

基于多实例运动学特征学习的动态手势识别研究

周彩秋¹, 杨余旺¹, 庞海波²

(1. 南京理工大学计算机科学与工程学院, 江苏 南京 210094;

2. 郑州大学软件与应用科技学院, 河南 郑州 450002)

摘 要: 在动态手势特征提取和识别方面, 利用运动学模式解决动态手势识别问题, 在光流场基础上计算出散度模式, 旋度模式, 对称模式, 反对称模式, 梯度张量第二、第三主不变模式, 应变张量第二、第三主不变模式以及自旋转张量第三主不变模式; 进一步提出一种基于多实例学习的方法, 将每一个动态手势的所有运动主模式构成一个动态手势词袋, 将未知类型动态手势的运动主模式与词袋空间中对应运动主模式进行相似度计算, 利用最近邻方法对手势进行识别。实验结果表明: 基于多实例运动学主模式学习的动态手势识别方法取得了较高的识别率。

关键词: 手势识别; 运动学特征; 时空轴降维; 多实例学习

中图分类号: TP391

文献标识码: A

Research of dynamic gesture recognition based on multi-instance learning of kinematics features

ZHOU Cai-qiu¹, YANG Yu-wang¹, PANG Hai-bo²

(1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2. School of Software and Applied Science and Technology, Zhengzhou University, Zhengzhou 450002, China)

Abstract: Compared to static gestures, dynamic gestures had some new characteristics. The problems of dynamic gestures recognition was spewed by using kinematics mode, such as divergence modes, curl modes, symmetric and ant-symmetric modes, the second and third principal invariant modes of the gradient tensor, the second and third principal invariant modes of the strain tensor and the third principal invariant modes of spin tensor; Further, a framework based on multi-instance learning was proposed, organize all these principle modes for each gesture were organized to a dynamic gestures bag-of-words, and the similarity between the mode of unknown type dynamic gestures and the all bag-of-words were calculated. Then, the nearest neighbor method was used to recognize the dynamic gestures. The experimental results show that the dynamic gestures recognition based on multi-instance kinematics features principal mode learning methods can obtain a higher recognition rate.

Key words: gesture recognition, kinematics features, temporal axis dimension reduction, multi-instance learning

1 引言

动态手势是运动的手势, 动态手势不同于静态手势, 在图像处理中它由一组序列图像组成,

并且与模型参数空间中的一条轨迹相对应。不同用户做的同一类型动态手势在空间和时间尺度上是不同的, 甚至同一个用户前后 2 次做的同一类型动态手势在空间和时间尺度上也有差异。因

收稿日期: 2017-06-10; 修回日期: 2017-09-21

通信作者: 杨余旺, yuwangyang@njust.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61640020, No.61402420); 河南省高等学校重点科研项目资助计划基金资助项目 (No.17A520014); 河南省科技攻关计划项目基金资助项目 (No.172102310496)

Foundation Items: The National Natural Science Foundation of China (No.61640020, No.61402420), Higher Education Institutions Program of Henan Province (No.17A520014), Scientific and Technological Research Project in Henan Province (No.172102310496)

此，如何表示同一类型手势，如何消除同一类型手势存在的时空差异是动态手势识别需要解决的关键问题^[1]。

动态手势的表示方法可以分为基于表观的表示方法和基于模型的动态手势识别表示方法。

基于表观的方法采用图像序列上的颜色信息或者灰度信息对人的运动进行分析^[2]。但是由于不同人种肤色有很大差别，因此该方法在应用过程中容易受到肤色的限制^[3]。其次，由于背景通常相差也比较大，因此在对颜色特征提取过程中，也需要进行动态的建模^[4]。

基于模型的动态手势识别通常包括手势建模、手势分析和手势识别。动态手势建模方法会直接影响手势分析和手势识别的复杂度和精确度。基于模型的动态手势建模通常需要计算图像运动参数，常见的手势模型参数包括目标轨迹、光流以及速度等。Molina 等^[5]先对手势进行分割，然后结合手势形态学轮廓和手势运动模型来识别动态手势，但是效果较差。Fan^[6]引入一种模板动态更新机制对手势进行跟踪和识别，实时性较差。Liu 等^[7]采用目标特征点的相关性生成关系，从而对手势进行跟踪和识别，实时性较差。

Wang 等^[8]提出基于动态贝叶斯网络的分层视觉模型对动态手势识别，通过利用快速差分颜色跟踪算法，对运动轨迹进行跟踪和识别，识别效果较好，但是过程比较复杂。Sgouropoulos 等^[9]提出一种基于自组织神经网络和隐马尔可夫模型的手势识别算法，该算法能够较好地适应复杂光照，但是计算过于复杂。Durbin 等^[10]利用光流场实现手势识别，但该方法容易受到复杂光照的影响。

