

doi: 10.3969/j.issn.1006-1576.2012.01.016

人体红外图像序列威胁行为理解

李永浩¹, 张西红², 赵欣³, 彭涛⁴

(1. 军械工程学院研究生管理大队, 石家庄 050003; 2. 军械工程学院计算机工程系, 石家庄 050003;
3. 河北测绘局制图院, 石家庄 050000; 4. 军械工程学院训练部教育技术中心, 石家庄 050003)

摘要: 为识别人体的威胁行为, 对提取的人体红外图像行为序列特征进行研究, 提取每一帧图像的 HOG 特征并且并联形成序列二维特征。为降低特征维数, 对二维图像进行离散余弦变换, 使用变换后的包含该二维图像主要信息的前两列近似代替该图像的主要特征, 最终使用最小二乘支持向量机对降维后的多段图像序列进行训练和测试, 并给出最终训练测试结果。结果表明, 该特征能实现对几种典型的威胁行为识别。

关键词: 红外图像; 威胁行为识别; 支持向量机; 方向梯度直方图; 离散余弦变换
中图分类号: TJ03 **文献标志码:** A

Threat Behavior Identifying to Infrared Image of Human

Li Yonghao¹, Zhang Xihong², Zhao Xin³, Peng Tao⁴

(1. *Administrant Brigade of Postgraduate, Ordnance Engineering College, Shijiazhuang 050003, China;*
2. *Dept. of Computer Engineering, Ordnance Engineering College, Shijiazhuang 050003, China;*
3. *Dept. of Mapping Institute, Surveying & Mapping Bureau of Hebei, Shijiazhuang 050000, China;*
4. *Education Technology Center, Dept. of Training, Ordnance Engineering College, Shijiazhuang 050003, China*)

Abstract: To identify threat behavior of human, researched the feature of infrared image, extracted HOG of every image, propose to extract the human infrared image sequence characteristics which are compressed by discrete cosine transform (DCT), and train support vector machine (SVM). Using the first and the second arrange of feature compressed by DCT, finally this feature recognize some typical threat behavior.

Key words: infrared image; threat behavior identifying; SVM; histograms of oriented gradients (HOG); DCT

0 引言

通过对人体红外图像的分析能够识别出威胁姿态, 及时发现潜在的敌人^[1-2], 这对打击敌人, 保存我单兵生命至关重要。但如果仅仅对人体威胁姿态进行识别, 还不能达到具体的实战要求, 因为姿态反映的是某一静态的特征, 而许多威胁必须要经过对动作整体的观察和理解才能达到具体的识别需要。目前对人体行为理解的常用行为识别方法有: 使用模板匹配法(基于时空特征的模式分析方法)和状态空间法(基于概率的方法)。模板匹配法虽然计算复杂度低、实现简单, 但是对于噪声和运动时间间隔的变化敏感^[3]。状态空间方法最有代表性的是 HMMs^[4], 虽然能克服模板匹配的缺点, 并且已得到了广泛的应用, 但通常涉及到复杂的迭代运算。因此, 笔者提出提取人体红外图像序列中每一帧图像的 HOG 特征, 将这些 HOG 特征并联组成二维图像。通过识别二维图像来实现对人体行为的理解, 实现了对人体威胁行

为的准确识别。

1 威胁行为分析

行为由一系列的姿态组成^[1], 威胁行为由一系列的人体复杂姿态构成。在文献[1]中, 通过对单副人体红外图像姿态的分析, 将人体的行为分为威胁行为和正常的人体行为, 将这些姿态通过一定顺序进行组合就形成了单兵个体的行为。为了全面反映单兵的威胁行为, 首先对可能存在的威胁行为进行了分析、提炼, 可以划分为以下几种典型的威胁行为:

1) 姿态的变化为正常姿态 \Rightarrow 正常姿态时, 此类行为可以判定为正常的人体行为, 主要有以下可能:

- ① 正面行走;
- ② 侧身移动(向左和向右);
- ③ 背对后退;
- ④ 正面跑动;
- ⑤ 转身(左右前后 8 种组合);
- ⑥ 蹲下到站立(双向);
- ⑦ 左跨步和右跨步;

收稿日期: 2011-08-23; 修回日期: 2011-09-20

基金项目: 国家自然科学基金“基于传感器网络的分布式目标检测”资助项目(60672143)

作者简介: 李永浩(1980—), 男, 河南人, 硕士, 从事图像处理与模式识别研究。

⑧ 弯腰直身(双向)。

2) 姿态的变化为威胁姿态⇒正常姿态时, 此类行为可以判定为溃败的人体行为:

- ① 持枪逃跑;
- ② 背枪逃跑;
- ③ 双手持枪逃跑;
- ④ 收枪逃跑;
- ⑤ 边攻击边逃窜;
- ⑥ 弃枪逃跑。

3) 姿态的变化为威胁姿态⇒威胁姿态时, 此类行为可以判定为冲击或者定位攻击的人体行为:

- ① 持枪正面冲击;
- ② 持枪侧面冲击;
- ③ 端枪冲击;
- ④ 低姿跃进携枪冲击;
- ⑤ 背枪跃进;
- ⑥ 屈身枪口向下跃进;
- ⑦ 高姿匍匐携枪冲击;
- ⑧ 侧高姿携枪冲击;
- ⑨ 抱枪冲击;

⑩ 侧姿匍匐冲击;

⑪ 卧姿匍匐冲击。

4) 姿态的变化为正常姿态⇒威胁姿态时, 此类行为可以判定为开始实施攻击的人体行为:

- ① 左持枪+右持枪;
- ② 跪姿持枪;
- ③ 立姿正面持枪;
- ④ 卧姿持枪;
- ⑤ 侧身持枪;
- ⑥ 连续持枪姿态微变(正面, 侧面, 蹲姿);
- ⑦ 持枪站到蹲下(双向);
- ⑧ 侧面到正面持枪转化(双向);
- ⑨ 侧面到蹲下持枪转化(双向)。

通过对以上典型行为分析, 笔者使用 4 个客体人作为采集对象, 在红外热成像仪下按照上面分析的行为进行动作。共采集样本 320 个, 每一个样本包含 30 帧的红外图像。使用 2 个客体的行为作为训练目标, 使用 2 个客体的行为作为测试目标。其中一个攻击样本的部分帧如图 1 所示。



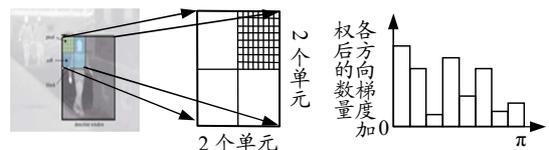
图 1 攻击行为红外人体图像

2 特征提取和分类方法研究

2.1 并联加权 HOG 特征表征运动行为

笔者使用图像加权的 HOG 特征作为每一帧图像的特征^[1]。HOG 特征的描述可以用图 2 简单描述^[5]。从图像零坐标开始, 每个相邻 8×8 的像素为一个单元, 每 4 个相邻的单元组成一个 2×2 的块。这样每个块就与周围的 8 个块至少有一个单元的重叠。将每个单元中的每个梯度进行方向投票加权, 如果设为 9 个方向, 那么每个单元就得到 9 维的特征向量, 将块中的特征向量串联起来, 每个块就形成了 4×9=36 维的特征向量, 求该块的 HOG 特征能量作为该块特征的权值。如果是一个 64×128 的图像则有 7×105 个块。将每个块的特征向量串联起来最终形成 105×36=3 780 维的特征向量, 这个向量就是 HOG 特征, 是对图像进行分类的依据。图 2 是 HOG 特征的获取过程。为了增加边缘部分在分类时的贡献率, 更好地表征姿态行, 对每一块的特征按

照能量的大小进行加权, 这套理论在文献[1]中判断威胁姿态取得了良好的效果。



(a) 红外视场中 (b) 对滑动窗口 (c) 一个单元HOG 分割获取识别对象 中块的描述 特征获取方式

图 2 HOG 的获取过程

笔者采集的特征描述如下: 一段视频包含 30 帧(3 s), 将 30 帧的加权 HOG 特征并联起来作为表征行为的方法。每一幅图像的 HOG 特征表示为: $I(n), n=1, \dots, 3\ 780$ 。并联起来的加权 HOG 组成了一幅二维矩阵:

$$I_N = \begin{bmatrix} I_n(1,1) & I_n(1,2) & \dots & I_n(1,m) \\ I_n(2,1) & I_n(2,2) & \dots & I_n(2,m) \\ \vdots & \vdots & & \vdots \\ I_n(n,1) & I_n(n,2) & \dots & I_n(n,m) \end{bmatrix} \quad (1)$$

其中: $n=3\ 780$, $m=30$ 。如图 3, 用图像表示该矩阵。该图像纵向表征了每一帧图像的姿态, 横向则表征了姿态的变化。所以该二维图像是能够比较客观的反映人体的行为, 通过对该二维图像的认识就能实现对人体行为的理解。

