



人体动作数据编码与 CNN 精确识别

胡青松^{1*}, 张亮¹, 丁娟², 李世银¹

(1. 中国矿业大学信息与控制工程学院 江苏 徐州 221116; 2. 吉林大学珠海学院 广东 珠海 519041)

【摘要】人体动作的精确识别面临多方面的挑战,特别是动作采集易受光照强度的影响、动作特征描述不清楚和易物理变形。为了降低这些不利因素的影响,提高动作的识别精确度,该文从3个步骤展开研究:首先,对 Kinect 提取的人体关节数据进行预处理,从而克服光照问题;随后,使用针对性编码方法对人体动作数据进行编码,进而利用卷积神经网络(convolutional neural network, CNN)自动提取人体动作特征,解决动作特征描述的难题;最后,使用 SoftMax 完成复杂动作的识别。实验表明,该文算法具有较高的识别准确率和泛化能力,其 F_1 值普遍在 0.8 以上;在单一属性测试中,复合属性数据比被复合的单一属性数据更有优势, F_1 值可达 0.916;混合属性测试的 F_1 值相比单一属性测试有所下降,下降幅度最高可达约 25%。

关键词 卷积神经网络; 特征提取; 人体动作识别; 动作数据编码

中图分类号 TP391.1 **文献标志码** A **doi**:10.12178/1001-0548.2019108

Data Encoding and CNN Accurate Recognition of Human Body Motion

HU Qing-song^{1*}, ZHANG Liang¹, DING Juan², and LI Shi-yin¹

(1. School of Information and Control Engineering, China University of Mining & Technology Xuzhou Jiangsu 221116;

2. Zhuhai College, Jilin University Zhuhai Guangdong 519041)

Abstract It is affected by many negative factors to recognize human motions accurately, such as the affecting of the light intensity during motion data collecting, the vagueness and transformation of the human motion features. To decrease the effect of these adverse factors and improve the accuracy of motion recognition, the following contents are investigated in this paper. Firstly, the human joint data collected by Kinect are preprocessed to overcome the illumination problem. Secondly, encoding methods are proposed to encode the preprocessed data, and then the encoded data are inputted into CNN to extract human motion features automatically, making the description of motion feature easier. Finally, the CNN completes motion classification with SoftMax. These experiments show the proposed algorithm can achieve a high recognition accuracy (most of the F_1 values are larger than 0.8) and can adapt to different data settings; the compound property data are better than single property data in single property tests and the F_1 value can be 0.916; the F_1 values in compound property tests are smaller than those of single property tests and the maximal decrease percentage can be 25%.

Key words convolutional neural network; feature extraction; human motion recognition; motion data encoding

人体动作识别^[1]旨在通过在线或离线的方式,从传感器所采集的数据中自动识别出人体正在执行的动作,它是计算机视觉^[2]、机器学习^[3]、模式识别^[4]和人工智能^[5]等技术交叉融合发展的结果,在新型人机交互、虚拟现实、增强现实和辅助培训等领域具有广阔的应用前景。

人体动作识别算法主要有模板匹配和机器学习

两大类。模板匹配算法将动作实例与模板库中的动作进行对比,模板库中与动作实例相似度最高的动作即为识别结果;文献[6]提出的时间自相似和动态规整的识别算法即属此类。模板匹配识别算法的缺点是随着模板库的增加,模板比对的开销会越来越大。机器学习算法是用一系列动作实例训练一个分类器,此分类器能够区分不同动作的共性与差

收稿日期: 2019-04-22; 修回日期: 2019-11-26

基金项目: 国家自然科学基金(51874299, 61771474); 江苏省科技成果转化专项资金(BA2016016)

作者简介: 胡青松(1978-), 男, 副教授, 主要从事物联网、移动目标定位、救灾通信等方面的研究. E-mail: hqsong722@163.com

异,进而利用分类器进行动作分类;文献[7]提出的方法即属此类,它结合时空特征与3D-SIFT描述子提取人体特征,使用SVM(support vector machine)算法识别人体动作。与基于模板匹配的认识算法类似,基于机器学习的识别算法也要求手动提取动作特征,因此操作繁琐,难以通用。

与普通的机器学习不同,深度学习算法^[8-10]能够自动提取目标特征,其中CNN在深度学习中应用得尤为广泛^[11],在很多情况下都表现出优秀的性能,适合处理人体动作识别等复杂任务。为此,本文提出基于CNN的精确人体动作识别算法,对人体骨骼数据进行编码处理,进而构建深度学习框架,从而准确识别数据所包含的人体动作。

1 动作数据的预处理

本文采用可开放获取的MSRC-12(the Microsoft research Cambridge-12)人体动作数据集。该数据集由微软剑桥联合实验室使用Kinect设备采集^[12],共包含594个人体动作序列,719 359帧,播放完毕约需6小时40分钟。由于Kinect通过发射红外光获取人体骨骼和关节的位置坐标,而不是采用可见光获取图像或视频,因此克服了光照强度对人体动作识别的影响。

