

(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
2016 年 7 月 7 日 (07.07.2016)



(10) 国际公布号
WO 2016/107103 A1

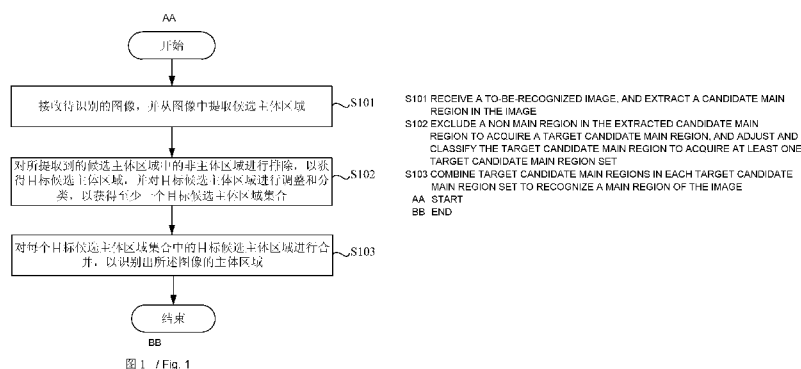
- (51) 国际专利分类号:
G06K 9/62 (2006.01)
- (21) 国际申请号: PCT/CN2015/082049
- (22) 国际申请日: 2015 年 6 月 23 日 (23.06.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410850511.2 2014 年 12 月 30 日 (30.12.2014) CN
- (71) 申请人: 百度在线网络技术(北京)有限公司
(BAIDU ONLINE NETWORK TECHNOLOGY
(BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
- (72) 发明人: 陈世佳 (CHEN, Shijia); 中国北京市海淀区
上地十街 10 号百度大厦三层, Beijing 100085 (CN)。
李旭斌 (LI, Xubin); 中国北京市海淀区上地十街 10
号百度大厦三层, Beijing 100085 (CN)。 文石磊
(WEN, Shilei); 中国北京市海淀区上地十街 10 号百
度大厦三层, Beijing 100085 (CN)。 余轶南 (YU,
Yinan); 中国北京市海淀区上地十街 10 号百度大厦
三层, Beijing 100085 (CN)。

- (74) 代理人: 北京清亦华知识产权代理事务所(普通
合伙) (TSINGYIHUA INTELLECTUAL PROP-
ERTY LLC); 中国北京市海淀区清华园清华大学照
澜院商业楼 301 室, Beijing 100084 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,
JP, KE, KG, KN, KP, KR, 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, SM, ST,
SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, 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)。

[见续页]

(54) Title: METHOD AND DEVICE FOR RECOGNIZING MAIN REGION OF IMAGE

(54) 发明名称: 图像主体区域的识别方法及装置



(57) Abstract: The present invention provides a method and device for recognizing a main region of an image. The method comprises: receiving a to-be-recognized image, and extracting a candidate main region of the image; excluding a non main region in the extracted candidate main region to acquire a target candidate main region, and adjusting and classifying the target candidate main region to acquire at least one target candidate main region set; and combining target candidate main regions in each target candidate main region set to recognize a main region of the image. By means of method and device for recognizing a main region of an image in embodiments of the present invention, a main region of an image can be accurately acquired, and main regions corresponding to multiple main bodies in an image containing multiple main bodies can be recognized.

(57) 摘要: 本发明提供了一种图像主体区域的识别方法及装置, 其中, 该方法包括: 接收待识别的图像, 并从图像中提取候选主体区域; 对所提取到的候选主体区域中的非主体区域进行排除, 以获得目标候选主体区域, 并对目标候选主体区域进行调整和分类, 以获得至少一个目标候选主体区域集合; 以及对每个目标候选主体区域集合中的目标候选主体区域合并, 以识别出图像的主体区域。本发明实施例的提供的图像主体区域的识别方法及装置, 可以准确地获得图像的主体区域, 且可以从包含多个主体的图像中识别出多个主体对应的主体区域。



WO 2016/107103 A1



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

图像主体区域的识别方法及装置

相关申请的交叉引用

- 5 本申请要求百度在线网络技术（北京）有限公司于 2014 年 12 月 30 日提交的、发明名称为“图像主体区域的识别方法及装置”的、中国专利申请号“201410850511.2”的优先权。

技术领域

本发明涉及计算机技术领域，尤其涉及一种图像主体区域的识别方法及装置。

10

背景技术

- 目前，图像信息的提取主要集中在三种粒度层面，其一是以图像整体为单位的粗粒度图像分类标注，其二是以获得图像中实体目标（例如行人、人脸、汽车等）为目的的目标检测方式，该方式需要训练专门类别的检测器进行检测。其三是精细到图像像素的细粒图像分割或者剖析，图像分割或者剖析能够对图像主体细节进行像素级的粒度描述，然而，上述三种
- 15 粒度层面的图像处理方式存在的问题是，图像分类标注不能准确和全面的分析包含多个主体的图像，所获得的分析结果不稳定；目标检测方式对于包含多个类别的目标主体时需要对图像进行多次遍历处理，所需计算量大；图像分割或者剖析所需分析时间较长，可在例如人脸、服饰、皮肤、箱包分割等特定场景上使用。

