

一种基于 RBF 网络生成人脸表情的算法

余重基,李际军

(浙江大学 计算机科学与技术学院,浙江 杭州 310027)

(palms@yeah.net)

摘要:提出了只需确定少量标定点就能高精度实现人脸表情生成的算法。此方法并不直接利用标定点,而是通过自动生成的扭曲控制点来进行 2D 图像扭曲变形;为了保证精度,最大程度地减少人工输入标定点的数目。根据脸部肌肉的解剖特点和人脸的对称性定义了一个信息冗余很小的脸部特征点集(FCPS);通过径向基函数神经网络(RBF Network)来确定 FCPS 的移动与表情变化的映射关系;RBF 网络输出的 FCPS 通过数据映射,自动生成一个矫正的扭曲控制点集;2D 径向基函数图像扭曲算法(RBF Warping Algorithm)利用这些扭曲控制点扭曲图像,产生指定的表情,有效地降低了扭曲控制点误差带来的负面影响;为了避免病态数据的影响并减少计算量,训练网络时采用了递推正则最小二乘算法。实验显示,该算法具有良好的实用性。

关键词:表情生成;标定点;扭曲控制点;脸部特征点集

中图分类号: TP391.9;TP391.41;TP182 **文献标识码:** A

An efficient algorithm of facial expression generation based on RBF network

YU Zhong-ji, LI Ji-jun

(College of Computer Science and Technology, Zhejiang University, Hangzhou Zhejiang 310027, China)

Abstract: The method to generate high-precision expression images by locating a small number of definite landmarks was proposed. The warping did not use landmarks as control points directly, but used a set of generated warping control points. Taking account of facial anatomy and symmetry, a novel set named Facial Characteristic Points Set (FCPS) was designed as landmarks with very low information redundancy. A radial basis function network (RBF Network) was employed to define the relationship between the movement of FCPS and the corresponding change of expressions. The FCPS, produced by RBF network, was processed by data-mapping component to generate a set of compensated warping control points. By using these generated warping control points, 2D RBF warping algorithm, which could effectively reduce the negative effects brought from the errors of generated warping control points, was employed to create assigned expression image. To prevent ill-posed data from affecting the computation and to reduce the computation complexity, recursive regularized least square (RRLS) algorithm was adopted in the neural network training.

Key words: expression generation; landmark; warping control point; facial characteristic points set

0 引言

目前基于人脸表情的应用正变得越来越重要,从简单的网络虚拟角色向导、远程虚拟通讯,到复杂的人脸自动识别、人脸图像数据库的自动查询等,都会涉及面部表情的变化。众所周知,人脸的表情是极为丰富的,因而表情识别遇到了很大的困难。而表情的人工生成则是建立在表情识别基础上的应用,困难就更大了。

对于一张毫无表情的人脸图像,根据需要使之产生不同的表情,称为表情的人工生成。最初的方法有纹理映射的 3D 人脸图像合成和 3D 肌肉组织模型等方法^[1],这些方法在操作时需要给定很多抽象的参数;Nur Arad 等人提出了基于径向基函数(Radial Basis Functions, RBF)来插值控制点(Anchor Points)的 2D 扭曲变形方法^[2],但他们根据经验主观地添加控制点,而没有对如何有效确定这些控制点进行深入研究。

研究。若将表情生成建立在人脸自动识别的基础上,则可以克服上述缺陷。H. Kobayashi 和 F. Hara 运用神经网络技术^[1],通过在输入图像上加入 30 个特定的点(称为标定点, Landmarks),不仅成功地识别了 6 种通用表情,并且还给出每种表情的强度(Strength)计算方法^[1]。

本文提出的算法并不直接利用标定点,而是通过训练 RBF 神经网络来确定标定点的移动与表情变化的映射关系;基于脸部肌肉的解剖特点,利用人脸的对称性提出了一个信息冗余很小的脸部特征点集(FCPS),在保证精度的前提下最大程度地减少训练网络时人工输入标定点的数目;为了避免训练时病态数据的影响及减少训练的计算量,采用了递推正则最小二乘算法;在表情生成过程中当 RBF 网络输出一个被高度压缩的 FCPS 时,通过数据映射,自动生成一个被矫正了大小、拍摄角度和头部小角度倾斜的扭曲控制点集;2D 的径向基函数图像扭曲算法(RBF Warping Algorithm)利用这些扭

