

文章编号 1004-924X(2016)08-2010-08

应用在线随机森林投票的动作识别

王世刚, 鲁奉军*, 赵文婷, 赵晓琳, 卢 洋
(吉林大学 通信工程学院, 吉林 长春 130012)

摘要:提出了基于在线随机森林投票识别人物动作类别的方法。建立了在线随机森林投票模型。通过在线训练和在线检测两部分进行了算法研究,提高了检测人物动作类别的准确率。基于人物动作在时间和空间上有重要信息,该方法首先通过提取图像立体块的 lab 色彩空间值、一阶差分、二阶差分以及大位移光流特征值在线训练随机森林;训练结束后,形成强分类器,利用分类器对检测图像进行投票,生成动作空间图;最后,在动作空间图中寻求最大值,判断检测图像的动作类别。验证结果表明在低分辨率的视频图像中,本方法能够确定人物的动作类别,对 Weizmann 数据库和 KTH 数据库的识别率分别为 97.3% 和 89.5%,对 UCF sports 数据库的识别率为 79.2%,动作识别准确率有所提高。该方法增加了光流能量场特征表述,将原始投票理论拓展至三维空间,并且采用向下采样的方式更新结点信息,能够判断人物动作类别,为智能视频技术提供了有效的补充信息。

关键词:动作识别;随机森林投票;大位移光流;动作空间图;智能视频

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/OPE.20162408.2010

Action recognition based on on-line random forest voting

WANG Shi-gang, LU Feng-jun*, ZHAO Wen-ting, ZHAO Xiao-lin, LU Yang

(College of Communication Engineering, Jilin University, Changchun 130012, China)

* Corresponding author, E-mail: lufengjun2012@sina.com

Abstract: An action recognition method for people is proposed based on on-line random forest voting to judge the action classification. The on-line random forest voting model is established and its algorithms are researched through the two parts consisting of on-line training and on-line detection to improve the precision of the action classification. As people action shows important information in both space and time, the method firstly trains the random forests in line by extracting 3D image features containing a lab color space, the first order difference, the second order difference and displacement optical flow. After training, a strong classifier is formed. Then, the classifier is used to vote for detection images to produce an action space map. Finally, by seeking the maximum in the map, the category of action in the detection images is complemented. Experimental results indicate that the method determines the category of people action in the low resolution video images. The accurate rates of the Weizmann data, the KTH data and the UCF sport data are 97.3%, 89.5%, and 79.2%, respectively. These results show that the accuracy of action recognition is improved. Moreover, the model proposed adds the feature representation of light flow energy field, expands the traditional forest voting theory

收稿日期:2015-11-09;修订日期:2015-12-14.

基金项目:教育部博士学科点专项科研基金资助项目(No. 20120061110091);吉林省科技发展计划资助项目(No. 20150204006GX);长春市科技局资助项目(No. 14KG007)

to a 3D space, and uses to update information. It improves the stability and the reliability and will be of potential application in the intelligent video surveillances.

Key words: action recognition; random forest voting; large displacement optical flow; action map; intelligent video

1 引言

人物的动作识别是计算机视觉的重要组成部分,通过分析视频图像中的人物特征判断动作类别,建立起特征数据与高层语义行为之间的关系。人物动作通常利用摄像机拍摄获得。动作识别系统在智能视频、人机交互技术、体育竞技场分析等方面有着广泛的应用。

动作识别系统将人物动作视为带有标记的特定视频图像序列,其主要包括视频图像预处理、人物特征提取、动作类别建模和分类器训练等多方面技术。其主要可分为时空特征方法和视频序列方法。早期的人物动作识别方法主要是针对单一静态背景视频序列的方法^[1],其主要将视频图像作为连续的观察序列,计算人物产生某种动作的特征似然度,并判断检测序列的后验概率,从而确定人物动作行为。

最近时空特征方法已成为主流研究方法^[2],其主要将兴趣点检测器和 BOW (bag-of-words) 的方法相结合,以增强人物动作的时空信息整合。视频图像在时间和空间方向上变化剧烈的特征点通常称为时空兴趣点,然后,通过聚类方法将特征向量形成时空单词集合,其经过训练得到时空单词模型,检测序列通过模型判断人物动作类别^[3-5]。但是对于低分辨率、运动模糊和镜头移动的视频序列,这些方法识别人物动作比较困难。而且,系统还需要克服低分辨率视频图像的影响,适应人物动作的突变性和复杂性,提高抗干扰能力的特点。

