

基于非负矩阵分解的航拍图像信息提取

郝红, 徐常青, 张新平

(浙江农林大学 环境与资源学院, 浙江 临安 311300)

摘要: 为研究非负矩阵分解(NMF)算法的优越性, 将此方法应用于航拍图像的信息提取。在对航拍图像进行信息提取时, 与传统的主成分分析方法进行了比较。通过选取不同的特征值 r 与 E_{igmm} 进行实验。结果表明: 当 $r=20, 50$ 时, 应用本研究算法的运行时间小于 $E_{\text{igmm}}=20, 50$ 时主成分分析的运行时间; 并对 $r=50, E_{\text{igmm}}=50$ 时进行了识别率比较, 应用此方法的识别率更高, 证明了非负矩阵分解算法在航拍图像应用中的时间短、识别率高的优越性。图 10 表 2 参 7

关键词: 森林经理学; 非负矩阵分解; 航拍图像; 主成分分析; 特征值

中图分类号: S757; TP751

文献标志码: A

文章编号: 2095-0756(2012)01-0072-06

Aerial image information extraction based on non-negative matrix factorization

HAO Hong, XU Chang-qing, ZHANG Xin-ping

(School of Environmental and Resource Sciences, Zhejiang A & F University, Lin'an 311300, Zhejiang, China)

Abstract: This study was on superiority of the non-negative matrix factorization (NMF) algorithm for application of information extracted with aerial images. First, NMF was used for aerial image information extraction, and then this data was compared with a principal component analysis (PCA) in which r (the number of rows or columns of basic matrix) and E_{igmm} (the number of eigenvalues) were given different values. Experimental results showed that the run time of NMF with $r = 20$ or 50 was less than that of PCA with an $E_{\text{igmm}} = 20$ or 50 . Also, the recognition rate of NMF with $r = 50$ was higher than that of an $E_{\text{igmm}} = 50$. The experiment showed that non-negative matrix factorization had advantages of a short time period with a high recognition rate. [Ch, 10 fig. 2 tab. 7 ref.]

Key words: forest management; non-negative matrix factorization (NMF); aerial image; principle component analysis (PCA); eigenvalue

图像识别是模式识别和图像处理等学科研究的一个热点课题, 但将非负矩阵分解(non-negative matrix factorization, NMF)方法用于航拍图片识别方面还较少研究。本研究所提的航拍图片的识别主要是指航拍图像中的人工区域(房屋、建筑物、道路等)与自然区域^[1]的分割的相关研究。主成分分析(principle component analysis, PCA)是用于信号恢复的经典算法, 也是多元数据分析中最重要的方法之一。识别算法中, 主成分分析是当前用得最多的方法^[2]。主成分分析是线性模型参数估计性能的一种常用方法。基本思想是将原来的回归自变量变换到另一组变量, 即所谓的“主成分”, 然后选择其中一部分重要成分作为自变量(此时丢弃了一部分不重要的自变量), 最后利用最小二乘方法对选取主成分后的模型参数进行估计。主成分分析用于图像识别主要是在二阶统计量基础上进行分析的, 但是主成分分析有一个优点同时

收稿日期: 2011-02-28; 修回日期: 2011-04-07

基金项目: 国家自然科学基金资助项目(10871230); 国家高技术研究发展计划(“863”计划)项目(2006AA12Z109); 浙江农林大学研究生科研创新基金资助项目(2112009044)

作者简介: 郝红, 从事环境科学中的模式识别研究。E-mail: 191967012@qq.com。通信作者: 徐常青, 教授, 从事模式识别等研究。E-mail: cqxurichard@zafu.edu.cn

也是缺点：它是完全无参数限制的。在主成分分析的计算过程中完全不需要人为的设定参数或是根据任何经验模型对计算进行干预，如果用户对观测对象有一定的先验知识，掌握了数据的一些特征，却无法通过参数化等方法对处理过程进行干预，可能会得不到预期的效果，效率也不高。并且它的数据量比较大，运行速度慢，在计算时常会用到奇异值分解(singular value decomposition, SVD)，这会使计算量变大。在此提出运用非负矩阵分解方法来对它进行分析，这从某种程度上克服了以上缺点。它通过数据矩阵的非负分解，对数据矩阵进行降维，产生基本矩阵和编码矩阵，其矩阵元素的非负性也保证了分解结果的可解释性，并且也压缩了数据存储量，加快了数据的处理速度。本研究将非负矩阵分解算法运用在无人机航拍图片的对象识别中，并与传统的主成分分析法进行比较，可以看出非负矩阵分解算法的优越性。

