

基于视觉注意力的视频图像卡通化技术*

张群, 陈锻生

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 提出了一种新型的视频图像卡通化处理方案。首先根据视觉显著性模型从输入的视频图像中计算出一张视觉注意力函数图, 并将视频图像转换到 $L^a \cdot b^*$ 色彩空间; 在视觉注意图的指导下对低对比度区域用迭代的可分离双边滤波器做平滑操作, 对高对比度区域用 DoG 算子作线条增强; 最后用颜色的软量化算子处理 L 亮度通道使之生成类似于卡通化的颜料块效果, 同时加深各边界之间的不连续性。实验表明, 本方案对高注意力区域保留更多的细节信息来突出感兴趣区域, 同时淡化了不感兴趣的区域, 解决了 Winnemöller 不能处理前景和背景对比度不明显的图像的情况, 而且在视觉上更能体现出卡通化的效果。

关键词: 图像卡通化; 视觉注意力; 双边滤波; 线条绘制; 颜色量化

中图分类号: TP391.41

文献标识码: A

文章编号: 1674-7720(2013)10-0045-03

Converting video image to cartoon-based visual attention

Zhang Qun, Chen Duansheng

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: To convert video image into cartoon effectively, a new approach is presented. Firstly, a visual attention function image is calculated based on the visual saliency model for the video image, then the color image is converted to the $L^a \cdot b^*$ color space. Under the guidance of visual attention, we smooth the low contrast areas using iterative separable bilateral filter and enhance the line drawing using DoG operator in the high contrast areas. Finally, soft color quantization operator is applied to luminance channel to generate cartoon like block effect, and the discontinuity between boundaries is more evident. The experiment show that this approach can reserve more details for high attention area to highlight the region of interest, and at the same time overlook uninterested area, solve a problem that Winnemöller can not handled in the case of the image which with similar foreground and background color shows better cartoon like effect visually.

Key words: image cartoon stylization; visual attention; bilateral filter; line drawing; color quantization

近年来, 视频或图像的卡通化作为一种新兴的文化产业受到越来越多国内外视觉领域专家学者的关注和研究, 其中很多国外学者提出了实现视频自动卡通化风格的方法。WINNEMÖLLER H 等人^[1]提出了一种实时的视频和图像卡通化框架, 该框架用双边滤波器来进一步减少低对比度区域的对比度, 同时用高斯差分线条来加深高对比度区域的对比度。DECARLO D 和 SANTELLA A^[2]在一个多尺度的系统下使用眼动仪来对图像进行简化。但是由于仪器价格昂贵而且测试起来比较复杂耗时, ITTI L 等人^[3]提出了基于显著性模型来检测感兴趣的区域, 他们认为除了颜色和亮度以外, 对比度本身也是一

个很重要的特征信息。HAREL J^[4]等人在参考文献[3]的基础上提出了一种计算视觉显著性的改进算法, 该算法不仅考虑了特征的相似性信息, 而且还考虑了特征的差异性信息, 从而可以对图像特征进行准确的计算。

本文在参考文献[1]的基础上使用参考文献[4]的计算视觉显著性模型来简化视频图像中的不重要特征, 同时突出重要的视觉特征。

1 提取图像的显著性特征图

显著性特征图是一种表示相关场景视觉显著性的局部图。参考文献[4]提出了精确的特征计算方法, 该算法采用拓扑结构而不再是 bottom-up 的方式来达到精确的特征计算, 同时还使用了马尔科夫链来权衡图像中两

*基金项目: 福建省科技计划重大项目 (2011H0028, 2011H6016)