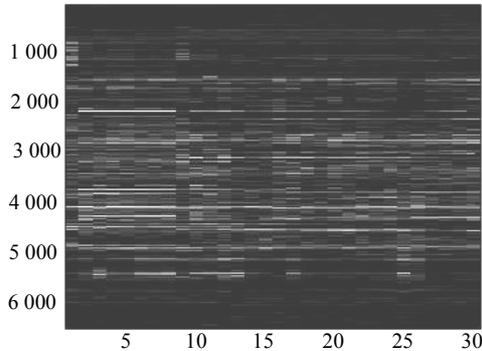


图 3 并行加权 HOG 特征二维图

2.2 行为表征降维处理

由于该二维图像包含的特征太大, 直接进行分类非常困难, 而且当维数过高时, 分类效果会减低, 甚至不能有效地分类^[6]。为解决此问题采用离散余弦变换的方式进行降维。离散余弦变换为:

$$D(u, v) = \alpha(u)\alpha(v) \cdot \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} I(x, y) \times \cos \frac{\pi(2x+1)u}{2N} \times \cos \frac{\pi(2y+1)v}{2M} \quad (2)$$

$$0 \leq u \leq N-1, 0 \leq v \leq M-1$$

式中: $\alpha(u) = \begin{cases} 1/\sqrt{N}, u=0 \\ \sqrt{2/N}, 0 \leq v \leq N-1 \end{cases}$,

$$\alpha(v) = \begin{cases} 1/\sqrt{M}, v=0 \\ \sqrt{2/M}, 0 \leq v \leq M-1 \end{cases}。$$

M 、 N 为图像长和宽的最大值; $D(u, v)$ 为变换结果, 也称为 $I(x, y)$ 的 DCT 变换系数。

对该二维图像进行离散余弦变换, 变换后生成与原图像大小一致的矩阵 $D(u, v)$ ^[7]。经过余弦变换后的图像, 绝大多数的信息集中在矩阵的前几列, 可以说前几列集中包含了原二维图像的绝大多数信息, 为了直观表示经过离散余弦变换后图像的能量分布情况, 对纵向能量进行了绝对值的求和, 结果如图 4。变换后的能量信息集中在前 2 列。由于离散余弦变换是可逆的, 笔者用前 2 列近似表征该 30 帧的威胁行为。

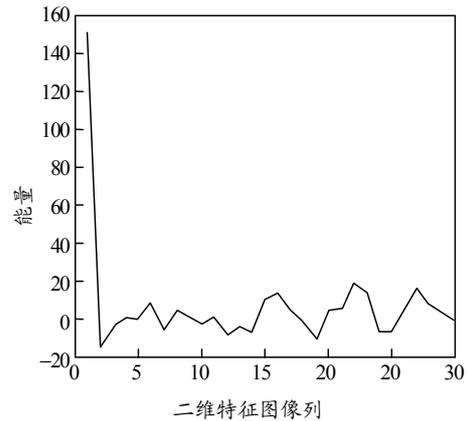


图 4 经过离散余弦变换后二维特征图像的纵向能量分布

2.3 支持向量机分类

考虑到支持向量机的处理小样本量的优势, 笔者采用最小二乘多分类支持向量机对采集的样本进行训练和测试^[7]。在该分类方法中对 4 个类别构造 4 个支持向量机, 每一个支持向量机分别将某一类的数据从其他类别中分离出来。在测试时, 取决策函数输出值最大的类别为测试样本的类别。其第 i 个 SVM 可通过解决下面的最优化问题得到:

$$\min_{w^i, \xi^i} \frac{1}{2} (w^i)^T w^i + C \sum_{j=1}^l \xi_j^i \quad (3)$$

其约束条件为:

$$\begin{cases} (w^i)^T \phi(x_i) \geq 1 - \xi_j^i, y_j = i \\ (w^i)^T \phi(x_i) \leq 1 + \xi_j^i, y_j \neq i \\ \xi_j^i \geq 0, j=1, \dots, l \end{cases} \quad (4)$$

求解式 (4) 后, 得 4 个决策函数:

$$\begin{aligned} &(w^1)^T \phi(x) + b^1 \\ &(w^2)^T \phi(x) + b^2 \\ &(w^3)^T \phi(x) + b^3 \\ &(w^4)^T \phi(x) + b^4 \end{aligned}$$

