第 5 卷(A 版) 第 10 期
2000 年 10 月

中国图象图形学报
Journal of Image and Graphics

Vol. 5(A), No. 10
Oct. 2000

# 基于主元分析法的行为识别

胡长勃 冯 涛 马颂德 卢汉清

(中国科学院自动化所模式识别国家重点实验室, 北京 100080)

**摘 要** 通过研究,建立了一个基于主元分析的识别人体行为的系统.其方法是通过在 H、S、I 颜色空间对皮肤颜色建立高斯模型,结合运动限制和区域连续性,系统地分割并跟踪人脸和双手.然后,在 PCA 框架下,表示脸和手的运动参数曲线,并和范例进行匹配,这种通过对行为在时空域变化的建模方法,能在行为主体和成象条件有变化的情况下识别行为.以太极拳式为例,来验证方法和系统的效果,实验结果证明了此方法误识率低,有一定的鲁棒性,可应用于建立基于运动语义识别的视频检索、高效视频编码、自动教练等场合.

**关键词** 行为识别, 区域跟踪, 主元分析

**中图法分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2000)10-0818-07

# PCA Based Human Activity Recognition

HU Chang-bo, FENG Tao, MA Song-de, LU Han-qing

(*National Lab of Pattern Recognition, the Institute of Automation, Chinese Academy of Sciences, Beijing* 100080)

**Abstract** This paper presents an approach to recognize human activity based on principal component analysis (PCA). By a Gaussian model of skin color in HSI color space, our system tracks human face and hands incorporating with the constraints of motion and region continuity. A constant acceleration motion estimation and Schwarz representation based shape matching are applied to create correspondence between frames. Then motion parameters of faces and hands are represented and matched with the parameter curves of exemplars in PCA framework. Through modeling the spatio-temporal variants of each type of activity, recognition can be achieved although subject and imaging condition are different from those of exemplars. Examples of Taiji postures recognition are studied and discussed to illustrate our method. The experiment shows that this activity recognition approach is of low confusion rate and robust in some degree. We believe this approach can be applied to develop an activity interpretation system. Applications fields of this work include indexing video based on motion semantic description, assistant exercise training, video surveillance efficient video coding, etc.

**Keywords** Activity recognition, Area tracking, Principal component analysis

## 0 引 言

人的动作识别在理论和应用上都很重要.它的应用非常广泛,如视觉监控、虚拟现实、人机交互、运动技能训练等场合.从视频上学习姿势时,如果计算机能从学习者的动作上判断出他需要哪一段视频,那么就可以大大方便学习者,并提高效率.动作识别的过程可以分为两个阶段:第一阶段是对人体运动的跟踪和分析;第二阶段是对运动的解释.

在第一阶段,通常用两种方法,其中,一种方法是应用人体模型[1],从 1D 到 3D 都有,而基于模型的方法有一个共同的特点,就是需要标定出人体的组成部分;第二种方法则不使用模型,而是把人的运动视为一个整体,如在文献[2]中,就介绍了一种通过计算相邻帧间的光流场,并把光流场分成一组窗口,再对每个窗口的运动幅值求和,以构成一个高维的特征矢量,来用于识别的方法.Davis 采用类似的

基金项目:国家自然科学基金资助项目(69805005)和国家"973"重点基础研究计划(G1998030500)

收稿日期:1999-11-01;改回日期:2000-03-01

方法，用图象序列的运动能量图象(MEI)和运动历史图象(MHI)来解释人的运动[3]，但其中的运动不是光流场，而是用相邻帧间差分得到的. 但这两种方法都有其缺点，如在模型法中，从身体各部分的运动提取出的观察向量，固然可以使我们能准确识别复杂的运动，但是，自动跟踪和标定肢体是异常困难的，特别是在复杂场景下，而且如何解决遮挡也是一个挑战，另外，计算量也需要考虑，由于这种方法的维数较高，从而增加了识别时的困难；在无模型法中，虽然运动的计算容易进行，但是运动的意义很难确定，如一个人挥手和一个树枝的摇摆运动是相同的，在没有足够的限制和先验知识的情况下，结论可能是荒谬的. 根据我们的观察和经验，通常只需识别人体的少部分，就足够能识别很多种运动，且能满足一些识别目的. 由于脸和手是人体的重要器官，它们有特定的皮肤颜色，且在正常情况下都是裸露的，因此对跟踪有很大帮助. 在系统中，若在 HSI 空间建立皮肤颜色的两维高斯模型，即可采用 $H$ 和 $S$ 分量来进行识别，因为 $H$、$S$ 对光照变化不敏感. 为了避免背景区域上相同颜色的干扰，首先需通过用减去背景的方法抽取出人体剪影，然后只在剪影中，通过对皮肤颜色区域聚类，来分割出脸和手. 接着通过运动估计和区域连续性在相继帧中找到对应的三块区域. 在发现和跟踪到脸和手以后，由它们的位置参数来构成运动矢量，然后转入运动解释阶段.

