

第十五章 三维运动估计

三维运动估计是指从二维图象序列来估计物体三维运动参数以及三维结构。具体地说，假定三维物体上一点 M 相对于摄像机坐标系从时刻 t_k 的位置 (x_k, y_k, z_k) 运动到时刻 t_{k+1} 的位置 $(x_{k+1}, y_{k+1}, z_{k+1})$ ，它在二维图象平面上的投影从 (x'_k, y'_k) 运动到 (x'_{k+1}, y'_{k+1}) ，然后，通过分析二维运动来恢复物体的三维运动及物体上感兴趣点的深度值。这一点类似于立体视觉的深度恢复，不过立体视觉是从立体图象对来恢复深度值，而三维运动分析是从图象序列中恢复参数。三维运动估计有着广泛的应用，如机器人视觉，自主导航，目标跟踪，图象监控，智能车辆高速公路系统，基于物体的图象压缩等。三维运动估计仍然是一个不适定问题，必须增加适当的约束才能得到有效解。

三维运动估计和分析也可以是基于场景的深度图像序列，其方法与基于二维图像序列完全不同。基于深度图像序列的三维运动估计是一个适定问题，求解方法要比基于二维图像序列要简单一些，主要问题是数据量十分巨大，因此许多研究人员一开始就以实现大规模集成电路（VLSI）作为三维场景估计的基础。我们将不讨论深度图像序列运动估计问题，感兴趣的读者可以查阅有关文献[Wheeler 1996, Jiar 1996]。

15.1 基于成象模型的对应点估计

首先回顾一下第十二章引入的三维刚体运动方程。假定三维场景中有一个刚性物体，其上面一点 M 从时刻 t_k 的位置 (x_k, y_k, z_k) 经过旋转和平移，运动到时刻 t_{k+1} 的位置 $(x_{k+1}, y_{k+1}, z_{k+1})$ 。设旋转矩阵和平移向量分别是 \mathbf{R}_k 和 \mathbf{T}_k ，则三维刚体运动模型重新表示为

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{bmatrix} r_{xx} & r_{xy} & r_{xz} \\ r_{yx} & r_{yy} & r_{yz} \\ r_{zx} & r_{zy} & r_{zz} \end{bmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} = \mathbf{R}_k \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \mathbf{T}_k \quad (15.1)$$

用欧拉角的形式表示上述旋转矩阵（见式(12.2)，(12.3)），并假定旋转角较小，则旋转矩阵可以表示为

$$\mathbf{R}_k = \begin{bmatrix} 1 & -\phi & \psi \\ \phi & 1 & -\theta \\ -\psi & \theta & 1 \end{bmatrix} \quad (15.2)$$

其中 θ, ψ, ϕ 分别表示绕 x, y, z 轴逆时针旋转小角位移。

15.1.1 正交投影模型

当物体深度变化范围不大时，正交投影是透视投影的一个很好的逼近。其它逼近方法还有弱透视投影，超透视投影，正交透视投影等[Dementhon 1992]。

设空间点 (x, y, z) 在图象平面上的投影为 (x', y') 。如果成象模型为正交投影（参见图1.8），则有

$$x' = x \quad y' = y \quad (15.3)$$

所以，式(15.1)可表示为

$$\begin{aligned} x_{k+1} &= x'_{k+1} = r_{xx}x'_k + r_{xy}y'_k + (r_{xz}z_k + t_x) \\ y_{k+1} &= y'_{k+1} = r_{yx}x'_k + r_{yy}y'_k + (r_{yz}z_k + t_y) \end{aligned} \quad (15.4)$$

上述方程包含有6个参数，即 $r_{xx}, r_{xy}, r_{yx}, r_{yy}, (r_{xz}z_k + t_x)$ 和 $(r_{yz}z_k + t_y)$ ，表示第 k 帧

图象像素 (x'_k, y'_k) 到第 $k+1$ 帧图象像素 (x'_{k+1}, y'_{k+1}) 的仿射映射关系。显然，正交投影模型是无法确定物体点到成像平面的距离，因为垂直于图像平面的一条直线，其上的所有点都将投影到该图像平面上一点（见图 1.8）。但是，如果在物体上选择一个参考点，并设该点的深度值为 z_{ref} ，则通过上式有可能估计出物体上其它点相对于这个参考点在垂直图像平面方向上的距离 z_{rel} 。实际上，我们无法得到真实的相对深度值，只能得到关于一个尺度系数 α 的相对深度值，即

$$z = z_{ref} + \alpha z_{rel} \quad (15.5)$$

从方程 15.4 中，我们看到 r_{xz} 和 r_{yz} 缩小 α ， z_k 放大 α ，方程仍然成立，因此产生多义性。文献[U1179]证明三帧或三帧以上图象上的四点就可以克服这个问题。

15. 1. 2 基于正交投影的运动估计

将(15.2)小角度旋转矩阵代入(15.1)，得到如下的正交投影模型：

$$\begin{aligned} x'_{k+1} &= x'_k - \phi y'_k + \psi z_k + t_x \\ y'_{k+1} &= \phi x'_k + y'_k - \theta z_k + t_y \end{aligned} \quad (15.6)$$

在该方程中，对于每一个给定的对应点 (x'_k, y'_k) 和 (x'_{k+1}, y'_{k+1}) ，共有 6 个未知参数，其中 5 个是全局运动参数 ϕ ， ψ ， θ ， t_x ， t_y ，一个是深度参数 z_k 。另外，这个方程是一个双线性方程，因为 z_k 与未知参数 ϕ 和 θ 是乘积关系。理论上，给定三点，就可以根据(15.6)列出 6 个方程，从而解出六个未知参数。但实际上，由于数值计算误差，常常需要多个点，这样有可能得到较好的结果。文献[Aizawa 1989] 基于上述正交投影模型提出了基于两帧图象的两步迭代法：首先，根据上一次迭代得到的深度估计值，确定运动参数，然后再使用新的运动参数更新深度估计值。具体实现见算法 15.1

算法 15.1 基于两帧图像的运动估计两步迭代算法

① 给定 n 个对应点坐标对 $\{(x'_{i,k}, y'_{i,k})\}$ $\{(x'_{i,k+1}, y'_{i,k+1})\}$ 和深度估计值 $\{z_{i,k}\}$ ， $i=1,2,\dots,n$ ，且 $n \geq 3$ ，这样方程(15.6)可重新写为

$$\begin{pmatrix} x'_{k+1} - x'_k \\ y'_{k+1} - y'_k \end{pmatrix} = \begin{pmatrix} 0 & z_k & -y'_k & 1 & 0 \\ -z_k & 0 & x'_k & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta \\ \psi \\ \phi \\ t_x \\ t_y \end{pmatrix} \quad (15.7)$$

n 个对应点对应着 $2n$ 个方程，而未知参数仅有 5 个，因此，可以通过最小二乘法来求解这 5 个运动参数。深度参数的初始估计值可以根据场景的先验模型来设置，深度估计值应在预先设定的范围内选定，这主要是为了避免解的不唯一性。

② 根据①得到的运动参数估计值，再对深度值 $\{z_{i,k}\}$ 进行估计。将式(15.7)重新写为

$$\begin{pmatrix} x'_{k+1} - x'_k + \phi y'_k - t_x \\ y'_{k+1} - y'_k - \phi x'_k - t_y \end{pmatrix} = \begin{pmatrix} \psi \\ -\theta \end{pmatrix} (z_k) \quad (15.8)$$

由于每一个深度值对应两个方程，即方程(15.8)是一个超定方程，因此，可以用最小二乘法来求解。

③ 重复上述两步，直到两次迭代值之差小于给定的某一个阈值。

请注意，在上述算法中，运动估计误差和深度估计误差有着密切的关系。由方程(15.7)和(15.8)可知，深度估计的随机误差会重复反馈到运动估计上。因此，当深度估计不够准确或深度的初始值设置不当时，都可能导致迭代算法的错误收敛或收敛在一个局部最小值。为了避免这种错误的收敛，[Bozdagi 1994]提出了改进的算法，该算法的基本思想是在每一次修正后，在深度估计值上加一个随机扰动。改进的算法如算法 15.1 所示。

算法 15.2 基于两帧图像运动估计扰动迭代算法:

- ① 初始化深度值 $\{z_{i,k}\}$, $i = 1, 2, \dots, n$, 置迭代计数器 $m = 0$ 。
- ② 在给定深度值下根据式 (15.6) 估计运动参数
- ③ 根据当前的运动估计和深度参数, 由公式 (15.6) 计算对应点的坐标 $(x'_{i,k+1}, y'_{i,k+1})$
- ④ 计算预估误差:

$$E_m = \frac{1}{n} \sum_{i=1}^n e_i \quad (15.9)$$

其中 $e_i = (x'_{i,k+1} - x'_{i,k+1})^2 + (y'_{i,k+1} - y'_{i,k+1})^2$, $x'_{i,k+1}$ 和 $y'_{i,k+1}$ 是已知的对应点真实坐标。

- ⑤ 如果 E_m 小于预定的误差阈值 E_T , 即 $E_m < E_T$, 则终止迭代, 否则, 置 $m = m + 1$ 。
- ⑥ 给深度参数赋一个扰动值

$$z_{i,k}^{(m)} \leftarrow z_{i,k}^{(m-1)} - \beta \frac{\partial e_i}{\partial z} + \alpha \Delta_i^{(m)} \quad (15.10)$$

其中 α 和 β 是常系数, $\Delta_i^{(m)} = N_i(0, \sigma_i^{2(m)})$ 是零均值高斯分布函数, 其方差 $\sigma_i^{2(m)} = e_i$ 。