收稿日期:2004-12-24;修订日期:2005-03-16

作者简介:余重基(1961-),男,浙江杭州人,硕士研究生,主要研究方向:计算机图形学;李际军(1967-),男,山东青岛人,副教授,博士,主要研究方向:计算机图形学、曲面造型、CAD/CAM。

曲控制点扭曲图像,产生指定的表情。RBF Warping 算法对扭曲控制点的精度要求不高,能有效地减少生成的扭曲控制点的误差带来的负面影响。

1 相关知识

心理学家把人的表情分为高兴、悲伤、愤怒、害怕、惊吓和厌恶 6 种基本表情,但人的表情是极为复杂的,在日常生活的人们很少出现纯粹的单一表情。从解剖学的角度,表情是脸部肌肉活动的结果。Ekman 和 Friesen 提出脸部活动编码系统(Facial Action Coding System, FACS)^[3],对头和脸的动作进行了分类编码,它把脸部的各部分的活动编码成 44 个活动单元(Action Units, AU),每个 AU 对应着单一或肌肉群的收缩。这些脸部的活动表达了四种信息:1)情绪;2)交际符号(如眨眼);3)与操作相关的活动(如咬嘴唇);4)示意(伴随着说话并用于加强的动作)^[3]。FACS 虽然很客观,但不能表达语义上的表情,如 AU20 表示双唇及部分颊窝褶皱被拉向侧后方,嘴也向两侧拉长,这并不能反映表达的是什么感情。一个表情总是一些 AU 的组合,K. R. Scherer 和 P. Ekman^[4]解释了人们已经观察到的约 7000 种组合。

表情的人工生成通常是和表情的识别联系在一起的。到目前为止,计算机通过 FACS 识别表情还是非常困难的,有的研究者建立了 6 种通用表情与部分 AU 之间的对应关系^[5],以期对表情人工生成提供帮助。但是由于 FACS 本身的限制(即除了只有 5 个 AU 有弱、中和强的强度外,AU 不能表达强度,而且在某个特定表情中,其最大强度都因人而异),它对计算机人工表情生成直接帮助不大。

综上所述,表情生成的关键是确定与表情密切相关的脸部特征,如眼、嘴等的位置变化与表情变化之间的映射关系,主观地操作 AU 并不能生成真实的表情。

Hongcheng Wang 和 Narendra Ahuja 则通过表情分解来建立表情、个人及面部特征之间的关系^[6],用高次奇异值分解(HOSVD)多线性生成法,从含有 7 种表情(他们认为有 7 种基本表情)的系列图像中建立表情次空间(Expression Subspace)和个人次空间(Person Subspace)。用这个次空间模型,既可以生成一个新人的新表情,也可以识别新人新表情;给定一个人,可以生成他的各种表情;给定一幅图像,可以识别表情或个人。这个算法中的表情没有量值变化,只能识别/模拟出高兴或是厌恶的表情,不能生成从微笑到大喜过望或从不喜欢到深恶痛绝等的表情。

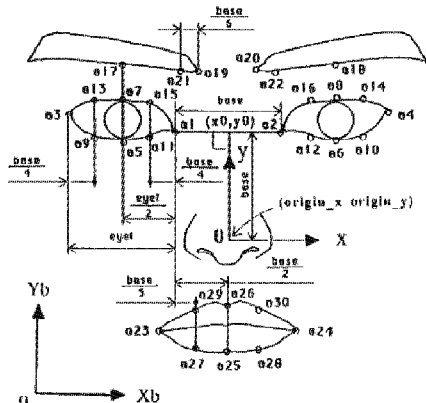


图1 H. Kobayashi 和 F. Hara 定义的脸部特征点集(FCPS)

