

# 基于预测的无人机视频快速匹配算法研究

杨德龙<sup>1</sup>, 吴云东<sup>2</sup>, 蔡国榕<sup>2</sup>, 陈水利<sup>3</sup>

(1. 集美大学理学院, 福建 厦门 361021; 2. 集美大学计算机工程学院, 福建 厦门 361021;

3. 集美大学诚毅学院, 福建 厦门 361021)

[摘要] 针对无人机视频数据, 提出一种快速逐帧匹配算法。该算法首先使用计算机在每一帧中均匀标记一定数量的点, 接着利用视频相邻帧之间的特点使用绝对差值法完成同名点的寻找, 之后引入内插理论建立预测模型, 缩小了搜索窗口和匹配窗口大小, 降低了算法时间复杂度。实验结果表明, 该算法可以高效地完成相邻帧之间的匹配。

[关键词] 相邻帧; 模板匹配; PID; 预测; 内插

[中图分类号] N 941.5

## Fast Matching Methods for UAV Video with Prediction

YANG De-long<sup>1</sup>, WU Yun-dong<sup>2</sup>, CAI Guo-rong<sup>2</sup>, CHEN Shui-li<sup>3</sup>

(1. School of Science, Jimei University, Xiamen 361021, China;

2. School of Computer Engineering, Jimei University, Xiamen 361021, China;

3. Chengyi University College, Jimei University, Xiamen 361021, China)

**Abstract:** This paper proposes a fast matching algorithm for UAV's video. Firstly, the algorithm use computer to sign some points which are uniform distribution on the image. Secondly, the algorithm take advantage of the adjacent frame's characteristics to find the correspond points. At last, we introducing the interpolation theory to establish the prediction model to reduce the size of the searching window and matching window, make the time complexity lower than before.

**Keywords:** adjacent frame; template matching; PID; prediction; interpolation

## 0 引言

利用无人机视频制作大场景静态图像, 有利于在军事侦察、抢险救灾等应用中掌握全局信息, 是无人机应用的关键技术之一<sup>[1]</sup>。当前, 对图像帧的拼接, 是获取大场景静态图像可行的技术途径, 为此, 需要针对视频帧的特点, 研究视频帧的快速匹配算法。

视频帧的时间间隔为毫秒级, 相邻帧之间具有尺度变化、旋转角度小、几何畸变不明显等特点, 可以视相邻帧为超高重叠度的立体像对, 这为帧与帧的匹配带来了极为有利的条件。对此, 在实现帧匹配的过程中, 应优先选用时间复杂度低的简单算法, 以满足数量庞大的视频帧匹配。同时, 视频往往是近距离地观测目标, 地物高差与航高的比值偏大, 造成帧图像之间视差连续性偏低, 属于大视差

[收稿日期] 2015-12-24

[修回日期] 2016-02-27

[基金项目] 国家科技支撑项目(2013BAC08B01-01)

[作者简介] 杨德龙(1988—), 男, 硕士生, 从事图像处理、模式识别研究。通信作者: 蔡国榕(1979—), 男, 副教授, 从事影像信息处理、模式识别研究, E-mail: guorongcai.jmu@gmail.com。

图像,不利于匹配。结合视频的成像过程,帧匹配可以视为一个大视差逐步移入的图像匹配问题。为此,有必要研究视差预测的方法,以减小搜索窗口,提高匹配效率。

现有匹配算法按匹配依据划分,有基于语义的匹配<sup>[2]</sup>、基于灰度的匹配<sup>[3]</sup>、基于特征的匹配<sup>[4]</sup>三类。基于语义的匹配目前研究成果较少,对于图像来说获取其结构信息比较困难,因此本文主要研究基于灰度的匹配和基于特征的匹配,从中选择最合适的算法用于视频帧的匹配。在基于灰度的匹配算法中,AD算法(绝对差值法)仅仅通过计算匹配窗口中像素灰度值的绝对差值之和完成图像匹配,时间复杂度是最低的<sup>[5-7]</sup>。而相关系数法是最常用的灰度匹配算法之一。虽然基于灰度的匹配算法效率高于基于特征的匹配,但是基于特征的匹配算法可靠性更高。由文献[8]知特征匹配算法中时间复杂度最低的是ORB(oriented FAST and rotated BRIEF)算法。本文首先对比ORB算法、相关系数法、AD算法这三种算法的可靠性和计算效率,从中选择合适的算法用于匹配相邻帧。

由于视频相邻帧之间的特点,可以引入预测理论以提高算法效率。PID由比例、积分、微分三部分构成一个系统,完成相应同名点的位移预测<sup>[9-10]</sup>。内插(interpolation)法可以根据前几帧同名点的位移预测下一帧的位移。因此本文对PID和内插法这两种预测模型的可靠性和效率展开讨论并进行实验验证,从中选择可靠性好、效率高的作为预测模型,并与合适的匹配算法相结合,以解决无人机视频数据的快速逐帧匹配问题。

