

文章编号 1004-924X(2018)06-1533-09

目标运动轨迹匹配式的红外-可见光视频自动配准

王洪庆¹, 许廷发^{1*}, 孙兴龙¹, 李相民¹, 刘太辉²

(1. 北京理工大学 光电学院, 北京 100081;
2. 北华大学 计算机科学技术学院, 吉林 吉林 132013)

摘要:为实现精确的红外-可见光视频序列的自动配准,提出了一种新的基于目标轨迹线匹配的配准方法。首先,利用运动目标检测技术提取目标前景,并由基于相关滤波器(KCF)的多目标跟踪算法对每个前景顶点进行跟踪,进而获取每个目标的运动轨迹。此后,为每条轨迹线建立归一化运动方向描述子与归一化运动幅度描述子,通过时序分析、方向描述子匹配及幅值描述子匹配建立分步约束的匹配机制,完成轨迹线匹配工作。最后,采用迭代更新的方式获取最佳全局配准矩阵,实现对异源视频的配准。在 LITIV 数据库上的 9 组视频上进行测试验证,实验的结果表明:本文配准算法的重叠率误差一般小于 0.2,接近或已超过手动的 Ground-Truth 矩阵。通过充分利用目标的运动信息,该算法实现了精确的红外-可见光图像序列配准。

关键词:红外图像;可见光图像;图像配准;多目标跟踪;特征匹配

中图分类号:TP394.1;TH691.9 **文献标识码:**A **doi:**10.3788/OPE.20182606.1533

Infrared-visible video registration with matching motion trajectories of targets

WANG Hong-qing¹, XU Ting-fa^{1*}, SUN Xing-long¹, LI Xiang-min¹, LIU Tai-hui²

(1. School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China;
2. College of Computer Science and Technology, Beihua University, Jilin 132013, China)
* Corresponding author, E-mail: xutingfa1123@163.com

Abstract: In order to realize accurate and automatic infrared-visible video registration, a novel registration method was proposed based on matching the motion trajectories of targets. First, the top pixel of each foreground was tracked by using the multi-target tracking algorithm based on KCF. In this way, the trajectory of each target was obtained. Then, the normalized motion orientation descriptor and the normalized motion magnitude descriptor were established for each trajectory. The stepwise constraint matching framework was structured by using time analysis, orientation descriptor matching and magnitude descriptor matching. Finally, the best registration matrix was obtained with iterative updating. The method proposed was validated with the nine pairs of videos in the LITIV database. The results indicate that the overlap error of the proposed method is smaller than 0.2, which is close to or better than the manual Ground-Truth matrix. By adequately using the motion information of tar-

收稿日期:2017-11-02;修订日期:2017-12-11.

基金项目:国家自然科学基金重大科学仪器专项资助项目(No. 61527802);国际合作对俄专项资助项目(No. 2014DFR10960)

get, the algorithm can realize precise infrared-visible image sequence registration.

Key words: infrared image; visible image; image registration; multi-target tracking; feature matching

1 引言

近年来,多传感器图像融合技术已经引起了人们的广泛关注,并在行人检测、视觉监控、目标跟踪和人脸识别等领域得到了广泛的应用^[1-3]。其中,红外与可见光图像因为能够为同一场景提供互补性的光谱信息,所以两者之间的融合已经成为信息融合领域的研究热点^[4]。图像配准旨在对齐在不同时间、不同位置或由不同传感器拍摄的两幅或者多幅图像,是图像融合的必要前提。精确的图像配准能够消除传感器空间位置差异的影响,从而有效地提高融合图像的质量。因此,本文致力于研究红外-可见光图像序列自动配准技术。

图像配准的关键在于在两幅或者多幅图像间寻找对应信息。然而,红外与可见光图像的成像原理不同,所以确定对应信息是一个十分困难的过程。红外图像记录了场景的热分布,而可见光记录了场景的反射光分布,两种图像存在明显的灰度差异及纹理差异。总体而言,红外可见光图像自动配准技术一直是一个具有挑战性的课题。

目前,已经存在配准算法主要可以分为两类:基于区域信息的配准方法及基于特征的配准方法。其中,基于区域信息的配准算法主要包含相关法,互信息法等。互相关法^[5]计算两个图像中每一个窗口对的互相关函数,并选取具备最大值的作为匹配窗口。互相关法的主要缺陷是实时性较差。此外,当待配准图像的纹理信息相对单一时,互相关法的可靠性差。互信息法^[6-7]通过计算图像的熵值来估计图像的信息关联程度,其在医学图像配准中得到了广泛的应用。然而,由于红外与可见光图像间的强度及纹理差异,基于互信息的配准算法难以实现高精度红外-可见光配准。

基于特征的图像配准方法包括点特征配准、边缘特征配准等。基于点特征的配准的核心在于在红外/可见光图像之中寻找可靠的特征点,角点^[8]以及 SIFT 特征点^[9]等已经被广泛地用于图像配准领域。然而,直接利用特征点配准红外-可见光的精度较低。边缘特征可以同时被红外可见