定义1 标定点(Landmark),即根据脸部特征确定的点称为脸部特征点(Facial Characteristic Point, FCP)。它们全体的集合

叫脸部特征点集(Facial Characteristic Points Set, FCPS)。

定义2 漠无表情的脸称为正常脸,带有表情的脸称为表情脸。

H. Kobayashi 和 F. Hara 提出的表情识别方法^[1]要求用户用鼠标在输入图片上确定 30 个标定位置,这些标定点位于对表情有重要影响的部位,如眼、眉、嘴等的特定位置,即 FCPS,如图 1 所示。

比较输入图像中 FCPS 和正常脸上 FCPS 的位置变化,可以识别出高兴、悲伤、愤怒、恐惧、惊异和厌恶等 6 种基本表情及其强度。

该算法把图像中的 30 个 FCP 作为径向基函数神经网络(RBF Network)的 60 个输入。在这个神经网络的输出层上有 6 个节点,分别对应上述 6 种表情。

2 表情自动生成算法

2.1 基本思想

该算法的首要任务是建立表情强度与 FCPS 移动的映射关系。一旦建立了这种映射关系,表情自动生成便有了依据。众所周知,上述映射关系是非线性的,应用这种非线性映射关系来自动生成扭曲控制点的工具中,RBF 神经网络成为最佳首选。在表情自动生成过程中,输入表情强度矢量,就可使 RBF 网络输出相应的 FCPS。通过 FCPS 生成更多的点作为变形过程的扭曲控制点,控制正常脸图像扭曲过程,产生表情脸图像。

把正常脸图像扭曲变形为表情脸图像,采用了基于点的 RBF 的扭曲算法变形过程^[7],距离控制点越近的像素受控制点位置变化的影响越大,距离越远控制点的影响越小。这个算法不仅使控制点具有了局部性,还大大降低了控制点输入位置的精度要求。在本算法中,扭曲控制点是由 FCPS 自动生成的,必然存在一定的误差,而基于点的 RBF 的扭曲变形算法对控制点的精度要求低,所以不会造成扭曲结果的失真。该算法还可以用少量的控制点,达到其他扭曲算法(如网格算法和特征矢量算法等)需要很多控制点才能实现的效果。

一般的 RBF 神经网络训练有两个相互独立的任务:确定基函数中心以及修正连接权矢量。在本算法中,网络基函数中心不必训练,唯一的任务就是确定连接权矢量。采用递推正则最小二乘算法(RRLS)来训练网络,无需矩阵运算,不仅减小了计算量,同时还避免了矩阵求逆矩阵运算中存在的病态问题^[8]。

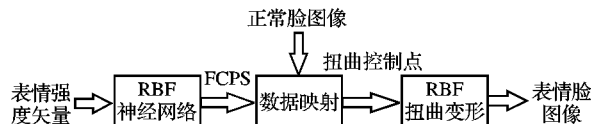


图2 表情自动生成系统结构

如图 2 所示,完成了某人正面面像的表情训练,即 FCPS 的移动与表情变化的映射关系被确立后,就可以生成表情了。输入 1 个表情矢量 $E(e_1, e_2, e_3, e_4, e_5, e_6)$, $e_i \in [0, 1]$,这 6 个分量分别是(happy, sad, angry, fear, surprised, disgusted),RBF 网络便输出相应的 FCPS。FCPS 是被标准化了的数据,不能直接应用到具体输入图像中,还要根据输入正常脸图像的实际大小和方向,映射成为扭曲控制点。变形过程对正常脸图像根据扭曲控制点进行扭曲变形,产生具有指定表情指定强度的人脸图像。

RBF 神经网络的结构如图 3 所示。每个隐含层神经元对

应一个中心,其径向基函数为高斯函数。表情矢量 E 被送到每个中心后,由径向基函数根据其为中心的欧氏距离作非线性变换,各中心通过连接权连接到输出层,把各隐含层神经元的量值线性组合的结果作为网络输出矢量,就是 FCPS。

