

# 一种新的结合稀疏编码的红外图像聚类分割算法

宋长新

(青海师范大学计算机学院,青海 西宁 810008)

**摘要:**聚类作为一种重要的图像分割方法得到了大量研究,提出了一种新的结合稀疏编码的红外图像聚类分割算法,扩展了传统的基于 K-means 聚类的图像分割方法。结合稀疏编码的聚类算法能有效融合图像的局部信息,而且易于利用像素之间的内在相关性,但是对于分割会出现过分割和像素难以归类的问题。为此,在字典的学习过程中,将原子的聚类算法引入其中,有助于缩减字典中原子所属类别的数目防止出现过分割;同时将稀疏编码系数同原子对聚类中心的隶属程度相结合来判断像素所属的类别。这种处理方式能更好地实现利用像素的内在相关性进行聚类分割,并在其中自然引入了局部空间信息,达到更好分离目标区域和背景区域的目的。实验结果表明,结合稀疏编码的 K-means 聚类分割算法能更好的实现复杂背景下红外图像重要区域的准确分割提取。

**关键词:**图像分割;K-means 聚类;稀疏编码;字典聚类

**中图分类号:**TP391 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2012.11.023

## New algorithm for infrared image segmentation based on clustering combined with sparse coding

SONG Chang-xin

(Department of Computer, Qinghai Normal University, Xining 810008, China)

**Abstract:** Clustering is an important method for image segmentation, and has got much research. A new algorithm for infrared image segmentation based on clustering combined with sparse coding is proposed. The traditional image segmentation method based on K-means clustering is extended. The clustering algorithm combined with sparse coding can fuse the local information of image. The inner relationships of pixels are used. But it produces the problem of over-segmentation and difficulty in pixels classification for segmentation. The clustering method is introduced for atoms in dictionary learning. The class number of atoms in dictionary is reduced in order to avoid over-segmentation. The class of pixels is estimated by combining the sparse coefficients and the degrees of membership with the atoms to cluster center. The usage of inner relationships of pixels and the local information can help to enhance the segmentation performance of background and target area. The experimental result shows that the important area can be separated well under complex background in infrared image by this method.

**Key words:** image segmentation; K-means clustering; sparse coding; dictionary clustering

### 1 引言

红外图像在军事和民用领域有大量应用,实际应用的前提是能将红外图像划分为有意义的区域,尤其是要能提取出感兴趣的重要区域,因此,红外图像的分割好坏在红外图像应用中具有着重要作用,

例如军事目标的检测、识别和精确定位的基础就是

基金项目:青海省自然基金项目(No. 2011-z-748)资助。

作者简介:宋长新(1971-),博士,教授,主要研究方向为图像处理及数据挖掘。E-mail: songchxin@163.com

收稿日期:2012-04-14

依赖于红外图像的准确分割。由于红外图像具有纹理少、对比度差、信噪比低及复杂背景的干扰等特点,这些问题都导致了红外图像分割问题的困难。人们提出了多种分割算法,如 C-V 模型<sup>[1]</sup>、直方图<sup>[2]</sup>、聚类<sup>[3-6]</sup>等红外图像分割算法。其中基于聚类的红外图像分割算法是一种重要的分割算法,分割过程中的聚类不需过多的人工干预,很适合于自动分割的应用领域,得到了大量研究。常用的聚类分割算法有 K-means 算法、模糊 C 均值算法(FCM)等,为了更准确地从复杂背景中分割重要区域,在上述基础上提出的空间约束 FCM 分割算法<sup>[3]</sup>、模糊核聚类分割算法<sup>[4]</sup>、多重模糊聚类分割算法<sup>[5]</sup>、权重 FCM 分割算法<sup>[7]</sup>等。

但是这些算法大都是对红外图像像素点直接进行聚类分割,忽略了图像有用的局部信息,而且很少考虑将像素之间的相关性用于聚类,同时这些约束性很难和聚类算法相融合,导致分割效果并不理想。这里,我们提出了一种结合稀疏编码的红外图像聚类分割新算法。稀疏编码是最近计算机视觉领域的研究热点之一,稀疏编码方法广泛应用于图像恢复、识别及检测等计算机视觉的各个方面<sup>[8-12]</sup>。稀疏编码对 K-means 聚类算法的扩展在对向量量化进行识别时具有明显改善<sup>[8-9]</sup>,但直接将上述稀疏编码用于聚类进行图像分割有 2 个困难:①容易产生过分割,难以得到有意义的区域;②造成像素归类的判断问题。为此,我们在字典学习过程中,将原子的聚类算法引入其中,有助于缩减字典中原子所属类别的数目,同时将稀疏编码系数同原子对聚类中心的隶属程度相结合来判断像素所属的类别。这种处理方式能更好地实现类别内像素点的内在联系性,并在其中自然引入了局部空间约束,达到更好分离目标区域和背景区域的目的。实验结果表明,所提算法能更好地实现复杂背景下红外图像重要区域的准确分割提取。

