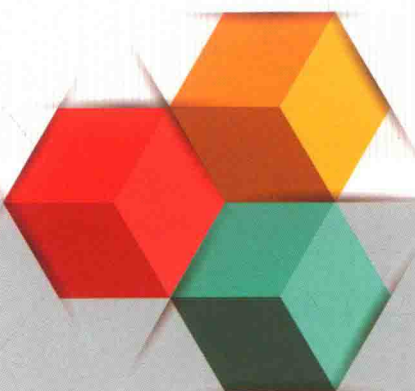


“十二五”国家重点图书出版规划项目
航空航天精品系列



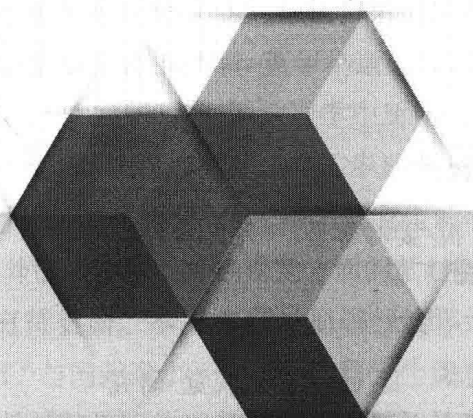
计算机视觉 运动分析

姜峰 刘绍辉 张盛平 杨炽夫 张健 著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

“十二五”国家重点图书出版规划项目
航空航天精品系列



计算机视觉 运动分析

姜峰 刘绍辉 张盛平 杨炽夫 张健 著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内 容 简 介

本书涉及计算机、信息论、心理学等专业知识,阐述了视觉运动分析理论,内容涵盖了从图像序列中估计2D和3D运动、差分运动分析及基于核函数的视觉跟踪、蒙特卡罗运动分析、铰链运动分析及人体姿态估计、多目标跟踪算法及基于学习的跟踪算法等。通过本书,读者可全面了解计算机视觉运动分析中最为前沿、重要的研究热点,如视频序列中铰链运动捕捉与分析、多目标的分析与跟踪诸方向的研究现状、方法原理等。

本书可作为高等工科院校自动化类高年级本科生或硕士研究生教材,也可作为从事控制系统分析、设计的工程技术人员及其他相关专业的师生学习的参考书。

图书在版编目(CIP)数据

计算机视觉 运动分析/姜峰等著. —哈尔滨:哈尔滨工业大学出版社,2018.6

ISBN 978-7-5603-7036-1

I. ①计… II. ①姜… III. ①计算机视觉-研究
IV. ①TP302.7

中国版本图书馆 CIP 数据核字(2017)第 274995 号

策划编辑 王桂芝 刘 威

责任编辑 刘 瑶 张凤涛

出版发行 哈尔滨工业大学出版社

社 址 哈尔滨市南岗区复华四道街10号 邮编 150006

传 真 0451-86414749

网 址 <http://hitpress.hit.edu.cn>

印 刷 哈尔滨圣铂印刷有限公司

开 本 787mm×1092mm 1/16 印张 11.75 字数 270千字

版 次 2018年6月第1版 2018年6月第1次印刷

书 号 ISBN 978-7-5603-7036-1

定 价 48.00元

(如因印装质量问题影响阅读,我社负责调换)

前 言

从图像处理与模式识别发展起来的计算机视觉的理论方法,主要是基于几何、概率和运动学计算、三维重构的视觉计算理论。计算机视觉要达到的最终目的是实现利用计算机对三维景物世界的理解,即实现人的视觉系统的某些功能。计算机视觉起步于1970年,以Marr的奠基性著作(Marr英年早逝,由其学生整理成书)为重要标志。目前,计算机视觉进入大发展时期,即物理学中所谓的“相变”期,大量的数学模型、算法不断涌现,一个大一统理论和框架正在逐步显形。

在我国,部分高校已将“计算机视觉”作为硕士研究生的学位必修课和博士研究生的选修课。目前,国内外已出版的计算机视觉方面的教材版本虽然比较多,但其中计算机视觉理论成果还有待更新。本书作者长期讲授哈尔滨工业大学“计算机视觉”研究生课程,并承担了“与国际高水平学者共建研究生课程”的共建项目,经与普林斯顿大学S. Y. Kung教授就编写体现最新教学思想及教学内容的“计算机视觉”教材进行了深入探讨,并通过多次授课、逐次修改,本书得以完成。

