

基于局部不变映射的特征 描述器算法

徐小明¹ 杨丹¹ 张小洪¹ 周小龙¹

摘要 提出了一种新的基于局部不变映射 (Locality preserving projections, LPP) 的描述器设计算法。该算法用 LPP 预先生成一个特征矩阵, 接着把特征点邻域内所有点的梯度组成一个高维的梯度向量, 然后通过特征矩阵将该梯度向量嵌入到一个低维的流形空间中, 生成一个维数很低的向量, 并把它作为该特征点的描述器。所提出的算法能保持描述器之间的几何结构不变: 原空间中邻接的描述器映射到低维空间后保持邻接, 而不相似的描述器映射后区分度更大, 所以该算法所生成的描述器能表现特征点之间的内在关系, 具有很强的鲁棒性。通过与 SIFT (Scale invariant feature transform), PCA-SIFT 的实验比较, 此算法更快速, 更具鲁棒性。

关键词 局部不变映射, 主成分分析, 特征点, 描述器
中图分类号 TP13

The Algorithm of Descriptor Based on Locality Preserving Projections

XU Xiao-Ming¹ YANG Dan¹ ZHANG Xiao-Hong¹
ZHOU Xiao-Long¹

Abstract This paper presents a novel algorithm to design the descriptor of image feature points based on locality preserving projections (LPP). Firstly, an eigenmatrix was pre-produced by LPP, and a high-dimensional gradient vector was constructed by the gradient vectors of all neighborhood points around feature points. Then, the high-dimensional gradient vector was embedded into a lower dimensional manifold space with the eigenmatrix, and a low-dimensional descriptor of the feature points was generated. The proposed algorithm can preserve invariability on the geometric structure: the eigenvectors which are neighboring each other in the original space will maintain the same attribute in low-dimensional space; on the contrary, the unsimilar eigenvectors become apart farther each other. Therefore, the description generated by our algorithm can show the interrelationship between features and has strong robustness. Moreover, the comparative experiments illustrated that the proposed algorithm is more rapid and accurate than SIFT and PCA-SIFT.

Key words Locality preserving projections (LPP), principle component analysis (PCA), feature point, descriptor

图像特征提取和描述是图像模式识别的一个关键技术, 其主要应用的目标识别、图像配准、图像拼接和内容检索等领域。从特征匹配的性能要求来看, 描述器性能的好坏很大程度上决定了该技术的优劣。

目前已经提出了许多描述器设计算法。Mikolajczyk 等在文献 [1-2] 中详细分析了各种描述器的性能。其中, 差分不

变描述器^[3]相对于其他描述器来说, 总体的性能较差; 可调滤波器^[4]和梯度矩^[5]虽然设计简单、维数低、匹配速度快, 但鲁棒性不够; 互相关矩阵表现中等, 但它的缺点是不稳定, 依赖于特征点和特征区域的检测; 其他的描述器如形状上下文^[6], 复系数滤波器^[7]等都对图像的变换不具有很强的鲁棒性。值得注意的是, SIFT (Scale invariant feature transform) 算法^[8]是目前最具鲁棒性的描述器算法。该算法主要由 4 部分组成: 1) 尺度空间极值检测, 在高斯差分尺度空间上检测出特征点; 2) 特征点定位, 这里分两部分: a) 对特征点的亚像素精确定位; b) 去除边缘响应点; 3) 分配主方向, 通过为特征点分配一个或多个主方向, 使得该算法对图像的旋转具有不变性; 4) 生成描述器, 以特征点为中心选择一块区域, 旋转到主方向, 接着将该区域分为 4×4 小块, 同时把梯度方向分为 8 个条目, 最后对小块里每个像素的梯度在对应的梯度方向上进行加权求和, 生成一个 128 维向量。把向量归一化, 以消除光照的影响, 最后的结果就作为该点的特征描述器。