## 2 稀疏编码

稀疏编码模型是指信号可以用一个过完备字典集中的少数几个基向量的线性组合来表示,其所用的基向量个数要尽可能少,也就是稀疏性<sup>[10-11]</sup>。假设数据  $X = [x_1, x_2, \dots, x_n] \in R^{m \times n}$ , 其中  $x_i$  是维数为  $m$  的第  $i$  样本,  $n$  为样本数,过完备字典  $D = [d_1, d_2, \dots, d_k] \in R^{m \times k}$  ( $k > d$ ), 其中字典  $D$  中的每个列  $d_k \in R^{m \times 1}$  称为原子(也可称基向量),为归一化向量,即  $d_k^T d_k = 1$ ,  $L$  为字典中的原子个数,稀疏编码模型就是在过完备字典  $D$  下求解信号的稀疏描述并

最小化重构误差,即:

$$F = \arg \min_{D, A} \left\{ \sum_{i=1}^n \|x_i - D\alpha_i\|_F^2 + \lambda \|\alpha_i\|_1 \right\} \quad (1)$$

其中,  $\|\cdot\|_1$  是 1-范数,表示稀疏性; $\alpha_i \in R^{L \times 1}$  是  $x_i$  的稀疏系数向量; $A = [\alpha_1, \alpha_2, \dots, \alpha_n] \in R^{L \times n}$ , 参数  $\lambda$  主要是控制重构误差和系数的稀疏性。可以看出上述稀疏性约束采用的是 1-范数,一般稀疏性用系数向量中的非零个数表示,即 0-范数,如果式(1)采用 0-范数约束则是一个 NP 难的非凸优化问题,难以求解,通常用 1-范数代替 0-范数,很多文献已经证明两者具有等价性<sup>[9-10]</sup>。上述代价函数同时对变量  $D$  和  $\alpha$  求解是非凸函数,但如果固定其中一个变量求解另一个变量,则是一个凸函数优化问题,因此,通常可以在固定  $D$  时学习  $\alpha$ ,在固定  $\alpha$  时学习  $D$ ,如此交替进行学习优化。

## 3 结合稀疏编码的聚类分割算法

### 3.1 稀疏编码观点下的 K-means 算法

K-means 聚类算法是一种重要的聚类算法,因其物理意义清晰且易于实现得到广泛的应用。K-means 聚类算法是一种无监督类别划分方法,通过最小化样本点和所属聚类中心的距离进行优化,其目标函数表示为:  $\min_V \sum_{i=1}^n \min_{c=1, \dots, J} \|x_i - \nu_c\|^2$ , 其中  $\nu_c$  为第  $c$  类聚类中心,  $J$  为类别数。该算法先对类别数进行确定,初始化聚类中心,按照目标函数迭代更新聚类中心。对于向量量化,前述 K-means 聚类目标函数也可表示为如下形式:

$$F = \min_{U, V} \sum_{i=1}^n \|x_i - Vu_i\|^2 \quad (2)$$

s. t.  $\text{card}(u_i) = 1, \|u_i\|_1 = 1$

其中,  $V = [\nu_1, \nu_2, \dots, \nu_J] \in R^{m \times J}$  是聚类中心形成的矩阵;  $U = [u_1, u_2, \dots, u_n] \in R^{J \times n}$  是样本对聚类中心的归属指标矩阵;其所有元素为非负值;  $\text{card}(u_i) = 1$  表示  $u_i$  中只有一个元素非零,  $\|u_i\|_1 = 1$  确保  $u_i$  所有元素的绝对值之和为 1,上述限制使得  $u_i$  只有一个元素为 1,其余均为 0,也就是  $u_i$  表示了样本  $x_i$  所属的聚类中心,式(2)等价于原始 K-means 目标函数。但  $\text{card}(u_i) = 1$  限制条件严格,从而使得对  $x_i$  的重构误差较大,损失部分信息;如果放松该限制,会造成样本属于过多聚类中心,这时采用反映非零元素个数的稀疏性约束代替  $\|u_i\|_1 = 1$ ,则结合稀疏编码的 K-means 算法目标函数为<sup>[7,13]</sup>:

$$F = \min_{U, V} \sum_{i=1}^n \|x_i - Vu_i\|^2 + \lambda \|u_i\|_1 \quad (3)$$

s. t.  $\|v_c\| \leq 1$

其中,  $\|v_c\| \leq 1$ , 该归一化项的约束防止所求的解产生奇异性。结合稀疏编码的 K-means 算法与原始 K-means 算法相比, 具有如下优点<sup>[7]</sup>: ①式(3)由于约束条件更为宽松, 因而较 K-means 具有更低的目标函数误差, 对样本有更低的重构误差; ②通过稀疏性能捕捉图像更显著的特征; ③图像的统计特性表明图像块本身就具有稀疏性。

### 3.2 新的结合稀疏编码的聚类分割算法(SCC)

从式(3)中可以看出, 样本  $x_i$  从单纯的像素点聚类扩展到以该像素点为中心的图像块聚类, 利于排除野点的干扰, 具有一定的空间约束性; 而且通过联合所有像素点所在像素块进行学习字典, 各个局部像素块通过字典挖掘了它们之间的一定相关性, 更为有利于利用像素的相似性进行聚类。但直接将上述稀疏编码用于聚类进行图像分割有 2 个困难: ①由于通常  $K > d$ , 字典  $D$  中的原子个数较多, 直接将  $D$  中的原子作为聚类中心会造成类别数过多, 容易产生过分割, 难以得到有意义的区域; ②得到的稀疏系数向量难以反映像素点所属的类, 造成像素归类的判断问题。

我们采用的思想是在字典学习过程中, 将原子的聚类算法引入其中, 有助于缩减字典中原子所属类别的数目, 防止产生过分割的问题; 同时将稀疏编码系数同原子对聚类中心的隶属程度相结合来判断像素所属的类别。但是字典中原子之间具有较大的相关性, 如果采用 K-means 聚类算法强制将原子划分为某一类别, 容易产生较大的聚类误差, 对后续的像素归类判断产生误导。这里我们采用模糊 C 均值聚类算法 (FCM) 进行原子的聚类。FCM 是在 K-means 算法的基础上引入了模糊隶属度的概念, 样本可以归属于多个聚类中心, 不再进行硬划分。定义样本  $x_i$  对第  $c$  类的模糊隶属度函数为  $w_{ci}$ , 且隶属度函数  $w_{ci}$  满足  $\sum_{c=1}^J w_{ci} = 1, 0 \leq w_{ci} \leq 1$ , FCM 的目标函数为:  $F = \min_{W, V} \sum_{i=1}^n \sum_{c=1}^J w_{ci}^p \|x_i - v_c\|^2$ , 式中  $p \geq 1$  是模糊权重指数, 一般取为 2。若式中的隶属度函数  $w_{ci}$  只取 0

或 1, 则为 K-means 聚类算法。通过迭代更新隶属度函数  $w_{ci}$  和聚类中心  $v_c$  最小化目标代价函数。根据以上分析, 我们将基于 FCM 的原子聚类算法引入到式(3)中进行字典学习, 给出如下目标函数:

$$F = \min_{U, V, W, Z} \left\{ \sum_{i=1}^n \|x_i - Vu_i\|^2 + \lambda \|u_i\|_1 + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\} \quad (4)$$

其中,  $\gamma$  为控制样本重构误差和原子聚类误差比例的参数;  $z_c$  表示原子的聚类中心。式(4)等号右边第一项是样本在字典下的重构误差, 反映字典和稀疏系数所含的样本信息; 第二项是稀疏性约束; 第三项表示原子的聚类, 反映原子的归类问题, 一般  $J \ll K$ , 即将原子分为  $J$  类。求解式(4)是一个联合优化问题, 如果同时优化全部 4 个参数比较困难, 我们选择交替优化方式:

(1) 固定  $Z, V, W$ , 通过拉格朗日对偶算法<sup>[14]</sup>进行稀疏编码求解  $U$ ;