本书系统地介绍了中层计算机视觉的基础知识、分析方法和典型应用。全书共分为6章:第1章介绍了从图像序列中估计2D和3D运动的经典方法,该方法是20世纪八九十年代中层计算机视觉研究的重点内容,作为本书的基础;第2章和第3章分别介绍了在差分运动分析中的两大流派,这是20世纪90年代计算机视觉运动分析的研究热点,其中第2章介绍了差分运动分析及基于核函数的视觉跟踪方法,第3章介绍了蒙特卡罗运动分析方法;第4章和第5章系统地介绍了近年在中层计算机视觉、运动分析中的研究热点,其中第4章介绍了铰链运动分析及人体姿态估计方法,第5章介绍了多目标跟踪算法,包括各种流派思想在两种典型应用中的代表性方法;第6章从基于学习的角度,介绍了几种重要的代表性跟踪方法。通过本书的阅读,读者可全面了解计算机视觉运动分析中最为前沿、重要的研究热点,如视频序列中铰链运动捕捉与分析,多目标的分析与跟踪诸方向的研究现状、方法及原理等。

本书结构合理,叙述深入浅出,可作为高等工科院校自动化类高年级本科生或硕士研究生教材,也可作为从事控制系统分析、设计的工程技术人员及其他相关专业的师生学习的参考书。

本书在编写过程中,石武楨和崔文学协助完成了部分文字、图表的编辑工作,在此表示感谢。

由于作者水平有限,难免存在疏漏或不妥之处,敬请读者批评指正。

作者
2018 年 4 月

目 录

第 1 章 从图像序列中估计 2D 和 3D 运动	1
1.1 运动场与光流场	1
1.2 光流计算	3
1.3 光流技术的研究难点及研究策略	13
1.4 3D 运动恢复	21
参考文献	29
第 2 章 差分运动分析及基于核函数的视觉跟踪	32
2.1 基于差分方法的光流计算	32
2.2 视觉跟踪概述	40
2.3 核函数跟踪算法	42
参考文献	54
第 3 章 蒙特卡罗运动分析	56
3.1 跟踪问题的形式化表示	56
3.2 卡尔曼滤波与广义卡尔曼滤波	57
3.3 特征表示与提取	60
3.4 目标跟踪方法评价指标	64
3.5 序列蒙特卡罗方法研究	65
参考文献	88
第 4 章 铰链运动分析及人体姿态估计	92
4.1 人体模型及观测似然函数	92
4.2 模拟退火的粒子滤波算法	94
4.3 非参数置信传播算法	96
4.4 人体运动估计	97
参考文献	105

第 5 章 多目标跟踪算法..... 107

5.1 运动场与光流场 107

5.2 数据关联算法 110

5.3 基于图模型的多目标跟踪算法 116

参考文献..... 142

第 6 章 基于学习的跟踪方法..... 145

6.1 背景知识 145

6.2 Semi-Boosting 跟踪方法 148

6.3 支持向量机跟踪方法 153

6.4 多示例学习跟踪方法 159

6.5 TLD 跟踪方法 167

参考文献..... 179

名词索引..... 180

第 1 章 从图像序列中估计 2D 和 3D 运动

运动图像分析是计算机视觉中的一项重要任务,其中的关键技术之一就是光流计算。光流不仅携带被观察物体的运动信息,还携带有被观察物的三维(3D) 结构、传感器参数、非刚性物体的局部弹性形变甚至流体运动的矢量结构特征等丰富的信息。通过光流,我们可以了解物体许多重要的运动特性,因为光流在运动图像计算与分析中扮演着非常重要的角色。

1.1 运动场与光流场

Gibson 于 1950 年首先提出光流的概念。运动可以用运动场来描述,反映真实世界的三维运动;光流场是运动场在二维(2D) 图像上的投影,携带有关运动和结构的丰富信息。研究光流场的目的是从图像序列中近似计算出不能直接得到的运动场。所谓光流是指图像中灰度模式的表面运动,它是物点的三维速度矢量在成像平面上的投影,表示物点在图像中位置的瞬间变化。心理学与神经生理学的大量实验表明,光流概念对认识人和动物的视觉感知机制原理具有重要的意义。

运动场由图像中每个点的运动(速度) 矢量构成。当目标在相机前运动或相机在一个固定的环境中运动时能够获得相应的图像变化,这些变化可用来恢复(获得) 相机和目标间的相对运动以及场景间多个目标的相互关系。在某一特定时刻,图像中像点 p_i 对应于目标表面上的点 p_0 ,如图 1.1 所示。

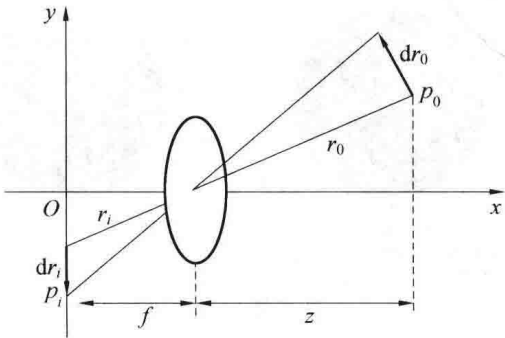


图 1.1 投影方程示意图

图中, f 为镜头焦距; z 为镜头中心到目标的距离; r_0 为物点到镜头中心的距离; r_i 为像点到镜头中心的距离。物点的运动矢量与其像点光流矢量根据投影关系联系在一起。假

