

一种针对大尺度运动的快速光流算法

卫保国, 李 晶

(西北工业大学 电子信息学院, 西安 710072)

摘要: 传统的光流法必须满足一致连贯性假设, 不适用于大幅度运动目标的跟踪。为克服复杂场景下, 大尺度运动中目标位移量超出特征跟踪窗口带来的光流计算问题, 提出了一种针对大尺度运动的快速光流计算方法。引入图像金字塔模型, 利用基于多尺度分层的金字塔结构和光流映射变换, 实现小窗口捕获大运动。同时, 采用非极大值抑制方法保留图像自相关矩阵的大特征值, 得到的强角点作为特征点, 降低了光流计算的时间开销。实验结果表明, 提出的方法在复杂场景和运动条件下, 可准确快速地计算出表征目标运动方向和速率的光流矢量, 具有较高的鲁棒性和实时性。

关键词: 光流计算; 图像金字塔; 强角点特征; 映射变换

中图分类号: TP317.4 **文献标志码:** A **文章编号:** 1001-3695(2012)09-3551-03

doi:10.3969/j.issn.1001-3695.2012.09.093

Fast optical flow algorithm for large-scale motion

WEI Bao-guo, LI Jing

(School of Electronics & Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Conventional optical flow computation method must satisfy the consistent and coherent hypothesis, which is not appropriate to large-scale motion. In order to overcome the difficulty of calculating optical flow on fast and great magnitudes motion whose target displacement exceeds feature tracking window, this paper introduced the multi-scale image pyramid model and vector mapping transform, adjusting the image resolution to capture large-scale motion within small calculating window. Meanwhile, by using non-maximum suppression method to select points with large eigenvalue of image correlation matrix as strong corner feature, it reduced the cost of optical flow calculation. Experimental results show that the proposed algorithm is robust to complicated circumstances and can obtain the direction and speed of optical flow vector efficiently.

Key words: optical flow computation; image pyramid; strong corner feature; mapping transform

光流是空间运动目标在观测成像面上的像素运动的瞬时速度。光流法是目标检测与跟踪的常用方法之一。由于光流法^[1]可提供目标运动和结构信息, 其具有以下明显的优点: a) 计算精度较高, 可准确检测到亚像素级位移; b) 光流计算可应用于相对复杂的运动中, 如平移、旋转、缩放等, 因此得到广泛的应用。

光流计算的前提是运动一致连贯性假设: a) 光照一致假设, 即亮度恒定, 场景中被跟踪部分的亮度不变; b) 小运动假设, 即目标小运动, 也即目标运动相对于帧率是缓慢变化的; c) 局部一致性假设, 即相邻的点具有相似的运动, 并保持相邻。由于此假设, 已有的光流计算方法^[1-3]均采用小窗口, 存在不适用于大幅度运动以及计算开销大等局限性。

对于超出计算窗口的大幅度运动, 如果增大计算窗口, 将违背运动一致连贯性假设, 造成无法计算光流或光流计算错误。为此, 文献[4]提出采用支持向量回归(support vector regression)计算光流, 以得到平滑的光流, 消除某些帧的光流缺失。但该方法不适用于大位移或不连续运动。文献[5]采用拉普拉斯金字塔, 利用差值平方和区域匹配方法实现对流速及其协方差矩阵的估计, 然后用变分法在方向平滑约束下迭代求解光流。这种方法精度较高, 可以有效地抑制对遮挡和运动不

连续处的过度平滑, 但对噪声较为敏感, 且运算量较大。本文利用图像多分辨率特征与金字塔分层结构, 将图像金字塔结构应用于光流计算。首先动态调整图像至不同分辨率层, 在满足运动一致连贯性条件下采样低分辨率图像层计算光流, 通过光流映射, 得到原始分辨率的图像层光流, 实现小窗口捕获大运动。同时, 为了减小光流计算开销, 选取强角点, 将计算的光流稀疏化, 提高算法鲁棒性和实时性。

