

(12) 发明专利申请

(10) 申请公布号 CN 102395966 A

(43) 申请公布日 2012. 03. 28

(21) 申请号 201080016836. 0

(74) 专利代理机构 北京律盟知识产权代理有限  
责任公司 11287

(22) 申请日 2010. 04. 14

代理人 宋献涛

(30) 优先权数据

61/169, 295 2009. 04. 14 US

12/537, 520 2009. 08. 07 US

(51) Int. Cl.

G06F 17/30(2006. 01)

G06K 9/20(2006. 01)

G06K 9/00(2006. 01)

(85) PCT申请进入国家阶段日

2011. 10. 14

(86) PCT申请的申请数据

PCT/US2010/031059 2010. 04. 14

(87) PCT申请的公布数据

W02010/120901 EN 2010. 10. 21

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 约瑟夫·于惠·黄 张勇 李翔尊

丹佛德尔·秋乐·山河

若泽·理查多·多斯桑托斯

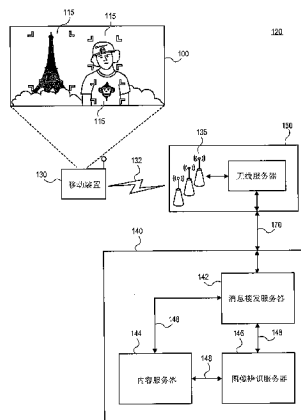
权利要求书 8 页 说明书 17 页 附图 6 页

(54) 发明名称

用于使用移动装置进行图像辨识的系统和方法

(57) 摘要

实施方案涉及用于实时的图像辨识和移动视觉搜索的系统和方法。例如蜂窝式电话等移动装置获取图像且预处理所述所获取的图像以基于所述所获取的图像中检测到的对象而产生视觉搜索查询。所述视觉搜索查询包含所述所获取的图像或从中提取的查询图像以及与所述检测到的对象相关联的元数据。所述移动装置将所述视觉搜索查询无线地传送到远程服务器，且响应于所述视觉搜索查询，所述远程服务器基于所述相关联的元数据而辨识所述查询图像中的对象。所述远程服务器随后基于所述所辨识的对象而产生信息内容，且将所述信息内容传送到所述移动装置以经由所述移动装置进行呈现。



1. 一种移动装置,其包括:  
到服务器的无线接口;以及  
处理器,其与所述无线接口通信,所述处理器经配置以起始处理,所述处理包括:  
获取图像;  
检测所述所获取的图像中的一个或一个以上对象;  
产生与所述检测到的对象中的至少一者相关联的元数据;  
基于所述至少一个对象从所述所获取的图像中提取查询图像;  
产生视觉搜索查询,其中所述视觉搜索查询包含所述查询图像和所述元数据;  
经由所述无线接口将所述视觉搜索查询传送到所述服务器;以及  
响应于所述视觉搜索查询而接收并呈现与所述至少一个对象相关联的信息内容。
2. 根据权利要求1所述的移动装置,其中产生元数据进一步包括:  
将所述检测到的对象中的一者或一者以上分类为一对象类别;  
突出显示所述经分类的对象;  
接收指示所述经分类的对象中的选定一者的输入;以及  
基于所述选定对象的所述对象类别而产生所述元数据。
3. 根据权利要求2所述的移动装置,其中将所述检测到的对象中的一者或一者以上分类进一步包括:  
基于所述检测到的对象中的至少一者而产生特征向量;  
将所述特征向量与所述对象类别中的训练图像的图像系数的集合进行比较,以确定所述对象类别中的与所述至少一个检测到的对象匹配的匹配的训练图像,其中图像系数的所述集合存储在所述移动装置中;以及  
基于所述匹配的训练图像而将所述检测到的对象分类。
4. 根据权利要求3所述的移动装置,其中所述处理器实时地执行所述获取、检测、突出显示和分类的操作。
5. 根据权利要求3所述的移动装置,其中所述对象类别包含标志、设计、面部、地标、服装、记号、自然对象或人造对象。
6. 根据权利要求1所述的移动装置,其中产生元数据进一步包括:  
突出显示所述检测到的对象;  
接收指示所述检测到的对象中的选定一者的输入;以及  
基于所述选定对象而产生所述元数据。
7. 根据权利要求6所述的移动装置,其中基于所述选定对象而产生所述元数据进一步包括:  
将所述选定对象分类为一对象类别;以及  
基于所述选定对象的所述对象类别而产生所述元数据。
8. 根据权利要求7所述的移动装置,其中将所述视觉搜索查询传送到所述服务器进一步包括:  
基于所述选定对象的所述对象类别而产生用于所述视觉搜索查询的目的地地址;以及  
根据所述目的地地址而将所述视觉搜索查询传送到所述服务器。
9. 根据权利要求1所述的移动装置,其中从所述所获取的图像中提取查询图像进一步

包括修剪所述所获取的图像、压缩所述所获取的图像、缩放所述所获取的图像,或  
将所述所获取的图像转换到灰度级。

10. 根据权利要求 1 所述的移动装置,其中产生元数据进一步包括:

基于所述检测到的对象中的至少一者而产生特征向量;

将所述特征向量与训练图像的图像系数的集合进行比较,以确定与所述检测到的对象匹配的匹配的训练图像,其中图像系数的所述集合存储在所述移动装置中;

基于所述匹配的训练图像而辨识所述检测到的对象;以及

基于所述匹配的训练图像而产生与所述所辨识的对象相关联的所述元数据。

11. 根据权利要求 1 所述的移动装置,其中所述移动装置包括上下文数据传感器,且其中获取图像进一步包括:

经由所述上下文数据传感器而获取与所述所获取的图像相关联的上下文数据,其中所述视觉搜索查询进一步包含所述上下文数据。

12. 根据权利要求 11 所述的移动装置,其中所述上下文数据包括全球定位系统 GPS 定位、辅助全球定位系统 A-GPS 定位、伽利略系统定位、塔式三边测量法定点、文本信息、听觉信息、加速计读数、陀螺仪读数,或温度读数。

13. 根据权利要求 1 所述的移动装置,其中所述信息内容包含名称、价格、制造商、评论、优惠券,或广告。

14. 一种执行图像辨识的方法,其包括:

通过移动装置获取图像;

检测所述所获取的图像中的一个或一个以上对象;

产生与所述检测到的对象中的至少一者相关联的元数据;

基于所述至少一个对象从所述所获取的图像中提取查询图像;

产生视觉搜索查询,其中所述视觉搜索查询包含所述查询图像和所述元数据;

无线地传送所述视觉搜索查询;以及

响应于所述视觉搜索查询而接收并呈现与所述至少一个对象相关联的信息内容。

15. 根据权利要求 14 所述的方法,其中产生元数据进一步包括:

将所述检测到的对象中的一者或一者以上分类为一对象类别;

突出显示所述经分类的对象;

接收指示所述经分类的对象中的选定一者的输入;以及

基于所述选定对象的所述对象类别而产生所述元数据。

16. 根据权利要求 15 所述的方法,其中将所述检测到的对象中的一者或一者以上分类进一步包括:

基于所述检测到的对象中的至少一者而产生特征向量;

将所述特征向量与所述对象类别中的训练图像的图像系数的集合进行比较,以确定所述对象类别中的与所述至少一个检测到的对象匹配的匹配的训练图像,其中图像系数的所述集合存储在所述移动装置中;以及

基于所述匹配的训练图像而将所述检测到的对象分类。

17. 根据权利要求 16 所述的方法,其中所述移动装置实时地执行所述获取、检测、突出显示和分类的操作。

18. 根据权利要求 14 所述的方法,其中产生元数据进一步包括:  
突出显示所述检测到的对象;  
接收指示所述检测到的对象中的选定一者的输入;以及  
基于所述选定对象而产生所述元数据。
19. 根据权利要求 18 所述的方法,其中基于所述选定对象而产生所述元数据进一步包括:  
将所述选定对象分类为一对象类别;以及  
基于所述选定对象的所述对象类别而产生所述元数据。
20. 根据权利要求 19 所述的方法,其中无线地传送所述视觉搜索查询进一步包括:  
基于所述选定对象的所述对象类别而产生用于所述视觉搜索查询的目的地地址;以及  
根据所述目的地地址而将所述视觉搜索查询传送到目的地。
21. 根据权利要求 14 所述的方法,其中获取图像进一步包括:  
经由所述移动装置的上下文数据传感器而获取与所述所获取的图像相关联的上下文数据,所述视觉搜索查询进一步包含所述上下文数据。
22. 一种用于执行图像辨识的系统,其包括:  
用于通过移动装置获取图像的装置;  
用于检测所述所获取的图像中的一个或一个以上对象的装置;  
用于产生与所述检测到的对象中的至少一者相关联的元数据的装置;  
用于基于所述至少一个对象从所述所获取的图像中提取查询图像的装置;  
用于产生视觉搜索查询的装置,其中所述视觉搜索查询包含所述查询图像和所述元数据;  
用于无线地传送所述视觉搜索查询的装置;以及  
用于响应于所述视觉搜索查询而接收并呈现与所述至少一个对象相关联的信息内容的装置。
23. 根据权利要求 22 所述的系统,其中所述用于产生元数据的装置进一步包括:  
用于将所述检测到的对象中的一者或一者以上分类为一对象类别的装置;  
用于突出显示所述经分类的对象的装置;  
用于接收指示所述经分类的对象中的选定一者的输入的装置;以及  
用于基于所述选定对象的所述对象类别而产生所述元数据的装置。
24. 根据权利要求 23 所述的系统,其中所述用于将所述检测到的对象中的一者或一者以上分类的装置进一步包括:  
用于基于所述检测到的对象中的至少一者而产生特征向量的装置;  
用于将所述特征向量与所述对象类别中的训练图像的图像系数的集合进行比较,以确定所述对象类别中的与所述至少一个检测到的对象匹配的匹配的训练图像的装置;以及  
用于基于所述匹配的训练图像而将所述检测到的对象分类的装置。
25. 根据权利要求 24 所述的系统,其中所述用于获取、检测、突出显示和分类的装置实时地执行。
26. 根据权利要求 22 所述的系统,其中所述用于产生元数据的装置进一步包括:  
用于突出显示所述检测到的对象的装置;

用于接收指示所述检测到的对象中的选定一者的输入的装置 ; 以及  
用于基于所述选定对象而产生所述元数据的装置。

27. 根据权利要求 26 所述的系统, 其中所述用于基于所述选定对象而产生所述元数据的装置进一步包括 :

用于将所述选定对象分类为一对象类别的装置 ; 以及  
用于基于所述选定对象的所述对象类别而产生所述元数据的装置。

28. 根据权利要求 27 所述的系统, 其中所述用于无线地传送所述视觉搜索查询的装置进一步包括 :

用于基于所述选定对象的所述对象类别而产生用于所述视觉搜索查询的目的地地址的装置 ; 以及

用于根据所述目的地地址而将所述视觉搜索查询传送到目的地的装置。

29. 根据权利要求 22 所述的系统, 其中所述用于获取图像的装置进一步包括 :

用于经由所述移动装置的上下文数据传感器而获取与所述所获取的图像相关联的上下文数据的装置, 所述视觉搜索查询进一步包含所述上下文数据。

30. 一种计算机程序产品, 其包含 :

计算机可读媒体, 所述计算机可读媒体存储在由计算机执行时致使所述计算机执行图像辨识的指令, 所述指令包括 :

用于致使所述计算机获取图像的至少一个指令 ;

用于致使所述计算机检测所述所获取的图像中的一个或一个以上对象的至少一个指令 ;

用于致使所述计算机产生与所述检测到的对象中的至少一者相关联的元数据的至少一个指令 ;

用于致使所述计算机基于所述至少一个对象而从所述所获取的图像中提取查询图像的至少一个指令 ;

用于致使所述计算机产生视觉搜索查询的至少一个指令, 其中所述视觉搜索查询包含所述查询图像和所述元数据 ;

用于致使所述计算机无线地传送所述视觉搜索查询的至少一个指令 ; 以及

用于致使所述计算机响应于所述视觉搜索查询而接收并呈现与所述至少一个对象相关联的信息内容的至少一个指令。

31. 根据权利要求 30 所述的计算机程序产品, 其中所述用于致使所述计算机产生元数据的至少一个指令进一步包括 :

用于致使所述计算机将所述检测到的对象中的一者或一者以上分类为一对象类别的至少一个指令 ;

用于致使所述计算机突出显示所述经分类的对象的至少一个指令 ;

用于致使所述计算机接收指示所述经分类的对象中的选定一者的输入的至少一个指令 ; 以及

用于致使所述计算机基于所述选定对象的所述选定对象类别而产生所述元数据的至少一个指令。

32. 根据权利要求 31 所述的计算机程序产品, 其中所述用于致使所述计算机将所述检

测到的对象中的一者或一者以上分类的至少一个指令进一步包括：

用于致使所述计算机基于所述检测到的对象中的至少一者而产生特征向量的至少一个指令；

用于致使所述计算机将所述特征向量与所述对象类别中的训练图像的图像系数的集合进行比较，以确定所述对象类别中的与所述至少一个检测到的对象匹配的匹配的所述训练图像的至少一个指令，其中图像系数的所述集合存储在所述计算机中；以及

