

# 搜救航拍序列图像关键帧提取研究

邹志远, 安博文, 潘胜达

(上海海事大学信息工程学院, 上海 201306)

**摘要:** 为提高搜救航拍中生成大视场红外图像的效率, 提出一种序列图像提取关键帧的方法。分析图像获取系统的运动参数和获取红外图像的内方位元素, 采用  $(\varphi, \omega, \kappa)$  转角系统作为图像的外方位元素, 根据摄影测量中的共线方程式计算序列图像间的重叠率, 并以相位相关法对计算出的重叠率进行验证。最后根据计算出的重叠率和采用的 Harris 拼接算法所要求的最佳重叠率, 来确定序列图像中的关键帧。试验结果表明, 计算出的重叠率可以作为关键帧提取的标准, 待拼接的图像数量减少, 生成大视场图像的效率提高 88.71%。

**关键词:** 红外序列图像; 关键帧; 共线方程; 相位相关; Harris

中图分类号: TP911.73 文献标识码: A 文章编号: 1001-8891(2014)09-0723-05

## Key Frame Extraction from Research and Rescue Aerial Image Sequence

ZOU Zhi-yuan, AN Bo-wen, PAN Sheng-da

(College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

**Abstract:** To improve the efficiency of wide field infrared image generation in research and rescue aerial photography, a method for key frame extraction from image sequence is proposed in this paper. Firstly, the motion parameters of image acquiring system and image's elements of exterior and interior orientation are considered. Secondly, the  $(\varphi, \omega, \kappa)$  angle system is adopted and the overlap rate between images is calculated according to the collinear equation. Finally, the overlap rate calculated and the best overlap rate offered by Harris algorithm are consider to determine the key frames. The result shows the overlap rate which is calculated can be indicator for key frame extraction, and the efficiency of wide field image is promoted to 88.71%.

**Key words:** infrared image, key frames, collinear equation, phase correlation, Harris

## 0 引言

在机载红外高分辨率搜救成像系统的开发中, 要求对摆正程期间获取的红外序列图像进行拼接, 组合成一幅大视场图像。序列图像帧间的相似性很大, 生成大视场图像过程中, 一般选取序列图像的子集进行拼接。子集图像之间应具有一定重叠率, 这样既可以保证拼接算法的实现, 又去除了冗余图像, 提高生成大视场图像的效率。把具有一定重叠率的图像称为航拍序列图像的关键帧。目前对序列图像的关键帧抽取主要有 4 种: ①镜头关键帧提取方法<sup>[1]</sup>: 最早由 Nagasak 等人提出, 该方法在固定位置提取关键帧, 如每个镜头取第 1 帧, 中间帧, 尾帧。该方法最大优点是速度快但伴随而来的缺点就是取得的关键帧代

表性不强, 且每镜头提取的关键帧数量相同。②基于图像内容提取关键帧<sup>[2]</sup>: 该方法基于信息论。通常通过图像的颜色, 纹理, 形状等特征来描述图像变化, 其计算量大, 且存在完全不同的图像具有相同特征的可能。③基于运动分析提取关键帧<sup>[3]</sup>: 由 Wolf 等人提出, 典型算法是光流法。该技术比前种方法的计算量还大, 这对实时性要求高的系统是致命的缺陷。④基于聚类提取关键帧<sup>[4]</sup>: 该方法把视频中每一帧看作聚类算法中的一个点, 提取图像特征, 然后用聚类算法计算聚类中心作为关键帧。该方法破坏了源图像序列的时间顺序, 计算量大, 且每个镜头的关键帧数目不易确定。

针对上述关键帧提取算法的优缺点, 结合机载搜救项目, 优先考虑算法的速度, 提出一种基于摄影测

收稿日期: 2014-03-03; 修订日期: 2014-06-18。

作者简介: 邹志远 (1989-), 男, 北京市延庆县人, 硕士, 主要研究领域为红外图像获取与处理。E-mail: canju521968@sina.com。

基金项目: 国家自然科学基金, 编号: 61171126; 上海自然科学基金, 编号: 11ZR1415200; 上海重点支撑项目, 编号: 12250501500。

