

# 基于自动编码器特征融合的图像行为识别算法

宋旭, 刘国英

(安阳师范学院 计算机与信息工程学院, 河南 安阳 455000)

**摘要:**为解决动作识别算法易受到光照、遮挡等环境变化影响,使算法的识别准确率与鲁棒性不佳等问题,提出一种基于递归神经网络与自动编码器特征融合的行为识别方法。分别通过递归神经网络(recurrent neural network, RNN)特征与自动编码器(auto encoder, AE)特征两种方法对行为动作进行特征提取,得到AE特征和RNN特征;引入特征相似度,对获得的AE特征和RNN特征进行融合,形成一种更加全面准确的AE-RNN特征;根据得到的AE-RNN特征,利用条件随机场分类学习,完成行为动作判断。实验结果表明,相比当前行为识别算法,所提算法具有更高的准确识别率,能够适应复杂背景和行为动作变化。

**关键词:**行为识别;自动编码器;特征相似度;递归神经网络;条件随机场

中图分类号: TP391 文献标识码: A 文章编号: 1000-7024(2019)05-1477-07

doi: 10.16208/j.issn1000-7024.2019.05.050

## Image behavior recognition research based on auto encoder feature fusion

SONG Xu, LIU Guo-ying

(School of Computer and Information Engineering, Anyang Normal University, Anyang 455000, China)

**Abstract:** The traditional motion recognition algorithm is easy to be influenced by the environmental changes such as illumination and occlusion, and the algorithm has low recognition rate and robustness. To improve the recognition rate and robustness, a method of behavior recognition based on feature fusion of recurrent neural network and automatic coding machine was proposed. Two kinds of methods were used to extract the behavior of the action using the recurrent neural network and the characteristics of the automatic coding machine, AE features and RNN features were obtained. Feature similarity was introduced, the AE features and RNN features were fused to form a more comprehensive and accurate AE-RNN feature. According to the characteristics of AE-RNN, conditional random fields classify learning was utilized to finish behavior judgment of behavior. Experimental results show that compared with the current commonly used behavior recognition algorithm, the proposed algorithm has higher accurate recognition rate, which can adapt to complex background and behavioral changes.

**Key words:** behavior recognition; auto encoder; feature similarity; recurrent neuron networks; conditional random fields

## 0 引言

多数视频数据的对象主要是人的活动,在智能监控、人机交互以及智能家居等领域广泛应用,人体行为识别是计算机视觉中重要研究方向<sup>[1,2]</sup>。随着视频拍摄设备不断发展,视频数据呈爆炸式增长,如何快速、准确与自动识别人体动作已成为当前的研究热点。由于行为动作具有复杂多变性,并且受各种环境因素的干扰,使得目前行为识别算法准确度不高,提取的行为特征存在冗余性和准确性不

够的问题<sup>[3,4]</sup>。

近年来,诸多学者提出了一系列的识别算法,如文献[5]通过时空兴趣点轨迹,通过KLT跟踪器对时空局部特征跟踪,并将轨迹当作处理单元,因此可较好描述动作变化。通过对于简单的背景具有一定的效果,但是在背景复杂时,提取的时空兴趣点特征有限,易导致识别率降低。文献[6]设计了一种姿势序列有限状态机动作识别方案,通过关节运动正则函数描述肢体动作轨迹,建立姿势序列有限状态机,完成动作识别。文献[7]提出了一种特征融

收稿日期: 2017-11-08; 修订日期: 2018-11-21

基金项目: 国家自然科学基金项目(41001251); 河南省高等学校重点科研项目计划基金项目(16B520001); 计算机科学与技术国家级特色专业建设基金项目(TS11576); 河南省科技计划基金项目(132102210212); 安阳师范学院科研培育基金项目(AYNU-KP-B08)

作者简介: 宋旭(1980-),男,河南安阳人,硕士,讲师,研究方向为图像处理、模式识别、信息安全; 刘国英(1979-),男,河南郑州人,博士,教授,研究方向为图像处理、模式识别。E-mail: Songxu1980cmt@126.com

合的行为识别算法，实验结果验证了其算法的合理性与优异性。该识别算法能较好对行为动作完成识别判断。

近年来，人体行为研究快速发展，计算机视觉受到了更严格的挑战。传统的简单场景的动作识别无法满足当代需求，需要面对真实场景下的复杂动作识别，因此，对行为动作识别算法提出了更高标准。