定 p_0 相对于摄像机的运动速度为 v_0 , 则这个运动会产生导致图像上对应的像点 p_i 产生运动, 速度为 v_i , 则这两个速度分别为

$$v_0 = \frac{dr_0}{dt}, \quad v_i = \frac{dr_i}{dt} \quad (1.1)$$

其中 r_0 和 r_i 由下式联系:

$$\frac{1}{f} r_i = \frac{1}{r_0 \cdot z} r_0 \quad (1.2)$$

对式(1.2)求导, 即可得到赋给每个像素点的速度矢量, 而这些矢量就构成了运动场。

视觉心理学认为人与被观察物体间发生相对运动时, 被观察物体表面带光学特征部位的移动给人们提供了运动和结构的信息。当相机与场景目标间有相对运动时, 所观察到的亮度模式运动称为光流(Optical Flow), 或者说物体带光学特征部位的移动投影到视网膜平面(即图像平面)上就形成了光流。

光流表达了图像的变化, 它包含运动目标的信息, 可以用来确定观察者相对目标的运动情况。光流有 3 个要素: 一是运动(速度场), 这是形成光流的必要条件; 二是带光学特征的部分(如有灰度的像素点), 它可以携带运动信息; 三是成像投影(从场景到图像平面), 因而能被观察到。

在理想情况下, 光流与运动场相对应, 但也有不对应的时候, 即光流场并不一定反映了目标的实际运动情况。如图 1.2(a) 所示, 光源不动, 物体表面均一, 且产生了自转运动, 但没有产生光流; 如图 1.2(b) 所示, 物体没有运动, 但是光源与物体发生相对运动, 却有光流产生。三维物体的实际运动在图像上的投影称为运动场。如果已知目标的运动场, 就可以利用投影关系恢复目标的运动, 但是恢复目标的运动场比较困难, 用二维图像恢复三维信息, 许多信息缺失, 因此只能使用光流场来近似目标的运动场。

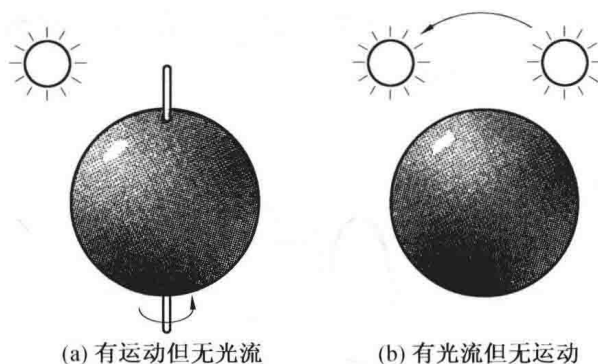


图 1.2 光流场与目标实际运动

光流包含了目标的重要信息。光流信息可以用来检测图像序列中的运动目标, 恢复目标的三维结构信息以及目标与摄像机之间的相对运动, 可以利用光流场的不连续性对图像进行分割, 还可用于机器人、自动导航及智能系统中。

光流法的核心就是求解出运动目标的光流, 即速度。根据视觉感知原理, 客观物体在

空间上一般是做相对连续运动的,在运动过程中,投射到传感器平面上的图像实际上也是连续变化的。为此可以假设瞬时灰度值不变,即灰度不变性原理,由此可以得到光流基本方程,即灰度对时间的变化率等于灰度的空间梯度与光流速度的点积。

1.2 光流计算

1.2.1 光流方程

光流可看作带有灰度的像素点在图像平面上运动而产生的瞬时速度场,Horn 和 Schunck 假设图像区域在时间和空间上都是连续且可导的,这是光流计算中的一个重要约束条件。对于一个图像序列,假设图像中一个像素点 (x, y) 在 t 时刻的亮度值为 $I(x, y, t)$,如果 $u(x, y)$ 和 $v(x, y)$ 表示 (x, y) 点处光流在 x 和 y 方向的运动分量,在足够小的一个时间 dt 内,点 (x, y) 移动到点 $(x + dx, y + dy)$,其中 $dx = udt, dy = vdt$,根据亮度恒定假设,即沿某运动轨迹曲线的各帧中相应的像素点具有相同的灰度值,即图像上对应点亮度不变,由此可得

$$I(x + dx, y + dy, t + dt) = I(x, y, t) \quad (1.3)$$

将式(1.3)按一阶 Taylor 公式展开可得

$$I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + O\partial^2 = I(x, y, t) \quad (1.4)$$