随机森林分类器由于在训练和检测阶段中分类效率比较高,处理速度快,在行为识别领域受到了高度关注^[6]。其分类方式主要在特征空间进行划分,在特定条件下对每个节点进行特征划分,这种方法的分类准确率取决于叶结点中数据类别的纯净度。

传统方法大多是基于人物的形态轮廓进行识

别^[7],对于复杂人体运动的分析描述并不准确。针对这问题,本文提出一种新颖的在线随机森林投票方法来确定人物的动作类别。本文引入了大位移光流能量场^[8]特征信息,从而可以明显地区分人物动作的外在客观条件及内在姿态复杂性。

光流算法是通过计算区域像素的变化来捕获物体运动信息的^[9],故可降低高层行为理解的复杂度,单像素的光流值计算较为困难,本文通过选取区域光流能量场作为特征提取的有用信息,以降低计算量。本文特征量值的优点如下:特征提取算法难度不高,可利用像素的底层信息作为特征的直接表述,与传统的基于轮廓外形特征的方法有明显区别,克服了传统特征表述信息表述不全面的缺点。随机森林能处理大量的高维数据,而且不会带来过度拟合问题,目前已经被广泛地应用到图像领域中^[10-11]。其在目标检测中^[12-13]已经成为一种重要的方法^[14],但是在人物动作识别中的应用较少,这主要因为其难以提取能够表述人物运动的特征信息。在目标检测领域,随机森林将目标色彩信息、区域像素点一阶差分及二阶差分和目标方向梯度直方图作为训练特征信息。本文通过增加了光流能量场特征表述,将原始投票理论拓展至三维空间,并且采用向下采样的方式更新结点信息。最后,通过实验验证了算法性能。

2 算法结构

动作识别过程主要分为在线训练部分和在线检测部分。在线训练部分主要通过提取图像立体块的 lab 色彩空间信息、差分以及大位移光流特征值训练随机森林,训练结束后,形成强分类器;在线检测部分首先遍历测试样本图立体块,再利用随机森林分类器对测试样本图立体块进行分类,叶子结点为所在立体块进行投票,在动作空间图中寻找最大值,判断人物动作类别。算法流程图如图 1 所示。

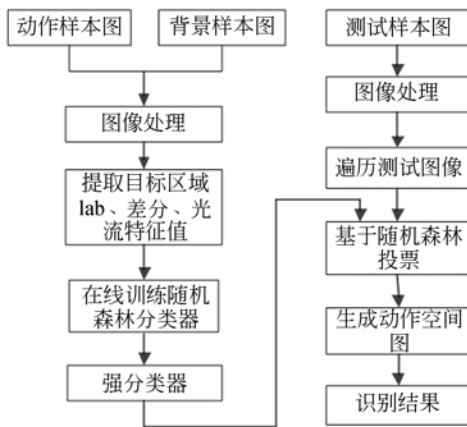


图 1 基于随机森林投票的动作识别算法流程图

Fig. 1 Flowchart of action recognition based on random forest voting algorithm

2.1 在线训练过程

2.1.1 图像处理

训练样本可分为正样本和负样本,正样本为已经标记边界框的人物的动作类别,如图 2 所示,图 2(a)被标记为滑板动作,图 2(b)被标记为跳水动作;负样本为无目标图像,如图 3 所示,图 3(a)和图 3(b)分别表示街道 1 和街道 2。