因此，为了解决提到的行为识别算法的精度和稳定性，基于自动编码器，设计了一种行为识别方案。通过获取图像的 AE 与 RNN 特征，提高对动作特征的描述能力，充分结合各自的互补信息，本文引入了特征相似度进行融合，获得的 AE-RNN 特征。再利用 AE-RNN 特征，通过分类训练，输出识别结果。最后在常用的数据集中对算法实施验证。

### 1 自动编码器特征提取

自动编码器 (auto encoder, AE) 是一种非监督学习算法，其无需对样本标记<sup>[8,9]</sup>。AE 主要通过输入数据、输出数据和中间数据构成一个网络。在训练时，每个样本，通过网络时在输出一个新的信息，AE 的作用就是保证使输出信息和输入信息一致。其结构如图 1 所示， $x$  表示输入数据， $y$  为中间数据， $z$  为输出数据，这些值组成了数据的特征。在整个网络中，为了得到输出  $z$ ，只需要利用到  $y$  数据而无需  $x$  数据，所以， $z$  只和数据  $y$  有关。因此，可通过  $z$  中输出数据来衡量  $y$  中是否包括了足够的输入数据  $x$ 。AE 可分为两部分：①输入数据和中间数据，可通过网络这信息压缩；②中间数据与输出数据，可将压缩信息恢复。

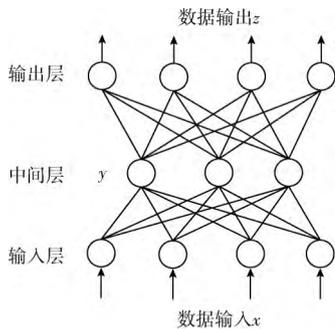


图 1 自动编码器结构

自动编码器的主要原理表示为：设 AE 中有  $d$  个输入和输出，中间层  $y$  的数量为  $h$ 。引入  $sigmoid$  方程对 AE 中每个单元激活，定义为

$$f(x) = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

在学习过程中，AE 先将输入数据  $x$  映射到中间层，获得中间层特征  $y$ ，称之为编码。接下来，中间层  $y$  被解码为输出特征  $z$ 。上述的描述的编码与解码可表示如下

$$y = \text{sigmoid}(W_y x + b_y) \quad (2)$$

$$z = \text{sigmoid}(W_z y + b_z) \quad (3)$$

式中： $W_y, W_z$  分别为输入-中间，中间-输出的联络矩阵。 $b_y, b_z$  为中间层，输出层每个单元的偏置。为了降低样本训练的参数，进行如下约束优化

$$W_y = W_z = W \quad (4)$$

通过式 (4) 中的约束，在训练时只需确定  $W, b_y, b_z$  这 3 个数值。

为了获得目标的最优化，需要通过  $W, b_y, b_z$  这 3 个参数构建输入与输出差异的最小值。因此，目标函数  $f$  定义为

$$f = \arg \min_{W, b_y, b_z} [\text{cost}(x, z)] \quad (5)$$

在给定的  $x$  下， $z$  的值与  $W, b_y, b_z$  密切相关， $\text{cost}(x, z)$  表示输入与输出差异。为了提高算法的计算速度，通过小批量数值对权重更新，表示如下

$$\text{cost} = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^d [x_{ik} \log(z_{ik}) + (1 - x_{ik}) \log(1 - z_{ik})] \quad (6)$$

其中， $d$  为输入值， $m$  为小批量值， $x_{ik}$  和  $z_{ik}$  分别为小批量中第  $i$  个输入、输出的第  $k$  个值。

目标函数与优化条件确定之后，引入随机梯度下降<sup>[10]</sup>进行求解， $W, b_y, b_z$  的优化函数表示如下

$$\begin{cases} W = W - \eta \frac{\partial \text{cost}(x, z)}{\partial W} \\ b_y = b_y - \eta \frac{\partial \text{cost}(x, z)}{\partial b_y} \\ b_z = b_z - \eta \frac{\partial \text{cost}(x, z)}{\partial b_z} \end{cases} \quad (7)$$

其中， $\eta$  为学习率。有了式 (7) 后，需要利用对罚方程对  $W, b_y, b_z$  求解偏导数。因此，式 (1) 的一阶和二阶方程表示为

$$f'(x) = f(x)[1 - f(x)] \quad (8)$$

$$f''(x) = f(x)[1 - f(x)][1 - 2f(x)] \quad (9)$$

AE 执行操作后，消除中间值和输出值，并将  $y$  当作下一个操作的输入。AE 在执行过程中，只需要根据特征  $y$ ，而  $y$  是在 AE 通过输入获得的。表明了特征  $y$  涵盖了输入数据的有效特征。对此，通过权重和偏置设计的非线性变换的 AE 是一种良好的特征提取方法。能够有效降低算法维度和降噪。