光传感器获取,所以许多的学者采用图像边缘或者边缘上的特征完成配准^[10-11]。但是边缘在红外与可见光图像之中并不完全一致,直接基于边缘特征进行配准仍难以保证配准的准确性。

由于静态特征难以实现高精度配准,采用运动特征完成配准得到了广泛的研究。该方法一般利用运动目标检测技术提取出目标前景,进而在前景上提取特征进行配准。其中, Bilodeau^[12-13]等人提出了基于目标运动轨迹的配准算法。此类算法利用目标跟踪技术确定场景中多个前景目标的运动轨迹,并采用随机一致性(RANSAC)算法匹配红外-可见光轨迹线。根据最佳匹配轨迹线对,继续利用随机一致性算法计算出配准矩阵。这种算法仅仅利用每个前景的顶点位置,不受前景轮廓检测误差的影响。

然而,基于目标运动轨迹的配准仍存在一些問題。首先,此类算法直接采用随机一致性算法匹配目标轨迹,匹配精度低,实时性差;其次,此类算法只利用一对匹配轨迹计算配准矩阵,严重降低了配准质量。针对这些问题,本文对传统算法进行改进,具体如下:

(1)通过建立归一化运动方向描述子与归一化幅值描述子,形成了逐步约束的轨迹线匹配机制,大幅度提高了轨迹线匹配的质量。

(2)同时采用多条轨迹线上的特征点对计算配准变换矩阵,并通过迭代更新的方式获取最佳配准矩阵,大幅度提高图像配准的精度。

实验结果表明,此算法能够利用异源图像中的运动特征信息,进而实现高精度的红外-可见光图像配准。

2 红外-可见光视频自动配准方法

2.1 原理框架

图 1 展示了红外-可见光视频自动配准算法的流程图,其主要包含多目标跟踪、轨迹线匹配和配准矩阵迭代更新 3 个部分。在多目标跟踪模块,首先对输入数据进行运动区域检测,并采用基于相关滤波器与前景信息相结合的多目标跟踪算法^[14]跟踪场景中的目标,获取红外-可见光视频

对所有目标的运动轨迹。在轨迹线匹配模块,首先为轨迹线建立归一化运动方向和运动幅值描述子,采用分步约束的方式匹配轨迹线。此后,基于轨迹点的前后时序关系,从匹配的轨迹线对上提取匹配点对。在矩阵迭代更新模块,我们利用随机一致性算法(RANSAC)算法计算当前配准变换矩阵,进而采用迭代更新策略获取适用于整个图像序列的最佳配准变换矩阵。

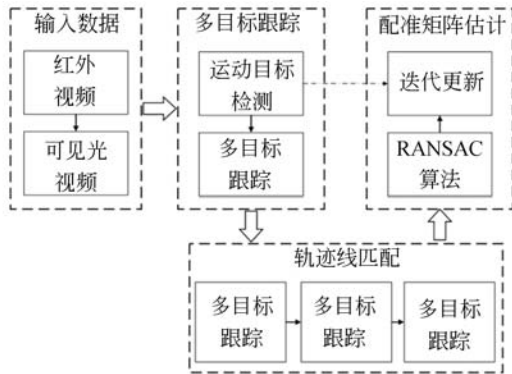


图 1 轨迹匹配式视频自动配准流程图

Fig. 1 Flow chart of proposed method

2.2 基于 KCF 的多目标跟踪

2.2.1 运动目标检测

本文基于多目标跟踪技术实现红外-可见光图像序列的自动配准,运动目标检测是本文方法的初始步骤。算法采用 St-Charles 等人提出的 PAWCS 算法^[14]提取运动前景。此算法利用颜色和二值化特征构建一个统计背景模型,并由一个负反馈框架实时更新背景模型,进而检测出运动目标。为了降低光线变换,阴影等噪声的影响,我们滤除面积过小的前景区。此外,为解决前景检测过程中的过度分割问题,合并两个距离较近(两者之间的距离小于 20 pixel)的前景块。通过这些处理,能够获取足够准确的前景检测结果来进行配准工作。图 2 给出了运动目标检测的结果。

2.2.2 多目标跟踪

在已存的算法中,L. M. Fuente 等人采用基于前景重叠率的多目标跟踪算法^[15]提取目标运动轨迹。然而,在环境变化、目标检测误差、遮挡等因素的影响下,简单的根据前景重叠率不能准确地获取目标在相邻帧间的关联性。因此,本文采用 Yang 等人提出的结合相关滤波器(KCF)与前景区域信息的多跟踪算法^[16]。在该算法中,前