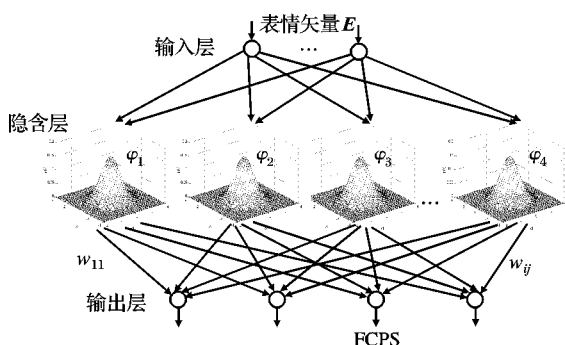


图3 RBF神经网络结构

2.2 控制点的选择

根据 RBF 扭曲算法的特点,扭曲控制点应该设置在表情变化时脸部发生位置改变的部分,变化越复杂的地方控制点越密集。既要保证一定的精度,又要尽可能地减少人工操作的工作量,如何选择扭曲控制点成为关注的焦点。

表情是脸部肌肉(肌群)活动的结果,脸部肌群的分布如图4所示。与表情有重大关系的眼区和口区肌肉分布较为密集且种类较多(有片状肌、条状肌和轮匝肌),应该设置较为密集的控制点,而其他部分肌肉不多,在表情变化时位置移动不大,可以少设点或不设点。

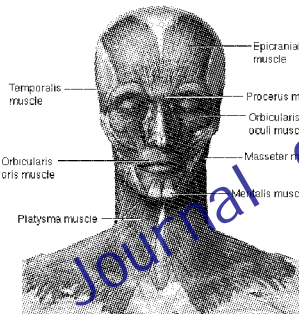


图4 脸部的主要肌群^[9]

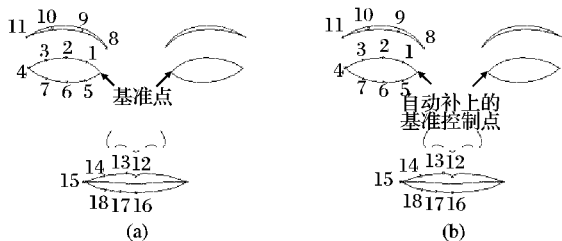


图5 FCPS和扭曲控制点

眉毛的运动受到前额片状肌和皱眉肌的影响,因而外侧主要作上下运动,而内侧则不仅作上下运动,还作横向移动,所以至少需要4个控制点。

上眼睑比下眼睑活动范围大,但考虑到在2D扭曲变形中,腮部的移动与下眼睑有很大的关系,所以上下眼睑上设置相同数目的控制点。

嘴的变化比眼睛更复杂,有开口与闭嘴的变化。我们的算法忽略了开口的变化,这样,我们只须关注口轮匝肌的活动就能满足要求。

鼻子上肌肉较少,对表情变化贡献不大,口部和下眼睑的控制

点足以影响鼻子的变化。事实上,真实的人脸中,鼻子的活动是由这些部分的活动带动的,鼻子自己活动的肌肉很小,可以忽略。

在选择扭曲控制点时,应该把握两个原则:一是控制点不宜过多,因为控制点是由 FCPS(其实质是标定点)通过数据映射生成的,过多的控制点就要求有更多的标定点,这不但会影响系统的运行速度,还会引入不必要的误差(因为标定点的选择和定位是手工操作,大量的标定点定位除了工作量大以外,还会因为照片的大小和方向差异,不可避免地引入误差);第二个原则是控制点必须尽可能地定位在肌肉群活动的特征点位置,对脸部活动影响大的地方应多设置控制点。考虑到本算法的扭曲变形采用 RBF 扭曲算法,活动的中心点必须设置点。根据这两条原则,本文设置了如图5(b)所示的36个扭曲控制点。

该算法最重要的特点在于简化人工操作。在训练网络时,并不需要把所有的控制点都作为训练样本,利用人脸的对称性,只需要确定一边的 FCP,就能很容易地推导出另外一侧的标定点。

为了进一步减少信息冗余,我们还可以减少网络输入。因为眼内角的表情变化时基本保存不变,所以我们可以以此为基准点,因此经过数据标准过程(详见下述)后,任何图像的这两点坐标都是 $(-0.5, 0)$ 和 $(0.5, 0)$ 。这样,我们就不必把它们输入网络进行训练了。