1 研究方法

非负矩阵分解(NMF)方法是基于一种“局部构成整体”思想的分解方法，由 Lee 等^[3]于 1999 年提出。非负矩阵分解方法是在原始数据矩阵 V 中所有元素均非负的条件下对 V 实现非负分解。传统矩阵分解(如矩阵特征分解、SVD、QR 分解等)得到的因子矩阵一般存在负元，因此，实际问题中往往缺失物理解释。非负矩阵分解方法提供了一种新的矩阵分解思路，具有可解释性和明确的物理意义以及占用存储空间少等优点。此方法一经提出就引起许多科学家和研究人员的广泛重视，在机器学习、计算机视觉和模式识别等研究领域有广泛应用。

1.1 非负矩阵分解的基本原理

给定一个 $n \times m$ 元素非负的矩阵 V ，非负矩阵分解算法是要寻找一个 $n \times r$ 的非负矩阵 W 和 $r \times m$ 非负矩阵 H 使得：

$$V \approx WH. \quad (1)$$

式(1)中： W 被称为基矩阵， H 被称为编码矩阵。自然数 r 满足不等式 $r \ll \frac{mn}{m+n}$ 。式(1)的最优解(W , H)依赖于代价函数的确定(如使得 2-范数 $\|V-WH\|$ 最小化等)。分解后的矩阵元素个数(数据存储量)远小于原矩阵数据量，从而对原始数据矩阵进行了降维。由于分解前后的矩阵仅包含非负元，原矩阵中的每个列向量可解释为基矩阵 W 的列向量加权和，加权向量为 H 中对应列向量。

为了得出最佳逼近结果，衡量分解的精确度，常用的代价函数为：

$$\|V-WH\|^2 = \sum_{ij} (V_{ij} - (WH)_{ij})^2. \quad (2)$$

对应的乘法迭代更新规则：

$$H_{a\mu} \leftarrow H_{a\mu} \frac{(W^T V)_{a\mu}}{(W^T W H)_{a\mu}}, \quad W_{ia} \leftarrow W_{ia} \frac{(V H^T)_{ia}}{(W H H^T)_{ia}}.$$

称为乘法迭代。非负矩阵分解算法有时会结合其他条件，如：概率模型下的非负矩阵分解算法、局部非负矩阵分解算法、稀疏性约束下的非负矩阵分解算法等^[2,4-5]。

非负矩阵分解算法要求分解得到的因子矩阵为元素非负。通常我们得到的分解结果中只有部分元素起作用，因此，使用稀疏度^[6]来控制其稀疏程度的表示是必要的。这样可以尽可能降低因子矩阵的维数。局部非负矩阵分解就是基于这种思想。它是如下约束化问题：

$$\min_{L_{\text{NMF}}} (W, H) = \sum_{ij} (V_{ij} \log \frac{V_{ij}}{(WH)_{ij}} - V_{ij} + (WH)_{ij}) + \alpha \sum_{ij} M_{ij} - \beta \sum_i N_{ii} \quad \text{s.t.} \quad \sum_i W_{ik} = 1, \quad \forall k = 1, 2, \dots, n.$$
 这里 $M = W^T W$, $N = H H^T$, α, β 为 Lagrange 乘子。给定一个 n 维实向量 x 。 x 的稀疏度定义为：

$$S(x) = \frac{\sqrt{n} - (\sum |x_i|)}{\sqrt{\sum x_i^2} - 1}.$$

$0 \leq S(x) \leq 1$, $S(x)=1$ 当且仅当 x 恰有一个非零元, $S(x)=0$ 当且仅当 x 的所有元素相等且不为 0。不同的稀疏度会得出不同的结果^[6]，所要做的就是对 NMF 施加约束条件以达到所要求稀疏度时的最佳分解矩阵。

1.2 非负矩阵分解的应用

非负矩阵分解算法一经提出，便被广泛用于诸多领域，如图像处理、文本聚类、数据挖掘、语音识别、生物医学工程和化学工程领域等。尤其是在图像处理领域，涉及到人脸识别、图像融合、图像检索、图像分类、图像压缩等诸多方面。在这些领域中，不管是单独应用非负矩阵分解算法，或是结合非负矩阵分解算法与不同方法的应用，实验都取得了较好的效果，但同时也提出了一些有待解决的问题，如因子矩阵的初始化以及数 r 的确定问题等^[7]。

