

基于多点模型的3D人脸姿态估计方法

胡步发 邱丽梅

(福州大学机械工程及自动化学院,福州 350002)

摘要 改进传统的活动形状模型法,准确地提取人脸特征点后,针对人脸形状特性,使用人脸的多个特征点作为人脸模型,通过最小二乘法优化求解,精确估计3D人脸空间姿态。实验结果表明,新方法不仅可以获得稳定的姿态解,而且与同类方法比较具有良好的姿态估计精确度。

关键词 人脸姿态估计 活动形状模型 最优化

中图法分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)07-1353-06

3D Face Pose Estimation Based on Multi-points Model

HU Bu-fa, QIU Li-mei

(College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350002, China)

Abstract The improved method of active shape model is used to extract face feature points accurately. Then in view of the characteristic of face shape, taking some face feature points as the face model, 3D space face pose is estimated via the least-squares method optimization. The experiment results show that the new method not only can achieve steady pose solution, but also has better estimation accuracy compared with the same kind methods.

Keywords estimation of face pose, active shape model(ASM), optimization

1 引言

人脸姿态估计是指确定输入图像中的人脸在3D空间中的姿态。姿态与视角有直接的对应关系,因此,姿态估计问题也称为人脸视角估计。人脸姿态估计在人脸识别、计算机游戏、虚拟现实和司机疲劳检测系统等方面都有着广泛的应用。对于现有人脸姿态估计的方法大体上可以分为以下两类:

(1)基于人脸外观的学习方法 假设3D人脸姿态与人脸图像的某些特性(图像密度、颜色、图像梯度值等)存在唯一的对应关系,用大量已知3D人脸姿态的训练样本,通过统计方法来建立这种关系^[1-6]。其中比较典型的有Darrell等人提出的利用本征空间进行姿态估计的方法^[1];Murase和Nayar提出的类似的特征空间方法^[2]。当然,基于人脸外

观的学习方法还经常使用人脸的颜色、灯光、头发及轮廓等等信息。Chen等人提出了用肤色和头发区域来建立头部模型^[3];Bruske等人提出了用神经网络的方法来进行人脸姿态估计^[4]。但是,因为这类方法大多数是利用插补,它们的精确度大大受限于训练样本的划分策略数量,结果人脸姿态的估计精度典型的都不少于10°。

(2)基于模型的方法 利用某种几何模型或结构来表示人脸的结构和形状,建立模型和图像之间的对应关系,然后通过几何或者其他方法实现人脸空间姿态估计^[7-13]。现有人脸姿态估计方法中,大多数姿态估计的方法属于基于模型的方法。Alter利用3点根据弱视投影估计人脸姿态^[7];Mazumdar等人使用两个外眉点和嘴的中心点,形成一个人脸T字型模型实现了人脸姿态的估计^[8];Kruger用Gabor小波变换来表示人脸各个特征点之间的关

系,进而估算出人脸方向^[9];Cipolla 等人利用合成图像模型来估计人脸姿态^[10];Ebisawal 以基于两瞳孔和鼻孔的 3D 检测来做人脸姿态估计^[11]。

这种方法的优点在于:人脸的几何结构通过投影模型清楚地揭示了 3D 人脸姿态和 2D 人脸图像的关系。当人脸特征点被精确定位后,这种方法简单易行,且具有高的精确度。

其中较直观的是使用物体表面的特征点与像平面上的对应像点建立方程组求解^[7,8],该方法所建立的模型简洁优美,效率高。但文献[7]、[8]利用 3 个特征点通过解析法来求解,容易造成姿态解不稳定或使得姿态估计的精度很受特征点定位精度的影响。对此,本文提出了一种基于多点模型的人脸姿态估计方法,通过改进传统的活动形状模型(ASM)法,准确地提取人脸特征点后,针对人脸形状特性,使用人脸的多个特征点作为人脸模型,通过最小二乘法优化求解,精确估计 3D 人脸空间姿态。实验结果表明,该新方法不仅可以获得稳定的姿态解,而且与同类方法比较具有良好的姿态估计精确度。

