利用形状估计的人脸特征点定位算法*

仲柔在，熊　磊，刘　畅
（空军工程大学 航空航天工程学院，西安 710038）

摘　要： 针对鲁棒级联姿势回归算法（robust cascaded pose regression，RCPR）缺乏形状约束，对复杂人脸图像的定位精度差、成功率低的问题，提出一种利用形状估计的分块特征点定位算法。为提高定位成功率和准确度，对人脸特征点进行分块，对每一块进行形状估计作为约束；为保证形状估计的精度和连续性，在传统核回归的基础上，学习得到图像特征与目标形状间的联合概率分布函数，称做匹配函数，并求取最大值作为形状估计；为提高算法性能，只需对部分点的位置进行回归，减少了回归器的数量，并引入了形状索引特征的采样先验。实验表明，算法对复杂人脸图像具有更高的定位准确度和鲁棒性，定位成功率可达 86%，同时计算速度可以实现实时处理。

关键词： 核回归；形状估计；人脸特征点定位；采样先验；形状约束
中图分类号： TP391.4　　**文献标志码：** A　　**文章编号：** 1001-3695(2017)07-2207-05
doi:10.3969/j.issn.1001-3695.2017.07.062

Facial landmarks localization based on pose estimation

Zhong Rouzai，Xiong Lei，Liu Chang
(*Institute of Aeronautics & Astronautics Engineering*，*Air Force Engineering University*，*Xi' an* 710038，*China*)

Abstract：To improve the localization accuracy and success rate of complex facial image which lack of shape constraint in RCPR algorithm，this paper proposed a novel part-based cascaded regression algorithm combining shape estimation. Firstly，it divided the facial landmarks into several areas and used the shape estimation as shape constraint. Secondly，in order to ensure the accuracy and continuity of the shape estimation，it learned a joint probability distribution function of image features and target shapes，known as matching function. The estimated shape of the target could be obtained by maximizing the matching function. Thirdly，to improve algorithm performance，it exerted a prior constraint of the distance between the pixels pair during sampling shape-indexed features，meanwhile it reduced the number of regressors. Experiments demonstrate that the algorithm outperforms in localization accuracy and achieves a better robustness for occlusion and complex facial image，a localization success rate of 86% is obtained and can realize real-time processing.

Key words：kernel regression；shape estimation；facial landmarks localization；sampling prior；shape constraint

0　引言

人脸特征点定位是人脸图像分析过程中极为关键的一个环节，为人脸图像的深入处理与分析提供了重要的前提，在人脸识别[1]、人脸跟踪[2]、表情分类[3]等领域内有着广泛的应用，具有重要的研究意义。

近年来回归技术在其他模式识别领域的成功应用开辟了新思路，可以将人脸特征点定位看做是一个回归问题，直接对人脸图像的外观进行投影学习回归函数，从而产生目标输出[4~12]。但传统的回归方法并未强调形状约束，只是确保初始形状符合人脸形状约束。对于复杂人脸图像，如图 1 所示，眼睛、眉毛和嘴巴部分因为表情（挑眉、眯眼、大笑）或者是遮挡（头发、眼镜等）会造成特征点定位偏差，超出人脸合理形状范围，因此选择添加约束。对整张人脸进行估计，只能判断出人脸是否旋转，不能了解每个动作导致的五官的具体变化，无法解决问题。因此本文选择对人脸特征点进行分块，对各个部分的形状分别进行估计并作为约束，进行下一步的定位任务。

图1　目标复杂人脸图像

目前，形状估计的方法分为：

a）基于模型的方法[13,14]。利用几何模型来表示物体的形状和结构，并通过提取特征，在模型与图像之间建立对应关系，从而实现形状估计。这种方法需要匹配多个局部特征，因此受光照、噪声等影响无法准确提取局部特征时鲁棒性较差。

b）基于学习的方法[15,16]。根据事先标定的不同形状下的训练样本学习对应关系，将学习得到的回归函数应用于测试，所得结果作为对应目标的形状估计。该方法采用全局特征，具有较好的鲁棒性，然而其只能将形状划分到事先定义的几个类别中，不能保证估计的精度和连续性。