用于致使所述计算机基于所述匹配的所述训练图像而将所述检测到的对象分类的至少一个指令。

33. 根据权利要求 32 所述的计算机程序产品，其中所述用于致使所述计算机获取、检测、突出显示和分类的至少一个指令致使所述计算机实时地获取、检测、突出显示和分类。

34. 根据权利要求 30 所述的计算机程序产品，其中所述用于致使所述计算机产生元数据的至少一个指令进一步包括：

用于致使所述计算机突出显示所述检测到的对象的至少一个指令；

用于致使所述计算机接收指示所述突出显示的对象中的选定一者的输入的至少一个指令；以及

用于致使所述计算机基于所述选定对象而产生所述元数据的至少一个指令。

35. 根据权利要求 34 所述的计算机程序产品，其中所述用于致使所述计算机基于所述选定对象而产生所述元数据的至少一个指令进一步包括：

用于致使所述计算机将所述选定对象分类为一对象类别的至少一个指令；以及

用于致使所述计算机基于所述选定对象的所述对象类别而产生所述元数据的至少一个指令。

36. 根据权利要求 35 所述的计算机程序产品，其中所述用于致使所述计算机无线地传送所述视觉搜索查询的至少一个指令进一步包括：

用于致使所述计算机基于所述选定对象的所述对象类别而产生用于所述视觉搜索查询的目的地址的至少一个指令；以及

用于致使所述计算机根据所述目的地地址而将所述视觉搜索查询传送到目的地的至少一个指令。

37. 根据权利要求 30 所述的计算机程序产品，其中所述用于致使所述计算机获取图像的至少一个指令进一步包括：

用于致使所述计算机经由所述移动装置的上下文数据传感器而获取与所述所获取的图像相关联的上下文数据的至少一个指令，所述视觉搜索查询进一步包含所述上下文数据。

38. 一种用于执行图像辨识的系统，其包括：

服务器，其经配置以：

从移动装置接收视觉搜索查询，其中所述视觉搜索查询包含图像和与所述图像中的至少一个对象相关联的元数据，

基于所述元数据而辨识与所述至少一个对象相关联的所述图像中的对象，

基于所述所辨识的对象而产生信息内容，以及

响应于所述视觉搜索查询而传送所述信息内容。

39. 根据权利要求 38 所述的系统,其中  
所述视觉搜索查询进一步包含与所述图像相关联的上下文数据;且  
进一步其中所述服务器经配置以基于所述所辨识的对象和所述上下文数据而产生所述信息内容。

40. 根据权利要求 39 所述的系统,其中所述上下文数据包括全球定位系统 GPS 定位、辅助式全球定位系统 A-GPS 定位、伽利略系统定位、塔式三边测量法定点、文本信息、听觉信息、加速计读数、陀螺仪读数,或温度读数。

41. 根据权利要求 38 所述的系统,其中所述服务器进一步经配置以:  
将所述图像与训练图像进行比较以确定与所述图像匹配的匹配的训练图像,其中所述训练图像是基于所述元数据而被选定,以及  
基于所述匹配的训练图像而辨识所述图像中的所述对象。

42. 根据权利要求 38 所述的系统,其中所述服务器进一步经配置以:  
基于所述元数据而检测所述图像中的所述对象,  
产生所述对象的特征向量,  
将所述特征向量与训练图像的图像系数进行比较以确定与所述对象匹配的匹配的训练图像,其中所述图像系数是基于所述元数据而被选定,以及  
基于所述匹配的训练图像而辨识所述对象。

43. 根据权利要求 38 所述的系统,其中所述对象包含标志、设计、面部、地标、服装、记号、自然对象或人造对象。

44. 根据权利要求 38 所述的系统,其中所述信息内容包含名称、价格、制造商、评论、优惠券,或广告。

45. 根据权利要求 38 所述的系统,其中所述服务器进一步经配置以:  
存储所述视觉搜索查询,以及  
使所述信息内容与所述视觉搜索查询相关联。

46. 一种执行图像辨识的方法,其包括:  
从移动装置接收视觉搜索查询,其中所述视觉搜索查询包含图像和与所述图像中的至少一个对象相关联的元数据;  
基于所述元数据而辨识与所述至少一个对象相关联的所述图像中的对象;  
基于所述所辨识的对象而产生信息内容;以及  
响应于所述视觉搜索查询而传送所述信息内容。

47. 根据权利要求 46 所述的方法,其中所述视觉搜索查询进一步包含与所述图像相关联的上下文数据,且进一步其中产生信息内容包括基于所述所辨识的对象和所述上下文数据而产生所述信息内容。

48. 根据权利要求 47 所述的系统,其中所述上下文数据包括全球定位系统 GPS 定位、辅助式全球定位系统 A-GPS 定位、伽利略系统定位、塔式三边测量法定点、文本信息、听觉信息、加速计读数、陀螺仪读数,或温度读数。

49. 根据权利要求 46 所述的方法,其中辨识所述图像中的对象进一步包括:  
将所述图像与训练图像的集合进行比较以确定与所述图像匹配的匹配的训练图像,其中基于所述元数据而选择训练图像的所述集合;以及

基于所述匹配的训练图像而辨识所述图像中的所述对象。

50. 一种执行图像辨识的系统,其包括:

用于从移动装置接收视觉搜索查询的装置,其中所述视觉搜索查询包含图像和与所述图像中的至少一个对象相关联的元数据,

用于基于所述元数据而辨识与所述至少一个对象相关联的所述图像中的对象的装置;

用于基于所述所辨识的对象而产生信息内容的装置;以及

用于响应于所述视觉搜索查询而传送所述信息内容的装置。

51. 根据权利要求 50 所述的系统,其中所述视觉搜索查询进一步包含与所述图像相关联的上下文数据,且进一步其中所述用于产生信息内容的装置包括用于基于所述所辨识的对象和所述上下文数据而产生所述信息内容的装置。

52. 根据权利要求 51 所述的系统,其中所述上下文数据包括全球定位系统 GPS 定位、辅助式全球定位系统 A-GPS 定位、伽利略系统定位、塔式三边测量法定点、文本信息、听觉信息、加速计读数、陀螺仪读数,或温度读数。

53. 根据权利要求 50 所述的系统,其中所述用于辨识所述图像中的对象的装置进一步包括:

用于将所述图像与训练图像的集合进行比较以确定与所述图像匹配的匹配的所述训练图像的装置,其中训练图像的所述集合是基于所述元数据而被选定;以及

用于基于所述匹配的训练图像而辨识所述图像中的所述对象的装置。

54. 一种计算机程序产品,其包括:

计算机可读媒体,所述计算机可读媒体存储在由计算机执行时致使所述计算机执行图像辨识的指令,所述指令包括:

用于致使所述计算机从移动装置接收视觉搜索查询的至少一个指令,其中所述视觉搜索查询包含图像和与所述图像中的至少一个对象相关联的元数据,

用于致使所述计算机基于所述元数据而辨识与所述至少一个对象相关联的所述图像中的对象的至少一个指令;

用于致使所述计算机基于所述所辨识的对象而产生信息内容的至少一个指令;以及

用于致使所述计算机响应于所述视觉搜索查询而传送所述信息内容的至少一个指令。

55. 根据权利要求 54 所述的计算机程序产品,其中所述视觉搜索查询进一步包含与所述图像相关联的上下文数据,且进一步其中所述用于致使所述计算机产生信息内容的至少一个指令包括用于致使所述计算机基于所述所辨识的对象和所述上下文数据而产生所述信息内容的至少一个指令。

56. 根据权利要求 55 所述的计算机程序产品,其中所述上下文数据包括全球定位系统 GPS 定位、辅助式全球定位系统 A-GPS 定位、伽利略系统定位、塔式三边测量法定点、文本信息、听觉信息、加速计读数、陀螺仪读数,或温度读数。

57. 根据权利要求 54 所述的计算机程序产品,其中所述用于致使所述计算机辨识所述图像中的对象的至少一个指令进一步包括:

用于致使所述计算机将所述图像与训练图像的集合进行比较以确定与所述图像匹配的匹配的所述训练图像的至少一个指令,其中训练图像的所述集合是基于所述元数据而被选



定;以及

用于致使所述计算机基于所述匹配的训练图像而辨识所述图像中的所述对象的至少一个指令。

## 用于使用移动装置进行图像辨识的系统和方法

[0001] 根据 35U. S. C. § 119 主张优先权

[0002] 本专利申请案主张Ricardo dos Santos、Yong Chang、Joseph Huang、Hsiang-Tsun Li 和 Dev Yamakawa 在 2009 年 4 月 14 日申请的标题为“用于使用移动装置进行图像辨识的系统和方法 (Systems and Methods for Image Recognition Using Mobile Devices)”的第 61/169, 295 号临时申请案的优先权, 所述临时申请案转让给或义务转让给与本申请案相同的实体, 且在此明确以引用的方式并入本文中。

### 技术领域

#### 背景技术

[0003] 本发明的技术大体上涉及用于使用移动装置来执行图像辨识和视觉搜索的方法和装置, 且更特定来说, 涉及平台和技术, 其用于预处理在移动装置上提取的图像以提取图像参数的缩减的集合, 所述图像参数的缩减的集合可被传送到网络辨识系统以识别所关注的对象, 且基于所述识别来搜寻相关内容。

[0004] 蜂窝式通信技术和移动通信装置中的进步 (例如, 相机和视频记录技术集成到此类通信装置上、电子邮件和短消息接发服务集成到蜂窝式通信网络中等) 将更大的灵活性、处理能力和通信能力添加到已到处存在的移动通信装置。结果, 此类移动通信装置已在消费者市场中变得更普遍, 且许多消费者如今依赖于其移动通信装置 (例如, 蜂窝式电话) 来拍照片和摄取视频、在其社交网络中交换消息、作出购买决策、进行金融交易, 以及实行其它行为。

[0005] 基于受广告内容和其投送的时机影响的既定接收方对内容的点击和转换速率, 可对准、投送并定价广告和其它信息内容。举例来说, 在日本, 大量消费者已使用其蜂窝式电话来拍摄所印刷广告中的条形码的照片来获得与所广告的产品或服务相关联的信息, 且如果相关的广告内容被迅速地发送到潜在的消费者的蜂窝式电话, 则此广告内容可能具有较高的转换速率。潜在的消费者使用其蜂窝式电话来拍摄所印刷的的广告的照片, 其随后将具有所印刷的的广告的照片的多媒体消息接发服务 (MMS) 消息发送到服务器。服务器执行所述图片与广告数据库的一对一匹配, 且在大约三十到六十秒之后, 服务器将含有与所印刷的广告相关联的 web 链接的短消息接发服务 (SMS) 消息发送到潜在的消费者。然而, 此类广告和信息内容对准和递送系统需要大量的带宽来传输所印刷的的广告的图片, 且花费大量资源来匹配图片与整个广告数据库。

#### 发明内容

[0006] 以下内容呈现一个或一个以上方面的简化概要以便提供对此些方面的基本理解。此概述并非所有所涵盖方面的广泛综述, 且既定不识别所有方面的关键或决定性要素, 也不描绘任何或所有方面的范围。其唯一目的是以简化形式来呈现一个或一个以上方面的一些概念以作为稍后所呈现的更详细描述的前言。

[0007] 根据一个或一个以上方面中的本发明的教导,提供用于执行图像辨识和移动视觉搜索的方法和设备,其中移动装置用户经由移动装置来获取图像并接收与所述图像相关联的信息内容。在本发明的教导的一个或一个以上实施方案中,移动装置可基于所获取的图像中的突出特征群集而检测、分类,和/或辨识一个或一个以上对象,且基于所述对象而产生视觉搜索查询。所述视觉搜索查询可包含所述所获取的图像或从中提取的查询图像以及与所述对象相关联的元数据。移动装置可将视觉搜索查询无线地传送到远程服务器,远程服务器可响应于所述视觉搜索查询而产生信息内容,且其后,移动装置可接收并呈现所述信息内容。

[0008] 根据一个实施方案,移动装置可检测对象并向用户突出显示所述对象,且接收指示至少一个选定对象的输入。对象可包含,例如,标志、设计、面部、地标、服装、记号、自然或人造对象等。移动装置可随后基于选定对象而产生特征向量,且将所述特征向量与训练图像的图像系数的集合进行比较以确定与选定对象匹配的匹配的训练图像。图像系数的所述集合可存储在移动装置中。移动装置可其后基于匹配的训练图像而分类和/或辨识选定对象,且基于匹配的训练图像而产生元数据。移动装置还可例如通过修剪所获取的图像、压缩所获取的图像、缩放所获取的图像以及将所获取的图像转换到灰度级而基于选定对象从所获取的图像中提取查询图像。

[0009] 根据一个实施方案,移动装置可包括传感器,所述传感器获取与所获取的图像相关联的上下文数据,且将所述上下文数据包含于视觉搜索查询中。上下文数据可包含(例如)全球定位系统(GPS)定位、辅助式GPS(A-GPS)定位、伽利略系统定位、塔式三边测量法定点、用户输入的文本或编辑信息、加速计读数、陀螺仪读数、温度读数等。

