

提高光流估计性能的渐进性高斯 多维预滤波方法的研究

付 昀, 徐维朴

(西安交通大学人工智能与机器人研究所, 710049 西安)

[摘要] 基于光流计算方法统一框架理论, 研究了一种利用高斯多维滤波器的渐进性和时空性提高光流估计性能的有效方法。在保持现有光流计算方法的前提下, 通过调节时间维和空间维的方差参数, 改变时空预滤波和光滑效果, 突出时间混叠和光流主信息, 从而提高重构视频序列的信噪比。试验中以标准的 Flower Garden 和 Football 序列的前 50 帧作为参考图像序列, 以 LK 算法为参考光流算法。结果显示, 滤波窗口为 5×5 时的最佳时间方差参数为 0.4, 最佳空间方差参数为 [1.6, 2.0]; 加入高斯多维预滤波前后利用光流场重构图像的平均峰值信噪比 PSNR 提高 2.572 dB, 提高幅度为 13.6%。

[关键词] 光流计算; 高斯多维滤波; 峰值信噪比; 运动估计

[中图分类号] TP391, TN713+.1 **[文献标识码]** A **[文章编号]** 1009-1742(2004)12-0056-06

1 前言

光流 (optical flow) 是图像亮度的模式运动, 即灰度像素点在图像平面上运动而产生的瞬时速度场。基于光流方程 (OFE) 的光流计算, 是一种二维运动的非参数场模型。其他的用于二维运动估计的非参数模型包括块运动匹配、像素递归、频域运动估计和贝叶斯法^[1]。与这些非参数模型相比较, 光流计算更侧重于时空的平滑性约束。从 20 世纪 50 年代提出至今, 已经出现了诸多成熟有效的光流计算方法, 从概念上可以将它们分为微分法、区域匹配法、基于能量的方法、基于相位的方法和神经动力学方法等。各种光流技术通常都可以归结为 3 个处理阶段^[2]: a. 利用低通或带通滤波器进行预滤波和光滑处理, 从而提取感兴趣的信号信息, 并尽可能的提高信噪比; b. 确定可以用来度量光流场的基本约束条件和定量关系; c. 结合上述的度量信息产生 2D 光流场并进行光滑处理。

光流计算的一个发展趋势就是统一框架理论的完善建立。该理论依据是: 已有的各种光流计算方法基本上可以统一在一个框架下, 并将光流信息分成保持信息和邻域信息 2 部分, 再通过 2 种信息的提取、分析和融合恢复光流场^[3]。所以从提高光流计算精确度的角度考虑, 一个研究方向是在不影响该统一框架的前提下, 利用其他的图像视频技术和信号处理方法进行光流准确性的提升研究。

高斯函数对于图像进行的滤波处理, 可以有效地表达时域和频域的图像特征, 包括降噪处理, 边缘检测, 图像增强和多分辨率分析等。特别对于时间域的处理可以提取主要信号信息, 提高信噪比, 突出时间域混叠现象和量化效果。为了在现有的光流求解算法的基础上进一步提升光流场的准确性, Barron 在文献 [2] 中提到了利用一种高斯 (Gauss) 时空多维滤波技术对图像序列进行预光滑处理, 并且在多种视频序列中进行应用, 但是未对这种技术给出深入的理论方法和实验研究报告。笔

[收稿日期] 2003-11-05; **修回日期** 2003-01-17

[基金项目] 国家自然科学基金资助项目 (60205001); 国家自然科学基金创新研究群体科学基金资助项目 (60024301)

[作者简介] 付 昀 (1979-), 男, 西安市人, 西安交通大学硕士研究生; 主要研究方向为计算机视觉与图像处理、模式识别与智能系统、人像处理与人脸识别等

者在此基础上通过严格的实验方法, 利用预处理环节对序列图像的时间维和空间维进行耦合的预滤波处理, 再结合通常的光流算法有效地提升序列图像的运动估计效果, 侧重于研究高斯多维时空预滤波对于光流估计的提升能力以及提升幅度。

2 光流算法

光流技术从概念上可以分为多种方法, 包括微分法、区域匹配法、基于能量的方法、基于相位的方法和神经动力学方法等。无论是真实的还是合成的图像序列都要求尽可能地避免时空混叠现象的产生。基于微分度量的 LK 算法的精确度, 稳定性和可靠性都很高, 因此选择该算法作为参考光流算法。