(a) 红外前景

(b) 可见光前景

(a) Infrared foreground (b) Visible foreground

图 2 运动目标检测的结果

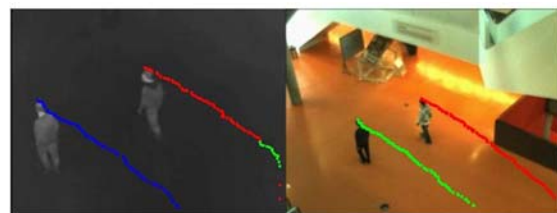
Fig. 2 Results of motion target detection

景区域被用于确定运动目标的尺寸,而相关滤波器被用于执行跟踪决策。对于第 t 帧图像,首先提取一系列的前景区域 FOR^t 。为了完成多目标跟踪并更新 KCF 滤波器,计算第 $t-1$ 帧所有相关滤波器在当前帧的输出。此后,通过计算前景区域与相关滤波器输出间的重叠率,确定滤波器和前景区域的关联性,实现从 $t-1$ 到 t 帧的目标跟踪。

通过采用多目标跟踪技术,目标在不同帧内的顶点位置 $[x_i, y_i, 1]^T$ 组成一条轨迹线,如下所示:

$$\mathbf{T}_i^t = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \quad (1)$$

其中: $L \in (\text{IR}, \text{visible})$, \mathbf{T}_i^t 表示第 i 条红外或可见光运动轨迹线, n 是轨迹线包含轨迹点的数目。图 3 给出了提取到的红外与可见光目标运动轨迹。



(a) 红外轨迹

(b) 可见光轨迹

(a) Infrared trajectories (b) Visible trajectories

图 3 目标运动轨迹

Fig. 3 Motion trajectories of targets

2.3 轨迹线匹配

2.3.1 轨迹线描述

Torabi^[12-13]等文献直接采用随机一致性算法

匹配轨迹,并通过多次计算重叠率误差确定一对最佳匹配轨迹。这种方式的计算量较大,同时对轨迹线的利用效率也较低。本文提出一种全新的轨迹线匹配策略,首先为每条轨迹线确定两个描述子:归一化方向描述子 \tilde{C}^l 及归一化幅值描述子 \tilde{r}^l ,其分别如公式(2)和公式(3)所示:

$$\tilde{C}^l = [\tilde{C}_1^l \quad \tilde{C}_2^l \quad \cdots \quad \tilde{C}_n^l], \quad (2)$$

$$\tilde{r}^l = [r_1^l \quad r_2^l \quad \cdots \quad r_n^l], \quad (3)$$

其中: $\tilde{C}_i^l (i \in (1, 2, \dots, n))$ 表示经过方向归一化后目标在某一运动时刻*i*的运动方向, $r_i^l (i \in (1, 2, \dots, n))$ 表示经过幅值归一化后目标在该运动时刻的运动幅值。在这里,方向归一化与幅值归一化的目的是使两个描述子都同时具有尺度不变性和旋转不变性。下面将具体介绍这两个描述子的确定方法。

对于一个目标,其在视频序列中形成的轨迹如公式(1)所示。通过确定目标在前后两帧的位置差,可以得到目标的运动矢量序列,如下所示

$$\Delta \mathbf{T}_i^l = \begin{bmatrix} \Delta x_1 & \Delta x_2 & \cdots & \Delta x_{n-1} \\ \Delta y_1 & \Delta y_2 & \cdots & \Delta y_{n-1} \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad (4)$$

其中 $[\Delta x_i, \Delta y_i, 0]^T$ 是目标在某时刻的运动矢量。在获取运动矢量序列之后,需要进一步确定目标的在相邻帧间的运动方向与运动幅值。在这里,目标的运动方向和运动幅值可分别由公式 $\theta = \arctan(\Delta y / \Delta x) \theta \in (0, 2\pi)$ 及 $r = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ 确定,从而获取运动方向序列以及运动幅值序列,如下所示

$$\begin{aligned} \theta_i^l &= [\theta_1^l \quad \theta_2^l \quad \cdots \quad \theta_{n-1}^l], \\ r_i^l &= [r_1^l \quad r_2^l \quad \cdots \quad r_{n-1}^l]. \end{aligned} \quad (5)$$

运动方向信息具有尺度不变性但不具备旋转不变性,运动幅值序列具备旋转不变性但不具有尺度不变性。为了建立同时具备尺度不变性及方向不变性的描述子,需要对两个描述子进行归一化处理。对于运动方向,首先进行量化编码,其方法如下所示:

$$C(\theta) = \text{floor}(\theta / 0.25 \times \pi) + 1, \quad (6)$$

其中 $C(\theta)$ 运动角度 θ 的方向编码。值得注意的是,如果目标的运动幅值为0,那么其方向编码值也为0。经过方向编码后,得到编码方向序列:

$$C_i^l = [C_1^l \quad C_2^l \quad \cdots \quad C_{n-1}^l], \quad (7)$$

其中 C_i^l 表示某一时刻的运动方向编码,其方向值在(0~8)之间。在编码后,只需要寻找到目标运动的主方向,即可完成方向归一化。对方向编码进行直方图统计,其峰值(编码为0的除外)对应的方向编码即为运动主方向的编码。设主方向编码为 C^{\max} ,在方向归一化时,主方向被定义为编码1,其他各方向的编码值依次变化,具体方式如下:

$$\tilde{C} = \begin{cases} 8, & \text{MOD}(C - C^{\max} + 1, 8) = 0 \\ \text{MOD}(C - C^{\max} + 1, 8), & \text{otherwise} \end{cases}. \quad (8)$$

