

基于 SIFT 特征的多帧图像超分辨重建

张 志^{1,2)} 叶 蓬¹⁾ 王润生¹⁾

¹⁾(国防科技大学 ATR 实验室,长沙 410073) ²⁾(空军装备研究院,北京 100085)

摘 要 精确的亚像素级图像配准是图像超分辨重建中的关键问题。在图像超分辨重建中广泛使用的基于像素特征的光流法,对于大幅度运动场的计算很难做到精确的亚像素级配准。本文考虑了一种基于 SIFT (scale invariant feature transform) 特征的鲁棒性多帧图像超分辨重建算法。首先提取输入的低分辨率待匹配图像对的 SIFT 关键点及其特征矢量,随后选取候选匹配关键点对,通过 RANSAC (random sample consensus) 鲁棒方法去除奇异值,并根据假设的平移性几何约束模型,获得图像对的平移运动配准参数,然后选取视场中心对应的或指定的图像帧为初始参考帧,再使用传统的超分辨重建框架获得最终的重建结果。仿真实验结果表明,提出的基于 SIFT 特征的图像超分辨重建方案是有效的,超分辨重建的图像质量在主观评价和客观指标上都获得了优于经典算法的效果。

关键词 超分辨 图像配准 SIFT 关键点 光流法 参考帧

中图法分类号: TP911.73

文献标识码: A

文章编号: 1006-8961(2009)11-2373-05

Multi-frame Image Super-resolution Reconstruction Based on SIFT

ZHANG Zhi^{1),2)}, YE Peng¹⁾, WANG Run-sheng¹⁾

¹⁾(ATR Laboratory, National University of Defense Technology, Changsha 410073) ²⁾(Equipment Academy of Airforce, Beijing 100085)

Abstract Accurate sub-pixel image registration is a key problem in image super-resolution reconstruction. Optical flow methods based on pixel feature, which are widely used in image super-resolution reconstruction, are difficult to achieve registration of sub-pixel accuracy for large motion field. This paper considered a robust multi-frame image super-resolution reconstruction method based on SIFT. Firstly, SIFT operator was used to pick up keypoints and their descriptors of input low-resolution image pairs which are to be registered. Then the candidate keypoint pair was selected, outliers were wiped off through RANSAC, and images pair displacement was computed at the basis of assumed transitional geometry constraint model. Secondly, initial reference frame was selected from vision center frame or specified image frame. Lastly, super-resolution reconstruction was done through conventional super-resolution reconstruction framework. Experimental results show that the proposed image super-resolution reconstruction method based on SIFT is feasible, and the quality of super-resolution reconstructed images is better than those of classical methods by both subjective evaluation and objective standards.

Keywords super-resolution, image registration, SIFT, keypoint, optical flow, reference frame

1 引 言

高分辨率图像能更详细地表示景物的细节信息,在计算机视觉、遥感、医学等诸多领域有着广泛

的应用。空间分辨率取决于光学传感器的空间密度及其引入的模糊误差,这些因素限制了图像中被观察到的物体或特征的最小空间尺寸,通过改进硬件设备性能的途径并不能无限制地提高获取图像的分辨率,而是存在一个物理极限,当前的图像传感器技

基金项目:国家自然科学基金项目(60302007);国防科技重点实验室基金项目(9140C8004010611,9140C8002010705)

收稿日期:2009-06-17;改回日期:2009-07-15

第一作者简介:张 志(1982 ~),男,国防科技大学电子科学与工程学院 ATR 实验室博士研究生。主要研究方向为图像处理、分析与理解。E-mail:zhzh.zhang@gmail.com

术已经较为接近这个物理极限。于是,通过软件途径来得到高分辨率图像的超分辨技术已经成为人们广泛研究的热点^[1-4]。

多帧图像超分辨重建是建立在准确的亚像素级运动估计的基础上的。传统的超分辨重建算法^[3]出于模型简化的考虑,大多使用在亮度不变假设下的基于像素特征的运动估计方法,如光流法。但是光流法利用略去泰勒级数展开的高阶项而建立的光流约束方程,对于大幅度的运动位移的误差相对较大,准的鲁棒性相对较差,从而直接影响最终的超分辨重建效果。