### 2 递归神经网络特征提取

递归神经网络 (recurrent neuron networks, RNN) 是表示节点此时的输入与之前时间的输出联系，有前向传播和反向传播<sup>[11,12]</sup>。前向传播 (forward propagation, FP) 是在一定的时间顺序利用激励函数求解获得，在每次执行计算时需加上时间。RNN 具有良好的动态性和存储性，能够较好对非线性函数逼近，对复杂特征之间的表示具有良好效果。RNN 的结构表示如图 2 所示。

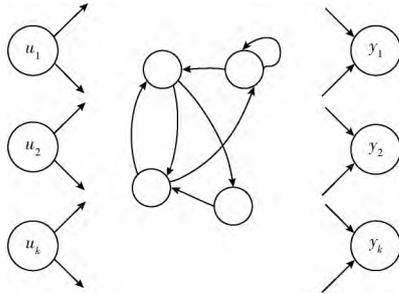


图 2 RNN 结构

从图 2 中看出, RNN 非单纯的从输入到输出操作, 这输入输出过程中构成了环运算, 在环运算中存在不定向操纵, 从中间层的输出成为下一时间的输入。为了更好地理解 RNN 结构, 进行一个简单案例说明, 将 RNN 的操作过程设计成一个计时器结构,  $u_1$  和  $u_2$  为输入量,  $y$  为输出量, 得到的结果如图 3 所示。从图 3 中看出,  $u_1$  等同于二进制开关, 在 RNN 开始的时候最大值为 1。  $u_2$  为离散变量, 在  $[0.1, 1]$  范围内变动, 表示开始的瞬间, 输出对应的长

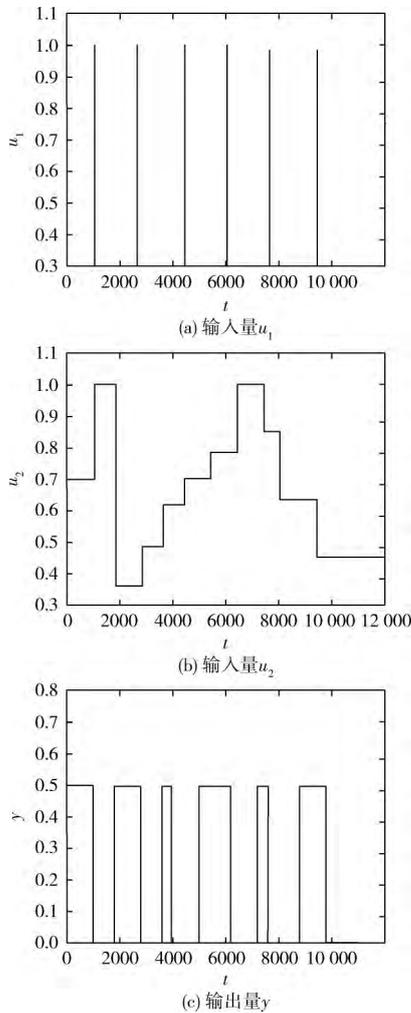


图 3 RNN 模拟计时器

度。在 RNN 中, 输出量  $y$  将在 0 到 0.5 范围内变换, 如图 3 (c) 所示。

在 RNN 中, 前向传播的中间层输出与下一个输入密切相关, 迭代时与时间相关, 设  $a'_h$  为  $t$  时的中间层输入量, 表示为

$$a'_h = \sum_{i=1}^I \omega_{ih} x'_i + \sum_{h'}^H \omega_{h'h} b_{h'}^{t-1} \quad (10)$$

其中,  $\omega_{ih}$  表示输入与中间层的权重,  $x'_i$  为时间  $t$  的输入量,  $\omega_{h'h}$  为中间层与上一个时间的中间层的权重,  $b_{h'}^{t-1}$  为  $t-1$  时的输出。

时间  $t$  中间层作为输入值可表示为

$$b_h = \theta_h(a'_h) \quad (11)$$

其中,  $\theta_h$  为激励函数。对于时间  $t$ , 输出层的输入值表示为

$$a'_o = \sum_{h=1}^H \omega_{oh} b'_h \quad (12)$$

其中,  $\omega_{oh}$  为中间层与输出层的权重。那么,  $t$  时的输出值可表示为

$$b'_o = \theta_o(a'_o) \quad (13)$$

根据 RNN 前向传播, RNN 的后向传播中, 令误差为  $L$ , 那么  $t$  时的误差表示为

$$\delta'_h = \theta'(a'_h) \left( \sum_{o=1}^O \delta'_o \omega_{oh} + \sum_{h'=1}^H \delta_{h'}^{t+1} \omega_{hh'} \right) \quad (14)$$

