

面向移动的卡通人脸生成系统^{*}

李龙生^{1,2} 周经野¹ 陈益强^{2,3} 刘军发^{2,3} 陈振宇¹

(湘潭大学信息工程学院 湘潭 411105)¹ (上海中科计算技术研究所 上海 201203)²

(中国科学院计算技术研究所 北京 100080)³

摘要 本文给出了一个面向移动数字娱乐的肖像漫画生成系统。该系统的输入是一张真实人脸照片,输出为具有娱乐效果的夸张人脸肖像。主要研究内容包括:第一,人脸照片的肖像插图画的生成。文中采用了 Unsharp Mask 锐化处理结合色彩缩放运算的方法,本方法不仅考虑到移动设备用户对计算速度的要求,而且能保留人脸的绝大部分特征,从而更为有效地生成人脸的肖像插图画。第二,人脸表情变形以产生人物的幽默漫画。对基于特征的变形方法进行了改进,使其更加自动化。此外,本文改进了传统的 ASM 方法,用来帮助去除人脸图像背景信息和更为方便地控制人脸上的特征点。整个系统在局域网和 PDA 上表现出了很好的娱乐效果。

关键词 图像处理,漫画,锐化,人脸变形,数字娱乐

Caricature Generation for Mobile Device

LI Long-Sheng^{1,2} ZHOU Jing-Ye¹ CHEN Yi-Qiang^{2,3} LIU Jun-Fa^{2,3} CHEN Zhen-Yu¹

(The Institute of Information Technology, Xiangtan University, Xiangtan 411105)¹

(Institute of Computing Technology, Chinese Academy of Sciences, Shanghai 201203)²

(Institute of computing Technology, Chinese Academy of Sciences, Beijing 100080)³

Abstract In this paper we propose a caricature generation for mobile digital entertainments device. The input of system is a facial photograph, the output is entertaining caricature. The main research work include: first, we create illustration from photographs of human faces. We apply Unsharp Mask sharpening and scale color processing. It's not only considering the speed of mobile device, but also reserving many facial features. These studies show that the facial illustrations generated using our techniques produce images that are effective. Second, generate the caricatures by image morphing. We improve the traditional feature-based metamorphosis for automation morphing. In addition, we improve the traditional ASM method for removing the background of facial photographs and controlling facial feature points. The system presents better entertainments.

Keywords Image processing, Caricature, Unsharp mask sharpening, Facial metamorphosis, Image entertainments

1 引言

在许多著名游览区,我们常常可以看到艺术家为游客绘制肖像画,这些肖像画中除了一部分写实的传统肖像画外,还有一部分以夸张变形为特点,寻求趣味性的肖像漫画,这些画以简单的线条、单纯的色彩,反映出模特的独一无二的神彩。在 Internet 网络上,我们也常常可以看到一些著名漫画家绘制的各国政要,演艺圈中的名人的漫画肖像,这些漫画也生动地反映出模特的性格、职业以及独一无二的神韵。

近年来,移动终端上的数字娱乐逐渐成为一种潮流。传统的人与人之间通过移动设备交流的方式是语音或者文本信息,随着 3G 移动网络和手机的发展,基于手机终端的数字媒体内容比简单的语音和文本具有更多的需求。将真实人脸照片变成肖像漫画也是近年来数字娱乐技术的研究问题之一。肖像漫画被用于手机终端,比传统的语音和文本更具有娱乐性,在这些技术领域,人们进行了大量的研究工作。

我们当前所做的工作主要基于两部分,一是基于真实照片生成肖像插图画的研究。二是对产生的肖像插图画进行夸

张变形。现在流行的主要有两种方法可以自动地生成照片的插图画。文[8]给出了一种产生黑白肖像画的方法。他首先使用 Blommaert 和 Martens' 的模型计算出一幅灰度图像,然后对其二值化,再与从真实照片产生二值图进行叠加,从而得到照片上人物的肖像插图画,这种方法的效果主要依赖于 Blommaert 和 Martens' 的模型对真实照片计算的结果,实验表明,这种方法对手机或 PDA 的照片的处理并不理想,且计算速度慢。另外一种方法来自文[3,4]所做的工作。他们主要是通过从照片中所提取到的人脸轮廓信息,然后用固定的线条构建出相应的肖像画。相对来说,前者比较接近真实照片,但包含了较多的冗余信息。后者完全使用线条构造人脸,人脸表征过于简单,失去了许多重要的反映人物特征的重要信息。