其中, $O\partial^2$ 是 dx, dy, dt 的二阶或二阶以上的高阶微量,可以忽略不计。考虑到速度关系, $dx = \frac{\partial x}{\partial t}dt = udt, dy = \frac{\partial y}{\partial t}dt = vdt$,将两边的 $I(x, y, t)$ 相抵消,整理可得

$$I_x u + I_y v + I_t = 0 \quad (1.5)$$

或

$$\nabla I \cdot \mathbf{V} + I_t = 0 \quad (1.6)$$

式中, I_x, I_y, I_t 表示点 (x, y) 的灰度值 $I(x, y, t)$ 分别沿 x, y, t 的偏导数,可通过图像序列求得; $\mathbf{V}$ 是速度场, $\mathbf{V} = (u, v)$; ∇I 是图像像素点 (x, y) 的空间灰度梯度。式(1.5)称为光流约束方程,或者称为光流基本等式,它表示图像的灰度空间梯度与光流速度之间的关系。在光流约束方程中, u 和 v 都是未知数,而且只有一个约束方程,所以不能唯一确定光流,需要加入其他约束条件才能确定 u 和 v 。

如果令 u 和 v 分别为二维坐标的横轴和纵轴,那么光流约束方程就是一条直线,所有满足该方程的 $\mathbf{V}$ 值都在这条直线上,如图1.3所示。

从图1.3中可以看出,约束直线与图像的灰度梯度垂直,因此只能决定梯度方向的分量,即等灰度轮廓的法线分量 $\mathbf{V}_n = f\mathbf{n}$,其中

$$\mathbf{n} = \frac{\nabla I}{\|\nabla I\|}, \quad f = \frac{-I_t}{\|\nabla I\|} \quad (1.7)$$

V_n 称为法线流,沿等灰度线轮廓的切线分量不能确定,这就是通常所说的孔径问题。因为光流约束方程只有一个,但有两个未知数,因此,需要增加其他约束才能解决这个问题。

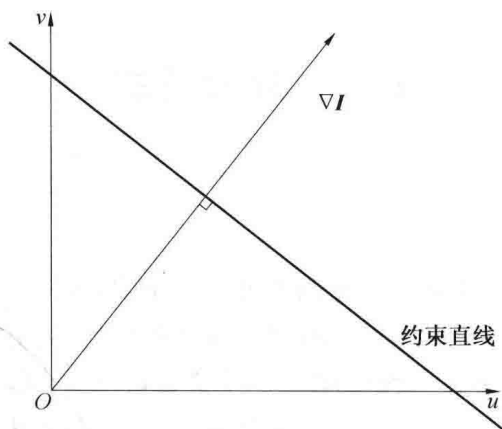


图 1.3 光流约束方程

1.2.2 经典光流计算方法

通过光流约束方程可以看出,光流约束方程只有一个,但有两个未知数,因此仅使用光流约束方程不能确定图像光流,需要引入其他约束条件,引入不同的约束条件,就会产生不同的光流计算方法。目前较为常用的光流计算方法有基于梯度的方法、基于匹配的方法、基于频域的方法和基于相位的方法。其中最常用的是基于梯度的光流算法,该方法也被称为微分法,主要是根据图像灰度的梯度函数得到图像中每个像素点的运动矢量。基于梯度的光流算法已得到广泛应用,本小节主要介绍基于梯度的光流算法中两种经典的光流算法:Horn - Schunck 算法和 Lucas - Kanada 算法。

1. Horn - Schunck 算法

Horn - Schunck 算法是在基于假设强度不变的条件下,引入全局光流平滑约束假设,假设在整个图像上光流的变化是光滑的,即物体运动矢量是平滑的或只是缓慢变化的,尤其对刚体来说,其各相邻像素点间的速度是相同的。Horn 和 Schunck 利用光流约束方程和全局光流平滑假设一起来计算 V ,该方法得到的是稠密光流场,由于相邻像素点运动速度相同,因而对于局部区域来说,其速度的空间变化率为 0,进而得出光流矢量的梯度接近于 0。Horn 引入的约束条件的基本思想是光流需要尽可能地平滑,使平滑约束项 E_s 极小化:

$$E_s = \iint \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right] dx dy \quad (1.8)$$

根据光流基本等式 $I_x u + I_y v + I_t = 0$,对光流误差 E_c 进行极小化,得

$$E_c = \iint (I_x u + I_y v + I_t)^2 dx dy \quad (1.9)$$

于是对光流场的求解可转化为对如下问题的求解:

$$\operatorname{argmin} \iint \left\{ (I_x u + I_y v + I_t)^2 + \lambda \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right] \right\} dx dy \quad (1.10)$$

其中 λ 是一个权重参数,如果导数 I_x, I_y, I_t 能比较精确地计算出来,那么 λ 可取较大值,否则 λ 取比较小的值。

将式(1.10)对 u 和 v 分别求导,可得

$$I_x^2 u + I_x I_y v = -\lambda^2 \nabla u - I_x I_t \quad (1.11)$$

$$I_y^2 v + I_x I_y u = -\lambda^2 \nabla v - I_y I_t \quad (1.12)$$

设 $\bar{u}$ 和 $\bar{v}$ 分别表示 u 邻域和 v 邻域中的均值,令 $\nabla u = u - \bar{u}, \nabla v = v - \bar{v}$,则式(1.11)和式(1.12)分别变为

$$(\lambda^2 + I_x^2) u + I_x I_y v = -\lambda^2 \bar{u} - I_x I_t \quad (1.13)$$

$$(\lambda^2 + I_y^2) v + I_x I_y u = -\lambda^2 \bar{v} - I_y I_t \quad (1.14)$$