虽然 SIFT 算法鲁棒性强, 但计算复杂性高, 而且由于描述器的维数太大, 导致了匹配速度较慢。基于此, YanKe 和 Rahul 对此算法进行了改进, 提出了基于主成分分析 (Principle component analysis, PCA)^[9] 的 SIFT 描述器 (PCA-SIFT) 算法^[10]。该算法的主要贡献在于通过对大量特征点的邻域梯度进行统计分析, 生成一个低维的特征空间; 然后把特征点的邻域梯度向量映射到这个低维空间中, 生成一个低维的描述器向量, 从而提高了匹配速度。但是由于 PCA 要求样本数据是椭圆分布的, 而且建立的是线性模型, 对于非线性的高维数据效果不好, 准确性较 SIFT 差。因此, 针对以上问题, 本文提出了一种新的描述器设计算法: 局部不变映射 (Locality preserving projections, LPP)^[11] 的描述器设计算法。该算法通过流形学习, 生成一个低维的特征空间, 而且考虑了特征点的邻接点信息, 保持了特征点的内在邻接属性不变和特征点的可区分性。因而, 设计算法在具有维数低, 计算简单等特点的同时, 还保证了图像在发生各种仿射变换时具有很强的鲁棒性和准确性。最后, 通过实验证明了 LPP 算法的有效性。

1 基于局部不变映射的描述器

LPP 描述器设计算法可概括为以下三个步骤: 1) 利用采集的数据预先计算一个特征空间; 2) 对于给定的一个特征点, 计算它邻域内的梯度, 生成一个高维的梯度向量; 3) 把这个梯度向量映射到我们生成的特征空间中, 最终的结果作为该特征点的描述器。

为了使实验结果更准确, 该算法的数据采集方法与 PCA-SIFT 采集方法保持一致: 用 SIFT 算法的前三步提取特征点, 然后以特征点为中心选取 41×41 的区域, 计算特征点邻域内的水平与垂直的梯度, 最后生成 $2 \times 39 \times 39 = 3042$ 维向量。

1.1 特征空间的计算

对特征空间的计算, 本文提出的 LPP 算法不同于 PCA, 而是利用了流形学习的思想。因为对于特定的目标, 它的每个高维特征描述器向量, 本质上位于一个低维的流形空间中, 描述器的内在维数就是相机的自由度, 描述器间存在的也是非线性关系。如图 1 中给出了一个描述器样本集在 PCA 映射下的例子, 从图中可以看出, 原空间中有些分开的描述器映射后会靠的很近, 甚至重合。在这种情况下, 使用 PCA 是不准确的, PCA 只关心样本的最大方差方向。但 LPP 又不同于

收稿日期 2007-08-03 收修改稿日期 2007-12-19
Received August 3, 2007; in revised form December 19, 2007
国家自然科学基金 (60604007), 重庆市自然科学基金 (CSTC2005BA2002) 资助
Supported by National Natural Science Foundation of China (60604007) and Natural Science Foundation of Chongqing (CSTC2005BA2002)
1. 重庆大学软件学院 重庆 400044
1. School of Software Engineering, Chongqing University, Chongqing 400044
DOI: 10.3724/SP.J.1004.2008.01174

典型的流形学习算法, 如 LLE^[12], Laplacian Eigenmap^[13]. 因为它们有一个共同的缺点, 只能定义于样本集中, 不具有延展性, 若要对 N 个描述器进行处理, 就需要解一个矩阵为 $N \times N$ 大小的特征值. 对于需要处理大量描述器的算法来说, 这无疑是不可取的. 而本文所提出的 LPP 描述器算法恰好结合了它们的优点: LPP 是线性变换, 计算简单而且可延展, 适合实际的应用, 这是其他非线性映射所不具备的; LPP 作为 Laplace-Beltrami^[13] 算子的近似, 能保持数据集的几何结构不变. 换句话说, 对于本文的描述器, 经过 LPP 算法处理后, 描述器间的相对关系并没有太多改变. 同时, 在 LPP 算法中, 通过为数据点分配合适的权重, 使得大多数描述器的区分性得到增强. 这使得后面的特征匹配比较准确, 也比较容易. 即通过设定一个阈值, 就可以很好地将它们区分开. 如图 1 中所示, 图中顶部序列 1) 是一个 5 维的数据样本集, 左半部分是通过 PCA 把数据集映射到 2 维空间中的结果. 可以看到, 数据集在映射后关系变的很混乱. 右半部分序列 2) 是为 LPP 选择相应的邻接点, 序列 3) 是 LPP 把数据集映射到 2 维空间中的结果. 可以看到, 数据集在映射后能保持它们之间的关系.

