

复杂背景下文本提取的彩色边缘检测算子设计*

张引, 潘云鹤

(浙江大学 人工智能研究所, 浙江 杭州 310027);

(浙江大学 CAD & CG 国家重点实验室, 浙江 杭州 310027)

E-mail: yinzh-zju@263.net

http://www.zju.edu.cn

摘要: 分析了当前彩色边缘检测方法的现状与复杂背景下文本提取的社会需求, 提出一个新的彩色图像边缘检测算子——ColorRoberts, 该算子全面作用于三维 RGB 彩色空间. 实验证明, 在文本区域提取的彩色图像边缘检测中, ColorRoberts 具有传统算子所无法比拟的优势, 检测出的文本区域完整、清晰, 字符形状边缘保持较好, 文本区域与背景易于进一步剥离.

关键词: 彩色图像分割; 边缘检测; 文本提取; 基于内容的图像检索

中图法分类号: TP391 **文献标识码:** A

20 多年来, 灰度图像分割已发展为一个非常成熟的图像处理领域^[1], 涌现出大量的算法; 而彩色图像处理的工作由于过去的硬件条件所限, 做得很少, 直到最近几年才逐渐得到发展. 人类视觉系统对色彩非常敏感, 人眼能分辨的灰度只有 20 多级, 而能区分的颜色有 35 000 多种^[2], 目标与背景的区别用色差比用亮度差更适合于人眼识别. 彩色图像的信息全部包含在彩色模型的各个分量之中, 灰度图像只包含明暗信息, 因此, 彩色图像不仅在视觉感受上比灰度图像更舒服, 而且重要的是能够提供更多的视觉信息, 更有利于图像分割或目标提取.

当前, 作为彩色图像分割的方法 5 大类^[3]之一的边缘检测法, 与其他几类一样, 主要实现途径是相应灰度图像分割方法的简单扩展. 不足的是, 这种扩展往往只是简单地对彩色分量独立地应用灰度算法, 然后合并分量处理结果, 而忽视了不同彩色分量之间存在的、应当在分割期间保持的相关性. 如文献[4~6]中算法的主要缺陷在于, 它们均基于每个彩色分量的灰度分布直方图, 对于复杂图像, 每个彩色分量的直方图可能是多峰的, 使得难于正确分割彩色图像. 为了避免由第 1 种途径所产生的缺陷, 另一种最简单的途径是仅处理彩色图像的程度方面, 即将彩色图像转换为灰度图像, 然后执行灰度分割. 这种方式虽然有效果, 但却忽略了彩色信息的作用, 并且由于亮度与色度等反映的图像边缘并非一直重合, 往往会出现多边缘和宽边缘的现象.

彩色图像和视频中复杂背景下文本的提取, 由于其广泛的应用前景, 近年来越来越受到人们的重视^[7~10], 特别是复杂文档的自动处理、数字图书馆的建设、基于内容的图像检索和 Web 多媒体搜索引擎的发展, 对其提出了更迫切的需求. 本文针对这一需求和现有彩色图像分割方法的局限性, 提出了一个较好的彩色图像边缘检测算子——ColorRoberts.

* 收稿日期: 1999-09-14; 修改日期: 2000-04-04

基金项目: 国家自然科学基金资助项目(69603009)

作者简介: 张引(1970—), 女, 甘肃永登人, 博士, 讲师, 主要研究领域为图形/图像、智能 CAD; 潘云鹤(1946—), 男, 浙江杭州人, 教授, 博士生导师, 中国工程院院士, 主要研究领域为智能 CAD, 人工智能, CAD/CG, 形象思维.

1 ColorRoberts 算子的设计

彩色图像边缘检测与一般灰度图像边缘检测具有相同的要求:(1)高精度的图像边缘定位;(2)噪声的抑制能力.但这两点往往是矛盾的.

