一种基于特征的约束匹配方法

林志泉 胡永健 杨 晖 汪 伟
(华南理工大学自动化科学与工程学院, 广州 510641)

摘 要 提出了一种基于特征的双向视差域约束匹配方法。首先提取图像特征点约简候选匹配像素集,然后进行双向唯一性匹配并在匹配过程中根据双目图像的投影特性设计视差域约束,最后设计了一个滑动平均邻域视差来剔除视差值偏差太大的匹配对。实验结果表明,该方法简单快速,匹配成功率高,可处理复杂场景,具有良好的实用价值。

关键词 立体匹配 特征匹配 视差 投影
中图法分类号: TP391.4 **文献标识码**: A **文章编号**: 1006-8961(2007)11-2104-05

A Feature-Constrained Matching Method

LIN Zhi-quan, HU Yong-jian, YANG Hui, WANG Wei
(*Collage of Automation Science and Engineering, South China University of Technology, Guangzhou* 510641)

Abstract A bidirectional disparity area feature-constrained match method is proposed in this paper. In this method, after feature points of two images extracted, disparity area constraint is designed to reduce candidate match point sets during the process of bidirectional uniqueness constraint correspondence according to the binocular projection characteristics. Then the slipping neighborhood disparity is designed to reject false match points. Experimental results show that the proposed method can be easily implemented for processing complicated scenes and get high precision matching quickly.

Keywords stereovision matching, feature correspondence, disparity, projection

1 引 言

机器人视觉技术最困难的问题就是找出每对图像点之间的对应关系,即解决立体匹配问题。由于噪声及透视失真等因素的影响,对匹配的估计是一个很困难的问题。国内外学者在立体匹配算法上已经做了大量的工作,但是寻找更为稳定的算法仍然是大家共同追求的目标。目前比较有代表性的立体匹配算法可分为基于特征[1,2]、基于区域[3]和基于相位[4]3 类方法。基于特征的匹配就是将参考图像对上分别由同一关键场景点投影所得像素点的对应关系找出来,特征点通常是图像中具有各向灰度梯度极值或者作为图像边缘线交点的角点或孤立点,因为特征点数量较少,所以匹配速度较快,但是得到的匹配对数也有限。基于区域的匹配法主要是依据左右图像对应点邻域内的灰度的相关性来进行匹配,特点是算法简单易于实现,并且能够得到浓密的视差图,但是由于要计算相关,得到的匹配点集还可能存在着错误匹配和多重匹配,需要对点集进一步约简。基于相位的匹配方法主要优点是计算能够以并行方式高效实现,但是相位对各种扰动较为敏感,一般只能得到景物的粗糙结构。

本文结合区域相关性匹配,提出了一种基于特征的双向视差域约束匹配方法。该方法首选利用 Harris 算子提取特征点约简候选匹配点集,提取时调整参数尽可能得到较多的特征点,然后进行双向唯一性匹配并在匹配过程中根据双目图像的投影特性设计视差域约简候选匹配点,最后设计了一个滑动平均邻域视差对匹配结果中视差偏差太大的匹配对进行剔除,进一步提高了匹配的准确率。整个匹配过程时间短,匹配准确率高,可满足实时性的

基金项目:国家自然科学基金项目(60572140);广东省自然科学基金项目(04020004)
收稿日期:2006-06-29;改回日期:2006-08-23

第一作者简介:林志泉(1976~),男,讲师。华南理工大学自动化科学与工程学院在读硕士研究生。主要研究方向为机器视觉、图像处理和数字水印技术。E-mail: zqlin@163.com

