

三维动态偏移场模型及其在视频跟踪技术中的应用

刘艳^① 邹谋炎^②

^①(中国科学院研究生院 北京 100190)

^②(中国科学院电子学研究所 北京 100190)

摘要: 基于二维动态偏移场模型, 该文提出了动态偏移场模型的三维描述形式, 实现了对目标沿景深方向运动状态的描述, 并为失真图像序列的动态偏移提供了由平面偏移估计到空间运动状态估计的渐进描述。在视频跟踪系统中, 较之其他视频跟踪技术, 基于三维偏移场模型理论, 该文实现了跟踪视频序列的稳定化和对被跟踪目标运动状态的估计。

关键词: 图像序列失真; 三维动态偏移场模型; 视频跟踪; 运动状态估计

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2008)11-2784-04

3-D Dynamic Displacement Field Model and Its Application to Video Tracking Systems

Liu Yan^① Zou Mou-yan^②

^①(Graduate University of Chinese Academy of Sciences, Beijing 100190, China)

^②(Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Based on the 2-D dynamic displacement field model, the dynamic displacement model is extended from 2-D to 3-D. It realizes the description for the motion state estimation of the target in the depth direction of the sight field. It provides a gradually description process for distorted image sequence from estimating plane displacements to estimating spatial motion state. Compared to other video tracking techniques, based on the 3-D dynamic displacement field model, it has been implemented to stabilize the tracking sequence and estimate the target's motion state.

Key words: Image sequence distortion; 3-D dynamic displacement field model; Video tracking; Motion state estimation

1 引言

对于多种原因引起的图像序列失真现象, 文献[1]中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。该模型在一个统一描述框架下, 可以描述目前可观测到的多种二维图像序列失真现象, 如抖动、偏移、旋转、缩放、透视变形、非线性失真和包含多种失真模式的复合失真等。在三维空间中, 除了要描述图像序列的三维动态失真之外, 还要描述运动目标沿图像景深距离方向的速度和位移随时间变化而变化的情况。多数图像序列失真模型^[1-7]主要是对运动目标在二维图像平面上偏移和变形的描述, 没有涉及到运动目标随时间变化所呈现出来的运动状态的改变。

基于文献[1]提出的二维动态偏移场模型, 本文提出了动态偏移场模型的三维描述形式, 实现了对目标在不同时刻沿景深方向运动速度和景深偏移等运动状态的描述。该模型采用二维图像参数来描述目标在三维空间景深方向上的运动

状态, 大大减少了三维模型描述的复杂性。实验结果表明, 在视频跟踪系统中, 基于三维动态偏移场模型, 本文实现了跟踪视频序列的稳定化和对被跟踪目标运动状态的估计。

2 三维动态偏移场模型的数学描述

文献[1]中介绍的动态偏移场模型是在二维图像平面上建立的空不变和空变二维偏移场模型, 而实际应用中通常还需要描述运动目标在三维空间的失真现象, 例如目标在三维空间的三维透视失真现象等。由一阶偏移场模型理论^[1]可知, 图像序列的透视失真可以通过二维图像平面上不共线4个点的动态偏移进行描述。相应地, 三维透视失真则表现为空间8面体的变形, 需要通过三维空间上的8个点的动态偏移来描述。而这无疑增加了模型描述的复杂性, 使得失真模型的建模和校正工作变得更加困难。另外, 二维动态偏移场模型描述了运动目标在二维图像平面上的偏移和变形, 没有涉及到运动目标随时间变化所呈现出来的运动状态的改变, 例如在三维空间中, 运动目标沿图像景深距离方向的速度和位移随时间变化而变化的情况。

为了减小模型描述和参数估计的复杂性, 基于二维动态

偏移场模型, 本文根据二维图像缩放因子与景深距离的关系, 采用二维图像参数来描述目标在三维空间景深方向上的运动状态, 提出了动态偏移场模型的三维描述形式。基于动态偏移场模型的一般数学表达形式^[1], 可以得到三维动态偏移场模型的数学表达形式如下所示:

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

$$d' = d + z(t) \quad (2)$$

其中 d 为 t 时刻相机与目标间的景深距离; d' 为 t 时刻偏移后的景深距离; $z(t)$ 为景深方向距离的动态偏移量, 可以采用文献[1]中抖动偏移和平动偏移的描述方法来描述其动态偏移情况; (i', j') 为 t 时刻像素点 (i, j) 偏移后的新位置坐标, 可以通过景深距离描述的缩放因子来确定, 详细描述方法如下。