1 基于强角点的光流场计算

光流是三维速度在二维图像平面中的投影, 它可以通过图像的亮度变化提取出来, 其主要假设是每一个像素点在其运动轨迹上都保持灰度值不变。光流场的计算原理是将二维速度场与灰度相联系, 建立光流约束方程并求解得到光流。目前光流的计算一般采用基于梯度的光流计算方法, 其典型的方法有 H-S 算法和 L-K 算法两种。H-S 算法计算的光流将图像中的每个像素都与速度关联, 属于稠密光流场, 因此该方法的计算时间开销大, 速度缓慢。L-K 算法^[6,7]只需要感兴趣点周围的小窗口的局部信息就可得到稀疏光流, 从而降低光流计算过程中的时间开销。这些感兴趣点要与目标的运动速度或者连续两帧之间的位移关联。本文选取强角点作为特征点, 研究对

收稿日期: 2011-12-22; 修回日期: 2012-02-13

作者简介: 卫保国(1970-), 男, 陕西乾县人, 副教授, 博士, 主要研究方向为图像处理、模式识别、计算机视觉(wbg@nwpu.edu.cn); 李晶(1985-), 男, 宁夏灵武人, 硕士, 主要研究方向为目标检测与跟踪。

象主要是针对刚体,但是稍加修改,也可以适用于非刚体光流计算。

1.1 强角点提取

角点是二维图像亮度变化剧烈的点或图像边缘曲线上曲率极大值的点,本文中的强角点指与其邻域反差明显的点。图像中的角点不仅在两个正交方向上有明显导数,而且具备平移和旋转的不变性。可以证明,角点位于图像二阶导数自相关矩阵的两个最大特征值的地方,表明在以此点为中心的周围存在至少两个不同方向的边缘(或是纹理),因此选取有效角点的问题就转换为计算并选取图像矩阵中大特征值对应点的代数问题。例如典型的 Harris 特征点就是对图像 $I(x, y)$ 的二维矩阵,计算该矩阵行列式值和其迹的差值,然后保留差值大于指定阈值的点。但是, Harris 角点检测算法在梯度平缓变化的情况下可能漏掉一些实际的角点,而且对噪声比较敏感。为了避免计算二阶矩阵的行列式和迹,减少算法运算量,本文算法步骤如下:

a) 计算输入图像像素点邻域二阶导数自相关矩阵,即图像 $I(x, y)$ 的二维 Hessian 矩阵为

$$H(p) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix}_p \quad (1)$$

并求出该矩阵对应的两个特征值。

b) 进行非极大值抑制模板平滑,即保留邻域中的局部最大值。通过将较小的特征值与判定阈值作比较筛选特征点。若小特征值大于阈值,则判定得到一个强角点,否则排除。

c) 为了提高算法的稳定性和实时性,可以使用距离测度来剔除一些距离较近的角点,确保所有选取的角点之间保持足够的距离。

1.2 光流的计算

在提取强角点之后,为实现对于运动目标的特征跟踪,把图像中的角点作为特征点并标记角点在图像序列中的位置,通过计算相邻图像两帧之间特征点的位移场,得到运动的光流场。

假定图像上点 $M = (x, y)^T$ 在时刻 t 的灰度值为 $I(x, y, t)$, 且灰度随 x, y, t 缓慢变化,可以得到光流场的基本方程为

$$I_x u + I_y v + I_t = 0 \quad (2)$$

其中: $u = \frac{dx}{dt}$ 、 $v = \frac{dy}{dt}$ 代表 x, y 方向上的光流, $I_x = \frac{\partial I}{\partial x}$ 、 $I_y = \frac{\partial I}{\partial y}$ 、

$I_t = \frac{\partial I}{\partial t}$ 分别代表图像灰度对于 x, y, t 的偏导,其向量形式为

$$\nabla I \cdot v_m + I_t = 0 \quad (3)$$

其中: $\nabla I = (I_x, I_y)$ 表示图像在点 m 处的梯度, $V_m = (u, v)$ 是点 M 处的光流。式(3)称为光流约束方程,本文在提取强角点之后,基于 L-K (lucas-kanade) 方法^[7] 计算光流。该方法不受目标仿射变换的影响,如图 1 所示。

