

基于 Huber 函数双边全变分的多帧文档 图像超分辨率重建*

梁风梅¹, 邢剑卿¹, 罗中良², 邓雪晴²
(1. 太原理工大学信息工程学院, 山西 太原 030024)
(2. 惠州学院计算机系, 广东 惠州 516007)

摘要: 针对低分辨率文档图像中噪声模型不确定、字符边缘和纹理走向复杂多变的问题, 提出 Geman&McClure (G&M) 范数替代 L1、L2 范数用于提高算法的鲁棒性, 设计了结合双边全变分 (BTV) 和 Huber 函数的正则化项, 采用 Lucas-Kanade 光流配准算法, 利用字符结构特征的先验信息, 使算法在重建过程中更加关注边缘细节与边缘方向信息。实验表明, 与 L1BTV、L2BTV 和无 Huber 函数的 G&MBTV 正则化 (下文简称 G&M 方法) 重建方法相比, 文中算法在混合噪声模型下能够显著平滑噪声、锐化边缘、提升文档图像字符的分辨率, 字符识别率提高 14.69% 的同时运算时间缩短了 29.34%。

关键词: 超分辨率重建; 文档图像; 正则化; Huber 函数; BTV

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 0529-6579 (2014) 04-0074-06

The Super-resolution Reconstruction of Multi-frames Document Images Based on Huber Function and Bilateral Total Variation

LIANG Fengmei¹, XING Jianqing¹, LUO Zhongliang², DENG Xueqing²

(1. College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China;

2. Department of Computer Science, Huizhou University, Huizhou 516007, China)

Abstract: It has been a key problem in document image super-resolution which the noise model is uncertain, edge and texture of characters tend to complex and changeable. Geman&McClure (G&M) norm instead of L1 and L2 norm was proposed and used to improve the robustness of the algorithm. A regularization item combining Bilateral Total Variation (BTV) and Huber function was designed with adopting Lucas-Kanade light flow registration algorithm which is highly compatible for the proposed algorithm. This algorithm uses full information of the characters structure characteristics to make the algorithm pays more attention to the edge details in the process of reconstruction, applies the edge direction information more efficiently, and promotes information fusion of a series of low resolution images. Finally, experiments corresponding to L1BTV, L2BTV and G&M BTV regularization algorithm without Huber function were carried out, the results show that the algorithm under the G&M norm with combination of BTV and Huber function is much more excellent than other algorithms. In the environment of mixed noise model, document image super-resolution reconstruction can smooth the noise significantly, sharpen the edge and improve the resolution of characters in document image. The recognition rate of characters was improved by 14.69%, and the operation time of the proposed algorithm was shorten by 29.34% at the same time.

Key words: super-resolution reconstruction; document image; regularization; Huber function; BTV

* 收稿日期: 2013-12-30

基金项目: 山西省自然科学基金资助项目 (2013011017-3)

作者简介: 梁风梅 (1969 年生), 女; 研究方向: 图像处理; E-mail: fm_liang@163.com

自从 Tsai 和 Huang^[1] 提出可以由一组低分辨率 (LR) 图像序列恢复出单帧高分辨率 (HR) 图像的理论, 图像超分辨率重建一直是图像处理领域研究的热点。

文档图像超分辨率重建主要用于光学字符识别软件 (OCR) 的预处理过程, 也可用于从模糊的监控视频中重建 HR 文档图像^[2], 或者对 web 页面的模糊文档进行实时重建。常见的图像插值方法远远达不到理想的重建质量, 近期国内外学者设计了一些主要针对文档图像的重建技术。Fan 等^[3] 利用局部约束一致性在重建过程进行线性正则化, 并自适应生成最佳匹配的 HR 文档图像块。Patanavijit^[4] 运用 M3SS 基于块仿射配准方法, 结合分块技术和仿射模型对降质文档图像重建。Nasonov 等^[5] 把差分稳定双边全变分和双峰惩罚函数引入正则化项, 使算法在文档图像重建中具有很强的去振铃和去噪能力。Kumar^[6] 在基于稀疏表示的重建基础上, 指出虽然字符形状不一致, 但是其边缘和笔画曲线等特征都是相似的, 最后通过字典训练出了稀疏分解的理想 HR 图像块。另外, 贝叶斯估计方法在文档图像重建中也受到普遍关注^[7]。