在选定拍摄设备的情况下, 可以拍摄到的景物范围和人眼的注视视野随即可以确定下来。由于相邻两帧的时间间隔很短, 可以假定相邻帧间人眼感兴趣的注视视野范围是保持不变的, 即以观测点(人眼或相机)为顶点, 以运动目标在图像平面的投影为底面的圆锥体积保持不变。如图 1 所示, l 为圆锥的底面半径, 即运动目标在图像平面投影的横向长度的一半; d 为圆锥的高, 即相机与目标间的景深距离。由上述假设可知, 视野范围 $V = (1/3)\pi l^2 d$ 保持不变。

取圆锥的纵剖面图, 可得相邻帧间相机与目标运动状态的投影图, 如图 2 所示。为简化表示, 用两个圆圈表示 t_1 和 t_2 两个时刻相机所在位置, 用两条横线表示目标在图像平面的投影直径, 其中实线表示 t_1 时刻目标在平面上的投影, 虚线表示 t_2 时刻目标在平面上的投影。 t_1 时刻, 目标投影的半径为 l_1 , 相机与目标的景深距离为 d_1 ; t_2 时刻, 目标投影的半径为 l_2 , 相机与目标的景深距离为 d_2 。由相邻帧间视野范围 V 保持不变可得:

$$d_2/d_1 = l_1^2/l_2^2 = 1/s^2 \quad (3)$$

其中 s 为相邻两帧图像间的缩放因子。式(3)表明, 前后两个时刻相机与目标的景深距离之比与前后两帧图像缩放因子平方成反比。假设前后两个时间间隔为 t , 相机的平均运动速度为 v_1 , 目标的平均运动速度为 v_2 , 由图 2 可知:

$$d_1 - v_1 t = d_2 - v_2 t \quad (4)$$

联合方程(3)和方程(4)可得

$$s^2 = \frac{d_1}{(v_2 - v_1)t + d_1} \quad (5)$$

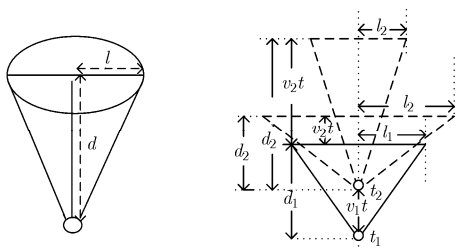


图 1 相邻帧间视野范围保持不变 图 2 相机与目标运动状态投影图

方程(3)和方程(5)建立起了目标图像缩放因子与相机和目标的运动参数之间的关系。分析可知, 当 $d_2 < d_1$ 或 $v_2 < v_1$ 时, $s > 1$, 说明目标与相机的距离拉近, 目标的运动速度减小, 目标尺寸放大; 反之, 目标与相机的距离拉长, 目标运动速度增加, 目标尺寸缩小, 如图 2 所示。

根据景深距离与目标图像缩放因子之间的关系, 采用帧间缩放因子调整目标图像所在区域像素点的位置坐标, 在二维图像平面上就可以仿真出目标大小随景深距离变化而变化的三维运动效果。假设目标图像区域为 S , 该区域像素点位置坐标经缩放因子 s 调整后, 表示为

$$i' = s \cdot i, \quad j' = s \cdot j, \quad (i, j) \in S \quad (6)$$

在此缩放处理的基础上, 在二维平面上对图像平面所有像素点的横、纵坐标分别添加二维动态偏移, 便可得到三维空间上目标距离相机远近变化的抖动偏移、平动偏移等偏移模式。

3 三维偏移场模型在视频跟踪中的应用

基于目标图像缩放因子与相机和目标的运动参数之间的关系, 下面主要讨论三维偏移场模型在视频跟踪中的应用。一般的视频跟踪系统^[8-13]主要包括对被跟踪目标的检测、识别、定位和跟踪, 而不涉及对被跟踪目标运动状态的估计。基于稳定后的视频序列, 本文采用空变失真校正技术逐帧估计目标图像的缩放因子, 根据缩放因子与运动参数间的关系, 基于初始运动参数, 实现了对运动目标速度和目标与相机间景深距离的较准确估计。