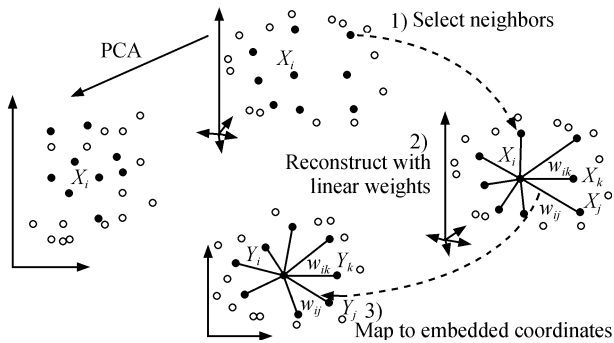


图 1 描述器数据集在 PCA 与 LPP 映射下的结果比较
Fig. 1 Data projected by PCA and LPP

具体的特征空间的计算比较简单, 仅仅需要求解一个稀疏矩阵的广义特征值, 步骤如下:

步骤 1. 采集多幅图片, 本文算法中共选择了 17 幅图片. 包括文献 [1-2] 中标准的 8 幅图片以及自然拍摄的 9 幅图片. 利用 SIFT 前三步, 生成 (我们的算法选择 40 000) 个特征点. 然后在特征点周围提取 41×41 区域, 计算每个像素的水平和垂直的梯度方向, 生成一个维向量, 把它作为特征矩阵的一列.

步骤 2. 为特征点数据集构造一个邻接图 G , G 是一个 $N \times N$ 的矩阵. 其中, G 的构造方法有两个选择: 对于任意的特征点 X_i , 如果特征点 X_j 在 X_i 的 k 邻接点之内, 那么, 就在 X_i, X_j 间输入一条边; 或者, 当 $\|X_i - X_j\| \leq \epsilon$, 那么就在 X_i, X_j 间输入一条边.

步骤 3. 选择特征点之间的权重, 构造特征点权重矩阵 W , 它是一个 $N \times N$ 的稀疏矩阵. W 构造方法有两种选择: 如果特征点 X_i, X_j 相邻, 那么 $W_{ij} = 1$; 否则 $W_{ij} = 0$; 或者如果特征点 X_i, X_j 相邻, 那么 $W_{ij} = e^{-(x_i - x_j) \times (x_i - x_j) / t}$.

步骤 4. 计算广义特征值和广义特征向量

$$X L X^T a = \lambda X D X^T a \quad (1)$$

其中, D 是对角矩阵, D_{ii} 是权重矩阵 W 第 i 列或第 i 行的总和, $D_{ii} = \sum_j W_{ji}$. $L = D - W$ 是拉普拉斯矩阵. X

是特征矩阵, 每一列是一个特征点的梯度向量, 行是特征点的个数. 通过解式 (1) 得到一个特征矩阵 A , 它的每一列是一个特征向量. 并且按特征值从小到大的顺序排列, $\lambda_0 < \lambda_1 < \dots < \lambda_{l-1}$.

需要说明的是, 在步骤 2 中, 算法采用第一种选择, 因为在实际应用中, 对于第二种选择中的参数很难确定, 而且 k 邻接点更能说明描述器的作用, 实验中一般选择 5. 步骤 3 中, 算法选择的是第一种方法, 虽然第二种选择效果要好于第一种, 但第二种选择中的 t 很难确定, 这里通过为邻接点分配一个权重比, 使得映射后的特征点保持相似性的同时又扩大区分性.

1.2 特征描述

有了特征空间后, 接下来要做的工作就是将特征点的梯度向量映射到低维的特征空间中. 步骤和上面类似, 对于一个给定的特征点, 生成一个 3 042 维的高维梯度向量. 为了消除光照的影响, 将其归一化. 然后通过上面求出的特征矩阵 A , 把梯度向量映射到低维的流形中, 映射后的向量就作为特征点的描述器. 在实验中, 特征空间取的维数是 20. 相对于 SIFT 的 128 维向量, LPP 算法在速度上有很大的提高.

$$x_i \rightarrow y_i = A^T x_i, A = (a_0, a_1, \dots, a_{l-1}) \quad (2)$$

式 (2) 给出了映射关系, 其中 y_i 是一个 L 维的向量, 也是该算法中的描述器向量.

2 实验与讨论

2.1 特征点对的匹配准则和评估方法

判断某两个特征点是否相似主要有三种方法: 一是基于欧几里德距离; 二是邻接点距离; 三是第一邻接点和第二邻接点的比值. 在本文中, 采用的匹配准则是基于距离阈值, 因为用 LPP, PCA 映射后的描述器都是在一个低维空间, 所以基于距离阈值产生的结果较为准确.

