

基于 SURF 和 TPS 的立体影像密集匹配方法

侯文广 吴梓翠 明跃

(华中科技大学 生命科学与技术学院, 湖北 武汉 430074)

摘要: 利用加速鲁棒特性(SURF)算法搜索影像的特征点, 基于薄板样条(TPS)建立对应域之间连续和光滑形变变换, 用于影像视差的估计, 在此基础上提出一种新的立体影像密集匹配方法. 首先, 基于 SURF 算法进行特征匹配; 然后, 利用极线和 TPS 变换约束选择稳定可靠的同名点, 计算影像之间的 TPS 变换矩阵, 估计对应点的位置; 最后, 基于影像灰度进行密集匹配, 并利用极线约束和 TPS 矩阵删除误匹配点. 为保证视差估计的可靠性, 利用种子点的外接凸边形对匹配区域进行限定. 以人脸三维重建中的影像匹配为例, 得到了稳定可靠且密集的同名点.

关键词: 影像匹配; 加速鲁棒特性(SURF); 尺度旋转不变特征(SIFT); 薄板样条(TPS); 视差

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-4512(2010)07-0091-04

Densely stereo image matching using SURF and TPS

Hou Wenguang Wu Zicui Ding Mingyue

(College of Life Science and Technology, Huazhong University of Science

and Technology, Wuhan 430074, China)

Abstract: Speeded-up robust features (SURF) has more chance of finding corresponding feature points, which can be used as seed points for densely matching. Thin-plate spline (TPS) transformation can establish a smooth and continuous transformation relationship between corresponding regions which can be used to estimate parallax in matching. Based on the two methods, the paper proposes a novel matching solution. Firstly, a set of reliable seed points was obtained by filtering the corresponding points obtained using SURF with TPS. Then, the corresponding points of each pixel in left image were estimated by the TPS matrix. Thirdly, the dense match can be carried out between two images. Since the seed points and TPS transformation could estimate the coordinates of corresponding point more accurately, so the dense match can be performed easily, reliably and quickly. At the same time, the matching region was confined which is the outer protruding polygon surrounded seed points. Finally, the errors corresponding points using TPS and epipolar restriction were deleted. This new method was clarified using some stereo images in human face reconstruction whose similarity is low very much. Using this method, a dense and reliable point cloud can be obtained.

Key words: image matching; speeded-up robust features (SURF); scale invariant feature transform (SIFT); thin-plate spline (TPS); parallax

立体影像匹配本质上是一个病态问题, 主要是由于成像过程中失去了目标的三维信息, 摄影环境和噪声等因素也进一步增加了匹配的难度^[1~3]. 近景影像的匹配更加困难, 原因是摄影距

离相对于物体自身起伏的变化较小, 立体影像的相似程度低, 局部区域的比例尺变化显著, 视差分布连续性差. 本研究以人脸三维重建中的影像匹配为应用背景, 人脸目标外形复杂, 是近距离立体

收稿日期: 2009-11-10.

作者简介: 侯文广(1974-), 男, 博士, E-mail: houwenguang99@163.com.

基金项目: 中国博士后科学基金资助项目(20070410282).

©1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

视觉的典型情况. 尺度旋转不变特征(SIFT)具有对图像旋转、尺度和亮度变化的不变性, 对视角变化、仿射变换、噪声也保持一定程度的稳定性, 缺点是数据量大、计算复杂度高^[4]. Yanke 采用 PCA 对特征向量进行降维, Delponete 利用 SVD 进行特征匹配等, 主要改进在于特征向量的匹配^[5]; 而加速鲁棒特性(SURF)算法对兴趣点及其特征向量的提取方法进行了改进, 尽管在旋转不变性方面有一定程度的降低, 但总体仍能取得良好的匹配效果, 且计算速度显著提高^[6]. 薄板样条(TPS)变换在保证约束点严格满足对应关系的基础上, 通过最小化弯曲能量函数, 得到对应域之间一种连续和光滑的映射关系^[7-9], 可以有效地应用于估计影像视差. 本研究的策略就是采用 SURF 进行特征匹配, 利用 TPS 变换估计视差, 然后基于影像灰度进行密集匹配^[10].