假设采用文献[14]的空变失真校正技术已经获得相邻帧间的目标图像缩放因子 s , 并已知相机与目标间的初始景深距离 d_1 和相机的平均运动速度 v_1 。由式(3)和式(5)可得被跟踪目标与相机的当前景深距离和目标当前的平均运动速度 v_2 为

$$d_2 = \frac{d_1}{s^2}, \quad v_2 = \left(\frac{1}{s^2} - 1 \right) \frac{d_1}{t} + v_1 \quad (7)$$

其中 t 为目标的运动时间。分析式(7)可知, 当 $s > 1$ 时, $d_2 < d_1$, $v_2 < v_1$, 说明目标尺寸放大, 目标与相机的距离拉近, 目标的运动速度减小; 反之, $d_2 > d_1$, $v_2 > v_1$, 说明目标尺寸缩小, 目标与相机的距离拉长, 目标运动速度增加, 如图 2 所示。根据目标图像的缩放变化, 通过式(7)和式(8)即可实现对其运动状态的估计。

4 实验结果

在三维动态偏移场模型的仿真中, 景深距离偏移的幅度和频率设定为 $\sigma_d = 5, f_d = 1\text{Hz}$, 如图 3(a)所示。基于该距离偏移序列, 由式(3)可得图像的缩放因子(图 3(b))。根据缩放因子, 对原始图像逐帧调整目标图像的大小, 可以得到目标距离相机远近变化的视频序列^[12], 如图 4 所示。在此仿真结果的基础上, 在二维平面上对图像平面所有像素点的横、纵坐标分别添加二维平动偏移量和抖动偏移量, 便可得到三维

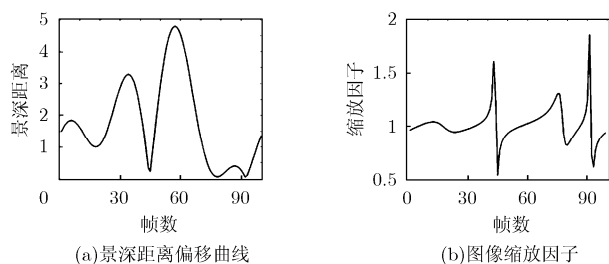


图3 景深距离偏移序列及图像缩放因子

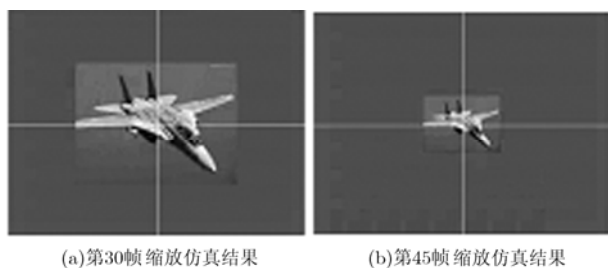


图4 三维动态偏移场模型仿真结果

空间上目标距离相机远近变化的抖动模型和平动偏移模型, 请参见视频结果^[15]。

以上述仿真结果作为实验序列, 采用空变失真校正方法逐帧估计目标图像的缩放因子(图5(a))。假设目标与相机的初始距离 $d_1 = 1.5$, 根据式(7)估计可得目标与相机的景深距离序列(图5(b))。与图3 仿真设定的景深距离序列和缩放因子相比, 空变失真校正方法能够较为准确地估计出目标图像的缩放因子; 基于该缩放因子, 在已知初始运动参数的情况下, 可以实现对目标与相机间景深距离(或目标运动速度)较为准确的估计。

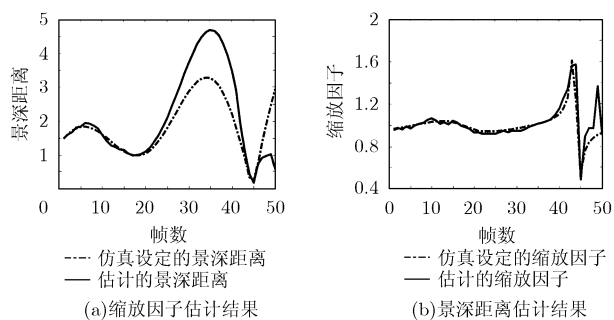


图5 目标图像缩放因子与景深距离偏移序列的估计结果

在视频跟踪系统中, 三维偏移场模型理论和运动状态估计方法在实现对目标运动状态较准确估计的同时, 还可以实现对抖动失真或平动偏移目标的稳定化。基于图4 三维偏移场模型的抖动模型和零阶模型, 采用基于偏移场模型的运动滤波方法^[1]实现了抖动的滤除和平动偏移的校正。从视频结果^[15]可以看到, 运动目标在图像平面中间呈现出稳定的远近变化现象。在CPU为 Pentium(R)4, 主频为2.4GHz, 内存

