



(12) 发明专利申请

(10) 申请公布号 CN 103582884 A

(43) 申请公布日 2014. 02. 12

(21) 申请号 201280018354. 8

重·U·李

(22) 申请日 2012. 04. 13

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287

(30) 优先权数据

代理人 宋献涛

61/475, 428 2011. 04. 14 US

13/312, 335 2011. 12. 06 US

(51) Int. Cl.

G06F 17/30 (2006. 01)

(85) PCT国际申请进入国家阶段日

G06K 9/62 (2006. 01)

2013. 10. 14

(86) PCT国际申请的申请数据

PCT/US2012/033620 2012. 04. 13

(87) PCT国际申请的公布数据

W02012/142483 EN 2012. 10. 18

(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 桑迪普·瓦达迪

奥努尔·C·哈姆西奇

尤里娅·列兹尼克 约翰·H·洪

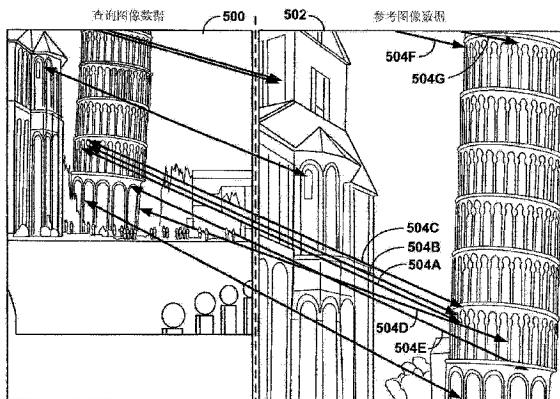
权利要求书9页 说明书17页 附图8页

(54) 发明名称

用于视觉搜索的稳健特征匹配

(57) 摘要

本发明揭示用于针对视觉搜索执行稳健特征匹配的技术。一种包括接口和特征匹配单元的设备可实施这些技术。所述接口接收查询特征描述符。所述特征匹配单元接着计算查询特征描述符与参考特征描述符之间的距离，且根据群集算法确定所述所计算距离的第一群组和所述所计算距离的第二群组，其中所计算距离的第一群组包括所述所计算距离中的两者或两者以上。所述特征匹配单元接着基于所述所计算距离的所述所确定第一群组和第二群组来确定所述查询特征描述符是否与所述参考特征描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配。



1. 一种用于利用视觉搜索装置执行视觉搜索的方法,所述方法包括:

利用所述视觉搜索装置来计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符中的每一者之间的距离,其中所述视觉搜索查询起始所述视觉搜索;

利用所述视觉搜索装置根据群集算法确定所述所计算距离中的一者或一者以上的第一群组和所述所计算距离的第二群组,其中所述所计算距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的相关联一者相对于所述所计算距离中被确定为在所计算距离的所述第二群组中的那些所述所计算距离接近所述查询特征描述符的那些所述所计算距离,且其中所述所计算距离的所述第二群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算距离中被确定为在所述所计算距离中的所述一者或一者以上的所述第一群组中的那些所述所计算距离远离所述查询特征描述符的那些所述所计算距离;以及

利用所述视觉搜索装置基于所述所确定的所述所计算距离的第一群组和所述所计算距离的所述第二群组来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配。

2. 根据权利要求 1 所述的方法,其中确定所述所计算距离中的一者或一者以上的所述第一群组和所述所计算距离的所述第二群组包含根据 k 均值群集算法、高斯拟合算法和图形切割算法中的一者或一者以上来确定所述所计算距离中的所述一者或一者以上的所述第一群组和所述所计算距离的所述第二群组。

3. 根据权利要求 1 所述的方法,其中确定所述查询特征描述符是否与所述多个参考特征描述符中的一者匹配包含:

计算所述所计算距离中被确定为在所述第一群组中的两者或两者以上的平均值,以产生第一群组距离平均值;

计算被确定为在所述第二群组中的所述所计算距离的平均值,以产生第二群组距离平均值;

将所述第一群组距离平均值除以所述第二群组距离平均值,以产生平均距离比率量度;

将所述平均距离比率量度与阈值进行比较;以及

基于所述比较来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的所述最小者相关联的所述一个参考特征描述符匹配。

4. 根据权利要求 1 所述的方法,其进一步包括:

确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算距离的那些参考特征描述符相关联的参考图像,使得所述组参考图像不包含重复的参考图像;

向被确定为在参考图像的所述唯一群组中的所述参考图像中的每一者指派选票;

基于所述所指派的选票将所述参考图像排序;以及

响应于所述视觉搜索查询而返回所述经排序的参考图像。

5. 根据权利要求 4 所述的方法,其中向被确定为在参考图像的所述唯一群组中的所述参考图像中的每一者指派选票包含以下各者中的一者或一者以上:

向所述参考图像中的每一者指派恒定选票;

指派与所述对应的所计算距离和所述查询特征描述符与所述多个参考特征描述符中的最接近一者之间的所计算距离相比的距离比率成比例的选票；以及

当所述第一群组的所述所计算距离从最小到最大被排序时，指派与所述第一群组内的所述对应的所计算距离的等级成比例的选票。

6. 根据权利要求 1 所述的方法，

其中所述查询特征描述符包括多个查询特征描述符中根据视觉搜索算法从查询图像数据提取且借助所述视觉搜索查询提供的一个查询特征描述符，且

其中所述方法进一步包括：

针对所述多个查询特征描述符中的每一者，计算所述多个查询特征描述符中的当前一者与所述多个参考特征描述符中的每一者之间的距离以产生所计算的多个距离；

针对所述多个距离中的每一者，根据所述群集算法确定所述所计算的多个距离的所述第一群组和所述所计算的多个距离的所述第二群组，其中所述所计算的多个距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所述所计算的多个距离的所述第二群组中的那些距离接近所述查询特征描述符的那些所述所计算距离，且其中所述所计算的多个距离的所述第二群组包含所述所计算的多个距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所计算的多个距离的所述第一群组中的那些距离远离所述多个查询特征描述符中的所述当前一者的那些距离；以及

利用所述视觉搜索装置基于所述所确定的所述所计算的多个距离的第一群组和所述所计算的多个距离的所述第二群组来确定所述多个查询特征描述符中的每一者是否与所述多个参考特征描述符中与所述所计算的多个距离中的最小者相关联的一个参考特征描述符匹配。

7. 根据权利要求 6 所述的方法，其进一步包括：

针对所述多个查询特征描述符中的每一者确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算的多个距离的那些参考特征描述符相关联的参考图像，使得所述组参考图像不包含重复的参考图像；

向被确定为在针对所述多个查询特征描述符中的每一者而确定的参考图像的所述唯一群组中的所述参考图像中的每一者指派选票；

在相对于所述多个查询特征描述符中的每一者指派所述选票之后，基于所述所指派的选票将所述参考图像排序；以及

响应于所述视觉搜索查询而返回所述经排序的参考图像。

8. 根据权利要求 1 所述的方法，其中根据基于局部特征的视觉搜索算法从查询图像提取所述查询特征描述符。

9. 根据权利要求 8 所述的方法，其中所述基于局部特征的视觉搜索算法包括尺度不变特征变换 SIFT 算法。

10. 根据权利要求 1 所述的方法，其进一步包括：

经由所述视觉搜索装置的接口从客户端装置接收所述查询特征描述符作为经压缩查询特征描述符；以及

利用所述视觉搜索装置从所述经压缩查询特征描述符重构所述查询特征描述符。

11. 一种用于执行视觉搜索的设备,所述设备包括:

用于计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符中的每一者之间的距离的装置,其中所述视觉搜索查询起始所述视觉搜索;

用于根据群集算法确定所述所计算距离中的一者或一者以上的第一群组和所述所计算距离的第二群组的装置,其中所述所计算距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的相关联一者相对于所述所计算距离中被确定为在所述所计算距离的所述第二群组中的那些所述所计算距离接近所述查询特征描述符的那些所述所计算距离,且其中所述所计算距离的所述第二群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算距离中被确定为在所述所计算距离中的所述一者或一者以上的所述第一群组中的那些所述所计算距离远离所述查询特征描述符的那些所述所计算距离;以及

用于基于所述所确定的所述所计算距离的第一群组和所述所计算距离的所述第二群组来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配的装置。

12. 根据权利要求 11 所述的设备,其进一步包括用于根据 k 均值群集算法、高斯拟合算法和图形切割算法中的一者或一者以上来确定所述所计算距离中的所述一者或一者以上的所述第一群组和所述所计算距离的所述第二群组的装置。

13. 根据权利要求 11 所述的设备,其进一步包括:

用于计算所述所计算距离中被确定为在所述第一群组中的两者或两者以上的平均值以产生第一群组距离平均值的装置;

用于计算被确定为在所述第二群组中的所述所计算距离的平均值以产生第二群组距离平均值的装置;

用于将所述第一群组距离平均值除以所述第二群组距离平均值以产生平均距离比率量度的装置;

用于将所述平均距离比率量度与阈值进行比较的装置;以及

用于基于所述比较来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的所述最小者相关联的所述一个参考特征描述符匹配的装置。

14. 根据权利要求 11 所述的设备,其进一步包括:

用于确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算距离的那些参考特征描述符相关联的参考图像以使得所述组参考图像不包含重复的参考图像的装置;

用于向被确定为在参考图像的所述唯一群组中的所述参考图像中的每一者指派选票的装置;

用于基于所述所指派的选票将所述参考图像排序的装置;以及

用于响应于所述视觉搜索查询而返回所述经排序的参考图像的装置。

15. 根据权利要求 14 所述的设备,其进一步包括以下各者中的一者或一者以上:

用于向所述参考图像中的每一者指派恒定选票的装置;

用于指派与所述对应的所计算距离和所述查询特征描述符与所述多个参考特征描述

符中的最接近一者之间的所计算距离相比的距离比率成比例的选票的装置；以及

用于当所述第一群组的所述所计算距离从最小到最大被排序时指派与所述第一群组内的所述对应的所计算距离的等级成比例的选票的装置。

16. 根据权利要求 11 所述的设备，

其中所述查询特征描述符包括多个查询特征描述符中根据视觉搜索算法从查询图像数据提取且借助所述视觉搜索查询提供的一个查询特征描述符，且

其中所述设备进一步包括：

用于针对所述多个距离中的每一者计算所述多个查询特征描述符中的当前一者与所述多个参考特征描述符中的每一者之间的距离以产生所计算的多个距离的装置；

用于针对所述多个距离中的每一者根据所述群集算法确定所述所计算的多个距离的所述第一群组和所述所计算的多个距离的所述第二群组的装置，其中所述所计算的多个距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所述所计算的多个距离的所述第二群组中的那些距离接近所述查询特征描述符的那些所述所计算距离，且其中所述所计算的多个距离的所述第二群组包含所述所计算的多个距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所计算的多个距离的所述第一群组中的那些距离远离所述多个查询特征描述符中的所述当前一者的那些距离；以及

用于基于所述所确定的所述所计算的多个距离的第一群组和所述所计算的多个距离的所述第二群组来确定所述多个查询特征描述符中的每一者是否与所述多个参考特征描述符中与所述所计算的多个距离中的最小者相关联的一个参考特征描述符匹配的装置。

17. 根据权利要求 16 所述的设备，其进一步包括：

用于针对所述多个查询特征描述符中的每一者确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算的多个距离的那些参考特征描述符相关联的参考图像以使得所述组参考图像不包含重复的参考图像的装置；

用于向被确定为在针对所述多个查询特征描述符中的每一者而确定的参考图像的所述唯一群组中的所述参考图像中的每一者指派选票的装置；

用于在相对于所述多个查询特征描述符中的每一者指派所述选票之后基于所述所指派的选票将所述参考图像排序的装置；以及

用于响应于所述视觉搜索查询而返回所述经排序的参考图像的装置。

18. 根据权利要求 11 所述的设备，其中所述查询特征描述符是根据基于局部特征的视觉搜索算法从查询图像提取的。

19. 根据权利要求 18 所述的设备，其中所述基于局部特征的视觉搜索算法包括尺度不变特征变换 SIFT 算法。

20. 根据权利要求 11 所述的设备，其进一步包括：

用于经由视觉搜索装置的接口从客户端装置接收所述查询特征描述符作为经压缩查询特征描述符的装置；以及

用于利用所述视觉搜索装置从所述经压缩查询特征描述符重构所述查询特征描述符的装置。

21. 一种经配置以执行视觉搜索的设备,所述设备包括:

接口,其经配置以接收查询特征描述符;

特征匹配单元,其:经配置以计算借助视觉搜索查询提供的所述查询特征描述符与多个参考特征描述符中的每一者之间的距离,其中所述视觉搜索查询起始所述视觉搜索;以及经配置以根据群集算法确定所述所计算距离中的一者或一者以上的第一群组和所述所计算距离的第二群组,其中所述所计算距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的相关联一者相对于所述所计算距离中被确定为在所述所计算距离的所述第二群组中的那些所述所计算距离接近所述查询特征描述符的那些所述所计算距离,且其中所述所计算距离的所述第二群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算距离中被确定为在所述所计算距离中的所述一者或一者以上的所述第一群组中的那些所述所计算距离远离所述查询特征描述符的那些所述所计算距离,且其中所述特征匹配单元进一步经配置以基于所述所确定的所述所计算距离的第一群组和所述所计算距离的所述第二群组来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配。

22. 根据权利要求 21 所述的设备,其中所述特征匹配单元经配置以根据 k 均值群集算法、高斯拟合算法和图形切割算法中的一者或一者以上来确定所述所计算距离中的所述一者或一者以上的所述第一群组和所述所计算距离的所述第二群组。

23. 根据权利要求 21 所述的设备,其中所述特征匹配单元进一步经配置以:计算所述所计算距离中被确定为在所述第一群组中的两者或两者以上的平均值,以产生第一群组距离平均值;计算被确定为在所述第二群组中的所述所计算距离的平均值,以产生第二群组距离平均值;将所述第一群组距离平均值除以所述第二群组距离平均值,以产生平均距离比率量度;以及将所述平均距离比率量度与阈值进行比较且基于所述比较来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的所述最小者相关联的所述一个参考特征描述符匹配。

