

基于改进的 snake 模型的嘴唇轮廓提取

余健仪, 郑胜林, 潘保昌, 邱海枫

(广东工业大学 信息学院 数字图像技术研究所, 广东 广州 510006)

摘要:为了更好地利用 snake 模型来提取彩色图像中的物体轮廓, 提出一种改进的 snake 算法。此方法首先自动生成 snake 的初始模型, 然后在 GVF - snake 的基础上重新设计了 snake 的外部能量函数, 采用色彩聚类算法对原始图像进行分割, 利用像素到聚类中心的距离增强图像并进行差分运算, 提取有意义区域的边缘梯度, 对 GVF 向量场进行了归一化处理并改进了平滑因子。实验结果证明, 改进后的算法, 特别是在处理彩色图像时, 大大优于原始方法, 提高了轮廓提取的精度且有较好的鲁棒性。

关键词:唇读识别; 轮廓提取; snake 模型; GVF; 能量函数

中图分类号: TP391

文献标识码: A

文章编号: 1673 - 629X(2008)08 - 0095 - 03

Lip Contour Extracting Based on Improved Snake Model

YU Jian-yi, ZHENG Sheng-lin, PAN Bao-chang, QIU Hai-feng

(Image Processing Lab., Information Sch., Guangdong University of Technology, Guangzhou 510006, China)

Abstract: In order to improve the effectiveness of snake active model to extract the object contour in color images, introduce an improved snake algorithm. First, develop a new approach to generate an initial contour automatically; then a new external force is redesigned basing on the GVF - snake. Use the color - based clustering algorithm to divide the image. Then use the distance between the pixels and the clustering center to enhance the image and use the difference to get the edge grads of interest. In addition, normalize the GVF field and improve the smooth factor. Comparing the snake model prototype with the new method, it indicates that the improvement is considerable, especially when color images are involved.

Key words: lip reading; contour extracting; snake model; GVF; energy function

0 引言

唇读, 主要是听力残障者从正常人说话中获取信息的一种技巧。他们凭借自身已有的语言知识、谈话内容和背景知识等众多相关因素, 再根据观察到说话人的口型变化情况, 映像成说话内容。唇读识别的研究具有重大的现实意义和理论研究价值, 它是一个有着广阔发展前景的崭新研究课题。

嘴唇特征的提取是唇读识别的核心技术。目前, 研究人员在这方面也做了不少研究, 主要提出了灰度积分法、弹性模板匹配法、snake 模型算法等。snake 方法最早是由 Kass^[1] 等人在 1987 年提出的。近年来 snake 模型在图像分析及计算机视觉等领域的应用非常广泛, 包括边缘检测、形状建模, 分割和运动跟

踪^[2,3]。

然而, 传统的 snake 模型算法存在三大缺陷^[4]: 一是初始模型的位置需要人工确定。二是 snake 模型传统上是以灰度图作为主要研究对象, 当这些方法应用于彩色图像时, 因针对灰度图而设计的外部能量函数中的固有缺陷, 致使在某些情况下会造成严重的错误。三是经典的外力场中, 在边界附近, 高斯外力迅速减小; 高斯外力抗干扰性差, 微小的噪声就会改变外力方向, 使活动轮廓的运动偏离正确的运动方向。

针对传统 snake 的缺陷, 在嘴唇轮廓提取的应用中, 提出了一种改进的 snake 模型算法。首先自动生成 snake 的初始模型; 然后在 GVF - snake 的基础上重新设计了 snake 的外部能量函数, 采用色彩聚类算法对原始图像进行分割后进行差分提取有意义区域的边缘梯度, 代替传统方法中常用的像素灰度的差分来近似图像梯度, 对 GVF 向量场进行了归一化处理并改进了平滑因子。

此改进算法一定程度上解决了传统算法的缺陷, 提高了原算法的适应性。

收稿日期: 2007 - 11 - 04

基金项目: 国家自然科学基金项目(60572078)

作者简介: 余健仪(1983 -), 女, 广东佛山人, 硕士研究生, 研究方向为图像处理、模式识别; 郑胜林, 副教授, 研究方向为图像处理、模式识别; 潘保昌, 教授, 博士生导师, 研究方向为图像处理、模式识别。

1 改进 snake 模型并提取嘴唇轮廓

1.1 初始轮廓的自动生成

目前,snake 初始轮廓一般采用手工勾勒,该方法工作繁琐且不易于实现自动化,提出一种简单有效的方法来确定 snake 模型的初始轮廓:

(1) 判断嘴唇的状态:首先对预处理后的二值图像做水平投影,通过判断投影曲线的主要波峰的个数确定嘴唇的状态,即一个波峰表示嘴唇闭合,两个波峰表示嘴唇张开。闭合时只需拟合外唇,张开时需要同时拟合内外唇。图 1(a)是嘴唇闭合时的水平投影,图 1(b)是嘴唇张开时的水平投影。

(2) 外唇模型:在水平和垂直投影曲线上确定嘴唇的上下左右四条边界线,得到一矩形目标区域并以此矩形的内切椭圆为 snake 模型的外唇初始轮廓,通常情况下由于嘴角处的边界不明显导致 snake 在嘴角的拟合效果不好,所以文中将预先提取左右外嘴角点,并让初始轮廓经过这两点。

(3) 内唇模型:通过投影对内唇的区域作一个估计。同样要先提取左右内嘴角点,定义一椭圆位于内唇区域内并经过这两点,作为内唇初始轮廓。

图 2 是使用上述方法自动生产的初始轮廓线。

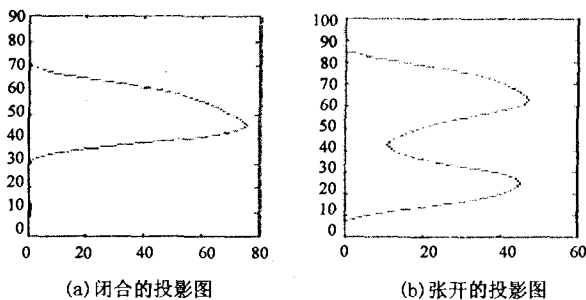


图 1 嘴唇投影图



图 2 内外唇初始轮廓

1.2 改进的 GVF - snake 模型

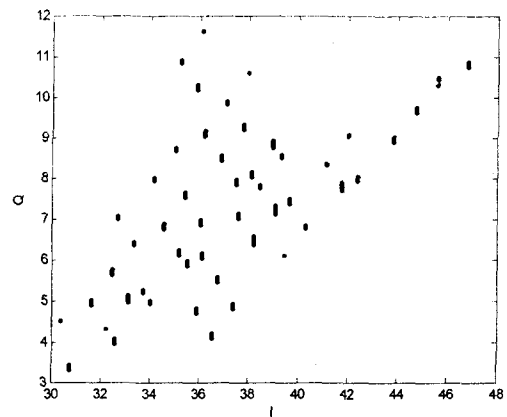
得到初始轮廓以后,将应用梯度矢量流 (gradient vector flow, GVF) 代替经典外力场,确定目标轮廓。GVF - snake 模型具有较大的捕捉范围^[5]。在 GVF 理论中,定义了灰度图像 $I(x, y)$ 的一个边缘图 $f(x, y) = -F_{\text{esl}}(x, y)$ 和一个 GVF 向量场 $V_{\text{GVF}}(u, v) = (u(x, y), v(x, y))$, 这个向量场可以最小化能量函数^[5]:

$$E = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V_{\text{GVF}} - \nabla f|^2 dx dy \quad (1)$$

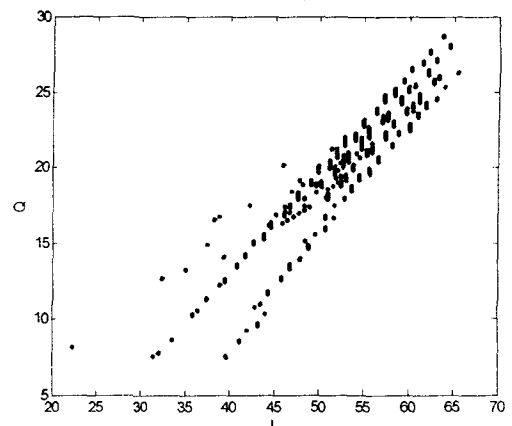
其中 u_x, u_y, v_x, v_y 是 u, v 分别对 x, y 的一阶偏导, f 是被处理图像 I 的边缘图, ∇f 是 f 的梯度场, μ 是规则化参数。