# 1 模型建立

## 1.1 匹配算法选择

### 1.1.1 算法时间复杂度分析

基于灰度的匹配算法的时间复杂度由搜索窗口和匹配窗口的大小决定<sup>[11-13]</sup>。令搜索窗口大小为 $M \times N$ ,匹配窗口大小为 $K_1 \times K_2$ 。AD算法和相关系数法的时间复杂度均为 $O(M \times N \times K_1 \times K_2)$ 。虽然它们的时间复杂度属于同一数量级,但是在实际计算中AD算法的计算次数是 $3n(M \times N \times K_1 \times K_2)$ ,相关系数法的计算次数是 $9n(M \times N \times K_1 \times K_2)$ ,其中 $n$ 是匹配点的数量,因此AD算法的计算效率高于相关系数法。

ORB算法具有旋转不变性,本文结合视频相邻帧旋转角度小的特点,不再计算ORB算法中FAST<sup>[13]</sup>角点以及BRIEF<sup>[14]</sup>的方向,降低了算法时间复杂度。对于每一个像素,判断其是否是FAST角点需要64次计算。模板匹配中,本文选取49个点,平均分布于图像中,为了与之对比,FAST角点的搜索范围取 $7M \times 7N$ ,则FAST角点提取部分总的计算次数是 $3136(M \times N)$ 。提取FAST角点之后,使用BRIEF算法进行特征描述,对于每一个FAST角点,特征描述子用一个 $n_d = 256$ 维的二进制字符串表示,即 $f_{n_d}(p) = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(p; x_i, y_i)$ ,其中 $\tau(p; x_i, y_i)$ 计算次数为2,  $(x_i, y_i)$ 是属于以像素 $p$ 为中心的一个小区域中的点对,  $i = 1, 2, \dots, 512$ ,故获取一个点的BRIEF特征需要进行23 707 648次计算。获取特征描述子之后通过计算匹配点与待匹配点特征描述子之间的汉明距离进行点匹配,其算法时间复杂度是 $O(n^2)$ ,其中 $n$ 是FAST角点的个数。

由以上分析知ORB算法时间消耗随着匹配点的增加,时间消耗呈平方级增长,而AD算法和相关系数法呈线性增长。因此完成相同数量的同名点寻找,ORB算法的时间消耗最大,相关系数法次之,AD算法最小。

### 1.1.2 匹配结果对比分析

图1a、b是提取自同一段视频中的连续两帧,地物信息以建筑物为主,分别对以上两帧使用ORB算法、相关系数法、AD算法计算匹配率,其中的相关系数法和AD算法在第一帧中等间距标定49个点作为待匹配点,匹配结果如图2所示。

