

人机交互中的人脸信息曲波分析

吴贤星, 赵杰煜, 沈明炘

(宁波大学计算机科学与技术研究所, 宁波 315211)

摘要: 视觉交互是自然人机交互的重要组成部分,而人脸特征提取则是视觉交互成功的关键。针对小波变换难以充分描述人脸曲线特征的缺点,为了更好地提取人脸特征,将更符合人类视觉特性的曲波变换用于人脸信息处理,提出了结合曲波变换与 Adaboost 方法的人脸检测优化方法和基于曲波变换与 SVM 进行表情分析的新方法,并开展了人脸检测、人脸识别与表情分析的对比实验。实验结果显示,曲波变换在人脸特征提取中具有明显优势,从而为自然人机交互的下一步工作打下了坚实基础。

关键词: 曲波变换; 人脸检测; 人脸识别; 表情分析; 人机交互

中图分类号: TP391 **文献标志码:** A **文章编号:** 1006-8961(2010)-1309-09

Face processing in human-computer Interaction using curvelet analysis

WU Xianxing, ZHAO Jieyu, SHEN Mingxin

(Institute of Computer Science and Technology, Ningbo University, Ningbo 315211)

Abstract: Vision interaction is one of important aspects of human-computer interaction, and the facial feature extraction is crucial to vision interaction. This paper applies the curvelet transform to the face processing to extract facial feature more effectively. It overcomes the weakness of the wavelet transform which is unable to efficiently extract curve features of face images. An optimized method based on Adaboost and curvelet transform is proposed for face detection. A new approach combining SVM and curvelet transform is designed for facial expression recognition. Experiments on face detection, face recognition and facial expression recognition are carried out. The results reveal that curvelet transform has distinct advantages in facial feature extraction, and lays a good foundation for the further work of the natural human-computer interaction.

Keywords: curvelet transform; face detection; face recognition; facial expression recognition; human-computer interaction

0 引言

自计算机诞生以来,人机交互已经从早期的命令行交互阶段,图形用户界面阶段发展到自然人机交互阶段。

视觉交互是自然人机交互技术中重要的一环,统计表明,在人类接受的信息中,视觉信息占60%以上^[1]。众所周知,图像和视频是人类接受视觉信

息的主要媒介,而视觉信息中的人脸信息处理是非常重要的部分,因为人机交互中的主体是人。人脸信息处理涉及到的技术主要包括人脸检测、人脸识别、表情识别等,这些都是视觉交互的重要组成部分。由于人的脸部包含了丰富的个人信息,因此对这些信息进行有效利用,能够使得人机交互过程变得更加自然、和谐。而要获得这些信息,关键就在于特征提取,好的处理方法不但能明显提高识别率,而且能降低系统的运算量。

基金项目: 国家自然科学基金项目(NSFC-60672071);浙江省自然科学基金重大项目(D1080807)。

收稿日期: 2010-03-20; **改回日期:** 2010-04-20

第一作者简介: 吴贤星(1986—),男,宁波大学计算机应用技术专业硕士研究生。主要研究方向为计算机视觉、人机交互等。

E-mail: iwuxx@163.com。

在人脸特征提取方面,目前的研究主要集中在以下几个方面:

1) 面部几何结构特征。提取该类特征的方法主要有手动标点的方法、全自动提取特征点(如 SUSAN 角点)方法与基于面部轮廓特征的方法^[2]等。几何特征在早期的人脸提取中用得较多,但是由于其精确程度不够,且受各种外界因素影响较大,因此限制了其应用。

2) 统计特征。统计特征提取方法是通过对大量样本的训练来获得特征的统计特性的,其基本思想是将人脸图像映射到特征空间。代表性方法主要有基于主成分分析的方法(PCA)^[3]、基于线性判别分析(LDA)的方法^[4]与基于独立成分分析的方法^[5]。该类方法通常都能获得一个较低维度的子空间,由于数据维度较小,故识别速度较快。但该类方法的识别率易受外界因素,如光照、角度及人脸形状差异等的影响。为了减弱这些影响,可在特征提取前,先进行图像的预处理工作,比如尺度归一化、直方图均衡化等。

3) 变换域特征。这类特征是人脸最重要的特征之一,提取该类特征的方法主要是先将图像变换到频率域,再在频率域进行相应操作和特征提取。常见的方法有小波变换、Gabor 小波^[6]、离散余弦变换^[7]和曲波变换^[8]等。变换域特征由于可以对图像进行单独操作,因此训练识别都较为方便。小波变换与曲波变换都是符合人类视觉特性的方法,在人脸特征提取中得到了广泛应用。

