

## 基于仿射模型的无人机视频实时压缩算法

沈秋\* 李小凡 孔繁镛 代侯西  
(南京航空航天大学航天学院 南京 210016)

**摘要:** 在现有算法无法满足无人机视频实时压缩传输需求的应用背景下, 该文提出一种新的低复杂度无人机(Unmanned Aerial Vehicle, UAV)视频实时压缩算法, 利用无人机视频的背景平面性、运动一致性等特点, 建立仿射模型描述视频的全局运动, 并根据模型参数进行视频压缩。实验结果表明, 与现有压缩算法 H.264 相比, 该算法在保证压缩性能的前提下, 大量减少了压缩时间, 可满足大多数情况下无人机视频传输的质量要求和实时性要求。

**关键词:** 视频压缩; 无人机视频; 全局运动估计; 仿射模型

中图分类号: TN919.81

文献标识码: A

文章编号: 1009-5896(2014)12-2855-06

DOI: 10.3724/SP.J.1146.2014.00080

## A Real-time Video Compression for UAV Based on Affine Model

Shen Qiu Li Xiao-fan Kong Fan-qiang Dai Yu-xi

(College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

**Abstract:** Since the existing algorithms can not meet the requirements of real-time compression and transmission for UAV(Unmanned Aerial Vehicle) videos, a new real-time compression algorithm with low complexity for UAV videos is proposed. Considering the plane background and unified motion of UAV videos, the proposed method establishes an affine model for global motion estimation and compression. The experimental results demonstrate that the proposed algorithm is able to reduce the total encoding time while maintaining the performance compared with H.264, which the quality requirement and real-time performance for UAV videos transmission can be satisfied in most cases.

**Key words:** Video compression; Unmanned Aerial Vehicle (UAV) videos; Global motion estimation; Affine model

### 1 引言

信息化时代的到来要求人们改变对偏远地区(例如高原、森林等)信息采集的传统方式。无人机(Unmanned Aerial Vehicle, UAV)由于经济环保、安全便捷, 现已越来越广泛地应用于土地监测、森林防火、抗震救灾、电力巡线等领域。然而无人机由于机载设备受限、要求视频实时传输等原因, 要求其采用的视频压缩算法复杂度低、压缩性能高。

目前, 无人机视频压缩主要使用 MPEG-2<sup>[1]</sup>和 H.264<sup>[2]</sup>等标准算法, 算法核心是基于块的运动补偿/变换的混合编码, 算法复杂度较高, 无法实现实时压缩和传输。运动估计是现有视频压缩算法的核心技术之一, 也是算法复杂度高的主要因素<sup>[3]</sup>, 以

H.264 为例, 运动估计的时间占整个视频编码时间的 60%~80%<sup>[4]</sup>。因此, 如何有效地进行运动估计是减少压缩编码时间的关键。

一方面, 无人机在执行侦察、勘测任务时, 大部分飞行高度都在 500 m 以上, 远大于拍摄距离, 故可以忽略地表起伏以及独立的运动目标, 将拍摄的背景视为平面<sup>[5]</sup>; 另一方面, 航拍视频中独立的运动目标在画幅中所占比例较小, 并且大部分拍摄任务在飞行状态比较平稳时进行, 因此, 可假设无人机视频的相邻帧之间存在有规律的、一致的全局运动; 其次, 无人机的飞行速度一般在 100~300 km/h, 而视频拍摄帧率在 25~30 帧/s, 两帧之间具有 70%~90%的重叠度<sup>[6]</sup>。鉴于以上几个特征, 在大部分情况下, 对无人机拍摄视频采用全局运动估计代替分块的运动估计, 可以在不严重影响视频质量的同时, 较大地降低算法复杂度。

现有的全局运动估计算法可分为基于像素域和基于压缩域的方法, 前者大多采用的是最小二乘的迭代方法以及快速的金字塔算法, 又可细分为基于

2014-01-13 收到, 2014-05-09 改回

国家自然科学基金(61201365), 南京航空航天大学基本科研业务费专项项目(NS2012071), 南京航空航天大学青年科技创新基金(NS2013085), 南京航空航天大学研究生创新基地(实验室)开放基金(kfjj201457)和中央高校基本科研业务费专项资金资助课题