将式(1.13)和式(1.14)联立可得

$$u = \bar{u} - I_x \frac{I_x \bar{u} + I_y \bar{v} + I_t}{\lambda^2 + I_x^2 + I_y^2} \quad (1.15)$$

$$v = \bar{v} - I_y \frac{I_x \bar{u} + I_y \bar{v} + I_t}{\lambda^2 + I_x^2 + I_y^2} \quad (1.16)$$

根据松弛迭代法可得

$$u^{(n+1)} = \bar{u}^{(n)} - I_x \frac{I_x \bar{u}^{(n)} + I_y \bar{v}^{(n)} + I_t}{\lambda^2 + I_x^2 + I_y^2} \quad (1.17)$$

$$v^{(n+1)} = \bar{v}^{(n)} - I_y \frac{I_x \bar{u}^{(n)} + I_y \bar{v}^{(n)} + I_t}{\lambda^2 + I_x^2 + I_y^2} \quad (1.18)$$

式(1.18)和式(1.17)即为 Horn - Schunck 光流计算方法。光流的初始估计 u_0 和 v_0 。在实际的应用中通常都取 0, n 为迭代次数,当相邻两次迭代结果小于预定的某一值时,迭代过程终止。一般至少需要迭代 20 次,才能求出精度较高的光流值。

2. Lucas - Kanada 算法

Lucas - Kanada 算法即 L - K 光流法,最初于 1981 年被提出。该算法假设在一个小的空间邻域内运动矢量保持恒定,使用加权最小二乘法估计光流。由于该算法应用于输入图像的一组点上时比较方便,因此被广泛应用于稀疏光流场。Lucas - Kanada 算法的提出是基于以下 3 个假设:

(1) 亮度恒定不变。

目标像素在不同帧间运动时外观上是保持不变的,对于灰度图像,假设在整个被跟踪期间,像素亮度不变。

(2) 时间连续或者运动是“小运动”。

图像运动相对于时间来说比较缓慢,实际应用中指时间变化相对图像中运动比例要

足够小,这样目标在相邻帧间的运动就比较小。

(3) 空间一致。

同一场景中同一表面上的邻近点运动情况相似,且这些点在图像上的投影也在邻近区域内。

在一个小的邻域内,通过对下式的加权平方和最小化来估计 V :

$$\sum_{(x,y \in \Omega)} W^2(X) (I_x u + I_y v + I_t)^2 \quad (1.19)$$

式中, $W^2(X)$ 为一个窗口权重函数,该函数使得邻域中心的加权比周围的加权大。对于 Ω 内的 n 个点 $X_1, X_2, \dots, X_n$, 设 $v = (u, v)^T$, $\nabla I(X) = [I_x, I_y]^T$, 式(1.19) 的解可用最小二乘法得

$$A^T W^2 A v = A^T W^2 b \quad (1.20)$$

式中

$$A = [\nabla I(X_1), \dots, \nabla I(X_n)]^T \quad (1.21)$$

$$W = \text{diag}[W(X_1), \dots, W(X_n)] \quad (1.22)$$

$$b = - \left[\frac{\partial I(X_1)}{\partial t}, \dots, \frac{\partial I(X_n)}{\partial t} \right]^T \quad (1.23)$$

最后求得式(1.20) 的解为

$$V = (A^T W^2 A)^{-1} A^T W^2 b \quad (1.24)$$

Lucas - Kanada 光流法得出的 V 估计值的可靠性由 $A^T W^2 A$ 的特征值来决定,而它的特征值又由图像空间梯度的大小来决定。设其特征值为 λ_1 和 λ_2 , 且 $\lambda_1 \geq \lambda_2$, 并设定一个阈值 τ , 如果 $\lambda_1 \geq \lambda_2 \geq \tau$, 那么可以求解出 V ; 如果 $\lambda_2 = 0$, 那么矩阵 $A^T W^2 A$ 是奇异的, 不能计算光流; 如果 $\lambda_1 \geq \tau$ 而 $\lambda_2 \leq \tau$, 那么无法得到 V 的完整信息, 只能得到光流的法线分量。

Horn - Schunck 算法和 Lucas - Kanada 算法都是比较经典的光流算法, Horn - Schunck 光流法属于稠密光流算法的一种, 实际上稠密光流在计算时并不容易, 例如对于一张白纸的运动来说, 当前帧中白色像素在下一帧中仍为白色, 只有边缘部分的像素才会产生变化, 稠密光流法对于那些运动不明确的像素, 通常在比较容易跟踪的像素之间采用某种差值法来解决这个问题, 因此稠密光流有相当大的计算量。而 Lucas - Kanada 光流法就是一种稀疏光流法, 它计算的是一组指定的点, 很好地解决了计算量大这个问题, 因而得到广泛应用。由于 Lucas - Kanada 光流法是利用点实现目标跟踪的, 因此它存在一定的缺陷, 后来人们对该光流法进行了改进。

