

基于邻域传递的鱼眼图像的准稠密匹配

许振辉¹ 张峰¹ 孙凤梅² 胡占义¹

摘要 针对鱼眼图像高畸变的特点, 提出了一种适合于鱼眼图像的准稠密匹配扩散算法. 该算法采用局部仿射模型来建立图像对应区域之间的变换关系, 并利用这个仿射变换来规范化对应区域, 最后在规范化的区域上进行匹配扩散. 在局部仿射变换的计算中, 采用邻域传递的思路, 边扩散边邻域更新, 使扩散始终在较准确的对应区域内进行. 实验表明, 在大畸变的鱼眼图像下, 本文提出的准稠密匹配扩散算法能够取得比较令人满意的匹配扩散结果.

关键词 鱼眼, 准稠密匹配, 扩散, 邻域传递
中图分类号 TP391

Quasi-dense Matching by Neighborhood Transfer for Fish-eye Images

XU Zhen-Hui¹ ZHANG Feng¹ SUN Feng-Mei² HU Zhan-Yi¹

Abstract In this paper, a novel quasi-dense matching algorithm through local neighborhood transfer is proposed for fish-eye images, characterized by severe projective distortions. A local affine model is used for the geometric transformation between two corresponding image patches. Patches are normalized with the known affine transformation before match propagation is applied to them. The local affine transformation is computed during the propagation by a new mechanism called neighborhood-transfer to ensure the propagation to be carried out in the correct neighborhood. Experiments with real fish-eye images show that our algorithm can provide satisfactory results.

Key words Fish-eye, quasi-dense matching, propagation, neighborhood-transfer

三维重建, 即从二维图像恢复三维物体可见表面的几何结构的过程, 一直是计算机视觉的重要研究内容. 近年来, 随着三维重建在数字地球、数字考古、数字娱乐等领域应用的不断发展, 物体表面结构的视觉效果已成为三维重建越来越关注的一个问题. 然而, 图像对应点的自动匹配这一问题并没有得到很好的解决, 严重制约着三维重建技术的有效利用. 目前比较流行的匹配算法主要是基于特征的, 如基于 Harris 角点^[1], SIFT (Scale-invariant feature transform)^[2], MSER (Maximally stable extremal regions)^[3], Hessian-Affine^[4] 等的匹配方法. 这些方法一般只能得到比较稀疏的匹配, 通过它们只能恢复出三维空间中的稀疏点云, 不能有效表达三维物体的空间形状, 视觉效果差. 因此, 欲重建出较精确的三维结构, 需要得到稠密或至少是准稠密的匹配点.

最典型的稠密匹配是标准的稠密立体匹配. 经过几十年的发展, 已经产生了很多的立体匹配算法^[5-8]. 但是, 传统的立体匹配算法只适用于短基线并且相机像平面平行的情况, 对于长基线并不适用. 其次, 基于立体匹配的重建算法很费时, 内存要求也很高. 一些算法如基于体元 (Volumetric) 的算法^[6-7] 需要离散化待重建空间并且需要预先定义深度的精度. 如果要得到更精确的重建, 必然需要增加深度的层数, 使得算法对时间和内存的要求很高.

基于准稠密的方法^[9-12] 克服了基于稠密匹配的算法的缺点, 在较准确描述三维场景的同时仍能保持高计算效率, 是解决三维重建问题的一个较好的方案. 如在 Lhuillier 和 Quan 的方法^[9-10] 中, 以文献 [1-2] 的方法得到的稀疏匹配作为种子, 在其周围扩散出新的匹配, 并将新匹配又作为种子, 迭代直至种子匹配为空, 最终通过采样得到数量足够描述三维场景的准稠密匹配. Lhuillier 等通过实验指出其准稠密匹配算法在种子匹配有很大外点比例的情况下仍然能够恢复出正确的极几何关系, 因此该方法比基于稀疏匹配的方法更加稳定. Zhang 的方法^[11] 与此类似, 但有一个主要不同点, 即每次迭代都要在所有种子点周围扩散一次, 而 Lhuillier 等^[9] 只在一个最好种子点周围扩散. 以上两种准稠密匹配方法的共同缺点是只适用于短基线的情况. Kannala^[12] 提出的基于局部仿射模型的宽基线匹配扩散方法是对文献 [9] 的方法的扩展. 该方法假设图

收稿日期 2008-07-14 收修改稿日期 2009-01-08
Received July 14, 2008; in revised form January 8, 2009
国家自然科学基金 (60835003, 60673104, 60773132), 国家科技支撑计划 (2006BAK31B04) 资助
Supported by National Natural Science Foundation of China (60835003, 60673104, 60773132) and National Key Technology Research and Development Program of China (2006BAK31B04)
1. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190
2. 北方工业大学理学院 北京 100041
1. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. Faculty of Sciences, North China University of Technology, Beijing 100041
DOI: 10.3724/SP.J.1004.2009.01159

像的局部对应区域之间的变换关系, 都可用一个仿射变换近似. 因此, 对于每对种子匹配, 只要计算出二者邻域之间的仿射矩阵, 并利用该仿射矩阵将邻域校正到同一视角下, 就可文献 [9] 的扩散策略进行扩散. 该方法在宽基线的情况下获得了一定的效果.

以上所有的扩散算法都是基于透视图像的, 迄今为止, 我们还未见到关于鱼眼图像的匹配扩散的报道. 然而, 鱼眼镜头因为视角广, 在过去的十几年中受到了广泛的关注. 在那些需要大视角的应用场合, 如大场景的重建, 鱼眼相机有着普通透视相机不能比拟的大视场优势. 这是因为普通相机视角较小, 如果用普通相机图像来重建较大场景, 势必需要几幅甚至几十幅图像才能覆盖整个场景. 寻找图像间的对应, 重建出各个局部场景, 然后融合, 整个重建过程会变得相当复杂, 算法的时间复杂度和空间复杂度也会很大. 同时由于误差积累, 重建精度也会受到影响. 然而如果用鱼眼相机重建, 整个大场景往往用几幅图像就可以基本覆盖, 重建过程会大大简化. 因此, 探索一种适合于鱼眼图像的点匹配准稠密扩散算法具有重要的应用价值.