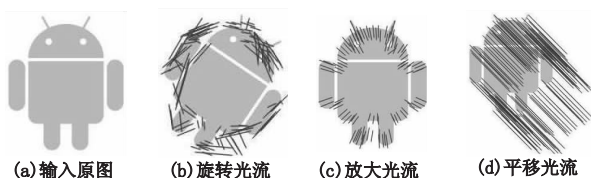


图1 目标在多种变换情况下光流计算结果

L-K 方法是基于局部约束的。假定以 p 点为中心的一个小区域内各点的光流相同,对区域内不同的点给予不同的权重,这样光流的计算就转换为最小化式(4):

$$\sum_{x \in \Omega} W^2(x) [\nabla I(x, t) \cdot v + I_t(x, t)]^2 \quad (4)$$

其中: Ω 代表以 p 点为中心的邻域; $W(x)$ 为窗函数,代表区域中各点的权重,离 p 点越近,权重越高。式(4)的解可由式(5)得到。

$$A^T W^2 A v = A^T W^2 b \quad (5)$$

对于邻域 Ω 内的 n 个点,其中:

$$A = (\nabla I(x_1), \nabla I(x_2), \dots, \nabla I(x_n))^T$$

$$W = \text{diag}(W(x_1), W(x_2), \dots, W(x_n))$$

$$b = -(I_t(x_1), I_t(x_2), \dots, I_t(x_n))^T$$

最后,方程的解为

$$v = (A^T W^2 A)^{-1} A^T W^2 b \quad (6)$$

当递归过程中 v 收敛,设图像大小为 $N \times N$,搜索相应窗口范围为 $W \times W$,则计算的时间复杂度为 $W^2 \log N$ 。

2 改进金字塔光流算法

通常的微分光流法仅适用于目标位移较小时,在较小的窗口内跟踪计算光流。但在实际应用中,通常存在大量快而不连贯的运动。运动较大的点会超出小跟踪窗口的范围,无法在窗口中找到选取的特征点,甚至造成丢失现象。但如果使用大的窗口来跟踪较大的运动,就会违背运动一致连贯性的假设并增加运算复杂度。为此,本文引入图像金字塔结构。对于在高分辨率图像层无法跟踪的大幅度运动,沿图像金字塔寻找合适的较低分辨率的图像层,在该层计算光流,然后将光流映射到原始图像层,实现对较大运动对应的光流计算。

2.1 图像金字塔模型

图像的金字塔结构^[8]是计算机视觉中一种多分辨率表示法,不同金字塔分辨率层可分析不同尺度的目标。同时,较低分辨率的高层(粗尺度)得到的信息还可用来指导对较高分辨率的低层(精尺度)的分析。本文采用的图像金字塔结构中,当前层是下层平滑后进行下采样结果,并规定原始图像(最底层)层数为 0。当图像分层到一定层次,目标在两帧间的运动量将变得足够小,就能够满足光流计算的约束条件,从而进行光流估计,如图 2 所示。

采用高斯低通滤波后下采样得到图像金字塔。图像与滤波器作卷积运算,然后将结果作隔行隔列取样,从而得到上一层的图像。假设第 $N-1$ 层的图像为 $I^{N-1}(x, y)$,则通过下采样得到上一层第 N 层的图像为 $I^N(x, y)$ 。如果在某一个 L 层相邻两帧间的目标运动幅度足够小,窗口中包含选取的特征点,就能够满足光流计算的前提假设,直接进行精确的光流计算。

同理,上采样则可以看做下采样的逆过程,首先对图像用零来填充新增的行或列,即隔行(隔列)插入零点,再与滤波器进行卷积操作,然后乘以系数 k (其中 k 为乘法因子),即得到下一层的图像。通常选取 $k=2$,保留奇数行和奇数列。

2.2 层间映射变换

采用分层策略,将跟踪目标的位移在图像金字塔内分层,在不同图像层进行光流计算,通过一系列层内光流迭代运算和层间映射变换,得到大幅度运动光流。

计算光流时,从图像金字塔高层到低层进行。上层的光

流值作为当前层光流计算的初始估计值。以此初始估计值为基础,计算光流增量,将其加到该层的光流估计值上,得到该层的光流值,并以此作为其下一层的光流计算估计值。将这一过程进行迭代,直至估计出原始图像的光流。如图3所示。