人脸变形方面,常用的有两种变形方法。一种是网格变形^[1],另一种是基于线对的变形^[2],基于网格的变形技术是图像变形技术中最早的方法之一。其算法输入是一幅源图像和两个二维控制点坐标数组 u 和 v , u 是源图像上控制点的坐标数组, r 则是经过位移后的控制点坐标数组。 u 和 r 具有相同

^{*} 本文得到国家自然科学基金(60303018)、北京科技新星计划基金(2005B54)以及北京工业大学多媒体与智能软件技术北京市重点实验室开放课题基金联合资助。

的控制点数和相同的网络拓扑结构。将源图像上的控制点从 u 变换到 r , 同时带动周围的像素点, 这就实现了图像变形。另一种是基于线对的变形: 文[2]中的这种方法使得我们可以在高层控制变形过程。我们在图像中选取相应的特征点线对, 算法利用线对将源图像中的特征点和目标图像中的特征点相对应, 特征点周围区域的像素点受到不同程度的影响。变形采用的是逆向映射, 即逐个扫描目标图像的像素, 根据其位置采用一定的算法对源图像进行采样, 从而产生高层想达到的变形效果。

2 系统概述

我们的系统运行在一个具有服务器和客户端的无线局域网中, 由三步组成。第一步是人脸的自动定位。这里我们采用的是改进的 ASM (Active Shape Model) 方法。传统的 ASM 方法其输出并不包含前额的 9 个特征点, 经改进之后, 可以得到人脸的整个形状。第二步是肖像插图画的产生。我们采用 Unsharp Mask 锐化处理结合色彩缩放运算的方法, 本方法不仅考虑到移动设备用户对计算速度的要求, 而且能保留人脸的绝大部分特征, 从而更为有效地生成人脸的肖像插图画。第三步是对肖像插图画进行变形生成具有一定娱乐效果的幽默漫画。

客户端是指移动设备 (例如手机、PDA 等)。在客户端获取人脸照片, 然后发送到服务器端生成相应的幽默漫画。然后幽默漫画就可在不同的客户终端之间传送。

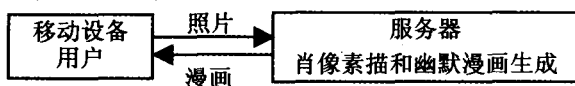
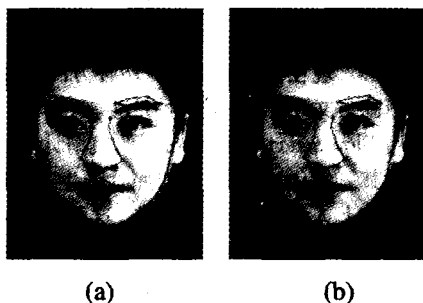


图1 系统运行框架

系统在 3G 网络运行时, 移动设备就是指 3G 手机, 而服务器则是指移动通信网上的服务器。手机总是和服务器交换数据, 首先拍照, 传送到服务器, 服务器处理成用户照片的肖



像漫画再返回给用户。

3 人脸肖像漫画自动生成方法

3.1 人脸头像区域的获取

一般来说, ASM 是最常用人脸定位的一种方法^[4,5]。然而, 传统的 ASM 仅仅考虑了脸颊、眼睛、鼻子和嘴巴这些部位的特征点 (见图 2(a))。我们可以通过扩展 ASM 的特征点, 将整个人脸完全覆盖, 在此, 我们加以改进; 在人脸的前额部分增加了一些特征点 (见图 2(b)), 这样的改进可以达到如下两个效果。