LK 算法^[4]是由 Lunas 和 Kanade 提出的一种

$$\begin{cases} \sum_{x,y \in B} \left(\frac{\partial E(x,y)}{\partial x} \hat{v}_x + \frac{\partial E(x,y)}{\partial y} \hat{v}_y + \frac{\partial E(x,y)}{\partial t} \right) \cdot \frac{\partial E(x,y)}{\partial x} = 0 \\ \sum_{x,y \in B} \left(\frac{\partial E(x,y)}{\partial x} \hat{v}_x + \frac{\partial E(x,y)}{\partial y} \hat{v}_y + \frac{\partial E(x,y)}{\partial t} \right) \cdot \frac{\partial E(x,y)}{\partial y} = 0 \end{cases} \quad (3)$$

其中 \hat{v}_x , \hat{v}_y 是块 B 中心像素点的运动矢量估计值。方程组式 (3) 的求解为

$$\begin{bmatrix} \hat{v}_x \\ \hat{v}_y \end{bmatrix} = \begin{bmatrix} \sum_B E_x^2 & \sum_B E_x E_y \\ \sum_B E_x E_y & \sum_B E_y^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_B E_x E_t \\ -\sum_B E_y E_t \end{bmatrix} \quad (4)$$

其中 $E_x = \partial E(x, y) / \partial x$, 则定义 $\nabla E(x, y) = [E_x, E_y]^T$ 。

在寻找模式运动速度的光流场方面, LK 算法是一种有效的方法, 它的原理和实现比较简单, 而且在找寻光流的准确性方面也有一定的精确度。但是, 其本身也有一定的适应范围, 对于相对运动较小的 2 幅图像来说, 能够求得比较好的效果, 但是对于 2 幅图像相对运动比较大, 或者说某一部分有较大差异的 2 幅图像来说, 求取的光流就不够准确。利用高斯时空多维滤波技术, 可以有效地提升相对运动较大的图像序列的光流场计算精度。

3 高斯时空多维滤波模型

3.1 高斯多维滤波核

高斯函数对图像进行的滤波处理可以有效地表达时域和频域的图像特征, 包括降噪处理、边缘检测、图像增强和多分辨率分析等。高斯函数的基本表达式为

光流估计算法。该算法假设在给定时刻的运动场中, 任何图像像素点的运动矢量都限制于某个唯一的像素运动区域块 B 中, 该区域是均匀的, 并可以看作一个运动点, 其整体光流运动可以由中心像素点的运动替代处理:

$$v(x, y) = [v_x, v_y]^T = v, x, y \in B \quad (1)$$

在满足一般性的光流约束方程 (OFE) 的情况下, 像素块 B 内的误差函数定义为

$$\epsilon = \sum_{x,y \in B} \left(\frac{\partial E(x,y)}{\partial x} v_x + \frac{\partial E(x,y)}{\partial y} v_y + \frac{\partial E(x,y)}{\partial t} \right)^2 \quad (2)$$

块中心像素的运动估计矢量就是使得上述误差函数达到最小的 $v(x, y)$ 值, 即满足 ϵ 对 x 方向和 y 方向运动矢量的一阶偏导数为 0, 由此得到方程组:

$$G(t) = A \exp[-(t - \mu)^2 / 2\sigma^2] \quad (5)$$

其中 μ, σ 是常数, 分别代表统计学中的均值和方差, A 一般是常系数。作为高斯滤波核函数考虑时, 通常令 $A = 1, \mu = 0$, 则式 (5) 表示为

$$G_t = G(t) = A \exp[-t^2 / 2\sigma^2] \quad (6)$$

高斯滤波核做卷积结果仍然是高斯核, 对于一般的具有各向同性的低通滤波核的迭代光滑, 都可以归结为高斯光滑, 高斯多维滤波核为

$$G_{t_1} * G_{t_2} * \dots * G_{t_n} = G_{t_1+t_2+\dots+t_n} \quad (7)$$

高斯多维滤波核中的不同维的滤波性能, 可以通过调节相应的方差参数 σ 进行控制。

3.2 高斯时空三维滤波核

为了利用多维高斯滤波核的滤波特性对序列图像 (包括视频序列或者人工合成图像) 进行处理, 提取主要的信号信息, 提高信噪比, 突出时间域混叠现象和量化效果, 可以同时考虑空间滤波和时间滤波^[2]。时空滤波核的函数表达式为

$$G^{\sigma, \tau}, \{x, y, t\} = C \exp[-((x^2 + y^2) / 2\sigma^2 + t^2 / 2\tau^2)] \quad (8)$$