4) 代数特征。常见的代数特征包括图像不变矩与奇异值特征^[9]等。

而作为提取图像变换域特征的小波变换,其影响十分巨大。在计算机视觉的各个方面均得到了广泛运用。

但小波变换并不是完美的,基于对人类视觉系统的研究和对图像的统计分析,理想的图像表示方法或者说特征提取方法应该满足多分辨率、局部化、临界采样、基函数方向性和各向异性 5 个条件^[10]。小波变换的基函数的支撑区间为正方形,由于其方向的有限性,致使小波变换只能够满足前 3 个条件,而对后 2 个条件则无能为力。

针对小波变换的这一缺点而发展起来的曲波变换,由于其具有多尺度、多方向性以及基函数各向异性方面的优势,能够很好地满足上面的 5 个条件,是更符合人类视觉特性的一种新的图像表示方法。近

年来,曲波变换在人机交互的视觉信息处理中得到了广泛的应用,比如人脸识别^[7-8]和手势识别等。

本文利用曲波变换来提取人脸特征,提出了一种结合曲波变换与支持向量机(SVM)的表情分析新方法。基于日本女性表情数据库(JAFFE)的人脸识别和表情识别实验以及在 Yale 数据库上的人脸识别实验均表明,曲波变换比小波变换在人脸特征提取上具有明显优势。在人脸检测方面,针对基于简单 Haar 小波特征的 Adaboost 算法的误检率较高的问题,本文给出了一种结合曲波变换的人脸检测优化方法,实验效果较好。

1 曲波变换

1.1 基本理论

曲波变换(curvelet transform)^[11]是多尺度几何分析方法或超小波(beyond wavelets)方法中的一种,由 Candès 和 Donoho 等人在 1999 年提出,其实质上是由脊波理论发展而来。2002 年,Donoho 等人提出了第 2 代曲波变换^[12],并于 2005 年提出了快速数值实现方法^[13-15]。相对于第 1 代曲波变换,第 2 代曲波变换具有概念更简洁、运算速度更快、冗余性更小等优点。第 2 代曲波变换引入了尺度、角度和位置 3 个参数,能够充分地各角度、尺度对图像的曲线特征进行提取。本文中采用的是第 2 代曲波变换。

曲波变换提出后,引起了学者的广泛兴趣,国内外学者做了很多相关研究,特别是 2005 年第 2 代曲波变换的数值实现方案的提出,更加推动了这一方法的发展与应用。如 Mandal 等人利用曲波变换提取出的特征与主成分分析结合用于人脸识别^[16],Aroussi 等人将曲波变换与线性判别分析等结合进行了人脸识别实验^[17],Zhang 等人提出了用曲波变换和小波变换相结合的人脸识别方法^[18],Majumdar 等人还将曲波变换应用到了手写体字符的识别中^[19]。

曲波变换的思想就是当对曲线进行分割时,将每一小段看作是直线段。在此基础上再对图像进行多尺度、多方向的曲波变换。

与小波变换类似,曲波变换的形式也是基函数与信号函数做内积的形式,于是曲波变换可表示为

$$c(j, l, k) := \langle f, \varphi_{j, l, k} \rangle \quad (1)$$

其中, $\varphi_{j, l, k}$ 表示曲波函数, j, l, k 分别表示尺度、方向

和位置参数。

曲波变换在频率域中实现,并采用频率域中的窗函数 U_j 来实现 φ 的表示。其定义如下:

$$U_j(r, \theta) = 2^{-3j/4} W(2^{-j}r) V\left(\frac{2^{\lfloor j/2 \rfloor}}{2\pi}\right) \quad (2)$$

其中 W 和 V 是满足如下条件:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1 \quad r \in \left(\frac{3}{4}, \frac{3}{2}\right) \quad (3)$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1 \quad t \in \left(-\frac{1}{2}, \frac{1}{2}\right) \quad (4)$$

的半径窗和角度窗。其中 $\lfloor j/2 \rfloor$ 表示对 $j/2$ 向下取整,由式(2)~(4)的定义可知,窗函数 U_j 即为极坐标下的一种“楔形”窗(如图1所示)。

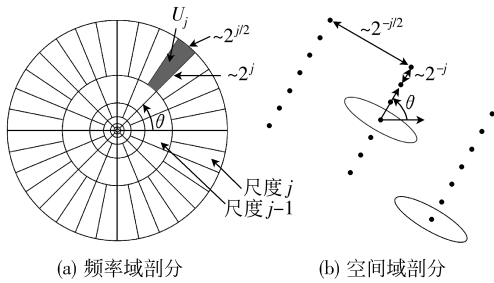


图1 曲波变换剖分示意图

Fig.1 Curvelet tiling of frequency and space domain

令该“楔形”窗上的曲波为 $\hat{\varphi}_j(x)$, 其 Fourier 变换为 $\hat{\varphi}_j(\omega) = U_j(\omega)$, 这个曲波称为母曲波, 知道了 j 尺度上的曲波变换, 那么其他 2^{-j} 尺度上的曲波均可通过旋转和平移来获得。定义在尺度为 2^{-j} 、方向为 θ_l 、平移参数为 k_1, k_2 的曲波函数为

$$\varphi_{j,l,k}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)})) \quad (5)$$