本研究将非负矩阵分解算法应用于无人机航拍图片处理中，以期获得更好的图像分割效果。

2 非负矩阵分解在航拍图像识别中的应用

2.1 试验数据

本研究选取的是 1 幅无人机航拍图片。为了节约存储空间，首先将彩色图像进行压缩，压缩后的图片数据矩阵大小为 90×80 ，之后在 Matlab 中将其化为灰度图像(图 1)。

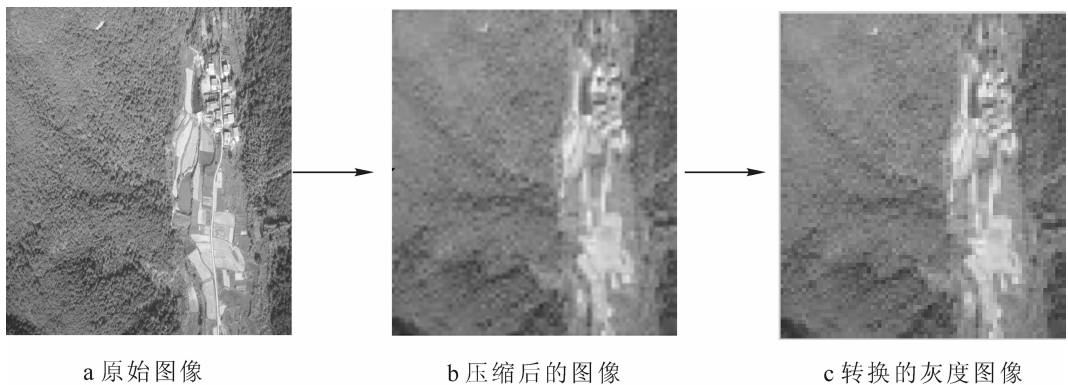


图 1 无人机航拍图片

Figure 1 Unmanned aerial images

2.2 试验结果及算法分析

本研究是在内存为 2.00 GB 的 Lenovo 机器上进行测试的，使用的 Matlab 版本为 MatlabR 2008 b。对于非负矩阵分解算法的实现，首先要确定 r 的值，以便对它们进行分解。本研究分别选取 $r = 10, 20, 50$ 。本研所得分解结果 W 和 H 均为数字矩阵，分别用相应的图形来表现，图 2 是 $r = 10$ 的分解结果。

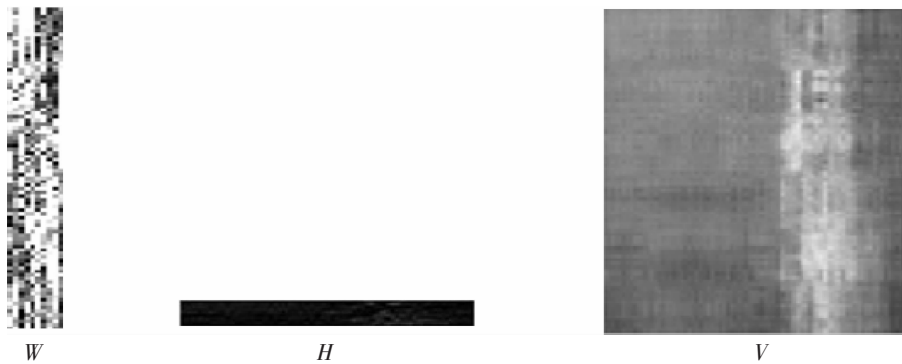


图 2 $r = 10$ 时非负矩阵分解结果

Figure 2 Images after NMF, corresponding to the case $r = 10$

在图 2 中， W 表示的是灰度图像的特征向量(如森林树木中的不同属性、建筑等)所组成的矩阵，对 W 中的列向量进行加权和即得到对应的 V 的列向量， H 为编码矩阵，加权向量便是 H 所对应的列向量，它的元素数值表示 W 列中每个特征所占比例的大小。经非负矩阵分解算法后对原始矩阵 V 进行存储，便可以将其分为 2 个部分 W 和 H 分别进行存储，以此节约空间。 $r = 20, 50$ 时所得结果分别如图 3 和

图 4 所示。不同 r 对应的运行时间 t 及 cputime 如下表 1 所示，很明显看出虽然 r 取值越大，所得效果越好，但同时算法所需运行时间和中央处理器占用时间越长。对原始图像进行阈值分割后可得到图 5，同样对非负矩阵分解算法后的 3 个图像进行分割，并与图 5 进行比较。由图 5 和图 6 可以比较出白色区域有可能是水泥路面或者建筑物，黑色区域为森林。从图 5 和图 6 可看出， r 越大， V 的识别效果也越好，