根据上述规则确定图5(a)的 FCPS,虽然只包含了18个基准点(2个基准点不用 RBF 网络处理),但比其他研究者(通常包含30个点)^[1]的 FCPS 表达了更多信息,也就是说这18点的信息冗余很小。

对于 RBF 神经网络而言,较少的扭曲标定点意味着在训练时可以减少病态数据对训练精度的影响。RBF 网络的连接权是通过把表情矢量与表情脸的 FCPS 作为样本对进行有监督学习而获得的。FCPS 的分量越多,训练中的线性方程组出现病态的可能性越大。

2.3 数据的标准化的

在训练网络时,为了矫正不同照片由于拍摄角度、拍摄距离以及拍摄姿势的不同而造成的失真,就必须对输入图像进行标准化。当表情变化时,人的两眼内侧眼角之间的距离基本保持不变,可以作为标准化的基本量。在标准化时,首先把坐标系的原点移到两眼之间的鼻根处,即两眼内侧眼角连线的中点处;然后设两眼内角连线的距离为 $base$,左边内眼角和右边内眼角的坐标分别为 (x_l, y_l) 和 (x_r, y_r) ,则 $base$ 可由如下(1)式计算出:

$$base = \sqrt{(x_l - x_r)^2 + (y_l - y_r)^2} \quad (1)$$

在实际应用时必须考虑图像中人物的脸是否倾斜,若两眼内侧眼角的连线相对于水平扫描线的倾角为 α 。图像上任意一点 P 在原坐标系中坐标为 (x_{old}, y_{old}) ,经坐标系旋转后的新坐标 (x_{new}, y_{new}) ,则:

$$\begin{cases} x_{new} = x_{old} \cos \alpha + y_{old} \sin \alpha \\ y_{new} = -x_{old} \sin \alpha + y_{old} \cos \alpha \end{cases} \quad (2)$$

在实际计算 α 时,若直接检测两眼内侧眼角的连线的倾角可能引入较大的误差,因为两眼内侧眼角比较靠近,在标定它们时如果在垂直方向上存在一个小误差,就会引入很大的误差。解决的方法是把的两眼内侧眼角连线的中点、FCP₁₂ 以及 FCP₁₆ 这3点拟合为一直线,通过这条直线的斜率来计算 α 。

最后把所有 FCP 坐标的单位都化为 *base* 的倍数形式。若 FCP_{*i*} (*i* = 1, 2, ..., 18) 的坐标为 (*x_i*, *y_i*), 则现在应该是 (*x_i*/*base*, *y_i*/*base*)。

2.4 数据映射过程

数据映射过程把 RBF 神经网络输出的 FCPS 映射为扭曲变形过程的扭曲控制点。数据映射过程有两个任务:还原数据标准化过程中被标准化了数据,这是数据标准化过程的逆过程;另一个任务是自动补全一些扭曲控制点。为了节省计算量,本文采取以下步骤:

1) 补上两眼内侧眼角控制点,位置分别在 (-0.5, 0) 和 (0.5, 0);

2) 以 *y* 轴为对称轴,生成另外一侧的 FCP,新点的坐标为 (-*x_i*, *y_i*), (*i* = 1, 2, ..., 18);

3) 旋转坐标系:

$$\begin{cases} x_{new} = x_{old} \cos \alpha - y_{old} \sin \alpha \\ y_{new} = x_{old} \sin \alpha + y_{old} \cos \alpha \end{cases} \quad (3)$$

4) 所有 FCPS 坐标值都乘以 *base*;