第一是可以为去除背景作好准备。一般来说, 人物的肖像漫画并不需要背景, 因此去除背景本身是肖像漫画的要求之一, 另外去除背景也可以防止照片背景上的物体也被扭曲变形, 因此我们就尽可能地将背景信息去除。只要我们能获得人脸的区域, 去除人脸以外的背景就相对比较容易。为了保留尽量多的头发信息, 我们构造了一个高斯函数。由于头发总是在离人脸区域最近的地方, 我们设计的函数可将远离人脸区域的像素点去除, 离得越远, 去除得越干净, 这样就达到了尽量多地保留头发信息的目的。大量试验表明, 这样处理的效果的确很好 (见图 2(c))。去除背景的函数是这样构造的:

$$f(d_i) = A \cdot L_i \cdot \exp(-B \cdot (d_i/D)^2), i=0,1,\dots,M \times N \quad (1)$$

$$d_i = \begin{cases} 0 & , P_i \in S \\ \min(\sqrt{(P_i x - P_j x)(P_i y - P_j y)}) & , j=0,1,\dots,n, P_i \notin S \end{cases} \quad (2)$$

其中: $f(d_i)$ 表示返回的灰度值; d_i 是像素点 i 到人脸区域的距离; L_i 是像素点 i 的原始灰度值; D 是人脸区域宽度的一半; A 和 B 是经验系数。 M 和 N 分别为图像的宽度和高度; S 为 ASM 方法获得的人脸区域; $(P_i x, P_i y)$ 是像素点 i 的坐标; n 是人脸区域外边缘的特征点的总个数。 $(P_j x, P_j y)$ 是我们定义的 118 个特征点中的任意一点。

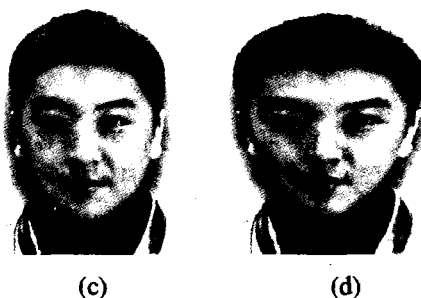


图2 改进的 ASM 方法。(a)是传统的 ASM 方法检测出的特征点, (b)是我们改进的 ASM 方法检测出的特征点, (c)去除背景后的图像, (d)是利用我们的方法将图像变形后的图像。

第二是可以全自动地完成人脸变形。我们采用了线对变形方法。然而这种方法的一个缺陷是需要手动地去指定特征点, 我们对此进行了改进。改进的 ASM 方法能够自动地获得特征点, 从而可以让计算机自动确定指定的线对所在的位置的坐标, 弥补了该方法的不足 (详见 4.2 节)。比如我们想获得如图 2(d)所示的变形效果, 改进的 ASM 方法能够提供前额部分的特征点, 从而可以达到自动变形的效果。

3.2 人脸插图画的生成

我们将彩色图像进行 Unsharp Mask 锐化处理, 再进行色彩缩放运算, 可以得到具有一定半径的图像的边缘细节。

Unsharp Mask 处理的基本原理是先通过高斯滤波得到原始图像的一个平滑版本, 然后在此基础上用原始图像减掉这部分平滑的部分, 从而达到对原始图像进行高通滤波的作用。常用的二维离散高斯滤波器的数学表达式如下:

$$g[i, j] = e^{-\frac{i^2 + j^2}{2\sigma^2}} \quad (3)$$

通过上式我们选取合适的 σ^2 , 就可以在 $n \times n$ 的窗口上设计高斯滤波模板 $w[i, j]$, 设原始图像的函数表达式为 $f[i, j]$, 这样通过在图像 $f[i, j]$ 上的卷积操作就可以达到对图像进行平滑处理的目的, 设平滑后的图像为 $h[i, j]$, $*$ 为卷积运算, 具体计算式为:

$$h[i,j]=f[i,j]*w[i,j]$$

$$= \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} w[k,l] \times f[i-k,j-l] \quad (4)$$

