

应用于稳像系统中的改进梯度光流法

赖丽君, 徐智勇, 张栩铤

(中国科学院光电技术研究所, 四川 成都 610209)

摘要: 由于传统的梯度光流法当运动不连续时运动场估计值与真实值有较大偏差, 因而不能直接应用于稳像系统中。引入金字塔多分辨率分层技术对传统的梯度光流法进行改进。首先, 在视频序列中选定细节丰富的区域作为计算区域; 其次, 利用结合金字塔多分辨率分层技术的光流法迭代求解相邻帧间的仿射参数; 最后, 采用帧间补偿方式, 并增加控制累积错误传播的措施, 在不丢失过多原有信息基础上, 实现了长时间稳像。实验表明: 改进的方法能够检测出剧烈的复杂抖动, 并能达到旋转精度小于 0.09° 、平移精度小于 0.07 个像素, 缩放精度小于 0.02 的高精度估计, 且补偿序列平均峰值信噪比值提高了 2.36 dB 以上。

关键词: 光流法; 稳像; 金字塔; 仿射变换

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA201645.0428004

Improved gradient optical flow for digital image stabilization

Lai Lijun, Xu Zhiyong, Zhang Xuyao

(Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China)

Abstract: Traditional gradient optical flow has large motion estimation deviation with real value, thus cannot be directly applied to image stabilization system. In order to improve traditional gradient optical flow, a pyramid multi-resolution coarse-to-fine search strategy was incorporated into this algorithm. Firstly, computing area affine transform parameters were selected as the final transform parameters. Then, in the compensation period, error control propagation method was selected to obtain long term stabilized sequence. The experiment results show the improved method can detect severe complex jitter, and can achieve the rotating precision less than 0.09° , translation precision less than 0.07, scaling precision less than 0.02, the compensation sequences' average PSNR raised 2.36 dB.

Key words: optical flow; stabilized sequence; pyramid; affine transformation

收稿日期: 2015-08-11; 修订日期: 2015-09-20

作者简介: 赖丽君(1989-), 女, 硕士生, 主要从事电子稳像技术方面的研究。Email: lailj864@163.com

导师简介: 徐智勇(1970-), 男, 研究员, 硕士生导师, 主要从事数字图像处理方面的研究。Email: xzy158@163.com

0 引言

电子稳像技术旨在利用数字图像处理的方法直接确定视频序列间的偏移量并进行补偿,减轻或消除视频序列中的随机抖动从而得到稳定、清晰的视频画面^[1]。

通常采用块匹配算法^[2-3],代表点匹配算法^[4],相位相关算法^[5]以及灰度投影算法^[6]来处理仅含平移抖动的视频序列。然而研究表明,人眼对旋转非常敏感,因此在稳像系统中准确获得抖动视频的旋转量具有重要意义。针对包含平移,缩放,旋转的复杂抖动视频序列,常采用特征提取的方法如 Sift 算法^[7-8], Harris 算法^[9], Surf 算法^[10],这些算法有一定的精度和鲁棒性,但是当提取的特征过于集中于局部区域时,误匹配率高,难以获得准确的序列间变换参数。对于帧间运动偏移量小于 1 个像素的缓慢运动传统的梯度光流法^[11-12]可以精确计算出仿射变换参数,但当运动偏移量超过 1 个像素时计算的变换参数具有较大误差,限制了其工程应用。文中针对此问题,在光流计算中引入了多分辨率金字塔分层技术,使其能够应用于电子稳像系统中。

1 梯度的光流法

Horn^[13]在图像灰度一致性假设基础上导出了光流场计算基本公式如公式(1)。认为在相邻连续两帧图像上,对应像素灰度值变化很小,可以认为是相等的,即

$$f(x+dx, y+dy, t+dt)=f(x, y, t) \quad (1)$$

式中: $f(x, y, t)$ 为图像点 (x, y) 在 t 时刻的灰度; $f(x+dx, y+dy, t+dt)$ 为像素点 (x, y) 运动到 $(x+dx, y+dy)$ 处的灰度。基于图像灰度一致性,假设有:

$$\frac{\partial f(x+dx, y+dy, t+dt)}{\partial t}=0 \quad (2)$$

对公式(2)进行泰勒级数展开忽略 2 次以上高阶项并简化可以得到光流场计算基本公式如公式(3),即

$$\frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t} = 0 \quad (3)$$

式中: $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial t}$ 分别为图像亮度在水平方向、垂直方向和时间上的梯度; dx/dt 和 dy/dt 分别为水平和垂直方向的光流。

