

基于混合投影函数的眼睛定位*

耿新, 周志华⁺, 陈世福

(南京大学 计算机软件新技术国家重点实验室, 江苏 南京 210093)

Eye Location Based on Hybrid Projection Function

GENG Xin, ZHOU Zhi-Hua⁺, CHEN Shi-Fu

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

+ Corresponding author: Phn: 86-25-3593163, Fax: 86-25-3300710, E-mail: zhouzh@nju.edu.cn

<http://cs.nju.edu.cn/people/zhouzh/>

Received 2002-09-20; Accepted 2002-12-04

Geng X, Zhou ZH, Chen SF. Eye location based on hybrid projection function. *Journal of Software*, 2003,14(8):1394~1400.

<http://www.jos.org.cn/1000-9825/14/1394.htm>

Abstract: A method is described to locate eyes accurately in this paper. Eye location is the key step of many face detection methods. After obtaining an eye window, the eye can be located by using projection functions. A novel projection function, i.e. hybrid projection function (HPF), is proposed in this paper. HPF combines integral projection function and variance projection function so that it can reflect the mean and the variance of the pixels in a particular direction of the image. Experimental results show that eyes can be located quite accurately by using hybrid projection function.

Key words: face detection; face recognition; eye location; projection function

摘要: 介绍一种精确定位眼睛的方法.眼睛定位是很多人脸检测算法的关键步骤.在获得眼睛窗口之后,可以利用投影方法来定位眼睛.提出一种混合投影函数,该函数结合了积分投影函数和方差投影函数,能够反映出图像在某个方向上灰度均值和方差的综合变化情况.实验结果表明,利用混合投影函数可以相当精确地确定眼睛中心的位置.

关键词: 人脸检测;人脸识别;眼睛定位;投影函数

中图法分类号: TP391 文献标识码: A

目前的人脸识别算法可以分成两大类^[1].一类是基于几何特征的方法,另一类是基于模板的方法.而无论哪种方法,都是建立在人脸区域的大小和位置已知的基础上的.人脸检测的目的就是确定给定的图像或图像序列中是否包含人脸,如果是,则给出人脸的位置和大小^[2].所以,对自动人脸识别系统来说,有效的人脸检测是进行精确识别的先决条件.

人脸检测是一项非常困难的工作,其复杂度在某种程度上甚至超过了人脸识别^[2].经过多年的研究,已经出

* Supported by the 10th Five Years Plan of Jiangsu Province of China under Grant No.BE2001028 (江苏省“十五”科技攻关计划)

第一作者简介: 耿新(1978—),男,江苏徐州人,硕士生,主要研究领域为模式识别,数据挖掘.

现了许多人脸检测的方法.而由于眼睛是面部最为显著的特征之一,所以眼睛定位成为许多人脸检测方法的关键步骤.一旦左右眼睛的位置确定下来,人脸所在的位置也就基本确定了.进一步,根据两眼之间连线的长度和方向,人脸区域的大小和方向也可以大致估计出来.

目前已经出现了一些有效的眼睛定位方法.例如,Bala 等人^[3]提出一种基于遗传算法和决策树的眼睛定位方法.该方法采用一种混合遗传结构使基本视觉规则不断进化,最终得到可用于眼睛定位的决策树形式的视觉规则.Reinders 等人^[4]提出一种基于神经网络的眼睛定位方法.该方法将搜索窗口的像素作为神经网络的输入,如果该窗口包含眼睛图像,则神经网络的输出较大.Wu 和 Zhou^[5]提出一种基于灰度对比的眼睛定位方法.该方法利用眼睛区域的灰度较大这一特点找出眼睛的位置.然而,上述方法仅能给出眼睛的大致位置,并不能精确定位眼睛的中心.因此,为了提高人脸检测的准确率,有必要研究眼睛精确定位的方法.

投影是一种有效提取图像特征的方法.通常,一幅二维图像可以由两个正交的一维投影函数来分析.维数的降低便于分析图像的特征,并且减小了计算量,所以投影成为一种重要的图像分析方法.到目前为止,已有很多学者将投影函数成功运用于定位面部特征.Kanade^[6]最早将积分投影函数成功地应用于人脸识别,他首先用拉普拉斯算子对原始灰度图进行二值化,然后用积分投影函数对二值图进行分析.Brunelli 和 Poggio^[1]对 Kanade 的算法作了改进,他们将积分投影函数应用于边界图分析,从而确定出面部各个特征的位置.方差投影函数的概念最早由 Feng 和 Yuen^[7]提出,并同时提出了一种利用方差投影函数定位眼睛的简单方法.后来,他们又提出一种多线索定位眼睛的方法^[8],其中应用了一种眼睛方差过滤器(eye variance filter),而这种过滤器正是利用方差投影函数产生的.由此可以看出,投影实际上是人脸识别中经常采用的一种定位技术.