第二阶段也有多种方法，如隐马尔科夫模型(HMM)[4]，主元分析(PCA)[5]以及概率模型[6]等. 我们采用 PCA 法来进行动作建模和识别，因为 PCA，或者说是特征空间表示法可以在一个大量缩小的维数上表示和构造动作模型. 在文献[5]的基础上，考虑了如何通过在时空域上的归一化，来识别因为不同的干扰和变形而有所不同的动作. 另外，针对实际问题，还考虑了不同视角的情况. 上述两个阶段，构成了本文提出的行为识别系统，其流程图见图 1.

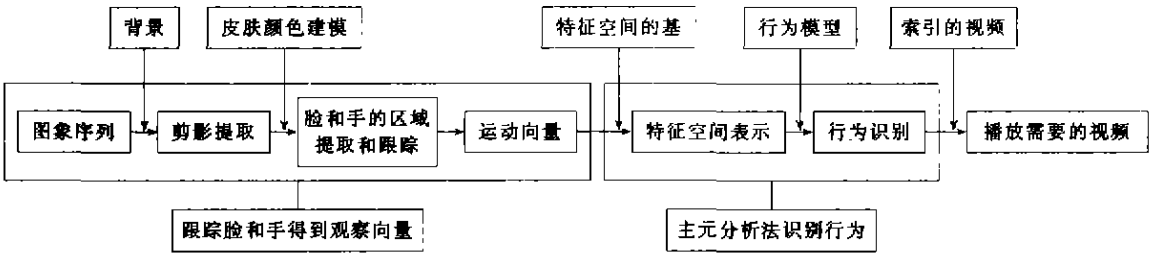



图1　系统框架

本文用太极拳的图象序列来验证该系统的性能，太极拳的动作完成，在很大程度上依赖于手的动作，MIT 也曾研究了太极辅助教练系统[7]. 他们是用两个自定标摄像机在 3 维空间重建和跟踪脸和手，然后应用 HMM 来建模和识别拳式，但要求表演者面向固定方向.

## 1　脸和手皮肤颜色建模和跟踪

### 1.1　HSI 空间中皮肤颜色的建模

在 Rits Eye 系统[8]中，脸的皮肤颜色是在 R、G、B 空间中用 3D 高斯模型来建模的. 我们则是用 H、S、I 颜色空间，对皮肤颜色的分布用一个 2D 高斯模型表示. H、S、I 空间的定义如图 2 表示. 在色度($H$)、亮度($I$)、饱和度($S$)3 个颜色分量中我们选用 $H$ 和 $S$ 是因为它们对光照变化不敏感. 按照距高斯分布中心的 Mahalanobis 距离，把皮肤区域和非皮肤区域分割开. 因为 $H$ 是圆心角，并归一到区间[0,1)，但由于 0 和 0.9 间的距离不是 0.9，而应是 0.1，因而这样在计算距离时会出现问题，考虑到大部分要考虑的象素分布在色度值 $h$ 附近(实验表明，$h$ 大约为 0.1)，为了避免上面的问题，在计算过程中，先对 $H$ 进行旋转，即把 $h$ 旋转到它的对称位置 $H'$ (见图 2)，即

$$H' = \mathrm{Mod}(H + 0.5 - h, 1) \qquad (1)$$

其中，Mod($x$,$y$)表示 $x/y$ 取余.