其中的 R_{θ_l} 是通过旋转 θ_l 而来, 表示旋转平移参数。

进而可得频率域中的连续曲波变换定义为

$$c(j, l, k) := \frac{1}{(2\pi)^2} \int \hat{f}(\omega) U_j(R_{\theta_l}\omega) \exp[i \langle x_k^{(j,l)}, \omega \rangle] d\omega \quad (6)$$

离散曲波变换的定义为

$$c^D(j, l, k) := \sum_{0 \leq t_1, t_2 < n} f[t_1, t_2] \overline{\varphi_{j,l,k}^D[t_1, t_2]} \quad (7)$$

推出:

$$c^D(j, l, k) = \int \hat{f}(S_{\theta_l}\omega) \tilde{U}_j(\omega) \exp[i \langle b, \omega \rangle] d\omega \quad (8)$$

其中 $\tilde{U}_j(\omega)$ 为笛卡尔坐标系下的局部窗, S_{θ_l} 为剪切

矩阵, b 是离散的位置参数。式(7)和式(8)中的上标“D”代表“离散”的意思。

1.2 实现方法

第2代曲波变换有2种数值实现方法,一种是基于非均匀采样的快速 Fourier 变换(unequally-spaced fast Fourier transforms),另一种是基于特殊选择的 Fourier 采样的 wrapping 算法(wrapping of specially selected Fourier samples)^[15]。第1种方法具体的实现步骤如下:

1) 对图像进行2维快速傅里叶变换,即可得到以下2维频域表示: $\hat{f}[n_1, n_2]$, $-\frac{n}{2} \leq n_1, n_2 \leq \frac{n}{2}$;

2) 在频率域,对由每一对 (j, l) (尺度,角度)组成的“楔形”窗进行重采样,即可得到采样值 $\hat{f}[n_1, n_2 - n_1 \tan \theta_l]$, $(n_1, n_2) \in P_j$, P_j 是图1中所示“楔形”支撑区间的点的位置集合;

3) 将内插后的采样值与窗函数相乘即得到

$$\tilde{f}_{j,l}[n_1, n_2] = \hat{f}[n_1, n_2 - n_1 \tan \theta_l] \tilde{U}_j[n_1, n_2] \quad (9)$$

4) 再对其进行2维快速傅里叶逆变换,即得到离散的曲波系数集合。

wrapping 算法就是在上述算法的第3步之后加上 wrapping 步骤。通过上面两种方法可以实现对图像的曲波变换并得到曲波系数。

2 基于曲波变换的人脸信息处理

2.1 结合曲波变换的人脸检测

检测是视觉交互的第1步,其效果的好坏将直接影响到交互效果。在人脸检测方面,主要有基于肤色的检测方法、基于形状的检测方法和基于特征的检测方法3类。基于特征的检测方法方面具有里程碑意义的工作是由 Viola 和 Jones 两位学者提出的基于简单 Haar 特征的 Adaboost 算法^[20],该算法采用的是简单的 Haar 小波特征,并结合 Adaboost 算法。该方法具有鲁棒性强、速度快等特点,在人机交互的人脸检测中是现在用得最广泛的方法之一。由于其采用的是简单的 Haar 特征,速度虽然很快,但误检率也比较高,经常将一个不是人脸的区域误判为人脸。为了克服这一缺点,本文对该算法进行了优化,即在利用 Adaboost 算法得到候选人脸区域后,进行曲波变换并将候选区域与曲波特征脸进行比对,在设定的相似度阈值内则判定为人脸,否则判

定为非人脸。该算法将 Haar 小波在分析图像“点奇异”性上的优势和曲波在提取“线”的特征上的优势进行了有机结合,从而提高了算法的正确率。

2.1.1 曲波特征脸向量的构建

由于只需要判断候选区域是不是人脸,所以只要构建一个由较多人脸样本构成的平均脸作为对比样本即可。由于曲波变换在提取曲线特征上具有优势,而且系数稀疏,故很适合用来做构建算法。人脸样本包括 Yale 数据库、ORL 数据库和 JAFFE 数据库的部分样本以及部分自建的人脸样本,总共 724 张。

