

基于奇异值分解的宽基线图像匹配算法

岳思聪 郑江滨 赵荣椿

(西北工业大学 西安 710072)

摘要 图像匹配是计算机视觉中许多应用研究的基础。窄基线匹配技术虽然较为成熟,但是解决能力有限,不能处理较大的尺度、旋转、亮度以及仿射变化引起的宽基线图像序列的匹配。针对宽基线图像序列匹配的特点,在分析传统SVD匹配算法不足的基础上,引入具有尺度和旋转不变性的特征,改进邻近矩阵的度量方式,设计了一种新的基于奇异值分解的宽基线自动匹配算法。通过对比实验表明,该算法性能优于基于SIFT距离的匹配器和原SVD匹配算法,对于存在较大的尺度、旋转、亮度等宽基线变化的图像序列,能够自动获得更多的正确匹配点对和较高的准确性,鲁棒性强,甚至对视角和仿射变换也有一定的适应性。

关键词 图像匹配,特征点对应,奇异值分解

Wide-baseline Image Correspondence Based on Scale and Rotation Invariant Feature

YUE Si-cong ZHENG Jiang-bin ZHAO Rong-chun

(Northwestern Polytechnical University, Xi'an 710072, China)

Abstract Image correspondence is a key problem in computer vision. Although many matching technique in short-baseline have been developed, the wide-baseline correspondence problem with large scale, rotation, illumination and affine transformations is still not tackled very well. The paper proposed a new SVD matching method to achieve large number of accurate point correspondences between uncalibrated image sequences of the same scene for wide baseline. Normalized cross correlation between scale and rotation invariant features was introduced and the proximity matrix was redefined to improve the performance of robustness and reliability. Experimental results show that the proposed SVD matching method can be used for severe scene variations and provide evidence of improved performance with respect to the SIFT distance matcher and the previous SVD matching algorithm.

Keywords Image matching, Feature point correspondence, Singular value decomposition

1 引言

图像匹配是计算机视觉中的核心研究课题,在基础矩阵估计、图像拼接、运动分析与跟踪、目标分类与识别以及三维结构重建等应用研究中都得到了运用。它通过在三维场景的多幅图像之间准确建立起图像点的对应关系,为后续的研究和应用奠定了基础。同时,图像匹配又是计算机视觉研究中的一个瓶颈问题,目前许多重要的理论研究都是在假设宽基线匹配问题已经解决的前提下展开的,因而实现精确的宽基线自动匹配成为当前计算机视觉中的研究热点。

解决图像对应点匹配问题的方法主要有:1)基于图像窗内灰度信息的方法,如最小方差算法^[1]、基于确定性退火的算法^[2]等;2)基于特征比较的算法,该方法是在若干假设或约束下通过获取点的局部或全局特征,如空间分布特征、变换域特征等,而后通过比较特征向量获取对应关系^[3-6]。基于灰度的匹配算法对噪声敏感,区分能力差,不能适应宽基线变换的匹配,因此基于特征的匹配技术研究得到了更多的关注。

用于图像匹配的不变量特征,可分为全局特征和局部特征。全局特征如互信息、Hausdorff距离、Hu不变矩等,这些

方法虽然经典,但它们要在图像的全部空间内进行搜索,计算量太大,而且不能反映局部特征。局部特征的提取则分为两步:提取兴趣点和计算特征描述符。其中以Lowe提出的SIFT特征最具代表性^[7],它对于尺度、旋转、噪声等都具有不变性,甚至对亮度、仿射变换也具有一定的不变性,在存在噪声和遮挡情况下也可鲁棒地识别目标。Mikolajczyk^[8]比较了各种局部特征。实验表明,对于尺度、旋转、噪声、仿射等变化,基于SIFT特征的距离匹配器的性能要优于其它局部特征。

