

## 一种新的图像序列失真模型：动态偏移场模型

邹谋炎<sup>①</sup> 刘艳<sup>①②</sup> 曹瑛<sup>①</sup> 卢晓鹏<sup>①</sup> 杜坤<sup>①</sup>

<sup>①</sup>(中国科学院电子学研究所 北京 100190)

<sup>②</sup>(中国科学院研究生院 北京 100039)

**摘要：**针对图像序列失真现象，该文提出一种新的图像序列失真模型：动态偏移场模型。该模型是一个充分参数化的模型，它为目前可观测到的多种失真现象提供了一个统一的描述框架。实验结果表明，动态偏移场模型能够简单地描述图像序列失真现象，并为失真校正工作提供了一个由简单到复杂渐进处理的研究平台。

**关键词：**图像序列失真；动态偏移场；抖动；失真校正；抖动抑制

中图分类号：TN911.73

文献标识码：A

文章编号：1009-5896(2008)09-2143-05

## Dynamic Displacement Field Model: A New Model for Describing Image Sequence Distortion

Zou Mou-yan<sup>①</sup> Liu Yan<sup>①②</sup> Cao Ying<sup>①</sup> Lu Xiao-peng<sup>①</sup> Du Kun<sup>①</sup>

<sup>①</sup>(Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China)

<sup>②</sup>(Graduate University of the Chinese Academy of Sciences, Beijing 100039, China)

**Abstract:** A Dynamic Displacement Field (DDF) model for describing image sequence distortion is proposed. The parameterized DDF model offers a unified framework that covers all the reported image sequence distortions. Experimental results show that the DDF model is a simple and effective model for describing image sequence distortion and provides a gradually upgraded and evolutionary process for image sequence distortion correction.

**Key words:** Image sequence distortion; Dynamic Displacement Field (DDF); Jitter; Distortion correction; Jitter suppression

### 1 引言

在工程应用中，由于多种因素引起的图像序列失真现象是一个十分棘手的问题。拥挤或运动环境中，由于拍摄环境的运动或不稳定因素，拍摄的视频图像序列出现各种各样的失真，如随机抖动、偏移、旋转、缩放、倾斜、变形等。高空摄影、水下摄影或通过其他非均匀介质表面摄影时，由于介质等外界环境因素的影响，拍摄图像序列也会出现各种变形失真现象。这些图像序列的失真严重影响了拍摄视频序列的质量。而对于视频监控和视频跟踪系统，由于拍摄平台的不稳定因素，经常会带来跟踪误差，在某些情况下甚至导致跟踪目标的丢失。因此，图像序列失真校正技术不仅对商业应用，而且在科学研究、工业和军事应用等方面都有着十分重要的意义。

近年来，对于图像序列失真的研究，人们提出了许多相关技术，如视频稳定化技术、视频跟踪技术、运动参数估计技术等<sup>[1-7]</sup>。针对特定问题通常需要采用特定的数学模型，基于合理的失真模型，对图像序列的失真进行合理的规律性描述，将更加有利于找出合适的失真校正方法。为此，研究

人员提出了许多失真模型<sup>[1-3]</sup>，而一种通用的失真模型对于工程应用来说是十分必要的。考虑到实际应用的需求，该通用模型应该具备以下几个特点：

(1)模型应能从简单到复杂、从低频到高频地描述各种图像序列的失真现象；

(2)为便于实现失真序列的校正，模型应该是充分参数化的；

(3)模型的描述应该在模型复杂性与描述精度之间取得适当折中。

基于以上几个特点，本文提出了动态偏移场模型作为一种新的图像序列失真模型，用来描述图像序列的失真现象。该模型是一个充分参数化的模型，它为图像序列失真提供了一个统一的描述框架，在此框架下动态偏移场模型可以描述目前可观测到的多种图像序列失真现象。考虑到工程应用中现有技术水平、经济条件和计算时间的限制，偏移场模型提供了一个由低阶到高阶逐步演进的描述方式，能够由简单到复杂地处理图像序列的失真问题。基于偏移场模型的失真校正技术，对于实现运动参数辨识和视频稳定化更加简单有效。实验结果表明，动态偏移场模型可以简单地描述图像序列的失真现象，并能应用到许多实际问题中，为图像序列的失真校正工作提供良好的研究平台。