5) 把坐标系原点移到左下角。

这样就得到了如图 5(b) 所示的扭曲控制点。

3 实验结果

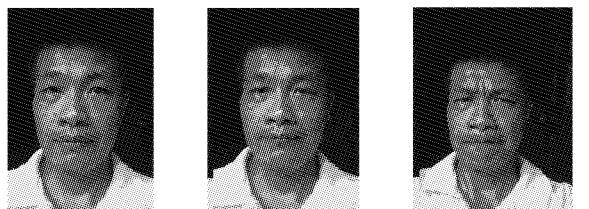
3.1 训练

本文选用了两个人的照片进行实验,一个是作者之一的照片,另一个是选自 Internet 照片。

为了检验泛化能力,本文把实验分为两组。在第一组中,分别将两个人的正常脸、中度高兴、极度高兴、中度悲伤、极度惊异、中度厌恶和极度厌恶共 13 幅照片作为 RBF 网络中心,而另一组则分别将两个人的正常脸、极度高兴、极度悲伤、极度愤怒、极度恐惧、极度惊异、和极度厌恶共 7 幅照片作为 RBF 网络中心。

3.2 生成实验

在 Windows XP 下用 VC6.0 实现这个算法,进行上述生成实验。图 6 是生成的部分表情照片。输入了表情矢量 $E(e_1, e_2, e_3, e_4, e_5, e_6)$, $e_i \in [0, 1]$, 系统便对给定的正常脸照片施加扭曲变形,产生相应表情。



(a) 正常脸 (b) 生成的0.6愤怒表情 (c) 输入的1.0愤怒照片



(a) 正常脸 (b) 生成的1.0惊异表情 (c) 输入的1.0惊异照片

图6 表情生成实验

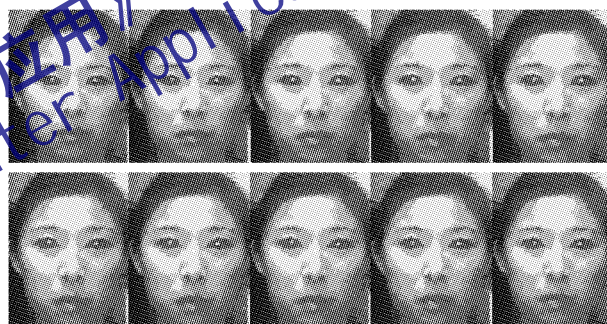
在生成的表情图像中,有明显特征变化。上图中,眉毛向内

聚拢,上眼睑呈三角形,下眼脸上抬,嘴唇压紧嘴角向外向下拉。下图中,眼睁开,嘴角向内压,上下唇分别向上下展开。但是,基于 2D 图像变形的表情生成不能处理皱纹和开口等图像。

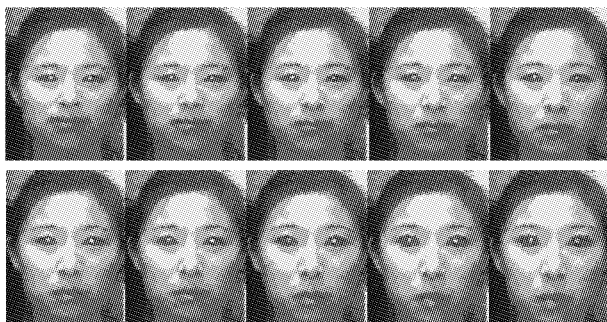
对 13 个中心与 7 个中心(中心信息详见表 1)的 RBF 神经网络用同样的样本进行训练。分别生成从 0.1 到 1.0 的 10 幅惊异表情,如图 7。