在曲波特征脸向量的构建中,先将所有图片缩放到 50×50 pixels。为了降低数据维度,算法只提取第 1 层系数,第 1 层曲波系数是低频系数,其代表了图像概貌。特征脸构建过程如下:

1) 先对所有样本图片的人脸区域进行提取,并缩放至 50×50 pixels,再进行一定的预处理,如直方图均衡化等;

2) 求取所有人脸样本的平均人脸图像;

3) 对平均脸进行曲波变换,提取第 1 层系数,并将其归一化到一个行向量,这个向量即为曲波特征脸向量。

构建出的平均脸如图 2 所示。



图 2 平均脸

Fig. 2 Average face

2.1.2 算法与实验

结合 Haar 小波与曲波变换的算法流程图如图 3 所示。

利用基于 Haar 小波特征的 Adaboost 方法检测到人脸后,先将人脸区域归一化到 50×50 pixels,并进行预处理;然后进行曲波变换抽取第 1 层系数,再将第 1 层系数转成行向量;最后将其和曲波特征脸向量进行比较,相似度可用两个向量之间的夹角的余弦来计算,即

$$S(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^m x_i y_i}{\sqrt{\sum_{i=1}^m x_i^2} \sqrt{\sum_{i=1}^m y_i^2}} \quad (10)$$

其中 \mathbf{X} 为提取到的人脸区域曲波特征向量, \mathbf{Y} 为曲

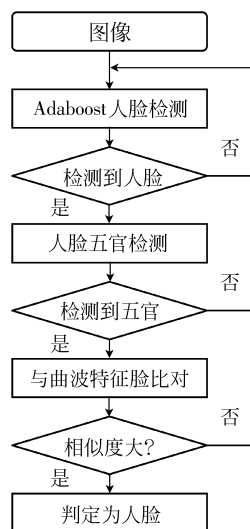
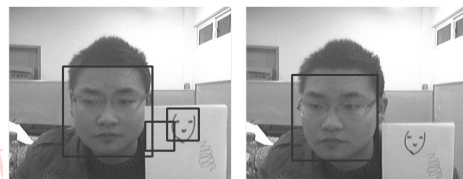


图 3 结合曲波变换的人脸检测算法流程图

Fig. 3 Flowchart of proposed face detection algorithm

波特征脸向量。实验中,相似度阈值设定为 0.75。经过基于曲波的检测优化后,就能够剔除大部分的误检人脸,效果较理想。为了进一步提高正确率,实验中还加入了对候选人脸区域进行人脸五官(如眼睛、鼻子和嘴巴)的检测。由于本方法具有一定的计算量,对处理速度有所影响,实验结果如图 4 所示。



(a) 原始 Adaboost 检测算法

(b) 本文检测算法

图 4 不同算法人脸检测结果

Fig. 4 Face detection results using different algorithms

2.2 基于曲波变换的人脸识别

人脸图像是包含曲线、直线等不同线特征较多的图像,从理论上讲,曲波变换应该比小波变换能够更好地提取人脸特征。国内外有很多学者对将曲波变换用于人脸识别进行了很多研究,并取得了不错的效果。本文将曲波变换、PCA 方法和 SVM 分类器进行结合,并开展了人脸识别实验。实验选择利用的人脸数据库是 Yale 数据库和 JAFFE 数据库。JAFFE 数据库有 10 人,每人有 21 幅左右的不同表情图像,共有 213 幅。实验中,将 JAFFE 数据库中

每个人的9张图像作为训练样本,剩余的图像作为测试样本。Yale数据库有15人,每人11张共165张人脸图像。实验中,将Yale数据库中每个人的前5张图像作为训练样本,将后6张图像作为测试样本,先将两个数据库人脸图像的人脸区域分割出来,并归一化到 100×100 pixels。

实验步骤如下:

- 1) 对分割出来的人脸图像进行曲波变换;
- 2) 提取曲波系数,并将曲波变换后的各层系数矩阵转化成行向量(不包括最后一层);
- 3) 对由各样本系数行向量组成的样本矩阵进行主成分分析,即得到映射矩阵和各训练样本在主成分空间的坐标;
- 4) 将得到的训练图片坐标信息输入到SVM进行训练;
- 5) 对测试图片进行曲波变换,提取曲波系数并求出坐标;
- 6) 将测试图片的坐标数据输入到SVM进行分类识别。

2.2.1 曲波变换参数的选择

由于曲波变换是对图像进行多尺度、多方向的变换,所以经过曲波变换后得到的是各层、各个角度的曲波系数。在此实验中,层数 s 的确定是根据文献[15]的以下公式: $s = [\log(\min(m, n)) - 3]$,其中 m, n 为图像的尺寸。