2 基于 ASM 的特征点提取

活动形状模型是一种基于统计学的、灰度和形状分离的可变形模型^[14],它的形状是由点的集合表示的。通过模型的变形使得 ASM 与图像的目标形状吻合,而模型的变化限制在一定的模式之中,这些变化模式是通过训练集合进行主分量分析获得的。具体的定位是采用形状模型和模型上点的局部灰度分布相结合的方法来实现的,而相似度的计算则是通过利用模型上点的当前局部灰度分布和经验灰度分布匹配的加权和来表示。在达到最佳匹配时,待检测图像的形状模型参数和模型上每点的局部灰度分布参数也就同时确定了。它可分为训练阶段和搜索阶段。

2.1 训练阶段

从由 N 幅训练图像构成的训练集中提取形状信息,最简单的办法是在每幅训练图片上手工挑选 n 个特征点,并且代表每幅图像各个特征部位轮廓的点的个数相同,将每幅图像对应的若干特征点的坐标按照一定的顺序放在形状向量里,得到的若干个形状向量就是 ASM 的点分布模型:

$$\mathbf{X}_i = (x_{i,1}, y_{i,1}, x_{i,2}, y_{i,2}, \dots, x_{i,n}, y_{i,n})^T \quad (1)$$

式中, $(x_{i,j}, y_{i,j})$ 是第 i 幅图像的第 j 个特征点的坐标, $i = 1, 2, \dots, N; j = 1, 2, \dots, n$ 。

为了能够比较来自不同形状相对应的点,首先需将训练集合中的形状进行校准,校准就是以某个形状为基础,对其他形状进行缩放、旋转和平移使得它们与基准形状尽可能接近。然后计算校准后形状向量的协方差矩阵 \mathbf{S} ,再通过主分量分析(PCA),取 \mathbf{S} 前 t 个特征值对应的特征向量,记为 $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_t)$,这样任何形状向量都可以用前 t 个特征向量的线性组合来表示:

$$\mathbf{x} = \bar{\mathbf{X}} + \mathbf{P} \cdot \mathbf{b} \quad (2)$$

式中, $\mathbf{b} = (b_1, b_2, \dots, b_t)^T$ 是控制前 t 个模式的系数, $\bar{\mathbf{X}}$ 表示特征点均值。不同的 \mathbf{b} 对应不同的形状。

其次,为了调整形状,还要分析和提取训练图像特征点周围的局部灰度信息,并计算其经验灰度分布。在每一幅图像上通过每一个特征点,沿着该点邻近的两点连线的垂直方向上,分别在特征点的两边取 k 个像素的灰度值,然后以法线段上某个点为中心,截取 $k \times k$ 的窗口,该窗口内像素的平均值就作为该中心点的灰度值,每一个点对应的灰度模板就是关于这 $2k + 1$ 个灰度值(包括特征点的灰度)的统计模型,即均值和方差。

对于第 i 幅图像的第 j 个标记点,其灰度层可表示为

$$\mathbf{g}_{i,j} = (g_{i,j_1}, g_{i,j_2}, \dots, g_{i,j_{(2k)}}, g_{i,j_{(2k+1)}})^T \quad (3)$$

式中, $i = 1, 2, \dots, N; j = 1, 2, \dots, n$ 。

对 $\mathbf{g}_{i,j}$ 求差分可得

$$d\mathbf{g}_{i,j} = (g_{i,j_2} - g_{i,j_1}, g_{i,j_3} - g_{i,j_2}, \dots, g_{i,j_{(2k+1)}} - g_{i,j_{(2k)}})^T \quad (4)$$

对 $d\mathbf{g}_{i,j}$ 进行标准化可得

$$\mathbf{y}_{i,j} = \frac{d\mathbf{g}_{i,j}}{\left(\sum_{l=1}^{2k} |d\mathbf{g}_{i,j_l}|\right)} \quad (5)$$

计算对于第 j 个特征点的灰度层差分均值和方差分别为

$$\bar{\mathbf{y}}_j = \frac{1}{N} \sum_{i=1}^N \mathbf{y}_{i,j} \quad (6)$$

$$\mathbf{C}_{y,j} = \frac{1}{N} \sum_{i=1}^N (\mathbf{y}_{i,j} - \bar{\mathbf{y}}_j)(\mathbf{y}_{i,j} - \bar{\mathbf{y}}_j)^T \quad (7)$$