则 x 所属类为 $\arg \max_{i=1,4} [(w^i)^T \phi(x) + b^i]$

“一对多” SVMs 简单、有效, 训练时间较短, 可用于大规模数据。但其缺点在于: 1) 当类别数较大时, 某一类的训练样本将大大少于其它类训练样本的总和, 这种训练样本间的不均衡将对精度产生影响; 2) 存在误分、拒分区域; 3) 泛化能力相对较差。笔者需要分类的样本只有 4 类, 类别较少, 但是特征的维数比较高, 达到 7 560, 所以使用该支持向量机比较适合。

3 训练测试实验过程描述

- 1) 采集红外人体图像视频样本;
- 2) 并将视频序列分为 2 个部分, 分别用于训练和测试;
- 3) 提取视频中每一帧的加权 HOG 特征;
- 4) 每一序列视频中的单幅图像加权 HOG 特征并联, 构建二维视频 HOG 特征;
- 5) 对所有的二维 HOG 特征进行离散余弦的变化, 用前 2 列近似代替整个二维 HOG 特征;
- 6) 使用训练样本对最小二乘支持向量机进行训练, 使用测试样本进行分类的测试。

4 Matlab 的实现

实验采集的样本大小为 128×64(像素), 每个视频样本包含 30 帧的图像。采集样本的方法使用节 2 中叙述的方法。一共有 320 个样本, 使用 2 人的行为作为训练样本, 另外 2 人的行为作为测试样本。使用最小二乘支持向量机进行训练分类, 特征使用文中提出的并联加权 HOG 特征, 并经过降维后作为行为表征。最终训练测试的结果如表 1。

表 1 最终训练测试结果

测试	结果				/%
	行为	实	和	分数量	
行为	40				100
实		37	1	2	92.5
和			37	3	92.5
		1	1	38	95

实验结果证明:

- 1) 使用并联加权 HOG 特征的方法能够非常好地表征人体的行为特征。
- 2) 采取离散余弦变换的方法对识别的特征进行降维处理, 极大地减少了运算量, 实验证明: 经过变换使用相对少维数就能实现对威胁行为的分类。
- 3) 使用最小二乘支持向量机进行威胁行为的分类, 能够很好地区分出几种典型的威胁行为。同时, 通过实验也发现了该方法的不足:
 - 1) 由于采集的样本量有限, 建立的库只是反映

了一些非常典型的威胁行为, 建立更加复杂的威胁行为库异常困难。

- 2) 对于最小二乘支持向量机多分类的拒分问题, 还没有很好的解决方法。
- 3) 单兵行为的冲击和边攻击边逃窜部分行为有非常高的相似性, 较难区分。

5 结束语

笔者采集了几种典型威胁姿态, 建立了典型威胁行为的样本库, 对行为库进行加权 HOG 特征的提取和降维处理来表征人体威胁行为。对建立的特征库使用最小二乘支持向量机进行训练, 并使用一部分新的样本进行分类验证。实验结果证明: 该特征使用最小二乘支持向量机进行分类能够达到非常好的效果, 能够对典型威胁行为进行有效的分类。

根据对威胁行为的理解, 笔者建议从以下方面继续探讨: 1) 进一步扩展威胁行为库, 增加训练样本, 增加实用性; 2) 解决该特征使用支持向量机分类的拒分问题; 3) 解决遮挡、粘连人体威胁行为的识别问题。

参考文献:

- [1] Li Yonghao, Zhang Xihong, Kang Jian. Threat Posture Detection Based on Weighted HOG of Human-body Infrared Image[C]. singapore: 2011 3rd international conference on computer and network Technology, 2011.
- [2] Jürgen Metzler, Dieter Willersinn. Towards a sensor based threat warning system for patrols in MOUT scenarios[J]. Automatic Target Recognition. 2007, 22(12): 1103-1109.
- [3] Bobick A F, Davis J W. The recognition of human movement using temporal templates[J]. IEEE Trans. PAMI, 2001, 15(3): 257-267.
- [4] Mun wai Lee. A Model-Based Approach for Estimating Human 3D poses in Static Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(6): 905-916.
- [5] Navneet Dalal, Bill Triggs. Histograms of Oriented Gradients for Human Detection[C]. San Diego, CA: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [6] , 李 . 用 特征识别人的 行为[J]. , 2008, 19(12): 1686-1689.
- [7] , . 用 特征分 和识别人的 [J]. 机 用, 2008, 44(32): 158-161.