

基于 SIFT 特征点的图像拼接

李艳丽 向辉

(山东大学计算机科学与技术学院 济南 250061)

摘要 本文提出一种基于 SIFT (Scale Invariant Feature) 特征点的多幅图像全自动无缝拼接方法。该方法针对手持相机拍摄的多幅图像进行全自动拼接,要求其中一幅是基准图像,其他的图像跟基准图像有重叠部分。允许图像有缩放变换、旋转变换和轻微的仿射变换,不受图像噪声和色差的影响,具有很强的鲁棒性。设计一种高效的算法求取 8 参数矩阵,提高匹配的准确性,用多段融合法避免加权法的叠影现象。实验结果验证了该算法的有效性和执行的快速性。

关键词 全景图; 图像配准; 8 参数矩阵; K-D 树; LM 迭代法; 多段融合

Image Mosaicing Based on SIFT

Li Yanli Xiang Hui

(Department of Computer Science, Shandong University, Jinan, 250061)

Abstract: This paper presents a totally automatic method of image mosaicing based on SIFT (Scale Invariant Feature), which is to stitch images from hand-hold camera, demanding one of them is a basic image and the other are overlapped with this basic image. The SIFT is invariant to scale, orientation, slight affine transformation, noise and illumination of the images. So it is very robust. An efficient method to get the 8-parameters projective matrix is designed. Multi-band blending is used at last. Experimental results verified the algorithm's efficiency.

Key Words: Panorama, Image Registration, 8-parameter Projective Matrix, K-D Tree, LM, Multi-band Blending

1. 引言

图像拼接有着广泛的应用前景,自动建立大型、高分辨率的图像拼接技术也一直是摄影测量学、计算机视觉、图像处理和计算机图形学的活跃研究领域。

不同的应用产生不同的图像拼接技术:在虚拟现实和房地产展示等应用中,为了得到 360 球状全景图,需要用鱼眼镜头和其他的硬件获取图像,然

后恢复鱼眼图像的扭曲,最后进行图像配准和色度调整^[1],利用特殊的播放软件来浏览;柱面全景图也有广泛的应用,利用直接法或相关改进算法生成柱状全景图像^[2, 3]的技术已经比较成熟;目前针对手持相机获取的图像的拼接技术也越来越受到关注。

这类技术多用特征点匹配法。对于特征点提取,很多商业软件,如 PTgui 要求用户对拼接图像手动选取特征点。但在一些应用上,如卫星探测和大屏幕的拼接,特征点的自动获取和匹配就是不可

作者简介

李艳丽(1982—),女,硕士研究生,研究方向:图像处理和计算机视觉。

向辉(1970—),男,副教授,博士,研究方向:多媒体信息安全和人机交互与虚拟现实。

避免的问题。目前有许多特征点的选取方法,常用的有 Harris 特征点^[4]、KLT 特征点^[5]和 SIFT 特征点^[6]。研究表明^[7], SIFT 特征点的性能较优。另外,由于在辅助扫描技术、屏幕拼接技术和手持相机获取多幅图片的拼接技术中,往往会有一幅基准图片。根据这方面的应用,本文就基于 SIFT 特征点,实现了一个鲁棒性很强的、快速的、融合很好的全自动的图片拼接系统。

本文综合考虑拼接速度和拼接质量等因素,从几个方面来设计算法。

首先,考虑到输入图片之间的缩放变换与旋转变换及色差与噪声的影响,采用对这些变换具有不变性的 SIFT 特征点;其次,考虑到图像匹配的准确性和匹配速度,利用 K-D 树^[8]进行特征点匹配,并设计了一个算法高效准确的求取旋转矩阵;最后,为避免加权法产生的图像叠影现象,采用多段融合^[9]来拼接图像。

2. 图像拼接总过程

基于特征点的图像拼接一般包括 4 个部分:特征点提取和描述、特征点匹配、图像转换矩阵的求取及图像的拼接和融合。

本文相应地分成 3 个部分来描述拼接过程。第 3 部分描述图像 SIFT 特征点的提取过程,讨论一下 SIFT 特征点的不变性。第 4 部分利用 K-D 树,建立特征点匹配,详细描述如何由 M 对特征点匹配对求取最优转换矩阵。第 5 部分描述多段融合法进行图像拼接的过程。

总的算法实现:

- (1) 导入 $M+1$ 幅要拼接的图片 $P_0, P_1, P_2, \dots, P_M$, 默认第一幅是基准图像。
- (2) 提取每一幅图像的 SIFT 特征点, 建立 K-D 树。
- (3) 对每一幅要匹配到基准图像上的图像 P_i :
 - ① P_i 的特征点在 P_0 的 K-D 树上查找最近邻居点;
 - ② P_0 的特征点在 P_i 的 K-D 树上查找最近邻居点;
 - ③ 仅保留 $P_i \leftrightarrow P_0$ 中一一对应的特征点对, 并根据两特征点的描述符误差对特征点对从小到大排序, 删除误差过大的特征点对;
 - ④ 获取 P_i 和 P_0 的转换矩阵 H_i 。

3. 提取 SIFT 特征点

图像的特征点是图像固有的属性, 特征点一般是角点。另外, 因为图像获取过程中有可能产生椒盐噪声和高斯噪声, 希望提取的特征点不受这些图像噪声的影响, 图像获取过程中有可能会因为曝光或天气原因导致图像过亮或者过暗的现象, 希望提取的特征点不受色差影响。也希望特征点描述符具有图像缩放不变性和旋转不变性。针对这些要求, SIFT 特征点的提取过程包括下面 4 个步骤:

3.1 提取缩放空间的极值点作为候选特征点, 这一步保证了特征点的缩放不变性, 并保证特征点不受亮度差和噪声的影响。

图像的缩放空间 (图 1 (a)) 是图像的高斯空间 (图 1 (b)) 的差值, 即 DOG (Different of Gaussian)。图像的高斯空间是对原始图像进行一系列高斯变换得到的。图像的高斯变换保证了特征点不受噪声影响, DOG 保证特征点不受亮度差影响, 在一系列缩放空间中提取的极值点 $D(x, y, scale)$ (图 2 (a)) 保证缩放不变性。

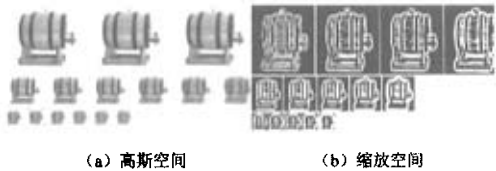


图 1 图像的高斯空间和缩放空间

Fig.1 Gaussian-Space and Scale-Space of the Image

3.2 精确定位特征点, 剔除响应弱的特征点和边线上的特征点。

获取的特征点应该是响应强的特征点, 需要剔除弱特征点。

先获取候选特征点处拟合函数

$$D(X) = D + \frac{\partial D^T}{\partial X} + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (1)$$

求导得到极值点:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2)$$

对应极值:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (3)$$

不断修正 X 求取局部最优点, 剔除 $D(\hat{X}) < 0.03$ 的弱特征点 (图 2 (b)), 同时获取候选特征点的精确位置, 缩放比例。

特征点是角点, 需要剔除边界上的特征点。

获取特征点的 Hessian 矩阵

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (4)$$

H 的特征值 α 和 β 代表 X 和 Y 方向的梯度, 剔除仅对一个方向梯度大的特征点 (图 2 (c))。经过这两步处理, 保留下的特征点具有很强的鲁棒性。



图 2 候选 SIFT 特征点的剔除结果

Fig.2 Result of Removing Keypoints With Low Contrast and Keypoints in Edges

3.3 确定特征点的主方向

这一步是为了进行特征点匹配用的, 找出主方向后, 在进行匹配的时候可以把图像所有旋转到主方向角度, 这样就达到缩放不变性了。对于每一个特征点, 考虑它周围一个窗口内的所有点, 最多点投票的方向就是主方向, 每一个点的对特征点方向的贡献是由一个以特征点为中心的高斯分布, 再乘上该点的梯度大小决定的。

计算特征点周围点 $L(x, y)$ 的梯度方向 $\theta(x, y)$ 和大小 $m(x, y)$:

$$\delta(x, y) = \tan^{-1} \{ [L(x, y+1) - L(x, y-1)] / [L(x+1, y) - L(x-1, y)] \}$$

$$\theta(x, y) = \{ (\delta(x, y) + \pi) \% 2\pi \} \times 36 \quad (5)$$

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (6)$$

用高斯窗口加权求和再投影得到方位直方图。

直方图的最高峰值点对应处的方向即为主方向, 如果直方图的高峰值 (高于最高峰值的 80%) 有好几个的话, 就可能出现多主方向的情形, 就复制同一个特征点, 使他们分别朝向不同的方向。这种情况不是很多。

