

基于单幅图像的人脸三维姿态估计

张志刚¹, 周明全^{1,2}, 耿国华¹

(1. 西北大学 信息科学与技术学院, 陕西 西安 710069;

2. 北京师范大学 信息科学与技术学院, 北京 100875)

(zzg.gg@163.com)

摘 要:针对单张人像的三维姿态计算, 结合面貌测量和射影几何的理论提出了一种方法: 首先在人面部的平面区域内, 选取眼角点, 口角点, 鼻翼点建立人脸模型; 然后根据人脸平面上两个相互垂直的特征线投影到照片上的灭点位置, 求出人脸平面的旋转方向。该方法特征点易于标定, 且无需任何的辅助设备和先验知识, 具有一定的实用性。

关键词:姿态估计; 颅像重合; 灭点

中图分类号: TP391 **文献标识码:** A

3D pose estimation based on single face image

ZHANG Zhi-gang¹, ZHOU Ming-quan^{1,2}, GENG Guo-hua¹

(1. College of Information Science and Technology, Northwest University, Xi'an Shaanxi 710069, China;

2. College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

Abstract: Aiming at 3D pose computation from single facial image, a method was proposed based on the theory of anthropometry in face and projective geometry: first, corners of eyes and mouth, and nose wing points were selected to establish facial model within facial plane; then, the orientation of face was determined based on the location of vanishing points in image, which were resulted from two perpendicular features lines in perspective project. This method can be implemented without any auxiliary equipments and prior knowledge, distinct feature points are easy to be located, and the practicability can be verified.

Key words: pose estimation; photographic superimposition; vanishing points

0 引言

颅像重合法是指对一张失踪人留下的头部照片负片与一个身源不明的颅骨像负片进行重叠影像, 根据人类头部颜面与颅骨的法医人类学解剖投影关系所确立的一系列鉴定指标, 以重叠时能否达到解剖关系的一致, 来确定颅骨与照片是否为同一人的一种身份认证技术。其中重合摄影角度是关键技术之一^[1], 即从一个人的头部二维照片中估计出三维的姿态, 以此调整颅骨的姿态使二者保持摄影角度的一致。由于人脸姿态的估计在许多领域都有着广泛的应用, 所以一直以来都是个研究的热点。一些学者提出对颅骨每隔 5° 拍摄一张照片, 然后选出同人像角度最接近的与之重合, 此类方法每次需拍摄至少十张不同角度的颅骨负片, 操作繁琐成本高, 且缺乏科学量化的标准, 易出现误差。文献[2]利用人脸的三维模型和仿射对应原理从视频图像序列中估计人脸的空间姿态, 但是需要先得到正面的平行投影; 文献[3~4]采用计算机变形技术生成基于单视图的多姿态图像样本库, 再将待测照片与样本图像进行匹配以估计出人脸的姿态。这些方法的结果很大程度上依赖于样本的数量, 且计算开销很大。另一类方法是利用面部特征模型进行姿态估计^[5~9]。文献[5]利用外眼点和嘴角点, 通过空间几何计算人脸姿态, 要求先得到两个外眼点的空间距离; 文献[6]使用两个外眉点和嘴中心, 利用视频流序列进行姿态估计, 需要同等状态下的人脸正视

参照图作为辅助信息; 文献[7]根据四个眼角点和鼻尖点构造的三角形的变化估计出姿态。这类方法的特点是: 计算结果依赖于特征定位的准确与否, 另外更重要的是它们都要求事先满足一定的条件, 如已知空间距离, 焦距等。而在颅像重合技术中, 除了一张失踪人留下的照片, 没有其他任何可参考的信息, 因此以上方法并不适用。我国学者在鉴定颅骨身源技术研究中提出了一个完整、系统的方法: 即通过活体测量利用生物统计技术, 用水平偏转指数和仰俯指数, 根据在透视成像条件下面部特征点间距的比值和角度偏转的回归关系, 来推算人像的角度^[1]。如图 1 所示。