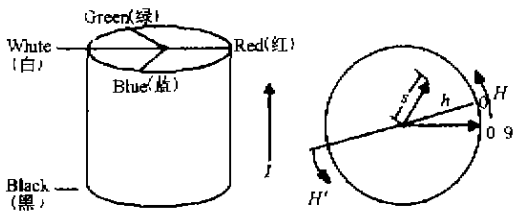



图 2　H、S、I 颜色空间

通用的 2D 高斯分布可表示为

$$p(x) = \frac{1}{2\pi}\|\Sigma\|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(x-\mu)^{T}\Sigma^{-1}(x-\mu)\right] \qquad (2)$$

其中，$\mu$，$\Sigma$ 为由手工标出的皮肤颜色的学习均值和

协方差矩阵

$$\mu = \frac{1}{n}\sum_{k=1}^{n} x_k,\ \Sigma = \frac{1}{n}\sum_{k=1}^{n}(x_k-\mu)(x_k-\mu)^{\mathrm{T}} \quad (3)$$

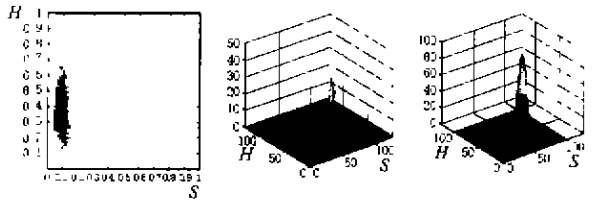
其中,$x=[H',S]^{\mathrm{T}}$.

如果一个象素的颜色矢量距中心点的 Mahalnobis 距离为

$$D=(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu) \quad (4)$$

若该值小于预定的阈值,则该象素属于皮肤,否则被分类成非皮肤.

实验中,我们是用人脸的图象来计算皮肤颜色分布的. 图 3 显示了在 $H$、$S$ 平面上皮肤颜色的分布、直方图和估计出的高斯分布.



(a) HS 平面上的皮肤颜色分布　(b) 直方图　(c) 估计出的高斯分布

图 3

### 1.2 剪影提取

为了避免背景区域杂乱点的影响,可用减去背景的方法提取人体剪影,然后用形态学算子将其滤波成封闭的单连通区域. 其中背景的构造过程是采用类似于文献[3]中的长期曝光法. 图 4 表示一个剪影提取的例子.

(a) 序列中一帧图象　(b) 提取剪影图象

图 4 剪影提取

### 1.3 脸和手的跟踪

在皮肤和非皮肤象素分类后,再将皮肤象素聚类成一个区域. 然后靠运动估计和区域限制来跟踪该区域.

设脸的位置在 $t$ 时刻是 $f(t)=[x(t),y(t)]^{\mathrm{T}}$,那么就可以在匀加速假设下,估计它在 $t+1$ 时刻的位置.

$$\begin{aligned} v(t) &= f(t)-f(t-1),\\ a(t) &= v(t)-v(t-1)\\ &= f(t)-2f(t-1)+f(t-2) \quad (5)\\ f(t+1) &= f(t)+v(t)+a(t)\\ &= 3f(t)-3f(t-1)+f(t-2) \end{aligned}$$

通过运动估计,可以确定后继帧中跟踪区域的初始位置,并预测是否有遮挡发生. 在预测位置邻域内,如果有多个候选区域存在,我们可用形状匹配来解决对应问题. 由此可见,匀加速运动和区域平滑的假设对我们面临的问题是一个合理的假设.

## 2 行为建模和识别

### 2.1 PCA 行为建模

选用 6 套杨氏太极拳作为识别的实例,这 6 套拳式分别是起手式、揽鹊尾、单鞭、提手上式、手挥琵琶和右下式,分别记作 O,G,S,Rh,P,Rb.

脸和手的位置参数被用来描述运动的状态,把脸的质心记作$(f_x,f_y)$,左右手的质心记作$(l_x,l_y)$和$(r_x,r_y)$. 因为绝对坐标依赖于人在图象中的位置,所以采用它们的相对差值作为观察参数.

$$\begin{aligned} fl_x &= f_x-l_x;\ fr_x = f_x-r_x\\ fl_y &= f_y-l_y;\ fr_y = f_y-r_y \end{aligned} \quad (6)$$

然后用 PCA 方法对实例建立行为模型. 主要步骤如下:

第 1 步　构造行为范例 $e_i^j$. $e_i^j$ 是一个 $nT$ 列向量,是行为 $j$ 的第 $i$ 个范例. $n$ 是参数个数,$T$ 是行为的持续时间,即序列数. 本文把所有范例归一为相同的长度.

第 2 步　构造行为矩阵 $A$. 把 $e_i^j$ 一列一列排起来构成 $A$. 设行为共有 $J$ 类,每个行为范例数为 $I$,那么,$A$ 的维数是 $nT\times IJ(IJ<nT)$.