[0010] 根据一个实施方案,移动装置可将视觉搜索查询无线地传送到图像辨识系统中的远程服务器。在接收到包含查询图像和与所述查询图像中的至少一个对象相关联的元数据的视觉搜索查询后,远程服务器可即刻基于所述相关联的元数据而辨识所述查询图像中的对象。举例来说,远程服务器可基于相关联的元数据而选择训练图像的集合,将所述查询图像与所述训练图像的集合进行比较以确定与所述查询图像匹配的匹配的训练图像,且基于所述匹配的训练图像而辨识图像中的对象。远程服务器可随后基于所辨识的对象而产生信息内容,且响应于视觉搜索查询而传送信息内容。信息内容可包含(例如)名称、价格、制造商、评论、优惠券和广告。

[0011] 根据一个实施方案,远程服务器可接收视觉搜索查询,所述视觉搜索查询除了查询图像和相关联的元数据之外还包含与查询图像相关联的上下文数据。在此实施方案中,远程服务器可基于所辨识的对象和上下文数据而产生信息内容,且其后响应于视觉搜索查询而将信息内容传送到移动装置。

[0012] 根据一个实施方案,在一方面,因为移动装置在无线地传送视觉搜索查询之前预处理所获取的图像,所以移动装置可提取并发送所获取的图像的相关部分而不是整个所获取的图像,且因而提高传送视觉搜索查询的速度,且降低通信带宽要求。另外,远程服务器可利用与查询图像相关联的元数据和/或上下文数据来辅助辨识查询图像中的关注对象,其使得远程服务器能够集中于视觉搜索的范围,且因而改进远程服务器以及整个图像辨识系统的准确性、速度和效率。此外,远程服务器可使用相关联的元数据和/或上下文数据来集中于或以其它方式裁剪信息内容,其可使得远程服务器且因此图像辨识系统能够响应于

视觉搜索查询而实时地或准实时地提供相关的信息内容。

[0013] 为了实现上述和相关目的,一个或一个以上方面包括在下文全面描述且在权利要求书中特定指出的特征。以下描述和附图详细陈述一个或一个以上方面的某些说明性特征。然而,这些特征仅指示其中可采用各种方面的原理的数种方式,且此描述意在包含所有此些方面和其等效物。

#### 附图说明

[0014] 并入本说明书中并构成本说明书的一部分的附图说明本发明的教示的各方面,且与描述一起用以阐释本发明的教示的原理。在图中:

[0015] 图 1 说明与本发明的教示的一个方面相一致的示范性图像辨识系统,其包含具有便携式图像传感器的移动装置和图像辨识系统的后端中的远程服务器;

[0016] 图 2 说明根据本发明的教示的一个实施方案的移动装置的示范性配置;

[0017] 图 3 说明与本发明的教示的一个实施方案相一致的用于促进并参与移动视觉搜索和图像辨识的图像辨识系统的后端的示范性配置;

[0018] 图 4 说明根据本发明的教示的另一实施方案的由移动装置执行以实现移动视觉搜索并促进图像辨识的处理流程图;

[0019] 图 5 说明根据本发明的教示的又一实施方案的由图像辨识系统执行以实现移动视觉搜索并促进图像辨识的处理流程图;以及

[0020] 图 6A 到图 6D 说明根据本发明的教示的再一实施方案的示范性移动视觉搜索的处理序列。

#### 具体实施方式

[0021] 现将详细参考本发明的教示的一个实施方案,其实例说明于附图中。只要可能,将在整个图式中使用相同的参考标号来指代相同或相似的部分。

[0022] 现在参考图式来描述各个方面。在以下描述中,出于阐释的目的,陈述大量特定细节以便提供对一个或一个以上方面的透彻理解。然而,可明显地看出,所述方面可在没有这些特定细节的情况下得以实践。

[0023] 在本描述中,词“示范性”用于指充当实例、例子或说明。任何在本文中被描述为“示范性”的方面或设计不一定被理解为比其它方面或设计优选或有利。事实上,词“示范性”的使用希望以具体方式来展现概念。

[0024] 另外,术语“或”意在指包含性的“或”而非排它性的“或”。即,除非另有指定或从上下文清楚地看出,否则短语“X 使用 A 或 B”意在指自然包含性排列中的任一者。即,以下例子中的任一者均满足短语“X 使用 A 或 B”:X 使用 A;X 使用 B;或 X 使用 A 及 B 两者。此外,如本申请案及所附权利要求书中所使用的冠词“一”通常应被理解成表示“一个或一个以上”,除非另有指定或从上下文清楚地看出其是针对单数形式。

[0025] 另外,将按照可包含许多装置、组件、模块等的系统来呈现各个方面或特征。应理解和了解,各种系统可包含额外装置、组件、模块等,和 / 或可不包含结合图所论述的装置、组件、模块等中的全部。也可使用这些方法的组合。

[0026] 本发明的教示的方面涉及用于经由移动装置来执行视觉搜索和图像辨识的系统

和方法。更特定来说,在一个或一个以上方面中,且例如如图 1 中大体上展示,提供在其中经由移动装置 130 来起始移动视觉搜索和图像辨识的平台和技术,且其在时间图像辨识系统 120 中执行。根据一个实施方案,且例如如图 1 和图 2 中大体上展示,移动装置 130 获取并预处理图像 100 以起始移动视觉搜索。移动装置 130 可基于图像 100 中的突出特征的群集或所关注的特征来检测一个或一个以上对象,且突出显示检测到的对象,例如图案 115,且将所述对象与训练图像进行比较以分类或辨识所述对象。移动装置可基于所分类或所辨识的对象而从所获取的图像中提取子图像。移动装置 130 还可基于所分类或所辨识的对象而产生元数据并获取与图像 100 相关联的上下文数据(例如,全球定位系统(GPS)位置)。移动装置 130 可产生包含所获取的图像或从中提取的子图像以及相关元数据和/或上下文数据的视觉搜索查询,且经由无线连接 132 和无线服务提供商 150 而将所述视觉搜索查询传送到远程服务器 140。在一个或一个以上实例中,所提取的子图像具有比所获取的图像的文件大小小的文件大小。因此,传送包含所提取的子图像而不是整个所获取的图像的视觉搜索查询。此图像缩减可提高传送视觉搜索查询的速度,并降低到服务器或其它目的地的通信带宽要求。

[0027] 根据一个实施方案,且例如如图 1 和图 3 中所展示,图像辨识系统 120 的远程服务器 140 可接收视觉搜索查询,且产生待经由移动装置 130 呈现的信息内容。在接收到包含查询图像和与所述查询图像相关联的元数据和/或上下文数据的视觉搜索查询后,远程服务器 140 可即刻基于相关联的元数据而辨识查询图像中的至少一个对象。远程服务器可基于所辨识的对象和相关联的上下文数据而产生信息内容,且随后将所述信息内容传送到移动装置 130。其后,移动装置 130 可响应于视觉搜索查询而呈现所述信息内容。

[0028] 受益于与查询图像相关联的元数据和/或上下文数据,远程服务器 140 可集中于视觉搜索的范围,且因而改进远程服务器 140 以及整个图像辨识系统 120 的准确性、速度和效率。此外,远程服务器 140 可使用相关联的元数据和/或上下文数据来裁剪信息内容,其可使得远程服务器 140 且因此图像辨识系统 120 能够响应于视觉搜索查询而实时地或准实时地提供相关的信息内容。

[0029] 移动装置 130 所俘获的图像 100 或图像 100 内的图案 115 可含有对应于一个或一个以上对象的突出特征的一个或一个以上群集(例如,特征、所关注的对象等)。对象可包含(例如,但无任何限制)标志、设计、面部、地标、服装(例如,t 恤、帽子、鞋子、口袋等)、记号(例如,路牌、酒店记号等)、条形码、广告、杂志、报纸、海报(例如,“一张”等)、广告牌、布告、油画、素描、其上显示或投影图像的背景幕、零售商品保养说明书、数字视频光盘(DVD)盒、张贴物、票、压缩光盘(CD)盒、棒球卡、苏打罐等,或其任何组合。在一个实例中,图像 100 或图案 115 可为二维的,即使图像 100 中所俘获的对象或对象的表面不平坦和/或是二维的也如此。图 1 展示图像辨识系统 120 的一个实施方案,其中由具有便携式图像传感器的移动装置 130 俘获图像 100 和一个或一个以上图案 115。

[0030] 图像辨识系统 120 可经提供以实现视觉搜索并递送与图像 100 中的对象和/或图像 100 内的图案 115 相关联的信息内容。与对象相关联的信息内容可包括视觉、听觉或感官内容,或使得可访问此类内容的位置描述符。举例来说,信息内容可呈以下形式:图像、文本、串流或非串流视频、串流或非串流音频、通用资源定位符(URL)、无线应用协议(WAP)页、超文本标记语言(HTML)页、可扩展标记语言(XML)文档、可执行程序、文件名称、因特网

协议 (IP) 地址、电话呼叫、装置或其它内容。可经由通信协议将信息内容传送到移动装置 130, 所述通信协议例如为 (无任何限制) 电子邮件 (e-mail)、多媒体消息接发服务 (MMS)、增强型消息接发服务 (EMS)、短消息接发服务 (SMS)、WAP 推送、应用程序推送 (例如, 推送注册等)、标准电话显示, 或例如传输控制协议 (TCP)、IP、用户数据报协议 (UDP)、超文本传递协议 (HTTP) 和文件传递协议 (FTP) 等标准因特网协议。

[0031] 如图 1 中所示, 图像辨识系统 120 包括移动装置 130, 移动装置 130 俘获、产生、获取或以其它方式复制图像 100 (其子图像包含包含一个或一个以上对象的图案 115), 且基于对象而产生视觉搜索查询。图像 100 是由移动装置 130 俘获的对象的电子表示。举例来说, 图像 100 可为包括二维像素信息阵列的数据结构。移动装置 130 的实例可包括任何移动电子装置, 例如 (无任何限制) 蜂窝式电话 (“手机”)、个人数字助理 (PDA)、数码相机, 或适于在无线接入网络 (例如使用 IEEE 802.16 标准 (WiMAX) 或 IEEE 802.11 标准 (Wi-Fi) 操作的无线接入网络) 上操作的无线电话, 或这些装置中的两者或两者以上的电耦合集合, 例如与 PDA 有线或无线通信的数码相机。

[0032] 移动装置 130 包括便携式图像传感器 (例如, 如图 2 中所示的图像传感器 200 等), 其可为能够产生图像 100 的任何电子装置。举例来说, 便携式图像传感器可包括电荷耦合装置 (CCD) 传感器或互补金属氧化物半导体 (CMOS) 传感器, 以及用以将光图案传达到传感器上并进而产生图像 100 的一组光学镜头。在一个实施方案中, 便携式图像传感器被建置到移动装置 130 上。在操作中, 用户在目标的大体方向上对准移动装置 130 的便携式图像传感器, 且移动装置 130 在俘获覆盖目标的区域后产生图像 100。移动装置 130 还可检索一个或一个以上所存储的图像或俘获视频的一个或一个以上帧以产生图像 100。举例来说, 作为使用便携式图像传感器产生图像的替代, 移动装置 130 可检索存储在 130 中的或经由通信协议 (例如, 电子邮件、MMS、EMS、SMS、HTTP、UDP 等) 发送的图像以产生图像 100。在一个实施方案中, 所检索的图像或所俘获的帧可包含来自先前进行的视觉搜索的视觉搜索结果和 / 或用户信息, 且移动装置 130 可独立于或结合 (例如, 重叠于其上) 图像 100 来显示视觉搜索结果和 / 或用户注解。如将在下文更详细地描述, 移动装置 130 可检测图像 100 中的对象, 且实时地或准实时地向用户突出显示或以其它方式指示对象中的一者或一者以上。关于在移动装置 130 中集成的方面, 可使用购自加利福尼亚州圣地亚哥市的高通公司的 BREW® 编程接口 (BREW® API) 来执行或增加对象检测。其它图像检测和辨识 API 或服务可用于将对象检测集成于移动装置 130 中, 例如可在 Java 平台、来自 Sun 微系统的微型编辑 (Java ME™)、来自 Symbian 有限公司的 Symbian™ OS, 来自 Adobe 系统的 Flash Lite™, 来自微软公司的 Windows Mobile™, 来自苹果公司的 iPhone™ OS, 来自开放手持机联盟的 Android™ 等下执行的 API 和服务。

[0033] 移动装置 130 还可包含用以检测位置、定位、定向、移动和 / 或与移动装置 130 在产生图像 100 时相关联的其它上下文数据的能力。对移动装置 130 的位置或定位的检测和识别可例如使用各种定位服务来执行, 例如全球定位系统 (GPS)、辅助式 GPS (A-GPS)、基于蜂窝式电话基站对经注册的蜂窝式电话的三角测量法或三边测量法的基站三角测量法或三边测量法、欧洲伽利略定位系统, 或其它定位或位置服务或技术。对移动装置 130 的定向或移动的检测和识别可例如使用各种服务来执行, 例如内置传感器 (例如, 如图 2 中所示的传感器 290 等), 所述内置传感器包含 (例如) GPS 单元、加速计、陀螺仪, 和 / 或其它定向

和移动检测服务或技术。移动装置 130 可进一步包含用户输入接口（例如，小键盘、麦克风等），所述用户输入接口可接收用户输入的文本或听觉信息并提供所述文本或听觉信息作为上下文数据。移动装置 130 还可包含其它类型的传感器，例如温度传感器，所述传感器可提供其它类型的上下文数据。如图 1 中所示，移动装置 130 可经由无线连接 132 和由在图像辨识系统 120 内操作的一个或一个以上无线服务器支持的一个或一个以上基站 135 与无线服务提供商 150 通信。无线服务提供商 150 可随后与一组资源通信，所述资源包含例如存储用户相关预订、配置、定位和其它信息的用户数据库。