- 20 随着互联网技术的发展，上述单一粒度的图像处理的方式在处理效率上难以满足更加多样化的网页（web）图片、社交（social）图片等的要求，为了适应时代的发展，现有的对图像进行处理的过程为：从图像中检测出主体所在的主体区域，然后基于主体区域对主体进行后续分析，例如，可对主体进行分类或者识别。在这个整个过程中获得图像中的主体区域是获得图像准确分析结果的关键，目前常用的图像主体区域检测方式有两种，第一种方式为基
- 25 于显著性区域的主体检测方式，该方式通过显著性能量函数对图像进行处理，以生成图像的能量分布图，并根据能量分布图获得显著性较高的区域，并将该区域作为图像的主体区域。第二种方式为基于深度学习的主体坐标回归，该方式获得待处理的图像，然后通过已训练好的深度神经网络预测主体外接矩形的四个点坐标，实现主体区域检测定位。

- 然而，上述两种方式存在的问题是：第一种方式所使用的显著性能量函数不具有通用性，
- 30 主体区域检测结果的准确性和有效性不足，第二种方式不仅计算复杂，时效性效果不佳，而

且该方式只能返回一个主体区域，很难处理包含多个主体的图像。

发明内容

本发明旨在至少在一定程度上解决相关技术中的技术问题之一。为此，本发明第一方面
5 实施例在于提出一种图像主体区域的识别方法，该方法可以准确地获得图像的主体区域，且
可以从包含多个主体的图像中识别出多个主体对应的主体区域。

本发明的第二方面实施例在于提出一种图像主体区域的识别装置。

本发明的第三方面实施例在于提出一种存储介质。

本发明的第四方面实施例在于提出一种图像主体区域的识别设备。

10 为了实现上述目的，本发明第一方面实施例的图像主体区域的识别方法，包括：接收待
识别的图像，并从所述图像中提取候选主体区域；对所提取到的候选主体区域中的非主体区
域进行排除，以获得目标候选主体区域，并对所述目标候选主体区域进行调整和分类，以获
得至少一个目标候选主体区域集合；以及对每个目标候选主体区域集合中的目标候选主体区
域合并，以识别出所述图像的主体区域。

15 根据本发明实施例的图像主体区域的识别方法，首先从待识别图像中提取候选主体区
域，然后对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，
并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合，最后对每
个目标候选主体区域集合中的目标候选主体区域进行合并，以识别出图像的主体区域，由此，
可以准确地获得图像的主体区域，且可以从包含多个主体的图像中识别出多个主体对应的主
20 体区域。

为了实现上述目的，本发明第二方面实施例的图像主体区域的识别装置，包括：提取模
块，用于接收待识别的图像，并从所述图像中提取候选主体区域；处理模块，用于对所提取
到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对所述目标候选
主体区域进行调整和分类，以获得至少一组目标候选主体区域集合；以及合并模块，用于对
25 每个目标候选主体区域集合中的目标候选主体区域合并，以识别出所述图像的主体区域。

根据本发明实施例的图像主体区域的识别装置，通过提取模块从待识别图像中提取候选
主体区域，然后通过处理模块对所提取到的候选主体区域中的非主体区域进行排除，以获得
目标候选主体区域，并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体
区域集合，以及对每个目标候选主体区域集合中的目标候选主体区域进行合并，以识别出图
30 像的主体区域，由此，可以准确地获得图像的主体区域，且可以从包含多个主体的图像中识
别出多个主体对应的主体区域。

为了实现上述目的，本发明第三方面实施例的存储介质，用于存储应用程序，所述应用

程序用于执行本发明第一方面实施例所述的图像主体区域的识别方法。

为了实现上述目的，本发明第四方面实施例的图像主体区域的识别设备，包括：一个或者多个处理器；存储器；一个或者多个模块，所述一个或者多个模块存储在所述存储器中，当被所述一个或者多个处理器执行时进行如下操作：接收待识别的图像，并从所述图像中提取候选主体区域；对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对所述目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合；以及对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出所述图像的主体区域。

附图说明

图 1 是根据本发明一个实施例的图像主体区域的识别方法的流程图。

图 2 是根据本发明一个实施例的图像主体区域的识别方法的流程图。

图 3 是未经过处理的原图的示例图。

图 4 是对图 3 提取候选主体区域后的结果的示例图。

图 5 是对图 4 中的候选主体区域排除非主体区域后的结果的示例图。

图 6 是对图 5 中的候选主体区域进行调整和分类后的结果的示例图。

图 7 是对图 6 中的目标候选主体区域进行合并后的结果的示例图。

图 8 是根据本发明一个实施例的图像主体区域的识别装置的结构示意图。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

