

MRF 随机场在文本图像清晰化算法中的应用

沈 剑, 路林吉

(上海交通大学 自动化系, 上海 200240)

摘 要:对于扫描或相机拍摄的低分辨率文本图片, 提出一种基于马尔科夫随机场的文本图像清晰化算法。该算法首先引入了马尔科夫随机场模型, 然后考虑平衡数据项与平滑项之间的关系从而确定出能量函数的表达式。利用改进过的置信度传播方法来解决马尔科夫随机场中的最大后验估计问题, 以达到最小化能量函数的目的。实验结果表明, 本算法能够较有效地提高低分辨率图像的清晰度, 一定程度上提高了图片的分辨率, 尤其对文字的边缘部分有很好的锐化效果。

关键词:图像复原; 马尔科夫随机场; 数据项约束; 双峰值模型; 置信度传播

中图分类号: TP39

文献标识码: A

文章编号: 1673-629X(2011)06-0009-03

Enhancement of Document Text Images Using MRF Approach

SHEN Jian, LU Lin-ji

(Automation Department of Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: Document text images captured by book scanner or camera usually require resolution enhancement for other usage, present a text image enhancement algorithm based on Markov random field. Firstly the algorithm introduces the MRF model, and then defines the energy function considering the relation between data constraint and smoothness constraint. The improved belief propagation was implemented to resolve the MAP (Maximum A Posterior) of MRF model, in order to minimize the energy function. The experiment result shows that the algorithm can efficiently enhance the low resolution text image and improve the image resolution to some extent. And the edge of the text has a good sharpening effect.

Key words: image restoration; Markov random field; data constraint; bimodality model; belief propagation

0 引言

图像在传送和转换过程中, 如成像、扫描、复制、传输、显示等, 会造成图像的降质, 造成图像模糊、失真、引入噪声等等。为了抑制噪声、改善图像质量, 图像复原技术具有很重要的意义^[1,2]。

马尔科夫随机场模型为解决图像恢复问题, 提供了统一的、鲁棒的框架。文中首先提出图像随机场模型, 并对能量函数的数据项和平滑项进行分析。引入双峰值模型对低分辨率灰度文本图片构造数据项, 并给出适应性平滑约束。运用置信度传播 (Belief Propagation, BP) 算法^[3], 通过分层技术减少 BP 算法的迭代次数, 提高消息的收敛速度。最后得到的效果图维持或增加了文本字符和背景之间的差别, 显著提高了对比度。

1 MRF 随机场的图像模型以及 BP 最大后验估计

用一个无向图表示 MRF^[4], 图 1 中白色点代表一幅图像中所有像素点的集合 P , 黑色点是有限标号集 L , 在这里代表像素的灰度值。一个标记 f 就是给像素集 P 中的某个元素 $p \in P$ 赋一个标号 $f_p \in L$ 。事实上, 黑色点是观测顶点, 与白色像素点组成了数据边。而白色像素顶点组成了平滑边。最大后验 (MAP) 估计问题在这里就是要寻求一个标记, 使得能量值最小的 MRF 随机场得以实现^[5]。

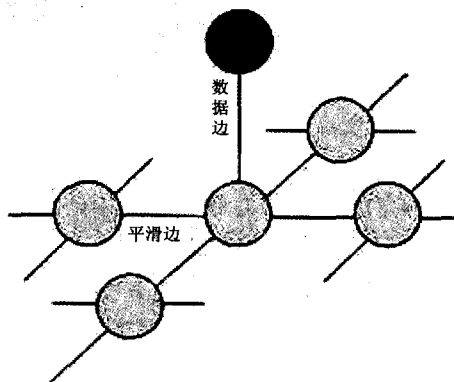


图 1 图像随机场模型

收稿日期: 2010-10-26; 修回日期: 2011-03-06

基金项目: 国家科技型技术创新基金项目 (07C26213101270)

作者简介: 沈 剑 (1986-), 男, 江苏南通人, 硕士研究生, 研究方向为计算机视觉与图像处理; 路林吉, 副教授, 博士, 研究方向为计算机视觉与图像处理、工业自动化、嵌入式系统。

