

基于纹理约束和参数化运动模型的光流估计

杨 波 徐光祐

(清华大学计算机系人机交互与媒体集成研究所, 北京 100084)

摘 要 提出了一种基于局部小平面运动的光流估计新方法。目的是获得精确致密的光流估计结果。与以往采用亮度一致性区域作为假设平面的算法不同, 本算法利用序列图像的纹理信息, 在纹理分割区域的基础上, 进行运动估计。该算法首先通过微分法计算粗光流, 可以得到参数化光流模型的初始估计, 然后通过区域迭代算法, 调整初始估计, 从而得到精细的平面分割及其对应的参数化光流模型。基于纹理信息的部分拟合算法被用于算法的每一步当中, 保证了纹理边缘位置的光流估计值的准确性。实验采用了标准图像序列, 结果表明, 可以得到更为精细的光流估计结果, 特别是对于那些有着丰富纹理信息的室外环境的图像序列, 而且在运动边界处的结果改善尤为明显。

关键词 光流 分割 纹理约束 参数化光流模型 部分拟合

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2004)06-0705-07

Optical Flow Estimation Based on Texture Constraints and Parametric Motion Model

YANG Bo, XU Guang-you

(Department of Computer Science and technology, Tsinghua University, Beijing 100084)

Abstract This paper presents a new method for estimating dense and accurate optical flow based on the motion of planar regions. Different from previous approach, textural information are exploited to organize and constrain the interpretation of the motion. Regions with similar texture properties are supposed to be planar patches. This hypothesis is more reasonable and effective than those hypotheses using the brightness smooth region as planar patches. Parametric flow models are estimated in these textually consistent regions in two steps that first compute a coarse fit according to the coarse optical flow. Then this initial fit is refined using a generalization of the standard area-based regression approaches. Partial fit based on the texture information is employed all through the process. Accurate optical flows are computed finally according to the refined parametric motion models. Experimental results on a variety of images sequence indicate that our method produce accurate flow estimates especially in outdoor environment containing rich texture content. Furthermore the incorporation of the textural information through partial fit and image segmentation provides precise localization of motion boundaries and makes this method overmatch the previous brightness smooth hypothesis based approaches.

Keywords optical flow, segmentation, texture constraint, parameterized flow models, partial fit

1 引 言

光流估计是图像序列分析中的基本问题之一, 目的就是希望通过图像亮度的时空纹理来估测图像的二维运动场。在光流估计的基础上, 可以进行许多进一步的研究, 包括被动的景物分析或者主动的导航等等。在这些研究工作中, 大多都要求有精确而致

密的光流估计结果作为前提条件。

现有的光流估计算法一般依赖于某种平滑假设^[1], 但该假设并不是总能满足, 特别是当景物中存在明显的深度变化、独立运动物体或不规则运动物体的时候。同时预处理中微分计算的精确性, 相邻帧图像运动较大等等也是光流估计中误差产生的原因。因此, 提出了许多改善光流估计精确性的方法。

解决光流估计误差的方法大致可以分为 3 类:

(1) 认为误差主要由于光流约束中的微分计算引起, 从而通过修正光流约束方程中的微分项来改善光流估计的精度;

(2) 认为局部区域平滑假设在运动边界不成立, 从而导致传统的最小二乘方法估测得到的光流不准确, 解决方法是通过引入鲁棒算法来改善光流估计的精度;

(3) 也是认为局部区域平滑假设在运动边界不成立, 解决方法是引入了参数化运动模型约束的概念和预分割的概念, 通过预分割区域上的参数化运动模型来约束光流估计。

实际图像中, 局部区域内像素的光流可能不同是微分法估算光流误差的主要来源, 而不是所谓的微分导致的误差, 因此, 第1类方法没有解决根本的问题。第2类方法找到了光流估计误差的主要来源, 采用的方法包括多分辨率法、最小中值二乘法 LmedS、邻域投票法等等。

本文提出的解决方法可以归为第3类。Ju 等人认为^[2], 如果可以将邻域设定为做一致运动物体的区域, 同时平滑的概念也由邻域光流值相近改为符合某种运动一致约束, 那么这些预分割的局部区域就能够对参数化运动模型估计提供很强的约束, 从而有助于提高光流估计的精确性。具体的实现中, 他们采用基于亮度分割结果的区域平滑约束和基于刚体的参数化运动模型约束^[2,3]。在此之前, 也有类似的方法^[4], 但使用的是均匀划分的图像区域。