下面参考附图描述本发明实施例的图像主体区域的识别方法及装置。

图 1 是根据本发明一个实施例的图像主体区域的识别方法的流程图，如图 1 所示，该图像主体区域的识别方法包括：

S101，接收待识别的图像，并从图像中提取候选主体区域。

具体地，可根据通用的主体检测模型从待识别的图像中提取不同大小尺寸、长宽比的候选主体区域，为了可以使得候选主体区域可以涵盖待识别的图像中各种类别、不同大小、不同尺寸的主体内容，通常通用的主体区域检测模型可以获得图像的大量的候选主体区域，候选主体区域在图像中通常用矩形框标示。

其中，通用的主体检测模型是一种提取图像的候选主体区域的模型，该模型可根据预先设定的图像特征（例如像素灰度、颜色、图像梯度、边缘、方向梯度直方图 HOG（Histogram of Oriented Gradient）特征、局部二值模式 LBP（Local Binary Pattern）特征等）对图像的特征进行提取，在提取待识别的图像后，可通过图像的机器学习算法（例如支持向量机 svm
5 （support vector machine）、Adaboost 算法（该算法是一种自适应集成学习算法）、随机森林（random forest）算法等）对所获得的特征进行分析，以获得待识别图像对应的候选主体区域。需要说明的是，该模型提取图像的候选主体区域的效率高，一般情况下，提取一种图像的候选主体区域所需的时间为 100ms。

10 S102，对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合。

在该实施例中，在获得待识别图像的大量的候选主体区域后，可使用不同级别的分类器对所提取到的候选主体区域进行识别，以排除候选主体区域中的非主体区域。

具体地，在获得待识别图像的大量的候选主体区域后，为了可以获得准确获得图像的主体区域，可将获得的候选主体区域通过级联的不同级别的分类器进行识别，逐步排除候选主体区域中的非主体区域。其中，不同级别的分类器具有不同的判定准则，可以通过所有级别的分类器的候选主体区域成为图像的主体区域的可能性很高。
15

例如，假定当前级联了三个分类器，在获得当前图像对应的大量的候选主体区域后，可将所获得的候选主体区域输入第一级深度学习分类器，第一级深度学习分类器基于其对应的判断准则对候选主体区域进行识别，排除不满足主体区域特性的非主体区域，并将获得的识别结果输入第二级深度学习分类器，第二级深度学习分类器基于其对应的判断准则再排除一些不满足主体区域特性的非主体区域，并输出对应的识别结果，在获得第二级深度学习分类器的识别结果后，可将第二级深度学习分类器所获得的识别结果再输入至第三级线性分类器，其中，该线性分类器也可以称为支持向量机（Support Vector Machine，简称 SVM），第三级线性分类器基于其对应的判断准则对输入的识别结果再进行识别，并将其不满足其判断
20 准则的非主体区域去除，以及输出对应的识别结果，所输出的识别结果为所有通过上述三级分类器的候选主体区域。在这里，为了可以快速过滤大量非主体，在前期可使用低级别的分类器对大量的候选主体区域进行粗筛选，后期为了可以准确排除候选主体区域集合中的少量难区分的非主体区域，可使用高级别的分类器进行识别，以获得准确对应的目标候选主体区域。其中，通过所有级别的分类器的候选主体区域即为目标候选主体区域。

30 在获得目标候选主体区域后，为了可以准确获得图像的主体区域，可针对每个目标候选主体区域，根据主体和当前目标候选主体区域之间的大小关系和位置关系对当前目标候选主体区域进行缩放处理和/或平移处理。

在对每个目标候选主体区域进行调整后,可获得调整前和调整后的每个目标候选主体区域的评分和所属的类别,并根据同一类别中目标候选主体区域的评分对对应的目标候选主体区域进行排序,并根据排序结果获得目标候选主体区域集合。

例如,可将获得的调整前和调整后的目标候选主体区域通过一个基于深度学习模型的分
5 类器,该分类器对每个目标候选主体区域进行分类,并输出每个目标候选主体区域属于各个类别的概率值,假定调整前和调整后的目标候选主体区域总共有四个,用 a、b、c 和 d 分别表示这四个目标候选主体区域,并且当前分类器有三个类别,分别用 l、m、n 表示,通过分类器所获得的各个目标候选主体区域对应的三种类别 l、m、n 的概率分布为: a 为 (0.1, 0.1, 0.8), b 为 (0.2, 0.2, 0.6), c 为 (0.7, 0.1, 0.2), d 为 (0.1, 0.85, 0.05), 在获得对应的
10 概率后,针对每个类别,可根据每个目标候选主体区域的概率对其进行排序,并根据排序结果获得目标候选主体区域对应的类别。其中,对上述四个目标候选主体区域 a、b、c 和 d 进行处理后,所获得的结果为: a 和 b 属于类别 n, 且 a 的概率高(置信高), 排于 b 之前, c 属于类别 l, d 属于类别 m。

S103, 对每个目标候选主体区域集合中的目标候选主体区域进行合并,以识别出图像的
15 主体区域。

在该实施例,在获得图像对应的至少一个目标候选主体区域集合后,针对每个目标候选主体区域集合,可根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。