- ⑦ 回到第②步

实验证明, 这种改进的迭代算法在初始深度值有 50% 的误差的情况下, 也能很好地收敛到正确的运动参数值。

15.1.3 透视投影模型

设空间点 (x, y, z) 在图象平面上的投影为 (x', y') 。如果成象模型为透视投影, 则

$$x' = F \frac{x}{z} \quad y' = F \frac{y}{z} \quad (15.11)$$

根据 (15.1) 式有

$$\begin{aligned} x'_{k+1} &= F \frac{x_{k+1}}{z_{k+1}} = F \frac{r_{xx}x_k + r_{xy}y_k + r_{xz}z_k + t_x}{r_{zx}x_k + r_{zy}y_k + r_{zz}z_k + t_z} \\ y'_{k+1} &= F \frac{y_{k+1}}{z_{k+1}} = F \frac{r_{yx}x_k + r_{yy}y_k + r_{yz}z_k + t_y}{r_{zx}x_k + r_{zy}y_k + r_{zz}z_k + t_z} \end{aligned} \quad (15.12)$$

由于成象系统的焦距是一个常数, 因此, 不乏一般性, 取 $F = 1$, 即规范化透视投影。上式右边分子分母同除以 z_k , 得到图象平面坐标表示式:

$$\begin{aligned} x'_{k+1} &= \frac{r_{xx}x'_k + r_{xy}y'_k + r_{xz} + t_x/z_k}{r_{zx}x'_k + r_{zy}y'_k + r_{zz} + t_z/z_k} \\ y'_{k+1} &= \frac{r_{yx}x'_k + r_{yy}y'_k + r_{yz} + t_y/z_k}{r_{zx}x'_k + r_{zy}y'_k + r_{zz} + t_z/z_k} \end{aligned} \quad (15.13)$$

按照图像平面坐标, 透视投影模型 (15.13) 是一个非线性方程。因为每一点对应的深度值 z 是一个自由参数, 因此, 这个模型适合于任意表面形状三维物体的运动估计。

15.1.4 外极线方程和基本矩阵

由方程 (15.13) 可见, 在透视投影情况下, 运动和结构参数之间的关系是非线性的。早期人们使用迭代方法求解运动参数, 但迭代过程往往不收敛, 比如两步迭代法。[Huang 1986] 证明, 使用八个对应点或更多对应点来求取外极线约束方程并估计运动参数是可以改进两步迭代法的收敛性能。本节主要讨论外极线方程及其性质, 关于估计运动参数, 将在下一节介

绍。

将方程 15.1 缩写为

$$\mathbf{M}_{k+1} = \mathbf{R}_k \mathbf{M}_k + \mathbf{T}_k \quad (15.14)$$

或

$$\mathbf{M}_{k+1}^T \cdot (\mathbf{T}_k \times \mathbf{R}_k \mathbf{M}_k) = 0 \quad (15.15)$$

令

$$\mathbf{E}_k = \mathbf{T}_k \times \mathbf{R}_k \quad (15.16)$$

其中“ \times ”表示矢量积。为了书写简单，在不引起混淆的情况下，可以把下标 k 去掉，即：

$$\mathbf{E} = \mathbf{T} \times \mathbf{R} \quad (15.17)$$

方程(15.15)可以重新写为

$$\mathbf{M}_{k+1}^T \mathbf{E} \mathbf{M}_k = 0 \quad (15.18)$$

下面引进一个反对称矩阵：

$$[\mathbf{T}]_x = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$$

因此式(15.17)可表示为

$$\mathbf{E} = [\mathbf{T}]_x \mathbf{R} \quad (15.19)$$

\mathbf{E} 是一个 3×3 矩阵，称为基本矩阵(essential matrix)，矩阵元素称为基本参数[Huang 1986]。下面首先看一下矩阵 \mathbf{E} 的一些性质[Faugeras 1993]：

- 1) $\det \mathbf{E} = 0$
- 2) $\mathbf{E}^T \mathbf{T} = 0$
- 3) $\mathbf{E} \mathbf{E}^T = (\mathbf{T}^T \mathbf{T}) \mathbf{I} - \mathbf{T} \mathbf{T}^T$
- 4) $\|\mathbf{E}\|^2 = 2\|\mathbf{T}\|_x^2$ ，其中 $\|\mathbf{E}\|^2 = \sum e_{ij}^2$ ，称为 Frobenius 模。

在性质 1) 中， \mathbf{E} 是反对称矩阵，因此该性质是成立的；

对于性质 2)， $\mathbf{E}^T \mathbf{T} = ([\mathbf{T}]_x \mathbf{R})^T \mathbf{T} = \mathbf{R}^T [\mathbf{T}]_x^T \mathbf{T} = -\mathbf{R}^T \mathbf{T} \times \mathbf{T} = 0$ ；

对于性质 3)， $\mathbf{E} \mathbf{E}^T = ([\mathbf{T}]_x \mathbf{R})([\mathbf{T}]_x \mathbf{R})^T = ([\mathbf{T}]_x \mathbf{R} \mathbf{R}^T)[\mathbf{T}]_x^T = (\mathbf{T}^T \mathbf{T}) \mathbf{I} - \mathbf{T} \mathbf{T}^T$ ；

性质 4) 显然是成立的。

用 $z_k z_{k+1}$ 除以等式(15.18)的两边得

$$\begin{pmatrix} \frac{x_{k+1}}{z_{k+1}} & \frac{y_{k+1}}{z_{k+1}} & 1 \end{pmatrix} \mathbf{E} \begin{pmatrix} \frac{x_k}{z_k} \\ \frac{y_k}{z_k} \\ 1 \end{pmatrix} = 0 \quad (15.20)$$

根据规范化透视投影方程(15.11)，得到外极线方程：

$$\begin{pmatrix} x'_{k+1} & y'_{k+1} & 1 \end{pmatrix} \mathbf{E} \begin{pmatrix} x'_k \\ y'_k \\ 1 \end{pmatrix} = 0 \quad (15.21)$$

$\mathbf{m}_k = (x'_k, y'_k)^T$ 和 $\mathbf{m}_{k+1} = (x'_{k+1}, y'_{k+1})^T$ 是物体上一点 \mathbf{M} 在第 k 帧和第 $k+1$ 帧图像上的投影点坐标，其齐次坐标分别为 $\tilde{\mathbf{m}}_k = (x'_k, y'_k, 1)^T$ 和 $\tilde{\mathbf{m}}_{k+1} = (x'_{k+1}, y'_{k+1}, 1)^T$ ，则外极线方程(15.21)可以表示为

$$\tilde{\mathbf{m}}_{k+1}^T \mathbf{E} \tilde{\mathbf{m}}_k = 0 \quad (15.22)$$

该方程的几何意义十分明显，从图 15.1 可见，三条线 $\mathbf{c}_{k+1}\mathbf{c}_k$ ， $\mathbf{c}_{k+1}\mathbf{m}_{k+1}$ 和 $\mathbf{c}_k\mathbf{m}_k$ 是共面的，将这三个矢量变换到第二个坐标系中，其表示式分别为 \mathbf{T}_k ， $\tilde{\mathbf{m}}_{k+1}$ 和 $\mathbf{R}_k\tilde{\mathbf{m}}_k$ 。

第 k 帧图像上一点 \mathbf{m}_k 对应矢量 $\mathbf{c}_k\mathbf{m}_k$ 表示的射线上所有点的投影，射线上任何一点可以表示为 $\lambda\tilde{\mathbf{m}}_k$ ， $\lambda \in [0, \infty)$ 。根据针孔模型，射线上任何一点在第 $k+1$ 帧图像平面上的投影是 $\lambda\mathbf{R}\tilde{\mathbf{m}}_k + \mathbf{T}$ ，而整个射线在第 $k+1$ 帧图像平面上的投影是一条直线，表示成为第 k 帧图像上一点 \mathbf{m}_k 在第 $k+1$ 帧图像平面上的外极线。这条直线可以用两点来表示，一点是 $\mathbf{e}_{k+1} = \mathbf{T}$ ， $\lambda = 0$ ，对应第 k 帧图像光学中心 \mathbf{c}_k 在第 $k+1$ 帧图像平面上的投影，常称为极点 (epipole)；另一点是无穷远点， $\mathbf{m}_{k+1,\infty} = \mathbf{R}\tilde{\mathbf{m}}_k$ 。则外极线可以表示为

$$\mathbf{e}_{k+1} \times \mathbf{m}_{k+1,\infty} = \mathbf{T} \times \mathbf{R}\tilde{\mathbf{m}}_k \quad (15.23)$$

第 k 帧图像上一点 \mathbf{m}_k 在第 $k+1$ 帧图像平面上的对应点 \mathbf{m}_{k+1} 应该位于该直线上。

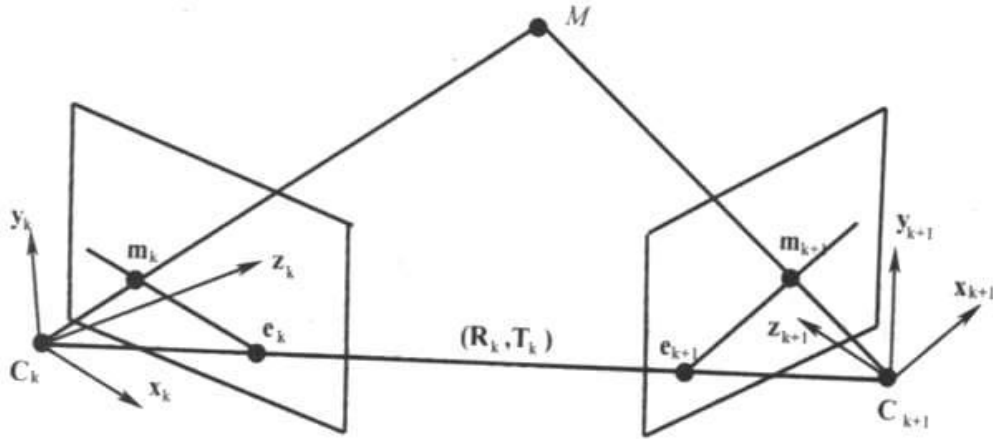


图 15.1 基于对应点的外极线约束恢复运动参数示意图

方程 (15.22) 是包含 9 个未知参数的齐次线性方程，齐次线性方程无解或有无穷解。因此，可以令基本矩阵 \mathbf{E} 的一个系数为 1，这样待估计的参数有八个。另外，矩阵 \mathbf{E} 是一个斜对称矩阵和一个旋转矩阵的乘积，它们并不是独立的。矩阵 \mathbf{E} 的前三个性质构成三个约束方程，这样，方程 (15.22) 只包含 5 个未知的独立的参数，这也和运动参数的自由度数量相一致，即三个旋转自由度，二个平移自由度（或三个关于一个比例系数的平移自由度），实际上，由方程 (15.15) 可见，平移量 \mathbf{T}_k 乘以任何不为零的系数都不影响方程成立，也就是说，当用同一个比例系数改变物体形状或运动平移量时，所得到的图像完全一样。因此，从运动恢复形状和从图像序列恢复运动参数，只能在关于一个比例系数的意义下进行。

