

基于总变分模型的图片文字干扰消除

董剑龙^{1*}, 王浩全²

¹太原工业学院教务处, 山西 太原 030008;

²中北大学仪器科学与动态测试教育部重点实验室, 山西 太原 030051

摘要 针对图片中文字对视觉的影响, 提出基于 TV 模型的图片文字干扰消除算法。采用 Roberts 算子与二值形态学对待处理图片中干扰文字区域进行检测与准确定位, 将已确定的干扰文字区域作为 TV 模型的约束条件进行求解即可有效消除文字的干扰。实验结果表明, 该方法可以高效地检测和消除图片中的干扰文字。

关键词 图像处理; TV 模型; 文字检测; 文字消除; 二值形态学

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.241004

Elimination of Interferential Text from Images Based on Total Variation Model

Dong Jianlong^{1*}, Wang Haoquan²

¹Academic Affairs Office, Taiyuan Institute of Technology, Taiyuan, Shanxi 030008, China;

²Key Laboratory of Instrumentation Science and Dynamic Measurement, Ministry of Education, North University of China, Taiyuan, Shanxi 030051, China

Abstract In this study, a text interference elimination algorithm is proposed based on the total variation (TV) model with respect to the visual influence of the text in an image. Roberts operator and binary morphology are applied to detect and accurately locate the interferential text area in an image. Subsequently, the interferential text can be effectively eliminated by considering the determined interferential text area as the constraint condition of the TV model and solving it. The simulation experiments denote that the proposed method is efficient and practical for the detection and elimination of interferential text from images.

Key words image processing; total variation model; text detection; text elimination; binary morphology

OCIS codes 100.2000; 100.3008; 100.4994; 100.5010

1 引言

在影视剧制作过程中, 字幕会以各种形式出现在视频影像中。某些老式的视频影像, 由于受当时技术水平的限制, 字幕常被固化在视频中, 无法通过现有技术将其消除或替换, 严重妨碍了不同语种之间的视频交流。

字幕会以不同的程度呈现在视频影像的某些画面中。季丽琴等^[1]提出的基于彩色边缘检测形态逻辑算子“与”(CEMA)和纹理修复的技术对含较少文字的图片具有较好的处理效果。基于区域匹配的方法主要用

来消除块状干扰^[2-3]。将插值法与图像修复相结合^[4-5], 使用方向梯度直方图^[6]对图片中的划痕、文字或块状干扰进行去除均能取得很好的效果, 但算法较为复杂。采用总体变分(TV)模型与其他方法结合^[6-11]可对纹理图片、老旧照片等进行修复。本文将 TV 模型与 Roberts 算子、二值形态学巧妙结合, 用于检测并确定文字所在区域, 然后利用已确定的矩形区域四个角的坐标约束 TV 模型, 有效消除了图片中文字的干扰。

2 文字干扰自动消除原理

图 1 表示受文字干扰图片的待消除区域及其邻

收稿日期: 2019-04-11; 修回日期: 2019-05-07; 录用日期: 2019-06-05

基金项目: 山西省回国留学人员科研资助项目(2017-089)、山西省自然科学基金(201801D121160)

* E-mail: 120126130@qq.com

域,其中 D 为含文字的区域, E 为含文字区域的外邻域,一般取环状。本文采用 TV 模型对图片中文字干扰进行消除的原理是在噪声条件约束下对区域 E 和 D 进行整体变分。

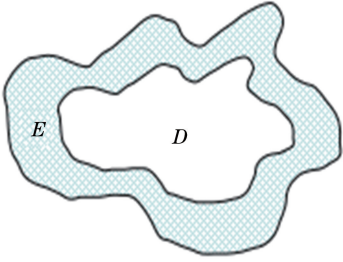


图1 含文字区域及其外邻域示意图

Fig. 1 Schematic of area containing texts and its outer neighborhood

设去除文字干扰后 $E \cup D$ 区域图片(即 E 与 D 的并集)为 $u(x, y)$,为使含文字区域的边界尽量平滑,定义代价函数为

$$R(u) = \int_{E \cup D} |\nabla u(x, y)| dx dy, \quad (1)$$

式中: ∇ 为梯度算子; x, y 分别为空间坐标变量。

噪声约束条件为

$$\frac{1}{A(E)} \int_E |u(x, y) - u_0(x, y)|^2 dx dy = \sigma^2, \quad (2)$$

式中: $A(E)$ 为区域 E 的面积; $u_0(x, y)$ 为 E 区域内受高斯白噪声污染的图片; σ 为高斯白噪声的标准偏差。

引入拉格朗日常数对 TV 文字干扰消除模型进行改进,可得

$$J_\lambda[u(x, y)] = \int_{E \cup D} |\nabla u(x, y)| dx dy + \frac{\lambda}{2} \int_E |u(x, y) - u_0(x, y)|^2 dx dy, \quad (3)$$

$$F[u(x, y)] = |\nabla u(x, y)| +$$

$$\frac{\lambda_f}{2} [u(x, y) - u_0(x, y)]^2,$$

$$\lambda_f = \begin{cases} \lambda, (x, y) \in E \\ 0, (x, y) \in D \end{cases}, \quad (4)$$