## 2 动态偏移场模型的数学描述

动态偏移场模型描述了图像平面上每个像素点相对于其原始参考的动态偏移,基本思想可以通过图1的几个例子给予直观的阐释。

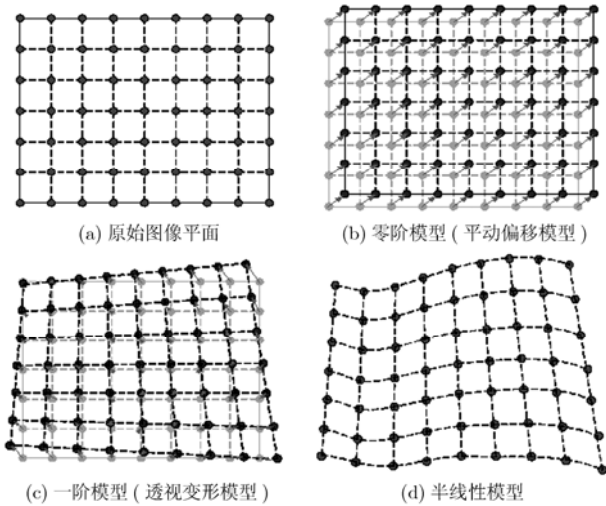


图1 动态偏移场模型示意图

根据动态偏移场模型的直观描述,下面给出偏移场模型的一般数学表达形式。设原始无失真图像  $A$  的尺寸为  $M \times N$ , 假定它的每个像素的位置在时间  $t$  由  $(i, j)$  偏移到  $(i', j')$ , 灰度值受到加性噪声  $\xi(i', j'; t)$  的扰动。由此,得到失真后的新图像  $B$ 。采用动态偏移场模型描述如下:

$$B(i', j'; t) = A(i, j; t) + \xi(i', j'; t) \quad (1)$$

$$i' = i + \sum_{k=1}^K x_k(i, j; t), \quad j' = j + \sum_{k=1}^K y_k(i, j; t), \quad i \in [1, M], \quad j \in [1, N] \quad (2)$$

其中变量  $t$  代表连续或离散时间量。 $x_k(i, j; t)$ ,  $y_k(i, j; t)$  分别是  $t$  时刻像素点  $A(i, j; t)$  在  $i$  方向和  $j$  方向的空间偏移量,它们既是时间序列,同时也是空间可变的。 $(x_k(i, j; t), y_k(i, j; t))$  表示一种偏移模式,式(2)中假定共有  $K$  种不同的偏移模式。

式(2)为复杂的动态失真提供了分解描述方式,其中下脚标  $k$  可以赋予特定的含义。例如,用  $(x_{jit}(i, j; t), y_{jit}(i, j; t))$  描述抖动偏移。在每一时刻,如果所有像素点的抖动偏移(在统计意义上)都相等,就可以说抖动(在统计意义上)是空间不变的,抖动偏移可以简化为  $(x_{jit}(t), y_{jit}(t))$ 。典型地,可以通过均值为零、方差为  $\sigma_{jit}^2$  的高斯概率分布模型从系综上来描述抖动偏移的统计特性,通过频谱函数或二阶统计函数(自相关函数)来描述其时间频率特性。同样地,可以采用  $(x_{sh}(t), y_{sh}(t))$  描述空间不变的平动偏移,并把该失真模型称之为零阶偏移场模型。一般地,平动偏移可以采用 ARMA 模型或者它的两种简化形式(AR 模型或 MA 模型)来建模。多数情况下,平动偏移的频谱限制在低频带内,而抖动偏移的频谱限制在高频带内。