要求。

2　图像匹配问题的数学描述

假设在双目匹配中,有由同一场景在存在有限相对偏移和旋转的相机模型下投影产生的两幅不同参考图像为 F_a 和 F_b,图像对应的特征点集分别为 $\boldsymbol{P}^a$ 和 $\boldsymbol{P}^b$:

$$\boldsymbol{P}^a = \{p_i^a = (x_i^a, y_i^a) \mid i = 1, \cdots, N^a\}$$

$$\boldsymbol{P}^b = \{p_j^b = (x_j^b, y_j^b) \mid j = 1, \cdots, N^b\}$$

其中,p_i^a、p_j^b 分别为图像 F_a、F_b 的特征点,(x_i^a, y_i^a)、(x_j^b, y_j^b) 为特征点在各自 2 维图像坐标系下的坐标值,N^a、N^b 为 F_a、F_b 上各自的特征点数。对 F_a、F_b 上由同一类特征点提取算法获得的特征点集 $\boldsymbol{P}^a$ 和 $\boldsymbol{P}^b$,这两个集合之间存在着弱相关关系[5],即从两集合在投影对象空间中对应的场景点角度而言,$\boldsymbol{P}^a \cap \boldsymbol{P}^b \neq \varnothing$。基于特征匹配算法的目标就是找出点集 $\boldsymbol{P}^a$ 和 $\boldsymbol{P}^b$ 中所有由同一场景点产生的特征点对,建立大小为 $N^a \times N^b$ 的匹配矩阵:

$$\boldsymbol{M} = \{m_{ij} \mid m_{ij} \in (0,1), i = 1, \cdots, N^a; j = 1, \cdots, N^b\}$$

m_{ij} 表示特征点 p_i^a 和 p_j^b 的匹配关系,如果特征点 p_i^a 和 p_j^b 匹配,那么 $m_{ij} = 1$,否则 $m_{ij} = 0$,按照匹配对的唯一性,有

$$\sum_{i=1}^{N^a} m_{ij} \leqslant 1 \text{ 且 } \sum_{j=1}^{N^b} m_{ij} \leqslant 1$$

3　图像匹配的相似度描述

图像匹配的相似度可用两点像素值之间的相关系数进行描述,本文采用的是规范化相关系数。计算前先对图像进行规范化处理,这样可以减少因图像光强不一致而带来的误差。然后再计算一幅图像以某一特征点为中心的一个小窗口内的所有像素灰度与另一幅图像中各个潜在对应特征点为中心的同样的小窗口内的像素灰度之间的相关值,具有最大相关系数的特征点就是匹配特征点。

若 p_i^a、p_j^b 像素值分别为 $f_a(x_i, y_i)$ 和 $f_b(x_j, y_j)$,窗口 S 大小为 $(2m+1) \times (2n+1)$ 且以两特征点为中心,则 p_i^a、p_j^b 的规范化相关系数 $r(i,j)$ 定义为

$$r(i,j) = \frac{\sum_{u=-m}^{m}\sum_{v=-n}^{n}[\bar{f}_a(x_i+u, y_i+v)][\bar{f}_b(x_j+u, y_j+v)]}{\left\{\sum_{u=-m}^{m}\sum_{v=-n}^{n}[\bar{f}_a(x_i+u, y_i+v)]^2 \sum_{u=-m}^{m}\sum_{v=-n}^{n}[\bar{f}_b(x_j+u, y_j+v)]^2\right\}^{1/2}}$$

其中,$\bar{f}_a(x_i+u, y_i+v)$、$\bar{f}_b(x_j+u, y_j+v)$ 分别是 $f_a(x_i, y_i)$ 和 $f_b(x_j, y_j)$ 规范化处理后的结果,即

$$\bar{f}(x+u, y+v) = f(x+u, y+v) - \frac{1}{(2m+1) \times (2n+1)} \sum_{u=-m}^{m}\sum_{v=-n}^{n} f(x+u, y+v)$$