可以看到, 第3类方法最能反映图像中运动的实际情况, 即邻域平滑约束应该限制在一个刚体的区域内。此外, 刚体内平滑处理也不是简单的邻域相似, 而是改为符合某种参数化运动模型, 从而能够处理同一物体由于深度不同导致的光流不同的情况, 如地面远近位置的光流变化。在此类方法中, Black 等人使用图像的亮度信息来进行预分割^[3], 他们认为: “图像上亮度一致的区域很可能对应于真实世界中的平面区域”。而本算法的前提假设为纹理一致性假设, 即“图像上纹理特性一致的区域很可能对应于真实世界中的平面区域”。研究发现, 文献^[3]中的亮度一致性假设与本文提出的纹理一致性假设相比, 在运动估计上存在某些缺陷。因为首先纹理区域通常是由一些很小的亮度一致的块构成, 如树叶、碎石等等, 而这些亮度一致的小区域在其对应尺度下不一定是平面的, 而且用来估算光流的信息量也太少。其次纹理区域包含有许多对运动估计有用的信息, 而亮度平滑的区域在运

动估计时, 只有区域边缘的那些像素可以用来进行运动估计, 更为不利的是这些地方的噪声又较大。综上所述, 这些内在的不足要求在运动估计中使用更为有效的纹理一致性平面假设。

2 算法基础

本算法分低层和高层两步来进行。低层的处理主要包括对图像进行预分割以及使用微分法来估计粗光流。而高层的处理主要是分割区域的参数化运动模型的估计。

2.1 光流估计

一般说来, 现有的光流估计方法可以归结为四大类: (1) 微分方法, 通过计算原图像或滤波(低通或带通)图像的时空偏导数来计算光流; (2) 区域匹配方法, 将光流 v 定义为能够在不同时间的图像区域间得到最佳匹配的偏移量, 并使相似性度量最大来得到最佳匹配; (3) 基于能量技术或者频域技术, 主要是基于在傅立叶变换空间设计的滤波器的能量输出来计算光流; (4) 基于相位的方法。其中, 微分法由于较为容易实现, 而且估计精度高, 所以使用最广泛^[1,4]。基于最小二乘的微分方法大致如下:

假设图像平面上 $x = (x, y)$ 位置处像素的亮度为 $I(x, y, t)$, 且假设景物中特定点的亮度在景物运动的时候不变, 于是有

$$\frac{dI}{dt} = 0 \quad (1)$$

偏微分展开后有

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

这就是局部光流约束方程。设 $v = (v_x, v_y) = \left(\frac{dx}{dt}, \frac{dy}{dt} \right)$, 通过在每个局部区域 Ω 用加权最小二乘法使以下目标函数最小, 可以估计出局部光流 v

$$\sum_{x \in \Omega} W^2(x) [\nabla I(x, t) \cdot v + I(x, t)]^2 \quad (3)$$

这里 $W(x)$ 是用作权值的窗口函数, 以使中间位置的点较边缘位置的点在拟合中所起的作用更大。方程(3)的解由下式给出

$$A^T W^2 A v = A^T W^2 b \quad (4)$$

给定某一时刻 t , 并设局部区域内有 n 个点 $x_i \in \Omega$, 则式(4)中各个变量分别为

$$A = [\nabla I(x_1), \dots, \nabla I(x_n)]^T \quad (5)$$

$$W = \text{diag}[W(x_1), \dots, W(x_n)] \quad (6)$$

$$b = - [I_x(x_1), \dots, I_x(x_n)]^T \quad (7)$$

假设 $A^T W^2 A$ 非奇异, 则方程(4)的解为

$$v = [A^T W^2 A]^{-1} A^T W^2 b \quad (8)$$

其中,

$$A^T W^2 A = \begin{bmatrix} \sum W^2(x) I_x^2(x) & \sum W^2(x) I_x(x) I_y(x) \\ \sum W^2(x) I_x(x) I_y(x) & \sum W^2(x) I_y^2(x) \end{bmatrix}$$

