过程中, 通过计算  $E(H)$  的 Jacobian 矩阵调节步长来寻找最优解, 这一点与梯度下降法类似, 并使用  $\mu$  保证主矩阵的正定性。LM 算法迭代过程如下:

$$\begin{cases} H(k+1) = H(k) + \Delta H \\ \Delta H = -(J^T(H)J(H) + \mu I)^{-1} J^T(H)E(H) \\ E(H) = \frac{1}{2} \sum_{i=1}^N e_i^2(H) \end{cases} \quad (6)$$

其中,  $J(H)$  是误差指标函数  $E(H)$  的 Jacobian 矩阵,  $\mu$  为比例系数,  $I$  为单位矩阵。

### 3 实验结果与分析

以本文算法为核心, 实现了一套可见光-热红外图像序列配准软件系统. 该系统以可见光-热红外同步图像序列为输入, 通过本文提出的配准算法计算图像序列间的变换模型参数, 并可针对两图像序列中的任意帧同步图像进行拼接和融合.

为了验证本文算法的有效性, 这里选择公开发布的俄亥俄州大学红外-可见光数据库中的 2 组同步序列和自建数据库中的 6 组同步序列作为实验数据. 实验数据的分辨率为 320 像素  $\times$  240 像素, 帧率为 20 帧/秒, 每组序列长度 400 帧, 8 组数据共 3200 帧. 实验数据所拍摄目标包括行人、车辆等, 拍摄的时段包括白天、傍晚和夜晚. 本文算法需要可见光-热红外图像序列为同步数据. 在自建数据库时, 通过在图像序列采集端之间使用延迟测量时间同步 (Delay measurement time synchronization, DMTS) 算法保证各采集端的系统时间同步, 并在此基础上使用时间对齐方式完成数据同步. 红外热像仪采用的是武汉高德公司生产的 IR913A 型非制冷焦平面红外热成像仪.

在选择和拍摄实验数据时, 考虑到配准问题的特点, 将数据分为平移变化、较大尺度变化、大尺度旋转变化和大视场角变化四类数据. 以下将从鲁棒性和精确度两个方面对算法进行分析.

#### 3.1 算法鲁棒性分析

这里首先分析了存在几何变化情况下的配准, 针对仅有平移和小角度变化的数据说明基本实验过程和实验结果, 使用较大尺度变化的实验数据、存在较大尺度旋转平移变化的数据以及大视场角变化下的数据说明算法的有效性; 最后又对其他典型影响因素进行了分析, 包括采集设备分辨率不同、帧频不同、可见光阴影、可见光-热红外图像运动目标轮廓及面积存在差异等情况.

#### 3.1.1 几何变化

##### 1) 平移变化

为测试算法在平移变化下的配准效果, 选择三组数据作为实验数据, 它们分别为俄亥俄州大学红外-可见光同步序列 I、自建同步序列 I 和自建同步序列 II. 此处使用自建同步序列 I (如图 5 所示) 作为输入图像序列来说明实验过程.

图 5 显示了自建同步序列 I 的部分图像. 经过第 2.1 节描述的运动检测方法, 该可见光-热红外同步序列得到的运动前景如图 6 所示. 虽然受到阴影、虚警、漏检等的影响, 该方法仍保证了重叠区域内两序列运动检测结果具有较高的相似性, 从而为 Co-motion 特征的应用奠定基础.

图 7 给出了运动状态序列筛选前后的对比结果, 其中累加图是指将图像序列中每个运动状态序列沿时间轴进行累加并最后二值化后的图像. 红外序列和可见光序列中均有近 50% 的运动状态序列被去除. 筛选后的红外运动状态累加图和可见光累加图已具有一定的相似性了. 其中, *low\_thd* 为 0.05, *hgh\_thd* 为 0.85.

图 8(a) 给出了基于运动统计特征产生的候选同名点对, 从图中可以看出候选同名点对集合中存在大量的错误匹配点对; 图 8(b) 为基于次优/最优方法剔除的干扰点, 它们中的绝大多数为错误匹配点对, 可以看出该方法的有效性; 图 8(c) 为经过级联 RANSAC 算法得到的同名点对, 级联 RANSAC 输出了精确的同名点对, 有效地抑制了外点. 10 组实验中次优/最优干扰点剔除方法采用的阈值均为 0.3.