为解决上述问题，本文提出了一种基于特征形状核回归模

收稿日期： 2016-05-04；**修回日期：** 2016-06-27　　**基金项目：** 国家自然科学基金资助项目（61379104，61372167）

作者简介： 仲柔在（1992-），男，江苏如皋人，硕士研究生，主要研究方向为图像处理、模式识别（zhongrouzai@163.com）；熊磊（1976-），男，江西南昌人，副教授，博士，主要研究方向为图像处理、模式识别；刘畅（1992-），男，山东济南人，硕士研究生，主要研究方向为图像处理、模式识别.

型的估计方法。通过训练,学习得到输入变量(图像特征)与输出变量(形状)的联合概率分布函数,称为匹配函数。进而通过求取匹配函数在输出空间的最大值,得到目标的形状估计。同时改进级联回归算法,将形状估计作为约束,并对特征采样添加先验信息,进行特征点的精确定位。

1　利用形状估计的级联回归算法

1.1　传统级联回归算法

为了使下文对于该方法的介绍更加明确,首先定义算法的基本概念。令 $S=[x_1,y_1,\cdots,x_L,y_L]^{\mathrm{T}}$ 为被拉伸为列向量的、在人脸图像 I 中的 L 个特征点的 (x,y) 坐标,即人脸形状。令 S^j 表示人脸形状 S 的当前估计(S^0 即为初始人脸形状),S_i 表示人脸图像数据库中对人脸形状的基准标定(面部特征点真实位置),$\bar{S}$ 表示训练样本集中 S_i 的平均值。

在回归之前,首先对训练样本中的人脸形状进行预处理,将当前的输入形状 S^j 投影到 $\bar{S}$,即对 S^j 进行一个相似变换,从而来确保在整个回归过程中人脸的形状不变性。这样一个相似变换可以通过最小化两者之间的绝对距离来实现。

$$|\bar{S}-(S_i^j R_i^j S^j+T_i^j)| \tag{1}$$

其中:S_i^j、R_i^j、T_i^j 分别表示所求相似变换的尺度变换矩阵、旋转变换矩阵和平移变换矩阵。

得到这样一个标准化的形状后,就可以开始采用递进式的回归策略,如算法 1 所示,训练 N 个有顺序的树结构的弱回归器 $R=(R^1,\cdots,R^N)$ 。

算法 1　级联回归的计算流程

输入:I,S

for $i=1$ to T do

$x^j=h^j(I,S^{j-1})$//计算形状索引特征

$\delta S=R^j(x^j)$　　//计算形状增量

$S^j=S^{j-1}+\delta S$　　//更新当前形状

end for

输出:最终形状 S^N

在回归的每个阶段中,对应的树状回归器 R^j 的每一个叶子节点的输出是一个形状增量 $\Delta S^j=R^j(x^j)$,其中 $x^j=h^j(I,S^{j-1})$ 表示形状索引特征。再将所得的形状增量 ΔS^j 与上一次迭代所得形状估计 S^{j-1} 相加,得到更新后的新形状 S^j 作为当前的形状估计。给定粗略的初始形状估计 S^0,最终对于人脸形状的估计 S^N 就可以依照上述方式通过迭代的方法求得。