量的关键帧提取方法。

## 1 成像模型及分析

### 1.1 模拟搜救系统的运动方式

本文的模拟搜救系统的运动方式为摆扫画幅成像方式<sup>[5]</sup>, 如图 1, 当成像平面平行摄影区域平面时, 定义此时相机摆扫为 0°, 则相机向左侧摆扫 30°、右侧摆扫 30°, 共 60°, 一个摆扫正程视为一个镜头。其中⊗表示相机前进方向。

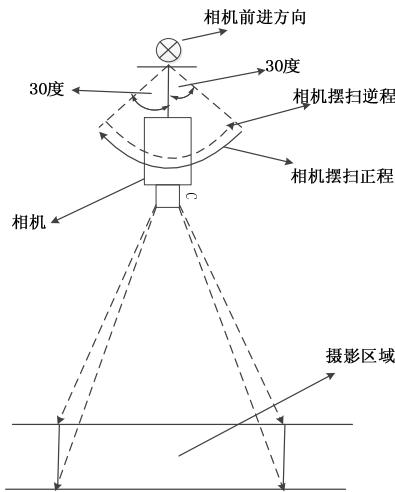


图 1 搜救模拟系统的运动方式

Fig.1 Movement of search-and-rescue simulating system

### 1.2 瞬时成像模型

对系统摆扫瞬时状态建模, 如图 2 所示。

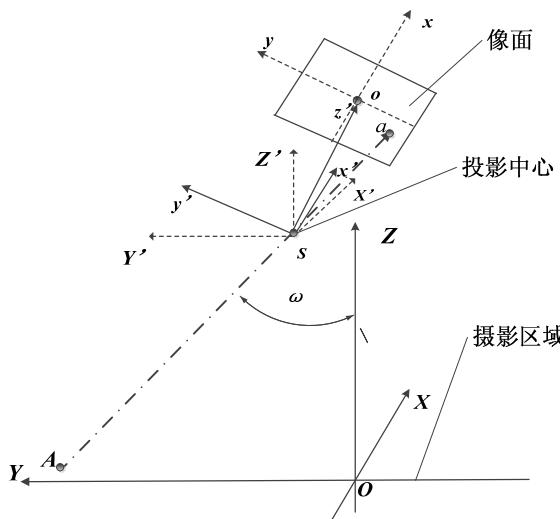


图 2 瞬时成像模型

Fig.2 Instantaneous imaging model

成像模型中定义 4 个坐标系。

#### 1) 像面坐标系: o-xy

描述物点在红外焦平面上位置的直角坐标系, 原点  $o$  为焦平面中心,  $x$  为转轴, 与相机前进方向同,  $y$  与  $x$  垂直。

#### 2) 像面过渡坐标系 $S-x'y'z'$ ( $z'=f$ 相机焦距)

用来推导物点和像点空间坐标关系建立的过渡坐标系。 $x'$ 、 $y'$  分别平行于  $x$ 、 $y$ 。 $z'$  指向焦平面中心。

#### 3) 摄影区域过渡坐标系 $S-X'YZ'$

用来推导物点和像点空间坐标关系建立的过渡坐标系, 原点  $S$  (投影中心) 随相机摆动而不同。其各坐标轴方向始终不变, 与像面平行摄影区域时的像面过渡坐标系的坐标轴方向相同且重合。

#### 4) 摄影区域坐标系: $O-XYZ$

用来描述物点在摄影区域位置的直角坐标系,  $O$  点为投影中心  $S$  在物面垂直对应点, 各坐标轴方向与摄影区域过渡坐标系相同。

为了推导像点  $a$  和摄影区域点  $A$  之间的对应关系, 需要推导坐标系 2 到坐标系 3 之间的转换关系, 本文选取  $(\varphi, \omega, \kappa)$  转角系统<sup>[6]</sup> 来表示图像的 3 个外方位角元素。其中  $\varphi$ 、 $\omega$ 、 $\kappa$  依次表示坐标系 2 绕其  $y'$ 、 $x'$ 、 $z'$  轴旋转所成角并令其对应的旋转变换矩阵依次为  $M_1$ 、 $M_2$ 、 $M_3$ 。则有坐标系 2 到坐标系 3 转换公式为:

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = M_1 M_2 M_3 \begin{pmatrix} x' \\ y' \\ f \end{pmatrix} \quad (1)$$

本文的模拟搜救系统只需考虑旋转变换矩阵  $M_2$ , 根据图 1 知, 在左侧 30° 和右侧 30° 成像过程具有对称性, 故只需考虑一侧成像方式。则式(1)变为:

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{pmatrix} \begin{pmatrix} x' \\ y' \\ f \end{pmatrix} \quad (2)$$

令  $A$  在摄影区域坐标系中  $O-XYZ$  中的坐标为  $A(X, Y, 0)$ ;  $a$  在像面坐标系  $o-xy$  中的坐标为  $a(x, y)$ , 则  $a$  在像面过渡坐标系中  $S-x'y'z'$  中的坐标为  $a(x, y, f)$ ;  $a$  在摄影区域过渡坐标系中  $S-X'YZ'$  中的坐标为  $a(X', Y, Z')$ ;  $S$  在摄影区域坐标系中  $S-XYZ$  中的坐标为  $S(0, 0, H)$ 。

$A, S, a$  三点共线方程<sup>[6]</sup> 为:

$$\frac{X-0}{X'} = \frac{Y-0}{Y'} = \frac{0-H}{Z'} \quad (3)$$

将式(2)带入式(3)中求出摄影区域坐标系任意物点  $A(X, Y)$  与像面坐标系像点  $a(x, y)$  之间关系:

$$\begin{cases} X = H \times \frac{x}{y \sin \omega - f \cos \omega} \\ Y = H \times \frac{y \cos \omega + f \sin \omega}{y \sin \omega - f \cos \omega} \end{cases} \quad (4)$$

式中:  $H$  为物点距离, 可看做投影中心  $S$  的距离;  $f$

为相机焦距。

### 1.3 简化成像模型

项目中采用的相机参数为  $x \in [-6, 6]\text{mm}$ 、 $y \in [-8, 8]\text{mm}$ 、 $f=100\text{ mm}$ 、帧率  $F=25\text{ 帧/s}$ , 摆扫正程时间为 3.5 s。式(4)中:

$$\frac{y \sin \omega - f \cos \omega}{-f \cos \omega} \geq 95.38\% \quad (5)$$

式(4)简化为式(6):

$$\begin{cases} X = H \times \frac{x}{-f \cos \omega} \\ Y = -H \times \frac{y}{f} - H \times \tan \omega \end{cases} \quad (6)$$

每幅图像对应物面区域为  $2X \times 2Y$  大小, 且  $X$ 、 $Y$  值随摆扫角度  $\omega$  的不同而不同。下面计算相邻两张图像内容所表示的物面区域大小差异。上述系统参数中提到摆扫视场角共为  $60^\circ$  历时 3.5 s, 每秒获取 25 张图像, 则计算出相邻两张图像之间的成像角差为  $0.6857^\circ$ 。则相邻图像之间  $X$  差异取决于式(7)中  $a$  的值:

$$\begin{cases} a = \frac{\cos(\omega_1 + \Delta\omega)}{\cos \omega_1} \\ \Delta\omega = 0.6857^\circ \end{cases} \quad (7)$$

式中:  $\omega_1$  为前一幅图像成像角度,  $\omega_1 + \Delta\omega$  为相邻图像的成像角度。图 3 为函数图像。

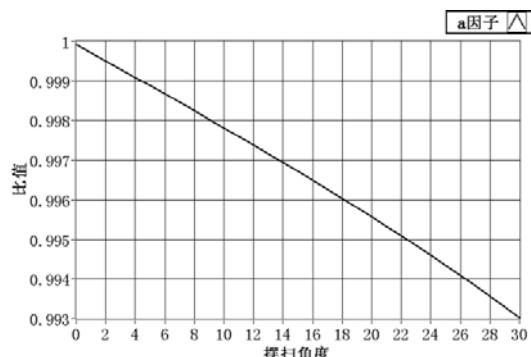


图 3  $a$  函数曲线图

Fig.3 Curve graph of  $a$

可见  $a$  值在摆扫范围内可认为是 1, 即认为相邻图像在  $X$  方向上所表示的区域范围无差别。

此外本文的搜救模拟系统对相机前进方向不做考虑, 进而相邻图像间的成像近似为相机焦平面绕投影中心  $S$  的旋转。此时相邻两张图像仅仅认为发生  $Y$  向平移的变化如图 4、图 5。