对于误差的权重偏导可表示为

$$\nabla L(\omega_{ij}) = \frac{\partial L}{\partial \omega_{ij}} = \sum_{t=1}^T \frac{\partial L}{\partial a'_j} \frac{\partial a'_j}{\partial \omega_{ij}} = \sum_{t=1}^T \delta'_j b'_i \quad (15)$$

其中,  $\delta'_j = \frac{\partial L}{\partial a'_j}$ ,  $\omega_{ij}$  为输入与输出权重。根据以上计算,

可得到权重的更新表达式

$$\omega_{ij} = \omega'_{ij} - \eta \nabla L(\omega_{ij}) \quad (16)$$

在每个时刻, 中间层根据外部链接表达当前量, 利用输入和输出层的值, 通过加权求和表达新的量。由于前一个时刻的值被用于表达后面时间的量, 因此, RNN 可保持之前时刻的记忆, 为后续判断提供帮助。在经过多次训练之后, 通过 RNN 对原始信号进行提取表达, 结果如图 4 所示, 从图 4 中看出, RNN 得到的其结果与原始信号较相似, 能够有效的对原始信息表达。

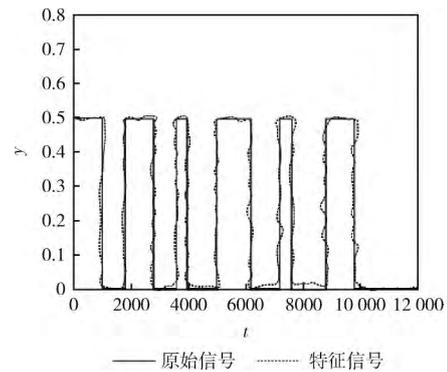


图 4 RNN 特征提取测试

### 3 本文动作识别算法

为了提高行为算法的识别率，需要在判断决策时提供更多的依据，因此，为了避免单一特征对行为动作描述力不强，提出了通过 2 种常见的特征提取，并利用一定的融合方法形成新的特征表示。因此，提出了一种自动编码器与递归神经网络特征融合的行为识别方法。本文提出的算法框架如图 5 所示。首先，分别通过自动编码器特征与递归神经网络特征对行为动作进行特征提取，得到 AE 特征和 RNN 特征；其次，引入特征相似度进行融合，将 AE 特征和 RNN 特征进融合形成一种新的 AE-RNN 特征；然后，利用得到的 AE-RNN 特征，通过条件随机场分类学习，完成行为决策判断。

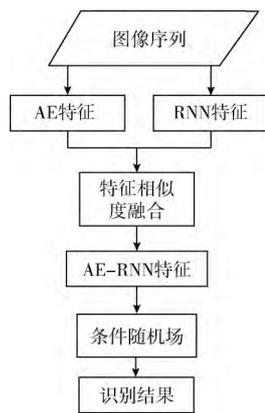


图 5 本文图像动作识别算法过程

#### 3.1 基于特征相似度的特征融合

不同特征具有不同的特征量，不同特征融合为一个特征时，特征维数是融合之前的总和。空间中每个维度具有相应的特征量，这种融合特征的内存储存巨大，运算效率低下。为了获得更准确的特征，将特征融合的参数进行优化。本文特征融合思想是基于特征相似度进行，一般来说，特征的空间位置表达了类型关系；相同的特征空间通过同类型为圆心，在一定半径范围内随机排列。特征类型不同，其相应的空间关系存在一定差异，通过特征向量和不同类型的特征的空间距离表示特征相似度，特征向量和同属特征模型相似度越高。

设训练样本的空间点作为中心，其对空间中的其它点产生引力  $F$ ，距离越近， $F$  越大。每个类型的吸引力只在特定的大小内有效。对于  $n$  个特征中，中心点  $x_i$  与特征  $x_j$  的吸引力  $F_{ij}$  表示为<sup>[13]</sup>

$$F_{ij} = e^{-d_{ij}^2/2\sigma^2} \tag{17}$$

式中： $\sigma$  为中心点的控制因子； $d_{ij}^2$  为特征之间的欧式距离， $d_{ij}^2 = x_i - x_j$ ；当  $x_i = x_j$ ， $d_{ij}^2 = 0$ ，特征  $x_i$  与  $x_j$  重叠， $F_{ij}$  值最大。 $d_{ij}^2$  越大， $F_{ij}$  越小。通过  $F_{ij}$  与  $d_{ij}^2$  的表达