其中 σ, τ 是常数, 分别表示空间域 (像素/帧) 和时间域 (像素/帧) 的方差, x, y 表示空间域中的二维坐标, t 表示时间域的一维坐标, C 是常系数。利用该时空滤波核可以对序列图像进行滤波和

平滑处理。图1表示了一个利用高斯时空滤波器进行图像处理的模型,其中, $E(x, y, t)$ 表示输入的图像序列,经过高斯时空滤波处理后,得到处理后图像序列 $Y(x, y, t)$ 为

$$Y(x, y, t) = E(x, y, t) * G^{\sigma, \tau}(x, y, t) \quad (9)$$

$$E(x, y, t) \rightarrow \boxed{G^{\sigma, \tau}(x, y, t)} \rightarrow Y(x, y, t)$$

图1 高斯时空滤波模型

Fig.1 Gaussian spatiotemporal filter model

对于数字图像,需要考虑离散卷积,从滤波效果和运算量角度考虑如何选择合适的滤波算子尺度。算子尺度太大,运算速度下降很大,无法满足实时性要求;算子尺度太小,尽管提高了运算速度,但是无法体现出高斯滤波的效果。式(9)的离散卷积表达式为

$$Y(x, y, t) = \sum_{h=0}^{T-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E(i, j, h) \cdot G_{\sigma, \tau}(x-i, y-j, t-h) \quad (10)$$

其中空间算子尺度大小为 $M \times N$, T 为图像序列的数目,即时间算子尺度。对算子进行归一化处理 $\sum \text{Mask}\{x, y, t\} = 1$, 则得到

$$Y(x, y, t) = \frac{1}{A(x, y, t)} E(x, y, t) * G^{\sigma, \tau}(x, y, t) \quad (11)$$

其中 $A(x, y, t) = \sum_{h=0}^{T-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} G_{\sigma, \tau}(x-i, y-j, t-h)$ 。

图2描述了取正方形滤波算子时,对于连续2帧 ($T=2$),大小为 $m \times n$ 的图像序列进行时空滤波的示意图。

3.3 高斯时空三维滤波的渐进性

光流算法适用于同一对象运动的序列图像或视

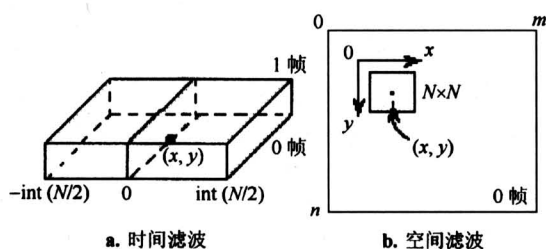


图2 正方形高斯时空多维滤波算子

Fig.2 Square Gaussian multidimensional filter operator

频图像,以及运动幅度不能太大的场景。对于对象运动变化较大的场合,常规光流算法无法有效计算模式运动场。利用高斯时空滤波对2幅相邻图像进行渐进处理,可以从时间域和空间域有效地缩小对象大幅运动的差异,使得光流算法得到一定效果的应用。这里的渐进性滤波是由式(10)中的参数 M, N, σ, τ 的调节实现的。在不同参数的条件下,特定图像与参考图像的渐进逼近效果见图3。通过分别调节时间和空间的参数都可以产生差异很小的渐变图像序列。由2幅图像生成的图像序列,可以在相邻图像间采用现有的光流算法进行运动估计。其中特定图像和参考图像是2幅不同性别的人脸图像。首先对它们进行了基于域 Warp 技术的形状对准处理^[5]。对这2幅不同的人脸图像,直接利用常规光流算法计算特定图像和参考图像间的光流场是失效的。



(2幅图像引自德国 Max-Planck 研究所人脸图像库)

图3 高斯时空滤波的渐进效果

Fig.3 The gradual effect of the Gaussian spatiotemporal filter (the two images are from MPI)