根据光流场灰度一致性假设条件可知该方法仅适用于运动连续的图像序列。而实际工程中视频序列很难满足运动连续性假设,这大大限制了梯度光流法在工程中的应用。文中将图像金字塔多分辨率分层技术引入光流法中,解决了视频序列间运动不连续时光流场计算误差大的缺点。

2 基于金字塔多分辨率光流算法的实现

2.1 运动模型参数求解

仿射模型适用于描述大部分室内外场景,不仅可以同时描述平移、旋转以及缩放三种形式,还能描述图像因不按中心旋转而引起的畸变。文中选择视频序列中当前帧的前一帧为参考帧,模型的公式描述如下所示:

$$\begin{pmatrix} X_i' \\ Y_i' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \end{pmatrix} + \begin{pmatrix} a_5 \\ a_6 \end{pmatrix} \quad (4)$$

式中: (X_i', Y_i') 和 (X_i, Y_i) 分别表示当前帧和参考帧的坐标,这里不妨设

$$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, T = \begin{pmatrix} a_5 \\ a_6 \end{pmatrix} \quad (5)$$

式中: A 包含了两帧之间的旋转、缩放以及畸变信息; T 包含了平移信息。通过求取矩阵 A 和 T 即可得到相邻两帧的仿射变换参数。

当两帧满足仿射变换关系时,图像灰度绝对误差最小。因此文中通过相邻两帧图像 $f(x, y, t)$ 和 $f(x, y, t-1)$ 间的最小误差函数来求得仿射变换参数,最小误差函数用 E 来表示即为:

$$E(a) = \sum_{x,y \in \Omega} [f(x, y, t) - f(a_1x+a_2y+a_5, a_3x+a_4y+a_6, t-1)]^2 \quad (6)$$

将公式(6)进行泰勒级数近似展开并化简可以得到:

$$E(a) \approx \sum_{x,y \in \Omega} [f_t - (a_1x+a_2y+a_5-x)f_x - (a_3x+a_4y+a_6-y)f_y]^2 = \sum_{x,y \in \Omega} [s - m^T a]^2 \quad (7)$$

式中: Ω 为图像计算区域,当误差 $E(a)$ 趋向于 0 时,仿射参数最为理想。其中 $a^T=(a_1, a_2, a_3, a_4, a_5, a_6)$, $s=f_t+xf_x+yf_y$, $m^T=(xf_x, yf_y, xf_y, yf_x, f_x, f_y, f_t)$ 分别表示图像强度 $f(x, y, t)$ 对 x, y, t 的偏导。对公式(7)进行差分处理可以得到:

$$\frac{dE(a)}{da} = \sum_{\Omega} 2[s - m^T a](-m^T) \quad (8)$$

令公式(8)中 $\frac{dE(a)}{da}=0$ 即可求得误差函数的极小值,

$$\text{可得 } a = \left[\sum_{\Omega} m^T \right]^{-1} \left[\sum_{\Omega} sm^T \right] \left[\sum_{\Omega} m^T \right]^{-1}.$$

从公式(8)中可以看出,要获得仿射变换参数的解,必须先求得图像强度 $f(x, y, t)$ 对 x, y, t 的偏导 f_x, f_y, f_t 。在数字图像处理中,常用差分的方法求梯度,这里采用前向有限差分和后向有限差分取平均来近似计算图像的梯度,即