基于上述所获得同名点对集合, 本文方法通过 LM 结合最小二乘模型估计获得了精确的模型参数. 为了清晰、直观地展示配准的效果, 本文采用拼接融合方法来展示所给算法的效果. 图 9 分别给出了自建同步序列 I、俄亥俄州大学同步序列 I 和自建同步

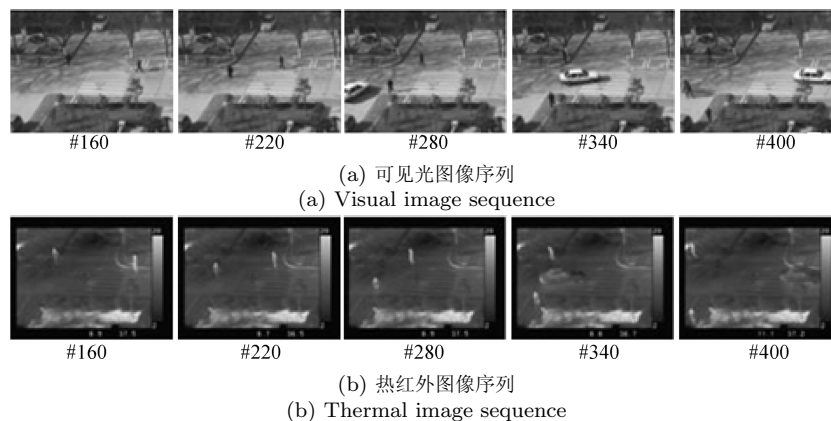


图 5 自建同步序列 I

Fig. 5 Image Sequence I in our database

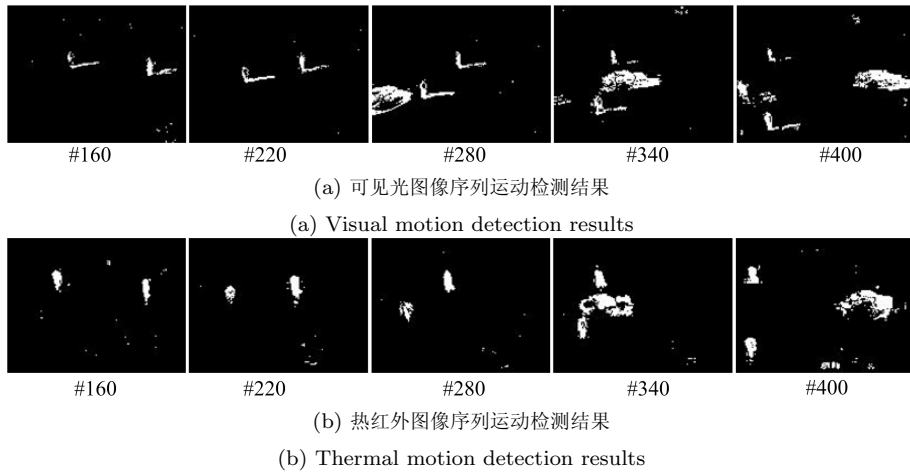


图 6 运动检测结果

Fig.6 Motion detection results

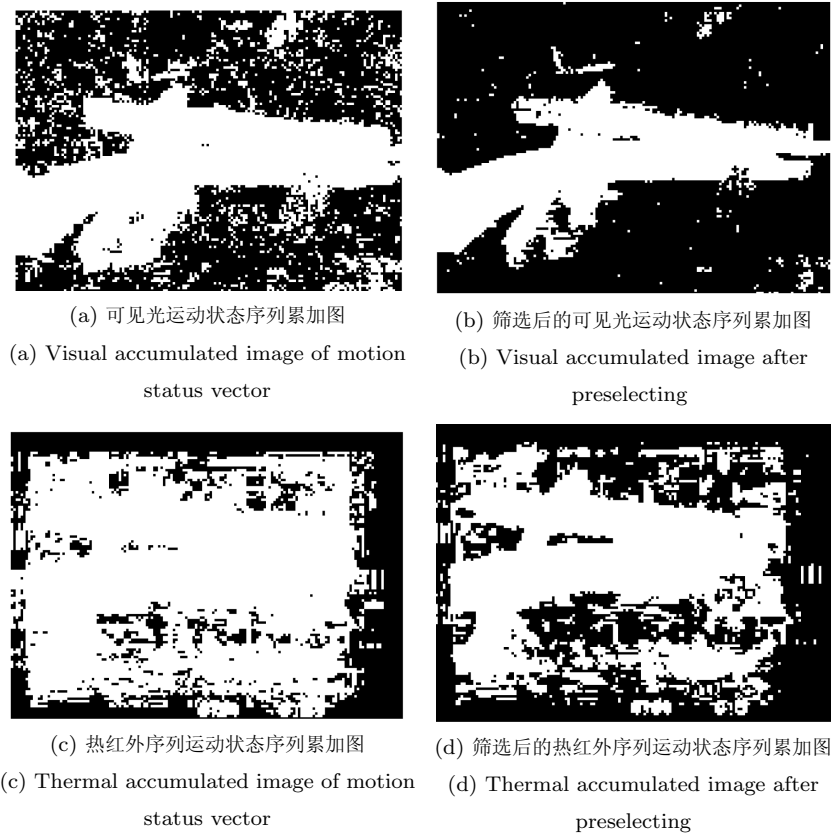


图 7 运动状态序列筛选前后的对比

Fig.7 Preselecting results of motion status vectors

序列 II 某一帧图像的拼接融合结果图, 从图中可以看出两幅异源图像得到了很好的对准, 特别是图 9(c) 中楼门前两根柱子的边缘衔接很自然.