高斯时空滤波器的这种渐进性支持了其从时间域和空间域对研究对象运动场的捕捉与控制能力。这种从图像的角度对模式运动场的改变可以提高光流计算效果。

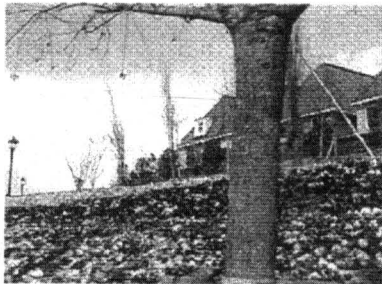
4 实验与分析

实验中,首先对标准 LK 光流算法进行研究,然后加入高斯时空预滤波后对相同视频序列进行处理,最后计算出客观评价的峰值信噪比,从而比较2种算法的光流计算效果。运动估计中的客观评价标准很多,如峰值信号噪声比 (PSNR)、熵 (entropy) 等。采用文献 [6] 中定义的 PSNR (略有改进) 对图像运动场计算结果进行客观比较,如

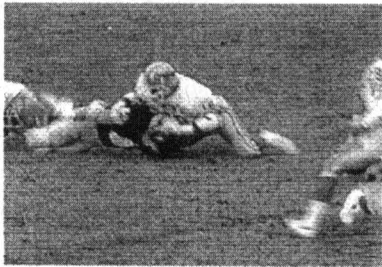
$$\text{PSNR} = 10 \lg \left[\frac{255^2(M-4)(N-4)}{\sum_x [F(x, n) - F_0(x, n)]^2} \right] \quad (12)$$

其中, $M \times N$ 表示帧图像的大小, $F(x, n)$ 表示生成帧, $F_0(x, n)$ 表示参考帧。采用 $(M-4) \times (N-4)$ 是考虑到高斯滤波窗无法对图像边界进行处理, 从而无法比较高斯时空滤波的预滤波后光流提升效果, 相应的生成帧与参考帧的像素距离也只计算到图像边界 2 个像素以内的所有灰度值。

在不同区域块大小的情况下, 利用 LK 算法分别对 Flower Garden 和 Football 序列 (如图 4 所示的两序列的第一帧图像) 的前 50 帧 352×40 图像进行相邻帧的光流场计算, 然后利用相邻帧的前帧图像与求解出的光流信息重构出后帧图像, 再求解出重构图像与实际图像的峰值信号噪声比, 最后对 50 次计算结果进行平均, 获得 PSNR 的平均值, 如表 1 所示。结合主观评价, 笔者认为, 分块大小为 5×5 时对于所研究的 2 图像序列的光流计算结果较好, 其中 Flower Garden 序列中的平均峰值信噪比 PSNR 为 18.870 dB, Football 序列中的平均峰值信噪比为 19.211 dB。



a. Flower Garden 第 1 帧



b. Football 第 1 帧

图 4 标准 Flower Garden 和 Football 视频序列的第一帧图像

Fig.4 The first frame of the standard video sequences: Flower Garden and Football

同样, 以前 50 帧 Flower Garden 视频序列为参

表 1 不同分块大小下 Flower Garden 和 Football 序列的平均 PSNR (dB) 结果

Table 1 Average PSNR (dB) of the two sequences under different block size

分块大小	3×3	5×5	7×7	9×9	11×11	13×13
PSNR /dB						
Flower Garden	18.845	18.870	18.726	18.545	18.384	18.434
Football	19.333	19.221	18.892	18.630	18.410	18.243