受文献 [12] 的启发, 我们将鱼眼图像的扩散思路归结为: 1) 计算每对匹配对应邻域之间的变换关系; 2) 利用这个变换关系来规范化对应的邻域; 3) 在规范化后的对应邻域上利用透视图像扩散中使用的视差梯度限制等来进行匹配扩散. 因此, 对鱼眼图像的匹配扩散而言, 最关键的问题是找到匹配点的正确对应邻域并且恢复出二者之间的变换关系. 然而对于给定的一对匹配点, 找出其邻域的正确对应是一个迄今未能很好解决的难题. 现在大部分的方法采用一种局部仿射模型, 即假设局部对应小区域之间近似满足一个仿射变换. MSER^[3], Harris-Affine, Hessian-Affine^[4], IBR (Image-based region), EBR (Edge-based region)^[13], 显著性区域 (Salient region)^[14] 等方法都属于这类方法, 但是它们只能在特定位置, 如灰度变化大的区域提取出对应区域并恢复出仿射变换矩阵, 不适用于我们的扩散算法, 因为扩散算法需要在任何图像点处都能得到邻域的变换关系. 另外一种与局部仿射模型相对的模型是任意变换模型, 如 Ling 和 Jacobs^[15] 提出的任意形变下不变特征的提取方法. 不过这种方法的时间复杂度很高且不稳定. 鉴于以上考虑, 我们采用了附加信息辅助下的局部仿射模型. 在我们的算法中, 这种附加信息就是对应点处极线角度的知识. 极几何关系可以通过图像初始的稀疏匹配来恢复^[16-18], 因此下文中我们假定两幅图像之间的极几何关系是已知的.

本文的结构如下: 第 1 节介绍传统准稠密扩散

算法及其相关的研究; 第 2 节具体介绍我们提出的基于鱼眼图像的扩散算法; 第 3 节给出扩散和重建的实验结果; 第 4 节总结全文.

1 透视图像的扩散算法

由于本文扩散方法的基本思路是对任意种子点对, 先在鱼眼图像上确定对应邻域, 并将邻域规范化成标准的透视图像对形式, 最后, 利用经典的针对透视图像的扩散算法进行扩散. 所以, 下面我们首先对透视图像的匹配扩散作一简单介绍.

Lhuillier 等的扩散算法^[9-10] 可以作为基于透视图像的扩散算法的代表. 它以两幅图像中的初始稀疏匹配为起始, 在其周围扩散出稠密的候选匹配, 然后通过采样得到较准确的准稠密匹配. 假设初始匹配集为 $(\mathbf{x}_i, \mathbf{x}'_i)$, 其中 \mathbf{x}_i 和 \mathbf{x}'_i 是两幅图像中的对应点. 将匹配按 ZNCC (Zero-mean normalized cross-correlation) 分数从高到低排成一个队列, 称为种子队列; 建立两个与图像同大小的标志矩阵, 矩阵中每个元素非 0 即 1, 分别标志对应像素是否找到匹配 (0 表示没有被匹配, 1 表示已被匹配, 初始为 0). 则 Lhuillier 等的扩散过程可概括为如下几个步骤:

步骤 1. 从种子队列中取出 ZNCC 分数最高的种子匹配, 并将其从种子队列中删除;

步骤 2. 在当前种子匹配的对应邻域内搜索出满足视差梯度限制的候选匹配;

步骤 3. 将这些候选匹配按 ZNCC 分数排成另一个临时队列, 称为局部队列; 按 ZNCC 分数从高到低的顺序从局部队列中逐个取出候选匹配; 如果候选匹配的 ZNCC 分数和置信度都超过既定的阈值, 并且它们没有被匹配过 (即相应的标志位为 0), 则将其判为正确匹配并保存下来, 接着将其放入种子队列中, 将匹配标志置为 1;

步骤 4. 迭代步骤 1~3 直至种子队列为空.

在步骤 2 中, 种子匹配点 $\mathbf{x}_i, \mathbf{x}'_i$ 的邻域分别被定义为

$$\mathcal{N}(\mathbf{x}) = \{\mathbf{u} | (\mathbf{u} - \mathbf{x}) \in [-N, N]^2\}$$

$$\mathcal{N}(\mathbf{x}') = \{\mathbf{u}' | (\mathbf{u}' - \mathbf{x}') \in [-N, N]^2\}$$

其中 N 为代表邻域大小的半径, 一般取为 2. 在邻域 $\mathcal{N}(\mathbf{x})$ 和 $\mathcal{N}(\mathbf{x}')$ 上, 满足二维视差梯度限制的候选匹配 $\mathcal{N}(\mathbf{x}, \mathbf{x}')$ 被定义为

$$\mathcal{N}(\mathbf{x}, \mathbf{x}') = \{\mathbf{u} \in \mathcal{N}(\mathbf{x}), \mathbf{u}' \in \mathcal{N}(\mathbf{x}'), \\ \|\mathbf{u}' - \mathbf{x}'\| - \|\mathbf{u} - \mathbf{x}\| \leq \varepsilon\}$$

其中 ε 是视差梯度的阈值, 一般取为 1.

在步骤 3 中, 置信度的限制是用来防止匹配扩散到灰度值过于一致、纹理信息非常少的区域 (如天

空), 它有利于提高扩散算法的可靠度. 图像 I 上某点 \mathbf{u} 的置信度被定义为

$$s(\mathbf{u}) = \max\{|I(\mathbf{u} + \boldsymbol{\delta}) - I(\mathbf{u})|, \boldsymbol{\delta} \in \{(\pm 1, 0), (0, \pm 1)\}\}$$

因此, 置信度限制就可以表示为 $s(\mathbf{u}) < t$ 或 $s'(\mathbf{u}') < t$, 其中 t 为置信度阈值, 一般取为 0.01. 另外, ZNCC 分数在固定大小的 $(2W + 1) \times (2W + 1)$ 的窗口上计算, 其中 W 为计算 ZNCC 分数的窗口半径.