但是, 上述方法都存在着算法复杂度高、噪声模型单一、实时性和鲁棒性差等缺点。鉴于此, 本文选用算法复杂度较低的正则化方法, 减小配准偏差的干扰, 从而提升重建效率。目标函数数据拟合项中的 L1 范数有利于保持边缘^[8], 主要针对拉普拉斯噪声模型; L2 范数有利于平滑噪声^[9], 主要针对高斯噪声模型。然而, 实际图像中总是夹杂着加性高斯噪声和拉普拉斯脉冲噪声^[10], 而且在重建过程中噪声及其分布模型都有可能发生突变, 所以上述方法都不能保证高质量的重建。本文提出利用不针对特定噪声模型的 G&M 范数作为数据拟合项, 并在 BTV 正则项的基础上引入了 Huber 函数, 使其在水平、垂直、对角线方向估计像素点的梯度值, 利用局部结构特征辨别边缘和平面区域, 克服 TV 和 BTV 正则化方法不能充分利用边缘方向信息的缺陷^[11]。

1 图像退化模型

通常把 HR 图像的降质过程作为观察模型, 通过降质模型建立 HR 和 LR 之间的关系, 最后生成求解 HR 的方程。设 HR 图像大小为 $L_1 N_1 \times L_2 N_2$, L_1 、 L_2 分别为垂直和水平方向上的下采样因子, 则 LR 图像的尺寸为 $N_1 \times N_2$ 。LR 图像是 HR 图像通过下采样、平移、模糊、加噪声等一系列操作获

得, 图像的降质模型为:

$$u_k = \mathbf{D}_k \mathbf{H}_k \mathbf{B}_k z + v \quad k = 1, 2, 3, \dots \quad (1)$$

k 为 LR 图像序列的帧数, u_k 为第 k 帧 LR 图像, \mathbf{D}_k 为下采样矩阵, 大小为 $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$; \mathbf{H}_k 为运动矩阵, 大小为 $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$; \mathbf{B}_k 为模糊矩阵, 大小为 $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$ 。 z 是原始图像信号, v 为加性高斯噪声。

多帧 LR 图像重建时需要融合图像之间的相关信息, 由于本身分辨率低且帧数多, 所以对运动估计的精确度要求很高。本文选用的 Lucas-Kanade (LK) 光流配准算法采取金字塔分层处理^[12], 可在精确计算 H_k 前提下减小计算量, 能够正确检测亚像素级的移动距离, 更重要的是, 它能准确反映运动边缘部分的光流信息, 而且对平面区域的光流变化较敏感。

令 $A_k = \mathbf{D}_k \mathbf{H}_k \mathbf{B}_k$, 线性算子 A_k 尺寸为 $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$, 有

$$u_k = A_k z + v \quad (2)$$

由 (2) 式可知, 图像的重建过程就是通过退化模型估计反求 z 的过程。

2 正则化设计

正则化过程中依据解的先验信息构建附加限制条件, 以保证重建解的确定性和唯一性。利用正则化方法分析退化模型, 可以构造出加入求解限制条件的目标方程:

$$\hat{z} = \operatorname{argmin} \left[\sum_{k=1}^N f_{G\&M}(z) A_k z - u_k + \lambda \Omega(z) \right] \quad (3)$$

$$N = L_1 N_1 \times L_2 N_2$$

式中 $\sum_{k=1}^N f_{G\&M}(z) A_k z - u_k$ 为数据拟合项, 用来衡量 HR 和 LR 之间的差距, 消除数据中配准误差产生的局部异常值。 $f_{G\&M}(z)$ 为 Geman&McClure 范数函数。

$$f_{G\&M}(z) = \frac{z^2}{1 + z^2} \quad (4)$$

λ 为正则化系数, λ 取值偏大, 则重建结果就偏于平滑, 不能突出体现边缘区域; λ 取值偏小, 则数据拟合项的值偏低, 对配准误差不敏感。所以, 对 λ 的适当选取是获取高质量重建图像的重要前提, 本文选取 λ 的经验值为 0.7。