目前基于模型的动态手势识别方法在进行手势建模、跟踪、分析和识别时容易受复杂光照影响，很难实现识别的精度和实时性的平衡。对于动态手势识别，如果能够提取动态手势高层语义，或能够对动态手势进行抽象层次描述，然后利用这些高层次语义或高层次描述区分不同动态手势，将有助于提高识别的精度和实时性。本文提出一种基于多实例运动学特征学习的动态手势识别方法，提取一些能够描述运动的特性，以期提高动态手势识别精度。

光流本身是一种描述运动的矢量场，运动学中的散度场、旋度场、对称场、反对称场、梯度

张量场、应变张量场和自旋张量场均能反映矢量场的不同运动特性^[11]。因此，本文在光流场基础上计算上述几种矢量场及其特征，用于描述动态手势并完成动态手势识别。

2 动态手势的运动学模式场计算

2.1 动态手势散度场的计算

图 1 和图 2 为动态手势“画圆”的第 3 帧以及利用块相关算法计算出的光流场三维图。



图 1 手势“画圆”连续帧光流图像

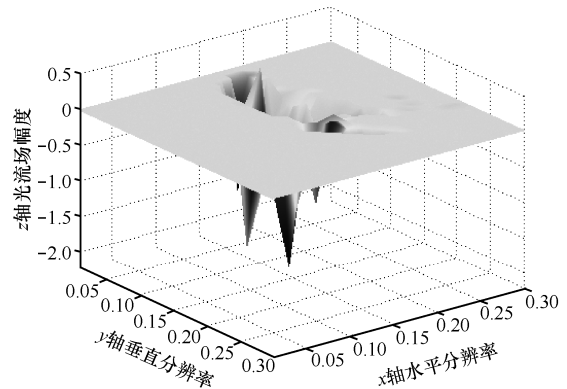


图 2 光流场三维图

设某矢量场为

$$A(x, y, z) = P(x, y, z)i + Q(x, y, z)j + R(x, y, z)k \quad (1)$$

其中， P, Q, R 是 A 在 3 个坐标轴上的投影，分别表示在 x, y, z 方向的分速度，并且具有一阶连续偏导数。 Σ 是场内一有向曲面， n 是 Σ 在点 (x, y, z) 处的单位法向量，则 $\iint A \cdot ndS$ 被称为矢量场 A 通过曲面 Σ 向着指定侧的通量。考虑到动态手势光流场的二维特性和时间特性，矢量场 A 的散度，即散度场可表示为

$$f_1(x, y, t_i) = \frac{\partial u(x, y, t_i)}{\partial x} + \frac{\partial v(x, y, t_i)}{\partial y} \quad (2)$$

其中， $\frac{\partial u(x, y, t_i)}{\partial x}$ 和 $\frac{\partial v(x, y, t_i)}{\partial y}$ 分别代表了点 (x, y, t_i) 在 t_i 时刻光流在 x 和 y 方向的分量 u 和 v 的偏导数。

动态手势“画圆”动作包含手和手臂的大范围动作，相应的散度特征不但能够体现人手的运动方向和运动力度，还能体现手臂的运动方向和运动力度。

2.2 动态手势旋度场的计算

设某向量场由式(1)给出，散度场由式(2)给出，在式(2)基础上计算旋度场，旋度场的计算式为

$$f_2(x, y, t_i) = \frac{\partial v(x, y, t_i)}{\partial x} - \frac{\partial u(x, y, t_i)}{\partial y} \quad (3)$$

其中， $\frac{\partial v(x, y, t_i)}{\partial x}$ 和 $\frac{\partial u(x, y, t_i)}{\partial y}$ 分别代表了点 (x, y, t_i) 在 t_i 时刻光流在 x 和 y 方向的分量 v 和 u 的偏导数。

旋度场能更好地描述局部小区域画圆运动或者画圆动作产生的流体域。“画圆”动作实验者的手和手臂完成画圆动作的过程中产生的光流具有较强的旋度特性。旋度特征可以用来区分具有旋转特性和非旋转特性的手势。