此算法的缺点是只适用于传统的立体图像对. 另外, 由于在计算 ZNCC 分数的时候直接在图像上取矩形窗口, 图像的方向也需一致, 为了将扩散算法应用在宽基线图像对上, Kannala^[12] 和 Megyesi^[19-21] 分别提出了各自的基于宽基线图像对的扩散算法. 这些算法都使用仿射变换来近似局部区域的对应关系, 然后通过恢复仿射矩阵来规范化对应区域, 并在规范化后的区域上作扩散. Kannala^[12] 使用一种代表局部形状的灰度矩来更新对应区域的仿射关系. 假设在两个满足仿射变换 A 的局部对应区域上, 灰度矩分别为 S 和 S' , 则它们满足关系 $S' = ASA^T$. 因此在 S, S' 已知的前提下, 可以求出仿射矩阵为 $A = S'^{1/2}RS^{-1/2}$, 其中 R 可以通过计算主方向来求取. 这种思想与文献 [4] 基本一致, 但后者用的是灰度梯度矩. 所以像文献 [4] 中一样, A 的准确求解过程应该是一种迭代直至收敛的过程. 但是 Kannala 计算时却只迭代一次, 即将初始仿射矩阵 A_0 代入求得 S_0, S'_0 后, 直接计算出 A , 这种近似在很多情况下误差很大. 通过实验我们发现, 这种仿射关系的求取方法不准确且不稳定. 其实, 即使采用迭代的方法来求解 A , 也不会对结果有多大的改进. 因为对于大部分对应区域来说, 迭代过程都不收敛, 就像文献 [4] 中情况一样, 这也是 Hessian-Affine 只能找到稀疏的区域匹配的原因之一. Megyesi^[19-21] 在扩散之前先将图像对校正, 使得两幅图像的极线都在水平方向, 虽然减少了计算量, 但是大大限制了算法的使用范围. 事实上, 对大畸变图像的校正本身就是一件困难的事情. 对于鱼眼图像而言, 视角可以大于 180° , 此时, 从理论上来说要对鱼眼图像进行整体校正已不可行.

2 鱼眼图像的匹配扩散

我们的扩散算法也采用局部仿射模型, 并且基于这样一个假设: 三维空间物体表面是平滑的. 因此, 我们的扩散算法可以归纳为: 先找到种子匹配的邻域之间的仿射变换矩阵 (利用邻域传递的方式), 并用此仿射矩阵规范化对应区域, 然后在规范化后的对应区域上用传统扩散算法扩散. 本文的扩散算法的具体步骤由算法 1 给出.

算法 1 (鱼眼图像的准稠密扩散算法).

步骤 1. 找到初始匹配 (即种子匹配, 假设有 n 对), 并得到以种子匹配为中心的邻域的仿射矩阵; 规范化对应邻域, 在规范化后的区域上计算种子匹配的 ZNCC 分数; 建立一个全局队列 $seed_list$, 将 n 个种子匹配及其仿射矩阵一起按 ZNCC 分数从高到低排列到这个队列中.

步骤 2. 从队列 $seed_list$ 中取出 ZNCC 分数最高的匹配作为种子匹配 (取出此匹配的同时也从队列 $seed_list$ 中将其删除); 提取并规范化以种子匹配为中心的对应该邻域. 规范化的目的是使得 ZNCC 的计算用到的窗口包含相同的区域, 并且使得扩散中的二维视差梯度应用具有合理性.

步骤 3. 在规范化后的区域上, 以第 1 节中介绍的方法进行扩散, 得到一系列新的匹配; 在此规范化的对应区域上计算这些新匹配的 ZNCC 分数并保存.

步骤 4. 利用邻域传递的方式在新匹配上更新仿射矩阵, 然后将新匹配及其邻域仿射矩阵按 ZNCC 分数从高到低加入全局队列 $seed_list$ 中, 同时将标志位矩阵 $match_map$ 中相应位置 1.

步骤 5. 重复步骤 2~4, 直至队列 $seed_list$ 为空.

步骤 6. 通过一定约束剔除外点, 然后用重采样方法得到比较均匀的准稠密匹配点.

算法 1 中, $seed_list$ 是全局种子队列; $match_map$ 是标志位矩阵, 共 2 个, 分别表示两幅图像. 从算法 1 中可以看到, 扩散的核心问题就是获得匹配的对应邻域, 包括两方面内容: 1) 种子匹配的对应邻域的获得; 2) 扩散出的新匹配的对应邻域的获得. 本文用邻域传递的思想来计算新匹配的对应邻域. 以下先介绍邻域传递的思想, 然后再具体介绍算法的其他部分.

2.1 邻域传递

假设 $(\mathbf{x}, \mathbf{x}')$ 是一对种子匹配, 它们的邻域近似满足一个已知的仿射变换 A_0 . 又知在它们的邻域内找到一对新匹配 $(\mathbf{u}, \mathbf{u}')$, 现在要用邻域传递的方法来计算 $(\mathbf{u}, \mathbf{u}')$ 的邻域满足的仿射变换矩阵 A .

我们知道, 一个 2×2 的仿射矩阵有 4 个自由度. 但是在极线方向已知的前提下, 可以通过沿极线方向取对应邻域的方法, 将自由度减少到 3. 这可以通过将仿射矩阵分解的方法来实现, 具体过程如图 1 所示. 图 1 中, l, l' 分别是与 \mathbf{u}' 和 \mathbf{u} 相对应的极线; α 和 β 分别是两条极线在 \mathbf{u} 和 \mathbf{u}' 处的方向角; R_1, R_2 分别代表旋转角度 α, β 的旋转矩阵; 矩阵 B 是区域 P 与 Q 之间的变换矩阵, 自由度为 3.

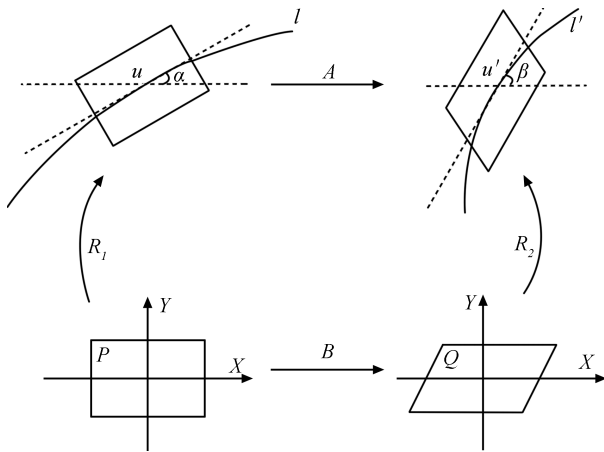


图1 仿射矩阵的分解

Fig.1 Decomposition of the affine matrix

首先在点 u 处取一个以 u 为中心的沿极线方向的矩形邻域 $\mathcal{N}'(u)$, 此矩形区域可以通过一个沿水平方向的矩形区域 P 旋转 α 角度得到. 由于 u 和 u' 是对应点, 则以它们为中心的对应邻域中包含的极线段对应的内容应该相同 (由于邻域取得很小, 邻域包含的极线段几乎为直线). 在变换满足仿射变换的假设下, $\mathcal{N}'(u')$ 只能通过 $\mathcal{N}'(u)$ 先沿极线 l 斜切和尺度变换、沿 l 垂直方向进行尺度变化后, 再旋转 $\beta - \alpha$ 角度得到. 因为仿射变换不改变直线之间的平行关系以及线段的比例关系, 而两对应极线段方向已知, 且分别平行于各自邻域的平行四边形的对应的两条边, 这些条件限制了仿射矩阵的形式, 减少了它的自由度. $\mathcal{N}'(u')$ 也可以通过将区域 P 沿 X 方向斜切和尺度变换, 沿 Y 方向尺度变换 (即图 1 中的 B 矩阵) 后, 再旋转 β 角度 (R_2 矩阵) 得到. 因此对于区域 P 上的任一点 x , 很容易得到 $AR_1x = R_2Bx$. 所以 A 与 B 的关系式可表示为