SVD匹配方法最初由Scott^[9]引入到图像匹配领域,通过构造图像特征点之间的邻近矩阵,并进行SVD分解来获得对应关系;后来Pilu^[10]提出在构造邻近矩阵时加入相似性度量的约束,以提高算法的鲁棒性。SVD算法具有简单高效的优点,不需要显性设定阈值参数即可得到对应关系,而且算法本身包含了最近邻、排异性和相似性的匹配3原则。但是在宽基线条件下,这些SVD匹配算法的性能明显下降,其原因主要是所采用特征不具有抵抗尺度、旋转、仿射等变换的能力,以及邻近矩阵中度量方式的选择和表达形式的问题。本文针对SVD匹配方法的不足,采用旋转、尺度不变的SIFT

特征,重新设计邻近矩阵的度量方式,实现了鲁棒的宽基线条件下的自动匹配。实验表明,本算法能够获得更多的正确匹配点对,对宽基线有更好的鲁棒性能,同时避免了阈值参数选择的难题。

2 SVD 匹配方法

Scott 和 Longuet-Higgins^[9] 最先将奇异值分解引入到图像匹配之中,他们利用奇异值分解的特性,提出了一种简单而又直接的 SVD 匹配方法,既满足最近邻匹配原则,又满足了排异性匹配原则,即一对一原则。

设 I_1 和 I_2 为同一场景图像序列中的两幅不同的图像,分别包含 m 和 n 个特征点 $A_i (i=1, \dots, m)$ 和 $B_j (j=1, \dots, n)$ 。SVD 匹配算法分为 3 个步骤:

1) 建立邻近矩阵 $G \in M_{m,n}$ 。

$$G_{ij} = e^{-r_{ij}^2 / 2\sigma^2} \quad i=1, \dots, m, j=1, \dots, n \quad (1)$$

矩阵采用了高斯权重函数, $r_{ij} = \|A_i - B_j\|$ 表示特征 A_i 和 B_j 之间的欧式距离,由此得到的邻近矩阵 G 是正定的。该邻近矩阵是整个匹配算法的核心,能够起到点对间匹配强度的度量作用。在 Scott 的 SVD 匹配中,直接使用了特征点之间的平面坐标距离作为度量,随着点对之间距离的增加, G_{ij} 的取值从 1 到 0 单调下降。参数 σ 则控制着匹配点对之间距离变化的范围,较小的 σ 产生的匹配必然在局部小邻域内,而较大的 σ 则允许匹配点对之间有较大的距离变化。

2) 对邻近矩阵 G 进行奇异值分解。

$$G = UDV^T \quad (2)$$

3) 计算匹配矩阵 P 。将式(2)分解得到的对角矩阵 D 中非零对角元 D_i 置为 1,记新矩阵为 E ,由下式计算得到匹配矩阵 P :

$$P = UEV^T \quad (3)$$

由上式可以看出,匹配矩阵 P 与 G 大小相同,但是经过计算后的矩阵 P 已经隐含了匹配信息。如果 P_{ij} 既是 i 行的最大值,又是 j 列的最大值,那么表示特征 A_i 和特征 B_j 匹配成功,否则 A_i 和 B_j 匹配失败。文献[9]证明了匹配矩阵 P 的行向量是互相正交的,而且奇异值分解能够使 $P^T G$ 的迹取得最大值,从而有效地找出最小均方距离意义下的点对映射关系。

但是由于图像坐标之间的距离度量有很大的局限性和不稳定性,导致 SVD 方法在实际图像匹配中失败。Pilu^[10] 在此基础上,采用以特征点为中心的图像窗为特征,计算点对之间的归一化相关度量,对邻近矩阵 G 进行加强,使邻近矩阵 G 的元素包含两个部分:高斯距离度量和相似性度量,从而将相似性匹配原则嵌入到邻近矩阵 G 中。Pilu 方法的邻近矩阵 G 定义为:

$$G_{ij} = \frac{C_{ij} + 1}{2} e^{-r_{ij}^2 / 2\sigma^2} \quad (4)$$

$$C_{ij} = \sum_l \frac{(A_i^l - \text{mean}(A_i))(B_j^l - \text{mean}(B_j))}{l \times \text{stdv}(A_i) \text{stdv}(B_j)} \quad (5)$$