$$S^N=S^0+\sum_{j=1}^{N}\Delta S^j \tag{2}$$

训练每个回归器 R^t 时,目标就是使得 $S_i^j=S_i^{j-1}+R^j(x_i^j)$ 能够使损失函数最小,即最小化当前形状和目标形状的距离。

$$R^j=\arg\min_R\sum_i d(R(x_i),S_i) \tag{3}$$

训练好 R^j 后,利用计算得到的 S_i^j 用于下一个回归器的训练。若回归器不能再减小误差,则停止训练。具体计算可见参考文献[12]。

1.2　改进的回归算法

在实际测试中,可以发现部分复杂人脸图像的最终定位结果偏差较大,远超出人脸合理形状范围。其原因在于级联回归器中的前几次回归,若产生较大偏离,则无法回归到正确位置上,最终结果偏离较大。因此本文对级联回归方法作出改进,对人脸特征点进行分块,对各个部分的形状分别进行估计并作为约束。算法具体流程为:

a)输入训练图集,进行特征形状核回归模型训练,获得目标的形状估计;

b)将估计的形状作为初始形状约束,初始坐标为所有训练图集的平均坐标,进行级联回归;

c)输出特征点的定位结果。

以 COFW(caltech occluded faces in the wild)人脸数据库为例,如图 2 所示,一共有 29 个特征点,共分为七块(左眉、右眉、左眼、右眼、鼻子、嘴巴、下巴),最容易受到干扰的部分是眉毛、眼睛和嘴巴,分别进行形状估计。其中,左眼以(9、11、13、14)四个点形成的四边形进行表示,17 号点则用(13、14)两点连线的中点代替。以最上方的点为一号点,逆时针排序。左眉、右眉、右眼部分类似,嘴巴部分则是用(23、24、25、28)表示,(26、27)两点在(25、28)两点连线上,与(25、28)两点的距离固定。

对其他数据库中的人脸图片,因为标注点的不同,同样采取分块的方式,但只对眼睛、眉毛和嘴巴部分进行姿态估计,对下巴、鼻子等不易发生形变和遮挡的部位,在确保初始形状合理的条件下,直接进行回归定位。

本文模型的输入变量为包含目标部分的人脸图像数据,输出变量为四边形中 1、2、3 号点对应的角度以及 1、2 号点连线和 2、3 号点连线的长度,以此来确定四边形的形状,也就是每一块的具体形状。

对每一块分别进行回归时,因为其形状已知,即四边形的形状已经确定,只需要确定任意两个点的位置,即可完成每个部分的整体定位。其初始位置设定为所有训练数据的平均值。因为不需要计算所有点的位置,可以更快降低估计误差从而收敛,减少了回归器的数量,提高了计算效率。

同时考虑到人脸的结构性,像素点间的距离越近,几何意义也更接近,其判别性更强,如眼角与眉毛附近的像素灰度差的判别性要强于眉毛与下颌附近的像素灰度差,如图 3 所示。

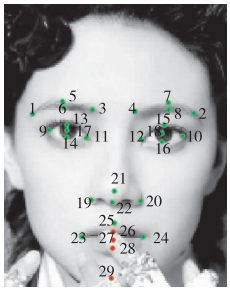


图2　COFW数据库目标

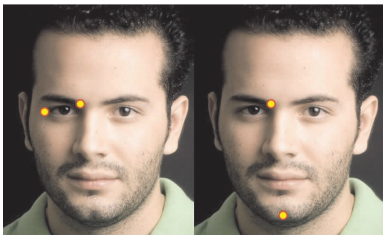


图3　距离不同的形状索引特征

因此,为了确保得到的形状索引特征的距离相近,在采样过程中加入指数先验进行距离约束。

$$P_j(u,v)\propto e^{-\alpha\|u-v\|+\beta j} \tag{4}$$

其中:$\|u-v\|$ 表示像素点对的距离;j 表示当前回归器序号;系数 α、β 用来控制距离。初始回归时,只需进行粗略定位,采样的像素点对距离较大;随着回归器的增多,为了对特征点进行精确定位,选择的距离不断减小,从而可以应对人脸表情的细微变化。