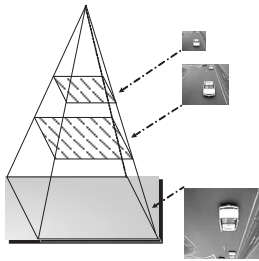


图2 三层图像金字塔结构模型

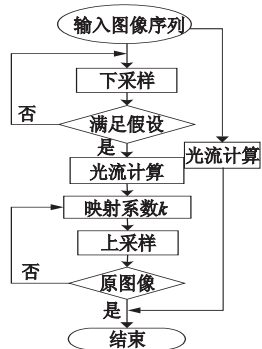


图3 光流计算中的层间映射变换

设 g_L 为 L 层对应的初始估计光流向量, Δf_L 为 L 层内计算的光流结果,假设经判断在金字塔第 $N-1$ 层满足光流计算条件,则:

最顶层的初始估计光流为

$$g_{N-1} = 0 \quad (7)$$

层间映射递推式为

$$g_{L-1} = k(g_L + \Delta f_L) \quad (8)$$

由上而下计算,得到

$$f = g_0 + \Delta f_0 = \sum_{L=0}^{N-1} k_L \Delta f_L \quad (9)$$

假设算法采用三层金字塔结构,在金字塔最顶层 $g_2 = 0$,经过式(6)光流法计算得出向量 Δf_2 ,代入式(8),当经过一次上采样计算,得到下一层的初始运动向量变为 $g_1 = k\Delta f_2$,经过对第2层的图像映射变换,再由式(6)和(8)计算出光流向量 Δf_0 和 g_0 , $f = g_0 + \Delta f_0$ 。以此类推迭代,最终得出最底层光流值为

$$v(x, y) = \sum_{L=0}^2 k_L \Delta f_L \quad (10)$$

2.3 算法流程

输入:视频帧 $I_i, I_k (k > i)$ 。

输出:强特征点位置及其光流 $v(x_1, y_1), v(x_2, y_2), \dots$ 。

a) 初始化。在初始图像中设定光流跟踪窗口宽度 W , 特征值阈值。

b) 在目标范围内提取目标的强特征点。

c) 选取一个强特征点。

d) 判断当前层当前帧是否满足光流计算条件,如果满足则按照式(6)计算光流,转至 f); 否则继续执行 e)。

e) 对 I_i, I_k 高斯低通下采样得到图像金字塔上一层,转至 d)。

f) 如果为原始层(第0层),则输出光流;否则上采样,并进行层间光流映射。

g) 如果所有强特征点处理完毕,结束;否则选取下一个强特征点,转至 d)。

依据上述递推过程,不难发现较高层上的计算误差会不断地积累并且向较低层传播和放大。当金字塔结构光流计算出最大位移量满足假设中的条件时,特征点跟踪成功,否则失败。因此,实际应用中,进行基于图像金字塔分层结构的光流计算时,分解的层数不宜过多。

3 实验结果及分析

实验1 本实验中采用图像序列为运动汽车,选取视频中相邻两帧图像完成角点提取和基于金字塔的光流跟踪算法,运动汽车选为目标。在选取的实验输入中,相邻帧间的背景较复杂,目标运动幅度大,且图像中含有噪声。实验环境为 VC6.0, Pentium 4 3.20 GHz 处理器,512 MB 内存,操作系统为 Window XP。

图4 为对快速运动汽车光流计算结果对比。由图4(c)的基于传统光流算法结果不难发现,背景中存在大量无关的错误特征点(图中大房屋的角点和静止汽车的角点),不但干扰特征点的选取、影响跟踪精度,而且由于特征点数目相对较多,在算法执行相同迭代次数的情况下,增加了运算开销,降低了算法实时性;(d)为本文算法的实验结果,对比(c)可知,(d)选取的特征点集中在目标(汽车)上,克服了传统光流无法计算快速不连贯运动导致的运动目标光流位移量超出跟踪窗口范围的缺点;(e)(f)分别为本文算法所选取特征点的光流矢量和光流跟踪的局部放大结果。