2.1 BTV 正则化

$\Omega(z)$ 为正则化项, 有助于算法从最后的解中剔除伪解, 加快迭代速率, 补偿先验知识。本文设计了结合 BTV 和 Huber 函数的正则化项, 力求在保证求得最优解的前提下通过正则化最大限度地运

用字符局部结构特征信息重建最佳效果。

$$\Omega(z) = \text{BTV}(z) + a\rho(z) \quad (5)$$

$$\text{BTV}(z) = \sum_{s,t=-p}^p r^{|s|+|t|} |S_x^s S_y^t z - z| \quad (6)$$

利用 BTV 双边滤波可以消除图像直方图双峰之间的噪声值, 即抑制噪声, 让重建后文档图像灰度值呈双峰分布。 r 为标量权重, 用于为正则化项之和引入空间衰减效应, $r = 0.8$ 。 S_x^s 表示 z 沿 x 轴方向平移 s 个像素点, S_y^t 表示 z 沿 y 轴方向平移 t 个像素点。参数 p 指定了方向的数目, p 值越高, 重建质量越高, 但会增加计算量, 权衡后本文选取 $p = 1$ 。

2.2 Huber 先验模型

文档图像具有梯度值分段光滑性质^[13], 梯度值在图像的边缘像素点处较大, 在平面像素点处较小, 因此, 可根据图像梯度统计信息设定一个阈值区分边缘和平面区域。(7)式中 a 为边缘梯度阈值。

$$\rho(z) = \begin{cases} z^2 & |z| \leq a \\ 2a|z| - a^2 & |z| > a \end{cases} \quad (7)$$

梯度值较大的像素点对应着边缘区域, 即 $|z| > a$, Huber 函数表达式为 $2a|z| - a^2$, 观察图中函数值偏小, 此时可加大 $\Omega(z)$ 项中 BTV (z) 的作用力度, 有利于消除字符边缘上的异常值。梯度值较小的像素点对应着平面区域, 即 $|z| \leq a$, Huber 函数表达式为 z^2 , 观察图中函数值偏大, 可增强自身 $\rho(z)$ 项的作用, 使正则化项在平面区域能更注重细节, 充分利用字符局部结构特征信息, 见图 1。

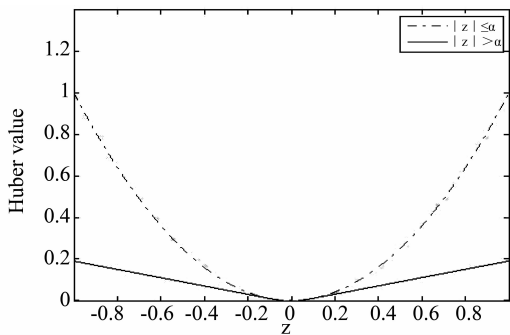


图 1 Huber 函数图

Fig. 1 The figure of Huber function

图 2 以中文字符“木”和英文字符“C”体现了中英文字符中典型的结构特征。Huber 函数在正则化项中补偿字符结构特征的先验信息, 以克服字符边缘纹理走向复杂多变的难点, 并分别在 0° 、 45° 、 90° 、 135° 这四个方向上利用局部近似微分估计出像素点的梯度值:

$$G_{0^\circ} = \frac{\partial \rho(z)}{\partial z}$$

$$G_{45^\circ} = \frac{\partial \rho(z+1)}{\partial z} - \frac{\partial \rho(z)}{\partial z}$$

$$G_{90^\circ} = \frac{\partial \rho(z+1)}{\partial z} + \frac{\partial \rho(z-1)}{\partial z}$$

$$G_{135^\circ} = \frac{\partial \rho(z-1)}{\partial z} - \frac{\partial \rho(z)}{\partial z} \quad (8)$$

$\frac{\partial \rho(z)}{\partial z}$ 为中心像素点偏微分, $\frac{\partial \rho(z+1)}{\partial z}$ 、 $\frac{\partial \rho(z-1)}{\partial z}$ 为中心像素点邻域偏微分。