24. 根据权利要求 21 所述的设备,其中所述特征匹配单元进一步经配置以:

确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算距离的那些参考特征描述符相关联的参考图像,使得所述组参考图像不包含重复的参考图像;

向被确定为在参考图像的所述唯一群组中的所述参考图像中的每一者指派选票;

基于所述所指派的选票将所述参考图像排序;以及

响应于所述视觉搜索查询而返回所述经排序的参考图像。

25. 根据权利要求 24 所述的设备,其中所述特征匹配单元经配置以通过以下各者中的一者或一者以上指派所述选票:向所述参考图像中的每一者指派恒定选票;指派与所述对应的所计算距离和所述查询特征描述符与所述多个参考特征描述符中的最接近一者之间的所计算距离相比的距离比率成比例的选票;以及当所述第一群组的所述所计算距离从最小到最大被排序时指派与所述第一群组内的所述对应的所计算距离的等级成比例的选票。

26. 根据权利要求 21 所述的设备,

其中所述查询特征描述符包括多个查询特征描述符中根据视觉搜索算法从查询图像

数据提取且借助所述视觉搜索查询提供的一个查询特征描述符，且

其中所述特征匹配单元进一步经配置以：

针对所述多个查询特征描述符中的每一者，计算所述多个查询特征描述符中的当前一者与所述多个参考特征描述符中的每一者之间的距离以产生所计算的多个距离；

针对所述多个距离中的每一者，根据所述群集算法确定所述所计算的多个距离的所述第一群组和所述所计算的多个距离的所述第二群组，其中所述所计算的多个距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所述所计算的多个距离的所述第二群组中的那些距离接近所述查询特征描述符的那些所述所计算距离，且其中所述所计算的多个距离的所述第二群组包含所述所计算的多个距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所计算的多个距离的所述第一群组中的那些距离远离所述多个查询特征描述符中的所述当前一者的那些距离；以及

基于所述所确定的所述所计算的多个距离的第一群组和所述所计算的多个距离的所述第二群组来确定所述多个查询特征描述符中的每一者是否与所述多个参考特征描述符中与所述所计算的多个距离中的最小者相关联的一个参考特征描述符匹配。

27. 根据权利要求 26 所述的设备，其中所述特征匹配单元进一步经配置以：

针对所述多个查询特征描述符中的每一者确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算的多个距离的那些参考特征描述符相关联的参考图像，使得所述组参考图像不包含重复的参考图像；

向被确定为在针对所述多个查询特征描述符中的每一者而确定的参考图像的所述唯一群组中的所述参考图像中的每一者指派选票，在相对于所述多个查询特征描述符中的每一者指派所述选票之后基于所述所指派的选票将所述参考图像排序；且

其中所述接口响应于所述视觉搜索查询而返回所述经排序的参考图像。

28. 根据权利要求 21 所述的设备，其中所述查询特征描述符是根据基于局部特征的视觉搜索算法从查询图像提取的。

29. 根据权利要求 28 所述的设备，其中所述基于局部特征的视觉搜索算法包括尺度不变特征变换 SIFT 算法。

30. 根据权利要求 21 所述的设备，

其中所述接口经配置以从客户端装置接收作为经压缩查询特征描述符的所述查询特征描述符；且

其中所述设备包含特征重构单元，所述特征重构单元经配置以利用视觉搜索装置从所述经压缩查询特征描述符重构所述查询特征描述符。

31. 一种包括指令的计算机可读媒体，所述指令在被执行时致使一个或一个以上处理器：

计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符中的每一者之间的距离，其中所述视觉搜索查询起始视觉搜索；

根据群集算法确定所述所计算距离中的一者或一者以上的第一群组和所述所计算距离的第二群组，其中所述所计算距离的所述第一群组包含所述所计算距离中指示所述多个

参考特征描述符中的相关联一者相对于所述所计算距离中被确定为在所计算距离的所述第二群组中的那些所述所计算距离接近所述查询特征描述符的那些所述所计算距离,且其中所述所计算距离的所述第二群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算距离中被确定为在所述所计算距离中的所述一者或一者以上的所述第一群组中的那些所述所计算距离远离所述查询特征描述符的那些所述所计算距离;以及

基于所述所确定的所述所计算距离的第一群组和所述所计算距离的所述第二群组来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配。

32. 根据权利要求 31 所述的计算机可读媒体,其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令:根据 k 均值群集算法、高斯拟合算法和图形切割算法中的一者或一者以上来确定所述所计算距离中的所述一者或一者以上的所述第一群组和所述所计算距离的所述第二群组。

33. 根据权利要求 31 所述的计算机可读媒体,其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令:

计算所述所计算距离中被确定为在所述第一群组中的两者或两者以上的平均值,以产生第一群组距离平均值;

计算被确定为在所述第二群组中的所述所计算距离的平均值,以产生第二群组距离平均值;

将所述第一群组距离平均值除以所述第二群组距离平均值,以产生平均距离比率量度;

将所述平均距离比率量度与阈值进行比较;以及

基于所述比较来确定所述查询特征描述符是否与所述多个参考特征描述符中与所述所计算距离中的所述最小者相关联的所述一个参考特征描述符匹配。

34. 根据权利要求 31 所述的计算机可读媒体,其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令:

确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算距离的那些参考特征描述符相关联的参考图像,使得所述组参考图像不包含重复的参考图像;

向被确定为在参考图像的所述唯一群组中的所述参考图像中的每一者指派选票;

基于所述所指派的选票将所述参考图像排序;以及

响应于所述视觉搜索查询而返回所述经排序的参考图像。

35. 根据权利要求 34 所述的计算机可读媒体,其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令:

向所述参考图像中的每一者指派恒定选票;

指派与所述对应的所计算距离和所述查询特征描述符与所述多个参考特征描述符中的最接近一者之间的所计算距离相比的距离比率成比例的选票;以及

当所述第一群组的所述所计算距离从最小到最大被排序时,指派与所述第一群组内的所述对应的所计算距离的等级成比例的选票。

36. 根据权利要求 31 所述的计算机可读媒体，

其中所述查询特征描述符包括多个查询特征描述符中根据视觉搜索算法从查询图像数据提取且借助所述视觉搜索查询提供的一个查询特征描述符，

其中所述计算机可读媒体进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令：

针对所述多个查询特征描述符中的每一者，计算所述多个查询特征描述符中的当前一者与所述多个参考特征描述符中的每一者之间的距离以产生所计算的多个距离；

针对所述多个距离中的每一者，根据所述群集算法确定所述所计算的多个距离的所述第一群组和所述所计算的多个距离的所述第二群组，其中所述所计算的多个距离的所述第一群组包含所述所计算距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所述所计算的多个距离的所述第二群组中的那些距离接近所述查询特征描述符的那些所述所计算距离，且其中所述所计算的多个距离的所述第二群组包含所述所计算的多个距离中指示所述多个参考特征描述符中的所述相关联一者相对于所述所计算的多个距离中被确定为在所计算的多个距离的所述第一群组中的那些距离远离所述多个查询特征描述符中的所述当前一者的那些距离；以及

基于所述所确定的所述所计算的多个距离的第一群组和所述所计算的多个距离的所述第二群组来确定所述多个查询特征描述符中的每一者是否与所述多个参考特征描述符中与所述所计算的多个距离中的最小者相关联的一个参考特征描述符匹配。

37. 根据权利要求 36 所述的计算机可读媒体，其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令：

针对所述多个查询特征描述符中的每一者确定参考图像的唯一群组以包含与所述多个参考特征描述符中从其根据参考特征描述符的对应群组计算被确定为在所述第一群组中的所述所计算的多个距离的那些参考特征描述符相关联的参考图像，使得所述组参考图像不包含重复的参考图像；

向被确定为在针对所述多个查询特征描述符中的每一者而确定的参考图像的所述唯一群组中的所述参考图像中的每一者指派选票；

在相对于所述多个查询特征描述符中的每一者指派所述选票之后，基于所述所指派的选票将所述参考图像排序；以及

响应于所述视觉搜索查询而返回所述经排序的参考图像。

38. 根据权利要求 31 所述的计算机可读媒体，其中所述查询特征描述符是根据基于局部特征的视觉搜索算法从查询图像提取的。

39. 根据权利要求 38 所述的计算机可读媒体，其中所述基于局部特征的视觉搜索算法包括尺度不变特征变换 SIFT 算法。

40. 根据权利要求 31 所述的计算机可读媒体，其进一步包括在被执行时致使所述一个或一个以上处理器进行以下操作的指令：

经由视觉搜索装置的接口从客户端装置接收作为经压缩查询特征描述符的所述查询特征描述符；以及

利用所述视觉搜索装置从所述经压缩查询特征描述符重构所述查询特征描述符。

41. 一种系统，其包括：

用以接收从一个或一个以上客户端装置接收的信息的输入,其中所述信息包括用以起始视觉搜索的查询特征描述符;

数据库,其包含指示多个参考查询描述符的信息;以及

视觉搜索服务器装置,其经配置以执行所述视觉搜索,所述视觉搜索服务器装置包括:

接口,其用以接收所述查询特征描述符;以及

特征匹配单元,其经配置以:

计算所述查询特征描述符与所述多个参考查询描述符中的每一者之间的距离;

根据群集算法确定所述所计算距离中的一者或一者以上的第一群组和所述所计算距离的第二群组,其中所述所计算距离的所述第一群组包含所述所计算距离中指示所述多个参考查询描述符中的相关联一者相对于所述所计算距离中被确定为在所述所计算距离的所述第二群组中的那些所述所计算距离接近所述查询特征描述符的那些所述所计算距离,且其中所述所计算距离的所述第二群组包含所述所计算距离中指示所述多个参考查询描述符中的所述相关联一者相对于所述所计算距离中被确定为在所述所计算距离中的所述一者或一者以上的所述第一群组中的那些所述所计算距离远离所述查询特征描述符的那些所述所计算距离;以及

基于所述所确定的所述所计算距离的第一群组和所述所计算距离的所述第二群组来确定所述查询特征描述符是否与所述多个参考查询描述符中与所述所计算距离中的最小者相关联的一个参考特征描述符匹配。

## 用于视觉搜索的稳健特征匹配

### 技术领域

#### 背景技术

- [0001] 本发明涉及图像处理系统,且更明确地说,涉及用图像处理系统执行视觉搜索。
- [0002] 在计算装置或计算机的背景下的视觉搜索涉及使计算机或其它装置能够在一个或一个图像内的其它对象和 / 或特征之中执行对对象和 / 或特征的搜索的技术。对视觉搜索的最近兴趣已导致使计算机能够在各种各样的不同图像条件(包含图像尺度、噪声、照度和局部几何失真的变化)下识别部分遮蔽对象和 / 或特征的算法。在此相同时间期间,移动装置已出现了特征相机,但其可具有受限的用户接口来输入文本或以其它方式与移动装置介接。移动装置和移动装置应用程序的开发者已试图利用移动装置的相机来增强与移动装置的用户交互。
- [0003] 为了说明一种增强,移动装置的用户可利用移动装置的相机来在商店购物时捕获任何给定产品的图像。移动装置可接着在用于各种参考图像的一组所存档特征描述符内起始视觉搜索算法,而基于匹配的参考图像来识别图像(其可称为“搜索图像”)中所展示的产品。在识别所述产品之后,移动装置可接着起始因特网的搜索,且呈现关于所识别产品的含有网页的信息,包含从附近商户和 / 或在线商户可购买所述产品的最低价格。以此方式,用户可避免必须经由键盘(其在作为用户与此相互作用的图像呈现于触摸屏上的意义上,通常为“虚拟”键盘)或其它输入机构与移动装置介接,但可仅捕获搜索图像来起始视觉搜索和后续网络搜索。
- [0004] 虽然存在配备有相机且对视觉搜索的接入权的移动装置可使用的若干应用程序,但用于实施视觉搜索的视觉搜索算法(例如尺度不变特征变换(SIFT)算法)在执行特征匹配方面可能是有缺陷的。特征匹配涉及视觉搜索算法的一方面,在此期间,对照从参考图像提取的参考特征描述符,匹配从搜索图像提取的搜索特征描述符。
- [0005] 为了说明这些缺陷,考虑SIFT算法,其在搜索特征描述符和参考特征描述符各自从搜索和参考图像的重复特征(例如跨建筑物重复的独特的拱门或窗户)提取时的例子中,可丢弃原本将与搜索特征描述符匹配的参考特征描述符。此外,SIFT算法通常响应于任何给定视觉搜索而仅返回单个图像,其中通过SIFT算法在算法上将此所返回的图像确定为“最佳匹配”。然而,用户无法以与SIFT算法相同的方式确定什么构成“最佳匹配”,这可导致用户受挫,因为单个SIFT最佳匹配结果无法匹配用户的预期。

### 发明内容

- [0006] 大体来说,本发明描述促进在执行视觉搜索时的特征匹配的技术。所述技术可通过提供能够提供搜索结果的等级列表的特征匹配算法来改进特征匹配。所述技术可通过提供适应重复特征的特征匹配算法来改进稳健性。不是在与当前搜索特征描述符匹配的两个或两个以上参考特征定位成彼此接近时(如由例如SIFT算法提供(在两个或两个以上匹配的参考特征描述符太靠近的前提下,这些匹配的参考特征描述符不可能是唯一的且因此

不可能促进搜索图像的分类)) 否决参考特征描述符，所述技术可利用群集算法来更适当地确定在执行特征匹配时的唯一性。此外，所述技术可促进匹配的参考图像的等级列表的产生和返回，而非如常规视觉搜索算法中常见地简单地返回单一参考图像。匹配的参考图像的等级列表可向用户提供选择用户认为“最佳匹配”的内容的机会，而非被迫接受关于构成最佳匹配的内容的算法确定。