2.3 动态手势对称场和反对称场的计算

对光流场的 u 和 v 这 2 个分量分别计算对称场和反对称场，对称场的计算式为

$$f_3(t_i) = u(t_i) + u(t_i)^T \quad (4)$$

$$f_4(t_i) = v(t_i) + v(t_i)^T \quad (5)$$

其中，符号 T 表示转置操作， $u(t_i)$ 和 $v(t_i)$ 和分别代表了时刻 t_i 光流场的 u 分量和 v 分量。反对称场的计算式为

$$f_5(t_i) = u(t_i) - u(t_i)^T \quad (6)$$

$$f_6(t_i) = v(t_i) - v(t_i)^T \quad (7)$$

图 3 和图 4 给出了图 2 对应动作“画圆”在 v 分量的对称场和反对称场。当动作图像为矩形时，沿对角线轴两侧有较高的值。否则，将在对角线轴两边出现一边值大，一边值小的情况。

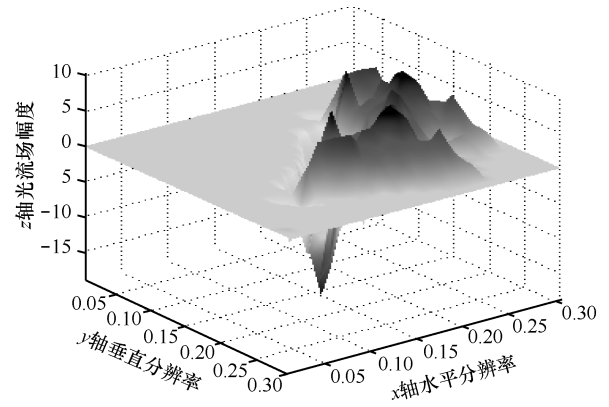


图 3 手势“画圆”v 对称场

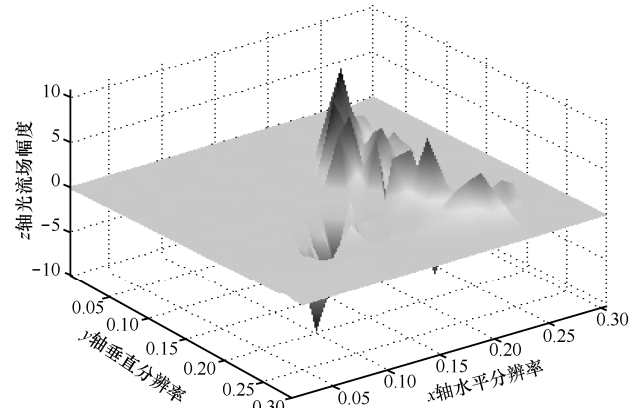


图 4 手势“画圆”v 反对称场

2.4 动态手势梯度张量场的计算

梯度张量特征是基于光流梯度产生的，在描述局部结构运动上，它比光流场具有更好的性能。光流梯度张量场的计算式为

$$\nabla U(X, t_i) = \begin{pmatrix} \frac{\partial u(X, t_i)}{\partial x} & \frac{\partial u(X, t_i)}{\partial y} \\ \frac{\partial v(X, t_i)}{\partial x} & \frac{\partial v(X, t_i)}{\partial y} \end{pmatrix} \quad (8)$$

从式(8)中可以看出，在光流场的每一个位置点 X ，梯度张量是一个 2×2 的矩阵。

梯度张量的第一主不变量计算式为

$$L(X, t_i) = -\text{trace}(\nabla U(X, t_i)) \quad (9)$$

第二主不变量计算式为

$$M(X, t_i) = \frac{1}{2}(L^2 - \text{trace}(\nabla U(X, t_i))^2) \quad (10)$$

第三主不变量计算式为

$$N(X, t_i) = \det(\nabla U(X, t_i)) \quad (11)$$

第一主不变量的值等于散度场。第二、三主不变量分别定义为 $f_7 = M(X, t_i)$ 和 $f_8 = N(X, t_i)$, 将这 2 个主不变量作为描述动态手势的特征。

2.5 动态手势应变张量场和自旋张量场的计算

应变张量场和自旋张量场代表了由刚体运动产生的偏差。其中应变张量场和自旋张量场可以通过式(12)和式(13)获得。

$$S(X, t_i) = \frac{1}{2}(\nabla U(X, t_i)) + (\nabla U(X, t_i))' \quad (12)$$

$$O(X, t_i) = \frac{1}{2}(\nabla U(X, t_i)) - (\nabla U(X, t_i))' \quad (13)$$

其中, S 代表应变张量场, O 代表自旋张量场。

由于 S 代表了应变张量场, 可以对 S 求 3 个主不变量 $L_s(X, t_i)$ 、 $M_s(X, t_i)$ 、 $N_s(X, t_i)$ 。 O 代表了自旋张量场, 也可以对 O 求 3 个主不变量 $L_o(X, t_i)$ 、 $M_o(X, t_i)$ 、 $N_o(X, t_i)$ 。

