基于稀疏特征竞争和形状相似性的超声图像序列分割方法

倪波^{1*}, 蔡贤涛², 邓丹君¹, 袁涌¹

1. 湖北理工学院计算机学院, 黄石 435003

2. 武汉大学计算机学院, 武汉 430072

* 通信作者. E-mail: nibo0906@sina.com

收稿日期: 2017-01-30; 接受日期: 2017-03-01; 网络出版日期: 2017-05-09

国家自然科学基金 (批准号: 61303215) 和湖北省教育厅青年人才项目 (批准号: Q20154404) 资助

摘要 从超声图像序列中精确地分割出目标区域对于提高基于超声图像引导的计算机辅助治疗的效果具有重要的意义. 活动轮廓模型被广泛地应用于医学图像的分割, 但是超声图像固有的低信噪比、灰度分布不均匀等特点导致活动轮廓容易搜索到错误的区域. 为此, 学者们提出了不同的目标特征或形状先验知识来约束活动轮廓的演化. 目前, 大多数目标图像特征的工作仅限于提取目标区域信息的特征, 目标形状先验也需要通过对大量训练样本中的目标形状进行学习, 而且训练样本中形状是否足够描述分割目标的形状信息也是值得怀疑的. 本文提出了一种基于稀疏特征竞争和形状相似性的超声图像序列分割方法. 该方法首先利用稀疏表达理论分别构建目标和背景的特征字典, 并根据目标和背景特征字典的重构误差构建一种基于稀疏特征竞争的目标搜索策略, 从而提高活动轮廓搜索真实边缘的鲁棒性. 然后, 对图像序列中目标形状的相似性进行建模, 并证明了目标形状的变化符合低秩属性, 该属性可看作是对图像序列中目标形状的先验知识进行无监督学习. 为了验证该方法的性能, 本文采用临床的超声图像序列作为训练和测试集合, 并与其他 3 种典型的分割方法在同一测试集合下进行比较. 实验结果显示, 针对超声图像中出现的边缘模糊、缺失等缺点, 该方法提供了更准确和鲁棒的分割结果, 从而提高了计算机辅助治疗的效率和效果.

关键词 超声图像序列分割, 活动轮廓模型, 稀疏特征竞争, 形状相似性

1 引言

高强度超声聚焦 (high intensity ultrasound focus, HIFU) 作为一种基于超声图像引导的计算机辅助治疗取得了良好的治疗效果^[1,2]. 如图 1 所示, 精确地从超声图像序列分割出病灶区域对提高 HIFU 治疗的效果具有重要意义. 目前这项工作是由临床医生在每一帧图像上对目标区域进行分割, 显然这

引用格式: 倪波, 蔡贤涛, 邓丹君, 等. 基于稀疏特征竞争和形状相似性的超声图像序列分割方法. 中国科学: 信息科学, 2017, 47: 752–770, doi: 10.1360/N112016-00235
Ni B, Cai X T, Deng D J, et al. Sparse feature competition and shapes similarity based ultrasound image sequence segmentation method (in Chinese). Sci Sin Inform, 2017, 47: 752–770, doi: 10.1360/N112016-00235

图1 HIFU 治疗的流程

Figure 1 The flowchart of HIFU therapy

是一个枯燥而又低效的过程. 因此, 一个精确而高效的超声图像序列的分割方法对提高 HIFU 治疗的效果具有重要的临床价值和研究意义. 然而, 有效的超声图像分割方法依然面临以下挑战: (1) 超声图像自身固有的低信噪比和灰度分布不均匀造成目标区域边缘模糊, 甚至边缘缺失; (2) 成像过程中病人体内软组织的形变, 从而引起病灶区域的几何轮廓的复杂性; (3) 部分病灶区域周边组织比较复杂, 容易与目标区域产生混淆, 产生误导性边缘特征. 这些因素导致许多方法不容易分割到真实的病灶区域.

研究者在医学图像分割领域做了大量卓有成效的工作^[3,4], 其中基于活动轮廓模型的分割方法受到了广泛关注. 通常情况下, 活动轮廓由标记点^[5,6]或者水平集函数^[7,8]表示. 然后, 通过构建不同的目标图像特征来驱动活动轮廓搜索目标的边缘. 早期的活动轮廓模型采用图像的梯度信息作为目标的搜索特征^[6,9,10]. 后来有学者使用目标区域的统计信息作为目标的图像特征. 例如, 文献^[11,12]使用图像的一阶或二阶的统计信息来引导活动轮廓区分目标和背景. 文献^[13~16]通过构建目标区域概率统计模型来对活动轮廓搜索至真实边缘提供置信区域. 文献^[17~20]则引入 Markov 随机场理论构建目标区域的纹理特征. Huang 等^[21,22]引入图论对图像信息进行建模. 以上图像特征在目标和背景的灰度差别较大和灰度分布均匀时效果较好, 但鉴于超声图像固有的缺陷, 目标边缘经常会产生误导性或缺失的边缘特征, 从而导致活动轮廓经常会搜索到错误的区域. 为了提高活动轮廓的鲁棒性, 有学者将目标形状的先验知识融入到活动轮廓分割框架. 活动形状模型^[5,23]就是其中最具代表性的方法之一. 该方法首先用一个标记点序列组成的向量作为目标形状的参数化表达模型, 则整个训练集就构成了一个形状空间. 然后通过主成分分析^[24]对该形状空间的形状变化进行降维, 从而将目标的先验形状表示为形状空间中的平均值和最显著性变化成分的线性组合. 在分割过程中, 活动轮廓的变化始终受到该形状空间的约束, 从而提高形变轮廓对噪声和目标误导性特征的鲁棒性. 后来, 形状先验作为一个活动轮廓演化的规则项被融入到水平集分割方法中^[25~27]. 文献^[28]通过引入基于流形的学习方法将训练集中的形状投影到一个非线性的空间, 从而提高活动轮廓在分割过程中表现形状细节的能力. Yan 等^[29]通过对图像序列中相邻图像之间的相似性对前列腺底部的形状进行约束. Zhang 等^[30]引入稀疏表达理论构建目标的形状字典从而生成目标的先验形状. 文献^[31,32]通过引入非线性动力学的方来对训练集中的目标形状的变化先验知识进行建模, 从而提高对目标细节变化的建模能力.

由此可知, 构建不同形式的图像特征和形状先验知识可以提高活动轮廓对噪声、缺失性信号和误导性特征等的鲁棒性. 但目前的工作也存在相应的局限: (1) 基于统计概率的目标图像模型先要对目标区域灰度的概率分布函数进行估计, 然后通过训练样本对该函数的参数进行估计. 由于超声成像质量的复杂性, 再加上目标周围复杂背景的干扰导致预估的分布模型往往不能提供足够鲁棒的目标信息; (2) 目前构建目标的形状先验基本属于基于离线的学习方法, 需要大量的训练数据并对样本数据进行标注, 而实际情况中搜集大量的临床样本是一件困难的事情, 而且有理由怀疑通过训练集得出的形状先验是否足够表达待分割目标形状的变化.

图 2 本文方法的流程

Figure 2 The flowchart of the proposed method

近年来, 有学者利用目标在序列图像中的变化特性提出了一些图像序列分割方法. 例如, Nieuwenhuis 等^[33] 将图像序列中目标的比例作为目标形状变化的先验知识用于图像序列分割. Wang 等^[34] 通过构建图像的多尺度空间, 然后利用目标在不同尺度空间中的相似性对活动轮廓演化至真实边缘进行优化. 文献^[35,36] 引入子空间聚类理论对图像序列中的前景目标进行分割.