(a) 滑板动作 (a) Skateboard
(b) 跳水动作 (b) Diving

图 2 正样本

Fig. 2 Positive samples



(a) 街道 1 (a) Street 1
(b) 街道 2 (b) Street 2

图 3 负样本

Fig. 3 Negative samples

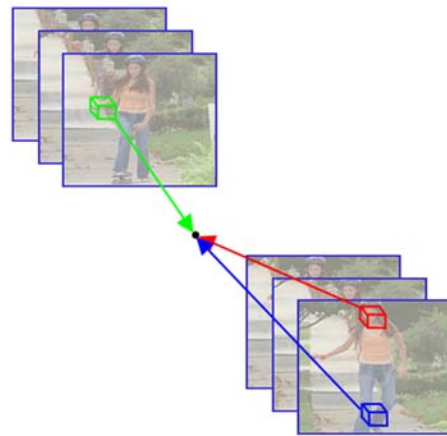


图 4 立体块特征信息

Fig. 4 Feature information of cuboids

人物动作在时间和空间中均含有信息,将图像中的连续动作采用时空立体块表示,立体块的大小为 $16 \times 16 \times 5$,它们分别表示立体块长和宽的像素长度以及图像帧数。如图 4 所示。

2.1.2 目标区域大位移光流特征值

光流法能够描述运动物体的特征信息,以图像像素点的变化表述物体的运动信息,在目标运动图像表达中有着重要的作用^[15]。大位移光流^[8]是由 Thomas Brox 等首先提出的,其主要用于描述图像像素点实时变化。文中提出了基于亮度、空间梯度和平滑区域能量变化最小的光流模型,而且又增加了平移不变的特征变换(Scale Invariant Feature Transform, SIFT)和颜色描述子匹配的内容,其能量模型公式如下:

$$E = E_{color} + \gamma E_{gradient} + \alpha E_{smooth} + \beta E_{match} + E_{desc}. \quad (1)$$

模型的最终目标是寻找式(1)中两帧图像最小能量值变化值,其中, E_{color} 表示相邻两帧图像亮度变化值,受光照影响, E_{color} 表达的信息有限,因此,增加了空间梯度能量最小值 $E_{gradient}$ 。然而, E_{color} 和 $E_{gradient}$ 较少表达描述性特征,单一估算这两部分的能量值并不能完善地表述光流场的特性,故又增加了平滑区域能量值 E_{smooth} 来避免光流场误算的问题。为了增加描述子特征的信息,将 SIFT 和色彩描述子作为匹配对象,能量的表达形式为 E_{match} 和 E_{desc} ,可变参数 α, β 和 γ 用以适应不同帧图像的运动剧烈程度。图 5 表示光流场强度,5(a)和 5(b)为两帧相邻图像,由图可知,两帧图像的人物运动变化较大,5(c)和 5(d)分别表示在 x, y 方向的光流场强度值,由图 5 可知大位移

光流特征值对于运动物体的信息表述更为准确。

2.1.3 训练

首先,提取 $16 \times 16 \times 5$ 彩色立体块 $\{P_i = (I_i, c_i, d_i)\}$ 的特征值,其中 P_i 表示三维立体块; I_i 表示第 i 立体块提取的特征值, $I_i = (I_i^1, I_i^2, \dots, I_i^F)$ 可容纳多通道特征值, I_i^f 表示在第 i 个立体块中特征通道 $f \in (1, 2, \dots, F)$ 对应的特征值; c_i 表示为动作类别标签,标记不同数字的立体块代表不同的动作类别,标记数字 0 的立体块代表负样本; d_i 表示立体块中心到达时空域动作中心的三维矢量值。本文选取的特征值为 lab 色彩空间值、差分以及大位移光流特征值,其中,特征值 lab 色彩空间可通过对图像 RGB 颜色空间的转换求得;特征值差分可通过像素矩阵求得;大位移光流特征值的计算过程如 2.1.2。这 3 类特征值可准确表述人物动作的运动信息。

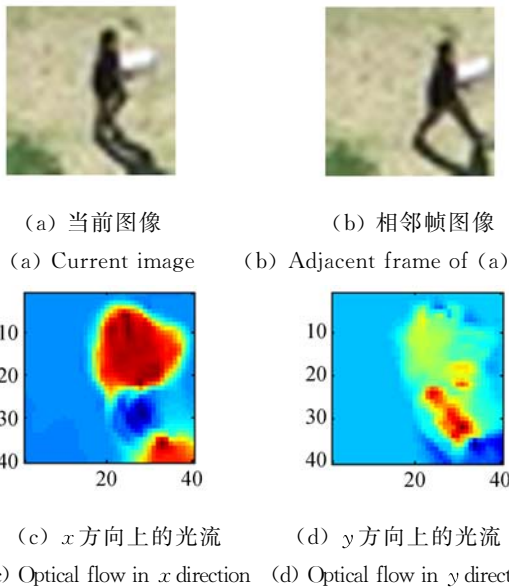


图 5 光流场强度

Fig. 5 Magnitude of optical flow field

其次,对提取的特征值进行分类,本文主要通过二值测试实现,具体过程如下:比较位置 $p \in R^3$ 和 $q \in R^3$ 中通道 f 上特征值的大小。在结点 B 的二值测试定义如下:

$$t_{B, f, p, q, \tau}(\Gamma) = \begin{cases} 0 & I^f(p) < I^f(q) + \tau \\ 1 & \text{其他} \end{cases}, \quad (2)$$

$I^f(p)$ 表示在特征通道 f 中位置 p 的特征值; $t_{B, f, p, q, \tau}(\Gamma)$ 表示在立体块中位置 p 和 q 特征值的比较值, τ 为设定阈值。在同一特征通道 f 下,比较不同位置处的特征值将立体块分为两大

类,每一类继续通过二值测试进行二分类,构成了二值测试集合 $\{t^k\}$,其中 k 表示集合的数量。

对二值测试集合中的立体块需要进行最佳分类,通过随机森林分类器进行分类,随机森林 F 由决策树 $\{T_1, T_2, \dots, T_N\}$ 构成, N 表示决策树的数目,每一棵决策树的函数形式为 $f(x; \Theta): X \rightarrow Y$, Θ 表示在树的分裂结点中定义参数, $x \in X \subseteq R^d$ 表示特征值。

在训练过程中,训练样本集合表示为 $\{x_i, y_i\} \in \{X, Y\}$,其中 $y_k \in \{1, 2, \dots, K\}$, K 表示动作类别的数目。随机森林的构建是一个递归过程,对于树中的每个结点,通过计算信息增益 ΔH 或者立体块中心偏移值 ΔD 寻求最佳分裂函数 $\zeta(\cdot)$,其中 ΔH 表示为:

$$\Delta H = H(I_p) - \frac{|I_l|}{|I_p|} H(I_l) - \frac{|I_r|}{|I_p|} H(I_r), \quad (3)$$

其中: I_p , I_l 和 I_r 分别表示双亲结点及其左右子结点的样本集合; $H(\cdot)$ 为熵值 $-\sum_{k=1}^K p_k \log p_k$, p_k 表示当前子结点中动作属于 k 类的概率, $H(I_l)$ 和 $H(I_r)$ 表示落入左右子结点中样本的动作类别熵值。

立体块中心偏移值用 ΔD 表示,其数学表达式为:

$$\Delta D = \sum_{l_i} \|d_i^l - \bar{d}^l\|^2 + \sum_{r_i} \|d_i^r - \bar{d}^r\|^2, \quad (4)$$

其中: d_i^l 和 d_i^r 分别表示落入左右子结点中第 i 个立体的块矢量积; \bar{d}^l 和 \bar{d}^r 分别表示落入左右子结点中立体块矢量积均值;

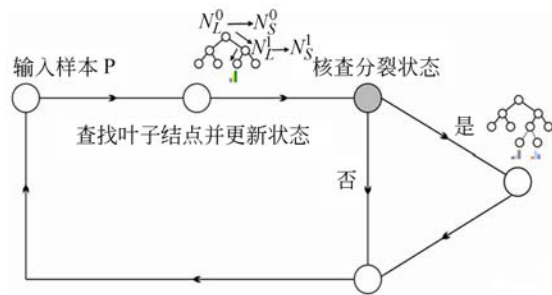


图 6 向下取样流程图

Fig. 6 Flowchart of subsamples

在线随机森林分类方法与原始方法^[16]不同,本文采用向下取样方式,如图 6 所示。从叶结点 N_i^0 开始分裂,首结点也称为根结点,0 表示当前随机树的深度。首先,将输入样本 P 分解成训练图像立体块,立体块依次到达根结点,当立体块数