所有求和均在邻域 Ω 内进行。

2.2 参数化运动模型估计

给定了粗光流估计结果, 同时假设已经将图像分割为一些纹理特性一致的区域 P_i , 就可以估测这些分割区域的参数化运动模型。一般情况下, 假设这些纹理一致性区域对应于真实世界中的平面区域, 特别是在室内环境和较大的室外环境。在这种情况下, 图像运动满足 6 参数仿射模型^[5]

$$v_x(x, y) = a_{x0} + a_{xx}x + a_{xy}y \quad (9)$$

$$v_y(x, y) = a_{y0} + a_{yx}x + a_{yy}y \quad (10)$$

其中 v_x 和 v_y 是光流的 x, y 分量, a_i 是仿射模型的变换参数。可以采用标准的最小二乘法来估计这些局部区域的仿射参数。因为仿射模型的 x 参数只与光流的 x 分量有关, 而 y 参数只与光流的 y 分量有关, 所以最小二乘拟合估计可以对每一个运动分量独立进行。令 $a_i^T = [a_{x0i} \ a_{xxi} \ a_{xyi} \ a_{y0i} \ a_{yxi} \ a_{yyi}]$ 表示第 i 个局部区域分别对应于 x, y 分量的仿射模型的 6 个参数, $\varphi^T = [1 \ x \ y]$ 为迭代变量, 则运动场方程(9)和(10)可以写为

$$v_x(x, y) = \varphi^T a_{xi} \quad (11)$$

$$v_y(x, y) = \varphi^T a_{yi} \quad (12)$$

因此, 给定运动场的 6 参数仿射运动模型 a_i 的线性最小二乘估计为

$$[a_{xi}, a_{yi}] = \left[\sum_{P_i} \varphi \varphi^T \right]^{-1} \sum_{P_i} (\varphi [v_x(x, y) \ v_y(x, y)]) \quad (13)$$

其中, 所有求和均在局部区域 P_i 内进行。

3 基于纹理特征的图像分割和粗光流估计

3.1 基于分形纹理特征的图像分割

Black 等人使用亮度平滑区域作为假设平面区域^[3]。但是, 对于那些包含强纹理的区域, 例如草地、花圃、干枯的河床以及崎岖不平的地面等等, 基于这种亮度一致性的假设进行分割会产生许多细小的区

域。而这些小区域一般情况下不满足平面假设的要求。此外, 当使用某种基于邻域的方法在一个有噪声的区域内估计运动模型时, 由于这些分割区域包含的像素不够多, 使得难以得到稳定而又准确的估计结果。本算法采用纹理一致性准则作为图像分割和平面假设的基础, 它较亮度平面假设有以下几点改进:

(1) 不会将强纹理区域分割成为许多细小的区域, 这样产生的纹理平滑区域内包含的像素足够多, 适于在其上进行最小二乘估计。

(2) 由纹理一致性准则分割得到的区域比使用亮度一致性约束得到的分割区域包含有更多的变化。这就更利于运动估计。相反, 对于那些由亮度一致性约束得到的分割区域, 只有区域的边缘部分才能用来进行运动估计, 而这些区域的边缘部分经常存在较大的噪声, 从而影响了运动估计的准确性。

(3) 当利用纹理特征进行图像分割时, 使用分形特征作为主要的纹理特征。对于基于仿射模型的运动估计来说, 分形纹理特征有一个特殊优点: 分形纹理特征对于纹理基元的大小变化不敏感^[6,7]。这就意味着对于同一块纹理, 本算法不会仅仅因为距摄像机的距离不同而将同一块纹理分割为由近及远的多个区域。而纹理基元由于远近不同而大小不同在真实世界中是常见的, 例如平铺开的地面。

算法中使用的纹理特征包括分形维数, Lacunarity, 以及其他一些常用的纹理特征例如对比度等等。分形纹理特征估计方法和基于纹理特征的图像分割算法的详细叙述参见文献^[7]。

3.2 基于部分拟合的粗光流估计

本算法采用微分法来估计粗光流。对于每一个局部区域, 使用最小二乘法, 基于方程(5)来估计光流。在粗光流估计中, 有运动平滑假设, 即此局部区域对应于真实世界中的一个作刚体运动的平面。以往的许多方法使用预先定义的具有相同形状和大小的矩形区域作为局部区域^[4]。这样, 就会有不满足假设的情况出现。本算法中用来计算光流的局部区域由纹理分割得来, 同时算法中也结合了部分拟合技术, 以便改进纹理边缘处光流估计。