*15. 1. 5 从基本矩阵估计运动参数

由于基本矩阵只有 5 个未知独立参数，因此可以采用所谓的 5-点算法 (5-points algorithm) 来求解基本矩阵，然后，基于基本矩阵恢复运动参数。不过由于 5-点算法比较复杂，且得到的解很不稳定，因此，实用价值不大，在这里不作讨论，感兴趣的读者可以参见文献[Faugerous 1990]。为了得到唯一解，一般选取 8 个以上的对应点通过最小二乘法来求解。[Longuet 1981]提出一种 8-点算法，通过 8 个点直接估计基本矩阵的 8 个未知参数，然后，在基本矩阵的基础上，估计运动参数。将基本矩阵 \mathbf{E} 表示为：

$$\mathbf{E} = \begin{pmatrix} e_{00} & e_{01} & e_{02} \\ e_{10} & e_{11} & e_{12} \\ e_{20} & e_{21} & e_{22} \end{pmatrix}$$

方程 (15.22) 可以重新写为

$$\begin{pmatrix} x'_{k+1}x'_k & x'_{k+1}y'_k & x'_{k+1} & y'_{k+1}x'_k & y'_{k+1}y'_k & y'_{k+1} & x'_k & y'_k & 1 \end{pmatrix} \begin{pmatrix} e_{00} \\ e_{01} \\ e_{02} \\ e_{10} \\ e_{11} \\ e_{12} \\ e_{20} \\ e_{21} \\ 1 \end{pmatrix} = 0 \quad (15.24)$$

设 $\mathbf{X} = (\mathbf{e}_0^T \quad \mathbf{e}_1^T \quad \mathbf{e}_2^T)^T$ ，将上式缩写为

$$\mathbf{a}_k^T \mathbf{X} = 0 \quad (15.25)$$

如果有 n 个对应点，每一个对应点生成一个方程，则共有 n 个方程

$$\mathbf{A}_{n,k}^T \mathbf{X} = 0 \quad (15.26)$$

其中 $\mathbf{A}_{n,k}^T = (\mathbf{a}_{1,k}^T \quad \mathbf{a}_{2,k}^T \quad \dots \quad \mathbf{a}_{n,k}^T)$ 。由于 \mathbf{E} 是关于一个比例系数的矩阵，因此，实际上只需要求解 8 个未知参数。所以，求解该方程的最小点数是 8 个。已知空间中有 8 个点及其在第 k 帧和第 $k+1$ 帧图像上的投影点坐标，则通过求解方程 (15.26) 得到基本矩阵参数。显然，只用 8 点求解基本矩阵，对噪声十分敏感，因此，一般采用更多的对应点并采用更鲁棒的方法来求解基本矩阵参数 [Faugerus 1993, pp273]。下面介绍一种极小化 $\mathbf{A}_{n,k}^T \mathbf{X}$ 范数的基本矩阵求解方法。为了避免平凡解 $\mathbf{X} = 0$ ，假定 $\mathbf{A}_{n,k}^T \mathbf{X}$ 的范数不为零。在实际中，由于 $\mathbf{E} = [\mathbf{T}]_{\times} \mathbf{R}$ ，很容易证明， $\|\mathbf{X}\|^2 = 2\|\mathbf{T}\|^2$ 。前面曾提到平移量 \mathbf{T} 是关于一个比例因子的量，因此，可以取 $\|\mathbf{T}\| = 1$ 。这样，求解基本矩阵参数问题就成为

$$\min_{\mathbf{X}} \|\mathbf{A}_{n,k}^T \mathbf{X}\|^2 \quad (15.27)$$

$$\text{约束条件: } \|\mathbf{X}\|^2 = 2$$

根据约束条件下的优化算法，引入一个拉格朗日乘子，得到目标函数：

$$F(\mathbf{X}) = \mathbf{X}^T \mathbf{A}_{n,k}^T \mathbf{A}_{n,k} \mathbf{X} + \lambda \mathbf{X}^T \mathbf{X} \quad (15.28)$$

该函数对 \mathbf{X} 微分，可得

$$\mathbf{A}_{n,k}^T \mathbf{A}_{n,k} \mathbf{X} = \lambda \mathbf{X} \quad (15.29)$$

上式说明 λ 是对称矩阵 $\mathbf{A}_{n,k}^T \mathbf{A}_{n,k}$ 的特征值，而 \mathbf{X} 是对应的最小特征向量。因为 $\mathbf{A}_{n,k}^T \mathbf{A}_{n,k}$ 是一个 9×9 的矩阵，有 9 个可能解，令 9 个特征值为 $\lambda_i, i=1,2,\dots,9$ ，并假定 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_9$ ，则 $F(\mathbf{X}) = 2\lambda$ ，由于 (15.28) 是求解 $F(\mathbf{X})$ 的极小值，故， $\lambda = \lambda_1$ 。此时的解是对称矩阵 $\mathbf{A}_{n,k}^T \mathbf{A}_{n,k}$ 的特征值且范数为 $\sqrt{2}$ 的特征向量。

求出基本矩阵后，可以进行运动参数 \mathbf{T} 和 \mathbf{R} 的估计。由基本矩阵的性质 2) 可知， \mathbf{T} 可以通过求下面均方问题极小化来求解：

$$\min_{\mathbf{T}} \|\mathbf{E}^T \mathbf{T}\|^2 \quad (15.30)$$

$$\text{约束条件: } \|\mathbf{T}\|^2 = 1$$

则 \mathbf{T} 是对应矩阵 $\mathbf{E}\mathbf{E}^T$ 的最小特征值的单位范数向量。

对于旋转矩阵 \mathbf{R} ，必须通过求解下面均方问题极小化得到

$$\min_{\mathbf{R}_k} \|\mathbf{E} - [\mathbf{T}]_x \mathbf{R}\|^2 \quad (15.31)$$

$$\text{约束条件: } \mathbf{R}^T \mathbf{R} = \mathbf{I} \quad \det(\mathbf{R}) = 1$$

由于 $\mathbf{E} - [\mathbf{T}]_x \mathbf{R} = (\mathbf{E} \mathbf{R}^T - [\mathbf{T}]_x) \mathbf{R}$, $\|\mathbf{E} - [\mathbf{T}]_x \mathbf{R}\|^2 = \|\mathbf{E} \mathbf{R}^T - [\mathbf{T}]_x\|^2$ 。

另外, $\mathbf{E} \mathbf{R}^T = [\mathbf{R} \mathbf{e}_0 \quad \mathbf{R} \mathbf{e}_1 \quad \mathbf{R} \mathbf{e}_2]$, 则上述问题变为:

$$\min_{\mathbf{R}} \sum_{i=0}^2 \|\mathbf{R} \mathbf{e}_i - \mathbf{t}_i\|^2 \quad (15.32)$$

$$\text{约束条件: } \mathbf{R}^T \mathbf{R} = \mathbf{I} \quad \det(\mathbf{R}) = 1$$

其中 \mathbf{t}_i 是矩阵 $[\mathbf{T}]_x$ 的第 i 行向量. 这一问题可以通过四元数表示法求解, 一般情况下, 解是唯一的.

15. 1. 6 从外极线方程估计运动参数

从基本矩阵估计运动参数是间接求解方法, 即首先求解基本矩阵 \mathbf{E} , 然后利用线性方法恢复运动参数 \mathbf{T} 和 \mathbf{R} . 由于间接求解方法采用的线性方法, 因此, 比较简单. 但为了求解有效的基本矩阵, 必须使用高阶多项式约束函数, 如果这样, 则丧失线性方法的简单性. 因此, 人们开始研究直接求解方法, 即通过外极线方程直接估计运动参数. 下面介绍一种直接的方法: Longuet-Higgins 准则[Faugeras 1993, pp275],

$$LH(\mathbf{R}, \mathbf{T}) = \sum_{i=0}^{n-1} (\tilde{\mathbf{m}}_{i,k+1}^T \cdot (\mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_{i,k}))^2 \quad (15.33)$$

方程(15.33)没有解析解, 必须通过非线性极小化方法求解.

