基于标号场的光流法二维运动估计 [1]

杨 勇 王 桥 吴乐南
(东南大学无线电系 南京 210096)

摘 要 该文提出了一种改进的二维运动估计光流算法. 在光流基本约束和全局平滑性约束的基础上, 引入二维运动标号场, 对 Horn-Schunck 算法进行了改进, 不仅加快了算法的收敛速度, 而且提高了运动估计的准确性, 减轻了遮挡边界问题的影响. 实验结果证明了该算法的有效性.

关键词 光流场, 标号场, 二维运动估计
中图号 TN911.73

1 引 言

二维运动估计是根据瞬时像素速度或帧到帧的像素来估计投影到图像平面上的运动, 它是数字视频处理的基础. 由于运动补偿是利用运动图像序列时空冗余信息的最有效途径之一, 因而基于运动估计的运动补偿也已成为数字视频压缩的国际标准 (H.261 , MPEG1-2) . 另外, 基于二维运动估计的运动分割和三维景物重构, 是计算机视觉理论 [1,2] 研究的一个重点课题.

基于图像梯度的方法是二维运动估计的主要方法之一. 自 Horn 和 Schunck[3] 于 1981 年提出了光流法开始, 出现了许多针对光流约束基本方程的正则化约束 [4−7] 及相应算法. 所有方法都必须解决时空偏导数的离散化及其计算这个问题. 然而, 在存在空间高频噪声和图像序列存在较大帧间位移量的情况下, 不同的时空偏导数离散化方法都存在相应的偏差. 本文的工作从基于时空梯度的局部最优估计方法开始, 对其进行了两个方面的改进: (1) 求解标号场 (Labeling Field) . 根据图像帧差将图像分为 "变化区域" 和 "未变化区域" 两部分; (2) 根据标号场对二维运动估计算法进行优化. 对 MPEG-4 标准测试图像的运动估计结果显示, 该算法不仅加快了算法的收敛速度, 而且具有较高的峰值信噪比.

2 基于标号场的运动估计算法

2.1 标号场

定义第 k 帧和第 $k+1$ 帧图像的强度分别为 $f(n_1,n_2,k)$ 和 $f(n_1,n_2,k+1)$, 两帧图像之间的帧差 $\mathrm{FD}_{k,k+1}$ 定义为

$$\mathrm{FD}_{k,k+1}(x_1,x_2)=f(x_1,x_2,k+1)-f(x_1,x_2,k) \tag{1}$$

理想情况下, "未变化区域" 的帧差为 0 , 然而由于观察噪声的存在, 在许多位置上, 其值并不为 0 . 因而, 通过对帧差的阈值化来检测 "变化区域" 和 "未变化区域", 即

$$z_{k,k+1}(x_1,x_2)=\begin{cases}1, & |\mathrm{FD}_{k,k+1}(x_1,x_2)|>T\\ 0, & |\mathrm{FD}_{k,k+1}(x_1,x_2)|\le T\end{cases} \tag{2}$$

其中 $z_{k,k+1}(x_1,x_2)$ 称为标号场 [8], T 为一最优阈值 [9] .

[1] 2000-06-27 收到, 2000-10-25 定稿
国家自然科学基金资助项目 (60072013, 60002008)

2.2 结合标号场的光流场

由光流约束方程：

$$f_x u + f_y v + f_t = 0 \tag{3}$$

其中 f_x, f_y, f_t 是 x 方向、y 方向和 t 方向的梯度分量. u 和 v 分别是 x 方向和 y 方向的速度分量. 由于上式有两个未知量，求解速度场必须附加相应的约束. Horn 和 Schunck 假设运动场不仅满足光流约束方程，而且还满足全局平滑性，提出了如下的平滑性约束，u, v 满足 (4) 式

$$L(u,v) = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \tag{4}$$

达到最小. 综合 (3) 和 (4) 式，光流场的计算归结为求解如下的变分问题：

$$(\hat{u},\hat{v}) = \arg\min_{u,v} \int\int \left[\lambda(f_x u + f_y v + f_t)^2 + \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right)\right] \times \mathrm{d}x\mathrm{d}y \tag{5}$$

其中 $\hat{u}, \hat{v}$ 表示最优估计. 其对应的欧拉方程为

$$\left.\begin{aligned} \Delta u &= \lambda f_x(f_x u + f_y v + f_t) \\ \Delta v &= \lambda f_y(f_x u + f_y v + f_t) \end{aligned}\right\} \tag{6}$$

其中 $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$ 是拉普拉斯算子.