图形、图像与多媒体

点之间边的权重和状态的转换。该算法不仅考虑了特征的相似性,而且考虑了特征的差异性。

参考文献[4]算法的主要步骤如下。

(1)输入一张视频图像 $I(x,y)$,假定特征映射是参考文献[3]的算法计算出来的:

$$M:[n]^2 \rightarrow R \tag{1}$$

(2)计算一个激活映射:

$$A:[n]^2 \rightarrow R \tag{2}$$

激活映射可以增强点的映射, $M[i,j]$ 如果与它的邻域中的值有明显差异,就会产生一个大的 A 值。为了更具有说服力,在 (i,j) 的周围区域中计算 $M[i,j]$ 的直方图,定义 (i,j) 的激活映射为:

$$A(i,j) = -\log(p(i,j)) \tag{3}$$

其中, $p(i,j)$ 定义为:

$$p(i,j) = \Pr\{M(i,j)|neighborhood\} \tag{4}$$

用一种原始的方法定义一对点之间的不相似性:

$$d((i,j) \parallel (p,q)) = |\log M(i,j) - \log M(p,q)| \tag{5}$$

其中, $d((i,j) \parallel (p,q))$ 表示两像素点的特征差异,每个像素点表示一个节点,两个被连接的边 (i,j) 和 (p,q) 将分配一个权值:

$$W_1((i,j),(p,q)) = d((i,j),(p,q)) * F(i-p, j-q) \tag{6}$$

其中, $F(a,b) = e^{-\frac{a^2+b^2}{2\sigma^2}}$ 。

(3)视觉注意力算法期望将感兴趣的区域集中在几个重要的位置上,在增强映射的基础上又添加了一个映射,将所有的像素点都视为节点,连接两个节点的权值定义为:

$$W_2((i,j),(p,q)) = A(p,q) * F(i-p, j-q) \tag{7}$$

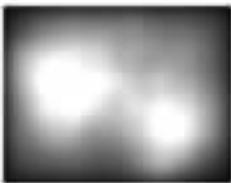
再一次应用步骤(2)中的方法构建马尔科夫链, A 值越大的地方会优先收敛聚集。两组图像实验结果的对比如图1和图2所示。

从参考文献[3]算法的显著性效果图可以看出,高视觉注意力区域不够明显而且聚集程度也不高;从参考文献[4]方法的显著性效果图可以看出,感兴趣区域点的分



(a) 原图

(b) 参考文献[3]方法显著性效果图



(c) 参考文献[4]方法显著性效果图

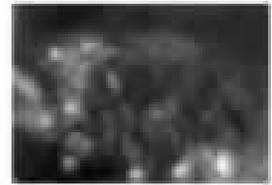


(d) 原图与参考文献[4]方法显著性效果图叠加效果

图1 第1组视觉显著性结果图



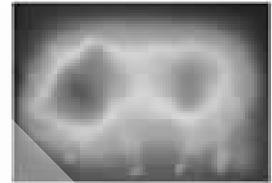
(a) 原图



(b) 参考文献[3]方法显著性效果图



(c) 参考文献[4]方法显著性效果图



(d) 原图与参考文献[4]方法显著性效果图的叠加效果

图2 第2组视觉显著性结果图

布更加集中,高注意力区域更加连贯,线条也更加明显,同时也能很好地描述主体的大致轮廓。将参考文献[4]方法的显著性函数图叠加到原图,其中中心区域和白色光环表示人眼停留在此区域的时间长短,中心区域表示停留时间比较长,可见,在前景和背景对比度不明显的情况下,参考文献[4]的算法也可以很好地检测出感兴趣的区域。

用提取的显著性特征图指导双边滤波器对图像中感兴趣区域和不感兴趣区域的模糊程度,通过控制双边滤波器^[5-6]的范围权值函数 $w(\cdot)$ 决定图像什么地方应该模糊什么地方应该强化保留。范围权值函数定义为:

$$w(x, \hat{x}, \sigma_r) = (1 - m(\hat{x})) - w'(x, \hat{x}, \sigma_r) + m(\hat{x}) \cdot u(\hat{x}) \tag{8}$$

$$w'(x, \hat{x}, \sigma_r) = e^{-\frac{1}{2} \left(\frac{\|f(\hat{x}) - f(x)\|}{\sigma_r} \right)^2} \tag{9}$$

$$u(\hat{x}) = f(\hat{x}) \cdot s(\hat{x}) \tag{10}$$

其中, \hat{x} 代表一个像素的位置, x 是它临域的像素点; σ_d 与模糊的半径相关, σ_d 值越大效果越模糊,如果 σ_d 值过大就会模糊掉图像的重要边缘信息; σ_r 决定了对比度的保留和模糊程度。式(10)用来决定对视觉感兴趣区域和非感兴趣区域的平滑程度。