第 3 步　$A$ 的奇异值分解(SVD)

$$A = U\Lambda V^{\mathrm{T}} \quad (7)$$

$U$ 是 $nT\times nT$ 正交阵,表示训练集中的主元方向;$\Lambda$ 是 $nT\times IJ$ 对角阵,有 $IJ$ 个对角元素 $\sigma_1,\sigma_2,\cdots,\sigma_k$ 按降序排列;$V^{\mathrm{T}}$ 是 $IJ\times IJ$ 正交阵,它编码了 $A$ 中主元方向的系数. 可以用最大的奇异值 $\sigma_1,\sigma_2,\cdots,\sigma_q$ 近似地表示 $e_i^j$.

第 4 步　用行为基近似地表示 $e$,即选择 $U$ 的前 $q$ 列作行为基

$$e \approx \sum_{i=1}^{q} c_i U_i \quad (8)$$

其中,$c_i=e\cdot U_i^{\mathrm{T}}$,因为 $U$ 是正交阵,故 $c_i$ 的意义是把

$e$ 投影到由 $q$ 个基向量张成的空间上. $U_i$ 称为行为基，而行向量 $\boldsymbol{c}^{\mathrm{T}}=[c_1,c_2,\cdots,c_q]$ 则用来表示行为 $e$.

### 2.2 行为识别

行为识别意味着把新观察的行为和范例相匹配，必须注意到由于成象条件和动作主体的变化，新观察的行为可能和范例不同. 此时可以用一类变换函数 $\Gamma$ 对其变化建模.

待识别行为记作 $\boldsymbol{D}(t):[1,T]\rightarrow R^n$，$[\boldsymbol{D}]$ 为把 $t$ 时刻的向量 $\boldsymbol{D}(t)$ 按顺序串连成的列向量. $[D]_j$ 表示向量 $[\boldsymbol{D}]$ 的第 $j(j=1,\cdots,nT)$ 个元素. 按照 2.1 节的方法，把 $[\boldsymbol{D}]$ 投影到行为基的特征空间上，恢复出系数向量 $\boldsymbol{c}$，就可以把行为近似成行为基的线性组合.

为了提高鲁棒性，可以用鲁棒性回归来最小化用特征空间表示的误差

$$\min_{c}\sum_{j=1}^{nT}\rho(e_j,\sigma) \tag{9}$$

$$e_j(c)=[D]_j-\sum_{i=1}^{q}c_iU_{ij} \tag{10}$$

然后用其来求解 $c$.

其中，Geman-Mclure 函数 $\rho_{GM}(x,\sigma)=x^2/(x^2+\sigma^2)$ 是 $x$ 的鲁棒性误差范数，$\sigma$ 是控制出格点影响的尺度参数.

在匹配问题上，本文引入了一个变换函数 $\Gamma(a,t)$，其中，$a$ 是参数，$\boldsymbol{D}(t)$ 在此变换下，表示为 $\boldsymbol{D}(t+\Gamma(a,T))$，然后，对变换 $\boldsymbol{D}(t+\Gamma(a,t))$ 进行泰勒展开，忽略高次项，则得到

$$\boldsymbol{D}(t+\Gamma(a,T))\approx\boldsymbol{D}(t)-\boldsymbol{D}_t(t)\Gamma(a,t) \tag{11}$$

其中，$\boldsymbol{D}_t(t)$ 是一阶导数. 为了在不同成象条件下，能够正确识别同一动作，就要求识别能在图象尺度、平移和旋转保持不变和动作主体有关的时间尺度和平移不变的情况下进行. 因此可把 $\boldsymbol{D}(t)$ 的时空尺度和时间平移变换进一步简化为 $S\cdot\boldsymbol{D}(at-L)$. 其中，$L$ 是时间平移，$S$ 是空间尺度因子，$a$ 是时间尺度因子、即快慢因子. 将其代入式(10)，则得到

$$e(c,a,L,S)=[S\cdot(\boldsymbol{D}_t(t)\Gamma(a,L,t)+\boldsymbol{D}(t))]_j-\sum_{i=1}^{q}c_iU_{ij} \tag{12}$$

其中，$\Gamma(a,L,t)=t+(a-1)t+L$.

还可以使用多尺度策略来使式(9)最小化，以逐渐减小尺度 $\sigma$，并交替对 $c$ 和 $a$ 使用梯度下降法，在给定系数 $c$ 后，解出变换参数 $a$；然后，由 $a$ 求解 $c$，直到收敛. 对于较大的变换，可以在大尺度 $\sigma$ 下收敛，以作为小尺度的初值.