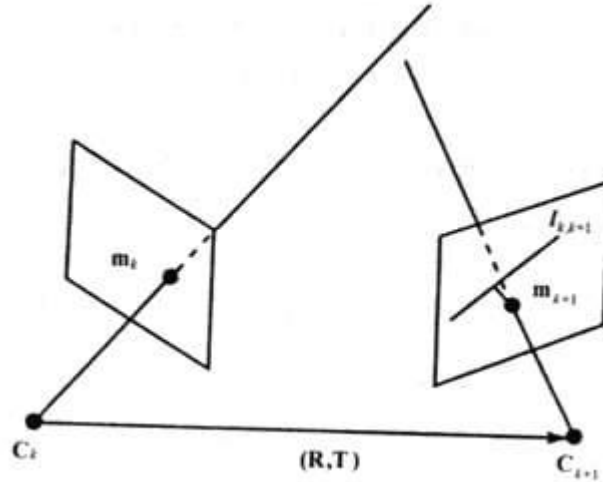


图 15.2 外极线与实际投影点的距离关系

Longuet-Higgins 准则的几何意义可以从图 15.2 解释. 已知空间一点 \mathbf{M} 在第 k 帧和第 $k+1$ 帧图像上对应的投影点坐标分别为 $\mathbf{m}_k = (x'_k, y'_k)^T$ 和 $\mathbf{m}_{k+1} = (x'_{k+1}, y'_{k+1})^T$. \mathbf{m}_k 在第 $k+1$ 帧图像平面上的外极线 $\mathbf{l}_{k,k+1}$ 表示为 $\mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_k$, 见式 15.23. 在理想情况下, 空间点 \mathbf{M} 在第 $k+1$ 帧图像上的投影点 \mathbf{m}_{k+1} 应该位于直线 $\mathbf{l}_{k,k+1}$ 上, 即 $\tilde{\mathbf{m}}_{k+1} \cdot (\mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_k) = 0$. 实际上, 由于投影点检测误差、数字计算误差以及系统误差等, \mathbf{m}_{k+1} 并不一定在外极线 $\mathbf{l}_{k,k+1}$ 上. \mathbf{m}_{k+1} 到 $\mathbf{l}_{k,k+1}$ 的距离可以定义为

$$d_{i,k+1} = \frac{|\tilde{\mathbf{m}}_{k+1} \cdot (\mathbf{T} \times \mathbf{R} \tilde{\mathbf{m}}_k)|}{\beta} \quad (15.34)$$

其中 β 是规范化系数，其值等于矢量 $\mathbf{R}^T (\mathbf{T} \times \mathbf{m}_k)$ 前两个坐标形成的矢量的范数。

如果有 n 个对应点，则 Longuet-Higgins 准则可以重新写为

$$\sum_{i=0}^{n-1} \beta_i d_{i,k+1} \quad (15.35)$$

上式是第 k 帧图像上的 \mathbf{m}_k 在第 $k+1$ 帧图像上的对应点与外极线 $\mathbf{l}_{k,k+1}$ 的距离。反过来，也可以求解第 $k+1$ 帧图像上的 \mathbf{m}_{k+1} 在第 k 帧图像上的对应点与外极线 $\mathbf{l}_{k+1,k}$ 的距离是

$$\sum_{i=0}^{n-1} \alpha_i d_{i,k} \quad (15.36)$$

为了可靠地求解运动参数，可对上面两个式子之和求极小化：

$$\sum_{i=0}^{n-1} (\alpha_i d_{i,k} + \beta_i d_{i,k+1}) \quad (15.37)$$

实践证明，上述算法优于解析方法。

15.2 三维运动估计的光流法

现在已经出现了许多用于三维结构估计的算法，比如，前面几章介绍的从立体视觉恢复三维结构，从明暗恢复结构，从运动恢复结构等。15.1 节曾讨论基于特征点的运动估计深度值的方法，由于特征点是稀疏的，因此要恢复出物体表面的完整结构，必须进行插值计算。本节首先讨论由稠密光流场实现 3D 平移运动表面结构估计的方法，然后讨论基于光流场的一般的运动结构参数的估计方法。

在光流场估计的基础上，从两幅正交投影或透视投影的图像中可以估计出三维物体的运动和结构。光流法与模型法的不同之处在于前者使用投影速度场模型，而后者使用的是位移场模型。光流法需要稠密的光流场，而不是选择并匹配明显的特征点。首先我们假定光流场是由一个物体产生的，对于多个物体，将在后面的运动场分割时讨论。

15.2.1 速度场正交投影模型与仿射流

在刚体旋转的角度很小时，可以根据式 (15.1) 和 (15.2) 得到如下刚体变换公式为

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{bmatrix} 1 & -\theta & \psi \\ \theta & 1 & -\phi \\ -\psi & \phi & 1 \end{bmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (15.38)$$

或

$$\begin{pmatrix} x_{k+1} - x_k \\ y_{k+1} - y_k \\ z_{k+1} - z_k \end{pmatrix} = \begin{bmatrix} 0 & -\theta & \psi \\ \theta & 0 & -\phi \\ -\psi & \phi & 0 \end{bmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (15.39)$$

上式两端同除以 $\Delta t = t_{k+1} - t_k$ ，得到速度变换公式：

$$\begin{bmatrix} \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{bmatrix} = \begin{bmatrix} 0 & -\dot{\theta} & \dot{\psi} \\ \dot{\theta} & 0 & -\dot{\phi} \\ -\dot{\psi} & \dot{\phi} & 0 \end{bmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{bmatrix} \dot{t}_x \\ \dot{t}_y \\ \dot{t}_z \end{bmatrix} \quad (15.40)$$

(1) 速度场正交投影模型

三维速度场在图像平面上的正交投影计算公式如下：

$$\begin{aligned} u &= \dot{x}' = \dot{x} \\ v &= \dot{y}' = \dot{y} \end{aligned} \quad (15.41)$$

根据式(15.40)，上式可表示为（为了书写简单，这里忽略了下标 k ）

$$\begin{aligned} u &= \dot{i}_x - \dot{\theta}y + \dot{\psi}z = \dot{i}_x - \dot{\theta}y' + \dot{\psi}z \\ v &= \dot{i}_y + \dot{\theta}x - \dot{\phi}z = \dot{i}_y + \dot{\theta}x' - \dot{\phi}z \end{aligned} \quad (15.42)$$

当图像平面离物体的距离越来越远时，或观测视角越来越小时，正交投影模型可以看作是透视投影模型的近似模型。

(2) 仿射流

进行刚体运动的平面在正交投影下生成仿射流。设平面方程为

$$z = a_0 + a_1x + a_2y \quad (15.43)$$

将(15.43)代入式(15.42)，得到含有六个参数的仿射流模型

$$\begin{aligned} u &= b_0 + b_1x' + b_2y' \\ v &= b_3 + b_4x' + b_5y' \end{aligned} \quad (15.44)$$

其中

$$\begin{aligned} b_0 &= t_x + a_0\dot{\psi} & b_1 &= a_1\dot{\psi} & b_2 &= a_2\dot{\psi} - \dot{\phi} \\ b_3 &= t_y - a_0\dot{\theta} & b_4 &= \dot{\phi} - a_1\dot{\theta} & b_5 &= -a_2\dot{\theta} \end{aligned}$$

注意到，如果已知三个或更多点的光流，就能建立包含六个未知量的六个或更多的方程，从而求解 (b_0, \dots, b_5) 。然而，由于正交投影的缘故，不可能从 (b_0, \dots, b_5) 中唯一地确定所有八个运动和结构参数。例如，在正交投影下 a_0 是不可观察的。

15. 2. 2 速度场透视投影模型与二次流

(1) 速度场透视投影模型

首先对透视投影模型(15.11)微分：

$$\begin{aligned} u &= \dot{x}' = F \frac{z\dot{x} - x\dot{z}}{z^2} = F \frac{\dot{x}}{z} - x' \frac{\dot{z}}{z} \\ v &= \dot{y}' = F \frac{z\dot{y} - y\dot{z}}{z^2} = F \frac{\dot{y}}{z} - y' \frac{\dot{z}}{z} \end{aligned} \quad (15.45)$$

再将(15.40)代入上式得（为了书写简单，这里忽略了下标 k ）：

$$\begin{aligned} u &= F \left(\frac{\dot{i}_x}{z} + \dot{\phi} \right) - \frac{\dot{i}_z}{z} x' - \dot{\theta}y' - \frac{\dot{\phi}}{F} x'y' + \frac{\dot{\psi}}{F} x'^2 \\ v &= F \left(\frac{\dot{i}_y}{z} - \dot{\phi} \right) - \frac{\dot{i}_z}{z} y' + \dot{\theta}x' + \frac{\dot{\phi}}{F} x'y' - \frac{\dot{\psi}}{F} y'^2 \end{aligned} \quad (15.46)$$

在规范化透视投影下，上式重新写为

$$\begin{aligned} u &= \frac{\dot{i}_x - x'\dot{i}_z}{z} + \dot{\psi}(1 + x'^2) - \dot{\theta}y' - \dot{\phi}x'y' \\ v &= \frac{\dot{i}_y - y'\dot{i}_z}{z} - \dot{\psi}(1 + y'^2) + \dot{\theta}x' + \dot{\psi}x'y' \end{aligned} \quad (15.47)$$

上式的透视投影速度与物体的 z 坐标有关。如果已知物体的表面模型，则有可能减缓对 z 坐标的依赖性。

(2) 二次流

二次流是最基本的光流场模型，因为它是透视投影下平面运动的准确模型；因此，二次流对曲面局部一阶泰勒级数逼近模型是有效的[Waxman 1987]。在规范化透视投影下，平面方程(15.43)可以写为

$$\frac{1}{z} = \frac{1}{a_0} - \frac{a_1}{a_0} x' - \frac{a_2}{a_0} y' \quad (15.48)$$

将 (15.48) 代入透视投影速度场模型 (15.47)，得到 8 个参数的二次流表示：

$$\begin{aligned} u &= b_1 + b_2 x + b_3 y + b_7 x^2 + b_8 xy \\ v &= b_4 + b_5 x + b_6 y + b_7 xy + b_8 y^2 \end{aligned} \quad (15.49)$$

其中

$$\begin{aligned} b_1 &= F\left(\frac{\dot{t}_x}{a_0} + \dot{\psi}\right) & b_2 &= -F\left(\frac{\dot{t}_x a_1}{a_0} + \frac{\dot{t}_z}{a_0}\right) & b_3 &= -(F\frac{\dot{t}_x a_2}{a_0} + \dot{\theta}) & b_4 &= F\left(\frac{\dot{t}_y}{a_0} - \dot{\psi}\right) \\ b_5 &= -(F\frac{\dot{t}_y a_1}{a_0} - \dot{\theta}) & b_6 &= -F\left(\frac{\dot{t}_y a_2}{a_0} + \frac{\dot{t}_z}{a_0}\right) & b_7 &= \left(\frac{\dot{t}_z a_1}{a_0} + \frac{\dot{\psi}}{F}\right) & b_8 &= \left(\frac{\dot{t}_z a_2}{a_0} - \frac{\dot{\psi}}{F}\right) \end{aligned}$$

如果我们知道平面上至少四个点的光流，就可以建立八个或更多的方程来求解八个未知参量 (b_1, \dots, b_8)。注意到 a_0 是表面点和摄像机之间的距离，它总是以比例因子的形式出现。类似的方法可以参见文献 [Diehl 1991, Anandan 1993]