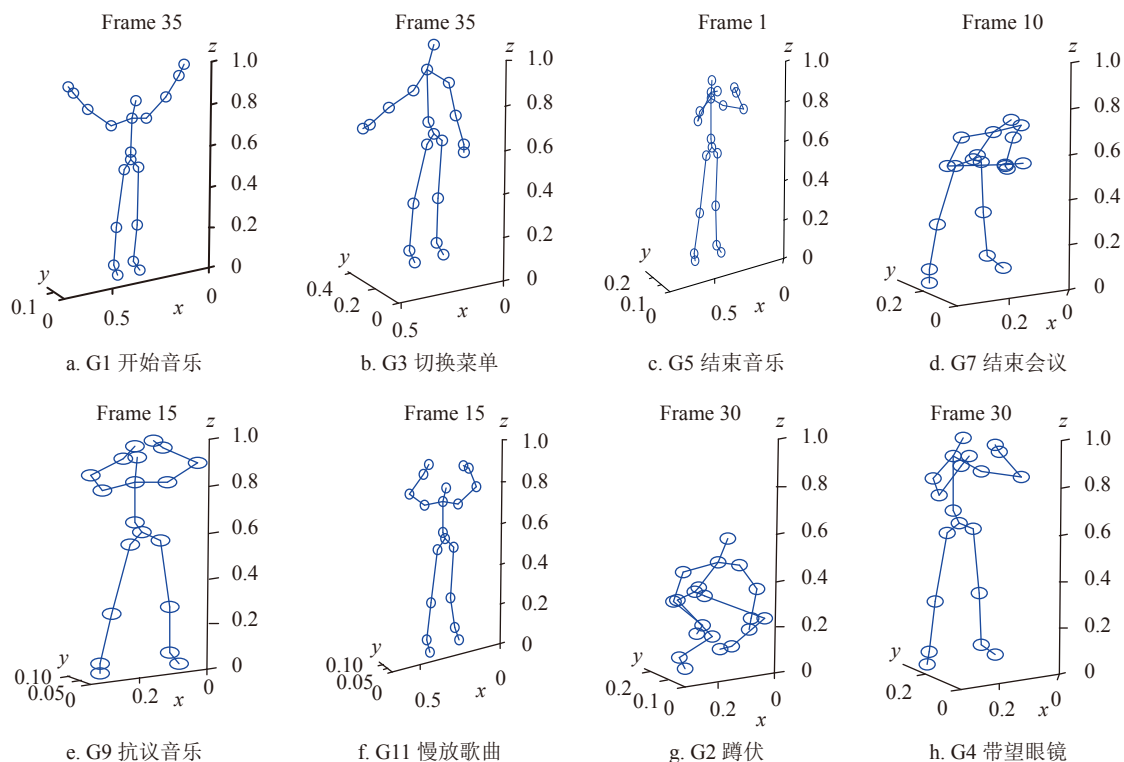
为了增强数据集的代表性和说服力,MSRC-12共采集了30名不同志愿者的动作。志愿者特征如

下:1)年龄特征:22~65岁,平均31岁;2)身高特征:1.52~2米,平均1.76米;3)男女比例:3:2;4)左、右撇子比例:1:14。

志愿者分别表演图1中的12种动作,共采集到6 244个动作实例。在这12种动作中,奇数动作是隐喻动作,意为把动作名称与隐含的寓意相关联;偶数动作是标准动作,动作名称直接描述了动作本身的含义。从图1可以看出,Kinect记录的是点和线组成的人体骨骼数据,其中的点代表人体关节,关节的不同位置组合便可表达不同的人体动作。

数据采集方式不同,测试者的动作也会受到一定影响。为了体现采集方式的影响,MSRC-12不但包括单一属性数据,也包括复合属性数据。单一属性数据包括:1)描述分解动作的文本,简称文本(Text);2)个体动作的有序静态图像集合,简称图像/Images);3)个体动作的视频,简称视频(Video)。复合属性数据包括:1)文本和图像的复合(Text+Images);2)文本和视频的复合(Text+Video)。

MSRC-12数据集中,每一个测试者的相同动作存储在一个.csv文件中,它记录了三维骨骼数据和该数据的记录时间。此外,每个动作还有一个.tagstream伴生文件,它记录了每一个动作的中心时间戳。把相同属性的数据追加到一个文件中,最终得到提取后的数据文件和相应的标签文件。