目达到 20 时,叶结点 N_L^0 开始转变为分裂结点 N_S^0 ,随机选取式(3)和式(4)两个测量标准,循环迭代寻求最佳分裂函数 $\zeta(\cdot)$ 。根据函数 $\zeta(\cdot)$ 可知,叶结点 N_L^0 将立体块分配给左右结点,如果右结点 N_L^1 中的立体块数量大于 20,则 N_L^1 继续分裂为 N_S^1 ,否则结点进入等待模式,当满足条件后,继续分裂。循环上述步骤,使每个叶结点都进行分裂,当达到树的最大深度($d_{Max}=15$)或者子结点中所含立体块数量小于最大立体块的数量($N_{max}=20$)时,停止分裂,在线随机森林构建完成。在原始方法中所有训练立体块都从根结点开始分裂,训练样本数量受限制,不能持续更新,而本文在线随机森林的分裂方式能够适应图像在线更新,不需要提前收集训练样本数量。

叶子结点存储 2 个信息,一是在训练中到达该叶子结点不同动作类别概率值 p_c^l ($\sum_c p_c^l = 1$),二是立体块中心偏移量的所有集合 $D_c^l =$

$$p(Q_c(y) | I(x)) = \sum_{l \in c} p(Q_c(y) | c(x) = l, I(x)) p(c(x) = l | I(x)) = p(Q_c(y) | c(x) = c, I(x)) p(c(x) = c | I(x)) = p(d(c, x) | c(x) = c, I(x)) p(c(x) = c | I(x)). \quad (5)$$

通过将立体块 $p(x)$ 遍历随机森林得到式(5)中的参数值。假设在决策树中,立体块处在的叶子 L 处。本文采用 parzen 窗估计法处理式(5)中的第一个参数因子,将 p_c^l 作为第二个参数处理^[16]。重写式(5):

$$p(Q_c(y) | I(x), T) = \left[\frac{1}{|D_c^l|} \sum_{d \in D_c^l} G((x - y) - d) \right] \cdot p_c^l, \quad (6)$$

其中 $G(\cdot)$ 表示三维高斯窗函数。

对整个随机森林 T 投票平均处理,得到(7)式:

$$P(Q_c(y) | I(x), T) = \frac{1}{|T|} \sum_t p(Q_c(y) | I(x), T_t). \quad (7)$$

2.2.2 动作空间图

为了能够准确定位目标的动作类别,将检测序列分成 $16 \times 16 \times 5$ 大小的立体块,立体块遍历构建的随机森林,叶子结点为到达的立体块进行动作类别投票,即计算立体块所在叶子结点中 $\frac{P_c^l}{|D_c^l|}$ 值。利用图像中每一像素点对应的立体块

$\{d_i\}_{c_i=c}$ 。存储的这 2 个信息将会为检测立体块进行投票。

2.2 在线检测过程

2.2.1 随机森林投票理论

在目标检测方面,随机森林投票方法应用较为广泛,其主要思想是,通过有监督学习方法,将二维图像块特征信息和随机森林投票建立映射联系,利用叶子结点判断目标类别。本文通过对其进行扩展,将随机森林的投票表决方法与三维图像立体块信息相结合进行描述。

以位置 $x \in R^3$ 为中心的立体块表示为 $p(x) = (I(x), c(x), d(c(x), x))$ 。其中: $I(x)$ 表示立体块特征, $c(x)$ 表示立体块的未知动作类别, $d(c(x), x)$ 表示立体块中心到未知动作中心矢量。 $Q_c(y)$ 表示图像中任意一点 $y \in R^3$ 处具有动作类别 c 的事件。根据贝叶斯公式推导出的条件概率 $p(Q_c(y) | I(x))$ 表示如下:

得到叶子投票信息,经过三维高斯平滑滤波,提取动作类别和时间信息,得到二维动作空间图,求解动作空间图最大值,从而确定人的动作类别,最后将得到的不同动作类别信息进行收集,评估序列动作类别检测的准确性。

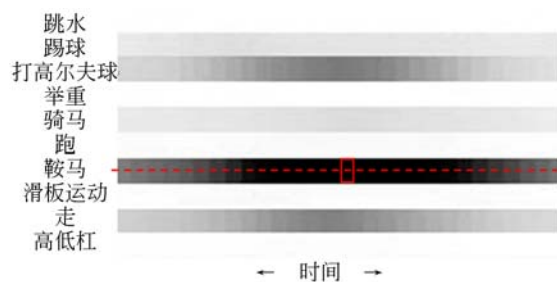


图 7 动作空间图

Fig. 7 Spatial graph of different actions

图 7(彩图见期刊电子版)表示 UCF sports 数据库的动作空间图,红色虚线表示鞍马运动连续帧图像的检测结果,红色小矩形框表示随机森林对鞍马运动其中一帧投票的结果。动作空间图能够准确判断此动作为鞍马运动。