2) 较大尺度变化

此处使用自建同步序列 III 来验证本文算法在

大尺度变化下的配准效果. 图 10 给出了自建同步序列 III 中某一帧图像的拼接融合结果图, 其中行人、汽车及大楼门窗得到了对准, 但拼接图中汽车和楼门柱边缘的衔接部分没有图 9 中的那么自然.

3) 较大尺度旋转变换



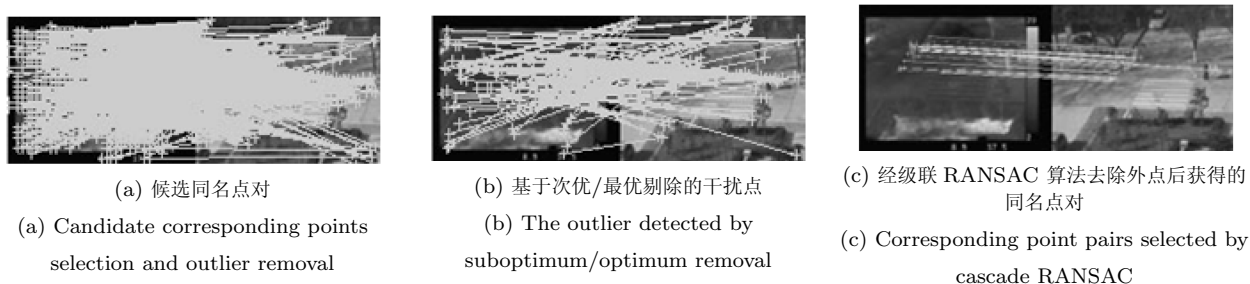


图 8 候选同名点对选取及外点去除

Fig. 8 Candidate corresponding points selection and outlier removal



图 9 平移变化下可见光 - 热红外序列某帧图像的拼接融合结果图

Fig. 9 Registration results under change of translation

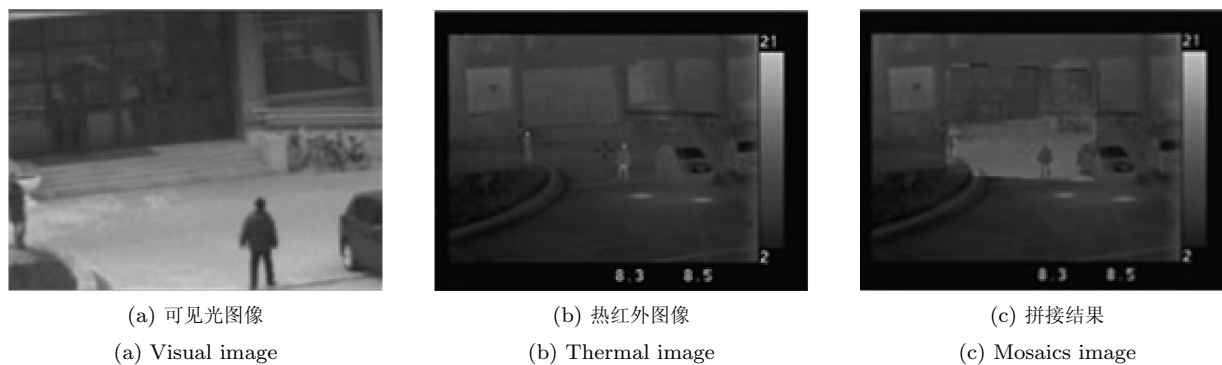


图 10 大尺度变化下的配准结果 (自建同步序列 III)