$$\begin{aligned} f_x &= \frac{1}{4} \{f(x+1, y, t) - f(x, y, t) + f(x+1, y+1, t) - \\ & \quad f(x, y+1, t) + f(x+1, y, t-1) - f(x, y, t-1) + \\ & \quad f(x+1, y+1, t-1) - f(x, y+1, t-1)\} \\ f_y &= \frac{1}{4} \{f(x, y+1, t) - f(x, y, t) + f(x+1, y+1, t) - \\ & \quad f(x+1, y, t) + f(x, y+1, t-1) - f(x, y, t-1) + \\ & \quad f(x+1, y+1, t-1) - f(x+1, y, t-1)\} \\ f_t &= \frac{1}{4} \{f(x, y+1, t) - f(x, y+1, t-1) + f(x, y, t) - \\ & \quad f(x, y, t-1) + f(x+1, y, t-1) - f(x+1, y, t-1) + \\ & \quad f(x+1, y+1, t) - f(x+1, y+1, t-1)\} \end{aligned} \quad (9)$$

直接将得到的 f_x, f_y, f_t 代入公式(8),当不满足光流场计算假设条件时,帧间运动矢量的估计值与实际值会有较大的偏差,因此文中引入了金字塔多分辨率分层技术。

2.2 金字塔多分辨率分层技术迭代求解运动参数

对每帧图像构建 L 层金字塔模型。将原始图像看作第 0 层,第 $L+1$ 层由第 L 层图像平滑后下采样而来。层越高图像分辨率越低,其运动矢量估计区域和搜索范围也随着层数的增加而成倍减少。下采样后顶层满足运动连续性假设。

应用梯度光流法计算出顶层间的仿射变换参数,根据得到的变换关系对参考帧进行插值处理,形成新的参考帧,再一次对新的参考帧和当前帧进行 $L-1$ 层金字塔分解,并在顶层进行运动矢量估计,直至 L 为 0。假设第 k 次分层的运动矢量估计值为 A_k, T_k ,第 $k+1$ 次的分层运动矢量估计值为 A_{k+1}, T_{k+1} 。则相邻两次分层估计值的变换关系满足:

$$p_2^k = A_k p_1^k + T_k = A_k (A_{k+1} p_1^{k+1} + T_{k+1}) + T_k = A_k A_{k+1} p_1^{k+1} + T_k \quad (10)$$

式中: p_1^k 表示参考帧(序列中当前帧的前一帧)第 k 次分层顶层某点的坐标,其对应应在当前帧第 k 次分层

顶层的坐标为 p_2^k ,其中 $A=A_k A_{k+1}, T=A_k T_{k+1} + T_k$ 是以低分辨率运动矢量为初值得到的高分辨率层的运动参数。

算法实现:

(1) 设定金字塔层数 $L=4$ 和计算区域。

(2) 对相邻两帧同时进行 L 层下采样。用公式(9)在顶层计算梯度 f_x, f_y 和 f_t ,并将求得的梯度值代入公式(8),求取参数 A, T ,更换 $T=T*2$ 。

(3) 对参数进行迭代计算 $A_{cum}=A*A_{cum}, T_{cum}=AT_{cum}+T$,当第一次计算时取迭代设初值为 $A_{cum} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, T_{cum} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$,如果 $L \geq 0$,用 A_{cum}, T_{cum} 对原始的参考帧进行插值处理形成新的参考帧,同时更新 $L=L-1$ 并转至步骤 2;若 $L < 0$,则输出最终参数 A_{cum} 和 T_{cum} 。采用金字塔多分辨率分层技术的仿射参数估计算法流程图如图 1 所示。

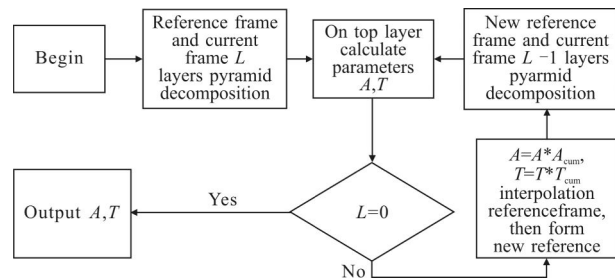


图 1 金字塔参数求解流程图

Fig.1 Flow chart of pyramid parameters solving

3 运动补偿

运动补偿方式有逐帧补偿和固定帧补偿两种。固定帧补偿只选用固定的某一帧作参考帧容易产生边缘发散^[6]。文中选用逐帧补偿的方式。

在逐帧补偿的过程中,后一帧是由前一稳定帧补偿得到,在补偿过程中出现的错误会向后序帧传播,带来累积误差,随着累积误差的增大,原始序列信息丢失严重,最终会导致失稳。因此文中在运动补偿的计算中增加了控制累积误差传播的方法,补偿过程描述如下:

(1) 计算原始序列和稳定序列累积和的均方误差,并设定阈值 threshold,为了不丢失过多原有图像信息,阈值一般选择为图像高度、宽度中较小者的 1/12,补偿公式如下:

$$\left| \sum_{j=1}^{i-1} u_{j,original} - \sum_{j=1}^{i-1} u_{j,stabilized} \right|^2 + \left| \sum_{j=1}^{i-1} v_{j,original} - \sum_{j=1}^{i-1} v_{j,stabilized} \right|^2 \leq threshold^2 \quad (11)$$

式中： $u_{j,original}$, $u_{j,stabilized}$, $v_{j,original}$, $v_{j,stabilized}$ 分别为稳定序列水平方向矢量,原始序列水平方向矢量,原始序列垂直方向矢量,稳定序列垂直方向矢量。

(2) 如果均方误差小于 $threshold^2$, 那么就用原始的第 k 帧补偿得到稳定的第 $k+1$ 帧。

(3) 如果均方误差大于 $threshold^2$, 那么就用稳定的第 k 帧补偿得到稳定的第 $k+1$ 帧。

从公式(11)可以看出:采用小的阈值,稳定序列不会出现突然的跳变,采用大的阈值,会较多地利用原始帧信息,累积误差可以得到有效的控制。

4 实验结果与分析

4.1 改进方法运动参数估计效果测试

为了测试改进方法对旋转、偏移、缩放运动估计的效果,文中从视频序列中抽出一帧并对其进行一系列的旋转、偏移、缩放变换,用改进的方法进行测量,并比较测量值与真实值,确定改进算法的运动估计精度和范围。由于序列中无前景物体干扰,全局运动参数近似等于局部运动参数,为了降低计算量,划定中间细节丰富区域作为计算区域如图 2 所示。实验结果如表 1~3 所示。



图 2 运动参数求解的区域划定

Fig.2 Zoning of the motion parameters

表 1 旋转测试对比

Tab.1 Rotating test contrast

Real rotation/(°)	Improved measure / (°)	Traditional measure/(°)
1	1.04	0.74
5	5.09	1.78
10	9.96	2.47

表 2 平移测试对比

Tab.2 Transition test contrast

Real transition (X,Y)/pixel	Improved measure (X,Y)/pixel	Traditional measure (X,Y)/pixel
(1, 1)	(1.01, 1.00)	(0.93, 0.94)
(15, 15)	(14.97, 15.01)	(0.31, 0.65)
(20, 20)	(19.93, 20.03)	(0.27, 1.23)

表 3 缩放测试对比

Tab.3 Scaling test contrast

Real scale/times	Improved measure /times	Traditional measure /times
1.5	1.50	1.00
2	1.98	1.38
2.5	1.78	1.04

表 1 所示为采用改进的算法和传统的梯度光流法算法对旋转变换图像进行算法效果测试的结果对比,可以看出:当真实的旋转量小于 1° 时,图像还具有连续性,传统算法可以较为准确地检测出实际的旋转量;随着旋转量逐渐增大,图像的不连续性越来越强,传统算法的误差也越来越大。而文中改进的算法可以准确地检测出旋转量小于等于 10° 的图像变换,精度能达到 0.09° 。

同表 1 的分析类似,从表 2 和表 3 中可以看出:文中的方法能检测出平移量小于等于 20 个像素,缩放小于 2 倍的图像间变换,且均能达到 0.07 个平移像素和 0.02 倍缩放的高精度估计。改进的方法在适用范围和精度上较传统的梯度光流法都有很大的提高。在稳像系统中,运动估计的估计范围以及精确度直接影响稳像系统的性能,文中的方法能检测出大范围的旋转、平移、缩放的复杂抖动,且满足稳像系统的精度要求。

