

## 基于反馈稀疏约束的非负张量分解算法

刘亚楠<sup>1,2</sup>, 涂铮铮<sup>1</sup>, 罗斌<sup>1\*</sup>

(1. 安徽大学 计算机科学与技术学院, 合肥 230039; 2. 合肥师范学院 计算机科学与技术系, 合肥 230601)

(\* 通信作者电子邮箱 luobin@ahu.edu.cn)

**摘要:** 为了充分利用图像本身的结构信息并充分压缩图像数据, 把得到的子空间中数据(反馈)的稀疏性作为约束项加入非负张量分解目标函数中, 即采用基于反馈稀疏约束的非负张量分解算法对图像集合进行降维。最后, 将该算法应用于手写数字图像库中, 实验结果表明所提出的方法能有效改善图像分类的准确性。

**关键词:** 非负矩阵分解; 稀疏约束; 张量分解

**中图分类号:** TP391.41 **文献标志码:** A

### Non-negative tensor factorization based on feedback sparse constraints

LIU Yanan<sup>1,2</sup>, TU Zhengzheng<sup>1</sup>, LUO Bin<sup>1\*</sup>

(1. School of Computer Science and Technology, Anhui University, Hefei Anhui 230039, China;

2. Department of Computer Science and Technology, Hefei Normal College, Hefei Anhui 230601, China)

**Abstract:** In order to fully use the structural information of the data, and compress the image data, the sparse constraints of the subspace (feedback) were applied to the object function of non-negative tensor factorization. Then this algorithm was used to reduce the dimension of the image sets. Finally, image classification was realized. The experimental results on the handwritten digital image database show that the proposed algorithm can effectively improve the accuracy of the image classification.

**Key words:** Non-negative Matrix Factorization (NMF); sparse constraint; tensor factorization

## 0 引言

非负矩阵分解 (Non-negative Matrix Factorization, NMF)<sup>[1]</sup> 被广泛应用于图像处理、计算机视觉、数据挖掘等领域中<sup>[2-3]</sup>, 它使分解后的所有分量均为非负值并且同时实现非线性的维数约简。然而在很多情况下, 需要处理的数据是多维(张量)数据, 如果在图像处理领域中采用 NMF 方法, 需要把图像数据拉直成向量形式, 在转换过程中会丢失图像数据本身的结构信息, 破坏图像的空间几何结构。为了避免这些问题, Welling 等<sup>[4]</sup> 把 NMF 推广到非负张量分解 (Non-negative Tensor Factorization, NTF), NTF 被广泛应用于图像处理与模式识别领域<sup>[5-6]</sup>。此外, 稀疏非负矩阵分解 (Sparse Non-negative Matrix Factorization, SNMF)<sup>[7-8]</sup> 把稀疏约束加入进来, 希望得到数据更有意义的分解, 同时稀疏非负张量分解 (Sparse Non-negative Tensor Factorization, SNTF) 作为 SNMF 自然的扩展也得到很多的关注<sup>[5,9]</sup>。

传统的 SNTF 的稀疏约束并不能保证降维后子空间中的数据是稀疏的(非零元素较多), 如果降维后子空间中的数据是稀疏的, 那么每个样本的信息量就会被再次压缩, 更便于后续的分类。由于降维后子空间中的数据可以看作是反馈信息, 本文提出一种基于反馈稀疏约束的非负张量分解的算法, 在保留图像结构信息的同时, 压缩图像的信息量, 并将其应用于手写数字数据库分类中, 得到了较好的分类结果。

## 1 背景介绍

### 1.1 张量几何相关运算

本文算法中用到的张量几何<sup>[10]</sup>的相关定义和定理如下:

**定义 1** 张量矩阵展开是将一个张量中的元素重新排列, 得到一个矩阵的过程, 张量  $X \in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N}$  的  $n$  模展开矩阵表示为  $X_{(n)} \in \mathbf{R}^{I_n \times (I_1 \cdots I_{n-1} I_{n+1} \cdots I_N)}$ 。