由于 S 的第一主不变量等于 ∇U 的散度的特征向量, 所以选择 S 的第二主不变量 $f_9 = M_s(X, t_i)$ 和第三主不变量 $f_{10} = N_o(X, t_i)$ 作为动态手势特征。

O 的第二主不变量等于旋度量, 所以选择 O 的第三主不变量 $f_{11} = N_o(X, t_i)$ 作为动态手势特征。图 5 和图 6 给出了“画圆”对应 f_{10} 和 f_{11} 的三维图。

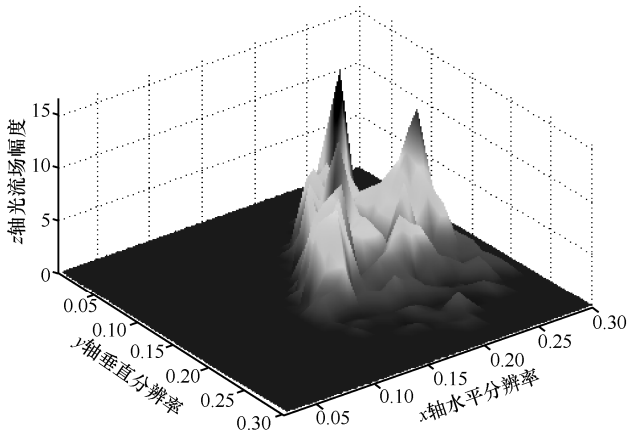


图 5 “画圆”对应 S 的第三主不变量三维图

对于不同的动作, 它们的矢量场值和表现是不同的, 而且也不同于前面所述的其他特征值和特性。因此, 可以用运动矢量场值描述不同运动的动态手势。

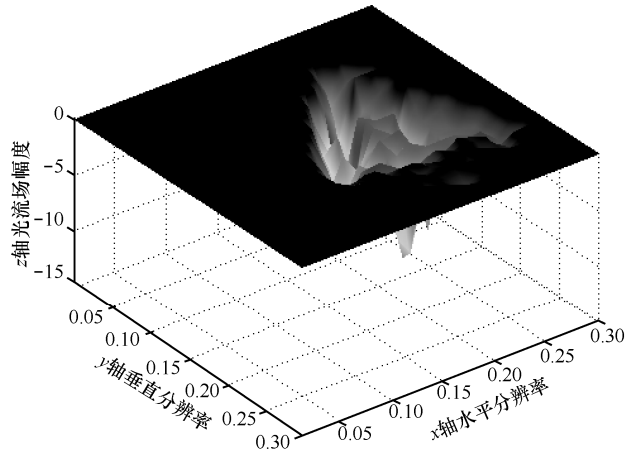


图 6 “画圆”对应 O 的第三主不变量三维图

3 基于多实例运动学模式学习模型的构建

3.1 时空轴 PCA 降维

考虑到时间轴上的相关性等信息, 采用下述步骤在时空轴上对运动学矢量场进行 PCA 降维, 获取运动学特征。

1) $U(X, t_i), i=1, \dots, M$ 代表视频第 t_i 帧矢量场。利用式(14)求出该矢量序列平均值。

$$\bar{U}(X) = \langle U(X, t_i) \rangle = \frac{1}{M} \sum_{i=1}^M U(X, t_i) \quad (14)$$

其中, $\bar{U}(X)$ 通常假定为 0, 符号 $\langle \cdot \rangle$ 为平均操作符。

2) 利用 PCA 方法求出时间独立正交基 $\phi_k(X)$ 和时间相关扩散系数 $\alpha_k(t_i)$, 以便重建 $U(X, t_i)$ 。保证当给定任何的 $m \leq M$ 时, $U(X, t_i)$ 在平均最小二乘截断误差意义上最优。