在复杂背景下的文本提取中,文本区域(如视频中的片头和字幕、图书封面图像中的书名、广告或商标图像中的产品信息、票据上的数据、信封上的邮编、汽车图像中的牌照等等)由于字符的存在,其空间频率变化较大.一个好的彩色边缘检测算子应能满足下面的设计准则:

- (1) 文本区域清晰、完整,与周围背景有较强的对比;
- (2) 基本保持字符的形状和边缘;
- (3) 背景区域淡化,以利于进一步的文本区域提取.

根据上述准则,我们设计了一个新的彩色图像边缘检测算子——ColorRoberts.其基本思想是:首先推广灰度图像边缘检测的 Roberts 算子到彩色图像边缘检测中,得到一幅边缘灰度图像,在此,我们没有将灰度 Roberts 算子简单用于彩色图像像素的各彩色分量上,或将彩色图像转换为灰度图像,而是采用欧氏距离将各彩色分量综合考虑,这样既保持了彩色图像中信息的完整性,又考虑到了彩色分量间的关联性,即全面作用于三维彩色空间;其次,通过采用灰度图像中的一类最佳边缘检测算子 LOG 算子,对提取出的彩色图像边缘灰度图像作二次边缘提取,进而根据二次边缘提取的结果确定最终二值边缘提取采用的阈值,从而完成彩色图像的边缘提取.

1.1 基于三维彩色空间和基于单维空间合成的彩色边缘检测方法

对灰度图像而言,任一灰度边缘检测算子若要检测一个像素点是否为边缘点,其结果与被检测的像素点相关邻域内的像素点的灰度差有密切关系:若灰度差大,则对检测的结果影响也大.对于彩色图像来说,任何边缘检测算子与像素点之间的色差有密切关系.传统的彩色图像边缘检测算法往往单纯地把经典的灰度边缘检测算子分别作用于彩色空间的三分量上分别提取边缘,并用适当的方式合并三分量的处理结果,如当有两个分量有边缘时即可判定彩色图像存在边缘(单维空间投票法),或当有一分量存在边缘时才继续另一分量的边缘检测,以证实前一分量的检测结果,均不能取得好的结果.彩色空间的3个分量只反映了色彩的一部分信息,色彩的全部信息是由三分量共同综合反映出来的.在 RGB 三维彩色空间中(如图1所示), D_1 和 D_2 为两个像素点,RGB三分量分别为 (R_1, G_1, B_1) 和 (R_2, G_2, B_2) .显然,图中像素点 D_1 和 D_2 存在色差 ΔD .

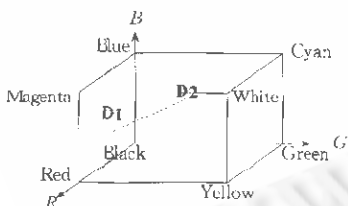


Fig.1 RGB ColorSpace
图1 RGB彩色空间

$$\Delta D = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2}$$

但是,若采用单维空间投票法,则最终边缘检测结果与像素点 D_1 和 D_2 分别在 R, G, B 这3个一维空间中的色差有关,更确切地说,与 R, G, B 空间中各自的强度差有关, D_1 和 D_2 相应三分量的色差分别为

$$R \text{ 分量: } \Delta R = |R_1 - R_2|,$$

$$G \text{ 分量: } \Delta G = |G_1 - G_2|,$$

$$B \text{ 分量: } \Delta B = |B_1 - B_2|,$$

即分别为 D_1 和 D_2 在RGB三维空间中向 R, G, B 三维投影得到的差值.图1中仅在 R 分量上存在差异.因此,从三分量的分别作用投票表决只能得出 D_1 和 D_2 无色差的结论,这显然是错误的.

无论单维法如何组合三分量的色差,都不能准确地体现 RGB 三维彩色空间的色差 ΔD . 本文提出的 ColorRoberts 算子充分体现了全面作用于彩色三分量的综合方法,充分利用了全面的色彩信息,所以进行彩色图像边缘检测的效果优于只分别作用在彩色三分量上的传统彩色边缘检测方法,实验结果见第 2 节.