为512MHz的PC机上的MATLAB6.5开发环境下进行上述仿真实验, 图4的仿真建模实验, 每帧图像处理时间为1s; 图5的运动状态估计实验, 每帧图像的处理时间为0.9s, 基本可以实现对目标运动状态的实时估计。

5 结论

基于二维动态偏移场模型, 本文提出了动态偏移场模型的三维描述形式, 实现了对目标沿景深方向运动状态的描述。该模型采用二维图像参数来描述目标在三维空间景深方向上的运动状态, 大大减少了三维模型描述的复杂性。该模型能够描述目标运动状态随时间变化而变化的特性, 为失真图像序列的动态偏移提供了由平面偏移估计到空间运动状态估计的渐进描述。在视频跟踪系统中, 基于三维偏移场模型理论, 实现了跟踪视频序列的稳定化和对运动目标速度和目标与相机间景深距离的较准确估计。

参考文献

- [1] 邹谋炎, 刘艳等. 一种新的图像序列失真模型: 动态偏移场模型. 电子与信息学报, 2008, 30(9): 2143-2147.
- [2] Zou Mou-yan and Liu Yan, *et al.* Dynamic displacement field model: A new model for describing image sequence distortion. *Journal of Electronics & Information Technology*, 2008, 30(9): 2143-2147.
- [3] Tico M, Alenius S, and Vehvilainen M. Method of motion estimation for image stabilization. Proc. ICASSP, Toulouse, France, 2006, 2: 277-280.
- [4] Yang S H, Jheng F M, and Cheng Y C. Two-dimensional adaptive image stabilization. *Electronics Letters*, 2007, 43(8): 446-448.
- [5] Ertürk S. Translation, rotation and scale stabilization of image sequences. *Electronics Letters*, 2003, 39(17): 1245-1246.
- [6] Chang Hung-chang, Lai Shang-hong, and Lu Kuang-rong. A Robust and efficient video stabilization algorithm. Proc. ICME, Taipei, Taiwan, 2004: 29-32.
- [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] Matsushita Y and Ofek E, *et al.* Full-frame video stabilization. Proc. CVPR, San Diego, CA, USA, 2005, 1: 50-57.
- [9] Choi J H and Lee K H, *et al.* Vehicle tracking using template matching based on feature points. Proc. IRI, Waikoloa Village, HI, USA, 2006: 573-577.
- [10] Comaniciu D and Meer P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. on Pattern Anal. Mach. Intell.*, 2002, 24(5): 603-619.

- [10] Hsieh J W and Yu S H, *et al.*. Automatic traffic surveillance system for vehicle tracking and classification. *IEEE Trans. on Intelligent Transportation Systems*, 2006, 7(2): 175-187.
- [11] Humphries M, Radev P, and Shirvaikar M. A realtime vehicle tracking system. Proc. Thirty-Seventh Southeastern Symposium, Alabama, USA, 2005: 357-361.
- [12] Idler C and Schweiger R, *et al.*. Realtime vision based multi-target-tracking with particle filters in automotive applications. IEEE Intelligent Vehicles Symposium, Tokyo Japan, 2006: 188-193.
- [13] Withopf D and Jahne B. Learning algorithm for real-time vehicle tracking. Proc. ITSC, Toronto, Canada, 2006: 516-521.
- [14] 刘艳, 邹谋炎. 基于偏移场模型的空变失真图像序列的视频稳定化技术. 中国科学院研究生院学报, 2008, 25(1): 111-116. Liu Yan and Zou Mou-yan. Video stabilization technique for spatially variable distorted image sequence based on dynamic displacement field model. *Journal of the Graduate School of the Chinese Academy of Sciences*, 2008, 25(1): 111-116.
- [15] 刘艳, 邹谋炎. 动态偏移场模型仿真结果. rar, <http://mwimlab.com/>.
- 刘 艳: 女, 1979 年生, 博士, 讲师, 研究方向为图像、视频处理技术.
- 邹谋炎: 男, 1941 年生, 教授、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.