计算出的规范化相关系数值的大小反映了两个特征点的相似度,因为同一场景点在两个图像上投影点的邻域内灰度值相似性会较大,反之则会较小。但实际上由于受各种干扰和光强等的影响,即使是严格找到了同一场景点的投影点,计算出来也几乎是不可能达到理想的最大值。为此,可以设定一个相关系数的阈值 T_r,满足 $r(i,j) \geqslant T_r$ 时,认为匹配成功,可用下式描述匹配矩阵系数:

$$m_{ij} = \begin{cases} 1 & r(i,j) \geqslant T_r \\ 0 & \text{其他} \end{cases}$$

其中,$i = 1, \cdots, N^a$; $j = 1, \cdots, N^b$。即使把 T_r 设得比较大,出现误匹配情况还是比较多的,当然提高 T_r 值的同时,有可能把一些正确匹配对当成非匹配对,降低 T_r 值则可能把非匹配对当成正确匹配对,因此单纯靠相似度很难达到理想的正确匹配率。

4　基于特征的匹配算法原理

4.1　特征提取

从上节可知,直接在全局图像域内计算各像素对的相似度判定匹配很难达到理想的正确匹配率,而且计算量大。因为实际应用中的机器视觉技术在很多情况下不需要知道每一像素对的匹配情况,而只需知道目标场景的特征点匹配情况,因此,可对待匹配像素集进行约简。特征点的提取算法有很多,Tissainayagam 等人[6]评估了不同算法提取出的特征点,认为 Harris 等人[7]在 1988 年提出了 Harris 算子具有优良的特性,本文采用了 Harris 算子提取特征点。在提取特征点时,凡满足 Harris 算子 $H(x,y)$ 大于某一阈值 T 的像素点均被认为是特征点,阈值 T 依赖于实际图像的属性如尺寸、纹理等。$H(x,y)$ 定义如下:

$$H(x,y) = Det(\boldsymbol{C}) - k \cdot Trace^2(\boldsymbol{C})$$

$$\boldsymbol{C} = G(\sigma) \times \begin{bmatrix} F_x^2 & F_{xy} \\ F_{xy} & F_y^2 \end{bmatrix}$$

$$Trace^2(\boldsymbol{C}) = F_x + F_y$$

其中,F_x、F_y 分别表示图像的灰度在 i,j 方向的偏导,

$G(\sigma)$表示以 σ 为系数的高斯模板，k 通常取 0.04。

4.2 根据双目投影特性设计的视差域约束

通过提取图像对特征点，第 1 步约简了待匹配点集，现在进行第 2 步约简待匹配点集。对于同一场景点在图像 F_a 和图像 F_b 的投影位置有一个小偏差，这个偏差的范围可用视差域描述，在图像坐标系下，可定义视差域 $D(l,k)$ 是以待匹配点为中心，大小为 $(2l+1)\times(2k+1)$ 的一个矩形域。在移动机器人的双目匹配中，立体图像对经过校准后，待匹配对理论上应位于同一极线上，无垂直方向的偏差，即 $k=0$，这时视差域包含的特征点集是在长度为 $2l+1$ 线段上的候选匹配特征点集。在实际中，k 很难为 0，但 k 值可取近似为 0。这样，在视差域约束下，候选匹配点数减少，需要计算的相关系数就大大减少了。下面的匹配方法会根据双目投影特性进一步约简候选匹配点，降低误匹配，提高匹配的准确率。

(1) $F_a\rightarrow F_b$ 匹配过程

假定左目图像为 F_a，右目图像为 F_b，两幅图像均经过校准。由成像规律可知，同一场景点在图像 F_b 中的投影相对图像 F_a 中的投影偏右，根据这一投影特性，可设计不同方向上的视差域。令 p_i^a 与 p_j^b 分别是图像 F_a 和 F_b 上的一对正确匹配对。如图 1(a)所示，选取图像 F_a 的特征点 p_i^a，其在图像 F_b 对应的位置为像素点 p'^a_i(不一定是图像 F_b 的特征点)。因为在图像 F_b 对应与 p_i^a 匹配的特征点 p_j^b 只可能出现在 p'^a_i 的右方，因此可定义 p_i^a 的视差域 $D_i^b(l,k)$ 只在 p'^a_i 的右方出现，如图 1(a)阴影部分所示。校正方向后的视差域大小缩小了一半，即从 $(2l+1)\times(2k+1)$ 缩小到 $(l+1)\times(2k+1)$。实际根据投影特性校正后的视差域可参看图 2 的第 2 组图，图中 $D_i^b(l,k)$ 准确地包含了 p_i^a 的正确匹配特征点 p_j^b，而且 $D_i^b(l,k)$ 内仅有 3 个候选匹配特征点，这样便舍弃了 p'^a_i 左方的视差域，进一步约简了候选匹配点集，这对减少计算量和降低误匹配率是很有意义的。