式，当  $d_{ij}^2$  较大时， $F_{ij}$  会迅速下降为 0。中心点对其它特征产生吸引力，中心点相互也会具有吸引力，中心点的引力构建了一个相似向量  $F$ ，对于  $n$  维空间的  $m$  个特征中  $F$  表示如下

$$F = \begin{bmatrix} F_{11} & F_{12} & \dots & F_{1m} \\ F_{21} & F_{22} & \dots & F_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ F_{n1} & F_{n2} & \dots & F_{nm} \end{bmatrix} \tag{18}$$

对于某个特征  $y$ ，对  $y$  和不同类型的特征之间的吸引力，通过不同吸引力建立了样本的特征相似度  $f$ ，从而表示了融合特征。图 6 为提出的特征融合结果。图 6 (a) 为样本图像序列，图 6 (b) 为提取的走路动作 AE 特征，图 6 (c) 为提取的走路动作 RNN 特征，图 6 (d) 为通过提出的融合技术将 AE 特征与 RNN 融合得到的 AE-RNN 特征。通过提出的融合技术，很好的将 2 种不同方式得到的特征进行融合，较完整、准确地反映了动作形态。

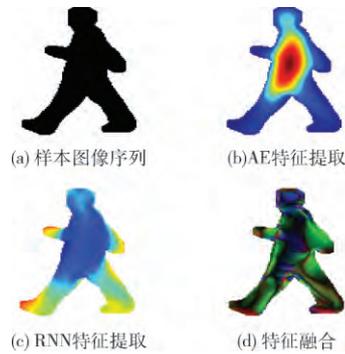


图 6 特征融合结果

#### 3.2 条件随机场分类学习

条件随机场分类 (conditional random field, CRF) 是在一定的输入变量中，求解输出随机变量的条件概率分布<sup>[14,15]</sup>。CRF 在条件概率函数  $P(Y|X)$  中， $X$  是输入值，为需要分类学习的观察序列， $Y$  是输出值，为标记序列。

设  $X$  和  $Y$  为随机值， $P(Y|X)$  为在一定  $X$  下得到  $Y$  的条件概率分布， $Y$  形成了由无向图  $G = (V, E)$  的 Markov 随机场，表示为

$$P(Y_v | X, Y_w, w \neq v) = P(Y_v | X, Y_w, w \sim v) \tag{19}$$

其中， $w \sim v$  为在  $G = (V, E)$  中与节点  $v$  有关联的点  $w$ ； $w \neq v$  为与节点  $v$  无关联的点； $Y_v, Y_w$  表示节点  $v, w$  对应的值。

在实际过程中，为简化模型，可令  $X$  与  $Y$  具有相同的结构，那么  $G = (V, E)$  成为一个线性链，表示为

$$G = (V = \{1, 2, \dots, n\}, E = \{(i, i + 1)\}), \tag{20}$$

$$i = 1, 2, \dots, n - 1$$

对此， $X = (X_1, X_2, \dots, X_n)$ ， $Y = (Y_1, Y_2, \dots, Y_n)$  表示为相邻 2 节点的集，如图 7 所示。

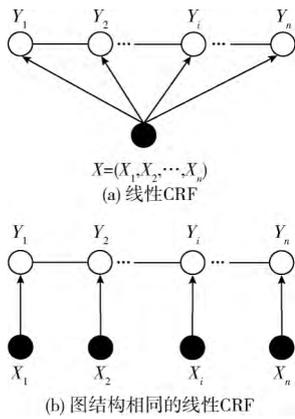


图 7 两种不同的 CRF 结构

根据图 7 中看出, 对于线性链 CRF,  $X = (X_1, X_2, \dots, X_n)$ ,  $Y = (Y_1, Y_2, \dots, Y_n)$  表线性链随机变量。在一定的  $X = (X_1, X_2, \dots, X_n)$  下,  $Y$  的条件概率分布  $P(Y|X)$  符合 Markov 条件, 组成了一个 CRF, 表示如下

$$P(Y_i | X, Y_1, \dots, Y_{i-1}, Y_{i+1}, \dots, Y_n) = P(Y_i | X, Y_{i-1}, Y_{i+1}) \quad (21)$$