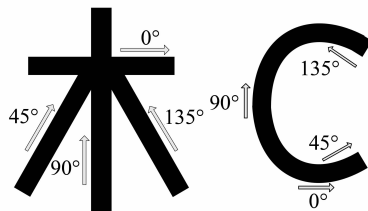


图 2 字符结构特征

Fig. 2 Structural features of characters

2.3 算法实现

将 (3) 式中的 \hat{z} 对 z 求微分:

$$\frac{d\hat{z}}{dz} = \sum_{k=1}^N f'_{G\&M}(z) \cdot A_k^T (A_k z - u_k) + \lambda \left[\sum_{s,t=-p}^p r^{|s|+|t|} (I - S_y^{-t} S_x^{-s}) \text{sign}(z - S_x^s S_y^t z) + aH'(z) \right] \quad (9)$$

本文采用最速下降法得到如下迭代式, 并求解 (3) 式的极小化问题, 加快收敛速度得到最终解实现重建。

$$\hat{z}_{n+1} = \hat{z}_n - \beta \left\{ \sum_{k=1}^N f'_{G\&M}(z) \cdot A_k^T (A_k \hat{z}_n - u_k) + \lambda \left[\sum_{s,t=-p}^p r^{|s|+|t|} (I - S_y^{-t} S_x^{-s}) \cdot \text{sign}(\hat{z}_n - S_x^s S_y^t \hat{z}_n) + aH'(\hat{z}_n) \right] \right\} \quad (10)$$

β 为梯度方向上的迭代步长, 在初始迭代阶段取偏大值以加快收敛速度, 在即将收敛阶段取偏小值以保证重建精确度。

3 实验结果及分析

实验在 matlabR2012 平台上进行。实验选用的原始 HR 图像大小为 128×128 。首先对文档 HR 图像进行下采样、平移、模糊、加噪声获取 30 帧 LR 图像。下采样因子为 2, 获得尺寸为 64×64 的 LR 图像, 然后对图像在水平和垂直方向上进行随机平移运动。为了模拟摄像头的 PSF 模糊效果,

将图像与核大小为 3×3 标准偏差等于 1 的对称高斯低通滤波器卷积运算。同时加入信噪比 (SNR) 为 12 dB 的高斯和拉普拉斯混合噪声，以验证算法的鲁棒性。所有方法迭代次数都为 20 次。

3.1 光流配准

首先将 LR 图像双线性插值到高分辨率栅格中，上采样因子为 2。分别选择序列图像中第一帧

和第二帧作为参考帧和配准帧，然后将配准结果作为和下一帧 LR 图像配准的参考帧。图 3 和图 4 分别为中英文字符的配准过程，观察 (c) 图可以发现 LK 光流配准算法能准确地检测出亚像素级的运动，并且对噪声不敏感，其光流场连续光滑。(d) 图为配准融合图像信息之后的结果。