例如,假定当前获得两类目标候选主体区域集合,针对每个目标候选主体区域集合,可
20 根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并,具体地,在对于同一类别的目标候选主体区域进行合并的过程中,若两个目标候选主体区域交叠,可按照两者的外接矩形进行合并,另外,在合并的过程中,也可以使用主体框聚类算法、基于 CRF(Conditional Random Fields, 条件随机场)模型的主体框分类划分算法等机器学习算法来对目标候选主体区域进行合并,以获得两种类别的主体在图像中对应的主体区域。

25 该实施例的图像主体区域检测方式可以准确获得图像中对应主体的主体区域,并且可以获得主体区域在图像中的位置关系,方便了后续基于所获得的主体区域对图片进行分类、识别、裁剪、搜索等,另外,还可以根据图像的主体区域对图像进行裁剪,并将裁剪后的图片应用在搜索结果页中,以在搜索结果页中图文并茂的展示搜索结果。

本发明实施的图像主体区域的识别方法,首先从待识别图像中提取候选主体区域,然后
30 对所提取到的候选主体区域中的非主体区域进行排除,以获得目标候选主体区域,并对目标候选主体区域进行调整和分类,以获得至少一个目标候选主体区域集合,最后对每个目标候选主体区域集合中的目标候选主体区域进行合并,以识别出图像的主体区域,由此,可以准

确地获得图像的主体区域，且可以从包含多个主体的图像中识别出多个主体对应的主体区域。

图 2 是根据本发明另一个实施例的图像主体区域的识别方法的流程图，该实施例以图 3 所示的图片为例详细描述图像主体区域的识别过程，如图 2 所示，该图像主体区域的识别方法包括：

S201，从图像中提取候选主体区域。

具体地，可根据通用的主体检测模型从图 3 中提取不同大小尺寸、长宽比的候选主体区域，对图 3 提取候选主体区域后的结果如图 4 所示，图 4 中用矩形框即表示候选主体区域，图 4 的每一个矩形框表示一个候选主体区域。

其中，通用的主体检测模型是一种提取图像的候选主体区域的模型，该模型提取图像的候选主体区域的效率高，一般情况下，提取一种图像的候选主体区域所需的时间为 100ms。

S202，对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域。

通过图 4 可以看出候选主体区域中包含大量的候选主体区域，为了可以准确获得图像的主体区域，可通过级联的不同级别的分类器对所提取到的大量的候选主体区域集合进行识别，并逐步排除候选主体区域中的非主体区域，以获得目标候选主体区域。其中，通过所有级别的分类器的候选主体区域成为图像的主体区域的可能性比较大，通过所有级别的分类器的候选主体区域可称为目标候选主体区域。

具体而言，在处理图 4 的过程中，通过图 4 可以看出，该图中包含两个主体，因此，可通过多级二分类的分类器对两个主体对应的候选主体区域中的非主体区域进行排除，图 4 的候选主体区域排除非主体区域后，所获得的结果如图 5，通过比较图 4 与图 5 可以看出，图 5 中仅包含了几个可能成为图像的主体区域的候选主体区域（即目标候选主体区域）。

S203，对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合。

在该实施例中，在获得目标候选主体区域后，为了可以准确获得图像的主体区域，针对每个目标候选主体区域，可根据主体和当前目标候选主体区域之间的大小关系和位置关系对当前目标候选主体区域进行缩放处理和/或平移处理。

通过图 5 可以看出，图 5 中所包含的目标候选主体区域中与图像的主体区域之间还存在一些偏差，为了可以准确获得图像的主体区域，可对图 5 中的目标候选主体区域进行调整，具体而言，针对图 5 中的每个目标候选主体区域，可根据目标候选主体区域与主体之间的大小关系和位置关键进行缩放、平移等以生成新的目标候选主体区域。

在对每个目标候选主体区域进行调整后，可获得调整前和调整后的每个目标候选主体区域的评分和类别，并根据同一类别中目标候选主体区域的评分对对应的目标候选主体区域进

行排序,并根据排序结果获得目标候选主体区域集合。其中,对图 5 中的目标候选主体区域进行调整,然后在通过分类器例如基于深度学习模型的分器对调整前和调整后的候选主体区域进行处理,处理后所获的结果如图 6 所示,图 6 示出了所获得的多个目标候选主体区域。

5 S204,对每个目标候选主体区域集合中的目标候选主体区域进行合并,以识别出图像的主体区域。

在该实施例中,在获得至少一个目标候选主体区域集合后,针对每个目标候选主体区域集合,针对每个目标候选主体区域集合,可根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。

10 对于图 6 中的目标候选主体区域,由于图 6 中包含两个主体,因此,对于图 6 来说,可以获得两个目标候选主体区域集合,在获得两个目标候选主体区域集合后,针对每个目标候选主体区域集合,可根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。具体地,在对图 6 中的目标主体区域进行合并的过程中,针对不同的主体,可分别对主体周围对应的目标候选主体区域进行合并,具体而言,可根据目标候选主体区域之间的大小关系和位置关系进行合并,若两个目标候选主体区域交叠,可按照两者的外接矩形进行合并,对图 6 中的目标候选主体区域进行合并后,所获得的图像的主体区域如图 7 所示,图 7 中所示的两个矩形框即为图 3 的主体区域。