3 实验结果及分析

本文实验的硬件运行环境为 Intel Pentium (R) Dual-Core CPU、物理内存 2G 的 PC 机;软件的运行环境为 Ubuntu 系统,借助 Opencv2.0 开源计算机视觉库训练随机森林分类器。

实验中从国际标准人物动作数据库中选用 3 组测试样本,其中包括单人静态背景下的 Weizmann 数据库和 KTH 数据库,以及单人动态背景下 UCF sports 数据库。

3.1 Weizmann 数据库、KTH 数据库和 UCF sports 数据库

Weizmann 数据库^[17]中含有 10 种不同的动作类别,其中每组动作由 9 个不同的表演者分别实现。本实验将每组动作中的 8 个视频序列作为训练序列,第 9 个视频序列作为检测序列,每次实验重复 9 次,保证每个表演者都被检测,最后将检测结果取平均值作为最终判断结果。Weizmann 数据库的检测结果如图 8 所示。

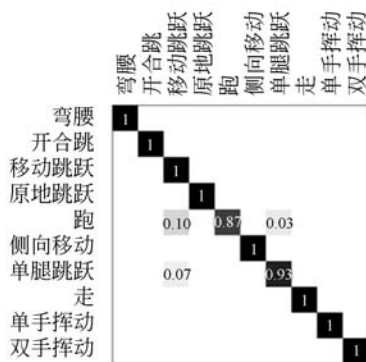


图 8 Weizmann 数据库混淆矩阵

Fig. 8 Confusion matrices for Weizmann dataset

KTH 数据库^[18]中含有 6 个不同的动作类别,这 6 个动作由 25 个表演者在 4 个不同的场景下拍摄所得。本文将同一人在 4 个不同场景下的同类动作视为动作完整图像序列,并将 20 个表演者的完整图像序列作为训练序列,剩下的 5 个用于检测,以每 5 组一轮换的方式进行训练和检测,最后将检测结果取平均值作为最终判断结果。KTH 数据库的检测结果如图 9 所示。

UCF sports 数据库^[19]含有 10 组不同的动作类别,每组动作由 15 个不同的视频序列组成。

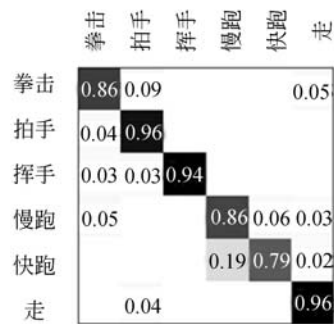


图 9 KTH 数据库混淆矩阵

Fig. 9 Confusion matrices for KTH dataset

本文在每类动作中依次选取 5 大组,大组轮换次数为 15;每大组中又包括 5 个小组,其中 4 个小组作为训练序列,剩下 1 组作为检测序列,小组轮换次数为 5,最后计算检测结果的平均值。UCF sports 数据库的检测结果如图 10 所示。

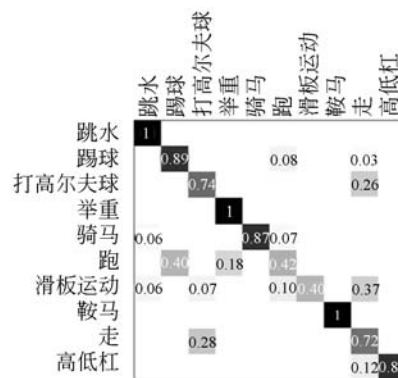


图 10 UCF sports 数据库混淆矩阵

Fig. 10 Confusion matrices for UCF sports dataset

3.2 动作分类比较结果

BOW 算法把已提取的人体局部轮廓作为特征信息,通过 K 均值聚类算法将特征分为 K 类;则在视频中的每帧图像形成长度为 K 的特征向量。本文主要与随机森林在词袋模型 BOW 中应用的方法进行对比。词袋中的特征向量通过随机森林(Random Forest, RF)和支持向量机(Support Vector Machine, SVM)进行分类识别,文中将这两种方法分别称为 BOW + RF 和 BOW + SVM。根据词袋外数据误差进行分析,又可形成 RF-1 方法^[7]。考虑到提取轮廓特征信息的维数比较大,RF-1 算法应用 PCA (Principal Component Analysis)算法对特征信息进行降维,其在随

