

(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

(43) 国际公布日  
2023 年 4 月 20 日 (20.04.2023)



(10) 国际公布号  
WO 2023/061102 A1

(51) 国际专利分类号:  
G06V 20/40 (2022.01) G06V 10/40 (2022.01)  
G06V 10/80 (2022.01) G06K 9/62 (2022.01)

厦 35 层, Guangdong 518057 (CN)。 陈宇 (CHEN, Yu); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(21) 国际申请号: PCT/CN2022/116947

(74) 代理人: 华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房 (部位: 自编 01-03 和 08-12 单元) (仅限办公用途), Guangdong 510623 (CN)。

(22) 国际申请日: 2022 年 9 月 5 日 (05.09.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:  
202111202734.4 2021 年 10 月 15 日 (15.10.2021) 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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, 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。

(71) 申请人: 腾讯科技 (深圳) 有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 胡益晖 (HU, Yihui); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 杨伟东 (YANG, Weidong); 中国广东省深圳市南山区高新区科技中一路腾讯大

(54) Title: VIDEO BEHAVIOR RECOGNITION METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 视频行为识别方法、装置、计算机设备和存储介质

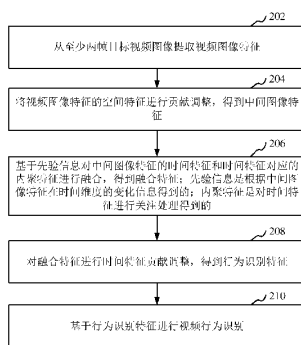


图 2

- 202 Extract video image features from at least two frames of target video images
- 204 Perform contribution adjustment on spatial features of the video image features, so as to obtain an intermediate image feature
- 206 Fuse, on the basis of prior information, a time feature of the intermediate image feature with a cohesion feature corresponding to the time feature, so as to obtain a fused feature, wherein the prior information is obtained according to change information of the intermediate image feature in a time dimension, and the cohesion feature is obtained by performing attention processing on the time feature
- 208 Perform time feature contribution adjustment on the fused feature, so as to obtain a behavior recognition feature
- 210 Perform video behavior recognition on the basis of the behavior recognition feature

(57) Abstract: A video behavior recognition method, which is executed by a computer device, and comprises: extracting video image features from at least two frames of target video images (202); performing contribution adjustment on spatial features of the video image features, so as to obtain an intermediate image feature (204); fusing, on the basis of prior information, a time feature of the intermediate image feature with a cohesion feature corresponding to the time feature, so as to obtain a fused feature, wherein the prior information is obtained according to change information of the intermediate image feature in a time dimension, and the cohesion feature is obtained by performing attention processing on the time feature (206); performing time feature contribution adjustment on the fused feature, so as to obtain a behavior recognition feature (208); and performing video behavior recognition on the basis of the behavior recognition feature (210).

(57) 摘要: 一种视频行为识别方法, 由计算机设备执行, 包括: 从至少两帧目标视频图像提取视频图像特征(202); 将视频图像特征的空间特征进行贡献调整, 得到中间图像特征(204); 基于先验信息对中间图像特征的时间特征和时间特征对应的内聚特征进行融合, 得到融合特征; 先验信息是根据中间图像特征在时间维度的变化信息得到的; 内聚特征是对时间特征进行关注处理得到的(206); 对融合特征进行时间特征贡献调整, 得到行为识别特征(208); 及基于行为识别特征进行视频行为识别(210)。

[见续页]

(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))。