2.2 搜索阶段

具体的定位是采用形状模型和模型上点的局部灰度分布相结合的方法来实现的,而相似度的计算则是通过利用模型上点的当前局部灰度分布和经验

灰度分布匹配的加权和(即马氏距离)来表示。在达到最佳匹配时,待检测图像的形状模型参数和模型上每点的局部灰度分布参数也就同时确定了。

在搜索之前,首先需要将待检测图像进行形状校准。由式(2)可以看出,选定不同的 \mathbf{b} 可以得到不同的形状,假定有初始形状 \mathbf{X} ,它可由基准形状 \mathbf{x} 通过旋转 θ 和缩放 s 得到,可用下式表示:

$$\mathbf{X} = M(s, \theta) \cdot \mathbf{x} + \mathbf{X}_c \quad (8)$$

其中,

$$\mathbf{X}_c = (x_c, y_c, x_c, y_c, \dots, x_c, y_c)^T, (x_c, y_c)$$

是图像模型的中心,

$$\mathbf{x} = \bar{\mathbf{X}} + \mathbf{P} \cdot \mathbf{b}$$

$$M(s, \theta) \cdot \mathbf{x} = \begin{pmatrix} s \cos \theta & -s \sin \theta \\ s \sin \theta & s \cos \theta \end{pmatrix} \mathbf{x}$$

它是旋转和缩放的几何变换。

由此可见,如果有合适的初始形状,就可以更快速、准确地达到最佳匹配。针对这点及为了适应多姿态变化图像特征点的准确提取,引入一种仿射变换不变性初始化原则(affine transform insensitive initialization algorithm,简称 ATIIA)^[15],对传统的ASM进行了改进。设 \mathbf{F} 和 \mathbf{L} 分别表示被选择的特征点 m 的位置和标志点 n 的位置,定义如下:

$$\begin{cases} \mathbf{F} = (x_1^F, y_1^F, x_2^F, y_2^F, \dots, x_m^F, y_m^F)^T \\ \mathbf{L} = (x_1^L, y_1^L, x_2^L, y_2^L, \dots, x_n^L, y_n^L)^T \end{cases} \quad (9)$$

假设特征点和标志点之间存在线性关系,则初始模型可以被表示为

$$\mathbf{X} = \mathbf{L} - \bar{\mathbf{L}} = \mathbf{B} \cdot (\mathbf{F} - \bar{\mathbf{F}}) \quad (10)$$

式中, $\bar{\mathbf{F}} = (\bar{x}_{f_1}, \bar{y}_{f_1}, \bar{x}_{f_2}, \bar{y}_{f_2}, \dots, \bar{x}_{f_m}, \bar{y}_{f_m})^T$

$$\bar{\mathbf{L}} = (\bar{x}_{l_1}, \bar{y}_{l_1}, \bar{x}_{l_2}, \bar{y}_{l_2}, \dots, \bar{x}_{l_n}, \bar{y}_{l_n})^T \quad (11)$$

\mathbf{B} 是一个 $2n \times 2m$ 的矩阵可以从训练中通过线性回归得到, \bar{x}, \bar{y} 表示特征点的 x 和 y 坐标的平均值。

证明:基于仿射变换有

$$\begin{aligned} \mathbf{L}' &= \mathbf{L} \cdot \mathbf{R} + \mathbf{T}_L \\ \bar{\mathbf{L}}' &= \bar{\mathbf{L}} \cdot \mathbf{R} + \mathbf{T}_L \\ \mathbf{F}' &= \mathbf{F} \cdot \mathbf{R} + \mathbf{T}_F \\ \bar{\mathbf{F}}' &= \bar{\mathbf{F}} \cdot \mathbf{R} + \mathbf{T}_F \end{aligned} \quad (12)$$