2　特征形状核回归模型及求解

2.1　特征形状核回归模型

回归分析是确定两种或两种以上变量间相互依赖的定量

关系的一种统计分析方法,包括逻辑回归、线性回归、非线性回归、非参数回归等。其主要内容是从一组数据出发,确定某些变量之间的定量关系表达式,这种表达式称为回归方程。利用所求得的回归方程进行预测和控制,其一般形式如下:

$$y_i = m(x_i) + \varepsilon_i \quad i = 1, \cdots, N \tag{5}$$

其中:$m(x)$ 为回归函数;变量 y 一般为标量;变量 x 可以为标量或矢量;ε 是随机误差,通常假设随机误差的均值为 0,方差为 σ^2。假设变量 X 和 Y 是随机的,可以认为回归函数 $m(x)$ 是 Y 在 $X = x$ 条件下的数学期望。

而核回归模型[17,18],对于给定训练数据集 $\{(x_i, y_i)\}_{i=1}^N$,$m(x)$ 可以表示为一系列核函数在训练数据集或其子集上的线性组合,即

$$m(x) = E(y|x) == \sum_{i=1}^{N} \alpha_i k(x, x_i) \tag{6}$$

其中:核函数 $k_i(x, x_i)$ 表示输入变量 x 之间的关系,α_i 可以通过线性方程求解得到。

传统模型的输出变量为一维变量。对于 N 维输出变量,一般采用 N 个回归函数分别估计每一维输出。但是在学习 1 号点角度的回归函数时,同一个输出值,可能会对应多幅图像,而这些图像的 2、3 号点的角度不同,这样的训练过程无法反映图像中该形状的完整信息。同时,在整个模型的训练过程中,每一张图片都被重复使用了 N 次,耗费了大量时间。

为解决上述问题,重新定义训练数据集 $\{(x_i, y_i)\}_{i=1}^N$,其中 y_i 为训练图像(或图像特征)x_i 对应的姿态向量。定义核函数 $k_Y(y, y_i)$ 表示输出变量 y 之间的关系,将式(6)中左侧输出加权到右侧,转换为

$$M(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} k_X(x, x_i) k_Y(y, y_j) \tag{7}$$

其中:$k_X(x, x_i)$ 和 $k_Y(y, y_j)$ 分别为定义输入和输出空间的核函数,表示输入变量之间与输出变量之间的关系;系数 β_{ij} 通过训练得到;$M(x, y)$ 为匹配函数,表示对于某个输入 x,其输出为 y 的可能性。当定义 $M(x, y)$ 取值在 0 ~ 1 时,可以将其看做是 x 和 y 的联合概率分布函数。式(7)进一步改写为

$$M(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} k_X(x, x_i) k_Y(y, y_j) = \sum_{i=1}^{N} [\beta_{i1}, \beta_{i2}, \cdots, \beta_{iN}] \begin{bmatrix} k_Y(y, y_1) \\ k_Y(y, y_2) \\ \vdots \\ k_Y(y, y_N) \end{bmatrix} k_X(x, x_i) \tag{8}$$

本文模型通过定义在输出空间的核函数 $k_Y(y, y_i)$ 加权到传统核回归模型的核函数上,可以同时表示输入变量之间和输出变量之间的关系,通过一次学习得到整个模型的回归函数,无须对每一维输出单独进行训练,大大提高了模型的训练效率。

2.2　模型求解

定义全部训练数据的匹配函数矩阵 $M \in R^{N \times N}$ 的元素为

$$M_{ij} = M(x_i, y_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{9}$$

定义系数矩阵 $A = (\beta_{ij}) \in R^{N \times N}$,则

$$M = K_X \times A \times K_Y \tag{10}$$

其中:M 为 N 阶单位阵;K_X、K_Y 均为 N 阶方阵;第 i 行第 j 列的元素分别是 $k_X(x_i, x_j)$ 和 $k_Y(y_i, y_j)$。于是,通过训练得到的系数矩阵:

$$A = K_X^{-1} \times M \times K_Y^{-1} \tag{11}$$

定义 $\psi(x) = [k_Y(x, x_1), k_Y(x, x_2), \cdots, k_Y(x, x_N)]^{\mathrm{T}}$,$\varphi(y) = [k_Y(y, y_1), k_Y(y, y_2), \cdots, k_Y(y, y_N)]^{\mathrm{T}}$,则匹配函数为

$$M(x, y) = \psi(x) \times A \times \varphi(y) \tag{12}$$

对于一幅未知图像 $\hat{x}$,其形状估计为

$$\hat{y} = \arg\max m(\hat{x}, y) \tag{13}$$

2.3　图像特征及核函数的选取

为有效提高速度和识别率,本文选取的图像特征为像素差特征(pixel-difference feature,PDF)[19]。即在同一图像中,两个先前已经定义好位置的像素点的灰度值之差。具体而言,每个像素灰度差特征 $h_{p_1, p_2}(I)$ 就是定义在图像中 p_1、p_2 位置上的两个像素点差值,其赋值为

$$h_{p_1, p_2}(I) = I(p_1) - I(p_2) \tag{14}$$

其中:$I(p)$ 表示图像 I 在位置 p 处的灰度值。

像素差特征采用线性插值的方法,依据目标的形状而不是原始图像坐标对像素点进行索引,如图 4 所示。其可以确保特征的几何不变性,简化了回归模型,在训练过程中收敛速度更快,在处理充足的训练样本时,计算仍然非常简便而且效果鲁棒。

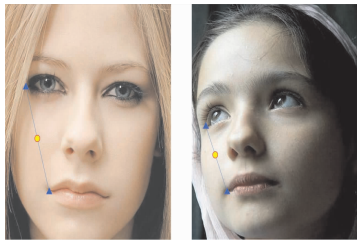


图4　线性插值索引

本文模型中,输入和输出空间核函数均采用 Gasussian 核函数:

$$\begin{cases} k_X(x_i, x_j) = \exp\left(\dfrac{-\| x_i - x_j \|_X^2}{2\sigma_X} \right) \\ k_Y(y_i, y_j) = \exp\left(\dfrac{-\| y_i - y_j \|_Y^2}{2\sigma_Y} \right) \end{cases} \tag{15}$$

其中:$\| \cdot \|_X$ 和 $\| \cdot \|_Y$ 是分别定义在输入和输出空间的 2 范数;x_i 为通过训练确定的固定的每幅图像的像素差组;y_i 为姿态向量,即输出的三个角度值和两个距离值。核函数参数设为

$$\begin{cases} \sigma_X = 0.1 \times \max\limits_{i,j} \| x_i - x_j \|_X^2 \\ \sigma_Y = 0.1 \times \max\limits_{i,j} \| y_i - y_j \|_Y^2 \end{cases} \tag{16}$$

需要强调的是,核函数的选取对于本文模型的性能将产生一定影响,具体情况将在第 3 章实验中进行分析。

3　实验分析

为验证本文方法对遮挡和形变条件下特征点定位的准确性,选择了三个当下流行的人脸数据库进行测试,分别是 LFPW(labeled face parts in the wild)、HELEN 和 COFW 数据库。

LFPW 数据库在文献[14]中被创立并提出。在测试对非约束条件下的人脸形状估计算法的性能中,LFPW 数据库已成为当前所用最多的人脸数据库之一。

HELEN 数据库由 2 330 个高分辨率的人脸图像构成,它的

特征点标定非常密集,每幅图片包含 194 个点,这也增加了在该数据库中的实验难度。本文选择其中 2 000 张图片作为训练样本,其余 330 张图片作为测试样本。

COFW 数据库包含了 1 007 张姿态表情变化较大、遮挡较为严重的图片,与 LFPW 数据库具有相同基准标定。本文采用 845 张 LFPW 数据库的图片和 500 张 COFW 数据库的图片作为训练样本,其余 507 张 COFW 数据库人脸作为测试样本。