3.4 描述特征点

特征点的描述和特征点的提取同样重要, 特征点的匹配就是进行特征点描述符的比较。前面获取特征点的缩放值 and 主方向都是为了将不同缩放图像或旋转方向图像的特征点建立在统一的框架中, 从而得到一样描述符。根据特征点的周围点得到的描述符, 描述符是一个 4×4 的窗口, 每一个窗口单元有 8 个方向, 即 $4 \times 4 \times 8 = 128$ 维。

SIFT 特征点 (x, y, σ, θ) 的描述符建立过程如下所述。

(1) 由缩放值 σ 确定特征点圆周大小。

(2) 对圆周窗口内所有点 (x', y', θ', m') :

① 绕特征点进行旋转变换和缩放变换, 旋转 θ , 缩小 σ , 得到在框架窗口 (4×4) 中的定位点 (x_r, y_r) ;

② 该点梯度方向 θ' 减去 θ , 得到在 8 个方向中的定位方向 (θ_r) ;

③ 将该点梯度值 m' 用高斯函数 $w(x_r, y_r)$ 处理后得到的 $m' \times w(x_r, y_r)$ 叠加到描述符 $FV(x_r, y_r, \theta_r)$

(3) $FV(x_r, y_r, \theta_r)$ $x_r = 0 \dots 3, y_r = 0 \dots 3, \theta_r = 0 \dots 7$ 为描述符信息。

经过以上 4 部分获取每一个特征点信息为 $(x, y, \sigma, \theta, m, FV)$ (如图 3 所示)。

ID66: X=114.0782, Y=223.0037, $\sigma=1.7253$, $\beta=-1.5370$ M=1.2289																																		
13	6	0	4	9	0	0	0	29	35	1	0	0	0	0	1																			
13	58	5	0	0	0	2	1	16	30	1	0	1	0	0	0																			
36	10	0	3	14	0	0	1	54	40	1	0	0	0	0	4																			
56	41	1	0	0	0	0	3	72	18	0	0	0	0	0	3																			

图 3 一个特征点的信息

Fig.3 the Information of a Keypoint

YanKe 和 Rahul Sukthankerd 对 SIFT 进行的改进, 提出了的 PCA-SIFT^[10], 将 128 维的描述符映射成 36 维, 设想很好, 但我们经过特征点匹配测

试,表明还是 SIFT 匹配的较准确。这里仍然采用了 SIFT 特征点。

4. 求取转换矩阵

对于相邻两幅图像,分别提取 SIFT 特征点后,如果一一对应查找最近邻居点需要 $O(n^2)$ 复杂度,通过建立 K-D 树^[8]降低复杂度, K-D 树是一棵多维空间的检索树,是二叉检索树的扩展,每一层将空间分成两个。树的顶层结点按一维进行划分,下一层结点按另一维进行划分,依次类推。建立的时间复杂度是 $O(n \log n)$,查找的时间复杂度是 $O(\log n)$ 。按特征点描述符建立图片的 K-D 树,特征点搜索 K-D 树中查找匹配点。

在特征点匹配对中,只保留两幅拼接图像的一一对应点对,对这 N 对特征点对按照特征点对描述符的差值从小到大排序,保留前 M 对,删除差值过大的特征点对。实验表明经过这两步的处理,保留下来的特征点对匹配得就相当精确了。

图片拼接需要计算他们之间的映射关系,根据摄影几何理论,固定相机位置,经图像的投影变换,一幅图像可以通过一系列矩阵变换转换成另一幅图像^[11,12,13,14]。

世界坐标系下的点 $P_w=(X_w, Y_w, Z_w)$ 投影到图像坐标系下的 $P_t=(X_t, Y_t, 1)$ 的转换关系:

$$P_t = TVRP_w \quad (7)$$

$$\text{其中, } T = \begin{bmatrix} 1 & 0 & C_x \\ 0 & 1 & C_y \\ 0 & 0 & 1 \end{bmatrix} \quad (8)$$

是图像坐标系下图像中心点阵对图像坐标原点的平移,如果图像中心点是图像坐标系的原点,则 $C_x = C_y = 0$, $T = I$ 。

$$V = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ 是 } f \text{ 控制的缩放变换。} \quad (9)$$

$$R = \begin{bmatrix} i_a \times i_b & j_a \times i_b & k_a \times i_b \\ i_a \times j_b & j_a \times j_b & k_a \times j_b \\ i_a \times k_b & j_a \times k_b & k_a \times k_b \end{bmatrix} \text{ 是旋转矩阵。} \quad (10)$$