1.2.1 对边缘图 f 的改进

对于嘴唇图像,由于受胡须、阴影、光照等影响较大,而且往往内唇线比外唇线模糊,其 ∇f 较小,如果直接利用灰度图像的边缘信息提取口型轮廓,会使 snake 线收敛到非唇线边界。对于唇色和肤色,色调是一个重要特征。为得到仅表示色度的分量,需要把 RGB 进行坐标变换。因为 YIQ 表示法的冗余信息最少,因此采用 YIQ 色彩空间^[6]。图 3(a) 表示肤色的 I 值与 Q 值较小,图 3(b) 表示唇色的 I 值与 Q 值较大,可以看出, I 空间有部分重叠, Q 空间有较好的可分性。



(a) 肤色的 IQ 值分布



(b) 唇色的 IQ 值分布

图 3 IQ 值分布

不同于以往要对大量的人脸图像的肤色和唇色分布情况进行统计并建立模型的方法,取 Q 作为样本的特征值,然后利用 k 均值算法^[2]对色彩自动聚类,从嘴唇区域的像素值信息中过滤掉肤色部分。

对每一像素点根据其特征值 Q 到嘴唇聚类中心

Q_0 的距离确定一增强系数增强嘴唇图像,即

$$\begin{cases} \mu = l \cdot \frac{(Q - Q_0)^2}{\text{thresh}^2} & |Q - Q_0| \leq \text{thresh} \\ \mu = 1 & |Q - Q_0| > \text{thresh} \end{cases} \quad (2)$$

其中 μ 为增强系数, $\mu \in [0, 1]$, Q_0 为唇色区域的聚类中心, l 为可调系数, thresh 为唇色区域的半径, 增强后的边缘图表达式为 $f = \text{gradient}(255\mu)$ 。

1.2.2 归一化 V_{GVF} 的初始值

为了突出梯度幅值的作用,需要对 GVF 向量场归一化处理。图像上任意一点梯度值分布在 $[0, 255]$ 的区间上,在人的视觉上,梯度值 125 与梯度值 150 之间存在着显著的差别,但是在没有经过邻域归一化处理时,这种差别在数值上反映得并不明显。邻域归一化是在图像中以某一点为中心的邻域内做归一化处理。令 v 为该点的梯度值, v_{max} 为邻域内图像的最大梯度值, v_{min} 为邻域内图像的最小梯度值,则在该点处的图像能量值可以由公式 $(v - v_{\text{min}})/(v_{\text{max}} - v_{\text{min}})$ 求得。在求力场扩散方程之前进行归一化处理,不但可使各点的力场受附近点的幅值影响大大减小,而且受附近力场的方向影响也会减弱。

1.2.3 对平滑因子 μ 的改进

观察式(1)可以发现^[7]:

- (1) 当 ∇f 较大时,能量 E 主要是由数据项 $|V_{\text{GVF}} - \nabla f|^2$ 来控制;
- (2) 如果要求能量 E 最小,则应取 $V_{\text{GVF}} = \nabla f$;
- (3) 当 ∇f 较小时,能量 E 主要是由平滑项 $\mu(u_x^2 + u_y^2 + v_x^2 + v_y^2)$ 来控制,它将产生一个缓慢变化的场;
- (4) 当 $\mu = 0$ 时,就退化为经典的图像梯度外力场情况。

可以根据 ∇f 的变化重新定义 μ , ∇f 较大时就减小 μ , 避免对 ∇f 的平滑过大,以至 snake 轮廓不能进入狭长的凹陷区域, ∇f 较小就增大 μ , 使得 GVF 场的作用范围也逐渐扩大,加快轮廓的逼近。

令 $\mu = g(|\nabla f|) = 1 - |\nabla f|^2/k = 1 - (f_x^2 + f_y^2)/k$, 则式(1)表示成

$$E = \iint g(|\nabla f|)(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V_{\text{GVF}} - \nabla f|^2 dx dy \quad (3)$$

根据欧拉公式有:

$$\nabla g(|\nabla f|) * \nabla u + g(|\nabla f|) \nabla^2 u - (u - fx)(f_x^2 + f_y^2) = 0$$

$$\nabla g(|\nabla f|) * \nabla v + g(|\nabla f|) \nabla^2 v - (v - fy)(f_x^2 + f_y^2) = 0 \quad (4)$$