为验证算法效能,选择在 Windows 7、i5-4440 3.1 GHz、8 GB内存的 PC 机上用 MATLAB 对比分析算法的优劣。

3.1 核函数选择实验

本文采取两种评价指标:a)平均误差,即估计值与真值误差绝对值的平均;b)正确率,即估计值与真值误差绝对值小于某一阈值的百分比。

为评价输入和输出变量的核函数的选取对本文方法的性能影响,选取 Gaussian 和 Sigmoid 核函数进行比较,实验在一维数据上进行,即只对单一角度结果进行计算。对比实验结果如表 1 所示。对于输入变量,采用 Gaussian 核函数的姿态估计结果优于 Sigmoid 核函数;对于输出变量,在平均误差方面,采用 Gaussian 核函数较好,而在 2° 和 5° 范围内的误差百分比则是 Sigmoid 核函数占优。

表 1 核函数选择实验结果

输入核函数	输出核函数	平均误差/°	正确率/%		
			<1°	<2°	<5°
Gaussian	Gaussian	0.53	87.53	88.24	98.81
Gaussian	Sigmoid	1.07	55.29	97.18	99.24
Sigmoid	Gaussian	1.24	75.47	87.18	92.15
Sigmoid	Sigmoid	1.53	54.36	92.12	94.92

结合两种核函数本身的性质来看,输入变量间的差别较大,采用 Gaussian 核函数进行表示较为合适,而输出变量的变化范围有限;采用 Sigmoid 核函数进行表示,可提高整体估计的准确性,但相较 Gaussian 核函数其估计精度有所下降。因此,本文选择输入和输出空间均采用 Gaussian 核函数。

3.2 图像特征选择实验

实验在二维数据上进行,即对两种角度结果进行计算,实验结果如表 2 所示。

表 2 图像特征选择实验结果

图像特征	平均误差/°	正确率/%		
		<1°	<2°	<5°
FD	13.32	3.51	14.62	29.83
	10.28	6.86	24.23	30.84
HOG	3.65	29.92	46.54	87.23
	2.27	45.20	77.32	91.24
PDF(1)	15.21	2.35	8.23	27.52
	10.25	4.03	9.85	28.93
PDF(2)	9.93	7.55	13.91	35.24
	6.54	8.75	37.22	51.76
PDF(3)	5.12	18.52	46.28	78.72
	3.98	22.45	51.27	83.41
PDF(4)	3.87	27.13	45.53	85.24
	2.65	36.82	76.42	90.28
PDF(5)	2.44	55.38	78.65	93.58
	1.12	75.28	95.32	99.24
PDF(6)	1.87	76.84	92.15	98.81
	0.53	92.65	99.24	100
PDF(7)	1.65	88.75	96.65	99.41
	0.48	98.62	99.80	100
PDF(8)	1.52	90.12	96.65	99.80
	0.43	98.82	99.80	100

选择像素差特征,在选取超过五组像素对后,其估计误差几乎完全控制在 2° 的范围内;选择梯度直方图,有近 90% 的估计误差控制在 5° 的范围内;而傅里叶描述子的估计结果效果很差,失真严重。

实验结果表明,采用不同的图像特征对姿态估计的结果有很大影响。这是因为对于傅里叶描述子、梯度直方图等一些具有不变性的图像特征而言,不能够在特征空间对目标的姿态变化进行准确描述,从而导致了姿态估计精度的下降。因此在进行人脸部位的姿态估计时,要选用对目标姿态变化比较敏感的图像特征。同时,考虑到计算速度和训练效率,本文选择采用六组像素差对作为图像特征。

3.3 特征点定位实验

为验证本文方法对存在遮挡以及大的形变的人脸特征点定位的性能,采取的评价指标为平均误差(error)、失败率(failure)以及测试速度(Fps)。其中,定义最终估计的特征点位置与数据库中标定的特征点间的距离与每幅图片中双瞳的距离之比为误差,而误差大于 0.1,则会被判定为失败,从而计算失败率。