15 在获得图 3 的主体区域后,可根据所获得的主体区域对图片进行描述、分类、识别、裁剪展示等,另外,还可以基于所获得的主体区域还可以根据图像的主体区域对图像进行裁剪,并将裁剪后的图片应用在搜索结果页中,以在搜索结果页中图文并茂的展示搜索结果。

20 本发明实施的图像主体区域的识别方法,首先从待识别图像中提取候选主体区域,然后对所提取到的候选主体区域中的非主体区域进行排除,以获得目标候选主体区域,并对目标候选主体区域进行调整和分类,以获得至少一个目标候选主体区域集合,最后对每个目标候选主体区域集合中的目标候选主体区域进行合并,以识别出图像的主体区域,由此,可以准确地获得图像的主体区域,且可以从包含多个主体的图像中识别出多个主体对应的主体区域。

为了实现上述实施例,本发明还提出一种图像主体区域的识别装置。

图 8 是根据本发明一个实施例的图像主体区域的识别装置的结构示意图。

30 如图 8 所示,该图像主体区域的识别装置包括:提取模块 100、处理模块 200 和合并模块 300,其中:

提取模块 100 用于接收待识别的图像,并从图像中提取候选主体区域集;处理模块 200

用于对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合；以及合并模块 300 用于对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出图像的主体区域。

5 处理模块 200 具体用于：针对候选主体区域集合中的每个候选主体区域，使用不同级别的分类器对当前候选主体区域进行识别，以排除非主体区域。

具体地，处理模块 200 可通过级联的不同级别的分类器对当前候选主体区域进行识别，以排除候选主体区域中的非主体区域，以获得目标候选主体区域。

10 在处理模块 200 获得目标候选主体区域后，针对每个目标候选主体区域，处理模块 200 可根据主体和当前目标候选主体区域之间的大小关系和位置关系对当前目标候选主体区域进行缩放处理和/或平移处理。

在处理模块 200 对目标候选主体区域进行调整后，处理模块 200 可获得调整前和调整后的每个目标候选主体区域的评分和所属的类别，并根据同一类别中目标候选主体区域的评分对对应的目标候选主体区域进行排序，并根据排序结果获得目标候选主体区域集合。

15 在处理模块 200 获得至少一个目标候选主体区域集合后，针对每个目标候选主体区域集合，合并模块 300 可根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。

包含提取模块 100、处理模块 200 和合并模块 300 的图像主体区域的识别装置识别图像主体区域的过程可参见图 1 或者图 2 以及对应的文字描述，此处不赘述。

20 本发明实施的图像主体区域的识别装置，通过提取模块从待识别图像中提取候选主体区域，然后通过处理模块对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合，以及对每个目标候选主体区域集合中的目标候选主体区域进行合并，以识别出图像的主体区域，由此，可以准确地获得图像的主体区域，且可以从包含多个主体的图像中识别出多个主体对应的主体区域。

为了实现上述目的，本发明还提出了一种存储介质，用于存储应用程序，该应用程序用于执行本发明任一实施例所述的图像主体区域的识别方法。

30 为了实现上述目的，本发明还提出了一种图像主体区域的识别设备，包括：一个或者多个处理器；存储器；一个或者多个模块，所述一个或者多个模块存储在所述存储器中，当被所述一个或者多个处理器执行时进行如下操作：

S101'，接收待识别的图像，并从所述图像中提取候选主体区域。

S102'，对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，

并对目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合。

S103'，对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出图像的主体区域。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、
5 或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

10 此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中，“多个”的含义是至少两个，例如两个，三个等，除非另有明确具体的限定。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或
15 更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分，并且本发明的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能，这应被本发明的实施例所属技术领域的技术人员所理解。

在流程图中表示或在此以其他方式描述的逻辑和/或步骤，例如，可以被认为是用于实现
20 逻辑功能的可执行指令的定序列表，可以具体实现在任何计算机可读介质中，以供指令执行系统、装置或设备（如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统）使用，或结合这些指令执行系统、装置或设备而使用。就本说明书而言，“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机
25 可读介质的更具体的示例（非穷尽性列表）包括以下：具有一个或多个布线的电连接部（电子装置），便携式计算机盘盒（磁装置），随机存取存储器（RAM），只读存储器（ROM），可擦除可编程只读存储器（EPROM 或闪速存储器），光纤装置，以及便携式光盘只读存储器（CDROM）。另外，计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质，因为可以例如通过对纸或其他介质进行光学扫描，接着进行编辑、解译或必要时以其他
30 合适方式进行处理来以电子方式获得所述程序，然后将其存储在计算机存储器中。