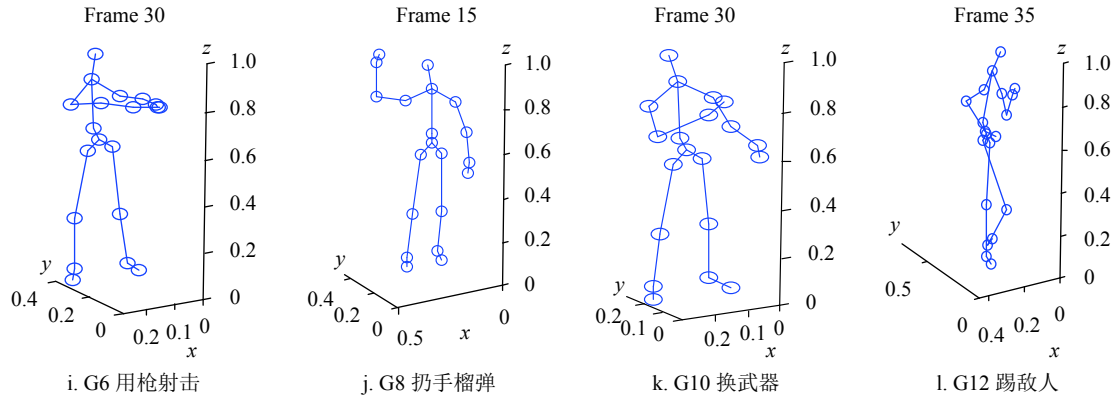


图 1 MSRC-12 数据集的动作实例

最后, 将提取的数据进行归一化处理。归一化的目的有两点: 一是保证所绘图形不失真; 二是便于后续的深度学习处理。因为归一化使得数据无量纲化, 从而加快了优化过程, 并能提高模型精度。本文把 x, y, z 轴分开处理, 用 M 表示 text_x.mat , M_x 、 M_y 和 M_z 分别表示 x, y, z 轴上的数据。归一化处理之前, 首先利用式 (1) 消除 x 轴的负值, 接着利用式 (2) 对 x 轴数据进行归一化。 y, z 轴数据采用同样方法处理, 不再赘述。

$$M_x = M_x - \min(M_x) \quad (1)$$

$$M_x = \frac{M_x - \min(M_x)}{\max(M_x) - \min(M_x)} \quad (2)$$

2 数据编码与动作识别

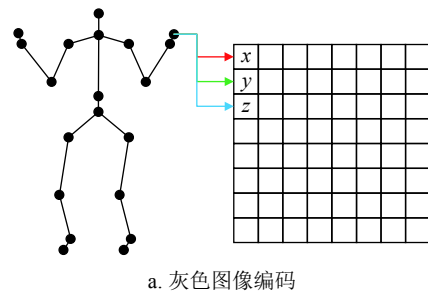
2.1 动作数据编码

本文提出的基于 CNN 的动作识别所处理的对象是图像, 因此在进行识别前需要对预处理完毕后的动作数据进行编码, 将代表动作的人体关节数据转化为图像形式。

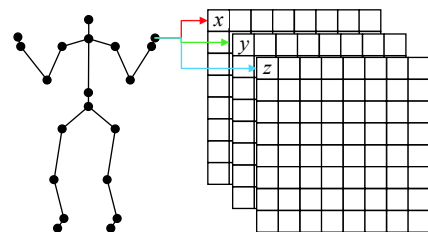
采用何种方式将动作数据转化成图像, 对 CNN 的性能具有较大影响。本文共提出 3 种编码方式, 每种编码方式均可构建灰度图像或彩色图像, 因此共有 6 种编码组合。灰度图像的构建方法如图 2a 所示, 选取一个关节点的数据, 将其 x, y, z 三轴分量依次放入同一个方阵中, 每个分量填充一个像素点。按照相同的方法将其他关节点的数据填入方阵, 所得到的方阵即是一个灰度图像方阵。彩色图像的构建方法与灰度图像类似, 所不同的是将关节点的三轴分量分别放入 3 个不同的方阵中, 如图 2b 所示。将所得的 3 个方阵作为彩色图像的 r, g, b 分量, 组成的图像即是一张彩色图像。

如图 3 所示, 3 种编码方法分别称为平铺编

码、回旋编码和“之”字形编码。平铺编码指的是将数据按照从左到右来回扫描进行平铺排列, 如图 3a; 回旋编码指的是将数据按照从外到内进行“回旋”排列, 如图 3b; “之”字形编码则是将数据排列成类似“之”字的编码方式, 如图 3c。

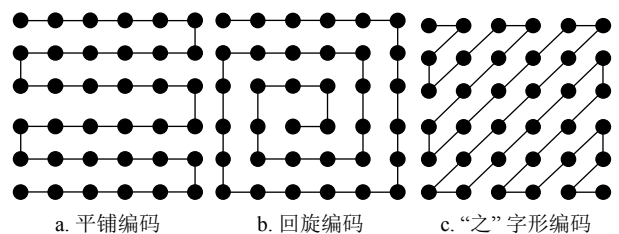


a. 灰色图像编码



b. 彩色图像编码

图 2 动作数据的灰度与彩色图像编码



a. 平铺编码 b. 回旋编码 c. “之”字形编码

图 3 动作数据的编码方法

MSRC-12 数据集的一个动作实例由连续的 35 帧构成。如果采用灰色图像编码, 容纳一帧动作的数据需要一个 46×46 的方阵, 不能填满的方阵元素用“0”填充。图 4 为同一个动作的 6 种编码图案, 第一行从左到右依次为采用平铺编码、回旋

编码、“之”字型编码得到的灰色图像；第二行从左到右分别为采用平铺编码、回旋编码、“之”字型编码得到的彩色图像。这些编码方式的动作识别能力将在第3节通过仿真的方式进行对比研究。