图3是提取自同一段视频中的连续两帧,地物信息以水面为主,分别对以上两帧使用ORB算法、

相关系数法、AD 算法计算匹配率，其中相关系数法和 AD 算法在第一帧中等间距标定 49 个点作为待匹配点，匹配结果如图 4 所示。

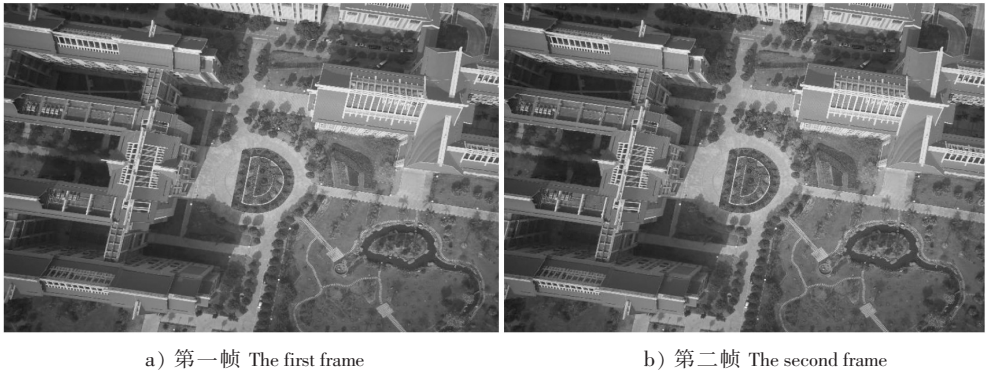
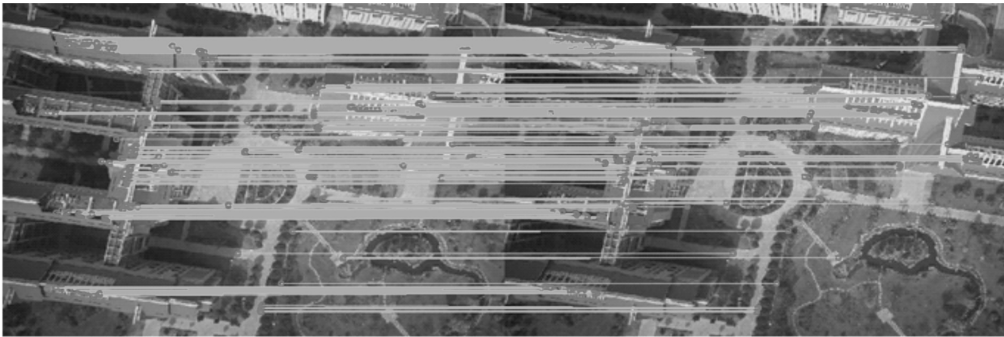
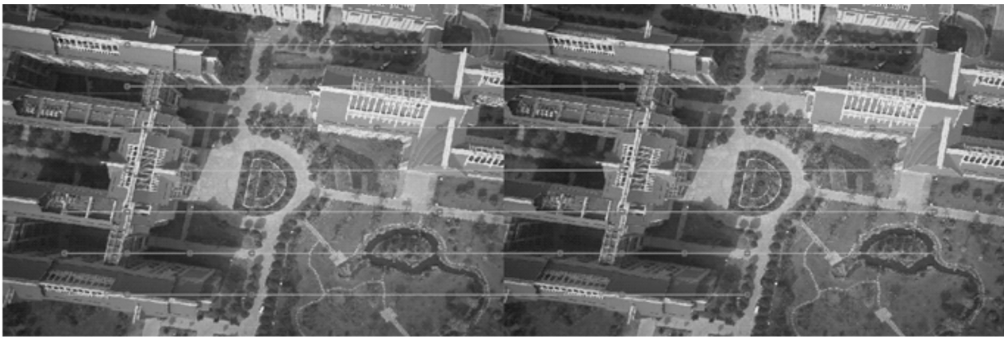


