

## 结合稀疏编码和空间约束的红外图像聚类分割研究\*

宋长新<sup>†</sup> 马克 秦川 肖鹏

(青海师范大学计算机学院, 西宁 810008)

(2012年5月3日收到; 2012年10月4日收到修改稿)

提出了结合稀疏编码和空间约束的红外图像聚类分割新算法, 在稀疏编码的基础上融合聚类算法, 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分割算法能有效融合图像的局部信息, 便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; 考虑到像素及其邻域像素具有类别属性一致性的特点, 引入了空间类别属性约束信息, 并给出了一种交替优化算法. 联合学习字典、稀疏系数、聚类中心和隶属度, 将稀疏编码系数同原子对聚类中心的隶属程度相结合, 构造像素归属度来判断像素所属的类别. 实验结果表明, 该方法能够有效提高红外图像重要区域的分割效果, 具有较好的鲁棒性.

**关键词:** 图像分割, 稀疏编码, 聚类, 空间约束

**PACS:** 07.05.Pj, 87.57.N-, 95.75.Mn

**DOI:** 10.7498/aps.62.040702

## 1 引言

红外图像的分割就是根据一致性准则将图像划分为互不重叠的各具特性的区域, 并区分出感兴趣目标的过程<sup>[1,2]</sup>, 是红外图像在军事和民用中应用的前提. 红外图像分割的好坏关系到目标检测、识别和精确定位等任务. 通常采取的算法有阈值化分割、基于边缘的分割及基于区域的分割等<sup>[1]</sup>. 由于红外图像具有纹理少、对比度差、信噪比低及复杂背景干扰等特点, 导致了红外图像分割问题的困难<sup>[3]</sup>. 为了克服这些困难, 提高分割的准确性, 人们提出了许多改进的分割算法, 如结合直方图或熵的阈值化分割<sup>[1,4]</sup>、基于 C-V 模型的水平集分割<sup>[5]</sup>、空间约束聚类分割<sup>[6,7]</sup>等分割算法. 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, 分割过程中不需过多的人工干预, 具有较好的抗噪性, 适于红外图像自动分割, 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, 主要采用的是词袋 (BOF: bag of features) 的思想<sup>[8]</sup>, 为了更准确地从复杂背景中分割重要区域, 后又在上述聚类算法基础上提出

了空间约束 FCM 分割算法<sup>[7,9,10]</sup>、模糊核聚类分割算法<sup>[3]</sup>等. 这些算法本质上都是  $K$ -means 算法, 对非凸的数据结构和彼此交叠严重的数据存在局限性<sup>[8]</sup>, 也较少考虑像素点在空间上的依赖关系, 导致分割效果并不理想.

本文结合稀疏编码算法提出了一种红外图像聚类分割算法. 稀疏编码是在超完备基上的一种稀疏表示, 是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、识别及检测等模式识别的各个方面<sup>[11-14]</sup>. 稀疏编码对  $K$ -means 聚类算法的扩展在向量量化构造 BOF 特征进行识别时具有明显改善<sup>[11,15,16]</sup>, 但直接将上述稀疏编码用于聚类进行图像分割容易产生过分割, 难以得到有意义的区域, 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. 这种处理方式能通过字典更好地反映类别内像素的内在联系性, 由于采用局部块进行处理, 自然地引入了局部信息, 而且这些约束条件与聚类算法相融合也较为自然, 在此基础

\* 青海省自然科学基金 (批准号: 2011-z-748) 和青海省 135 高层次人才培养基金资助的课题.

<sup>†</sup> 通讯作者. E-mail: songchxin@163.com

上进一步根据图像特点对图像中像素的邻域像素施加空间类别属性约束提高分割质量. 实验结果表明, 所提算法能更好地实现复杂背景及干扰下红外图像重要区域的准确分割提取.

## 2 稀疏编码基本概念

所谓稀疏编码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<sup>[12]</sup>, 这些稀疏系数及其对应的字典可以反映信号的主要特征和内在结构. 假设维数为  $m$  的  $n$  个样本形成的数据  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ , 其中  $\mathbf{x}_i$  是第  $i$  个样本, 有过完备字典  $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K] \in \mathbb{R}^{m \times K}$  ( $K > m$ ), 其中  $\mathbf{D}$  中的每个列向量  $\mathbf{d}_k \in \mathbb{R}^{m \times 1}$  称为原子(也可称基向量), 满足约束条件  $\mathbf{d}_k^T \mathbf{d}_k \leq 1$ ,  $K$  为字典中的原子个数, 稀疏编码模型就是通过最小化下面重构误差求解信号的稀疏描述和过完备字典, 即

$$\arg \min_{\mathbf{D}, \mathbf{A}} \left\{ \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{D} \cdot \boldsymbol{\alpha}_i\|_F^2 + \lambda \|\boldsymbol{\alpha}_i\|_1 \right\}, \quad (1)$$

其中  $\|\cdot\|_F$  表示 Frobenius 范数(简称 F 范数),  $\|\cdot\|_1$  是 1 范数, 表示稀疏性约束,  $\boldsymbol{\alpha}_i \in \mathbb{R}^{K \times 1}$  是  $\mathbf{x}_i$  的稀疏系数向量,  $\mathbf{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n] \in \mathbb{R}^{K \times n}$ , 参数  $\lambda$  平衡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 (1) 式采用 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>[12]</sup>. 上述代价函数同时对变量  $\mathbf{D}$  和  $\boldsymbol{\alpha}$  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, 一般分为两步求解: 字典学习和稀疏编码, 在固定字典  $\mathbf{D}$  时求解稀疏系数  $\boldsymbol{\alpha}$ , 在固定稀疏系数  $\boldsymbol{\alpha}$  时学习字典  $\mathbf{D}$ , 多次迭代进行优化, 直到收敛. 当前稀疏编码应用的难点在于根据问题设计反映任务需要的字典.