式中: λ 为拉格朗日常数; λ_f 为分段函数。将(3)式中的积分内容合并后,引入 $F[\cdot]$ 函数,可得

$$J_\lambda[u(x, y)] = \int_{E \cup D} F[u(x, y)] dx dy, \quad (5)$$

$J_\lambda[u(x, y)]$ 取极值时的 $u(x, y)$ 即为 $F[u(x, y)]$ 取极值时的 $u(x, y)$ 。

结合梯度的概念和欧拉-拉格朗日方程^[12]求解(4)式中 $F[u(x, y)]$ 的解 $u(x, y)$,即求解(6)式中

的 $u(x, y)$,可消除图片中文字的干扰。

$$-\nabla \cdot \left[\frac{\nabla u(x, y)}{|\nabla u(x, y)|} \right] + \lambda_f [u(x, y) - u_0(x, y)] = 0. \quad (6)$$

3 数值实现

(6)式为偏微分方程,在实际图片中无法表示,本文用近似的数值差分方程替代。图2为受文字干扰处像素及其邻域点示意图, O 为待处理像素点, $P_0 = \{N, S, W, E\}$ 为 O 点相邻4个像素点 N, S, W, E 构成的集合, $p = \{n, s, w, e\}$ 为 O 点和4个相邻像素的半像素点 n, s, w, e 的集合。

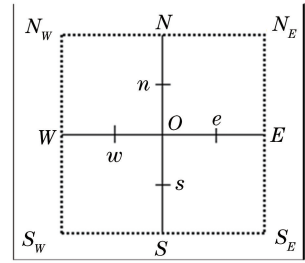


图2 受文字干扰处像素及其邻域点示意图

Fig. 2 Schematic of pixels with text interference and their neighborhood points

令 $\mathbf{V} = (v^{(1)}, v^{(2)}) = \nabla u / |\nabla u|$,其散度为

$$\nabla \cdot \mathbf{V} = \frac{\partial v^{(1)}}{\partial x} + \frac{\partial v^{(2)}}{\partial y} \approx \frac{v_e^{(1)}(1) - v_w^{(1)}(1)}{h} + \frac{v_s^{(2)}(2) - v_n^{(2)}(2)}{h}, \quad (7)$$

式中: \mathbf{V} 为梯度方向; $v^{(1)}$ 为梯度沿水平方向的分量; $v^{(2)}$ 为梯度沿垂直方向的分量; $\nabla \cdot \mathbf{V}$ 为待处理像素处的曲率; $v_e^{(1)}, v_w^{(1)}, v_s^{(2)}, v_n^{(2)}$ 分别为半像素点 e, w, s, n 的梯度值; h 为步长。

设(7)式中的步长 h 取1,下面以半像素点 e 为例计算梯度值。

$$v_e^{(1)}(1) = \frac{1}{|\nabla u_e|} \left[\frac{\partial u}{\partial x} \right]_e \approx \frac{1}{|\nabla u_e|} \frac{u_E - u_O}{h}, \quad (8)$$

$$\nabla u_e = \left(\left[\frac{\partial u}{\partial x} \right]_e, \left[\frac{\partial u}{\partial y} \right]_e \right) = \left(\frac{u_E - u_O}{h}, \frac{u_S + u_{S_E} - u_N - u_{N_E}}{4h} \right), \quad (9)$$

$$|\nabla u_e| =$$

$$\frac{1}{h} \sqrt{(u_E - u_O)^2 + [(u_N + u_{N_E} - u_S - u_{S_E})/4]^2}, \quad (10)$$