当空间分辨率 N 远高于观察样本数目 M 时, 特征向量可以表示成一组数据向量的线性组合, 如式(15)所示。

$$\phi_k = \sum_{i=1}^M v_i^k U(X, t_i), k=1, 2, \dots, M \quad (15)$$

3) 系数 v_i^k 为

$$Cv = \lambda v \quad (16)$$

其中, $v = (v_1^k, v_2^k, \dots, v_M^k)$ 是第 K 个特征向量, 矩阵 C 代表 $M \times M$ 对称矩阵, 如式(17)所示。

$$C_{ij} = \frac{1}{M} (U(X, t_i) \cdot U(X, t_j)) \quad (17)$$

其中, 符号“ \cdot ”为标准向量内积。

4) 求解 $M \times M$ 维矩阵 C 的特征向量得到

$N \times M$ 矩阵 \mathbf{R} 的特征向量, 当 $N \gg M$ 时, 计算效率较高。

5) 利用式(15)和特征向量 \mathbf{v} , 计算出光流的主时域模式, 利用式(18)重建矢量场信息。

$$\alpha_k(t_i) = (U(X, t_i) \cdot \phi_k(X)), i = 1, \dots, M \quad (18)$$

3.2 多实例运动学模式学习模型的构建

计算光流场运动学特征正交基获得光流场运动学模式。用式(19)求特征向量代替求矩阵 \mathbf{C} 的核。

$$\mathbf{C}_k(t_i, t_j) = \frac{1}{m} (f_k(t_i) \cdot f_k(t_j)) \quad (19)$$

其中, k 表示计算的运动学特征的序号, $f_k(t_i)$ 代表第 k 个运动学特征在 t_i 时刻的矢量场, 符号 $f_k(t_i)$ 是 $f_k(X, t_i)$ 向量化后的描述式, 可以看出, 向量化后的符号消除了空间依赖关系。

利用上述矩阵 $\mathbf{C}_k(t_i, t_j)$, 可以导出时间独立特征向量或运动学模式以及时间依赖扩散系数。

基于多实例学习方法构建动态手势运动学主模式模型, 其中, 每一个实例可以通过时空轴 PCA 的方法获得运动学主模式。下面利用数学定义来更加明确地阐明上述过程。

给出一个正的样本和一个负的样本。对每个动态手势求出 11 种运动学矢量场, 每种运动学矢量场可以通过 PCA 方法求出该运动学矢量场的多个模式。

利用 PCA 方法计算每一种运动学特征的 2 个主要模式, 对于每一个样本, 共计可以得到 11 种运动学特征的 22 个主模式, 即 $n_i^+ = 22, i = 1$ 。

其中, $1^+ = 1$ 表示只有一个正样本, 而 $s_{11}^{+f_1}$ 表示该样本散度特征的第 1 个主要模式, $s_{12}^{+f_1}$ 表示散度特征的第 2 个主要模式, $s_{11}^{+f_2}$ 表示旋度特征的第 1 个主要模式, $s_{12}^{+f_2}$ 表示旋度特征的第 2 个主要模式, 以此类推。

将上述 2 个样本的所有运动学模式串联之后就得到了向量 \mathbf{C} , 根据 $e = \left\{ 1, \dots, \left(\sum_{i=1}^L n_i^+ + \sum_{i=1}^L n_i^- \right) \right\}$ 可以计算出向量 \mathbf{C} 中的运动学模式个数 ($e = 22 \times 2 = 44$)。由于每一个成员 $S_e^{f_k}$ 是一个运动学主模式, 那么一

个运动学主模式属于词袋 B 的条件概率是 $P(S_e^{f_k} | B)$, 式(20)定义了所有模式属于词袋 B 的条件概率值。

$$[P(S_1^{f_1} | B), P(S_2^{f_1} | B), \dots, P(S_e^{f_1} | B)] \quad (20)$$

在式(20)中, 符号 B 没有显示上标 + 和下标 -, 这表明还不知道该词袋的标签。式(20)定义的向量被称作基于多实例运动学模式学习的模型库或者特征空间 f_c , 其中, 每一个词袋代表一个模型或者空间中的一个点。

第 e 个模式 $S_e^{f_k}$ 的条件概率 $P(S_e^{f_k} | B)$ 独立于该词袋的标签, 可以被写成为

$$P(S_e^{f_k} | B) \propto d(s_e^{f_k} | B) = \max_j \exp \left(\frac{-\|s_{ij}^{f_j} - s_e^{f_k}\|}{\sigma^2} \right) \quad (21)$$

$$\forall f^i = f^k$$

令 B_i 的第 z 个运动模式等于集合 C 中的第 z 个模式, 因为在 C 中的第 z 个模式来自于 B_i 的第 z 个模式。它们的相似度可以用函数 d 近似表示, 即一个模式属于某一个样本模式的条件概率近似, 可用式(21)中的 d 表示, 对该式求解即可计算出每一个条件概率值。

如果要计算 B_i 的散度场的第一个模式 $s_{11e}^{+f_1}$ ($s_{11e}^{+f_1}$ 表示散度特征的第一个模式) 和矩阵 \mathbf{C} 中的运动学特征的运动模式的相似度, 那么需要计算它和 \mathbf{C} 中 $s_{11}^{f_1}$ 和 $s_{23}^{f_1}$ 之间的距离, 即计算它与所有运动散度特征的第一主模式之间的距离。