## 3 空间约束下结合稀疏编码的聚类算法

### 3.1 稀疏编码观点下的 $K$ -means 算法

$K$ -means 聚类算法是一种无监督类别划分方法, 通过最小化样本和所属聚类中心的距离进行优化, 其目标函数表示为:  $\min_V \sum_{i=1}^n \min_{c=1, \dots, J} \|\mathbf{x}_i - \mathbf{v}_c\|_F^2$ , 其

中  $\mathbf{v}_c$  为第  $c$  类聚类中心,  $J$  为类别数, 是向量量化构造 BOF 特征的一个重要步骤.  $K$ -means 聚类算法的目标函数等价于如下形式:

$$\min_{\mathbf{U}, \mathbf{V}} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{V} \mathbf{u}_i\|_F^2, \quad \text{s.t. } \text{card}(\mathbf{u}_i) = 1, \quad \|\mathbf{u}_i\|_1 = 1, \quad (2)$$

式中  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_J] \in \mathbb{R}^{m \times J}$  是聚类中心形成的矩阵;  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in \mathbb{R}^{J \times n}$  是样本对聚类中心的归属指标矩阵, 其所有元素为非负值; 函数  $\text{card}(\mathbf{x})$  表示  $\mathbf{x}$  中的非零元素个数,  $\text{card}(\mathbf{u}_i) = 1$  表示  $\mathbf{u}_i$  中只有一个元素非零,  $\|\mathbf{u}_i\|_1 = 1$  确保  $\mathbf{u}_i$  所有元素的绝对值之和为 1, 上述限制使得  $\mathbf{u}_i$  只有一个元素为 1, 其余均为 0, 也就是  $\mathbf{u}_i$  确定了样本  $\mathbf{x}_i$  所属的聚类中心,  $\mathbf{u}_i$  中的非零元素所在的位置对应了  $\mathbf{x}_i$  所属的聚类中心, 可以看出 (2) 式与原始  $K$ -means 目标函数是等价的. 但  $\text{card}(\mathbf{u}_i) = 1$  限制条件严格, 从而使得对  $\mathbf{x}_i$  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  $\|\mathbf{u}_i\|_1 = 1$ , 则在稀疏编码观点下的  $K$ -means 聚类算法目标函数为<sup>[11,17]</sup>

$$\min_{\mathbf{U}, \mathbf{V}} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{V} \mathbf{u}_i\|_F^2 + \lambda \|\mathbf{u}_i\|_1, \quad \text{s.t. } \|\mathbf{v}_c\|_2^2 \leq 1, \quad (3)$$

其中  $\|\mathbf{v}_c\|_2^2 \leq 1$ , 该归一化项的约束防止产生奇异解. 稀疏编码观点下的  $K$ -means 算法与原始  $K$ -means 算法相比, 具有如下优点<sup>[11]</sup>: 1) (3) 式由于约束条件更为宽松, 较  $K$ -means 具有更低的目标函数误差, 对样本有更低的 reconstrution 误差, 可保留样本更多的信息; 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, 更符合图像的特性. 可以看出采用 (3) 式的样本只与几个聚类中心有关(对应于稀疏系数的非零元素), 与其他的聚类中心不存在关系.

### 3.2 结合稀疏编码的聚类分割算法 (SCC)

从 (3) 式中可以看出, 样本  $\mathbf{x}_i$  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建立了某种联系, 挖掘了它们之间一定的内在相关性, 有利于利用像素的相似性

进行聚类; 稀疏系数可以捕捉图像更显著的特征, 抑制图像中的非结构信息, 如杂点、噪声和其他一些外来干扰. 但直接将上述稀疏编码用于聚类进行图像分割有两个困难: 1) 由于  $K > m$ , 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, 容易导致过分割, 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 FCM 算法进行原子的聚类. 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, 不再是硬划分. 定义样本  $\mathbf{x}_i$  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:  $\min_{\mathbf{W}, \mathbf{V}} \sum_{i=1}^n \sum_{c=1}^J w_{ci}^p \|\mathbf{x}_i - \mathbf{v}_c\|_F^2$ , 式中  $p \geq 1$  是隶属度指数, 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  $\mathbf{v}_c$  最小化目标代价函数. FCM 与 (3) 式不同之处在于: FCM 算法中样本通过欧氏距离与所有聚类中心有联系, 而 (3) 式中样本通过稀疏系数只与几个聚类中心有关, 这种基于稀疏系数的相关性正如上所述, 反映的是图像内在结构的相似性. 根据以上分析, 我们将基于 FCM 的原子聚类算法引入到 (3) 式中进行字典学习, 给出如下目标函数:

$$\begin{aligned} \min_{\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}} \left\{ \sum_{i=1}^n \left( \|\mathbf{x}_i - \mathbf{V}\mathbf{u}_i\|_F^2 + \lambda \|\mathbf{u}_i\|_1 \right) \right. \\ \left. + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|\mathbf{v}_l - \mathbf{z}_c\|_F^2 \right\}, \\ \text{s.t. } \|\mathbf{v}_c\|_2^2 \leq 1, \quad \sum_{c=1}^J w_{cl} = 1, \\ 0 \leq w_{cl} \leq 1, \end{aligned} \quad (4)$$

其中  $\gamma$  为控制样本重构误差和原子聚类误差比例的参数,  $\mathbf{z}_c$  表示原子的聚类中心. 上式中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. 如果  $\gamma$  过大, 主要强调原子的聚类, 则学习的字典  $V$  将有明显的结构信息, 相对来说会造成样本重构误差过大, 弱化字典  $V$  反映的样本信息; 反之, 如果  $\gamma$  过小, 则学习的字典  $V$  将更多地反映样本信息, 从而弱化结构信息, 难以取得较好的聚类效果.