受此启发, 本文在活动轮廓分割框架内提出了一种基于稀疏特征竞争和形状相似性的超声图像序列分割方法. 如图 2 所示, 该方法首先利用稀疏表达理论^[37,38] 构建基于目标和背景的过完备特征字典, 并利用目标边缘的灰度特征相对于目标和背景特征字典的重构误差, 构建了一种基于稀疏特征竞争的轮廓搜索策略. 由于该搜索策略分别利用了目标和背景的特征, 从而提高了活动轮廓搜索到真实边缘区域的鲁棒性. 而且目标和背景特征的性能取决于其过完备库的容量而无需假设其特征符合某种先验模型. 然后, 对图像序列之间目标形状的变化进行无监督学习, 证明了组成图像序列形状的矩阵符合低秩属性, 并将该属性作为目标形状变化的先验知识对活动轮廓的演化进行约束. 最后, 通过一种邻近梯度优化算法完成对超声图像序列的分割. 实验结果证明, 该方法提高了活动轮廓对噪声、误导性特征和缺失边缘的鲁棒性, 而且该方法是对整个图像序列进行分割, 从而提高了制定 HIFU 术前计划的效率.

2 基于稀疏特征竞争的轮廓模型

稀疏表达作为一种新的信号处理理论在人脸识别、图像复原和去噪等领域被广泛应用, 并取得了良好的效果^[39~41]. 近年来, 该理论也被用作对医学图像进行处理, 如 Chen 等^[42] 利用低剂量计算机断层扫描 (low-dose computed tomography, LDCT) 图像中目标组织和伪影特征提出了一种区别性稀疏

图 3 (网络版彩图) 目标和背景的局部区域的示意图
Figure 3 (Color online) Object and background feature regions

编码, 然后利用该区分性稀疏表示字典对 LDCT 图像进行重构从而减少 LDCT 图像的伪影特征. 文献 [43~46] 则通过构建目标区域的纹理模型实现对目标的分割. 本文利用该理论构建目标和背景的特征信息, 并设计一种基于稀疏特征竞争的搜索机制来引导活动轮廓搜索真实边缘. 稀疏表达的本质是通过一个过完备矩阵 (又称特征字典), 将输入的信号表示为字典中列向量 (又称基向量) 的近似线性组合, 如式 (1) 所示:

$$\arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{D} \cdot \mathbf{x} + \mathbf{e}, \quad \|\mathbf{e}\|_2 \leq \varepsilon, \quad (1)$$

其中 $\mathbf{y} \in \mathbb{R}^m$ 表示输入的信号, 过完备矩阵 $\mathbf{D} = [\mathbf{a}_1, \dots, \mathbf{a}_n]^T \in \mathbb{R}^{m \times n}$, $m \ll n$ 表示一个特征字典, 字典中的每一列 $\{\mathbf{a}_i\}_{i=1}^n$ 被称做原子信号, $\mathbf{x} \in \mathbb{R}^{n \times 1}$ 是信号 $\mathbf{y}$ 的表示系数, $\|\mathbf{x}\|_0$ 为 l_0 范数, 它表示系数 $\mathbf{x}$ 中非零项的个数 (即稀疏度), $\|\mathbf{e}\|_2$ 表示信号 $\mathbf{y}$ 在字典 $\mathbf{D}$ 中投影后的残差. 这样输入信号 $\mathbf{y}$ 可以通过一个满足一定稀疏度的系数在特征字典 $\mathbf{D}$ 上近似重构. 由式 (1) 可知, 字典的构造对于实现信号的稀疏表达至关重要. 目前, 主要采用基于学习的字典构造方法, 这种方法需要建立信号实例的训练集合, 这些实例要与实际应用中的信号相似, 并利用它们构造出具备先验知识的冗余字典.

如图 3 所示, 为了更好地描述目标边缘周围的信息, 对训练样本中目标边缘的像素点进行等距离均匀采样, 并定义为目标点, 背景点为沿目标点法向量方向的像素点. 在训练集中分别采集 M 个目标点 $\{\mathbf{u}_i\}_{i=1}^M$ 和 N 个背景点 $\{\mathbf{v}_i\}_{i=1}^N$. 以每个目标点 $\mathbf{u}_i$ 为中心定义一个正方形区域 $\mathbf{p}_o^i \in \mathbb{R}^{k \times k}$ 使其包含 $k \times k$ 个像素点. 将 $\mathbf{p}_o^i$ 向量化: $\mathbf{q}_o^i \in \mathbb{R}^{k^2 \times 1}$ 构建目标区域特征集合 $\mathbf{A}_o = [\mathbf{q}_o^1, \dots, \mathbf{q}_o^M] \in \mathbb{R}^{k^2 \times M}$, $k^2 \ll M$. 对背景点采取同样的方式构建背景区域的特征集合: $\mathbf{A}_b = [\mathbf{q}_b^1, \dots, \mathbf{q}_b^N] \in \mathbb{R}^{k^2 \times N}$.

使用 K-SVD 算法^[47] 分别构建 $\mathbf{A}_o$ 和 $\mathbf{A}_b$ 的特征字典 $\mathbf{D}_o \in \mathbb{R}^{k^2 \times M}$ 和 $\mathbf{D}_b \in \mathbb{R}^{k^2 \times N}$. K-SVD 算法具有良好的自适应性, 而且可以有效抑制训练信号中的 Gauss 白噪声, 特别是在对图像信号进行字典训练时具有令人满意的性能, 目标函数被表示为

$$\min_{\mathbf{D}, \{\mathbf{x}_i\}_{i=1}^M} \sum_{i=1}^M \|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}_i\|_0 \leq \delta, 1 \leq i \leq M, \quad (2)$$

其中 $\{\mathbf{y}_i\}$ 表示训练信号集, δ 表示系数的稀疏度, M 表示训练集的容量. 式 (2) 的求解是一个 NP 难问题, 要实现上式的最优化一般通过交替迭代的方式: 首先固定字典 $\mathbf{D}$ 计算稀疏系数 $\mathbf{x}_i$, 然后根据 $\mathbf{x}_i$ 实现对 $\mathbf{D}$ 中列的更新, 如此交替反复, 最终完成对字典的更新. 在 K-SVD 算法中通常分别使用垂直匹配追踪算法 (orthogonal matching pursuit, OMP)^[48] 实现稀疏系数的求解和奇异值分解算法 (singular value decomposition, SVD) 完成字典中原子的更新.

算法 1 显示了以 $\mathbf{A}_o$ 为训练集构建目标特征字典 $\mathbf{D}_o \in \mathbb{R}^{k^2 \times M}$ 的过程, 其中矩阵 $\mathbf{X}$ 的列表示稀疏系数 $\mathbf{x}_1, \dots, \mathbf{x}_M$. 通过 l_0 范数可以控制系数的稀疏度, 从而减轻活动轮廓演化过程中计算的负担. 当算法 1 收敛时, 目标的特征字典可以表示为 $\mathbf{D}_o = [\mathbf{a}_1, \dots, \mathbf{a}_M]$, 其中 $\mathbf{a}_i \in \mathbb{R}^{k^2 \times 1}$, M 表示字典中原子的数量. 同样的方法, 构建了背景的特征字典 $\mathbf{D}_b \in \mathbb{R}^{k^2 \times N}$. 如图 4 所示, 这两个特征字典包含了图像中目标和背景的内在特征结构, 它们将被整合进活动轮廓分割框架中用于目标边缘的搜索.

算法 1 基于 K-SVD 构建目标特征字典

输入:

- 1: 训练信号集 $\mathbf{A}_o = [\mathbf{q}_o^1, \dots, \mathbf{q}_o^M] \in \mathbb{R}^{k^2 \times M}$.
- 2: 初始化.
- 3: (1) 随机选择 M 个样本构造初始的特征字典 $\mathbf{D}_o^{(0)} = [\mathbf{a}_1^{(0)}, \dots, \mathbf{a}_M^{(0)}]$, 并对 $\mathbf{D}_o^{(0)}$ 的各列减去训练集合的均值: $\mathbf{D}_o^{(0)} = \mathbf{D}_o^{(0)} - \frac{1}{M} \sum_{i=1}^M \mathbf{q}_o^i$.
- (2) $k = 0$.