对于初始编码为0的方向,在方向归一化之后,方向编码仍为0。经过归一化之后,就获得了同时具备旋转不变性以及尺度不变性的方向描述子。

对运动幅值描述子进行归一化处理,此时采用运动幅值的总和作为归一化权重,归一化运动幅值描述子可以为:

$$\tilde{r}_i^l = \frac{1}{\sum_{i=1}^N r_i^l} [r_1^l \quad r_2^l \quad \cdots \quad r_{n-1}^l], \quad (9)$$

经过幅值归一化之后,即可得到同时具备旋转不变性及尺度不变性的幅值描述子。

2.3.2 轨迹线匹配

本节提出一种快速精确轨迹匹配策略,其旨在于为任意一条红外(可见光)轨迹 T_i^R 找到对应的可见光(红外)运动轨迹 T_i^{VIS} 。然而,并不是所有的轨迹都可用于提取配准特征,若轨迹线的长度过短(小于10帧),说明在跟踪目标时发生了跟踪丢失,或者目标快速进入被观测场景又迅速离开。此类轨迹对确定配准矩阵不存在任何意义,应当予以滤除。此后,本文采用分步约束的方式匹配红外-可见光轨迹,具体如下:

由于多个轨迹线存在时间段并不一致,首先需要对其轨迹线进行时序分析。对于一条红外轨迹 T_i^R ,设其起始和终止时刻分别为 t_1 和 t_2 。对于任意一条可见光轨迹,如果其在 $t_1 \sim t_2$ 时间段内存在的时间小于10帧,则说明其不是候选匹配轨迹。得到候选可见光轨迹集为 $\mathbf{T}^{\text{VIS}} = [T_1^{\text{VIS}} \quad T_2^{\text{VIS}} \quad \cdots \quad T_{n-1}^{\text{VIS}}]$ 。

在经过时序分析后,继续利用归一化运动方向及归一化运动幅值描述子匹配红外-可见光轨迹。对于运动方向描述子而言,量化编码使其具

有了更强的抗干扰能力,采用它进行轨迹匹配可以有效地剔除运动方向一致性较差的候选轨迹。但是,编码也导致方向描述子不能辨别两个相近的运动方向,难以可靠地确定最佳匹配轨迹。对于未经过幅值编码的运动幅值描述子,其抗干扰能力较弱,直接利用它进行匹配的计算量也较大。但是其能够区分两个相近的运动幅值,可以从运动规律相似的轨迹集中确定最佳匹配轨迹。考虑这些因素,利用归一化运动方向描述子进行轨迹粗匹配,剔除运动方向匹配程度较差的候选轨迹。并由归一化运动幅值描述子完成最后的匹配工作,确定最佳的可见光匹配轨迹。对于红外轨迹 T_i^{LR} 的每一条候选可见光轨迹 T_k^{VIS} ,假设共存时间段为 $t_s \sim t_k$ 。在此时间段内,确定两条轨迹的运动方向描述子,并作归一化直方图统计,分别得到红外轨迹与可见光轨迹的运动方向直方图,即 $\mathbf{H}_{IR} = [h_0^{LR}, h_1^{LR}, \dots, h_g^{LR}]$, $\mathbf{H}_{VIS} = [h_0^{VIS}, h_1^{VIS}, \dots, h_g^{VIS}]$ 。然后,计算两个直方图的欧式距离。两者的欧式距离 D_H 越小,说明两个轨迹线的匹配度越高。设定距离阈值 D_T (一般情况下, $D_T = 0.5$)。如果 $D_H \leq D_T$,则仍视 T_k^{VIS} 为 T_i^{LR} 的一条候选匹配轨迹。否则,则舍弃该可见光轨迹。经过处理后,得到候选的可见光轨迹集为 $\hat{\mathbf{T}}^{VIS} = [\hat{T}_1^{VIS}, \hat{T}_2^{VIS}, \dots, \hat{T}_{n-1}^{VIS}]$ 。

最后,对于每条候选可见光轨迹 \hat{T}_k^{VIS} ,在公共运动时间 $t_s \sim t_k$ 内,确定其归一化运动幅值描述子 r_k^{VIS} 与当前红外轨迹 T_i^{LR} 运动幅值描述子 r_i^{LR} 间的幅值距离,计算方式如下:

$$D_A = \frac{1}{t_k - t_s} \text{Sqrt} \left(\sum_{i=1}^{M=t_k-t_s} (r_i^{LR} - r_i^{VIS})^2 \right), \quad (10)$$