图 1 第一段视频中的连续两帧

Fig.1 Two consecutive frames in the first period of video



a) ORB 算法匹配结果 The matching result of ORB algorithm



b) 相关系数法匹配结果 The matching result of correlation coefficient algorithm



c) AD 算法匹配结果 The matching result of AD algorithm

图 2 第一组匹配结果

Fig.2 The first group of matching results



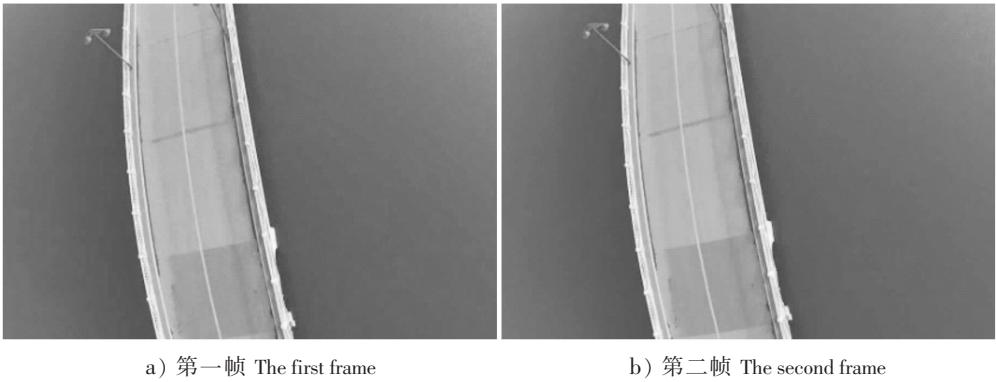


图 3 第二段视频中的连续两帧

Fig.3 Two consecutive frames in the second period of video

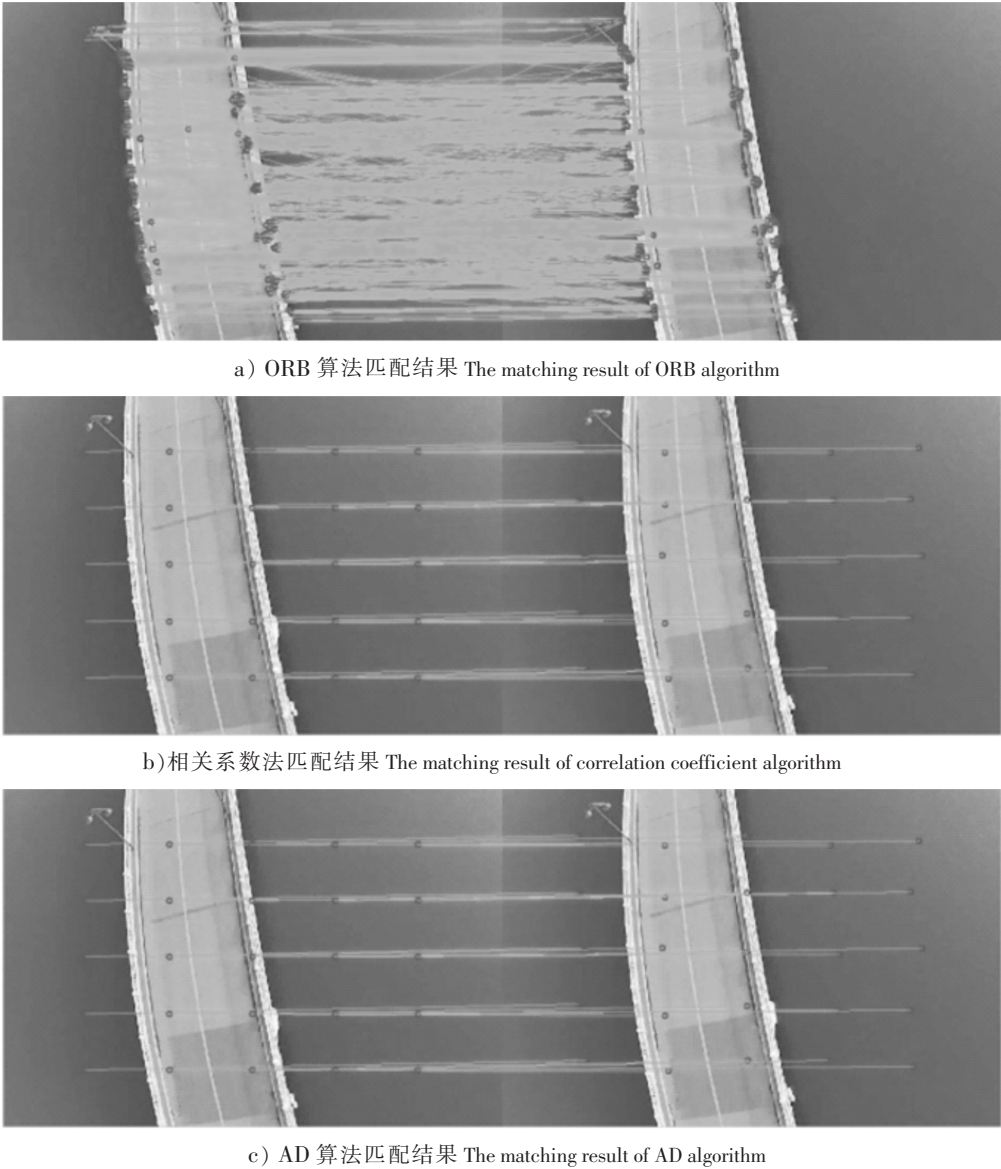


图 4 第二组匹配结果

Fig.4 The second group of matching results

图 5 是提取自同一段视频中的连续两帧，地物信息以草坪为主，分别对以上两帧使用 ORB 算法、相关系数法、AD 算法计算匹配率，其中的相关系数法和 AD 算法在第一帧中等间距标定 49 个点作为待匹配点，匹配结果如图 6 所示。