主迭代:

- 4: **repeat**
- 5: $k + 1$.
- 6: (1) 计算稀疏编码:
- 7: 针对式 (2) 利用 OMP 算法计算系数矩阵 $\mathbf{X}_k$.
- 8: (2) 字典更新阶段:
- 9: **for** $j_0 = 1, \dots, M$ **do**
- 10: //更新 $\mathbf{D}_o^k$ 中的第 j 列
- 11: 定义使用原子 $\mathbf{a}_j$ 的信号集 $\Omega_j = \{i = 1 \leq i \leq M, \mathbf{X}_k[j_0, j] \neq 0\}$.
- 12: 计算信号残差 $\mathbf{E}_{j_0} := \mathbf{A}_o - \sum_{j \neq j_0} \mathbf{a}_j \mathbf{x}_j^T$, 其中 $\mathbf{x}_j$ 表示 $\mathbf{X}_k$ 的第 j 行.
- 13: 提取 $\mathbf{E}_{j_0}$ 的第 i 列来构建 $\mathbf{E}_{j_0}^R$, 其中 $i \in \Omega_{j_0}$.
- 14: 对 $\mathbf{E}_{j_0}^R = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ 进行 SVD 分解.
- 15: 更新字典原子 $\mathbf{a}_{j_0} = \mathbf{u}_1$, 其中 $\mathbf{u}_1$ 是 $\mathbf{U}$ 的第一列.
- 16: 用 $\mathbf{V} \times \mathbf{\Lambda}(1, 1)$ 的第 1 列更新 $\mathbf{X}_{j_0}^T$ 的非零元素.
- 17: **end for**
- 18: **until** $\sum_{i=1}^M \|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_2^2 \leq \text{阈值}$.

输出:

- 19: 特征字典 $\mathbf{D}_o$ 和稀疏系数 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$.
-

给出信号集 $\mathbf{Q} = \{\mathbf{q}_i\}_{i=1}^n \in \mathbb{R}^{k^2 \times n}$, 为了提高输入信号对目标特征字典的匹配性, 将输入信号减去训练特征集合 $\mathbf{A}_o$ 的均值: $\mathbf{Q} = \mathbf{Q} - \frac{1}{M} \sum_{i=1}^M \mathbf{q}_o^i$, 然后利用式 (3) 计算 $\mathbf{Q}$ 在特征字典 $\mathbf{D}_o$ 上的稀疏表

图 4 稀疏特征字典示例图

Figure 4 The examples of sparse feature dictionaries. (a) Object sparse feature dictionary; (b) background sparse feature dictionary

达系数 $\mathbf{X}_o$. 同理, 利用式 (4) 计算 $\mathbf{Q}$ 在背景特征字典 $\mathbf{D}_b$ 上的稀疏表达系数 $\mathbf{X}_b$:

$$\min_{\mathbf{X}_o} \|\mathbf{Q} - \mathbf{D}_o \cdot \mathbf{X}_o\|_2 \leq \epsilon \quad \text{s.t.} \quad \|\mathbf{X}_o\|_0 \leq T, \quad (3)$$

$$\min_{\mathbf{X}_b} \|\mathbf{Q} - \mathbf{D}_b \cdot \mathbf{X}_b\|_2 \leq \epsilon \quad \text{s.t.} \quad \|\mathbf{X}_b\|_0 \leq T. \quad (4)$$

根据稀疏表达系数 $\mathbf{X}_o$ 和 $\mathbf{X}_b$, 可通过式 (5) 和 (6) 由 OMP 算法分别计算信号集与目标和背景特征字典的重构误差:

$$\mathbf{e}_o = \|\mathbf{Q} - \mathbf{D}_o \cdot \mathbf{X}_o\|_2, \quad (5)$$

$$\mathbf{e}_b = \|\mathbf{Q} - \mathbf{D}_b \cdot \mathbf{X}_b\|_2. \quad (6)$$

用一个标记点序列来表示活动轮廓的形状, 即 $\mathcal{C} = [\mathbf{m}_1, \dots, \mathbf{m}_n]^T \in \mathbb{R}^{2n}$, 其中 $\mathbf{m}_i = [x_i, y_i]^T$ 表示轮廓上标记点的像素坐标. 根据式 (5) 和 (6), 定义 $\mathbf{e} = \frac{\mathbf{e}_o - \mathbf{e}_b}{\mathbf{e}_o}$ 表示活动轮廓上标记点是离目标的特征字典重构误差大还是离背景局部的特征字典的重构误差大. 换言之, 当 $\mathbf{e}$ 的值越大时, 说明活动轮廓上采样点越远离真实的边缘; 反之, 说明采样点越靠近背景区域. $\mathbf{e}$ 作为一种基于稀疏特征重构误差的搜索策略被整合进活动轮廓分割框架, 如式 (7) 所示:

$$f(\mathcal{C}) = -(1 - \lambda) \int_{(x,y) \in \mathcal{C}} \mathbf{e}(x,y) dx dy + \lambda \int_0^1 \left\| \frac{d\mathcal{C}}{dq} \right\| dq, \quad (7)$$

其中 $\left\| \frac{d\mathcal{C}}{dq} \right\|$ 表示活动轮廓的曲率, λ 表示其曲率光滑的权值.

3 基于低秩属性的形状先验

如图 1 所示, HIFU 设备提供的是目标的断层扫描图像序列, 序列之间目标形状的变化具有一定的相关性. 该序列中的目标形状可构成一个矩阵 $[\mathcal{C}_1, \dots, \mathcal{C}_n]$, 该矩阵代表了该图像序列中目标的形状空间. 通常情况下, $\mathcal{C}_i$ 和 $\mathcal{C}_j$ 之间的相关性可以通过计算相对应的标记点的平均欧式距离来表示. 平均欧式距离有时无法准确描述两者之间的相关性. 例如, 当一个轮廓形状经过放大或缩小变化得到另

一个形状时, 其平均欧式距离显示两者之间属于不同的目标, 这显然与所描述的实际情况是不符合的. 用矩阵的秩作为该形状空间中不同形状之间变化的度量. 当矩阵的秩为 1 时, 矩阵中的所有轮廓表示同一个目标形状, 如果矩阵是一个满秩矩阵, 则矩阵中所有的形状轮廓不具备相关性. 假设图像序列中目标形状轮廓可由其中一张图像中的目标轮廓经过线性变化而产生, 并且组成形状的矩阵的秩满足低秩属性, 具体的推导过程如下:

$$C_i = \mathcal{M}C_1 + \mathcal{T}, \quad \forall C_i \in \{C_1, C_2, \dots, C_n\}, i \neq 1, \quad (8)$$

其中 $\mathcal{M}$ 表示线性变换的矩阵, $\mathcal{T}$ 表示一个位移向量. 根据文献 [49] 中描述, 形状矩阵可以表示为式 (9) 的形式:

$$[C_1, \dots, C_n]^T = \Psi[\Phi_1, \dots, \Phi_n], \quad (9)$$

其中

$$\Psi = \begin{bmatrix} C_1^x & 0 & C_1^y & 0 & 1 & 0 \\ 0 & C_1^x & 0 & C_1^y & 0 & 1 \end{bmatrix} \in \mathbb{R}^{2n \times 6}$$

和

$$\Phi_i = [m_1^i, m_2^i, m_3^i, m_4^i, t_1^i, t_2^i]^T \in \mathbb{R}^{6 \times 1}.$$

C_1^x 和 $C_1^y \in \mathbb{R}^n$ 分别表示向量 C_1 的 x 和 y 坐标半部, m_j^i 和 t_j^i 分别表示矩阵 $\mathcal{M}$ 和向量 $\mathcal{T}$ 中的元素. 根据定理 1 [50] 可知, 因为矩阵 Ψ 的列最多只有 6, 所以无论矩阵 $\mathcal{M}$ 和向量 $\mathcal{T}$ 如何变化, 矩阵 $[C_1, \dots, C_n]^T$ 是一个低秩矩阵.