式中,

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} s \cdot \cos \theta & s \cdot \sin \theta \\ -s \cdot \sin \theta & s \cdot \cos \theta \end{bmatrix} \\ \mathbf{T}^L &= [t_x^L, t_y^L, t_x^L, t_y^L, \dots, t_x^L, t_y^L]^T_{2n} \\ \mathbf{T}^F &= [t_x^F, t_y^F, t_x^F, t_y^F, \dots, t_x^F, t_y^F]^T_{2m} \end{aligned} \quad (13)$$

则

$$\begin{aligned} \mathbf{L}' - \bar{\mathbf{L}}' &= (\mathbf{L} \cdot \mathbf{R} + \mathbf{T}_L) - (\bar{\mathbf{L}} \cdot \mathbf{R} + \mathbf{T}_L) \\ &= (\mathbf{L} - \bar{\mathbf{L}}) \cdot \mathbf{R} = \mathbf{B} \cdot (\mathbf{F} - \bar{\mathbf{F}}) \cdot \mathbf{R} \\ &= \mathbf{B} \cdot [(\mathbf{F} \cdot \mathbf{R} + \mathbf{T}_F) - (\bar{\mathbf{F}} \cdot \mathbf{R} + \mathbf{T}_F)] \\ &= \mathbf{B} \cdot (\mathbf{F}' - \bar{\mathbf{F}}') \end{aligned} \quad (14)$$

\mathbf{R}, \mathbf{T} 分别代表旋转和平移矩阵。

可见,式(14)表明初始化模型对仿射变换是稳定的。利用此初始模型,ASM的搜索可以在较少的循环下取得较好的结果,且更适应多姿态的变化。

对于第 j 个特征点,任意取灰度层中长度为 $2k+1$ 的一部分,记为 $\mathbf{p}(i)$, i 是该部分中心点的位置序号,将 $\mathbf{p}(i)$ 标准化,记为 $\mathbf{h}(i)$ 。定义第 j 个特征点的决策函数,即局部灰度向量和经验灰度向量之间的马氏距离为

$$f_j(i) = (\mathbf{h}(i) - \bar{\mathbf{y}}_j)^T \cdot \mathbf{C}_{y,j}^{-1} \cdot (\mathbf{h}(i) - \bar{\mathbf{y}}_j) \quad (15)$$

寻找使得 $f_j(i)$ 达到最小的 i ,即特征点新的位置所在。根据式(8)可得到新的形状 \mathbf{X} 。将 \mathbf{X} 作为新的初始形状,重复计算直到形状的改变不大,即可获得待检测图像的轮廓,从而获得所要的特征点位置。

3 人脸姿态估计

人脸姿态有6个自由度的变化,即沿 X, Y, Z 轴的平移和绕 X, Y, Z 轴的旋转。对沿 X, Y 的平移,在图像上表现为人脸的位置变化,可以通过统一坐标系实现;对沿 Z 轴的平移,在图像上表现为比例的变化,可以通过比例归一化实现。所以这里只研究人脸绕 X, Y, Z 轴的旋转问题,旋转角分别为 α, β 和 γ 。

3.1 模型的建立

选用左右眼角和嘴角,鼻尖,下巴顶点等共6个点来构成脸的模型。模型坐标系的原点选在鼻尖处,这样,脸的模型实际由5个点的坐标构成

$$(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_5) = \begin{pmatrix} -x_e & x_e & x_m & -x_m & 0 \\ y_e & y_e & -y_m & -y_m & -y_e \\ z_e & z_e & z_m & z_m & z_e \end{pmatrix} \quad (16)$$

式中, \mathbf{p}_i 为模型点在自身坐标系中的坐标, x_e 表示外眼角间距的一半, x_m 为嘴角间距的一半, y_e 为下巴定点到鼻尖的距离, z 表示相应的深度距离,不同的人脸的这3个值是不同的,其差异部分反映了模型与真实人脸的结构差异。