在积分投影函数(integral projection function,简称 IPF)和方差投影函数(variance projection function,简称 VPF)的基础上,本文提出一种混合投影函数(hybrid projection function,简称 HPF).该函数综合考虑了图像在某个方向上的均值与方差,既能反映出不同的图像灰度,又能反映出不同的灰度变化率,具有较强的适应性和准确性.本文通过将混合投影函数应用于眼睛精确定位验证了这一点.

本文第 1 节介绍了研究背景,第 2 节介绍混合投影函数,第 3 节介绍一种利用混合投影函数定位眼睛的方法,第 4 节给出实验结果,最后给出结论.

1 研究背景

1.1 积分投影函数

最为常用的投影函数是积分投影函数^[6].假设 $I(x,y)$ 表示点 (x,y) 处的像素灰度值,在区间 $[x_1,x_2]$ 和 $[y_1,y_2]$ 内的垂直积分投影函数和水平积分投影函数分别表示为 $S_v(x)$ 和 $S_h(y)$,则

$$S_v(x) = \int_{y_1}^{y_2} I(x,y)dy, S_h(y) = \int_{x_1}^{x_2} I(x,y)dx. \quad (1)$$

平均积分投影函数 $M_v(x)$ 和 $M_h(y)$ 分别表示为

$$M_v(x) = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x,y)dy, M_h(y) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x,y)dx. \quad (2)$$

当图像某一列像素灰度均值发生变化时,这种变化会在 $S_v(x)$ 和 $M_v(x)$ 的值上反映出来.同理,当图像某一行像素灰度均值发生变化时,这种变化会在 $S_h(x)$ 和 $M_h(x)$ 的值上反映出来.因此,可以通过对积分投影函数值的分析提取图像中的特征.例如图 1(b),在水平方向上 3 个不同区域的交界处,积分投影函数对应的曲线发生了跃变,从而准确反映出了图像中的区域特征.

但是,在均值相同的情况下,积分投影函数会失效.如图 1(a)所示,很显然,在水平方向上有 3 个不同的区域,但是积分投影函数对应的曲线并不能反映出这种变化.

1.2 方差投影函数

积分投影函数只考虑了均值的变化,却不能反映均值相同情况下灰度值的变化.如果用方差代替均值,就可以避免这种情况的发生,这就是方差投影函数^[7].设在区间 $[x_1,x_2]$ 和 $[y_1,y_2]$ 内的垂直方差投影函数和水平方差投

影函数分别表示为 $\sigma_v^2(x)$ 和 $\sigma_h^2(y)$, 则

$$\sigma_v^2(x) = \frac{1}{y_2 - y_1} \sum_{y_i=y_1}^{y_2} [I(x, y_i) - M_v(x)]^2, \sigma_h^2(y) = \frac{1}{x_2 - x_1} \sum_{x_i=x_1}^{x_2} [I(x_i, y) - M_v(y)]^2. \quad (3)$$

其中, $M_v(x)$ 和 $M_h(y)$ 由式(2)求出.