定理1 针对 $A \in \mathbb{R}^{m \times k}$ 和 $B \in \mathbb{R}^{k \times n}$, 矩阵的秩满足以下性质:

$$\text{rank}(A) + \text{rank}(B) - k \leq \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.$$

本质上说, C_1 可以看作是图像序列中目标形状组成的形状空间 $\{C_1, \dots, C_n\}$ 的子空间, Φ_i 可以看作是 C_1 的变化系数. 因此, 形状矩阵 $[C_1, \dots, C_n]$ 的秩可以看作描述形状变化的自由度. 使用形状矩阵的秩来约束活动轮廓的演化, 从而提高活动轮廓搜索对超声图像中错误信息的鲁棒性.

4 本文的分割方法

根据 HIFU 给出的超声图像序列 $I_1, \dots, I_n$, 使用式 (7) 对每一帧的图像进行分割. 在分割过程中, 形状矩阵 $[C_1, \dots, C_n]$ 的低秩属性作为形状先验被整合进图像序列的分割, 如式 (10) 所示:

$$\min_{\mathcal{C}} \sum_{i=1}^n f(C_i), \quad \text{s.t.} \quad \text{rank}(\mathcal{C}) \leq K, \quad (10)$$

其中 $f(C_i)$ 代表式 (7), $\mathcal{C} = [C_1, \dots, C_n]$, $\text{rank}(\cdot)$ 表示计算矩阵 $\mathcal{C}$ 的秩, K 表示设置的阈值. 近来, 矩阵的核范数常被用作计算矩阵秩的近似凸代理 [51, 52], 则式 (10) 可转化为

$$\min_{\mathcal{C}} \sum_{i=1}^n f(C_i) + \beta \|\mathcal{C}\|_*, \quad (11)$$

其中 $\|C\|_*$ 表示矩阵 C 的核范数, 即 C 的奇异值的和, β 表示矩阵秩的权重系数. 使用邻近梯度 (proximal gradient, PG) 算法^[53,54] 对式 (11) 进行优化. PG 算法主要用来解决以下优化问题:

$$\min_{\mathbf{X}} F(\mathbf{X}) + \beta \mathbf{R}(\mathbf{X}), \quad (12)$$

其中 $F(\mathbf{X})$ 是一个差分方程, $\mathbf{R}(\mathbf{X})$ 是一个凸惩罚项. 针对式 (11), $F(\mathbf{X}) = \sum_{i=1}^n f_i(C_i)$ 和 $\mathbf{R}(\mathbf{X}) = \|C\|_*$. PG 算法的优化过程主要通过每一次迭代中利用前一次 $\mathbf{X}'$ 和式 (13) 对 $F(\mathbf{X})$ 实现二次逼近,

$$\begin{aligned} Z_{\mu}(\mathbf{X}, \mathbf{X}') &= F(\mathbf{X}') + \langle \nabla F(\mathbf{X}'), \mathbf{X} - \mathbf{X}' \rangle + \frac{\mu}{2} \|\mathbf{X} - \mathbf{X}'\|_F^2 + \lambda \mathbf{R}(\mathbf{X}), \\ &= \frac{\mu}{2} \left\| \mathbf{X} - \left[\mathbf{X}' - \frac{1}{\mu} \nabla F(\mathbf{X}') \right] \right\|_F^2 + \lambda \mathbf{R}(\mathbf{X}) + \text{const}, \end{aligned} \quad (13)$$

其中 $\langle \cdot, \cdot \rangle$ 表示内积, $\|\cdot\|_F$ 表示 Frobenius 范数, μ 表示一个常数. 在每次迭代过程中, 式 (13) 中的 $\mathbf{X}'$ 可以通过以下公式计算:

$$\begin{aligned} \mathbf{X}^{k+1} &= \arg \min_{\mathbf{X}} Z_{\mu}(\mathbf{X}, \mathbf{X}'), \\ &= \arg \min_{\mathbf{X}} \frac{1}{2} \left\| \mathbf{X} - \left[\mathbf{X}^k - \frac{1}{\mu} \nabla F(\mathbf{X}^k) \right] \right\|_F^2 + \frac{\lambda}{\mu} \mathbf{R}(\mathbf{X}). \end{aligned} \quad (14)$$

当 $F(\mathbf{X})$ 保持可微, 则式 (12) 将以 $O(\frac{1}{k})$ 收敛于一个稳定的点. 式 (15) 的优化问题可由引理 1 推导得出^[55].

引理1 已知 $\mathbf{X} \in \mathbb{R}^{m \times n}$, 式 (15) 的解:

$$\min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2 + \alpha \|\mathbf{X}\|_1, \quad (15)$$

由 $\mathbf{X}^* = \mathcal{D}_{\alpha}(\mathbf{Z})$ 给出, 其中

$$\mathcal{D}_{\alpha}(\mathbf{Z}) = \sum_{i=1}^{\min(m,n)} (\sigma_i - \alpha) \mathbf{u}_i \mathbf{v}_i^T, \quad (16)$$

其中 $\mathbf{u}_i$ 和 $\mathbf{v}_i$ 分别是矩阵 $\mathbf{Z}$ 的左右奇异向量, σ_i 是奇异值. $\mathcal{D}_{\alpha}(\mathbf{Z})$ 代表奇异值分解操作. 因此, 算法的更新步骤可以表示为

$$\mathbf{X}^{k+1} = \mathcal{D}_{\frac{\lambda}{\mu}} \left(\mathbf{X}^k - \frac{1}{\mu} \nabla F(\mathbf{X}^k) \right). \quad (17)$$

剩下的任务就是对 $\nabla F(\mathbf{X}^k) = [\nabla f_1(C_1^k), \dots, \nabla f_n(C_n^k)]$ 进行计算, 以式 (7) 为例

$$\nabla f_i(p) = (1 - \lambda) \mathbf{e}(C(p)) \mathbf{N}(p) - \lambda K(p) \mathbf{N}(p), \quad (18)$$

其中 p 表示轮廓的标记点, $\mathbf{N}(p)$ 和 $K(p)$ 分别表示标记点 p 的法向量和曲率. 为了提高 PG 算法的收敛速度, 文献 [54] 对 $\mathbf{X}$ 的更新进行了改进, 用式 (19) 计算的中间值 $\mathbf{Y}^k$ 对式 (13) 中 $\mathbf{X}^{k-1}$ 和 $\mathbf{X}^k$ 进行线性插值,

$$\mathbf{Y}^k = \mathbf{X}^k + \frac{t^{k-1} - 1}{t^k} (\mathbf{X}^k - \mathbf{X}^{k-1}), \quad (19)$$

其中 t^k 为随机确定的. 算法 2 描述了式 (11) 的整个流程, 即首先活动轮廓根据图像特征进行演化, 然后由形状的相似性规则项对其形状的演化进行进一步约束.