最后,我们用原图像与这个平滑版本做如下的减法运算,就可以得到较好的高通滤波效果,实验中我们用大量的图片处理,都可以得到较好的高通效果。设最后高通的结果图片的表达式为 $t[i,j]$,我们的减法运算为:

$$t[i,j]=(1+depth) \times f[i,j]-(depth-\eta) \times h[i,j] \quad (5)$$

式中 $depth$ 为调整从原始图片中减去的平滑图片的量, $depth$ 越大,图像越锐化, η 是调整用经验系数,我们的实验中 η 一

般取 0.3, $depth$ 一般取 4.0, σ 一般取 3.3~5.0 之间。

通过以上方法处理,我们得到了图像的高通滤波结果,但为了得到更好的高通滤波效果,我们对处理后的图片再进行了色彩缩放处理,色彩缩放的基本原理就是对图像的 r, g, b 三个颜色分量进行同等程度的放缩处理,在保证一定平滑度的基础上让图像得到更好的高通效果。处理的结果如图 3 (d) 所示。可以看到,用我们的算法产生的肖像插图,能更好地保留图人像的特征,又可以去除一些图像上的冗余信息,这在移动设备的带宽限制下更能体现其应用价值。

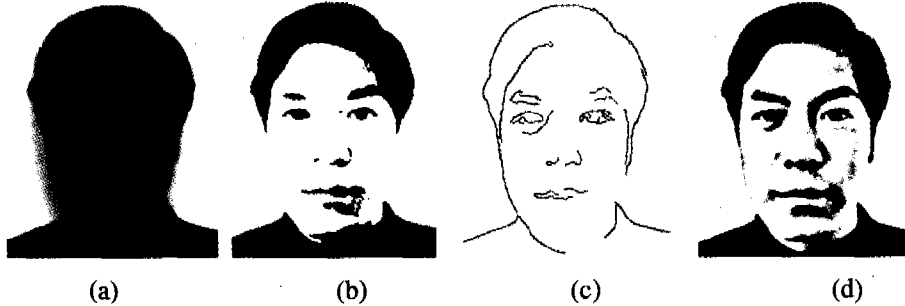


图 3 三种方法的比较,(a)为原始图像,(b)为当阈值等于 125 时产生的二值图像,(c)是利用 Canny 算法直接从照片提取出的边缘,(d)是本文算法产生的人脸肖像插图。

3.3 人脸变形

本文将网格变形和基于线对的变形两种算法的性能进行了比较。结果表明它们在不同的方面性能表现不一。当我们从计算速度的角度来衡量时,网格变形计算速度更快。因为在网格变形中,人脸被划分成很多网格区域,每个网格点都有属于自己的局部影响区域,计算是在区域级别上进行的。表 1 给出了我们的实验中两种算法变形所需的平均时间。

表 1 两种算法计算速度的比较

算法名称	计算时间
网格变形	0.25s (有 118 个网格点)
基于线对的变形	0.85s (有三条线对)

然而,当我们比较它们的变形质量时,基于线对的变形方法则更胜一筹。因为在基于线对的变形算法中,图像的变形是点对点进行的,这就保证了源图像和目标图像之间像素的转换足够的平滑。而在网格变形中,源图像转换到目标图像所进行的映射是以像素块为单元的,因此这种转换其边缘部分不够自然。基于这一特点,我们选择基于线对变形的方用来对肖像插图变形,从而生成幽默漫画。通过基于线对的变形,我们可将指定的受控的像素点精确地从原始位置变形到我们所期望的目标位置。控制点周围的其它像素点则可以根据该点非常平滑地计算出变形所产生的位移量。这样我们就可直觉地从高层控制变形的过程,因此可以很方便地生成幽默漫画。



图 4 实验结果。上排为真实人脸图像,下排为本文算法产生的人脸肖像漫画。

此外,基于线对的变形方法还有另外一个缺点,即在常规情况下,我们必须手动指定表现人脸特征的特征线对,如眼

睛、嘴巴、鼻子等。但在我们的系统中,首先执行的 ASM 已将人脸准确地定位,也就是说特征线对已经自动获得了。例

如,如果想将鼻子变大,我们不必指出鼻子的起始和结束位置。我们要做的仅仅是将鼻子起始点和结束点的序号输入到变形模块。ASM已经在人脸图像中将鼻子定位,然后把特征点序号映射为相应像素点的坐标。对于一个给定的人脸,我们给出了一些幽默漫画生成的类型,比如大嘴巴、胖脸、大鼻子、小眼睛等等。用户可以根据变形的风格自由地进行选择。