查全率和查准率在信息检索中是一个非常重要评估方法, 它们的定义如下:

$$\text{查全率} = \frac{\text{正确命中的数量}}{\text{数据库中对应正确命中的总数}}$$

$$\text{查准率} = \frac{\text{正确命中的数量}}{\text{总的命中数量}}$$

不过查全率和查准率很难同时达到最佳, 一般查全率高, 查准率就低, 反之, 查准率高, 查全率就低, 实际中常把它们的平均值作为衡量指标. 在实验中, 要做的是验证基于 LPP 的描述器的有效性, 也就是在保证匹配点对多的同时还要尽可能地保证正确率. 所以在本文中, 借用这个方法来自估描述器的性能, 同时定义它们如下:

$$\text{Recall} = \text{查全率} = \frac{\text{正确匹配数}}{\text{对应的正确匹配总数}}$$

$$1 - \text{precision} = \text{查错率} = \frac{\text{错误的匹配数}}{\text{总的匹配数}}$$

2.2 实验结果

在本文中, 我们选取四组不同的图像来完成对三种算法的评估. 其中 LPP 与 PCA 选取的特征向量是 20 维, SIFT 采取的还是通用的 128 维. 图 2 显示了选择的图像集, 它们分别代表了不同的图像变化: (a) 光照变化; (b) 噪声影响;

(c) 视角变化; (d) 尺度和方向变化. 下面我们对这四组图像分别进行实验和讨论.

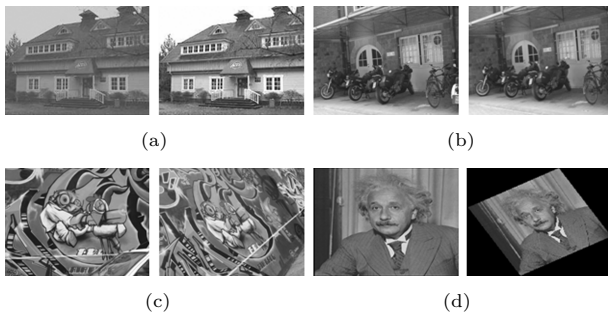


图 2 实验中所用的图像集

Fig. 2 Data set used in experiment

图 3 是图 2 中对应图片的实验结果, 从图 3(a) 可以看出, 三个算法对图像的光照变化都有很好表现. 在查错率小的时候 LPP 算法优于其他两个算法, 不过随着查错率的增加, LPP 的表现不如 SIFT、PCA-SIFT, 其中在查错率大于 0.23 以后, SIFT 是最优的. 总的说来, 在光照变换的情况下, 三者都有满意的效果. 图 3 (b) 是对噪声的影响进行评估. 根据实验, LPP 和 PCA 对噪声的鲁棒性相近, LPP 略好于 PCA. 在查错率小于 0.8 时, 它们都好于 SIFT, 但查错率在 0.8 以后, SIFT 明显优于前两个. 图 3 (c) 是关于尺度和方向的变化, 对此, 三个算法的性能很接近, 很难准确地说明哪个算法更优. 从实验数据分析, LPP 略好于其他两种算法, SIFT 在小的查错率下的表现不好, 不过在查错率大于 0.6 时, SIFT 要好于 PCA. 这也说明 PCA 在对样本集进行线性变换时, 丢失了一些重要的数据信息, 使得一些不匹配的点产生了错误的匹配. 而 LPP 恰好保存了特征点的 k 邻接点, 使得它们的局部结构保持不变. 图 3 (d) 是视角变化的实验结果, 从图中可以看出, 在小的查错率的情况下 (小于 0.5), LPP 优于 PCA、SIFT, 在更大的查率下, LPP 表现不