令 p_i^a 在视差域 $D_i^b(l,k)$ 内的候选匹配特征点集为

$$P_i^{ab}=\{p_{in}^{ab}=p_n^b \mid n=j_q(i),q=1,2,\cdots,N_i^b\}$$

其中，$i=1,\cdots,N^a$，N_i^b 为 $D_i^b(l,k)$ 内特征点数，$j_q(i)$ 是 $D_i^b(l,k)$ 内特征点的下标，$j_q(i)\in\{1,\cdots,N^b\}$。进行匹配的时候，分别求取 p_i^a 与点集 P_i^{ab} 内特征点的相关系数 $r(i,n)$，若相关系数的最大值大于或等于阈值 T_r，则认为 $F_a\rightarrow F_b$ 方向的匹配成功。为

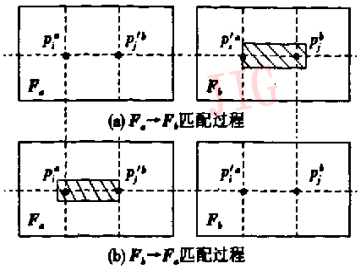


图 1 视差域约束

Fig. 1 Disparity area constraint

$F_a\rightarrow F_b$方向的匹配过程建立一个大小为 $N^a\times N^b$ 的匹配矩阵

$$M^a=\{m_{ij}^a \mid m_{ij}^a\in(0,1),i=1,\cdots,N^a;j=1,\cdots,N^b\}$$

匹配系数 m_{ij}^a 可用下式求取：

$$m_{ij}^a=\begin{cases}1 & \max\limits_j(r(i,n))\geqslant T_r\\ 0 & \text{其他}\end{cases}$$

其中，r 为规范化相关系数，$n=j_q(i)$，$q=1,2,\cdots,N_i^b$；$i=1,\cdots,N^a$。

(2) $F_b\rightarrow F_a$ 匹配过程

这个匹配过程上的视差域校正方向与上一步匹配过程正好相反，因为这时的投影特性是：同一场景点在图像 F_a 中的投影相对图像 F_b 的投影偏左。这样若选取图像 F_b 的特征点 p_j^b，对应在图像 F_a 的位置为像素点 p'^b_j(不一定是图像 F_a 的特征点)，则可定义视差域 $D_j^a(l,k)$ 在 p'^b_j 的左方，如图 1(b)阴影部分所示，校正方向后的视差域大小同样也从 $(2l+1)\times(2k+1)$ 缩小至 $(l+1)\times(2k+1)$。同理可得到 p_j^b 在视差域 $D_j^a(l,k)$ 的特征点集

$$P_j^{ba}=\{p_{jm}^{ba}=p_n^a \mid n=i_q(j);q=1,2,\cdots,N_j^a\}$$

其中，$j=1,\cdots,N^b$，N_j^a 为 $D_j^a(l,k)$ 内特征点数，$i_q(j)$ 是 $D_j^b(l,k)$ 内特征点下标，$i_q(j)\in\{1,\cdots,N^a\}$。匹配求取方法同 $F_a\rightarrow F_b$ 过程相同。为 $F_b\rightarrow F_a$ 方向的匹配过程建立一个大小为 $N^b\times N^a$ 的匹配矩阵

$$M^b=\{m_{ji}^b \mid m_{ji}^b\in(0,1),i=1,\cdots,N^b;j=1,\cdots,N^a\}$$

匹配系数 m_{ji}^b 可用下式求取：

$$m_{ji}^b=\begin{cases}1 & \max\limits_i(r(j,n))\geqslant T_r\\ 0 & \text{其他}\end{cases}$$