#### 4 实验测试

在 KTH 测试集<sup>[16]</sup>与 Weizmann 数据集<sup>[17]</sup>中实施识别性能验证。其中, Weizmann 测试集含有 10 个动作类别, 分别为: Bend, Jack, Jump, P-jump, Run, Side, Skip, Walk, Wave1, Wave2。Weizmann 测试集中的分辨率为  $180 \times 144$ , 25 Frame/s。图 8 为 KTH 数据集。测试条件为: Core I3, 3.50 GHz CPU, 4 GB 运行 RAM, Win7 操作系统。为了突出所提方案的优势, 将文献 [5] 算法、文献 [6] 算法和文献 [7] 算法视为对比组, 分别记为 A 算法、B 算法、C 算法。借助 10 次 3 重交叉法进行实验。



图 8 KTH 数据集

#### 4.1 评价指标

对于待查询的目标  $Q_i$ , 其相应的查准率  $P_k(Q_i)$  为

$$P_k(Q_i) = \frac{N_r}{N} \quad (22)$$

式中:  $N_r$  代表准确判别的数量;  $N$  是查询目标的总数。

同样, 对于目标  $Q_i$ , 其查全率  $R_k(Q_i)$  函数为

$$R_k(Q_i) = \frac{N_r}{N_c} \quad (23)$$

式中:  $N_c$  代表样本总数。

根据  $N_r$ 、 $N$  与  $N_c$ , 相应的识别率 AC 表示为

$$AC = \frac{N_r}{N + N_c} \quad (24)$$

#### 4.2 实验结果

不同识别方案在 KTH 和 Weizmann 数据集集中的平均识别率如图 9 所示。

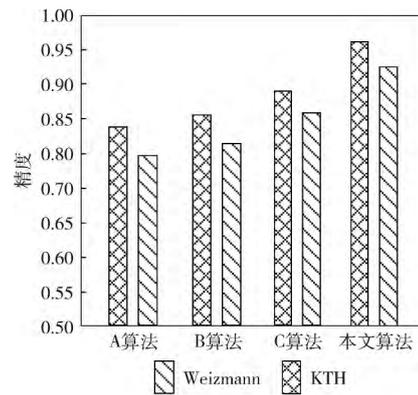


图 9 KTH 与 Weizmann 数据集识别率测试

由图 9 发现, 对于不同的两个数据库, 所提方案的准确识别精度最为理想, 均高于 90%。

另外, 统计 KTH 数据集集中的所提方案对应的混淆矩阵, 见表 1。表中第  $i$  行第  $j$  为第  $i$  行识别为第  $j$  种动作的可能。同行的概率和为 1, 表 1 为 KTH 测试集中识别率的混淆矩阵。从表中看出, Walk 与 Run 的动作正确率相对较低, Walk 容易被误认为 Run 与 Punch; Run 主要容易被误识别为 Walk。主要是因为这 3 种动作相似动作较多, 较容易产生误判。而 HC 识别率最高, 可达 100%。

表 1 KTH 数据集的混淆矩阵

	本文方法					
	Walk	Jog	Run	Punch	Wave	HC
Walk	0.95	0.01	0.04			
Jog	0.02	0.96	0.03			
Run	0.04	0.01	0.94	0.01		
Punch		0.02		0.97	0.02	
Wave			0.01	0.03	0.98	
HC						1

表 2 为 Weizmann 测试集中得到的混淆矩阵。从表 2 中得出, Run 和 Side 的正确率相对较低, Run 容易被误判为

Side 与 Jump, Side 易被误判为 Jump。主要是这 3 种动作具有较高的相似性, 而 Jump 与 Wavel 的识别相对最高。

表 2 Weizmann 数据集的混淆矩阵

	Bend	Jack	Jump	P-jump	Run	Side	Skip	Walk	Wavel	Wave2
Bend	0.95				0.01			0.04		
Jack		0.94	0.01				0.02			0.03
Jump			1.00							
P-jump		0.03		0.95				0.02		
Run		0.01	0.04		0.89	0.05			0.01	
Side		0.01	0.04		0.02	0.92		0.01		
Skip			0.01			0.02	0.96		0.01	
Walk		0.02		0.03				0.93		0.02
Wavel									1.00	
Wave2			0.03						0.02	0.95

