

(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 2022/257578 A1

(43) 国际公布日
2022 年 12 月 15 日 (15.12.2022)

(51) 国际专利分类号:
G06V 20/62 (2022.01)

(21) 国际申请号: PCT/CN2022/085317

(22) 国际申请日: 2022 年 4 月 6 日 (06.04.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110632180.5 2021年6月7日 (07.06.2021) CN

(71) 申请人: 京 东 科 技 信 息 有 限 公 司 (JINGDONG TECHNOLOGY INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。

(72) 发明人: 陶大程(TAO, Dacheng); 中国北京市北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, Beijing 100176 (CN)。 何玥(HE, Yue); 中国北京市北京经济技术开发区科创十一街 18 号院 2 号楼 6 层 601, 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, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC,

(54) Title: METHOD FOR RECOGNIZING TEXT, AND APPARATUS

(54) 发明名称: 用于识别文本的方法和装置

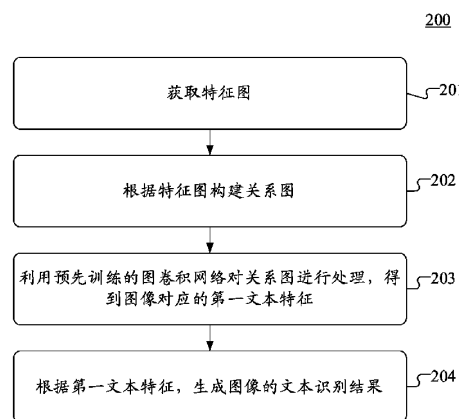


图 2

- 201 Obtain feature map
202 Construct relationship graph according to feature map
203 Utilize pre-trained graph convolutional network to perform processing on relationship graph, and obtain first text feature corresponding to image
204 Generate text recognition result for image according to first text feature

(57) Abstract: A method for recognizing a text, and an apparatus. A specific embodiment of the method comprises: obtaining a feature map (201), the feature map being obtained by means of performing text instance segmentation on an image having a text to be recognized; constructing a relationship graph according to the feature map (202), wherein in the relationship graph, each node represents a pixel in the feature map, and each edge represents that a similarity measure of spatial semantic features of two connected nodes is greater than a target threshold, and spatial semantic features comprise a type feature and a position feature of a pixel represented by

[见续页]

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))。

a node; utilizing a pre-trained graph convolutional network to perform processing on the relationship graph, and obtaining a first text feature corresponding to the image (203); and generating a text recognition result for the image according to the first text feature (204). The present embodiment provides a graph-based method for recognizing a text.

(57) 摘要: 用于识别文本的方法和装置, 包括: 获取特征图(201), 特征图通过对呈现有待识别文本的图像进行文本实例分割得到; 根据特征图构建关系图(202), 其中, 关系图中的节点表示特征图中的像素点, 边表示所连接的两个节点的空间语义特征的相似度大于目标阈值, 空间语义特征包括节点指示的像素点的位置特征和类别特征; 利用预先训练的图卷积网络对关系图进行处理, 得到图像对应的第一文本特征(203); 根据第一文本特征, 生成图像的文本识别结果(204)。提出了一种基于图的文本识别方法。