



(12) 发明专利

(10) 授权公告号 CN 113094547 B

(45) 授权公告日 2022. 01. 18

(21) 申请号 202110365963.1

(22) 申请日 2021.04.06

(65) 同一申请的已公布的文献号
申请公布号 CN 113094547 A

(43) 申请公布日 2021.07.09

(73) 专利权人 大连理工大学
地址 116024 辽宁省大连市甘井子区凌工
路2号

(72) 发明人 黄万鸿 韩兰灵 江波 刘玉琴

(74) 专利代理机构 大连理工大学专利中心
21200

代理人 关慧贞

(51) Int. Cl.

G06F 16/732 (2019.01)

G06F 16/78 (2019.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

(56) 对比文件

CN 111339942 A, 2020.06.26

CN 108763295 A, 2018.11.06

CN 103324685 A, 2013.09.25

CN 111914643 A, 2020.11.10

CN 111259804 A, 2020.06.09

WO 2019074316 A1, 2019.04.18

耿治萌等. 日语文本语料库的开发与利用.
《中国教育信息化》. 2015, (第01期),

刘晓林. 基于深度学习的人体动作识别研
究. 《中国优秀硕士学位论文全文数据库 信息科
技辑》. 2020, I138-2119.

Langling Han等. Implementation of
CALL-based online Japanese learning
system. 《IEEE》. 2011, 585-588.

Rong Liu等. Si-GCN: Structure-induced
Graph Convolution Network for Skeleton-
based Action Recognition. 《IEEE》. 2019, 1-8.

审查员 邓慧丽

权利要求书2页 说明书6页 附图2页

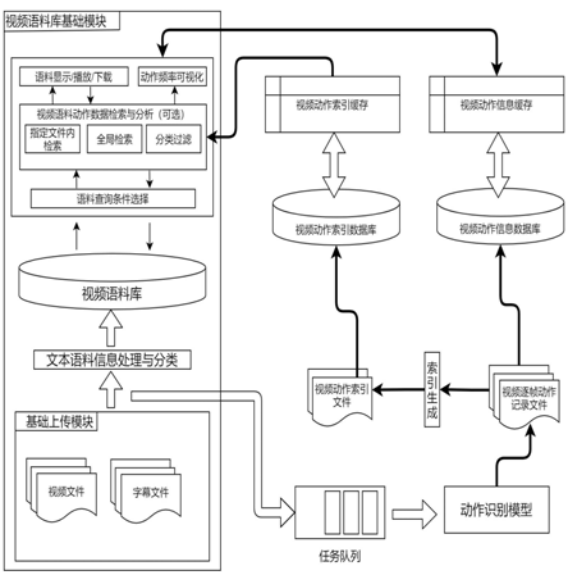
(54) 发明名称

日语在线视频语料中特定动作视频片断检
索方法

(57) 摘要

本发明日语在线视频语料中特定动作视频片断检索方法属于视频片断检索领域, 涉及一种基于深度学习技术的在线日语多模态语料库中对包含特定动作的视频片断的快速检索方法。检索方法采用了深度学习与统计学习技术、动作和字幕与视频匹配的定位技术, 缓存与索引技术, 数据可视化技术。检索方法分为视频语料的上传、视频语料的逐帧动作解析与特征提取、索引建立、动作检索四大步骤。首先在视频语料上传结束后, 使动作解析与数据处理服务器对其进行解析与特征提取, 并对最终结果进行索引, 用户通过WEB应用进行检索。本发明实现了日语视频语料中特定动作的视频片断检索的快速查询、准确定位、下载, 为日语学习研究提供方便的检索

服务。



1. 一种日语在线视频语料中特定动作视频片断检索方法, 该检索方法采用的网络系统由计算机, 视频语料库和服务服务器构成, 其特征是, 该检索方法采用的网络系统还具有动作解析与数据分析服务器, 该服务器与计算机连接, 并通过因特网使用; 动作解析与数据处理服务器与视频语料库服务器分离, 两个服务器之间通过HTTP协议进行通信; 检索方法通过视频语料库上传模块, 进行视频的上传以及逐帧解析; 在视频语料库基础模块中添加了动作检索模块、动作识别与特征解析模块; 动作识别与特征解析模块采用人体关节点特征识别技术以及图卷积网络深度学习技术; 检索方法采用了索引与缓存技术, 统计学习与分析技术、数据可视化技术、动作片断与字幕和视频匹配的定位技术, 全局语料动作检索以及特定语料动作检索技术, WEB技术, 实现了包含特定动作的视频片断的快速查询; 动作数据库的建立、缓存与索引的建立, 视频特征的分析与可视化来完成; 检索方法的具体内容如下:

一、视频语料的上传;

对于视频语料的上传, 存在两个分支:

1) 一个分支为原有视频语料库上传分支, 正常完成视频语料的上传任务, 在上传完成后, 上传完成的视频语料将会被放置于某个文件夹下, 此时生成一个随机序列号UUID, 方便对会话进行追踪, 并使用该UUID创建视频动作解析请求对象, 置入动作解析后端模块的任务队列中, 在该上传语料的动作解析任务完成之前, 可以正常使用语料库检索到该语料, 但对于该语料的动作检索暂时不可用, 全局语料动作检索中暂不收录该语料, 在数据库中为视频语料添加了动作检索的有效标识; 在某语料的动作解析任务全部完成之后, 该标识将置为有效, 之后对于该语料的动作检索以及分析变为可用;

2) 另一个分支为新添加的视频语料动作解析与数据处理分支; 在用户提交上传请求后; 用户所选择上传的视频语料数据会被并行地上传到负责视频语料动作解析与数据处理服务器中, 在上传完毕后, 会生成一个动作数据解析的逻辑请求对象, 其中包含视频语料的定位信息以及用于会话跟踪的UUID信息, 置入视频语料动作解析与数据处理服务器的任务队列中;

二、动作的逐帧解析与记录;

在动作解析模块后端中, 为了保持系统在请求压力较大的环境下的稳定性, 当任务队列中存在未完成的任务时, 服务器将每次从任务队列中取出一个任务, 根据任务的具体信息, 定位到视频语料文件, 对于输入的视频, 使用Openpose逐帧进行对人物以及关节点坐标与特征识别的解析, 并生成 $[K, n, 18, 3]$ 的输出, 其中, K —帧数, n —人物数量, 18 —关节关键点数, 3 —二维坐标以及置信度; 之后该输出作为图卷积网络的输入, 对于每帧中动作的识别, 采用以该帧为中心 $[input_size, n, 18, 3]$ 的数据作为图卷积网络的输入; 其中, $input_size$ 根据实际情况调节; 最终图卷积网络的输出与全连接层相连, 产生 $[n, A]$ 的输出, 代表每个人在该帧做动作列表中某个动作的置信度, 其中, A —某个动作的置信度; 由于该值未被归一化, 使用softmax函数归一化后, 作为最终的动作概率输出; 之后获得的完整特征集合将作为图卷积网络预训练模型的输入, 最终得到逐帧的动作预测结果集合, 对预测结果使用归一化技术进行处理, 并使用阈值过滤, 最终将结果集合存储于使用Java语言编写的特制的数据库中, 生成视频逐帧动作记录文件;

三、动作数据索引与缓存建立技术;

为了高效地对包含指定动作的视频片断进行检索, 需要建立动作数据索引, 在视频的

逐帧动作解析完成后,将会将视频语料标识符作为索引键值,对所出现的动作片断,位置,时长,特征向量作为索引具体项进行索引的建立;同时为了提高检索效率,重复利用数据访问的时间以及空间局部性,减少I/O操作代价,在主存中为索引以及具体项建立缓存;

四、语法解析功能;

该功能采用基于自动机的语法识别算法进行实现;编写了通用有限自动机结构模块,并为了在日语能力水平考试中高频出现的217条语法分别构造了相应的语法解析自动机;为该语法识别模块输入一个合法的句子后,能够输出相应的识别语法编号以及简单描述;

五、有特定动作的视频片断检索技术;

对于特定视频语料的动作检索,将会从索引缓存中提取对应索引,并从动作数据缓存中检索相应的动作数据,最终按出现的时间顺序显示在WEB前端页面中;同时每个动作的出现位置使用Kuromoji日语分词技术进行上下文文本语料的提取,使用者能够查看该动作出现时的上下文文本;这些数据可以保存,作为学习或分析研究使用,或作为其他语义分析模型的输入使用;

六、动作数据可视化与相似度分析技术;

对于每个视频语料,使用统计学技术提取其出现的动作分布,再使用分布相似度度量算法以及数据降维技术对动作分布进行可视化以及相近动作分布的视频显示;利用Web前端echart框架,结合使用python编写的数据处理后端,对该语料中某个动作的出现分布进行可视化;使用概率作为分布值,将对动作分布核密度估计后的结果绘制显示在前端页面上;对于每个动作出现的时间点,对出现的词汇上下文进行分词,取附近的k个词进行词频统计,并且使用预测概率作为统计的权重;最终,为每种动作出现时的上下文单词,使用echart散点图进行数据可视化,并显示在web页面中,其中,散点大小由词频决定;鼠标移动到散点上后,将显示浮点标签,显示该散点所代表单词的具体含义;单击后,将跳转到该单词的检索页。