下面，将标号场引入 (6) 式的离散化求解过程. 首先，考虑拉普拉斯算子的离散化. 考虑到对于“变化区域”内的像素，应当加强其邻域中具有较大运动矢量的像素对它的影响，而对于“未变化区域”内的像素，应当削弱其邻域中具有较大运动矢量的像素对它的影响. 特别地，对于遮挡 (Occluding) 边界附近的像素，通过加权处理，可以减轻遮挡边界问题对它的影响. 因而，对拉普拉斯算子进行如下的估计：

$$\left.\begin{aligned} \Delta u(i,j) &= u(i,j) - u^*(i,j) = u(i,j) - \frac{1}{A}\left(\varepsilon_{i,j}\sum_{i,j\in N_1} u^{(n-1)}(i,j) + (1-\varepsilon_{i,j})\sum_{i,j\in N_2} u^{(n-1)}(i,j)\right) \\ \Delta v(i,j) &= v(i,j) - v^*(i,j) = v(i,j) - \frac{1}{A}\left(\varepsilon_{i,j}\sum_{i,j\in N_1} v^{(n-1)}(i,j) + (1-\varepsilon_{i,j})\sum_{i,j\in N_2} v^{(n-1)}(i,j)\right) \end{aligned}\right\} \tag{7}$$

其中 $(\bullet)^{(n-1)}$ 表示第 $n-1$ 次迭代结果 (计算时取 $u^{(0)}, v^{(0)}$ 为 0)，以下类同. A 为 (i,j) 的邻域 N 内像素的个数，ε 为一个加权系数，ε, N_1 和 N_2 分别满足

$$\varepsilon_{i,j} = \begin{cases} 0.5+\alpha, & z_{k,k+1}(i,j) = 1, \\ 0.5-\alpha, & z_{k,k+1}(i,j) = 0, \end{cases} \quad \text{且 } 0 \le \alpha \le 0.5 \tag{8}$$

$$N_1 = \{(i_1,j_1) | z(i_1,j_1) = 1, \quad \text{且}(i_1,j_1) \in N\} \tag{9}$$

$$N_2 = \{(i_2,j_2) | z(i_2,j_2) = 0, \quad \text{且}(i_2,j_2) \in N\} \tag{10}$$

由 (7) 式，得到如下的结论：

(1) 对于完全处于未变化区域内的像素，此时如取 $\alpha = 0.5$，则 $\varepsilon_{i,j} = 0$，

$$\Delta u(i,j) = u(i,j) - u^*(i,j) = u(i,j) - \frac{1}{A}\left(\sum_{i,j\in N_2} u^{(n-1)}(i,j)\right) \tag{11}$$

结果与 Horn-Schunck 方法相同.

(2) 对于完全处于变化区域内的像素，此时如取 $\alpha = 0.5$，则 $\varepsilon_{i,j} = 1$，

$$\Delta u(i,j) = u(i,j) - u^*(i,j) = u(i,j) - \frac{1}{A}\left(\sum_{i,j\in N_1} u^{(n-1)}(i,j)\right) \tag{12}$$

结果与 Horn-Schunck 方法相同.

(3) 对于处于遮挡边界上的运动像素点，此时 $\varepsilon_{i,j} > 1 - \varepsilon_{i,j}$，其邻域中具有较大运动矢量的像素对其影响得到加强.