[0007] 在一个实例中，一种用于利用视觉搜索装置执行视觉搜索的方法，所述方法包括利用视觉搜索装置计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符的每一者之间的距离，其中所述视觉搜索查询起始视觉搜索。所述方法还包括利用视觉搜索装置根据群集算法确定所计算距离的一者或一者以上的第一群组和所计算距离的第二群组，其中所计算距离的第一群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离，且其中所计算距离的第二群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符的那些所计算距离。所述方法进一步包括利用视觉搜索装置基于所确定的所计算距离的第一群组和所计算距离的第二群组确定查询特征描述符是否与所述多个参考特征描述符中的与所计算距离中的最小者相关联的一个参考特征描述符匹配。

[0008] 在另一实例中，一种用于执行视觉搜索的设备包括用于计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符的每一者之间的距离的装置，其中所述视觉搜索查询起始视觉搜索。所述设备还包括用于根据群集算法确定所计算距离的一者或一者以上的第一群组和所计算距离的第二群组的装置，其中所计算距离的第一群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离，且其中所计算距离的第二群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符的那些所计算距离。所述设备进一步包括用于基于所确定的所计算距离的第一群组和所计算距离的第二群组确定查询特征描述符是否与所述多个参考特征描述符中与所计算距离中的最小者相关联的一个参考特征描述符匹配的装置。

[0009] 在另一实例中，一种用于执行视觉搜索的设备包括：接口，其接收查询特征描述符；以及特征匹配单元，所述特征匹配单元：计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符的每一者之间的距离，其中所述视觉搜索查询起始视觉搜索；根据群集算法确定所计算距离的一者或一者以上的第一群组和所计算距离的第二群组，其中所计算距离的第一群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离，且其中所计算距离的第二群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符的那些所计算距离；以及基于所确定的所计算距离的第一群组和所计算距离的第二群组确定查询特征描述符是否与所述多个参考特征描述符中与所计算距离中的最小者相关联的一个参考特征描述符匹配。

[0010] 在另一实例中，一种包括指令的计算机可读媒体，所述指令在被执行时致使一个

或一个以上处理器：计算借助视觉搜索查询提供的查询特征描述符与多个参考特征描述符的每一者之间的距离，其中所述视觉搜索查询起始视觉搜索；根据群集算法确定所计算距离的一者或一者以上的第一群组和所计算距离的第二群组，其中所计算距离的第一群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离，且其中所计算距离的第二群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符的那些所计算距离；以及基于所确定的所计算距离的第一群组和所计算距离的第二群组确定查询特征描述符是否与所述多个参考特征描述符中与所计算距离中的最小者相关联的一个参考特征描述符匹配。

[0011] 在另一实例中，一种系统包括：客户端装置，其借助搜索查询发射查询特征描述符以起始视觉搜索；数据库，其存储多个参考查询描述符；以及视觉搜索服务器装置，其执行视觉搜索。所述视觉搜索服务器装置包括：接口，其借助搜索查询接收查询特征描述符；以及特征匹配单元，所述特征匹配单元：计算查询特征描述符与所述多个参考特征描述符的每一者之间的距离；根据群集算法确定所计算距离中的一者或一者以上的第一群组和所计算距离的第二群组，其中所计算距离的第一群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离，且其中所计算距离的第二群组包含指示所述多个参考特征描述符中的相关联一者相对于被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符的那些所计算距离；以及基于所确定的所计算距离的第一群组和所计算距离的第二群组确定查询特征描述符是否与所述多个参考特征描述符的与所计算距离中的最小者相关联的一个参考特征描述符匹配。

[0012] 一个或一个以上实例的细节陈述于附图和以下描述中。其它特征、目标和优点将从所述描述和图式以及从权利要求书中显而易见。

## 附图说明

[0013] 图 1 是说明实施本发明中所描述的稳健特征描述符匹配技术的图像处理系统的框图。

[0014] 图 2 是更详细地说明图 1 的特征匹配单元的框图。

[0015] 图 3A 和 3B 是说明实施本发明所描述的特征匹配技术中的视觉搜索服务器的示范性操作的流程图。

[0016] 图 4 是说明特征提取单元借以确定高斯差 (DoG) 金字塔以供执行关键点提取中使用的过程的图。

[0017] 图 5 是说明在确定高斯差 (DoG) 金字塔之后的关键点的检测的图。

[0018] 图 6 是说明特征提取单元借以确定梯度分布和定向直方图的过程的图。

[0019] 图 7 是说明特征提取单元借以确定梯度分布和定向直方图的过程的图。

## 具体实施方式

[0020] 图 1 是说明实施本发明中所描述的稳健特征匹配技术的图像处理系统 10 的框图。在图 1 的实例中，图像处理系统 10 包含客户端装置 12、视觉搜索服务器 14 和网络 16。客户

端装置 12 在此实例中表示移动装置,例如膝上型计算机、所谓的笔记本型计算机、个人数字助理 (PDA)、蜂窝式或移动电话或手持机(包含所谓的“智能电话”)、全球定位系统 (GPS) 装置、数码相机、数字媒体播放器、游戏装置,或能够与视觉搜索服务器 14 通信的任何其它移动装置。虽然本发明中相对于视觉搜索服务器 14 来描述,但本发明中所描述的技术不应限于视觉搜索服务器。而是,所述技术可由能够实施基于局部特征的视觉搜索算法的特征匹配方面的任何装置来实施。

[0021] 视觉搜索服务器 14 表示服务器装置,其接受通常呈传输控制协议 (TCP) 连接的形式的连接,且以其自己的 TCP 连接来响应,以形成借以接收查询数据且提供识别数据的 TCP 会话。视觉搜索服务器 14 可表示视觉搜索服务器装置,其中视觉搜索服务器 14 执行或以其它方式实施基于局部特征的视觉搜索算法,以识别一个或一个以上参考图像内的一个或一个以上特征或对象。

[0022] 网络 16 表示公共网络,例如因特网,其互连客户端装置 12 与视觉搜索服务器 14,但网络 16 还可为私人网络。通常,网络 16 实施开放系统互连 (OSI) 模型的各个层以促进通信或数据在客户端装置 12 与视觉搜索服务器 14 之间的传送。网络 16 通常包含任何数目的网络装置,例如交换机、集线器、路由器、服务器,以实现数据在客户端装置 12 与视觉搜索服务器 14 之间的传送。虽然展示为单个网络,但网络 16 可包括经互连以形成网络 16 的一个或一个以上子网络。这些子网络可包括服务提供商网络、接入网络、后端网络或公共网络中通常使用的任何其它类型的网络,以实现遍及网络 16 的数据传送。虽然在此实例中描述为公共网络,但网络 16 可包括公众通常不可接入的私人网络。

[0023] 如图 1 的实例中所示,客户端装置 12 包含特征提取单元 18、特征压缩单元 20、接口 22 和显示器 24。特征提取单元 18 表示根据特征提取算法(例如尺度不变特征变换算法,或任何其它用于提取特征的特征描述提取算法)执行特征提取的单元。通常,特征提取单元 18 对图像数据 26 进行操作,可使用包含于客户端装置 12 内的相机或其它图像捕获装置(图 1 的实例中未展示)在本地捕获图像数据 26。或者,客户端装置 12 可借助于从网络 16、在本地经由与另一计算装置的有线连接,或经由任何其它有线或无线形式的通信下载此图像数据 26 来存储图像数据 26,而不自己捕获此图像数据。

[0024] 虽然下文更详细地描述,但概括地说,特征提取单元 18 可通过高斯模糊图像数据 26 来提取特征描述符 28,以产生两个连续的经高斯模糊的图像。高斯模糊通常涉及以所定义的尺度用高斯模糊函数来卷积图像数据 26。特征提取单元 18 可递增地卷积图像数据 26,其中通过尺度空间中的常数将所得经高斯模糊图像彼此分离。特征提取单元 18 可堆叠这些经高斯模糊的图像,以形成可称为“高斯金字塔”或“高斯差金字塔”的内容。特征提取单元 18 将两个连续堆叠的经高斯模糊的图像进行比较,以产生高斯差 (DoG) 图像。DoG 图像可形成称为“DoG 空间”的内容。

[0025] 基于此 DoG 空间,特征提取单元 18 可检测关键点,其中关键点指代特定样本点周围的像素区或小块,或图像数据 26 中从几何观点看潜在感兴趣的像素。通常,特征提取单元 18 将关键点识别为所构造的 DoG 空间中的局部最大值和 / 或局部最小值。特征提取单元 18 接着基于曾检测到关键点的小块的局部图像梯度的方向来为这些关键点指派一个或一个以上定向或方向。为了表征这些定向,特征提取单元 18 可依据梯度定向直方图来界定定向。特征提取单元 18 接着将特征描述符 28 界定为位置和定向(例如,借助于梯度定向

直方图)。在界定特征描述符 28 之后,特征提取单元 18 将此特征描述符 28 输出到特征压缩单元 20。通常,特征提取单元 18 使用此过程将一组特征描述符 28 输出到特征压缩单元 20。

[0026] 特征压缩单元 20 表示压缩或以其它方式相对于特征提取单元 18 用来界定这些特征描述符的数据量来减少用以界定特征描述符(例如特征描述符 28)的数据量的单元。为了压缩特征描述符 28,特征压缩单元 20 可执行一种形式的量化(称为类型量化)来压缩特征描述符 28。在此方面中,不是整体发送由特征描述符 28 界定的直方图,而是特征压缩单元 20 执行类型量化来将直方图表示为所谓的“类型”。通常,一种类型为直方图的经压缩表示(例如,在所述类型表示直方图的形状而不是整个直方图的情况下)。所述类型通常表示符号的一组频率,且在直方图的上下文中,可表示直方图的梯度分布的频率。换句话说,类型可表示产生特征描述符 28 中的对应一者的来源的真实分布的估计。在此方面中,所述类型的编码和发射可被视为等效于编码和发射分布的形状,因为其可基于特定样本来估计(即,在此实例中,其为由特征描述符 28 中的对应一者界定的直方图)。

[0027] 在给定特征描述符 28 和量化等级(其可在本文中在数学上表示为“n”)的情况下,特征压缩单元 20 为特征描述符 28 中的每一者计算具有参数  $k_1, \dots, k_m$ (其中  $m$  表示维度的数目)的类型。每一类型可表示具有给定公分母的一组有理数,其中所述有理数总计为一。特征描述符 28 可接着使用字典式枚举来将此类型编码为索引。换句话说,对于具有给定公分母的所有可能类型,特征压缩单元 20 基于这些类型的字典式排序将索引有效地指派给这些类型中的每一者。特征压缩单元 20 进而将特征描述符 28 压缩为单个经字典式布置的索引,且以查询数据 30 的形式将这些经压缩的特征描述符输出到接口 22。

[0028] 接口 22 表示能够经由网络 16 与视觉搜索服务器 14 通信的任何类型的接口,包含无线接口和有线接口。接口 22 可表示无线蜂窝式接口,且包含必要的硬件或其它组件,例如天线、调制器等,以经由无线蜂窝式网络与网络 16 通信,且经由网络 16 与视觉搜索服务器 14 通信。在此例子中,尽管图 1 的实例中未展示,但网络 16 包含无线蜂窝式接口 22 借以与网络 16 通信的无线蜂窝式接入网络。显示器 24 表示能够显示图像(例如图像数据 26,或任何其它类型的数据)的任何类型的显示单元。显示器 24 可例如表示发光二极管(LED)显示装置、有机 LED(OLED)显示装置、液晶显示器(LCD)装置、等离子体显示装置或任何其它类型的显示装置。

[0029] 视觉搜索服务器 14 包含接口 32、特征重构单元 34、特征匹配单元 36 和特征描述符数据库 38。视觉搜索服务器 14 的接口 32 可类似于客户端装置 12 的接口 22,因为接口 32 可表示能够与网络(例如网络 16)通信的任何类型的接口。特征重构单元 34 表示对经压缩特征描述符进行解压缩以从经压缩特征描述符重构特征描述符的单元。特征重构单元 34 可执行与特征压缩单元 20 所执行的操作相逆的操作,因为特征重构单元 34 执行量化的逆操作(通常称为重构),以从经压缩特征描述符重构特征描述符。

[0030] 特征匹配单元 36 表示执行特征匹配以基于所重构的特征描述符来识别图像数据 26 中的一个或一个以上特征或对象的单元。特征匹配单元 36 可存取特征描述符数据库 38 以执行此特征识别,其中特征描述符数据库 38 存储界定特征描述符且使这些特征描述符中的至少一些与包含从图像数据 26 提取的对应特征或对象的参考图像相关联的数据。这些参考图像还可与识别参考图像的一个或一个以上主题、特征或对象的识别数据相关联。

数据库 38 可使用经压缩 k 维树 (KD 树) 来存储此数据。

[0031] 在基于经重构特征描述符 (例如经重构特征描述符 40 (其在本文中还可称为“查询数据 40”, 因为此数据表示用以执行视觉搜索或查询的识别搜索查询数据)) 成功识别从图像数据 26 提取的特征或对象后, 特征匹配单元 36 即刻返回一个或一个以上匹配参考图像以及任何相关联的识别数据, 以作为查询结果数据 42。

[0032] 最初, 客户端装置 12 的用户与客户端装置 12 交互以起始视觉搜索。所述用户可与由显示器 24 呈现的用户接口或其它类型的接口交互, 以选择查询图像数据 26, 且接着起始视觉搜索以识别一个或一个以上特征或对象, 其为存储为查询图像数据 26 的图像的焦点。举例来说, 查询图像数据 26 可指定地标的图像, 例如比萨斜塔。用户可能已使用客户端装置 12 的图像捕获单元 (例如, 相机) 捕获此图像, 或者从网络 16 下载此图像, 或经由与另一计算装置的有线或无线连接在本地下载此图像。在任一情况下, 在选择查询图像数据 26 之后, 在此实例中, 用户起始视觉搜索以识别地标。