[0034] 在一个实施方案中，图像辨识系统 120 可进一步包括远程服务器 140，远程服务器 140 与移动装置 130 和无线服务提供商 150 联合操作以实现视觉搜索并实时地、准实时地或以其它方式递送与图像 100 中的对象相关的信息内容。远程服务器 140 包括一个或一个以上服务器 142、144 和 146，其可通过跨越例如局域网 (LAN)、内联网或因特网等一个或一个以上通信网络的连接 148 而耦合。举例来说，远程服务器 140 可包含消息接发服务器 142 中的一者或一者以上，以处置与无线服务提供商 150 和 / 或移动装置 130 的通信，且响应于视觉搜索查询而向移动装置 130 递送信息内容或提供对信息内容的访问权，所述视觉搜索查询可包含图像数据、元数据和 / 或与图像 100 相关联的上下文数据；远程服务器 140 可包含内容服务器 144 以存储并提供信息内容；且远程服务器 140 可包含图像辨识服务器 146 以确定递送何信息内容和 / 或递送信息内容的方式。在一个实施方案中，消息接发服务器 142、内容服务器 144 和图像辨识服务器 146 可驻留于不同物理位置处，且经由因特网 148 上的连接而以通信方式耦合。举例来说，消息接发服务器 142 和图像辨识服务器 146 可物理上驻留于由还管理无线服务提供商 150 的蜂窝式电话公司管理的位置处。同时，内容服务器 144 可物理上驻留于广告销售网络、销售提供商、内容提供商、媒体提供商，或其它提供商或待递送到移动装置 130 的内容的源处。

[0035] 远程服务器 140 可经由一个或一个以上通信连接 170 耦合到无线服务提供商 150，所述一个或一个以上通信连接 170 可包含有线电链路（例如，T1 或 T3 线路等）、无线链路、光学链路或其它通信耦合模式。无线服务提供商 150 可将蜂窝式电话或其它数字通信服务提供给电子装置（例如，移动装置 130）的用户。举例来说，无线服务提供商 150 可为蜂窝式电话服务提供商（例如，Sprint Nextel 公司等）、个人通信服务 (PCS) 提供商，或其它无线服务提供商。无线服务提供商 150 可包含一个或一个以上无线服务器和基站 135 的网络。移动装置 130 可使用无线连接 132 上的多层次（例如，客户端 - 服务器等）软件架构经由基站 135 与无线服务提供商 150 的无线服务器通信。因而，移动装置 130 可经由无线服务提供商 150 与远程服务器 140 通信，且远程服务器 140 可经由无线服务提供商 150 将相关的信息内容递送到移动装置 130。递送信息内容可包含向图像辨识系统 120 的用户呈现信息内容。举例来说，可将信息内容传输到移动装置 130 以例如在视觉显示器上或在音频扬声器上向用户呈现。

[0036] 现将参考图 2 来说明与本发明的教导的一个或一个以上实施方案相一致的移动装置 130 的示范性配置。移动装置 130（如图 1 中所示）可包含：至少一个天线 202（例如，发射接收器或包括输入接口等的此类接收器的群组），其接收信号（例如，关于移动呼叫起始或其它信号交换、信号交换响应、移动应用数据传递、数据事件、数据事件响应、信号交换终止等）；以及接收器 204，其对所接收信号执行若干动作（例如，滤波、放大、下变频等）。

天线 202 可例如发射或接收对信号交换请求、数据事件请求等的响应。天线 202 和接收器 204 还可与解调器 206 耦合,解调器 206 可解调所接收的信号,并将其提供给处理器 208 以进行处理。移动装置 130 可另外包含存储器 210,其包含一个或一个以上计算机可读媒体,其在操作上耦合到处理器 208 且可存储待执行的指令和待发射、接收、处理的数据等。

[0037] 处理器 208 可分析由天线 202 和 / 或移动装置 130 的用户输入接口 (未描绘) 接收的信息,且 / 或产生供发射器 218 经由调制器 216 发射的信息。另外,处理器 208 可控制和 / 或参考移动装置 130 的一个或一个以上资源或组件,所述资源或组件例如包含图像传感器 200、解调器 206、存储器 210、调制器 216、发射器 218、图像检测单元 250、图像辨识单元 260 和传感器 290。处理器 208 还可执行运行时环境 212 (例如高通公司的 BREW<sup>®</sup>、来自 Sun 微系统的 Java ME<sup>™</sup>、来自 Symbian 有限公司的 Symbian<sup>™</sup> OS,来自 Adobe 系统的 Flash Lite<sup>™</sup>,来自微软公司的 Windows Mobile<sup>™</sup>,来自苹果公司的 iPhone<sup>™</sup> OS,来自开放手机联盟的 Android<sup>™</sup> 等),以及应用程序集合 214,或其它软件、模块、应用程序、逻辑、代码等。

[0038] 在一个实施方案中,移动装置 130 包含存储器 210 以存储计算机可读数据 (例如,如图 1 中所示的图像 100、图像系数库 262 等) 和计算机可执行软件指令 (例如,图像检测 / 辨识软件 270、运行时环境 212、应用程序集合 214 等)。存储器 210 可包括固态存储器 (例如,只读存储器、随机存取存储器、快闪存储器等)、磁性硬盘驱动器、光学可读媒体 (例如,压缩光盘 (CD) 或数字视频光盘 (DVD)) 等中的一者或一者以上。移动装置 130 还可包括至少一个处理器 208 来执行存储在存储器 210 中的软件指令。所述指令经执行以配置处理器 208 来控制和 / 或执行 (例如) 图像传感器 200、图像检测单元 250 和图像辨识单元 260 的功能,如下文将例如关于图 4 更详细地描述。

[0039] 在一个实施方案中,将图像感测能力和图像检测和 / 或辨识功能性展示为涉及由移动装置 130 的图像传感器 200、图像检测单元 250 和图像辨识单元 260 进行的处理。举例来说,图像传感器 200 可包括 CCD 传感器或 CMOS 传感器,以及一组光学镜头,其将光图案传达到传感器上且进而产生图像 100。在操作中,用户在目标的大体方向上对准移动装置 130 的图像传感器 200,且图像传感器 200 在俘获了覆盖目标的区域后产生图像 100。移动装置 130 还可检索一个或一个以上所存储的图像或俘获视频的一个或一个以上帧以产生图像 100。在一个实施方案中,图像传感器 200 被建置到移动装置 130 中。然而,图像检测和图像辨识的功能性可完全驻留于移动装置 130 中、远程服务器 140 中,或其任何组合中。举例来说,图像检测单元 250 和图像辨识单元 260 可实施为存储于移动装置 130 的存储器 210 中并可由处理器 208 执行的图像处理软件的一个或一个以上集合 (例如,图像检测 / 辨识软件 270 等)。

[0040] 在一个实施方案中,图像检测 / 辨识软件 270 可为移动装置 130 及其组件提供到图像传感器 200、图像检测单元 250 和 / 或图像辨识单元 260 的功能性的接口。图像检测 / 辨识软件 270 可包含用于检测图像中的一种或一种以上对象类别和 / 或基于突出特征群集而辨识图像中的对象的算法。所述算法可包含 (例如) 标度不变特征转变 (例如, SIFT、SIFT++、LTI-lib SIFT 等)、加速稳健特征 (例如, SURF、SURF-d 等)、扩张现实 (例如, BazAR 等),和所属领域的技术人员已知的其它图像检测和辨识算法。图像检测 / 辨识软件 270 还可包含用于基于对应于图像中的对象的突出特征群集来检测图像中的一个或一个以上对象的类别或对其进行分类的算法,例如生物视觉皮质网络 (例如,体系最大化架构、



HMAX 等),和所属领域的技术人员已知的其它对象分类算法。对象类别可包含(例如)自然对象,例如面部、动物、植物、陆地特征等。对象类别还可包含(例如)人造对象,例如标志、设计、建筑物、地标、服装、记号、交通工具等。虽然术语“对象的类别”和“对象类别”用于描述共享某些特性的对象集合,但可互换地使用所属领域的技术人员已知的其它相似术语,例如对象的等级、对象的种类、对象的类型等。

[0041] 在一个实施方案中,移动装置 130 可使用一个或一个以上算法来检测图像中的对象,使用相同或不同算法来检测对象的类别,且/或使用相同或不同算法来辨识所述对象。在一个实施方案中,移动装置 130 可基于检测到的对象类别来选择辨识算法。举例来说,移动装置 130 可使用 HMAX 来检测并分类图像 100 中的对象,且随后使用 SIFT 来辨识图像 100 中的被分类为人造对象的对象。

[0042] 在一个实施方案中,图像检测/辨识软件 270 可包含用于检测标志的算法。标志出现在几乎每个用于销售目的的产品上,且标志检测算法可通过检测标志图案及其在图像内的边界而促进移动视觉搜索。标志可具有高度的对比度但有限的亮度和/或色彩水平,且因而标志图案的亮度和/或色度直方图可具有两个主峰。基于这些检测到的特性,可通过例如使用表 1 中所示的公式 1 获取亮度(或在 RGB 分量被伽马压缩的情况下,为 luma)和色度分量的直方图,而有效地检测标志图案。

[0043] 表 1- 公式 1 :亮度和色度分量获取

[0044]

联合分量=Y + Cb + Cr, 其中 Y 是亮度或 luma 分量, Cb 是蓝色色差色度分量, 且 Cr 是红色色差色度分量;  
Bin = (联合分量) \* 归一化\_因子, 其中归一化\_因子用于将联合分量转化为 bin 索引;  
直方图[bin] = 直方图[bin] + 1

[0045] 亮度和色度分量的直方图可具有任何数目的 bin。在一个实例中,16bin 直方图提供充分的分辨率来区分标志图案的主峰。在获取亮度和色度分量的直方图之后,标志检测算法可定位直方图中的最强的峰,通常为两个。标志检测算法可确保直方图的不同 bin 处的表示为 (peak1, bin1) 和 (peak2, bin2) 的两个最强峰满足表 2 中所提供的准则。

[0046] 表 2

[0047]

Peak1 + Peak2 > 峰\_阈值 \* 图像\_大小  
Peak1 + Peak2 < 差\_阈值 \* 图像\_大小  
Abs(bin1 - bin2) > Bin\_索引\_阈值

[0048] 在检测到标志图案后,标志检测算法可使用一维 (1-D) 投影算法来检测标志图案边界。1-D 投影算法可使用(例如)表 3 中所提供的公式 2 来获得最大联合分量和最小联合分量在 X 和 Y 方向上的增量。

[0049] 表 3- 公式 2 :1-D 投影算法

[0050]

X 投影=每一列的 Max (联合分量) - Min (联合分量);  
Y 投影=每一行的 Max (联合分量) - Min (联合分量)

[0051] 标志检测算法可基于 X 投影和 Y 投影来确定标志图案边界。在一个实例中,由于标志图案的 X 投影和 Y 投影波形尚的显著增量,标志检测算法可有效地且在高置信度的情况下确定标志图案边界。可由图像检测单元 250 使用存储在图像检测 / 辨识软件 270 中的标志检测算法,以检测和 / 或定位图像 100 内的一个或一个以上标志,如下文将例如关于图 4 更详细地描述。

[0052] 在一个实施方案中,图像处理软件可存取图像系数库 262,图像系数库 262 可存储可能的图像候选者或训练图像的图像系数。训练图像中的每一者可具有唯一地表示所述训练图像的系数或图像系数的对应向量。图像系数可包含形成对应训练图像的签名的编号集合,且图像系数的大小大体上对应于训练图像类别。举例来说,标志 (BREW GAMING MONKEY™ 标志等) 的图像系数可具有约 22x18x32 字节或约 12 千字节的大小,且人面部的图像系数可具有大于一兆字节的大小。训练图像可使用分类算法 (例如, HMAX、K-最近邻、支持向量机、神经网络、随机化树,或所属领域的技术人员已知的其它分类算法) 基于包含于其中的对象而被分类。训练图像的图像系数可存储在图像系数库 262 中,且还可根据训练图像中的对象的类别、元数据 (例如,对象类别、商标等),和 / 或与训练图像相关联的上下文数据 (例如, GPS 位置、位置识别符等) 而被编制索引。存储在图像系数库 262 中的训练图像的图像系数可由图像检测单元 250 和图像辨识单元 260 用来分类、辨识,或以其它方式识别图像 100 和 / 或图案 115 内的一个或一个以上对象,如下文将例如关于图 4 更详细地描述。