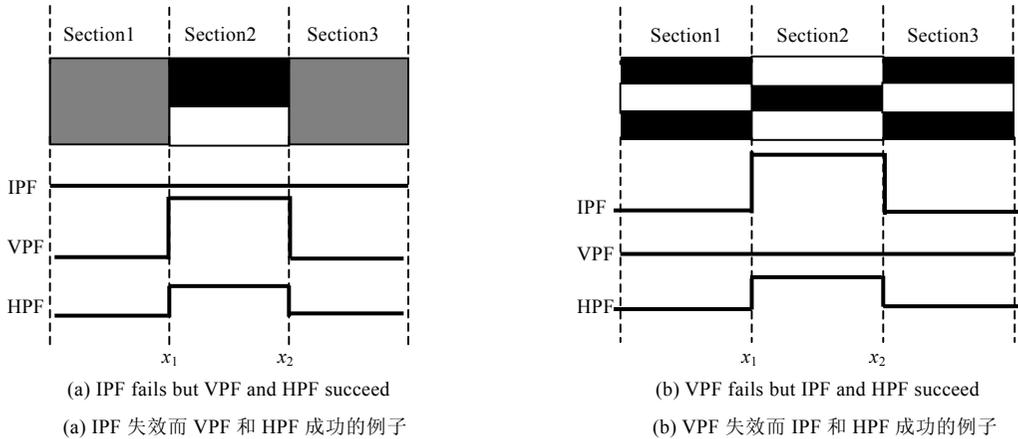


Fig. 1 Hybrid projection function reflects changes between different sections

图 1 混合投影函数反映出不同区域的变化

当图像某一列像素灰度的方差发生变化时,这种变化会在 $\sigma_v^2(x)$ 的值上反映出来.同理,当图像某一行像素灰度的方差发生变化时,这种变化会在 $\sigma_h^2(y)$ 的值上反映出来.此外,Feng 和 Yuen^[7]还证明了 VPF 对随机噪声并不敏感.因此,同样可以利用 VPF 对图像特征进行分析.在图 1(a)中,垂直 VPF 对应的曲线在不同区域的边界处发生了较大的跃变,从而反映出了图像在水平方向上的区域变化.

同样地,由于只考虑了灰度值的变化情况,而没有考虑灰度平均值,在方差相同的情况下,VPF 也会失效.如图 1(b)所示,水平方向上有 3 个明显不同的区域,但是 VPF 对应的曲线并不能反映这种变化,而积分投影函数对应的曲线则在不同区域的边界处发生了跃变,从而反映出水平方向上的不同区域.

2 混合投影函数

考虑到积分投影函数与方差投影函数在一定程度上具有互补性,例如图 1(a)中 IPF 失效 VPF 成功,而图 1(b)中 VPF 失效 IPF 成功.本文提出一种混合投影函数.垂直混合投影函数 $H_v(x)$ 和水平混合投影函数 $H_h(y)$ 分别定义为

$$H_v(x) = \frac{1}{2} \sigma_v^2(x) + \frac{1}{2} M_v'(x), H_h(y) = \frac{1}{2} \sigma_h^2(y) + \frac{1}{2} M_h'(y). \quad (4)$$

其中 $\sigma_v'(x)$, $\sigma_h'(y)$, $M_v'(x)$, $M_h'(y)$ 分别是 $\sigma_v(x)$, $\sigma_h(y)$, $M_v(x)$, $M_h(y)$ 规范化到 [0,1] 区间后的结果:

$$\sigma_v'(x) = \frac{\sigma_v(x) - \min(\sigma_v(x))}{\max(\sigma_v(x)) - \min(\sigma_v(x))}, \sigma_h'(y) = \frac{\sigma_h(y) - \min(\sigma_h(y))}{\max(\sigma_h(y)) - \min(\sigma_h(y))}. \quad (5)$$

$$M_v'(x) = \frac{M_v(x) - \min(M_v(x))}{\max(M_v(x)) - \min(M_v(x))}, M_h'(y) = \frac{M_h(y) - \min(M_h(y))}{\max(M_h(y)) - \min(M_h(y))}. \quad (6)$$

HPF 的值反映了图像在某个方向上灰度均值和方差变化情况的综合,所以具有更强的适应性.假设 Δ 表示一个很小的值,当 HPF 的值从 $H_v(x_0)$ 到 $H_v(x_0 + \Delta)$ 发生了显著变化时,则表示在 $x = x_0$ 处是两块不同区域的边界.具体地,给定阈值 T ,找出所有满足以下条件的连续区间 X :

$$\forall x \in X, |H_v(x + \Delta) - H_v(x)| > T. \quad (7)$$

对离散的数字图像来说,一般 $\Delta = 1$.然后找出所有满足条件的 X 区间中的最大值点 x_0 ,即 x_0 应满足以下条件:

$$H_v(x_0) = \max_{x \in X} (|H_r(x+\Delta) - H_r(x)|), \tag{8}$$

则 $x=x_0$ 处为不同区域的边界.如图 1 所示, x_1, x_2 即为图像在水平方向上的区域边界点.由图 1 可以看出,不论是在积分投影函数失效的情况下(图 1(a)),还是在方差投影函数失效的情况下(图 1(b)),混合投影函数均能反映出水平方向上不同区域的变化.