$$\begin{cases} B = R_2^{-1}AR_1 \\ A = R_2BR_1^{-1} \end{cases} \quad (1)$$

上式中 R_1 和 R_2 已知, 只要求出 B 就能计算出 A . 根据前面的讨论, B 矩阵有三个未知参数, 有如下形式

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

这三个参数有着不同意义, 它们的变动带来的效果如图 2 所示. 从图 2 可以看出, 参数 a 控制沿 X 方向的尺度变化, 参数 b 控制沿 X 方向的斜切大小, 参数 c 控制沿 Y 方向的尺度变化.

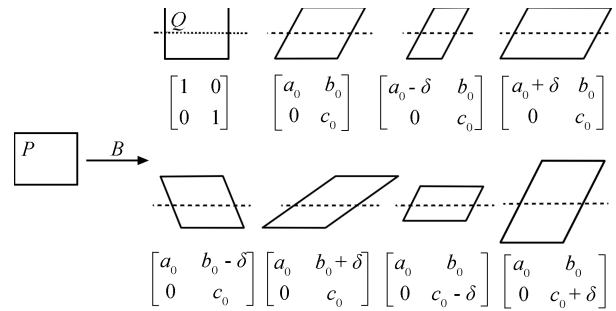


图2 矩阵 B 的三个参数的变化效果

Fig.2 Effect of changing three parameters of matrix B

所谓邻域传递, 实质上是指邻域对应的仿射矩阵的传递, 这里指从种子匹配 (x, x') 的邻域的仿射矩阵 A_0 到新匹配 (u, u') 的邻域的仿射矩阵 A 的传递. 为了减小搜索空间, 我们用从 B_0 到 B 的传递的方法来实现 A_0 到 A 的传递. 通过式 (1), 可计算得到 B_0 , 设其三个参数为 a_0, b_0, c_0 . 由平滑性可知, 新匹配 (u, u') 的 B 与 B_0 相差很小, 它的三个参数可表示为: $a = a_0 + \delta_1, b = b_0 + \delta_2, c = c_0 + \delta_3$, 其中 δ_i 为一微小变化量 (在实际计算时, 我们取 $\delta_1 = \delta_2 = \delta_3$), 这样就将 (x, x') 的邻域传递到了 (u, u') . 因此, 仿射矩阵的计算步骤为:

步骤 1. 由已种子匹配 (x, x') 的邻域的仿射矩阵 A_0 得到 B_0 (由式 (1) 进行计算).

步骤 2. 分别给 a_0, b_0, c_0 一系列微小变动量, 得到很多候选矩阵 B ; 然后由式 (1) 计算出 A , 规范化新匹配 (u, u') 的邻域, 在此规范化后的邻域上计算 ZNCC 分数, 将取得最高分数时的 A 作为正确值.

通过以上两个步骤, 在已知种子匹配 (x, x') 的仿射矩阵 A_0 的前提下, 就可以得到 (u, u') 的邻域的仿射矩阵 A 了. 由于 B_0 已知, 而且物体表面平滑, B 与 B_0 相差很小, 所以在实际应用中只要搜索很少的变化量就可以得到比较理想的结果. 本邻域传递算法充分利用了初始仿射矩阵 A_0 . 新的仿射矩阵 A 只需在与其相近的 A_0 的基础上进行较小范围参数空间内的搜索便可得到, 因此算法比较稳定. 另外, 算法中区域之间的相似度采用 ZNCC 分数度量, 充分利用了图像灰度信息, 保证了最后获得的矩阵 A 在搜索空间中的优化, 提高了算法的准确性. 正是由于以上两点, 使得本算法的仿射矩阵求取稳定性高于 Kannala^[12] 的方法. Megyesi^[19] 的仿射矩阵求取算法与本文较类似, 但是该算法需要预先校正图像, 这大大限制了算法的应用范围. 而本文算法并不需要预先校正图像, 其有更广的应用范围.

2.2 种子匹配及对应的仿射矩阵的计算

从第 2.1 节中可知, 邻域传递的前提是需要知

道种子匹配的邻域的仿射矩阵. 关于种子匹配邻域仿射矩阵的计算, 我们采用了下面两种方法.

2.2.1 利用仿射协变区域来得到种子匹配

为了获得初始稀疏匹配, 可以采用仿射协变区域检测子 MSER^[3] 或 Hessian-Affine^[4], 以及旋转不变描述子 SIFT^[2]. 首先用 Hessian-Affine 或者 MSER 在两幅对应图像中分别提取出仿射协变区域, 并用 2×2 的对称二阶矩阵表示. 假设一对仿射协变区域为 M_L 和 M_R , 则它们重心的坐标 \mathbf{x} 和 \mathbf{x}' 可以认为是一对匹配. 分别以 \mathbf{x} 和 \mathbf{x}' 为中心建立局部坐标系, 则对应区域的坐标之间满足一个仿射变换关系 B , 即 $\mathbf{x}_R = B\mathbf{x}_L$. 由文献 [4] 知

$$M_R = B^{-T} M_L B^{-1} \quad (2)$$

并且在 M_L 和 M_R 已知的前提下, 可以用式 (2) 估计出仿射矩阵, 即

$$B = M_R^{-1/2} R M_L^{1/2} \quad (3)$$

这里 R 是一个任意的旋转矩阵, 它可以通过在规范化的区域上计算 SIFT 的主方向^[2] 来得到. 这里的规范化与前一节中所述的情况稍有不同, 具体可参考文献 [4, 22]. 这里也可以直接利用对应的极线方向来计算 R , 这个方法更简便且更精确. 得到初始匹配之后, 有些是不可靠的, 需要利用一些约束来去除外点. 首先利用极线约束, 然后利用 ZNCC 分数. 经过以上两步之后剩下的种子匹配就比较可靠了.