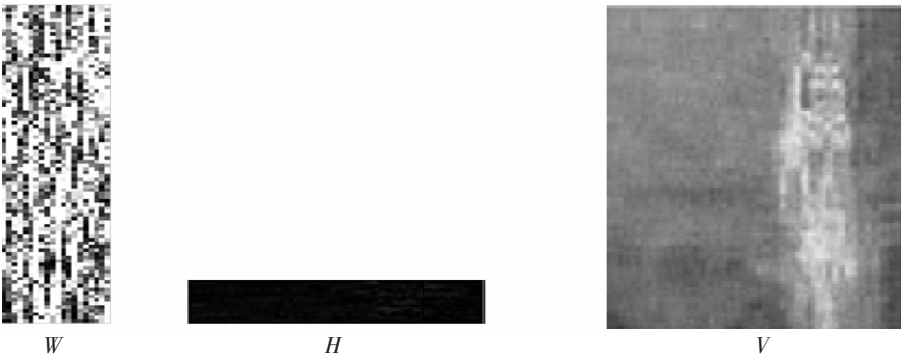


图 3 $r=20$ 时非负矩阵分解结果

Figure 3 Images after NMF, correspond ing to the case $r = 20$

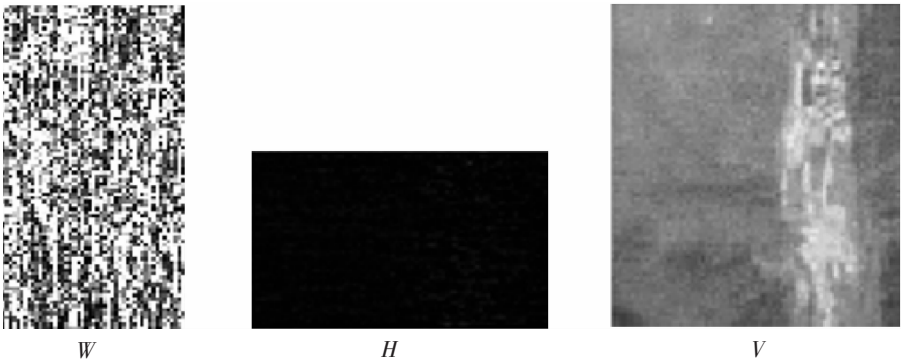


图 4 $r=50$ 时非负矩阵分解结果

Figure 4 Images after NMF, correspond ing to the case $r = 50$

表 1 不同 r 对应的非负矩阵分解算法运行时间与中央处理器占用时间

Table 1 Running time and CPU time corresponding to NMF algorithm for different r

r	运行时间/s	中央处理器占用时间/s
10	28.016 0	9.687 7
20	30.351 0	13.502 4
50	40.106 0	28.907 0

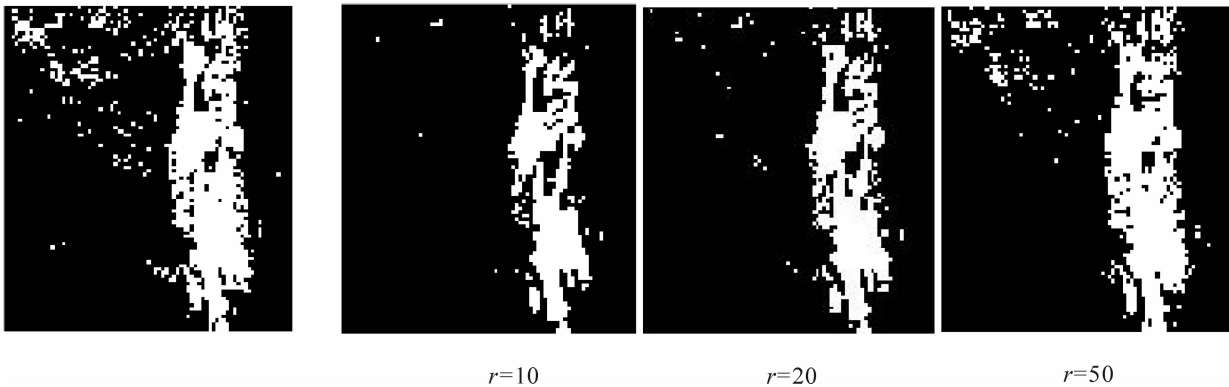


图 5 原始图像分割图

Figure 5 Segmentation image of original image

图 6 经非负矩阵分解的图像分割图

Figure 6 Segmentation image after NMF

但是 W 和 H 的存储空间也越大。但即便是 $r = 10$ ，也可分辨出人工区域和自然区域。

原始矩阵经非负矩阵分解算法后对因子矩阵进行分割，因为 W 的每列元素(森林或建筑物周围)的特征(自上而下)很相似，分割效果不明显，所以只对 H 进行阈值分割，不同 r 值可得到不同的分割效果，在 r 的 3 个值中 $r = 20$ 时效果最好，将 H 分为了 3 个部分，如图 7 所示。