Fig. 10 Registration results under the change of large scale (Sequence III)

本文使用自建同步序列 IV 和自建同步序列 V 来说明算法在大尺度旋转变化下的配准效果. 同步序列 IV 为傍晚拍摄, 同步序列 V 为夜间拍摄. 这两组同步序列某一帧的拼接融合结果如图 11 和图 12 所示. 图中的可见光图像和热红外图像间存在着很大的尺度、旋转和平移变化, 即便在这种情况下, 拼接融合的结果图中的行人、窗户及花坛边缘仍旧得到了很好的对准, 说明了本文算法在大尺度旋转情况下的有效性.

4) 大的视场角变化

由于目前公开的红外 - 可见光同步序列数据大

多是由于融合, 它们几乎都是同一视角拍摄. 因此, 本文采用自建数据 VI 来说明实验效果. 在拍摄同步序列 VI 时, 热红外相机朝向某建筑的门口拍摄, 而可见光相机斜向该建筑的东南方向拍摄. 图 13 给出了实验结果, 从拼接结果可以看出, 场景中花坛的边缘、运动着的行人得到了很好的对准. 在中远场景下本文算法可以适应于较大的视场角变化.

通过平移、大尺度、较大尺度旋转和大的视场角变化四类实验的分析可知, 本文算法具有很好的鲁棒性, 不易受这四类变化的影响. 需要说明的是, 在大尺度变化下, 配准算法的精确度稍有所下降, 这



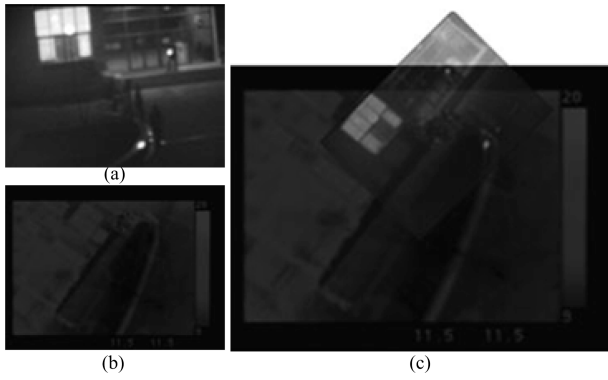


图 11 较大尺度旋转变换下的配准结果 (自建同步序列 IV)  
((a) 可见光图像; (b) 热红外图像; (c) 拼接结果)

Fig. 11 Registration results under the change of large scale and rotation (Sequence IV)

((a) Visual image; (b) Thermal image; (c) Mosaics image)

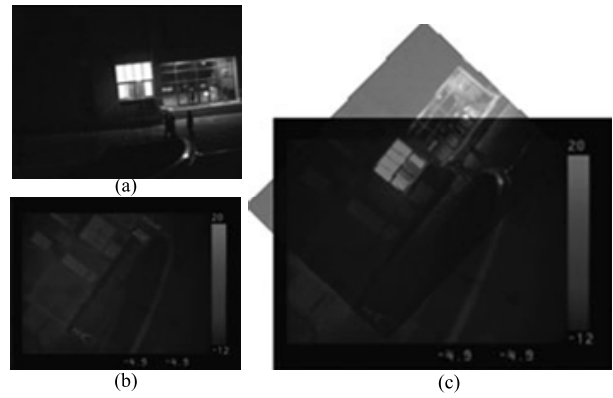


图 12 较大尺度旋转变换下的配准结果 (自建同步序列 V)  
((a) 可见光图像; (b) 热红外图像; (c) 拼接结果)

Fig. 12 Registration results under the change of large scale and rotation (Sequence V)

((a) Visual image; (b) Thermal image; (c) Mosaics image)