日语在线视频语料中特定动作视频片断检索方法

技术领域

[0001] 本发明属于视频片断检索领域,涉及一种基于深度学习技术的在线日语多模态语料库中对包含特定动作的视频片断的快速检索方法。

背景技术

[0002] 近年来,随着互联网技术的发展,越来越多的外语学习者尝试选择通过视频语料进行外语的学习,视频语料在外语学习过程中受到高度关注,因为视频语料能够弥补单纯文本语料的不足,提供更真实的语境。而动作在视频语料中扮演着重要的角色,因为动作往往是语境与文化的体现,寻找视频中出现的动作,不论对外语学习者或是研究者来说都具有重要意义。对于日语学习者来说,有助于对某个语言下的特定语境以及文化的理解,加深对词汇、语法、语用的印象,对于研究者来说,有助于更方便地分析某个视频片断的语义,分析在某种语境下的动作特征,以及对特定文化进行分析。而在深度学习技术出现以前,提取视频动作特征及片断是一件复杂的事,随着近年深度学习技术的发展,识别并提取视频中的动作变得在工程实践中具有可行性。

[0003] 目前,虽然存在日语视频语料库,例如姜国海等人的专利,ZL201310216448.2“日语在线视频语料的视频片断检索方法”,但均不具有动作检索与特征提取功能,失去了动作所带来的丰富的语义信息,使得语境分类,语用研究,文化分析等工作较难进行。为了捕获这些动作所带来的语义信息,从而更好地进行日语研究以及教学,实现一个高效的日语语境下的包含特定动作的视频片断检索系统是必要的。

发明内容

[0004] 本发明的目的是克服现有技术的缺陷,发明一种基于深度学习的在线日语视频语料的包含指定动作的视频片断检索方法。该方法在基础视频语料库模块上,添加了动作检索与特征分析功能,采用了深度学习计算机视觉技术,在Openpose人体关节点特征识别技术的基础上使用深度学习图卷积网络(GCN)作了进一步的动作识别分类,能够对视频语料中点头,鞠躬等动作进行检测和动作片断提取与记录,还采用了统计学习与分析技术、特征比较技术,动作与字幕和视频匹配的定位技术,数据可视化技术,基于日语分词的动作上下文文本语料采样技术,基于自动机的语法解析技术,全局语料动作检索以及特定语料动作检索技术,WEB相关技术,如索引技术,缓存技术,消息队列技术等,数据的空间与时间局部性的利用技术等,实现了视频片断的快速查询,保障了动作检索系统的高效稳定运行。

[0005] 本发明采用的技术方案是日语在线视频语料中特定动作视频片断检索方法,该检索方法采用的网络系统由计算机,视频语料库和服务器构成,其特征是,该检索方法采用的网络系统还具有动作解析与数据分析服务器,该服务器与计算机连接,并通过因特网使用;两种服务器之间通过HTTP协议进行通信;检索方法通过视频语料库上传模块,进行视频的上传以及逐帧解析;检索方法采用了动作检索模块、动作识别与特征解析模块;动作识别与特征解析模块采用人体关节点特征识别技术以及图卷积网络深度学习技术;检索方法采用

了索引与缓存技术,统计学习与分析技术、数据可视化技术、动作片断与字幕和视频匹配的定位技术,全局语料动作检索以及特定语料动作检索技术,WEB技术,实现了包含特定动作的视频片断的快速查询;动作数据库的建立、缓存与索引的建立,视频特征的分析与可视化来完成;检索方法具体内容如下:

[0006] 一、视频语料的上传

[0007] 对于视频语料的上传,存在两个分支:

[0008] 1) 一个分支为原有视频语料库上传分支,正常完成视频语料的上传任务,在上传完成后,上传完成的视频语料将会被放置于某个文件夹下,此时生成一个随机序列号UUID,方便对会话进行追踪,并使用该UUID创建视频动作解析请求对象,置入动作解析后端模块的任务队列中,在该上传语料的动作解析任务完成之前,可以正常使用语料库检索到该语料,但对于该语料的动作检索暂时不可用,全局语料动作检索中暂不收录该语料,在数据库中为视频语料添加了动作检索的有效标识。在某语料的动作解析任务全部完成之后,该标识将置为有效,之后对于该语料的动作检索以及分析变为可用;