其中 $\text{mean}()$ 是均值, $\text{stdv}()$ 是标准差, C_{ij} 是图像窗之间的归一化互相关度量,取值范围为 $[-1, 1]$ 。因此 G_{ij} 仍然在 $[0, 1]$ 之间取值,两点之间距离越近,局部图像相似度越高,则 G_{ij} 值越大。相似性度量能够消除一些坏点引起的错误匹配,虽然实际匹配效果有一定提高,但是其选取固定大小和方向的图

像窗,非常容易受到噪声、尺度、旋转和仿射变换的影响,使得算法无法适应实际图像序列中的宽基线变化。

3 基于不变量特征的 SVD 匹配

尽管 Pilu 对 SVD 匹配算法进行了改进,但是匹配效果仍然不佳。当存在较大的尺度、旋转和光照变化时,错误匹配会急剧增加,若是存在仿射变形的情况,匹配结果则更差。究其原因,从度量点对匹配强度的邻近矩阵不难发现,它们均使用了图像坐标之间的欧式距离作为邻近矩阵的匹配度量,而且是相似性与距离乘积的度量方式,这就导致邻近矩阵非常容易受到点对间平面距离大小的影响。例如匹配点对 A_i 和 B_{j_1} 的相似度稍大于 A_i 和 B_{j_2} ,但 A_i 和 B_{j_2} 之间距离却远小于 A_i 和 B_{j_1} 之间的距离,这样通过 SVD 分解很可能产生 $P_{ij_1} < P_{ij_2}$,导致 A_i 和 B_{j_2} 的错误匹配。也就是说,造成匹配性能下降的根本原因是所选取的特征和邻近矩阵的度量方式,而不是 SVD 匹配方法自身存在缺陷。同时也说明了邻近矩阵是 SVD 匹配算法的核心。

本文针对以上不足,引入具有尺度旋转不变的 SIFT 特征,用 SIFT 特征的归一化互相关度量取代空间距离度量,而邻近矩阵不采用 Pilu 改进的两种度量相乘的方式,而使用 Scott 算法中单一度量的高斯权重模型,以获得对尺度、旋转和光照具有鲁棒性的图像匹配。

3.1 SIFT 特征

SIFT(scale invariant feature transform)特征描述子^[7]是 Lowe 在总结了现有的不变量特征技术的基础上提出的一种基于尺度空间的、对图像缩放、旋转甚至仿射变换保持不变性的图像局部特征描述算子。生物视觉理论认为,生物大脑视觉皮层中的复杂神经元细胞对图像梯度的方向和空域频率信息响应非常强烈,相反地,对梯度的位置信息敏感度不高。因此 SIFT 特征描述子利用特征点邻域图像窗口内梯度的方向统计直方图来构造特征描述向量,有效提高了特征描述子对图像尺度变换、旋转变换、视角变换和仿射变换的鲁棒性。

SIFT 特征提取包括 4 步:

(1) 尺度空间极值检测,以初步确定关键点位置和所在尺度。通过构造尺度空间的高斯金字塔,并搜索高斯差分图像(DoG)的局部极值作为关键特征点。

(2) 通过拟合三维二次函数以精确确定关键点的位置和尺度,定位精度可以达到亚像素级,同时去除低对比度的关键点和不稳定的边缘响应点,以提高抗噪声能力。

(3) 确定关键点的方向。在关键点所在的尺度上,利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数,使算子具备旋转不变性。实际计算时,在以关键点为中心的邻域窗口内采样,用直方图统计邻域像素的梯度方向。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。

(4) 生成 SIFT 特征向量。首先将坐标轴旋转为关键点的方向,以确保旋转不变性。然后在对应尺度的图像上以关键点为中心取 8×8 的窗口,每 2×2 的小块上计算 8 个梯度方向的直方图,即可形成一个种子点。每个关键点由 4×4 共 16 个种子点来描述,这样对于一个关键点就形成 128 维的 SIFT 特征向量。此时 SIFT 特征向量已经去除了尺度变化、旋转等几何变形因素的影响。这种邻域方向信息联合的思想