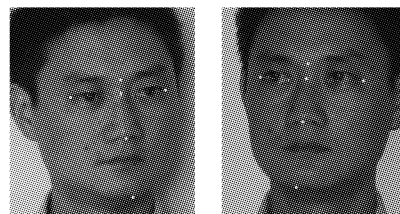


图 1 人像偏转与特征间距变化示意图

该方法无需任何辅助设备, 仅测量出四个数值求得两个指数, 即可计算出偏转角度, 但易受样本数量、分布等影响, 且一些特征位置不易准确定位, 对表情变化较敏感。在上述研究基础上提出一种仅从单幅照片中估计人脸三维姿态的方法: 这种方法计算简便, 且没有对先验知识的要求, 所选特征点易于标定, 只要照片没有残缺, 特征点可见, 均可适用。

收稿日期: 2006-10-30; 修订日期: 2007-01-15 基金项目: 国家自然科学基金资助项目(60573179)

作者简介: 张志刚(1970-), 男, 陕西西安人, 博士研究生, 主要研究方向: 数字图像处理、计算机视觉; 周明全(1954-), 男, 陕西黄陵人, 教授, 博士生导师, 主要研究方向: 计算机图形学、数字图像处理、科学计算可视化; 耿国华(1955-), 女, 山东莱州人, 教授, 博士生导师, 主要研究方向: 模式识别、工程数据库。

1 姿态估计的面部模型

人脸姿态的变化有六个自由度,即沿 X 、 Y 、 Z 轴的平移和绕 X 、 Y 、 Z 轴的旋转,如图2所示。

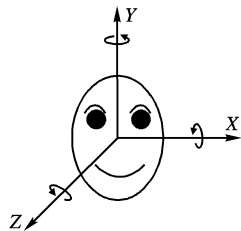
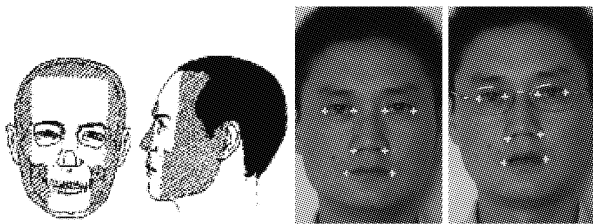


图2 人脸姿态变化示意图

对沿 X 、 Y 轴的平移,在图像上表现为人脸位置的变化,对沿 Z 轴的平移,则表现为比例的变化,这些都可通过适当的方法进行处理,故姿态变化主要是由绕 X 、 Y 、 Z 轴的旋转组成。根据人脸测量统计的知识,在人面部存在着一个平面区域(相对平行区),这部分由发际点、眼角点、口角点和颌下点连线而构成平面^[1]。由于该区域内特征器官基本处在一个刚性的平面上,且在正视状态下该平面与 Z 轴垂直,两个外眼角点的连线与 X 轴平行,因此人脸的姿态估计也就转化成为求该人脸平面的旋转角度问题。本文采用两个外眼角点,两个内眼角点,两个鼻翼点,和两个口角点作为面部模型的特征点。如图3所示,图3(a)中的白色区域为相对平行区,图3(b)为区域内的模型特征点。



(a) 脸部相对平行区

(b) 面部模型特征点

图3 人脸部的特征模型

应用这些面部模型的特征点进行姿态估计时,可利用以下的面貌形态特征:

- 1) 人脸部的眼角点在空间上连成一条直线,并且当人脸处于正视状态时,此直线平行于 X 轴;
- 2) 在空间上,两个口角点的连线与眼角点的连线平行;
- 3) 两条内眼角点和鼻翼点的连线(左右各一条)相互平行,且垂直于口角点连线和眼角点连线;
- 4) 以上四条直线在一个平面上。

以上这些特征点在可见的状态下都具有相当显著的特征,因此相比于鼻尖点等特征点更容易精确地定位,而且它们具有很好的表情无关性和稳定性,即便有人脸模式(如发型、眼镜等)的干扰,也还是容易被分辨出来。