首先, 在每一个像素位置计算纹理特征, 之后, 基于这些特征分割图像。在粗光流估计时, 首先将图像分割为一系列具有相同大小的互不重叠的区域。区域的尺寸应该尽可能的小, 以使光流估计限制在一个小的局部区域内, 同时避免光流的估计在运动边缘进行。另一方面, 较大的局部区域在有噪声的情况下能

够提供更为精确的参数估计。因此,需要将局部区域的初始大小设定在一个合理的区域内。但是,对于那些具有较多运动边缘的图像序列,光流估计仍旧有可能在跨运动边缘的情况下进行。本算法使用部分拟合算法解决这一问题。所谓部分拟合就是在局部区域内,只有那些与区域中心像素具有相似纹理特征的像素才被用来进行运动估计,局部区域内的其他像素不参与估计光流时的最小二乘计算。部分拟合方法减少了其他临近纹理区域内的像素对本区域光流估计的影响,提高了纹理边缘位置处的光流估计结果。

4 基于参数化运动模型的精确光流估计

要得到致密而精确的光流估计结果,即要求在图像的每个像素上都得到精确的光流估计结果,传统的逐点计算的方法会面临计算量大的问题。此外,传统的方法还需要一定的邻域光流平滑处理,这就有可能导致在运动边界处的光流估计产生误差。使用参数化运动模型,一个独立运动物体只用一个参数化运动模型来表征其运动,光流值则通过此运动模型来估算。这样既减少了计算量,而且由参数化运动模型计算光流的连续性也可以保证邻域光流平滑约束的满足。

4.1 基于直接法估计参数化运动模型

由2.2节可知,参数化模型 θ 可以通过粗光流来估计(其中 a_i 为6参数仿射运动模型)。但是,由于粗光流估计的不精确,影响了这些参数化运动模型的估计。因此有必要直接根据相邻2帧图像灰度匹配后的残差最小化来校正前面估计得到的运动模型^[8]。直接法估计运动模型的最小化目标函数如下:

$$\begin{aligned} \min_{\theta} \sum_i \rho(r_i; \sigma) \\ r_i = I(p_i, t) - I(p_i - u(p_i; \theta), t - 1) \quad (14) \\ \rho(r; \sigma) = \frac{r^2}{r^2 + \sigma^2} \end{aligned}$$

其中, $\rho(r; \sigma)$ 是基于一个局部区域内的残差 r 定义的目标函数, σ 是尺度因子, p_i 表示第 i 个图像像素。选择这个目标函数是因为:首先它避免了使用平方误差最小,即最小二乘法拟合时噪声的影响,对于那些 $|r_i| \geq \sigma/\sqrt{3}$ 的点,可以被认为是外点而排除在最小化计算之外从而提高了算法的鲁棒性;其次,对于内点, $\frac{\dot{\rho}(r)}{r} = \frac{2\sigma^2}{(\sigma^2 + r^2)^2}$ 随着 r_i 的增大而迅速变小,而它是后面高斯牛顿法等式右边残差的权值,有

利于减小误差点的影响。一般情况下, σ 可以取值为 $\sigma = 1.4826 \times \text{median} |r_i|$ ^[8]。本算法使用高斯-牛顿法来解最小化问题。相应像素位置 i 处的图像梯度和Hessian矩阵可计算如下:

$$\begin{aligned} g_k &= \sum_i \frac{\partial \rho}{\partial r_i} \frac{\partial r_i}{\partial \theta_k} \\ H_{kl} &= \sum_i \frac{\partial^2 \rho}{\partial r_i^2} \frac{\partial r_i}{\partial \theta_k} \frac{\partial r_i}{\partial \theta_l} \end{aligned} \quad (15)$$

其中, k, l 是参数的序号。因为式(14)的 ρ 不是一个平方形式,所以 $\frac{\partial^2 \rho}{\partial r^2}$ 有可能取负值,这样在牛顿迭代时就有可能得到一个非下降的方向。因此一般使用 $\frac{\dot{\rho}(r)}{r}$ 作为 $\frac{\partial^2 \rho}{\partial r^2}$ 的近似,以保证它处处为正。这样,高斯牛顿方程就变为

$$\sum_i \sum_i \frac{\dot{\rho}(r_i)}{r_i} \frac{\partial r_i}{\partial \theta_k} \frac{\partial r_i}{\partial \theta_l} \delta \theta_k = - \sum_i \frac{\dot{\rho}_i}{r_i} r_i \frac{\partial r_i}{\partial \theta_k} \quad (16)$$