[0033] 响应于起始视觉搜索, 客户端装置 12 调用特征提取单元 18 来提取至少一个特征描述符 28, 且通常提取若干个特征描述符 28, 其描述通过查询图像数据 26 的分析而找到的所谓的“关键点”中的一者。特征提取单元 18 将此查询特征描述符 28 转发到特征压缩单元 20, 其继续压缩查询特征描述符 28 且产生查询数据 30。特征压缩单元 20 将查询数据 30 输出到接口 22, 其经由网络 16 将查询数据 30 转发到视觉搜索服务器 14。

[0034] 视觉搜索服务器 14 的接口 32 接收查询数据 30。响应于接收到查询数据 30, 视觉搜索服务器 14 调用特征重构单元 34。特征重构单元 34 基于查询数据 30 重构查询特征描述符 28, 且输出经重构的特征描述符 40。特征匹配单元 36 接收经重构的查询特征描述符 40, 且基于查询特征描述符 40 执行特征匹配。特征匹配单元 36 通过针对查询特征描述符 40 中的每一者来存取特征描述符数据库 38 且遍历由特征描述符数据库 38 存储为数据的参考特征描述符以识别大体上匹配的特征描述符, 而执行特征匹配。在基于经重构的查询特征描述符 40 成功地识别从图像数据 26 提取的特征后, 特征匹配单元 36 即刻输出与匹配的参考特征描述符相关联的一个或一个以上匹配参考图像以及任何相关联的识别数据, 以作为查询结果数据 42。接口 32 接收此查询结果数据 42, 并经由网络 16 将查询结果数据 42 转发到客户端装置 12。

[0035] 客户端装置 12 的接口 22 接收此查询结果数据 42, 且通常将此查询结果数据 42 转发到调用搜索查询的任何应用程序。就是说, 客户端装置 12 通常执行一个或一个以上应用程序, 其可调用视觉搜索且管理所返回的查询结果数据 (例如查询结果数据 42) 的呈现。此应用程序可与显示器 24 介接, 以呈现查询结果数据 42。在一些例子中, 应用程序可执行因特网搜索或其它操作, 以基于由查询结果数据 42 界定的识别数据来检索额外信息, 以扩充查询结果数据 42。在此例子中, 查询结果数据 42 的识别数据可包括地标的名称 (即在此实例中为比萨斜塔)、建造比萨斜塔的建造者的名字、比萨斜塔的完工日期, 以及与此地标有关的任何其它信息。

[0036] 虽然视觉搜索可促进与具有较小屏幕且 / 或有限输入机构 (例如, 提供可能限制或阻挠用户对文本的输入的极少触觉反馈的虚拟键盘) 的客户端装置的用户交互, 且通常在促进各种应用方面提供若干益处, 但在一些例子中, 一些视觉搜索算法, 例如尺度不变特征变换 (SIFT) 算法在执行特征匹配方面可能是有缺陷的。为了说明这些缺陷, 在查询特征

描述符和参考特征描述符各自从搜索和参考图像的重复特征（例如跨建筑物（例如比萨斜塔）重复的独特的拱门或窗户）提取时的例子中，SIFT 算法可丢弃原本将会与搜索或查询特征描述符匹配的参考特征描述符。此外，SIFT 算法通常响应于任何给定视觉搜索而仅返回单个图像，其中通过 SIFT 算法在算法上将此所返回的图像确定为“最佳匹配”。然而，用户无法以与 SIFT 算法相同的方式确定什么构成“最佳匹配”，这可导致用户受挫，因为单个 SIFT 最佳匹配结果无法匹配用户的预期。

[0037] 更具体地说，虚拟搜索算法（包含 SIFT 算法）使用可被称为“相似性量度”的东西实施特征匹配。SIFT 算法使用距离比率测试来测量查询特征描述符与参考特征描述符中的每一者有多相似。此距离比率测量查询特征描述符中的当前一者的最近参考特征描述符与查询特征描述符的所述当前一者的第二近（再次在特征描述符空间中）特征描述符有多远。如果最近与第二近参考特征描述符非常近（如阈值所识别），则 SIFT 算法确定查询特征描述符中的当前一者不是唯一的，且因此匹配是不可靠的。此距离比率相似性量度可由以下等式（1）表示：

[0038]

$$\text{如果 } \frac{d(q, nn_1)}{d(q, nn_2)} < T, \text{ 接受作为匹配; 否则拒绝匹配。} \quad (1)$$

[0039] 在等式（1）中， $q$  表示查询描述符、 $nn_1$  表示数据库中的最近参考特征描述符（其还可称为“最近邻域”或“nn”），且  $nn_2$  表示数据库中的第二近参考特征描述符（其可称为“第二近邻域”）， $d$  表示 L2 规范运算， $T$  表示阈值。

[0040] 虽然 SIFT 距离比率相似性量度通常提供合适结果，但 SIFT 距离比率相似性量度在某些例子中可能无法正确地识别匹配的参考特征描述符。举例来说，在其中图像含有重复结构（例如，图 7 的实例中所示的比萨斜塔的重复窗结构）的例子中，SIFT 距离比率相似性量度可确定考虑到其两个最近参考特征描述符（其也是从描述比萨斜塔的图像数据提取的）之间的距离可能较小，从此些重复的图像特征提取的查询特征描述符不是唯一的。在此例子中，SIFT 距离比率相似性量度可能否决潜在正确的匹配。

[0041] 根据本发明中所描述的技术，视觉搜索服务器 14 的特征匹配单元 36 使用最稳健的距离比率相似性量度，其可适应重复结构，且另外促进这些和其它例子中的较准确匹配。此外，所述技术可提供参考图像的经排序等级，以适应算术确定中的差异以及用户对“最佳匹配”的感知，而不是像常规系统中常见的那样，简单地返回单个“最佳匹配”参考图像。

[0042] 在实施根据本发明中所描述的技术的距离比率相似性量度中，特征匹配单元 36 可首先计算经重构的查询特征描述符 40 中的一者与存储到特征描述符数据库 38 的多个参考特征描述符中的每一者之间的距离。特征匹配单元 36 可接着使用群集算法来确定所计算距离中的一者或一者以上的第一群组，以及所计算距离的第二群组。常见的群集算法可包括 k 均值群集算法（其中，在此实例中，将 k 设定为二，以产生第一和第二群组）、高斯拟合算法和图形切割算法。

[0043] 在一些例子中，群集算法可包括将最近或最小距离群集到第一群组中，且将多个次近或次小距离群集到第二群组中的算法，其中所述第二群组的所述多个次近或次小距离的数目可（例如）包含次近距离中的两者、三者或四者。虽然本文中相对于较正式的群集算法（例如 k 均值群集算法）来描述，但可相对于本发明中所描述的技术而利用能够确定

两个或两个以上群组的任何算法,其中第一群组包含一个或一个以上距离,且第二群组包含两个或两个以上距离。

[0044] 使用群集算法,特征匹配单元 36 可确定所计算距离中的一者或一者以上的第一群组,使得此第一群组包含所计算距离中指示存储到特征描述符数据库 38 的多个参考特征描述符中的相关联一者相对于所计算距离中被确定为在所计算距离的第二群组中的那些所计算距离接近查询特征描述符的那些所计算距离。还通过使用此群集算法,特征匹配单元 36 可确定所计算距离的第二群组,使得此第二群组包含所计算距离中指示存储到特征描述符数据库 38 的多个参考特征描述符中的相关联一者相对于所计算距离中被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符 40 中的当前一者的那些所计算距离。

[0045] 视觉搜索服务器 14 的特征匹配单元 36 可接着基于所确定的所计算距离的第一群组和所计算距离的第二群组来确定查询特征描述符 40 中的当前一者是否与多个参考特征描述符中与所计算距离中的最小者相关联的一个参考特征描述符匹配。举例来说,特征匹配单元 36 可首先计算被确定为在第一群组中的所计算距离的平均值,以产生第一群组距离平均值,且同样地,计算被确定为在第二群组中的所计算距离的平均值,以产生第二群组距离平均值。在计算第一和第二群组距离平均值之后,特征匹配单元 36 可即刻将第一群组距离平均值除以第二群组距离,产生平均距离比率量度。特征匹配单元 36 可接下来将平均距离比率量度与阈值进行比较,且基于所述比较确定查询特征描述符 40 中的当前一者是否与多个参考特征描述符中与所计算距离中的最小者相关联的一个参考特征描述符匹配。

[0046] 通过以此方式群集或以其它方式分组所计算距离,且接着求出每一群组的距离的平均值,所述技术可适应从邻近重复特征提取的特征描述符的匹配,而不茫然地否决彼此靠近且两者均与查询特征描述符匹配的任何参考特征描述符,如常规上常见的那样。使用群集算法对距离量度进行分组还可实现较近特征描述符与较远特征描述符之间的较清楚区分,使得视觉搜索服务器 14 的特征匹配单元 36 可使用距离来正确地识别彼此不同或唯一的特征描述符。通过求出距离量度的平均值,特征匹配单元 36 可提供这两个群组彼此相距多远的相对接近近似,以便正确地应用阈值,且进而避免否决将被视为匹配的参考特征描述符的内容。

[0047] 为了实施所述技术的匹配图像方面的经排序或分级列表,视觉搜索服务器 14 的特征匹配单元 36 利用所计算距离的第一群组,其与所计算距离的第二群组相比相对接近查询特征描述符 40 中的当前一者。对于被确定为在第一群组中且使用上文所述的较稳健距离相似性量度被确定为匹配的所有所计算距离,特征匹配单元 36 可将选票指派给与对应于第一群组的所计算距离的匹配参考特征描述符相关联的独特图像。就是说,如果对应于第一群组的所计算距离的参考特征描述符中的多者是从同一参考图像提取,那么特征匹配单元 36 将单个选票指派给所述参考图像。

[0048] 举例来说,特征匹配单元 36 可确定唯一参考图像的群组,使得参考图像群组不包含任何重复的参考图像。特征匹配单元 36 可首先考虑从其计算出被确定为在第一群组中的所计算距离的参考特征描述符,其中这些参考特征描述符可称为“第一群组参考特征描述符”。每一参考特征描述符与在最初建立特征描述符数据库 38 时从其提取所述参考特征描述符的参考图像(其还可存储到特征描述符数据库 38 或图 1 的实例中未展示的某一其

它数据库、存储器或存储单元)相关联。特征匹配单元 36 可接着将参考图像的初始群组确定为与第一群组参考特征描述符相关联的那些参考图像。特征匹配单元 36 可接着将任何重复的参考图像从参考图像的初始群组去除,以产生唯一参考图像的群组,使得参考图像的所述群组不包含任何重复的参考图像。

[0049] 特征匹配单元 36 可接着将选票指派给参考图像中被确定为在唯一参考图像的群组中的每一参考图像。所述选票可为常数,例如一。或者,所述选票可与当前所计算距离与查询特征描述符 40 中的当前一者与最近参考特征描述符之间的所计算距离相比的距离比率成比例。作为另一替代,当第一群组的所计算距离从最小到最大排序时,选票可与第一群组内的等级成比例(例如,如果第一群组参考特征描述符中的一者是第五接近查询特征描述符 40 中的当前一者,那么选票将为 1 / 5)。特征匹配单元 36 接着以此方式为查询特征描述符 40 中的每一者指派选票。在已以此方式将选票指派给参考或目标图像之后,特征匹配单元 36 可接着基于所收集的选票对参考或目标图像进行分级,且响应于呈查询结果数据 42 的形式的查询数据 30 而提供目标图像的等级或经排序列表。

[0050] 虽然本发明中描述各种组件、模块或单元,以强调经配置以执行所揭示技术的装置的功能方面,但这些单元不一定要求由不同硬件单元来实现。相反,如上文所述,各种单元可组合在硬件单元中,或通过互操作硬件单元的集合结合存储到计算机可读媒体的合适软件和 / 或固件来提供,所述互操作硬件单元包含如上文所述的一个或一个以上处理器。在此方面中,本发明中对单元的参考意在建议可或可不实施为单独硬件单元和 / 或硬件与软件单元的不同功能单元。注意,如本文所使用的短语“计算机可读媒体”仅适用于制造商,且不适用于暂时传播信号。

[0051] 图 2 是更详细地说明图 1 的实例中所展示的特征匹配单元 36 的框图。在图 2 的实例中,特征匹配单元 36 包含距离计算单元 50、分组单元 52、匹配单元 54 和结果产生器单元 56。距离计算单元 50 表示基于存储到特征描述符数据库 38 的参考特征描述符 51 和查询特征描述符 40 来计算距离 62A 到 62N(“距离 62”)的单元。更具体地说,距离计算单元 50 可表示将查询特征描述符 40 中的每一者的距离 62 计算为从查询特征描述符 40 中的当前一者减去参考特征描述符 51 中的每一者的绝对值的单元。距离计算单元 50 将距离 62 输出到分组单元 52。

[0052] 分组单元 52 表示接收距离(例如距离 62)或两个或两个以上值的任何其它非空集合且根据群集算法 64 将这些值分组为至少两个群组的单元。群集算法 64 可表示 k 均值群集算法(其中在此实例中,将 k 设定为二,以产生第一和第二群组)、高斯拟合算法和图形切割算法,以及任何其它能够从一组值产生至少两个群组的常见群集算法中的一者或一者以上。在图 2 的实例中,分组单元 52 接收距离 62,且根据群集算法 64 计算距离 62 中的不同距离的群组 66A、66B(“群组 66”)。分组单元 52 可确定所计算距离 62 中的两者或两者以上的群组 66A,使得此第一群组 66A 包含所计算距离 62 中指示存储到特征描述符数据库 38 的参考特征描述符 51 中的相关联一者相对于所计算距离 62 中被确定为在所计算距离 62 的群组 66B 中的那些所计算距离接近查询特征描述符 40 中的当前一者的那些所计算距离。同样根据群集算法 64,分组单元 52 可确定所计算距离 62 的群组 66B,使得此第二群组 66B 包含所计算距离 62 中指示参考特征描述符 51 中的相关一者相对于所计算距离中被确定为在所计算距离的第一群组中的那些所计算距离远离查询特征描述符 40 中的当前一