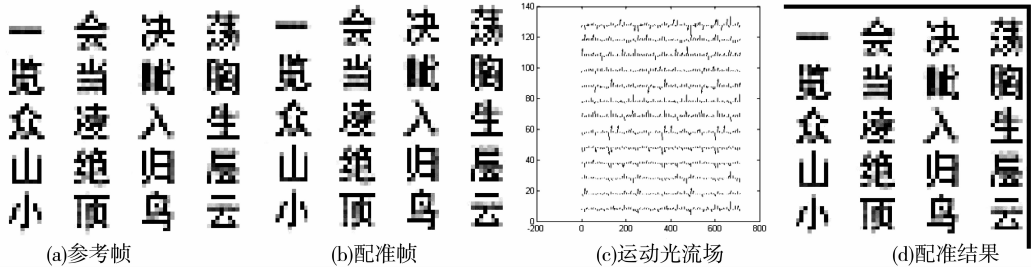


图 3 中文字符文档图像配准

Fig. 3 Registration of Chinese characters image

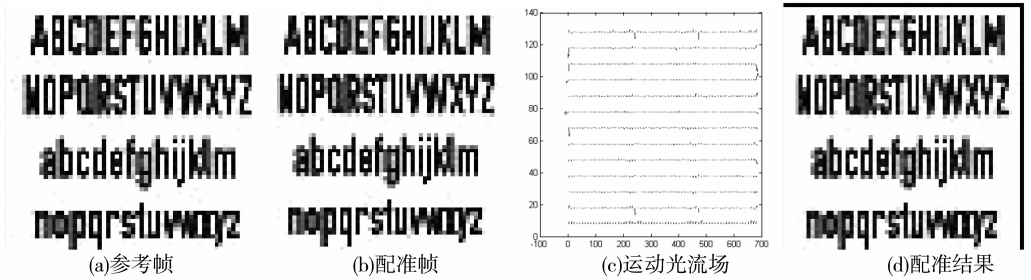


图 4 英文字符文档图像配准

Fig. 4 Registration of English characters image

3.2 重建效果分析

图 5 和图 6 中 L1BTV 和 L2BTV 能在一定程度上锐化字符边缘，但是由于其方法只能限于消除特定的模糊噪声，所以重建结果受到噪声的污染较为严重，导致重建图像出现明显的振铃效应，无法正确分辨具有复杂笔画的字符，如图 5 (b) 和 (c) 中的汉字“眦”。在图 6 中英文字符间距较小，采用 L1BTV 和 L2BTV 方法重建时由于 LR 图像边缘连续性不稳定，导致出现某些英文字符连笔

的现象，如 (b) 和 (c) 图中的字母“Q”和“R”等。(d) 图运用的是 G&M 方法，有效地消除了噪声，但是由于正则化项中没有加入 Huber 函数，不能充分利用边缘方向上的信息，导致字符笔画不均匀，甚至也出现笔画严重变形、连笔等现象。本文方法充分利用字符结构特征的先验信息，在平滑平面区域同时锐化边缘，克服了字符笔画复杂和间距小等难点，重建结果具有较高的分辨率，基本上已经消除了噪声，能够清晰地辨别出字符。

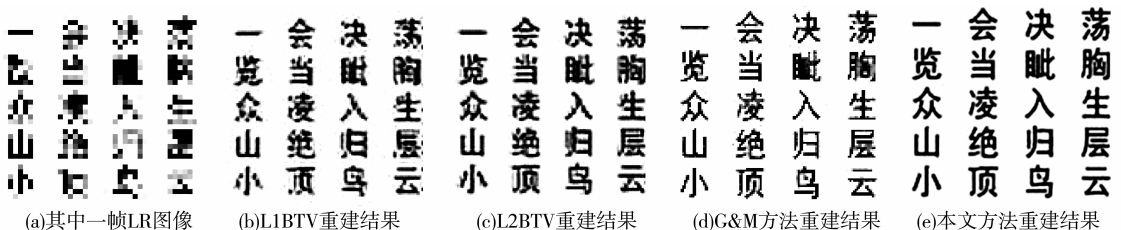


图 5 中文字符文档图像重建结果比较

Fig. 5 Reconstruction results of Chinese characters image

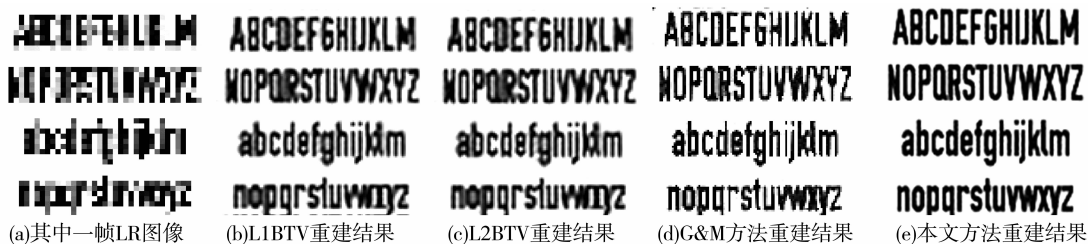


图 6 英文字符文档图像重建结果比较

Fig. 6 Reconstruction results of English characters image

3.3 重建质量评价

本文用 PSNR 值客观评价三种方法的重建结果质量, PSNR 由 (11) 式求得, 具体值见表 1。

$$\text{PSNR} = 10 \log_{10} \frac{L_1 N_1 \times L_2 N_2 \times 255^2}{\|\hat{z} - z\|^2} \quad (11)$$