\*通信作者: 沈秋 shenqiu@nuaa.edu.cn

像素<sup>[7]</sup>、基于特征<sup>[5, 8, 9]</sup>、基于块匹配<sup>[10, 11]</sup>3种算法;后者利用压缩码流中的相关信息进行快速的全局运动估计<sup>[12, 13]</sup>,相较于前者,计算量少,但压缩码流中运动矢量由于局部运动影响,会使全局运动估计产生偏差。现有利用全局运动估计进行视频压缩的算法,其全局运动矢量只包含水平和垂直两个分量,适合简单的平移运动,不能应对旋转、缩放等复杂情况,预测后的编码残差大,码率高而性能低。

本文结合无人机运动的特点,提出了一种基于六参数仿射变换模型的全局运动估计算法,可以精准地描述视频图像的平移、旋转、缩放等运动,并利用模型参数指导视频的帧间预测和运动补偿,在保持视频压缩性能的同时,极大地降低计算复杂度。

## 2 无人机视频的全局运动模型

按照摄像机运动方式的不同,全局运动模型有对应于平移的二参数模型,对应于平移和缩放运动的三参数模型<sup>[14]</sup>,对应于平移、缩放和旋转的四参数模型,六参数的仿射模型<sup>[15, 16]</sup>,八参数的投影模型以及十二参数的抛物模型。参数越多,描述的模型越复杂,描述的精度也越高,但是模型的计算复杂度也就越高。

如图1所示,无人机航拍高度比较高,与拍摄的距离比起来,地表的起伏可以忽略,可以将拍摄画面视为在2维平面当中。图1中 $H$ 表示拍摄高度, $L$ 表示拍摄范围的长度,阴影部分表示物体平面内拍摄的重合区域, $P_{1w}$ 表示重合区域内的一点, $P_1, P_1'$ 分别是对应物点 $P_{1w}$ 的图像平面内的像点。

当无人机做水平运动时,其运动模型可以用二参数的平移模型来表达。

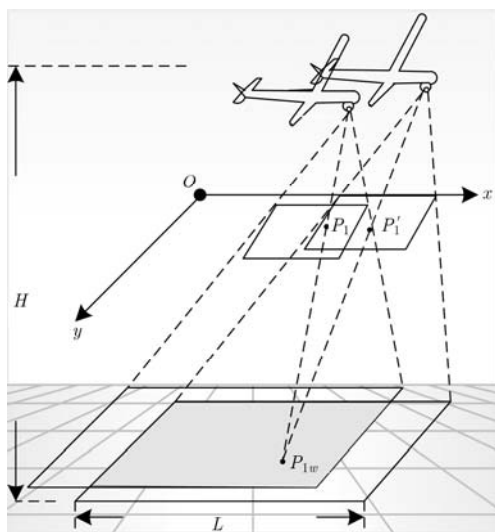


图1 航拍视频成像图

$$\left. \begin{aligned} x_{n-1}^{(i)} &= x_n^{(i)} + a_n \\ y_{n-1}^{(i)} &= y_n^{(i)} + b_n \end{aligned} \right\} \quad (1)$$

其中, $(x_{n-1}^{(i)}, y_{n-1}^{(i)})$ 为参考帧中的第 $i$ 个特征点坐标, $(x_n^{(i)}, y_n^{(i)})$ 为当前帧中的对应特征点坐标, $(a_n, b_n)$ 为常数,代表全局运动参数。

然而无人机在飞行过程中免不了要进行各种姿态调整,这就会带来比平移更复杂的全局运动,如旋转、缩放等,如图2所示,是一个无人机航拍视频中某一帧的运动矢量图的一部分。图像上半部分与下半部分的运动方向不一致,运动矢量的大小也不相同,一般的平移模型已经不能满足这种情况。