图 4、5 中  $PQ=2Y_1$ , 为前图像在摄影区域  $Y$  方向表示的跨度,  $P'Q'=2Y_2$  为后图像在摄影区域  $Y$  方向表示的跨度。则由图 4、图 5 可知相邻图像的重叠率:

$$\rho_0 = \frac{PQ - H \tan \Delta\omega}{PQ} = 1 - f \times \frac{\tan \Delta\omega}{2y} \approx 92.51\% \quad (8)$$

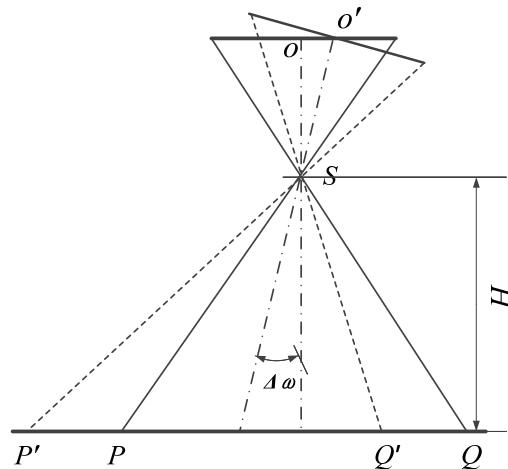


图 4 相邻图像成像示意图

Fig.4 Model of adjacent image

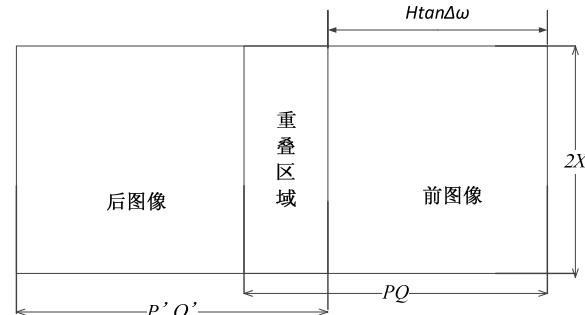


图 5 相邻图像重叠示意图

Fig.5 The overlap rate of adjacent image

### 1.4 关键帧的确定

文献 7 指出 Harris 拼接算法要求重叠区域 30% 以上, 考虑到海面图像的内容、纹理、特征少, 本文要求最佳重叠率  $\rho_{best}=60\%$ , 将  $\rho_{best}$  和  $\rho_0$  带入式(9), 得出序列图像关键帧的间隔数  $N=6$ :

$$N = \left\lfloor \lg \frac{\rho_{best}}{\rho_0} \right\rfloor \quad (9)$$

对图像获取系统在一个正程期间获取的红外序列图像进行子集筛选(关键帧的提取), 设  $N_p$  为前一关键帧帧号,  $N_m$  后一关键帧帧号。以序列图像第 1 帧为起始关键帧, 则一个镜头(正程期间)内获取的关键帧图像之间的帧号应满足:

$$N_m - N_p = 6 \quad (10)$$

即对序列图像间隔 6 帧抽取图像构成子集。

### 1.5 提取流程

图 6 中  $t$  表示摆扫正程的时间,  $f$  表示相机焦距,  $F$  表示相机的帧率。在 LABVIEW2012 开发环境下对获取的红外序列图像进行关键帧提取。

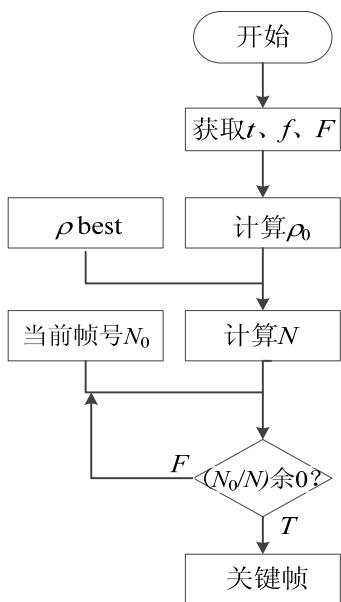


图 6 算法流程图 Fig.6 Flowchart of algorithm

## 2 实验结果及分析

实验硬件环境如下：计算机 CPU Intel(R)