为了恢复基于使用二阶光流模型的三维运动和结构参数，Waxman 等人 [Waxman 1987] 定义了 12 个运动变形参数，它们与光流的一阶和二阶偏导数有关。虽然可以得到解析解，但是光流估计的偏导数对噪声十分敏感。

15. 2. 3 延伸焦点

对纯平移运动来说，所有的光流矢量在图象平面上的投影表现为从某一点延伸出去，或是从远处汇聚到某一点，人们将此点称为延伸焦点 (focus of expansion, FOE)，或汇聚焦点。这个点实际上是物体运动方向与图像平面的交点。物体仅作纯平移运动时，物体上一点在时刻 t_k 处的三维坐标是

$$\begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} = \begin{pmatrix} x_0 + \dot{t}_x \Delta t \\ y_0 + \dot{t}_y \Delta t \\ z_0 + \dot{t}_z \Delta t \end{pmatrix} \quad (15.50)$$

其中 (x_0, y_0, z_0) 是时刻 t_0 的三维坐标， $\Delta t = t_k - t_0$ ， $\dot{t}_x, \dot{t}_y, \dot{t}_z$ 是沿 x, y, z 轴方向匀速运动的速度。在规范化透视投影下，这个点在图像平面上的坐标由下式给出：

$$\begin{pmatrix} x'_k \\ y'_k \end{pmatrix} = \begin{pmatrix} \frac{x_0 + \dot{t}_x \Delta t}{z_0 + \dot{t}_z \Delta t} \\ \frac{y_0 + \dot{t}_y \Delta t}{z_0 + \dot{t}_z \Delta t} \end{pmatrix} \quad (15.51)$$

当 $\Delta t \rightarrow \infty$ 时，上式变为：

$$\mathbf{e} = \lim_{\Delta t \rightarrow \infty} \begin{pmatrix} x'_k \\ y'_k \end{pmatrix} = \begin{pmatrix} \frac{\dot{t}_x}{\dot{t}_z} \\ \frac{\dot{t}_y}{\dot{t}_z} \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (15.52)$$

上式说明，矢量 $(\dot{t}_x / \dot{t}_z \quad \dot{t}_y / \dot{t}_z \quad 1)^T$ 指示物体平移运动的瞬时方向。当物体作匀速平移运动时，物体上的所有点将从图像平面上一个固定点 \mathbf{e} 延伸出去，如图 15.3 所示。计算延伸焦点的方法有光流估计法 [Lawton 1983]、直接搜索方法 [Jain 1983] 等。

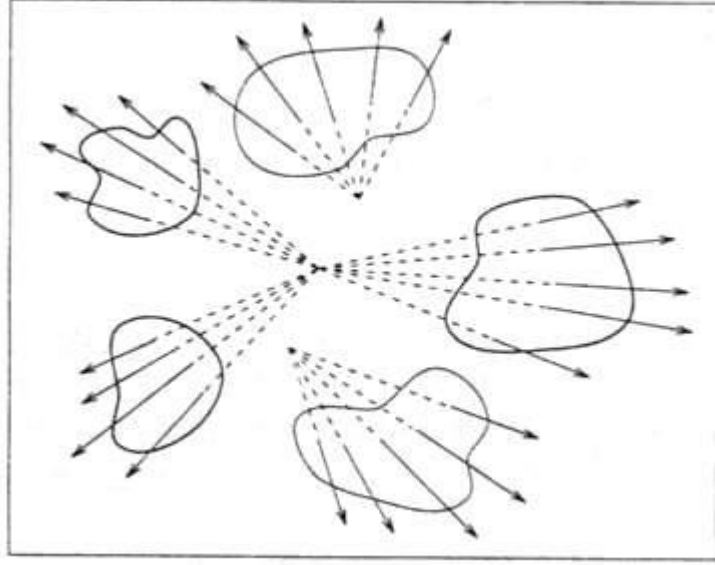


图 15.3 具有不同运动速度的物体对应的延伸焦点示意图

根据 FOE 可以确定空间点的相对深度。下面是确定三维场景点相对深度的步骤：

- ① 通过第 k 帧和第 $k+1$ 帧图像估计光流矢量，并确定 FOE 的位置。
- ② 第 $k+1$ 帧图像上第 i 个点 $\mathbf{m}_{i,k+1} = (x'_{i,k+1}, y'_{i,k+1})^T$ 对应的深度值 $z_{i,k+1}$ 可以相对于一个参考深度值 z_{ref} 表示为

$$\frac{z_{i,k+1}}{z_{ref}} = \frac{d_i}{\Delta d_i} \quad (15.53)$$

上式是利用透视投影相似三角原理得到，其中 d_i 是第 k 帧图像上的第 i 点 $\mathbf{m}_{i,k} = (x'_{i,k}, y'_{i,k})^T$ 与 FOE 之间的距离。 Δd_i 是图像点从第 k 帧到第 $k+1$ 帧的位移，如图 15.4 所示。 z_{ref} 对应三维场景点从第 k 帧到第 $k+1$ 帧沿 z 轴方向的位移，这个值无法确定。

- ③ 第 $k+1$ 帧图像上第 i 点和第 j 点相对深度可以通过 (15.53) 来计算：

$$\frac{z_{i,k+1}}{z_{j,k+1}} = \frac{d_i \Delta d_j}{d_j \Delta d_i} \quad (15.54)$$

如果物体点的运动不是简单的匀速平移，而是其它类型的运动，如，变速运动、旋转运动、变形运动等，则 FOE 可能不存在。

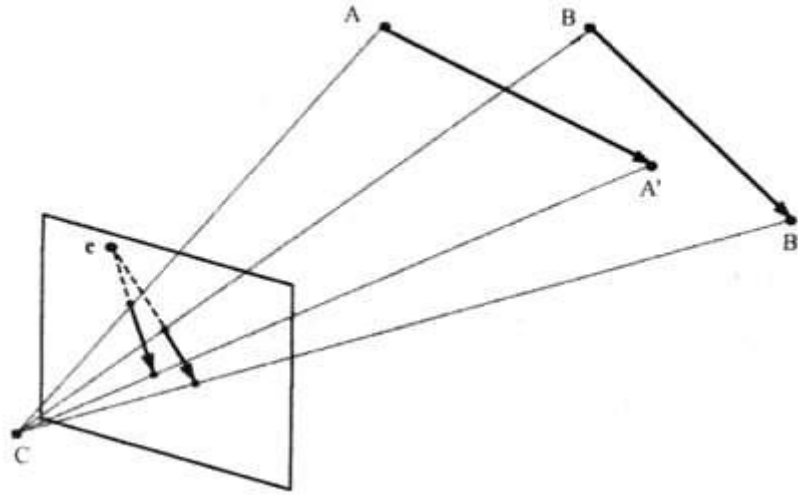


图 15.4 根据 FOE 确定空间点相对深度示意图

15. 2. 4 代数法

三维运动估计的光流法由两步组成：光流场估计和基于光流场的三维运动估计。光流估计方法已经在第 14 章介绍了。基于光流场的三维运动恢复有许多方法，各种方法之间的主要差别在于光流场估计所用的假设不同，或三维运动恢复所用的准则不同。这些方法大致可以分为两类：代数法和优化法。本节主要讨论光流法恢复三维运动参数的代数法，下一节将讨论优化法。

将方程组 (15.46) 的两个式子合并，并消除参数 z

$$\begin{aligned} & -ve_1 + ue_2 - x'(\dot{\phi} + \dot{\theta}e_1) - y'(\dot{\psi} + \dot{\theta}e_2) - x'y'(\dot{\psi}e_1 + \dot{\phi}e_2) \\ & + (x'^2 + y'^2)\dot{\theta} + (1 + y'^2)\dot{\phi}e_1 + (1 + x'^2)\dot{\psi}e_2 = uy' - vx' \end{aligned} \quad (15.55)$$

其中 e_1, e_2 是延伸焦点坐标 (见式 (15.52))。通过求解上式可以得到平移速度和一个关于比例系数的深度值，这个比例系数等于 i_z 。一般来说，同一个光流场可能对应多个解。方程 (15.55) 有 5 个未知参数， $\dot{\phi}, \dot{\psi}, \dot{\theta}, e_1, e_2$ 。在无噪声光流数据情况下，使用 5 个光流矢量，至少对应有 10 个解，而 6 个或 6 个以上的光流矢量几乎总是能唯一地确定三维运动。

(1) 解析法

式 (15.55) 可以表示成矢量矩阵的形式

$$[-v \quad u \quad -x' \quad -y' \quad -x'y' \quad x'^2 + y'^2 \quad 1 + y'^2 \quad 1 + x'^2] \mathbf{H} = uy' - vx' \quad (15.56)$$

其中

$$\mathbf{H} = [e_1 \quad e_2 \quad \dot{\phi} + \dot{\theta}e_1 \quad \dot{\psi} + \dot{\theta}e_2 \quad \dot{\psi}e_1 + \dot{\phi}e_2 \quad \dot{\theta} \quad \dot{\phi}e_1 \quad \dot{\psi}e_2]^T \quad (15.57)$$

给出在 8 个图象点处的光流矢量，建立八个线性方程就可以求解 \mathbf{H} 。通常，需要的图象点多于 8 个，并采用最小二乘法来估计 \mathbf{H} 中的 8 个参数，这样可以来缓解光流估计中误差的影响。从 (15.57) 中很容易看出，当所有的点都共面时，将生成一个二次流场，其中，系数矩阵中各元素之间相关。但所选的 8 个图像点不共面时，可以唯一地估计出参数 h_i 。

五个运动参数 $\dot{\phi}, \dot{\psi}, \dot{\theta}, e_1, e_2$ 可以依次从 8 个变量 h_i ($i=1, \dots, 8$) 中恢复出来。这样，每一个点的深度 z 就可以从模型 (15.47) 中估计出来。注意，在光流估计有误差的情况下，从 h_i 估计的运动参数不是唯一的，这是由于 8 个中间变量 h_i 中仅有 5 个是独立的。例如，从 h_1, h_2, h_6, h_7 和 h_8 估计的运动参数并不需要一定满足 h_3, h_4 和 h_5 。