增强了算法抗噪声的能力,同时对于含有定位误差的特征匹配也提供了较好的容错性。

3.2 基于 SIFT 互相关度量的 SVD 匹配

SIFT 特征是一种非常有效的局部特征变换, Mikolajczyk 在文献[8]中比较了各种局部特征描述的匹配性能。他指出,对于旋转、尺度、视角等多种变化, SIFT 算子的性能要优于其它特征。这正是宽基线匹配所需的不变性特征。但是 Mikolajczyk 的实验也指出, SIFT 特征仅对线性光照变换具有不变性,对于一般的光照变化比较敏感,其性能明显下降。因此,本文选取归一化互相关度量(NCC)来弥补此缺陷, NCC 既能够度量两个特征向量的相关程度,又可以消除亮度变化的影响。

邻近矩阵 G 中点对间的平面坐标距离度量在理论上是有效的。因为当其他相似条件完全相同时,认为距离上最近的点就是正确匹配且合理的,这满足了理论上的最近邻匹配原则。但是这在实际图像匹配中不一定有效,一是除了距离之外的其他条件完全相同的情况出现的可能性非常小,二是随着基线的增大,正确匹配点对之间的距离变化范围会随之增大,有些理论匹配点对之间很可能出现较大的平面位置差异,这样最近邻原则就不再适用。因此本文使用 SIFT 特征的归一化互相关度量取而代之。此时邻近矩阵 G 表示为

$$G_{ij} = e^{-\frac{1-c_{ij}^2}{2}} \quad i=1, \dots, m, j=1, \dots, n \quad (6)$$

$$C_{ij} = \frac{(S_{1i} - \text{mean}(S_{1i}))(S_{2j} - \text{mean}(S_{2j}))}{l \times \text{std}v(S_{1i}) \text{std}v(S_{2j})} \quad (7)$$

其中 $\text{mean}(S)$ 是 S 的均值, $\text{std}v(S)$ 是 S 的标准差, C_{ij} 是 SIFT 特征之间的归一化互相关度量,它描述了两个点特征向量之间的相似程度。SIFT 特征自身具有尺度和旋转不变性,加之抗噪声强的优点,使得邻近矩阵的匹配度量更加稳定,对宽基线的多种变化具有较强的适应能力,进而提高了整个匹配算法的鲁棒性。

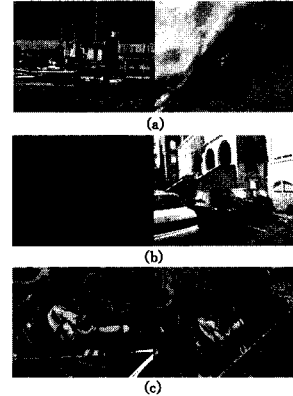
高斯权重函数实质上起到单调映射函数的作用,通过高斯权重模型使得相似度量之间的差异更加明显,有利于正确匹配的生成。基于上式定义的邻近矩阵 G , 对其进行 SVD 分解,按照 2 节中的方法计算匹配矩阵 P , 最终由 P 确定特征点的对应关系。如果第一幅图像中的 A_i 和第二幅图像中的 B_j 匹配,那么 P_{ij} 必须同时满足两个条件。条件(1): P_{ij} 既是 i 行的最大值,又是 j 列的最大值;条件(2): 设 P_{ij} 和 P_k 分别是 i 行的最大值和次大值, P_{ij} 和 P_{hj} 分别是 j 列的最大值和次大值,那么 $0.5P_{ij} > P_k$ 且 $0.5P_{ij} > P_{hj}$ 。

4 实验结果及分析

为了验证本文提出的新 SVD 算法(N-SVD)的性能,我们对包含旋转、尺度(焦距)、光照和仿射等不同基线条件变化的场景序列进行匹配测试,并与文献[8]中的基于 SIFT 距离的匹配器(S-Dist)和 Pilu 的 SVD 匹配算法(P-SVD)进行对比实验。部分实验图像如图 1 所示,所有实验均以每个序列的第 0 帧图像作为参考图像,后续图像与其进行匹配,同时使用鲁棒估计的单应 H 来判断匹配点对是否正确。