假设模型点 $p_i(x_i, y_i, z_i)$ 对应的像点为 $q_i(\mu_i, \nu_i)$, \mathbf{R} 为姿态旋转矩阵, \mathbf{t} 为平移矩阵, f 为像机焦距。根据透视成像可得:

$$\mu_i = f \frac{\mathbf{r}_1 \cdot \mathbf{p}_i + t_1}{\mathbf{r}_3 \cdot \mathbf{p}_i + t_3}$$

$$\mathbf{R} = (r_1, r_2, r_3)^T = \mathbf{R}(\alpha) \cdot \mathbf{R}(\beta) \cdot \mathbf{R}(\gamma) = \begin{pmatrix} \cos\beta \cdot \cos\gamma & \cos\beta \cdot \sin\gamma & -\sin\beta \\ -\cos\alpha \cdot \sin\gamma + \sin\alpha \cdot \sin\beta \cdot \cos\gamma & \cos\alpha \cdot \cos\gamma + \sin\alpha \cdot \sin\beta \cdot \sin\gamma & \sin\alpha \cdot \cos\beta \\ \sin\alpha \cdot \sin\gamma + \cos\alpha \cdot \sin\beta \cdot \cos\gamma & -\sin\alpha \cdot \cos\gamma + \cos\alpha \cdot \sin\beta \cdot \sin\gamma & \cos\alpha \cdot \cos\beta \end{pmatrix}$$

相应的透视成像模型可表示如下:

$$\mathbf{q}_i = \mathbf{C} \cdot [\mathbf{R}(\alpha, \beta, \gamma) \cdot \mathbf{p}_i + \mathbf{t}] \quad (18)$$

姿态估计问题可表述为已知像机参数 \mathbf{C} , 模型点 $\{p_i, i=1, 2, \dots, n\}$ 和其在像机参数 \mathbf{C} 中的像点 $\{q_i, i=1, 2, \dots, n\}$, 求姿态参数 (α, β, γ) 。

求解方式为

$$(\alpha, \beta, \gamma) = \operatorname{argmin} \left\{ \sum_{i=1}^n \left\| \mathbf{q}_i - \mathbf{C} \cdot [\mathbf{R}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \cdot \mathbf{p}_i + \mathbf{t}] \right\|^2 \right\} \quad (19)$$

3.2 最优化求解

对式(19), 求其最小二乘优化解。通过线性化调整最小二乘^[16], 过程如下: 在检测图像坐标时总是存在误差的, 则

$$\begin{cases} \mu_i = \hat{\mu}_i + \varepsilon_i \\ \nu_i = \hat{\nu}_i + \delta_i \end{cases} \quad (20)$$

$i = 1, 2, \dots, n$

式中, $(\hat{\mu}_i, \hat{\nu}_i)$ 是被测图像点, $(\varepsilon_i, \delta_i)$ 是相应的测量误差。类似的有

$$\begin{cases} \alpha = \alpha^0 + \Delta\alpha \\ \beta = \beta^0 + \Delta\beta \\ \gamma = \gamma^0 + \Delta\gamma \end{cases} \quad (21)$$

式中, $\alpha^0, \beta^0, \gamma^0$ 是近似值, $\Delta\alpha, \Delta\beta, \Delta\gamma$ 是它们相应的误差。

假设:

$$\begin{aligned} F_{i1} &= \mu_i - f \cdot \frac{\mathbf{r}_1 \cdot \mathbf{p}_i + t_1}{\mathbf{r}_3 \cdot \mathbf{p}_i + t_3} \\ F_{i2} &= \nu_i - f \cdot \frac{\mathbf{r}_2 \cdot \mathbf{p}_i + t_2}{\mathbf{r}_3 \cdot \mathbf{p}_i + t_3} \end{aligned} \quad (22)$$

在此, 不考虑平移量 \mathbf{t} , 通过牛顿第一定律进行线性化得:

$$F_{i1} = F_{i1}^0 + \varepsilon_i + a_{i1} \cdot \Delta\alpha + a_{i2} \cdot \Delta\beta + a_{i3} \cdot \Delta\gamma \quad (23)$$

$$F_{i2} = F_{i2}^0 + \delta_i + b_{i1} \cdot \Delta\alpha + b_{i2} \cdot \Delta\beta + b_{i3} \cdot \Delta\gamma$$