2 人脸姿态估计

图4所示为面部模型所在的空间坐标系。其中 O 为摄影的焦点,也是空间坐标系 O - XYZ 的原点,以视线的方向作为 Z 轴, Z 轴与相机光轴重合。像平面坐标系为 u - v ,其原点为光轴与像平面的交点且与像平面的中心点重合。 f 是焦距。 $E1$ 、 $E2$ 、 $E3$ 、 $E4$ 分别代表了空间中人脸上的两个外眼角点和两个内眼角点, $N1$ 、 $N2$ 是两个鼻翼点, $M1$ 、 $M2$ 是口角点,它们在像平面上投影点分别对应为 $e1$ 、 $e2$ 、 $e3$ 、 $e4$ 、 $n1$ 、 $n2$ 、 $m1$ 、 $m2$ 。

由于 $E1$ 、 $E2$ 的连线与 $M1$ 、 $M2$ 的连线平行,因此这两条直

线若与像平面不平行,则它们在像平面上的投影线 $e1e2$ 和 $m1m2$ 将汇聚为一灭点,即图4中的 $vp1$ 。设向量 $M = \{i, j, k\}$ 代表 $E1$ 、 $E2$ 连线的方向向量,则平面 $OE1E2$ 与像平面的交

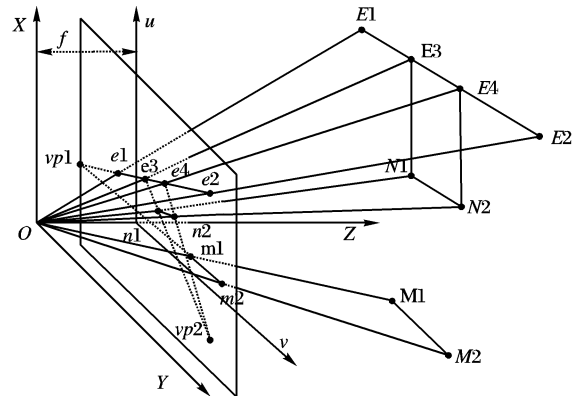


图4 空间坐标系与人脸面部模型

线,和平面 $OM1M2$ 与像平面的交线可分别通过以下两组方程式求得:

$$\begin{cases} \{x, y, z\} \cdot (\overrightarrow{OE1} \times \{i, j, k\}) = 0 \\ z = f \end{cases} \quad (1)$$

$$\begin{cases} \{x, y, z\} \cdot (\overrightarrow{OM1} \times \{i, j, k\}) = 0 \\ z = f \end{cases}$$

由(1)求得两条交线后,可得到灭点 $vp1(x, y, z)$ 的坐标为:

$$\begin{cases} x = f \times \frac{i}{k} \\ y = f \times \frac{j}{k} \\ z = f \end{cases} \quad (2)$$

由(2)可得:

$$M = \left\{ \frac{x \times k}{f}, \frac{y \times k}{f}, k \right\} \quad (3)$$

上式说明根据灭点在像平面的坐标和焦距,就可以求出空间一组平行线的方向向量。同理内眼角点和鼻翼点的连线 $E3$ 、 $N1$ 与 $E4$ 、 $N2$ 相互平行,它们也会在像平面上产生一个灭点,即图中的 $vp2$ 点。类似地设向量 $N = \{m, n, l\}$ 代表 $E3$ 、 $N1$ 连线的方向向量,则 $vp2(x', y', z')$ 的坐标和该向量分别为:

$$\begin{cases} x' = f \times \frac{m}{l} \\ y' = f \times \frac{n}{l} \\ z' = f \end{cases} \quad (4)$$

$$N = \left\{ \frac{x' \times l}{f}, \frac{y' \times l}{f}, l \right\} \quad (5)$$