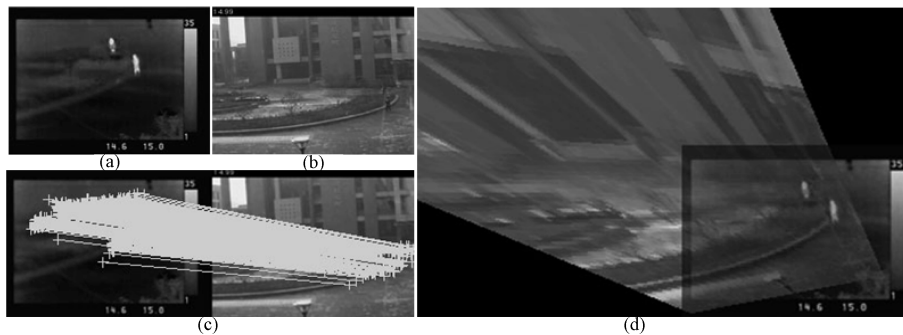


图 13 大视场角变化下的配准结果 (自建同步序列 VI)

((a) 可见光图像; (b) 热红外图像; (c) 同名点对; (d) 拼接结果)

Fig. 13 Registration results under the change of large viewing angle (Sequence VI)

((a) Visual image; (b) Thermal image; (c) Corresponding points; (d) Mosaics image)

是由于在较大尺度情况下提取的候选同名点对中大量多对一映射所造成的影响. 尽管经过外点去除, 但要从大量正确的多对一映射中选择出精确的同名点对仍比较困难.

### 3.1.2 其他典型影响因素

#### 1) 采集设备帧频不同

本文对运动状态序列的相似性度量是以沿时间序列逐点求取海明距离再求和的方式进行的, 如第 2.1 节式 (3) 所示. 因此要求参与相似性度量的两个运动状态序列中各点对齐, 也即要求可见光图像序列和热红外图像序列数据对齐. 在采集设备的帧频不一样的情况, 应用本文算法时, 需要采取时间序列预处理方法使得两图像序列数据对齐.

#### 2) 采集设备分辨率不同

本文算法并不要求可见光图像序列和热红外图像序列具有相同的图像分辨率. 现有的可见光-热

红外同步数据大多分辨率是相同的, 因此这里采用模拟数据来说明在不同分辨率下本文算法的可行性. 将自建同步序列 III 中的热红外序列进行裁剪后再降采样获得同步序列 III', III' 中可见光图像分辨率为  $320 \text{ 像素} \times 240 \text{ 像素}$ , 热红外图像分辨率为  $70 \text{ 像素} \times 50 \text{ 像素}$ . 同步序列 III' 的拼接结果与同步序列 III 的配准结果相比较, 无明显变化, 因此本文算法可以适应于采集设备分辨率不相同的情况.

#### 3) 可见光阴影

相对于可见光-可见光图像序列配准, 可见光-热红外图像序列配准还需面临阴影的影响. 在存在可见光阴影的情况下, 因为运动阴影伴随着运动目标, 在目标经过场景的某一时刻, 会在不同区域 (目标区域和阴影区域) 产生相似的运动状态; 统计分析目标经过场景的整个时段, 存在两种情况, 目标和阴影在整个时段内作用的场景区域不重叠或重叠.

对于目标和阴影作用的场景区域不重叠的情况,

红外像素点运动状态与可见光中目标和阴影引起的运动都是相似的, 从而引入了多对一候选同名点对, 这些候选同名点对中正确和错误的匹配点对并存, 如第 2.2 节所述, 本文不采用“多对一映射去除”, 使用级联 RANSAC 来处理。

对于目标和阴影作用的场景区域重叠的情况, 该区域内的可见光像素点存在着由阴影和目标引起的两段运动. 相对于红外序列中的运动, 可见光像素点的运动状态均存在由阴影带来的误差, 因此可见光像素点运动状态相似性度量仍然主要取决于目标所形成的运动特征. 尽管会引入一些错误的匹配点对, 但可见光阴影并不影响全局的运动状态相似性度量。

由上述的分析可知, 本文算法不易受可见光阴影的影响. 自建同步序列 I 的可见光图像序列中行人和汽车都存在着阴影, 其实验结果 (图 9 (a)) 从一定程度上验证了这个问题. 为了进一步说明, 此处使用俄亥俄州大学同步序列 II 进行再次实验验证, 该序列的场景中有云朵从天空飘过, 存在很大的光照变化, 具有较严重阴影, 包括行人自身的阴影和云朵引起的阴影. 图 14 给出了俄亥俄州大学同步序列 II 其中一帧图像对的拼接结果, 结果表明可见光阴影并未影响到配准结果。