[0009] 2) 另一个分支为新附加的视频语料动作解析与数据处理分支。在用户提交上传请求后。用户所选择上传的视频语料数据会被并行地上传到负责视频语料动作解析与数据处理服务器中,在上传完毕后,会生成一个动作数据解析的逻辑请求对象,其中包含视频语料的定位信息以及用于会话跟踪的UUID信息,置入视频语料动作解析与数据处理服务器的任务队列中。

[0010] 二、动作的逐帧解析与记录

[0011] 在动作解析模块后端中,为了保持系统在请求压力较大的环境下的稳定性,当任务队列中存在未完成的任务时,服务器将每次从任务队列中取出一个任务,根据任务的具体信息,定位到视频语料文件,对于输入的视频,使用Openpose逐帧进行对人物以及关节点坐标与特征识别的解析,并生成 $[K, n, 18, 3]$ 的输出,其中, n —人物数量, 18 —关节关键点数, 3 —二维坐标以及置信度, K —帧数。之后该输出作为图卷积网络的输入,对于每帧中动作的识别,采用以该帧为中心 $[input_size, n, 18, 3]$ 的数据作为图卷积网络的输入。其中, $input_size$ 根据实际情况调节。最终图卷积网络的输出与全连接层相连,产生 $[n, A]$ 的输出,代表每个人在该帧做动作列表中某个动作的置信度,其中, A —某个动作的置信度。由于该值未被归一化,使用softmax函数归一化后,作为最终的动作概率输出。之后获得的完整特征集合将作为图卷积网络预训练模型的输入,最终得到逐帧的动作预测结果集合,对预测结果使用归一化技术进行处理,并使用阈值过滤,最终将结果集合存储于使用Java语言编写的特制的数据库中,生成视频逐帧动作记录文件;

[0012] 三、动作数据索引与缓存建立技术

[0013] 为了高效地对包含指定动作的视频片断进行检索,需要建立动作数据索引,在视频的逐帧动作解析完成后,将会将视频语料标识符作为索引键值,对所出现的动作片断,位置,时长,特征向量等作为索引具体项进行索引的建立。同时为了提高检索效率,重复利用数据访问的时间以及空间局部性,减少I/O操作代价,在主存中为索引以及具体项建立缓存。

[0014] 四、语法解析功能

[0015] 该功能采用基于自动机的语法识别算法进行实现。为此,编写了通用有限自动机

结构模块,并为在日语能力水平考试(JLPT)中高频出现的217条语法分别构造了相应的语法解析自动机。为该语法识别模块输入一个合法的句子后,能够输出相应的识别语法编号以及简单描述。

[0016] 五、有特定动作的视频片断检索技术

[0017] 对于特定视频语料的动作检索,将会从索引缓存中提取对应索引,并从动作数据缓存中检索相应的动作数据,最终按出现的时间顺序显示在WEB前端页面中。同时每个动作的出现位置使用Kuromoji日语分词技术进行上下文文本语料的提取,使用者能够查看该动作出现时的上下文文本。这些数据可以保存,作为学习或分析研究使用,或作为其他语义分析模型的输入使用。

[0018] 六、动作数据可视化与相似度分析技术

[0019] 对于每个视频语料,使用统计学技术提取其出现的动作分布,再使用分布相似度量算法以及数据降维技术对动作分布进行可视化以及相近动作分布的视频显示;利用Web前端echart框架,结合使用python编写的数据处理后端,对该语料中某个动作的出现分布进行可视化;使用概率作为分布值,将对动作分布核密度估计后的结果绘制显示在前端页面上;对于每个动作出现的时间点,对出现的词汇上下文进行分词,取附近的k个词进行词频统计,并且使用预测概率作为统计的权重;最终,为每种动作出现时的上下文单词,使用echart散点图进行数据可视化,并显示在web页面中,其中,散点大小由词频决定;鼠标移动到散点上后,将显示浮点标签,显示该散点所代表单词的具体含义;单击后,将跳转到该单词的检索页。

[0020] 本发明具有有益效果是该方法运用了人体关节点特征识别技术,图卷积网络等深度学习技术,采用了索引与缓存技术,统计学习与分析技术、数据可视化技术、动作片断与字幕和视频匹配的定位技术,全局语料动作检索以及特定语料动作检索技术,WEB技术,实现了在视频语料库检索界面对包含指定动作片断的视频语料进行快速查询,以及精确定位,以及对相似动作频率分布视频的探索。为日语学习者提供了更丰富的语境信息,为日语研究者提供了对语言,文化进行分析的工具。