对于方向的选择,为了降低数据维度,本文实验中,对第2层方向数设定为8,第3层方向数设定为16。图像经过曲波变换后的第1层曲波系数是低频系数,代表了图像的概貌。最后1层是高频系数,体现了图像的细节。第2层有8个方向的曲波系数,第3层有16个方向的曲波系数,这两层是中高频系数,主要包含了图像的边缘特征。第1层和第4层无须设定方向。曲波系数示例及变换效果如图5所示(样本图片来源JAFPE数据库,为了方便观看,图片经过放缩)。

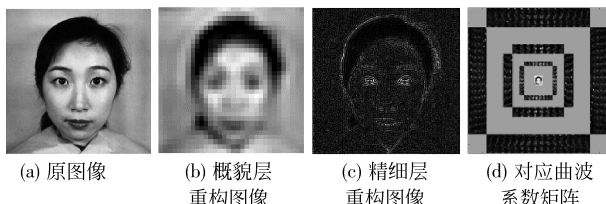


图5 曲波变换示例

Fig. 5 An example of curvelet transform

2.2.2 曲波系数的主成分分析

对于一个分辨率为 100×100 pixels的图像,经过曲波变换后,得到的曲波变换系数的数量是较大的,因此算法首先将每一层的曲波系数矩阵变形成行向量,第1层是289维,第2层是3960维,第3层是14416维;然后再连接成一个行向量,维数是18665维。每个图像变换后的系数都进行这样的操作。

如果将这些系数全部作为特征值提取出来进行识别实验是不科学的,不但计算量大,而且影响识别率,所以要进行一定的数据降维。本文选取主流的PCA方法作为降维方法,PCA方法能够有效降低数据维度,并且有利于识别率的提高。

2.2.3 分类器的选择

SVM克服了神经网络中可能遇到的易陷入局部极小和过学习问题,是统计学习理论中新发展起来的模式识别算法。SVM对于样本量较少、训练和测试样本较均衡的情况的分类效果较好,实验数据库中的人脸样本也较为均衡,故在本实验中选取了SVM作为分类器。核函数的选择对于SVM也十分重要,常用的有多项式核、Sigmoid核和径向基核等,本文实验选取径向基核。

2.2.4 实验结果

JAFPE数据库上的人脸识别结果及比较如表1所示,Yale数据库上的人脸识别结果如表2所示,在Yale数据库上利用曲波各层的系数分别进行识别实验的结果如表3所示。

表1 JAFPE数据库中的人脸识别结果比较

Tab. 1 Comparative study on JAFPE database

方法	识别率/%
小波变换 + k - NN ^[18]	92.5
曲波变换 + k - NN ^[18]	94.5
本文算法	99.2

表2 Yale数据库中的人脸识别结果比较

Tab. 2 A comparison of the recognition rate on Yale database

方法	识别率/%
小波变换 + PCA ^[21]	84.0
曲波变换 + PCA + k - NN ^[16]	83.9
本文算法	93.3

表 3 曲波变换各层识别结果 (Yale 数据库)

Tab.3 Recognition results on each level (Yale database)

系数层	识别率/%
第 1 层(概貌层)	86.7
第 2 层(细节层 1)	81.7
第 3 层(细节层 2)	76.7
第 4 层(精细层)	65.6
本文算法	93.3

由表 1—表 3 可以看出,通过曲波变换后,提取出的特征提高了识别率。本文方法在识别率上比文献[16]和文献[21]的方法均提高了 9% 左右,效果明显。曲波变换由于善于提取图像线的特征(如直线、曲线),因而在人脸特征的提取上展示出了较大的优势。分别利用曲波变换各层的系数进行的识别实验表明,概貌层识别率最高,而随着图像曲波变换系数层频率的提高,识别率反而有所下降(表 3),这也符合人类的视觉特性。因为人类识别一个人,并不用通过很精细的细节来实现,而是通过一个概貌就能大致上识别。

2.3 基于曲波变换的表情识别

表情分析是视觉交互中不可或缺的部分,也是人机交互技术中的难点之一。

人类的表情丰富,而且各种表情包含大量的曲线特征,对这些特征进行有效提取无疑将有利于表情识别率的提高。本文将曲波变换用于人脸表情特征的提取,提出了一种结合曲波变换与 SVM 的人脸表情识别新方法,算法步骤如下:

1) 先对数据库图片进行人脸检测,并归一化到 100×100 pixels, 然后进行直方图均衡化;

2) 再对预处理后的人脸图片进行曲波变换,提取各层曲波系数,将各层系数组成一个行向量(不包括最后一层);