应当理解，本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件

来实现。例如，如果用硬件来实现，和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列（PGA），现场可编程门阵列（FPGA）等。

- 5 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本发明各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以
10 采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用，也可以存储在一个计算机可读取存储介质中。

上述提到的存储介质可以是只读存储器，磁盘或光盘等。尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域
15 的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

权 利 要 求 书

1、一种图像主体区域的识别方法，其特征在于，包括：

接收待识别的图像，并从所述图像中提取候选主体区域；

5 对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对所述目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合；以及
对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出所述图像的主体区域。

2、根据权利要求 1 所述的方法，其特征在于，所述对所提取到的候选主体区域中的非
10 主体区域进行排除，包括：

使用不同级别的分类器对所提取到的候选主体区域进行识别，以排除所述非主体区域。

3、根据权利要求 1 或 2 所述的方法，其特征在于，所述对所述目标候选主体区域进行调整，包括：

针对每个目标候选主体区域，根据主体和当前目标候选主体区域之间的大小关系和位置
15 关系对所述当前目标候选主体区域进行缩放处理和/或平移处理。

4、根据权利要求 1-3 任一项所述的方法，其特征在于，所述对所述目标候选主体区域进行分类，包括：

获得调整前和调整后的每个目标候选主体区域的评分和所属的类别，并根据同一类别中
目标候选主体区域的评分对对应的目标候选主体区域进行排序，并根据排序结果获得所述目
20 标候选主体区域集合。

5、根据权利要求 1-4 任一项所述的方法，其特征在于，所述对每个目标候选主体区域集合中的目标候选主体区域进行合并，以识别出所述图像的主体区域，包括：

根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。

6、一种图像主体区域的识别装置，其特征在于，包括：

25 提取模块，用于接收待识别的图像，并从所述图像中提取候选主体区域；

处理模块，用于对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对所述目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合；以及

合并模块，用于对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出所
30 述图像的主体区域。

7、根据权利要求 6 所述的装置，其特征在于，所述处理模块，具体用于：

使用不同级别的分类器对所提取到的候选主体区域进行识别，以排除所述非主体区域。

8、根据权利要求 6 或 7 所述的装置，其特征在于，所述处理模块，具体用于：

针对每个目标候选主体区域，根据主体和当前目标候选主体区域之间的大小关系和位置关系对所述当前目标候选主体区域进行缩放处理和/或平移处理。

9、根据权利要求 6-8 任一项所述的装置，其特征在于，所述处理模块，还用于：

5 获得调整前和调整后的每个目标候选主体区域的评分和所属的类别，并根据同一类别中目标候选主体区域的评分对对应的目标候选主体区域进行排序，并根据排序结果获得所述目标候选主体区域集合。