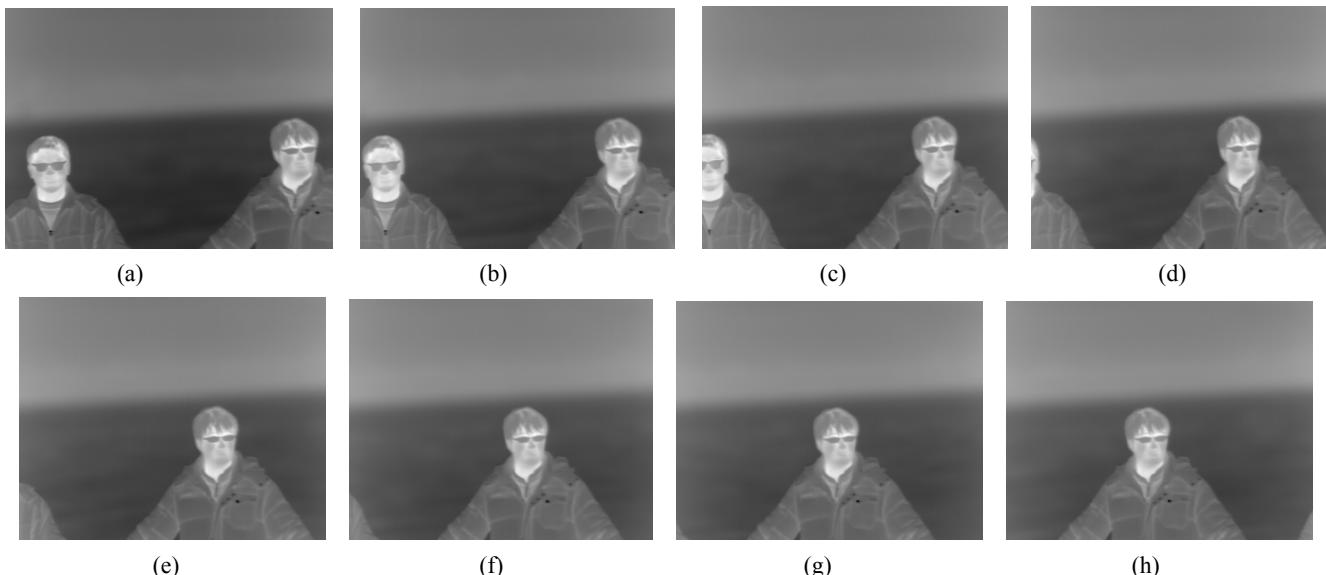


图 7 部分序列图像

表 1 平移量及重叠率

Table 1 Translation and overlap rate

序列号	平移量	重叠率 $\rho_1$	误差 $e = \rho_1 - \rho_0$
1-2	(38.47, 0.50)	93.99	1.48%
2-3	(37.46, 0.49)	94.14	1.63%
3-4	(38.43, 0.46)	93.99	1.48%
4-5	(39.40, 1.40)	93.84	1.33%
5-6	(38.42, 0.49)	93.99	1.48%
6-7	(38.47, 0.48)	93.99	1.48%
7-8	(39.41, 0.46)	93.84	1.48%
1&7	(231.44, 5.45)	63.83	—

Core(TM) i5-2500；内存为 4 GB、系统为 32 位 winXP 系统；采用烟台艾睿光公司 SZ007 型长波红外相机。实验地点位于上海市浦东新区滴水湖北岸，实验过程采用 PL40 步进电机搭载红外相机来模拟机载搜救过程中的摆扫成像方式，通过 LABVIEW 开发环境下开发的图像获取程序控制步进电机摆动进而带动红外相机对湖面进行摆扫成像，图像获取系统的参数为：摆扫范围 60°，历时 3.5 s，相机帧频为 25，成像分辨率为 640×480，位深为 14 位，噪声等效温差为 50 mK，焦距为 100 mm。通过对站在岸边的实验人员进行摆扫成像后，提取红外图像序列来完成实验分析。图 7 为部分序列图像。