### 3.3 空间约束下的 SCC 分割算法

图像分割中的空间约束是一个重要的信息, 文献 [8—10, 15,16] 相关结论表明, 考虑空间约束能有效提高实验效果, 对于红外图像来说, 其像素及其邻域中的像素更具有类别属性一致性的特点. 我们将图像像素的空间类别属性约束引入 (4) 式中, 给出下面的目标函数:

$$\begin{aligned} \min_{\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}} \left\{ \sum_{i=1}^n \left( \|\mathbf{x}_i - \mathbf{V}\mathbf{u}_i\|_F^2 + \lambda \|\mathbf{u}_i\|_1 \right) \right. \\ \left. + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|\mathbf{v}_l - \mathbf{z}_c\|_F^2 \right. \\ \left. + \frac{\rho}{N_R} \sum_{i=1}^n \sum_{c=1}^J r_{ci} \sum_{r \in N_{i,t}} \|\mathbf{x}_r - \mathbf{z}_c\|_F^2 \right\}, \\ \text{s.t. } \|\mathbf{v}_c\|_2^2 \leq 1, \quad \sum_{c=1}^J w_{cl} = 1, \\ 0 \leq w_{cl} \leq 1, \end{aligned} \quad (5)$$

其中  $\rho$  为空间约束性惩罚因子.  $N_{i,t}$  以  $\mathbf{x}_i$  为中心窗口, 大小为  $(2t+1) \times (2t+1)$  的邻域.  $N_R$  为邻域中像素的个数.  $r_{ci}$  是样本  $\mathbf{x}_i$  对于聚类中心  $\mathbf{z}_c$  的归属度, 其他参数同 (4) 式所述. 上式第四项是空间类别属性约束, 反映某像素点及其邻域像素应尽可能属于相同的一类. 为了求解上式, 首先需要定义样本对于聚类中心的归属度  $r_{ci}$ , 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  $\mathbf{x}_i$  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程度, 所以我们定义样本  $\mathbf{x}_i$  对于聚类中心  $\mathbf{z}_c$  的归属度为

$$r_{ci} = \sum_{l=1}^K (|\mathbf{u}_{i,l}| \cdot w_{cl}^p) = \mathbf{w}_c^T \cdot |\mathbf{u}_i|, \quad (6)$$

其中  $\mathbf{w}_c = [w_{c1}^p; w_{c2}^p; \dots; w_{cK}^p]$ ,  $|\cdot|$  为元素的绝对值. 我们将 (6) 式代入到 (5) 式中, 有

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}} \left\{ \sum_{i=1}^n \left( \|\mathbf{x}_i - \mathbf{V}\mathbf{u}_i\|_F^2 + \lambda \|\mathbf{u}_i\|_1 \right) \right.$$

$$\begin{aligned}
 & + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \\
 & + \frac{\rho}{N_R} \sum_{i=1}^n \sum_{c=1}^J \sum_{l=1}^K (|u_{i,l}| \cdot w_{cl}^p) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \Big\}, \\
 \text{s.t. } & \|v_c\|_2^2 \leq 1, \quad \sum_{c=1}^J w_{cl} = 1, \\
 & 0 \leq w_{cl} \leq 1.
 \end{aligned} \tag{7}$$

如果同时优化全部 4 个参数比较困难, 这里采用交替优化迭代方法求解  $Z, U, V, W$ . 随机选取样本初始字典后, 首先求解稀疏系数; 然后更新字典; 最后计算聚类中心和隶属度. 具体求解算法如下.

1) 稀疏编码: 固定  $Z, V, W$ , 即

$$\begin{aligned}
 \min_{u_i} & \left\{ \|x_i - V u_i\|_F^2 + \lambda \|u_i\|_1 \right. \\
 & \left. + \frac{\rho}{N_R} \sum_{c=1}^J \sum_{l=1}^K (|u_{i,l}| \cdot w_{cl}^p) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right\} \\
 = \min_{u_i} & \left\{ \|x_i - V u_i\|_F^2 \right. \\
 & \left. + \lambda \|u_i\|_1 + \lambda_1 \sum_{l=1}^K (\varpi_l \cdot |u_{i,l}|) \right\},
 \end{aligned} \tag{8}$$

其中  $\lambda_1 = \frac{\rho}{N_R}$ .  $\varpi_l = \sum_{c=1}^J w_{cl}^p \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2$ . 可以看出上式转化为交叠群 Lasso 模型 [18]. 通过加速梯度下降算法进行求解  $U$ , 具体求解算法可以参考文献 [18].

2) 字典学习: 固定  $Z, U, W$ , 学习字典  $V$ , 此时的优化目标函数为

$$\begin{aligned}
 F_V = \min_V & \left\{ \sum_{i=1}^n \|x_i - V u_i\|_F^2 + \lambda \|u_i\|_1 \right. \\
 & \left. + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\}.
 \end{aligned} \tag{9}$$

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:

$$\begin{aligned}
 F_{v_l} = \min_{v_l} & \left\{ \sum_{i=1}^n \|x_i - V u_i\|_F^2 + \lambda \|u_i\|_1 \right. \\
 & \left. + \gamma \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\} \\
 = \min_{v_l} & \left\{ \|X - V U\|_F^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\} \\
 = \min_{v_l} & \left\{ \left\| X - \sum_{h \neq l} v_h q_h - v_l q_l \right\|_F^2 \right\}
 \end{aligned}$$

$$\left. + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\}, \tag{10}$$

其中  $q_t$  是稀疏系数  $U$  的行向量 ( $t = 1, \dots, J$ ),  $U = [q_1; q_2; \dots; q_J]$ . 令与原子  $v_l$  无关的项  $S = X - \sum_{h \neq l} v_h q_h$ , 则 (10) 式可以写成

$$F_{v_l} = \min_{v_l} \left\{ \|S - v_l q_l\|_F^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\}.$$

我们知道对 F 范数有  $\|A\|_F^2 = \text{tr}(AA^T)$ , 其中  $\text{tr}(\cdot)$  表示矩阵的迹. 对于约束  $\|v_l\|_2^2 \leq 1$ , 即  $v_l^T v_l \leq 1$ , 可通过拉格朗日对偶法 [19] 进行求解, 但需花费大量运算时间, 这里将约束强化为  $\|v_c\|_2^2 = 1$ , 即  $v_l^T v_l \leq 1$ , 能有效提高运算速度. 根据拉格朗日乘子法有

$$\begin{aligned}
 F_{v_l, \eta} & = \|S - v_l q_l\|_F^2 + \gamma \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 + \eta (1 - v_l^T v_l) \\
 & = \text{tr} \left( (S - v_l q_l) (S - v_l q_l)^T \right) \\
 & \quad + \gamma \sum_{c=1}^J w_{cl}^p \text{tr} \left( (v_l - z_c) (v_l - z_c)^T \right) + \eta (1 - v_l^T v_l) \\
 & = \text{tr} \left( S S^T + \gamma \sum_{c=1}^J w_{cl}^p z_c z_c^T - \left( S q_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c \right) v_l^T \right. \\
 & \quad \left. - v_l \left( q_l S^T + \gamma \sum_{c=1}^J w_{cl}^p z_c^T \right) \right) \\
 & \quad + v_l \left( q_l q_l^T + \gamma \sum_{c=1}^J w_{cl}^p - \eta \right) v_l^T + \eta,
 \end{aligned} \tag{11}$$