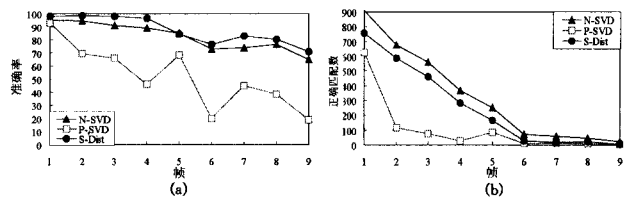
图 2 给出了 Boat 序列的实验结果曲线,该序列由 10 帧图像组成,包含了不同程度的旋转和尺度变换。从图中可以看出,本文算法的正确匹配数均高于 S-Dist 和 P-SVD。尤其当后几帧存在较大的尺度和旋转变换时,本文方法的正确匹配数是 S-Dist 和 P-SVD 的几倍。特别是最后一帧图像, S-Dist 和 P-SVD 仅能获得 5 对左右

许多应用的需要,而本文算法能够得到 24 对正确匹配点对,可见本文算法对于大的尺度和旋转变换有很好的适应能力。就准确性而言, S-Dist 最高,平均达到 90%;本文方法略低,平均 85%左右; P-SVD 则较差,仅有 50%。本文算法在保证准确率和 S-Dist 基本相当的情况下,能够提供更多的正确匹配点对。



(a) 为 Boat 序列的第 0 帧和第 9 帧,包含了较大的尺度和旋转变换;
(b) 图像为 Car 序列的第 0 帧和第 5 帧,包含了较大的亮度变化;
(c) 为 Graffiti 序列的第 0 帧和第 3 帧,包含了较大的视角和仿射变换。

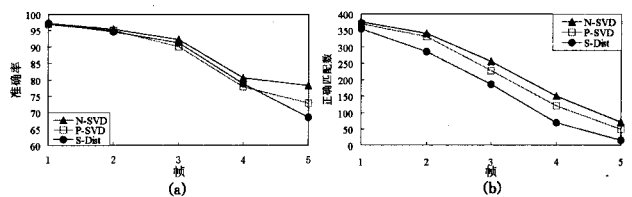
图 1 部分测试序列



Boat 序列包含 10 帧图像,随着帧数增加,尺度旋转变换逐渐增大。
(a) 为 3 种算法的匹配准确率, (b) 为正确匹配数。

图 2 Boat 图像序列实验结果曲线

对于包含亮度变化的 Car 图像序列的对比实验结果曲线如图 3 所示。Car 序列由 6 帧图像组成,亮度由暗到亮。在 3 种方法中,本文算法依然获得了最多的正确匹配数,同时本文方法的准确率也是最高的。与上一组序列的实验结果不同的是,本组中 P-SVD 的性能较 S-Dist 要好,主要是因为 SIFT 特征对于一般亮度变化不具有不变性,而 P-SVD 算法因为运用了归一化互相关度量,所以邻近矩阵的抵抗光照变化的能力得到了增强,因此得到优于 S-Dist 的匹配结果。而在 boat 序列中, SIFT 匹配器对于尺度和旋转具有较好的不变性,因此匹配结果好于使用了图像坐标距离的 P-SVD 方法。



Car 序列包含 6 帧图像,随着帧数增加,亮度逐渐增强。
(a) 为 3 种算法的匹配准确率, (b) 为正确匹配数。

图 3 Car 图像序列实验结果曲线

Graffiti 序列的对比实验结果如表 1 所示。该序列由 6 帧图像组成,由不同的拍摄视角产生出不同程度的仿射变换。

(下转第 265 页)

[10] Cook J, Wolf A. Discovering models of software processes from event-based data [J]. ACM Transactions on Software Engineering and Methodology, 1998, 7(3): 215-249

[11] Reiss S P, Renieris M. Encoding program executions [C]//Proceedings of the 23rd International Conference on Software Engineering. 2001: 221-230

[12] White L, Cohen E. A domain strategy for computer program testing [J]. IEEE Transactions on Software Engineering, 1991, 17:

[13] Sourceforge, sourceforge.net/projects/elevatorsim

[14] Ghosh S, Mathur A. Interface mutation to assess the adequacy of tests for components and systems [C]//Proceedings of Technology of Object-oriented Languages and Systems. IEEE Computer Society, 2000: 37-46