其中, k, l 是运动模型参数的序号, i 是图像像素的序号。本算法采用仿射运动模型,也就是假设图像的每一部分均可以被看作是真实世界中的一个平面区域的投影。在平面假设和6参数仿射运动模型的理想情况下,可以估计景物中的物体表面所有的刚体运动。方程(16)中的 $\frac{\partial r}{\partial \theta}$ 的计算方法如下:

给定光流场与运动模型的对应公式

$$u(p(x, y); \theta) = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix} \theta = M \theta \quad (17)$$

在高斯迭代算法中,方程(17)给出了 $\delta \theta$ 和 δu 之间的线性关系 $\delta u(p; \theta) = M \delta \theta$ 。因此,在每一次迭代中残差 r 对于未知的 θ 的偏微分为

$$\frac{\partial r_i}{\partial \theta} = \frac{\partial \delta u}{\partial \theta} \frac{\partial r_i}{\partial \delta u} = M^T \nabla I(p, t) \quad (18)$$

将上式与方程式(16)结合在一起,就可以得到每一步迭代新的高斯牛顿方向。之后,再根据初始的运动模型参数,通过线性搜索,就可以得到一个新的更为精确的运动模型 θ 。

4.2 精确光流估计

当通过迭代处理得到较为精确的分割区域的参数化运动模型之后,便可以直接由此6参数运动模型,通过方程式(9)和式(10)计算出每个像素位置处的光流。这里已不再需要在邻域内对计算得到的光流作进一步的平滑处理,因为方程式(9)和式(10)对于 x, y 坐标是连续的,这样计算得到的某像素位置处的光流相当于已经与该像素邻域内同样性质像素的光流进行了平滑处理。这种隐含的平滑是一种保

真平滑,将不会弱化边缘信息。

5 试验结果

精细光流估计算法的处理流程如图 1 所示。

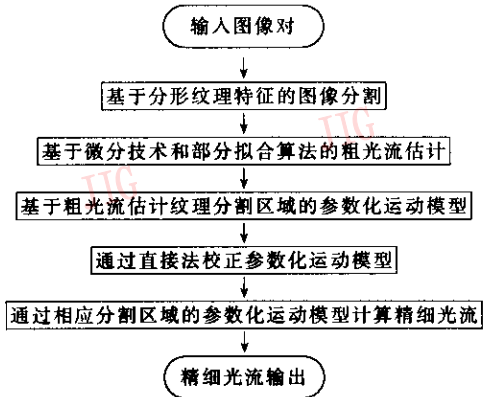


图 1 精细光流估计算法流程图

首先,基于分形纹理特征来分割图像。之后,使用微分法估计粗光流。在粗光流估计中使用基于纹理特征的部分拟合技术将对估计结果有所改善。接下来,基于粗光流,可以估计纹理分割区域的参数化运动模型,然后再用直接法校正上面得到的参数化

运动模型。因为粗光流的不准确,由它得到的参数化运动模型有较大误差,而基于直接法的运动参数校正将会对最后的估计有很大改进。最后,根据每个分割区域对应的参数化运动模型计算出精细的光流。因为在光流估计中同时使用了运动一致性约束和纹理一致性约束,本算法估测出的光流精确度大大好于粗光流估计结果,特别是对于有丰富纹理的图像系列。

为了说明算法的性能,试验中使用了多个标准测试图像序列。这些图像序列包含有摄像机运动的情况、有独立运动的物体情况、含有室外自然景物的情况以及合成图像的情况。图 2 为 flower garden 序列的试验结果,其中包含有摄像机运动的情况以及不同深度景物的情况。由图可见,在树干和草地交界处,光流估计的准确性有了较大的提高。图 3 为 Yosemite 序列这一合成图像的试验结果,由图可见,在图像上方云的区域以及图像左下部分区域,光流估计结果的准确性有了较大提高。图 4 为包含有平移和发散运动的 tree 图像序列,由图可见,本算法得出的光流估计结果的精确度较粗光流估计结果的精确度有较大提高,特别是在图像边缘部分。