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

$$\begin{aligned}
 v_{l,0} & = \left( S q_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c \right) \\
 & \quad \times \left( q_l q_l^T + \gamma \sum_{c=1}^J w_{cl}^p - \eta \right)^{-1}.
 \end{aligned} \tag{12}$$

然后, 归一化

$$v_l = \frac{v_{l,0}}{\|v_{l,0}\|_2} = \frac{\left( S q_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c \right)}{\left\| \left( S q_l^T + \gamma \sum_{c=1}^J w_{cl}^p z_c \right) \right\|_2},$$

如此对  $V$  中的原子逐一进行更新.

3) 更新聚类中心: 固定  $V, W, U$ , 通过

$$\min_Z \left\{ \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right\}$$

$$\begin{aligned}
 & + \frac{\rho}{N_R} \sum_{i=1}^n \sum_{c=1}^J (w_c^T \cdot |u_i|) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \Big\} \\
 \Rightarrow I_{z_c} & \\
 = \min_{z_c} & \left\{ \sum_{l=1}^K w_{cl}^p \|v_l - z_c\|_F^2 \right. \\
 & \left. + \frac{\rho}{N_R} \sum_{i=1}^n (w_c^T \cdot |u_i|) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right\}. \quad (13)
 \end{aligned}$$

求解  $Z$ , 通过  $\frac{\partial I_{z_c}}{\partial z_c} = 0$ , 可以得到此时  $Z$  的优化公式为

$$z_c = \frac{\sum_{l=1}^K w_{cl}^p v_l + \frac{\rho}{N_R} \sum_{i=1}^n (w_c^T \cdot |u_i|) \sum_{r \in N_{i,t}} x_r}{\sum_{l=1}^K w_{cl}^p + \frac{\rho(2t+1)^2}{N_R} \sum_{i=1}^n (w_c^T \cdot |u_i|)}. \quad (14)$$

4) 更新隶属度: 固定  $Z, V, U$ , 通过

$$\min_W \left\{ \sum_{l=1}^K \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right.$$

$$\begin{aligned}
 & \left. + \frac{\rho}{N_R} \sum_{i=1}^n \sum_{c=1}^J \sum_{l=1}^K (|u_{i,l}| \cdot w_{cl}^p) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right\} \\
 \Rightarrow \min_{w_{cl}} & \left\{ \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 \right. \\
 & \left. + \frac{\rho}{N_R} \sum_{c=1}^J \sum_{i=1}^n (|u_{i,l}| \cdot w_{cl}^p) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right\}, \quad (15)
 \end{aligned}$$

$$\text{s.t. } \sum_{c=1}^J w_{cl} = 1, \quad 0 \leq w_{cl} \leq 1.$$

求解  $W$ , 根据拉格朗日乘子法有

$$\begin{aligned}
 W_{w_{cl}} & = \sum_{c=1}^J w_{cl}^p \|v_l - z_c\|_F^2 + \frac{\rho}{N_R} \sum_{c=1}^J \sum_{i=1}^n (|u_{i,l}| \cdot w_{cl}^p) \\
 & \times \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 + \vartheta \left( \sum_{c=1}^J w_{cl} - 1 \right),
 \end{aligned}$$

由  $\frac{\partial W_{w_{cl}}}{\partial w_{cl}} = 0$ , 此时  $W$  的优化公式为

$$w_{cl} = \frac{\left( \|v_l - z_c\|_F^2 + \frac{\rho}{N_R} \sum_{i=1}^n (|u_{i,l}|) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right)^{-1/(p-1)}}{\sum_{c=1}^J \left( \|v_l - z_c\|_F^2 + \frac{\rho}{N_R} \sum_{i=1}^n (|u_{i,l}|) \sum_{r \in N_{i,t}} \|x_r - z_c\|_F^2 \right)^{-1/(p-1)}}. \quad (16)$$

通过上述的多次迭代优化过程, 可以求出字典  $V$ , 稀疏系数  $U$ , 字典聚类中心  $Z$  及隶属度  $W$ . 最后完成聚类分割, 需要判断像素点所属的类, 计算样本  $x_i$  对于聚类中心  $z_c$  的归属度:  $r_{ci} = \sum_{l=1}^K (|u_{i,l}| \cdot w_{cl}^p) = w_c^T \cdot |u_i|$ . 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  $l_i = \arg \max_c \{r_{1i}, r_{2i}, \dots, r_{Ji}\}$ , 可得到最终聚类结果, 从而完成红外图像的聚类分割.

## 4 实验结果与分析

为了验证本文所提红外图像分割算法的性能, 我们采用机载对地面飞机和道路的红外图像进行实验, 分别如图 1(a) 和图 3(a) 所示, 并采用  $K$ -means, FCM, 模糊核聚类 (KFCM) 和结合空间信息的模糊聚类 (SFCM) 分割算法结果作为对比. 首先, 我们讨论一下具体的参数选择, FCM, KFCM, SFCM 和本文算法中的隶属度指数  $p$  统一取为 2, FCM, KFCM 及 SFCM 算法的迭代次数为 100, 本文算法中稀疏性参数  $\lambda$  取 0.001,  $\gamma = 0.2$ ,  $\eta = 0.5$ .