式中 $u_E, u_O, u_N, u_{N_E}, u_S, u_{S_E}$ 分别为像素点 $E, O,$

N, N_E, S 和 S_E 对应的像素值。其中 N_E 和 S_E 为待处理像素点 O 的对角邻域像素。

同理可计算得到其余的半像素点梯度值 $v_w^{(1)}, v_s^{(2)}, v_n^{(2)}$, 将其代入(6)式后可得

$$\sum_{P \in \Lambda_O} \frac{1}{|\nabla u_p|} (u_O - u_P) +$$

$$\lambda_f(O)(u_O - u_O^{(0)}) = 0, P \in \Lambda_O, p \in \Lambda, (11)$$

式中: Λ_O, Λ 分别为待处理图片中像素点集和半像素点集; u_O 表示当前待处理的像素值; u_P 为待处理像素邻域内像素点的值; $u_O^{(0)}$ 表示上一次迭代得到的结果。定义 w_p 为扩散系数, h_{Op} 和 h_{OO} 为权系数, 即

$$w_p = \frac{1}{|\nabla u_p|}, h_{Op} = \frac{w_p}{\sum_{p \in \Lambda} w_p + \lambda_f(O)},$$

$$h_{OO} = \frac{\lambda_f(O)}{\sum_{p \in \Lambda} w_p + \lambda_f(O)}, (12)$$

将这三个系数代入(11)式可得到

$$u_O^{(n)} = \sum_{P \in \Lambda_O, p \in \Lambda} h_{Op}^{(n-1)} u_P^{(n-1)} + h_{OO}^{(n-1)} u_O^{(n-1)}, (13)$$

式中: 上标 (n) 表示迭代次数; $u_O^{(n)}$ 表示第 n 次迭代后的像素值; $h_{Op}^{(n-1)}$ 和 $h_{OO}^{(n-1)}$ 表示上一次迭代更新后的权系数; $u_P^{(n-1)}$ 和 $u_O^{(n-1)}$ 表示上一次迭代后的像素值。通过(13)式可求出整个图片待处理的像素值, 从而完成对文字干扰的消除。

4 算法步骤

1) 输入存在文字干扰的图片, 对干扰文字进行检测与定位, 确定待处理区域;

2) 根据图片中文字区域的大小合理设置迭代次数;

3) 用 O 表示待处理域中的一个未知像素点, 根据(10)式求解点 O 半像素邻域点 n, s, w, e 的梯度模值 $|\nabla u_p|$;

4) 将步骤 3) 中所求得的梯度模值 $|\nabla u_p|$ 代入(12)式求出 w_p, h_{Op} 以及 h_{OO} ;

5) 将 w_p, h_{Op} 和 h_{OO} 代入(13)式求得待处理像素点 O 的像素值, 重复上述步骤对待处理域中其余像素的像素值进行求解;

6) 待处理域中所有未知像素的像素值求解完成后, 将迭代次数减 1;

7) 重复步骤 3)~6) 进行多次迭代, 直到迭代次数减为 0;

8) 计算并输出处理过程所用时间, 输出文字干

扰消除后的图片。

5 仿真结果及分析

下面利用计算机 (I7-3.2 GHz 处理器、8 GB 内存、240 GB SSD 硬盘、128 位 2 GB 显卡、64 位 WIN10 操作系统), 在 VS2015 平台上对图 3 所示的大小为 435 pixel \times 333 pixel 的人物图片中的文字干扰进行消除。



图 3 受文字干扰的图片

Fig. 3 Image with text interference

5.1 干扰文字的检测与定位

由图 3 可知, 干扰文字所在区域边缘信息相对密集。为降低文字与背景像素值相近时的误检率, 本文先采用 Roberts 算子对图 3 进行边缘检测, 然后进行二值化处理, 以排除边缘检测结果中部分干扰因素, 仿真结果如图 4 所示。



图 4 干扰文字的检测。(a)对图 3 边缘检测的结果;

(b)对图 4(a)二值化的结果

Fig. 4 Detection for text interference. (a) Detection result of edge in Fig. 3; (b) result of binarization of Fig. 4(a)

为获取干扰文字所在区域, 对图 4(b)依次进行闭运算、开运算、三次水平膨胀和一次水平腐蚀处理, 得到如图 5(a)所示的连通域图, 设定阈值将包含像素较少的连通域去除, 得到如图 5(b)所示的结果。

对图 5(b)所示的连通域求外接矩形, 记录矩形区域 4 个顶点的坐标。读取待修复图像对应的灰度图, 设定阈值, 将灰度图中的前景和背景分离, 处理结果如图 6 所示。

将上述所得矩形区域的 4 个顶点坐标作为(13)式的限定条件, 将图 1 与图 6 中矩形限定区域内的

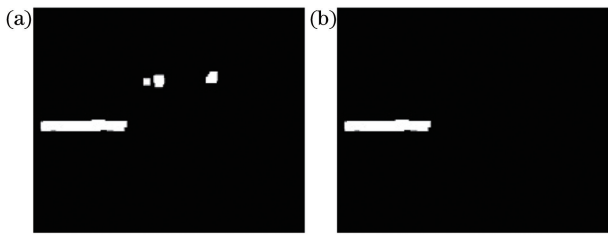


图 5 干扰文字的定位。(a)形态学处理后的连通域图；(b)去除像素少的连通域后的结果

Fig. 5 Location of interference text. (a) Connected domain processed by morphology; (b) result after removing connected domain with few pixels