考样本, 首先对相邻 2 帧图像进行一定时空维方差参数 σ, τ 下的高斯时空预滤波处理, 然后对滤波后的 2 帧图像利用 LK 算法进行光流场的计算, 利用相邻帧的前帧图像与求解出的光流场重构后帧图像, 求解出重构图像与实际图像的峰值信号噪声比, 最后对 50 次计算结果进行平均, 获得 PSNR 的平均值。实验中, 高斯时空预滤波窗口大小为 5×5 , 参数 σ, τ 分别从 0.1 开始以步长 0.1 递增, LK 算法中的分块大小为 5×5 。实验结果见表 2, 显示了部分不同参数 σ, τ 下的平均 PSNR 结果比较。表 2 中未列出的其他不同参数 τ 情况下的平均 PSNR 结果均低于未加入高斯时空预滤波时利用 LK 算法计算的峰值信噪比 18.870 dB。图 5 描绘了空间方差 σ 从 0.1 变化到 1.0、时间方差 τ 从 0.1 变化到 0.5 的 PSNR 曲线图 (称为有效曲线), 数据取自表 2。由于 $\tau = 0.6$ 时 PSNR 均低于未加入高斯时空预滤波时利用 LK 算法计算的 PSNR 值, 所以未画出。从图 5 可以看出当 $\tau = 0.4$ 时 PSNR 曲线增长最快, 而且当 $\sigma = 0.8$ 时, 该曲线已经超过其他有效曲线。表 2 也列出了时间方差 σ 在 0.1~0.6、空间方差 σ 从 1.0 变化到 2.0 的 PSNR 结果, 当 $\tau = 0.4$ 时 PSNR 增长仍然最快 (见图 6)。从图 6 可以看出, 当空间方差 $\sigma > 1.5$ 时曲线的增长速度减慢并且趋于稳定, 增长幅度小于 0.05 dB。为了与未加入高斯时空预滤波时 LK 算法的计算结果进行比较, 求出表 2 中 $\sigma = [1.6, 2.0]$ 区间的 PSNR 平均值为 21.442 dB, 与未加入高斯时空预滤波时利用 LK 算法计算的峰值信噪比 18.870 dB 相比, 提升 2.572 dB, 提升幅度为 13.6%。图 7 描绘了加入高斯时空预滤波前后的利用 LK 算法计算的 Flower Garden 序列前两帧图像序列的光流场, 比较图 7a 和图 7b 可以看出主要运动信息的增强效果以及时间混叠效果。

表2 不同时空方差参数下 Flower Garden 序列的平均 PSNR (dB) 结果

Table 2 Average PSNR (dB) of Flower Garden under different σ and τ (1)

σ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\tau=0.1$	18.840	18.870	18.879	19.184	19.773	20.248	20.524	20.651	20.712	20.759
$\tau=0.2$	18.810	18.844	18.878	19.184	19.773	20.248	20.524	20.651	20.712	20.759
$\tau=0.3$	18.809	18.810	18.855	19.167	19.763	20.253	20.542	20.680	20.754	20.812
$\tau=0.4$	18.595	18.596	18.626	18.921	19.526	20.094	20.505	20.765	20.939	21.064
$\tau=0.5$	18.031	18.032	18.052	18.253	18.667	19.127	19.510	19.784	19.982	20.106
$\tau=0.6$	17.383	17.384	17.385	17.486	17.669	17.833	17.955	18.038	18.107	18.142
σ	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
$\tau=0.4$	21.161	21.237	21.296	21.346	21.382	21.402	21.431	21.441	21.463	21.472
$\tau=0.1$	20.800	20.834	20.858	20.891	20.912	20.934	20.944	20.948	20.960	20.964
$\tau=0.2$	20.800	20.834	20.858	20.891	20.912	20.934	20.943	20.948	20.960	20.964
$\tau=0.3$	20.854	20.902	20.921	20.954	20.977	21.007	21.015	21.026	21.042	21.046
$\tau=0.5$	20.209	20.278	20.322	20.362	20.380	20.389	20.407	20.406	20.410	20.409
$\tau=0.6$	18.166	18.174	18.174	18.168	18.158	18.151	18.131	18.124	18.118	18.103