**定义 2** 张量乘法。张量与矩阵的乘法由  $n$  模乘积所定义。具体地, 张量  $X \in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N}$  和矩阵  $U \in \mathbf{R}^{J_n \times I_n}$  的  $n$  模乘积表示为  $X \times_n U$ , 这是一个  $I_1 \times I_2 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N$  阶张量, 定义如下:

$$(X \times_n U)_{i_1 i_2 \cdots i_n \cdots i_N} = \sum_{j_n} a_{i_1 i_2 \cdots i_n \cdots i_N} \cdot u_{j_n i_n} \quad (1)$$

### 1.2 非负张量分解

现有的张量分解模型一般分为两类: 标准分解模型<sup>[11]</sup>和 Tucker 分解模型<sup>[12]</sup>。这两种分解模型都是对矩阵奇异值分解的高阶推广, 并且第一种分解模型是第二种分解模型的特例。

非负 Tucker 模型的目标是把非负张量  $X \in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N}$  分解为

$$X \approx G \times_1 U_1 \times_2 U_2 \times \cdots \times_N U_N \quad (2)$$

其中:  $G \in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N}$ ,  $U_n \in \mathbf{R}^{J_n \times I_n}$  ( $J_n \leq I_n, n = 1, 2, \cdots, N$ )。该分解可通过求解最优化问题式(3)得到:

$$\min \|G \times_1 U_1 \times_2 U_2 \times \cdots \times_N U_N - X\|^2$$

**收稿日期:** 2013-04-17; **修回日期:** 2013-05-29。 **基金项目:** 国家自然科学基金资助项目(61073116); 高校省级优秀青年人才基金重点资助项目(2011SQRL129ZD); 安徽大学“211”工程创新团队项目。

**作者简介:** 刘亚楠(1984-), 女, 山东济宁人, 讲师, 博士研究生, 主要研究方向: 模式识别、图像处理; 涂铮铮(1982-), 女, 安徽六安人, 讲师, 博士研究生, 主要研究方向: 模式识别、图像处理、视频信号处理; 罗斌(1963-), 男, 安徽合肥人, 教授, 博士生导师, 主要研究方向: 模式识别、图像处理。

$$\text{s. t. } \mathbf{G} \geq 0, \mathbf{U}_n \geq 0, n = 1, 2, \dots, N \quad (3)$$

### 1.3 稀疏非负张量分解

SNTF 一般描述为最优化问题<sup>[5]</sup>:

$$\min_{\mathbf{U}_1, \dots, \mathbf{U}_N} \|\mathbf{G} \times_1 \mathbf{U}_1 \times_2 \dots \times_N \mathbf{U}_N - \mathbf{X}\|^2 + \sum_{1 \leq n \leq N} \lambda_n \|\mathbf{U}_n\|_1 \quad (4)$$

$$\text{s. t. } \mathbf{U}_n \geq 0, 1 \leq n \leq N$$

在该问题中,  $L_1$  范数惩罚项  $\sum_{1 \leq n \leq N} \lambda_n \|\mathbf{U}_n\|_1$  作为稀疏约束, SNTF 的目标是希望 NTF 的分解因子  $\mathbf{U}_n (1 \leq n \leq N)$  是稀疏的。

## 2 基于反馈稀疏约束的非负张量分解算法

对图像集合三阶张量  $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\} \in \mathbf{R}^{I_1 \times I_2 \times N}$  ( $\mathbf{X}_n \in \mathbf{R}^{I_1 \times I_2}, n = 1, 2, \dots, N$ ) 来说, 由式(4)可以得到  $\mathbf{U}_1, \mathbf{U}_2$ , 由  $\mathbf{U}_1, \mathbf{U}_2$  可得  $\mathbf{X}_i$  对应子空间中的  $\mathbf{Y}_i (i = 1, 2, \dots, N)$ :

$$\mathbf{Y}_i = \mathbf{U}_1^T \times \mathbf{X}_i \times \mathbf{U}_2 \quad (5)$$