4) 可见光和热红外图像运动目标的轮廓和面积差异较大情况

可见光和热红外图像运动目标轮廓和面积存在较大差异的情况, 可能由热辐射效应、较大的尺度变化、大视场角变化、可见光能见度低、可见光运动目标与背景相似、场景温度与运动目标温度差异较小等多种因素引起. 在上述的 10 组实验中, 可见光和热红外图像的运动目标轮廓和面积都存在着一一定的差异. 在自建同步序列 I 中, 由于热辐射效应热红外图像中目标边缘模糊, 运动目标轮廓不同于可见光; 自建同步序列 III 中可见光能见度低, 目标轮廓不清晰; 自建同步序列 IV 中可见光序列和热红外序列间存在着很大的尺度差异; 自建同步序列 VI 中, 由于视场角差异, 热红外序列中的目标面积明显大于可见光序列。

在本文算法中, 运动目标轮廓、面积存在差异时, 会造成可见光和热红外图像运动前景提取不一致, 从而使得同一个点的两个运动状态序列存在差异. 同时较大的面积差异, 也会引入大量的多对一候选同名点对. 但是由于本文采用的是统计特征, 从统计意义上度量两个运动状态序列的相似性, 并使用鲁棒的外点去除方法, 去除了大量错误点的干扰. 自建同步序列 I、III、IV、VI 的实验均取得了较好的结果, 因此本文算法不依赖于精确运动前景检测, 对运动目标轮廓和面积的差异有一定的抗干扰能力。

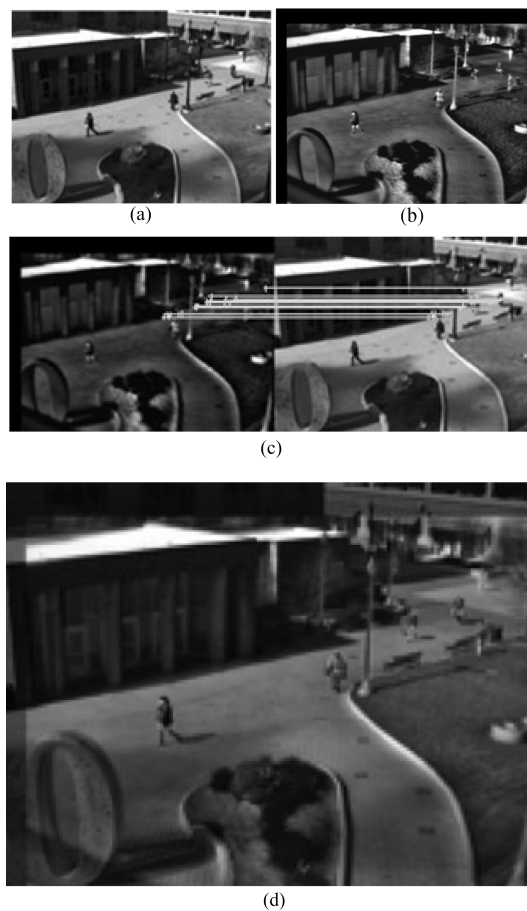


图 14 较严重阴影情况的配准结果 (俄亥俄州大学同步序列 II) ((a) 可见光图像; (b) 热红外图像; (c) 同名点对; (d) 拼接结果)

Fig. 14 Registration results in the heavy visual shadow situation (Ohio State University Sequence II) ((a) Visual image; (b) Thermal image; (c) Corresponding points; (d) Mosaics image)

### 3.2 算法精确度和效率分析

由第 3.1 节 10 组实验的拼接结果可知, 拼接图中的行人、汽车、窗户等目标物均得到了较好的对准, 边缘衔接自然, 取得了理想的效果。

相对于图像特征的可见光-热红外图像配准, 本文算法具有更好的鲁棒性和适应性, 不易受到阴影、几何特征不显著以及可见光-热红外灰度纹理差异等因素影响; 相对于 Ju 等<sup>[11]</sup>提出的方法, 本文算法不需要人为参与, 不依赖于准确的运动检测, 属于自动配准算法。

针对可见光-热红外配准问题, 目前鲜有基于运动特征的图像序列自动配准方法, 大多数仍关注于单幅图像配准. 且可见光-热红外图像序列没有标准库, 很难找到与之比较的同类算法. 为了进一步评价算法精确度, 此处采用手动从多幅图像对中选择同名点对的方法获得模型参数, 并以此模型参数

为标准数据,使用“平均坐标误差”对本文算法的精确度进行分析。“平均坐标误差”是指以手动选择的基准点集合  $BasePts$  为输入,分别根据标准模型参数  $H_{Truth}$  和本文算法获得的模型参数  $H_{Alg}$  计算出匹配点集合  $MatchPts_T$  和  $MatchPts_A$ ,并计算这两个集合中对应坐标点间绝对误差的平均值.精确度分析如表 1 所示,从表 1 中可知大尺度变化相对于平移、旋转变化的分析一致.但 8 组实验结果的横坐标和纵坐标平均误差分别为 0.86 和 1.48 个像素,该精度已经足以满足多源相机监控系统中协同检测、跟踪处理的要求.