若将式(12)中的各项同除以 $S$，则系数 $c_i$ 变成 $c_i/S$，其它形式保持不变，所以实际上可以在有比例因子的情况下恢复系数向量，即恢复 $\boldsymbol{c}'_i=c_i/S$.

恢复出系数向量 $\boldsymbol{c}'$ 后，再计算它与范例行为的系数向量 $\boldsymbol{m}$ 的欧氏距离 $d$，

$$d^2=\sum_{i=1}^{q}(\boldsymbol{c}'_i/\|\boldsymbol{c}'\|-\boldsymbol{m}_i/\|\boldsymbol{m}\|)^2 \tag{13}$$

由此可以容易看到：① $d$ 和 $S$ 是无关的，因为系数进行了归一，② $d$ 和图象平移和旋转无关，因为采用了相对坐标；③ $d$ 和动作快慢及平移无关，因为系数向量是在 $\Gamma(a,L,t)$ 变换下进行恢复. 总之，$d$ 和图象平面上的尺度、平移和旋转以及时间轴的尺度和平移无关.

但是 $d$ 和动作者相对摄像机的三维旋转是有关的. 在我们的问题里，只有围绕垂直轴的旋转(即改变朝向)需要考虑. 在没有深度信息的情况下，不可能计算出 $d$ 和面向角度的关系. 但可以采用增加不同角度的范例来弥补.

我们把观察到的行为归类为和它的距离最小的范例行为.

## 3 实验和讨论

### 3.1 皮肤颜色建模

在学习皮肤颜色高斯参数实验中，选择了 8 个表演人，利用手工提取他们的脸部图象块，其每个图象块的尺寸是 $20\times20$ 象素，且每个个体取 50 幅图象，所以实验总象素数为 $20\cdot20\cdot50\cdot8=160\,000$. 然后在两组光照条件下进行对比实验. 其中第 2 组图象较暗，因为把照明灯光关闭了一半.

第 1 组的计算均值为

$$\mu_1=[0.111\,17\quad 0.238\,95]^{\mathrm{T}}$$

协方差矩阵 $\Sigma_1=\begin{bmatrix}0.000\,70 & -0.000\,08\\ -0.000\,08 & 0.008\,02\end{bmatrix}$

第 2 组的计算均值为

$$\mu_2=[0.113\,12\quad 0.241\,91]^{\mathrm{T}}$$

协方差矩阵 $\Sigma_2=\begin{bmatrix}0.000\,60 & -0.000\,06\\ -0.000\,06 & 0.007\,68\end{bmatrix}$

对比两组实验结果，可以发现用此方法进行皮肤颜色识别受光照影响较小. 因此使用该学习的均值和协方差，可以对象素点进行分类和聚类. 图 5 是提取脸和手的例子.

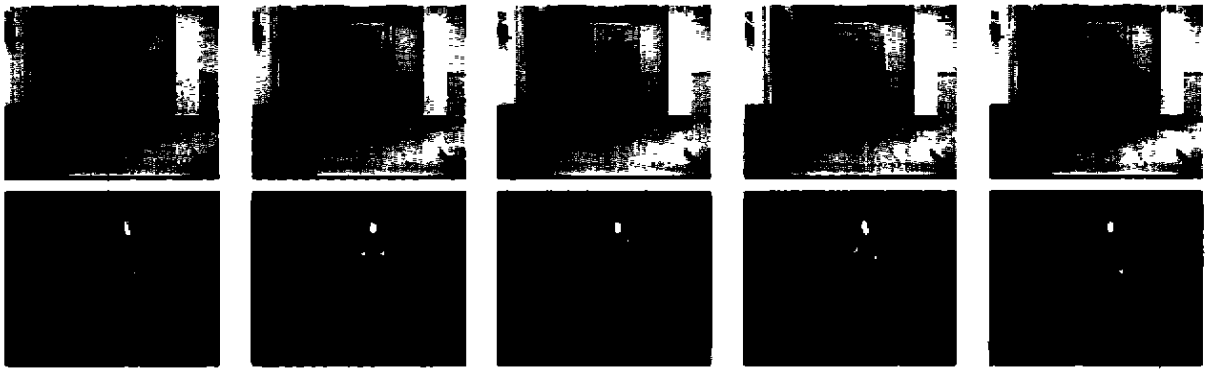


图 5 跟踪脸和手的区域(本图为起手式)

### 3.2 太极拳式识别

在行为建模和识别实验中,要求动作者面向角由 −30°到 30°,每 10°做一个范例. 实验不关心其它方向的旋转,因为摄像机是固定的,而动作者是站在地面上. 由于每个动作有 7 个角度上的范例,而且由 2 人表演,所以共有范例 $7\times2\times6=84$. 由于每个序列先归一为 100 帧图象,因此矩阵 $A$ 的维数是 $400\times84$.