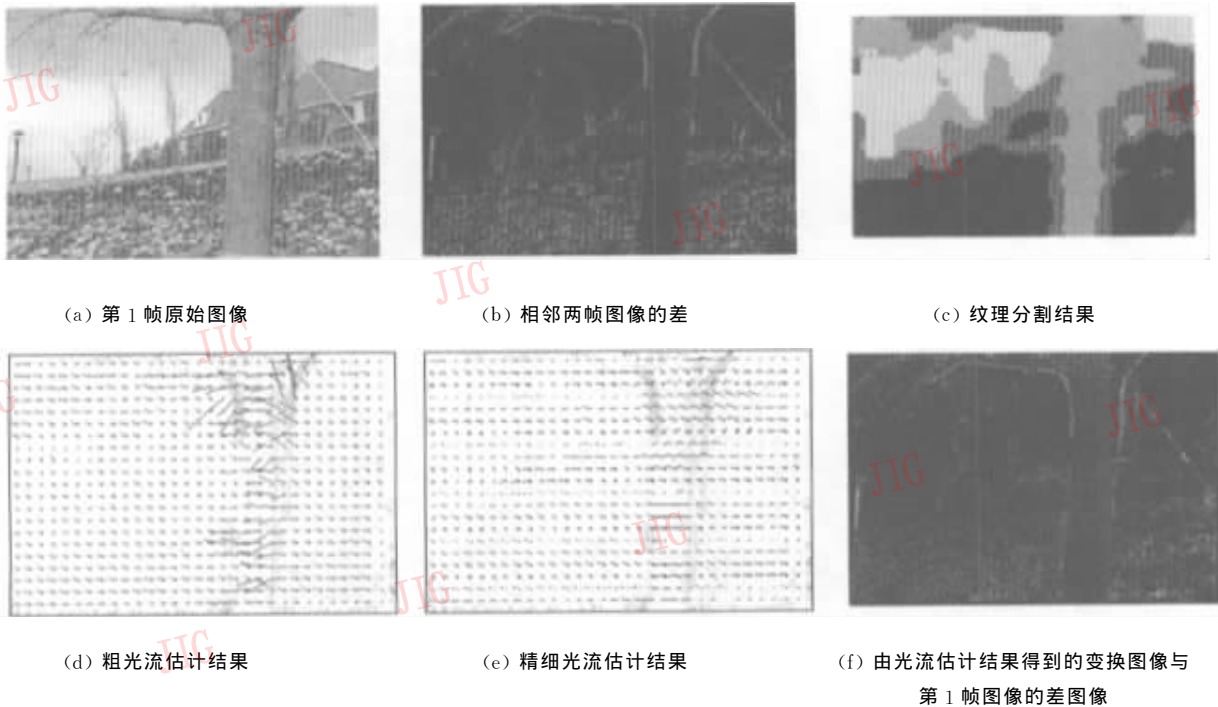


图 2 flower garden 图像序列试验结果
(flower graden 图像序列是 MPEG 压缩标准测试图像序列)