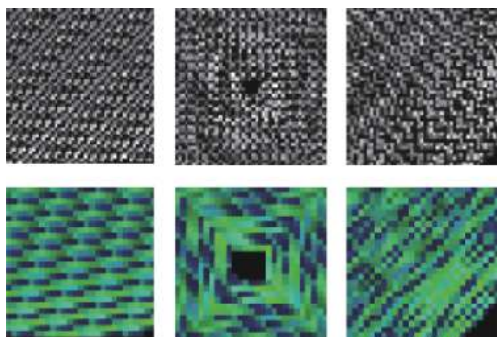


图4 动作数据的6种编码实例

2.2 基于CNN的人体动作识别

CNN是一种高效的深度学习方法，通常包括输入层、卷积层、池化层、全连接层和分类器，如图5所示。它通过局部卷积来提取图像区域特征，并通过共享卷积权值和池化操作来降低数据维度，最终再通过传统机器学习方法完成样本分类。

在一个基于CNN的学习实例中，可以根据需

要设置多个卷积层和池化层，在全连接层将得到的特征图片进行全连接，所得的特征向量经由分类器输出，即可得到识别结果。相比其他算法，图像从输入层输入之后，CNN能够自动提取图像特征，识别准确率很高。

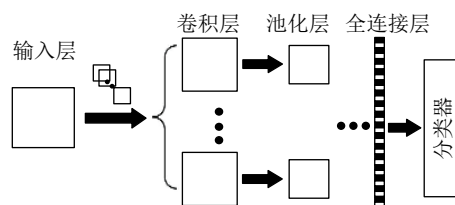


图5 CNN学习框架

为了提高识别能力，本文构造的CNN学习框架包括7层结构，如图6所示。下面以单通道灰色图像为例对各层予以说明：

第一层(输入层)：输入图像，其尺寸为 46×46 。

第二层(卷积层1)：设置6幅 3×3 的卷积核，并将卷积步长设定为1(若无特殊说明，下文的卷积步长都设定为1)，分别与输入图像卷积，得到6幅 44×44 的特征图。

第三层(池化层1)：对卷积层1得到的特征图进行均匀池化，得到6幅 22×22 的特征图。

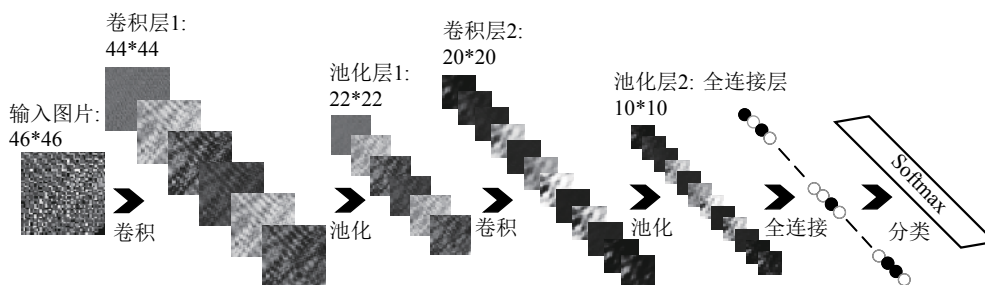


图6 7层结构的动作识别CNN框架

第四层(卷积层2)：其处理方式与卷积层1相同，但是卷积核不同，这里设置了12个 3×3 的卷积核，得到12幅 20×20 的特征图。

第五层(池化层2)：对卷积层2输出的特征图进行与池化层1相同的处理，得到 10×10 的特征图。

第六层(全连接层)：将池化层2输出的特征图展开成一维向量。

第七层(分类器)：采用深度学习中广泛使用的Softmax函数^[13]对人体动作进行分类，得到动作识别结果。

为了进行性能对比，彩色图像编码的人体动作数据采用相同的识别框架，每一层的卷积核数目和卷积步长等参数均保持不变。对CNN进行训练时，对网络参数利用梯度下降算法进行优化。学习

速率不同，训练结果具有较大差异，本文根据反复实验将学习速率确定为 $\alpha=0.3$ 。

3 识别性能分析

3.1 评价指标

评价深度学习网络模型的指标很多，其中准确率(accuracy)因其直观性而得到了广泛使用^[14]，它表示被正确分类的样本数据比例。与此相反，错误率则表示被错误分类的样本数据比例。显然，错误率=1-准确率，是网络对样本拟合程度的间接表示。由于样本数据中一般都存在噪声，因此拟合程度不宜设置得过高，否则将导致噪声也被拟合，最终导致网络过拟合。用 C 表示样本数量，可用式(3)计算准确率：

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{C} \quad (3)$$

式中, TP 即 true positive, 表示被分类为正确又确实为正确的样本数量; TN 即 true negative, 表示被分类为负又确实为负例的样本数量。

另外一个常用的评价指标是精确率 (precision), 表示被分类为正确又确实为正确的样本占所有正例的比例, 采用式 (4) 计算:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

式中, FP (false positive, 假正) 表示分类为正的负例。

与精确率相关的一个指标是召回率 (recall), 表示真正的样本被分类为正例的比例, 其计算式为:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

式中, FN (false negative, 假负) 表示分类为负的正例。精确率和召回率比准确率更适合评判深度学习网络的性能。

另一个常用的深度学习网络性能评价指标是 F_1 参数, 其计算方法如下:

$$\frac{2}{F_1} = \frac{1}{\text{precision}} + \frac{1}{\text{recall}} \quad (6)$$