1.2.3 光流法的改进

当多个相同目标存在时, 光流法依然可以跟踪其选定目标并进行分析。但是光流法本身都是在一定的假设之上成立的, 所以有很多的限制条件, 比如图像序列目标的特性、场景中照明、光源的变化以及目标运动速度的影响等多种因素影响光流算法的有效

性。具体到某种光流法,又会加上一定的约束条件,比如对于 Lucas - Kanada 光流法来说,不能跟踪运动尺度较大的目标。因此后来人们提出了很多关于光流法改进的方法。

1. 金字塔 Lucas - Kanada 光流法

Lucas - Kanada 光流法本身存在一个固有缺陷,即该算法不能准确跟踪在相邻帧间运动尺度过大的目标,只适用于小图像灰度模式的运动估计,其原因在于光流估计方法的两个应用条件:① 灰度守恒约束:任何物体点的观测亮度随时间变化是恒定的;② 速度平滑性约束:图像平面内的邻近点移动方式相似。在实际应用中,当目标运动速度比较大时,难以满足光流估计方法的约束条件。在现实生活中,普遍存在运动尺度大且不连贯的运动,而该算法由于此原因在实际应用中的跟踪效果并不是很理想,对于大的运动通常需要大的窗口来捕获,但是大的窗口又会使运动连贯的假设不成立,而图像金字塔能够解决此问题,即先在较大的空间尺度上对目标进行跟踪,然后通过对图像金字塔降级采样,直到满足约束条件对图像像素处理来修正初始运动速度的假定。

图像金字塔是一个图像集合,该集合中全部图像都源于同一个原始图像,而且通过对原始图像连续降级采样获得,直到达到设定的终止条件才停止。图像的金字塔结构是一种多分辨率表示方法,不同金字塔分辨率层可分析不同尺度的目标。

基于图像金字塔的光流基本思想是:首先在金字塔的最高层计算光流,然后用所得结果作为下一层的起始点,重复这个操作直到图像分解到一定小的程度,即满足光流计算的约束条件,就可以直接进行光流估计。这样可大大降低不满足运动假设的可能性,从而实现快速运动目标的跟踪,该方法称为金字塔 Lucas - Kanada 光流法(LKP)。金字塔图像示意图如图 1.4 所示。

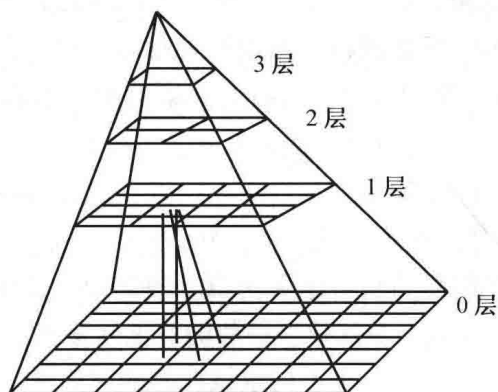


图 1.4 金字塔图像示意图

金字塔基层 $f=0$ 为原始图像,假设基层上目标像素在相邻帧间的运动距离是 D , $f=1$ 层图像通过将原图像降采样至原尺寸的 $1/2^N$ (一般情况下 $N=1$)获得,则 $f=1$ 层中相邻帧间目标像素的运动距离为 $D/2^N$ 。按照这样的计算方法,金字塔中任一 $(f+1)$ 层($f \geq 1$)图像都可通过 f 层图像按相同比例减采样得到。当 f 达到某一设定值时(设定值一般为 $3 \leq f \leq 5$),最高层相邻帧间的目标运动距离可达到亚像素级,此时满足光流约束条件,利用 Lucas - Kanada 算法可进行光流估计。其简化流程图如图 1.5 所示。

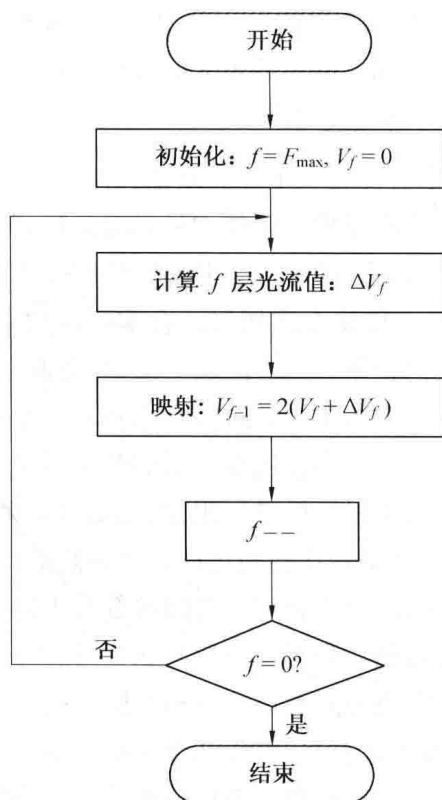


图 1.5 金字塔 Lucas - Kanada 光流法简化流程图

2. 能量法

能量法也称为频率法,该方法设计了一组频率调谐滤波器,图像区域的移动速度与某一滤波器的调谐速度越相近,该滤波器的输出值就越高。设图像 $I(x, y, t)$ 在点 (x, y) 处的运动速度为恒定值 (u, v) , 则图像 $I(x, y, t)$ 可由零时刻图像表示为