1.2 基于三维彩色空间 ColorRoberts 算子描述

ColorRoberts 算子描述如下:

步骤 1. 输入彩色图像 I , 其中像素点 (i, j) 的 R, G, B 彩色分量分别为 $I(i, j, k), k=1, 2, 3$.

步骤 2. 定义彩色梯度算子如下:

对于像素点 (i, j) ,

$$\text{扩展 Roberts}(i, j) = \sqrt{\Delta_1^2 + \Delta_2^2},$$

$$\Delta_1 = \text{Eud}(i+n, j+n; i, j),$$

$$\Delta_2 = \text{Eud}(i, j+n; i+n, j),$$

其中 $\text{Eud}(i_1, j_1; i_2, j_2)$ 定义为像素点 (i_1, j_1) 与 (i_2, j_2) 之间的彩色值欧氏距离,

$$\text{Eud}(i_1, j_1; i_2, j_2) =$$

$$\sqrt{(I(i_1, j_1, 1) - I(i_2, j_2, 1))^2 + (I(i_1, j_1, 2) - I(i_2, j_2, 2))^2 + (I(i_1, j_1, 3) - I(i_2, j_2, 3))^2}, \quad (1)$$

I 经扩展 Roberts 处理后得到一灰度边缘图像 I_1 .

步骤 3. 对 I_1 采用 LOG 算子作二次边缘提取, 得到 I_1 的二值边缘图像 I_2 .

步骤 4. 计算 I_2 图像密度 $\text{ImageDensity}(I_2) = \frac{\#\{I_2(i, j)=1, (i, j) \in I_1\}}{\#(I_1)}$, 其中 $\#(\cdot)$ 为像素总数求和算子.

步骤 5. 计算灰度边缘图像 I_1 的灰度直方图 $\text{Hist}(i), i=0, \dots, 255$.

步骤 6. 在灰度直方图上, 根据 $\text{ImageDensity}(I_2)$ 确定灰度边缘图像 I_1 的二值化图像阈值 TH, 具体做法如下:

For $i=0$ to 255 do

{

$$\text{density} = \sum_{k=i+1}^{225} \text{Hist}(k)$$

if $\text{density} \leq \lambda \cdot \text{ImageDensity}(I_2)$ then

{

TH = $i+1$;

Break;

}

}

其中 $1.2 < \lambda < 1.5$, λ 为密度因子.

步骤 7. 根据 TH 对灰度边缘图像 I 作二值化处理, 得到二值边缘图像 I_e .

步骤 8. I_e 即为彩色图像 I 的二值边缘提取图像.

算子执行过程如图 2 所示, 其中图 2(a) 为彩色图像原图, 图 2(b) 为算子提取的边缘灰度图像, 图 2(c) 为采用 LOG 算子对图 2(b) 作二次边缘提取后的二值边缘图像, 图 2(d) 是图 2(b) 所示的边缘灰度图像的直方图, 图 2(e) 为最终得到的二值边缘提取图像. 由图 2(e) 可见, 提取出的边缘图像

中文本区域与背景易于分割,且文本区域完整、清晰,字符形状保持较好,基本符合最初提出的彩色图像边缘检测算子设计准则,满足后续工作提取文本区域的要求。



(a) Video frame

(a) 视频帧



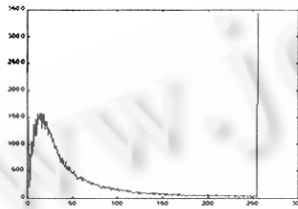
(b) Gray edge extractor of ColorRoberts

(b) ColorRoberts 灰度边缘提取



(c) The second edge extraction of ColorRoberts

(c) ColorRoberts 二次边缘提取



(d) Histogram of Fig. 2(b)