表 1 本文算法精度分析表

Table 1 Accuracy analysis of the proposed method

实验数据	坐标误差平均值 ( $\overline{\Delta x}$ , $\overline{\Delta y}$ )
俄亥俄州大学同步序列 I	(0.56, 0.64)
俄亥俄州大学同步序列 II	(0.39, 0.44)
自建同步序列 I	(0.74, 1.30)
自建同步序列 II	(0.65, 1.29)
自建同步序列 III	(0.70, 1.85)
自建同步序列 IV	(0.99, 1.92)
自建同步序列 V	(1.34, 1.91)
自建同步序列 V	(1.50, 2.56)
平均误差	(0.86, 1.48)

本文提出的异源图像序列配准为离线处理方法,算法在时间序列间的相似性度量方面耗时较长.在硬件配置为 AMD Athlon64 处理器(主频 1.8 GHz),1 GB 内存的机器上,运行环境为 Matlab 7.4.0,对于实验中的 8 组图像序列,分辨率为 120 像素  $\times$  160 像素时,可见光-热红外图像序列配准的平均运行时间为 6 分钟.运动状态序列间的相似性度量是其中较耗时的操作.

## 4 结论

针对可见光-热红外图像序列配准问题,提出了一种基于 Co-motion 的可见光-热红外图像序列自动配准方法.该方法具有以下特点: 1) 以同步图像序列为处理对象,而非单幅图像配准问题,可推广到多路同步图像序列,适用于多源图像传感器监控系统; 2) 实现了适合于异源图像序列的同名点对提取和外点去除; 3) 通过保留候选同名点对中多对一映射和有效的外点去除方法,实现了大尺度变化下的异源图像配准; 4) 通过引入 Co-motion 运动统计特征,避开了异源图像相似图像特征提取和精确运

动检测的难题.通过大量图像序列的测试,证实了本文算法在大的平移、尺度、旋转、视场角变化情况下均能取得好的结果.在进一步的研究中,拟研究如何自动确定适合配准的帧数.

## References

- 1 Krotosky S J, Trivedi M M. Person surveillance using visual and infrared imagery. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, **18**(8): 1096–1105
- 2 Chen Y X, Han C Z. Night-time pedestrian detection by visual-infrared video fusion. In: Proceedings of the 7th World Congress on Intelligent Control and Automation. Chongqing, China: IEEE, 2008. 5079–5084
- 3 Zhao J P, Yang H M, Ding Y. Medical image registration algorithm research based on mutual information similarity measure. In: Proceedings of the International Symposium on Photo Electronic Detection and Imaging. Beijing, China: SPIE, 2007. 6–9
- 4 Lu X, Zhang S, Su H, Chen Y. Mutual information-based multimodal image registration using a novel joint histogram estimation. *Computerized Medical Imaging and Graphics*, 2008, **32**(3): 202–209
- 5 Gao Feng, Wen Gong-Jian, Lv Jin-Jian. An optimal algorithm for IR/visual image registration based on main-line-pairs. *Chinese Journal of Computers*, 2007, **30**(6): 1014–1021  
(高峰, 文贡坚, 吕金建. 基于干线对的红外与可见光最优图像配准算法. 计算机学报, 2007, **30**(6): 1014–1021)
- 6 Hrkac T, Kalafatic Z, Krapac J. Infrared-visual image registration based on corners and Hausdorff distance. In: Proceedings of the 15th Scandinavian Conference. Aalborg, Denmark: Springer, 2007. 383–392
- 7 Kim Y S, Lee J H, Ra J B. Multi-sensor image registration based on intensity and edge orientation information. *Pattern Recognition*, 2008, **41**(11): 3356–3365
- 8 Lei Lin, Jiang Yong-Mei, Kuang Gang-Yao. A method of the remote sensing image registration based on image classification. *Journal of National University of Defense Technology*, 2004, **26**(2): 35–40  
(雷琳, 蒋咏梅, 匡纲要. 一种基于图像分类的遥感图像配准方法. 国防科技大学学报, 2004, **26**(2): 35–40)
- 9 Caspi Y, Simakov D, Irani M. Feature-based sequence-to-sequence matching. *International Journal of Computer Vision*, 2006, **68**(1): 53–64
- 10 Lee L, Romano R, Stein G P. Monitoring activities from multiple video streams: establishing a common coordinate frame. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 758–767
- 11 Ju H, Bhanu B. Fusion of color and infrared video for moving human detection. *Pattern Recognition*, 2007, **40**(6): 1771–1784
- 12 Szlavik Z, Havasi L, Sziranyi T. Estimation of common groundplane based on co-motion statistics. In: Proceedings of the International Conference on Image Analysis and Recognition. Porto, Portugal: Springer, 2004. 347–354