其中，r 为规范化相关系数，$n=i_q(j)$，$q=1,2,\cdots,N_j^a$，$j=1,\cdots,N^b$。因为图像中的每个特征点最多只能与另一个图像的一个特征点匹配，即两幅图像中的匹配唯一，因此若两个方向的匹配均一致，则认为

匹配成功。综合以上两个方向的匹配结果,去除多义匹配(匹配不唯一),匹配系数可计算如下:

$$m_{ij}=\begin{cases}1 & \text{如果 } m_{ij}^{a}=m_{\mu}^{b}\\ 0 & \text{其他}\end{cases}$$

其中,$i=1,\cdots,N^a$;$j=1,\cdots,N^b$。

4.3 滑动平均邻域视差约束

为了剔除误匹配,算法最后再对匹配结果设计了一个滑动平均邻域视差进行约束,目的是去除视差偏差太大的匹配对,可有效地提高匹配的准确率。定义滑动平均邻域视差是在以 p_i^a 为中心的窗口 W_i 内所有 N_i^w 对匹配对的平均欧氏距离(视差),即

$$\bar{d}_i=\frac{1}{N_i^w}\sum_{n=1}^{N_i^w}d_n=\frac{1}{N_i^w}\sum_{n=1}^{N_i^w}\|p_n^a-p_n^b\|$$

设定邻域视差阈值 $D_i=\omega\times\bar{d}_i$,$w$ 为预设的常系数,可由实验确定。当匹配对的视差超出视差阈值 D_i 时,去除该匹配对。窗口对每一对匹配对均进行滑动计算,得到最终匹配矩阵 M 的系数:

$$m_{ij}=\begin{cases}1 & d_i<D_i\\ 0 & \text{其他}\end{cases}$$

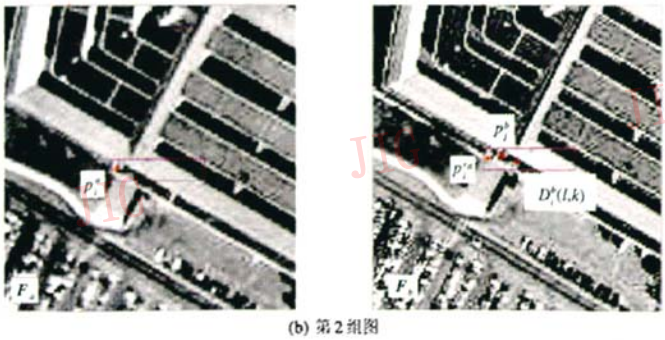
其中,$i=1,\cdots,N^a$;$j=1,\cdots,N^b$。

5 实验结果与分析

为了评价本文算法的性能,对多组室内和室外标准测试图像进行了测试。算法在一台配置为 2.4GHz Pentium CPU,256M 内存,操作系统为 Windows XP 的微机上用 MATLAB 实现,特征点匹配的准确率都达到或超过 93%,匹配速度是区域相关法的 3 倍,实验结果令人满意。

图 2 的第 1 组图是 512×512 的 Pentagon 室外 8 位灰度图匹配结果(用"O"表示),第 2 组图显示了 Pentagon 图像 $F_a\rightarrow F_b$ 方向一对特征点的匹配过程,"+"号处表示特征点对(p_i^a 和 p_j^b),用"O"包围住的特征点表示找到的匹配特征点(p_j^b 点)。从第 2 组图像的右图可观察到,根据 $F_a\rightarrow F_b$ 匹配方向投影特性设计的视差域 $D_i^b(l,k)$ 准确包含了 p_i^a 的正确匹配特征点 p_j^b(这里取 $l=15,k=3$),而且视差域内的候选匹配特征点数量是相当少的(根据特征点的疏密而不同,一般在 0~5 个范围内),这里有 3 个候选匹配特征点,表明有效地减少了计算量和误匹配。匹配最后提取出 324 对匹配对,其中 307 对正确匹配,匹配的正确率为 94.7%,第 1 组图的圆点"O"表