设 X 为随机变量,其期望与方差分别为 $E(X)$ 和 $\sigma^2(X)$. η 为独立的随机噪声,满足正态分布 $N(0, \sigma^2(\eta))$, 则

$$\begin{aligned} \frac{1}{2}\sigma^2(X+\eta) + \frac{1}{2}E(X+\eta) &= \frac{1}{2}(E(X+\eta - E(X+\eta))^2 + E(X)) \\ &= \frac{1}{2}(E(X - E(X))^2 + E(\eta^2) + E(X)) \\ &= \frac{1}{2}(\sigma^2(X) + E(X) + \sigma^2(\eta)). \end{aligned} \tag{9}$$

而在一般情况下,随机噪声的方差 $\sigma^2(\eta)$ 远远小于 $\sigma^2(X)$, 即 $\sigma^2(\eta) \ll \sigma^2(X)$, 所以

$$\frac{1}{2}\sigma^2(X+\eta) + \frac{1}{2}E(X+\eta) \approx \frac{1}{2}\sigma^2(X) + \frac{1}{2}E(X). \tag{10}$$

由此可见,HPF 对随机噪声并不敏感.

3 眼睛定位

眼睛区域具有两个显著的特征.其一,眼睛区域与周围区域相比较暗,即灰度值较低(应用积分投影函数定位眼睛时利用了这一特点);其二,眼睛区域的灰度变化率较大(应用方差投影函数定位眼睛时利用了这一特点).因此,精确定位眼睛位置要综合考虑这两个特征.HPF 既能反映灰度的变化,又能反映灰度变化率的变化,所以适合于眼睛定位的应用.

由于眉毛、嘴唇等非眼睛的面部特征会对 HPF 产生较大影响,所以如果将 HPF 直接应用于整幅面部图像,很难取得眼睛的准确位置.因此,要利用 HPF 定位眼睛中心,首先必须得到一幅仅包含眼睛的图像,即眼睛窗口.

利用目前已知的许多眼睛定位方法都可以大致标识出眼睛的坐标.扩展该坐标周围的一块图像作为眼睛窗口,然后在其中精确定位眼睛的中心位置.设左右眼睛粗糙定位的坐标分别为 $l(x_1, y_1)$ 和 $r(x_2, y_2)$, l 点和 r 点的距离为 d , 则分别以 l 点和 r 点为中心,长为 $4/5d$, 宽为 $2/5d$ 的矩形区域作为眼睛窗口,如图 2 所示.其中,小圆圈代表粗糙定位的眼睛位置,矩形框为眼睛窗口,十字叉为精确定位的眼睛中心位置.

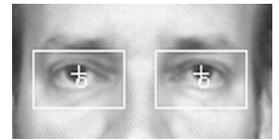


Fig.2 Eye window
图 2 眼睛窗口

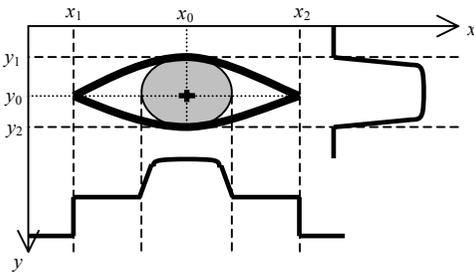


Fig.3 Standard eye model
图 3 标准眼睛模型

为了定位眼睛中心,首先建立标准的眼睛模型,如图 3 所示.从图中可以看出,要定位眼睛的中心位置,需要确定左右眼角的 x 轴坐标和上下眼睑的 y 轴坐标. x_1 和 x_2 分别为左右眼角的 x 轴坐标, y_1 和 y_2 分别为上下眼睑的 y 轴坐标.垂直 HPF 在 x_1 和 x_2 处会有明显变化,而水平 HPF 在 y_1 和 y_2 处会有明显变化,令

$$x_o = \frac{x_1 + x_2}{2}, y_o = \frac{y_1 + y_2}{2}, \tag{11}$$

则点 $O(x_o, y_o)$ 即为眼睛的中心点.

图 4 是混合投影函数在眼睛窗口中定位眼睛中心的情况.其中,图 4(a)表示水平混合投影函数用于确定上下眼睑的位置,图 4(b)表示垂直混合投影函数用于确定左右眼角的位置.其中,黑色曲线表示混合投影函数,灰色曲线表示混合投影函数的变化率,白色直线标出了变化率峰值的位置. y_1, y_2 表示上下眼睑的坐标, x_1, x_2 表示左右眼角的坐标.可以看出,水平 HPF 在上下眼睑处发生了显著变化,而垂直 HPF 在左右眼角处发生了显著变化.