- 13 Szlavik Z, Havasi L, Sziranyi T. Image matching based on co-motion statistics. In: Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission. Thessaloniki, Greece: IEEE, 2004. 584–591
- 14 Szlavik Z, Sziranyi T, Havasi L, Benedek C. Optimizing of searching co-motion point-pairs for statistical camera calibration. In: Proceedings of the International Conference on Image Processing. Genoa, Italy: IEEE, 2005. 1178–1181
- 15 Benedek C, Havasi L, Sziranyi T, Szlavik Z. Motion-based flexible camera registration. In: Proceedings of the International Conference on Video and Signal Based Surveillance. Como, Italy: IEEE, 2005. 439–444
- 16 Szlavik Z, Sziranyi T, Havasi L. Video camera registration using accumulated co-motion maps. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2007, **61**(5): 298–306
- 17 Levenberg K. A method for the solution of certain nonlinear problems in least squares. *Quarterly of Applied Mathematics*, 1944, **2**(2): 164–168
- 18 Marquardt D W. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 1963, **11**(2): 431–441
- 19 Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981, **24**(6): 381–395
- 20 Brisson N M, Zaccarin A. Moving cast shadow detection from a Gaussian mixture shadow model. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 643–648
- 21 Yang Gen, Zhang Yan-Ning, Xi Run-Ping. Detection algorithm for infrared moving objects based on gray delaminating. *Infrared and Laser Engineering*, 2007, **36**(6): 980–983 (杨根, 张艳宁, 郗润平. 红外运动目标的灰度分层检测算法. *红外与激光工程*, 2007, **36**(6): 980–983)



张秀伟 西北工业大学计算机学院博士研究生. 主要研究方向为视频内容分析和多源信息协同处理技术. 本文通信作者. E-mail: weirui20011@163.com  
(ZHANG Xiu-Wei Ph.D. candidate at the College of Computer Science, Northwest Polytechnical University. Her research interest covers video

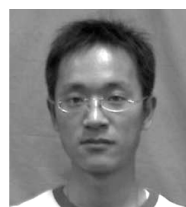
content analysis and multi-sensor information cooperative processing. Corresponding author of this paper.)



张艳宁 西北工业大学计算机学院教授. 主要研究方向为数字图像处理与模式识别. E-mail: ynzhang@nwpu.edu.cn  
(ZHANG Yan-Ning Professor at the College of Computer Science, Northwest Polytechnical University. Her research interest covers digital image processing and pattern recognition.)



杨涛 西北工业大学计算机学院博士后. 主要研究方向为视频内容分析、视频目标检测与跟踪、运动视频配准与稳像. E-mail: yangtaonwpu@163.com  
(YANG Tao Postdoctor at the College of Computer Science, Northwest Polytechnical University. His research interest covers video content analysis, visual object detection and tracking, and motion video registration and stabilization.)



张新功 西北工业大学计算机学院硕士研究生. 主要研究方向为多源信息协同处理、运动视频匹配与稳像. E-mail: hermas@163.com  
(ZHANG Xin-Gong Master student at the College of Computer Science, Northwest Polytechnical University. His research interest covers multi-sensor information cooperative processing, and motion video registration and stabilization.)



邵大培 西北工业大学计算机学院硕士研究生. 主要研究方向为视频内容分析和图像识别. E-mail: sdp1599180@163.com  
(SHAO Da-Pei Master student at the College of Computer Science, Northwest Polytechnical University. His research interest covers video content analysis and image recognition.)