$$I(x, y, t) = I(x + ut, y + vt, 0) \quad (1.25)$$

对式(1.25)进行傅里叶变换,得

$$\hat{I}(\zeta, \eta, \omega) = \hat{I}_0(\zeta, \eta) \delta(\omega + \zeta u + \eta v) \quad (1.26)$$

式中, ω 是三维傅里叶变换的时间频率; ζ 与 η 是三维傅里叶变换的空间频率; $\hat{I}_0(\zeta, \eta)$ 是 $I(x, y, 0)$ 的傅里叶变换; $\delta(\zeta, \eta)$ 是狄拉克函数的傅里叶变换。

Heeger 提出的能量法构建了 12 个不同空域比例的滤波器,分别对应不同的空域方向与时间频率。在只有平移运动的理想情况下,这些滤波器的响应将集中于某一频率平面上。与运动图像的时空频率 (ω, ζ, η) 相对应的 Gabor 能量滤波器的响应,在图像中只有白噪声时,可表示为

$$R(u, v) = \left[\frac{-4\pi^2 \sigma_x^2 \sigma_y^2 \sigma_t^2 (\omega + \zeta u + \eta v)}{(u\sigma_x \sigma_t)^2 + (v\sigma_y \sigma_t)^2 + (\sigma_x \sigma_y)^2} \right] \quad (1.27)$$

式中, σ_x, σ_y 与 σ_t 为 Gabor 滤波器的高斯标准差。

3. 相位法

Fleet与Jepson于1990年首次提出将图像相位应用于光流场计算,由于相位法较传统的差分法有更好的精确度与稳定性,因而引起了部分学者的关注,并随之提出了其他相位光流法。

Fleet提出的相位法定义了代表瞬时运动的速度分量与带通速度调谐滤波器输出的等相位轮廓相垂直,利用带通速度调谐滤波器根据尺度、速度和方向来分离输入信号。图像作为输入信号,通过带通速度调谐滤波器的响应为

$$R(x, y, t) = \rho(x, y, t) \exp[i\varphi(x, y, t)] \quad (1.28)$$

式中, $\rho(x, y, t)$ 与 $\varphi(x, y, t)$ 分别是 $R(x, y, t)$ 的幅度和相位。由于图像中的二维速度位于等相位轮廓的法线方向上,因此可由式(1.29)估计运动矢量。

$$V_n = sn, \quad s = \frac{-\varphi_t(x, y, t)}{\|\nabla\varphi(x, y, t)\|}, \quad n = \frac{\nabla\varphi_y(x, y, t)}{\|\nabla\varphi(x, y, t)\|} \quad (1.29)$$

式中, $\nabla\varphi(x, y, t) = [\varphi_x(x, y, t), \varphi_y(x, y, t)]^T$ 。由此可以发现,相位法实际是对图像相位而非图像亮度进行的差分。实际上,相位法与等相位相关技术一致。图像的相位梯度为

$$\begin{cases} \varphi_x(x, y, t) = \frac{\text{Im}[R^*(x, y, t)R_x(x, y, t)]}{|R(x, y, t)|^2} \\ \varphi_y(x, y, t) = \frac{\text{Im}[R^*(x, y, t)R_y(x, y, t)]}{|R(x, y, t)|^2} \end{cases} \quad (1.30)$$

式中, R^* 是 R 的共轭复数。

4. 彩色法

在基于BCM模型进行的光流场估计中,为了克服孔径问题,需要引入附加的约束条件。若用彩色图像的各颜色分量构建光流场模型,则可通过丰富的色彩信息克服孔径问题。1989年,Ohta首次提出将彩色图像上的颜色模型应用于BCM模型;其后,Lai利用颜色信息研究非形变物体的运动;Golland通过实验分析了基于颜色恒定假设的光流场模型在不同光照条件下的鲁棒性;Barron通过实验证实了利用颜色信息可获得更精确的光流场;Andrew将色彩信息与Lucas-Kanade方法相结合求解光流场;Madjidi将改进的彩色光流法应用于水下图像;Arshad借助可测距的移动机器人与色彩信息获得光流场矢量;Simon结合彩色光流场与运动分割技术提高了光流场的准确度。随着彩色图像应用面的不断扩大,彩色光流法的实现具有现实条件与意义。