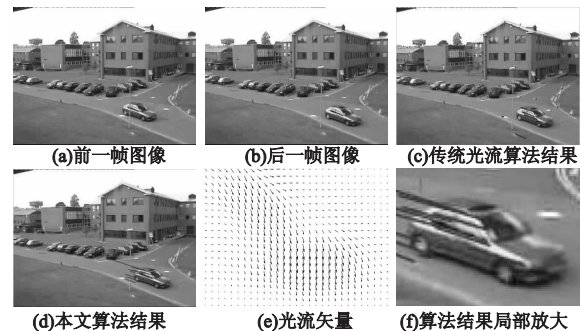


图4 光流算法和图像金字塔光流算法结果比较

实验2 为了验证本文算法在非刚体目标中的光流计算效果,进行了人脸跟踪的实验。其中,初始帧人脸范围采用目标检测算法得到,而在后续帧跟踪人脸时,采用本文算法的光流计算结果重新确定人脸范围。图5 为本文算法对人脸动作的光流计算结果,(b)为采用人脸检测算法提取人脸范围内的特征点,即按照1.1节方法得到的用于计算光流的强角点。实验采用图像分辨率为 270×350 (依次为序列的1、2、7、13、15、17、19、22和28帧),可以看出,改进的算法可较好地跟踪复杂光照和背景情况下快速移动的人脸。图中白色方框为算法跟踪结果,其中箭头代表目标移动矢量,箭头长度表征目标移动速率大小。此算法对于复杂背景中的点均无误检,适应多变运动条件下跟踪,实验结果较为准确地反映了目标运动的速度大小和方向信息。通过大量的实验表明,本文对 CIF 格式的视频跟踪时间为每秒 30.4 帧,达到了实时性要求。

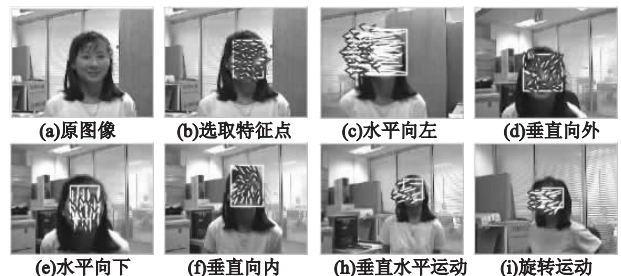


图5 在序列“seq_mb”中的光流计算结果

4 结束语

针对大尺度运动,本文提出了一种基于特征点提取的金字塔光流算法。利用金字塔分层结构,在不同级别的图像分辨率层次中,动态调整光流计算窗口的尺寸,使光流的计算符合运动一致连贯性假设。通过选择强角点作为特征点,使光流计算过程稀疏化,降低时间开销。实验证明,本文方法具有良好的性能;通过预测运动值提高了运算准确度,同时增强了其抗噪能力。今后将考虑建立光流计算的分层或多阶段模型;随着生物视觉研究的不断深入,基于神经机制的光流乃至计算机视觉算法也值得深入研究。

参考文献:

- [1] BARRON J L, FLEET D J, BEAUCHEMIN S S, *et al.* Performance of optical flow techniques[C]//Proc of IEEE Computer Society Conference on Computervision and Pattern Recognition. 1992:236-242.
- [2] 屠大维,江济良. 改进的光流运动图像分析方法及其应用[J]. 光

学精密工程,2011,19(5): 1159-1164.

- [3] LIAO Shih-kuan, LIU Baug-yu. An edge-based approach to improve optical flow algorithm[C]//Proc of the 3rd International Conference on Advanced Computer Theory and Engineering. 2010:45-51.
- [4] COLLIEZ J, DUFRENOIS F, HAMAD D. Optic flow estimation by support vector regression[J]. **Engineering Applications of Artificial Intelligence**, 2006, 19(7): 761-768.
- [5] 张泽旭,崔平远,崔祐涛. 一种基于方向平滑约束优化的多尺度光流算法[J]. 高技术通讯, 2007, 17(2): 153-158.
- [6] LUCAS B, KANADE T. An iterative image registration technique with an application to stereo vision[C]//Proc of the 7th International Joint Conference on Artificial Intelligence. 1995: 674-679.
- [7] LEE Y S, MOKRI S S, HUSSAIN A, *et al.* Motion detection using Lucas Kanade algorithm and application enhancement[C]//Proc of International Conference on Electrical Engineering and Informatics. 2009:537-542.
- [8] BOUGUET J Y. Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm[R]. Hillsboro: Intel Corporation Microprocessor Research Labs, 2003.