这样获得词袋 B_i 中所有特征模式与矩阵 \mathbf{C} 中的所有同类特征模式之间的距离关系如下。

$$\mathbf{m}(B_i) = [p(s_{11}^{f_1}, B_i), d(s_{21}^{f_1}, B_i), d(s_{31}^{f_1}, B_i), \dots, d(s_{c1}^{f_1}, B_i)] \quad (22)$$

每个动态手势被映射到一个 e 维向量 $\mathbf{m}(B_i)$ 。然后, 对 e 维向量 $\mathbf{m}(B_i)$ 利用最近邻分类器判断所属的手势类别。

4 实验验证和分析

4.1 动态手势库建立

本文研究的动态手势集选择了参考文献[2]中的 12 种动态手势, 动态手势如图 7 所示。

第一行 6 个动态手势图形代表了 6 种不同画图操作, 从左向右分别是画圆、画三角形、画矩形、

画弧线、画水平直线和画垂直线。

第二行动态手势图形代表了 6 种不同编辑操作，从左向右分别是移动、复制、撤销、交换、删除和关闭。

动态手势样本通过在线视频捕捉的方式获得。选择 5 个实验者，让每个实验者依次将 12 种动态手势分别以不同时间尺度和空间尺度执行 5 遍，摄像头记录全部 5 个实验者手势运动视频，视频尺寸为 320×240。将实验者视频按照手势动作进行分割，每个实验者的每个动态手势作为一个视频，共记录 300 个视频，并按照手势类型分为 12 类，每类包含 25 个相同类型动态手势的视频。

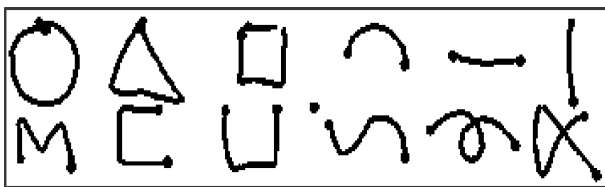


图 7 本文定义动态手势集

4.2 实验设计

实验时将视频帧调整到 240×240，分辨率缩小到 120×120 和 60×60 进行验证。在参数设置上，当分辨率缩小到 120×120 和 60×60 时，相应的块大小为 6×6 和 4×4；分别选择 1~6 个运动学特征的模式进行实验。采用“交叉验证”法随机选择 225 个训练视频和 75 个测试视频进行实验，利用最近邻分类器，实现对动态手势的分类和识别。作为比较，也选择动态手势的光流场主模式、散度场主模式、旋度场主模式、对称场主模式、反对称场主模式、梯度张量场主模式、应变张量场主模式和自旋张量场主模式分别进行实验验证。本文将与文献[12]中提到的 DTW、IDTW 及相关信息法进行比较。

4.3 实验结果和分析

图 8 给出了当运动矢量场主模式取值为 1~6，视频分辨率调整到 240×240 时获得的平均识别率。

当对每一个动态手势视频的每一种运动特征选择一个主模式时，能够获得平均识别率为 0.803 2，随着运动特征主模式的增加，平均识别率有所增加，当对每一个运动特征选择 4 个主模式时能够获得的平均识别率为 0.937 6，当继续增加运动特征主模式个数时，对平均识别率的提高影响很小。

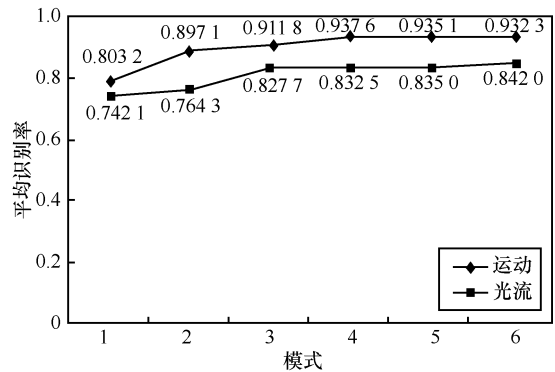


图 8 光流和运动学模式平均识别率

图 9 给出当分辨率为 240×240，对应运动散度特征主模式和运动旋度特征主模式取值为 1~6 时，获得的平均识别率。对比图 8 和图 9，当光流模式、散度模式、旋度模式分别取第一主模式时，散度模式和旋度模式获得的平均识别率明显高于光流模式。单独散度矢量场的散度模式、旋度矢量场的旋度模式所得的平均识别率不如全部矢量场的运动学模式，即综合运动学模式具有最高的识别率。