从世界坐标系 (i_a, j_a, k_a) 到摄像机坐标系 (i_b, j_b, k_b) 的旋转,是一个有 6 个自由度的单位正交矩

阵,计算机视觉中多用 $R = e^{[\theta]_k}$ 来计算^[11,12,13,14],

$$[\theta]_k = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix} \quad (11)$$

从而固定转轴,获取两幅图像 l 和 k 的关系:

$$P_l = V_l R_l R_k^{-1} V_k^{-1} P_k \quad (12)$$

总的矩阵可以表示成一个 8 参数矩阵

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = X' \sim HX = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (13)$$

$$\begin{cases} -m_0x - m_1y - m_2 + m_6xx' + m_7yy' + x' = 0 \\ -m_3x - m_4y - m_5 + m_6xy' + m_7yy' + y' = 0 \end{cases} \quad (14)$$

已知 M 对 $X' \sim X$, 需要求出这个 8 参数矩阵,问题转化成从 $M \times 2$ 个方程中求取最优的 8 参数的线性方程组的求解问题。存在大量的解线性方程组的方法。有直接法和迭代法,直接法:准确,可靠,理论上得到的解是精确的,如高斯消元法、直接分解法和三对角阵的追赶法,可以利用直接法对四对特征点对求 8 个参数,很明显,会出现局部最优的问题,这种方法也行不通了;迭代法:速度快,但有误差如牛顿迭代法。

有些算法^[15]利用 RANSAC^[16,11,12]来预先估计矩阵,再用 LM 算法^[17]求精确解,这里经过前面的处理(剔除不是相互对应点对和差值较大的点对),保留下来的特征点对经实验表明已经匹配的比较准确。

参考 RANSAC,设计的匹配算法过程如下所述:

(1) 按照已经排好顺序的特征点对共 $M/4!$ 个提取前 k 个最优 4 点组合,这 4 点中任意两点都不能太近,任意 3 个点都不共线^[12]。

(2) 对每一个 4 点组合,计算转换矩阵 H ,设定一个误差阈值 T ,误差小于 T 的为内点,统计所有特征点匹配对,计算内点个数,内点误差和。这个转换矩阵可以用直接法求,这里用的也是 LM 算法,初始矩阵是单位矩阵。实验表明这种方法可行。

(3) 保留内点个数最多的组合,如果这样的组合多于两个,就保留内点误差和最小的组合,保留

的信息有内点个数 n ，内点信息，对应的初始转换矩阵。

(4) 根据以上保留下来的信息，利用 LM 迭代算法求取转换矩阵。

相关参数和参数经验值如下所述：

mixthreshdis——如果特征点对之间的距离大于这个阈值，就剔除此特征点对，设为 (10000.0~14000.0)；

mixcount——获取的前面最优组合的个数，设为 (100~400)；

threshinner——决定是否为内点的阈值，设为 (10~30)。

这个算法简单有效，速度比较快。常用的 RANSAC 算法是不断随机选取 4 点对得到 8 参数模型，本算法考虑到点对之间的误差信息，保留一一对应的点对，剔除误差大的点对，经过这两步处理已经将大多数的外点排除掉了，并且优先考虑误差小的特征点对，实验表明用这种方法提高了计算的速度，更重要的是保证了 8 参数模型的精确性。

5. 图片融合

经过上边的处理，每幅要拼接的图像跟基准图像都有了转换矩阵，通过反向求解可以将拼接图像映射到基准图像上。

先将图像 P_1, P_2, \dots, P_M 定位到基准图像上 P_0 ，计算出输出图像 P_{out} 的位置和长宽，下边分别处理 R, G, B 3 分量，由 P_{out} 的每个像素反向求解获取在其他图像上的对应点，理想情况下，各个对应点的像素值是一样的。但实际情况不是这样的，因为图像之间可能有色差，图像匹配过程中也会有细微的偏差，一幅图像的边界也往往会发暗，相机内部参数调整得不对也会产生图像扭曲的情况。因此，好的图像融合对拼接后图像的效果就相当重要了。

常用的融合算法有线性法和加权法。

柱面图像的拼接^[2]多用简单的线性法：

$$p(x, y) = p_1(x_1, y_1) \times w + p_2(x_2, y_2)(1 - w) \quad (15)$$

很多图像拼接系统^[3]用的是加权法：

$$p(x, y) = \sum (p_i(x', y') \times w_i(x') \times w_i(y')) / \sum (w_i(x') \times w_i(y')) \quad (16)$$