在灰度文本图像中,平滑区域标号变化不大,而在边界处非常明显,因此通常考虑基于标号问题的能量函数:

$$E(f) = \sum_{p \in P} D_p(f_p) + \sum_{\{p,q\} \in N} V_{\{p,q\}}(f_p - f_q) \quad (1)$$

BP 算法的基本过程如下^[6]:构造(1)式中的能量函数,在满足一定条件下建立随机场中 MRF 置信度与能量值之间的关系,进而求出置信度的最小值,达到求解图模型 MAP 问题的近似值的目的。

$$m_{p \rightarrow q}^i(f_q) = \min_{f_p} (V(f_p - f_q) + D_p(f_p) + \sum_{s \in N(p) \setminus q} m_{s \rightarrow p}^{i-1}(f_p)) \quad (2)$$

其中, $m_{p \rightarrow q}^i$ 表示第 i 次迭代时 p 节点传递给相邻节点 q 的消息, $N(p) \setminus q$ 表示不包含 q 的 p 邻域。

T 次迭代后为每个节点计算信任度向量:

$$b_q(f_q) = D_q(f_q) + \sum_{p \in N(q)} m_{p \rightarrow q}^T(f_q) \quad (3)$$

节点 q 最终的标号 f_q^* 是向量 $b_q(f_q)$ 中最小的元素对应的标号。可以求出本算法的计算复杂度为 $O(nk^2T)$ 。其中, n 为像素点个数, k 是每个像素可能的标号数, T 表示消息传播过程中迭代的次数。最小化置信度可以获得网络中的最优化解。

2 自适应的数据项和平滑项

数据项在这里表示一对对应的像素点灰度值是否有能够相容的局部特征。灰度的文本图像在前景和背景区域都是平滑的,只是在边缘有尖锐的转变。所以一般灰度文本图像都有双峰值分布^[7],如图 2 所示。

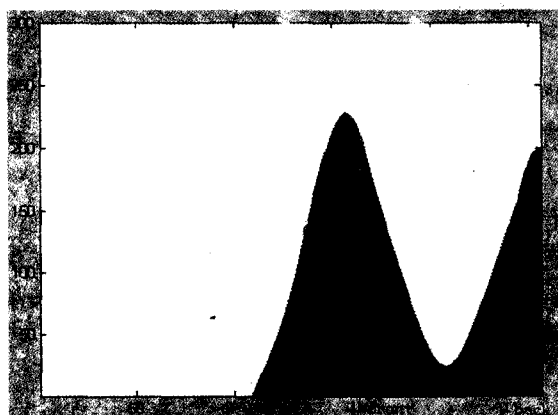


图 2 双峰值分布的灰度直方图

(μ_{back} 代表白色背景, μ_{front} 代表黑色文字)

在黑色文字和白色背景峰值之间存在少量的灰度值,是从黑到白过渡的灰度像素点。中间灰度值像素点的数量与文档图像的模糊程度有关。为了获得一个尖锐的双峰图像,可以借鉴类似直方图拉伸的思想,定义一个能量函数,把过渡段的灰度像素点拉至邻近的二峰值之一^[8]。

$$B_p(f_p) = (f_p - \mu_{\text{back}})^2 (f_p - \mu_{\text{front}})^2 \quad (4)$$

其中: μ_{front} 和 μ_{back} 为前景和背景的尖峰值; $B_p(f_p)$ 表示把标签 f_p 赋值给像素 p 后,与输入数据之间的视差,即双峰数值代价。

文中将数据项定义为:

$$D_p(f_p) = [I(p) - f_p]^2 + B_p(f_p) \quad (5)$$

在文本图像中,尖锐的边缘对应沿着边缘聚集的相对较大强度的文本图像梯度,而相对平滑的区域中只有较弱强度或没有文本图像梯度。高分辨率和低分辨率图像间的关系本质上依赖于边缘方向的平滑性。文本字符边缘通常包含分段平滑的曲线,而曲线的交点即为字符的拐点。好的增强方法需要区分文本图像中的平滑曲线和拐角,所以必须保持沿着文本字符边缘的平滑性,以及穿越边缘的尖锐的非连续性。在平坦区域,能量函数更多地依赖于平滑项同时减少数据项的惩罚量。在图像的边缘处,情况刚好相反。能够自适应地平衡数据项与平滑项间的关系是很关键的。