如果由 SNTF 得到  $\mathbf{U}_1, \mathbf{U}_2$ , 那么  $\mathbf{U}_1, \mathbf{U}_2$  是稀疏的, 但是它们的稀疏性并不能保证  $\mathbf{Y}_i (i = 1, 2, \dots, N)$  是稀疏的, 如果希望降维后的  $\mathbf{Y}_i$  是稀疏的, 需要修改稀疏约束项; 另外, 由于  $L_1$  范数不可导, 本文选择可导的  $L_{2,1}$  范数惩罚项  $\|\mathbf{Y}_i\|_{2,1}$  作为稀疏约束, 构造目标函数:

$$\Psi = \|\mathbf{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 - \mathbf{X}\|^2 + \sum_{i=1}^m \lambda_i \|\mathbf{Y}_i\|_{2,1} = \|\mathbf{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 - \mathbf{X}\|^2 + \sum_{i=1}^m \lambda_i \|\mathbf{U}_1^T \mathbf{X}_i \mathbf{U}_2\|_{2,1}$$

$$\text{s. t. } \mathbf{G} \geq 0, \mathbf{U}_n \geq 0, n = 1, 2, 3 \quad (6)$$

对式(6)求偏导可得:

$$\frac{\partial \Psi}{\partial \mathbf{G}} = \mathbf{B} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T \quad (7)$$

$$\frac{\partial \Psi}{\partial \mathbf{U}_1} = \mathbf{B}_{(n)} \mathbf{A}_{(n)}^n \mathbf{G}_{(n)}^T + \left( \sum_{i=1}^m \lambda_i \mathbf{X}_i \mathbf{U}_2 \right) (\mathbf{SD})^T \quad (8)$$

$$\frac{\partial \Psi}{\partial \mathbf{U}_2} = \mathbf{B}_{(n)} \mathbf{A}_{(n)}^n \mathbf{G}_{(n)}^T + \mathbf{SD} \left( \sum_{i=1}^m \lambda_i \mathbf{X}_i^T \mathbf{U}_1 \right) \quad (9)$$

$$\frac{\partial \Psi}{\partial \mathbf{U}_3} = \mathbf{B}_{(3)} \mathbf{A}_{(3)}^{(3)} \mathbf{G}_{(3)}^T \quad (10)$$

其中:  $\mathbf{B} = \mathbf{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 - \mathbf{X}$ ;  $\mathbf{A}_{(n)}^n = \mathbf{U}_N \otimes \dots \otimes \mathbf{U}_{n+1} \otimes \mathbf{U}_{n-1} \otimes \dots \otimes \mathbf{U}_1 (N = 3, n = 1, 2)$ ;  $\mathbf{B}_{(n)}$  和  $\mathbf{G}_{(n)}$  为对应张量  $\mathbf{B}$ 、 $\mathbf{G}$  的  $n$  模展开矩阵;  $\mathbf{S} = \mathbf{U}_1 \sum_{i=1}^m \lambda_i \mathbf{X}_i \mathbf{U}_2$ ;  $\mathbf{D}$  为对角矩阵, 其第  $i$  个对角元素为矩阵  $\mathbf{S}$  第  $i$  行的 2 范数的倒数。

由式(7)可得:

$$(\mathbf{G})_{ij} \leftarrow (\mathbf{G})_{ij} \frac{(\mathbf{X} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T)_{ij}}{(\mathbf{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T)_{ij}} \quad (11)$$

由式(8) ~ 式(10)可得:

$$(\mathbf{U}_1)_{ij} \leftarrow (\mathbf{U}_1)_{ij} \frac{(\mathbf{X}_{(1)} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \mathbf{A}_{(1)}^n \mathbf{G}_{(1)}^T)_{ij}}{(\mathbf{G}_{(1)} \mathbf{A}_{(1)}^n \mathbf{G}_{(1)}^T + \left( \sum_{i=1}^m \lambda_i \mathbf{X}_i \mathbf{U}_2 \right) (\mathbf{SD})^T)_{ij}} \quad (12)$$

$$(\mathbf{U}_2)_{ij} \leftarrow (\mathbf{U}_2)_{ij} \frac{(\mathbf{X}_{(2)} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \mathbf{A}_{(2)}^n \mathbf{G}_{(2)}^T)_{ij}}{(\mathbf{G}_{(2)} \mathbf{A}_{(2)}^n \mathbf{G}_{(2)}^T + \mathbf{SD} \left( \sum_{i=1}^m \lambda_i \mathbf{X}_i^T \mathbf{U}_1 \right))_{ij}} \quad (13)$$

$$(\mathbf{U}_3)_{ij} \leftarrow (\mathbf{U}_3)_{ij} \frac{(\mathbf{X}_{(3)} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \mathbf{A}_{(3)}^3 \mathbf{G}_{(3)}^T)_{ij}}{(\mathbf{G}_{(3)} \mathbf{A}_{(3)}^3 \mathbf{G}_{(3)}^T)_{ij}} \quad (14)$$

