

文章编号: 1000-582X(2011)12-121-05

梯度优化的多尺度视频光流估计

胡学刚^{a,b}, 李文羽^b

(重庆邮电大学 a. 系统科学研究中心; b. 计算机科学与技术学院, 重庆 400065)

摘要:提出一种基于梯度优化的不同运动幅度视频图像光流估计的新算法。先用 Loggabor 滤波器对原视频图像进行相位、尺度滤波,再用所得的特征图像来计算时空梯度,最后根据时空梯度计算光流。该算法模型同时运用由粗到精的图像金字塔方法对视频图像分层处理。理论分析和实验结果表明,该算法适用于大幅度的视频运动光流估计,不仅能得到适合人眼视觉分辨率特性的图像,而且使时空梯度更加优化,光流计算更准确。并且在时间复杂度上与传统光流计算方法相当,在计算精度上优于 Horn 和 Schunck、段先华等人提出的算法。

关键词:光流; Loggabor 滤波; 时空梯度; 金字塔方法

中图法分类号: TP391.41

文献标志码: A

Multi-scale optical flow estimation of the video based on gradient optimization

HU Xue-gang^{a,b}, LI Wen-yu^b

(a. Research Center of System Science; b. College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, P. R. China)

Abstract: A new algorithm based on gradient optimization is proposed for optical flow estimation of video images with different motion ranges. The original video images are transformed by using Loggabor filtering on phases and measures, and then the spatio-temporal gradient is calculated by using the obtained feature images. The optical flow is calculated with the spatio-temporal gradient. The video images are layered and processed with coarse-to-fine image pyramid method. The theoretical analysis and experimental results show that the algorithm is suitable for the video optical flow motion estimation of the significant range. It can not only obtain the video images following the human visual resolution characteristics, but also optimize the spatio-temporal gradient, while the optical flow calculation is more accurate. Besides, the time complexity of this algorithm is equivalent to that of the traditional optical flow method, and the accuracy of the algorithm is superior to the methods suggested by Horn-Schunck, Duan, et al.

Key words: optical flow; Loggabor filtering; spatio-temporal gradient; pyramid method

运动估计是视频系列图像处理的一项重要研究内容^[1-2],计算精确的光流场成为视频图像应用中非常重要的工具^[3-4]。自 Horn 和 Schunck^[5]首先提出

计算系列图像光流的时空微分方法以来,光流运动估计相关理论和技术迅速发展^[6-9]。近年来,段先华等提出基于相位微分技术的方法^[10],王嘉等提出基

收稿日期: 2011-06-20

基金项目: 国家自然科学基金资助项目(11071226); 重庆市教委科研基金资助项目(KJ100505)

作者简介: 胡学刚(1965-),男,重庆邮电大学教授,主要从事偏微分方程与图像处理方向研究,(Tel)15923324881;
(E-mail)huxg62461211@126.com。

于三参数模型的全局光流估计方法^[11], Chamorro-Martínez 等提出基于运动模式识别的局部光流估计方法^[12], 从新的角度对光流运动估计进行研究, 引起众多学者的关注。其中文献[7-10]运用相位滤波来计算光流比较精确、鲁棒, 但他们应用的滤波器中的 Gabor 传递函数不仅是在线性频率尺度上的高斯函数, 不具有适合人眼视觉分辨特性, 而且包含有 DC 分量, 不能构造互成正交的滤波器, 也不能保证视频图像处理的质量。此外, 当相邻两帧中目标存在较大的不同幅度运动的情况下, 计算光流也不准确。

图像处理归根结底是要为人服务的, 那么自然应该使用与人类视觉系统认识特性一致的计算机制。人们已经证明人类视觉系统具有非线性, 而这种非线性具有对数性质, 并且发现当使用滤波器的传递函数为在对数频率尺度上的高斯函数时, 可以对图像进行更有效的编码^[13-14]。此外, 有零 DC 响应滤波器的条件很重要, 否则就不能构造互成正交的滤波器, 就不能在亮度大小跨几个数量级的条件下处理图像, 也就不能保证所处理的视频图像的质量。