其中 Sqrt 代表着平方根函数。在计算每个候选可见光轨迹与当前红外轨迹的幅值距离后,取幅值距离最小的可见光轨迹作为匹配的可见光轨迹 T_i^{VIS} 。此时,可以得到红外-可见光运动轨迹匹配集 $\{(T_{m1}^{LR}, T_{m1}^{VIS}), (T_{m2}^{LR}, T_{m2}^{VIS}), \dots, (T_M^{LR}, T_M^{VIS})\}$ 。

获取轨迹匹配集后,可以从每个匹配轨迹对上提取数配准需要的特征点对。对于每一过匹配轨迹对 (T_i^{LR}, T_i^{VIS}) ,在同一时刻的红外轨迹点及可见光轨迹点即可以构成一对红外-可见光匹配点对。将所有轨迹线上提取到的匹配点对汇聚在

一起,形成匹配特征集,其将被用于计算全局配准变换矩阵。

2.4 最佳全局配准矩阵估计

根据红外和可见光视频中的对应匹配特征点,通过采用随机一致性算法(RANSAC)求解当前帧的全局配准矩阵。然而,为了获取对于此视频最优的变换矩阵,并克服目标在三维空间内运动的影响(在一帧中多目标近似在一个平面上),需要同时利用不同帧的配准矩阵。因此,本文采用迭代更新的方式确定最佳变换矩阵。假设当前帧的配准矩阵为 \mathbf{T}_{cur} ,利用此矩阵对红外(可见光)目标进行插值变换,并求取变换后红外-可见光前景的重叠率误差 E_{cur} ;假设此前的最佳变换矩阵为 \mathbf{T}_{t-1}^{be} ,求取变换后的红外-可见光前景的重叠率误差 E_{t-1}^{be} ,迭代更新的方式为

$$\mathbf{T}_t^{be} = (1 - \beta) * \mathbf{T}_{t-1}^{be} + \beta * \mathbf{T}_{cur}, \text{ if } E_{cur} < E_{t-1}^{be}, \quad (11)$$

其中, β 是矩阵更新因子,一般情况下, $\beta = 0.2$ 。通过采用矩阵迭代更新的方式,可以为视频对确定一个帧宽型配准矩阵,其不但可以配准当前帧内的目标,更可以配准整个视频序列中的目标。随着对视频的每一帧不断的进行配准,配准的精度也逐渐的提升。整体的配准流程如算法 1 所示。

算法 1 配准流程

输入:红外视频图像序列:

$\{IR_i\}, i=1:\text{length}(IR_i)$

可见光视频图像序列:

$\{VIS_i\}, i=1:\text{length}(IR_i)$

(1)对于每一对红外-可见光图像 $IR_i, VIS_i (i=1:\text{length}(IR_i))$,循环以下步骤。

(2)提取红外与可见光图像的运动前景 F_{LR}, F_{VIS} 。

(3)确定红外及可见光目标的运动轨迹集合 $\{T^{LR}\}, \{T^{VIS}\}$ 。

(4)通过式(4)-(9)为每个运动轨迹计算归一化运动方向描述子 \hat{C}_i^L 和归一化运动幅值描述子 r_i^L 。

(5)通过分析轨迹长度、共存时间以及分别计算归一化运动方向描述子和归一化运动幅值描述子的匹配程度,确定红外-可见光运动轨迹匹配集合 $\{(T_{m1}^{LR}, T_{m1}^{VIS}), (T_{m2}^{LR}, T_{m2}^{VIS}), \dots, (T_M^{LR}, T_M^{VIS})\}$ 。

(6)利用随机一致性算法和公式(11)确定当前帧的最佳配准矩阵 \mathbf{T}_i^{be} 。

输出:

(7) 针对整个红外-可见光视频图像序列, 确定每一帧图像对的最佳配准矩阵 $\{T_i^*\}$

3 实验与结果

3.1 实验设置

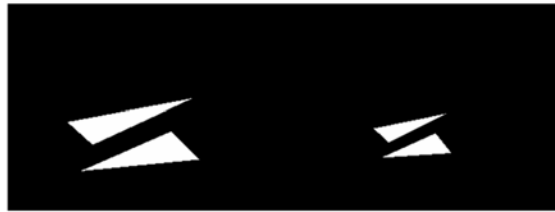
为了验证本文提出算法的有效性, 采用公开通用 LITIV 数据库^[13]作为我们的实验数据。此数据库包含 9 组红外-可见光视频序列, 它们的分辨率是 240×320 , 而帧数在 200 与 1 200 之间。在所有的红外和可见光图像序列中, 拍摄到的运动目标为行人。由于目标在较远的距离被拍摄, 所以场景中所有目标近似位于同一深度平面内, 满足平面假设条件。此外, 该数据库为每个视频对提供 ground-truth 仿射变换矩阵, 其是通过手动挑选匹配点对确定的。在对实验数据进行分析时, 通过与此矩阵作比较, 可以有效的突出本文算法的精确性。