其中 ∇^2 是 Laplace 算子, f_x, f_y 是边缘图对 x, y 的偏导, k 是调节系数。式 $1 - |\nabla f|^2/k$ 和式 $|\nabla f|^2$ 是此消彼长的关系,因此在边缘附近后项会较大,前项会较小。将式(4)进行离散化处理,因 $\nabla g(|\nabla f|) * \nabla u \approx 0, \nabla g(|\nabla f|) * \nabla v \approx 0$, 可以得到 GVF 向量场 $V_{\text{GVF}}(x, y)$ 的两个分量 $u(x, y), v(x, y)$ 分别为:

$$u_{n+1} = u_n + (1 - (f_x^2 + f_y^2)/k) \nabla^2 u_n - (u_n - fx)(f_x^2 + f_y^2)$$

$$v_{n+1} = v_n + (1 - (f_x^2 + f_y^2)/k) \nabla^2 v_n - (v_n - fy)(f_x^2 + f_y^2)$$

其中 n 是迭代次数。根据 GVF 理论将 f_x, f_y 进行多次迭代,可以得到新的向量场分量 u 和 v , 在它们的作用下,snake 模型的初始轮廓可以不必充分靠近真实边缘,并可以进入凹形区域。

2 实验结果

图 4(a)是采用原算法的实验结果,图 4(b)是运用文中方法的实验结果,图 4(c)~(e)是各种口型的提取结果。比较图 4(a)与图 4(b)的实验结果可以看出:文中算法由于图像能量项利用了彩色信息,通过改变平滑因子增强了外部能量的自适应性,不但加快了收敛的速度,而且使轮廓提取的效果更加精确,最终 snake 曲线与目标轮廓吻合更好。

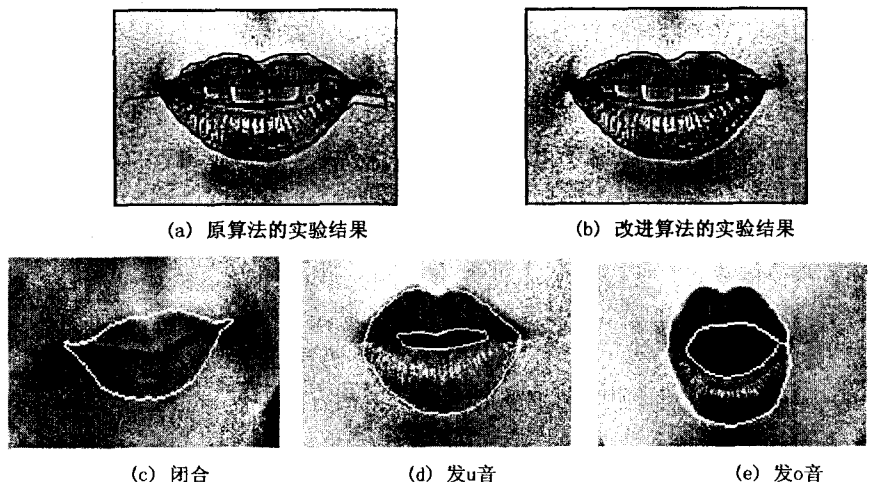


图 4 实验结果

3 结束语

针对彩色嘴唇图像的特点,为了更好地提取其轮廓,在一般的 snake 算法的基础上,提出了一种改进的 snake 模型,首先通过估计嘴唇的大致区域和形状,自

文中选用 $\Delta t_i = 0.01, \Delta t_s = 0.1, A = 20, B = 2$, 在 256×256 的灰度图像上进行实验; 图 1 是一幅含噪声的破损图像, (a) 是原图像, 经过 200 次迭代后修复成 (b), 经过 400 次迭代后修复为 (c), 经过 500 次迭