空间约束目的是使得邻域内的所有像素尽量具有相同的类别, 对于 SFCM 和本文算法空间约束的邻域大小选择, 正如文献 [8—10] 中所论述的, 选择较大的邻域进行空间约束, 如  $7 \times 7, 9 \times 9$  等, 虽能使得分割结果更为平滑完整, 较好地消除野点及噪声, 但难以避免会丢失图像中固有的边缘或较小目标等信息. 一般来说, 如果目标较大, 且较为平滑完整, 则可适当地选择较大邻域. 对于本文的图像, 我们选择大小为  $3 \times 3$  的邻域用于 SFCM 和本文算法的实验; 其中对于空间约束惩罚系数  $\alpha$ , 选取过小的  $\alpha$  将难以起到空间约束的作用, 选取过大的  $\alpha$  将抹杀图像像素的固有属性, 这里我们在 SFCM 和本文算法中选取空间约束惩罚系数  $\alpha = 4$ . KFCM 算法所用的核函数及其参数在后面实验中给出. 另外, 本文算法由于采用像素邻域的图像块向量代替了原来的像素值, 可以看出如果取较大的图像块, 则容易抹杀了图像的细节信息, 本文算法实验中选取图像块大小为  $3 \times 3$ . 如文献 [11, 12, 20] 所示, 并对数据进行归一化处理, 字典大小可以通过冗余性指数  $N$  和图像块向量维数  $m$  确定, 即  $K = N \times m$ , 冗

余性指数  $N$  反映了稀疏系数的稀疏性,  $N$  越大则越稀疏, 本文中  $N$  取值范围可设为 5—10.  $K$ -means, FCM, KFCM, SFCM 及本文算法中如何选择类别数  $J$  是一个普遍的难题, 与诸多文献类似, 我们通过实验根据不同的问题选择不同的类别数. 针对多组实验, 本文算法的迭代次数统一选为 50, 其收敛性将在后面予以讨论.

实验中计算机为 i5-2.5 GHz Intel 酷睿处理器,

4 GB 内存, 仿真工具为 Matlab 7.10. 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. 本实验主要观察分割算法能否分割出重要区域, 保持重要区域的完整性, 抑制非重要区域对重要区域的误导. 可以看出图 1(a) 和图 2(a) 所示的重要区域分别为飞机和道路.

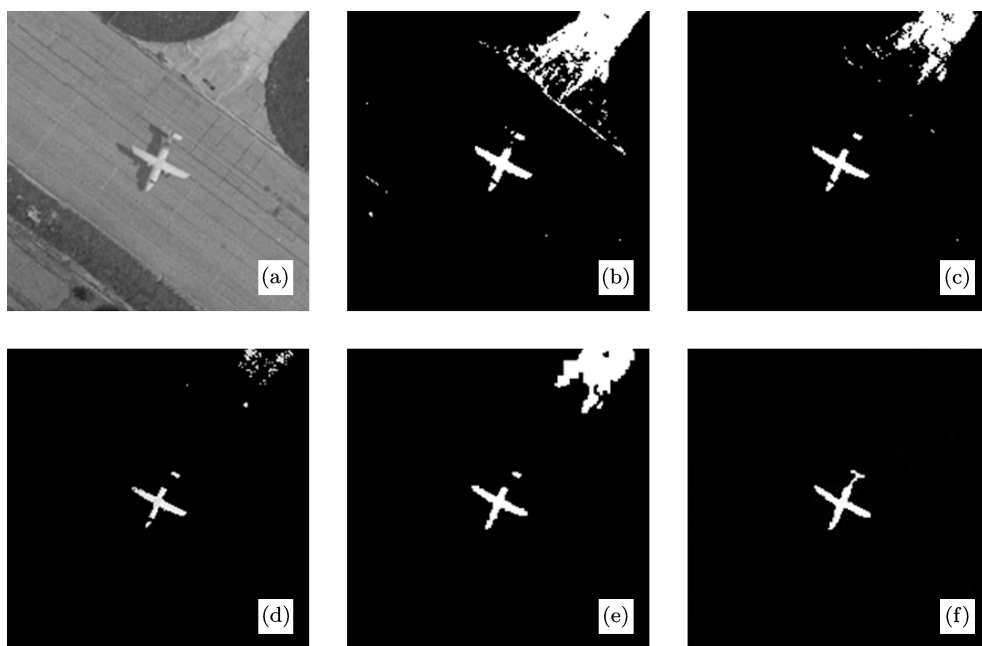


图 1 分割结果对比 (a) 原始图像; (b)  $K$ -means 分割; (c) FCM 分割; (d) KFCM 分割; (e) SFCM 分割; (f) 本文算法

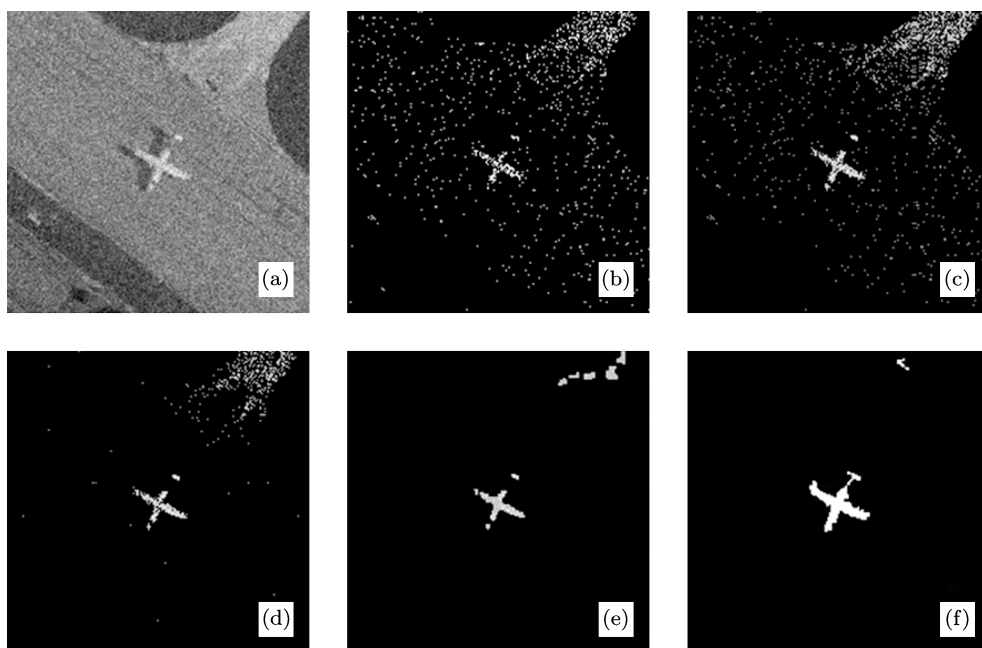


图 2 含噪图像分割结果对比 (a) 噪声图像; (b)  $K$ -means 分割; (c) FCM 分割; (d) KFCM 分割; (e) SFCM 分割; (f) 本文算法