为了衡量算法的精度, 估计算法在实验数据库上的配准误差。此时需要为每一对测试视频确定一对二值化图像, 方法如文献^[13]所述: 对于红外可见光图像序列, 首先在红外图像中手动选取一些易识别的点, 并在可见光图像中寻找它们的对应点。然后, 连接这些点建立多边形对和手动二值化图像。对于实验采用的 LITIV 数据库, Charles^[17]等人已经为其中的每个视频对都构建

了一对二值化图像, 第一个视频对(LITIV1)的二值化图像如图 4 所示。本文就采用 Charles^[17]等人构建的二值化图像计算配准误差。在获得二值化图像后, 根据算法获取的配准矩阵对二值化图像进行插值变换, 并将变换后图像的重叠率误差作为配准精度误差。重叠率的计算方式为:

$$E = 1 - \frac{S_i \cap S_v}{S_i \cup S_v}, \quad (12)$$

其中: S_i 与 S_v 是变换后的红外二值化图像, S_v 是可见光二值化图像。



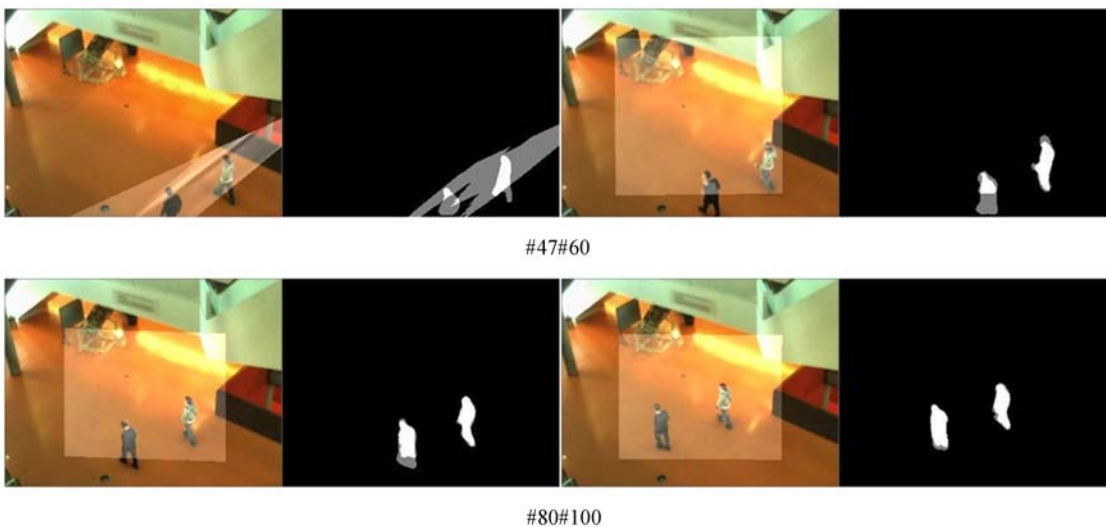
(a) 红外前景 (b) 可见光前景
(a) Infrared binary image (b) Binary image

图 4 LITIV1 序列的二值化图像对

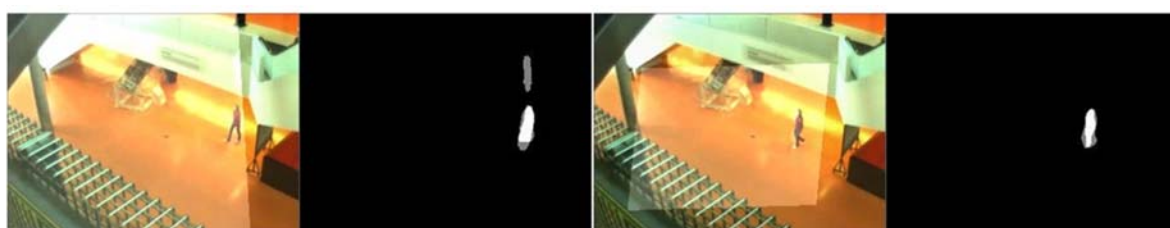
Fig. 4 Binary image pair for LITIV1 sequence