具体地, 本文算法思路如算法 1。

算法 1 基于反馈稀疏约束的非负张量分解算法。

1) 构造图像集合三阶张量  $\mathbf{X}$ ;

2) 初始化  $\mathbf{G}, \mathbf{U}_n (n = 1, 2, 3)$ ;

3) 根据式(11) ~ (14) 分别迭代更新  $\mathbf{G}, \mathbf{U}_n (n = 1, 2, 3)$ , 直到满足迭代截止条件(本文的截止条件选为迭代 500 次);

4) 根据式(5)得到子空间降维后的数据  $\mathbf{Y} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N\} \in \mathbf{R}^{I_1 \times I_2} (J_1 \leq I_1, J_2 \leq I_2)$ 。

## 3 实验结果与分析

为了验证本文算法的有效性, 采用手写数字数据库 USPS Handwritten Digits 和 Binary Alphanumerics<sup>[13]</sup>。

第一个数据库包含 0 ~ 9 等十个数字共 10 类, 每类有 1100 幅图像, 图像为  $20 \times 20$  的灰度图像, 从中随机选取 400 幅图像做两组实验。第一组实验是随机选取 100 幅图像做训练, 其余图像做测试, 得到对于不同维数 ( $10 \times 10, 11 \times 11, 12 \times 12, 13 \times 13, 14 \times 14, 15 \times 15$ ) 下图像的分类结果, 如图 1(a) 所示; 第二组实验是从每类中随机选取 50、100、150、200、250 幅图像做训练, 其余图像做测试, 得到不同训练样本数下的分类结果(选取维数为  $10 \times 10$ ), 如图 1(b) 所示。

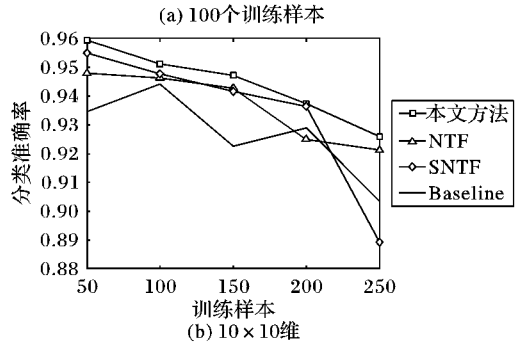
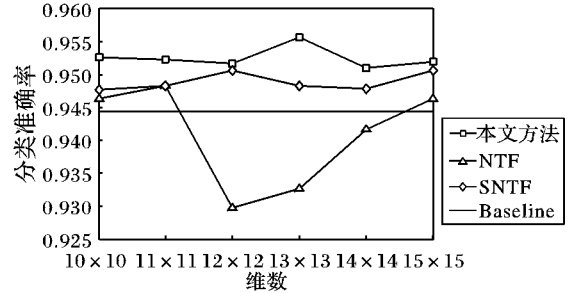


图1 USPS Handwritten Digits 分类准确度比较

第二个数据库包含 0 ~ 9 等 10 个数字, 以及“A”~“Z”等 26 个英文字母, 共 36 类, 每类有 39 幅图像, 图像为  $20 \times 16$  的灰度图像。对该数据库同样做两组实验, 第一组实验是从每类中随机选取 10 幅图像做训练, 其余图像做测试, 得到对于不同维数 ( $10 \times 10, 11 \times 11, 12 \times 12, 13 \times 13, 14 \times 14, 15 \times 15$ ) 下图像的分类结果, 如图 2(a) 所示; 第二组实验是从每类中随机选取 5、10、15、20、25 幅图像做训练, 其余图像做测试, 得到不同训练样本数下的分类结果(选取维数为  $10 \times 10$ ), 如图 2(b) 所示。