表1 两组实验的 RBF 网络中心

作为中心的表情矢量		说明
13 个中心	7 个中心	
(0.0, 0.0, 0.0, 0.0, 0.0, 0.0)	(0.0, 0.0, 0.0, 0.0, 0.0, 0.0)	neutron
(0.4, 0.0, 0.0, 0.0, 0.0, 0.0)	-	Happy 0.4
(0.9, 0.0, 0.0, 0.0, 0.0, 0.0)	(0.9, 0.0, 0.0, 0.0, 0.0, 0.0)	Happy 0.9
(0.0, 0.5, 0.0, 0.0, 0.0, 0.0)	-	Sad 0.5
(0.0, 1.0, 0.0, 0.0, 0.0, 0.0)	(0.0, 1.0, 0.0, 0.0, 0.0, 0.0)	Sad 1.0
(0.0, 0.0, 0.7, 0.0, 0.0, 0.0)	-	Angry 0.7
(0.0, 0.0, 0.9, 0.0, 0.0, 0.0)	(0.0, 0.0, 0.9, 0.0, 0.0, 0.0)	Angry 0.9
(0.0, 0.0, 0.0, 0.5, 0.0, 0.0)	-	Fear 0.5
(0.0, 0.0, 0.0, 0.8, 0.0, 0.0)	(0.0, 0.0, 0.0, 0.8, 0.0, 0.0)	Fear 0.8
(0.0, 0.0, 0.0, 0.0, 0.6, 0.0)	-	Surprised 0.6
(0.0, 0.0, 0.0, 0.0, 0.9, 0.0)	(0.0, 0.0, 0.0, 0.0, 0.9, 0.0)	Surprised 0.9
(0.0, 0.0, 0.0, 0.0, 0.0, 0.6)	-	Disgusted 0.6
(0.0, 0.0, 0.0, 0.0, 0.0, 0.9)	(0.0, 0.0, 0.0, 0.0, 0.0, 0.9)	Disgusted 0.9



(a) 7个中心RBF网络生成的表情序列
(自上到下, 自左到右分别是0.1到1.0的惊异表情)



(b) 13个中心RBF网络生成的表情序列
(自上到下, 自左到右分别是0.1到1.0的惊异表情)

图7 对比实验

3.3 不同中心数的分析

在视觉上,从 0.1 到 0.4 两个网络存在一定的差异,而 0.5 以上则两者没有明显的差异。尤其引人注目的是 0.6, 两幅表情几乎没有什么不同。而两组中心的设置在惊异这个表情上,正是在 0.6 的中心存在差别(详见表 1)。这说明 RBF 神经网络具有最大的泛化能力。

另一个令人关注的现象是输出结果的非线性特征。在 7 个中心的网络中(图 7(a)),五官位置呈现 U 型波动;同样地,在 13 个中心的这一组中(图 7(b)),可以明显地看出输出

以0.6为转折呈现~型波动现象,且波动性更为明显。

4 结语

本算法利用了人脸的对称性,输入一侧的标定点,可以有效地自动生成另一侧的标定点。这个方法可以极大地减小冗余信息的输入,提高效率。基于本文提出的从FCPS到扭曲控制点的自动数据映射方法,进一步减少了人工操作工作量,同时也保证了非常好的精度。

本算法是基于RBF神经网络的2D图像扭曲表情人工生成方法,有很好的泛化能力,只需要很少的样本进行训练。这样不仅对一幅图像的标定工作量小,而且因为只需少量的样本就能确立脸部活动与表情的映射关系,更是大规模地减少了人工操作的工作量。由于采用了对扭曲控制点误差敏感的RBF Warping算法,在定位FCPS时,人工操作变得更为轻松。比较图1与图5(a),后者(本算法的)FCPS不仅简洁,而且在定位时没有前者那样的苛刻要求。如果引入某些自动脸部特征(如眼睛)检测方法,甚至可以实现全自动的表情识别与生成。尽管目前这些自动检测法还存在着一定的误差,但只需采取适当的措施,把误差控制在一定范围,就有可能应用本算法来实现全自动人工表情生成。

综上所述,本算法只需少量的图像作为样本,每个样本只有较少的FCP需要定位,且每个FCP的定位精度要求不高,这几个方面都大大减轻了表情生成的人工操作工作强度并减少了计算机计算量,运行速度较快。

然而,基于2D图像扭曲的表情人工生成也存在一些问题,即这类方法无法解决脸部细节的变化问题,如皱眉,开口等。作为进一步研究,脸部特征的自动检测可以实现自动FCPS定位,另外,还应进一步研究如何产生逼真的人工褶皱,

如何产生人工开口,逼真的模拟出牙齿和舌头等因开口而显示的部位。

参考文献:

- [1] LEUNG MYY, HUNG YH, KING I. Facial Expression Synthesis by Radial Basis Function Network and Image Warping[A]. IEEE International Conference on Neural Networks[C], 1996, 3: 1400 - 1405.
- [2] ARAD N, DYN N, REISFELD D, et al. Image warping by radial basis function: Application to facial expressions[J]. CVGIP: Graphical Models and Image Processing, 1994, 56(2): 1161 - 172.
- [3] CHANG Y, HU CB, TURK M. Manifold of Facial Expression[A]. Proceeding of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures(AMFG'03)[C]. IEEE, 2003.
- [4] PANTIC M, ROTHKRANTZ LJM. Facial Action Recognition for Facial Expression Analysis From Static Face Images[J]. IEEE Transactions on Systems, Man, and Cybernetics - PART B: CYBERNETICS, 2004, 34(3).
- [5] ZHANG YM, JI Q. Facial Expression Understanding in Image Sequences Using Dynamic and Active Visual Information Fusion[A]. Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)[C], 2003.
- [6] WANG HC, AHUJA N. Facial Expression Decomposition[A]. Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03)[C], 2003.
- [7] GOMES J, DARSÁ L, COSTA B, et al. Warping and Morphing of Graphical Objects[M]. Morgan Kaufmann Publishers, 1998.
- [8] 王正欧, 朱涛, 王书新, 等. 基于RBF网络的高技术项目投资评估决策模型的研究[J]. 系统工程理论与实践, 2003, (3).
- [9] <http://www.thailabonline.com/muscle/face.htm> [EB/OL], 2004 - 10.

(上接第1607页)

实验还比较了利用文献[1]的方法计算外参的结果与利用本文优化算法计算得到的结果。图3中三个图分别为外参平移向量 t_x, t_y, t_z 的时间距离曲线。从图中可以看出,直接使用 H 计算出来的平移参数易受噪声影响,抖动较大,这一点在图3(c)中表现的尤为明显。而利用本文中优化算法得到的结果相对平滑。

本文利用OpenGL模拟摄像机投影过程,并实现了一个简单的增强现实系统(图4)。该系统可以将OpenGL绘制的虚拟物体绘制在视频中,该虚拟物体随“平板图案”做出相同的运动。图5演示了本文算法在“毛笔书法笔迹输入”系统中的一个应用。该系统使用本文中的平板图案定位算法来对纸张进行定位,并利用跟踪结果重现写字的全过程。

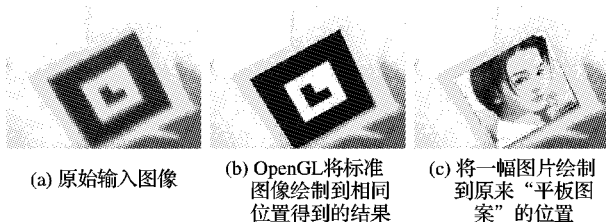


图4 利用OpenGL实现的增强现实结构

从实验结果可以看出,本文算法可以在“运动模糊”,“目标较小”,以及“部分遮挡”等复杂情况下对平板图案进行持续跟踪,计算摄像机参数结果准确、平滑,适用于基于摄像头

的增强现实应用。

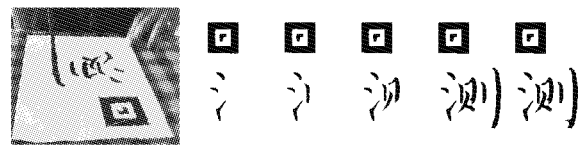


图5 毛笔书法笔迹输入

参考文献:

- [1] MALIK H, ROTH G, MCDONALD C. Robust 2D Tracking for Real-time Augmented Reality[A]. Proceedings of Vision Interface (VI)[C]. Calgary, Alberta, Canada, 2002.
- [2] ZHANG ZY. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330 - 1334.
- [3] 卢开澄. 图论及其应用[M]. 北京: 清华大学出版社, 1981.
- [4] SHI JB, TOMASI C. Good Features to Track[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C], 1994. 593 - 600.
- [5] BOUGUET J - Y. Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm[R]. Intel Corporation, 2000.
- [6] OpenCV, Open Source Computer Vision Library [CP/OL]. <http://www.intel.com/research/mrl/research/opencv/>, 2005 - 01 - 01.