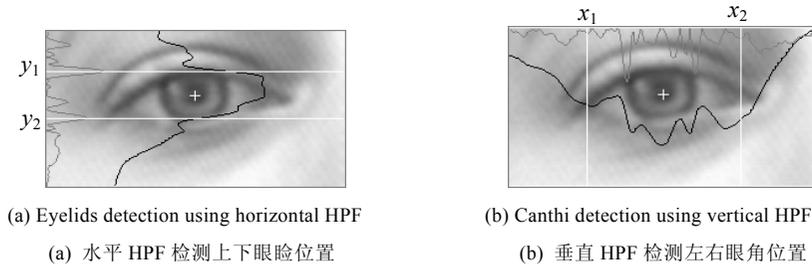


Fig.4 Eye center location in the eye window using HPF

图4 HPF用于在眼睛窗口中定位眼睛中心

4 实验测试

4.1 数据集

实验采用两种不同的数据集,即 BioID 人脸库^[9,10]和 NJUFace 人脸库.其中 BioID 人脸库共包含 23 个不同人物的 1 521 张正面人脸灰度照片,照片中的人物大部分为欧美人,每张照片的解析度为 384×268 .人脸图像在不同的季节、地点采集,光线、背景、人脸大小均有较大变化.图 5(a)给出了 BioID 人脸库中一些典型的人脸照片.NJUFace 是南京大学计算机软件新技术国家重点实验室建立的人脸库,共包括 359 张人脸照片,所有人物均为中国人.每张照片的解析度为 380×285 .同一人物的不同照片在面部表情和头部方位等方面均有较大差异.图 5(b)给出了 NJUFace 人脸库中一些典型的人脸照片.

需要说明的是,本文实验所采用的数据集中每幅图像上大多只有一张人脸,而一般测试人脸检测算法均选择较为复杂(如一幅图像中包括多张人脸且大小不同)的图像库.这是因为本文实验的目的是比较在已获得眼睛窗口的基础上,用不同投影函数精化眼睛定位的相对效果,而图像的复杂化主要是增加了获得眼睛窗口的难度,并不是本文的研究重点.



(a) Typical examples in BioID

(a) BioID 人脸库中的典型例子



(b) Typical examples in NJUFace

(b) NJUFace 人脸库中的典型例子

Fig.5 Typical examples in the face databases

图5 人脸库中的典型例子

4.2 度量标准

本文实验采用了 Jesorsky 等人的度量标准^[11].设手工标定的左右眼睛准确位置分别为 E_l 和 E_r ,检测到的左右眼睛位置分别为 E_l' 和 E_r' , d_l 为 E_l 到 E_l' 的欧氏距离, d_r 为 E_r 到 E_r' 的欧氏距离, d_{lr} 为 E_l 到 E_r 的欧氏距离,则检测的相对误差定义为

$$err = \frac{\max(d_l, d_r)}{d_{lr}} \quad (12)$$

一般来讲, d_{lr} 约为眼睛宽度的两倍. 这样, $err=0.25$ 就意味着 d_l, d_r 中较大的一个约为眼睛宽度的一半, 低于这个误差被认为是可以接受的^[11]. 所以, 如果 $err < 0.25$, 则认为检测结果是正确的, 否则错误.

4.3 测试结果

实验采用 Wu 和 Zhou^[5]提出的基于灰度对比的眼睛定位方法来进行眼睛的粗糙定位. 在此基础上, 分别采用积分投影函数、方差投影函数和混合投影函数 3 种方法定位实验数据集中的眼睛位置. 实验从两个方面考察 3 种定位方法. 其一为正确率, 即 $err < 0.25$ 的照片在照片总数中所占的比率; 其二为准确定位的平均相对误差, 即在 $err < 0.25$ 的条件下, err 的平均值.

需要指出的是, 由于眼睛区域的灰度值较低且变化较大, 因此, $M_v'(x), M_h'(y)$ 会降低, 而 $\sigma_v'(x), \sigma_h'(x)$ 会提高, 为了使二者在眼睛区域的变化趋势一致, 应对式(4)略作调整. 即将 $M_v'(x), M_h'(y)$ 分别换成 $M_v''(x)=1-M_v'(x), M_h''(y)=1-M_h'(y)$.

实验结果见表 1.