机森林中使用决策树的数目为 500。另外,还选用文献[5]基于时空兴趣点的人体动作识别方法进行对比。

表 1 各方法在 Weizmann、KTH 和 UCF sports 数据库的比较结果

Tab. 1 Comparison of classification results on Weizmann, KTH and UCF sports data base for differen algorithms (%)

| method | Weizmann | KTH | UCF sports |
|----------|----------|------|------------|
| BOW+SVM | 53.76 | 47.2 | 30.8 |
| BOW+RF | 87.10 | 76.5 | 53.7 |
| RF-1 | 91.40 | 82.3 | 60.1 |
| Ref. [5] | 92.7 | 81.0 | 66.4 |
| 本文 | 97.3 | 89.5 | 79.2 |

本文不需要对特征维数进行降维,构造的决策树数目为 15 棵,将本文算法与其他算法的实验

结果进行比较,如表 1 所示。由表 1 可知,本文方法在 Weizmann 数据库的识别率达 97.3%,KTH 数据库识别率达 89.5%,较为复杂的 UCF sports 数据库识别率达 72%。结果表明,本文算法识别效果高于其他算法。

4 结 论

针对人物动作识别的问题,本文提出了一种应用在线随机森林的动作识别方法,重点介绍了大位移光流特征的提取方法,并构建了随机森林和在线检测方法。实验结果表明:在低分辨率的视频图像中,本文方法对简单动作数据库 Weizmann 和 KTH 的识别率分别为 97.3% 和 89.5%,对复杂动作数据库 UCF sports 的识别率为 79.2%。该方法能够判断人物动作类别,将为智能视频技术提供了有效的补充信息。

参考文献:

- [1] SCHULDT C, APTEV I, CAPUTO B. Recognizing human actions: a local SVM approach [C]. *IEEE, Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, the United Kingdom, 2004*:32-36.
- [2] DOLLAR P, ABAUD V R, COTTRELL G, et al.. Behavior recognition via sparse spatio-temporal features [C]. *Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Beijing, P. R. China. 2005*:65-72.
- [3] 付朝霞,王黎明. 基于时空兴趣点的人体行为识别[J]. *微电子学与计算机*, 2013,30(8): 28-35.
FU ZH X, WANG L M. Human action recognition based on space-time interest point [J]. *Microelectronics & Computer*, 2013,30(8):28-35. (in Chinese)
- [4] 王博,李燕. 视频序列中的时空兴趣点检测及其自适应分析[J]. *计算机技术与发展*, 2014,24(4): 49-56.
WANG B, LI Y. Space-time interest points detection in video sequence and its adaptive analysis [J]. *Computer Technology and Development*, 2014, 24(4):49-56. (in Chinese)
- [5] 王世刚,孙爱滕,赵文婷,等. 基于时空兴趣点的单
- 人行为及交互行为识别[J]. *吉林大学学报(工学版)*, 2015,45(1): 304-308.
WANG SH G, SUN A M, ZHAO W T, et al.. Single and interactive human behavior recognition algorithm based on spatio-temporal interest point [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2015, 45(1): 304-308. (in Chinese)
- [6] 鲍一丹,陈纳,何勇,等. 近红外高光谱成像技术快速鉴别国产咖啡豆品种[J]. *光学 精密工程*, 2015, 23(2): 349-355.
BO Y D, CHEN N, HE Y, et al.. Rapid identification of coffee bean variety by near infrared hyperspectral imaging technology [J]. *Opt. Precision Eng.*, 2015, 23(2): 349-355. (in Chinese)
- [7] 蔡加欣,冯国灿,汤鑫,等. 基于局部轮廓和随机森林的人体行为识别[J]. *光学学报*, 2014, 34(10): 1015006(1-10).
CAI J X, TANG G C, TANG X, et al.. Human action recognition based on local image contour and random forest [J]. *Acta Optica Sinica*, 2014, 34(10): 1015006(1-10). (in Chinese)
- [8] BROX T, MALIK J. Large displacement optical flow: descriptor matching in variational motion estimation [J]. *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(3):