2.2.2 利用 SIFT 和三参数空间搜索得到仿射矩阵

利用寻找仿射协变区域方法的缺点之一是找出的初始匹配非常少, 对某些图像对而言, 甚至仅有几对匹配. 如果场景简单, 几对种子匹配就完全可以可靠地扩散到整幅图像. 但是如果场景比较复杂, 很少的几对种子匹配往往不能可靠扩散. 如图像中含有两个垂直的墙面, 如果只有一个墙面上有种子匹配, 则此种子匹配就不能扩散到另一面墙上. 因为墙面交界处两边的仿射矩阵变化很大, 邻域传递将受阻, 从而导致扩散失败.

这里, 我们采取了另外一种方法: 1) 利用 SIFT 找到一些初始匹配, 一般至少都有几百对; 2) 利用较严格的极线约束来去除外点; 3) 利用第 2.1 节中的仿射矩阵的计算方法来得到这些匹配邻域的仿射矩阵, 并用其规范化局部邻域, 计算 ZNCC 分数, 分数高于一定阈值的点才保留.

经过以上三步就可以得到初始匹配以及与之对应的仿射矩阵. 需要指出的是, 这里的第 3 步中, 参数 a, b, c 都要在较大的一个范围内搜索. 这是因为种子匹配的邻域的信息是完全未知的, 并不像前面介绍的邻域传递过程那样已有一个较好的初始值

(事实上, 这也正是邻域传递的价值所在). 但是, 由于种子匹配较少 (一般限制在几十到一百多对之间), 时间的花费是可以接受的.

2.3 局部区域的规范化

在算法 1 中, 扩散是在种子点 \mathbf{x} 和 \mathbf{x}' 的大小为 $(2N+1) \times (2N+1)$ 的邻域 $\mathcal{N}'(\mathbf{x})$ 和 $\mathcal{N}'(\mathbf{x}')$ 上进行的^[9]. 匹配的好坏程度用 ZNCC 分数做标准, 而 ZNCC 用到的矩形框大小为 $(2W+1) \times (2W+1)$. 匹配 $(\mathbf{x}, \mathbf{x}')$ 周围邻域的具体规范化过程如下: 1) 在左图提取一块以 \mathbf{x} 为中心的大小为 $(2N+1) \times (2N+1)$ 的沿极线方向的矩形区域; 2) 通过已知的仿射变换矩阵 A , 直接变换矩形区域到右图像为一个平行四边形区域; 3) 将两块区域都用双线性插值变换成沿水平方向的 $(4N+1) \times (4N+1)$ 大小的新区域; 4) 重采样新区域为 $(2N+1) \times (2N+1)$ 大小. 如果仿射矩阵 A 准确, 那么理论上两片最后得到的区域应该代表相同的场景内容. 注意, 这里先将区域插值成 $(4N+1) \times (4N+1)$ 大小, 然后重新采样回 $(2N+1) \times (2N+1)$ 是为了防止出现插值后的锯齿效应.

2.4 外点的剔除和重采样

通过算法 1 的步骤 2~5 扩散出了比较稠密的匹配. 三维重建算法一般需要的是分布尽量均匀的准稠密点, 因此需要将这些已知的较稠密匹配采样成准稠密匹配. 另外, 在扩散得到的这些稠密匹配中, 存在着一些外点 (即误差较大的匹配), 采样之前需要先将它们剔除. 这里我们借助空间物体表面的平滑性约束, 并利用鲁棒算法 RANSAC (Random sample consensus)^[23] 对外点进行了剔除. 这部分内容具体可参考文献 [9-10].

3 实验结果

本节报道三个实际场景的鱼眼图像对的匹配扩散结果以及重建结果. 场景包括两个简单场景与一个复杂场景, 其中第二个简单场景图像包含较大程度的透视变化. 算法程序用 C++ 实现, 程序运行平台为主频 3 GHz, 内存 1 GB 的 PC 机. 以下三个实验中, 算法的主要参数都设置为: 邻域半径 N 取 3, ZNCC 计算窗口半径 W 取 3, 邻域传递过程 ZNCC 阈值取 0.9, 置信度 t 取 0.01, 视差梯度阈值 ε 取 1.

3.1 简单场景

我们人工设置了一个简单场景: 用一个布满纹理的盒子平面放于一块白板之上, 并且将种子匹配也限制在中间的盒子处. 因为有置信度的限制, 扩散只能在盒子表面进行, 而不能蔓延到白板区域内. 此场景图像包含旋转与小尺度变化, 图像分辨率为

3 008 × 2 000. 鱼眼图像的初始种子匹配采用第一种方法 (参见第 2.2.1 节) 来获取, 即先利用 MSER 得到区域匹配后再计算得到. 此处一共有 3 对种子匹配, 其仿射矩阵计算用时 15 ms, 平均每对用时 5 ms. 图像对其初始种子匹配如图 3 所示. 图 3 中, 匹配点用“+”号标出, 匹配区域用椭圆表示.

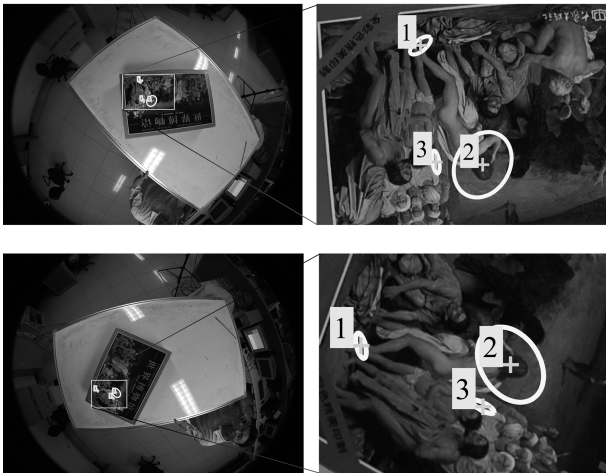


图 3 鱼眼图像对其初始匹配

Fig. 3 A pair of fish-eye images and their initial matches

这 3 对种子匹配经过扩散后, 得到 312 325 对稠密匹配, 用时 1 211 s, 平均每对匹配用时 3.877 ms. 将稠密匹配重采样后得到 3 052 对准稠密匹配, 匹配结果如图 4 所示. 从图中可以看出, 扩散得到的准稠密匹配比较均匀和准确.