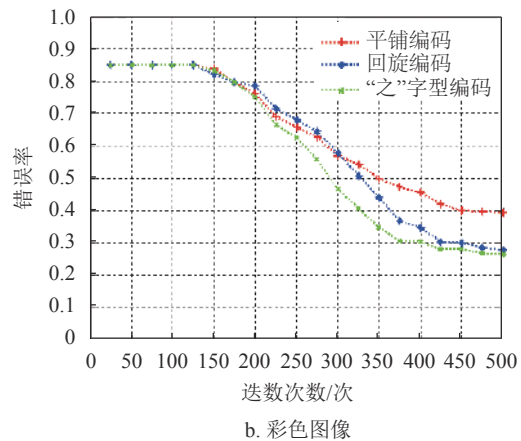
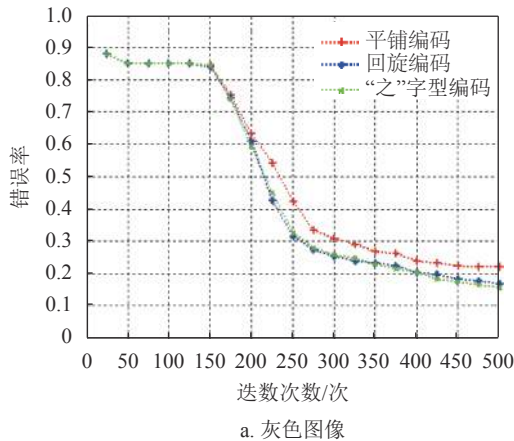


图7 Text 属性数据不同编码方案的错误率

图8为 Images 属性数据集的实验结果。Images 属性数据集的动作总数为 1 249 组, 从中随机选取 149 组作为测试数据, 剩下的作为训练数据。从图8可以得到与图7类似的结论; 对比图8和图7可知, Images 属性数据的识别错误率比 Text 属性数据稍低。

图9给出了 Images+Text 属性数据集的实验结果。Images+Text 属性数据集的动作总数为 1 248 组, 从中随机选取 148 组作为测试数据, 剩下的作

为训练数据。从总体趋势看, 采用不同编码方式的 Images+Text 属性数据的错误率及其变化趋势与 Text 属性数据类似。Video 属性数据以及 Video+Text 属性数据得到的结果与此类似, 此处不再赘述。

3.2 编码方法对识别结果的影响

对每种编码方法, 采用如下步骤迭代 500 次测试其性能: 1) 选取数据: 从整个数据集随机选取约 1/10 的样本作为测试数据, 其他数据作为训练数据; 2) 训练网络: 对 CNN 进行训练, 每训练 25 次使用测试数据测试一次; 3) 得到结果: 重复执行步骤 1) 和 2) 10 次, 取其均值。

图7给出了 Text 属性数据在不同编码方案下的错误率。Text 属性数据集的动作总数为 1 304 组, 从中随机选取 154 组作为测试数据, 剩下的作为训练数据。图7a为采用灰色图像编码的结果, 可以看出, “之”字型编码和回旋编码的识别错误率相近, 而平铺编码的识别错误率较高, 其原因是平铺编码所成图像十分相似, 而回旋编码和“之”字型编码却差异明显。图7b为采用彩色图像编码的结果, 可以看出, 3种编码的识别错误率从低到高分别为“之”字编码、回旋编码和平铺编码。从图7a和图7b可知, 各种编码方式的识别错误率都随着迭代次数的增加而逐渐减小, 最后趋于稳定; 灰色图像编码的识别性能优于彩色图像编码。

综上所述, 灰色图像编码的性能优于彩色图像编码; 在灰色图像编码方式中, 回旋编码的识别错误率最低, 收敛速度最快。因此, 本文采用回旋灰色图像编码测试动作识别模型的性能。