(4) 对于处于遮挡边界附近的非运动像素点，此时 $\varepsilon_{i,j} < 1 - \varepsilon_{i,j}$，其邻域中具有较大运动矢量的像素对其影响被削弱.

其次，讨论系数 λ 的物理意义. 在 (5) 式中，第一项是光流基本约束项，它决定了运动估计结果偏离基本等式的误差；第二项是光滑性约束项，它决定了运动估计结果偏离光滑性要求的误差. 系数 λ 是两种偏差间的一个折衷，即加权系数. 已经知道，标号场将运动图像分成了“变化区域”和“未变化区域”两部分，在变化区域内，与偏离光滑性要求的误差相比，偏离基本等式的误差是运动估计结果误差的主要源项. 而在未变化区域内，两者相比，偏离光滑性要求的误差占主要地位. 根据以上分析，对加权系数 λ 进行修正，令

$$\lambda = \lambda_0^{\mu} \tag{13}$$

其中 λ_0 为一常数，且 $\lambda_0 > 1$，

$$\mu(i,j) = \begin{cases} 1, & z_{k,k+1}(i,j) = 1 \\ -1, & z_{k,k+1}(i,j) = 0 \end{cases} \tag{14}$$

综合以上因素，将 (7) 式和 (13) 式代入 (6) 式，最终得到其相应的离散化求解形式为

$$\left.\begin{aligned} u^{(n-1)} &= u^{*(n)} - \lambda_0^{\mu}\frac{(f_x u^{*(n)} + f_y v^{*(n)} + f_t)}{1 + \lambda_0^{\mu}(f_x^2 + f_y^2)} f_x \\ v^{(n-1)} &= v^{*(n)} - \lambda_0^{\mu}\frac{(f_x u^{*(n)} + f_y v^{*(n)} + f_t)}{1 + \lambda_0^{\mu}(f_x^2 + f_y^2)} f_y \end{aligned}\right\} \tag{15}$$

3 实　验

为方便起见，称新算法为 LBOF(Labeling-Based Optical Flow) 方法，用标准测试图像序列“销售员 (Saleman)”和“特雷弗 (Trevor)”进行二维运动估计，将结果与 Horn-Schunck 方法 (简称 Horn 方法) 进行比较. 图 1 和图 2 分别是“销售员”序列中第 9 帧和第 10 帧图像. 图 3 和图 4 分别是用 Horn 方法和 LBOF 方法进行二维运动估计.

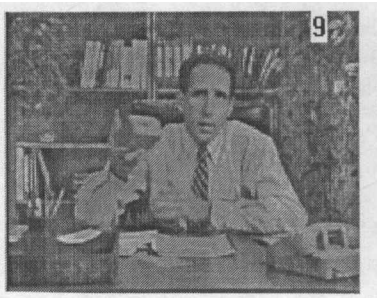

图 1 第 9 帧原图像

图 2 第 10 帧原图像

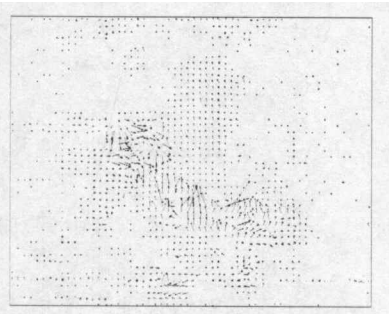

图 3 Horn 方法运动估计结果

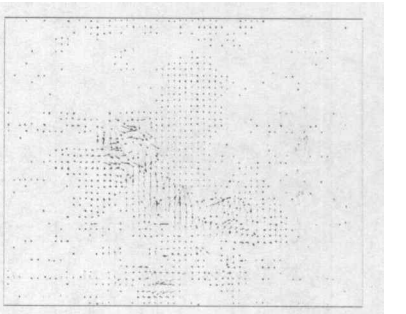

图 4 LBOF 方法运动估计结果

在奔腾 II 250、内存 64 兆的微机上实现时，用 Horn 方法需要 3011ms，峰值信噪比 (PSNR) 为 31.807dB；而用 LBOF 方法需要 2276 ms， PSNR 为 31.813dB． LBOF 算法不仅具有较快的收敛速度，而且 PSNR 也有所提高．