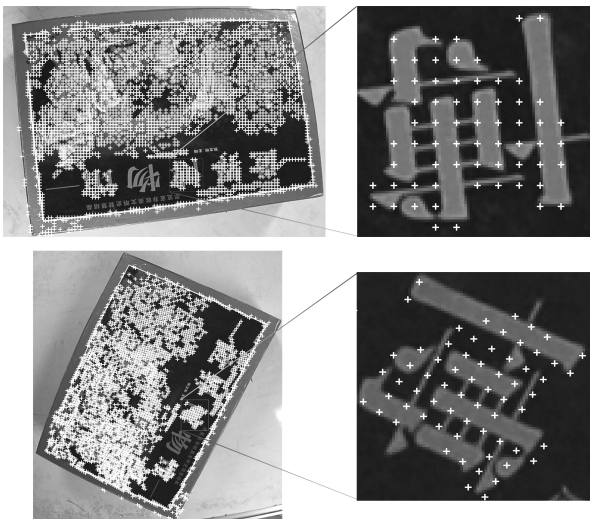


图 4 扩散并重采样后的准稠密匹配 (图中用“+”标出)

Fig. 4 Matches (marked with “+”) obtained via propagation algorithm

用得到的匹配重建出三维点, 然后利用三维点云生成三维网格, 贴上纹理, 得到最后的重建结果.

值得注意的是重建结果没有经过任何后处理以及优化, 是用扩散生成的匹配直接计算得到的. 重建结果见图 5.

从图 5 (b), 5 (c) 和 5 (e) 可以看到, 虽然重建出的三维点绝大部分位于一个平面上 (这与盒子表面是平面相符合), 但是在靠近上下边缘的地方有少部分点误差较大. 仔细观察发现, 误差较大的地方正好是纹理很少的地方, 并且正好沿极线方向, 这样误差较大点不仅不能通过极线约束来剔除. 这些点需要通过后续的算法进一步优化以得到更好的结果.

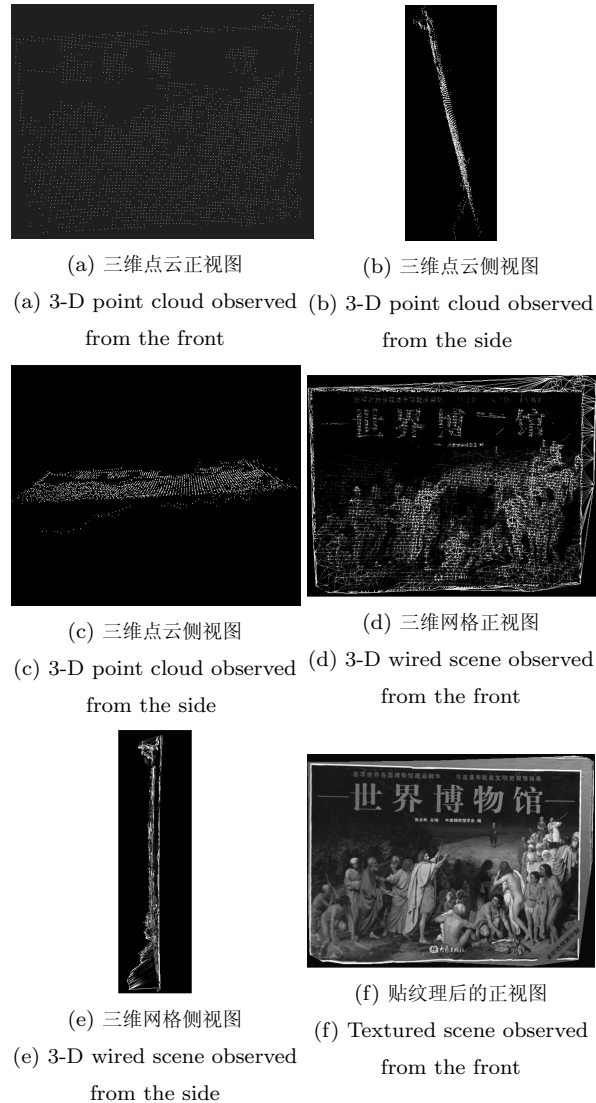


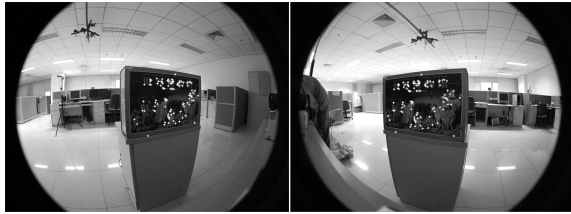
图 5 重建结果

Fig. 5 The reconstructed scene

3.2 透视变化场景

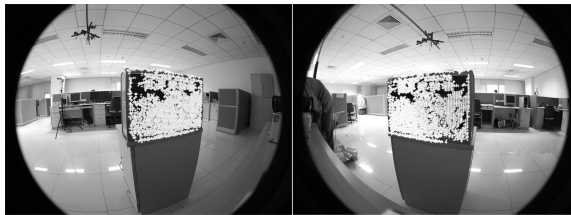
我们在两个不同视角下拍摄同一个盒子, 得到两幅具有较大透视变化的图像, 其中图像分辨率为 3 008 × 2 000. 利用第 2.2.2 节的方法, 采用 SIFT 提取图像对的初始种子匹配并计算其对应邻域的仿射

矩阵. 此处共有 104 对初始匹配, 仿射矩阵计算用时 5.51 s, 平均每对需时 53 ms. 扩散得到 221 922 对稠密匹配, 用时 849 s, 平均每对匹配需时 3.826 ms. 将稠密匹配重采样后得到分布比较均匀的 3 125 对准稠密匹配. 图像的初始匹配以及扩散后的准稠密匹配如图 6 所示.



(a) 鱼眼图像对其初始匹配

(a) A pair of fish-eye images and their initial matches



(b) 扩散并重采样后的准稠密匹配

(b) Matches obtained via match propagation

图 6 扩散前后的匹配

Fig. 6 Matches before and after propagation

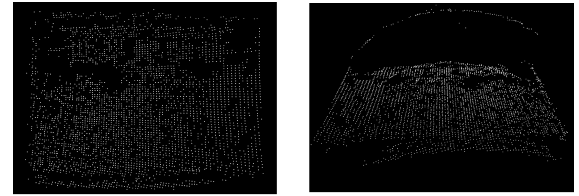
用重采样得到的准稠密匹配重建出三维点, 利用三维点云生成三维网格并贴上纹理, 得到如图 7 所示的重建结果.