者的那些所计算距离。分组单元 52 将距离 62 的群组 66 输出到匹配单元 54。

[0053] 匹配单元 54 表示确定查询特征描述符 40 中的当前一者是否唯一地匹配参考特征描述符 51 中的一者或一者以上的单元。为了确定查询特征描述符 40 中的当前一者是否唯一地匹配参考特征描述符 51 中的一者或一者以上, 匹配单元 54 通常计算所提供群组中的每一者的值的平均值。在图 2 的实例中, 匹配单元 54 通过分别求出群组 66A 的距离 62 和群组 66B 的距离 62 的平均值来计算群组平均值 68A、68B。接下来, 匹配单元 54 将群组平均值 68A 除以群组平均值 68B, 且将结果与阈值 70 进行比较。就是说, 匹配单元 54 通常仅相对于从使用群集算法 64 形成的距离 62 的群组 66 计算的群组平均值, 而执行与 SIFT 算法当前执行 (且上文展示为等式 (1)) 的比较类似的比较。

[0054] 如果将群组平均值 68A 除以群组平均值 68B 的结果大于阈值 (意味着与群组 66B 相关联的参考特征描述符 51 接近与群组 66A 相关联的那些参考特征描述符), 那么匹配单元 54 输出指示特征描述符 40 中的当前一者不唯一地匹配参考特征描述符 51 中的任一者的匹配指示符 72。本文使用术语“唯一匹配”来指代执行视觉搜索的上下文中的特征匹配的方面, 其中匹配可能存在, 但可能不促进查询图像数据的识别, 且因此不将所述匹配视为唯一。如果匹配促进查询图像数据的识别, 那么所述匹配是唯一的。举例来说, 可从比萨斜塔的查询图像提取比萨斜塔上的拱门, 以作为可唯一地识别此地标的查询特征描述符。从比萨斜塔的参考图像中的这些拱门提取的参考特征描述符可能匹配。不是像考虑这些拱门的重复性质以及唯一性的狭窄定义的常规 SIFT 算法中常见的那样否决这些匹配的参考图像特征描述符, 而是匹配单元 54 可适当地确定这些唯一地匹配, 而不管归因于匹配单元 54 对其操作的数据的群集性质而引起的重复。此唯一匹配还可称为上文所述的相似性量度。唯一匹配是优选的, 因为其比非唯一匹配更快速地促进查询图像数据的识别, 这是匹配单元 54 否决非唯一匹配的原因。

[0055] 在任何情况下, 如果将群组平均值 68A 除以群组平均值 68B 的结果小于阈值 (意味着与群组 66B 相关联的参考特征描述符 51 远离与群组 66A 相关联的那些参考特征描述符), 那么匹配单元 54 输出指示特征描述符 40 中的当前一者唯一地匹配参考特征描述符 51 中与距离 62 中的最小者相关联的一个参考特征描述符的匹配指示符 72。匹配单元 54 输出此匹配指示符 72, 且将群组 66A 转发到结果产生器单元 56。在一些例子中, 如果匹配指示符 72 指示匹配, 那么匹配单元 54 仅转发群组 66A, 因为当匹配单元 54 已确定不存在匹配时, 结果产生器单元 56 无需考虑群组 66A。

[0056] 结果产生器单元 56 表示产生查询结果数据 42 的单元。结果产生器单元 56 包含列表形成单元 58 和选票指派单元 60。列表形成单元 58 表示基于群组 66A 形成唯一参考图像 74 的列表的单元。更具体地说, 列表形成单元 58 可在匹配指示符 72 发信号通知匹配时处理群组 66A 以确定参考特征描述符 51 中从其计算群组 66A 的距离 62 的那些参考特征描述符。列表形成单元 58 接着检索从其提取参考特征描述符 51 中的这些参考特征描述符中的每一者的参考图像 53, 且形成初始参考图像列表 76。列表形成单元 58 可接着将任何冗余图像从初始参考图像列表 76 去除, 以产生唯一参考图像列表 74。

[0057] 尽管描述为形成初始参考图像列表 76, 但列表形成单元 58 可不一定形成此列表 76, 而是可替代地仅检索参考图像 53 中的唯一参考图像。就是说, 列表形成单元 58 可在检索从其提取参考特征描述符 51 中的所确定参考特征描述符的参考图像 53 时, 首先确定其

是否已经检索图像 53 中的此图像,且接着如果其尚未检索到图像 53 中的此图像,那么仅检索图像 53 中的此图像。通过避免存储图像 53 的拷贝或副本,列表形成单元 58 可避免消耗存储器,这可使实施者能够降低存储器要求。出于此原因,所述技术不应限于借以计算唯一参考图像列表 74 的任何一种方式。由此,不管列表形成单元 58 如何计算唯一参考图像列表 74,列表形成单元 58 均将唯一参考图像列表 74 输出到选票指派单元 60。

[0058] 选票指派单元 60 表示将选票指派给参考图像 53 中的一者或一者以上的单元。通常,选票指派单元 60 通过使这些图像 53 中的每一者与选票清点表 78 中的条目相关联来将选票指派给由唯一参考图像列表 74 识别的图像 53。选票指派单元 60 可更新与图像 53 相关联的每一条目,以递增每一条目的选票计数。选票指派单元 60 可以任何数目的方式(包含上文所述的那些方式)来确定选票。就是说,所述选票可为常数,例如一。或者,所述选票可与当前所计算距离与查询特征描述符 40 中的当前一者与最近参考特征描述符之间的所计算距离相比的距离比率成比例。作为另一替代,当第一群组的所计算距离从最小到最大排序时,选票可与第一群组内的等级成比例(例如,如果第一群组参考特征描述符中的一者是第五接近查询特征描述符 40 中的当前一者,那么选票将为 1 / 5)。

[0059] 距离计算单元 50、分组单元 52、匹配单元 54 和结果产生器单元 56 可以上文所述的方式继续,直到所有查询特征描述符 40 均已经历此特征匹配过程为止。在处理查询特征描述符 40 中的最后一者之后,结果产生器单元 56 基于选票清点表 78 构造查询结果数据 42。举例来说,结果产生器单元 56 可检索参考图像 53 中接收到最多选票、第二多选票、第三多选票等等直到某一限制或阈值的那些参考图像,且输出这些图像 53 以作为查询结果数据 42(以及描述或以其它方式与特征描述符数据库 38 中的这些图像中的每一者相关联的元数据或其它数据)。或者,结果产生器单元 56 可形成查询结果数据 42,以作为图像 53 中接收到选票的任一图像(再次,连同描述或以其它方式与特征描述符数据库 38 中的这些图像中的每一者相关联的元数据或其它数据)。

[0060] 图 3 是说明视觉搜索服务器(例如图 1 的实例中所示的视觉搜索服务器 14)在实施本发明中所描述的特征匹配技术时的示范性操作的流程图。最初,视觉搜索服务器 14 经由接口 32 接收通常界定经压缩特征描述符的查询数据 30(90)。视觉搜索服务器 14 的特征重构单元 34 从由查询数据 30 所界定的经压缩特征描述符重构查询特征描述符 28,以产生经重构特征描述符 40(92)。特征重构单元 34 将经重构的查询特征描述符 40 输出到特征匹配单元 36。

[0061] 特征匹配单元 36 用距离计算单元 50 接收这些查询特征描述符 40(如图 2 的实例中所示)。距离计算单元 50 选择查询特征描述符 40 中的一者(94)。距离计算单元 50 接着以上文所述的方式计算查询特征描述符 40 中的选定一者与存储到特征描述符数据库 38 的参考特征描述符 51 中的一者或一者以上之间的距离(96)。在图 2 的实例中将这些距离展示为距离 62。距离计算单元 50 将这些距离 62 输出到分组单元 52。

[0062] 在接收到这些距离 62 后,分组单元 52 即刻根据群集算法 64 确定所计算距离 62 的第一和第二群组 66A、66B(98)。第一群组 66A 包含与第二群组 66B 的一组非零距离 62 不同的一组非零距离 62。就是说,将距离 62 被划分为第一和第二群组 66,使得距离 62 中无一者与群组 66 的两者相关联。分组单元 52 将群组 66 输出到匹配单元 54。

[0063] 响应于接收到群组 66,匹配单元 54 计算第一群组距离 66A 的平均值以及第二群

组距离 66B 的平均值 (100)。在图 1 的实例中将这些平均值展示为群组平均值 68A 和群组平均值 68B, 其中群组平均值 68A 表示群组 66A 的距离 62 的所计算平均值, 且群组平均值 68B 表示群组 66B 的距离 62 的所计算平均值。匹配单元 54 接着将群组平均值 68A 除以群组平均值 68B, 将除法的结果与阈值 70 进行比较, 以确定除法的结果是否小于阈值 70 (102 到 106)。如果结果小于阈值 70 (“是” 106), 那么匹配单元 54 产生匹配指示符 72, 其识别查询特征描述符 40 中的选定一者与参考特征描述符 51 的所确定最近一者匹配 (在距离 62 方面) (108)。匹配单元 54 将匹配指示符 72 和第一群组 66A 转发到结果产生器单元 56。

[0064] 结果产生器单元 56 接收匹配指示符 72 和群组 66A。响应于接收到匹配指示符 72 和群组 66A, 结果产生器单元 56 调用列表形成单元 58。列表形成单元 58 通过首先以上文所述的方式识别与距离 62 的第一群组 66A 相关联的参考特征描述符 51 来产生初始参考图像列表 76 (110)。列表形成单元 58 接下来再次以上文所述的方式识别与参考特征描述符 51 中的所识别参考特征描述符相关联的参考图像 53 (112)。列表形成单元 58 可将所识别的参考图像 53 存储为初始参考图像列表 76。也如上文所述, 列表形成单元 58 接着基于所识别的参考特征描述符 51 (或初始参考图像列表 76) 产生唯一参考图像 53 的列表, 其在图 2 的实例中将中此列表展示为唯一参考图像列表 74 (114)。结果产生器单元 56 可接下来调用选票指派单元 60, 其以上文所述的方式将选票指派给唯一参考图像 53 的列表 (或唯一参考图像列表 74) (116)。选票指派单元 60 可维持选票清点表 78, 以将选票指派给由唯一参考图像列表 74 识别的那些参考图像 53。

[0065] 特征匹配单元 36 接着以上文所述的方式确定其是否已完成了处理所有的查询特征描述符 40 (118)。或者, 如果匹配单元 54 确定将群组平均值 68A 除以群组平均值 68B 的结果不小于阈值 70 (返回参看图 3A, “否” 106), 那么特征匹配单元 36 可同样地以上文所述的方式确定其是否已完成了处理所有的查询特征描述符 40 (118)。不管在何上下文中其确定其是否已完成了处理所有的查询特征描述符 40, 如果未完成处理所有查询特征描述符 40 (“否” 118), 那么单元 50 到 60 以所描述的方式操作, 以处理剩余的查询特征描述符 40 (94 到 118)。

[0066] 然而, 如果完成了处理所有的查询特征描述符 40 (“是” 118), 那么选票指派单元 60 可以上文所述的方式基于所指派的选票 (如由选票清点表 78 界定) 产生参考图像 53 的经分级列表 (120)。结果产生器单元 56 可接着产生查询结果数据 42 以包含参考图像 53 的经分级列表 (122)。结果产生器单元 56 可接着经由接口 32 将查询结果数据 42 发射到客户端装置 12 (124)。

[0067] 图 4 是说明已确定为在特征描述符提取中使用的高斯差 (DoG) 金字塔 204 的图。图 1 的特征提取单元 18 可通过计算高斯金字塔 202 中的任何两个连续经高斯模糊图像的差来构造 DoG 金字塔 204。图 1 的实例中展示为图像数据 26 的输入图像  $I(x, y)$  经逐渐高斯模糊, 以构造高斯金字塔 202。高斯模糊通常涉及使原始图像  $I(x, y)$  以尺度  $c\sigma$  与高斯模糊函数  $G(x, y, c\sigma)$  卷积, 使得高斯模糊函数  $L(x, y, c\sigma)$  被界定为  $L(x, y, c\sigma) = G(x, y, c\sigma) * I(x, y)$ 。此处,  $G$  为高斯核,  $c\sigma$  表示用于模糊图像  $I(x, y)$  的高斯函数的标准偏差。随着  $c$  变化 ( $c_0 < c_1 < c_2 < c_3 < c_4$ ), 标准偏差  $c\sigma$  变化, 且获得逐渐模糊。西格玛  $\sigma$  是基础尺度变量 (本质上是高斯核的宽度)。当初始图像  $I(x, y)$  与高斯  $G$  递增地卷积以产生经模糊图像  $L$  时, 经模糊图像  $L$  在尺度空间中通过常数因子  $c$  分离。

[0068] 在 DoG 空间或金字塔 204 中,  $D(x, y, a) = L(x, y, c_n \sigma) - L(x, y, c_{n-1} \sigma)$ 。DoG 图像  $D(x, y, \sigma)$  是尺度  $c_n \sigma$  和  $c_{n-1} \sigma$  下两个邻近经高斯模糊图像  $L$  之间的差。 $D(x, y, \sigma)$  的尺度位于  $c_n \sigma$  与  $c_{n-1} \sigma$  之间的某处。随着经高斯模糊的图像  $L$  的数目增加, 且针对高斯金字塔 202 而提供的近似接近连续空间, 所述两个尺度也接近为一个尺度。经卷积的图像  $L$  可通过八元组分组, 其中八元组对应于标准偏差  $\sigma$  的值的两倍。此外, 选择乘数  $k$  (例如  $c_0 < c_1 < c_2 < c_3 < c_4$ ) 的值, 使得每八元组获得固定数目的经卷积图像  $L$ 。接着, 可每八元组从邻近经高斯模糊的图像  $L$  获得 DoG 图像  $D$ 。在每一八元组之后, 高斯图像通过因子 2 进行下取样, 且接着重复所述过程。