为了找到边缘方向,按 x 轴和 y 轴方向计算图像梯度场,分别表示为: T_x 和 T_y 。

定义适应性的平滑项如下:

$$V(f_p - f_q) = \begin{cases} \min[\lambda T_x(f_p - f_q)^2, \rho], & \text{如果}(p, q) \text{ 沿着 } x \text{ 轴} \\ \min[\lambda T_y(f_p - f_q)^2, \rho], & \text{如果}(p, q) \text{ 沿着 } y \text{ 轴} \end{cases} \quad (6)$$

其中 λ 控制增长率,为平滑项的适应性因子,用来调节不同灰度值的邻接标号像素对标签和像素点间的差值,即视差的惩罚力度; ρ 控制平滑项的增长幅度,为截断常量,它允许纹理相似区域内的标签与像素差值在小范围内变化。

亮度变化和视差变化之间有一定的内在联系,视差的连续性往往表现在亮度的连续性,而亮度边缘的地方有较多的视差不一致性^[9]。所以,在领域内亮度相似的区域取较大的 λ 值惩罚视差的不连续,反之取较小的 λ 值奖励视差的不连续,如下(7)式所示。

$$\lambda = \begin{cases} \frac{|I(x) - I(y)|}{255}, & |I(x) - I(y)| \geq \theta \\ 1, & \text{else} \end{cases} \quad (7)$$

基于 MRF 模型的方法是一种隐式的边缘导向的方法。在此方法中,边缘像素的边缘方向由此方向的强度连续性所指示。用切向区域来测量每个方向的强度连续性。这些值从像素强度变化,即从图像梯度场获得。该方向相应的强度连续性可被用于边缘方向信息来形成几何空间约束的规律,这可被归纳为沿着边缘方向的平滑性和穿越边缘方向的尖锐性。

3 多尺度的 BP 算法

在 MRF 模型下最小化全局能量函数。置信度传播(Belief Propagation 即 BP)利用迭代的原理寻找图像

的全局能量最小值。一般传统的 BP 算法需要对消息进行大量的迭代运算以达到收敛,而这将消耗大量的计算时间。可以用多尺度的技巧来减少迭代运算的次数^[10]。基本想法是将原始图像按栅格图来处理,保持原有的领域结构顺序,将栅格图按金字塔分层采样^[11],由低分辨率图到高分辨率图进行传输,能够增强不同分辨率栅格层中不同像素块之间的交互能量,并且加快了收敛的速度,减少消息传播中的迭代运算次数。

多尺度 BP 算法首先求出低分辨率图上的消息,然后利用求出的消息初始化下一个分层上的消息。在符合栅格图层次结构和保持能量函数不变的前提下,多尺度的方法在保证计算正确性的同时加快了 BP 算法的收敛速度。

设 $\tau^0, \tau^1, \dots, \tau^l, \dots, \tau^{level-1}$ 表示由精到粗的栅格图集^[12],其中 τ^0 代表原始大小的图像,而任意一中间层 τ^l 中的节点是由 τ^{l-1} 层中的 2×2 的节点块组成,如图 3 所示。

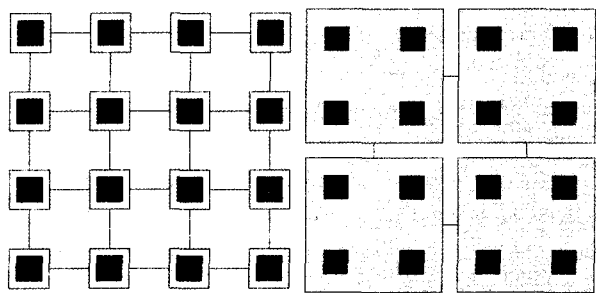


图 3 1 层上一个节点相当于 $l-1$ 层上 2×2 的节点块

f^l 是 τ^l 上一个像素点的标记, l 层上能量函数可以写为:

$$E(f^l) = \sum_{(i,j) \in \tau^l} D_{i,j}^l(f_{i,j}^l) + \sum_{(i,j) \in \tau^{l-1}} V^l(f_{i,j}^l - f_{i,j}^{l-1}) + \sum_{(i,j) \in \tau^{l-1}} V^l(f_{i,j}^l - f_{i,j+1}^l) \quad (8)$$