10、根据权利要求 6-9 任一项所述的装置，其特征在于，所述合并模块，具体用于：

根据目标候选主体区域之间的大小关系及位置关系对目标候选主体区域进行合并。

10 11、一种存储介质，其特征在于，用于存储应用程序，所述应用程序用于执行权利要求 1 至 5 中任一项所述的图像主体区域的识别方法。

12、一种图像主体区域的识别设备，其特征在于，包括：

一个或者多个处理器；

存储器；

15 一个或者多个模块，所述一个或者多个模块存储在所述存储器中，当被所述一个或者多个处理器执行时进行如下操作：

接收待识别的图像，并从所述图像中提取候选主体区域；

对所提取到的候选主体区域中的非主体区域进行排除，以获得目标候选主体区域，并对所述目标候选主体区域进行调整和分类，以获得至少一个目标候选主体区域集合；以及

20 对每个目标候选主体区域集合中的目标候选主体区域合并，以识别出所述图像的主体区域。

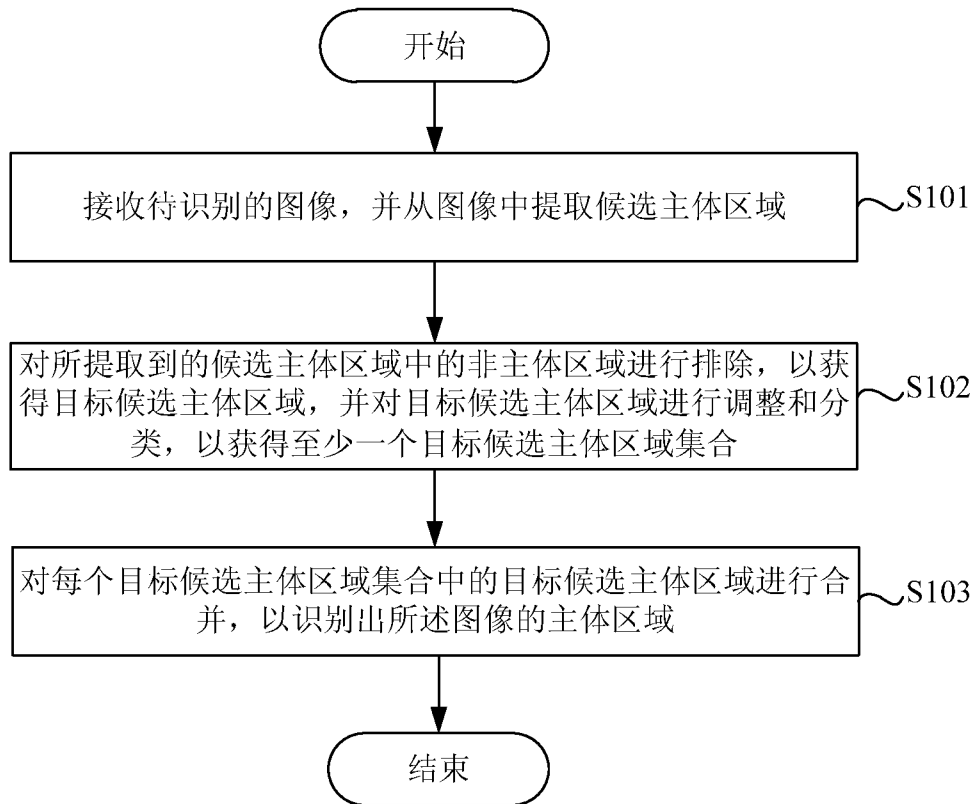


图 1

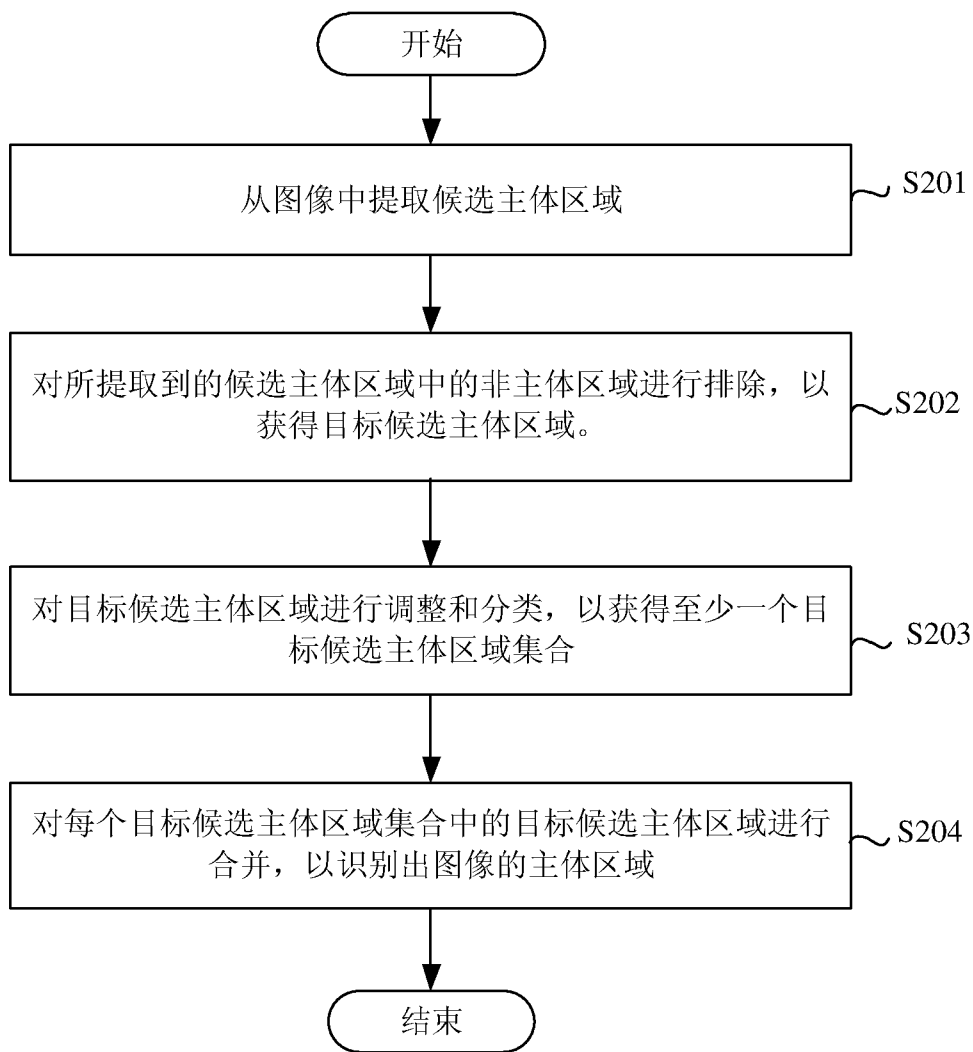


图 2



图 3

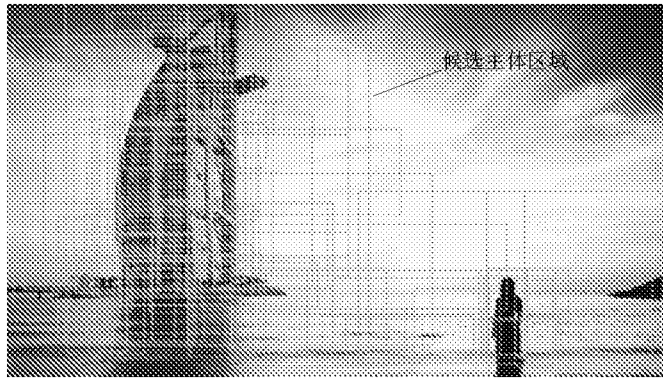


图 4

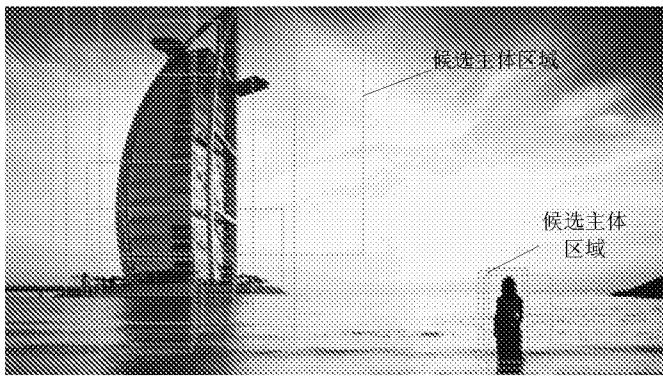


图 5



图 6

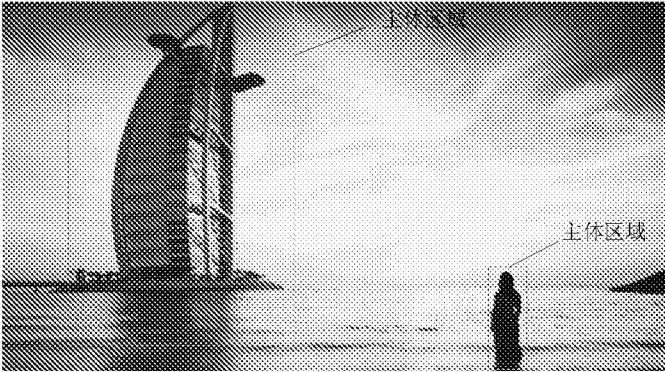


图 7



图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/082049**A. CLASSIFICATION OF SUBJECT MATTER**

G06K 9/62 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K 9/-, G06F 17/-, G06T 7/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

VEN, DWPI, CNABS, CNKI: main body, combine, translation, image, area, region, identif+, recognition, recogniz+, classify, extract, zoom

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 104573715 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 29 April 2015 (29.04.2015), claims 1-10	claims 1-10
X	CN 102013021 A 13 April 2011 (13.04.2011), description, paragraphs 0007-0039	claims 1-12
A	CN 101894361 A (YUNNAN UNIVERSITY), 24 November 2010 (24.11.2010), the whole document	claims 1-12
A	CN 102622761 A (XIDIAN UNIVERSITY), 01 August 2012 (01.08.2012), the whole document	claims 1-12

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 16 September 2015 (16.09.2015)	Date of mailing of the international search report 25 September 2015 (25.09.2015)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer TIAN, Jing Telephone No.: (86-10) 62089417