算法 2 基于 PG 算法解决式 (11)

```

初始化:  $\mathbf{X}^0 = \mathbf{X}^{-1}, t_0 = t_{-1} = 1.$ 
1: for  $k = 0 \rightarrow \text{iteration}_{\max}$  do
2:    $\mathbf{Y}^k = \mathbf{X}^k + \frac{t^{k-1}-1}{t^k}(\mathbf{X}^k - \mathbf{X}^{k-1}).$ 
3:   for  $i = 1 \rightarrow n$  do
4:      $y_i^k \leftarrow -\frac{1}{\mu} \nabla f_i(y_i^k).$ 
5:   end for
6:    $\mathbf{X}^{k+1} = \mathcal{D}_{\frac{\lambda}{\mu}}(\mathbf{Y}^k).$ 
7:    $t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2}.$ 
8:   if  $\|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F < \text{threshold}$  then
9:     retrun
10:  end if
11: end for

```

5 实验与分析

5.1 材料、评价和参数设置

采集来自不同病人的 12 个二维超声心动图像序列和 10 个子宫肌瘤图像序列作为实验的数据集, 并随机选择 5 个图像序列 (3 个心动图像, 2 个子宫肌瘤图像) 作为测试集合用于验证该方法的性能. 每个图像序列包含 30 帧图像, 每帧图像的分辨率为 800×600 个像素, 像素的尺寸为 $0.15 \text{ mm} \times 0.15 \text{ mm}$. 由于心脏收缩会造成超声成像过程中目标的边缘模糊和泄露, 目标区域的灰度分布不均匀造成边缘与周边组织的混淆产生误导性特征等, 子宫肌瘤作为一种软组织在超声成像过程中因受到呼吸运动和腹腔其他组织的挤压也会造成超声图像产生以上的特点, 因此这些图像具有一定的代表性. 使用有经验的临床医生手动分割的结果作为分割的 ground truth. 所有实验均在 MATLAB 2015a 环境下进行, 计算机硬件配置为 Intel (R) Core (TM) i7-6700 3.40 GHz, 内存 16 G.

使用 TP (the true position) 和 FP (the false position) 对计算得到的分割区域和手动分割区域之间的重叠误差进行定量分析, 其定义如下:

$$\text{TP} = \frac{|\Omega_m \cap \Omega_a|}{|\Omega_m|}, \quad (20)$$

$$\text{FP} = \frac{|\Omega_m \cup \Omega_a - \Omega_m|}{|\Omega_m|}, \quad (21)$$

其中, Ω_m 和 Ω_a 分别表示手动分割轮廓和计算分割轮廓上的像素集合, TP 的值越高表示越多的目标区域被覆盖, FP 值越低表示计算分割的轮廓越匹配目标的轮廓. 此外, 还使用 AMED (average minimum Euclidean distance) 和 HD (Hausdorff distance) 这两个指标分布对计算分割轮廓和手动分割轮廓之间的距离差异进行度量, 其定义如下:

$$\text{AMED}(A, B) = \frac{1}{2m} \sum_{i=1}^m \text{MED}(a_i, B) + \frac{1}{2n} \sum_{j=1}^n \text{MED}(b_j, A), \quad (22)$$

$$\text{HD}(A, B) = \max \left\{ \max_{i \in \{1, \dots, m\}} \text{MED}(a_i, B), \max_{j \in \{1, \dots, n\}} \text{MED}(b_j, A) \right\}, \quad (23)$$

表 1 Lui 等方法^[45]与式 (7) 分割结果的量化比较Table 1 The segmentation quantitative comparisons between the results by Lui et al.'s method^[45] and Eq. (7)

Method	Metrics	TP (%)	FP (%)	AMED (mm)	HD (mm)
Lui et al. [45]	Mean±SD	97.73±3.23	8.91±1.10	3.17±1.38	8.17±3.88
	Median±SD	95.34±4.15	6.38±1.25	2.46±1.38	9.28±2.21
Eq. (7)	Mean±SD	97.13±2.23	6.98±0.61	2.43±1.76	5.12±2.48
	Median±SD	96.84±3.46	5.41±0.32	2.19±1.58	4.21±1.88

其中 $A = \{a_1, \dots, a_m\}$ 和 $B = \{b_1, \dots, b_n\}$ 分别表示两个轮廓上有序的点集, MED (minimum Euclidean distance) 定义如下:

$$\text{MED}(a_i, B) = \min_{j \in \{1, \dots, n\}} \text{dis}(a_i, b_j), a_i \in A. \quad (24)$$

AEMD 值越低, 说明两个轮廓对应标记点之间的平均距离越小, HD 值表示两个轮廓中相距最远的标记点之间的距离.

为实现分割结果的可再现性, 对该方法中的相关初始化工作和参数进行了设置, 其中式 (10) 中的活动轮廓均被初始化为待分割区域内部的一个椭圆, 该椭圆由 24 个有序的标记点组成的向量进行参数化表达. 目标特征字典 D_o 和背景特征字典 D_b 的长度分别被设置为 $M = N = 100$. 控制图像块大小的参数 k , 式 (3) 和 (4) 中的参数 T 、式 (7) 中的参数 λ 、式 (11) 中的参数 β 和式 (13) 中的参数 μ 均根据实验样本使用在实验过程中效果最好的经验值. 其中参数 k , T 和 β 的选择以及对分割结果的影响将在第 5.5 小节做进一步的阐述.

5.2 稀疏特征竞争活动轮廓的性能分析

本小节主要验证第 2 节提出的基于稀疏特征竞争的活动轮廓模型 (式 (7)) 的性能, 使用 Lui 等^[45]方法与式 (7) 对测试集的分割结果进行比较. Lui 等^[45]方法采用稀疏表达对目标区域的纹理特征进行建模并将该纹理特征整合进活动轮廓模型从而提高活动轮廓搜索目标的能力. 表 1 给出了两种方法分割结果的量化比较, 从该表可以看出这两种方法在各项测量指标的均值都取得了不错的效果, 而且 Lui 等^[45]方法的 TP 均值要优于式 (7) 的 TP 均值, 但从 FP 均值上看 Lui 等^[45]要比式 (7) 的均值要高, 说明 Lui 等^[45]方法的分割结果有可能覆盖到错误的区域. 式 (7) 的 AMED 和 HD 值均比 Lui 等^[45]方法的要低, 说明式 (7) 分割轮廓的标记点更加接近手动分割的轮廓. 表 1 也显示式 (7) 在各项指标的中值上均要优于 Lui 等^[45]方法, 而且式 (7) 的标准差普遍低于 Lui 等^[45]方法. 这说明在不同的图像质量下, 式 (7) 相比较于 Lui 等^[45]方法的总体分割结果没有出现较大的变化, 分割的结果比较稳定. 图 5 分别显示了两种方法在超声心动图像和子宫肌瘤图像的分割结果示例, 其中虚线表示手动分割的结果, 实线表示计算的结果. 该图显示在目标与背景信息容易混淆的情况下 (白色矩形框), Lui 等^[45]方法的活动轮廓容易搜索到错误的区域, 而式 (7) 由于采用目标和背景的纹理信息因此分割效果更加鲁棒.