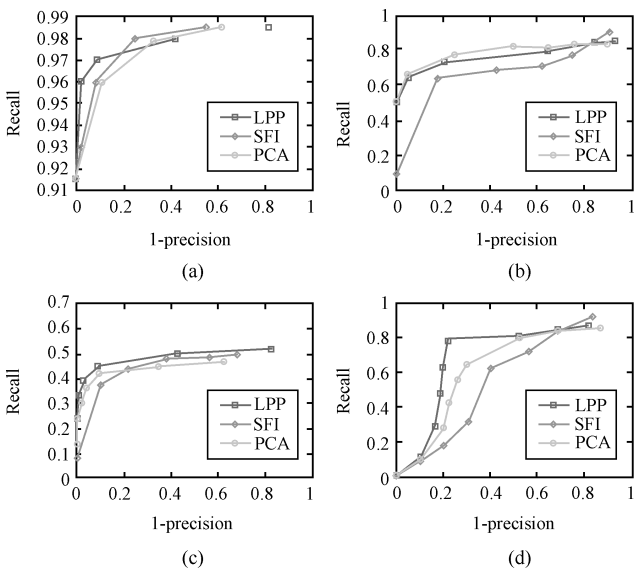


图 3 使用 LPP, PCA, SIFT 对图 2 中的图像集进行性能估计

Fig. 3 Descriptors computed by LPP, PCA, and SIFT (Test images were from Fig. 2.)

不如 SIFT, 这是由于在图像发生视角变化时, LPP 局部保存了 k 邻接点的信息, 同时也保存了由于视角变化带来的错误信息. 所以随着查错率的增加, 其查全率反而降低. PCA 在视角变化中的表现介于 SIFT 和 LPP 之间.

在图 4 中, 本文综合了各种影响来对三种算法进行评估. 在这一组图像中, 物体发生了方向、尺度、视角上的变化, 目标物体之间有一定的噪声影响, 为了加大噪声, 统一把右边的图像预先进行下采样. 根据实验表明, 在限定同样对应匹配点时, LPP 准确率优于 PCA 和 SIFT.

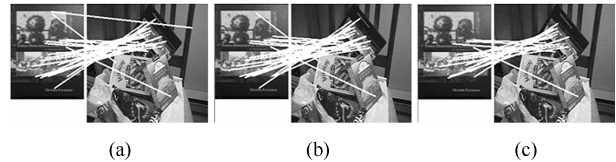


图 4 PCA-SIFT, SIFT, LPP-SIFT 三种算法在尺度、视角变化下的特征匹配效果, 正确率分别为 17/21, 19/21, 20/21

Fig. 4 The matching examples using PCA-SIFT, SIFT and LPP-SIFT for image changes, including scale, zoom and viewpoint (21 features were selected, and correct matches were 17, 19, and 20, respectively.)

在图 5 中, 图片间发生了视角变化. 从匹配的结果可以看到, PCA-SIFT 的表现要低于另外两种算法, LPP-SIFT 在三者中表现最好, 略优于 SIFT.

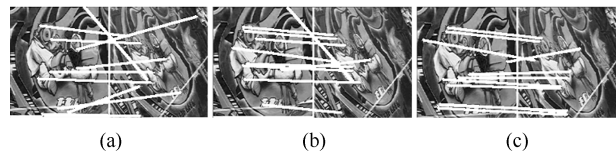


图 5 PCA-SIFT, SIFT, LPP-SIFT 三种算法在视角变化下的特征匹配效果, 正确率分别为 6/10, 8/10, 9/10

Fig. 5 The matching examples using PCA-SIFT, SIFT and LPP-SIFT for viewpoint changes (10 features were selected, and correct matches were 6, 8, and 9, respectively.)

图 6 是对不同卫星拍摄的图片进行测试, 实验结果表明, PCA-SIFT 与 SIFT 获得了近似的性能, 但它们都低于 LPP-SIFT.

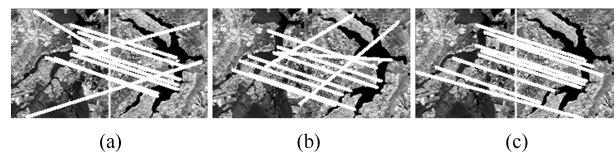


图 6 PCA-SIFT, SIFT, LPP-SIFT 三种算法在尺度、视角变化下的特征匹配效果, 正确率分别为 7/10, 7/10, 9/10

Fig. 6 The matching examples using PCA-SIFT, SIFT and LPP-SIFT for different sensors images (10 features were selected, and correct matches were 7, 7, and 9, respectively.)

从上面的实验结果可以看出, 对于查错率不大的情况下, LPP 算法的鲁棒性是最强的, 在实际应用中, 一般不容许大的错误率出现, 因此可以认为, LPP 算法的鲁棒性要强于 SIFT 与 PCA-SIFT.