(2) 固定  $Z, U, W$ , 求解字典  $V$ , 此时的优化目标函数为:

$$F_V = \min_V \left\{ \sum_{i=1}^n \|x_i - Vu_i\|^2 + \lambda \|u_i\|_1 + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\} \quad (5)$$

我们对  $V$  中的原子逐一进行求解, 固定其他原子, 则对原子  $v_l$  有如下表示:

$$F_{v_l} = \min_{v_l} \left\{ \sum_{i=1}^n \|x_i - Vu_i\|^2 + \lambda \|u_i\|_1 + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\} = \min_{v_l} \left\{ \|X - VU\|^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\} = \min_{v_l} \left\{ \| - \sum_{h \neq l} v_h q_h - v_l q_l \|^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\} \quad (6)$$

其中,  $q_l$  是稀疏系数  $U$  的行向量 ( $l = 1, \dots, J$ ),  $U = [q_1, q_2, \dots, q_J]$ , 令与原子  $v_l$  无关的项  $S = X - \sum_{h \neq l} v_h q_h$ , 则式(6)可以写成  $F_{v_l} = \min_{v_l} \left\{ \|S - v_l q_l\|^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \right\}$ , 在  $v_l^T v_l = 1$  约束下, 根据拉格朗日乘子法有:

$$F_{v_l, \eta} = \|S - v_l q_l\|^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 + \eta (v_l^T v_l - 1) = \text{tr} \left( \begin{array}{l} SS^T + \gamma \sum_{c=1}^J w_{cl}^p z_c z_c^T - (Sq_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c) v_l^T \\ - v_l (q_l S^T + \gamma \sum_{c=1}^J w_{cl}^p z_c^T) + v_l (q_l q_l^T + \gamma \sum_{c=1}^J w_{cl}^p - \eta) v_l^T + \eta \end{array} \right) \quad (7)$$

其中,  $\eta$  为调节参数, 将上式对  $v_l$  微分  $\frac{\partial F_{v_l, \eta}}{\partial v_l} = 0$ , 有:

$$v_{l,0} = (Sq_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c) v_l^T \cdot (q_l q_l^T +$$

$$\gamma \sum_{c=1}^J w_{cl}^p - \eta)^{-1} \quad (8)$$

然后,归一化  $v_l = \frac{v_{l,0}}{\|v_{l,0}\|_2}$ , 如此对  $V$  中的原子逐一进行更新;

(3) 固定  $Z, V, U$ , 通过  $\min_W \{ \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \}$  求解  $W$ , 与传统的 FCM 算法相似, 此时  $W$  的优化公式为:

$$w_{cl} = \frac{(\|v_l - z_c\|^{-2})^{1/(p-1)}}{\sum_{c=1}^J (\|v_l - z_c\|^{-2})^{1/(p-1)}} \quad (9)$$

(4) 固定  $V, W, U$ , 通过  $\min_Z \{ \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|^2 \}$  求解  $Z$ , 与传统的 FCM 算法相似, 此时  $Z$  的优化公式为:

$$z_c = \frac{\sum_{l=1}^K w_{cl}^p v_l}{\sum_{l=1}^K w_{cl}^p} \quad (10)$$

通过上述的迭代优化, 可以求出字典  $V$ 、字典聚类中心  $Z$ 、稀疏系数  $U$  及隶属度  $W$ 。

如果要完成聚类分割, 需要判断像素点所属的类, 由于稀疏系数反映了样本  $x_i$  对字典中各个原子的权重大小, 而字典中各个原子对聚类中心有不同的隶属度程度, 所以我们定义样本  $x_i$  对于聚类中心  $z_c$  的归属度为:

$$r_{ci} = \sum_{l=1}^K (|u_{i,l}| \cdot w_{cl}) = (|u_i|)^T \cdot w_c \quad (11)$$

其中,  $w_c = [w_{c1}^p, w_{c2}^p, \dots, w_{cK}^p]$ ,  $|\cdot|$  为元素的绝对值。根据样本对于各个聚类中心的归属度按照最大化原则进行分类, 即  $I_i = \arg \max_c \{r_{1i}, r_{2i}, \dots, r_{ji}\}$ , 可得到最终聚类结果, 从而完成红外图像聚类分割。