将特征点划分为七块(分别是左眉、右眉、左眼、右眼、鼻子、嘴巴和下巴)。每个分块的回归器在训练过程中采用固定的参数设置:第一层回归器的数量为 10,第二层回归器的数量为 300。每个回归器选用的形状索引特征的个数及所构建弱回归器的层数为 6。

同时,选择 ESR[7]、RCPR[8]、SDM[9]、LBF[10] 以及 cGPRT[11] 作为对比算法,不同数据库下各种方法平均误差对比如表 3 所示。

表 3 不同数据库下各种方法平均误差对比

	ESR	RCPR	SDM	LBF	cGPRT	本文算法
LFPW	0.038	0.035	0.035	**0.033**	0.035	0.034
HELEN	0.075	0.065	0.059	0.054	**0.046**	0.053
COFW	0.112	0.085	0.074	0.065	0.064	**0.058**

从表 3 中看出,本文算法在 LFPW 和 HELEN 此类无遮挡数据库上优势不是很大,但在 COFW 数据库上,对于存在遮挡及形变较大的图片,相较于其他算法定位精度和成功率有一定提高。

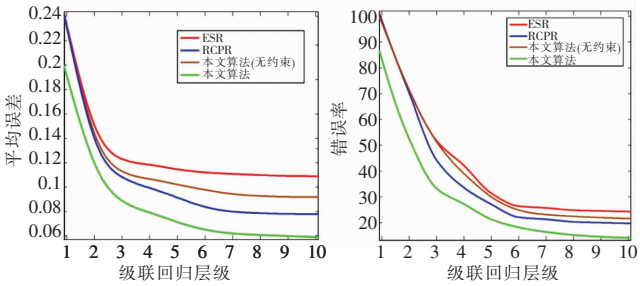
由于本文算法主要目的之一是处理存在遮挡的图像,而 LFPW 与 HELEN 数据库中存在遮挡的图像较少,所以在这两个数据库中的实验主要是用于验证分块回归和初始粗定位的性能。下面详细介绍在 COFW 数据库上的对比实验。因为部分算法没有获得相应代码,只有最终误差结果,选择 ESR、RCPR 以及本文算法只进行分块而不添加形状约束作为对比算法。OCFW 数据库实验结果对比如表 4 所示。COFW 数据库实验结果如图 5 所示。

表 4 COFW 数据库实验结果对比

method	error	failures/%	Fps
ESR	0.112	24.32	110
RCPR	**0.085**	**19.74**	**180**
本文算法(无约束)	0.092	21.63	320
本文算法	0.058	14.12	280
human	0.033	/	/

根据实验结果可知,只对图像进行分块操作后,本文算法相较于 ESR 算法在定位精度方面有了很大提高,但仍落后于 RCPR 算法。加入了姿态估计作为形状约束后的本文算法的定位精度大大提高,在回归未开始时就已经明显领先于其他算

法并最终输出最准确的结果,特别是将失败率降至 14%,体现出极强的鲁棒性,能够可靠地处理存在遮挡以及大的形变的人脸图像。同时,本文算法的处理速度达到 280 fps,可以满足实时处理的要求。部分数据库下一步实验结果如图 6 所示。



(a)平均误差　　(b)错误率

图 5　COFW 数据库实验结果

(a)部分 LFPW 数据库结果

(b)部分 HELEN 数据库结果

(c) 部分 COFW 数据库结果

图 6　实验结果

4　结束语

本文提出利用形状估计的分块特征点定位算法,通过学习得到图像特征与目标形状间的匹配函数,并求取匹配函数的最大值,得到目标的形状估计。与现有经典的 RCPR 和 cGPRT 方法的仿真对比及分析表明:

a)提高了算法对于复杂人脸图像和遮挡的定位精度。对人脸特征点进行分块,利用得到的形状估计作为约束,同时只需对部分点的位置进行回归,减少了回归器的数量,大幅提高了计算效率。