附图说明

[0021] 图1为日语在线视频语料中特定动作视频片断检索方法流程图。

[0022] 图2为视频语料的动作解析与数据输出流程框图

具体实施方式

[0023] 结合附图和技术方案详细说明本发明的具体实施。

[0024] 本实施例中采用的网络系统由计算机,视频语料库和语料库服务器以及动作解析与数据分析服务器构成,两个服务器之间通过HTTP协议进行通信。对于视频语料的查询,用户通过因特网与语料库服务器进行交互,上传等操作模块作为视频语料库基础模块。在此之上添加与视频语料库服务器独立的动作解析与数据处理服务器。服务器所处操作系统均为Windows 10x64,动作解析与数据处理服务器基于Python实现,而视频动作片断检索功能基于Java EE技术实现。基于JDK 1.8以及SpringBoot开源框架进行检索端的编写。开发环境使用PyCharm以及IntelliJ IDEA。对于动作上下文语料分词提取,使用了开源免费的

Kuromoji工具包,使用Java语言进行检索以及分词与语法解析处理。

[0025] 前端数据可视化的图表绘制使用了开源框架echarts,动作解析数据处理。后端采用pandas,keras,pytorch,openpose,mmskeleton等开源库进行数据分析以及机器学习算法处理。

[0026] 本发明基于已有的日语视频语料库以及深度学习相关技术,对所上传视频语料进行解析分类以及数据处理,还为处理结果建立索引,依据索引表对视频语料内指定动作的出现帧位置进行视频语料内指定动作的检索。应用缓存技术对磁盘I/O访问次数进行优化,提高检索效率,利用数据可视化相关技术,对动作出现分布进行可视化,应用分词以及帧定位技术提取动作出现的上下文语境,通过在线方式提供给互联网的使用者。图1为日语在线视频语料中特定动作视频片断检索方法流程图,图2为视频语料的动作解析与数据输出流程框图,实现动作检索方法的具体过程如下:

[0027] 一、视频上传时对视频进行动作解析及数据处理。

[0028] 为提高检索效率,采用了在上传时而非检索时对视频进行动作解析以及数据处理的方式。采用java语言编写了数据缓存模块以及文件系统访问模块,搭建了建议的数据库对动作数据文件进行读写以及缓存操作。在原先视频语料库已有的上传模块的基础上,拓展了处理流程。在上传结束之后生成带唯一标识请求号的请求对象置于动作解析与数据处理服务器的任务队列中,该任务队列使用Redis分布式缓存数据库进行简单实现,见流程图1。并在在文件系统数据库文件中添加语料数据项以及索引项,暂时设置动作检索有效位为无效。若被置换的数据被更改则写回文件系统。在该语料被动作解析与数据处理服务器处理完毕后,动作检索有效位将被置为有效。在未被处理完成之前,该语料能够被已有的视频语料库基础模块正常收录与检索,但对于该语料的动作检索暂时无效。

[0029] 二、动作解析功能和特征识别

[0030] 对于输入的视频,使用Openpose逐帧进行对人物以及关节点坐标与特征识别的解析,并生成 $[K,n,18,3]$ 的输出,其中, n —人物数量,18—关节关键点数,3—二维坐标以及置信度, K —帧数。之后该输出作为图卷积网络的输入,对于每帧中动作的识别,采用以该帧为中心 $[input_size,n,18,3]$ 的数据作为图卷积网络的输入。其中, $input_size$ 可以根据实际情况调节。最终图卷积网络的输出与全连接层相连,产生 $[n,A]$ 的输出,代表每个人在该帧做动作列表中某个动作的置信度,其中, A —某个动作的置信度。由于该值未被归一化,使用softmax函数归一化后,作为最终的动作概率输出。

[0031] 例如,在日语语境下,鞠躬动作往往蕴含着丰富的语义信息,并且这些语义信息可能随着鞠躬角度,时间等特征的变化而变化。为了捕获这些动作所带来的语义信息,从而更好地进行日语研究以及教学,实现一个高效的日语语境下的包含特定动作的视频片断检索系统。对于学生或日语爱好者来说,可以利用该检索系统进行查询,从而知道某个动作发生时,往往会存在着怎样的语境;或在某种语境下,其中的人物往往随着哪些动作。这构建起了文本与视频语料之间的桥梁,使得日语学习者能更能融入相应的语境以及文化进行日语的学习。例如,学习者可以知道,在哪些语句下往往伴随着鞠躬动作,或发生鞠躬动作时,往往伴随着哪些语句,此后学习者在实践使用相应的语句时,也能够自然地去做出相应的鞠躬动作。另一方面,对于研究者来说,能够获取动作出现时的上下文分词以及语法解析结果,之后对词性,词的形态素,词频以及语法等进行分析,从而用于相关的研究工作中。