(2) Heeger-Jepson 法

为了避免引进多余的变量, Heeger 和 Jepson[Heeger 1992]提出了一个两步法, 简称为 H-J 法。他们用矢量-矩阵的形式将一般光流方程 (15.46) 表示为

$$\mathbf{v}(x', y') = p(x', y')\mathbf{A}(x', y')\mathbf{V} + \mathbf{B}(x', y')\mathbf{\Omega} \quad (15.58)$$

其中

$$\mathbf{A} = \begin{bmatrix} F & 0 & -x' \\ 0 & F & -y' \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -x'y'/F & F+x'^2/F & -y' \\ -(F+y'^2/F) & x'y'/F & x' \end{bmatrix}$$

$$\mathbf{v}(x', y') = \begin{pmatrix} u(x', y') \\ v(x', y') \end{pmatrix}, \quad \mathbf{V} = [i_x \quad i_y \quad i_z]^T, \quad \mathbf{\Omega} = [\dot{\phi} \quad \dot{\psi} \quad \dot{\theta}]^T, \quad p = \frac{1}{z}$$

已知 n 个不同空间点 $(x'_1, y'_1), \dots, (x'_n, y'_n)$ 的光流矢量, 根据方程 (15.58), 可以得到 n 对方程, 缩写为

$$\mathbf{v} = \mathbf{A}(\mathbf{V})\mathbf{p} + \mathbf{B}\mathbf{\Omega} \\ = \mathbf{C}(\mathbf{V})\mathbf{q} \quad (15.59)$$

其中

$$\mathbf{A}(\mathbf{V}) = \begin{bmatrix} \mathbf{A}(x'_1, y'_1)\mathbf{V} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \mathbf{A}(x'_n, y'_n)\mathbf{V} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}(x'_1, y'_1) \\ \vdots \\ \mathbf{B}(x'_n, y'_n) \end{bmatrix}$$

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}(x'_1, y'_1) \\ \vdots \\ \mathbf{v}(x'_n, y'_n) \end{pmatrix}, \quad \mathbf{C}(\mathbf{V}) = [\mathbf{A}(\mathbf{V}) \quad \mathbf{B}], \quad \mathbf{q} = \begin{bmatrix} \mathbf{p} \\ \mathbf{\Omega} \end{bmatrix}$$

为了得到运动和结构参数 \mathbf{V} 和 \mathbf{q} 的最小二乘估计, 可以对下式求极小化

$$E(\mathbf{V}, \mathbf{q}) = \|\mathbf{v} - \mathbf{C}(\mathbf{V})\mathbf{q}\|^2 \quad (15.60)$$

显然, 相对于 \mathbf{V} 和 \mathbf{q} 使 (15.60) 达到最小值, 需要求出对每一个变量的偏导数, 并令它们等于零, 然后求解方程组。每一个方程对应一个多维空间中的曲面, 这些曲面的交点就是方程解。

计算 (15.60) 相对于 \mathbf{q} 的偏导数, 并令它等于零, 得到

$$\hat{\mathbf{q}} = [\mathbf{C}(\mathbf{V})^T \mathbf{C}(\mathbf{V})]^{-1} \mathbf{C}(\mathbf{V})^T \mathbf{v} \quad (15.61)$$

它是关于 \mathbf{V} 的曲面。显而易见, 由于 \mathbf{V} 也是一个未知量, 方程 (15.61) 不能被直接用来求 \mathbf{q} 的估值。然而, 因为实际的解位于这个曲面上, 因此可以将其代入 (15.60), 然后仅求关于 \mathbf{V} 的极小化。对于仅含有 \mathbf{V} 的非线性极小化函数, 可以通过搜索过程找出使 (15.60) 达到极小值的 \mathbf{V} 值。为了得到全局极小值, 必须对每一步估计的 \mathbf{V} 值进行扰动, 然后, 再求解方程 (15.61), 因此, 这种方法的计算量很大。

为此, Heeger 和 Jepson 提出了一个有效的方法来搜索仅含有 \mathbf{V} 的极小化函数, 并用一些常用的代数方法求解, 而不必在每一步计算 (15.61)。进一步说, 由于估计 \mathbf{V} 的算法仅仅取决于一个比例因子, 不失一般性地, 可以将搜索空间限制在单位球面。单位球面可以用球面坐标的两个角来描述。一旦确定了 \mathbf{V} 的最佳估计, 就可以从 (15.61) 中计算 \mathbf{q} 的最小二乘估计。实验表明这个方法对于光流估计误差相当稳定。

15.2.5 优化法

优化法的基本思想是独立地对三维运动和结构参数估计值进行扰动, 直到估计的三维运动在图像平面上的投影速度与实际的光流场相一致。从一个合理的初始值或通过其它方法得到的初始估计开始, 优化算法可以成功地跟踪长序列图像中的小变化量运动和结构参数。为了防止优化算法发散或收敛到代价函数的局部最小值, 通常需要某种在三维运动和结构参数上的平滑约束条件。

Morikawa 和 Harashima [Morikawa 1991] 使用正交速度场模型（见式 15.42）提出了一种估计三维运动的优化法。假定场景中只有单一的刚体运动，则运动参数是全局参数。然而，深度参数随空间变化。设 $\Omega(k) = (\dot{\phi}(k), \dot{\phi}(k), \dot{\theta}(k))^T$ ， $\mathbf{V}(k) = (\dot{i}_x(k), \dot{i}_y(k))^T$ 。若给出三维运动和深度参数的初始值，我们利用

$$\begin{aligned}\Omega(k) &= \Omega(k-1) + \Delta\Omega \\ \mathbf{V}(k) &= \mathbf{V}(k-1) + \Delta\mathbf{V} \\ z(x, y)(k) &= z(x, y)(k-1) + \Delta z(x, y)\end{aligned}\quad (15.62)$$

逐步修正上一次的估计结果。在这里，问题简化为寻找逐帧递增的参数。运动平滑度约束条件意味着修正项（递增的参数）应该很小。因此，根据下面的函数引进一个平滑测度：

$$\|P\|^2 = \frac{\alpha}{N} \|\Delta\Omega\|^2 + \frac{\beta}{N} \|\Delta\mathbf{V}\|^2 + \gamma \sum_{i=1}^N (\Delta z_i)^2 \quad (15.63)$$

其中 α, β, γ 是比例系数， $\|\bullet\|$ 是 L_2 的范数， N 是图像点的数量。

投影速度场估计值与实际光流场的均方差和三维运动参数平滑约束条件共同构成代价函数

$$E = \sum_i [(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2] + \|P\|^2 \quad (15.64)$$

其中

$$\begin{aligned}\hat{u}_i &= (\dot{i}_x + \Delta\dot{i}_x) + (\dot{\psi} + \Delta\dot{\psi})(z_i + \Delta z_i) - (\dot{\theta} + \Delta\dot{\theta})y' \\ \hat{v}_i &= (\dot{i}_y + \Delta\dot{i}_y) + (\dot{\theta} + \Delta\dot{\theta})x' - (\dot{\phi} + \Delta\dot{\phi})(z_i + \Delta z_i)\end{aligned}$$

优化法可以通过基于梯度法或模拟退火法来实现。因为这个过程使用正交速度场，所以能够找到一个满足附加常数和比例系数的深度参数，同时可以确定关于这一比例常数倒数的两个旋转角度。在正交模型下估计运动和结构的类似方法见 [Kanatani 1986]

15. 2. 6 几种方法比较

基于光流的运动和结构恢复不需要建立特征点的对应，只依赖于估计的光流场，但是光流估计本身是“不适定”问题，对于观测噪声非常敏感。文献 [Tekalp 1995] 对三种基于光流算法的性能进行了比较，它们是：i) 解析法，ii) Heeger-Jepson 法，iii) 优化法。其中前两种方法是基于透视投影光流场方法，后一种是基于正交投影光流场方法。表 15.1 给出了使用光流模型进行 3D 运动和深度参数估计的仿真结果，在仿真中共使用了 30 个模拟点，其中 x, y 值均匀地在区间 $[-25, 25]$ cm 之间， z 值均匀分布在 $[70, 100]$ cm 范围内的。令焦距参数 $f = 50$ mm。为了测试这些方法对于光流估计误差的敏感度，它们在光流矢量模拟数据的基础上分别增加 3% 和 10% 的误差。

表 15.1 运动参数的比较（角度单位：弧度，平移单位：像素）

真实值	解析法		Heeger-Jepson 法		
	无误差	3%误差	无误差	3%误差	10%误差
$\Omega_1=0.007$	0.0070	0.0133	0.0064	0.0094	0.0036
$\Omega_2=0.010$	0.0099	0.0058	0.0105	0.0066	0.0008
$\Omega_3=0.025$	0.0250	0.0232	0.0249	0.0251	0.0249
$V_2/V_1=1.80$	1.7999	2.9571	1.7872	1.9439	2.7814
$V_3/V_1=0.48$	0.4799	10.121	0.4693	0.5822	0.8343
深度误差	1.73E-5	0.824	1.41E-3	0.0181	0.0539
匹配误差	2.17E-7	0.0152	2.71E-5	5.99E-4	0.0018

在表 15.1，“深度误差”对应于深度估计中的规范化均方根误差：

$$\text{深度误差} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{z_i - \hat{z}_i}{z_i} \right)^2} \quad (15.66)$$

其中 $n = 30$ ，匹配误差是指三维运动和深度参数估计值生成的光流场 $\hat{\mathbf{v}}$ 与输入的光流场 \mathbf{v} 的偏差：

$$\text{匹配误差} = \frac{1}{n} \sqrt{\sum_{i=1}^n (u - \hat{u})^2 + (v - \hat{v})^2} \quad (15.67)$$

从表 15.1 容易看出解析解法对于噪声较敏感，这主要是因为它引进了三个冗余的中间未知量，以便使线性公式（15.56）成立。可以看出八个中间未知量中只要有五个就足以求解运动参数，在有噪声的情况下，五个中间未知量的不同组合可能产生不同的运动估计。虽然最小二乘方可以得到中间未知量，但最后的运动估计不是真正最小二乘方估计。H-J 法总可以找到运动和深度参数，使最小二乘方误差函数达到最小。因此，H-J 法是一种光流估计的鲁棒方法。注意，在 H-J 法中，估计值的正确性依赖于搜索过程中使用的步长（单位球面角度的增量）。表 15.1 所用的步长为 2° 。