$L_1 N_1 \times L_2 N_2$ 为图像尺寸, \hat{z} 和 z 分别为重建结果和原始图像。从表 1 中可知本文方法 PSNR 值略高于其他三种方法, 说明本文方法重建结果更接近原始 HR 图像, 同时也明显缩短了重建时间, 验证其具有较高的实时性。

表 1 四种方法重建结果的 PSNR 值与重建时间

Table 1 PSNR and reconstruction time of the four reconstructions

字符类型	L1BTV	L2BTV	G&M 方法	本文方法
中文字符/dB	27.31	27.72	28.43	29.30
英文字符/dB	27.48	27.53	28.32	28.96
平均重建时间/s	2.33	2.45	2.03	1.73

图 7 为汉字和英文字符的识别率对比折线图, 识别率是利用汉王 OCR 识别软件测试得出, 图中字符识别率是各个迭代阶段中英文字符识别率的平均值。观察图中可知本文方法在各个迭代阶段的识别率均高于其他三种方法, 充分说明本文方法的优越性。

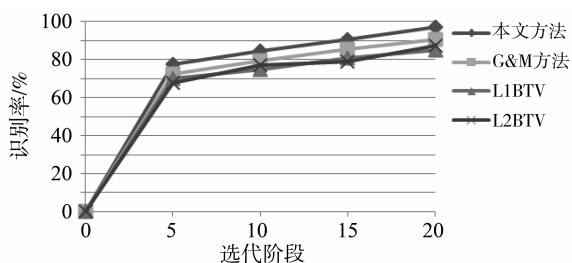


图 7 各个迭代阶段的字符平均识别率比较图

Fig. 7 The comparison of characters average recognition rate in specified iteration stages

4 结 论

本文提出一种基于 Geman&McClure 范数的 Huber 函数 BTV 正则化算法, 并采用高精度的 LK 光流配准方法对中文字符和英文字符 LR 图像进行运动估计。与 L1BTV、L2BTV、G&M 方法得到的重建结果进行对比, 验证了本文方法针对低分辨率文档图像的特征, 能够利用字符结构特征的先验模型, 克服了由于模糊造成严重的边缘不连续现象。在同等条件下, 本文方法比其他算法能够更高效地保持边缘细节、抑制多种未知噪声污染、消除振铃效应, 显著提高 LR 图像中的字符识别率, 缩短运算时间, 适用于 LR 文档图像超分辨率重建。

参考文献:

- [1] HUANG T S, TSAI R Y. Multi-frame image restoration and registration [J]. Advances in Computer Vision Image Processing, 1984, 1:317 - 339.
- [2] 万雪芬, 韩芳. 视频监控图像的超分辨率复原研究 [J]. 激光杂志, 2014, 35(3): 5 - 8.
- [3] FAN W, SUN J. Local consistency constrained adaptive neighbor embedding for text image super-resolution [C] // Proceedings of 10th IAPR International Workshop on Document Analysis System (DAS), 2012: 90 - 94.
- [4] PATANAVIJIT V. Video enhancement using a robust iterative SRR based on a Geman&McClure stochastic estimation with a general observation model [C] // International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTICON), 2010: 875 - 879.
- [5] NASONOV A V, KRYLOV A S. Text images super-resolution and enhancement [C] // 5th International Congress on Image and Signal Processing (CISP), 2012: 617 - 620.
- [6] KUMAR V, BANSAL A. Sparse document image coding for restoration [C] // 12th International Conference on Document Analysis and Recognition (ICDAR), 2013: 713 - 717.