由于 $E1$ 、 $E2$ 、 $E3$ 、 $E4$ 、 $N1$ 、 $N2$ 、 $M1$ 、 $M2$ 都在一个平面上,若记该人脸平面的法线向量为 F ,可得:

$$F = M \times N \quad (6)$$

将(3)(5)代入(6)得:

$$F = \left\{ y - y', x' - x, \frac{xy' - x'y}{f} \right\} \quad (7)$$

又 M 垂直于 N ,故它们的内积为零,即:

$$\left\{ \frac{x \times k}{f}, \frac{y \times k}{f}, k \right\} \cdot \left\{ \frac{x' \times l}{f}, \frac{y' \times l}{f}, l \right\} = 0 \quad (8)$$

将(8)代入(7)式,经整理得:

$$F = \left\{ y - y', x' - x, \frac{xy' - x'y}{\sqrt{-(xx' + yy')}} \right\} \quad (9)$$

上式只与两个灭点的坐标有关,这样就可以求出人脸平面法线的方向向量,从而确定三维旋转姿态。具体而言,设 α 、 β 、 γ 分别代表人脸相对 X 、 Y 、 Z 轴的旋转角度,则它们可通过下式计算得到:

$$\begin{cases} \alpha = \arctan\left[\frac{(x' - x) \times \sqrt{-(xx' + yy')}}{xy' - x'y}\right] \\ \beta = \arctan\left[\frac{(y - y') \times \sqrt{-(xx' + yy')}}{xy' - x'y}\right] \\ \gamma = \arctan[K(e1e2)] \end{cases} \quad (10)$$

其中 $K(e1e2)$ 表示外眼角连线在像平面上的投影线的斜率,即关于 Z 轴的旋转角度就是外眼角连线与水平轴的夹角。

3 实验与结论

测试图片包括各个方向不同程度偏转的姿态,考虑到具体的应用环境,人脸偏转角度都在正负90度之间。图5所示为使用本方法进行姿态估计的结果,其中 α 、 β 、 γ 代表人脸分别相对 X 、 Y 、 Z 轴的旋转角度,意义与(10)式同。

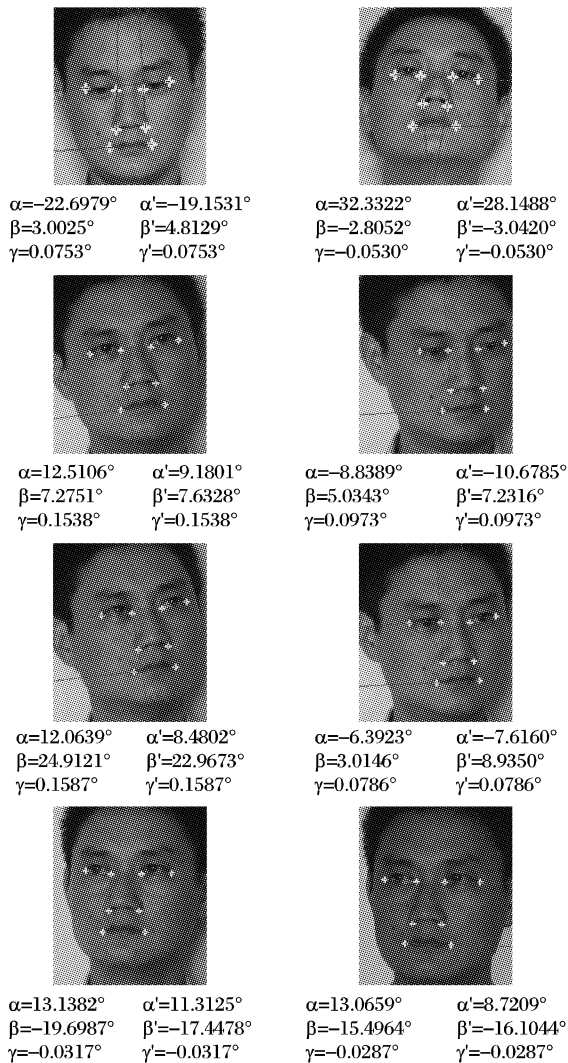


图5 姿态估计实验结果

从实验结果上看,估计的角度近似于实际的偏转角度。再将本算法与其他方法进行比较,由于文献[1]亦根据单幅照片去估计面部的姿态,无需其他预设条件,故将二者进行了对比,实验结果如图5所示,其中 α' 、 β' 、 γ' 是采用文献[1]所提出方法计算出的人脸分别相对 X 、 Y 、 Z 轴的旋转角度,可以看出二者的计算结果比较相近。从方法原理上分析,文献[1]操