[0053] 移动装置 130 中的图像处理软件可进一步包括图像编辑软件,所述图像编辑软件可用于修剪、压缩、缩放、转换到灰度级,或以其它方式处理由图像传感器 200 俘获的图像 100,以提取或以其它方式产生包含图案 115 的子图像。举例来说,可基于检测到的、经分类的和 / 或所辨识出的对象来修剪或以其它方式处理图像 100。替代地或另外,可根据从移动装置 130 的用户接收到的或由所述用户指定的指令或根据先前已由移动装置 130 接收到的计算机可读指令来修剪或以其它方式处理图像 100。可在任何合适的编程语言和 / 或开发环境 (例如, BREW®、Java ME™、Symbian™ OS、Flash Lite™、Windows Mobile™、iPhone™ OS、Android™) 中编写图像处理软件。替代地或另外,图像检测单元 250 和图像辨识单元 260 可实施为移动装置 130 中的硬件。所述硬件可包括包含无源和 / 或有源电子组件的电子电路。举例来说,在一个实施方案中,所述硬件可实施于至少一个专用集成电路 (ASIC) 中。

[0054] 现将参看图 3 来说明与本发明的教示的实施方案相一致的图像辨识系统 120 的后端 300 的示范性配置,图像辨识系统 120 包含可促进和 / 或参与图像辨识和视觉搜索的远程服务器 140 和无线服务提供商 150。在一个实施方案中,后端 300 可包含无线服务提供商 150,无线服务提供商 150 具有:接收器 310,其经由接收天线 306 接收来自一个或一个以上移动装置 (例如,如图 1 中所示的移动装置 130 等) 的一个或一个以上信号;以及发射器 322,其经由发射天线 308 将由调制器 320 调制的一个或一个以上信号发射到移动装置。接收器 310 可接收来自接收天线 306 的信息,且可进一步包含接收与未接收或不能破译的数据包相关的反馈数据的信号接收方 (未图示)。另外,接收器 310 在操作上与解调所接收的信息的解调器 312 相关联。处理器 314 可分析由解调器 312 提供的经解调的符号和信息。

[0055] 处理器 314 进一步耦合到存储器 316,存储器 316 可存储促进和 / 或参与移动装置、无线服务提供商 150 和 / 或远程服务器 140 之间的远程通信的一个或一个以上应用程序

序 318。举例来说,应用程序 318 可包含主要应用程序,所述主要应用程序经配置以起始信号交换并将数据事件请求(例如,关于诊断信息、数据分析等)发送到移动装置上操作的接收方应用程序。或者,应用程序 318 可包含次要应用程序,所述次要应用程序可接收信号交换请求并鉴定移动装置上的起始应用程序。应用程序 318 可进一步包含用于产生和/或验证识别应用程序 318 到移动装置上的对应应用程序或移动装置上的对应应用程序到应用程序 318 的识别符,或此些识别符的识别特定往返行程通信的递增的规则。此外,所述规则可指定用于重新传输未确认的传输、重新起始信号交换请求和/或响应以及终止信号交换等的政策。因此,应用程序 318 可加入与驻留于移动装置上的一个或一个以上应用程序(例如,如图 2 中所示的应用程序集合 214 等)的移动通信,和/或与执行本文中所陈述的各种动作和功能相关的任何其它合适活动。

[0056] 在一个实施方案中,后端 300 可包含远程服务器 140,远程服务器 140 与移动装置(例如,移动装置 130 等)和无服务提供商 150 联合操作以实现图像辨识和视觉搜索。远程服务器 140 可包含消息接发服务器 142 以处置与移动装置和/或无线服务提供商 150 的通信,且响应于视觉搜索查询而向移动装置递送信息内容或提供对信息内容的访问权。举例来说,消息接发服务器 142 可接收视觉搜索查询,所述视觉搜索查询可包含图像 100 或从图像 100 中提取的一个或一个以上子图像(例如,图案 115 等),以及元数据和/或与图像 100 相关联且由移动装置产生的上下文数据,且随后将所述视觉搜索查询传输到图像辨识服务器 146。对于另一实例,消息接发服务器 142 可接收视觉搜索结果,所述视觉搜索结果可包含由内容服务器 144 响应于视觉搜索查询而产生的与图像 100 或所提取的子图像相关的信息内容,且随后将所述视觉搜索结果传输到无线服务提供商 150 以用于传输到移动装置。

[0057] 远程服务器 140 可包含图像辨识服务器 146 或与其通信以基于图像数据、元数据,与图像 100 相关联的上下文数据,和/或针对先前为了类似的视觉搜索查询而提供的搜索结果的用户反馈来辨识或以其它方式识别图像 100 内的一个或一个以上对象或所提取的子图像中的一者或一者以上。针对搜索结果的用户反馈可例如包含关于搜索结果的准确性或相关性的二进制响应(例如,是/否、真/假、好/坏等)或缩放的响应(例如,从 1 到 10 的比例)、针对搜索结果的用户注解、响应于搜索结果的用户后续动作(例如,点击搜索结果中提供的链接或广告等等)。图像辨识服务器 146 可进一步基于至少一个所辨识的对象、元数据,和/或与图像 100 相关联的上下文数据,以及针对先前提供的搜索结果的任何用户反馈而产生语义搜索结果。在一个实施方案中,图像辨识服务器 146 包含处理器 360 和存储器 362,存储器 362 包含一个或一个以上计算机可读媒体,所述一个或一个以上计算机可读媒体在操作上耦合到处理器 360 且可存储待执行的指令和待发射、接收、处理的数据等。存储器 362 可包括固态存储器、磁性硬盘驱动器、例如 CD 或 DVD 等光学可读媒体等中的一者或一者以上。存储在其中的指令经执行以配置处理器 360 来与图像辨识系统 120 的其它组件联合地控制和/或执行(例如)视觉搜索和图像辨识。举例来说,存储器 362 可存储图像辨识软件 364 以及图像数据和系数库 366。图像辨识软件 364 可存取图像系数库 366,图像系数库 366 可存储图像数据和/或可能的图像候选者或训练图像的系数并为其编制索引。可使用分类算法(例如,HMAX、K-最近邻、支持向量机、神经网络、随机化树,或所属领域的技术人员已知的其它分类算法)基于包含于训练图像中的对象而将训练图像分类。图

像系数库 366 可根据训练图像中的对象的类别、元数据（例如，对象类别、商标等），和与训练图像相关联的上下文数据（例如，GPS 位置、位置识别符等）而为训练图像编制索引。训练图像中的每一者具有唯一地表示训练图像的数据和 / 或系数向量，所述数据和 / 或系数向量可存储在图像数据和系数库 366 中且由图像辨识软件 364 使用来辨识图像 100 或图案 115 内的一个或一个以上对象，如下文将例如关于图 5 更详细地描述。

[0058] 远程服务器 140 可进一步包含内容服务器 144 或与其通信，以存储、编制索引和提供信息内容，例如产品信息（例如，名称、价格、制造商、规格、评论、广告、优惠券、促销等）、到产品信息的引导链接、动作链接（例如，用于比较商店、保存到意愿列表、与朋友共享、即时购买等的到线上零售商的链接）、名人信息（例如，与名人相关联的姓名、个人简介、产品和 / 或服务）。地标信息（例如，与地标相关联的名称、历史、产品和 / 或服务），或其任何组合。内容服务器 144 可响应于（例如）由图像辨识服务器 146 基于至少一个对象、元数据和 / 或与图像 100 相关联的上下文数据而产生的语义搜索查询而提供相关信息内容。

[0059] 在一个实施方案中，内容服务器 144 包含处理器 340 和存储器 342，存储器 342 包含一个或一个以上计算机可读媒体，所述一个或一个以上计算机可读媒体在操作上耦合到处理器 340 且可存储待执行的指令和待发射、接收、处理的数据等。存储器 342 可包括固态存储器、磁性硬盘驱动器、例如 CD 或 DVD 等光学可读媒体等中的一者或一者以上。存储在其中的指令经执行以配置处理器 340 来与图像辨识系统 120 的其它组件联合地基于图像 100 或图案 115 而搜寻和提供相关信息内容。举例来说，存储器 342 可存储用于搜索引擎 344 和信息内容数据库 346 的指令。

[0060] 搜索引擎 344 可响应于来自移动装置和 / 或图像辨识服务器 146 的搜索查询而定位和提供相关信息内容。在所展示的实施方案中，在接收搜索查询之前，内容服务器 144 可执行对信息内容数据库 346 和 / 或耦合到远程服务器 140 的其它计算机可读数据存储媒体的蠕动，以定位存储在其中的信息内容并为其编制索引。因而，搜索引擎 344 可响应于搜索查询通过存取索引而定位相关信息内容。因此，信息内容服务器 144 可基于由图像辨识服务器 146 产生的语义搜索查询而确定将何信息内容递送到移动装置和 / 或如何递送信息内容，例如信息内容和通信协议等的形式，如下文将例如关于图 5 更详细地描述。

[0061] 图 4 和图 5 说明根据本发明的教导的一个或一个以上方面的方法和 / 或流程图。为了易于阐释，将方法描绘和描述为一系列动作。应理解并了解，本创新不受所说明的动作及 / 或动作次序的限制。举例来说，动作可以各种次序及 / 或同时发生，且与本文中未呈现并描述的其它动作一起发生。此外，可能不需要所有所说明的动作来实施根据所主张的标的物的方法。此外，所属领域的技术人员将理解并了解，方法可替代地经由状态图或事件而表示为一系列相关状态。此外，应进一步了解，下文中及贯穿本说明书所揭示的方法能够存储于制品上以促进将所述方法运输并传送到计算机。如本文中所使用的术语“制品”既定涵盖可从任何计算机可读装置、载体或媒体存取的计算机程序。

[0062] 图 4 说明根据本发明的教导的一个或一个以上实施方案的可由移动装置 130（如图 1 和图 2 中所展示）执行以使用图像辨识系统 120（如图 1 中展示）来搜索并促进图像辨识的处理的流程图。在 410 中，移动装置 130 可通过获取图像（例如，如图 1 中所示的图像 100，如图 6A 中所示的图像 600 等）而起始视觉搜索和图像辨识。举例来说，移动装置 130 的用户在目标的大体方向上对准移动装置 130 的图像传感器 200，且移动装置 130 可俘

获、产生、获取或以其它方式复制表示所述目标的图像。移动装置 130 还可检索一个或一个以上所存储的图像或俘获视频的一个或一个以上帧以产生图像。举例来说,作为使用图像传感器 200 产生图像的替代,移动装置 130 可检索存储在移动装置 130 中的或经由通信协议(例如,电子邮件、MMS、EMS、SMS、HTTP、UDP 等)发送的图像以产生图像。在一个实施方案中,所检索的图像或所俘获的帧可包含来自先前进行的视觉搜索的视觉搜索结果和/或用户信息,且移动装置 130 可独立于或结合(例如,重叠于其上)图像 100 来显示视觉搜索结果和/或用户注解。

[0063] 接下来,在 415 中,移动装置 130 可基于对应于所获取的图像中的对象的突出特征群集而检测一个或一个以上对象的存储和位置。在一个实施方案中,移动装置 130 可在不具有来自用户的肯定输入或其它动作(例如,按下快门)的情况下开始检测对象;而是,移动装置 130 可将连续获取的图像进行比较以确定图像传感器 200 何时静止或已静止了阈值时间周期,且因此可开始检测对象。在另一实施方案中,移动装置 130 可在来自用户的肯定输入或其它动作之后开始检测对象。

[0064] 在一个实施方案中,可通过使用存储在移动装置 130 中的一个或一个以上图像检测算法(例如,存储在如图 2 中所示的图像检测/辨识软件 270 和图像检测单元 250 中且由其执行的检测算法等)来执行或增强对象检测。可例如通过将关键点(例如,线、边、脊、隅角、滴、T 形交叉点或其它突出特征)定位在图像上来检测对象,且随后基于与关键点中的每一者相关联的点或区来产生关键点向量。通过使用关键点向量,移动装置 130 可定位图像中的对象,且随后针对对象中的每一者,移动装置 130 可产生唯一地表示对应对象的特征向量。可使用其它图像检测算法,包含(例如)HMAX、SIFT、SIFT++、LTI-lib SIFT、SURF、SURF-d、BazAR,或所属领域的技术人员已知的其它图像检测算法。

[0065] 在一个实施方案中,可针对各种类别的对象执行对象检测,例如,标志、设计、面部、地标、服装、记号、对象等。在一个方面中,可仅针对一个或一个以上预先选择或用户选择的类别的对象来执行对象检测。举例来说,对象检测可利用存储在图像检测/辨识软件 270 中的标志检测算法来检测和/或定位图像中的仅标志或类似于标志的图案。替代地或另外,移动装置 130 可通过实时地产生对应于检测到的对象的特征向量,且将所述特征向量与用于存储在图像系数库 262 中的选定类别的对象的可能的图像候选者或训练图像的图像系数进行比较,而确定检测到的对象的类别。训练图像中的每一者具有唯一地表示训练图像中的特征的对应系数向量。在一个实施方案中,移动装置 130 可通过计算向量之间的距离(例如,曼哈拉若尼斯(Manhalanobis)距离、欧几里得距离等)而将检测到的对象的特征向量与选定类别的训练图像的系数向量进行比较,以确定检测到的对象的类别。

[0066] 在一个实施方案中,移动装置 130 可检测所获取的图像中的对象,之后增强所获取的图像以供人类观看(例如,增强图像的锐度、亮度和色彩的动态范围等),且在移动装置 130 的取景器或显示器上显示经增强的图像。虽然经增强的图像在美学上更令用户愉悦,但此类增强可妨碍或甚至阻止移动装置 130 准确地并有效地检测图像中的对象。