结论和将来的工作 本文实现了一个新的基于无线局域网的手机动画系统。通过 PDA 拍照获得人脸正面图像,传送到服务器。通过改进的 ASM 获取人脸特征点。去除背景后,使用 Unsharp Mask 锐化处理结合色彩缩放运算的方法将照片转变为插图,然后采用 ASM 与特征线对变形技术相结合的方式,实现人脸幽默漫画。本系统在 3G 手机上可以实现多媒体娱乐彩信。

本系统的下一步工作是在真实的 3G 网络上进行试验,同时开展几个具有挑战性的问题的研究。其一是如何为人脸幽默肖像渲染颜色,这也是 NPR(Non Photorealistic Rendering)领域比较活跃的一个问题。其二是为动画增加手势合成,用夸张的人脸表情和同步手势来共同表达交流的内容,可以获得更准确的交流和更精彩的娱乐效果。

参考文献

- 1 Wolberg G. Image Morphing, A Survey, Visual Computer, 1998, 14: 360~372
- 2 Beier T, Neely S. Feature-based image metamorphosis. In: Proc. SIGGRAPH 92. In Computer Graphics, 1992. 35~42
- 3 Chen Hong, Liu Ziqiang, Rose Chunk, Xu Yingqing, Shum H Y, Salesin D. Example-Based Composite Sketching of Human

- Portraits. In: Proc. of the 3rd international symposium on Non-photorealistic animation and rendering table of contents, Annecy, France, 2004
- 4 Liang Lin, Chen Hong, Xu Ying-Qing, Shum H Y. Example-based Caricature Gen-eration with Exaggeration. In: IEEE Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, 2002
- 5 Cootes T F, Taylor C J, Cooper D, Graham J. Active shapemod-els—their training and application. Computer vision and image understanding, 1995, 61(1): 38~59
- 6 Omedas P, Berrizbeitia F, Szijártó G, Kiss B, Takács B. Model-based Facial Anima-tion for Mobile Communication. In: 1st Ibero-American Symposium on Computer Graphics (SIACG), Guim arães, Portugal, July 2002
- 7 Gutierrez M, Vexo F, Thalmann D. The Mobile Animator: Interactive Char-acter Animation in Collaborative Virtual Environ-ments. Virtual Reality Lab (VRlab), Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne Switzer-land CH-1015
- 8 Gooch B, Reinhard E, Gooch A. Human Facial Illustrations; Creation and Psychophysical Evaluation. ACM Transactions on Graphics, 2004, 23(1): 27~44
- 9 Miao Jun, Gao Wen, Ma C, Chen Yiqiang. Image Scale-Smoothing, Scale-Differentiating and Binarizing: A New Frame for General-Scale-Edge (GSE) Extraction. In: Proc. IEEE International Conference on System, Man and Cybernetics, Tucson, Oct. 2001. 2005~2010
- 10 Gao Wen, Chen Yiqiang, Wang Rui, Shan Shiguang, Jiang Dalong. Learning and Synthesizing MPEG-4 Compatible 3-D Face Animation from Video Sequence. IEEE Transactions on Circuits and Systems for Video technology, 2003, 13(11): 1119~1128
- 11 Chen Yiqiang, Gao Wen, Wang Zhaoqi, Yang Changshui, Jiang Dalong. Text to Avatar in Multimodal Human Computer Inter-face, Asia-Pacific Human Computer Interface (APCHI2002), 2002, 2: 636~643
- 12 Chen Yiqiang, Gao Wen, Wang Zhaoqi, Miao Jun, Jiang Dalong. Mining Audio/Visual Database for Speech Driven Face Anima-tion. In: Proc. IEEE International Conference on System, Man and Cybernetics (SMC2001), 2001. 2638~2643

(上接第 162 页)

的结论完全一样。

定理 8 与引理 3 是分别基于两种不同的等价关系 R 与 $\Gamma(\cdot)$ 得到的属性约简方法。下面,我们将证明,条件 $C_R = C_r$ 成立时,选择 $\Theta = C_R$,则这两种约简方法是等价的。

引理 4 $C_R = C_r$ 时, $\overline{C_R} \cup \{\emptyset\} = C_r$

证明: $C_R = C_r$ 时,

$B \in C_R \Leftrightarrow B \in C_r$

$\Leftrightarrow B = \cup \{E_i \in A/r, E_i \cap B \neq \emptyset\}$

$\Leftrightarrow \sim B = \cup \{E_j \in A/r, E_j \not\subset B\} = \cup \{E_j \in A/r, E_j \cap \sim B \neq \emptyset\}$