将改进的方法应用于图 2 所在的真实序列中,图 3 和图 4 分别表示图 2 所表示的视频序列在补偿前后的光流场,从补偿前后的对比可以看到原始序列中无序不规则的运动得到了很好的减弱,实验证明改进的方法在真实的图像序列中依然具有有效性。

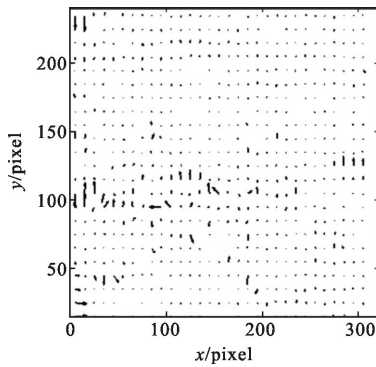


图 3 补偿前的光流场

Fig.3 Optical flow field before compensation

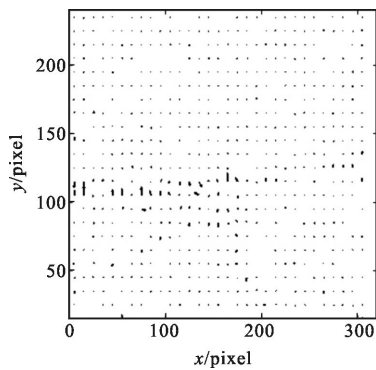


图 4 补偿后的光流场

Fig.4 Optical flow field after compensation

4.2 稳像有效性评估与分析

前文中对改进的光流法算法进行了有效性测试,这里将该算法应用于稳像系统,对真实的抖动视频进行稳像。图 5 所示为选取的四组不包含前景运

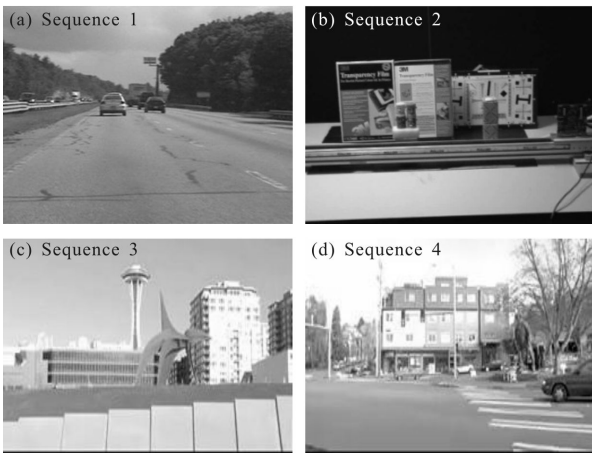


图 5 视频序列

Fig.5 Video sequence

动干扰的视频。从序列 1 中取出前 130 帧图像并绘制其补偿前后各个方向的运动矢量图,图 6 为水平

方向运动的矢量图,图 7 为垂直方向运动的矢量图,图 8 为旋转运动的角度图。

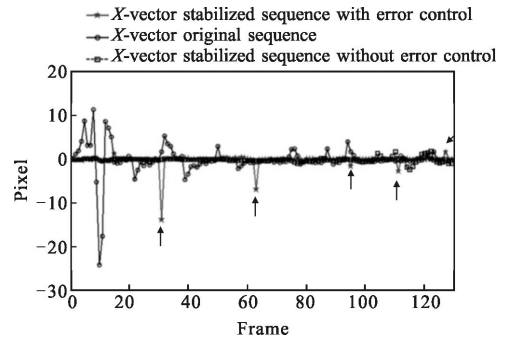


图 6 序列 1 水平方向运动矢量

Fig.6 Horizontal motion vector of sequence 1

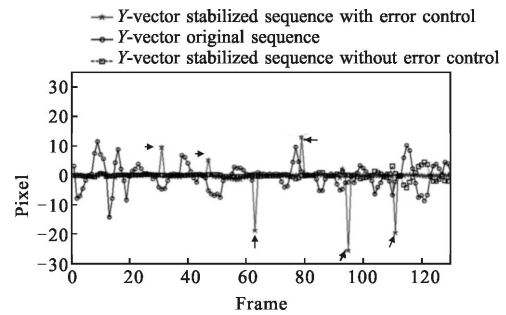


图 7 序列 1 垂直方向运动矢量。

Fig.7 Vertical motion vector of sequence 1

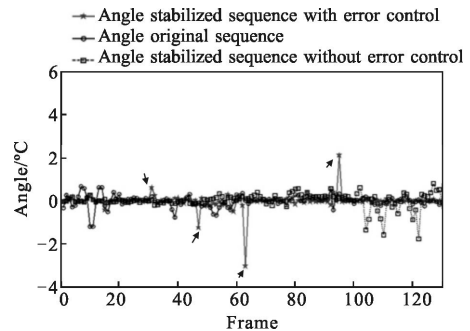


图 8 序列 1 旋转运动矢量

Fig.8 Rotational motion vector of sequence 1

从图 6~8 可以看出原始序列具有较大的平移抖动,而补偿后的图像序列有效地减弱了抖动。图中黑色箭头标注出了运动修正的地方,这些地方引入了原始帧做补偿,在控制了累积误差传播的同时也引入了突然的偏移运动,在不影响视觉效果的前提下,保证了原始帧补偿后得到的结果序列与原始序列不会出现太大的偏差,达到了持续稳像的目的。从 100 帧以后的运动矢量数据可以看出,未加入累积误差控制的补偿序列随着稳定帧数的增多,波动越来越厉