[0067] 在 420 中,移动装置 130 可通过将指示符重叠于图像上而突出显示或以其它方式指示图像的检测到的对象。举例来说,指示符可包含各种形式的扩张的现实图形,例如,围绕如图 1 中所示的图案 115 以及围绕如图 6A 和图 6B 中所示的图案 610 到 620 的指示符、方框、靶心超链接等。如果移动装置 130 在 415 中已确定检测到的对象的类别,则移动装置

130 可仅突出显示经分类为一个或一个以上预先选定或用户选定的类别的检测到的对象。接下来,在 425 中,移动装置 130 可接收来自用户的输入以选择经突出显示的对象中的至少一者,例如,如图 6B 中所示的选定图案 610。用户输入可包含肯定输入或来自用户的经由用户输入接口的其它动作。用户输入还可包含用户静止地握住移动装置 130,使得图像传感器 200 集中于检测到的对象中的一者上并持续阈值时间周期。

[0068] 接下来,在 430 中,移动装置 130 可分类、辨识或以其它方式确定选定对象的特性。移动装置 130 可任选地精细化对应于选定对象的所产生的特征向量。移动装置 130 可通过将选定对象的特征向量与用于存储在图像系数库 262 中的一种或一种以上特征的训练图像的图像系数进行比较而确定选定对象的类别。如果移动装置 130 已将选定对象分类(在 415 中),则移动装置 130 可保持选定对象的类别而不进一步将选定对象分类。在一个实施方案中,移动装置 130 可将选定对象的特征向量与存储在图像系数库 262 中的训练图像的图像系数进行比较,以辨识或以其它方式确定选定对象的特性。在一个实施方案中,移动装置 130 可通过计算向量之间的距离(例如,曼哈拉若尼斯 (Manhalanobis) 距离、欧几里得距离等)而将选定对象的特征向量与训练图像的系数向量进行比较,以找到与选定对象匹配的训练图像。如果移动装置 130 找到与选定对象匹配的训练图像,则移动装置 130 可基于匹配的训练图像来辨识选定对象。特征向量的维度的数目直接与使特征向量匹配所需的时间和处理功率相关,且因此可能需要最小化特征向量的维度的数目。然而,特征向量应具有足够维度以具有区别且还对噪声、检测错误以及几何和光度变形稳健。

[0069] 在 435 中,移动装置 130 可基于与选定对象匹配的训练图像而产生与选定对象相关联的元数据。举例来说,如果移动装置 130 使选定对象与 BREW GAMING MONKEY™ 标志的训练图像匹配,则移动装置 130 可产生指示所述选定对象为 BREW GAMING MONKEY™ 标志或包含 BREW GAMING™ 产品的元数据。或者,如果移动装置 130 不可使选定对象与训练图像匹配,则移动装置 130 可产生包含选定对象的特征向量的元数据。

[0070] 在 440 中,移动装置可获取与所获取的图像相关联的上下文数据。移动装置 130 可在获取图像或处理所述图像以检测对象时获取与移动装置 130 相关联的位置、定位、定向、移动和 / 或其它上下文数据。举例来说,上下文数据可包含在曾获取图像的地方的 GPS 位置。对于另一实例,上下文数据可包含在曾获取图像时移动装置 130 的定向(例如,在广告牌处向上、在杂志处向下等)或环境温度。对于又一实例,上下文数据可包含用户输入的文本或听觉信息,例如类似于“在 U. S. Open 处”的文本或话音消息,类似于背景噪声的被动信息,和类似于“左边的女士是谁?”或“左边的女士穿的什么”的指示。在一个实施方案中,移动装置 130 可在 440 中独立地获取上下文数据,或与在 410 到 435 中所执行的处理中的任一者同时进行。

[0071] 接下来,在 445 中,移动装置 130 可基于所获取的图像而产生视觉搜索查询,且将所述视觉搜索查询传送到图像辨识系统 120 的后端 300。视觉搜索查询可包含到后端 300 中的处理器或服务器或在其中运行的进程的目的地地址,且可基于选定对象的类别而裁定所述目的地地址。在一个实施方案中,视觉搜索查询可包含所获取的图像或基于选定对象从所获取的图像中提取的子图像,以及与所获取的图像或所提取的子图像相关联的元数据和 / 或上下文数据。移动装置 130 可基于选定对象来修剪、压缩、缩放、转换到灰度级或以其它方式处理所获取的图像以提取或以其它方式产生至少一个子图像。

[0072] 举例来说,如图 1 和图 6C 中所说明,如果在 430 中将选定对象辨识为 BREW GAMING MONKEY™ 标志,则移动装置 130 可修剪或以其它方式处理所获取的图像以提取包含所述标志或其上附加了所述标志的对象的子图像(例如,t 恤 630、广告、优惠券、帽子、一双鞋等)。替代地或另外,移动装置 130 可根据从移动装置 130 的用户接收到的或由所述用户指定的指令,或根据先前已由移动装置 130 接收到的计算机可读指令来修剪或以其它方式处理所获取的图像。在产生了视觉搜索查询(其包含所获取的图像或所提取的子图像以及与所获取的图像或所提取的子图像相关联的元数据和 / 或上下文数据)之后,移动装置 130 可将视觉搜索查询传送到图像辨识系统 120 的后端 300。所提取的子图像具有比所获取的图像的文件大小小的文件大小。因此,传送包含所提取的子图像而不是整个所获取的图像的视觉搜索查询。此图像缩减可再次提高传送视觉搜索查询的速度。另外,传送包含所提取的子图像而不是整个所获取的图像的视觉搜索查询还可降低到服务器或其它目的地的通信带宽要求。

[0073] 在 450 中,移动装置 130 可从后端 300 接收响应于视觉搜索查询的视觉搜索结果并将其呈现给用户。移动装置 130 还可存储视觉搜索结果和 / 或使所述视觉搜索结果于所述视觉搜索查询相关联,且可接收并存储来自用户的关于所述视觉搜索结果的注解。其后,移动装置 130 可经由通信协议来发射视觉搜索结果、视觉搜索查询和 / 或用户注解。视觉搜索结果可包含与所获取的图像中的选定对象相关联的信息内容。举例来说,如果选定对象是标志(例如,如图 6B 中所示的选定图案 610 中的标志等),则所述信息内容可包含产品信息(例如,产品商标 650 和产品类型 660,如图 6D 中所示),到产品信息的引导链接(例如,信息链接 670)、相关产品(例如,相关产品 690 和广告 695)、用于比较商店、保存到意愿列表、与朋友共享或即时购买的到线上零售商的链接(例如,购买链接 680)等,或其任何组合。如果选定对象是名人的面部,则信息内容可例如包含名人的名字、他或她的个人简介,与名人相关联的产品和 / 或服务,以及其它相关信息,或其任何组合。如果选定对象是地标,则信息内容可包含地标的名称、历史、与地标相关联的产品和 / 或服务,以及其它相关信息,或其任何组合。在一个实施方案中,移动装置 130 可从后端 300 接收关于视觉搜索结果的反馈请求,例如让用户对搜索结果的准确性或相关性进行评价的确认对话框、让用户对搜索结果进行注解的输入对话框等。对各种类别和类型的图像、元数据、上下文数据、视觉搜索查询和结果、信息内容和用户反馈机制的以上列举仅用于说明目的,且并不打算以任何方式限制本发明的教导。

[0074] 图 5 说明根据本发明的教导的一个实施方案的可由后端 300(如图 3 中所示)执行以使用图像辨识系统 120(如图 1 中所示)来实现视觉搜索并促进图像辨识的处理的流程图。在 510 中,后端 300 中的远程服务器 140(如图 1 和图 3 中所示)可经由无线连接 132 和无线服务提供商 150 或所属领域的技术人员已知的其它数据发射手段来接收视觉搜索查询。如上文所描述,视觉搜索查询可包含含有至少一个所关注的对象的图像,和与所述图像相关联的元数据和 / 或上下文数据。出于说明的目的,基于图像 600(如图 6C 中所示)而产生的示范性视觉搜索查询可包含 t 恤 630 的图像、指示所述图像与 BREWGAMING™ 相关联的元数据,和指示曾在特定 GPS 位置处获取所述图像的上下文数据。

[0075] 接下来,在 515 中,远程服务器 140 可基于所述视觉搜索查询而辨识或以其它方式识别图像中的所关注的对象。远程服务器 140 可利用与图像相关联的元数据和 / 或上下文

数据,以及与先前针对类似的视觉搜索查询而提供的搜索结果相关联的任何用户反馈来辅助辨识所关注的对象,其使得远程服务器 140 能够集中于或以其它方式限制视觉搜索的范围,且因而改进图像辨识系统 120 的准确性、速度和 / 或效率。在一个实施方案中,远程服务器 140 可执行存储在图像辨识服务器 146 中的图像辨识软件 364,以执行所述图像与存储在图像数据和系数库 366 中的图像数据(例如,图像光栅数据、图像系数等)的一对一的匹配。远程服务器 140 可基于与所述图像相关联的元数据和 / 或上下文数据而集中于一对一的匹配。举例来说,在接收到基于图像 600 而产生的示范性视觉搜索查询之后,远程服务器 140 可集中于 t 恤 630 与和 BREW GAMING™ 相关联的所存储的图像数据的一对一的匹配。

[0076] 作为一对一的匹配的替代或除其之外,在 515 中,远程服务器 140 可执行图像辨识软件 364 以检测图像中的所关注的至少一个对象,且计算唯一地表示所关注的对象的特征向量。远程服务器 140 可通过将特征向量与存储在图像数据和系数库 366 中的可能的图像候选者或训练图像的图像系数进行比较,而基于所计算的特征向量来辨识所关注的对象。在一个实施方案中,远程服务器 140 可通过计算向量之间的距离(例如,曼哈拉若尼斯(Manhalanobis)距离、欧几里得距离等)而使所计算的特征向量与训练图像的系数向量匹配,以辨识所关注的对象。远程服务器 140 可随后基于匹配的训练图像来辨识所关注的对象。远程服务器 140 可基于与所述图像相关联的元数据和 / 或上下文数据而集中于向量匹配。举例来说,在接收到基于图像 600 而产生的示范性视觉搜索查询之后,远程服务器 140 可集中于从 t 恤 630 计算的特征向量与和 BREW GAMING™ 相关联的所存储的图像系数的匹配。

[0077] 在 520 中,远程服务器 140 可响应于视觉搜索查询而基于所辨识的所关注的对象产生包含信息内容的视觉搜索结果。远程服务器 140 可基于所辨识的对象、元数据,和 / 或与图像相关联的上下文数据,以及与先前针对类似的视觉搜索查询而提供的搜索结果相关联的任何用户反馈而执行语义搜索,以检索与所辨识的对象相关联和 / 或相关的信息内容。通过使用相关联的元数据和 / 或上下文数据来集中于或以其它方式限制语义搜索的范围,远程服务器 140 和因此图像辨识系统 120 可响应于视觉搜索查询提供更准确和 / 或更相关的信息内容。

[0078] 在一个实施方案中,远程服务器 140 可执行存储在内容服务器 144 中的搜索引擎 344 以执行针对存储在信息内容数据库 346 中的信息内容的语义搜索。远程服务器 140 可基于与所述图像相关联的元数据和 / 或上下文数据而集中于语义搜索。举例来说,在接收到基于图像 600 而产生的示范性视觉搜索查询并辨识出 t 恤 630 包含 BREW GAMING MONKEY™ 标志的图像之后,远程服务器 140 可执行对 BREW GAMING™ 的语义搜索以检索相关信息内容,例如产品信息(例如,产品商标 650 和产品类型 660,如图 6D 中所示),到产品信息的引导链接(例如,信息链接 670)、相关产品(例如,相关产品 690)、用于比较商店、保存到意愿列表、与朋友共享或即时购买的到线上零售商的链接(例如,购买链接 680)等,或其任何组合。对于进一步实例,远程服务器 140 可利用相关联的上下文数据(例如, GPS 位置、用户输入的文本或听觉信息等)集中于语义搜索,以基于 GPS 位置来检索相关信息内容,例如用于在所述 GPS 位置附近的商店处的相关产品 690 的广告 695(如图 6D 中所示)、在对应于所述 GPS 位置的附近商店处可得的优惠券和促销等。对各种类型的搜索查询、图像、所关注的对象、元数据、上下文数据、视觉搜索查询和结果和信息内容的以上列举仅用于说明目的,



且并不打算以任何方式限制本发明的教导。

[0079] 接下来,在 525 中,远程服务器 140 可经由无线连接 132 和无线服务提供商 150 或所属领域的技术人员已知的其它数据发射手段来将包含相关信息内容的视觉搜索结果传送或以其它方式提供给移动装置 130。移动装置 140 还可传送对关于视觉搜索结果的用户反馈的请求,例如让用户对搜索结果的准确性或相关性进行评价的确认对话框、让用户对搜索结果进行注解的输入对话框等。在 530 中,远程服务器 140 可出于任何目的而记录视觉搜索查询,所述目的例如为费用收集、报告、数据开采、用户或产品简档、未来的销售等。此外,远程服务器 140 可出于任何目的的记录与对应的视觉搜索查询相关联的或与视觉搜索查询无关的视觉搜索结果。远程服务器 140 还可出于任何目的的记录对视觉搜索结果的用户反馈和 / 或使用户反馈与视觉搜索结果相关联,所述目的例如为训练图像辨识软件 364 和 / 或搜索引擎 344、费用收集、报告、数据开采、用户或产品简档、未来的销售等。另外,远程服务器 140 可高速缓冲视觉搜索结果以响应于与上述视觉搜索查询等同或几乎等同的未来的视觉搜索查询而快速地提供搜索结果并最小化冗余处理。此外,远程服务器 140 可与由图像辨识系统 120 进行的对视觉搜索查询的处理相关联的统计数据,例如搜索时间、视觉搜索结果中的信息内容的相关性方面的置信度水平等。