2 图像超分辨重建的数学模型

考虑尺寸为 $L_1 N_1 \times L_2 N_2$ 的高分辨图像 \mathbf{X} , 将其写为向量形式 $\mathbf{X} = [x_1, x_2, x_3, \dots, x_N]^T$, 其中, $N = L_1 N_1 \times L_2 N_2$, 低分辨图像的尺寸为 $N_1 \times N_2$, 参数 L_1 和 L_2 分别为观测模型水平方向和垂直方向的下采样因子。令第 k 个低分辨图像的向量表示为 $\mathbf{F}_k = [f_{k,1}, f_{k,2}, f_{k,3}, \dots, f_{k,M}]^T$, 其中 $k = 1, 2, 3, \dots, K, M = N_1 \times N_2$ 。观测到的低分辨图像 \mathbf{F}_k 是由高分辨图像 \mathbf{X} 通过变形、模糊、采样算子和噪声降质得到的。可以将观测模型写为式(1):

$$\mathbf{F}_k = \mathbf{D}\mathbf{M}_k\mathbf{B}_k\mathbf{X} + \mathbf{n}_k \quad 1 \leq k \leq K \quad (1)$$

式中, \mathbf{B}_k 代表 $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$ 的模糊矩阵, \mathbf{M}_k 是一个尺寸为 $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$ 的变形矩阵, 包括全局的或者是局部的平移旋转等, \mathbf{D} 为 $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$ 采样算子矩阵, \mathbf{n}_k 是加性噪声向量。详细的数学观测模型见文献[2]。

3 基于像素特征的运动估计

光流法是多帧超分辨图像重建方法中最常用的基于像素特征的运动估计方法^[5]。光流场可看成像素点在图像平面上运动而产生的瞬时速度场。令 $f(x, y, t)$ 为时刻 t 图像点 (x, y) 的灰度值, 利用亮度一致性假设并略去泰勒级数高阶项可得到光流约束方程表示为

$$f_x u + f_y v + f_t = 0 \quad (2)$$

式中, f_x, f_y 和 f_t 分别表示图像中像素灰度沿 X, Y 和 T 方向的梯度, u, v 为图像点在相应方向上的移动速度。光流方程表明, 运动图像中某一点的灰度时间变化率是灰度空间变化率与该点空间运动速度的乘积。具体的计算过程见文献[6]。而在光流方程的建立过程中, 由于使用泰勒级数近似的方法, 对于偏移量较大的运动不可避免地引入较大误差, 很难达到精确的亚像素级配准。而多帧图像超分辨图像重建是建立在准确的亚像素级运动估计的基础上, 因此在这种情况下, 就需要寻找一种更加稳健的基于特征的运动估计方法去获取精确的运动估计, 从而获得较好的图像超分辨重建效果。

4 基于 SIFT 特征的超分辨重建算法

本节考虑了基于 SIFT 特征的多帧图像超分辨重建问题。首先提取输入的低分辨待匹配图像对的 SIFT 关键点及其特征矢量; 随后在特征空间中以欧氏距离为相似性度量, 采用最近邻比的方法进行特征矢量的匹配, 得到候选匹配特征对, 应用 RANSAC 鲁棒方法去除奇异值, 并利用超分辨重建中通常所