综合考虑无人机视频运动特点和计算复杂度,本文选用六参数仿射运动模型作为无人机视频全局运动模型。设 $P_{n-1}^i(x_{n-1}^{(i)}, y_{n-1}^{(i)})$ 是参考帧 $F_{n-1}$ 的第 $i$ 个特征点, $P_n^i(x_n^{(i)}, y_n^{(i)})$ 是当前帧 $F_n$ 中的对应的匹配特征点,则 $P_n^i(x_n^{(i)}, y_n^{(i)})$ 到 $P_{n-1}^i(x_{n-1}^{(i)}, y_{n-1}^{(i)})$ 的变换关系用齐次坐标式表达为

$$\begin{bmatrix} x_{n-1}^{(i)} \\ y_{n-1}^{(i)} \\ 1 \end{bmatrix} = \mathbf{T}_{n,n-1} \cdot \begin{bmatrix} x_n^{(i)} \\ y_n^{(i)} \\ 1 \end{bmatrix} \quad (2)$$

其中

$$\mathbf{T}_{n,n-1} = \begin{bmatrix} a_{n0} & b_{n0} & c_{n0} \\ a_{n1} & b_{n1} & c_{n1} \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

为仿射模型变换矩阵。该模型不仅可以描述平移运动,还可以描述旋转和缩放,并且算法复杂度不高,只需3对以上匹配点对就可以求出 $\mathbf{T}_{n,n-1}$ 中的6个未知参数。这样,全局运动参数估计就变成求解超定方程组的问题,可以用最小二乘法计算。设有 $N$ 对匹配点,则有

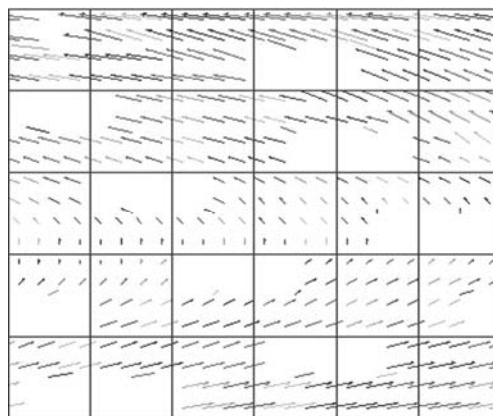


图2 航拍视频运动矢量图

$$\begin{aligned} \mathbf{X}_{n-1} &= \begin{bmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \\ \mathbf{1} \end{bmatrix} \\ \mathbf{Y}_{n-1} &= \begin{bmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \\ \mathbf{1} \end{bmatrix} \end{aligned} \quad (4)$$

其中  $\mathbf{X}_n = (x_{n1}, x_{n2}, \dots, x_{nN})$  表示当前帧中匹配点的  $x$  坐标组成的向量,  $\mathbf{Y}_n = (y_{n1}, y_{n2}, \dots, y_{nN})$  表示当前帧中匹配点的  $y$  坐标组成的向量,  $\mathbf{X}_{n-1}$  与  $\mathbf{Y}_{n-1}$  的意义同理。则由最小二乘法可得极小解为

$$\begin{cases} \begin{bmatrix} a_{n0} & b_{n0} & c_{n0} \end{bmatrix} = \mathbf{X}_{n-1} \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \\ \begin{bmatrix} a_{n1} & b_{n1} & c_{n1} \end{bmatrix} = \mathbf{Y}_{n-1} \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \end{cases} \quad (5)$$

其中  $\mathbf{A} = \begin{bmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \\ \mathbf{1} \end{bmatrix}$ 。

### 3 基于仿射模型的无人机视频实时压缩算法

本文利用六参数仿射模型对无人机视频的全局运动进行描述, 并根据全局运动参数对视频进行压缩, 避开了视频压缩编码中最耗时的运动估计部分, 达到大幅降低计算复杂度的目的。算法主体框架与标准算法 H.264 相同(如图 3 所示), 主要改进的核心步骤包括两个部分, 即全局运动模型参数估计和基于全局运动模型的运动补偿, 如图中虚线框内所示。

#### 3.1 全局运动模型参数估计

本文采用基于特征点匹配的全局运动模型参数估计, 主要步骤包括:

(1)对当前帧  $F_n$  和参考帧  $F_{n-1}$  进行特征点提取与匹配, 采用的是计算复杂度低的 ORB 算法<sup>[7]</sup>, 并利用随机抽样一致算法(RANdOm SAmple Consensus, RANSAC)方法<sup>[8]</sup>去除误匹配点对;

(2)由匹配点对坐标利用最小二乘法计算得出全局运动参数  $T_{n,n-1}$  (如式(4)和式(5));

(3)估计出的运动参数需要加入到压缩码流中, 可以加在视频图像头参数列表中, 也可以利用补充增强信息单元携带该信息。

#### 3.2 基于全局运动模型的运动补偿

该步骤直接利用上述模型参数对当前帧  $F_n$  中的每个子块换算出运动矢量: 将每个宏块(大小为  $16 \times 16$ )划分为 16 个子块(大小为  $4 \times 4$ ), 以每个子块的左上顶点为坐标, 利用式(2), 得到在前一帧中的匹配点坐标  $(x', y')$ , 则运动矢量为

$$\mathbf{MV} = (x', y') - (x, y) \quad (6)$$

直接利用换算出的运动矢量对图像进行运动补偿和预测, 并对帧间预测补偿后的残差压缩编码得到压缩比特流, 而该码流中不再携带每个分块的运动矢量。

### 4 实验结果与分析

实验以 H.264 的参考软件 JM12.1 为基础进行算法改进和测试, 设置参考帧个数为 1, 采用 I-P-P 编码方式, 运动估计采用全搜索方式。对 9 个无人机航拍视频片段进行测试, 其中, 序列 1, 2, 3, 5, 9 有镜头旋转, 序列 1, 9 纹理复杂, 序列 7, 8 有镜头缩放; 视频分辨率为  $352 \times 288$ , 帧率为 25 帧/s, 每个序列的总编码帧数为 25 帧。计算机系统为 Windows7(主频 2.3 GHz, 内存 16 G)。

实验中分别对 H.264 压缩算法全搜索方式(以下用 H.264 全搜索代指)、基于二参数平移模型的全局运动估计压缩算法(以下用二参数平移模型代指)和本文提出的算法(以下用六参数仿射模型代指)进行测试和比较。

#### 4.1 实验结果

**4.1.1 压缩时间比较** 各算法平均压缩用时如表 1 所示, 本文算法和基于二参数平移模型的算法所需视频压缩时间均远远少于 H.264 全搜索方式, 只有 H.264 全搜索方式的 4%左右, 对于无人机视频实时传输应用具有重要的实践意义。并且由于基于二参数平移模型的算法不能准确地表达全局运动, 导致编码残差较大, 对应的编码时间略长, 因此, 其压缩时间也略大于本文算法。

**4.1.2 压缩性能比较** 本文对于压缩视频的性能采用率失真曲线图衡量。量化参数分别取 16, 22, 28, 34, 40, 46, 这里列举 4 个序列的率失真曲线图如图 4~图 7 所示。

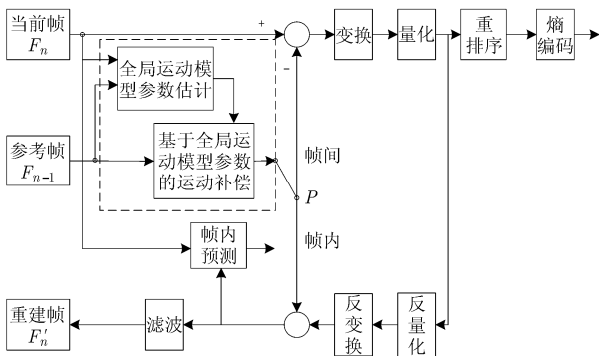


图 3 基于仿射模型的无人机压缩算法编码框架图

表 1 压缩时间比较

序列	H.264 全搜索(s)	二参数平移模型(s)	六参数仿射模型(s)
序列 1	28.398	1.321	1.188
序列 2	27.088	1.008	0.954
序列 3	28.037	1.239	1.199
序列 4	27.611	1.206	1.087
序列 5	28.423	1.118	1.068
序列 6	27.560	1.124	1.014
序列 7	27.904	1.277	1.112
序列 8	29.236	1.531	1.386
序列 9	29.772	1.470	1.312