在文献[7-10]光流计算的基础上, 结合 Loggabor 滤波和金字塔方法, 提出了一种适应性更强的基于梯度优化的不同运动幅度光流估计新算法。该算法先用 Loggabor 滤波器对原视频图像进行相位、尺度滤波, 再用所得的新特征图像来计算时空梯度, 最后根据时空梯度计算光流; 并且运用由粗到精的图像金字塔方法对视频图像分层处理。理论分析和实验结果表明, 该算法适用于大幅度的视频运动光流估计, 不仅能得到适合人眼视觉分辨率特性的图像, 而且使平均匹配残差减小, 光流计算更准确。同时在时间复杂度上与传统光流计算方法相当, 在计算精度上优于 Horn 和 Schunck、段先华等人提出的算法。

1 计算优化的时空微分梯度

1.1 Loggabor 滤波器的设计

为了构造 Loggabor 滤波器, 将坐标原点设在图像区域中心, ω_0 为滤波器的中心频率, θ_0 为滤波器的初始方向角; ρ 和 θ 分别是极径和极角, σ_ρ 和 σ_θ 分别是半径标准差和角标准差, 在频域上定义 Loggabor 函数滤波器如下

$$\varphi(\rho, \theta) = e^{-\left\{ \frac{[\log(\rho/\omega_0)]^2}{2[\log(\sigma_\rho/\omega_0)]^2} \right\}} e^{-\left[\frac{(\theta-\theta_0)^2}{2\sigma_\theta^2} \right]}, \quad (1)$$

由于对数函数在原点的奇异性, 规定 $\varphi(0, 0) = 0$ 。为了保证滤波器的形状恒定, 对于不同的中心频率 ω_0 , σ_ρ/ω_0 的值必须保持不变。通常情况下, σ_ρ/ω_0 取 0.74 时, 滤波器大致可以获得 1 倍频率的带宽; 取 0.55 可以获得 2 倍频的带宽; 取 0.41 时可以获得 3 倍频的带宽。

上面定义的 Loggabor 函数滤波器有 2 个重要特性: 一是 Loggabor 函数没有 DC 分量; 一是该滤波器的传递函数图像在高端有一个延长的尾巴, 与人类视觉系统在对数频率尺度上对称性的细胞响应一致。

1.2 时空微分梯度的计算

设 $I(x, y, t)$ 表示输入图像; $I_{\varphi(o, s)}(x, y, t)$ 是滤波输出图像; \cdot 表示卷积; o 和 s 分别是滤波器的方向和尺度。用 Loggabor 滤波器对输入相邻帧图像进行 Loggabor 变换, 即

$$I_{\varphi(o, s)}(x, y, t) = I(x, y, t) \cdot \varphi(\rho, \theta), o = 1, 2, \dots, n; s = 1, 2, \dots, m, \quad (2)$$

这里 n 为滤波的初始方向, m 为滤波的尺度。把所有方向和尺度的滤波器滤波输出图像相叠加, 作为 Loggabor 滤波器滤波的方向、尺度特征图像, 即

$$I_\varphi(x, y, t) = \sum_{o=1}^n \sum_{s=1}^m I_{\varphi(o, s)}(x, y, t). \quad (3)$$

一般地, 图像 $I(x, y, t)$ 的梯度可以用 (I_x, I_y, I_t) 来计算, 由于 Loggabor 滤波后得到的新特征图像 $I_\varphi(x, y, t)$ 是在原视频帧图像 $I(x, y, t)$ 上进行对数频率滤波的结果, 因此新的图像 $I_\varphi(x, y, t)$ 具有适合人眼视觉的自然特性, 能在时域、空域和方向上获得最佳的分辨率。于是采用如下的公式来计算特征图像的梯度

$$I_x = \frac{\partial I_\varphi}{\partial x}, I_y = \frac{\partial I_\varphi}{\partial y}, I_t = \frac{\partial I_\varphi}{\partial t}. \quad (4)$$

后面的实验表明, 这样的梯度计算光流场更准确。

取特征图像的相邻两帧, 如果 $I_{i,j,k}$ 表示特征图像在点 (x, y, t) 的灰度值 $I(x, y, t)$, $I_{i,j,k+1}$ 表示特征图像在 $(x, y, t+dt)$ 点的灰度值 $I(x, y, t+dt)$, 那么特征图像的时空微分梯度的数值计算可根据式(4)用如下差分公式计算

$$I_x = (I_{i+1,j,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i,j+1,k} + I_{i+1,j,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i,j+1,k+1})/4; \quad (5)$$

$$I_y = (I_{i,j+1,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i+1,j,k} + I_{i,j+1,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i+1,j,k+1})/4; \quad (6)$$

$$I_t = (I_{i,j,k+1} - I_{i,j,k} + I_{i+1,j,k+1} - I_{i+1,j,k} + I_{i+1,j,k+1} - I_{i+1,j,k} + I_{i+1,j+1,k+1} - I_{i+1,j+1,k})/4. \quad (7)$$