[0032] 三、动作数据索引与缓存功能

[0033] 作为提高系统检索效率的一种手段,在本方法中还使用了索引以及缓存技术。在每上传一个文件,并由动作解析与数据处理服务器处理完成后,服务器将会为从深度学习模型输出动作数据进行索引的建立。其中,索引键值为语料标识ID,对应索引项为语料标识ID所对应的动作数据文件。为了减少文件系统操作所带来的开销,在本方法中使用了缓存技术,对索引以及具体动作数据项进行缓存,其中缓存大小可根据实际需求设置,若缓存满后,尝试调入一个新的语料动作数据时,将会执行LRU置换算算法置换出最近最少使用的缓存项。并且使用预加载以及按需加载策略,在服务器启动时,首先在缓存中装载部分数据,其余数据在存在请求时从文件系统调出。

[0034] 四、语法解析功能

[0035] 该功能采用基于自动机的语法识别算法进行实现。为此,编写了通用有限自动机结构模块,并为在日语能力水平考试(JLPT)中高频出现的217条语法分别构造了相应的语法解析自动机。为该语法识别模块输入一个合法的句子后,能够输出相应的识别语法编号以及简单描述。

[0036] 例如为该模块输入句子“結果のいかにかわらず、必ず報告してください”后,模块将会返回(id_i, “[体言]+の+いかにかわらず”)以及(id_j, “動詞+て+ください”)。其中语法编号(ID)以及文字描述均存储在本系统的MySQL数据库中。

[0037] 五、包含特定动作的视频片断检索功能

[0038] 该功能与原视频语料库的检索功能配合使用,所谓包含特定动作的视频片断检索是指在服务器端存放完整的视频,对于特定视频语料的动作检索,用户通过检索相应的词汇,可以显示出相应的视频语料,当进入某个视频语料后,会对该视频中所出现的动作通过索引与缓存模块进行检索,并且在视频语料播放与详情页中显示每个动作所出现的时间位置以及概率,并且用户可下载格式为csv的动作数据文件,方便后续进行研究与分析工作。对于全局动作检索,用户可以在选择动作类型后,单击检索,之后会对全局语料进行检索显示出出现该动作最多的k个视频语料。单击语料项后,可以进入视频语料详情页。

[0039] a). 特定视频语料的动作检索

[0040] 用户通过浏览器访问本应用提供的Web网站,按照通常的视频语料库使用方式,选择“词汇查询”,在检索框中输入要查询的日语词汇,即可查询到相应的视频语料。在搜索结果将会附加上各个视频语料中每种动作的出现次数。当用户选择一个视频语料单击进入详情页后,除了视频对应的文本语料信息,以及播放界面等基础视频语料库前端部件之外,还会存在一个动作列表,在其中显示所出现的动作以及发生时间,以及判别概率。当在动作列表中单击某个项目后,将会获取该动作时间点的上下文语料,使用分词技术将分词结果显示,用户可以单击分词后的词汇进行词汇检索的跳转。此外,也使用语法解析功能对该动作出现时间点对应的句子,以及上下文前后两个句子进行语法解析并显示在前端页面中。

[0041] b). 全局视频语料动作检索

[0042] 基本操作流程为,用户在浏览器中打开该应用的Web主页,之后在搜索框旁边的选项中选择动作检索,并且选择某个动作,之后,系统将会对存在的所有语料中,出现该动作的语料进行检索,并返回按动作频率排序的视频语料搜索结果。类似特定视频语料的动作检索功能,搜索结果中,除了显示视频语料的概要信息,如标题,摘要外,还会显示各个动作

出现的次数,在用户单击某个视频语料后,将同样进入的视频语料的详情页。

[0043] c).数据分布可视化功能

[0044] 在视频语料详情页中,除了显示动作列表外,还利用Web前端echart框架,结合使用python编写的数据处理后端,对该语料中某个动作的出现分布进行可视化。使用概率作为分布值,将对动作分布核密度估计后的结果绘制显示在前端页面上。同时,对于每个动作的出现时间点对出现的词汇上下文进行分词,取附近的k个词进行词频统计,并且使用预测概率作为统计的权重。最终,为每种动作出现时的上下文单词,使用echart散点图进行数据可视化,并显示在web页面中。其中散点大小,由词频决定。鼠标移动到散点上后,将显示浮点标签,其中显示该散点所代表单词的具体含义,单击后,将跳转到该单词的检索页。