对于图 1(a) 的分割实验,  $K$ -means, FCM, KFCM, SFCM 及本文算法中采用的类别数  $J$  取为 8, 本文算法所采用图像块大小为  $3 \times 3$ , 字典大小为 50. 从图 1 的对比中可以看出,  $K$ -means 分割算法飞机轮廓并不完整, 含有大量的非相关区域; FCM 较  $K$ -means 分割算法在抑制非相关区域较  $K$ -means 效果好, 但飞机轮廓也不完整; KFCM 由于通过核方法提高了特征的区别性 (采用高斯核, 所用核参数  $\sigma = 10$ ), 较  $K$ -means 和 FCM 分割算法大量抑制了非相关区域, 由于没有考虑到空间信息, 飞机轮廓并不完整; SFCM 分割算法能较好抑制孤立的杂点, 使得分割结果具有较好的平滑性, 由于特征的较弱区别性, 并不能去除大量的非相关区域; 而本文所提的算法由于考虑了图像局部信息、像素之间的内在相关性以及空间类别属性约束信息, 飞机轮廓保持比较完整, 同时非相关区域被较好抑制, 分割效果较为理想.

为了进一步验证算法的性能, 对加入噪声的图像进行分割实验, 在图 1(a) 中加入标准方差为 20 的高斯噪声, 如图 2(a) 所示,  $K$ -means, FCM, KFCM, SFCM 及本文算法中采用的类别数  $J$  取为 8, 本文算法所采用图像块大小为  $5 \times 5$ , 字典大小为 100. 从图 2 中可以看出,  $K$ -means, FCM 分割算法中飞机轮廓被大量杂点干扰, 难以判断; KFCM (采用高斯核, 所用核参数  $\sigma = 10$ ) 和 SFCM 具有一定的噪声抑制性, 飞机轮廓也受到干扰, 并不完整, 同时混有大量非相关区域; 而本文所提的算法飞机轮廓较为完整, 由于所用图像块较大, 噪声抑制很好, 使得不属于飞机的像素被分为飞机像素, 所以飞机轮廓有所扩大, 但相比于  $K$ -means, FCM, KFCM 和 SFCM 分割算法, 本文算法取得了较好的结果. 为了更加证实所提算法有效性, 下面我们对机载对地道路红外图像进行实验.

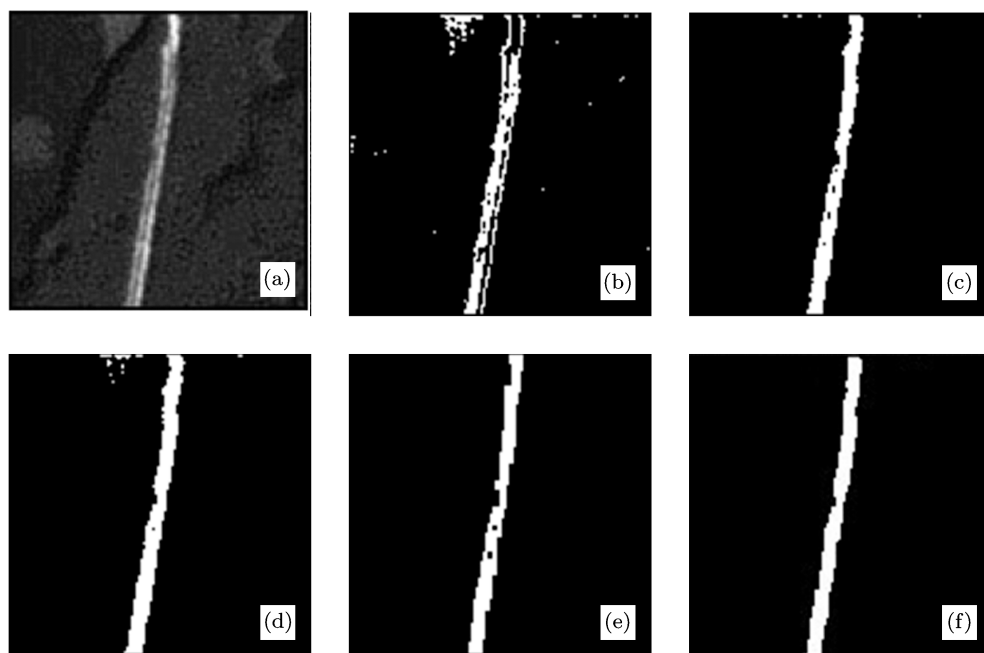


图 3 分割结果对比 (a) 原始图像; (b)  $K$ -means 分割; (c) FCM 分割; (d) KFCM 分割; (e) SFCM 分割; (f) 本文算法

对于图 3(a) 的分割实验,  $K$ -means, FCM, KFCM, SFCM 及本文算法中采用的类别数  $J$  取为 3, 本文算法所采用图像块大小  $3 \times 3$ , 字典大小 50.  $K$ -means, FCM, SFCM 和 KFCM (采用高斯核, 所用核参数  $\sigma = 2$ ) 算法虽能分割出道路的大致区域, 但道路区域不够清晰、完整, 同时也将非道路区域分割为道路; 本文所提算法分割得到的道路区域清晰、完整, 更为接近实际情况, 而且基本消除

了非道路区域及杂点的干扰. 我们将噪声加入到红外图像中进行实验, 在图 3(a) 中加入标准方差为 20 的高斯噪声, 如图 4(a) 所示,  $K$ -means, FCM, KFCM, SFCM 及本文算法中采用的类别数  $J$  取为 3, 本文算法所采用图像块大小  $3 \times 3$ , 字典大小 50.  $K$ -means, FCM 分割算法包含大量杂点, 影响道路的判断, 而且道路区域也不够完整; KFCM (采用高斯核, 所用核参数  $\sigma = 2$ ) 和 SFCM 分割算法可以较

好地抑制部分野点, 但道路区域也不够完整; 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、杂点及噪声的干扰, 分割效果较为理想.