表 1 是对 LPP-SIFT 与 PCA-SIFT 匹配速度的比较,

因为 PCA-SIFT 已经证明比 SIFT 速度快, 因此本文只对 LPP-SIFT 与 PCA-SIFT 比较, 如表 1, 初始选择了 2000 个关键点. 图片集是来自图 2 中的右上图. 为获得同样的匹配对数 (采用邻接比法, 邻接比为 0.65), LPP-SIFT 需要 20 维向量, PCA-SIFT 需要 48 维向量. 从表 1 中可以看到, 对于描述器的生成, LPP-SIFT 略快于 PCA-SIFT, 但 LPP-SIFT 匹配时间只有 PCA-SIFT 的一半. 因此, 在同样的特征提取算法下, LPP-SIFT 总的的时间少于 PCA-SIFT.

表 1 LPP-SIFT 与 PCA-SIFT 的速度比较
Table 1 Comparison of runing time of LPP-SIFT and PCA-SIFT

算法	特征点数目	匹配点对	描述器耗时间 (s)	匹配时间 (s)
LPP-SIFT (20)	2 000	232	1.57	0.46
PCA-SIFT (48)	2 000	232	1.63	0.97

3 结论

本文提出了一种新的描述器算法, 该算法主要有三个优点: 1) 通过 LPP (一种线性的方法) 来解决非线性数据的降维问题, 从而弥补了 PCA 对此类问题的不足, 提高了匹配的精度. 使得鲁棒性要好于 SIFT; 2) 同样的性能下, LPP-SIFT 匹配速度要快于 PCA-SIFT; 3) 与 PCA-SIFT 类似, 该算法只需在训练阶段进行一次的特征空间计算, 因此, 在实际应用中较为方便.

基于统计学习的描述器算法, 大多数情况下都取得了很好的效果, 但另一方面, 其性能常受到样本选取的影响, 因此, 如何尽量降低样本带来的误差, 是下一步将要研究的课题.

References

- Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1615–1630
- Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, Schaffalitzky F. A comparison of affine region detectors. *International Journal of Computer Vision*, 2005, **65**(1-2): 43–72
- Koenderink J, van Doorn A J. Representation of local geometry in the visual system. *Biological Cybernetics*, 1987, **55**(6): 367–375
- Freeman W T, Adelson E H. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, **13**(9): 891–906
- van Gool L J, Moons T, Ungureanu D. Affine/photometric invariants for planar intensity patterns. In: *Proceedings of the 4th European Conference on Computer Vision*. Cambridge, England: Springer, 1996. 642–651
- 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
- Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets, or “How do I organize my holiday snaps?”. In: *Processing of the 7th European Conference on Computer Vision*. Copenhagen, Denmark: Springer, 2002. 414–431
- Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110

- Jolliffe I T. *Principal Component Analysis*. London: Springer-Verlag, 2002
- Ye K, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE, 2004. 506–513
- He X F, Niyogi P. Locality Preserving Projections, Technical Report TR-2002-09, Computer Science Department, University of Chicago, USA, 2002
- Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding. *Science*, **290**(5500): 2323–2326
- Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in Neural Information Processing Systems 14*. New York: Springer, 2002

徐小明 重庆大学软件学院硕士研究生. 主要研究方向为数字图像处理与计算机视觉. 本文通信作者. E-mail: xxm9136@sina.com
(XU Xiao-Ming Master student at the School of Software Engineering, Chongqing University. His research interest covers digital image processing and computer vision. Corresponding author of this paper.)

杨丹 重庆大学教授, 主要研究方向为数字图像处理与计算机视觉. E-mail: dyang@cqu.edu.cn
(YANG Dan Professor at Chongqing University. His research interest covers digital image processing and computer vision.)

张小洪 重庆大学软件学院副教授. 主要研究方向为数字图像处理与计算机视觉. E-mail: xhongz@yahoo.com.cn
(ZHANG Xiao-Hong Associate professor at the School of Software Engineering, Chongqing University. His research interest covers digital image processing and computer vision.)

周小龙 重庆大学软件学院硕士研究生. 主要研究方向为数字图像处理与计算机视觉. E-mail: shaolongchou@126.com
(ZHOU Xiao-Long Master student at the School of Software Engineering, Chongqing University. His research interest covers digital image processing and computer vision.)