[0080] 当本文中所述的实施方案以软件、固件、中间件、微代码和 / 或程序代码或代码段实施时,其可存储于例如存储组件等计算机可读存储媒体中。代码段可表示过程、函数、子程序、程序、例程、子例程、模块、软件包、类,或者指令、数据结构或程序语句的任何组合。可通过传递和 / 或接收信息、数据、自变量、参数或存储器内容而将一代码段耦合到另一代码段或硬件电路。可使用包含存储器共享、消息传递、权标传递、网络发射等的任何合适手段来传递、转发或发射信息、自变量、参数、数据等。对于软件实施方案来说,可通过执行本文中所描述的功能的模块(例如,程序、函数等)来实施本文中所描述的技术。软件代码可存储于存储器单元中且由处理器执行。存储器单元可实施于处理器内或处理器外部,在实施于处理器外部的情况下,存储器单元可经由此项技术中已知的各种手段以通信方式耦合到处理器。

[0081] 前述描述是说明性的,且所属领域的技术人员可想到配置和实施方案方面的变化。举例来说,可使用通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其经设计以执行本文中所描述的功能的任何组合来实施或执行结合本文中所揭示的实施例而描述的各种说明性逻辑、逻辑块、模块和电路。通用处理器可以是微处理器,但在替代方案中,处理器可以是任何常规处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP 与微处理器的组合、多个微处理器、一个或一个以上微处理器与 DSP 核心的联合,或任何其它此类配置。

[0082] 在一个或一个以上示范性实施方案中,可以硬件、软件、固件或其任何组合实施所描述的功能。对于软件实施方案来说,可通过执行本文中所描述的功能的模块(例如,程序、函数等)来实施本文中所描述的技术。软件代码可存储于存储器单元中且由处理器执行。存储器单元可实施于处理器内或处理器外部,在实施于处理器外部的情况下,存储器单元可经由此项技术中已知的各种手段以通信方式耦合到处理器。如果以软件实施,则可将所述功能作为一个或一个以上指令或代码存储于计算机可读媒体上或经由计算机可读媒

体来传输。计算机可读媒体包括计算机存储媒体与通信媒体两者,通信媒体包括促进将计算机程序从一处传递到另一处的任何媒体。存储媒体可为可由计算机存取的任何可用的有形媒体。以实例而非限制的方式,此类有形计算机可读媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置,或可用于载送或存储呈指令或数据结构形式的所要程序代码且可由计算机存取的任何其它媒体。同样,可恰当地将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或例如红外线、无线电和微波等无线技术从网站、服务器或其它远程源传输软件,则所述同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电和微波等无线技术包括于媒体的定义中。上述各者的组合也应包括在计算机可读媒体的范围内。

[0083] 本文中所述的技术可用于各种无线通信系统中,例如,CDMA、TDMA、FDMA、OFDMA、SC-FDMA和其它系统。术语“系统”和“网络”常常可互换使用。CDMA系统可实施例如通用陆地无线电接入(UTRA)、cdma2000等无线电技术。UTRA包含宽带CDMA(W-CDMA)和CDMA的其它变体。此外,cdma2000涵盖IS-2000、IS-95和IS-856标准。TDMA系统可实施例如全球移动通信系统(GSM)等无线电技术。OFDMA系统可实施例如演进式UTRA(E-UTRA)、超移动宽带(UMB)、IEEE 802.11(Wi-Fi)、IEEE 802.16(WiMAX)、IEEE 802.20、Flash-OFDM等。UTRA和E-UTRA是全球移动通信系统(UMTS)的部分。3GPP长期演进(LTE)是使用E-UTRA的UMTS的版本,其在下行链路上采用OFDMA且在上行链路上采用SC-FDMA。UTRA、E-UTRA、UMTS、LTE及GSM被描述于来自名为“第三代合作伙伴计划”(3GPP)的组织的文献中。另外,来自名为“第三代合作伙伴计划2”(3GPP2)的组织的文献中描述了cdma2000和UMB。另外,所述无线通信系统可另外包括对等(例如,移动装置到移动装置)专用网络系统,其经常使用不成对的未经许可的频谱、802.xx无线LAN、蓝牙(BLUETOOTH)及任何其它短程或长程无线通信技术。

[0084] 结合本文中所揭示的实施例而描述的方法或算法的步骤可直接以硬件、以由处理器执行的软件模块或以所述两者的组合来体现。虽然已通过实例描述了方法,但可以与所说明的次序不同的次序或同时地执行所述方法的步骤。软件模块可驻留于RAM存储器、快闪存储器、ROM、EPROM、EEPROM、寄存器、硬盘、可装卸磁盘、CD-ROM,或此项技术中已知的任何其它形式的计算机可读存储媒体中。示范性存储媒体耦合到处理器,使得处理器可从存储媒体读取信息以及将信息写入到存储媒体。在替代方案中,存储媒体可与处理器成一体式。处理器及存储媒体可驻留于ASIC中。ASIC可驻留于移动装置中。在替代方案中,处理器和存储媒体可作为离散组件驻留于移动装置中。描述为单数或集成的其它资源可在一个实施方案中为复数或分布的,且描述为多个或分布的资源可在若干实施方案中为组合的。相应地,本发明的教示的范围既定仅受所附权利要求书限制。