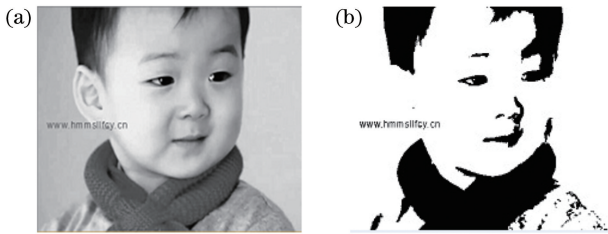


图 6 形态学处理去除小连通域后的图片。

(a)待修复图像对应的灰度图；(b)二值化后的灰度图

Fig. 6 Images obtained by morphological processing and removing small connected domain. (a) Gray scale corresponding to image to be repaired; (b) gray scale after binarization

黑色像素点处于相同位置的像素标记为白色,完成对干扰文字的定位,如图 7 所示。

5.2 干扰文字的消除

根据第 4 部分的算法步骤对图 7 中所标记的文字区域进行消除,结果如图 8 所示。从图 8(a)可以看出迭代 5 次对干扰文字无能为力,但迭代次数增加至 60 次时,该方法可以很彻底地消除干扰文字,只是在文字颜色与图中人物肤色接近处稍有失真。

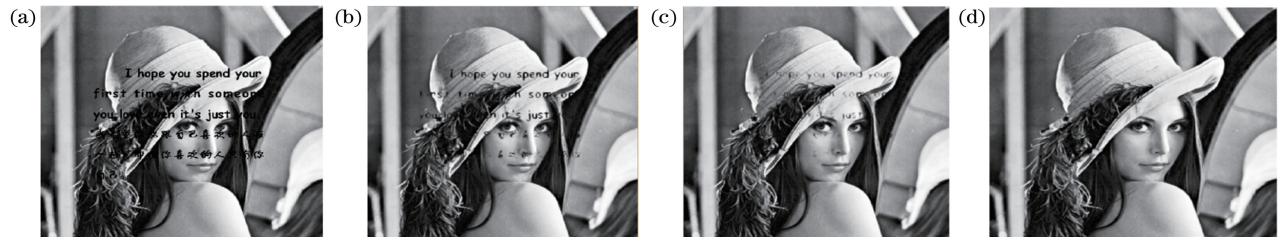


图 9 采用 TV 模型迭代不同次数消除文字干扰的结果。(a)待处理图片；(b)迭代 50 次；(c)迭代 100 次；(d)迭代 300 次
Fig. 9 Results after eliminating interference text in image by using TV model with different iterations. (a) Image to be processed; (b) 50 iterations; (c) 100 iterations; (d) 300 iterations

6 结 论

结合 TV 模型,通过 Roberts 算法和二值形态学对图片中的干扰文字区域进行检测与定位,对干

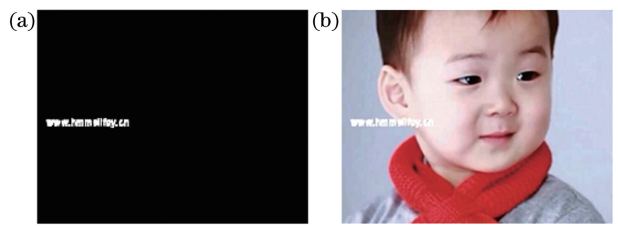


图 7 定位干扰文字。(a)掩模图；(b)标记文字区域的待处理图片

Fig. 7 Location of interference text. (a) Mask pattern; (b) image with text to be processed

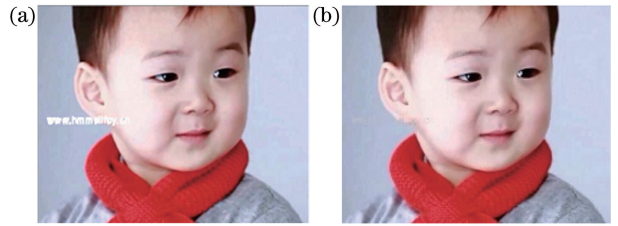


图 8 采用 TV 模型消除图片中干扰文字后的结果。

(a)迭代 5 次；(b)迭代 60 次

Fig. 8 Result after eliminating interference text in image by using TV model. (a) 5 iterations; (b) 60 iterations