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2015/082049

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104573715 A	29 April 2015	None	
CN 102013021 A	13 April 2011	CN 102013021 B	31 October 2012
CN 101894361 A	24 November 2010	CN 101894361 B	20 February 2013
CN 102622761 A	01 August 2012	CN 102622761 B	14 May 2014

A. 主题的分类		
G06K 9/62 (2006.01) i		
按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献 (标明分类系统和分类号)		
G06K 9/-, G06F 17/-, G06T 7/-		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))		
VEN, DWPI, CNABS, CNKI: 图像, 图象, 区域, 识别, 主体, 合并, 分类, 缩放, 平移, image, area, region, identif+, recognition, recogniz+, classify, extract, zoom,		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 104573715 A (百度在线网络技术北京有限公司) 2015年 4月 29日 (2015 - 04 - 29) 权利要求1-10	权利要求1-10
X	CN 102013021 A 2011年 4月 13日 (2011 - 04 - 13) 说明书第0007-0039段	权利要求1-12
A	CN 101894361 A (云南大学) 2010年 11月 24日 (2010 - 11 - 24) 全文	权利要求1-12
A	CN 102622761 A (西安电子科技大学) 2012年 8月 1日 (2012 - 08 - 01) 全文	权利要求1-12
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2015年 9月 16日		2015年 9月 25日
ISA/CN的名称和邮寄地址		授权官员
中华人民共和国国家知识产权局 (ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国		田竞
传真号 (86-10)62019451		电话号码 (86-10)62089417

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/082049

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	104573715	A	2015年 4月 29日	无	
CN	102013021	A	2011年 4月 13日	CN	102013021 B 2012年 10月 31日
CN	101894361	A	2010年 11月 24日	CN	101894361 B 2013年 2月 20日
CN	102622761	A	2012年 8月 1日	CN	102622761 B 2014年 5月 14日