害,最终会导致失稳。

峰值信噪比(PSNR)表征了两幅图像的相似度,相似度越高,其峰值信噪比值越大,稳像效果越好。这里采用 PSNR 对文中算法进行客观评价。考虑到摄像机的运动会产生边缘信息的丢失,因此将图像边缘剪切后进行峰值信噪比计算,其计算公式如下:

$$P_{SNR}(f_i, f_{i-1}) = 10 \lg \frac{255}{MSE(f_i, f_{i-1})} \quad (12)$$

式中: MSE 为图像 f_i 和 f_{i-1} 间像素的均方误差,反应了图像间的差异大小。

根据公式(12)计算序列 1 的峰值信噪比如图 9 所示。从图中可以很直观看出:经过稳像处理后的视频序列的峰值信噪比比原始序列有显著提高。图中黑色箭头所指的位置运动修正的地方,与运动矢量图一致。加入原始帧进行补偿由于引入了突然的偏移运动,会引起相邻两帧图像出现突然的跳变,两帧图像的相似度降低,峰值信噪比也会降低。但从整体上来看,加入累积错误控制的视频序列其峰值信噪比比原始序列有很大提升,从 100 帧以后可以看出,加入累积错误控制后,能长时间保持高信噪比值。

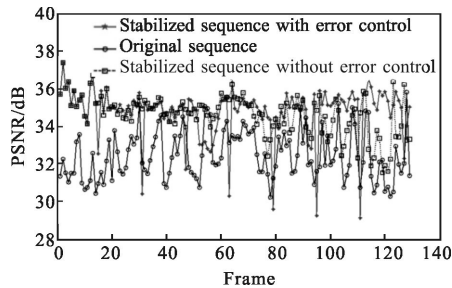


图 9 序列 1 峰值信噪比曲线
Fig.9 PSNR curve of sequence 1

对图 5 中的序列 2~4 进行了仿真实验,并计算了其平均峰值信噪比,实验结果如表 4 所示。可以看出:

表 4 四组序列的平均峰值信噪比计算结果

Tab.4 Average PSNR results of four sequences

Seq.	Motion feature	PSNR/dB			Gain
		Origin	With control	Without control	
1	Translation and small rotation	32.45	34.81	34.66	2.36
2	Translation and large rotation	32.59	35.06	35.01	2.47
3	Large rotation and translation	30.33	32.70	32.52	2.37
4	Small rotation and translation	28.12	30.52	30.51	2.40