1 算法流程

尽管 SURF 具有对旋转和尺度不变的优点^[6], 但其匹配得到的同名点远不能满足生成密集数字表面模型的要求, 仅可以作为密集匹配的种子点; TPS 变换是建立对应区域之间柔性变换关系的有效算法, 保证约束点(种子点)的严格对应关系^[7], 将主要用于未知点视差的估计. 本研究的匹配方法主要分为以下几个步骤: **a.** 建立金字塔影像(视实际情况而定); **b.** 基于 SURF 算法的立体影像特征匹配, 得到一定数量的同名点; **c.** 对初步得到的同名点进行筛选, 主要有极线约束条件和 TPS 变换约束, 可以采用随机采样一致性准则(RANSAC); **d.** 将筛选后的同名点作为种子点, 计算立体影像之间的 TPS 变换矩阵; **e.** 以其中一幅影像为基准设为左影像, 计算种子点的外接凸多边形, 限定匹配区域; **f.** 根据 TPS 变换矩阵, 估计待匹配点的视差, 得到同名点的预测区域; **g.** 给定匹配窗口的大小, 在同名点的预测区域内进行相关计算, 寻找相似性极大值的对应点, 若相似度大于给定的阈值, 则作为同名点; **h.** 基于极线约束条件和 TPS 变换约束条件剔除错误匹配的点; **i.** 将同名点作为参考点传递到下一层的影像匹配, 适应于建立金字塔影像的情况.

2 匹配策略

2.1 基于 SURF 的特征匹配

基于 SURF 的特征匹配主要包括如下步

骤^[5,6]: **a.** 在尺度空间中检测兴趣点; **b.** 计算特征点的主方向; **c.** 生成特征点的描述向量; **d.** 基于特征向量的同名点匹配. SURF 与 SIFT 的主要不同之处在于: 生成尺度空间时采用方框滤波近似代替二阶高斯滤波; 利用 Haar 小波代替直方图计算特征点主方向和生成特征向量. SURF 显著增加了特征提取的速度, 并且仍保持了所提取的特征向量具有旋转和尺度的不变性, 从而使其成为特征匹配的理想算法.

2.2 TPS 变换

TPS 是一种建立对应域之间柔性变换的方法, 以一定数量的点作为约束条件, 在这些约束点严格满足对应的条件下, 内插得到对应域之间一种连续和光滑的变换关系, 其原理简述如下^[9]: 设 x_1 和 x_2 分别表示立体像对左右两幅影像中同名点的坐标, 其维数为 $n \times 2$, 可计算得到

$$I = \begin{bmatrix} K & Q \\ Q^T & 0 \end{bmatrix}, \quad (1)$$

式中: K 是一个 $n \times n$ 的矩阵, $k_{ij} = U(|x_1^i - x_1^j|)$, $U(r) = r^2 \log r$; Q 是一个 $n \times 3$ 的矩阵,

$$Q = [1, x_1]. \quad (2)$$

令 $X' = [x_2, 0, 0, 0]^T$, 于是变换矩阵 W 为

$$I^T W^T = X'. \quad (3)$$

对于左图像中的任意一点 x , TPS 变换到右影像中的坐标为

$$x' = W U_d(x), \quad (4)$$

式中: $U_d(x) = (U(x - x_1), U(x - x_2), \dots, U(x - x_n), 1, x^T)$.

2.3 种子点的确定

若影像的内外参数已知, 则利用核线约束容易删除部分误匹配的点; 否则, 可以采用 RANSAC 方法来计算立体像对之间的极线关系, 同时删除错误匹配点^[11,12]. 对于内参数未知的情况, 若同时考虑到影像畸变的影响, 则极线约束条件的计算相对复杂. 显然, 极线约束不能剔除所有错误匹配的对应点, 故在极线约束基础上, 采用 TPS 变换作进一步的筛选, 算法如下: **a.** 建立左影像中同名点的 $k-d$ 树, k 正比于同名点的数量且大于 3; **b.** 对点 x_i , 设其邻域的点集为 X_i (不包括当前点 x_i), 对应于右影像中的同名点集为 X'_i , 利用 X_i 和 X'_i 计算 TPS 变换矩阵; **c.** 利用 **b** 中的 TPS 变换矩阵, 计算 x 在右影像中对应的点, 表示为 x'_i ; **d.** 计算 $dx_i = |x'_i - x_i|$, x'_i 为右影像中与点 x_i 对应的同名点; **e.** 对 $dx_i (i = 1, 2, \dots, N)$ 按照降序排列, 若均小于给定的阈值或者迭代次数超限, 则退出; 否则, 继续下一步; **f.** 将部分差

异值较大的同名点剔除; g. 返回 a.

该算法基于邻近区域内映射关系的一致性原理, 通过判断 TPS 的预测结果与已知同名点的差值, 达到剔除误匹配点的目的.