图 6 是在主体面向 0°时的 6 套太极拳的观察向量元素的曲线图. 图 7 是主体在两个不同面向的角度下,做起手式和揽鹊尾的观察曲线图.

为了确定需要多少个基可以表示足够的信息,做了行为基的信息表示能力曲线(如图 8). 图上纵轴是信息百分比,横轴是基的个数. 用 6 个基的信息比例是 79.48%, 而 7 个基时为 80.23%,其间只有 0.75%的提高. 且 80%的信息量已经能比较好的描述原信号了,因此选择 $q=6$.

6 个太极拳式范例间的距离见表 1,而学习者在做起手式(0°,10°,−30°面向下)时,和范例间的距离矩阵见表 2. 其中用了 6 个基. 识别结果由表 3 的误识矩阵表示. 两组对比实验分别由专家和学习者进行. 从结果看,对专家的识别识别率知识略高于学习者.

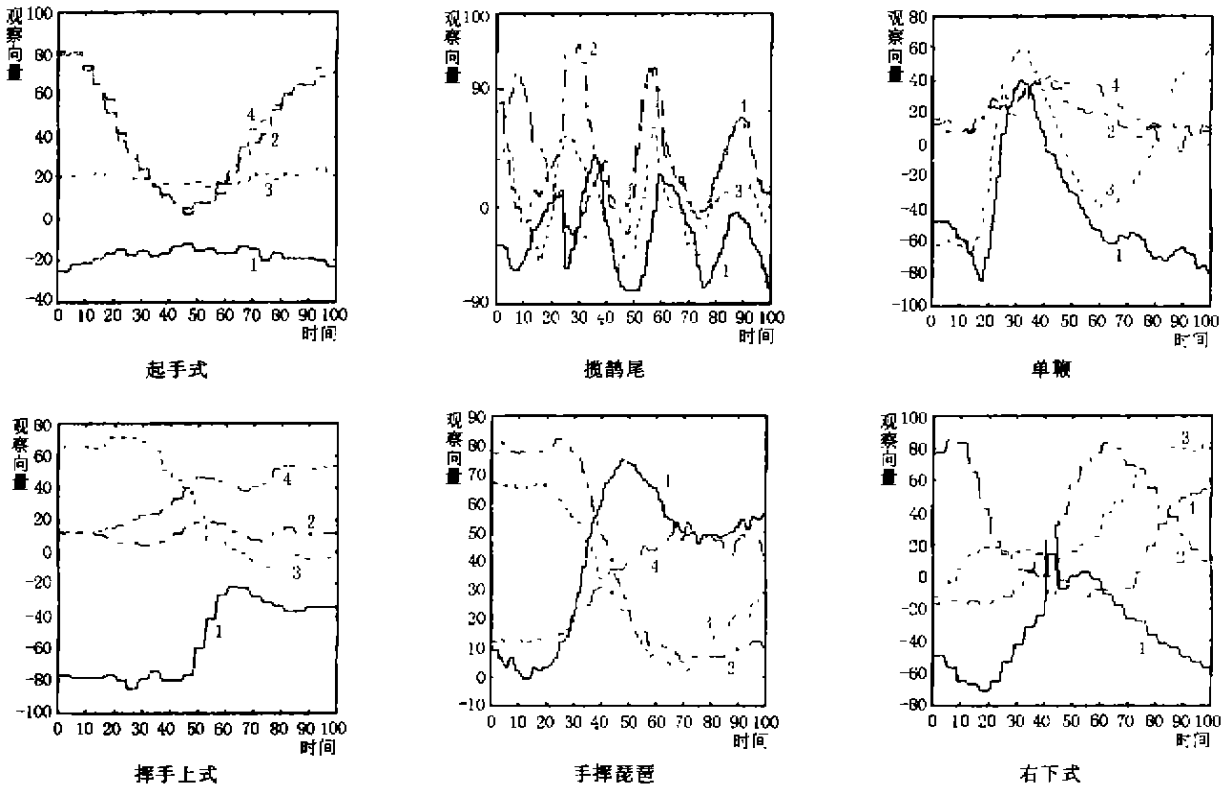



图 6 面向 0°表演的 6 式太极拳观察向量元素曲线图

(图中,1:$f_1$;2:$f_2$;3:$f_3$;4:$f_4$)