对色彩系统的分析是彩色光流法的基础。生物学家研究发现,人的视觉系统具有3种类型的亮度检测器,因此常规的摄像机也设置为3种类型的颜色检测器。当人眼获得相同的光谱输入 $S(\lambda)$ 时,若检测器的灵敏度函数 $D_R(\lambda)$ 、 $D_G(\lambda)$ 、 $D_B(\lambda)$ 不同,则视觉系统可获得不同颜色的图像。人眼中的亮度检测器获得的3种输出可近似表示为

$$R = \int_{\Omega} S(\lambda) D_R(\lambda) d\lambda, \quad G = \int_{\Omega} S(\lambda) D_G(\lambda) d\lambda, \quad B = \int_{\Omega} S(\lambda) D_B(\lambda) d\lambda \quad (1.31)$$

式中,光的波长 λ 的取值范围是 Ω ,具体值在 $[400 \text{ nm}, 700 \text{ nm}]$ 内。

设物体表面的光照波长为 λ , 物体表面点为 $r = (x, y, z)$, 则点 r 获得的光照用表面光照能量函数 $I(\lambda, r)$ 表示, 点 r 上的反射光照能量函数以 $\hat{I}(\lambda, r)$ 表示, $\hat{I}(\lambda, r)$ 一般不等于入射光。假设物体表面某点 r 的反射光照能量函数 $\hat{I}(\lambda, r)$ 仅由物体材料和点 r 所在表面的几何特性决定, 则 $\hat{I}(\lambda, r)$ 可描述为

$$\hat{I}(\lambda, r) = R(\lambda, \varphi, \theta, \gamma, r) I(\lambda, r) \quad (1.32)$$

式中, φ, θ, γ 分别为入射光角度、观测者角度及相位角度; $R(\lambda, \varphi, \theta, \gamma, r)$ 为表面反射函数, 由其完全定义点 r 处的表面反射特性。实验表明, $R(\lambda, \varphi, \theta, \gamma, r)$ 可表示为光谱因子 $\rho(\lambda, r)$ 与点 r 处的几何因子 $c(\varphi, \theta, \gamma, r)$ 的乘积, 其中光谱因子 $\rho(\lambda, r)$ 表示物体表面的颜色特征, 几何因子 $c(\varphi, \theta, \gamma, r)$ 表示物体表面的几何特征, 则式(1.32)可改写为

$$\hat{I}(\lambda, r) = \rho(\lambda, r) c(\varphi, \theta, \gamma, r) I(\lambda, r) \quad (1.33)$$

将式(1.31)与式(1.33)结合, 将亮度检测器的3种输出重新表示为

$$\begin{cases} R = \int_{\Omega} c[\varphi, \theta, \gamma, \rho(\lambda)] I(\lambda) D_R(\lambda) d\lambda \\ G = \int_{\Omega} c[\varphi, \theta, \gamma, \rho(\lambda)] I(\lambda) D_G(\lambda) d\lambda \\ B = \int_{\Omega} c[\varphi, \theta, \gamma, \rho(\lambda)] I(\lambda) D_B(\lambda) d\lambda \end{cases} \quad (1.34)$$

根据图像进行计算时, 不能保证获得摄像机的检测器参数, 此时可由颜色分量的比值代替 RGB 系统的具体值。由此获得彩色光流场的颜色守恒方程为

$$\begin{cases} \frac{\partial F_1}{\partial x} u + \frac{\partial F_1}{\partial y} v + \frac{\partial F_1}{\partial t} = 0 \\ \frac{\partial F_2}{\partial x} u + \frac{\partial F_2}{\partial y} v + \frac{\partial F_2}{\partial t} = 0 \end{cases} \quad (1.35)$$

式中, F_1 与 F_2 由图像采用的颜色模型决定。若为 RGB 系统, 则 F_1, F_2 可为规范化 RGB 中的任意两个分量。

在式(1.35)中有两个方程, 理论上可获得 u, v 的唯一解。然而当方程组线性相关或颜色分量间有相同变化时, 仍存在孔径问题。

5. 几何代数域法

光流概念与三维场景在二维平面上的投影有关, 且光流模型涉及时空关系, 因而可在几何代数域中分析、求解光流场。由于几何代数中的张量光流分析法不易受噪声影响, 因此受到越来越多学者的关注。

在基于结构张量的方法中, 将局部光流计算式改写为

$$\begin{aligned} E(u, v) &= \sum_{(x, y) \in \Omega} W^2(x, y) (I_x u + I_y v + I_t)^2 \\ &= \sum_{(x, y) \in \Omega} W^2(x, y) [\nabla_3 I(x, y, t) \mathbf{v}_3^T]^2 \\ &= \mathbf{v}_3 \sum_{(x, y) \in \Omega} W^2(x, y) [\nabla_3 I(x, y, t)^T \nabla_3 I(x, y, t) \mathbf{v}_3^T] \mathbf{v}_3^T \end{aligned}$$