$\Leftrightarrow \sim B \in C_r = C_R \Leftrightarrow B \in \overline{C_R}$ 或 $B = \{\emptyset\}$ 。

定理 12 $C_R = C_r$ 时,定理 8 与引理 3 描述的属性约简方法等价。

证明 选择 $\Theta = C_R$ 。由引理 4 知,

$C_0 = \{D' \subseteq A; D' \cap D \neq \emptyset (\forall D \in \overline{C_R})\}$

$= \{D' \subseteq A; D' \cap D \neq \emptyset (\forall D \in C_r, D \neq \emptyset)\}$ 。

因为对于任意的 E_i , 有 $E_i \in C_r, E_i \neq \emptyset$, 所以, $D' \cap E_i \neq \emptyset$, 即 C_0 中的元素 D' 包含每个 E_i 的至少一个元素。因此, C_0 中的极小元 D'_0 (即约简 B) 必然包含每个 E_i 的一个元素。即 $D'_0 = B = \{e_1, e_2, \dots, e_n\}$ 。

定理 8 与引理 3 在适当条件下的等价性为我们提供了 $C_R = C_r$ 时两种不同的约简方法。从实际计算来讲,定理 8 更简单易行;从理论意义来讲,引理 3 更具有普遍性。

小结 本文主要研究了以下几个问题:信息系统属性集及其幂集上的三个等价关系 $r, R, \Gamma(\cdot)$ 之间的联系,以及由 r 与 R 这两个等价关系所生成的闭集族 C_R, C_r 与不可辨识属性矩阵 Ω 间的包含关系; C_R 和 C_r 相等的充分必要条件;以及

在 $C_R = C_r$ 时信息系统属性集的约简方法,并证明了该方法与文[7]所给出方法的等价性。利用这些结果,我们可以把对信息系统的幂集的研究转移到属性集上来,从而简化信息系统的知识发现。如在 $C_R = C_r$ 时,利用定理 8 描述的方法,单独考察属性集中的一个属性 a_i , 确定 $[a_i]_r, [a_i]_R$, 相同的属性作为一个集合就是 A/r 中的元素 E_i 。从每个 E_i 中任取一个元素综合起来就是 A 的约简。这种方法显然比直接研究与属性集 A 等价的子集要简单易行。

参考文献

- 1 Pawlak Z. Rough sets. International Journal of Computer and Information Science, 1982, 11: 341~356
- 2 王珏,王任,苗夺谦,等. 基于 Rough set 理论的“数据浓缩”. 计算机学报, 1998, 21(5): 393~400
- 3 王国胤. Rough 集理论与知识获取. 西安: 西安交通大学出版社, 2001
- 4 张文修,吴伟志,梁吉业,等. 粗糙集理论与方法. 北京: 科学出版社, 2001
- 5 Kryzkievicz M. Comparative study of alternative types of knowledge reduction in inconsistent systems. International Journal of Intelligent Systems, 2001, 16: 105~120
- 6 张文修,米据生,吴伟志. 不协调目标信息系统的知识约简. 计算机学报, 2003, 26(1): 12~18
- 7 张文修,梁怡,吴伟志. 信息系统与知识发现. 北京: 科学出版社, 2003
- 8 Stepaniuk J. Approximation Spaces, Reducts and Representatives. In: Polkowski L, Skowron A. ed. Rough Sets in Knowledge Discovery 2, Physica-Verlag, 1998. 127~146
- 9 Novotny M, Pawlak, Z. Algebraic theory of independence in information systems. Fundamenta Informaticae, 1991, 14: 454~476
- 10 Novotny M. Dependence Spaces of Information Systems. In: Or-towska E ed. Incomplete Informations; Rough Set Analy-sis. Physica-Verlag, 1998. 193~246
- 11 Wei Ling, Zhang Wen-xiu. Attribute Reduction Based on Equivalence Relation Defined on Attribute Set and its Power Set. Advances in Soft Computing. Berlin Heidelberg: Springer-Verlag, 2005. 317~326