(d) 图 2(b) 的直方图



(e) Binary edge extraction of ColorRoberts

(e) ColorRoberts 二值边缘提取

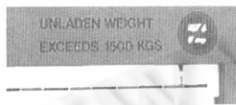
Fig. 2

图 2

2 实验分析与比较

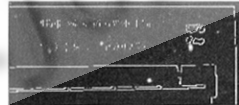
2.1 ColorRoberts 与基于单维空间的投票法比较

我们分别将 5 种著名的灰度图像边缘检测算子——Roberts 算子、Prewitt 算子、Sobel 算子和两种最佳算子 LOG 算子与 Canny 算子作用于彩色值各分量上,采用基于单维空间的投票法进行彩色图像边缘检测比较实验。由实验结果(如图 3 所示)可以看出,包括两种最佳算子在内的这 5 种灰度图像的著名算子在检测彩色图像的字符边缘时缺损很大,甚至有的字符区域未检测出来,而本文提出的 ColorRoberts 算子检测出的字符区域清晰而完整。



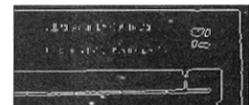
(a) Road sign color image

(a) 路标彩色图像



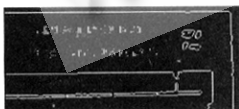
(b) Gray Roberts

(b) 灰度 Roberts



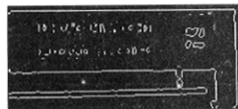
(c) Gray Prewitt

(c) 灰度 Prewitt



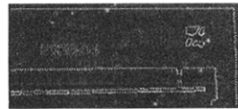
(d) Gray Sobel

(d) 灰度 Sobel



(e) Gray LOG

(e) 灰度 LOG



(f) Gray Canny

(f) 灰度 Canny



(g) ColorRoberts

(g) ColorRoberts 算子

Fig. 3

图 3

造成上述情况的原因在于:人眼从背景中辨认字符是由于字符与背景的颜色差距形成了较强

的对比. 字符笔划边缘与背景的对比构成了其强烈的空间频率变化. 颜色矢量由 3 个分量组成. 单纯地对各个分量进行边缘检测, 不能充分体现出颜色的差距, 因而不能准确地检测出图像边缘, 如图 4 所示. 图 4 为图 3(a)彩色三分量子图像, 显然, 这 3 幅图在视觉上的对比度不如图 3(a)的彩色对比. 这种情况也正好符合人眼的视觉特性, 即人眼对颜色的分辨率远远高于灰度.

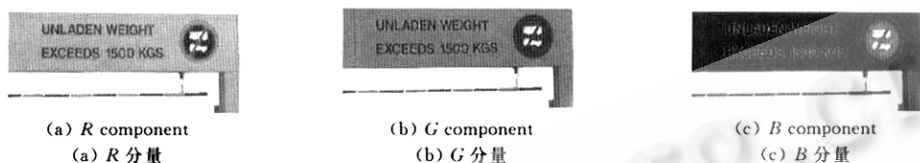


Fig. 4
图 4

ColorRoberts 算子由于对彩色值三分量同时作用, 避免了上述情况的发生, 可以较好地检测出文本区域边缘. 理论分析与实验结果完全吻合.

2.2 ColorRoberts 与加权 ColorRoberts 比较

在标准 CIE 匀色空间中, 色差是用两个颜色的距离来表示的. 但是, 显示器采用的 RGB 空间是显示器的设备空间, 与 CIE 系统的真实三原色不同, 两者之间的转换涉及到诸多复杂的问题. 我们在 ColorRoberts 算子中仍采用了 RGB 彩色空间的距离来表示色差. 但是, RGB 空间在视觉上是不均匀的, 因此, 在计算颜色的色差时, 应采用加权 RGB 空间中的距离来表示, 即加权:

$$Eud(i_1, j_1; i_2, j_2) =$$

$$\sqrt{W_1(I(i_1, j_1, 1) - I(i_2, j_2, 1))^2 + W_2(I(i_1, j_1, 2) - I(i_2, j_2, 2))^2 + W_3(I(i_1, j_1, 3) - I(i_2, j_2, 3))^2}. \quad (2)$$