仿射运动模型<sup>[1-3]</sup>描述了两条平行线间的映射和诸如平移、旋转、缩放等运动产生的效果,但不能完全描述拍摄中图像序列的失真。相比较而言,透视失真模型更具一般性和适用性。该模型可以通过逐像素点定义的偏移序列  $(x_{pr}(i, j; t), y_{pr}(i, j; t))$  来表示,将其称之为—阶偏移场模型。一般情况下,透视偏移可以通过 8 个参数进行描述,并且这 8 个参数可以通过偏移图像和原始图像上 4 对对应点来确定(见第 4 节)。这样,只需描述偏移序列在 4 对对应点处的动态特性和统计特性即可。原始图像序列中 4 个点的动态偏移可以采用 4 个概率模型或 4 个 ARMA(Autoregressive Moving Average)模型进行建模。如果采用 4 个概率函数建模,那么这些点集的统计偏移均值将构成一个图像失真的透视偏移平面。通常情况下,采用概率模型时需要估计相对原始参考的实际失真,这就要根据失真的物理机理来估计原始参考。对于视频稳定化技术,原始参考并不是必需的,只需选取一个相对参考帧,使得后续各帧相对该参考帧保持稳定即能实现图像序列的稳定化。这时,采用 ARMA(AR, MA)模型将是比较合适的。同样地,可以通过一系列对应点对上图像偏移量的统计特性或确定性特性来描述半线性模型  $(x_{semi}(i, j; t), y_{semi}(i, j; t))$  或其他不同类型的非线性模型。

在实际应用中,图像序列的失真是由很多因素引起的。对于存在多种失真模式的图像序列,根据式(2)可以得到其复合偏移场模型。一种典型情况是,主要包含高频抖动和低频平动偏移的失真模式。这种情况下,动态偏移场模型可以表示为

$$i' = i + x_{jit}(t) + x_{sh}(t), \quad j' = j + y_{jit}(t) + y_{sh}(t), \quad i \in [1, M], \quad j \in [1, N] \quad (3)$$

同样地,由式(4)描述的动态偏移场可以描述透视失真加抖动的失真现象。

$$i' = i + x_{jit}(t) + x_{pr}(i, j; t), \quad j' = j + y_{jit}(t) + y_{pr}(i, j; t), \quad i \in [1, M], \quad j \in [1, N] \quad (4)$$

由上述分析可知,动态偏移场模型采用按阶分解的描述方法,在统一模型框架下可以描述包括抖动、偏移、旋转、缩放、透视变形、非线性变形等失真模式。对图像序列失真建模时,需要选择合理的偏移场模型形式。通常,这是比较困难的。影响考虑的因素主要包括失真的物理原因、应用需求和经济条件的限制等。对多数普通应用来说,要求处理成本比较低,这时,由式(3)和式(4)描述的偏移场模型将是比较合适的。由此可见,偏移场模型的建模方法可以使一些复杂失真得到一定程度的简化,同时按阶分解描述方法也是一个参数化的建模过程,对于失真校正工作有着重要的意义。

## 3 基于偏移场模型的抖动抑制技术

图像序列抖动失真是最常见的失真现象,文献[4-7]提出了许多消除抖动失真的方法,最具代表性的是采用 Kalman 滤波技术或其变形形式的滤波方法。对于跟踪一个视频目标的低频运动, Kalman 滤波器不失为一个强有力的处理工具。

但对于抖动失真的抑制, Kalman 滤波的处理效果并不令人满意。这主要因为 Kalman 滤波器不是根据频域特性设计的,或者说它不是一个最优的频域滤波器。由偏移场模型理论可知,抖动失真是图像偏移的一种高频变化。因此,在频域中对偏移量滤波应能抑制抖动。而这正是传统滤波器可以实现的,滤波器的通带和阻带可以根据应用需求而定。本文基于偏移场模型理论,采用传统的数字滤波技术实现了抖动失真的抑制,具体方法如下。