[0069] 特征提取单元 18 可接着使用 DoG 金字塔 204 来识别图像  $I(x, y)$  的关键点。在执行关键点检测时, 特征提取单元 18 确定图像中的特定样本点或像素附近的局域区或小块是否为潜在感兴趣的小块 (在几何上讲)。通常, 特征提取单元 18 识别 DoG 空间 204 中的局部极大值和 / 或局部极小值, 且使用这些极大值和极小值的位置作为 DoG 空间 204 中的关键点位置。在图 4 中所说明的实例中, 特征提取单元 18 识别小块 206 内的关键点 208。可通过以相同尺度将 DoG 空间 204 中的每一像素 (例如, 用于关键点 208 的像素) 与其八个相邻像素及两侧上相邻尺度中的每一者中的九个相邻像素 (邻近小块 210 和 212 中) (共计 26 个像素 ( $9 \times 2 + 8 = 26$ )) 进行比较来实现找到局部极大值和极小值 (也称为局部极值检测)。如果关键点 208 的像素值是小块 206、210 和 208 中的所有 26 个所比较像素之中的最大值或最小值, 那么特征提取单元 18 将此选为关键点。特征提取单元 18 可进一步处理关键点, 使得更准确地识别其位置。在一些例子中, 特征提取单元 18 可丢弃所述关键点中的一些关键点, 例如低对比度关键点和边缘关键点。

[0070] 图 5 是更详细地说明关键点的检测的图。在图 5 的实例中, 小块 206、210 和 212 中的每一者包含  $3 \times 3$  像素区。特征提取单元 18 首先将以同一尺度 (例如, 小块 206) 对所关注像素 (例如, 关键点 208) 与其八个相邻像素 302 及关键点 208 的两侧上相邻尺度中的每一者中的邻近小块 210 和 212 中的九个相邻像素 304 和 306 进行比较。

[0071] 特征提取单元 18 可基于局部图像梯度的方向来为每一关键点指派一个或一个以上定向或方向。通过基于局部图像特性来将一致定向指派给每一关键点, 特征提取单元 18 可相对于此定向表示关键点描述符, 且因此实现对图像旋转的不变性。特征提取单元 18 接着计算经高斯模糊的图像  $L$  中的关键点 208 周围的相邻区且 / 或关键点尺度下的每一像素的量值和方向。位于  $(x, y)$  处的关键点 208 的梯度的量值可表示为  $m(x, y)$ , 且  $(x, y)$  处的关键点的梯度的定向或方向可表示为  $\Gamma(x, y)$ 。

[0072] 特征提取单元 18 接着使用所述关键点的尺度来选择经高斯平滑的图像  $L$ , 其具有与关键点 208 的尺度最接近的尺度, 使得所有计算均以尺度不变方式来执行。对于此尺度下的每一图像样本  $L(x, y)$ , 特征提取单元 18 使用像素差来计算梯量度值  $m(x, y)$  和定向  $\Gamma(x, y)$ 。举例来说, 可根据以下等式 (2) 来计算量值  $m(x, y)$  :

$$[0073] m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2)$$

[0074] 特征提取单元 18 可根据以下等式 (3) 来计算方向或定向  $\Gamma(x, y)$  :

[0075] 错误!无法从编辑域代码来创建对象。 (3)

[0076] 在等式 (3) 中,  $L(x, y)$  表示尺度  $\sigma$  下经高斯模糊的图像  $L(x, y, \sigma)$  的样本, 尺度  $\sigma$  也是关键点的尺度。

[0077] 特征提取单元 18 可以比 DoG 空间中的关键点的平面高的尺度为高斯金字塔中位于上方的平面或以比所述关键点低的尺度为高斯金字塔的位于下方的平面一致地计算关键点的梯度。不管怎样,对于每一关键点,特征提取单元 18 计算围绕关键点的矩形区域(例如小块)中处于同一尺度的梯度。此外,在经高斯模糊图像的尺度中反映图像信号的频率。但是,SIFT 和其它算法(例如经压缩梯度直方图(ChG)算法)仅使用小块(例如,矩形区域)中的所有像素处的梯度值。在关键点周围界定小块;在块内界定子块;在子块内界定样本,且此结构对于所有关键点保持相同,甚至在关键点的尺度不同时也是如此。因此,虽然图像信号的频率随着高斯平滑滤波器在同一八元组中的连续应用而改变,但在不同尺度下识别的关键点可用相同数目的样本来取样,而不管图像信号的频率的改变,其由所述尺度表示。

[0078] 为了表征关键点定向,特征提取单元 18 可通过使用例如 SIFT 来产生梯度定向直方图(见图 6)。每一相邻像素的贡献可由梯量度值和高斯窗来加权。直方图中的峰值对应于支配定向。特征提取单元 18 可相对于关键点定向测量关键点的所有特性,这提供对旋转的不变性。

[0079] 在一个实例中,特征提取单元 18 为每一块计算经高斯加权梯度的分布,其中每一块为 2 个子块乘以 2 个子块,总共 4 个子块。为了计算经高斯加权梯度的分布,特征提取单元 18 形成具有若干柱的定向直方图,其中每一柱覆盖关键点周围的区域的一部分。举例来说,定向直方图可具有 36 个柱,每一柱覆盖 360 度定向范围的 10 度。或者,直方图可具有 8 个柱,各自覆盖 360 度范围的 45 度。应清楚,本文所述的直方图译码技术可适用于任何数目个柱的直方图。

[0080] 图 6 是说明特征提取单元(例如特征提取单元 18)借以确定梯度分布和定向直方图的过程的图。此处,将二维梯度分布( $dx, dy$ )(例如,块 406)转换为一维分布(例如,直方图 414)。关键点 208 位于环绕关键点 208 的块 406(也称为小块、单元或区)的中心。针对金字塔的每一层级而预计的梯度展示为每一样本位置 408 处的小箭头。如图所示,样本 408 的区形成子块 410,其还可称为柱 410。特征提取单元 18 可使用高斯加权函数来将权重指派给子块或柱 410 内的每一样本 408。由高斯加权函数指派给样本 408 中的每一者的权重从柱 410 的质心和关键点 208 平滑地下降。高斯加权函数的目的是以窗位置的小改变来避免描述符的突然改变,且将较少重点给予远离描述符的中心的梯度。特征提取单元 18 用直方图的每一柱中的 8 个定向来确定定向直方图 412 的阵列,从而得出维度特征描述符。举例来说,定向直方图 413 可对应于子块 410 的梯度分布。

[0081] 在一些例子中,特征提取单元 18 可使用其它类型的量化柱星座(例如,具有不同的沃洛诺依(Voronoi)单元结构)来获得梯度分布。这些其它类型的柱星座可同样地使用一种形式的软分柱,其中软分柱指代重叠柱,例如当使用所谓的 DAISY 配置时所界定的那些柱。在图 6 的实例中,界定三个软柱,然而,可使用多达 9 个或 9 个以上软柱,其中质心通常位于关键点 208 周围的圆形配置 402 中。

[0082] 如本文中所使用,直方图为映射  $k_i$ ,其对属于被称为柱的各种不相交类别的观察、样本或发生(例如梯度)的数目进行计数。直方图的图表只是表示直方图的一种方式。因此,如果  $k$  为观察、样本或发生的总数,且  $m$  为柱的总数,那么直方图  $k_i$  中的频率满足表达为等式(4)的以下条件:

$$[0083] n = \sum_{i=1}^m k_i, \quad (4)$$

[0084] 其中  $\Sigma$  为求和运算符。

[0085] 特征提取单元 18 可通过具有是关键点的尺度的 1.5 倍的标准偏差的由经高斯加权函数界定的其梯量度值来对添加到直方图 412 的每一样本进行加权。所得直方图 414 中的峰值对应于局部梯度的支配方向。特征提取单元 18 接着检测直方图中的最高峰值，且接着检测在最高峰值的某一百分比（例如 80%）内的任一其它局部峰值（其还可用以还创建具有所述定向的关键点）。因此，对于具有类似量值的多个峰值的位置，特征提取单元 18 提取在同一位置和尺度但不同定向处创建的多个关键点。

[0086] 特征提取单元 18 接着使用称为类型量化（其将直方图表达为类型）的一种形式的量化来量化直方图。以此方式，特征提取单元 18 可提取每一关键点的描述符，其中此描述符可由呈类型形式的经高斯加权梯度的分布的位置 ( $x, y$ )、定向和描述符来表征。以此方式，图像可由一个或一个以上关键点描述符（也称为图像描述符）来表征。

[0087] 虽然图 4 到 6 中所示的前述实例提供由 SIFT 算法执行的特征提取的一个实例，但可相对于根据任何其它视觉搜索算法（例如经压缩梯度直方图 (CHoG) 算法）提取的特征描述符来实施所述技术。所述技术因此不应在此方面仅限于使用 SIFT 算法提取的特征描述符，而是可相对于根据任何视觉搜索算法提取的特征描述符来实施。

[0088] 图 7 是说明特征匹配单元（例如图 1 的特征匹配单元 36）在根据本发明中所描述的技术执行特征匹配时的实例操作的图。在图 7 的实例中，展示查询图像数据 500 和参考图像数据 502，其中将查询图像数据 500 展示为在参考图像数据 502 左边。黑线展示从查询图像数据 500 提取的查询特征描述符（其在可用时由左箭头表示）与从参考图像数据 502 提取且存储到特征描述符数据库 38 的参考特征描述符（其由右箭头表示）之间的关系。

[0089] 查询图像数据 500 展示由客户端装置（例如客户端装置 12）捕获或以其它方式获得的比萨斜塔的图像。参考图像数据 502 展示存储到特征描述符数据库 38 的比萨斜塔的另一图像。查询图像数据 500 和参考图像数据 502 展示作为不同尺度且从不同（尽管类似）角度捕获的比萨斜塔的图像。视觉搜索服务器 14 的特征匹配单元 36 在此实例中接收表示由左箭头表示的特征的查询特征描述符 40。特征匹配单元 36 确定匹配的参考特征描述符 51（参看图 2），其表示由右箭头表示的对应特征。

[0090] 在此实例中，比萨斜塔包含呈环绕比萨斜塔的重复拱门的形式的许多重复结构。具体表示为 504A 到 504C 的三条黑线说明在特征描述符空间中可彼此靠近的对应查询和参考特征描述符之间的潜在匹配。不是像常规视觉搜索算法中将常见的那样因为不是唯一的而否决这些匹配，而是特征匹配单元 36 可针对由这些线 504A 到 504C 的左手箭头表示的对应查询特征描述符 40 中的每一者来确定相对于彼此分组在第一群组 66A 中的距离 62（再次，参看图 2）。

[0091] 举例来说，考虑由线 504A 表示的匹配，特征匹配单元 36 可确定由线 504A 的左箭头表示的查询特征描述符 40 与由线 504A 到 504C 以及 504D 到 504G 中的每一者的右箭头表示的参考特征描述符 51 匹配。特征匹配单元 36 可以上文所述的方式，相对于与线 504A 到 504G 相关联的参考特征描述符 51 来确定相对于与线 504A 相关联的查询特征描述符 40 的距离 62。接下来，特征匹配单元 36 可对相对于与线 504A 相关联的查询特征描述符以及

与线 504A 到 504C 相关联的参考特征描述符，以及包含所有其余距离的另一群组而计算的距离 62 进行分组。特征匹配单元 36 可接着求出第一和第二群组距离 66 的平均值，将第一群组平均值 68A 除以第二群组平均值 68B，并将除法的结果与阈值 70 进行比较。特征匹配单元 36 在此实例中确定唯一匹配，且进而确定展示为线 504A 的匹配。

[0092] 视觉搜索算法（例如 SIFT）的常规特征匹配方面可能已通过不使用群集算法来否决由线 504A 表示的匹配。可能因为从由线 504A 的左箭头表示的特征提取的查询特征描述符 40 与从由线 504A 的右箭头表示的特征提取的参考特征描述符 51 之间的距离会是最小而发生否决此匹配。在此实例中，次最小距离会是从由线 504A 的左箭头表示的特征提取的查询特征描述符 40 与从由线 504B 的右箭头表示的特征提取的参考特征描述符 51 之间的距离。此次最小距离可几乎与上文所述的最小距离一样小。因此，当根据视觉搜索算法的常规特征匹配方面将最小距离除以次最小距离时，可产生接近一的数字，其可大于阈值，阈值通常被设定为一的某一分数，例如 0.5。

[0093] 因此，通过根据本发明中所描述的技术使用群集算法来确定群组，特征匹配单元 36 可将参考重复特征图案的所有这些相对小的距离 62 分组到第一群组 66A 中，而所有其它距离 62 可被分组在第二群组（群组 66B）中。特征匹配单元 36 接着求出这些距离的平均值，以便促进这些群组中的每一者的比较，且比较此平均值以避免否决可能将被视为唯一匹配的匹配。因此，所述技术可改进特征匹配，尤其是相对于具有重复特征或方面的图像。

[0094] 此外，为了避免否决潜在匹配，同时保持对接受错误匹配的控制，本发明的稳健距离比率技术提供上文所述的更具区别的相似性量度。可通过使用所存储的中间数据，同时使用 k 维树（KD 树）最佳柱优先来查询第一最近邻域，而用很少到无开销计算成本来执行稳健距离比率技术。

[0095] 在一个或一个以上实例中，所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件实施，那么可将功能作为计算机可读媒体上的一个或一个以上指令或代码而加以存储或传输。计算机可读媒体可包含计算机数据存储媒体。数据存储媒体可为可由一个或一个以上计算机或一个或一个以上处理器存取以检索指令、代码和 / 或数据结构以供实施本发明中所描述的技术的任何可用媒体。如本文所使用的“数据存储媒体”指代制造商，且不指代暂时传播信号。作为实例（而非限制），所述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，快闪存储器，或可用于存储呈指令或数据结构的形式的所要程序代码且可由计算机存取的任何其它媒体。如本文中所使用，磁盘和光盘包括压缩光盘（CD）、激光光盘、光学光盘、数字多功能光盘（DVD）、软磁盘及蓝光光盘，其中磁盘通常以磁性方式再现数据，而光盘使用激光以光学方式再现数据。上文的组合也应包含在计算机可读媒体的范围内。