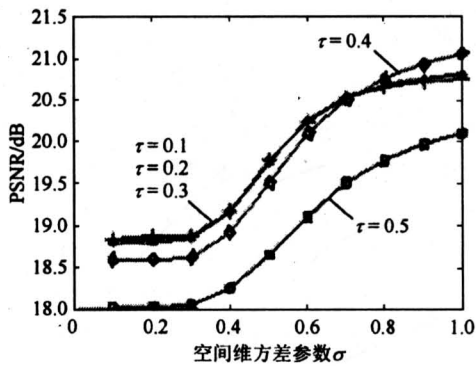


图5 不同时空参数下的 PSNR 曲线

Fig.5 PSNR (dB) versus σ and τ

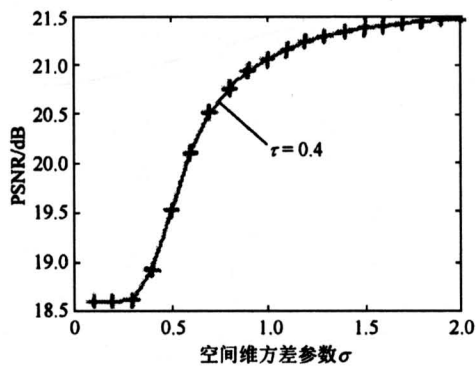


图6 时间方差为 0.4 时的 PSNR 曲线

Fig.6 PSNR (dB) versus σ when $\tau=0.4$



a. LK 算法计算出的前 2 帧的光流场



b. 加入预滤波后的光流场 ($\tau=0.4, \sigma=2.0$)

图7 加入高斯时空预滤波前后利用 LK 算法计算的 Flower Garden 序列前 2 帧的光流场

Fig.7 The optical flow field of the first two frames in Flower Garden influenced by Gaussian spatiotemporal filter

5 结论

研究了渐进性高斯多维预滤波对光流估计的提升性能, 实验分析表明在不改变现有光流计算方法的前提下, 通过调节时间维和空间维的方差参数, 改变时空预滤波和光滑效果, 突出主信息和时间混叠, 可以较大幅度提高峰值信噪比 PSNR。以标准的 Flower Garden 为测试序列, 以 LK 算法为参考光流算法。结果显示, 滤波窗口为 5×5 时的最佳时间方差参数为 0.4, 最佳空间方差参数为 [1.6, 2.0]。进一步的工作包括对更多视频序列进行实验验证, 总结出渐进性高斯多维预滤波对于不同视频条件下的光流估计的提升能力与针对性。

致谢: 感谢德国 Max-Planck 研究所 (MPI, Tübingen, Germany) 提供的高加索人脸图像库; 感谢张强, 张婷, 卓峰等智能人像研究小组成员给予的帮助。

参考文献

[1] 泰卡尔普 A.M. 数字视频处理 (Digital Video

Processing) [M]. 崔之祜, 江春, 陈丽鑫译. 北京: 电子工业出版社, 1998

- [2] Barron J L, Fleet D J, Beauchemin S S. Performance of optical flow techniques [J]. International Journal of Computer Vision, 1994, 12(1): 43~77
- [3] Singh A, Allen P. Image-flow computation: an estimation-theoretic framework and a unified perspective [J]. CVGIP: Image Understanding, 1992, 56: 152~177
- [4] Lucas B D, Kanade T. An iterative image registration technique with application to stereo vision [A], Proceedings of the 7th International Joint Conference on Artificial Intelligence[C]. Vancouver, 1981. 674~679
- [5] 付 昀. 线性人脸对象类模型的匹配提升技术[R]. 西安: 西安交通大学人工智能与机器人研究所, 2003
- [6] Nam J Y, Seo J S, Kwak J S, et al, New fast-search algorithm for block matching motion estimation using temporal and spatial correlation of motion vector [J]. IEEE Transaction on Consumer Electronics, 2000, 46(4): 934~942

The Improving Characteristics of the Gradual Gaussian Multidimensional Pre-filter for Optical Flow Estimation

Fu Yun, Xu Weipu

(Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China)

[Abstract] Based on the unified estimation-theoretic framework, an effective method of using the gradual Gaussian multidimensional pre-filter to improve the optical flow estimation is presented. The pre-filtering and smoothing effect, which attenuate the temporal aliasing and the interesting signal structure of the optical flow field, are altered with adjusting the spatiotemporal standard deviation parameters. The first 50 frames of the standard Flower Garden and Football video sequence are tested as the reference image sequences, and the LK algorithm as the reference optical flow computing method. Experimental results in objective evaluation show that the optimum temporal standard deviation parameter is 0.4, the optimum spatial standard deviation parameter is in a range of 1.6~2.0 under the condition that the pre-filtering window size is 5×5 pixels. After pre-filtering the image sequence by the Gaussian multidimensional filter, the average PSNR of the reconstructed frames enhance 2.572 dB, higher than that using the standard optical flow computing method by nearly 13.6 %.

[Key words] optical flow computing; Gaussian multidimensional filter; PSNR; motion estimation