(a) 第 1 组图



(b) 第 2 组图

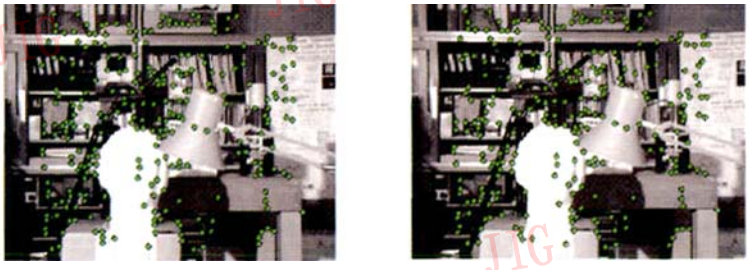
图 2　Pentagon 图的匹配结果

Fig. 2　Correspondence results of pentagon images

示了最终的匹配结果。采用区域匹配方法计算时需执行时间为 5.521s，而本文算法执行时间为 1.846s，明显提高了匹配速度。

图 3 的第 1 组图是 84×288 的 Tsukuba 室内彩色图，对彩色图先转换成灰度图再进行处理。对 Tsukuba 图，实验最后提取出 267 对匹配对，其中 252 对正确匹配，匹配的正确率为 94.3%，匹配结果如图 3 的第 2 组图所示。

(a) 第 1 组图



(b) 第 2 组图

图 3 Tsukuba 图的匹配结果

Fig. 3 Correspondence results of tsukuba images

6 结 论

区域相关的匹配方法可以得到整体的匹配，而且在自身纹理表现为小边缘的情况下可匹配很好，但计算量大，运算速度太慢。相位匹配在奇异点附近的不稳定性导致其也不能成为一种完美的匹配方法。本文正是在深入分析现有的立体匹配算法和机器视觉技术的应用特点的基础上，介绍了一种新的基于特征的匹配算法，可以大大提高匹配的效率。算法在匹配过程中根据双目图像的投影特性设计视差域约简候选匹配点，在算法的最后还设计了一个滑动平均视差邻域计算式来剔除视差值偏差太大的错误匹配对。从实验结果可以看出，本文算法结构简单，稳定可靠，在匹配精度和匹配效率方面都具有比较高的性能，可以处理复杂场景，具有较高的实用价值。

参考文献(References)

1 You Su-ya, Liu Jian, Wan Fa-guan. A two-stage stereo matching algorithm based on neighborhood support criterion [J]. Signal Processing, 1994, 10(2): 118 ~ 123. [游素亚, 柳健, 万发贯. 一种基于邻域支持准则的双层立体匹配算法[J]. 信号处理, 1994, 10(2): 118 ~ 123.]

2 Kim D H, Yoon Y I, Choi J S. A quasi-feature based image mosaic algorithm using modified block matching criteria[J]. IEEE of Japan, 2002, 122C(5): 892 ~ 898.

3 Zhou J, Shi J Y. A robust algorithm for feature point matching[J]. Computer & Graphics, 2002, 26(3): 429 ~ 436.

4 You Su-ya, Liu Jian, Xu Guang-you. Utilizing the ability of human visual phase discrimination for stereo vision [J]. Acta Electronica Sinica, 1996, 24(10): 72 ~ 75. [游素亚, 柳健, 徐光佑. 利用视觉相位鉴别能力求解立体视觉匹配[J]. 电子学报, 1996, 24(10): 72 ~ 75.]

5 Zhen Z, Wang H, Teoh E K. Analysis of gray level corner detection [J]. Pattern Recognition Letters, 1999, 20(2): 149 ~ 162.

6 Tissainayagam P, Suter D. Assessing the performance of corner detectors for point feature tracking applications[J]. Image and Vision Computing, 2004, 22(8): 663 ~ 679.

7 Harris C, Stephens M. A combined corner and edge detector[A]. In: Proceedings of the 4th Alvey Vision Conference [C], Manchester, England, 1988: 147 ~ 151.