2 金字塔光流运动估计

光流参数的计算基于如下 2 个假设:1)沿运动方向图像灰度恒定不变;2)图像运动场基本平滑变化。考虑系列图像中的连续两帧图像,设 t 时刻第一帧图像上 (x, y) 点经过 Δt 时刻运动到第二帧图像 $(x + \Delta x, y + \Delta y)$ 点,灰度值相应地由 $I(x, y, t)$ 变为 $I(x + \Delta x, y + \Delta y, t + \Delta t)$ 。显然当 $\Delta x \rightarrow 0, \Delta y \rightarrow 0, \Delta t \rightarrow 0$ 时可得到较高的计算精度。然而,在实际应用中,通常相邻两帧图像具有不同幅度的运动,存在运动速度大于 1 个像素/帧,因此采用传统的光流法不能得到精确的光流场。

基于上述原因,并考虑到图像噪声的影响,采用一种基于 Loggabor 滤波器和图像金字塔的梯度光流计算方法。该方法的基本思想是构造 Loggabor 滤波的特征图像序列的一个金字塔^[15],如图 1 所示。金字塔表示中,较高的层是下层平滑后的采样形式,原始图像层数等于零。当图像分解到一定的层后,相邻帧间图像运动量将变得足够小,满足光流计算的约束条件,可以进行光流估计。在实际计算时,由高层到低层进行,当某一级的光流增量计算出来后,将加到其初始值上,再进行投影重建,作为其下一层的光流计算初值。这一过程不断进行,直至估计出原始图像的光流。

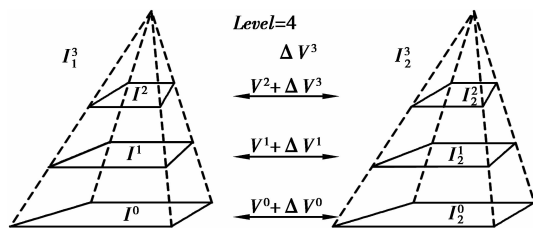


图 1 4 层图像序列金字塔框架

设 I_i^0 表示大小为 $w \times h$ 的原始相邻图像的灰度 ($i=1, 2$),它位于金字塔的第 0 层。依次生成 N 层高斯金字塔的第 $1, 2, \dots, N-1$ 层,分别表示为 $I_1^1, I_2^1, \dots, I_1^{N-1}$ 。对于其中的第 L ($0 \leq L \leq N-1$) 层图像,其灰度采样计算公式为

$$I_i^L(x, y) = \frac{1}{4} I_i^L(2x, 2y) + \frac{1}{8} [I_i^L(2x + 1, 2y) + I_i^L(2x - 1, 2y) + I_i^L(2x, 2y + 1) + I_i^L(2x, 2y - 1)] + \frac{1}{16} [I_i^L(2x + 1, 2y + 1) + I_i^L(2x + 1, 2y - 1) + I_i^L(2x - 1, 2y + 1) + I_i^L(2x - 1, 2y - 1)], \quad (8)$$

且 (x, y) 的定义域如下: $0 \leq x \leq \frac{1}{2^L} w - 1, 0 \leq y \leq \frac{1}{2^L} h - 1$ 。

依据由粗至精的计算方式,较高层的误差将不断的向较低层传播和扩大。因此,在进行基于分层结构的光流计算时,分解层数不能太多。一般地, $3 \leq N \leq 5$ 。设图像金字塔框架中第 L 层相邻两帧图像满足光流计算基本条件,可建立如下极小化方程

$$\epsilon(u^L, v^L) = \iint \{ (I_x^L u^L + I_y^L v^L + I_t^L)^2 + \lambda^2 [(u^L - \bar{u}^L)^2 + (v^L - \bar{v}^L)^2] \} dx dy, \quad (9)$$