本文选用 SNTF<sup>[5]</sup>、NTF、Baseline(仅采用图像原始特征, 不做任何降维处理, 即基线)与本文提出的基于稀疏约束的 NTF 方法作对比, 每一种方法均采用最近邻分类器进行分类。对这三种方法来说, 两幅图像之间的距离:

$$d(\mathbf{X}_i, \mathbf{X}_j) = \|\mathbf{X}_i - \mathbf{X}_j\|_{v_1 v_2} = \|\mathbf{U}_1^T \mathbf{X}_i \mathbf{U}_2 - \mathbf{U}_1^T \mathbf{X}_j \mathbf{U}_2\| \quad (15)$$

所有实验结果均为 10 次分类结果取平均。从图 1、2 中可以看出本文算法优于其他对比算法。

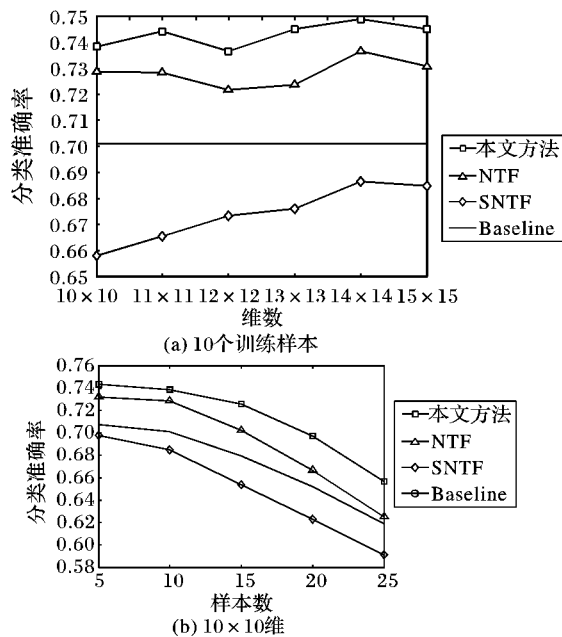


图2 Binary Alphadigits 分类准确度比较

#### 4 结语

为了保持图像数据空间的几何结构,选用非负张量分解进行降维,同时,结合子空间中数据的稀疏约束,提出了一个基于反馈稀疏约束的非负张量分解的目标函数,利用拉格朗日方法对该目标函数进行参数的求解,最后对图像进行分类实验,与其他算法相比本文算法可以有效提高图像分类的准确率。

#### 参考文献:

- [1] LEE D D, SEUNG H S. learning the parts of objects by non-negative matrix factorization[J]. *Nature*, 1999, 401(6755): 788 - 791.
- [2] CAI D, HE X F, HAN J W. Graph regularized non-negative matrix factorization for data representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1548 - 1560.

(上接第 2870 页)

用型隐写分析方法,应用颜色梯度矩阵提取梯度能量、梯度均值、梯度方差、梯度熵等 16 维向量构造隐写分类特征。实验结果表明,本文算法能以平均 81.1% 的准确率识别 RGB 彩色图像隐写,说明了隐写分析中引入颜色通道间相关性作为检测特征的有效性。

进一步的研究工作主要包括两方面:一是对图像颜色梯度的其他特征量进行深入研究,进一步提高隐写分析的准确率,并将其应用于其他类型的隐写分析;二是对彩色图像进行分块划分,提取代表性分块进行颜色梯度分析,提高分析效率,将算法应用于实时检测。

#### 参考文献:

- [1] KONG X W, LIU W F, YOU X G. Secret message location steganalysis based on local coherences of Hue[C]// *Proceedings of the 2005 Advances in Multimedia Information Processing*, LNCS 3768. Berlin: Springer, 2005: 301 - 311.
- [2] 陈铭, 张茹, 刘凡凡. 基于区域相关性的 LSB 匹配隐写分析[J]. *通信学报*, 2010, 31(3): 1 - 11.
- [3] LEE K, JUNG C H, LEE S J, *et al.* Color cube analysis for detection of LSB steganography in RGB color images[C]// *Proceedings of the 2005 Computational Science and Its Applications*, LNCS 3481. Berlin: Springer, 2005: 537 - 546.