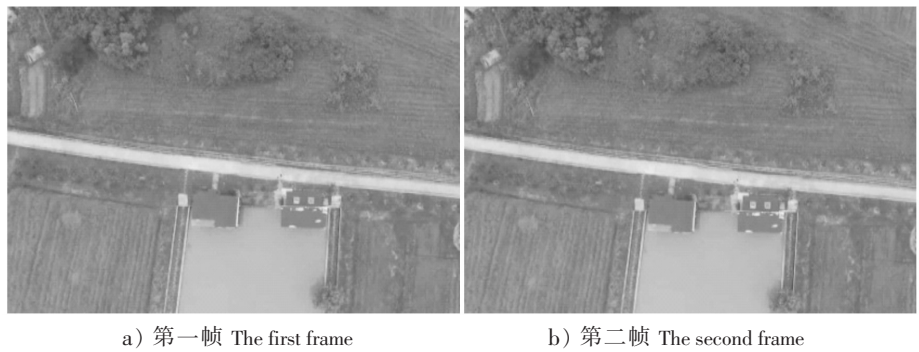


图 5 第三段视频中的连续两帧

Fig.5 Two consecutive frames in the third period of video

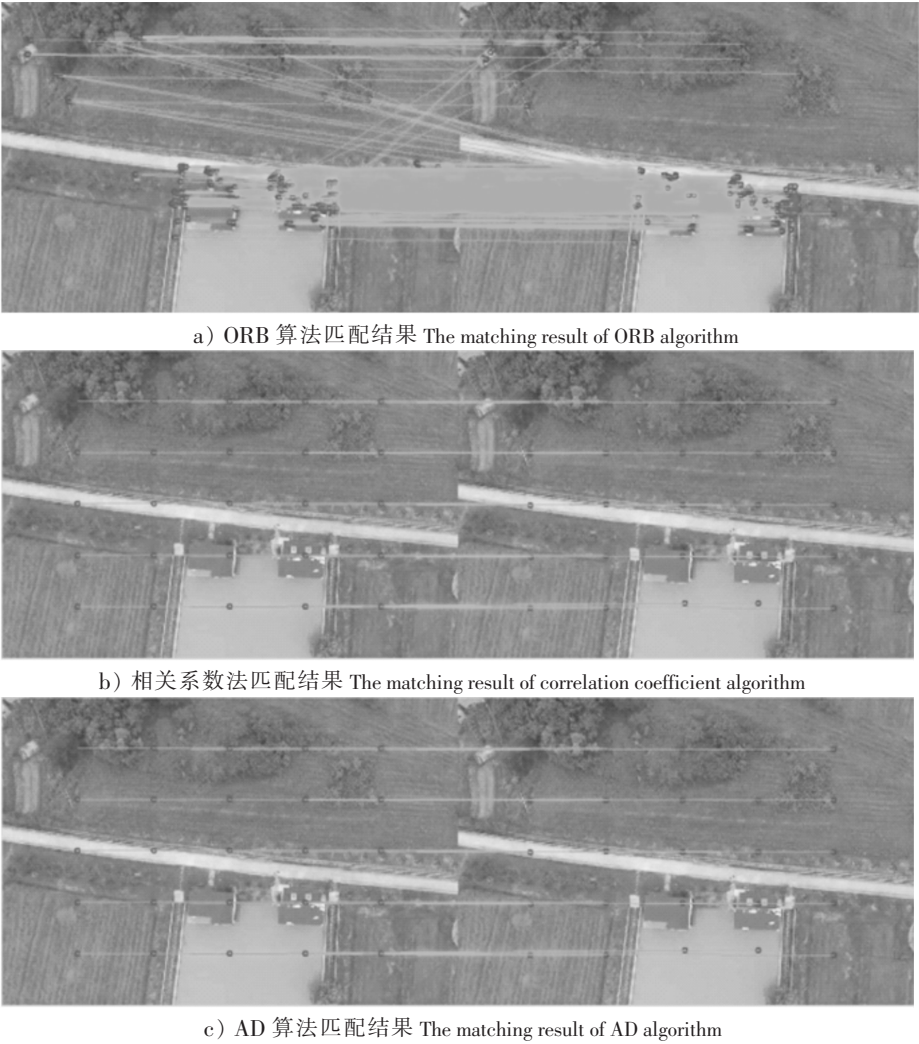


图 6 第三组匹配结果

Fig.6 The third group of matching results

经过实验结果对比,三组实验中 ORB 算法、相关系数法和 AD 算法的匹配时间分别为 431、85、65 ms,387、72、59 ms,398、78、60 ms,因此可得 AD 算法复杂度最低,时间消耗最小。但是当地物信息以水面、草坪为主时,该区域纹理信息匮乏,相关系数法和 AD 算法均告失败。因此,当待处理图像纹理信息丰富时,AD 算法效率最高,且可以满足可靠性要求。

1.2 预测模型

视频由多幅图像按时间序列排列组成,其成像可以视为一个大视差逐步移入的过程。当地物高度较低,不考虑高度对成像过程产生的影响,此时像点坐标在连续多帧中的变换接近于线性变换。但由于在匹配过程中必须考虑地面建筑物的高度对像点坐标变化的影响,即不能默认所有情况下像点坐标都是线性的。

考虑到一些地物高度会对成像位置产生影响,因此像点坐标的变化也是非线性的。同时除了投影变化,摄影角在拍摄过程中会不可避免地出现微小抖动,产生倾斜误差,而摄影角的变化无法得知,这进一步增大了像点坐标变化的非线性性。

由于影响地物像点坐标变化的因素较多且一些数据不容易获取,如摄影角度的变化,无法计算像点坐标变化函数等,为了完成匹配工作,扩大搜索窗口和匹配窗口是一种可行性方法,但是这样增加了算法的时间复杂度。为了降低搜索窗口和匹配窗口大小,本文引入 PID 理论建立动态预测模型,通过前几帧地物的像点坐标的变化来预测该地物在下一帧中的像点坐标。