5.3 不同方法分割图像序列的性能分析

本小节主要验证该方法对图像序列的分割性能, 选择 Li 等^[56]和 Qin 等^[57]方法与该方法进行了比较. Li 等^[56]方法使用子空间聚类对图像序列灰度信息之间的相关性进行建模, 并结合图割理论将序列图像中的前景目标从背景信息分割出来. Qin 等^[57]方法是一种基于活动轮廓模型的分割方法,

图 5 Lui 等^[45]方法与式 (7) 方法分别对超声心动图像与子宫肌瘤图像的分割实例

Figure 5 Segmentation results on a single image by Lui et al.'s method^[45] and Eq. (7). (a) Ultrasonic echocardiography image; (b) ultrasonic uterine fibroid image

它使用稀疏表达对形状先验进行离线学习, 并结合增量学习和主成分分析来构建目标的表观模型, 从而驱动活动轮廓对序列图像中的目标进行分割. 图 6 用箱形图显示了这 3 种方法在 5 个不同的图像序列上分割结果的量化比较, 其中 Seq1~Seq3 为超声心动图像序列, Seq4~Seq5 为子宫肌瘤图像序列. 根据 4 个指标观察, 虽然 Li 等^[56]方法比 Qin 等^[57]方法在 5 个序列中的分割结果均出现比该方法效果好的分割结果, 但从均值上观察, 该方法在 5 个不同的序列中的效果均要优于 Li 等^[56]和 Qin 等^[57]方法. 而且在不同序列中, 该方法的中值和均值比较接近, 整个箱形的上分位和下分位的幅度均小于前两种方法, 这说明该方法对各帧图像的分割结果比较稳定, 没有因图像质量的变化而出现分割结果大的分化. 这一点可以在图 7 得到进一步说明. 图 7 则分别显示了不同方法在一个超声心动图像和子宫肌瘤图像序列上对每一帧图像分割结果的 AMED 值. 从总体上观察, 该方法在 2 个序列图像中 AMED 值相对于其他两种方法, 不仅在整个序列的分割过程中保持比较稳定而且维持在一个较低的水平. Li 等^[56]方法的分割结果波动较大, 说明该方法容易受到边缘模糊、误导性边缘等影响. Qin 等^[57]方法采用稀疏表达来构建目标的形状先验, 因此其分割结果要比 Li 等^[56]方法的波动小. 这说明针对超声图像, 通过引入目标形状先验比单纯的依靠图像信息可以使得分割方法更加鲁棒. 但 Qin 等^[57]方法是通过增量学习来构建目标的表观模型, 导致在前一帧分割结果的误差容易传导到后面的分割结果, 因此分割结果的 AMED 值也呈现增长的趋势. 综上所述, 该方法在搜索目标边缘的过程中通过图像序列之间形状先验和目标与背景的灰度特征差异特性的约束, 不仅取得了较好的分割精度, 而且对超声图像中固有缺点具有较好的鲁棒性. 图 8 分别显示了不同方法在超声心动图像序列和子宫肌瘤图像序列中均匀选择的 8 帧图像上分割的示例, 其中虚线表示手动分割的结果, 实线表示计算的结果.

5.4 算法的鲁棒性分析

为了进一步分析该方法的鲁棒性, 请有经验的临床医生将测试序列中的样本按照目标的轮廓形变、噪声以及目标边缘模糊等情况分为 High, Medium 和 Low 3 个等级, 不同等级的数量如表 2 所示. 然后, 使用不同的测量指标对不同等级样本的分割效果进行统计分析 (如表 3 所示). 从表 3 的统计结

图 6 在不同的指标下不同方法分割结果的量化比较

Figure 6 Comparisons of TP, FP, AMED and HD for the segmentation results on five ultrasound image sequences by three methods

图 7 不同方法分别在一个超声子宫肌瘤和心动图像序列上分割结果的 AMED 值比较

Figure 7 Comparisons of AMED for the segmentation results on ultrasound image sequence by three methods. (a) Ultrasonic uterine fibroid image sequence; (b) ultrasonic echocardiogram image sequence

果可以观察, 在 Medium 和 Low 组, 该方法比 Li 等^[56] 和 Qin 等^[57] 表现出更好的分割效果. 而且, 随着图像质量从 High 至 Low 组, 该方法的各项测量指标虽然有所降低, 但是相对于其他两种方法, 该方法在各项测量指标的均值和标准差上的变化幅度相对较小. 这说明该方法相对于其他方法具有较强的

图 8 不同方法在超声图像序列中分割结果的示例

Figure 8 Segmentation results on ultrasound image sequence by three methods. (a) Ultrasonic echocardiography image sequence; (b) ultrasonic uterine fibroid image sequence

表 2 各测试序列中不同等级的样本数量
Table 2 Sample size of test sequences on different levels

Image	High	Medium	Low
Ultrasonic echocardiography image	15	54	21
Ultrasonic uterine fibroid image	16	28	16
Total	31	82	37

鲁棒性. 此外, 对实验结果进行分析显示该方法的初始化轮廓大小对分割结果精度的敏感度不大.

表 3 不同方法对不同图像质量分割结果的 TP, FP, AMED 和 HD 值比较

Table 3 The mean and standard deviation (SD) of TP, FP, AMED and HD measurements of the three methods on contour detection in the high, medium and low quality groups

Method	Level	Mean $\pm$ SD			
		TP (mm)	FP (%)	AMED (mm)	HD (mm)
Li et al. [56]	H	94.35 $\pm$ 3.92	4.15 $\pm$ 1.43	1.93 $\pm$ 1.14	2.84 $\pm$ 1.65
	M	92.26 $\pm$ 5.15	5.41 $\pm$ 1.57	3.85 $\pm$ 1.47	4.57 $\pm$ 1.57
	L	87.24 $\pm$ 7.19	8.27 $\pm$ 4.08	5.12 $\pm$ 2.03	6.54 $\pm$ 3.04
Qin et al. [57]	H	96.32 $\pm$ 1.18	3.22 $\pm$ 0.85	1.78 $\pm$ 0.56	2.18 $\pm$ 1.23
	M	94.34 $\pm$ 1.94	4.07 $\pm$ 1.74	2.98 $\pm$ 0.89	3.58 $\pm$ 1.06
	L	89.26 $\pm$ 2.75	6.31 $\pm$ 1.95	4.41 $\pm$ 1.48	4.21 $\pm$ 2.65
Ours	H	96.28 $\pm$ 1.91	2.12 $\pm$ 0.55	1.32 $\pm$ 0.21	1.54 $\pm$ 0.98
	M	95.16 $\pm$ 1.66	2.86 $\pm$ 0.68	1.82 $\pm$ 0.25	2.09 $\pm$ 0.95
	L	94.51 $\pm$ 1.58	3.32 $\pm$ 0.67	2.84 $\pm$ 0.38	2.97 $\pm$ 1.15

5.5 主要参数的选择与分析

本小节主要讨论几个主要参数的选择和对分割结果的影响, 即式 (3) 和 (4) 中的参数 T 、式 (11) 中的参数 β 和第 2 节中定义目标或背景正方形区域大小的参数 k . 采用 AMED 和 HD 测量不同的 T , β 和 k 值对分割结果的影响. 参数 T 主要是控制系数的稀疏度. 如图 9(a) 所示, 参数 T 取值过大或过小都会降低分割结果的精度. 根据式 (3) 和 (4) 可知, 当稀疏度 T 取值过大时, D_o 和 D_b 会给输入信号提供更多的信息表达能力, 从而削弱目标特征和背景特征之间的差异性. 这样会导致一个真实的目标区域在对背景特征字典上进行重构时误差不是很大, 而在对目标特征字典进行重构时误差也不够小. 当 T 取值过小时, 目标特征字典或背景特征字典都没能提供各自足够的特征信息, 这同样也会削弱输入信号在目标和背景之间的区分能力. 参数 k 表示局部图像块的大小, 如图 9(b) 所示, 参数 k 的取值过大或者过小也会降低分割的性能. 由于超声图像中目标边缘附近具有明显的灰度明暗变化, 当 k 的值被设置为过大, 容易造成在构建目标特征字典的时候会包含有部分的非目标信息. 同理, 参数 k 的取值过小, 图像块没有捕捉足够目标或背景的局部区域信息, 也会降低目标或背景特征字典各自的表达能力. 这样也会造成局部稀疏特征竞争模型降低活动轮廓搜索目标区域的判别能力. 式 (11) 中的参数 β 被用来控制图像序列中目标形状变化的权重. 如图 9(c) 所示, β 取值过大或过小也会降低分割结果的精度. 事实上, 当 β 值被设置的过大, 活动轮廓在图像序列分割过程中形状容易趋同, 目标的一些细节特征容易没有被分割, 当 β 被设置为过小, 会造成图像序列中目标形状变化的约束变小, 从而导致活动轮廓演化到错误的区域.