从图 7(b) 及 7(c) 中可以看出, 盒子上、下边缘的误差较大, 而中间纹理丰富部分较为准确, 原因与第一个实验中分析的相同.

3.3 复杂场景

这里使用的是一对真实室内场景图 (图像分辨率为 3008×2000), 通过置于天花板顶部的两个鱼眼相机拍摄得到. 因为此处的场景较为复杂, 需要较多的种子匹配点才能获得较好的扩散结果, 所以我们采用 SIFT 来提取初始种子匹配 (参见第 2.2.2 节). 此处一共有 150 对初始种子匹配. 初始匹配仿射矩阵计算用时 10.96 s, 平均每对用时 73 ms. 经扩散得到 1 249 055 对稠密匹配, 用时 4 419 s, 平均每对匹配用时 3.538 ms. 将稠密匹配重采样后得到 12 209 对准稠密匹配. 图像初始匹配及扩散后的准稠密匹配如图 8 所示. 从图 8(c) 可以看出, 本文的扩散算法能得到比较均匀的准稠密匹配.

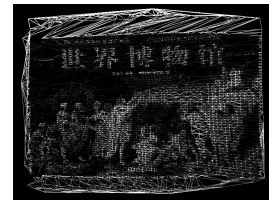
用重采样后的匹配进行重建得到三维点云, 利用点云进行三角化并贴上纹理得到三维效果, 结果如图 9 所示. 从重建的效果看, 本文的扩散方法可以得到比较可靠的匹配结果.



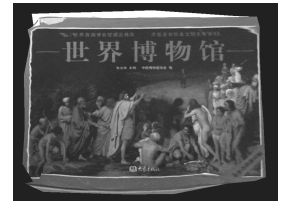
(a) 三维点云正视图

(b) 三维点云侧视图

(a) 3-D point cloud observed from the front (b) 3-D point cloud observed from the side



(c) 三维网格正视图

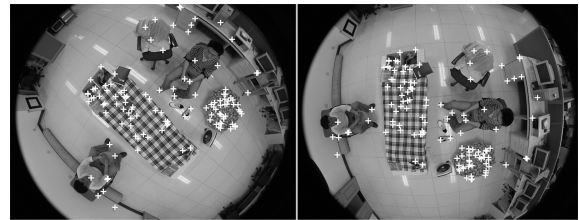


(d) 贴纹理后正视图

(c) 3-D wired scene observed from the front (d) Textured scene observed from the front

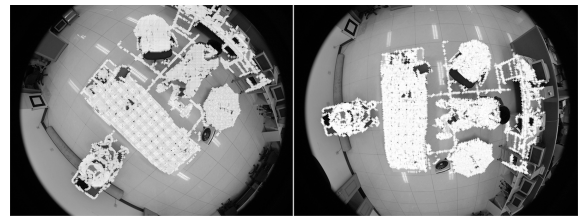
图 7 重建结果

Fig. 7 The reconstructed scenes



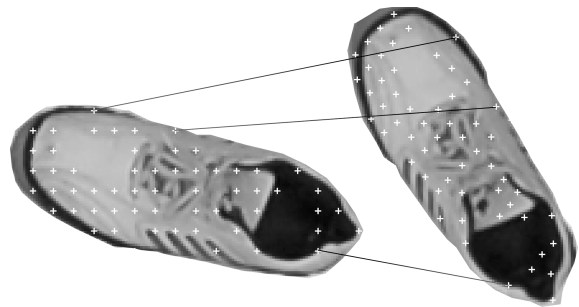
(a) 鱼眼图像对其初始种子匹配

(a) A pair of fish-eye images and their initial matches



(b) 扩散并重采样后的准稠密匹配

(b) Matches obtained via match propagation



(c) 图像局部放大

(c) A closer view of two corresponding local image patches

图 8 扩散前后的匹配

Fig. 8 Matches before and after propagation

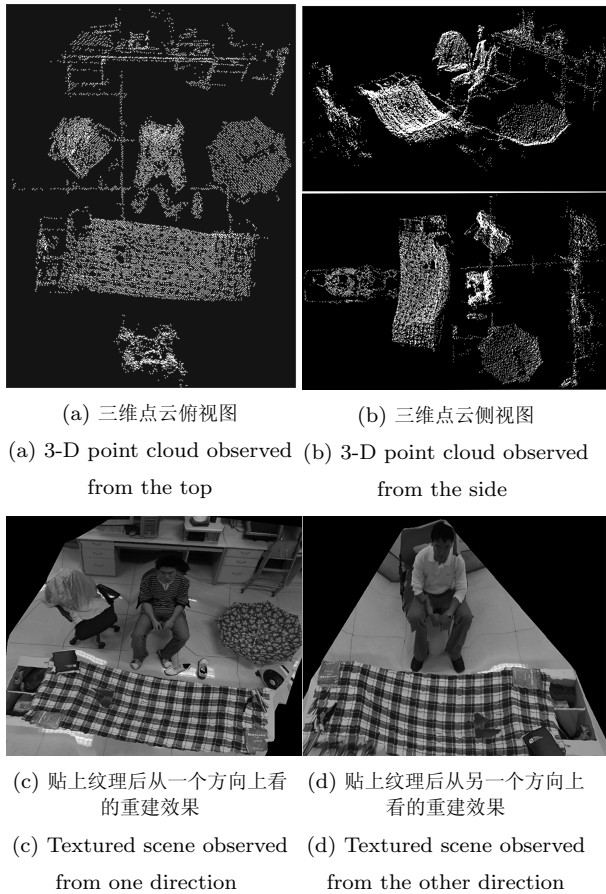


图 9 重建结果

Fig. 9 The reconstructed scenes

从以上的实验可以看出, 本文的扩散算法能够较好地应用于大畸变鱼眼图像对的匹配, 即使在图像之间存在大角度旋转变化、较大程度透视变化的情况下, 都能取得较好效果. 不过从实验 (第 1 个和第 2 个实验) 中也可以看出, 图像中沿极线方向的重复纹理对匹配结果影响较大, 因为重复纹理容易导致仿射矩阵计算不准确. 另外, 需要注意到种子匹配只能沿着平滑变化表面扩散. 这是因为非平滑的变化将导致仿射变换关系突变, 从而使得邻域传递中止, 因此没有种子匹配存在的独立平面上往往不能扩散出新匹配 (如第 3 个实验中地面上的风扇). 所以, 对于有很多平面的场景, 需要增加种子匹配的数量, 使得在每个需要扩散平面上都有种子匹配.