[0045] d).动作数据下载功能

[0046] 在视频语料详情页中,用户可以选择动作数据下载,可以对动作出现时间与预测概率的CSV文件的下载,以及GCN网络为该语料提取的对应动作特征数据的下载,以便进行进一步的分析或研究工作。

[0047] 在已有的基础视频语料库模块上,拓展了上传模块的功能,添加了动作解析以及动作数据处理与存储模块,能够在上传操作中,将所上传的语料信息作为请求内容的一部分,对动作解析与处理模块发起解析请求,为了保障系统的稳定性,使用消息队列结构,动作解析与数据处理服务器,每次从消息队列中取出一个请求进行处理。

[0048] 处理具体包括对视频语料的逐帧关节点特征提取与定位,以及通过图卷积网络结构以多个时间点上提取的人物关节位置与特征作为图卷积网络的输入进行动作的预测。预测结果使用概率归一化技术进行归一化,并使用阈值过滤,最终将预测结果以及特征记录在文件系统中。为了提高系统的效率,使用了缓存与索引机制,通过特定的语料ID作为索引值,能够快速找到该语料中所出现的动作,并且对索引存储以及动作数据存储均设计了缓存模块进行管理以减少文件系统操作带来的开销,在服务器启动时,将会进行预缓存机制,将部分数据载入内存中,未载入的数据使用按需加载机制。本系统与用户的交互主要使用WEB应用端进行,用户本系统中web应用指定的网址,可以对系统进行访问。本系统WEB端在原视频语料库WEB端上拓展,用户在特定语料的详情页中能够看到所出现的动作列表以及可视化后的动作分布。并可以选择下载动作数据。同时使用日语分词技术,采用开源的kuromoji工具包,对动作所出现的上下文语料进行分词操作,用户单击特定的动作出现时间点后,可以显示该动作的上下文分词过的文本语料,并单击进行词汇搜索的跳转。同时,在详情页面还对动作分布,以及分词过后的各动作出现的上下文以预测概率为权重,进行词频统计后的结果使用echart进行数据可视化,并可以单击可视化的散点进行特定单词检索的跳转。方便学生更好地进行日语的学习,以及方便研究人员对于特定语境下的语用,文化等的研究。

[0049] 本发明实现了日语视频语料的包含特定动作的视频片断检索的快速查询、准确定位以及对于动作数据的可视化以及下载,为日语学习者、研究者、爱好者提供方便的检索服务,以便更好地对特定语境下的语言,文化进行分析与学习。