3) 对提取到的训练图片系数行向量进行主成分分析,即得到主成分空间坐标及映射关系;

4) 将训练数据输入 SVM 进行训练,并建立 SVM 模型;

5) 提取待识别图像的曲波系数,并转换到主成分空间,然后将其坐标输入 SVM 进行识别。

2.3.1 算法参数设定

由于表情识别中的输入图像还是人脸图像,并且图像大小和人脸识别实验一样,所以曲波变换参数的选择还是和人脸识别实验一样,层数按上一小节中的公式选取,各层方向数亦和上一小节一样。

PCA 降维中的能量取 100%。SVM 分类器核函数选径向基核。

2.3.2 实验

实验是在 JAFFE 数据库上进行的,该表情库共有 10 个人的 7 种表情图像,每种表情有 3~4 张图片,共 213 张,JAFFE 数据库的示例图片及预处理后的图片如图 6 所示。



(a) JAFFE 数据库部分图像



(b) 预处理后的图像

图 6 JAFFE 数据库部分图片与预处理后的图片

Fig.6 Sample images from JAFFE database and pre-processed images

本文进行了两种方式的识别实验。实验方式如表 4 所示,实验结果如表 5 所示,各类表情的识别率及对比如表 6 所示。

表 4 实验方式

Tab.4 Experiment methods

实验方式	方式 1	方式 2
每人训练样本数	1~2	2
每人测试样本数	2	1~2
总训练样本	76	137
总测试样本	137	76

表 5 实验结果对比

Tab.5 Comparative study

方法	识别率/%
Gabor + PCA + k - NN ^[22]	90.7
Gabor + FUSION + 神经网络 ^[23]	92.6
本文实验方式 1	88.3
本文实验方式 2	94.7

由于人的表情变化具备较多的曲线特征,因此用曲波变换来提取特征无疑更加适合。从表 5,表 6 可以看出,基于曲波变换的识别率比传统的 Gabor 小波 + PCA 方法高出 4% 左右,比 Gabor 小波融合方法高出 2%,效果较明显,这说明曲波变换用于表情识别是适宜的、科学的。实验再次证明了曲波变换在人脸特征提取中的优势。

表 6 各类表情的识别率对比

Tab.6 A comparative study of different expressions recognition

表情	识别率/%		
	Gabor + PCA ^[23]	Gabor + FUSION ^[23]	本文实验 2
愤怒	60	95	100
厌恶	67	88	90.91
害怕	80	100	91.67
高兴	56	100	91.67
中性	67	75	100
悲伤	56	90	90.91
惊奇	89	100	100
平均	67.85	92.57	94.74

注:文献[22-23]采用的实验方式均为方式 2。

3 视觉交互实验

3.1 虚拟人自然交互平台介绍

本文的视觉交互实验是在虚拟人平台上进行的,智能 3 维虚拟人平台由宁波大学计算机科学与技术研究所研发,该虚拟人具备中英文发音、语音识别、3 维表情动作展现和视觉信息处理等功能。虚拟人主要由动作模块、视觉模块和听觉模块 3 个模块组成。其中动作模块主要负责 3 维虚拟人模型的表情动作、说话时的口型动作和视觉跟踪时的眼睛动作等;视觉模块主要负责虚拟人的视觉信息处理,比如人脸检测、人脸识别、表情分析、简单场景感知和视觉跟踪等;听觉模块主要负责听觉的模拟实现,现在主要有基于问答库的语音问答和中英文发音等功能。该平台最大的特色就是将多通道交互进行了有机融合,可以接受语音,视觉和文字等交互信息输入,由于虚拟人通过感知多通道交互信息,可展现不同的状态、动作,从而使得虚拟人具备一定的自然交互能力,其内部运行机制示意图如图 7 所示。

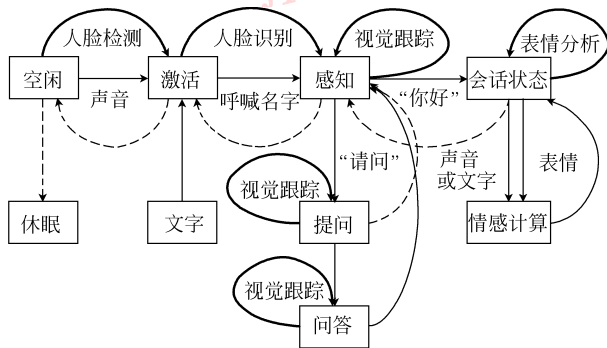


图 7 虚拟人内部运行机制示意图

Fig.7 The illustration of virtual human's operational mechanism

图 7 中的虚线表示“超时”事件,加粗的曲线表示该部分是视觉交互的内容。虚拟人通过接收和处理各种交互信息(如视觉信息、语音和文字等)来进行状态的切换。虚拟人开始的时候处在“空闲”状态,当检测到人脸或较大声音后则进入“激活”状态,然后通过人脸识别或直接语音呼叫虚拟人名字就可以切换到用户所喜好的 3 维人头模型,模型示例如图 8 所示。