最后需要指出两点: 1) 本文的方法没有与文献中的其他方法进行比较, 这是因为文献中的其他方法都不能直接应用于鱼眼图像的扩散. 如文献 [19–21] 中的方法需要在扩散前将图像校正, 使两幅图像的极线都在水平方向上, 而对大畸变的鱼眼图像而言, 校正本身就非常困难甚至理论上不可行 (如视角大于 180°). 2) 本文仅用三维重建的结果对扩散算法进行了定性评价, 并没有给出定量评价. 这

主要是因为对扩散而言, 有大量扩散其实并没有“真值”, 如何定量评价从理论上来说, 本身就是一个没有解决的问题, 这也是为什么在文献中基本上都没有对扩散匹配定量评价报道的原因. 另外, 对鱼眼图像而言, 边缘部分畸变大, 中间部分畸变小, 对这样的畸变严重不一致的大场景图像进行扩散的定量评价从理论上来说也是一个有待研究的问题.

4 结论

本文研究了在鱼眼图像上的准稠密扩散方法. 本算法采用局部仿射模型来建立图像对应区域之间的变换关系, 通过一种新的仿射矩阵计算方式 (称之为邻域传递), 计算每对匹配对应邻域的仿射矩阵并以此来规范化对应邻域, 最后在规范化的区域上进行匹配扩散. 实验表明, 在畸变较大的鱼眼图像下, 本文方法可以取得比较令人满意的扩散结果. 不过本文方法在计算初始种子匹配的对应邻域时, 本质上使用的是穷举 (Brute force) 的方法, 计算时间较长. 我们下一步的工作将探索更高效的计算匹配点对应邻域的方法.

References

- Harris C, Stephens M J. A combined corner and edge detector. In: Proceedings of the 4th Alvey Vision Conference. Manchester, UK: Springer, 1988. 147–151
- Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
- Matas J, Chum O, Martin U, Pajdla T. Robust wide baseline stereo from maximally stable extremal regions. In: Proceedings of the British Machine Vision Conference. London, UK: Springer, 2002. 384–393
- Mikolajczyk K, Schmid C. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 2004, **60**(1): 63–86
- Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 2002, **47**(1-3): 7–42
- Seitz S M, Dyer C R. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 1999, **35**(2): 151–173
- Kutulakos K N, Seitz S M. A theory of shape by space carving. *International Journal of Computer Vision*, 2000, **38**(3): 199–218
- Kolmogorov V, Zabih R. Multi-camera scene reconstruction via graph cuts. In: Proceedings of the 7th European Conference on Computer Vision. London, UK: Springer-Verlag, 2002. 82–96
- Lhuillier M, Quan L. Match propagation for image-based modeling and rendering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(8): 1140–1146
- Lhuillier M, Quan L. A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(3): 418–433

- 11 Zhang Z Y, Deriche R, Faugeras O, Luong Q T. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 1995, **78**(1-2): 87–119
- 12 Kannala J, Brandt S S. Quasi-dense wide baseline matching using match propagation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 1–8
- 13 Tuytelaars T, Van Gool L. Matching widely separated views based on affine invariant regions. *International Journal of Computer Vision*, 2004, **59**(1): 61–85
- 14 Kadir T, Zisserman A, Brady M. An affine invariant salient region detector. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 345–457
- 15 Ling H B, Jacobs D W. Deformation invariant image matching. In: Proceedings of the 10th International Conference on Computer Vision. Washington D. C., USA: IEEE, 2005. 1466–1473
- 16 Micusik B, Pajdla T. Estimation of omnidirectional camera model from epipolar geometry. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, USA: IEEE, 2003. 485–490
- 17 Micusik B, Pajdla T. Omnidirectional camera model and epipolar geometry estimation by RANSAC with bucketing. In: Proceedings of Scandinavian Conference on Image Analysis (SCIA). Goteborg, Sweden: Springer, 2003. 83–90
- 18 Svoboda T, Pajdla T. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 2002, **49**(1): 23–37
- 19 Megyesi Z, Chetverikov D. Affine propagation for surface reconstruction in wide baseline stereo. In: Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK: IEEE, 2004. 76–79
- 20 Megyesi Z, Chetverikov D. Enhanced surface reconstruction from wide baseline images. In: Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission. Thessaloniki, Greece: IEEE, 2004. 463–469
- 21 Megyesi Z, Kos G, Chetverikov D. Surface normal aided dense reconstruction from images. In: Proceedings of Computer Vision Workshop. Prague, Czech Republic: Springer, 2006. 64–69
- 22 Baumberg A. Reliable feature matching across widely separated views. In: Proceedings of the Conference on Computer Vision and Pattern Recognition. Hilton Head Island, USA: IEEE, 2000. 774–781

- 23 Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981, **24**(6): 381–395



许振辉 中国科学院自动化研究所硕士研究生. 2006 年获浙江大学电气工程及其自动化专业学士学位. 主要研究方向为特征提取和图像匹配. 本文通信作者.
E-mail: sagittarix@gmail.com

(**XU Zhen-Hui** Master student at the Institute of Automation, Chinese Academy of Sciences. He received his

bachelor degree from Zhejiang University in 2006. His research interest covers feature detection and image matching. Corresponding author of this paper.)



张峰 中国科学院自动化研究所博士研究生. 主要研究方向为图像处理 and 三维重建. E-mail: fzhang@nlpr.ia.ac.cn

(**ZHANG Feng** Ph. D. candidate at the Institute of Automation, Chinese Academy of Sciences. His research interest covers image processing and 3D reconstruction.)



孙凤梅 北方工业大学理学院副教授. 主要研究方向为光学和智能信号处理.

E-mail: fmsun@163.com

(**SUN Feng-Mei** Associate professor at the Faculty of Sciences, North China University of Technology. Her research interest covers optics and intelligent signal processing.)



胡占义 中国科学院自动化研究所研究员. 主要研究方向为摄像机标定、三维重建和视觉机器人导航.

E-mail: huzy@nlpr.ia.ac.cn

(**HU Zhan-Yi** Professor at the Institute of Automation, Chinese Academy of Sciences. His research interest covers camera calibration, 3D reconstruction,

and vision guided robot navigation.)