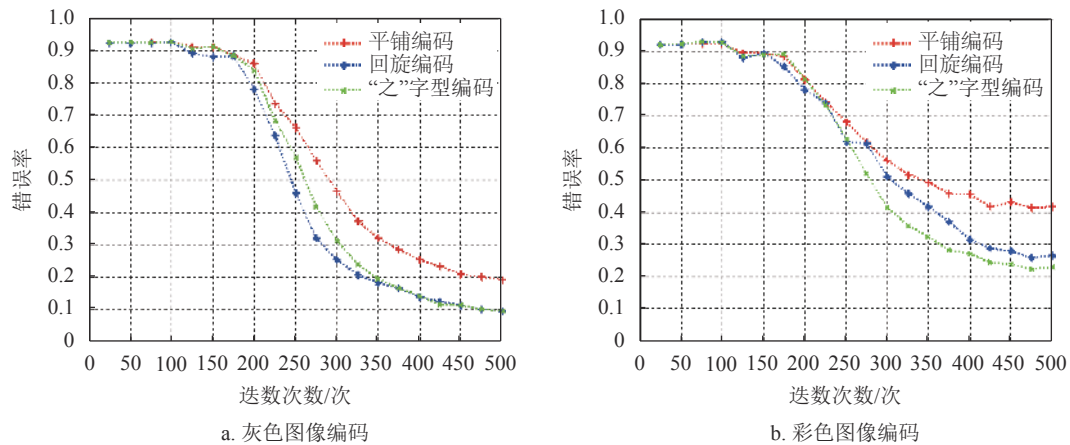


图8 采用不同编码方案的 Images 属性数据错误率

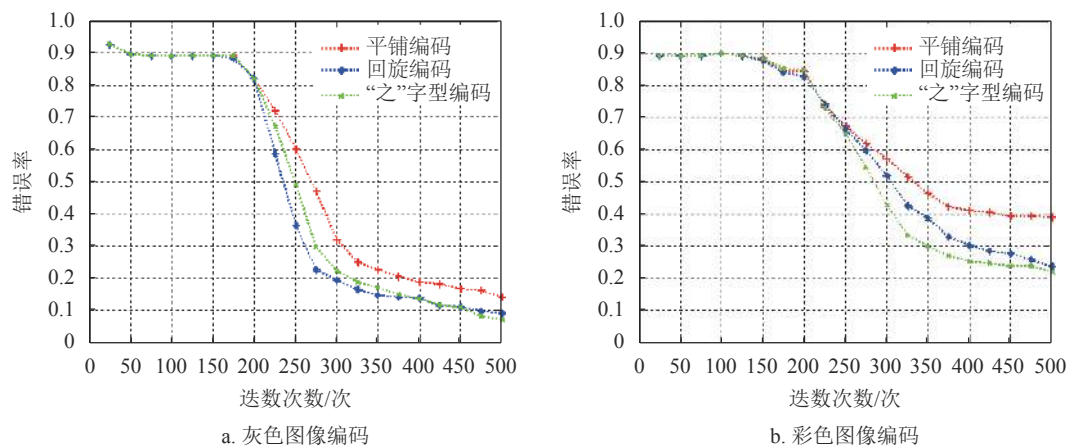


图9 采用不同编码方案的 Images+Text 属性数据错误率

3.3 动作识别模型的性能分析

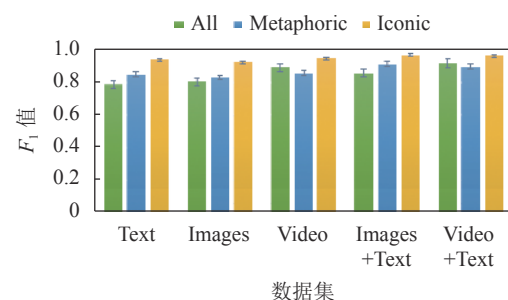
采用十折交叉验证法测试模型性能, 步骤如下: 1) 预处理: 将数据集等分为 10 份; 2) 选取数据: 选取其中 1 份作为测试集, 其他的作为训练集; 3) 得到结果: 记录实验结果, 然后重复步骤 2)、3) 10 次, 取 10 次结果的均值。

采用两种实验方案测试动作识别模型的性能, 一种方案称为单一属性测试, 其训练集和测试集都选择相同属性的数据, 比如都采用 Text 属性数据; 另一种方案称为混合属性测试, 其训练集和测试集分别选择不同属性的数据, 比如训练集使用 Text 属性数据, 测试集使用 Images 属性数据。

单一属性测试: 使用 F_1 作为评判标准, 5 种属性的数据各得到一个 F_1 值, 如图 10 所示。其中, Metaphoric 表示 6 个隐喻动作, Iconic 表示 6 个标准动作, All 则表示所有动作 (隐喻动作+标准动作)。如前所述, 每个 F_1 值都是重复测试 10 次得到的均值。

从图 10 可知, 在单一属性测试中, Video 属性数据拥有最好的性能, Images 属性数据紧随其

后, Text 属性数据的性能最差; 复合属性数据比复合的单一属性数据好, Video+Text 属性数据的 F_1 值最高。

图10 所有属性数据测试 10 次平均 F_1 值

混合属性测试: 针对所有动作、隐喻动作和标准动作, 分别测试训练集属性和测试集属性的所有可能组合, 其结果见图 11~图 13。

图 11 给出了全部动作数据的混合测试结果。从图 11 可以看出, 相比单一属性测试, 混合属性测试的 F_1 值有所下降, 这与文献 [11] 的结论相吻合。但是, 如果训练集采用的是复合属性数据, 并选择复合属性的任一成员属性作为测试集, 其 F_1 值