假定图像序列是时间上的帧序列,帧频记为  $f_s$ 。根据 Nyquist 采样定理,该图像序列的最高抖动频率是  $f_s/2$ 。如果  $f_j$  以上的图像偏移变动频率都视为抖动频率,那么抖动抑制滤波器的阻带范围为  $[f_j, f_s/2]$ 。抖动抑制滤波器可以采用数字滤波器的设计技术进行设计,它可以是基于偏移场模型的有限冲激响应(FIR)滤波器、无限冲激响应(IIR)滤波器(或 ARMA 滤波器)、或其他频域滤波技术。这里唯一的问题就是必须找到一种方法能够估计  $t = 0, 1, 2, \dots$  时的双序列  $(x_{jit}(t), y_{jit}(t))$ 。要确定每个  $t$  时的  $(x_{jit}(t), y_{jit}(t))$  就需要知道图像的原始参考,但对于失真图像序列,寻找原始参考帧通常是比较困难的。实际上,为了抑制抖动,可以采用一个固定的相对参考,使得后续各帧相对该帧保持稳定。事实上,相对参考与原始参考的差值将产生一个固定偏移  $(x_{jit}(0), y_{jit}(0))$ , 但该偏移并不影响抖动自身的估计。通过灰度投影<sup>[3]</sup>等方法可以估计得到帧间偏移量  $\Delta x_{jit}(t) = x_{jit}(t) - x_{jit}(t-1)$  和  $\Delta y_{jit}(t) = y_{jit}(t) - y_{jit}(t-1)$ 。根据相对参考,可以得到抖动偏移序列为

$$x_{jit}(t) = x_{jit}(0) + \sum_{k=1}^t \Delta x_{jit}(k), y_{jit}(t) = y_{jit}(0) + \sum_{k=1}^t \Delta y_{jit}(k) \quad (5)$$

随后,采用典型的数字滤波技术即可实现抖动的抑制,抖动抑制的效果可以通过估计抑制处理前后抖动偏移量的方差来评价。

#### 4 基于偏移场模型的图像序列失真校正

平动偏移校正的目的是稳定运动目标以便于观察,可以通过由估计的  $(x_{sh}(t), y_{sh}(t))$  平移到固定参考来实现,该固定参考的选择应使目标处于图像的中心位置。假设已经得到  $t = 0, 1, 2, \dots$  时估计的偏移时间序列  $(x_{sh}(t), y_{sh}(t))$ , 且当前时间为  $t = n$ , 固定偏移为  $(x_{sh}(0), y_{sh}(0))$ 。简单的校正方法是将当前帧在水平和垂直方向分别移动  $\Delta x_{sh}(n) = x_{sh}(n) - x_{sh}(0)$  和  $\Delta y_{sh}(n) = y_{sh}(n) - y_{sh}(0)$ 。实际应用中,如果估计的  $(x_{sh}(t), y_{sh}(t))$  存在误差,这种简单的平移可能会带来附加的抖动。采用低通滤波器对误差序列  $\Delta x_{sh}(t)$  和  $\Delta y_{sh}(t)$  分别进行滤波处理,即可抑制抖动的影响,低通滤波器的当前输出  $(\Delta \hat{x}_{sh}(n), \Delta \hat{y}_{sh}(n))$  即是该帧的运动量。这里,滤波器可以选择 FIR 或 IIR 低通数字滤波器,带宽可以根据目标的动态特性进行选择。

对于存在透视失真的图像序列,选择原始(无失真)图像

作为其参考图像通常是不可能的。考虑到精确校正通常难以实现,可以采用稳定目标的透视失真作为一种替代的解决方法。对目标的透视失真进行校正时,需要估计目标上不共线的4个点的动态偏移。定义与目标(图像)相关的4个点,可以以目标上的4个小窗口为基础来定义。采用图像配准技术<sup>[8, 9]</sup>估计4个小窗口的偏移量,4个窗口的4个中心点的偏移就可显示出目标的透视失真情况。像这样不共线的4个点的集合被称为目标透视失真的表示集(Indication Set of Projective Distortion, ISPD)。

透视变换采用双线性函数可以表示为

$$p = \frac{Au + Bv + C}{Gu + Hv + 1}, q = \frac{Du + Ev + F}{Gu + Hv + 1} \quad (6)$$