b)引入了形状索引特征的采样先验,使算法能够更加高效地采样到判别性强的特征来构建回归器,从而提升了算法的性能。

c)选择合适的图像特征后,本文算法可提供精确的形状估计,从而获得优于现有算法的定位成功率和准确度,具有较强的鲁棒性以及较快的计算速度,能够达到实时处理的要求。

为高速鲁棒处理视频序列中的人脸特征点信息,后续还需深入研究人脸特征点的先验信息以获取人脸特征点候选点以及加权组合方式,从而构建有效的人脸特征点定位模型。

参考文献:

[1] Ayinde O, Yang Y H. Face recognition approach based on rank correlation of Gabor-filtered images[J]. Pattern Recognition, 2002, 35(6): 1275-1289.

[2] Cao Chen, Hou Qiming, Zhou Kun. Displaced dynamic expression regression for real-time facial tracking and animation[J]. ACM Trans on Graphics, 2014, 33(4): 1-10.

[3] Martinez A, Du Shichuan. A model of the perception of facial expressions of emotion by humans: research overview and perspectives[J]. The Journal of Machine Learning Research, 2012, 13(1): 1589-1608.

[4] Forsyth D. Object detection with discriminatively trained part-based models[J]. IEEE Trans on Software Engineering, 2010, 32(9): 1627-1645.

[5] Dollar P, Welinder P, Perona P. Cascaded pose regression[C]//Proc of Computer Vision and Pattern Recognition. 2010: 1078-1085.

[6] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]//Proc of Computer Vision and Pattern Recognition. 2014: 1867-1874.

[7] Cao Xudong, Wei Yichen, Wen Fang, *et al*. Face alignment by explicit shape regression[J]. International Journal of Computer Vision, 2014, 107(2): 177-190.

[8] Burgos-Artizzu X P, Perona P, Dollár P, *et al*. Robust face landmark estimation under occlusion[C]//Proc of IEEE International Conference on Computer Vision. [S. l.]: IEEE Computer Society, 2013: 1513-1520.

[9] Xiong Xuehan, De La Torre F. Supervised descent method and its applications to face alignment[C]//Proc of Computer Vision and Pattern Recognition. 2013: 532-539.

[10] Ren Shaoqing, Cao Xudong, Wei Yichen, *et al*. Face alignment at 3000 FPS via regressing local binary features[C]//Proc of Computer Vision and Pattern Recognition. 2014: 1685-1692.

[11] Lee D, Park H, Yoo Chang D. Face alignment using cascade Gaussian process regression trees[C]//Proc of Computer Vision and Pattern Recognition. 2015: 4204-4212.

[12] 仲柔在,熊磊,刘畅. 基于相对距离分布聚类的人脸特征点定位算法[J]. 空军工程大学学报:自然科学版, 2016, 17(1): 77-82.

[13] Zhu Xiangxin, Ramanan D. Face detection, pose estimation, and landmark localization in the wild[C]//Proc of Computer Vision and Pattern Recognition. 2012: 2879-2886.

[14] Belhumeur P, Jacobs D, Kriegman D, *et al*. Localizing parts of faces using a consensus of exemplars[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2930-2940.

[15] Cristinacce D, Cootes T F. Feature detection and tracking with constrained local models[C]//Proc of British Machine Vision Conference, 2006: 929-938.

[16] Cootes T F, Edwards G J, Taylor C J. Active appearance model[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681-685.

[17] 张浩鹏,姜志国. 基于姿态加权核回归的航天器姿态估计[J]. 北京航空航天大学学报, 2014, 40(4): 494-499.

[18] Torki M, Elgammal A. Regression from local features for viewpoint and pose estimation[C]//Proc of IEEE International Conference on Computer Vision. 2011: 2603-2610.

[19] Ozuysal M, Calonder M, Lepetit V, *et al*. Fast keypoint recognition using random ferns[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2010, 32(3): 448-461.