柏子游等人^[11]根据人眼对 R, G, B 三色分量的分辨程度, 经验选取 $W_1=4, W_2=8, W_3=1$. 我们也采用此值, 通过将式(1)替换为式(2), 使 ColorRoberts 算子成为加权 ColorRoberts 算子. 通过对比实验我们发现(如图 5 所示), 虽然 I 经加权扩展 Roberts 处理后得到的灰度边缘图像 I_1 , 确实比未加权的扩展 Roberts 处理后的结果图像的细节更加清晰、丰富, 但是, 由于我们设计

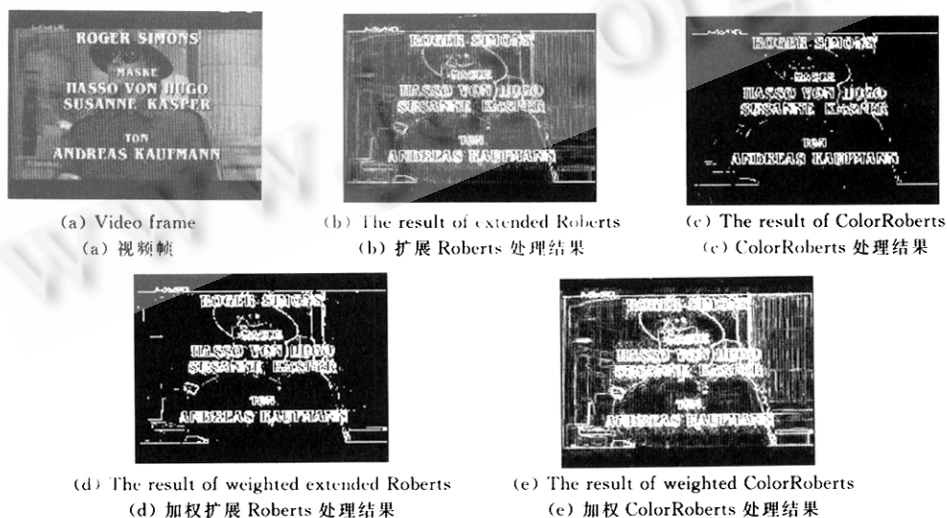


Fig. 5
图 5

ColorRoberts 算子的目的是为了提取复杂背景下的文本,背景中本身包含有丰富的非文本对象,所以对于背景图像不太复杂的情况,加权 ColorRoberts 算子和不加权 ColorRoberts 算子均可用于文本提取.而对于背景图像复杂的情况,加权 ColorRoberts 算子的最终结果是同时突出了背景中细节丰富的区域,在二值化边缘图像 I_e 中引入了更多的不利于文本提取的背景噪声.因此,我们采用不加权的 ColorRoberts 算子.

3 结 论

本文主要讨论了复杂背景下文本提取中的彩色图像边缘检测的有关问题,提出了彩色图像边缘检测算子设计的 3 个准则,并提出了一个彩色图像边缘检测算子 ColorRoberts.该算子直接作用于 RGB 颜色空间上,算法简单,适用性强,计算复杂度与原来的灰度边缘检测算子相似.虽然 RGB 彩色空间的色彩可达 $224 = 16777216$ 种,常见的彩色图像处理总是要先进行彩色量化、降维,而 ColorRoberts 算子避免了在如此巨大的空间上的运算以及繁琐的降维工作.

传统的彩色图像边缘检测算子大多是将灰度图像边缘检测算子分别作用于颜色空间的各个分量上,不能充分体现颜色的距离.与之相比,ColorRoberts 算子全面作用在颜色的 3 个分量上,符合人眼视觉特性,在文本区域提取的彩色图像边缘检测中效果优于传统算子,检测出的文本区域完整、清晰,字符形状边缘保持较好,文本区域与背景易于进一步剥离.实验表明,该算子适用于彩色视频中片头字幕提取、广告信息提取和汽车牌照定位等多种应用领域,图 6~8 为部分实验结果.