从图 4~图 7 分析可知,在低码率情况下,运动矢量数据量在码流中所占比重较大,而本文算法由于传输的是整帧的全局运动信息,不含有每个宏块的运动矢量,所以码率比 H.264 全搜索方式低,压缩性能与 H.264 全搜索方式相当;而高码率时,由

于码流中编码残差的数据量所占比重较大,而运动矢量的取代无法显著地减小数据量,并且全局运动的精度偏差造成图像整体上存在一定程度的偏移,所以残差较大,因此,性能比 H.264 全搜索方式低 4%左右。

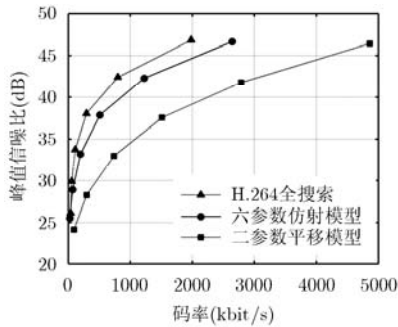


图 4 序列 1 性能比较图

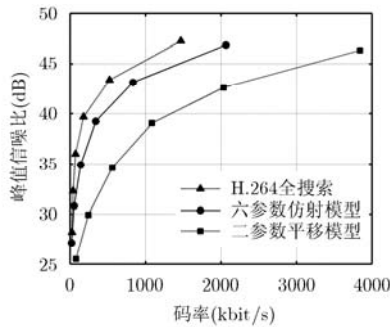


图 5 序列 4 性能比较图

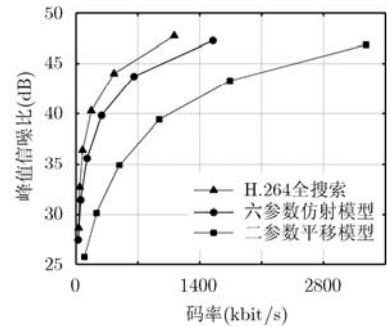


图 6 序列 6 性能比较图

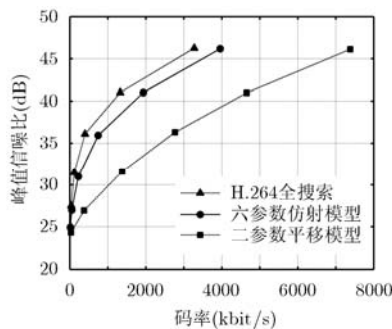


图 7 序列 9 性能比较图

样的码率下本文算法重建帧与 H.264 全搜索方式重建帧在细节方面具有一样良好的表现,与原始视频帧相比没有明显的损失,而二参数平移模型方式则细节损失严重,进一步说明了本文算法的有效性。

**4.1.3 主观质量及细节损失比较** 当压缩比为 102:1 时,选取序列 1 的重建视频中第 3 帧作为比较。图 8~图 11 分别展示了原始帧及重建帧以及它们的局部细节放大图和使用零交叉方式(zerocross)提取出来的边缘局部细节放大图。从图中可以看出,在同