其中  $(u, v)$  和  $(p, q)$  分别是发生透视变换前后两帧图像像素点的位置坐标,  $A, B, C, D, E, F, G, H$  为8个未知变换参数,  $G \neq 0$  且  $H \neq 0$ 。假设采用图像配准技术已经估计得到目标上4对对应点在参考帧和当前帧的位置坐标,将其代入式(7)可得到8个变换参数。一旦得到8个变换参数,当前帧的校正便可通过式(7)所示的反变换实现。

$$\begin{bmatrix} pG - A & pH - B \\ qG - D & qH - E \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} C - p \\ F - q \end{bmatrix} \quad (7)$$

目标透视失真的稳定化,可以通过以下方法来实现。假设已逐帧估计得到目标的 ISPD, 那么由 ISPD 可以产生4个时间序列  $\{(p_i, q_i; t), i = 0, 1, 2, 3, t = 0, 1, 2, \dots\}$ 。一种稍微复杂的透视失真稳定化方法是,先采用低通滤波器对4个时间序列进行滤波,然后以4个滤波输出序列  $\{(\hat{p}_i, \hat{q}_i; t), i = 0, 1, 2, 3, t = 0, 1, 2, \dots\}$  为基础来实现失真校正。滤波器的带宽可以根据目标的动态特性进行选择,处理方法与时变平动偏移失真的校正方法相同。

#### 5 动态偏移场模型的数学仿真

基于偏移场模型理论,可以很容易地实现图像序列失真的仿真。

首先,考虑抖动失真的产生方法。假定抖动失真的系综分布满足均值为零、方差为  $\sigma_{jit}$  的高斯分布,频谱限定在  $[f_j, f_s/2]$ 。两个长为  $N_t$  的抖动序列  $(x_{jit}(t), y_{jit}(t))$  产生方法如下:

- (1)根据仿真时间确定需要仿真处理的图像帧总量  $N_t$ , 即随机偏移量的样本总数  $N_t$ ;
- (2)产生两个长为  $N_t$ , 均值为零, 方差分别为  $\sigma_{jitx}$  和  $\sigma_{jity}$  的 Gauss 随机序列  $s_x(t)$  和  $s_y(t)$ ;
- (3)对两个序列分别作离散傅里叶变换, 变换结果为  $S_x(\mu)$  和  $S_y(\mu)$ ,  $\mu = 0, 1, \dots, N_t - 1$ ;
- (4)对序列  $S_x(\mu)$  和  $S_y(\mu)$  在频域上滤除  $\left[0, f_j \frac{N_t - 1}{f_s}\right]$  和  $\left[(f_s - f_j) \frac{N_t - 1}{f_s}, N_t - 1\right]$  范围的谱线, 使其满足  $[f_j, f_s/2]$  的频谱限制;

(5)对由(4)得到的两个频谱，分别作离散傅里叶反变换得到两个长为  $N_t$  的独立时间序列  $x_{jit}(t)$  和  $y_{jit}(t)$ ， $(t = 0, 1, 2, \dots, N_t - 1)$ ，即满足仿真要求的抖动偏移序列  $(x_{jit}(t), y_{jit}(t))$ 。

按照上述方法，根据被仿真目标的指定动态特性可以实现空不变平动偏移序列和透视失真序列的仿真。

最后，产生信噪比为 SNR(dB)的加性图像噪声  $\xi(i', j'; t)$ 。首先产生一个大小为  $M \times N$  且满足  $[0, 1]$  均匀分布的噪声矩阵，将此噪声矩阵乘以比例系数  $c = 255 \cdot \sqrt{12 \cdot 10^{-\frac{SNR}{20}}}$ ，即可得到信噪比为 SNR(dB)的加性噪声。

在实际应用中，图像序列失真的实际仿真可联合以上方法得到。

### 6 实验结果

基于动态偏移场模型的建模理论和校正方法，下面给出几种低阶偏移场模型的仿真结果和视频稳定化结果<sup>[10]</sup>。为验证动态偏移场模型在描述图像序列失真时的通用性和有效性，本文对航空扫描地图视频序列(图 2(a))进行了 200 帧的仿真。在仿真中，为便于观察地图的失真情况，在图 2 中添加了矩形框作为参考。抖动失真的幅度和频率以及平动偏移的幅度和频率初始化为  $\sigma_{jitx} = 2, f_{jx} = 6\text{Hz}, \sigma_{jity} = 2, f_{jy} = 7\text{Hz}$ ， $\sigma_{shx} = 10, f_{shx} = 2\text{Hz}, \sigma_{shy} = 12, f_{shy} = 1\text{Hz}$ 。图像序列的采样频率为 15Hz，加性噪声的信噪比为 10dB。