作简便,对特征点定位的稳定性较强,但易受表情(如张嘴)的影响,本算法的特征模型为面部的两对平行线,相对比较稳定,故受表情变化的影响较小。另外需要指出的是本方法仅适用于人像自然小角度的偏转,当人脸姿态变化较大时,由于面部模型中所需的部分特征点变得不可见,则无法进行计算,而这也是此类方法共同的不足之处,因此如何建立更有效的面部模型,以适用于更大角度的偏转姿态计算将是进一步的研究重点。

本文针对颅像重合技术中欲从一张照片估计出人脸的三维姿态提出了一种方法:充分利用人脸的测量统计知识和结构特征,确定了人脸平面上的两对相互垂直的平行线这一稳定的特征,从而在像平面与人脸空间模型之间建立了对应关系,根据射影几何原理和灭点位置估计出人脸的三维姿态。本方法简便快捷,不需要任何参考信息和辅助设备便能实现,也适用于类似的其他应用环境中。

对于该方法的误差来源,可总结为以下几个方面:

- 1) 图像噪声、脸上的遮盖物等都可能使某些计算所需的特征点定位不准,这样只能估计特征点的位置,从而造成误差;
- 2) 特征点标定者本身技术的娴熟程度也影响到计算的结果;
- 3) 面部特征可能受到表情的影响而产生误差:眼角连线 and 鼻翼眼角连线是相对比较稳定的特征点(线),而嘴角可能会因为不对称的歪斜,或笑容等而发生变化;
- 4) 由于没有其他的参考信息无法对相机进行定标,因此本文算法中假设相机光轴通过像平面中心且与之垂直,只是一种理想的状况,相机本身的畸变也会产生误差;
- 5) 若照片残缺,或经过剪裁、合成,如从合影照中分离而来,则因无法确定其原像的中心位置而无法估计。

因此下一步的研究重点是建立更合理更稳定的面部模型。另外,通过对图像的处理,使特征点更加明显,易于标注,以尽量减少标定者本身的因素对计算结果的影响。

参考文献:

- [1] 兰玉文. 颅像重合法研究与应用[M]. 北京: 群众出版社, 1999.
- [2] 梁国远, 查红彬, 刘宏. 基于三维模型和仿射对应原理的人脸姿态估计方法[J]. 计算机学报, 2005, 28(5): 792 - 800.
- [3] LEE M, RANGANATH S. Pose-invariant face recognition using a 3D deformable model[J]. Pattern recognition, 2003, 36(8): 1835 - 1846.
- [4] 朱长仁, 王润生. 基于单视图的多姿态人脸识别算法[J]. 计算机学报, 2003, 26(1): 104 - 109.
- [5] HO S, HUANG H. An analytic solution for the pose determination of human faces from a monocular image[J]. Pattern Recognition Letters, 1998, 19(11): 1045 - 1054.
- [6] YILMAZ A, SHAH M. Automatic feature detection and pose recovery for faces[A]. The 5th Asian Conference on Computer Vision [C]. Melbourne, 2002. 284 - 289.
- [7] HORPRASERT T, YACOOB T, DAVIS L. Computing 3D head orientation from a monocular image sequence[A]. IEEE International Conference on Automatic Face and Gesture Recognition[C]. Vermont, 1996. 242 - 247.
- [8] GEE A, CIPOLLA R. Estimating gaze from a single view of a face [A]. Proceedings of the 12th International Conference on Pattern Recognition[C]. Jerusalem, 1994. 758 - 760.
- [9] WANG J, SUNG E. Pose determination of human faces by using vanishing points[J]. Pattern Recognition, 2001, 34(12): 2427 - 2445.