5.6 迭代次数和运行时间

表 4 显示了 3 种分割方法的平均计算时间和迭代次数. Qin 等 [57] 方法在对图像序列进行分割时是利用前一帧的分割结果来对下一帧进行初始化, 从而缩小了活动轮廓搜索目标的范围, 因此该运行时间比该方法的收敛的时间少. 但该方法比 Li 等 [56] 方法的平均运行时间短, 这是因为 Li 等 [56] 方法采用图割理论需要对整幅图像的信息进行计算, 因此运行时间比其他两种方法长, 且没有迭代次数.

图 9 不同的参数 T , k 和 β 值对实验结果影响

Figure 9 AMED and HD comparisons for segmentation results on all test images when changing the values of (a) T , (b) k , and (c) β

表 4 不同方法的运行时间和迭代比较

Table 4 Running time and iterations of the three methods

Method	Running time (s)	Number of iterations
Li et al. [56]	12.5	—
Qin et al. [57]	9.3	24
Ours	11.3	32

5.7 本文方法的局限

由上面的讨论可知, 该方法利用图像序列中目标形状变化满足线性变化时, 图像序列的形状矩阵才满足低秩属性, 并以此作为形状变化的先验知识用于活动轮廓演化时的约束. 当目标组织在成像过程中受外力的影响, 图像序列之间目标形状发生复杂的非线性变化时, 该形状先验就不能完全反映每一帧中目标形状的细节. 利用图像序列中目标相似性作为形状先验, 虽然可以避免超声图像中误导性特征对活动轮廓的影响, 但同时也会让一些目标的细节部分没有被分割到, 如何取得这两者之间的平衡, 本文的做法是通过对本方法进行实验来对参数 β 进行经验的选择, 以及构建更加完备的目标和背景特征字典让活动轮廓对局部细节特征更加敏感.

6 结论与展望

综上所述,在活动轮廓分割框架下提出一种基于稀疏特征竞争和形状相似性的形状先验知识从而实现图像序列中目标形状的分割.实验证明,基于稀疏特征竞争的目标搜索策略的活动轮廓由于综合利用目标和背景之间图像信息的误差,因此比单纯利用目标区域纹理信息的活动轮廓模型的分割效果更加鲁棒.而且,对图像序列之间目标形状变化的先验知识进行无监督学习比通过训练样本对目标形状的先验知识进行有监督的学习更能反映分割目标形状的信息,从而进一步提升活动轮廓在边缘模糊或泄露的情况下分割的鲁棒性,而这一点对提高基于超声引导的计算机辅助治疗的效果是非常有价值的.在未来,希望引入 CUDA 并行计算提高方法的计算效率,并将其用于基于 3D 的超声图像分割以及软组织形变的仿真实验.

致谢 感谢鄂东医疗集团中心医院邱晓明主任等提供超声心动图像和子宫肌瘤图像,对病灶区域进行标注,以及帮助我们了解在 HIFU 治疗方面的知识.

参考文献

- 1 Zhang L, Wang Z B. High-intensity focused ultrasound tumor ablation: review of ten years of clinical experience. *Front Medicine China*, 2010, 4: 294–302
- 2 Leung J H, Yu S C, Cheung E C, et al. Safety and efficacy of sonographically guided high-intensity focused ultrasound for symptomatic uterine fibroids: preliminary study of a modified protocol. *J Ultrasound Medicine Off J Am Inst Ultrasound Medicine*, 2014, 33: 1811–1818
- 3 Saini K, Dewal M L, Rohit M. Ultrasound imaging and image segmentation in the area of ultrasound: a review. *Int J Adv Sci Tech*, 2010, 24: 41–60
- 4 He B S, Zhu F, Shi Y G. Medical image segmentation. *Int J Comput Sci Eng*, 2013, 3: 1209–1218
- 5 Cootes T F, Taylor C J, Cooper D H, et al. Active shape models-their training and application. *Comput Vis Image Und*, 1995, 61: 38–59
- 6 Xu C, Prince J L. Snakes, shapes and gradient vector flow. *IEEE Trans Image Process*, 1998, 7: 359–369
- 7 Malladi R, Sethian J A, Vemuri B C. Shape modeling with front propagation: a level set approach. *IEEE Trans Pattern Anal Machine Intell*, 1995, 17: 158–175
- 8 Chan T F, Vese L A. Active contours without edges. *IEEE Trans Image Process*, 2001, 10: 266–277
- 9 Mcinerney T, Terzopoulos D. T-snakes: topology adaptive snakes. *Med Image Anal*, 2000, 4: 73–91
- 10 Goldenberg R, Kimmel R, Rivlin E, et al. Fast geodesic active contours. *IEEE Trans Image Process*, 2001, 10: 1467–1475
- 11 Yuan H E, Luo Y P, Dong-Cheng H U. Unsupervised texture segmentation based on geodesic active regions. *J Softw*, 2007, 18: 592–599
- 12 Yezzi A, Tsai A, Willsky A. A statistical approach to snakes for bimodal and trimodal imagery. In: *Proceedings of International Conference on Computer Vision*. Kerkyra: IEEE Computer Society, 1999. 898–903
- 13 Adam A, Kimmel R, Rivlin E. On scene segmentation and histograms-based curve evolution. *IEEE Trans Pattern Anal Machine Intell*, 2009, 31: 1708–1714
- 14 Sarti A, Corsi C, Mazzini E, et al. Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. *IEEE Trans Ultrason Ferroel Freq Control*, 2005, 52: 947–960
- 15 Liao X, Yuan Z, Zheng Q, et al. Multi-scale and shape constrained localized region-based active contour segmentation of uterine fibroid ultrasound images in HIFU therapy. *Plos One*, 2014, 9: 103334
- 16 Bo N, He F, Yuan Z Y. Segmentation of uterine fibroid ultrasound images using a dynamic statistical shape model in HIFU therapy. *Comput Med Imag Graph*, 2015, 357: 302–314
- 17 Cross G R. Markov random field texture models. *IEEE Trans Pattern Anal Machine Intell*, 1983, 5: 25–39
- 18 Guo C E, Zhu S C, Wu Y N. Modeling visual patterns by integrating descriptive and generative methods. *Int J Comput Vision*, 2003, 53: 5–29
- 19 Tu Z, Zhu S C. Image segmentation by data-driven markov chain monte carlo. *IEEE Trans Pattern Anal Machine Intell*, 2002, 24: 657–673