[15] Delamaro M E, Maldonado J C, Mathur A P. Interface mutation: An approach for integration testing [J]. IEEE Transactions on Software Engineering, 2001, 27(3): 228-247

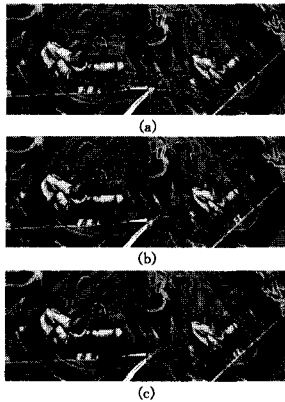
(上接第 225 页)

随着视角和仿射变形的增大, 3 种算法的性能均明显下降。由于最后两帧图像的仿射变形太大, 导致错误匹配过多, 无法鲁棒地估计单应, 因而不能准确判断正确匹配的数目, 因此仅列出前 4 帧的匹配结果。从表中不难看出, 在 3 种方法中, 本文算法依旧获得了最多的匹配特征点对, 说明本文算法对仿射和视角等宽基线匹配的适应能力有一定的提高, 但是仍然不能够抵抗更大的仿射变形。第 3 帧与参考帧之间匹配结果如图 4 所示。

表 1 Graffiti 序列的实验结果

匹配方法	N	P	S	N	P	S	N	P	S
帧	1			2			3		
准确率	0.92	0.88	0.95	0.76	0.37	0.65	0.43	0.08	0.45
正确匹配数	679	406	575	302	64	187	45	5	5

N, P, S 分别代表 N-SVD, P-SVD 和 S-Dist 3 种算法。由于最后两帧的仿射变形太大, 导致 3 种方法都失效, 表中只给出了前 4 帧的匹配结果。



Graffiti 序列包含较大的视角变化, 这两帧图像之间约有 50° 的拍摄视角。从 (a) 到 (c) 依次为 N-SDV, S-Dist, P-SVD 算法的匹配结果。

图 4 Graffiti 序列第 0 帧和第 3 帧的匹配结果

结束语 本文设计了一种基于尺度旋转不变特征的 SVD 宽基线自动匹配算法, 该算法弥补了原 SVD 方法的不足, 利用 SIFT 特征和归一化互相关度量, 增强了邻近矩阵的度量能力, 实现了算法对旋转、尺度缩放和光照变化等因素的不变性, 使得 SVD 算法在宽基线匹配中更加鲁棒。对比实验

表明, 该算法比基于 SIFT 距离的匹配器和 Pilu 改进的 SVD 算法的性能有较大的提高, 对于大的尺度、旋转和光照变化有较强的鲁棒性, 即使对较大的仿射变换也有一定的适应能力。同时, 本算法避免了参数选择的难题, 通过 SVD 分解产生的匹配矩阵能够获得简约的一对一匹配结果, 将相似性和排异性匹配原则自然地融为一体, 简单有效地实现了宽基线自动匹配。

参考文献

[1] Umeyama S. Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(4): 376-380

[2] Chui L, Rangarajan A. A new point matching algorithm for non-rigid registration. Computer Vision and Image Understanding, 2002, 89(2): 114-141

[3] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522

[4] Shapiro S, Brady M. Feature-based correspondence: an eigenvector approach. Image and Vision Computing, 1992, 10(5): 283-288

[5] Zitova B, Flusser J. Image registration methods: a survey. Image and Vision Computing, 2003, 21(11): 977-1000

[6] Carcassoni M, Hancock R. Spectral correspondence for point pattern matching. Pattern Recognition, 2003, 36(1): 193-204

[7] Lowe D. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110

[8] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630

[9] Scott G, Longuet-Higgins H. An algorithm for associating the features of two patterns//Proceedings of the Royal Society-Biological Sciences (Series B). London, UK, 1991, 244(1309): 21-26

[10] Pilu M. A direct method for stereo correspondence based on singular value decomposition//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Juan, Puerto Rico, 1997: 261-266