#### 4 仿真结果与分析

采用文中提出的红外图像分割算法进行实验, 所用图像为机载对地飞机和道路红外图像, 如图 1 (a) 和图 2 (a) 所示, 并采用 K-means 和 FCM 算法分割结果作为对比, 对比重要区域的分割效果 (分别为飞机和道路)。图 1 所用参数为: 聚类类别数为 8、稀疏参数  $\lambda$  取 0.001、模糊权重指数  $p=2$ 、图像块大小为  $3 \times 3$ 、字典大小为 50,  $\gamma=0.1$ , K-means、FCM 和本文算法聚类分割重要区域结果分别如图 1 (b) ~ 图 1 (d) 所示。图 2 所用参数为: 聚类类别数为 3、稀疏参数  $\lambda$  取 0.001、模糊权重指数  $p=2$ 、图像块大小为  $3 \times 3$ 、字典大小为 50,  $\gamma=0.2$ , K-means、FCM 和本文算法聚类分割重要区域结果分别如图 2

(b) ~ 图 2 (d) 所示。

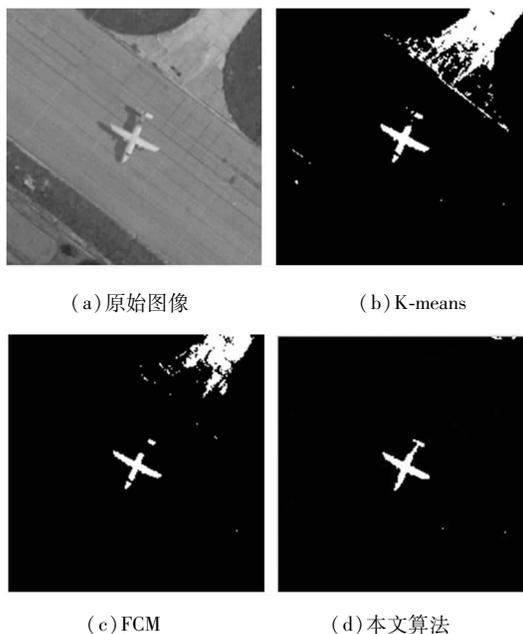


图 1 实验仿真结果

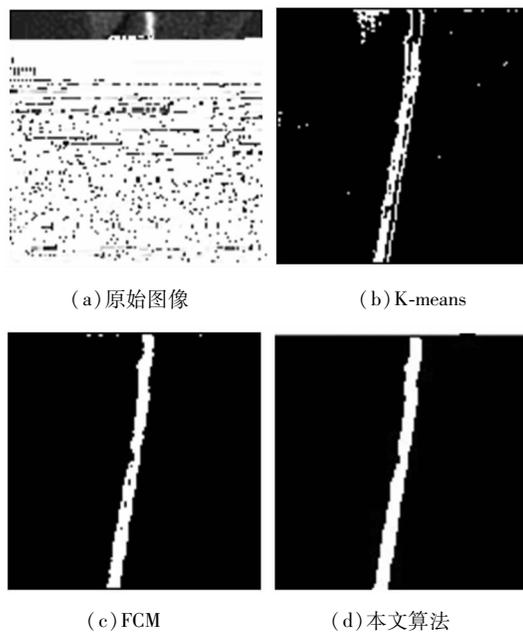


图 2 实验仿真结果

红外图像分割的效果一般无法进行定量评价, 一个广泛采用的评价原则就是看能否分割出期望的或者重要的区域, 并且尽量区分背景区域和目标区域。从图 1 的对比中可以看出, K-means 和 FCM 算法分割结果并不理想, K-means 算法有大量背景杂点, FCM 算法较 K-means 算法只能抑制部分杂点, 但是飞机轮廓都不完整, 会严重影响后续的目标判断; 而本文所提的算法由于考虑了图像局部信息和像素之间的内在相关性, 飞机轮廓保持比较完整, 同时杂点被大量抑制, 但由于边缘部分局部信息的限

制性,仍存在一点杂点干扰(图1(d)右上边缘所示)。对比图2结果,K-means和FCM算法虽分割出道路的大致区域,但道路区域不够清晰、完整;本文所提算法分割得到的道路区域更为接近实际情况,而且基本消除了背景及杂点的干扰,分割效果比前两种算法理想,为后续的检测识别跟踪奠定了基础。