- [3] OUHSAIN M, HAMZA B A. Image watermarking scheme using nonnegative matrix factorization and wavelet transform[J]. *Expert Systems with Applications*, 2009, 36(2): 2123 - 2129.
- [4] WELLING M, WEBER M. Positive tensor factorization[J]. *Pattern Recognition Letters*, 2001, 22(12): 1255 - 1261.
- [5] LIU J, LIU J, WONKA P, *et al.* Sparse non-negative tensor factorization using columnwise coordinate descent[J]. *Pattern Recognition*, 2012, 45(1): 649 - 656.
- [6] CONG F Y, PHAN A H, ASTIKAINEN P. Multi-domain feature of event-related potential extracted by nonnegative tensor factorization: 5 vs. 14 electrodes EEG data[C]// *Proceedings of the 10th International Conference on Latent Variable Analysis and Signal Separation*. Berlin: Springer-Verlag, 2012: 502 - 510.
- [7] HOYER P O. Nonnegative matrix factorization with sparseness constraints[J]. *Journal of Machine Learning Research*, 2004, 5(12): 1457 - 1469.
- [8] CICHOCKI A, ZDUNEK R, AMARI S. Csiszar's divergences for nonnegative matrix factorization family of new algorithms[C]// *Proceedings of the 6th International Conference on Independent Component Analysis and Blind Signal Separation*, LNCS 3889. Berlin: Springer-Verlag, 2006: 32 - 39.
- [9] MURAKAMI T, KROONENBERG P M. Three-mode models and individual differences in semantic differential data[J]. *Multivariate Behavioral Research*, 2003, 38(2): 247 - 283.
- [10] LATHAUWER L D. Signal processing based on multilinear algebra [D]. Katholieke, Belgium: University of Leuven, 1997.
- [11] HARSHMAN R A. Foundations of the PARAFAC procedure: models and conditions for an "explanatory" mul-modal factor analysis [J]. *UCLA Working Papers in Phonetics*, 1970, 16: 1 - 84.
- [12] TUEKER L R. Some mathematical notes of three-mode factor analysis[J]. *Psychometrika*, 1966, 31(3): 279 - 311.
- [13] ROWEIS S. Data for Matlab hackers [DB/OL]. [2013-02-01]. <http://www.cs.nyu.edu/~roweis/data.html>.

- [4] LYU S, FRAID H. Steganalysis using color wavelet statistics and one-class support vector machines[C]// *Proceedings of the 2004 SPIE Symposium on Electronic Imaging*. San Jose, CA: SPIE, 2004: 35 - 45.
- [5] 陈丹, 陈原, 王育民. 一种针对彩色图像的泛盲掩密分析方案[J]. *电子与信息学报*, 2005, 27(10): 1542 - 1549.
- [6] TU Y L, GONG S R. Universal steganalysis using color correlation and feature fusion[C]// *Proceedings of the 2008 International Symposium on Information Science and Engineering*. Piscataway: IEEE, 2008: 107 - 111.
- [7] 毛家发, 林家骏, 戴蒙. 基于图像攻击的隐藏信息盲检测技术[J]. *计算机学报*, 2009, 32(2): 318 - 327.
- [8] 孙文颀, 刘婷婷, 张新鹏, 等. 彩色图像通用隐写分析的多类统计特征[J]. *中国图象图形学报*, 2008, 13(10): 1914 - 1917.
- [9] 毛家发, 林家骏. 基于净图描述的通用隐写分析技术[J]. *计算机学报*, 2010, 33(3): 569 - 579.
- [10] PALUS H. Representations of color images in different color spaces [M]. London: Chapman & Hall, 1998: 67 - 92.
- [11] PHILIP G. Adding images to your site[EB/OL]. [2013-01-20]. <http://philip.greenspun.com>.
- [12] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[EB/OL]. [2013-01-20]. <http://www.csie.ntu.edu.tw>.