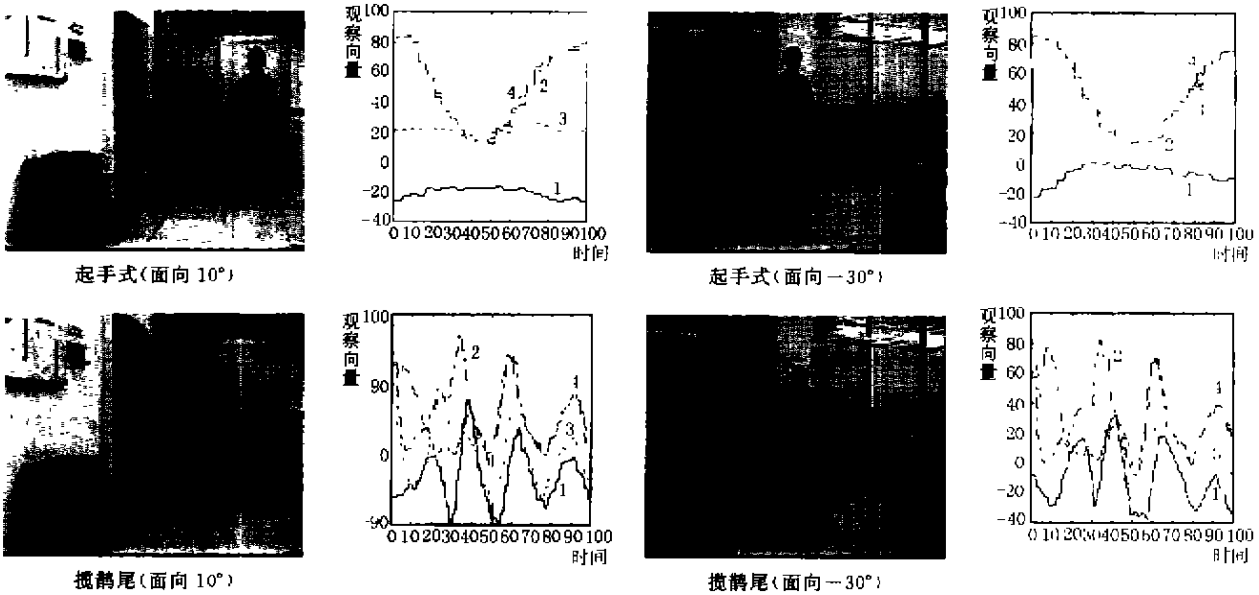



图 7　两个不同面向的起手式和揽鹊尾(角度指开始时的面向角度)观察曲线图

(图中，1：$fl_x$；2：$fl_y$；3：$fr_x$；4：$fr_y$)

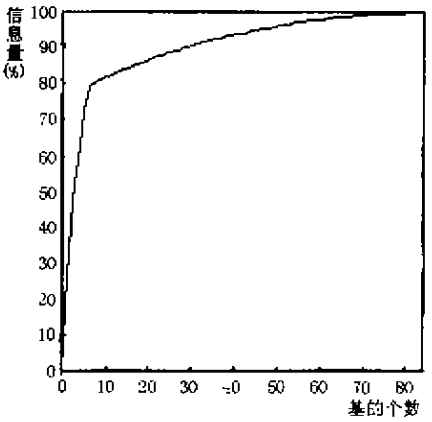



图 8　行为基的信息表示能力曲线

表 1　6 个范例间的距离矩阵

| 拳式 | O | G | S | Rh | P | Rb |
|---|---|---|---|---|---|---|
| O | 0 | 1.63 | 2.55 | 2.17 | 2.37 | 2.01 |
| G | | 0 | 2.45 | 1.86 | 2.66 | 2.00 |
| S | | | 0 | 2.04 | 2.34 | 2.02 |
| Rh | | | | 0 | 2.62 | 1.86 |
| P | | | | | 0 | 2.42 |
| Rb | | | | | | 0 |