由表 3 发现, 对于 Weizmann 数据库中的复杂行为, 所提技术的准确识别精度维持在 86% 以上。特别是 Walk、Wavel、Wave2 这 4 种行为, 其精度约为 95%。但是, 其它 3 种识别方案不稳定, 且精度要低于所提技术。原因是

所提方案的混合特征能够更好的表达行为动作特征, 具有更加全面准确的判断依据, 并且通过 CRF 对样本进行分类学习, CRF 对全部特征全局归一化, 能够有效获得全局最优解, 从而提高了识别率。

表 3 Weizmann 数据库中识别率对比/%

	Bend	Jack	Jump	P-jump	Run	Side	Skip	Walk	Wavel	Wave2
A 算法	78	77	82	78	83	76	82	82	76	82
B 算法	85	83	82	84	85	83	87	85	86	86
C 算法	87	87	88	87	89	85	88	86	87	84
本文算法	93	92	96	88	86	88	90	95	98	96

图 10 为在 Weizmann 数据集中, 所提方案与 3 种对照组技术的查准-查全曲线。根据该数据发现, 所提方案的呈现出更理性的曲线, 对于任意的  $R_k(Q_i)$ , 其  $P_k(Q_i)$  均为最大, 说明了本文算法性能相对其它算法更优。

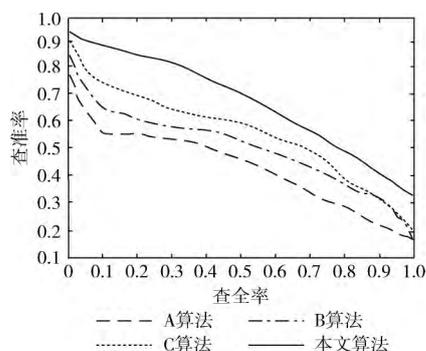


图 10 不同算法的查准-查全曲线

通过在 KTH 和 Weizmann 测试集实验得出, 在 KTH 上得到的识别率高于 Weizmann 数据集, 主要是在 KTH 的

动作类型比 Weizmann 简单, 动作种类简单, 而 Weizmann 中的动作复杂, 相互之间会有干扰, 从而降低了算法的识别率。本文算法在能够有效对行为动作识别, 相对其它算法在识别率上有较大提高。通过定量分析指标也可看出, 本文得到的查全率、查准率和正确识别率有较大提高。原因是所提方案对两种特征相似度进行融合, 形成一个更精确更全面的特征描述, 通过条件随机场分类决策。而其它算法一般只通过某一个的特征对图像序列描述, 得到的特征存在片面性和不完整性, 无法完整、准确地表征图像序列的信息, 从而导致了某些行为动作无法正确识别。

### 5 结束语

为了提高在行为识别中特征的描述能力, 本文提出了利用递归神经网络特征与自动编码器特征两种特征对视频序列进行特征提取。过权重和偏置设计的非线性变换的 AE 是一种良好的特征提取方法, 能够有效降低算法维度和降噪。RNN 特征能够有效表示在复杂背景下的行为特征, 并

通过特征相似度将 AE 与 RNN 特征融合, 形成一个更具特征描述的新特征, 为后续的条件随机场分类决策提供更多更全面的判断依据。通过在常用的 KTH 和 Weizmann 数据集上的实验验证了本文算法具有较高的行为识别率, 鲁棒性较强。