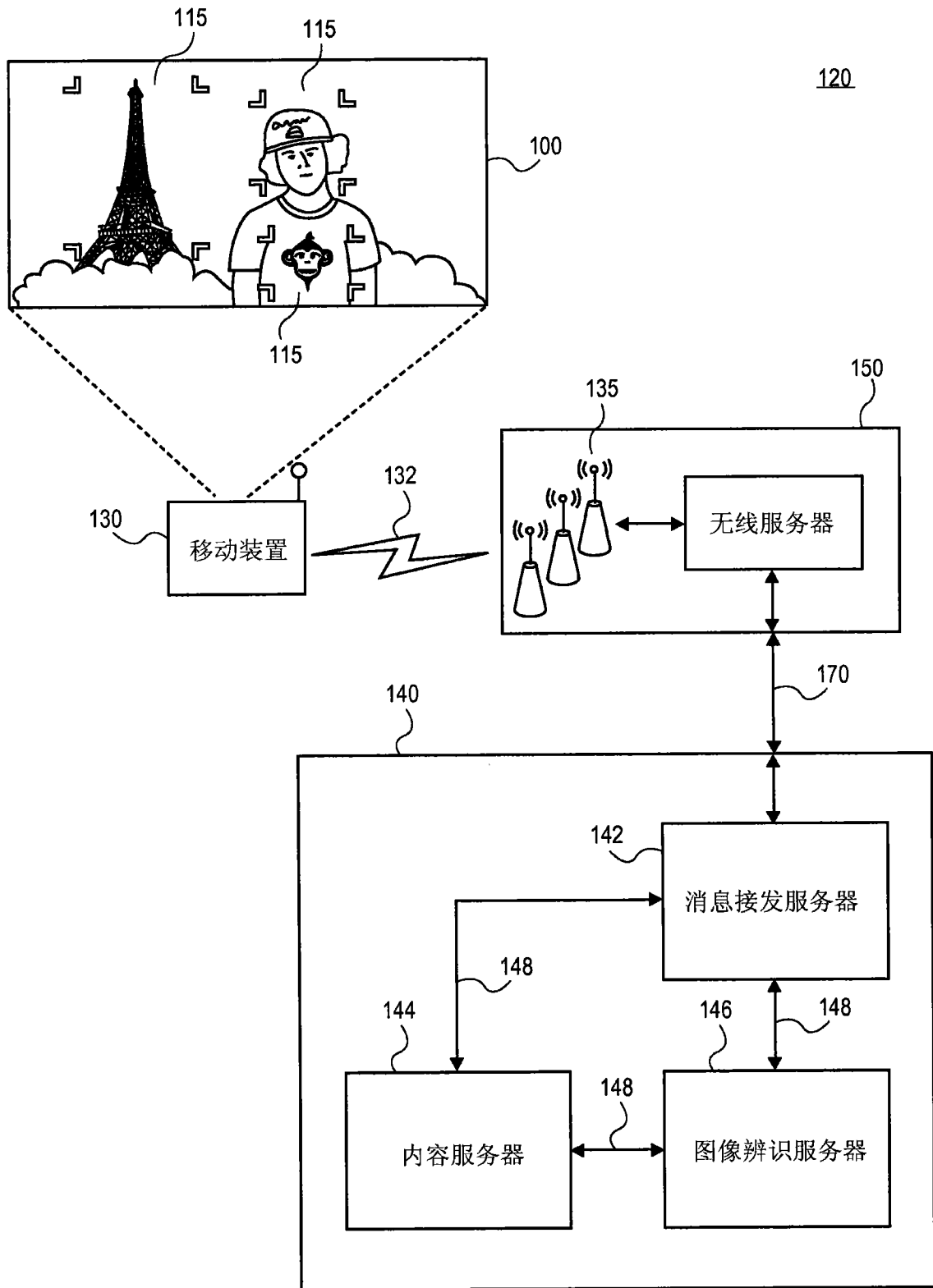


图 1

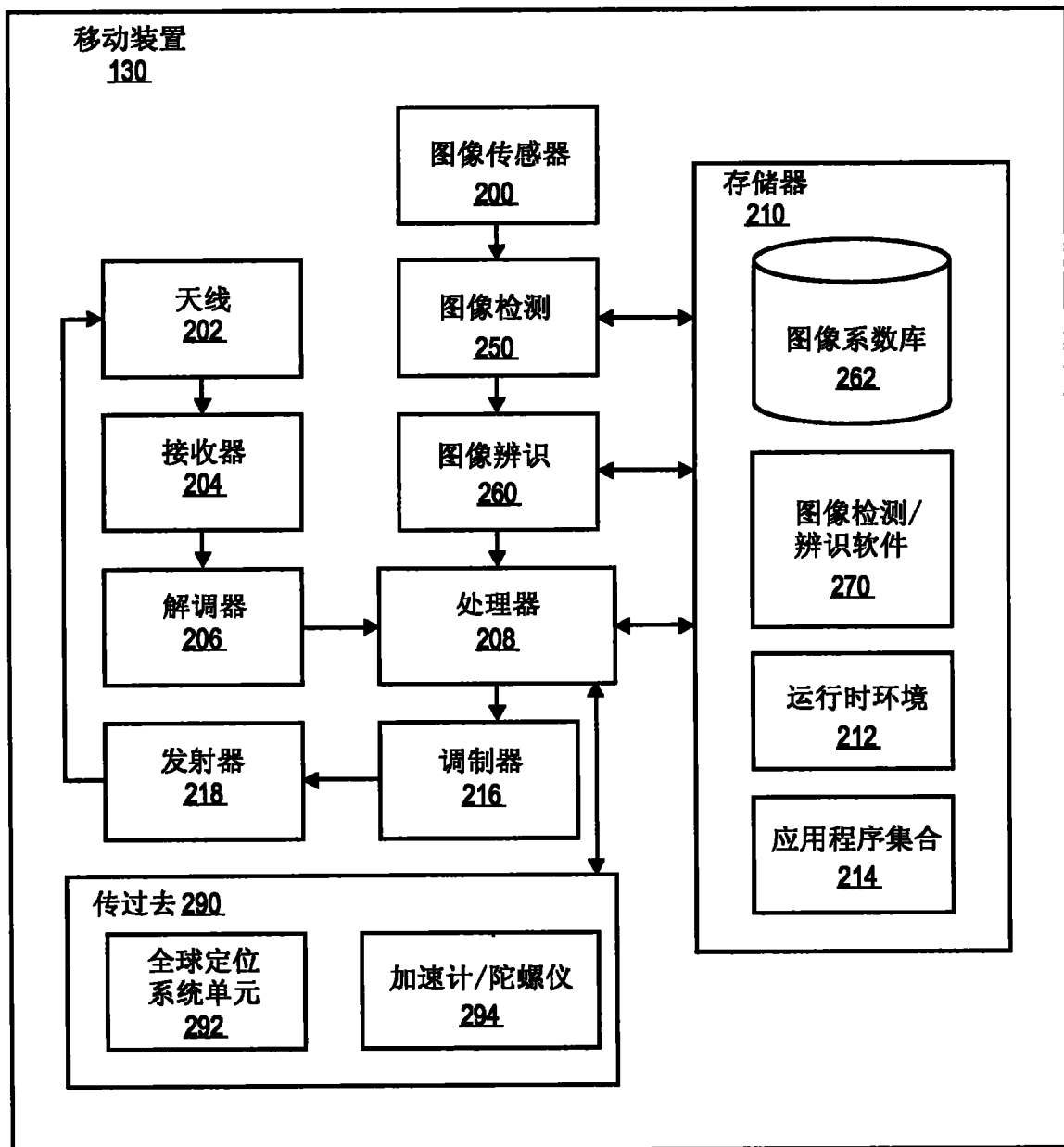


图 2

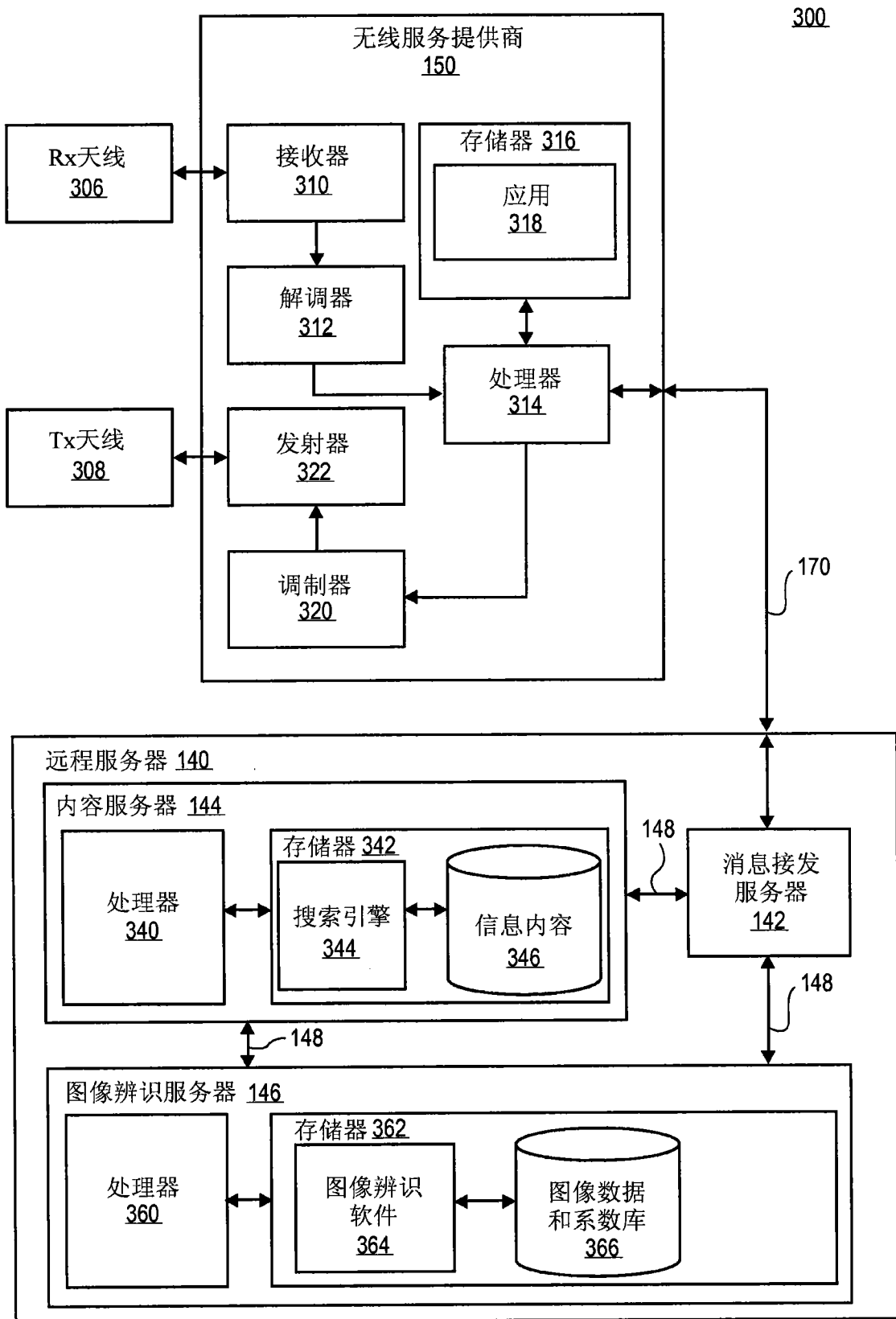


图 3

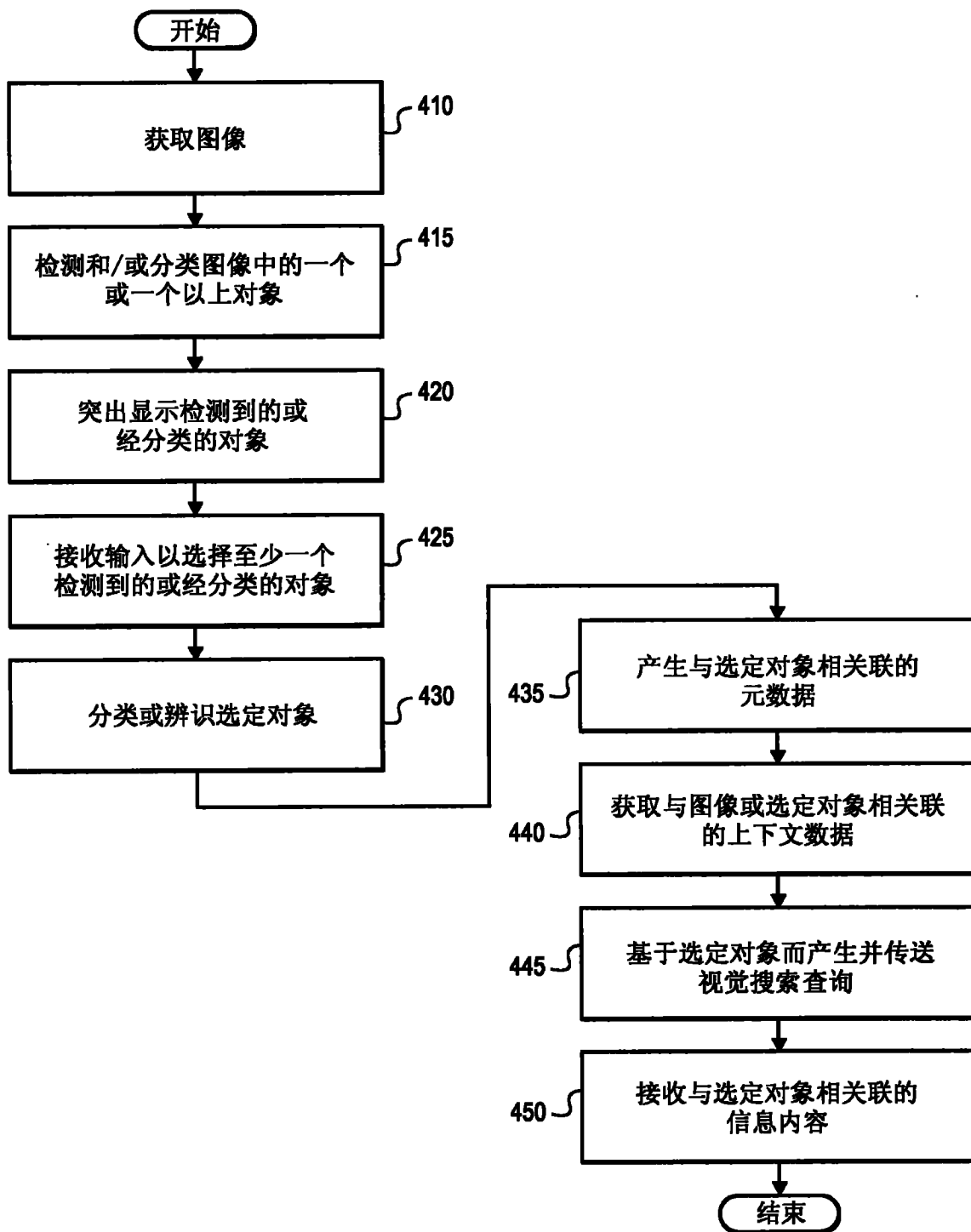


图 4

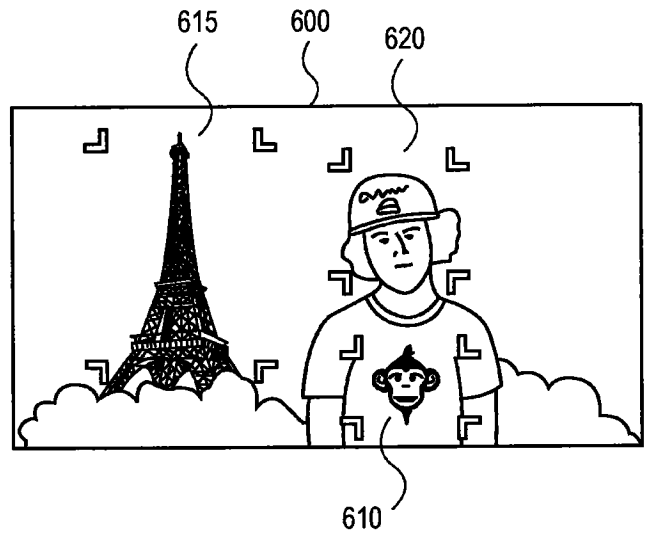
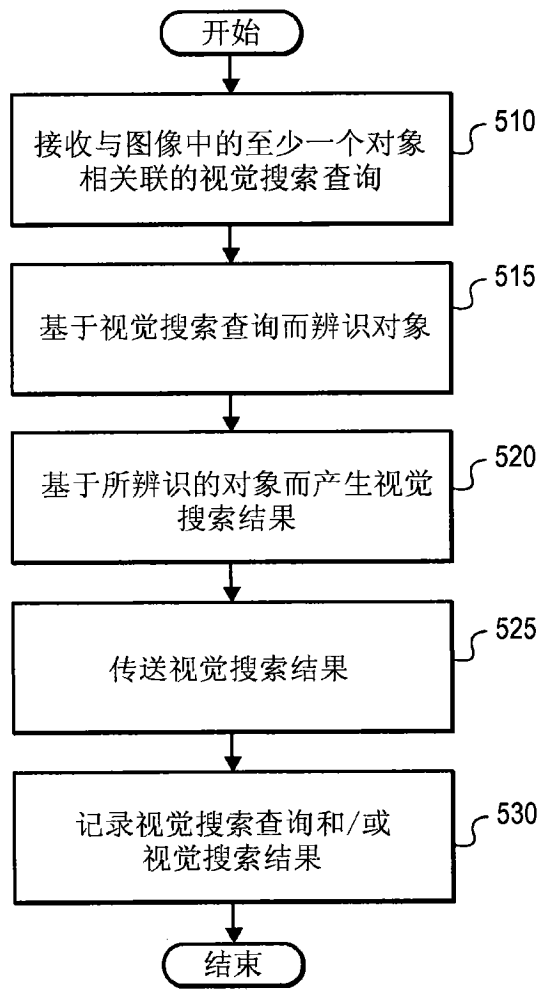


图 6A

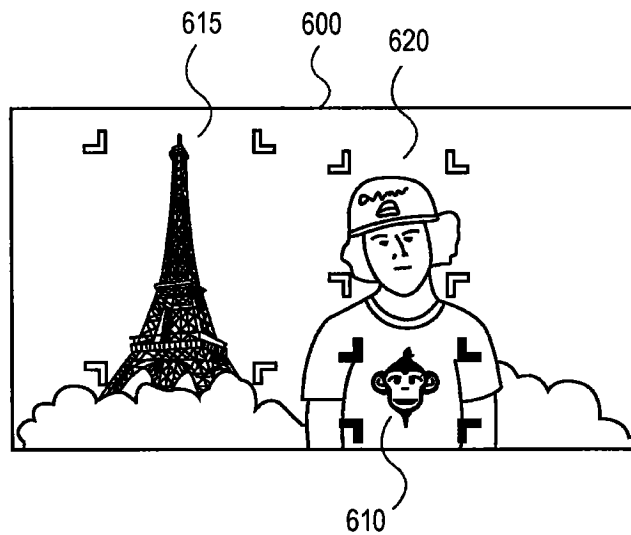


图 6B

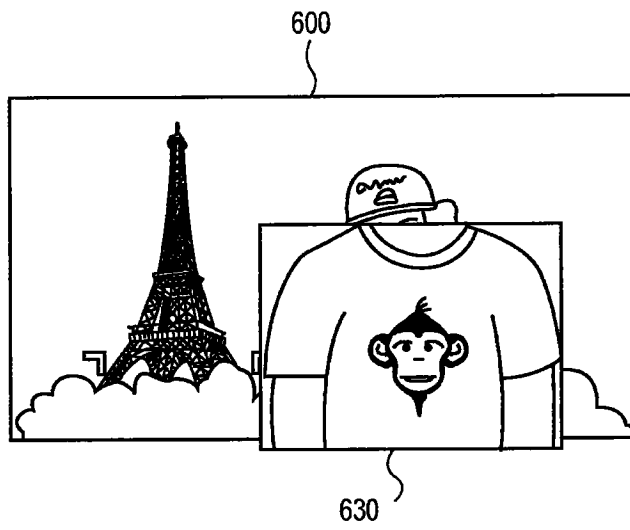


图 6C

650                      600

660 — brew 游戏  
 670 — BREW Gaming™: The Game Has Changed.™  
 670 — 服装  
 680 — 色彩和式样  
 680 — 在Cool-Shirt.com处\$15  
 690 — 由BREW Gaming™供电的移动手持机  
 670 — 评论  
 670 — 规格  
 695 — 在您附近的先锋地位电子处\$199-\$499

图 6D