抖动偏移场模型的视频仿真结果(图 2(b))显示，图像序列除呈现高频抖动外，还伴有图像整体灰度的变化。零阶偏移场模型中，图像序列发生了低频的平动偏移和图像整体灰度变化。如图 2(c)所示，矩形框偏移了中央位置，且随时间变化呈现水平和垂直方向的低频随机偏移。偏移后的空缺，在图像上以黑色显示出，如图 2(c)上方的黑色偏移空缺。在图 2(d)零阶偏移场模型加抖动的仿真结果中，图像在发生缓慢平动偏移的同时呈现出高频抖动。这种失真情况，更加符

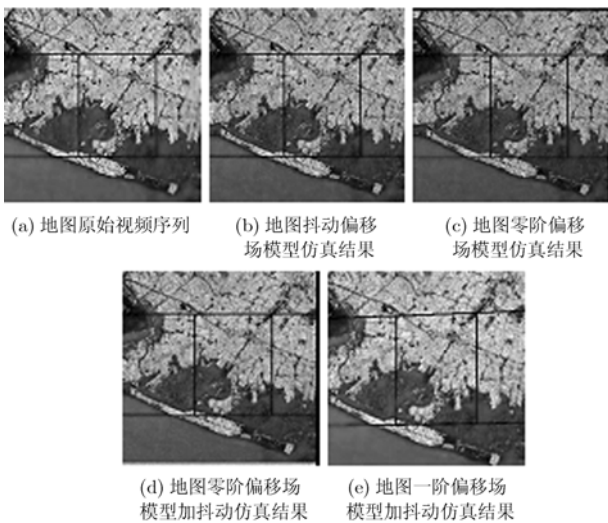


图 2 动态偏移场模型的地图仿真结果

合实际应用中因拍摄环境运动或不稳定造成的相机拍摄视频序列的失真。对于高空摄影图像序列的失真，一阶偏移场模型加抖动可以给予较为准确的描述。从图 2(e)的一阶偏移场模型加抖动的视频仿真结果可看到，相对于原始无失真图像，矩形框和地图发生了不同程度的拉伸和透视变形失真，且伴有高频抖动和图像整体灰度的变化。

下面，给出上述仿真结果的稳定化结果(图 3)，说明该模型在失真校正中的重要作用。在稳定化实验中，选择第 1 帧作为相对参考帧，后续图像帧相对该参考帧实现稳定化，基于偏移场模型的 FIR 滤波器截止频率设定为 0.5Hz。透视失真校正时，选择地图原始序列(图 2(a))作为参考图像。视频结果<sup>[10]</sup>显示，高频的抖动被滤除(图 3(a)和 3 (c))，低频的平动偏移得到估计和补偿(图 3(b)和 3 (c))，透视失真得到校正(图 3(d))，抖动的图像序列变得稳定。