Table 1 Experimental results about the three projection functions on BioID and NJUFace

表 1 3 种投影函数在 BioID 和 NJUFace 上的实验结果

	Detection rate		Average relative error of correct detections	
	BioID (%)	NJUFace (%)	BioID	NJUFace
IPF	93.69	92.48	0.098 1	0.081 1
VPF	94.41	95.54	0.087 3	0.076 9
HPF	94.81	95.54	0.087 0	0.073 8

从表 1 可以看出, 对于 BioID 人脸库, 混合投影函数在检测正确率和准确定位的平均相对误差两个指标上都明显优于积分投影函数和方差投影函数. 这说明混合投影函数在 BioID 人脸库上显著优于积分投影函数和方差投影函数. 对于 NJUFace 人脸库, 混合投影函数仍显著优于积分投影函数. 虽然混合投影函数的检测正确率与方差投影函数相同, 但其准确定位的平均相对误差小于方差投影函数. 这说明在 NJUFace 人脸库上, 混合投影函数略好于方差投影函数. 综合 3 种投影函数在两个数据库上的表现, 可以看出, 混合投影函数最优, 方差投影函数其次. 该结果也支持了 Feng 和 Yuen^[7]的发现, 即在眼睛检测方面, 方差投影函数与传统的积分投影函数相比, 可以获得更高的正确率和更低的误差.

5 结束语

本文提出了一种混合投影函数并将其应用在人脸检测中眼睛位置的精化, 取得了较好的效果. 实验结果表明, 在眼睛定位方面, 与积分投影函数和方差投影函数相比, 混合投影函数具有更强的适应性和更高的准确性.

References:

- [1] Brunelli R, Poggio T. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993, 15(10):1042~1052.
- [2] Yang MH, Kriegman DJ, Ahuja N. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(1):34~58.
- [3] Bala J, DeJong K, Huang J, Vafaie H, Wechsler H. Visual routine for eye detection using hybrid genetic architectures. In: Backer E, Gelsema ES, eds. *Proceedings of the 13th International Conference on Pattern Recognition*. Los Alamitos: IEEE CS Press, 1996, 3:606~610.
- [4] Reinders MJT, Koch RWC, Gerbrands JJ. Locating facial features in image sequences using neural networks. In: Essa I, ed. *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition*. Los Alamitos: IEEE CS Press, 1996. 230~235.
- [5] Wu JX, Zhou ZH. Efficient face candidates selector for face detection. *Pattern Recognition*, 2003, 36(5):1175~1186.
- [6] Kanade T. *Picture processing by computer complex and recognition of human faces* [Ph.D. Thesis]. Kyoto: Kyoto University, 1973.

- [7] Feng GC, Yuen PC. Variance projection function and its application to eye detection for human face recognition. Pattern Recognition Letters, 1998,19(9):899~906.
- [8] Feng GC, Yuen PC. Multi cues eye detection on gray intensity image. Pattern Recognition, 2001,34(5):1033~1046.
- [9] Alattar AM, Rajala SA. Facial features localization in front view head and shoulders images. In: Rodriguez J, ed. 1999 IEEE International Conference on Acoustics, Speech and Signal Processing (Icassp). Los Alamitos: IEEE CS Press, 1999,6:3557~3560.
- [10] The BioID Face Database, BioID-Technology Research, 2001. <http://www.bioid.com>.
- [11] Jesorsky O, Kirchberg KJ, Frishholz RW. Robust face detection using the Hausdorff distance. In: Bigunand J, Smeraldi F, eds. Audio and Video based Person Authentication. Lecture Notes in Computer Science, Berlin: Springer-Verlag, 2001. 90~95.

几何造型与处理 2004 国际会议 征文通知

Geometric Modeling and Processing(GMP)是国际上有重要影响的学术会议, 每两年举办一次, 会议论文集由 IEEE Computer Society Press 出版, 优秀论文还将在 Computer Aided Design 和 Computer Aided Geometric Design 两个代表该领域国际最高学术水平的刊物上以专辑形式发表。GMP2004 将于 2004 年 4 月 13 日~15 日在清华大学召开。大会主席孙家广院士, 程序委员会主席是维也纳技术大学的 Helmut Pottmann 教授和清华大学计算机系胡事民教授。

征稿范围:

曲线曲面	CAGD 的数学基础	实体造型
多分辨率造型	形状表示技术	数字几何处理
几何约束求解	几何不确定性	几何计算中的偏微分方程方法

同时,也欢迎讨论以下应用领域的几何计算问题的论文:

反向工程	图像处理	计算机视觉
机器人技术	快速成型	网格生成
医学成像	科学计算可视化	

投稿截止时间:2003 年 9 月 30 日。

有关投稿要求, 请浏览会议网页: <http://cg.cs.tsinghua.edu.cn/GMP2004/>

联系人: 严寒冰、蔡晓华, 电话: 010-62782052, Email: gmp2004@tsinghua.edu.cn