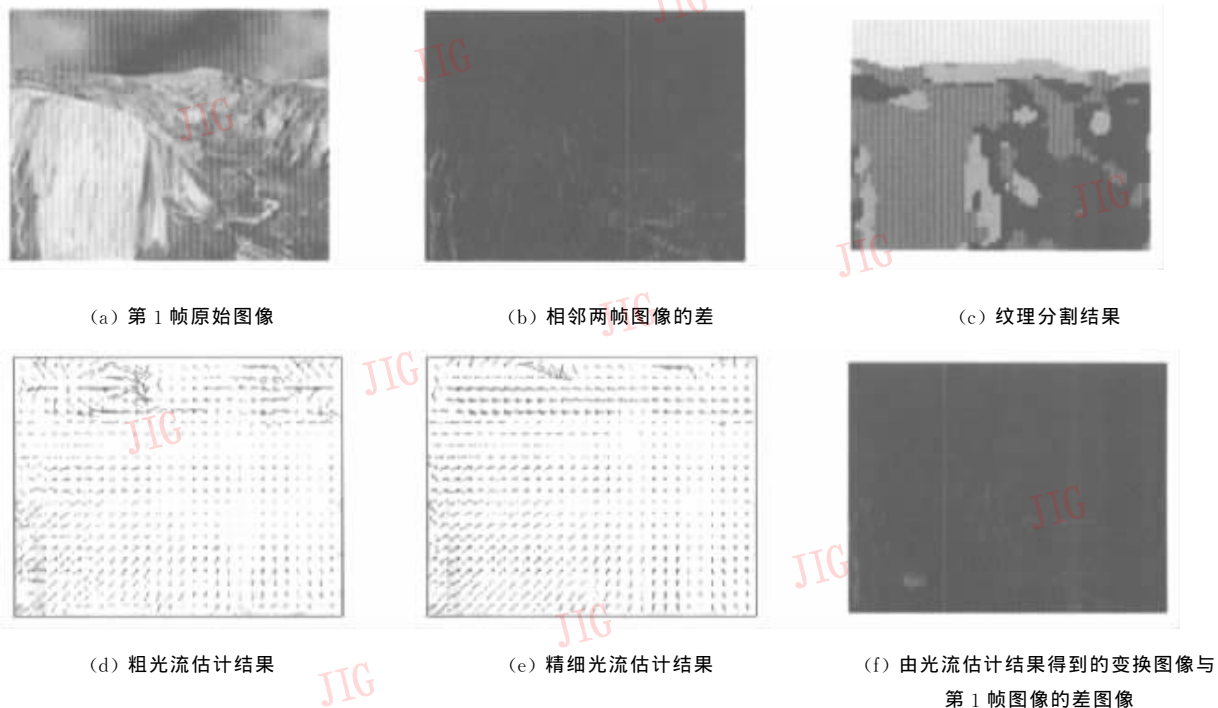


图 3 Yosemite 图像序列实验结果

(yosemite 图像序列由 SRI 的 Lynn Quam 制作, 其中上方云的运动是向右 2 个像素, 而图像的其他部分的运动为发散运动, 左下角的速度约为 5 个像素)