式中, $a_{i1} = \left(\frac{\partial F_{i1}}{\partial \alpha} \right)^0$, $a_{i2} = \left(\frac{\partial F_{i1}}{\partial \beta} \right)^0$, $a_{i3} = \left(\frac{\partial F_{i1}}{\partial \gamma} \right)^0$,

$$\nu_i = f \frac{\mathbf{r}_2 \cdot \mathbf{p}_i + t_2}{\mathbf{r}_3 \cdot \mathbf{p}_i + t_3} \quad (17)$$

式中,

$$\mathbf{t} = (t_1, t_2, t_3)^T$$

$$\begin{pmatrix} \cos\beta \cdot \sin\gamma & \cos\beta \cdot \cos\gamma + \sin\alpha \cdot \sin\beta \cdot \sin\gamma & -\sin\beta \\ \cos\alpha \cdot \cos\gamma + \sin\alpha \cdot \sin\beta \cdot \sin\gamma & \sin\alpha \cdot \cos\beta & \sin\alpha \cdot \cos\beta \\ -\sin\alpha \cdot \cos\gamma + \cos\alpha \cdot \sin\beta \cdot \sin\gamma & \cos\alpha \cdot \cos\beta & \cos\alpha \cdot \cos\beta \end{pmatrix}$$

$$b_{i1} = \left(\frac{\partial F_{i2}}{\partial \alpha} \right)^0, b_{i2} = \left(\frac{\partial F_{i2}}{\partial \beta} \right)^0, b_{i3} = \left(\frac{\partial F_{i2}}{\partial \gamma} \right)^0$$

其中, 上标“0”表示 α, β, γ 取 $(\alpha^0, \beta^0, \gamma^0)$ 时所对应的函数值, 当 $F_{i1} = F_{i2} = 0$, 线性化方程可被表示成矩阵形式:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} \\ b_{n1} & b_{n2} & b_{n3} \end{pmatrix} \cdot \begin{pmatrix} \Delta\alpha \\ \Delta\beta \\ \Delta\gamma \end{pmatrix} = \begin{pmatrix} -F_{i1}^0 \\ -F_{i2}^0 \\ \vdots \\ -F_{in1}^0 \\ -F_{in2}^0 \end{pmatrix} - \begin{pmatrix} \varepsilon_1 \\ \delta_1 \\ \vdots \\ \varepsilon_n \\ \delta_n \end{pmatrix} \quad (24)$$

对方程(24)采用奇异值分解来求解, 计算 $\mathbf{\Delta} = (\Delta\alpha, \Delta\beta, \Delta\gamma)$ 来更新 $\mathbf{\Lambda} = (\alpha^0, \beta^0, \gamma^0)$, 直到误差达到设定的阈值, 结束循环, 输出估计的姿态结果。

4 实验

实验采用 CMU PIE 人脸数据库, 该数据库包含了 68 人, 共 41 368 幅人脸图像。对其中 $\beta = 0^\circ, \pm 22.5^\circ, \pm 45^\circ$ 5 个姿态进行了测试, 结果表示形式为 (α, β, γ) , 单位为度, 部分结果如图 1(a) ~ (i) 所示。

在一般应用系统中, 对绕 Y 轴旋转的斜视图像应用较多, 且因为 CMU PIE 人脸数据库中图像的 β 值较具体明确, 所以通过测试 β 来检测算法的精确程度。根据对称性, 本文给出 $\beta = 0^\circ, 22.5^\circ, 45^\circ$ 时的绝对平均误差表, 如表 1 所示。