$p_i(x', y')$ 是第 i 幅图像在对应点 (x, y) 值， w_i

为线性权重函数，中心为 1，边界为 0。实验显示，如果使用加权法，拼接过程的偏差会使得产生图像重影现象，对于多幅图像的融合就造成图像多重叠区的模糊。为了避免叠影现象，本系统使用的是多段融合法^[9]。

多段融合是将输入图像分解成一系列不同频段的带通层，在带通层上进行图片拼接得到拼接图像的带通层，由拼接图像的带通层组合得到最终拼接图片。因为是在不同频段进行的图片拼接，所以能避免图片的叠影。图片的分解和组合是一个互逆的过程，即原始图片分解能得到带通层，带通层组合能得到原始图片，所以用这种方法得到拼接图片能保证正确性。

具体实现如下所述。

(1) 获取每一幅图像的低通层： G_0, G_1, \dots, G_n ， G_0 是原始图片，

$$G_L = \sum_{m,n=-2}^2 w(m, n) G_{L-1}(2i+m, 2j+n) \quad (15)$$

$w(m, n) = w(m) \times w(n)$ 为 5×5 窗口的加权函数。

(2) 获取每一幅图像的带通层： L_0, L_1, \dots, L_{n-1} ，

L_{n-1} ，

$$\begin{aligned} L_i(x, y) &= G_i(x, y) - \text{EXPAND}[G_{i+1}](x, y) \\ &= G_i(x, y) - 4 \sum_{m,n=-2}^2 G_{i+1}(\frac{2i+m}{2}, \frac{2j+n}{2}) \end{aligned} \quad (16)$$

(3) 图像融合操作在带通空间中用加权法实现，针对当前 L_k 层：

$$\begin{aligned} L_{kout}(x, y) &= \sum_{i=0}^{i=n-1} [L_{kii}(x', y') \times w_i(x') \times w_i(y')] / \sum_{i=0}^{i=n-1} [w_i(x') \times w_i(y')] \end{aligned} \quad (17)$$

获取的输出图像的对应带通空间 $L_{0out},$

$L_{1out}, \dots, L_{n-1out}$ 。

(4) 最后得到拼接图片

$$G_{out} = \sum_{k=0}^n L_{kout} \quad (18)$$

从图 4 可以看出，多段融合解决了加权法产生的叠影现象。



(a) 加权法

(b) 多段融合法

图 4 加权法和多段融合比较

Fig.4 Comparison of Linear-blending and Multi-band Blending

6. 实验结果

对大量图像实验的结果表明,该方法对手持相机拍摄的多幅图像的拼接行之有效且具有一定的鲁棒性,能给出很好的匹配结果。

图5给出了两幅图片特征点提取和匹配的实例。由此可见,提取的特征点具有很强的鲁棒性,经过剔除误差过多的特征点对,保留一一对应的特征点对这两步,保留下来的特征点对匹配相当精确。

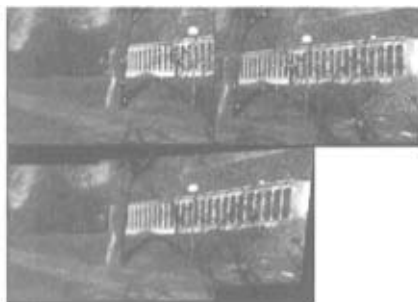


图5 特征点匹配和图片拼接结果

Fig.5 the result of Feature Matching and Image mosaicing

图6给出了6幅400×300图片的拼接结果,第一幅图片为基准图片,在2GHz的PC上运行了25s,产生出来的是739×455像素点的拼接图片。大部分的计算时间都花在了提取图片上的SIFT特征点上。