### 2.1 重叠率的验证

截取获取的红外序列图像的部分，以相位相关法测出相邻图像间的重叠率，对式 8 计算出的重叠率进行验证，结论见表 1。

表 1 可知，序列中相邻图像在  $y$  方向平移量  $\Delta y = 39$  像素， $x$  方向平移量在  $\Delta x = 1$  像素波动。 $\Delta x$  的产生是因为图像获取系统没有严格的水平放置，且  $\Delta x$

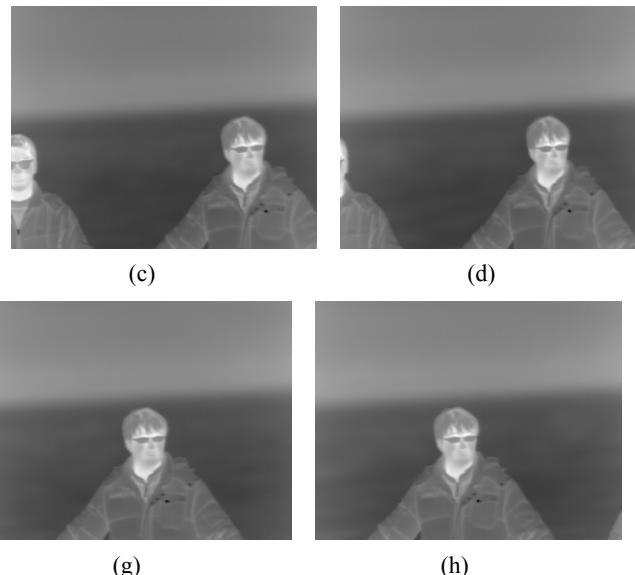


Fig.7 Sequence of infrared image

对重叠率的影响可以忽略不计。表 1 中式 8 计算出的重叠率与相位相关法计算出的重叠率之间的误差用  $e$  表示，并有  $e < 2\%$ ，即对于实际图像来说相差 12.8 像素内。按照式(9)计算出的间隔数  $N$ ，抽取的 2 个关键帧 1、7，重叠率差为  $63.83 - \rho_{best} = 3.83\%$ ，拼接中，稍大于最佳重叠率的 60% 对拼接来说不是坏事，因此根据摄影计算得出的重叠率  $\rho_0$  可以作为提取关键帧的指标。

### 2.2 大视场图像的生成

一个镜头内获取  $t \times F = 88$  帧图像，经过式(9)计算的间隔数  $N$  进行关键帧提取，并去掉首关键帧和末

关键帧(因为首末两帧图像因电机突然加速或减速,成像效果差)则最终关键帧子集序列有13帧,帧号如式(11):