2.4 匹配区域的限定

稳定可靠的种子点为密集匹配提供视差估计, 但对于远离种子点的区域, 其视差预测的准确性一般不会很好, 因此, 对匹配区域进行限定, 仅对种子点所包围的凸多边形区域内的点进行密集匹配, 对应的算法是: 对左影像中的种子点进行 Delaunay 划分; 对位于三角网边界上的顶点(边界是指仅属于一个三角形的边)再进行 Delaunay 划分; 于是可以得到三角形数量相对较少的三角网, 这些三角网组成的区域为由种子点组成的外接凸多边形; 计算种子点的最小外接矩形; 通过判断外接矩形内的点是否位于新构建的三角网的某个三角形内, 达到限定匹配区域的目的. 在应用中, 即使将匹配区域限定在此范围内, 其结果仍然存在相当大的冗余.

2.5 密集匹配

首先, 利用种子点计算 2 幅影像的 TPS 变换矩阵(表示为 M)以及极线约束关系; 以左影像为基础, 对位于待匹配区域中的 x , 通过 TPS 变换计算其在右影像中对应的候选点 x' ; 给定匹配窗口 w , 在一定的匹配区域 W 中, 计算影像相关系数; 相关系数最高的点将作为匹配同名点 x_r ; 要求相关系数高于确定门限 T ; 利用 TPS 的逆变换关系 M^{-1} , 计算右影像中点 x_r 对应于左影像中的像点 x_1 , 判断 $|x_1 - x|$ 是否小于给定的阈值 d_T ; 最后根据极线约束条件进行误匹配点的剔除.

该计算过程如图 1 所示, 设左影像中的三角网边缘包围的区域为限定的匹配区(所有的种子点均在该三角网中), 图中深色的矩形区域为匹配窗口, 若 $|x_1 - x| < d_T$, 且 x 和 x_r 满足极线约束, 则 $x - x_r$ 为最终确定的同名点. 本算法只能达到像素级的精度, 若要得到更高的匹配精度, 可以参照相关的匹配方法, 同时, 其他相容性约束条件均可在匹配和噪声剔除中应用.

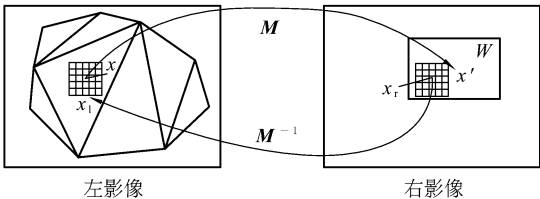


图 1 基于 TPS 变换的密集匹配过程示意图

3 实验与分析

本实验数据来自香港理工大学工业中心开发的三维人脸重建系统, 图 2 中为真实人脸不同角度的 4 个立体像对, 由于人脸本身纹理信息的变化不够丰富, 在获取影像的同时, 投影一定的随机纹理于脸上, 以利于立体影像匹配的实现, 从而提高匹配的效率和可靠性. 图中立体像对(a), (b), (c)和(d)主要用于获取人脸右侧、左侧、正面和下巴的三维模型, 对于这些立体像对而言, 相似程度低, 特征之间存在明显的旋转和尺度变化, 影像之间的视差分布相对复杂.

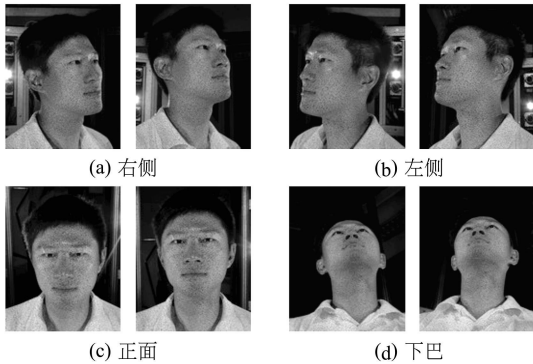


图 2 立体像对

需要指出的是, 依据光束法平差的原理, 本系统利用了一个标准的三维检校设备来求解各个相机的内外参数, 故在特征匹配之前极线约束条件是已知的, 采用该标准检校设备是因为各个像对匹配后计算得到的点云将位于同一个坐标系中, 无需进行点云的拼接; 本系统的摄影机均为普通相机(Canon 630 型号), 畸变大, 当采用 RANSAC 方法计算极线约束时, 未知参数多, 约束条件多, 计算相对复杂.