图 7 $r = 20$ 时 H 的分割图

Figure 7 Segmentation image of H , corresponding to $r = 20$

由图 5~7 看出，本研究选取的 3 个值中， $r = 20$ 时，不仅识别效果好同时也压缩了存储空间。对灰度图像数据运用主成分分析方法识别，得特征值(E_{ignum})总数为 80 个，选取累计贡献率 0.85，则只需选取 18 个即可，为方便与以上 3 种情况进行比较，分别再对特征值个数为 20 和 50 时进行运算。比较结果如图 8~9 所示。不同特征值个数对应的运行时间 t 及中面处理器占用时间如表 2 所示。

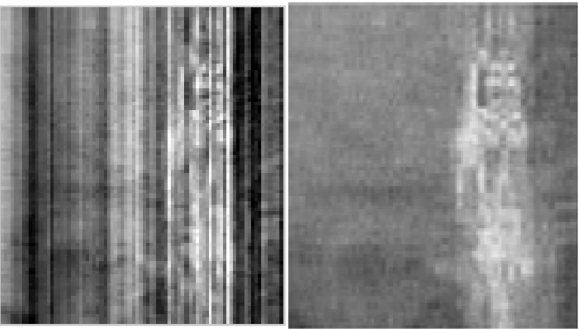


图 8 $r = 20$ 与 $E_{\text{ignum}} = 20$ 时主成分分析(左图)与非负矩阵分解算法(右图)识别图比较

Figure 8 Comparison between PCA (left) and NMF (right) corresponding to $r = 20$ and $E_{\text{ignum}} = 20$

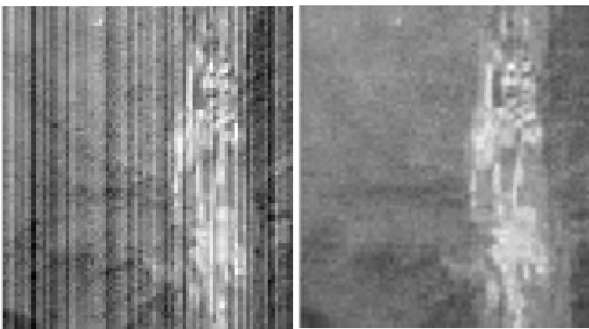


图 9 $r = 50$ 与 $E_{\text{ignum}} = 50$ 时主成分分析(左图)与非负矩阵分解算法(右图)识别图比较

Figure 9 Comparison between PCA (left) and NMF (right) corresponding to $r = 50$ and $E_{\text{ignum}} = 50$

表 2 主成分分析方法在不同特征值个数时对应运行时间及中央处理器占用时间

Table 2 Running time and CPU time of PCA method in different eigenvalues

E_{ignum}	运行时间/s	中央处理器占用时间/s
18	28.131 0	16.452 5
20	29.642 7	20.423 4
50	41.382 0	36.263 5

由图 8~9 明显看出：由于主成分分解后的结果不限制正负性，出现的正负结果在相互抵销后，使得图片模糊不清。 $r = 20, 50$ 时应用非负矩阵分解算法的效果比 $E_{\text{ignum}} = 20, 50$ 时应用主成分分析的效果好。表 1 和表 2 中， $r = 20, 50$ 时与 $E_{\text{ignum}} = 20, 50$ 相比较其运行时间也是小于后者，但是很明显，前者的中央处理器占用时间总是小于后者所对应的。2 种算法在 r 与 E_{ignum} 取相同值时的识别率如图 10 所示。当选取 $r = 50, E_{\text{ignum}} = 50$ 时进行了识别率比较，2 种方法的识别率分别为 83.48%和 72.60%。