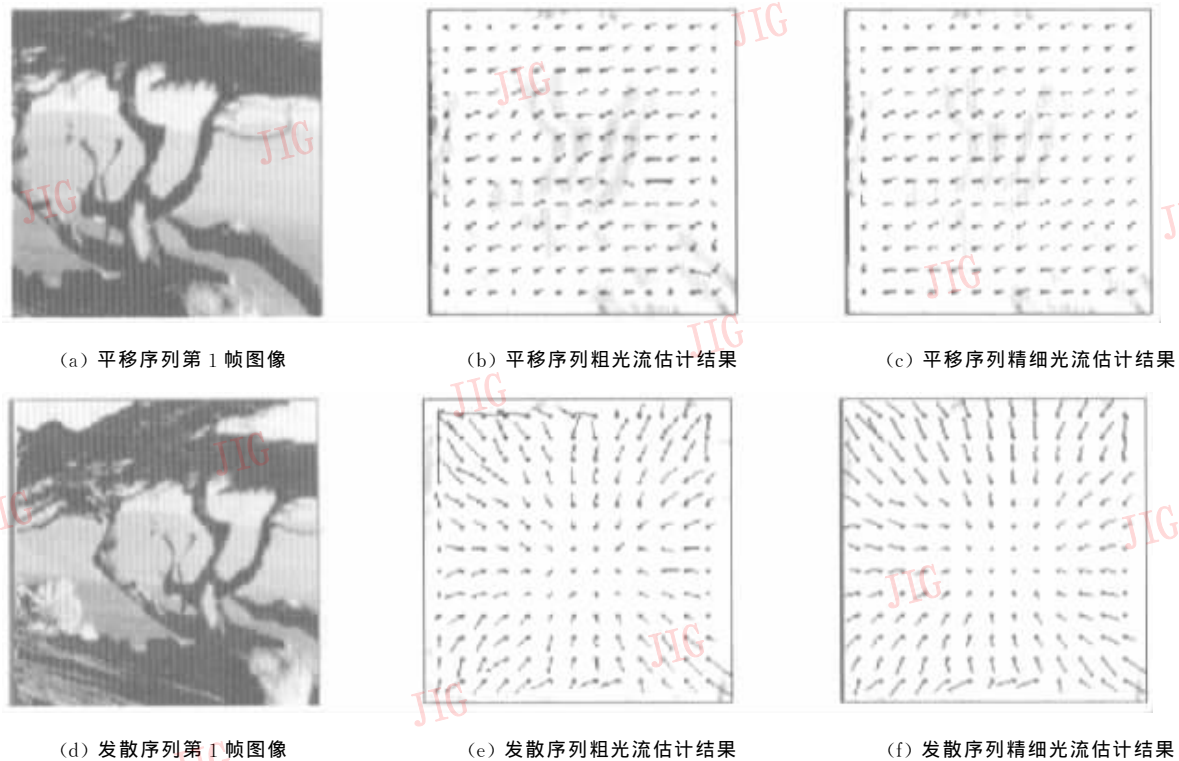


图 4 平移和发散 tree 图像序列实验结果

(平移和发散 tree 图像序列由 David Fleet 制作, 目的是为了模拟摄像机相对于一个有丰富纹理的平面的运动。在平移图像序列中, 摄像机垂直于视线, 沿摄像机 x 轴运动, 速度从左到右为 1.73~2.26 像素/帧。在发散序列中, 摄像机沿着视线运动, 发散焦点在图像中心, 图像运动速度左边为 1.29 像素/帧, 右边为 1.86 像素/帧)

6 结 论

本文提出了一种基于局部小平面运动参数模型的光流估计新方法。主要是利用图像序列中单幅图像的纹理信息,在纹理分割区域的基础上,进行参数化运动模型的估计,从而精确地得到致密的光流估计结果。

本算法特别适用于自然景物环境等有丰富纹理背景的图像序列。它的主要优点在于:(1)使用纹理一致性约束来分割图像,并将分割结果作为初始局部区域来进行运动估计。由于局部区域内有足够的变化信息,优于以往的局部区域的大小形状均预先定义好的算法。同时就分割区域是否满足平面假设来说,本算法也比以往的基于亮度分割图像的光流估计算法更为合理。(2)在光流估计和参数化运动模型估计中均使用了基于纹理一致性的部分拟合算法,可以明显地改善运动边界处的估计结果。

通过与其他已有方法的比较发现,此算法还可以在以下方面有所改进:使用高斯金字塔来提高运算速度;在估计得到光流之后反过来计算参数化模型,从而求得运动分割,来迭代修正纹理分割的结果;采用变阶参数的运动模型,从而避免过拟合和欠拟合的问题。在后续的视频图像序列分层表示的研究中,将在算法中引入这些改进措施。

本算法计算出来的精确而致密的光流对进一步的研究工作非常有用,例如基于分层模型的视频表示,自运动估计,以及景物内物体的三维结构分析等等。

参 考 文 献

1 Barron J L, Fleet D J, Beauchemin S S. Systems and experiment performance of optical flow techniques[J]. International Journal of Computer Vision, 1994, 12(1):43~77.

2 Ju S X, Black M J, Jepson A D. Skin and bones: multi-layer, locally affine, optical flow and regularization with transparency [A]. In: Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], San Francisco, CA, USA, 1996:307~314.

3 Black M J, Jepson A. Estimating optical flow in segmented images using variable-order parametric models with local deformations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(10):972~986.

4 Wang J Y A, Adelson E H. Representation moving images with layers [J]. IEEE Transactions on Image Processing, 1994, 3(5): 625~638.

5 Irani M, Anandan P. Video indexing based on mosaic representation [J]. Proceedings of the IEEE, 1998, 86(5): 905~921.

6 Keller J M, Chen S. Texture description and segmentation through fractal geometry [J]. Computer Vision, Graphics, and Image Processing, 1989, 45(1): 150~166.

7 Yang B, Xu G, Zhu Z. Natural environment image segmentation based on the fractal characters [A]. In: International Symposium on Multi-spectral Image Processing [C], 武汉, 1998, 3545: 450~453.

8 Sawhney H S, Ayer S. Compact representations of videos through dominant and multiple motion estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8):814~830.



杨 波 1973 年生,1995 于清华大学计算机系获学士学位,现为清华大学计算机系博士研究生。主要研究方向为图像和视频分析,基于内容检索。



徐光祐 1940 年生,教授,博士生导师。1963 年毕业于清华大学自动控制系,现为 IEEE 高级会员。研究兴趣为计算机视觉,分布式多媒体和人机交互技术。已发表专(译)著 5 部,在国内外重要刊物上发表论文近 80 篇。