式中 \bar{u}^L, \bar{v}^L 分别为速度分量 u^L 邻域和 v^L 邻域中的平均值; λ 为 Lagrange 乘子。实际上,式(9)是泛函的极值问题,依据变分原理,得到如下 2 个 Euler 方程

$$[(I_x^L)^2 + \lambda^2] u^L + I_x^L I_y^L v^L = \lambda^2 \bar{u}^L - I_x^L I_t^L, \quad (10)$$

$$[(I_y^L)^2 + \lambda^2] v^L + I_x^L I_y^L u^L = \lambda^2 \bar{v}^L - I_y^L I_t^L. \quad (11)$$

根据式(10)、(11)可计算出 u^L, v^L ;可用如下松弛迭代方程求解,这里 i 是迭代次数

$$\begin{cases} u_{(i)}^L = 0; \\ u_{(i+1)}^L = u_{(i)}^L - \frac{I_x^L (I_x^L u_{(i)}^L + I_y^L v_{(i)}^L + I_t^L)}{\lambda^2 + (I_x^L)^2 + (I_y^L)^2}, \end{cases} \quad (12)$$

$$\begin{cases} v_{(i)}^L = 0; \\ v_{(i+1)}^L = v_{(i)}^L - \frac{I_y^L (I_x^L u_{(i)}^L + I_y^L v_{(i)}^L + I_t^L)}{\lambda^2 + (I_x^L)^2 + (I_y^L)^2}. \end{cases} \quad (13)$$

把式(5)~(7)代入式(12)、(13)求解第 L 层光流场;迭代结束的条件为满足预先给定迭代次数或估计误差。

令 $V^L = V^L(u^L, v^L)$ 为 L 层初始光流猜测值,由 $N-1$ 层到 $L+1$ 层递推计算得到。最顶层初始光流猜测值 $V^{N-1} = 0, \Delta V^L$ 为第 L 层计算得到的光流结果,上下层间光流传递过程为

$$V^{N-1} = 0, V^{L-1} = 2(V^L + \Delta V^L). \quad (14)$$

上述递推过程自顶向下进行,直到第 0 层计算完毕并得到最终的光流计算结果

$$V = V^0 + \Delta V^0 = \sum_{L=0}^{N-1} 2^L \Delta V^L. \quad (15)$$

3 实验结果及分析

为了说明算法的效果,这里基于 MATLAB7.1 平台下,选用两组图像进行计算光流场的实验。实验对象分别选取背景和树木都发生不同幅度运动的 landscape 图像和向右平移和旋转的 hand 图像(如图 2 和图 3),并将研究方法 with 文献[5]和[10]的算法进行比较,从客观指标和主观视觉效果 2 个方面进行评价。

运动估计评价的客观指标通常选用平均匹配残差,该指标能客观的反映光流场计算的准确性^[16]。平均匹配残差可定义为

$$\text{mean} = \frac{1}{w \times h} \sum_{i=1}^w \sum_{j=1}^h |I_1(i, j) - I_2(i, j)| \quad (16)$$

其中 $I_1(i, j)$ 、 $I_2(i, j)$ 是相邻两帧图像, w, h 分别是图像的长和宽。为了估计算法的性能,所以选用平均匹配残差 mean 作为一个测度。

取 12 个方向, 6 个尺度的 Loggabor 滤波器, 最小波长等于 8, $\sigma_\theta = 1.4$, $\sigma_o/\omega_o = 0.57$, 通过对原始系列图像进行滤波得到特征图像; 然后取 4 层金字塔, $\lambda = 200$, 每层中迭代 3 次来计算光流场。根据式 (16), 2 组试验的匹配残差如表 1 所示。

表 1 实验结果的 mean 对比表

图像	Horn 光流法	段光流法	研究算法
landscape	25.2267	18.5763	12.6244
hand	5.5989	2.6710	2.4368

由表 1 可知, 算法得到的匹配残差比文献[5]中 Horn&schunck 算法和文献[10]中段先华等人的算法平均残差明显减少, 说明研究算法计算光流精度更高。原因主要有两点, 一是用对数滤波器对图像进行滤波后获得的特征图像具有较高的适合视觉特性的空间分辨率, 空间梯度比较好, 再用时空梯度计算光流较精确。二是运用图像金字塔框架分层计算光流, 把图像大幅度的运动通过采样变为符合光流假设条件的平滑运动, 光流计算越精确。