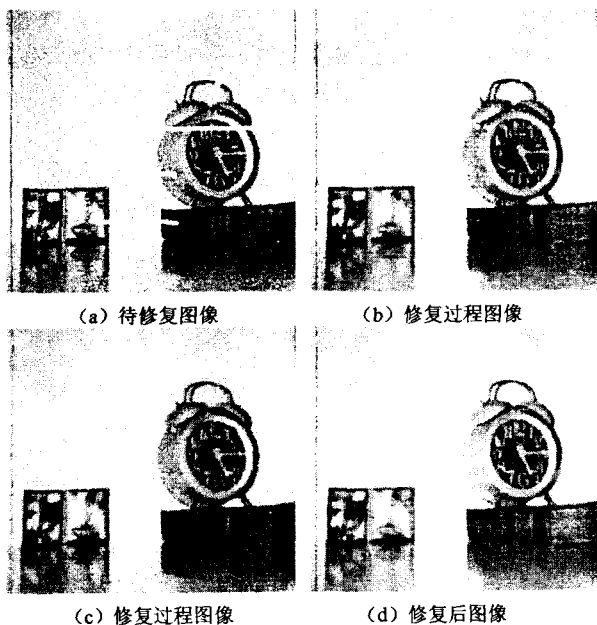


图 1 含噪声的破损图像修复过程图

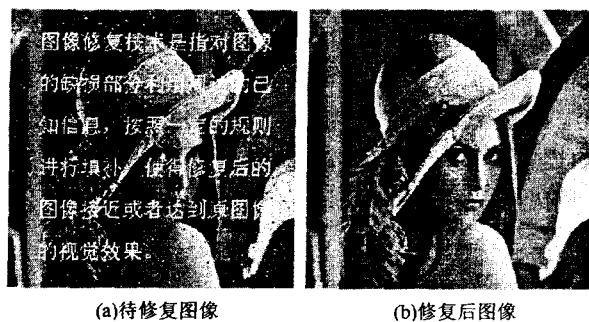


图 2 含椒盐噪声的图像修复

代修复完成 (d)。图 2 是对含椒盐噪声的 Lena 图进行的处理。

4 结束语

本方法信息的传播是沿着等照度线方向的, 图像的平滑也是各向异性的, 因此边缘修复的效果较好, 这点可以从图 1, 图 2 中看出。采用文中提出的模型进行数字图像修复同时去噪是一种有效的方法, 利用 PDE 方法处理图像既方便又有效。

参考文献:

- [1] Bertalmio M, Sapiro G, Caselles V, et al. Image Inpainting [C]//In Proc. ACM Conf. Comp. Graphics (SIGGRAPH 2000). New Orleans: [s. n.], 2000: 417 - 424.
- [2] Bertalmio M. Processing of Flat and non - Flat Image Information on Arbitrary Manifolds Using Partial Differential Equations[D]. USA: University of Minnesota, 2001.
- [3] Chan T F, Shen J H. Non texture inpainting by curvature driven diffusion (CDD) [D]. Los Angeles: Univ. of California, 2000.
- [4] Rudin L, Osher S, Fatemi E. Nonlinear Total Variation Based Noise Removal Algorithms[J]. Physical D, 1992, 60(1 - 4): 259 - 268.
- [5] Chan T F, Shen J. Mathematical Models for Local Nontexture Inpaintings[J]. SIAM Journal on Applied Mathematics, 2002, 62: 1019 - 1043.
- [6] Perona P, Malik J. Scale - space and Edge Detection Using Anisotropic Diffusion[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1990, 12: 629 - 639.
- [7] Alvarez L, Lions P L, Morel J M. Image selective smoothing and edge detection by nonlinear diffusion[J]. SIAM J. Numer. Anal, 1992, 29(3): 845 - 866.

(上接第 97 页)

动生成 snake 的初始模型, 解决了手工设置初值的问题; 然后在 GVF - snake 的基础上重新设计了 snake 的外部能量函数, 采用色彩聚类算法对原始图像进行分割后进行差分提取有意义区域的边缘梯度, 代替传统方法中常用的像素灰度的差分来近似图像梯度, 充分利用了图像的彩色信息, 解决了 snake 原型应用于彩色图像时出现的失真问题, 对 GVF 向量场进行了归一化处理并改进了平滑因子, 解决了经典力场中活动轮廓的收敛范围小、外力抗干扰性差等缺点。

参考文献:

- [1] Kass M, Witkin A, Terzopoulos D. Snakes: active contour models[J]. International Journal of Computer Vision, 1988, 2(3): 321 - 331.
- [2] Gonzalez R C, Woods R E. 数字图像处理[M]. 第 2 版. 北京: 电子工业出版社, 2003.
- [3] 边肇祺, 张学工. 模式识别[M]. 第 2 版. 北京: 清华大学出版社, 2003.
- [4] 李培华, 张田文. 主动轮廓线模型(蛇模型)综述[J]. 软件学报, 2000, 11(6): 51 - 75.
- [5] Xu C Y, Prince J I. Snakes, shapes, and gradient vector flow [J]. IEEE Transactions on Image Processing, 1998, 7(3): 359 - 369.
- [6] 姚鸿勋, 刘明宝, 高文, 等. 基于彩色图像的色彩坐标变换的面部定位与跟踪法[J]. 计算机学报, 2000, 23(2): 158 - 165.
- [7] 赵利刚, 鲍苏苏. 基于改进的活动轮廓图像分割[J]. 计算机与数字工程, 2006, 5(6): 58 - 76.