1.2.1 PID 预测模型的建立

绝对差值算法计算精度只能达到像素级,因此本文所建立的 PID 控制器是离散形式的。PID 系统的输入为控制偏差:  $e(t) = r(t) - y(t)$ ,输出是比例  $P$ 、积分  $I$  和微分  $D$  的线性组合:  $u(t) = K_p e(t) + K_i \sum_{i=1}^k e(i) + K_d \Delta e(t)$ ,其中  $t$  是 PID 系统中输入帧的序号;  $e(t)$  是像元在  $x$  轴方向的位移偏差;  $r(t)$  是像元在  $x$  轴方向的预测位移;  $y(t)$  是像元在  $x$  轴方向的实际位移;  $u(t)$  是 PID 控制器输出的值,即像元在  $x$  轴的最终预测位移;  $K_p$  是比例系数;  $K_i$  是积分系数;  $T_d$  是微分系数;  $\Delta e(t)$  是采样周期内的微分。

引入 PID 算法的目的是预测同一像元在相邻帧中的位移,单位是像素,方向沿  $x$  轴和  $y$  轴两个方向。在预测过程中可以认为  $x$  轴方向和  $y$  轴方向的位移是相互独立的,因此本文中仅以  $x$  轴方向的预测位移为例说明,  $y$  轴方向同理可得。

1.2.2 PID 预测模型参数的设置

本文采用理论计算整定法确定控制系统参数,再通过实验验证。  
为了保证系统的稳定性,参数整定原则是按比例系数从小到大,执行周期从大到小调节。由于模型建立背景是离散系统,因此令比例系数取值如下:

$$K_p \in \{0.1 \quad 0.2 \quad 0.3 \quad 0.4 \quad 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1.0\}$$
  
在参数整定过程中首先假设只存在比例项,针对  $K_p$  调节。在实际计算中如果比例项  $P = K_p \times e(t)$  的值不是整数,则四舍五入取整。通过数据分析,同一像元在相邻帧中沿  $x$  轴方向的位移约为 10 个像素,在此基础上令初始预测值为 0,选择  $K_p = 0.3$  时可以使得系统保持稳定,同时保持足够的调节力度。具体过程见表 1。当  $K_p = 0.5$  时,虽然误差  $e(t)$  迅速降低到 0,但是实际应用中像元不可能总是保持同样的运动轨迹变化,此时容易造成过调,导致系统的不稳定。当  $K_p = 0.3$  和  $K_p = 0.4$  时,均可以使系统的稳态误差为 1 像素且性能接近。但是,  $K_p$  值越小,系统越稳定,故本文选择  $K_p = 0.3$ 。  
由表 1 知理论上系统的稳态误差约为 1,像元位移的预测值约为 10,故一个执行周期选择 10,即执行 10 次匹配,使  $u(t) = \sum_{i=1}^{10} e(t)$ ,令  $K_i = 0.1$ ,恰好可以消除系统的稳态误差。当经过比例项和积分项调节之后,系统的偏差可以快速地进入稳定状态,偏差变化缓慢,此时微分接近 0,微分项无意义,因此本文仅仅使用比例项和积分项建立预测模型。



表 1 比例项系数  $K_p$  整定过程  
Tab.1 Proportional coefficient  $K_p$  setting process

$K_p$	$e(1)$	$e(2)$	$e(3)$	$e(4)$	$e(5)$	$e(6)$	$e(7)$
0.1	9	8	7	6	5	4	4
0.2	8	6	5	4	3	2	2
0.3	7	5	3	2	1	1	1
0.4	6	4	2	1	1	1	1
0.5	5	2	1	0	0	0	0

1. 2. 3 内插模型的建立

结合相邻帧之间的特点，进一步假设相邻帧之间不存在旋转、几何畸变、光照差异等情况，即相邻两帧是平移变换。首先使用相关系数算法计算两帧之间一定数量的同名点，为了保证可靠性，对应同名点之间的相关系数值要足够高。因为两帧之间是平移变换，所以图像中其他点的位移可以使用已知点位移的线性组合表示。

当地物信息主要是建筑物等存在大量角点及边缘时，对应同名点的相关系数值较高，精确匹配时可以达到 0.95，故选取相关系数阈值为 0.95。当计算所得同名点之间的相关系数值大于阈值时可以考虑作为表示内插点的基准点。一幅图像中任意一点可以由不在同一直线上的三个点表示，基准点的最低数量是 3。为了保证算法的高效性，基准点数量不能太多。本文选取待匹配图像拐角处的 4 个点和图像中心附近的 4 个点，通过计算如果符合以上标准的则选作基准点，最终能选择 3~8 个点作为基准点，如果符合要求的点少于 3 个，则在有明显角点位置继续标记，如图 7 所示。