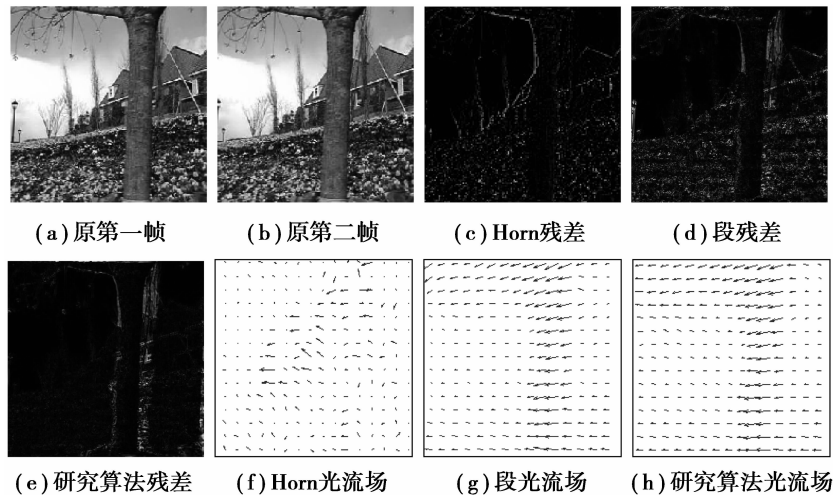


图 2 Landscape 系列图像的运动估计光流场及匹配残差图

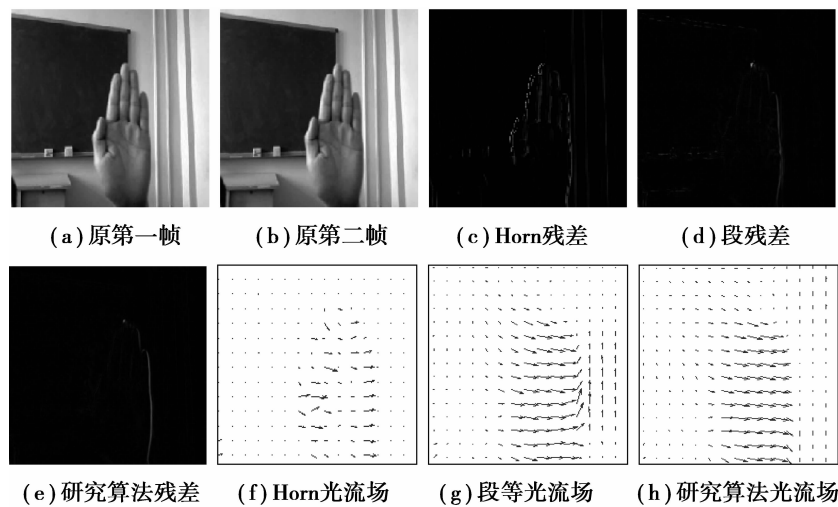


图 3 hand 图像的运动估计光流场及匹配残差图

观察图 2、图 3 中的(c)、(d)、(e)残差图,亮度区域越大且亮度越亮,残差越大。对比可以看出,算法残差较小,且较大残差的分布更为随机,说明算法计算的光流矢量更准确,使图像恢复的精确性得以提高。

由图 2 中风景合成相邻图像(a)、(b)知树离摄像机镜头较近,背景较远,当摄像机角度向右转动时,树向左位移较大,背景向左位移较小。从(f)、(g)、(h)3种方法的光流场可以看出,(f)描述不同运动幅度运动有较大误差,(g)有一定的改善,而研究方法的的光流场能较好的描绘出树的运动速度较大,背景运动速度较小,且光流较准确;从图 3 中挥手实际系列图像相邻两帧(a)、(b)可以看出,手是向右移动且向右下有一定转动。比较(f)、(g)、(h)光流场可知,算法的光流场能更形象的描绘出手的挥动,且光流更精确。

4 结 论

给出了一种新的基于梯度优化的不同幅度运动的视频图像光流估计算法。用 Loggabor 对数滤波器对原始系列图像进行滤波,得到符合人眼对数视觉特性和较高分辨率的特征图像,再用新特征图像计算梯度;对有不同幅度运动或有较大运动不连续的系列图像,用金字塔分层的方法使较大运动在低分辨率图像上变得平滑,从而准确计算出光流场。实验表明该算法与 Horn 和 Schunck 光流法、段先华等光流法相比,平均匹配残差明显减小,光流运动估计参数计算也更精确。随着计算机硬件性能的提高,使用该光流估计算法在线实时处理图像是下一步的研究工作。