TWA 谱分析检测的影响并得出结论。T 波交替变异现象详细的发生机制和完善的检测技术仍有待于今后进一步研究。

参考文献:

- [1] FURLANELLO F, GALANTI G, MANETTI P. Micro-volt T-wave alternans as predictor of electrophysiological testing results in professional competitive athletes [J]. *Ann Noninvasive Electrocardiol*, 2004, 9(3): 201 - 206.
- [2] VERRIER R L, IKEDA T. Ambulatory ECG-based T-wave alternans monitoring for risk assessment and guiding medical therapy: Mechanisms and clinical applications[J]. *Progress in Cardiovascular Diseases*, 2013, 56(2):172 - 183.
- [3] 王歆月. 微伏级 T 波电交替及其对心源性猝死的预测价值的临床应用[J]. *医学理论与实践*, 2013, 26(23):3107 - 3109.
- [4] 王娟, 黄忠朝, 刘正春. 基于增强的谱分析和奇异值分解的 T 波交替检测[J]. *浙江大学学报:工学版*, 2012, 46(1):177 - 181.
- [5] ARMOUNDAS A A, MELA T, MERCHANT F M. On the estimation of T-wave alternans using the spectral fast fourier transform method [J]. *Heart Rhythm*, 2012, 9(3): 449 - 456.
- [6] BURATTINI L, BINI S, BURATTINI R. Correlation method versus enhanced modified moving average method for automatic detection of T-wave alternans[J]. *Computer Methods and Programs in Biomedicine*, 2010, 98(1): 94 - 102.
- [7] 魏珑, 赵捷, 徐方舟, 等. 基于相关分析方法的非稳态心电 T 波交替检测研究[J]. *现代生物医学工程进展*, 2009, 9(16):3131 - 3134.
- [8] 严抗辉, 万向奎, 李明贵, 等. 非稳态 T 波交替检测:一种改进的相关分析方法研究[J]. *生物医学工程学杂志*, 2013, 30(4):860 - 865.
- [9] STRUMILLO P, RUTA J. Poincaré mapping for detecting abnormal dynamics of cardiac repolarization [J]. *IEEE Eng Med Biol Mag*, 2002, 21(1): 62 - 65.
- [10] 张石, 余黎煌, 徐中强. 心电 T 波电交替检测算法综述[J]. *中国生物医学工程学报*, 2010, 29(3):446 - 453.
- [11] CUESTA F, MICÓ T P, ABOY M. Enhanced modified moving average analysis of T-wave alternans using a curve matching method: a simulation study[J]. *Medical & Biological Engineering & Computing*, 2009, 47(3): 323 - 331.
- [12] MARTÍNEZ J P, OLMOS S. Methodological principles of T wave alternans analysis: a unified framework [J]. *IEEE Transactions on Bio-medical Engineering*, 2005, 52(4):599 - 613.
- [13] 赵捷, 华玫. 基于 T 波最大值的 T 波电交替分析法[J]. *生物医学工程学杂志*, 2005, 22(5):951 - 955.
- [14] NARAYAN S M, BOTTERON G W, SMITH J M. T-wave alternans spectral magnitude is sensitive to ECG beat alignment strategy[J]. *IEEE Trans on Computer in Cardiology*, 1997, 24(3): 593 - 596.
- [7] KATHERINE D, GREGORY K M. Bayesian super-resolution of text in video with a text-specific bimodal prior [J]. *International Journal on Document Analysis and Recognition*, 2005, 7: 159 - 167.
- [8] SONG H H, ZHANG D, WANG P K, et al. An adaptive L1-L2 hybrid error model to super-resolution [C] // 17th IEEE International Conference on Image Processing, 2010:2821 - 2824.
- [9] 郭昌. 小波变换与 HMT 模型的图像插值算法[J]. *中山大学学报:自然科学版*, 2012, 51(3):55 - 59.
- [10] 杨克伟. 基于光流法的迭代反投影超分辨率重构算法[J]. *现代计算机:专业版*, 2014(3):31 - 34.
- [11] XU Z G, SU X Q, ZHANG Z P. Multi-frame image super-resolution by Total-Variation regularization [J]. *Journal of Information and Computational Science*, 2012, 9(4):945 - 953.
- [12] TONG W. Formulation of Lucas-Kanade digital image correlation algorithms for non-contact deformation measurements: a review [J]. *Strain*, 2013, 49(4):313 - 334.
- [13] BANERJEE J, JAWAHAR C V. Super-resolution of text images using edge-directed tangent field [C] // *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems*, 2008:76 - 83.

(上接第 78 页)