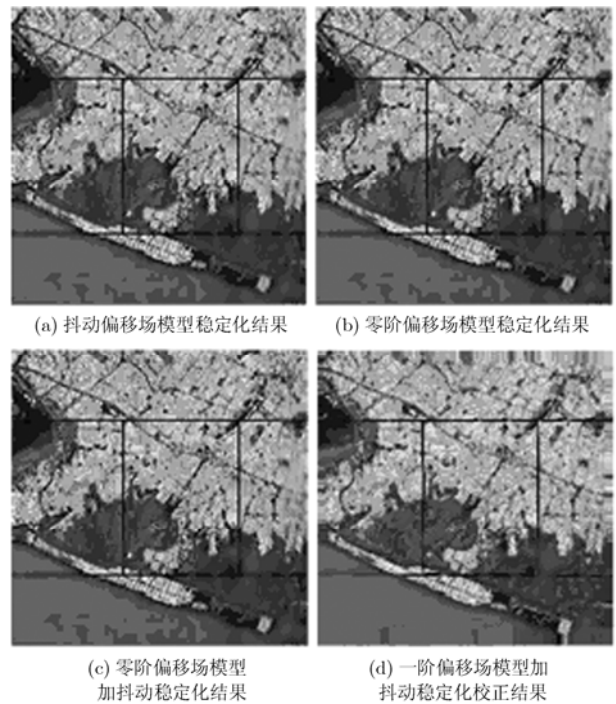


图 3 地图偏移场模型稳定化结果

基于偏移场模型理论，本文对实际拍摄视频序列进行了稳定化处理，并给出了该稳定化结果与商业稳定化软件 Steadyhand 的稳定化效果比较，稳定化效果请参见稳定化的视频结果<sup>[10, 11]</sup>。从稳定化视频结果可以看出，Steadyhand 软件不能实现对透视失真的校正，且由于其边界处理使得图像序列出现缩放及不连续现象；而本文的稳定化方法较好的实现了失真视频序列的稳定化和失真校正。

### 7 结束语

本文提出的动态偏移场模型作为一种新的图像序列失真模型，能够描述目前可观测到的多种图像序列失真现象，如随机抖动、低频偏移、透视变形、非线性失真和包含多种

失真模式的复合失真。偏移场模型的数学描述形式和仿真实现技术,为图像序列的动态失真提供了一个由低阶到高阶逐步演进的分解描述方式。本文对偏移场模型描述主要体现在二维空间,也可以将其类似地推广到三维空间。基于偏移场模型的失真校正技术,对于实现运动参数辨识和视频稳定化更加简单有效。实验结果表明,动态偏移场模型能够简单地描述图像序列失真现象,并为失真校正工作提供一个由简单到复杂渐进处理的研究平台。在实际工程中,采用动态偏移场模型描述图像序列失真现象是十分有意义的。

### 参考文献

- [1] Hoetter M. Differential estimation of the global motion parameters zoom and pan. *Signal Processing*, 1989, 16(2): 249-265.
- [2] Wu S F and Kittler J. A differential method for simultaneous estimation of rotation, change of scale and translation. *Signal Processing: Image Communication*, 1990, 2(1): 69-80.
- [3] Morimoto C and Chellappa R. Automatic digital image stabilization. Proc. ICPR, Osaka, Japan, 1997, 1: 60-65.
- [4] Uomori K, *et al.*. Automatic image stabilizing system by full-digital signal processing. *IEEE Trans. on Consumer Electronics*, 1990, 36(3): 510-519.
- [5] Ertürk S. Translation, rotation and scale stabilization of image sequences. *Electronics Letters*, 2003, 39(17): 1245-1246.
- [6] Wang Demin and Wang Limin. Global motion parameters estimation using a fast and robust algorithm. *IEEE Trans. on Circuits and Systems for Video Technology*, 1997, 7(5): 823-826.
- [7] Liang Y and Tyan H, *et al.*. Video stabilization for a camcorder mounted on a moving vehicle. *IEEE Trans. on Vehicular Technology*, 2004, 53(6): 1636-1648.
- [8] Brown L G. A survey of image registration techniques. *ACM Computing Surveys*, 1992, 24(4): 325-376.
- [9] Olson C F and Huttenlocher D P. Automatic target recognition by matching oriented edge pixels. *IEEE Trans. on Image Processing*, 1997, 6(1): 103-113.
- [10] <http://mwimlab.com/>动态偏移场模型的仿真及稳定化视频结果.rar.
- [11] <http://mwimlab.com/>采用Steadyhand软件的稳定化视频结果.rar.

邹谋炎: 男, 1941 年生, 研究员, 博士生导师, 研究方向为信号处理、图像处理理论和算法、无线宽带通信技术。

刘 艳: 女, 1979 年生, 博士生, 研究方向为图像、视频处理技术。

曹 瑛: 女, 1980 年生, 博士生, 研究方向为图像、视频处理技术。

卢晓鹏: 男, 1980 年生, 博士生, 研究方向为图像、视频处理技术。

杜 坤: 男, 1982 年生, 博士生, 研究方向为图像、视频处理技术。