[0096] 代码可由一个或一个以上处理器执行，例如一个或一个以上数字信号处理器（DSP）、通用微处理器、专用集成电路（ASIC）、现场可编程逻辑阵列（FPGA），或其它等效集成或离散逻辑电路。因此，如本文中所使用的术语“处理器”可指上述结构中的任一者或适合于实施本文中所描述的技术的任一其它结构。另外，在一些方面中，本文所述的功能性可提供于经配置以用于编码和解码的专用硬件和 / 或软件模块内，或并入在组合式编解码器中。并且，可将所述技术完全实施于一个或一个以上电路或逻辑元件中。

[0097] 本发明的技术可在各种各样的装置或设备中实施，包含无线手持机、集成电路

(IC) 或一组 IC(例如, 芯片组)。本发明中描述各种组件、模块或单元, 以强调经配置以执行所揭示技术的装置的功能方面, 但不一定要求由不同硬件单元来实现。相反, 如上文所述, 各种单元可组合在编解码器硬件单元中, 或通过互操作硬件单元的集合结合存储在计算机可读媒体上的合适软件和 / 或固件来提供, 所述互操作硬件单元包含如上文所述的一个或一个以上处理器。

[0098] 已描述了各种实例。这些和其它实例在所附权利要求书的范围内。

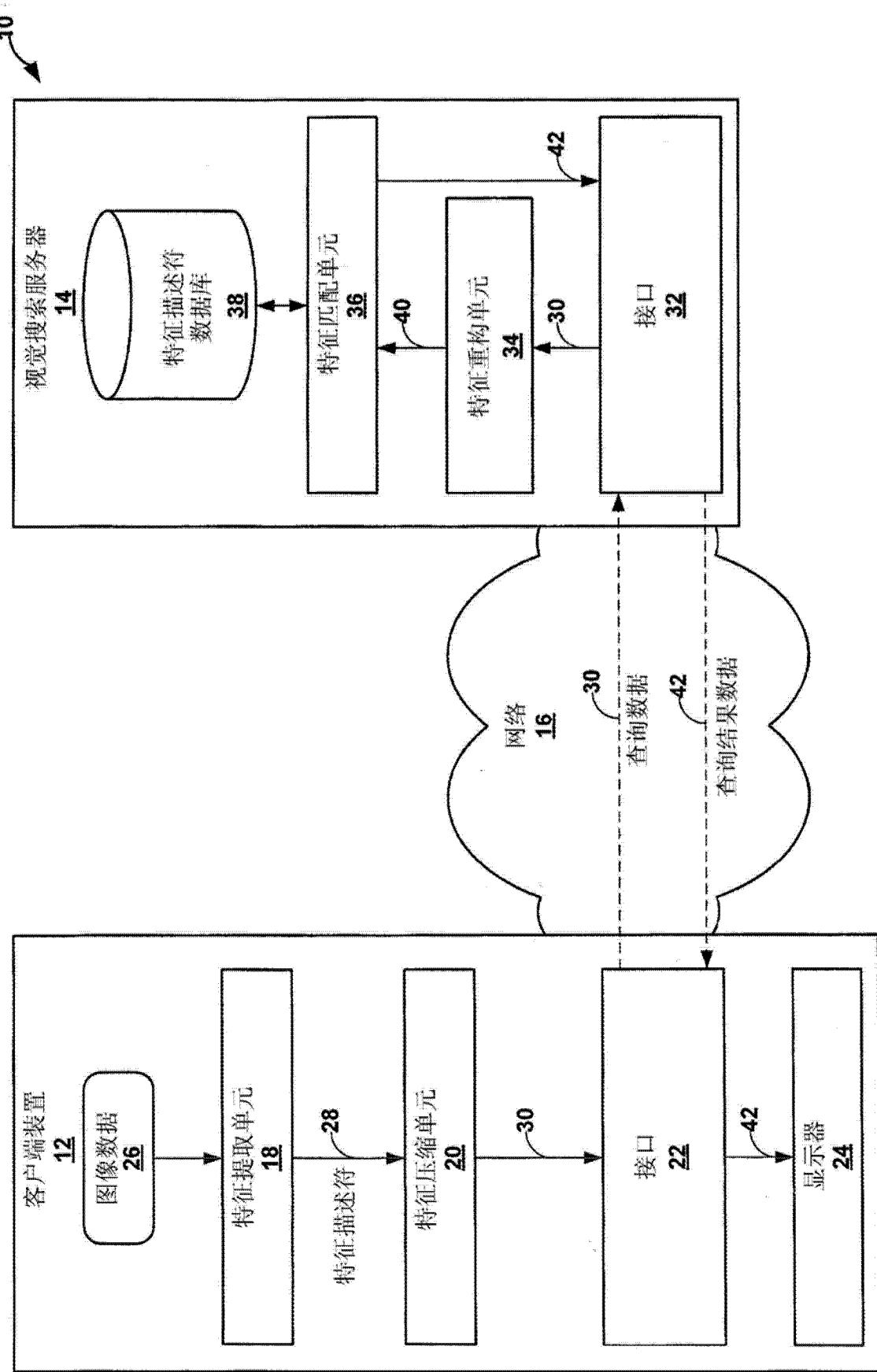
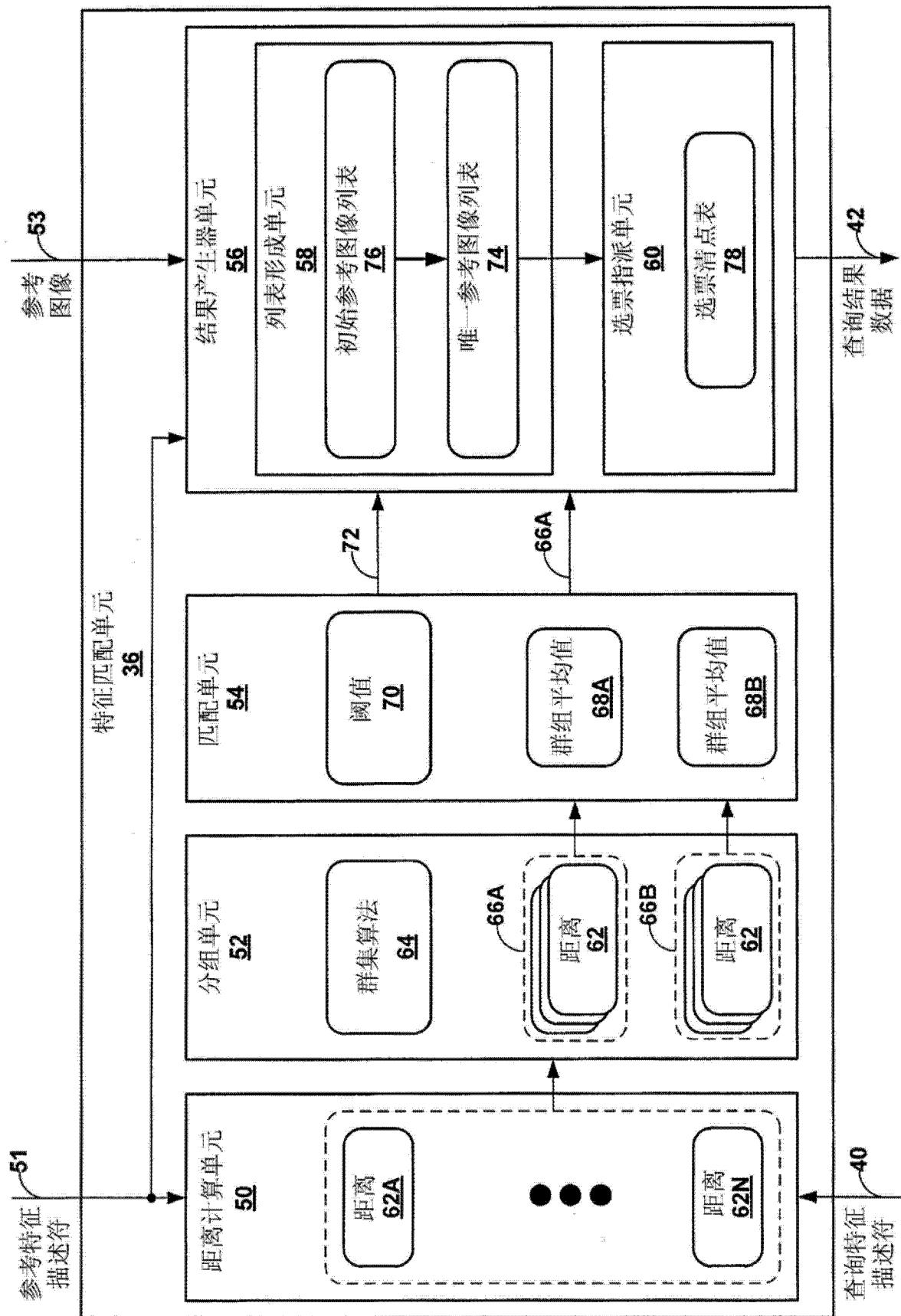