图 8 3 维人头模型示例

Fig.8 Virtual human's 3D head models

3 维人头模型是主要的信息输出窗口和交互界面。人脸识别或语音识别成功后,虚拟人先进入“感知”状态,然后虚拟人通过语音识别“你好”和“请问”来进入不同的交互状态。虚拟人处在“感知”、“提问”和“回答”状态时,会进行视觉跟踪。视觉跟踪时,虚拟人的眼球将会跟随运动目标的移动而转动,造成虚拟人“盯着你看”的逼真效果。进入会话状态后,虚拟人将对 7 种基本表情进行分析识别,然后将识别结果和文字或声音等信息将输入情感计算模块进行情感计算,计算结果将通过表情动作、语音和文字等多方面形式展现出来。由于用户可与虚拟人进行语音交互、视觉交互和情感交互,而且这些交互并不是独立的,而是融合在一起的,从而使整个交互过程变得自然。

3.2 视觉交互实验

虚拟人视觉机制中的相关人脸信息处理均利用本文算法实现。在 CPU 频率为 2.7GHz,内存为 2GB 的电脑上实验,对于 640 × 480pixels 的视频流,检测、识别时间均小于 45ms,可见算法能够满足实时的要求。图 9 中的虚拟人处于“空闲”状态(等待交互),没有“表情”,也不“说话”。虚拟人在检测到人脸并进行人脸识别后,会以文字和声音的形式向交互者问好,并展现相应的口型动作(如图 10 所示)。

如果用户注册时有记录喜好的人头模型,那么

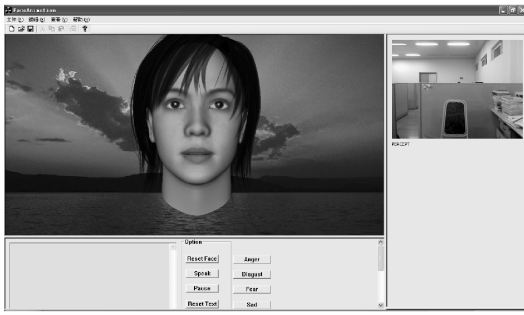


图 9 “空闲”状态下的虚拟人
Fig. 9 Virtual human under “idle” state



图 10 “感知”状态下的虚拟人
Fig. 10 Virtual human under “perception” state

虚拟人则会切换到对应的模型(如图 11 所示)。这样只要用户注册过,那么每次都可以与自己喜爱的虚拟人进行交互。虚拟人在进行视觉跟踪时,眼睛会实时盯着跟踪目标(如图 12 所示),这样设计的目的是为了能够更好地模拟人的视觉。而表情分析则是和情感计算结合在一起的,将表情分析的结果(7 种基本表情)输入情感计算模块进行计算,使得虚拟人具备一定的情感感知能力,甚至具有情感(如



图 11 人脸识别成功后可以切换到用户喜爱的模型
Fig. 11 Virtual human’s model will switch to user’s favor one after face recognition

图 13 所示)。由于情感计算不属于本文主题的范围,故本文不作进一步讨论。通过加入视觉机制,就能够使虚拟人和人们的交互更加自然、和谐。



图 12 虚拟人对视觉对象的跟踪
Fig. 12 Object tracking of virtual human



图 13 虚拟人通过分析交互者的表情来进行情感交互
Fig. 13 Virtual human will do emotion interaction by analyzing human’s facial expression

4 结 论

视觉交互是自然人机交互中的重要内容,而人脸特征提取则是视觉交互中不可忽视的一个方面。针对小波变换只能有效提取图像“点奇异”特征的缺陷,将更符合人类视觉特性的曲波变换引入到人机交互的人脸信息处理中。通过实验可以看出,曲波变换在这方面具有较大的优势和潜力。本文的主要工作是提出了一种结合曲波变换的人脸检测优化方法和一种基于曲波变换的表情识别新方法,并开展了人脸检测、人脸识别、表情分析的对比实验和在虚拟人平台的视觉交互实验。实验结果表明,曲波变换在人脸特征提取上比小波变换具有明显优势,识别率平均可提高 7% 左右。下一步的工作主要是研究曲波变换和其他算法的结合,并设计更加自然

和谐的视觉交互机制,针对曲波变换算法本身不足的改进也是下一步的研究重点之一。虽然曲波变换还处于初步发展中,但随着理论的进一步完善,必将在人机交互的人脸信息处理中发挥更大的作用。