3 结论

本研究将非负矩阵分解算法应用于无人机航拍

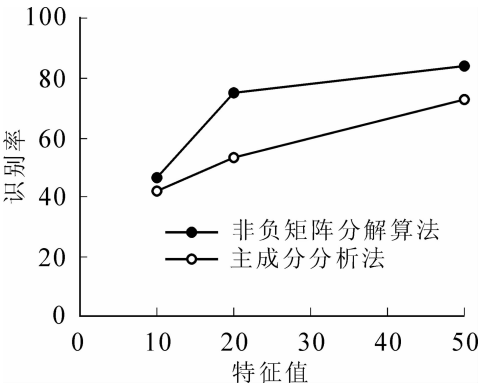


图 10 主成分分析与非负矩阵分解算法识别率

Figure 10 Recognition rate of PCA and NMF

图片的识别。从表 1 和表 2 看出, $r = 20, 50$ 时与 $E_{\text{igmm}} = 20, 50$ 相比较其运行时间也是小于后者, 但是很明显, 前者的中央处理器占用时间总是小于后者所对应的中央处理器占用时间。由图 2 和图 3 看出, 即便是选取这 3 种选择中的最小值时 ($r=10$), 也可以分辨出人工区域和自然区域。但由图 6 和图 7 知, 选取 $r = 20$ 时的效果好, 且同时压缩了储存空间。图 6 所示 6 种方法的识别率和表 3 所示 2 种方法在 $r = 50, E_{\text{igmm}} = 50$ 的识别率看出, 航拍图片的识别应用方法中, 非负矩阵分解算法方法要优于主成分分析法。

4 非负矩阵分解中有待解决的问题

非负矩阵分解算法是近年来发展起来的矩阵分解方法, 它克服了传统矩阵分解的许多问题^[7], 如因子矩阵的物理意义等, 并且被广泛应用在许多领域, 但非负矩阵分解算法在应用过程中也存在许多缺点和问题。针对本研究的应用, 总结了以下几个问题: ①非负矩阵分解算法代价函数有很多, 如何根据具体问题选择一个恰当的代价函数? ②非负矩阵分解算法可以带有多种约束条件(如基本矩阵 W 或编码矩阵 H 的稀疏度等), 并产生不同的推广形势下的非负矩阵分解算法。如何根据具体问题选择一个合适的非负矩阵分解算法, 使得分解效果最优? ③为了实现分解后的结果能达到降维的目的, 选择合适的维数是必要的, 那么如何选取才能达到降维目的并且使存储空间最小? ④在稀疏度约束下, 稀疏度选取多大才更合适?

参考文献:

- [1] 曹国, 杨新. 基于水平集方法的航拍图片人工区域的分割检测[J]. 模式识别与人工智能, 2006, **19** (4): 526 – 530.
CAO Guo, YANG Xin. Man-made objects detection from aerial images based on level set method [J]. *PR & AI*, 2006, **19** (4): 526 – 530.
- [2] 张永鹏, 郑文超, 张晓辉. 非负矩阵分解及其在图像压缩中的应用[J]. 西安邮电学院学报, 2008, **13** (3): 58 – 61.
ZHANG Yongpeng, ZHENG Wenchao, ZHANG Xiaohui. Nonnegative matrix factorization and its applications in data compression [J]. *J Xi'an Univ Posts Telecommun*, 2008, **13** (3): 58 – 61.
- [3] LEE D D, SEUNG H S. Learning the parts of objects by non-negative matrix factorization [J]. *Nature*, 1999, **401** (6755): 788 – 791.
- [4] 刘维湘, 郑南宁, 游屈波. 非负矩阵分解及其在模式识别中的应用[J]. 科学通报, 2006, **51** (3): 241 – 250.
LIU Weixiang, ZHENG Nanning, YOU Qubo. Non-negative matrix factorization and its application in pattern recognition [J]. *Sci Bull*, 2006, **51** (3): 241 – 250.
- [5] LEE D D, SEUNG H S. Algorithms for non-negative matrix factorization [G] //LEEN T, DIETTERICH T, TRESP V. *Advances in Neural Information Processing Systems*. Cambridge: MIT Press, 2000: 556 – 562.
- [6] HOYER P O. Non-negative matrix factorization with sparseness constraints [J]. *J Mach Learn Res*, 2004, **5**: 1457 – 1469.
- [7] 徐泰燕, 郝玉龙. 非负矩阵分解及其应用现状分析[J]. 武汉工业学院学报, 2010, **29** (1): 109 – 114.
XU Taiyan, HAO Yulong. The current research situation analysis of non-negative matrix factorization and applications [J]. *J Wuhan Polytech Univ*, 2010, **29** (1): 109 – 114.