- 20 Lian X C, Li Z, Lu B L, et al. Max-margin dictionary learning for multiclass image categorization. In: *Proceedings of European Conference on Computer Vision*, Verlag, 2010. 157–170
- 21 Huang Q H, Lee S Y, Liu L Z, et al. A robust graph-based segmentation method for breast tumors in ultrasound images. *Ultrason*, 2012, 52: 266–275
- 22 Huang Q, Bai X, Li Y, et al. Optimized graph-based segmentation for ultrasound images. *Neurocomputing*, 2014, 129: 216–224
- 23 Ginneken B V, Frangi A F, Staal J J, et al. Active shape model segmentation with optimal features. *IEEE Trans Med Imag*, 2002, 21: 924–933
- 24 Jolliffe I T. *Principal Component Analysis*. Berlin: Springer, 1986. 41–64
- 25 Paragios N. A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Trans Med Imag*, 2003, 22: 773–776
- 26 Li C, Xu C, Gui C, et al. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans Image Process*, 2010, 19: 3243–3254
- 27 Li C, Huang R, Ding Z, et al. A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans Image Process*, 2011, 20: 2007–2016
- 28 Zhang L, Zhang Q, Zhang L, et al. Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding. *Pattern Recogn*, 2015, 48: 3102–3112
- 29 Yan P, Xu S, Turkbey B, et al. Adaptively learning local shape statistics for prostate segmentation in ultrasound. *IEEE Trans Biomed Eng*, 2011, 58: 633–641
- 30 Zhang S, Zhan Y, Dewan M, et al. Sparse shape composition: a new framework for shape prior modeling. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Colorado, 2011. 1025–1032
- 31 Shepherd T, Prince S J, Alexander D C. Interactive lesion segmentation with shape priors from offline and online learning. *IEEE Trans Med Imag*, 2012, 31: 1698–1712
- 32 Ni B, He F Z, Yuan Z Y. A novel ultrasound image segmentation based on dynamical statistical shape model. *Sci Sin Inform*, 2015, 45: 1449–1465 [倪波, 何发智, 袁志勇. 一种基于动力学统计形状模型的超声图像分割方法. *中国科学: 信息科学*, 2015, 45: 1449–1465]
- 33 Nieuwenhuis C, Strekalovskiy E, Cremers D. Proportion priors for image sequence segmentation. In: *Proceedings of the IEEE International Conference on Computer Vision*. Sydney: IEEE, 2013. 2328–2335
- 34 Wang W, Zhu L, Qin J, et al. Multiscale geodesic active contours for ultrasound image segmentation using speckle reducing anisotropic diffusion. *Opt Laser Eng*, 2014, 54: 105–116
- 35 Tierney S, Gao J, Guo Y. Subspace clustering for sequential data. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 2014. 1019–1026
- 36 Wu F, Hu Y, Gao J, et al. Ordered subspace clustering with block-diagonal priors. *IEEE Trans Cybern*, 2015, 46: 1–11
- 37 Donoho D L. Compressed sensing. *IEEE Trans Inf Theory*, 2006, 52: 1289–1306
- 38 Cohen A, Dahmen W, DeVore R. Compressed sensing and best k-term approximation. *J Am Math Soc*, 2009, 22: 211–231
- 39 Mairal J, Elad M, Sapiro G. Sparse representation for color image restoration. *IEEE Trans Image Process*, 2008, 17: 53–69
- 40 Zhao C, Wang X, Cham W K. Background subtraction via robust dictionary learning. *EURASIP J Image Video Process*, 2011, 2011: 1–12
- 41 Ying P, Liu J, Lu H. Dictionary learning based superpixels clustering for weakly-supervised semantic segmentation. In: *Proceedings of IEEE International Conference on Image Processing*, Quebec, 2015. 4258–4262
- 42 Chen Y, Shi L, Feng Q, et al. Artifact suppressed dictionary learning for low-dose CT image processing. *IEEE Trans Med Imag*, 2014, 33: 2271–2292
- 43 Lecellier F, Fadili J, Jehan-Besson S, et al. Region-based active contours and sparse representations for texture segmentation. In: *Proceedings of the 19th International Conference on Pattern Recognition*, Tampa, 2008. 1–4
- 44 Gao Y, Bouix S, Shenton M, et al. Sparse texture active contour. *IEEE Trans Image Process*, 2013, 22: 3866–3878
- 45 Lui D, Scharfenberger C, Fergani K, et al. Enhanced decoupled active contour using structural and textural variation energy functionals. *IEEE Trans Image Process*, 2014, 23: 855–869
- 46 Nieuwenhuis C, Hawe S, Kleinstenber M, et al. Co-sparse textural similarity for interactive segmentation. In: *Proceedings of European Conference on Computer Vision*. Zurich: Springer, 2014. 285–301
- 47 Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*, 2006, 54: 4311–4322

- 48 Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans Image Process*, 2006, 15: 3736–3745
- 49 Horn R A, Johnson C R. *Matrix Analysis*. Cambridge: Cambridge University Press, 2012
- 50 Schme J. Matrix algebra useful for statistics. *Technometrics*, 1985, 27: 89
- 51 Candès E J, Recht B. Exact matrix completion via convex optimization. *Found Comput Math*, 2009, 9: 717
- 52 Candès E J, Li X, Ma Y, et al. Robust principal component analysis. *J ACM*, 2011, 58: 11
- 53 Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imag Sci*, 2009, 2: 183–202
- 54 Nesterov Y. Gradient methods for minimizing composite objective function. *Core Discuss Papers*, 2015, 38: 768–785
- 55 Cai J F, Candès E J, Shen Z. A singular value thresholding algorithm for matrix completion. *SIAM J Optim*, 2010, 20: 1956–1982
- 56 Li S, Li K, Fu Y. Temporal subspace clustering for human motion segmentation. In: *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 2015. 4453–4461
- 57 Qin X, Tian Y, Yan P. Feature competition and partial sparse shape modeling for cardiac image sequences segmentation. *Neurocomputing*, 2015, 149: 904–913

Sparse feature competition and shapes similarity based ultrasound image sequence segmentation method

Bo NI^{1*}, Xiantao CAI², Danjun DENG¹ & Yong YUAN¹

1. *Computer School, Hubei Polytechnic University, Huangshi 435003, China;*

2. *Computer School, Wuhan University, Wuhan 430072, China*

* Corresponding author. E-mail: nibo0906@sina.com

Abstract Segmentation of the object region in ultrasound image sequences plays a crucial role in many clinical applications. Active contour models are widely used in medical image segmentation. Yet, the inherent limitations of ultrasound imaging, such as its low signal-to-noise ratio or local fuzzy boundary caused by inhomogeneous distribution of intensity, have prevented the classical active contours from yielding satisfactory results. To cope with these limitations, many methods have been proposed for modeling image features or shapes prior to constraining active contours. However, most features are limited to constructing the object region's image features, and the object shapes priors usually are learnt from a large set of examples. This paper proposes a sparse feature complete and shapes similarity based method for segmentation of ultrasound image sequences. In this method, a novel contour searching strategy, named sparse feature competition, is proposed by exploiting the reconstruction errors based on both the object and the background feature dictionaries, which is used to alleviate the defects in ultrasound images. By proving that the variation in the object shape has the low-rank property in a linear space, the similarity between the object shapes in the image sequence is also exploited as a shapes prior, which can be interpreted as an unsupervised approach to the shapes prior modeling. To validate the performance of our method, sequences of clinical images were used as the training and test set. The proposed method was compared with three well-known methods on the same test set. The results demonstrate that the proposed method can consistently improve the performance of active contour models and increase the robustness against image defects; consequently, it is likely to improve the efficiency of computer-assisted therapy.

Keywords ultrasound image sequence segmentation, active contour model, sparse feature competition, shapes similarity

Bo NI was born in 1980. He received his Ph.D. degree in computer science from Wuhan University, Wuhan, China, in 2016. He has been a teacher with Hubei Polytechnic University, Hubei, China. His main research interests include image processing, machine learning, pattern recognition and computer vision.

Xiantao CAI was born in 1981. He received his Ph.D. degree in computer science from Wuhan University, Wuhan, China, in 2009. Presently, he is a senior researcher at the School of Computer Science, Wuhan University, Wuhan, China. His research interests include computer aided design, computer supported cooperative work, and cloud manufacturing.

Danjun DENG was born in 1981. She received her Master's degree from the Wuhan University of Technology, Wuhan, China, in 2011. Currently, she is a lecturer at Hubei Polytechnic University, Hubei, China. Her research interests include information retrieval and medical image processing.

Yong YUAN was born in 1983. He graduated from the Beijing Institute of Petrochemical Technology, Beijing, China, in 2005. He received his Master's degree from Wuhan University, Wuhan, China, in 2009. His research interests include software engineering and machine learning.