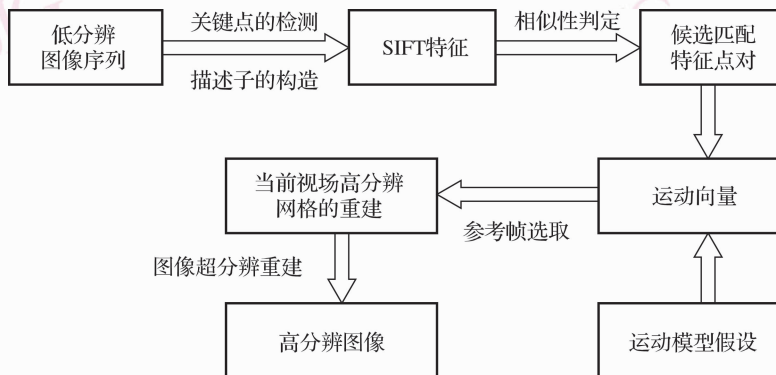


图 1 基于 SIFT 特征的超分辨重建算法框架

Fig. 1 Image super-resolution reconstruction framework based on SIFT

假设的纯平移运动的几何约束模型^[2],获得图像对的平移运动配准参数;然后选取视场中心对应的或者指定的图像帧为初始参考帧,再使用传统的超分辨重建框架^[3]获得最终的高分辨图像的重建结果。具体的框图如图 1 所示。

SIFT 特征的构造提取方法包括关键点的检测和描述子的构造两部分^[7]。

为了使特征具有尺度不变性,关键点检测是在多尺度空间完成的。将输入图像通过不同尺度的高斯核函数连续滤波和下采样,形成高斯金字塔图像,然后再对相邻尺度的两个高斯图像相减得到 DOG (difference of Gaussians) 金字塔多尺度空间表示。对 DOG 尺度空间每个点与相邻尺度和相邻空间位置的点逐个进行比较,得到的局部极值位置即为关键点所处的位置和对应的尺度。另外,通过曲面拟合的方法可以对关键点进行进一步的精确定位,以获得亚像素级的定位精度。

为了使特征具有旋转不变性,在构造 SIFT 描述子之前首先为每个关键点赋予一个主方向。关键点的尺度用来选择对应的高斯滤波图像,从而所有的计算具有尺度不变特性。在实际的计算当中,在对应的高斯滤波图像上,在以关键点为中心的邻域窗口内采样,并用直方图统计邻域像素的梯度方向。求梯度方向直方图时,用梯度幅值和圆形高斯加权窗对梯度方向加权,以突出中心区域的作用,抑制边缘区域的影响。梯度直方图的范围是 $0 \sim 360^\circ$,其中每 10° 一个柱,总共有 36 个柱。直方图的峰值则代表了该关键点处邻域梯度的主方向,即作为该关键点的方向。至此,图像关键点的检测已经实现,每个关键点都获得 3 个重要信息:空间位置、尺度和主方向。通过这些信息,可以为关键点构建一个局部坐标系,在局部坐标系内对图像特征的描述对空间位置、尺度及旋转变换具有不变性。

由此就可以对关键点进行不变特征向量的描述。SIFT 描述子的构造过程为:对每一个关键点,在其局部不变坐标系中,取以关键点为中心的 16×16 像素大小的邻域,再将此邻域均匀地分为 4×4 个子区域,对每个子区域计算梯度方向直方图(直方图均匀分为 8 个方向)。然后对 4×4 个子区域的 8 方向梯度直方图根据位置依次排序,这样就构成了一个 128 维的 SIFT 描述子向量。

当两幅图像的 SIFT 特征向量构造完成后,下一步采用关键点特征向量的欧氏距离来作为两幅图像

中关键点的相似性判定度量。取图像对中的一幅图像的某个关键点,并找出其与另一幅图像中欧氏距离最近的前两个关键点,在这两个关键点中,如果最近的距离除以次近的距离小于某个比例阈值(一般选取 $0.4 \sim 0.6$ 之间),则接受这一对匹配点。若降低这个比例阈值,SIFT 匹配关键点数目会减少,但更加稳定。依次选取获得匹配关键点对,然后根据运动模型假设,计算原始低分辨图像序列的运动向量序列。考虑当前视场的中心点选取参考帧,进行高分辨网格的重建,最后根据图像超分辨重建框架可获得最终的超分辨重建图像。