3.2 配准结果

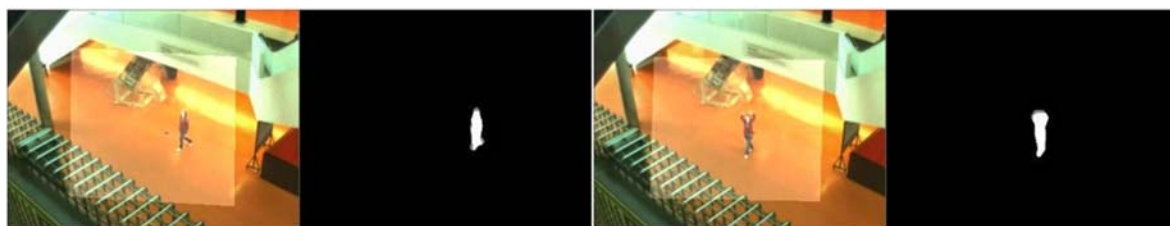
图 5 给出了本文方法对 LITIV 数据库内前三个图像序列的配准情况。在图中, 针对每一对测试图像给出两幅配准结果图像, 左侧图像通过叠加配准后的红外图像和可见光图像给出了算法对源图像的配准结果, 右侧图像给出了算法对前景图像的配准结果。在前景图像的配准结果中,



(a) LITIV1 图像序列配准结果
(a) Registration results on LITIV1



#53#70

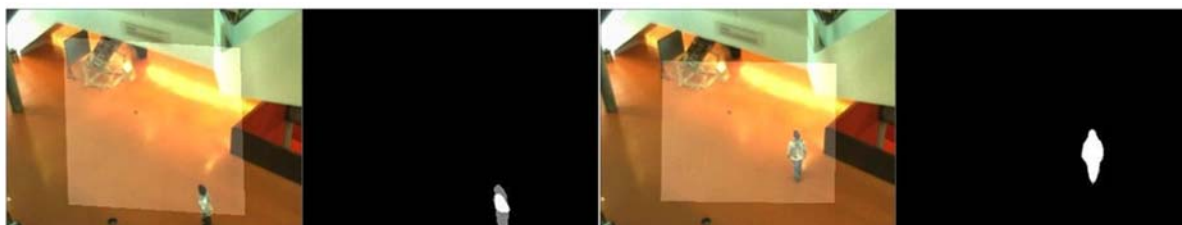


#90#110

(b) LITIV2 图像序列配准结果
(b) Registration results on LITIV2



#182#250



#370#410

(c) LITIV3 图像序列配准结果
(c) Registration results on LITIV3

图 5 LITIV 数据库上的配准结果(前三个视频序列)

Fig. 5 Registration results on LITIV dataset (for the first three sequences)

白色部分表示经本文算法配准后红外与可见光前景的重叠区域,灰度部分表示红外与可见光前景的不重叠区域,黑色部分表示背景区域。红外与可见光前景在配准后的重叠程度直观的体现了算法的配准能力。针对每个图像序列,给出的第一对图像展示了目标刚进入观测场景时算法的配准表现。可以看出,尽管在目标刚进入场景时,本文

提出的算法没有立刻给出精确的配准结果,但随着目标的运动,本文算法能够快速地计算出精确的全局配准矩阵,进而实现高精度的红外-可见光图像序列配准。对于前两个图像序列,在目标出现后的 30 帧内,即可得到可接受的配准结果。对于 LITIV3,算法消耗更多的时间来精确的配准红外可见光图像,这是因为目标在观测场景中的

运动区域较小,所获取的运动轨迹不具备代表性。

表 1 最小重叠率误差
Tab. 1 Minimum overlap error

Sequence pair	Proposed	Ground-truth
LITIV-1	0.165	0.149
LITIV-2	0.100	0.078
LITIV-3	0.113	0.080
LITIV-4	0.127	0.221
LITIV-5	0.103	0.150
LITIV-6	0.097	0.088
LITIV-7	0.055	0.136
LITIV-8	0.197	0.260
LITIV-9	0.097	0.134

表 1 给出了本文方法和手动发现的 Ground-truth 矩阵针对于每个序列的最小重复率误差。对比本文算法与 Ground-truth 矩阵可以得知,本文算法与 Ground-truth 矩阵的配准精度十分接近,并在某些图像序列上超过了 Ground-truth 矩阵。主要原因有以下两点:首先,Ground-truth 矩阵是通过人工在红外与可见光图像中挑选匹配点对构建形成的。在这个过程中,人为的误差是不可避免的,这也是 Ground-truth 矩阵配准误差不可

为 0 的原因。此外,Ground-truth 矩阵只能完美的配准完全平面场景。尽管已经将观测场景近似为平面场景,但是目标运动过程中的深度变化是实际存在的。此时,本文的算法通过迭代更新的方式致力于配准当前前景,能够比 Ground-truth 矩阵达到更高的配准精度。

4 结 论

为实现精确的红外-可见光图像序列配准,本文提出了一种基于目标运动轨迹匹配的自动配准算法。与传统的轨迹线匹配不同,本文为目标轨迹建立了归一化运动方向描述子以及归一化运动幅值描述子,并采用分步约束的匹配方式提高了轨迹线的匹配精度。根据匹配轨迹线得到的特征点对,采用随机一致性(RANSAC)算法和迭代更新的方式得到了最佳全局配准矩阵。实验结果显示,针对每个视频对,本文算法的最小重叠率误差均小于 0.2,并在某些情况下超过手动发现的 Ground-truth 矩阵。其可以实现精确的红外可见光视频自动配准。