参考文献:

- [1] 黄志鹏,唐治德. 基于 H. 264 标准的运动估计参数性能分析[J]. 重庆大学学报:自然科学版,2006, 29(10): 33-37.
HUANG ZHI-PENG, TANG ZHI-DE. Performance analysis of motion estimation parameters based on H. 264 standards[J]. Journal of Chongqing University: Natural Science Edition, 2006, 29(10): 33-37.
- [2] ISO/IEC JTC1/SC29/WG11 Moving picture experts group. MPEG2 test model [S]. Beijing: MPEG China, 1993.
- [3] 章毓晋. 图像工程(第二版)[M]. 北京:清华大学出版社, 2005.
- [4] 李文羽,胡学刚. 一种基于两步运动估计的系列图像匹配算法[J]. 计算机应用研究, 2009, 26(9): 3563-3565.
LI WEN-YU, HU XUE-GANG. A image matching algorithm based on two-step motion estimation [J]. Application Research of Computer, 2009, 26(9): 3563-3565.
- [5] HORN B K P, SCHUNCK B G. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1-3): 185-203.
- [6] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision[C]// Proceedings of the 7th International Joint Conference on Artificial Intelligence, August 24-28, 1981, Vancouver, BC, Canada. Vancouver: Morgan Kaufmann Publishers, 1981, 2:674-679.
- [7] FLEET D J, JEPSON A D. Computation of component image velocity from local phase information[J]. International Journal of Computer Vision, 1990, 5(1): 77-104.
- [8] GAUTAMA T, VAN HULLE M A. A phase-based approach to the estimation of the optical flow field using spatial filtering [J]. IEEE Transactions on Neural Networks, 2002, 13(5): 1127-1136.
- [9] KIM Y H, MARTINEZ A M, KAK A C. Robust motion estimation under varying illumination[J]. Image and Vision Computing, 2005, (23): 365-375.
- [10] 段先华,王平安,夏德深. 基于相位微分技术的光流估计[J]. 数据采集与处理, 2005, 20(3): 249-253.
DUAN XIAN-HUA, WANG PING-AN, XIA DE-SHEN. Phase differential technique based on estimation of optical flow [J]. Journal of Data Acquisition & Processing, 2005, 20(3): 249-253.
- [11] 王嘉,王海峰,刘青山,等. 基于三参数模型快速全局运动估计[J]. 计算机学报, 2006, 29(6): 921-927.
WANG JIA, WANG HAI-FENG, LIU QING-SHAN, et al. Fast global motion estimation based on 3-parameter global motion model[J]. Chinese Journal of Computers, 2006, 29(6): 921-927.
- [12] CHAMORRO M J, FERNANDEZ V J. A new approach to motion pattern recognition and its application to optical flow estimation [J]. IEEE Transactions on Systems, Man, and Cybernetics-part C: Applications and Reviews, 2007, 37(1): 39-51.
- [13] FIELD D J. Relations between the statistics of natural images and the response properties of cortical cells[J]. Journal of the Optical Society of America, 1987, 4(12): 2379-2394.
- [14] 肖志涛,国澄明,于明,等. Log Gabor 函数在人类视觉系统特性研究中的应用[J]. 信号处理, 2002, 18(5): 399-402.
XIAO ZHI-TAO, GUO CHENG-MING, YU MING, et al. A new aligned with human visual system characteristics wavelet and its application [J]. Signal Processing, 2002, 18(5): 399-402.
- [15] 江志军,易华容. 一种基于图像金字塔光流的特征跟踪方法[J]. 武汉大学学报:信息科学版, 2007, 32(8): 680-683.
JIANG ZHI-JUN, YI HUA-RONG. An image pyramid-based feature detection and tracking algorithm [J]. Geomatics and Information Science of Wuhan University, 2007, 32(8): 680-683.
- [16] 郭丰俊,杨新,PHILIPS W,等. 基于 HUBER 函数的序列图像位移场估计算法[J]. 高技术通讯, 2001, 11(1): 38-42.
GUO FENG-JUN, YANG XIN, PHILIPS W, et al. A displacement field estimation algorithm based on Huber function[J]. High Technology Letters, 2001, 11(1): 38-42.

(编辑 侯 湘)