参考文献 (References)

- [1] Zhang Youwei. Human-Computer Nature Interaction [M]. Beijing: National Defense Industry Press, 2004. [张有为. 人机自然交互 [M],北京:国防工业出版社,2004.]
- [2] Pardas M, Losada M. Facial parameter extraction system based on active contours [C]//Proceedings of IEEE 2001 International Conference on Image Processing. Washington, DC, USA: IEEE Computer Society, 2001, 1:1058-1061.
- [3] Turk M, Pentland A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, 3(1) :71-86.
- [4] Chellappa R, Eremad K. Discriminant analysis for recognition of human face images [J]. Journal of Optical Society of American, 1997, 8(14) :1724-1733.
- [5] Comon P. Independent component analysis, a new concept? [J]. Signal Processing, 1994, 36(3) :287-314.
- [6] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: A literature survey [J]. ACM Computer Survey, 2003, 35(4) :399-458.
- [7] Hafeed Z M, Levine M D. Face recognition using the discrete cosine transform [J]. International Journal of Computer Vision Special issue: Research at McGill University, 2001, 43(3) : 167-188.
- [8] Mandal T, Majumdar A, Wu Q M J. Face recognition by curvelet based feature extraction [C]//Lecture Notes in Computer Science, Berlin, Germany: Springer-Verlag, 2007, 4633: 806-817.
- [9] Hong Z Q. Algebraic feature extraction of image for recognition [J]. Pattern Recognition, 1991, 24(3) :211-219.
- [10] Majumdar A, Bhattacharya A. A comparative study in wavelets, curvelets and contourlets as feature sets for pattern recognition [J]. The International Arab Journal of Information Technology, 2009, 6(1) :47-51.
- [11] Candès E J, Donoho D L. Curvelets-a surprisingly effective nonadaptive representation for objects with edges [C]//Curves and Surfaces. Nashville, TN, USA: Vanderbilt University Press, 2000:105-120.
- [12] Candès E J, Donoho D L. New tight frames of curvelets and optimal representations of objects with piecewise- c^2 singularities [J]. Communications on Pure and Applied Mathematics, 2004, 57(2) : 219-266.
- [13] Candès E J, Donoho D L. Continuous curvelet transform: I. resolution of the wavefront set [J]. Applied and Computational Harmonic Analysis, 2005, 19(2) :162-197.
- [14] Candès E J, Donoho D L. Continuous curvelet transform II. discretization and frames [J]. Applied and Computational Harmonic Analysis, 2005, 19(2) :198-222.
- [15] Candès E J, Demanet L, Donoho D L, et al. Fast discrete curvelet transforms [J]. Multiscale Modeling and Simulation, 2006, 5(3) :861-899.
- [16] Mandal T, Wu Q M J, Yuan Yuan. Curvelet based face recognition via dimension reduction [J]. Signal Processing, 2009, 89(12) :2345-2353.
- [17] Aroussi M El, Ghouzali S, Hassouni M El, et al. Curvelet-based feature extraction with B-LDA for face recognition [C]//Proceedings of IEEE International Conference on Computer Systems and Applications, Washington. DC, USA: IEEE Computer Society, 2009:444-448.
- [18] Zhang Zhong, Wang Guanghui, Lin Xiang, et al, Face recognition based on wavelet-curvelet-fractal technique [C]//Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2009, 5627:532-540.
- [19] Majumdar A, Chaudhuri B B. Curvelet-based multi SVM recognizer for offline handwritten Bangla: a major Indian script [C]//Proceedings of the IEEE Ninth International Conference on Document Analysis and Recognition. Washington, DC, USA: IEEE Computer Society, 2007, 1:491-495.
- [20] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE Computer Society, 2001, 1:511-518.
- [21] Feng G C, Yuen P C, Dai D Q. Human face recognition using PCA on wavelet subband [J]. Journal of Electronic Imaging, 2000, 9(2) :226-233.
- [22] Deng Hongbo, Jin Lianwen. Facial expression recognition based on local gabor filter Band and PCA + LDA [J]. Journal of Image and Graphics, 2007, 12(2) :321-329. [邓洪波,金连文. 一种基于局部 Gabor 滤波器组及 PCA + LDA 的人脸表情识别方法 [J],中国图象图形学报,2007,12(2) :321-329.]
- [23] Liu Weifeng, Wang Zengfu. Facial expression recognition based on fusion of multiple gabor features [C]//Proceedings of the IEEE 18th International Conference on Pattern Recognition. Washington, DC, USA: IEEE Computer Society, 2006, 3: 536-539.