表 1 β 绝对平均误差

Tab. 1 Absolute average error of β

β 值	0°	22.5°	45°
相应的绝对平均误差	1.56°	2.37°	4.11°

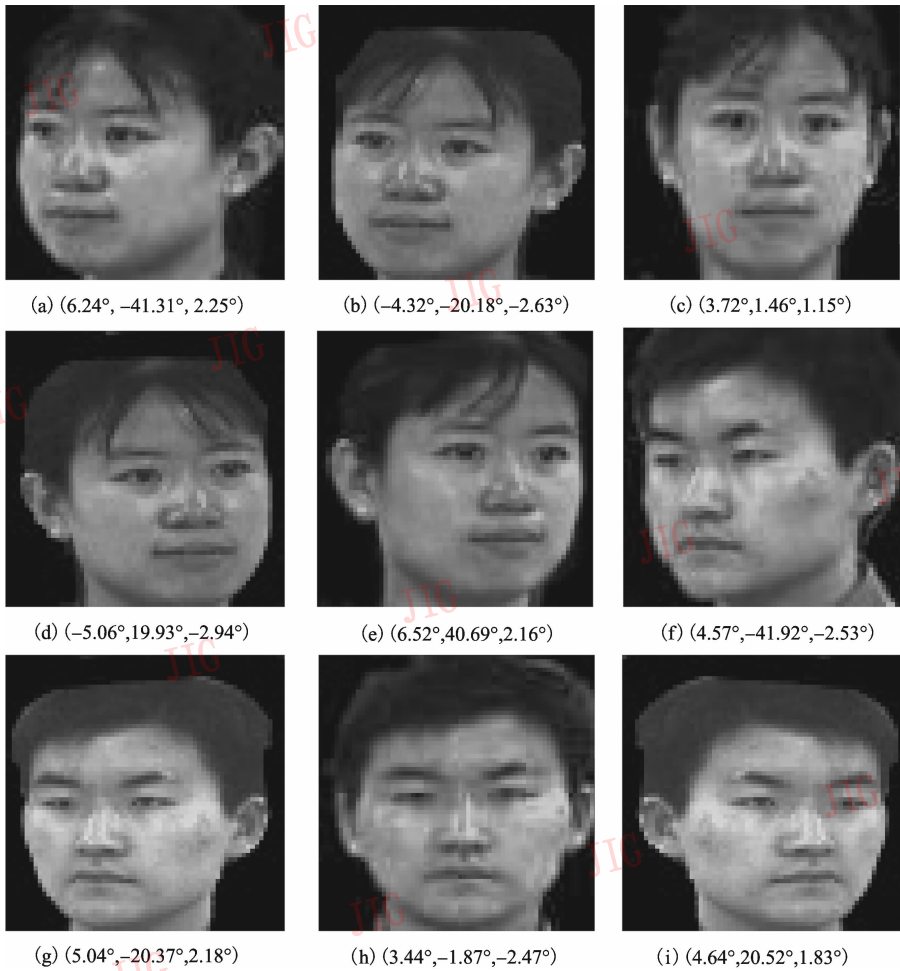


图 1 部分估计结果

Fig. 1 Some of estimated results

特征点定位往往是存在误差的,假设用上述的 ASM 方法所得的特征点定位误差为 0 像素,且以此时所得的估计结果为基准,图 2 显示了当特征点定

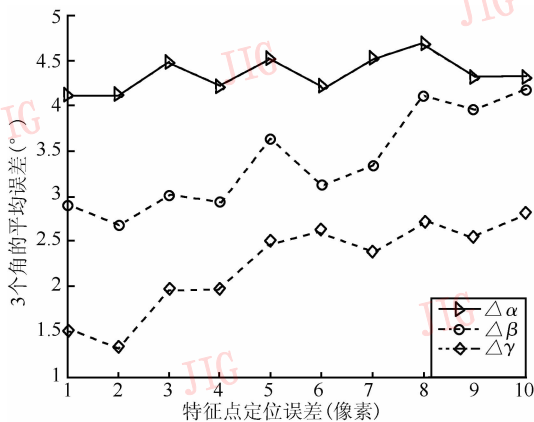


图 2 当特征点定位有误差时姿态估计误差的变化

Fig. 2 The change of estimated error with the error of feature points localization

位有误差时姿态估计的性能。其中特征点定位误差用定位值偏离基准位置的像素点距离表示。

曲线表明,虽然随着特征点的定位误差的增加,3 个角度的误差都有增大的趋势, α 和 β 的误差增长比较明显,但仍低于 5° 。

表 2 给出了用本文方法与文献[7]、[8]方法测得的 β 估计结果比较。

表 2 估计结果对比

Tab. 2 Contrast of estimated results

方法	姿态估计结果
文献[7]	姿态解不稳定
文献[8]	β 角的平均估计误差为 4.35°
本文	β 角的平均估计误差为 2.68°

可见,本文算法不仅可以获得稳定的姿态解,而

且具有良好的姿态估计精度。

5 结 论

理论研究与在 CMU PIE 上的实验结果表明:

(1) 利用本文算法得到的 β 估计值绝对对平均误差约为 2.68° ;

(2) 本文算法不仅可以获得稳定的姿态解,而且当特征点定位有误差时姿态估计误差仍低于 5° ;

(3) 本文算法与同类方法比较具有更好的姿态估计精确度。

可以预见,通过用图像序列较丰富的图像信息,或通过 2D 人脸图像建立人脸 3D 模型来估计人脸 3D 姿态,将使算法更适应“自遮挡”、光照变化和多表情变化情况。

参考文献 (References)

- 1 Darrell Trevor, Baback Moghaddam, Pentland A P. Active face tracking and pose estimation in an interactive room [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], San Francisco, CA, USA, 1996: 67 ~ 72.
- 2 Murase H, Nayar S K. Visual learning and recognition of 3D objects from appearance [J]. International Journal of Computer Vision, 1995, 14 (1): 5 ~ 24.
- 3 Chen Q, Wu H, Shimada T. A robust algorithm for 3D head pose estimation [A]. In: Proceedings of IEEE International Conference on Multimedia Computing and Systems [C], Nara, Japan, 1999: 697 ~ 702.
- 4 Bruske J, Abraham Mumm E, Pauli J, *et al.* Head pose estimation from facial images with subspace neural networks [A]. In: Proceedings of Internal Neural Networks and Brain Conference [C], Beijing, China, 1998: 528 ~ 531.
- 5 Khotanzad A, Liou J H. Recognition and pose estimation of unoccluded three-dimensional objects from a two-dimensional perspective view by banks of neural networks [J]. IEEE Transactions on Neural Network, 1996, 7(4): 897 ~ 906.
- 6 Li S Z. Learning multi-view face subspaces and facial pose estimation using independent component analysis [J]. IEEE Transactions on Image Process, 2005, 14(6): 705 ~ 712.
- 7 Alter T D. 3-D pose from 3 points using weak-Perspective [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(8): 802 ~ 808.
- 8 Mazumda Debasis R, Dutta Santanu, Mitra Soma. Automatic feature detection of a face and recovery of its pose [A]. In: Communicated to Journal of IETE [C], Washington, 2003: 505 ~ 511.
- 9 Kruger Volker, Bruns Sven, Somme Gerald R. Efficient head pose estimation with gabor wavelet networks [M]. Bristol, Great Britain: Great, 2000: 11 ~ 14.
- 10 Gree A, Cipolla R. Estimation gaze from a single view of face [A]. In: Proceedings of IEEE International Conference on Pattern Recognition [C], Seattle, WA, USA, 1994: 758 ~ 760.
- 11 Ebisawal Yoshinobu. Face pose estimation based on 3D detection of pupils and nostrils [A]. In: Proceedings of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems [C], Giardini Naxos, Italy, 2005: 6.
- 12 Nikolaidis Athanasios, Pitas Ioannis. Facial feature extraction and pose determination [J]. Pattern Recognition, 2000, 33 (6): 1783 ~ 1791.
- 13 Hu Yu-xiao, Chen Long-bin, Zhou Yi, *et al.* Estimating face pose by facial asymmetry and geometry [A]. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04) [C], Seoul, Korea, 2004: 2122 ~ 2128.
- 14 Wan Kwork Wai, Lam Kin Man, Chong Kit. An accurate active shape model for facial feature extraction [J]. Pattern Recognition Letters, 2005, 26(12): 2409 ~ 2423.
- 15 Liu Yang-hua, Li Yang, Tao Lin-mi, *et al.* Multi-view face alignment guided by several facial feature points [A]. In: Proceedings of the Third IEEE International Conference on Image and Graphics (ICIG'04) [C], San Francisco, DC, USA, 2004: 2244 ~ 2247.
- 16 Haralick Robert M, JOO Hyonam, Lee Chung-Nan, *et al.* Pose estimation from corresponding point data [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19(6): 1426 ~ 1446.