## 5 结束语

本文提出了一种新的结合稀疏编码的红外图像聚类分割算法,并给出了一种交替优化算法,联合学习字典、稀疏系数和隶属度函数,在此基础上完成分割。本文所提算法充分利用了像素的内在相关性和局部信息,实验结果也表明,本文方法能够有效提高红外图像重要区域的分割效果,很大程度上降低了背景对目标分割的干扰,适于进行复杂背景下红外目标的准确分割,利于后续图像处理任务。但是该算法也存在一些需要解决的问题,如字典大小及约束性参数的选择等。我们进一步的方向是考虑将核方法引入到目标函数中,提高分割的效果。

## 参考文献:

- [1] Li Zhiguo, Hao Xinran. Automatic segmentation method of infrared images based on C-V model[J]. *Laser & Infrared*, 2011, 41(3): 356-358. (in Chinese)  
李志国,郝欣然. 基于C-V模型的红外图像自动分割方法研究[J]. *激光与红外*, 2011, 41(3): 356-358.
- [2] Chen Liang, Ding Guohui, Guo Lei. Image thresholding based on mutual recognition of histogram[J]. *J. Infrared Millim. Waves*, 2011, 30(1): 80-84. (in Chinese)  
陈亮,丁国辉,郭雷. 基于直方图互确认的图像阈值化分割[J]. *红外与毫米波学报*, 2011, 30(1): 80-84.
- [3] Wu Jin, Li Juan, Liu Jian, et al. Infrared image segmentation via fast fuzzy C-Means with spatial Information[C]. *International Conference on Robotics and Biomimetics*, 2004: 742-745.
- [4] Song Changxin. Improved weighed KFCM based on spatially constrained for infrared image segmentation[J]. *Microelectronics & Computer*, 2009, 26(5): 60-63. (in Chinese)  
宋长新. 改进的空间约束加权模糊核聚类红外图像分割[J]. *微电子学与计算机*, 2009, 26(5): 60-63.
- [5] Dong Yanfei, Ren Qimei, Zhang Junfeng. Algorithm for infrared target segmentation based on multiple fuzzy clustering[J]. *Laser & Infrared*, 2010, 40(5): 554-557. (in Chinese)  
董燕飞,任琦梅,张俊峰. 一种基于多重模糊聚类的红外目标分割算法[J]. *激光与红外*, 2010, 40(5): 554-557.
- [6] Liu Cuiyin, Zhang Xiuqiong, Yin Xing, et al. K-means feature divergence genetic for infrared image segmentation[J]. *Laser & Infrared*, 2011, 41(11): 107-111. (in Chinese)  
柳翠寅,张秀琼,银星,等. 基于特征散度K-means红外图像分割遗传算法[J]. *激光与红外*, 2011, 41(11): 107-111.
- [7] Jianchao Fan, Min Han, Jun Wang. Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation[J]. *Pattern Recognition*, 2009, 42(11): 2527-2540.
- [8] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]. *CVPR*, 2009: 1794-1801.
- [9] Zhuang L S, Gao H Y, Liu C, et al. Nonnegative sparse locally linear coding[J]. *Journal of Software*, 2011, 22(Suppl. (2)): 89-95. (in Chinese)  
庄连生,高浩瀚,刘超,等. 非负稀疏局部线性编码[J]. *软件学报*, 2011, 22(增刊(2)): 89-95.
- [10] J Wright, Y Ma, J Mairal, et al. Sparse representations for computer vision and pattern recognition[J]. *Proc. IEEE*, 2010, 98(6): 1031-1044.
- [11] Zhao Jiajia, Tang Zhengyuan, Yang Jie, et al. Infrared small target detection based on image sparse representation[J]. *J. Infrared Millim. Waves*, 2011, 30(2): 156-162. (in Chinese)  
赵佳佳,唐峥远,杨杰,等. 基于图像稀疏表示的红外小目标检测算法[J]. *红外与毫米波学报*, 2011, 30(2): 156-162.
- [12] Ignacio Ramirez, Pablo Sprechmann, Guillermo Sapiro. Classification and clustering via dictionary learning with structured incoherence and shared features[C]. *CVPR*, 2010: 3501-3508.
- [13] Shenghua Gao, Ivor Waihung Tsang, Liangtien Chia. Kernel sparse representation for image classification and face recognition[C]. *ECCV*, 2010: 1-14.
- [14] H Lee, A Battle, R Raina, et al. Efficient sparse coding algorithms[C]. *NIPS*, 2006.