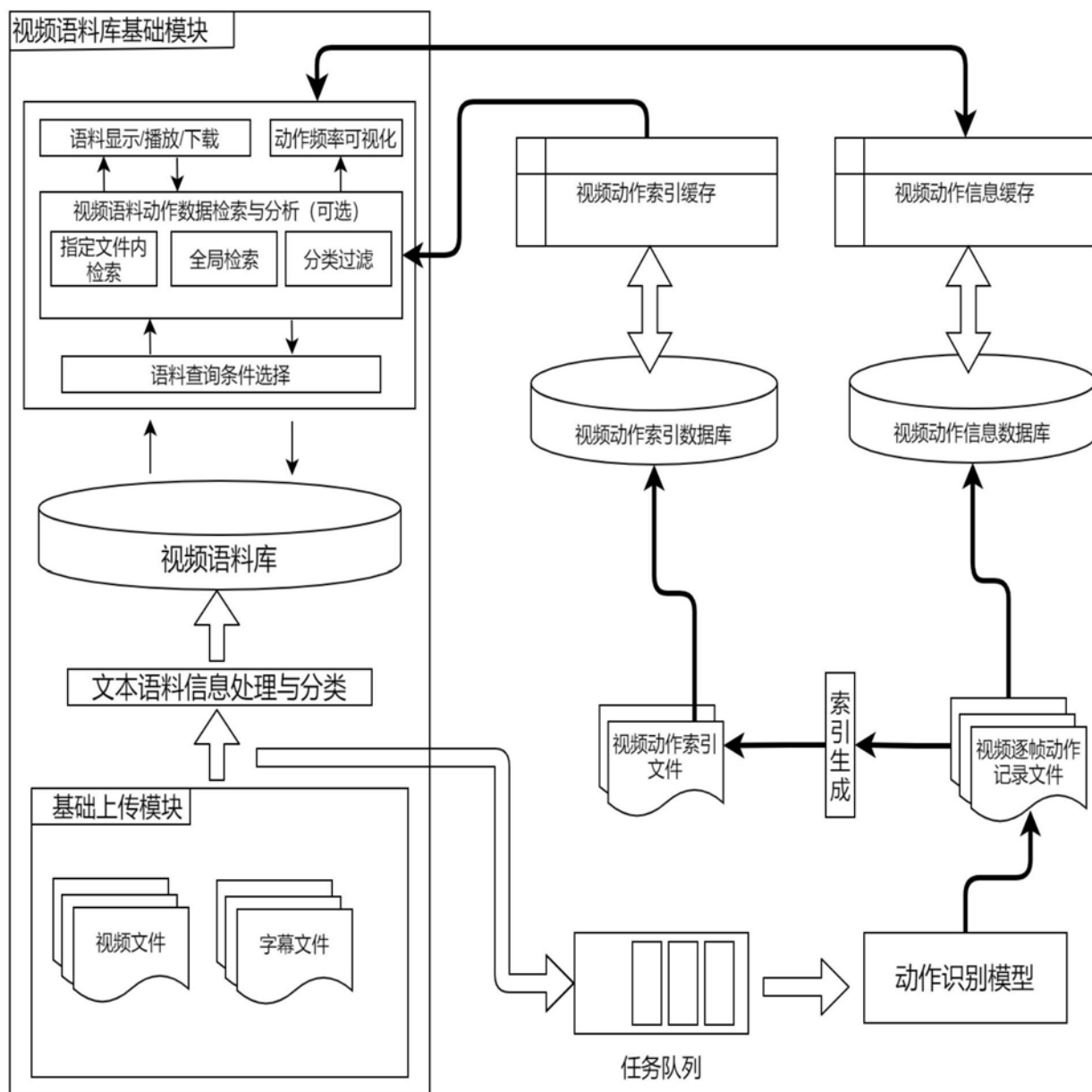


图1

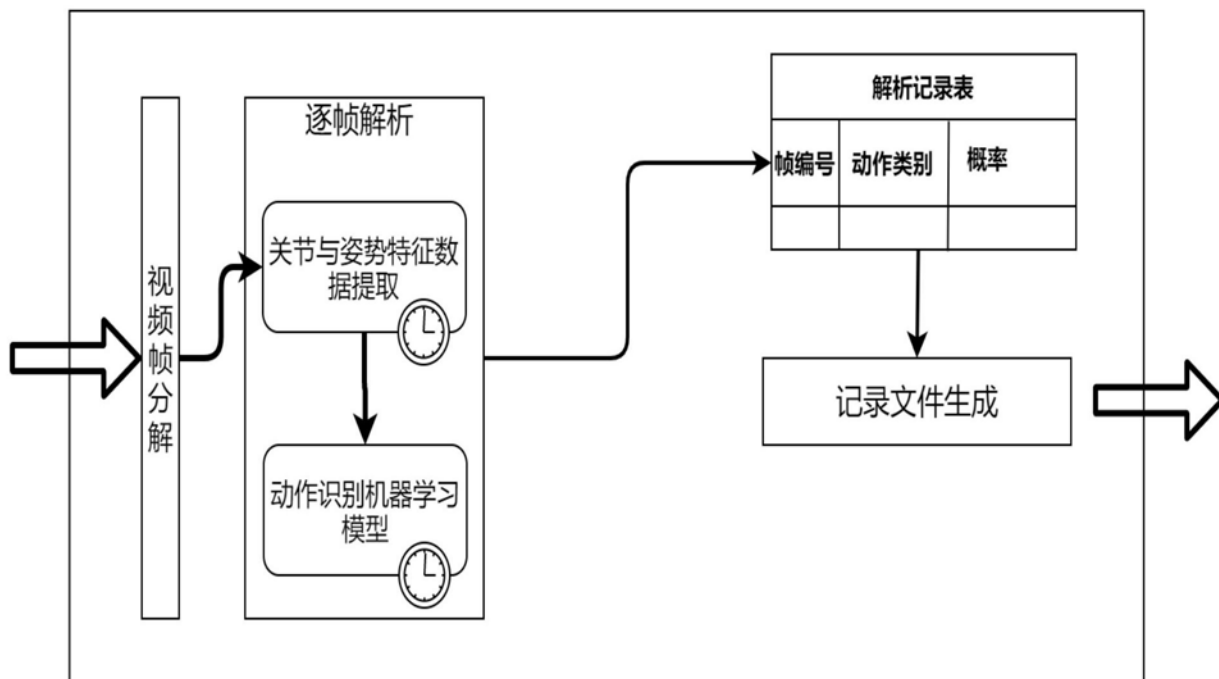


图2