这里仅对第一个立体像对的匹配情况进行描述. 首先, 采用 SURF 算法提取特征点和特征向量, 分别提取了 11 210 和 17 081 个特征点, 特征点描述向量的维数为 64, 详细算法参考文献[5, 6]. 其次, 采用最近邻向量匹配法进行特征匹配(系数值为 0.5), 共得到 2 158 个同名点. 然后, 剔除多对一或者一对多的同名点(共 69 个点); 根据极线约束条件, 剔除不位于同名极线上的点, 若同名点到对应核线的距离超过 2 个像素, 则将其剔除, 共剔除 261 个同名点; 再利用 2.4 节方法做进一步剔除, 选择 $k=25$, 预测值与匹配点差值的门限设为 5 像素, 剔除 186 个点, 最后共得到 1 642 个可靠的同名点. 对最终的同名点进行仔细检查,

未发现错误的同名点, 尽管在采用 TPS 变换剔除误匹配点中可能会删除部分正确点, 但由于 SURF 算法得到特征点数量较多, 因此不会对后续匹配产生明显的影响。

利用种子点计算两幅影像之间的 TPS 变换关系后, 采用 2.5 节中的方法计算左影像种子点的外接凸多边形; 计算位于该凸多边形内像点 x 的对应点, 作为右影像中的预测点; 匹配窗口大小为 5×5 , 搜索区为以预测点为中心的 10×10 窗口; 采用相关系数法计算相似性最高的点, 且相关系数大于 0.85, 作为可能的同名点; 利用 TPS 的逆变换计算该同名点对应于左影像中的像点, 若该计算结果位于以 x 为中心的 10×10 窗口内, 则利用极线约束关系做进一步判断, 剔除误匹配点。本实验的图像大小为 $2\,400 \times 3\,000$, 为了提高计算速度, 每隔 5 个像素按上述方法进行匹配, 得到 9.7×10^4 个同名点, 其三维空间点云如图 3(a) 所示, 其他立体像对的匹配结果如图 3(b) ~ (d) 所示。

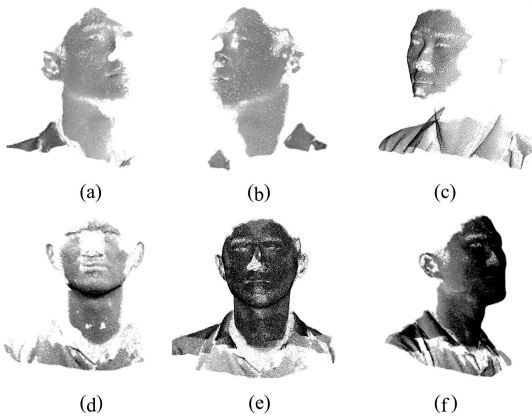


图 3 三维点云

a—人脸右侧; b—人脸左侧; c—人脸正面; d—下巴及嘴; e, f—合并后的人脸点云

从三维点云分析, 总体的匹配效果是良好的, 由于采用了多重约束准则, 因此有效地抑制了误匹配的情况。空间点云合并后, 如图 3 中的 (e) 和 (f) 所示, 共得到 37×10^4 个三维空间点, 一方面, 数据的冗余程度相当大; 另一方面, 点云中也可能存在一些空洞, 特别是在鼻子的右侧, 得到的同名点数

量相对较少, 原因是在立体像对图 2(a) 中, 鼻子处的影像变形相当严重, 而且采用了相对较高的相似性评价标准。匹配的结果也为本系统硬件的调整提供了一些参考依据。

参 考 文 献

- [1] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *Int J Comput Vision*, 2002, 47(1-3): 7-42.
- [2] Zhu Q, Zhao J. Triangulation of well-defined points as a constraint for reliable image matching [J]. *ISPRS J Photogramm*, 2005, 71(9): 1 063-1 069.
- [3] 明 洋. 特殊航空影像自动匹配的关键技术研究 [D]. 武汉: 武汉大学遥感信息工程学院, 2009.
- [4] Lowe D. Distinctive image features from scale-invariant keypoints [J]. *Int J Comput Vision*, 2004, 60(2): 91-110.
- [5] 张锐娟, 张建奇, 杨 翠. 基于 SURF 的图像配准方法研究 [J]. *红外与激光工程*, 2009, 38(1): 160-165.
- [6] Bay H, Ess A, Tuytelaars T. SURF: speeded up robust features [J]. *Comput Vis Image Und*, 2008, 110(3): 346-359.
- [7] Bookstein F. Principal wraps: thin-plate splines and decomposition of deformations [J]. *IEEE Trans on PAMI*, 1989, 11(6): 567-585.
- [8] 孙冬梅, 裴正定. 基于薄板样条实现非刚性图像匹配方法 [J]. *电子学报*, 2002, 30(8): 1 104-1 107.
- [9] Hutton T. Dense surface models of human face [D]. London: Eastman Dental Institute, University College London, 2004.
- [10] 张祖勋, 张剑清. 数字摄影测量学 [M]. 武汉: 武汉大学出版社, 2002.
- [11] Fischler M, Bolles R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. *Communications of the ACM*, 1981, 24(6): 381-395.
- [12] Zhang Z, Deriche R, Faugeras O. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry [J]. *Artif Intell*, 1995, 78: 87-119.