D^l 、 V^l 分别是 l 层上的数据项和平滑项。数据项 $D_{i,j}^l$, 假设 l 层格图上 (i,j) 处像素节点标号为 α , 将所有像素数据项相加得到:

$$D_{i,j}^l(\alpha) = \sum_{u=0}^{s-1} \sum_{v=0}^{s-1} D_{si+u, sj+v}(\alpha) \quad (9)$$

通过将相邻下一栅格层上相邻四个数据代价相加即可求得本层相应节点的数据项 $D_{i,j}^l$, 得到了数据代价, 避免每次都在初始栅格图上进行计算, 同时也提高了运算速度。节点块内部的不连续代价为 0, l 层各图上不连续代价定义为:

$$V^l(\alpha - \beta) = \varepsilon V\left(\frac{\alpha - \beta}{\gamma}\right) \quad (10)$$

ε 表示邻接的两个像素块在边界处像素点的个数, γ 表示这两个像素块之间的中心距离。

按照式(6)对应的平滑项, 有:

$$V^l(f_p - f_q) = \begin{cases} \min[\lambda^l T_x(f_p - f_q)^2, \rho], & \text{如果}(p, q) \text{ 沿着 } x \text{ 轴} \\ \min[\lambda^l T_y(f_p - f_q)^2, \rho], & \text{如果}(p, q) \text{ 沿着 } y \text{ 轴} \end{cases} \quad (11)$$

适应性因子 λ^l 的确定方法同式(7), 不同的是:

$$|I(x) - I(y)|_l = \frac{1}{\omega} \sum_{x \in p^l, y \in q^l, x \in N(y)} |I(x) - I(y)| = \frac{1}{2} \sum |I(x) - I(y)|_{l-1} \quad (12)$$

p^l, q^l 是 τ^l 层的像素点或节点块, ω 表示相邻两像素块交界上的像素点个数。

每一个栅格层上的消息都表示不同大小像素块间的关系, 且对应着同一个标号集的函数。用 l 层栅格图上某个节点块最终发出的消息, 初始化 $l-1$ 层栅格图上构成 l 层这个节点块四个小块的消息:

$$m_{p_{i,j}^{l-1} \rightarrow q_{i,j}^{l-1}}^{(l-1,1)} = m_{p-q}^{(l,T)} \quad (13)$$

如此以往, 最终传递到 τ^0 。每层消息在上、下、左、右四个方向上传播。

4 实验结果

实验中选取原始文本图像分辨率为 200×38 , 如图 4(a) 所示。低分辨率图片是对应高分辨率图片高斯模糊以后得到的, 这里采用的高斯模糊, 5×5 的卷积核, 标准差为 1.4。对于低分辨率图片分别采用双线性插值和三次样条插值, 以及文中的算法进行实验。BP 算法每层格图迭代 4 次, 重复迭代五层格图。从实验结果可以看到, 文字的边缘有强烈的锐化效果。当文字周围有一些孤立噪声点时, 可以通过中值滤波将噪声点滤除^[13]。

resolution

(a) 高分辨率图像

resolution

(b) 低分辨率图像

resolution

(c) 双线性插值

resolution

(d) 三次样条插值

resolution

(e) 文中提出的算法

图 4 实验结果

5 结束语

文中提出的低分辨率文本图像增强方法, 能够区

(下转第 15 页)

图5为实际检测中调试前波浪旋转架旋转半径为1.5米时所绘制的实时波高采样曲线,图6则为同一采集系统调试后所绘制的相对应的实时波高曲线。

通过对比,可以发现调试后的波高数据采集电路已经能够很好地满足系统的性能要求。当采样结束后,该检测系统会计算出采样区间内的有效波高和相对应的有效波周期,并显示在应用程序界面上,如图4中所示,该数据可以进一步验证波高数据采集电路的性能^[12]。

5 结束语

该数字化检测系统可以以100毫秒甚至更小的采样间隔采集数据,因此所绘制的实时波高曲线更加平滑,有效波周期可以精确到0.1秒,便于检测。经过多次试验验证,收效甚好。

参考文献:

- [1] 叶安乐,李凤歧.物理海洋学[M].青岛:青岛海洋大学出版社,1992.
 - [2] Huang Min-Chih, Chen Jia-Yuan. Wave direction analysis from data buoys ocean engineering[J]. Ocean Engineering, 1998, 25(8): 621-637.
 - [3] 唐原广,李琛.多功能波浪浮标的研制[J].气象水文海洋仪器, 2004(3/4): 12-15.
 - [4] 唐原广,赵曙东.基于单片机的倾角测量系统设计[J].微计算机信息, 2007, 3-2: 96-97.
 - [5] 戴佳,戴卫恒.51单片机C语言应用程序设计实例精讲[M].北京:电子工业出版社,2006.
 - [6] 龚建伟,熊光明. Visual C++/Turbo C 串口通信编程实践[M].北京:电子工业出版社,2004.
 - [7] 李长林,高洁. Visual C++ 串口通信技术与典型实例[M].北京:清华大学出版社,2006.
 - [8] Han Kun Yean. Flooding undation analysis resulting from le-nee-break[J]. Journal of Hydraulic Research, 1998, 136(5): 747-759.
 - [9] 杨高波,亓波.精通 MATLAB7.0 混合编程[M].北京:电子工业出版社,2006.
 - [10] 唐原广,王金平.基于VB和Matlab接口的海浪谱分析软件设计[J].计算机技术与发展, 2008, 18(9): 37-39.
 - [11] Heron S F, Heron M L. A Comparison of Algorithms for Extracting Significant Wave Height form HF Radar Ocean Backscatter Spectra[J]. Journal of Atmospheric and Oceanic Technology, 1998, 15(6): 1157-1163.
 - [12] 赵进平,朱光文.海洋监测仪器设备成果标准化[M].北京:海洋出版社,2004.
-
- (上接第11页)
- 分文本图像中的平滑曲线和拐角,达到保持沿着边缘的平滑性,以及穿越边缘的尖锐的非连续性的效果。在平坦区域,能量函数更多地依赖于平滑项同时减少数据项的惩罚量。在图像的边缘处,情况刚好相反。提出的基于MRF模型的能量函数,利用图像分层技术,构建了由低分辨率到高分辨率的金字塔栅格图层,有效地减少了置信度传播算法中的消息迭代运算次数,提高了运算速度并求出最优解。本算法较好地恢复了文本字符的信息,且无须先验算法训练集,可以运用到文本图像处理的实时清晰化处理中。
- 参考文献:
- [1] 苏秉华,金伟其,牛丽红,等.超分辨率图像复原及其进展[J].光学技术, 2001, 27(1): 6-9.
 - [2] 张红英,彭启琮.数字图像修复技术综述[J].中国图像图形学报, 2007, 12(1): 1-10.
 - [3] Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 25(7): 787-800.
 - [4] 方恒,吴怀宇.基于MRF和颜色空间的立体图像匹配算法[J].计算机技术与发展, 2008, 18(12): 28-31.
 - [5] Geman S, Geman D. Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images[J]. IEEE Trans Pattern Anal Machine Intell, 1984, 6(6): 721-741.
 - [6] Weiss Y, Freeman W T. On the Optimality of Solutions of the Max-Product Belief Propagation Algorithm in Arbitrary Graphs[J]. IEEE Transactions on Information Theory, 2001, 47(2): 736-744.
 - [7] Donaldson K, Myers G K. Bayesian super-resolution of text in video with a text-specific bimodal prior[J]. IJDAR, 2005, 7(2-3): 159-167.
 - [8] Thouin P D, Chang C I. A method for restoration of low-resolution document images[J]. IJDAR, 2000, 2(4): 200-210.
 - [9] 何华君,卢朝阳,焦卫东,等.基于梯度和MRF模型的视差估计算法[J].西安电子科技大学学报, 2007, 34(3): 373-376.
 - [10] 耿英楠,赵岩,陈贺新.基于运动估计的置信度传播立体视频匹配算法[J].吉林大学学报(信息科学版), 2010, 28(4): 329-333.
 - [11] Willsky A S. Multi-Resolution Markov Models for Signal and Image Processing[J]. Proc of the IEEE, 2002, 90(8): 1396-1458.
 - [12] Borgefors G. Distance Transformation in Digital Images[J]. Graphics and Image Processing, 1986, 34(3): 344-371.
 - [13] 刘丽梅,孙玉荣,李莉.中值滤波技术发展研究[J].云南师范大学学报, 2004, 24(1): 23-27.