5 仿真实验及结果

为了验证该算法的有效性,给出了仿真实验,并与 Farsiu 的经典图像超分辨重建算法^[8]和三次样条插值算法进行了比较。

选取 512×512 像素的标准 Lena 图像作为高分辨图像背景,使用文献[4]提出的生成模型,模拟生成实验图像序列。假设采用高斯模糊模型,使用 3×3 的空间不变高斯低通滤波器,标准离差设为 1,下采样因子定为 4,对所有低分辨率图像均加入加性高斯噪声,满足信噪比(SNR)为 30 dB。运动向量的设计包括了遍历一个 4×4 的正方形区域和一个抖动剧烈初始帧,运动位移的具体量如图 2 所示。因为要考虑边界问题和模拟相机运动参数,所以设定输出的低分辨图像序列的每一帧尺寸为 100×100 像素,输出的帧数设为 80 帧。基本光流算法采用 OpenCV 中采用的经典 Lucas-Kanade 光流法的金字塔算法去配准原始输入低分辨图像序列,所使用的 SIFT 算法是 Lowe 经典算法^[7]。

根据图 1 的算法框架,对 80 帧低分辨图像序列进行运动估计,获得初始的运动向量。两种方法所获得的运动向量值与真实值之间的误差如图 2 所示(以水平方向为例)。

由图 2 可以看出,对于运动偏移量较大的运动向量,光流法的计算结果和真实值的差距较大,因而引入较大配准误差,这样就导致了最终的超分辨重建图像效果比较差。而基于 SIFT 特征的配准方法获得了相对较好的效果。再利用根据当前的视场选取参考帧,最后进行超分辨重建,得到的实验结果如图 3 所示。

图 3(a) 显示的是 Lena 高分辨图像的真实值,

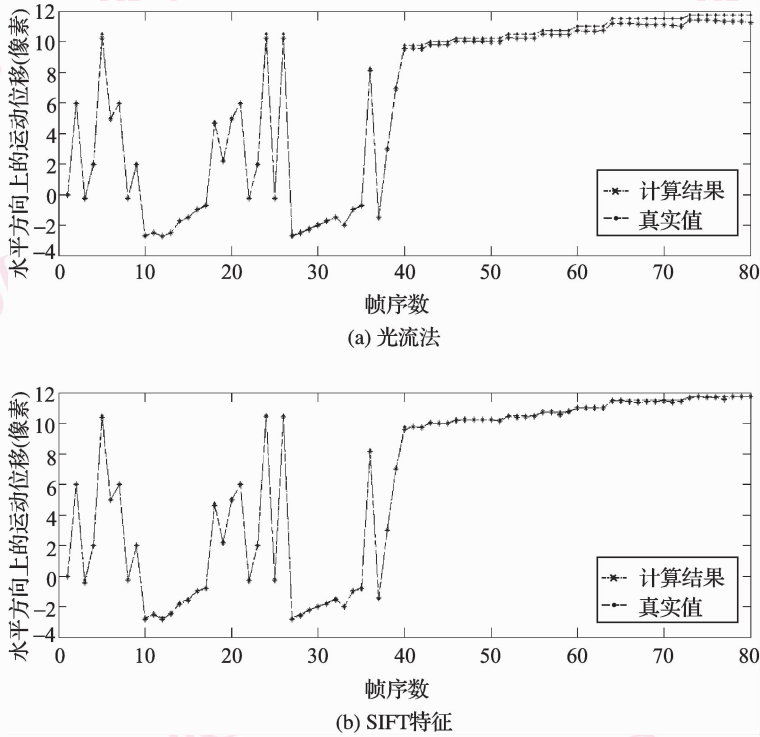


图 2 光流法与 SIFT 特征获取运动向量的计算误差对比

Fig. 2 Comparison of motion estimation error between the optical flow and SIFT



图 3 仿真实验结果

Fig. 3 Experimental results

图 3(b) 是三次样条插值所得到的结果, 图 3(c) 是使用 Farsiu 提供的软件包中的 Bilateral S&A 方法得到的结果^[9], 图 3(d) 是本文提出的方法得到的结果。