#### 参考文献:

- [1] CHEN Wenqiang, XIAO Guoqiang, LIN Xiao. On a human behaviors classification model based on attribute-Bayesian network [J]. Journal of Southwest China Normal University (Natural Science Edition), 2014, 39 (3): 7-11 (in Chinese). [陈文强, 肖国强, 林霄. 一种基于属性贝叶斯网络的行为识别模型 [J]. 西南师范大学学报 (自然科学版), 2014, 39 (3): 7-11.]
- [2] DONG Ke, GAN Zhaohui, JIANG Min. Behavior recognition based on motion capturing data [J]. Computer Engineering and Design, 2016, 37 (3): 762-768 (in Chinese). [董珂, 甘朝晖, 蒋旻. 基于动作捕获数据的行为识别 [J]. 计算机工程与设计, 2016, 37 (3): 762-768.]
- [3] Feng Z, Yang B, Li Y. Realtime oriented behavior driven 3D freehand tracking for direct interaction [J]. Pattern Recognition, 2013, 46 (2): 590-608.
- [4] Wu D, Shao L. Silhouette analysis-based action recognition via exploiting human poses [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23 (2): 236-243.
- [5] Derpanis K G, Sizintsev M, Cannons K J. Action spotting and recognition based on a spatiotemporal orientation analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35 (3): 527-510.
- [6] LIN Shuiqiang, WU Yadong, YU Fang. Posture sequence finite-state machine method for motion recognition [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26 (9): 1403-1411 (in Chinese). [林水强, 吴亚东, 余芳. 姿势序列有限状态机动作识别方法 [J]. 计算机辅助设计与图形学学报, 2014, 26 (9): 1403-1411.]
- [7] SHAO Yanhua, GUO Yongcai, GAO Chao. Human action recognition using multi-feature fusion [J]. Journal of Optoelectronics • Laser, 2014, 25 (9): 1818-1823 (in Chinese). [邵延华, 郭永彩, 高潮. 基于特征融合的人体行为识别 [J]. 光电子 • 激光, 2014, 25 (9): 1818-1823.]
- [8] SHEN Cheng'en, HE Jun, DENG Yang. Pam filtering based on modified stack auto-encoder [J]. Journal of Computer Applications, 2016, 36 (1): 158-163 (in Chinese). [沈承恩, 何军, 邓扬. 基于改进堆叠自动编码器的垃圾邮件分类 [J]. 计算机应用, 2016, 36 (1): 158-163.]
- [9] ZHAO Jun, ZHAO Yan, YANG Yong. Facial expression recognition auto-encoders method based on stacked denoising and feature reduction [J]. Journal of Chongqing University of Posts and Telecommunications, 2016, 28 (6): 844-848 (in Chinese). [赵军, 赵艳, 杨勇. 基于降维的堆积降噪自动编码器的表情识别方法 [J]. 重庆邮电大学学报, 2016, 28 (6): 844-848.]
- [10] CHEN Zhenhong, LAN Yanyan, GUO Jiafeng. Distributed stochastic gradient descent with discriminative aggregating [J]. Chinese Journal of Computers, 2015, 38 (10): 2054-2063 (in Chinese). [陈振宏, 兰艳艳, 郭嘉丰. 基于差异合并的分布式随机梯度下降算法 [J]. 计算机学报, 2015, 38 (10): 2054-2063.]
- [11] CAI Ruichu, XIE Weihao, HAO Zhifeng. Abnormal crowd detection based on multi-scale recurrent neural network [J]. Journal of Software, 2015, 26 (11): 2884-2896 (in Chinese). [蔡瑞初, 谢伟浩, 郝志峰. 基于多尺度时间递归神经网络的人群异常检测 [J]. 软件学报, 2015, 26 (11): 2884-2896.]
- [12] Vataya Chunwijitra, Ananlada Chotimongkol. A hybrid input-type recurrent neural network for LVCSR language modeling [J]. EURASIP Journal on Audio, Speech, and Music Processing, 2016, 16 (1): 1-12.
- [13] WANG Chaoyun, JIANG Gangyi, YU Mei. Manifold feature similarity based perceptual image quality assessment [J]. Acta Automatica Sinica, 2016, 42 (7): 1113-1124 (in Chinese). [王朝云, 蒋刚毅, 郁梅. 基于流形特征相似度的感知图像质量评价 [J]. 自动化学报, 2016, 42 (7): 1113-1124.]
- [14] QIAN Sheng, CHEN Zonghai, LIN Mingqiang. Saliency detection based on conditional random field and image segmentation [J]. Acta Automatica Sinica, 2015, 41 (4): 711-724 (in Chinese). [钱生, 陈宗海, 林名强. 基于条件随机场和图像分割的显著性检测 [J]. 自动化学报, 2015, 41 (4): 711-724.]
- [15] Zhang CY, Hong XG, Peng ZH. Extracting Web entity activities based on SVM and extended conditional random fields [J]. Journal of Software, 2014, 23 (10): 2612-2627.
- [16] FU Ying, GUO Jingyun. Dynamic time warping-based human action recognition [J]. Electronic Measurement Technology, 2014, 37 (3): 69-72 (in Chinese). [傅颖, 郭晶云. 基于动态时间规整的人体动作识别方法 [J]. 电子测量技术, 2014, 37 (3): 69-72.]
- [17] ZHANU Feiyan, LI Junfeng, SHEN Junmin. Human behavior recognition based on the statistical characteristics of gradient and optical flow [J]. Journal of Optoelectronics • Laser, 2015, 26 (8): 1593-1601 (in Chinese). [张飞燕, 李俊峰, 沈军民. 基于梯度和光流统计特性的人体行为识别 [J]. 光电子 • 激光, 2015, 26 (8): 1593-1601.]