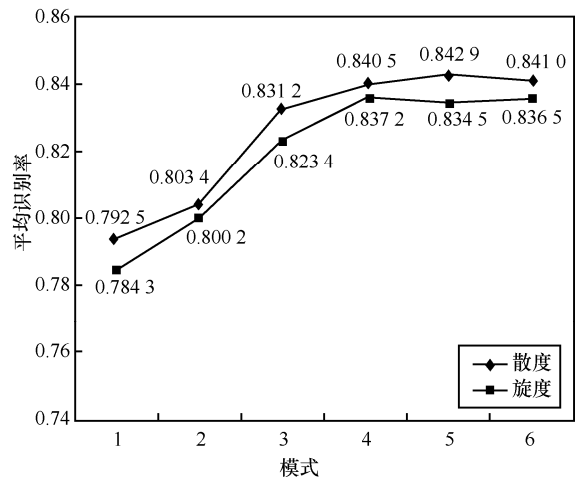


图 9 散度和旋度主模式平均识别率

图 10 给出了这些运动矢量场单独取 4 个主模式时而取得的平均识别率，同时给出了全部主模式组合时取得的平均识别率。可以看出，当图像分辨率为 120×120 时，每种运动矢量场取 4 个主模式，获得平均识别率为 0.897 1，同样取 4 个主模式，分辨率为 60×60 时，获得的平均识别率为 0.853 2。

比较图 8 和图 10 中运动矢量场的平均识别率，图像分辨率降低导致平均识别率明显下降，这是因为当分辨率降低时，一些关键信息丢失，导致计算出的运动特征模式描述能力明显减弱。

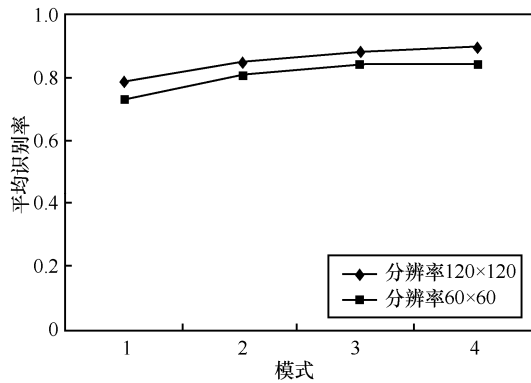


图 10 不同分辨率运动学主模式平均识别率

图 11 显示了分辨率为 240×240 时的识别情况，水平轴和垂直轴分别代表运动矢量场类型和平均识别率，从左到右分别是散度场，旋度场， u 分量对称场， v 分量对称场， u 分量反对称场， v 分量反对称场，梯度张量场第二、第三，应变张量场第二、第三和自旋张量场第三主模式。

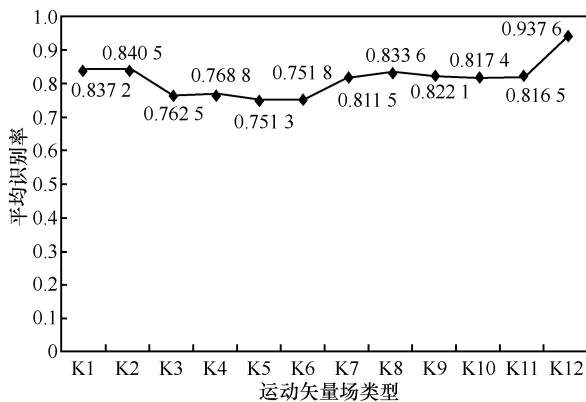


图 11 不同运动主模式的平均识别率

由图 11 可知，K3、K4、K5 和 K6 特征的主模式取得的平均识别率比较低，因为它们分别利用光流场的水平和垂直分量，并且只用到了运动矢量场的部分信息，因此平均识别率相对较低。另外，单独使用一个特征所取得的平均识别率远低于所有特征组合取得的平均识别率，当利用所有运动特征 4 个主模式进行识别时取得平均识别率为 0.937 6。由于每一类运动特征主模式所描述的动态手势运动特性不同，所以它们提供的运动信息不同。