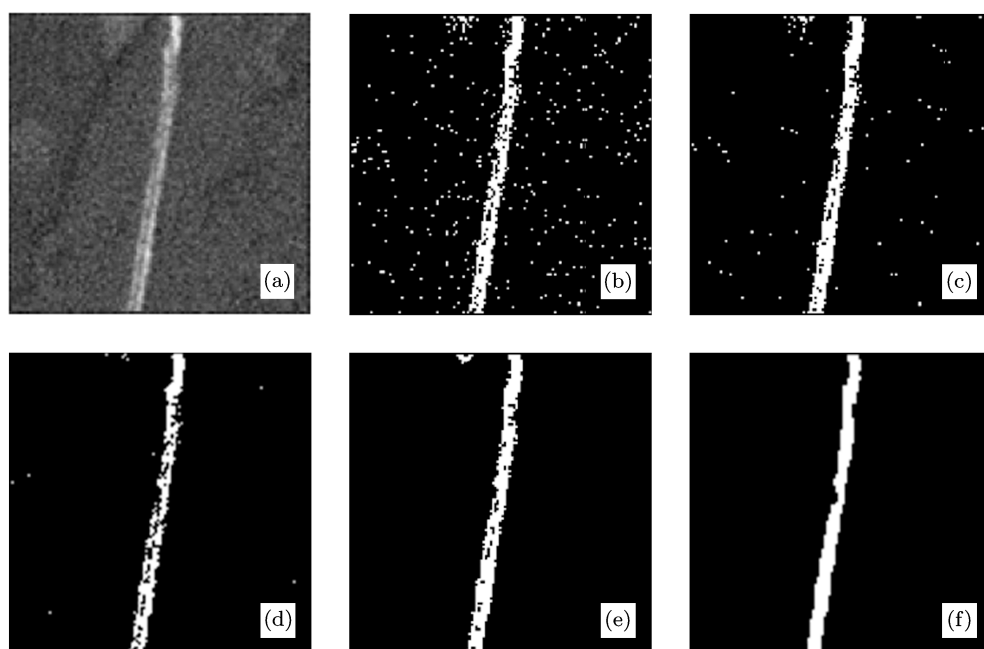


图4 含噪图像分割结果对比 (a) 噪声图像; (b)  $K$ -means 分割; (c) FCM 分割; (d) KFCM 分割; (e) SFCM 分割; (f) 本文算法

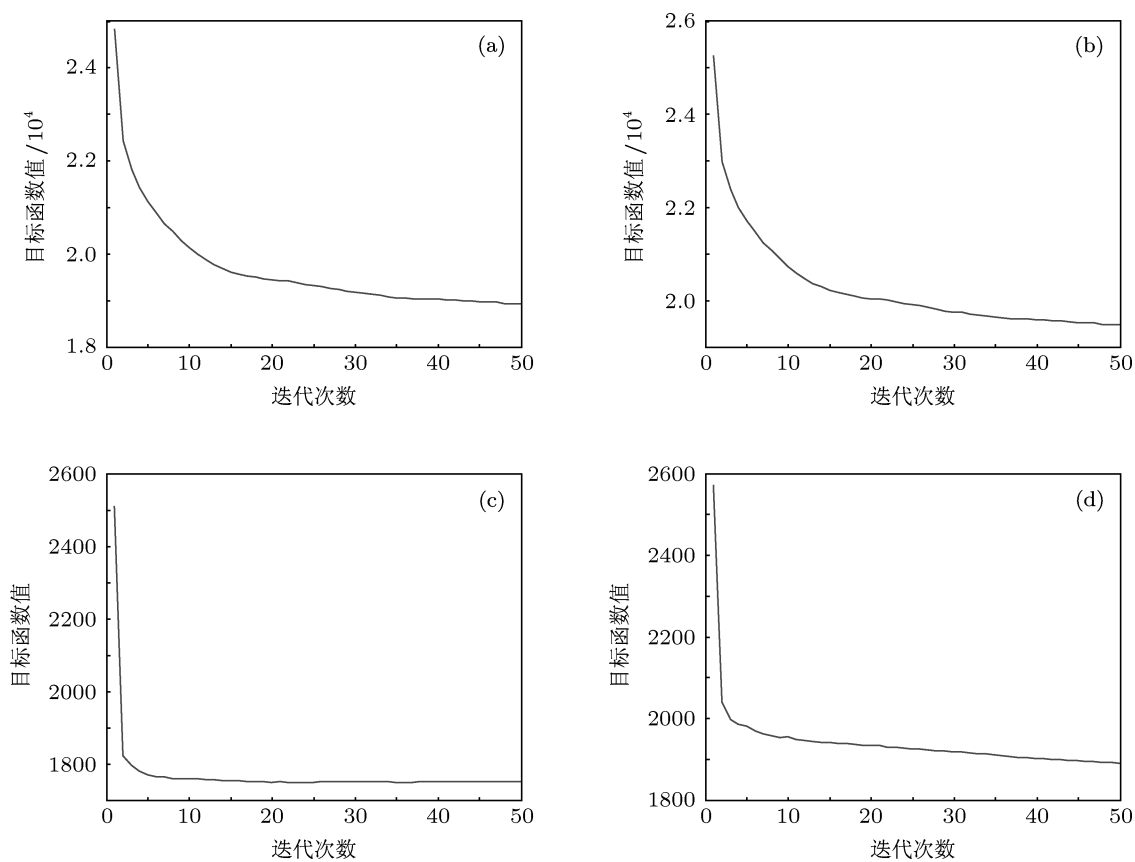


图5 本文算法目标函数收敛性曲线 (a) 图1; (b) 图2; (c) 图3; (d) 图4

从实验结果可以看出, 和传统的聚类算法相比, 考虑了图像局部信息、像素之间的内在相关性以及空间类别属性约束信息能有效地改进红外图像分割结果, 同类区域内部能呈现较好的一致性, 对平滑完整的目标分割效果较好, 所提算法在噪声环境下也具有更好的鲁棒性. 针对四组仿真实验, 我们补充了目标函数的收敛性曲线, 如图 5 所示, 可以看出本文算法经过 50 次迭代都可满足收敛要求.