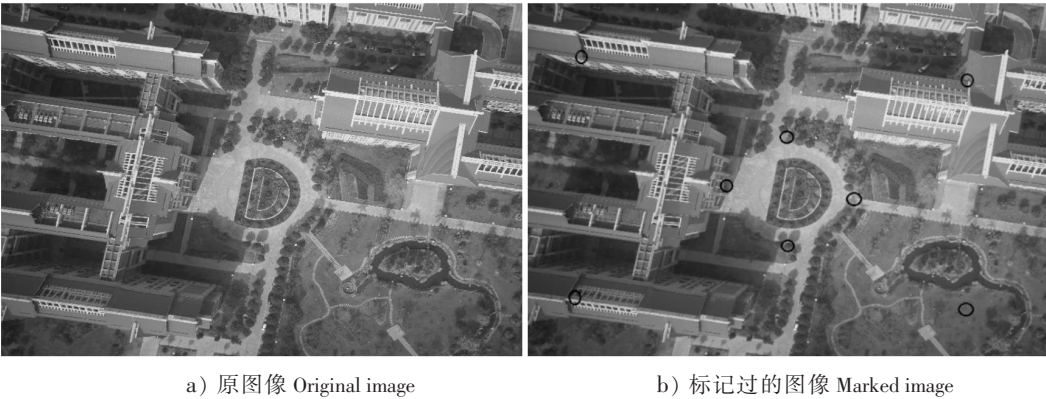


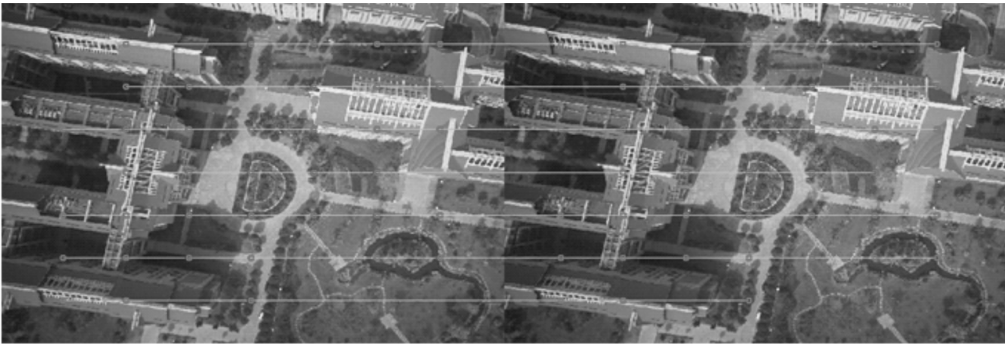
图 7 基准点的标记  
Fig.7 Datum mark

图 7a 中标记的 8 个点经过计算符合要求的有 5 个，图 7b 中任意一个点都可以由计算出的点的坐标线性表示，即  $(\Delta x, \Delta y) = \sum_{i=1}^n (a_{i1}, a_{i2}) \begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix}, n = 5$ ，其中  $(\Delta x, \Delta y)$  是图 7b 中的任意一点的预测位移， $a_{i1}, a_{i2}$  是线性表达的系数， $(\Delta x_i, \Delta y_i)$  是基准点的位移。因为两帧之间只存在平移变换，本文取基准点的平均位移作为对图像任一点位移的预测值，即  $a_{i1} = a_{i2} = 1/n$ 。

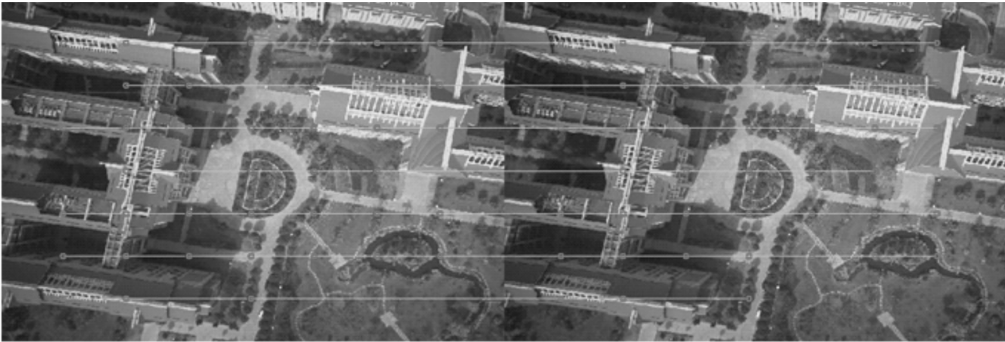
与 PID 预测模型相比，使用内插法预测同名点的位置只需要精确计算少量基准点的位置，不用逐点预测，当需要大量同名点时在时间消耗上其具有较大优势。