实验表明，利用运动模式构造的动态手势词袋对本文研究的前 10 种动态手势具有较高的顽健性，但是，对于最后 2 种复杂手势顽健性较低。

在时间性能上，计算一个时长大约 2 s，分辨率为 120×120 的动态手势视频的光流和运动矢量场，

利用 PCA 方法在时间轴上计算所有运动特征的模式时间性能不能满足实时性需求，这主要是代码没有进行优化和算法本身复杂度导致的。

文献[12]利用动态时间规整（DTW, dynamic time warping）、改进的动态时间规整（IDTW, improved dynamic time warping）及相关信息法（CI, correlated information）对 12 种动态手势进行识别取得了一定的效果，文献[12]对 12 种动态手势识别时分别考虑了图像分辨率为 40×30、320×240，以及采样点个数分别为 80 和 120 的情况。为了便于与本文方法的比较，将图像分辨率统一调整为 240×240，采样点个数取 120，实验结果如表 1 所示。

表 1 与文献[12]中 3 种方法的准确率比较

方法	准确率
DTW	0.896 3
IDTW	0.923 4
CI	0.914 6
本文方法	0.937 6

从表 1 可以看出，本文方法与文献[12]的基于时间序列的手势轮廓模型中动态手势识别的 3 种方法相比，本文方法在识别准确率上具有一定的优势。但是在实际使用中，本文方法计算复杂，从而时间性能较低。

5 结束语

利用本文提出的基于多实例运动学模式学习的方法进行动态手势识别取得了较高的识别率，相比传统光流场特征、散度特征和旋度特征的识别性能都有较大提高，同时，相比 DTW, IDTW 和 CI 等方法也有一定的提升。但是，本文提出的方法对于复杂手势顽健性较低，而且该方法需要较大的时间开销，不能满足实时性要求，如何在提高顽健性的基础上降低时间开销，仍然需要做进一步的研究。

参考文献：

[1] 周航. 基于计算机视觉的手势识别系统研究[D]. 北京: 北京交通大学, 2008.
 ZHOU H, Research on gesture recognition system based on computer vision[D]. Beijing: Beijing Jiaotong University, 2008.

[2] 顾立忠. 基于表观的手势识别及人机交互技术研究[D]. 上海: 上海交通大学, 2008.
 GU L Z, Research on surface based hand gesture recognition and hu-

- man-computer interaction technology[D]. Shanghai: Shanghai Jiao-tong University, 2008.
- [3] KHAN R, HANBURY A, THINGER J et al. Color based skin classification [J]. Pattern Recognition Letters, 2012, 33(2):157-163.
- [4] SUN H M. Skin detection for single images using dynamic skin color modeling [J]. Pattern Recognition, 2010, 43(4):1413-1420.
- [5] MOLINA L, ESCUDERO-Viñolo M, SIGNORIELLO A, et al. Real-time user independent hand gesture recognition from time-of-flight camera video using static and dynamic models [J]. Machine Vision and Applications, 2013, 24(1):187-204.
- [6] FAN L. A feature-based object tracking method using online template switching and feature adaptation [C]//The 6th International Conference on Image and Graphics. 2011: 707-713.
- [7] LIU Y, ZHOU W, YIN H G et al. Tracking based on SURF and super-pixel[C]//The 6th International Conference on Image and Graphics. 2011:714-719.
- [8] WANG W, TUNG C L. Dynamic hand gesture recognition using hierarchical dynamic Bayesian networks through low-level image processing[C]//The 7th International Conference on Machine Learning and Cybernetics. 2008: 3247-3253.
- [9] SGOUROPOULOS K, STERGIOPOULOU E, PAPAMARKOS N. A dynamic gesture and posture recognition [J]. Journal of Intelligent & Robotic Systems, 2014, 76 (2) :283-296.
- [10] DURBIN P A, REIF B. Statistical Theory and Modeling for Turbulent Flows[M]. Second edition. John Wiley & Sons, Ltd., Chichester. 2011.
- [11] ALI S, SHAH M. Human action recognition in videos using kinematic features and multiple instance learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 32 (2) :288-303.
- [12] 庞海波, 李占波, 丁友东. 基于时间序列手势轮廓模型的动态手势识别[J]. 华南理工大学学报(自然科学版). 2015, 43(1):140-146.
- PANG H B, LI Z B, DING Y D. Dynamic hand gesture recognition

based on time series gesture contour model[J]. Journal of South China University of Technology (Natural Science Edition) 2015, 43(1): 140-146.

作者简介:



周彩秋 (1982-), 女, 黑龙江哈尔滨人, 南京理工大学博士生, 主要研究方向为物联网技术及安全、模式识别等。



杨余旺 (1966-), 男, 江苏南京人, 博士, 南京理工大学教授, 主要研究方向为物联网安全、网络编码、大数据等。



庞海波 (1979-), 男, 河南安阳人, 博士, 郑州大学讲师, 主要研究方向为计算机视觉和模式识别。