表 2　学习者做起手式时和范例间的距离矩阵

| 起手式 | O | G | S | Rh | P | Rb |
|---|---|---|---|---|---|---|
| 0° | 0.18 | 1.71 | 2.47 | 2.10 | 2.20 | 1.93 |
| 10° | 0.22 | 1.74 | 2.50 | 2.10 | 2.19 | 1.92 |
| −30° | 0.58 | 1.95 | 2.45 | 2.20 | 1.84 | 1.95 |

表 3　误识矩阵

| 拳式 | 专家 | | | | | | 学习者 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | O | G | S | Rh | P | Rb | O | G | S | Rh | P | Rb |
| O | 7 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 1 | 0 | 0 |
| G | 0 | 4 | 0 | 1 | 0 | 2 | 0 | 3 | 1 | 2 | 1 | 2 |
| S | 0 | 1 | 4 | 1 | 1 | 0 | 0 | 2 | 3 | 1 | 0 | 1 |
| Rh | 1 | 0 | 0 | 6 | 0 | 0 | 1 | 0 | 0 | 5 | 0 | 1 |
| P | 0 | 0 | 1 | 0 | 6 | 0 | 0 | 0 | 1 | 0 | 6 | 0 |
| Rb | 0 | 1 | 0 | 1 | 0 | 5 | 1 | 0 | 0 | 1 | 0 | 5 |

动作特征的提取是实时进行的，一套动作的识别时间约为 1s.

## 4　结　论

本文提出并用实验模拟了一个行为识别框架，此框架不仅仅用于太极拳学习，还容易扩展到其它如芭蕾，体操等身体技能的培训场合. 若进一步提高跟踪算法的鲁棒性，则可以扩展为自动建立基于人行为语义信息的视频检索方法和视觉监控.

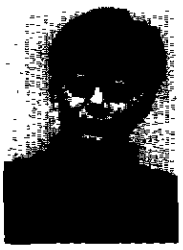
目前本系统的泛化存在一些问题. 一方面的问题是对脸和手的跟踪受光照突变和遮挡的影响，很难彻底解决. 另一个问题是旋转问题需要准确的解决. 一个可能的途径是利用人体测量特征的知识先求解人的姿态或者在有些应用背景下，用多摄像机立体视觉的方法解出三维坐标.

### 参 考 文 献

1　Aggarwal J K, Cai Q. Human motion analysis: a review. Computer Vision and Image Understanding, 1999, 73(3): 428～440

2 Polana R, Nelson R. Low level recognition of human motion (or how to get your man without finding his body parts). In: Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated, Austin 1994, 83～88.
3 Davis J W. Appearance-based motion recognition of human action. Tr387, 1997, MIT Media Lab.
4 Starner T, Pentland A. Visual recognition of American signal language using hidden markov models. In: International Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, 1995: 189～194.
5 Yacoob Y, Black M J. Parameterized modeling and recognition of activities. Computer Vision and Image Understanding, 1999, 73(2): 232～247.
6 Ivanov Y, Bobick A. Probabilistic parsing in action recognition. Tr430, 1997, MIT Media Lab.
7 Becker D A. Sensei: a real-time recognition, feedback and training system for T'ai Chi gesture. Tr426, 1997, MIT Media Lab.
8 Xu Gang, Sugimoto Takeo. Rits eye: a software-based system for realtime face detection and tracking using pan-tilt-zoom controllable camera. In: Proc. of 14th International Conference on Pattern Recognition, Brisbane, Australia, 1998, 1194～1197.

**胡长勃** 1970年生，1998年进入中科院自动化所模式识别实验室，攻读计算机视觉和模式识别方向的博士学位。目前的研究兴趣是图象处理、运动分析、视觉跟踪、人的行为识别和视觉监控。



**冯 涛** 1973年生，1997年进入中科院自动化所模式识别实验室，攻读计算机视觉和模式识别方向的博士学位。目前的研究兴趣是立体视觉、视觉跟踪、运动分割、图象理解和基于图象的图形绘制。

**马颂德** 1946年生，IEEE高级会员，1968年毕业于清华大学，1986年获法国国家博士学位，现为中科院自动化所所长，博士生导师，"973"国家项目"图象、语音、自然语言理解与知识发掘"首席科学家。长期从事计算机视觉、模式识别和图象处理等方面的工作。

**卢汉清** 1960年生，1982年毕业于哈尔滨工业大学，1988年在华中理工大学获博士学位，现为中科院自动化所研究员。目前的研究兴趣为图象处理和应用，多媒体信息系统以及图象和视频数据库。