文中的稳像算法对包含旋转、平移、缩放的视频序列均有效,其峰值信噪比均提高了 2.36 dB 以上。

5 结论

文中将金字塔多分辨分层技术引入传统的梯度光流法中,克服了当视频帧间运动不连续时运动参数估计误差大的缺点;采用迭代求解的方式,减少了搜索范围,加快了计算速度;采用增加累积错误控制措施的帧间补偿方式,避免了帧间补偿无法长时间稳像的缺点。文中对多种不同特征的图像视频进行了仿真验证,实验结果表明:改进的方法可以准确的检测出旋转量度,平移量个像素,缩放小于 2 倍的图像间变换,旋转精度达到 0.09° ,平移精度达到 0.07 个像素,缩放精度达到 0.02。对实际抖动的视频序列稳像处理后平均 PSNR 值提高了 2.36 dB 以上。

参考文献:

- [1] Wang Zhiming, Xu Xiaogang. A survey on electronic image stabilization [J]. *Journal of Image and Graphics*, 2010, 15 (3): 470-480. (in Chinese)
王志民, 徐晓刚. 电子稳像技术综述 [J]. 中国图像图形学报, 2010, 15(3): 470-480.
- [2] Zhu Juanjuan, Fan Jing, Guo Baolong. Adaptive electronic image stabilization algorithm resistant to foreground moving object [J]. *Acta Photonica Sinica*, 2015, 44 (6): 45-52. (in Chinese)
朱娟娟, 范静, 郭宝龙. 抗前景干扰的自适应电子稳像算法[J]. 光子学报, 2015, 44(6): 45-52.
- [3] Liu K X, Qian J, Yang R K. Block matching algorithm based on RANSAC algorithm [C]//2010 International Conference on IASP, 2010: 400-405.
- [4] Zhong Ping. Application of improved representative point match algorithm in electronic image stabilization [J]. *Optical Technique*, 2005, 31(5): 742-745. (in Chinese)
钟平. 一种改进的代表点匹配算法在稳像技术中的应用 [J]. 光学技术, 2005, 31(5): 742-745.
- [5] Sanjeev Kumar, Haleh Azartash, Mainak Biswas. Real-time affine global motion estimation using phase correlation and its application for digital image stabilization [J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3406-3418.
- [6] Qi Hong, Ren Hong, Tian Shaoqing, et al. An improved

- electronic image stabilization method based on gray projection [J]. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2012, 2(5): 44–49. (in Chinese)
- 齐红, 任红娥, 田少卿, 等. 一种基于灰度投影的改进电子稳像方法[J]. 中山大学学报(自然科学版), 2012, 2(5): 44–49.
- [7] Battiato S, Gallo G, Puglisi G, et al. SIFT features tracking for video stabilization [C]//IEEE International Conference on Image Analysis and Processing, 2007: 825–830.
- [8] Lee K -Y, Yu Y, Chuang B -Y, et al. Video stabilization using robust feature trajectories [C]//IEEE 12th International Conference on Computer Vision, 2009: 1394–1404.
- [9] Zhou Yajun, Tao Shengxiang, Zhang shu, et al. Image stabilization algorithm based on Harris[J]. *Laser & Infrared*, 2009, 39(5): 555–557. (in Chinese)
- 周亚军, 陶声祥, 张舒, 等. 一种基于 Harris 角点特征的稳像算法[J]. 激光与红外, 2009, 39(5): 555–557.
- [10] Chen Xiaolu, Dai Ming, Ge Ming, et al. Fast smoothing point-feature trajectories for digital image stabilization [J]. *Infrared and Laser Engineering*, 2014, 43 (6): 1988–1993. (in Chinese)
- 陈晓露, 戴明, 葛明, 等. 快速平滑特征轨迹电子稳像[J]. 红外与激光工程, 2014, 43(6): 1988–1993.
- [11] Nishizaka Y, Yamaguchi T, Harada H. Optical flow estimation using compensation method [C]//ICCVS -SICE, Fukuoka, 2009: 2680–2684.
- [12] Chang Huangchang, Lai Shanghong, Lu Kuangrong. A robust real time video stabilization algorithm [C]//J vis Commun, ImageR -17, 2006: 659–673.
- [13] Horn B K P, Schunck B G. Determining optical flow [J]. *Artificial Intelligence*, 1981, 17(13): 185–203.