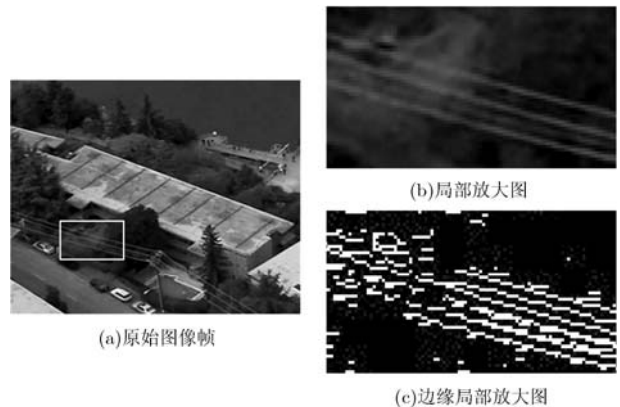


图 8 原始图像帧

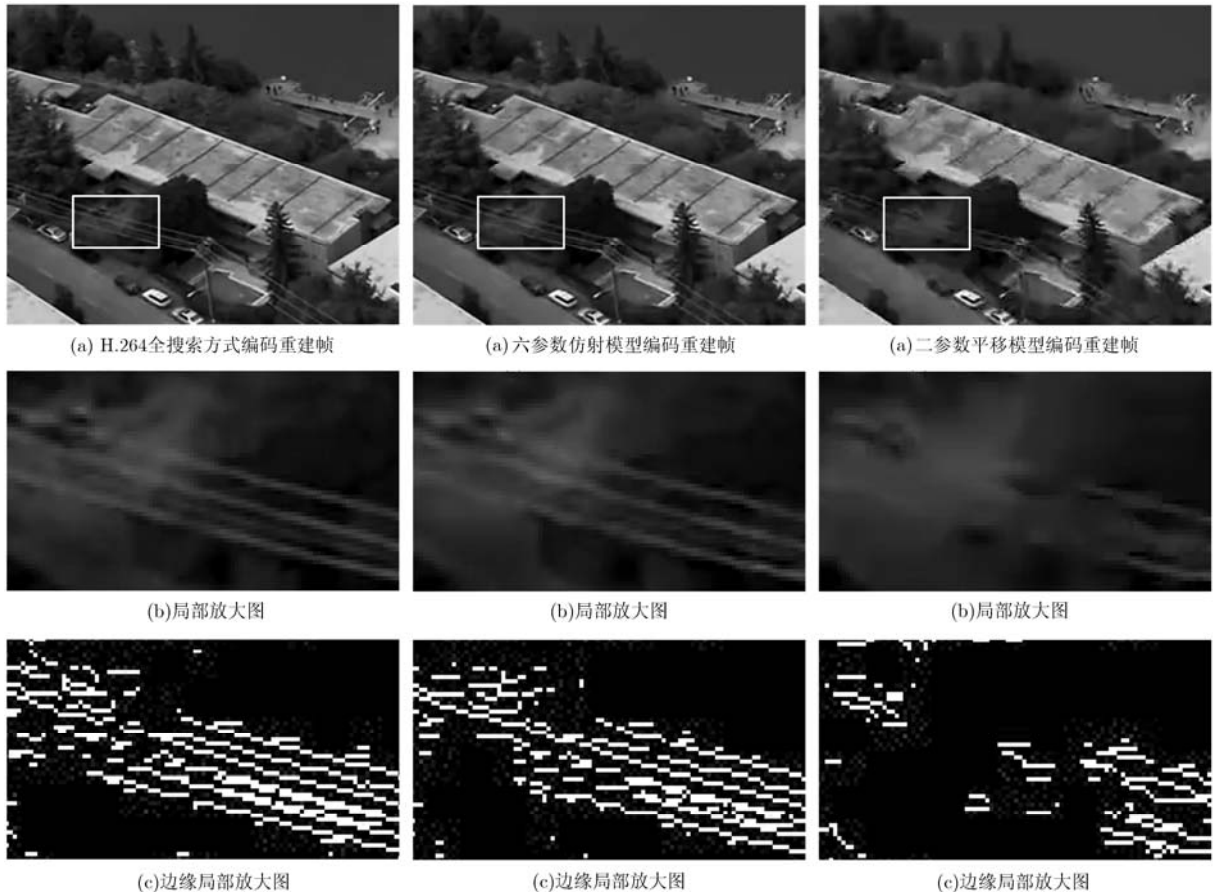


图 9 H.264 全搜索方式编码重建

图 10 六参数仿射模型编码重建

图 11 二参数平移模型编码重建

#### 4.2 实验结果分析

本实验分别从压缩时间、客观重建质量和主观重建质量 3 个方面，对视频压缩算法性能进行了全面的比较与分析。结果显示本文算法比 H.264 全搜索方式算法复杂度降低了 95% 以上；而基于二参数平移模型的算法虽然在压缩时间上和本文算法相当，但由于其无法适应旋转、缩放运动，造成运动矢量误差较大，所以压缩性能要远低于本文的算法，不能满足无人机视频实时传输的质量要求。因此，与传统方法相比，本文提出的基于六参数仿射模型的算法可以在压缩时间和压缩质量两方面取得最优的综合性能。