参考文献:

- [1] HERMOSILLA G, GALLARDO F, FARIAS G, *et al.*. Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems [J]. *Sensors*, 2015, 15: 17944-17962.
- [2] KROTOSKY S J, TRIVEDI M M. Person surveillance using visual and infrared imagery [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, 18(8): 1096-1105.
- [3] MA J, ZHAO J, MA Y, *et al.*. Non-rigid visible and infrared face registration via regularized Gaussian fields criterion [J]. *Pattern Recognition*, 2015, 48(3): 772-784.
- [4] 陈继红, 刘素真. 可见光与红外 CPCT 彩色图像融合研究与实现[J]. *现代电子技术*, 2017, 40(9): 5-9.
- [5] CHEN J H, LIU S Z. Research and implementation of visible-light and infrared CPCT color image fusion [J]. *Modern Electronics Technique*, 2017, 40(9): 5-9. (in Chinese)
- [6] ROCHE A, MALANDAIN G, PENNEC X, *et al.*. The correlation ratio as a new similarity measure for multimodal image registration [C]. *Proceedings of the Springer International Conference on Medical Image computing and Computer-Assisted Intervention, Cambridge, MA, USA: MIC-CAI'98*, 1998: 1115-1124.
- [7] VIOLA P, WELLS W M. Alignment by maximization of mutual information [J]. *International Journal of Computer Vision*, 1997, 24(2): 137-154.
- [8] LEGG P A, RDSIN P L, MARSHALL D, *et al.*. Feature neighbourhood mutual information for multi-modal image registration: an application to eye fundus imaging [J]. *Pattern Recognition*, 2015, 48(6): 1937-1946.
- [9] 车芳, 韩俊刚, 郭志全. SMT-PAAG 下的 Harris 角点检测与匹配算法[J]. *电子技术应用*, 2017, 43(4): 138-140.
- [10] CHE F, HAN J G, GUO ZH Q. Harris corner detection and matching algorithm at the SMT-PAAG [J]. *Application of Electronic Technique*, 2017, 43

- (4):138-140. (in Chinese)
- [9] 张宏伟,樊祥,朱斌,等. 引入外点剔除机制的双波段红外图像的配准[J]. 红外与激光工程, 2015, 44(s1):23-28.
- ZHANG HW, FAN X, ZHU B, *et al.*. Dual-band infrared image registration with the introduction of outliers rejection mechanism [J]. *Infrared and Laser Engineering*, 2015, 44(s1):23-28. (in Chinese)
- [10] KONG S G, HEO J, BOUGHORBEL F, *et al.*. Multiscale fusion of visible and thermal IR Images for illumination-invariant face recognition[J]. *International Journal of Computer Vision*, 2007, 71(2):215-233.
- [11] COIRAS E, SANTAMARIA J, MIRAVET C. Segment-based registration technique for visual-infrared images[J]. *Optical Engineering*, 2000, 39(1):202-7.
- [12] BILODEAU G A, TORABI A, MORIN F. Visible and infrared image registration using trajectories and composite foreground images[J]. *Image & Vision Computing*, 2011, 29(1):41-50.
- [13] TORABI A, MASSE G, BILODEAU G A. Feedback scheme for thermal-visible video registration, sensor fusion, and people tracking [J]. *Journal of Biomedical Materials Research Part A*, 2010, 2(1):15-22.
- [14] ST-CHARLES P L, BILODEAU G A, BERGEVIN R. A self-adjusting approach to change detection based on background word consensus[C]. *IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA: IW-CACV*, 2015:990-997.
- [15] FUENTES L M, VELATIN S A. People tracking in surveillance applications [J]. *Image&Vision Computing*, 2006, 24(11):1165-1171.
- [16] YANG Y, BILODEAU G A. Multiple object tracking with kernelized correlation filters in urban mixed traffic[EB/OL]. (2017-4-24). <https://arxiv.org/abs/1611.02364>. (accessed on 24 March 2017).
- [17] ST-CHARLES P L, BILODEAU G A, BERGEVIN R. Online multimodal video registration based on shape matching[C]. *IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, USA: CVPRW*, 2015:26-34.

作者简介:



王洪庆(1987—),男,吉林长春人,博士研究生,2012年于北京理工大学获得硕士学位,主要从事目标跟踪和图像处理等方面的研究。E-mail: wang_hongqing@126.com

导师简介:



许廷发(1968—),男,黑龙江肇东人,教授,博士生导师,2000年在东北师范大学物理系获得硕士学位,2004年在中国科学院长春光学精密机械与物理研究所获得博士学位,2006年华南理工大学电子与信息学院完成博士后工作。主要从事红外/可见目标跟踪与识别技术、电子稳像技术、光电图像实时处理与DSP/FPGA系统设计技术、高光谱遥感影像处理技术等方面的研究。E-mail: xutingfa1123@163.com