2 实验结果

在绝对差值算法过程中建立 PID 预测模型，可以根据视频帧之间像元的运动规律预测同一像元在不同帧中可能的位置，缩小了模板匹配的搜索窗口和匹配窗口大小，有效提高了算法的计算效率。使用 PID + AD 算法和内插 + AD 算法对待匹配数据进行匹配，两种算法的匹配结果（见图 8、图 9）非常接近，可靠性相似。



a) PID+AD 算法匹配结果 PID+AD algorithm matching results



b) 内插+AD 算法匹配结果 Interpolation+AD algorithm matching results

图 8 第一组实验匹配结果

Fig.8 The first group experiments of matching results



a) PID+AD 算法匹配结果 PID+AD algorithm matching results



b) 内插+AD 算法匹配结果 Interpolation+AD algorithm matching results

图 9 第二组实验匹配结果

Fig.9 The first group experiments of matching results



对比图 8、图 9 可知两种算法匹配效果接近，每一对同名点所在区域的相关系数值均大于 0.9，但是图 8 所对应算法的时间消耗显著小于图 2c 所对应的模板匹配算法，且其中内插 + AD 算法的时间消耗小于 PID + AD 算法，三种算法的参数如表 2 所示。因此，本文选用内插法作为预测模型。

表 2 三种算法的参数对比  
Tab.2 The contrast for parameters of the three algorithms

算法 Algorithm	搜索窗口 Searching window	匹配窗口 Matching window	时间消耗 Time cost/ms
AD	31 * 31	11 * 11	85
PID + AD	11 * 11	5 * 5	52
内插 Interpolation + AD	11 * 11	5 * 5	43

3 结论

基于灰度的匹配算法主要是通过计算搜索窗口内每个像元的相关系数值进行匹配，其算法的时间复杂度比基于特征的匹配算法低。当使用基于灰度的匹配算法得到大量同名点时会相应提高算法的可靠性，而时间消耗呈线性增加，最终可以在保证可靠性的前提下取得较高的匹配效率。经过实验对比，AD 算法的计算效率最高，故本文选择 AD 算法用于相邻帧匹配。本文建立的内插预测模型缩小了 AD 算法过程中匹配窗口和搜索窗口的大小，提高了计算效率。因此，内插 + AD 算法可以实现无人机视频的快速匹配。

[ 参 考 文 献 ]

[1] HAITAO X, LEI T. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosystems Engineering*, 2010, 108(2): 174-190. DOI:10.1016/j.biosystemseng.2010.11.010.

[2] ZHANG C J, LIU J, TIAN Q. Beyond visual features: a weak semantic image representation using exemplar classifiers for classification. *Neuro Computing*, 2013, 120: 318-324. DOI:10.1016/j.neucom.2012.07.056.

[3] PUAL E, ULRICH S. Cerebellar grey matter deficits in first-episode schizophrenia mapped using cortical pattern matching. *Neuro Image*, 2010, 53(4): 1175-1180. DOI:10.1016/j.neuroimage.2010.07.018.

[4] CHOI O, KWEON I. Robust feature point matching by preserving local geometric consistency. *Computer Vision and Image Understanding*, 2009, 113(6): 736-742. DOI:10.1016/j.cviu.2008.12.002.

[5] 唐璘, 李青. 一种快速的模板匹配算法. *计算机应用*, 2010, 30(6): 1559-1561, 1564.

[6] 高军, 李学伟, 张建, 等. 基于模板匹配的图像配准算法. *西安交通大学学报*, 2007, 41(3): 307-311.

[7] 陈皓, 马彩文, 陈岳承, 等. 基于灰度统计的快速模板匹配算法. *光子学报*, 2009, 38(6): 1586-1590.

[8] BELLO E D, SALVADEO P A. An image descriptors extraction hardware-architecture inspired on human retina//IX Southern Conference on Programmable Logic. Buenos Aires: IEEE, 2014: 1-6. DOI:10.1109/SPL.2014.7002205.

[9] 何芝强. PID 控制器参数整定方法及其应用研究. 杭州: 浙江大学, 2005.

[10] 李晓丹. 模糊 PID 控制器的设计研究. 天津: 天津大学, 2005.

[11] 孙远, 周刚慧, 赵立初, 等. 灰度图像匹配的快速算法. *上海交通大学学报*, 2000, 34(5): 702-704.

[12] 刘锦峰. 图像模板匹配快速算法研究. 长沙: 中南大学, 2007.

[13] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection// *Computer Vision-ECCV 2006*. Berlin: Springer, 2006: 430-443. DOI:10.1007/11744023\_34.

[14] MICHAEL CALONDER, VIUCENT LEPETIT, CHRISTOPH STRECHA, et al. Brief: binary robust independent elementary features//*Computer Vision-ECCV 2010*. Heraklion, Greece: Springer, 2010. 6314: 778-792. DOI:10.1007/978-3-642-15561-1\_56.

(责任编辑 朱雪莲 英文审校 黄振坤)