- 500-513.
- [9] 徐富元, 顾国华, 陈钱, 等. 转动红外探测器下地面远距离运动目标检测方法[J]. 红外与激光工程, 2014, 43(4): 1080-1086.
XU F Y, GU G H, CHEN Q, *et al.*. Ground target detection method on rotating infrared detector [J]. *Infrared and Laser Engineering*, 2014, 43(4): 1080-1086. (in Chinese)
- [10] 向涛, 李涛, 李旭东, 等. 基于随机森林的层次行人检测算法[J]. 计算机应用研究, 2015, 32(7): 2196-2199.
XIANG T, LI T, LI X D, *et al.*. Random forests for hierarchical pedestrian detection [J]. *Application Research of Computers*, 2015, 32(7): 2196-2199. (in Chinese)
- [11] 屠大维, 江济良. 改进的光流运动图像分析方法及其应用 [J]. 光学精密工程, 2011, 19(5): 1159-1164.
TU D W, JIANG J L. Improved algorithm for motion image analysis based on optical flow and its application [J]. *Opt. Precision Eng*, 2011, 19(5): 1159-1164. (in Chinese)
- [12] 刘翔, 杨鑫, 王蕾. 改进的基于码本模型目标检测算法[J]. 液晶与显示, 2014, 29(6): 999-1002.
LI X, YANG X, WANG L. Object detection algorithm based on improved codebook model [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(6): 999-1002. (in Chinese)
- [13] 林培杰, 郑柏春, 陈志聪, 等. 面向多区域视频监控的运动目标检测系统[J]. 液晶与显示, 2015, 30(3): 484-491.
LI P J, ZHENG B CH, CHEN ZH C, *et al.*. System of multi-regions moving object detection in video surveillance [J]. *Chinese Journal of Liquid Crystals and Displays*, 2015, 30(3): 484-491. (in Chinese)
- [14] 胡梦婕, 魏振忠, 张广军. 基于对象性测度估计和霍夫森林的目标检测方法[J]. 红外与激光工程, 2015, 44(6): 1936-1941.
HU M J, WEI ZH ZH, ZHANG G J. Object detection method based on objectness estimation and Hough forest [J]. *Infrared and Laser Engineering*, 2015, 44(6): 1936-1941. (in Chinese)
- [15] 赵海, 陈星池, 王家亮, 等. 基于四轴飞行器的单目视觉避障算法[J]. 光学精密工程, 2014, 22(8): 2232-2241.
ZHAO H, CHEN X CH, WANG J L, *et al.*. Obstacle avoidance algorithm based on monocular vision for quad-rotor helicopter [J]. *Opt. Precision Eng.*, 2014 22(8): 2232-2241. (in Chinese)
- [16] JUERGEN G, VICTOR L. Class-specific Hough forests for object detection [C]. *IEEE Transactions on Computer Vision and Pattern Recognition, Miami, the United States 2009*: 1022-1029.
- [17] BLANK M, GORELICK L, SHECHTMAN E, *et al.*. Actions as space-time shapes [C]. *IEEE International Conference on Computer Vision, Beijing, China, 2005*.
- [18] SCHULDT C, LAPTEV I, CAPUTO B. Recognizing human actions: a local SVM approach [C]. *Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 2004*.
- [19] RODRIGUEZ M, AHMED J, SHAH M. Action mach-a spatio-temporal maximum average correlation height filter for action recognition [C]. *2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, 2008*.

作者简介:



王世刚(1962—),男,吉林长春人,教授,博士生导师,1983年于东北大学获得学士学位,1997年于吉林工业大学获得硕士学位,2001年于吉林大学获得博士学位,研究方向为数字图像处理技术及应用。E-mail: wangshigang@vip.sina.com

通讯作者:



鲁奉军(1991—),男,吉林德惠人,硕士研究生,2013年于吉林大学获得学士学位,主要从事数字图像处理、机器学习等方面的研究。E-mail: lufengjun2012@sina.com