本文对图像序列“销售员”和“特雷弗”中的多帧图像，分别用上述两种方法进行运动估计，结果如表 1 所示．结果均显示 LBOF 方法对二维运动估计算法的改善是显著和有效的．

表 1 LBOF 方法与 Horn 方法的比较

图像序列		Horn 方法		LBOF 方法	
名称	帧序	收敛时间 (ms)	PSNR(dB)	收敛时间 (ms)	PSNR(dB)
销售员	12	3503	31.544	1973	31.596
	57	3005	32.326	1990	32.442
	110	3362	33.501	1850	33.741
	213	4167	33.478	1916	33.621
特雷弗	77	5042	31.136	2795	31.215
	90	4523	31.603	2164	31.708
	114	5245	31.627	3128	31.668
	130	4821	29.392	2756	29.478

4 结 论

标号场将运动图像区分为“未变化区域”和“变化区域”两部分，将它引入基于光流场的二维运动估计算法，对算法进行改进．首先，对拉普拉斯算子进行修正．结合标号场，使得每一位置像素点的运动估计向有利的方向“漂移”．其次，对基本等式中加权因子 λ 进行修正．根据各个像素点运动估计的误差来源，使得不同的位置具有不同的 λ，从而加速了算法的收敛．通过以上的改进，不仅加快了运动估计算法的收敛速度，而且提高了运动估计结果的准确性，减轻了遮挡边界问题对运动估计结果的影响．

参 考 文 献

[1] 高文，陈熙霖，计算机视觉 — 算法与系统原理，北京，清华大学出版社与广西科学技术出版社，1999，119-208.

[2] 吴立德，计算机视觉，上海，复旦大学出版社，1993，83-100.

[3] B. K. P. Horn, B. G. Schunck, Determining optical flow, Artificial Intelligence, 1981, 17(1-3), 185-203.

[4] D. Terzopoulos, Regularization of inverse visual problems involving discontinuities, IEEE Trans. on PAMI, 1986, PAMI-8(4), 413-424.

[5] H. H. Nagel, On the estimation of optical flow: Relation between different approaches and some new results, Artificial Intelligence, 1987, 33(3), 299-324.

[6] H. H. Nagel, W. Enkelmann, An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences, IEEE Trans. on PAMI, 1986, PAMI-8(5), 565-593.

[7] S. V. Fogel, Estimation of velocity vector fields from time-varying image sequences, CVGIP: Image Understanding, 1991, 53(3), 253-287.

[8] A. Murat Tekalp, Digital Video Processing, Englewood Cliffs, NJ, Prentice Hall PTR, 1998, 200-201.

[9] A. Rosenfeld, C. K. Avinash, Digital Picture Processing, New York, NY, Academic Press, 1976, 263-276.

TWO-DIMENSION MOTION ESTIMATION BASED ON LABELING FIELD AND OPTICAL FLOW FIELD

Yang Yong　Wang Qiao　Wu Lenan

(*Radio Engineering Dept., Southeast University, Nanjing* 210096, China)

Abstract In this paper, an improved algorithm combining optical flow field with labeling field is presented for estimating the two-dimension motion. Based on the basic constraint and the global smooth constraint, a labeling field is introduced, by which Horn-Schunck algorithm is modified. This method not only fastens the convergence of the algorithm, but also improves the accuracy and weakens the effect of the occluding boundary on the motion estimation. The experimental results show that it is a suitable method.

Key words Optical flow field, Labeling field, Two-dimension motion estimation

杨　勇：　男，1973 年生，博士生，主要研究数字图像处理和数字视频处理.

王　桥：　男，1966 年生，教授，主要研究偏微分算子微局部分析、相空间调和分析及其在信号处理和量子信息论中的应用.

吴乐南：　男，1952 年生，教授、博士生导师，从事信号处理理论和多媒体应用技术等研究.