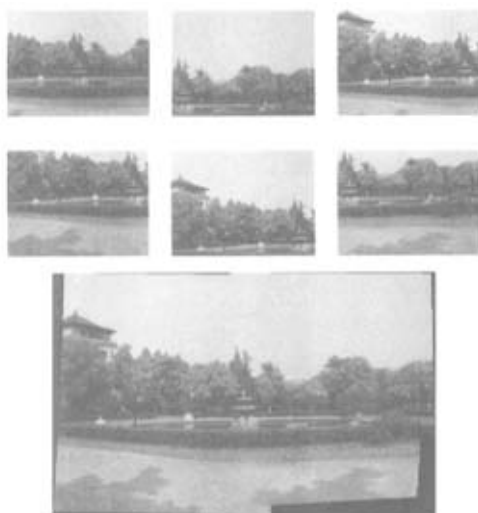


图6 实验结果

Fig.6 Experimental result

7. 结论

本文根据一些应用需求,提出了一种含有基准图片的多幅图片的全自动拼接方法,提取的SIFT特征点具有很强的鲁棒性,从而允许图像有缩放变换、旋转变换和轻微的仿射变换,也不受图像噪声和色差的影响;建立K-D树降低了算法复杂性;8参数旋转矩阵满足图像投影转换模型,并考虑特征点误差信息,设计了一种高效的算法来获取内点和最优转换矩阵,相对RANSAC算法提高了LM迭代参数的准确性;最后的多段融合解决了图像重影现象。

由于特征点的鲁棒性和匹配的准确性,该系统适用用大多有一幅基准图片的多幅图片的自动拼接。

参考文献

- [1] Zhao Yili, Chen Mingwei, Xu Dan. A system for creating and rendering spherical panorama using fisheyes images[C]. Chinagraph'2004, 2004: 202~205 (赵毅力, 陈明伟, 徐丹. 一个使用鱼眼图像进行球面全景图合成和绘制的系统. 第五届中国图形学大会. 2004:202~205)
- [2] Zhang Ju, Liu Jinyi. Constructing virtual environments based on cylinder panoramic images[C]. The Conference of China Virtual Reality and Vision 2004(CCVRV'2004). 2004: 151~155 (张菊, 刘金义. 基于柱面全景图的虚拟环境构建技术研究. 第四届全国虚拟现实与可视化学术会议, 2004 151~155)
- [3] Cao Junjie, Feng Jingbo, Su Zhixun. A panoramic image mosaic algorithm[J], 2003, 43(5): 180~182 (曹俊杰, 封靖波, 苏志勋. 全景图拼接算法. 2003, 43(5): 180~182)
- [4] C.J. Harris, M. Stephens. A combined corner and edge detector[C]. Proceedings of the 4th Alvey Vision Conference, 1988: 147~151
- [5] J. Shi, C. Tomasi. Good Feature to Track[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1994: 593~600
- [6] David G Lowe. Distinctive image features from scale-invariant keypoints [J]. International Journal of

- Computer Vision, 2004, 60(2): 91~110
- [7] K. Mikolajczyk , C. Schmid. A performance evaluation of local descriptors[C]. International Conference on Computer Vision & Pattern Recognition, 2003:275~263
 - [8] Andrew W .Moore. An introductory tutorial on kd-trees [R]//Andrew Moore. Efficient Memory-based Learning for Robot Control, PhD.Thesis.No.209, University of Cambridge. 1991:Chapter 6
 - [9] Burt , Adelson. A multiresolution spline with application to image mosaics[J]. ACM Trans on Graphics, 1983, 2(4):217~236
 - [10] Y. Ke, R. Sukthankar. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors[C]. Conference on Computer Vision and Pattern Recognition (CVPR).2004:506~503
 - [11] . David A. Forsyth , Jean Ponce .Computer vision –A modern approach[M].Prentice Hall,2003. (David A.Forsyth,Jean Ponce. 计算机视觉：一种现代的方法. 北京，清华大学出版社，2004)
 - [12] Richard Hartley , Andrew Zisserman. Multiple View Geometry in Computer Vision[M]. Cambridge University Press.2003:87-127,561~634
 - [13] Ma Songde, Zhang Zhengyou. Computer Vision-computer theory and algorithm(M),Science Press,1998:52~70. (马颂德, 张正友. 计算机视觉—计算理论与算法基础. 北京：科学出版社，1998:52~70)
 - [14] Richard Szeliski. Image Alignment and Stitching:A Tutorial1[R], Preliminary draft, January 26, 2005 Technical Report MSR-TR-2004~92
 - [15] M. Brown, D. G. Lowe. Recognizing Panoramas[C]. Proceedings of the 9th International Conference on Computer Vision(ICCV2003), 2003:1218~1225
 - [16] M.A. Fischler,R.C. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography[J]. Communications of the ACM.1981, 24(6):381~395
 - [17] K.Madsen , H.B.Nielsen , O.Tingleff .Methods for non-linear least squares problems 2 edition[R] .IMM 2004