

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2022/213857 A1

(43) 国际公布日
2022 年 10 月 13 日 (13.10.2022)

(51) 国际专利分类号:
G06K 9/00 (2022.01)

(21) 国际申请号: PCT/CN2022/083988

(22) 国际申请日: 2022 年 3 月 30 日 (30.03.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110380638.2 2021 年 4 月 9 日 (09.04.2021) CN

(71) 申请人: 京东科技控股股份有限公司
(JINGDONG TECHNOLOGY HOLDING CO., LTD.)
[CN/CN]; 中国北京市北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。

(72) 发明人: 邱钊凡(QIU, Zhaofan); 中国北京市北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室,

Beijing 100176 (CN)。潘滢炜(PAN, Yingwei); 中国北京市北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。姚霆(YAO, Ting); 中国北京市北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。梅涛(MEI, Tao); 中国北京市北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。

(74) 代理人: 北京英赛嘉华知识产权代理有限公司(INSIGHT INTELLECTUAL PROPERTY LIMITED); 中国北京市朝阳区建外大街光华东里 8 号中海广场中楼 7 层, Beijing 100020 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT,

(54) Title: ACTION RECOGNITION METHOD AND APPARATUS

(54) 发明名称: 动作识别方法和装置

200

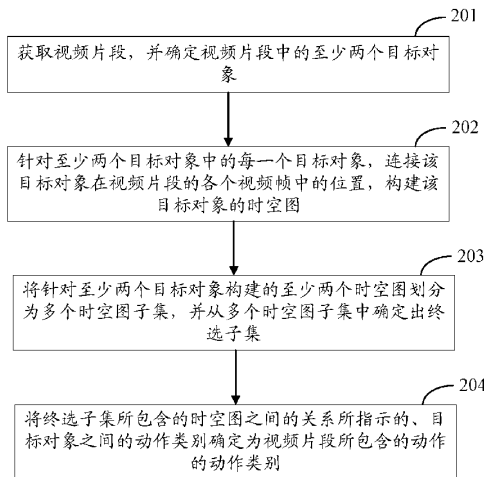


图 2

201 Acquire a video clip, and determine at least two target objects in the video clip
202 For each target object in the at least two target objects, connect positions of the target object in various video frames of the video clip, so as to construct a spatiotemporal graph of the target object
203 Divide at least two spatiotemporal graphs, which are constructed for the at least two target objects, into a plurality of spatiotemporal graph subsets, and determine a finally selected subset from the plurality of spatiotemporal graph subsets
204 Determine an action category of the action between the target objects that is indicated by a relationship between the spatiotemporal graphs included in the finally selected subset as the action category of an action included in the video clip

(57) Abstract: Disclosed in the present application are an action recognition method and apparatus. The method comprises: acquiring a video clip, and determining at least two target objects in the video clip; for each of the at least two target objects, connecting positions of the target object in various video frames of the video clip, so as to construct a spatiotemporal graph of the target object; dividing at least two spatiotemporal graphs, which are constructed for the at least two target objects, into a plurality of spatiotemporal graph subsets, and determining a finally selected subset from the plurality of spatiotemporal graph subsets; and determining an action category of the

[见续页]

JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

action between the target objects that is indicated by a relationship between the spatiotemporal graphs included in the finally selected subset as the action category of an action included in the video clip.

(57) 摘要: 本申请公开了动作识别方法和装置。其中, 该方法包括: 获取视频片段, 并确定视频片段中的至少两个目标对象; 针对至少两个目标对象中的每一个目标对象, 连接该目标对象在视频片段的各个视频帧中的位置, 构建该目标对象的时空图; 将针对至少两个目标对象构建的至少两个时空图划分为多个时空图子集, 并从多个时空图子集中确定出终选子集; 将终选子集所包含的时空图之间的关系所指示的、目标对象之间的动作类别确定为视频片段所包含的的动作的动作类别。