#### 5 结束语

本文在现有压缩算法的基础上，针对无人机的特殊需求提出了一种新的视频压缩编码算法，利用无人机视频的背景平面性、运动一致性等特点，建立仿射模型描述视频的全局运动，并根据模型参数进行视频压缩。算法计算复杂度低且性能与现有 H.264 算法相当，达到了实时性压缩目的，对于机载设备受限的无人机而言具有显著的实际应用价值。该算法对于无人机应用的大多数情况具有良好的性能，但也具有一定的局限性，当无人机运动变

化剧烈或飞行高度较低、运动目标较大时，该算法模型不能很好地表达视频中的复杂运动，从而造成性能下降。因此，作者下一步的研究重点是进一步地优化和完善全局运动模型以及实现运动目标的分割与独立编码，从而提高本文算法的适用性和有效性。

#### 参考文献

- [1] 毛伟勇. 可用于无线传输的高清 MPEG-2 实时编码器的设计与实现[D]. [硕士学位论文], 上海交通大学, 2008.  
Mao Wei-yong. Design and implementation of high definition real-time MPEG-2 encoder used for wireless transmission[D]. [Master dissertation], Shanghai Jiao Tong University, 2008.
- [2] 陈坤. 微型无人机图像传输系统研究[D]. [硕士学位论文], 天津大学, 2011.  
Chen Kun. Research of image transmission system in micro-UAV[D]. [Master dissertation], Tianjin University, 2011.
- [3] 刘英哲, 王进祥. H.264 中一种基于搜索范围自适应调整的运动估计算法[J]. 电子与信息学报, 2013, 35(6): 1382-1387.  
Liu Ying-zhe and Wang Jin-xiang. Motion estimation algorithm based on adaptive search range adjustment for H.264[J]. *Journal of Electronics & Information Technology*,