参考文献[1]双边滤波效果如图3所示,利用显著特征图改进的效果图如图4所示。从图3和图4的对比中可以发现,参考文献[1]的双边滤波对整体都有均匀的模糊效果,但是利用显著特征图指导双边滤波就可以在感兴趣的区域模糊程度相对低一些,从而保留更多的细节信息,同时对于不感兴趣的背景区域(如图中的地面和



图3 参考文献[1]双边滤波的效果图



图4 利用显著特征图改进的效果图

墙)都能去掉更多的细节信息。

2 颜色的软量化

为了减少颜色的数目并生成具有类似颜料块的效果,同时也为了避免视频帧与帧切换时造成的不稳定性,使用一个颜色软量子化方程^[1],以增强视频图像的卡通化效果。定义一个量子化函数:

$$Q(\hat{x}, q, \varphi_q) = q_{\text{nearest}} + \frac{\Delta q}{2} \tanh(f(\hat{x}) - q_{\text{nearest}})$$

其中, Δq 表示相邻色阶之间的宽度, q_{nearest} 是最接近 $f(\hat{x})$ 的那个色阶, φ_q 控制着从一个色阶跳跃到下一个色阶的变化尖锐程度。实验中取 $q=9, [\Lambda_s=0, \Omega_s=2][\Lambda_\varphi=3, \Omega_\varphi=14]$ 。

3 实验与分析

两组卡通化效果图如图5和图6所示。因为参考文献[1]假设感兴趣区域是图像的亮度和高对比度区域,对于第2组猫的图像来说,处理前景和背景的颜色差异比较相近的图像就失败了,不能很好地分辨出感兴趣区域。而本文的算法先是采用参考文献[4]算法检测出图像的感兴趣区域,然后再采用不同的权值对图像进行不同程度的滤波,所以在处理这种前景与背景对比度不明显的图像时也能达到很好的效果,并且弥补了参考文献[1]算法不能处理图像前景与背景对比度不明显的情况,从而在视觉效果上更能体现出图像的个性化和艺术性。

本文提出了一种新型的卡通化方案,在具有参考文献[1]的实时卡通化效果的基础上增加了视觉显著性模型,实现了视频图像的非均匀卡通化处理。实验表明,本文的方案在卡通化一些前景和背景对比度不明显的图像时可以取得很好的效果。

参考文献

[1] WINNEMÖLLER H, OLSEN S C, GOOEH B. Real-time



(a) 输入的原图 (b) 参考文献[1]方法的卡通化效果图 (c) 本文非均匀卡通化效果图

图5 第1组图像的卡通化效果图



(a) 输入的原图 (b) 参考文献[1]方法的卡通化效果图 (c) 本文非均匀卡通化效果图

图6 第2组图像的卡通化效果图

video abstraction[C]. Proceedings of ACM SIGGRAPH'06, ACM, 2006:1221-1226.

[2] DECARLO D, SANTELLA A. Stylization and abstraction of photographs[C]. Proceedings of ACM SIGGRAPH, 2002: 769-776.

[3] ITTI L, KOEH C. Computational modeling of visual attention[C]. Nature Reviews Neuro-Science 2, 2001:194-203.

[4] HAREL J, KOCH C, PERONA P. Graph-based visual saliency[J]. 2006. Nips, 2006.

[5] PHAM T Q, VAN VLIET L J. Separable bilateral filtering for fast video preprocessing[C]. Proceedings of IEEE International Conference on Multimedia and Expo, IEEE Computer Society, 2005:454-457.

[6] PHAM T Q, VLIET L J. Separable bilateral filtering for fast video preprocessing[C]. IEEE International Conference on Multimedia & Expo, CD1-4.

(收稿日期:2013-02-21)

作者简介:

张群,女,1988年生,硕士研究生,主要研究方向:图像处理,计算机视觉。

陈锻生,男,1959年生,博士,教授,主要研究方向:图像模式识别。