为进一步说明本文所提方法能有效消除文字,依照上述过程对大小为 554 pixel×554 pixel、含有较小且文字较多的 Lena 图片进行了迭代次数分别为 50、100 和 300 的处理,结果如图 9 所示。从图 9 可以看出:随着迭代次数不断增加,图片中文字的干扰逐渐减弱,当迭代达到 300 次时,消除文字干扰对图中细节没有任何影响,如图 9(d)所示。获得图 9(b)~(d)效果的耗时分别为 0.488345, 0.954017, 2.744955 s。可见,随着迭代次数增加,消除图片中文字干扰的时间延长,故需要合理选择迭代次数,在保证消除文字干扰效果的前提下尽量缩短算法的运行时间。

扰文字区域和不含文字区域进行区分,避免了消除文字干扰后图片中色彩的失真或信息丢失。对受不同程度文字干扰的图片进行处理后可以看出,采用本文方法对图片中文字干扰的消除效果显著,但对

较密集和较多文字的图片处理难度明显有所增大, TV模型的迭代次数和运行时间显著增多。可见,若要在改善图片中文字干扰效果的同时缩短运行时间,就得进一步加快TV模型算法的收敛速度。

参 考 文 献

- [1] Ji L Q, Wang J J. Automatic text detection and removal in video images [J]. *Journal of Image and Graphics*, 2008, 13(3): 461-466.
季丽琴, 王加俊. 视频字幕的自动检测与去除 [J]. *中国图象图形学报*, 2008, 13(3): 461-466.
- [2] Telea A. An image inpainting technique based on the fast marching method [J]. *Journal of Graphics Tools*, 2004, 9(1): 23-34.
- [3] Zhang X Q, Gao Z H. Image inpainting algorithm based on block matching [J]. *Journal of Optoelectronics · Laser*, 2012, 23(4): 805-811.
张显全, 高志卉. 一种块匹配的图像修复算法 [J]. *光电子·激光*, 2012, 23(4): 805-811.
- [4] Wang H J, Wang J L, Wang M H, *et al.* Efficient image inpainting based on bilinear interpolation downscaling [J]. *Optics and Precision Engineering*, 2010, 18(5): 1234-1241.
王昊京, 王建立, 王鸣浩, 等. 采用双线性插值收缩的图像修复方法 [J]. *光学精密工程*, 2010, 18(5): 1234-1241.
- [5] Cai Z C, Yao F F, Tang Z S. Digital image inpainting with Kriging method [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2013, 25(9): 1281-1287.
蔡占川, 姚菲菲, 唐泽圣. 基于克里金插值法的图像修复 [J]. *计算机辅助设计与图形学学报*, 2013, 25(9): 1281-1287.
- [6] Alilou V K, Yaghmaee F. Non-texture image inpainting using histogram of oriented gradients [J]. *Journal of Visual Communication and Image Representation*, 2017, 48: 43-53.
- [7] Wang W W, Li L, Han Y. Image inpainting based on nonlocal curvature-driven diffusion [J]. *Acta Optica Sinica*, 2010, 30(6): 1634-1638.
王卫卫, 李莉, 韩雨. 基于非局部曲率驱动扩散的图像修复 [J]. *光学学报*, 2010, 30(6): 1634-1638.
- [8] Li F, Shen C M, Liu R H, *et al.* A fast implementation algorithm of TV inpainting model based on operator splitting method [J]. *Computers & Electrical Engineering*, 2011, 37(5): 782-788.
- [9] Fuchs M, Müller J. A higher order TV-type variation problem related to the denoising and inpainting of images [J]. *Nonlinear Analysis: Theory Methods & Applications*, 2016, 154: 122-147.
- [10] Du S S, Han C. An improved image inpainting algorithm based on total variation model [J]. *Laser & Optoelectronics Progress*, 2018, 55(7): 071005.
杜闪闪, 韩超. 基于总变分模型的改进图像修复算法 [J]. *激光与光电子学进展*, 2018, 55(7): 071005.
- [11] Wang B, Hu L L, Xue R Y. Image restoration based on degradation conversion and separable total variation model [J]. *Acta Optica Sinica*, 2014, 34(10): 1011002.
王斌, 胡辽林, 薛瑞洋. 基于退化转换及可分离全变差模型的图像修复 [J]. *光学学报*, 2014, 34(10): 1011002.
- [12] Bourdin L, Cresson J, Greff I, *et al.* Variational integrator for fractional Euler-Lagrange equations [J]. *Applied Numerical Mathematics*, 2013, 71: 14-23.