Fig. 6 Video frame and its binary edge image

图 6 视频帧及边缘二值图

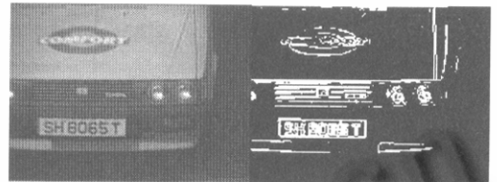


Fig. 7 Color car image containing license plate and its binary edge image

图 7 包含牌照的彩色汽车图像及边缘二值图



Fig. 8 Color ad. image of LINCOLN car and its binary edge image

图 8 林肯车的彩色广告图像及边缘二值图

References:

- [1] Weeks, A. R., Felix, C. E., Myler, H. R. Edge detection of color images using the HSL color space. In: Dougherty, E. R., ed. Proceedings of the SPIE 2424—Nonlinear Image Processing VI. San Jose, CA: SPIE, 1995. 291~297.
- [2] Meng, Zhang-rong. The requirement analysis on selections among various color models. Journal of Image and Graphics, 1996,1(3):238~241 (in Chinese).
- [3] Gauch, J., Hsia, C-W. A comparison of three color image segmentation algorithms in four color spaces. In: Maragos, P., ed. Proceedings of the SPIE 1818—The International Society for Optical Engineering Visual Communications and Image Processing (Part 3). Boston, MA: SPIE, 1992. 1168~1179.
- [4] Liu, J., Yang, Y. H. Multiresolution color image segmentation. IEEE Transactions on Pattern Analysis and Machine

- Intelligence, 1994, 16(2):689~700.
- [5] Celenk, M. A color clustering technique for image segmentation. *Computer Vision, Graphics and Image Processing*, 1990, 52(3):145~170.
- [6] Huang, C-L., Cheng, T Y., Chen, C-C. Color images' segmentation using scale space filters and markov random fields. *Pattern Recognition*, 1992, 256(10):1217~1229.
- [7] Tang, Y. Y., Lee, S-W., Suen, C. Y. Automatic document processing; a survey. *Pattern Recognition*, 1996, 29(12):1931~1952.
- [8] Zhuang, Yue-ting. Intelligent multimedia information analysis and retrieval with applications to visual design [Ph. D. Thesis]. Hangzhou: Zhejiang University, 1998 (in Chinese).
- [9] Jain, A. K., Yu, B. Automatic text location in images and video frames. *Pattern Recognition*, 1998, 31(12):2055~2076.
- [10] Zhong, Y., Jain, A. K. Locating text in complex color images. *Pattern Recognition*, 1995, 28(10):1523~1535.
- [11] Bai, Zi-you, Zhang Yong, Yu, Lie. A color segmentation method for truecolor ferrographic image. *Journal of Pattern Recognition and Artificial Intelligence*, 1999, 12(2):241~244 (in Chinese).

附中文参考文献:

- [2] 孟章荣. 各种颜色模型选用需求分析. *中国图像图形学报*, 1996, 1(3):238~241.
- [8] 庄越挺. 智能多媒体信息分析与检索的研究[博士学位论文]. 杭州:浙江大学, 1998.
- [11] 柏子游, 张勇, 虞烈. 一种彩色图像的彩色分割方法. *模式识别与人工智能*, 1999, 12(2):241~244.

Design of a New Color Edge Detector for Text Extraction Under Complex Background*

ZHANG Yin, PAN Yun-he

(Institute of Artificial Intelligence, Zhejiang University, Hangzhou 310027, China);

(State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310027, China)

E-mail: yinzhang@zju.edu.cn

http://www.zju.edu.cn

Abstract: In this paper, a new color edge detector ColorRoberts is presented after analyzing the status of current color edge detection methods and the social requirement to text extraction under complex background. ColorRoberts operates on the whole color space RGB. For text extraction under complex background, it shows that ColorRoberts is better than those classical edge detectors. The detected text edges are clear, integrated and easy to separated from the background.

Key words: color image segmentation; edge detection; text extraction; content-based image retrieval

* Received September 14, 1999; accepted April 4, 2000

Supported by the National Natural Science Foundation of China under Grant No. 69803009