从图 3 可以看出, 三次样条插值的方法由于仅仅是单帧图像进行的空间扩展, 不可避免地导致边界的模糊, 效果一般; Farsiu 的方法是使用光流法进行图像配准, 使得对某些帧配准时误差比较大, 导致最终的重建图像效果较差, 甚至图像质量劣于仅仅通过插值方法获得的结果, 这个解释和文献^[10]的结论是一致的, 也说明了精确的亚像素级图像配准在超分辨率重建中的重要作用。而本文提出的方法在效果上优于以上两种比较的方法, 具体的定量比较见表 1。

表 1 仿真实验结果的定量评价

Tab. 1 Quantitative evaluation of experimental results

	MSE	PSNR (dB)	SSIM
Farsiu 方法	30.201 2	18.530 3	0.499 02
三次样条插值方法	26.640 7	19.619 9	0.524 15
本文方法	5.614	33.145 3	0.918 17

使用的比较参数考虑了均方误差 (mean square error, MSE)、峰值信噪比 (peak signal to noise ratio, PSNR)、结构相似度 (structural similarity, SSIM)^[11-12] 3 个参数。SSIM 参数是 Wang 等人最近提出的图像质量评价度量, 用来度量两幅图像间的结构性失真。

从图 3 和表 1 可以看出,本文方法无论是在统计性和结构误差的指标上,还是主观视觉评价上均优于以上两种传统方法。

6 结 论

分析了使用光流法在图像超分辨重建中存在的局限,考虑了一种基于 SIFT 特征的鲁棒性多帧图像超分辨重建算法。仿真实验结果表明,提出的方案是有效的,超分辨重建的图像质量在主观评价和客观指标上都获得了优于经典算法的效果。

参考文献 (References)

- 1 Wu Wei, Yang Xiao-min, Chen Mo, *et al.* Novel method of face hallucination[J]. *Optics and Precision Engineering*, 2008, **16**(5): 815-821. [吴炜, 杨晓敏, 陈默等. 一种新颖的人脸图像超分辨率技术[J]. *光学精密工程*, 2008, **16**(5):815-821.]
- 2 Park S C, Park M K, Kang M G. Super-resolution image reconstruction: A technical overview[J]. *IEEE Signal Processing Magazine*, 2003, **20**(3):21-36.
- 3 Farsiu S, Robinson D, Elad M, *et al.* Advances and challenges in super-resolution[J]. *International Journal of Imaging Systems and Technology*, 2004, **14**(2):47-57.
- 4 Yau A C, Bose N K, Ng M K. An efficient algorithm for superresolution in medium field imaging[J]. *Multidimensional Systems and Signal Processing*, 2007, **18**(2-3):173-188.
- 5 Baker S, Matthews I. Lucas-Kanade 20 years on: A unifying framework[J]. *International Journal of Computer Vision*, 2004, **56**(3):221-255.
- 6 Wang Run-sheng. *Information Fusion* [M]. Beijing: Science Press, 2007. [王润生. *信息融合* [M]. 北京: 科学出版社, 2007.]
- 7 Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- 8 Farsiu S, Robinson M D, Elad M, *et al.* Fast and robust multiframe super resolution[J]. *IEEE Transactions on Image Processing*, 2004, **13**(10):1327-1344.
- 9 Farsiu S, Robinson D, Milanfar P. MDSP Resolution Enhancement Software [EB/OL]. <http://www.soe.ucsc.edu/~milanfar>, 2009-03-20.
- 10 Zhao W Y, Sawhney S. Is super-resolution with optical flow feasible? [A]. In: *Proceedings of the 7th European Conference on Computer Vision (ECCV02)* [C], Copenhagen, Denmark, 2000:599-613.
- 11 Wang Z, Bovik A C, Sheikh H R, *et al.* Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, **13**(4):600-612.
- 12 Wang Z, Wu G, Sheikh H R, *et al.* Quality-aware images[J]. *IEEE Transactions on Image Processing*, 2006, **15**(6):1680-1689.