Morikawa 法是一种基于正交投影光流模型方法，它的性能依赖于给定光流矢量偏离正交模型（15.42）的程度。偏离量和物体深度方向的尺度与物体各点投影中心平均深度之间的比值 R 有关（回顾一下 15.1.1 节的正交投影模型）， R 可定义为

$$R = \frac{\Delta z}{2z_{ref}}$$

取 $R \approx 18\%$ 时，Morikawa 法没能正常地工作。表 15.2 所示的是完全正交投影光流场和 Morikawa 法在 $R \approx 1\%$ 时的透视光流估计性能的比较。Morikawa 算法面临的另一个重要问题是如何对未知参数初始值。在上面的模拟中，所有参数的初始化选在它们真实值的 $\pm 5\%$ 上下范围内。一般来说，代价函数有多个最小值，而且结果非常依赖初始点。建议使用 Morikawa 法跟踪小变化量运动图像参数，这里第一帧的初始值可以通过其它方法（如 14 章中的两步迭代算法）确定。

表 15.2 运动参数的比较（角度单位：弧度，位移单位：像素）

真实值	正交	透视, $R \approx 1\%$
$\Omega_1 = 0.007$	0.00713	0.00677
$\Omega_2 = 0.010$	0.01013	0.00977
$\Omega_3 = 0.025$	0.02513	0.02477
$V_1 = -.100$	0.09987	0.10023
$V_2 = 0.180$	0.17987	0.18023
深度误差	0.00011	0.00023
匹配误差	0.00322	0.00489

15. 3 光流分割

一般来说，真实场景图像序列包含有许多个运动的物体，为了分析和理解这样的动态场景，首先需要对运动图像序列中各个独立的运动物体进行标记，这种标记过程称为三维运动分割。运动图像分割可以直接使用时空图像亮度和梯度信息来分割独立的运动区域，比如，使用第 14 章所述的差分图像来分割图像，也可以采用参数模型，利用最小二乘法拟合运动区域，从而将图像分割成小的区域[Hoetter 1988, Diehl 1991]。本节主要讨论基于稠密光流场的不连续性来分割运动图像，即首先估计运动图像稠密光流场，然后将相似的光流矢量合并，对应一个独立的运动物体。

15. 3. 1 Hough 变换分割方法

Hough 变换是众所周知的聚类方法，这种方法可以在量化参数空间中表决出最具有代表性的特征值。对于使用 6 个参数的仿射光流模型 (15.44) 的光流分割，可以直接应用 Hough 变换法。首先，确定每一个参数的最小值和最大值，并将 6 维特征空间 a_1, \dots, a_6 量化成某些参数状态，然后，每一个光流矢量 $\mathbf{v}(\mathbf{x}') = [u(\mathbf{x}') \quad v(\mathbf{x}')]^T$ “投票”支持一个已量化的参数集，使的下式达到极小：

$$\gamma^2(\mathbf{x}) \doteq \alpha^2(\mathbf{x}) + \beta^2(\mathbf{x}) \quad (15.68)$$

其中 $\alpha(\mathbf{x}') = u(\mathbf{x}') - a_1 - a_2 x' - a_3 y'$ 和 $\beta(\mathbf{x}') = v(\mathbf{x}') - a_4 - a_5 x' - a_6 y'$ 。如果一个参数集得到的赞成票数大于预定数量，则很有可能表示候选的运动。由此确定分类号 K 和用于标记每一个光流矢量的对应参数集。不过，这一方法的缺陷是需要的计算量很大。

为了减小计算量，[Adiv 1985]提出了一个使用改进的 Hough 变换算法。这个算法分两个阶段，在算法的第一阶段，将光流矢量连通集组合在一起，形成与单一参数集一致的成分。这里给出几种简化方法以减轻计算量，包括：1) 把参数空间分成两个不相连的子集合 $\{a_1, a_2, a_3\} \times \{a_4, a_5, a_6\}$ ，来实现两个三维 Hough 变换，2) 多分辨率 Hough 变换，其中在每一个分辨率水平上，参数空间将在上一个分辨率水平上的估计值周围进行到量化，3) 多路传输 Hough 变换技术，其中，与候选参数最一致的光流矢量首先得到组合。在算法的第二阶段中，将这些在算法的第一阶段生成的成分（它在最小二乘方意义下的二次流模型 (15.49) 相一致）合并成块。目前已经提出了几种合并准则。最后，可以将未合并的光流矢量归并到它们邻近的块中。

总之，改进的 Hough 变换分割方法建立在这样的基础上，即首先将光流矢量聚类成一个个小组，每一个小组与一个由运动小平面产生的光流场一致。然后基于某一个合并准则将这些小组合并成块。

15. 3. 2 贝叶斯分割方法

显然，光流分割的成功与否与光流场估计的正确性密切相关，因此，为了得到最佳分割结果应该同时进行光流场的估计和分割。本节讨论基于运动场表示的联合贝叶斯光流分割方法。在最大后验概率(maximum a posteriori probability, MAP)形式下，光流和分割场的相互依赖性可通过 Gibbs 分布表示[Derin 1987]。利用最高置信度优先(highest confidence first, HCF)和迭代条件模式(iterated conditional mode, ICM)算法求解最后的优化问题，即找到一个密集系列的运动矢量，一系列分割标记和一系列映射参数的估计值。

(1) 运动场模型

假定场景中有 k 个独立运动的物体，这里由每一个物体引起的二维运动可通过一个参数模型近似。那么， $\mathbf{v}_i(\mathbf{x}') = (u_i(\mathbf{x}'), v_i(\mathbf{x}'))$ 表示图像中第 i 个光流矢量， $l(\mathbf{x}')$ 表示每一个像素对应的分割标记，它的取值范围为 $1, 2, \dots, k$ 。光流矢量 $\mathbf{v}_i(\mathbf{x}')$ 可以表示成参数化光流场 $\tilde{\mathbf{v}}_i(\mathbf{x}') = (\tilde{u}_i(\mathbf{x}') \quad \tilde{v}_i(\mathbf{x}'))^T$ 和非参数化光流场残差 $\mathbf{e}_i(\mathbf{x}')$ 之和：

$$\mathbf{v}_i(\mathbf{x}') = \tilde{\mathbf{v}}_i(\mathbf{x}') + \mathbf{e}_i(\mathbf{x}') \quad (15.69)$$

其中残差 $\mathbf{e}_i(\mathbf{x}')$ 是由局部运动或建模误差引起的。参数化光流场 $\tilde{\mathbf{v}}_i(\mathbf{x}')$ 可以表示为

$$\begin{aligned} \tilde{u}_i(\mathbf{x}') &= a_1 x' + a_2 y' - a_3 + a_7 x'^2 + a_8 x' y' \\ \tilde{v}_i(\mathbf{x}') &= a_4 x' + a_5 y' - a_6 + a_7 x' y' + a_8 y'^2 \end{aligned} \quad (15.70)$$

运动场的参数分量显而易见是依赖于分割标记 $l(\mathbf{x}')$ 。

(2) 分割过程表示

假定第 k 帧图像表示为

$$g_k(\mathbf{x}') = f_k(\mathbf{x}') + n_k(\mathbf{x}') \quad (15.71)$$

相对于光流 $\mathbf{v}_1, \mathbf{v}_2$ 和分割标记 l 求后验概率密度函数(probability density function, pdf)的最大值：

$$p(\mathbf{v}_1, \mathbf{v}_2, l | g_k, g_{k+1}) = \frac{p(g_{k+1} | g_k, \mathbf{v}_1, \mathbf{v}_2, l) p(\mathbf{v}_1, \mathbf{v}_2 | l, g_k) p(l | g_k)}{p(g_{k+1} | g_k)} \quad (15.72)$$

上式右边第一个条件概率密度函数 $p(g_{k+1} | g_k, \mathbf{v}_1, \mathbf{v}_2, l)$ 提供了在第 k 帧情况下，当前位移和分割估计与第 $k+1$ 帧相一致程度的测度。它用 Gibbs 分布模型表示成

$$p(g_{k+1} | g_k, \mathbf{v}_1, \mathbf{v}_2, l) = \frac{1}{Q_1} \exp\{-U_1(g_{k+1} | g_k, \mathbf{v}_1, \mathbf{v}_2, l)\} \quad (15.73)$$

其中 Q_1 是分割函数（常量），且

$$U_1(g_{k+1}|g_k, \mathbf{v}_1, \mathbf{v}_2, l) = \sum_{\mathbf{x}} [g_k(\mathbf{x}') - g_{k+1}(\mathbf{x}' + \mathbf{v}(\mathbf{x}')\Delta t)]^2 \quad (15.74)$$

称为 Gibbs 势能。式 (15.72) 分子中的第二项是在给定运动分割和图像条件下，位移场的条件概率密度函数，它用 Gibbs 分布作为模型：

$$p(\mathbf{v}_1, \mathbf{v}_2|l, g_k) = p(\mathbf{v}_1, \mathbf{v}_2|l) = \frac{1}{Q_2} \exp\{-U_2(\mathbf{v}_1, \mathbf{v}_2|l)\} \quad (15.75)$$

其中 Q_2 是常量，同时

$$U_2(\mathbf{v}_1, \mathbf{v}_2|l) = a \sum_{\mathbf{x}} \|\mathbf{v}(\mathbf{x}') - \bar{\mathbf{v}}(\mathbf{x}')\|^2 + \beta \sum_{\mathbf{x}_i} \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_i}} \|\mathbf{v}(\mathbf{x}_i') - \mathbf{v}(\mathbf{x}_j')\|^2 \delta(l(\mathbf{x}_i') - l(\mathbf{x}_j')) \quad (15.76)$$