与使用相同属性数据测试的 F_1 值相当甚至更高, 反之, 则并不理想。

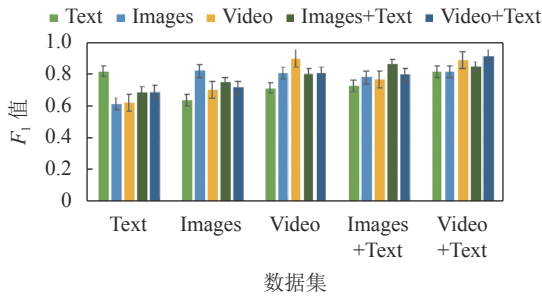


图 11 全部动作数据混合测试 10 次平均 F_1 值

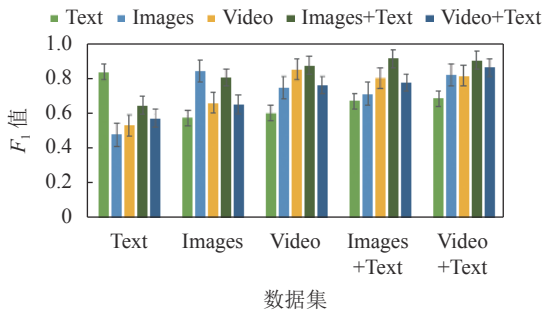


图 12 隐喻动作数据混合测试 10 次平均 F_1 值

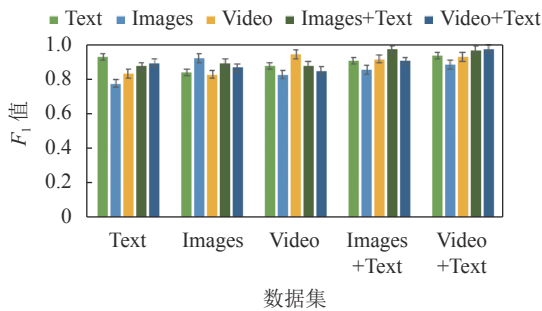


图 13 标准动作数据混合测试 10 次平均 F_1 值

图 12 给出了隐喻动作数据的混合测试结果, 其 F_1 值的整体走势与图 11 类似。不过, 相对于全部动作, 隐喻动作的 F_1 值普遍偏小, 意味着隐喻动作识别难度更大。

图 13 给出了标准动作数据的混合测试结果, 其结论与图 11 相同。不过, 相对于全部动作, 标准动作的 F_1 值整体偏大, 意味着标准动作更易识别。

图 14 显示了本文算法与文献 [12]、文献 [15] 的性能对比。公平起见, 各算法均采用相同属性的数据。文献 [12] 采用随机森林法进行训练和识别, 其特征向量包括 35 个关节角度、35 个关节角速度和 20 个关节点, 共 $35+35+20 \times 3=130$ 维。由于一组动作需用 35 帧描述, 因而特征向量共 $130 \times 35=4550$ 维。文献 [15] 旨在建立标准的人体动作模板, 它先建立人体模板字典, 再提取人体关节特征, 最后利用稀疏字典编码建立人体动作模板, 进

而利用模板匹配的方式识别人体动作。从图 14 可以看出, 本文算法的 F_1 值普遍大于 0.8, 遥遥领先于文献 [12], 相比文献 [15] 也有较大优势。

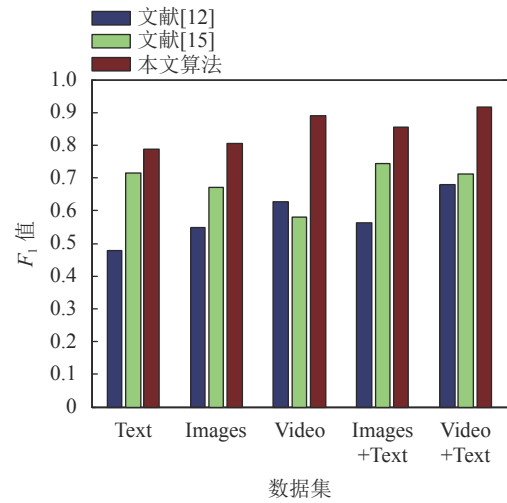


图 14 本文算法与文献 [12] 和文献 [15] 的识别性能对比

除了利用 MSRC-12 数据集进行性能测试之外, 本文还在 MSR Action3D 数据集上进行了性能测试, 以验证算法对不同数据集的普适性。本文算法在 MSR Action3D 数据集 (由 567 组动作组成) 上的平均识别准确率为 76.6%, 高于文献 [16] 提出的 Bag of 3D points 算法, 其平均识别准确率为 74.7%。另外, 本文还利用 Kinect 2.0 自采了图 1 中的 12 个动作各 10 组, 本文算法在自采数据集上的测试结果达到了 100% 的平均识别准确率。这些结果充分说明, 本文算法具有较强的场景适应能力。

4 结束语

人体动作识别在新型人机交互等众多领域具有广阔的应用前景。本文提出了一种基于卷积神经网络的精确人体动作识别算法, 用以对 Kinect 等传感器采集的动作数据集进行训练和识别。为了使用 CNN 框架处理动作数据, 本文设计了平铺编码、回形编码以及“之”字形编码, 这些编码方法可将动作数据编码成灰度图像或彩色图像。实验表明, 本文所提出的人体动作识别方法具有较高的识别率和较强的泛化能力。接下来, 将把本文算法应用于工程领域, 用以判断特殊岗位的工作人员 (比如煤矿井下的检修工人或矸石挑选工人) 是否执行了违规操作, 从而降低生产事故的发生率。

参考文献

- [1] JI S W, XU W, YANG M. 3D convolutional neural networks for human action recognition[J]. IEEE

- Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(1): 221-231.
- [2] ZHANG H B, ZHANG Y X, ZHONG B N, et al. A comprehensive survey of vision-based human action recognition methods[J]. *Sensors*, 2019, 19(1005): 1-20.
- [3] MURPHY K P. Machine learning: A probabilistic perspective[M]. Cambridge: The MIT Press, 2012.
- [4] WENG J W, WENG C Q, YUAN J S, et al. Discriminative spatio-temporal pattern discovery for 3D action recognition[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(4): 1077-1089.
- [5] RUSSELL S, NORVIG P. Artificial intelligence: A modern approach[M]. The 3rd edition. Chennai: Pearson Education India, 2015.
- [6] WANG J, ZHENG H C. View-robust action recognition based on temporal self-similarities and dynamic time warping[C]//In Proceeding of the 2012 IEEE International Conference on Computer Science & Automation Engineering (CSAE). Zhangjiajie: IEEE, 2012: 498-502.
- [7] 陈胜娣, 何冰倩, 陈思宇, 等. 基于时空兴趣点的人体动作识别[J]. *成都信息工程大学学报*, 2018, 33(2): 143-148.
CEHN Sheng-ti, HE Bing-qian, CHEN Si-yu, et al. Human action recognition based on Spatio-Temporal interest point[J]. *Journal of Chendu University of Information Technology*, 2018, 33(2): 143-148.
- [8] DENG L, YU D. Deep learning: Methods and applications[M]. Boston: Now Publishers Inc, 2014.
- [9] 张亮. 基于 Kinect 的人体动作识别算法研究与系统设计[D]. 徐州: 中国矿业大学, 2019.
ZHANG Liang. Algorithm research and system design of human motion recognition based on Kinect[D]. Xuzhou: China University of Mining and Technology, 2019.
- [10] SCHMIDHUBER J. Deep learning in neural networks: An overview[J]. *Neural Networks*, 2015, 61: 85-117.
- [11] 龚丁禧. 稀疏自组合时空卷积神经网络动作识别方法及其并行化[D]. 厦门: 厦门大学, 2014.
GONG Ding-xi. Action recognition method based on sparse auto-combination Spatio-Temporal convolutional neural network and its MapReduce implementation[J]. Xiamen: Xiamen University, 2014.
- [12] FOTHERGILL S, MENTIS H M, KOHLI P, et al. Instructing people for training gestural interactive systems[C]//In Proceeding of The 30th ACM Conference on Human Factors in Computing Systems. Austin: ACM, 2012: 1737-1746.
- [13] JANG E, GU S X, POOLE B. Categorical reparameterization with gumbel-softmax[C]//In proceeding of ICLR 2017: 5th International Conference on Learning Representations. Toulon: International Conference on Learning Representations, 2017.
- [14] LITJENS G, SÁNCHEZ C I, TIMOFEEVA N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis[J]. *Scientific Reports*, 2016, 6: 26286.
- [15] 刁俊方. 基于 Kinect 的人体动作识别技术研究[D]. 重庆: 重庆大学, 2015.
DIAO Jun-fang. Human action recognition research based on Kinect[D]. Chongqing: Chongqing University, 2015.
- [16] LI W Q, ZHANG Z Y, LIU Z C. Action recognition based on a bag of 3D points[C]//In proceeding of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. San Francisco: IEEE, 2010: 25-30.

编辑 刘飞阳

(上接第 472 页)

- [10] PETERSON L. K-nearest neighbor[J]. *Scholarpedia*, 2009, 4(2): 1883.
- [11] SVETNIK V, LIAW A, TONG C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling[J]. *Journal of Chemical Information & Computer Sciences*, 2003, 43(6): 1947.
- [12] YOCKEY H P. Information theory with applications to biogenesis and evolution[M]. McGraw-Hill: Biogenesis Evolution Homeostasis, 1973.
- [13] DRUCKER H. Improving regressors using Boosting techniques[C]//Proceedings of the Fourteenth International Conference on Machine Learning. [S.l.]: ACM, 1997: 107-115.
- [14] 孙嘉. 图像特征提取(纹理特征)[EB/OL]. [2018-07-26]. https://blog.csdn.net/weixin_42805841/article/details/81227078.
SUN Jia. Image feature extraction (texture feature)[EB/OL]. [2018-07-26]. https://blog.csdn.net/weixin_42805841/article/details/81227078.
- [15] GRIETHUYSEN J J M, FEDOROV A, PARMAR C, et al. Computational radiomics system to decode the radiographic phenotype[J]. *Cancer Research*, 2017, 77(21):104-107.
- [16] HARALICK R M, SHANMUGAM K, DINSTEN I H. Textural features for image classification[J]. *Studies in Media and Communication*, 1973, SMC-3(3): 610-621.
- [17] GALLOWAY M M. Texture analysis using gray level run lengths[J]. *IEEE Trans Comput Imaging*, 1975, 4(2): 172-181.
- [18] THIBAUT G, FERTIL B, NAVARRO C, et al. Texture indexes and gray level size Zone matrix application to cell nuclei classification[C]//10th International Conference on Pattern Recognition and Information Processing. [S.l.]: PRIP, 2009: 140-145.
- [19] SUN C, WEE W G. Neighboring gray level dependence Matrix for texture classification[J]. *Computer Vision, Graph Image Process*, 1983, 2(3): 341-352.
- [20] COMON P. Independent component analysis, a new concept?[J]. *Signal Processing*, 1994, 36(3): 287-314.
- [21] WOLD S. Principal component analysis[J]. *Chemometrics & Intelligent Laboratory Systems*, 1987, 2(1): 37-52.

编辑 刘飞阳