- 2013, 35(6): 1382-1387.
- [4] 姚玉兵. H.264 视频编码运动估计算法研究[D]. [硕士学位论文], 华东理工大学, 2011.  
Yao Yu-bing. Research of motion estimation based on H.264 video coding[D]. [Master dissertation], East China University of Science and Technology, 2011.
- [5] 王云丽, 张鑫, 高超, 等. 航拍视频拼图中基于特征匹配的全局运动估计方法[J]. 航空学报, 2008, 29(5): 1218-1225.  
Wang Yun-li, Zhang Xin, Gao Chao, *et al.* Feature matching based global motion estimation in aerial video mosaicing[J]. *Acta Aeronautica et Astronautica Sinica*, 2008, 29(5): 1218-1225.
- [6] 田金文, 谢清鹏, 谭毅华, 等. 无人机序列图像压缩方法研究[J]. 华中科技大学学报(自然科学版), 2005, 33(12): 76-78.  
Tian Jin-wen, Xie Qing-peng, Tan Yi-hua, *et al.* Methods for compressing sequence image of unmanned aerial vehicle[J]. *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, 2005, 33(12): 76-78.
- [7] 王思珺, 赵建, 韩希珍. 基于仿射变换的快速全局运动估计算法[J]. 液晶与显示, 2012, 27(2): 263-266.  
Wang Si-jun, Zhao Jian, and Han Xi-zhen. Fast global motion estimation algorithm based on affine transformation[J]. *Chinese Journal of Liquid Crystal and Displays*, 2012, 27(2): 263-266.
- [8] 于明, 孙炜焯, 阎刚, 等. 动态场景下基于精确背景补偿的运动目标检测[J]. 计算机应用与软件, 2013, 30(10): 139-141.  
Yu Ming, Sun Wei-ye, Yan Gang, *et al.* Moving objects detection based on exact background compensation in dynamic scene[J]. *Computer Applications and Software*, 2013, 30(10): 139-141.
- [9] 王天召, 徐克虎, 陈金玉, 等. 基于均值聚类 and 几何关系的运动背景估计算法研究[J]. 光电技术, 2013, 33(4): 244-248.  
Wang Tian-zhao, Xu Ke-hu, Chen Jin-yu, *et al.* Research on algorithm of moving background estimation based on means clustering and triangulation[J]. *Optoelectronic Technology*, 2013, 33(4): 244-248.
- [10] 吴浩, 邓宏彬, 何少阳. 基于分块灰度投影的无人飞行器视频稳像方法[J]. 北京理工大学学报, 2013, 33(4): 385-389.  
Wu Hao, Deng Hong-bin, and He Shao-yang. A UAV video stabilization method based on sub-block gray projection[J]. *Transactions of Beijing Institute of Technology*, 2013, 33(4): 385-389.
- [11] 席志红, 初守艳. 全局运动中结合改进 C-1BT 变换的平移运动估计[J]. 应用科学学报, 2013, 31(6): 601-606.  
Xi Zhi-hong and Chu Shou-yan. Translational motion estimation combined with improved C-1BT transform in global motion[J]. *Journal of Applied Sciences — Electronics and Information Engineering*, 2013, 31(6): 601-606.
- [12] 王喆, 刘贵忠, 钱学明. 一种高效的基于 H.264/AVC 压缩域信息的全局运动估计方法[J]. 电子学报, 2011, 39(3A): 19-23.  
Wang Zhe, Liu Gui-zhong and Qian Xue-ming. An efficient global motion estimation algorithm on H.264/AVC compression domain[J]. *Acta Electronica Sinica*, 2011, 39(3A): 19-23.
- [13] 王丽. 视频流中运动对象提取与分割的研究[D]. [硕士学位论文], 电子科技大学, 2011.  
Wang Li. Research on moving object extraction and segmentation among a video stream[D]. [Master dissertation], University of Electronic Science and Technology of China, 2011.
- [14] 王嘉, 王海峰, 刘青山, 等. 基于三参数模型的快速全局运动估计[J]. 计算机学报, 2006, 29(6): 920-926.  
Wang Jia, Wang Hai-feng, Liu Qing-shan, *et al.* Fast global motion estimation based on 3-parameter global motion model[J]. *Chinese Journal of Computers*, 2006, 29(6): 920-926.
- [15] 郭丽, 龚声蓉, 崔志明. 一种快速全局运动估计算法[J]. 中国图象图形学报, 2007, 12(12): 2119-2126.  
Guo Li, Gong Sheng-rong, and Cui Zhi-ming. A fast global motion estimation algorithm[J]. *Journal of Image and Graphics*, 2007, 12(12): 2119-2126.
- [16] Bhaskaranand M and Gibson J D. Low-complexity video encoding for UAV reconnaissance and surveillance[C]. Proceedings of the Military Communications Conference, Baltimore, MD, USA, 2011: 1633-1638.
- [17] Rublee E, Rabaud V, Konolige K, *et al.* ORB: an efficient alternative to SIFT or SURF[C]. Proceedings of the Computer Vision (ICCV), Barcelona, Spain, 2011: 2564-2571.
- [18] 张波, 罗海勇, 刘冀伟, 等. 视觉传感器网络中基于 RANSAC 的稳健定位算法[J]. 通信学报, 2013, 34(8): 62-69.  
Zhang Bo, Luo Hai-yong, Liu Ji-wei, *et al.* RANSAC based robust localization algorithm for visual sensor network[J]. *Journal on Communications*, 2013, 34(8): 62-69.
- 沈秋: 女, 1982年生, 讲师, 研究方向为视频压缩与处理等.  
李小凡: 男, 1990年生, 硕士生, 研究方向为数字图像处理等.  
孔繁镛: 男, 1980年生, 讲师, 研究方向为多光谱遥感图像传输与处理等.  
代俣西: 女, 1991年生, 硕士生, 研究方向为三维数字视频处理等.