是相应的 Gibbs 势能， $\|\bullet\|$ 代表欧几里德距离， $\mathbf{N}_{\mathbf{x}}$ 是在位置 \mathbf{x}' 的邻域集合。式 (15.76) 中的第一项强调对非参数化光流场残差 $\mathbf{e}_i(\mathbf{x}')$ 范数极小化的估计。式 (15.76) 的第二项只是在光流估计值上强加了一个分段局部平滑约束条件，没有引进任何附加变量。可以看出，这一项仅对和位置 \mathbf{x} 共享同样的分割标记的 $\mathbf{N}_{\mathbf{x}}$ 邻域中的那些像素起作用。因此，仅在单一物体产生的光流矢量上强加了一个空间平滑条件。参数 α 和 β 用来确定两项对 Gibbs 势能的影响比例系数。

式 (15.72) 中的第三项是分割场的先验概率模型，由下式给出：

$$p(l|g_k) = p(l) = \frac{1}{Q_3} \sum_{\omega \in \Omega} \exp\{-U_3(l)\} \delta(l - \omega) \quad (15.77)$$

其中 Ω 表示离散随机矢量 l 的采样空间， U_3 由下式给出

$$U_3(l) = \sum_{\mathbf{x}_i} \sum_{\mathbf{x}_j \in \mathbf{N}_{\mathbf{x}_i}} V_C(z(\mathbf{x}_i'), z(\mathbf{x}_j')) \quad (15.78)$$

其中 $\mathbf{N}_{\mathbf{x}_i}$ 表示标记场的邻域系，同时

$$V_C(l(\mathbf{x}_i'), l(\mathbf{x}_j')) = \begin{cases} -\gamma & l(\mathbf{x}_i') = l(\mathbf{x}_j') \\ +\gamma & \text{其它} \end{cases} \quad (15.79)$$

Q_3 由下式给出：

$$Q_3 = \sum_{\omega \in \Omega} \exp\{-U_3(l = \omega)\} \quad (15.80)$$

虽然区域边界通常与亮度边缘是相重合的，但通常忽略标记对于图像亮度的依赖关系。

(3) 算法

求后验概率密度函数(15.72)式的最大值等价于求代价函数的最小值，

$$E = U_1(g_{k+1}|g_k, \mathbf{v}_1, \mathbf{v}_2, l) + U_2(\mathbf{v}_1, \mathbf{v}_2|l) + U_3(l) \quad (15.81)$$

上式是由势能函数(15.74)，(15.76)，(15.78)构成的。相对于所有的未知量直接求 (15.81) 的最小值是一个相当困难的问题，因为运动和分割场构成一个较大的未知量的集合。为此，我们通过下面的两步迭代法来求解 (15.81) 的最小值[Chang 1994]:

1. 给出参数 $\mathbf{a}_i, i = 1, \dots, k$ 和 l 的最有效的估值，修正光流场 $\mathbf{v}_1, \mathbf{v}_2$ 。这步涉及求改进的代价函数的最小值

$$E_1 = \sum_{\mathbf{x}} [g_k(\mathbf{x}') - g_{k+1}(\mathbf{x}' + \mathbf{v}(\mathbf{x}')\Delta t)]^2 + \alpha \sum_{\mathbf{x}} \|\mathbf{v}(\mathbf{x}') - \tilde{\mathbf{v}}(\mathbf{x}')\|^2 + \beta \sum_{\mathbf{x}_i} \sum_{\mathbf{x}_j \in \mathbf{N}\mathbf{x}_i} \|\mathbf{v}(\mathbf{x}'_i) - \mathbf{v}(\mathbf{x}'_j)\|^2 \delta(l(\mathbf{x}'_i) - l(\mathbf{x}'_j)), \quad (15.82)$$

它由 (15.81) 中包含 $\mathbf{v}(\mathbf{x}')$ 的所有项构成。第一项表明 $\mathbf{v}(\mathbf{x}')$ 与观察结果的接近程度，第二项和第三项在运动估计值上强加先验约束条件，使得运动估计值与参数化光流模型相一致，并在每一个区域中平滑变化。为了求这个能量函数的最小值，可以利用由 Chou 和 Brown[Chou 1991]提出的 HCF 法，该方法属于代数求解方法，可以有效地解决邻域相互作用的多变量问题的优化。

2. 修正分割场 l ，假设光流场 $\mathbf{v}(\mathbf{x}')$ 是已知的。这步涉及到求既包含 $\tilde{\mathbf{v}}(\mathbf{x})$ 又包含 l 的 (15.81) 中所有项的最小值，由下式给出：

$$E_2 = \alpha \sum_{\mathbf{x}} \|\mathbf{v}(\mathbf{x}') - \tilde{\mathbf{v}}(\mathbf{x}')\|^2 + \sum_{\mathbf{x}_i, \mathbf{x}_j \in \mathbf{N}} \sum_{\mathbf{x}_i} V_c(l(\mathbf{x}'_i), l(\mathbf{x}'_j)) \quad (15.83)$$

上式的第一项定量地表示出 $\tilde{\mathbf{v}}(\mathbf{x}')$ 和 $\mathbf{v}(\mathbf{x}')$ 相一致的程度。第二项与当前配置的分割标记的先验概率有关。可以利用 ICM 方法对 E_2 优化[Chang 1993]。在每一个区域里通过最小二乘方估计来修正映射参数 a_i (见式(15.70))

利用具有全局平滑限制条件的贝叶斯法能够找到光流的初始估值。给定这个估值，利用与 Wang 和 Adelson 的[Wang 1994]相似过程初始化分割标记。随机参数 α, β 和 γ 的确定是有关设计的问题。一种方案是选择能够提供在一定动态范围内正确的数值，以致于代价函数 (15.79) 中的每一项具有相等的重要性。然而，因为我们用两步实现优化，比例 α/γ 也是很重要的。我们建议选择 $1 \leq \alpha/\gamma \leq 5$ ，这点依赖于用分段参数模型能恰当表示运动场的程度和我们是否有足够的类别个数。

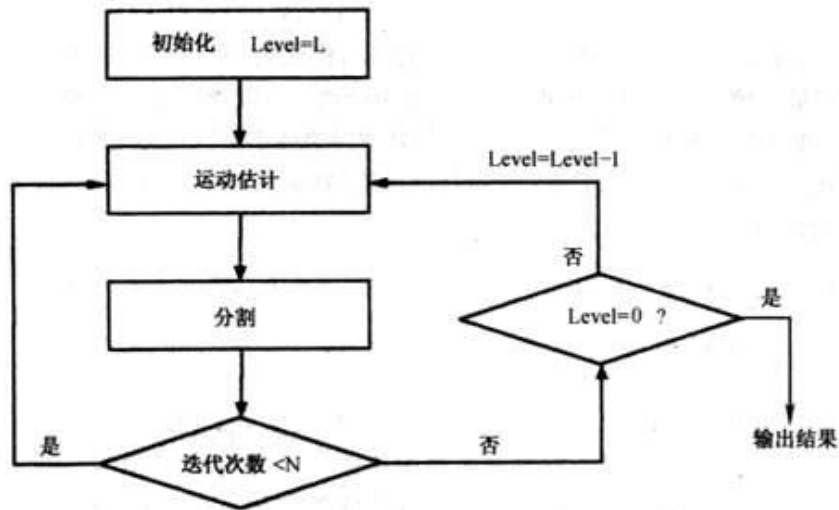


图 15.5 同时进行分割和估计的 MAP 法的方框图

通过形成图像 g_k 和 g_{k+1} 的连续的低通滤波模型，该算法的分级实现成为可能。由此，可在不同的分辨率下估计 $\mathbf{v}_1, \mathbf{v}_2$ 和 l 的数量。每一级的结果被用来初始化紧接着的低一级。图 11.5 描述出分级算法的框图。请注意：分割标记的 Gibbs 模型已经被扩展到包括由 Kato 等人[Kato 1993]给出的一定比例的邻域内。

(4) 讨论

有一点很重要：同时进行估计和分割的形式不仅适用于出现多个运动物体下的三维运动和结构估计，而且提供了改进的光流估计。几种已知的运动分析法可被表述成这个形式的特例。如果我们仅保留(15.81)中的第一项和第三项，同时假设所有的位置具有相同的分割标记，那么我们得到具有全局平滑限制条件的贝叶斯运动估计。Iu [Iu 93]提出的运动估计的算法也用这两项，但却用局部分离舍弃函数代替 $\delta(\cdot)$ 函数。

由 Stiller[Stiller 1994]提出的运动估计值和区域标注的方法涉及 (15.81) 式中所有的项，除了 (15.76) 式中的第一项。进一步说，Stiller 算法中的分割标记只不过作为供流量场上的分段平滑度限制条件之用的标记，同时它也不希望加强流矢量与参数分量的一致性。

练习题

15. 1 请阐述三维运动估计的点对应法和光流法的区别。
15. 2 当物体没有平移运动时，即 $(t_x, t_y, t_z) = (0, 0, 0)$ ，我们就无法用式 (15.13) 估计深度值。在这种情况下，至少需要多少对应点才能唯一地确定旋转矩阵 \mathbf{R} ？
15. 3 请说明基于平行投影模型和透视投影模型估计三维运动的异同。
15. 4 试解释外极线方程的几何意义。
15. 5 若已知 8 个或 8 个以上位于同一空间平面的三维点的光流矢量时，是否可以唯一地确定运动参数，为什么？
15. 6 试讨论贝叶斯光流分割法中各个尺度系数 $\alpha, \lambda, \gamma, \psi$ 的选择准则。

计算机作业

15. 1 从任意一个场景图像序列中选出相邻两帧图像，假设图像的变化是由摄像机运动引起的，两帧图像的最大变化不超过十几个像素。试根据利用点对应法和外极线方程编制一个程序，该程序具有如下功能：
 - (1) 可以自动在另一幅图像中找到对应点
 - (2) 可以估计出该图像对的基本矩阵，并可以画出任意点的外极线图可以完成图像的。
 - (3) 可以对图像实现校准，使得图像的外极线尽量平行
15. 2 试编制一个程序，该程序可以自动在两帧图像中找到对应点，并把局外点自动消除(考虑使用松弛方法)。