$$[6 \ 12 \ 18 \ 24 \ \cdots \ 72 \ 78] \quad (11)$$

图8为按照式(11)对获取的一个序列图像进行抽取后的关键帧子序列。

图9为对关键帧子序列进行拼接的大视场图像,图10为对全序列图像进行拼接的大视场图像。

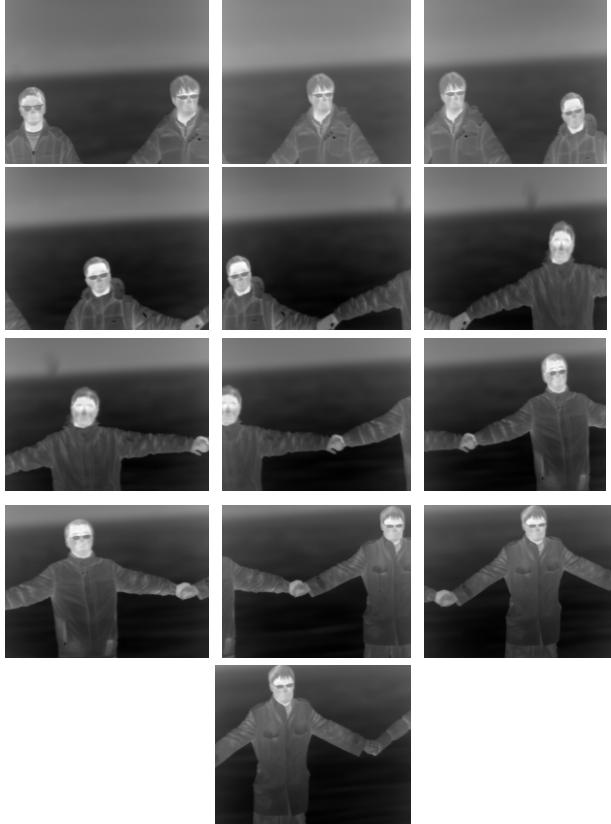


图8 关键帧子序列

Fig.8 Sequence of key frame



图9 关键帧子序列大视场生成

Fig.9 Wide filed base on sequence of key frame



图10 全序列大视场生成

Fig.10 Wide field based on whole sequence image

在拼接过程中,将生成的新图像作为参考图像,然后依次读取序列图像,拼接图像在重叠区域进行Harris角点检测<sup>[8]</sup>。采用Normalized Cross Correlation粗匹配<sup>[9]</sup>、RANSAC细匹配<sup>[10]</sup>、渐入渐出融合方法。最终比较见表2。

由表2和图9、10可知,根据摄影测量计算的关

键帧,保证大视场图像质量下,减少了待拼接图像的数量,同时减少了图像融合的次数,生成大视场图像的效率提高了88.71%。在实验过程中,图像获取系统未严格放置水平、 $\Delta x$ 的误差累加、人物之间未很好的保持静止状态等因素导致图9、图10大视场图像向右下方偏移。

表2 大视场图像效率比较

Table 2 Efficiency comparison of wide field image

	总时间/s	融合次数
全序列	124	87
关键帧子序列	14	12

### 3 结束语

根据摄影测量学计算了序列图像中相邻图像之间的重叠率,以此重叠率作为序列图像关键帧抽取的指标,避免了对序列图像进行子集选取时的盲目性,在保证拼接算法所要求重叠率前提下,去除了冗余图像,使大视场图像呈现的效率提高了88.71%。方法本质上属于镜头关键帧提取方法,但解决了传统镜头关键帧提取方法存在的固定帧数目、对变化快的序列内容不适合提取关键帧的缺点。同时方法还存在缺陷,当相机摆扫角度较大时,相邻关键帧之间的重叠率因为受几何变形的影响,用固定间隔N来计算重叠率已不准确,间隔数目N应该是变化的。这也是后期研究中,要加强的地方。

### 参考文献:

- [1] 韩龙. 移动环境下视频拼接技术的研究[D]. 安徽: 中国科学技术大学, 2010.
- [2] 孙淑敏, 张建明, 孙春梅. 基于改进 K-means 算法的关键帧提取[J]. 计算机工程, 2012, 38(23): 169-172.
- [3] 张建明, 蒋兴杰, 李广翠, 等. 基于粒子群的关键帧提取算法[J]. 计算机应用, 2011, 31(2): 358-361.
- [4] 罗森林, 马舒洁, 杨静, 等. 基于子镜头聚类方法的关键帧提取技术[J]. 北京理工大学学报, 2011, 31(3): 348-352.
- [5] 张德新. 面阵航侦 CCD 相机系统设计及其图像拼接技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2010.
- [6] 王佩军, 徐亚明. 摄影测量学[M].2 版. 武汉: 武汉大学出版社, 2010: 22-45.
- [7] 冯宇平. 图像快速配准与自动拼接技术研究[D]. 长春: 中国科学院研究生院长春精密机械与物理研究所, 2010.
- [8] 姚伟, 胡茂海, 秦晅, 等. 改进的基于特征点匹配的图像拼接算法[J]. 红外技术, 2010, 32(5): 288-290.
- [9] 龚平, 刘相滨, 周鹏. 一种改进的Harris角点检测算法[J]. 计算机工程与应用, 2010, 46(11): 173-175.
- [10] 穆柯楠, 惠飞, 曹健明, 等. 一种基于几何约束的 RANSAC 改进算法[J]. 计算机工程与应用, 2013, 32(10): 123-125.