## 5 结论

本文提出了一种结合稀疏编码和空间约束的红外图像聚类分割新算法, 扩展了传统的聚类分割算法, 考虑到像素之间的相关性, 引入了空间类别

属性约束信息, 并给出了一种交替优化算法, 联合学习字典、稀疏系数、聚类中心和隶属度, 通过稀疏系数和隶属度构造像素归属度完成分割. 本文所提算法充分利用了像素的内在相关性、局部信息以及空间类别属性约束信息, 实验结果也表明, 本文方法能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杂波对感兴趣区域的分割干扰, 适于复杂背景下红外目标的准确分割. 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等采用的是经验值, 如何自动选取这些参数有待进一步的研究. 同时, 为了更好地提高分割效果, 将本文算法与结构字典<sup>[20]</sup>及有监督的判别性字典学习算法<sup>[21]</sup>相结合也是将来的研究方向.

- 
- [1] Wu Y Q, Zhang J K 2010 *Acta Phys. Sin.* **59** 5495 (in Chinese) [吴—全, 张金矿 2010 物理学报 **59** 5495]
- [2] Liang Y M, Zhai H C, Chang S J, Zhang S Y 2003 *Acta Phys. Sin.* **52** 11 (in Chinese) [梁艳梅, 翟宏琛, 常胜江, 张思远 2003 物理学报 **52** 11]
- [3] Song C X 2009 *Microelectronics and Computer* **26** 60 (in Chinese) [宋长新 2009 微电子学与计算机 **26** 60]
- [4] Tang Y G, Di Q Y, Zhao L X 2009 *Acta Phys. Sin.* **58** 15 (in Chinese) [唐英干, 邸秋艳, 赵立兴 2009 物理学报 **58** 15]
- [5] Ren J J, He M Y 2008 *J. Infrared Millim. Waves* **27** 72 (in Chinese) [任继军, 何明— 2008 红外与毫米波学报 **27** 72]
- [6] Fan J C, Han M, Wang J 2009 *Pattern Recognition* **42** 2527
- [7] Wu J, Li J, Liu J, Tian J W 2004 *International Conference on Robotics and Biomimetics* Shenyang China August 22–26, 2004 p742
- [8] Jia J H, Jiao L C 2010 *J. Infrared Millim. Waves* **29** 69 (in Chinese) [贾建华, 焦季成 2010 红外与毫米波学报 **29** 69]
- [9] Chen S C, Zhang D Q 2004 *IEEE Trans. Syst. Man Cybern.* **34** 1907
- [10] Zhao F, Jiao L C, Liu H Q 2011 *Frontiers of Computer Science in China* **1** 45
- [11] Yang J C, Yu K, Gong Y H, Huang T S 2009 *IEEE Conference on Computer Vision and Pattern Recognition* Miami, USA, June 20–25, 2009 p1794
- [12] Wright J, Ma Y, Mairal J, Sapiro G, Huang T, Yan S 2010 *Proc. IEEE* **98** 1031
- [13] Zhuang L S, Gao H Y, Liu C, Yu N H 2011 *Journal of Software* **22** (Suppl. 2) 89 (in Chinese) [庄连生, 高浩渊, 刘超, 俞能海 2011 软件学报 **22** (增 2) 89]
- [14] Zhao J J, Tang Z Y, Yang J, Liu E Q, Zhou Y 2011 *J. Infrared Millim. Waves* **30** 156 (in Chinese) [赵佳佳, 唐峥远, 杨杰, 刘尔琦, 周越 2011 红外与毫米波学报 **30** 156]
- [15] Wang J J, Yang J C, Yu K, Lv F J, Huang T S, Gong Y H 2010 *CVPR* San Francisco, USA, June 13–18, 2010 p3360
- [16] Liu L Q, Wang L, Liu X W 2011 *ICCV* Barcelona, Spain, November 6–13, 2011 p2486
- [17] Gao S H, Tsang W H, Chia L T 2010 *ECCV* Crete, Greece, September 5–11, 2010 p566
- [18] Yuan L, Liu J, Ye J P 2011 *NIPS* Sierra, Nevada, December 16–17, 2011 p232
- [19] Lee H L, Battle A, Raina R, Andrew Y N 2007 *NIPS* Vancouver, B.C., Canada, December 3–6, 2007 p801
- [20] Ramirez I, Sprechmann P, Sapiro G 2010 *CVPR* San Francisco, CA June 13–18, 2010 p3501
- [21] Dahl A L, Larsen R 2011 *BMVC* Dundee, UK, August 29–September 2, 2011 p77.1

# Infrared image segmentation based on clustering combined with sparse coding and spatial constraints\*

Song Chang-Xin<sup>†</sup> Ma Ke Qin Chuan Xiao Peng

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

(Received 3 May 2012; revised manuscript received 4 October 2012)

## Abstract

A new algorithm for infrared image segmentation is proposed based on clustering combined with sparse coding and spatial constraints. The clustering algorithm is fused on the basis of sparse coding. 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 between pixels are used. However, the problem of over-segmentation and difficulty in pixels classification for segmentation arise. The clustering method is introduced for atoms into dictionary learning. The class number of atoms in dictionary is reduced in order to avoid over-segmentation. The spatial class property information is also introduced by considering the property of the pixel, and the pixels in the neighbor region should have class coherent constraints. An alternate optimization algorithm is proposed to learn the dictionary, sparse coefficients, cluster center and degrees of membership jointly. Then the classes of pixels are estimated by constructing pixel ownership degrees, combining the sparse coefficients and the degrees of membership with the atoms to cluster center. The experimental results show that the important area can be separated well, and the proposed method has good robustness.

**Keywords:** image segmentation, sparse coding, clustering, spatial constraints

**PACS:** 07.05.Pj, 87.57.N-, 95.75.Mn

**DOI:** 10.7498/aps.62.040702

---

\* Project supported by the Natural Science Foundation of Qinghai Province, China (Grant No. 2011-z-748) and the Fund for 135 High-Level Fostering Talents of Qinghai Province, China.

<sup>†</sup> Corresponding author. E-mail: songchxin@163.com