图 1



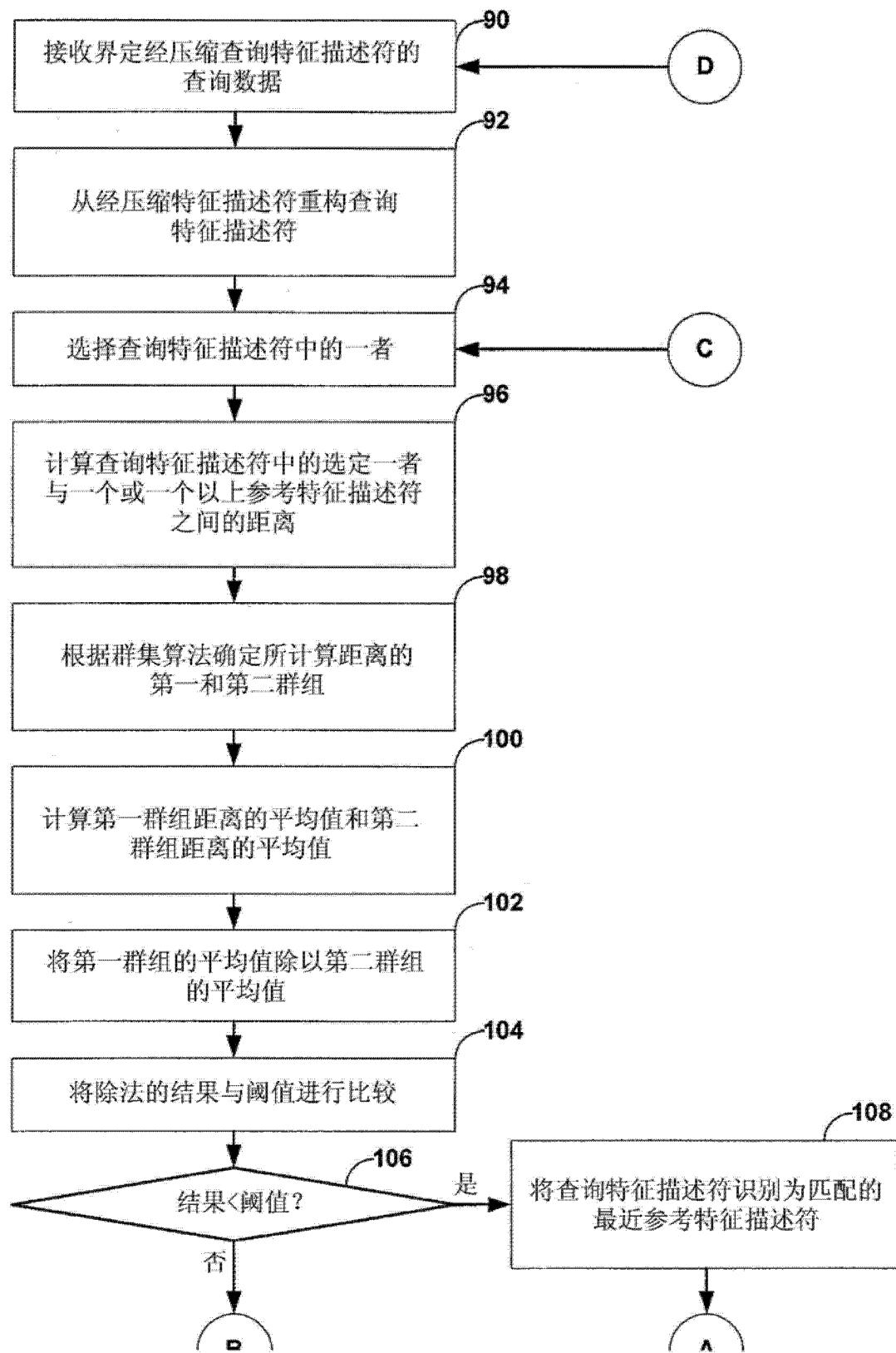


图 3A

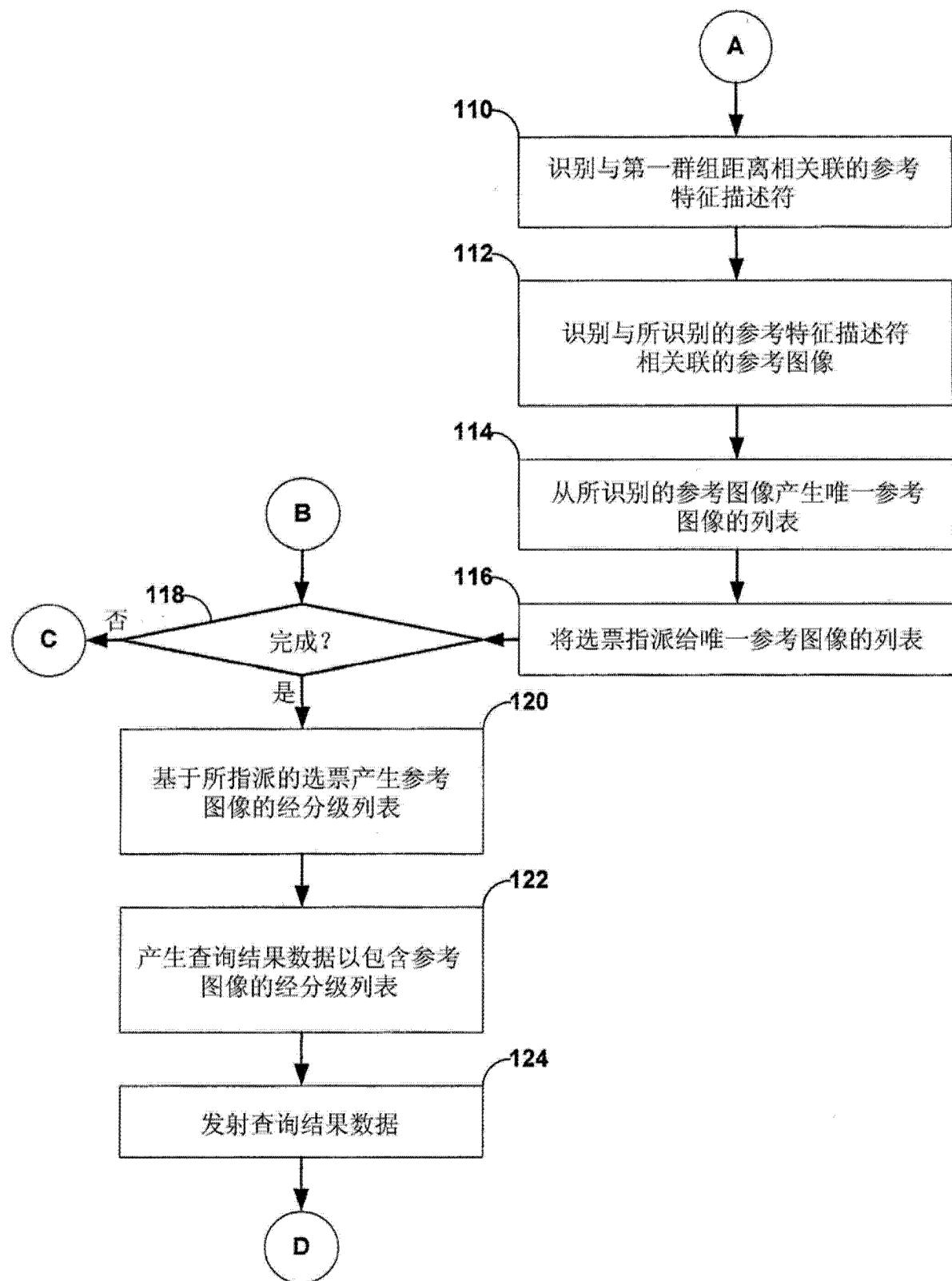


图 3B

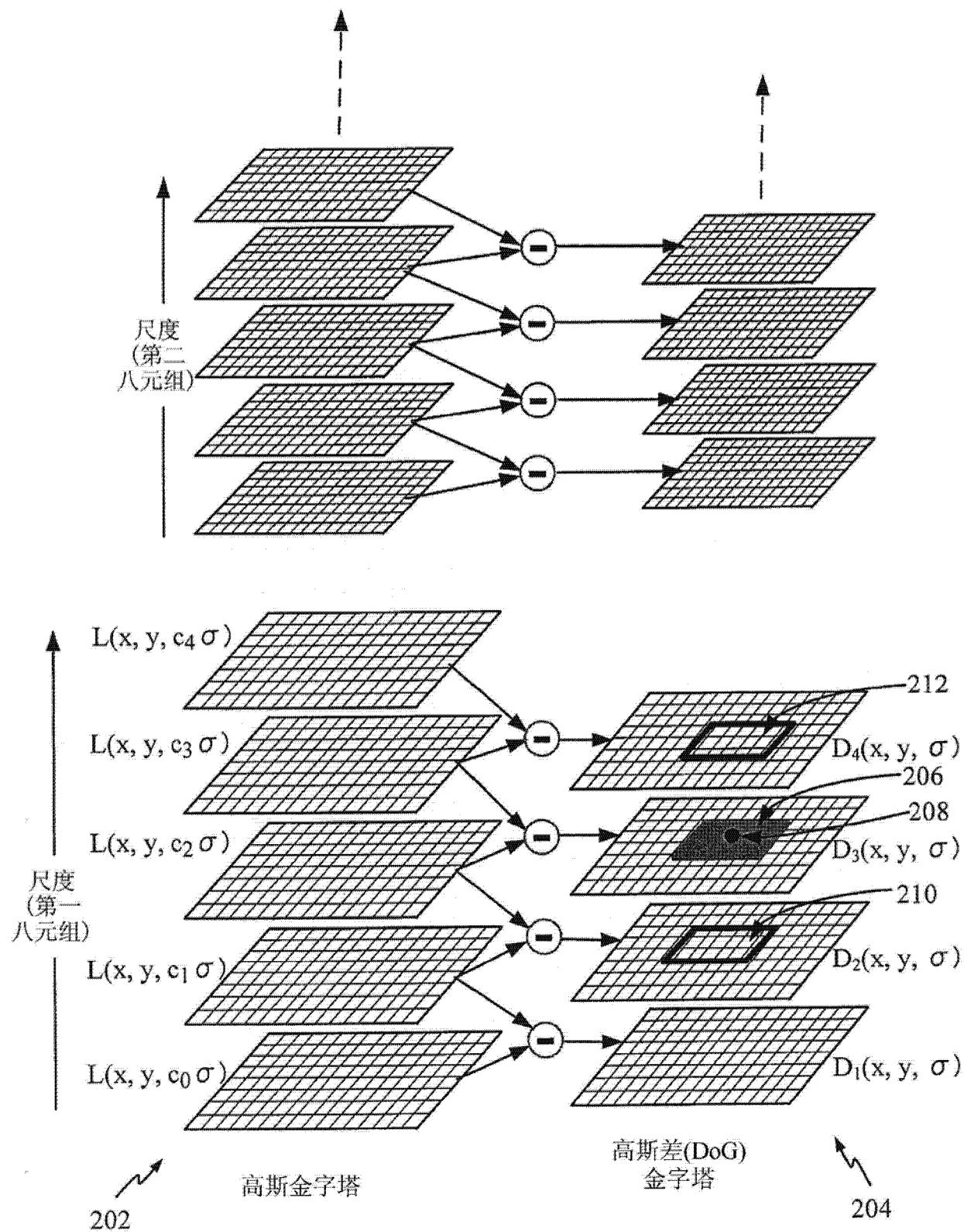


图 4

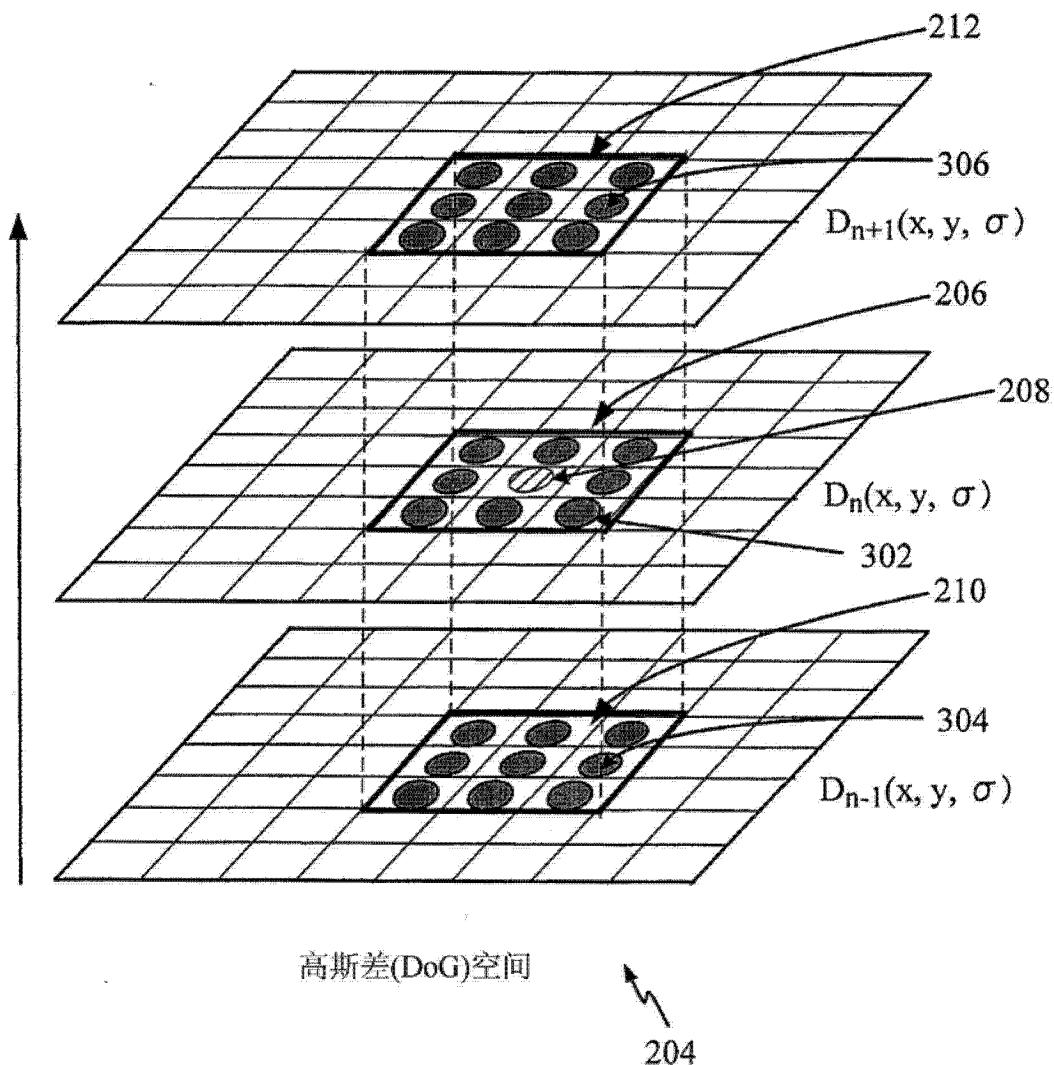


图 5

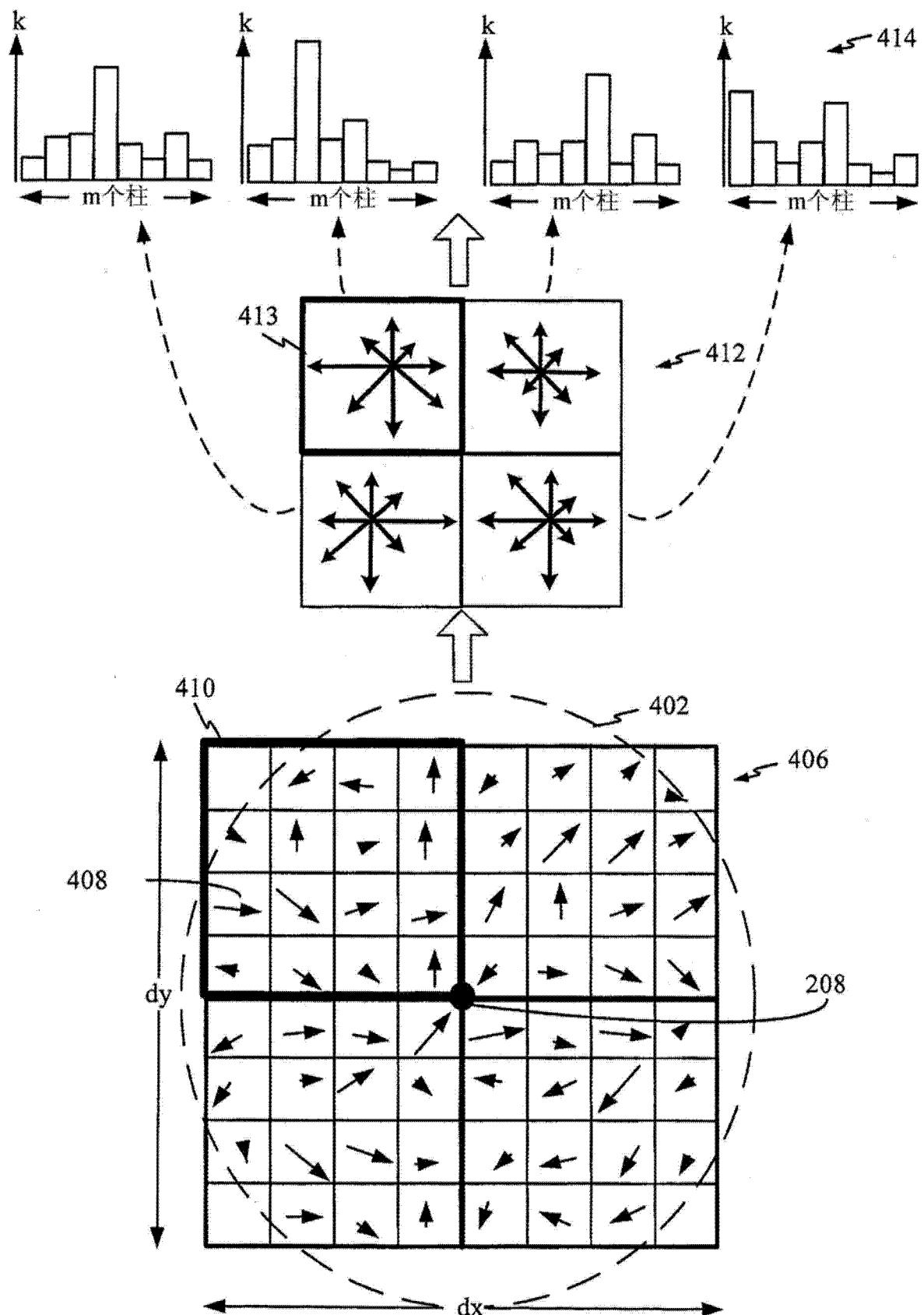


图 6

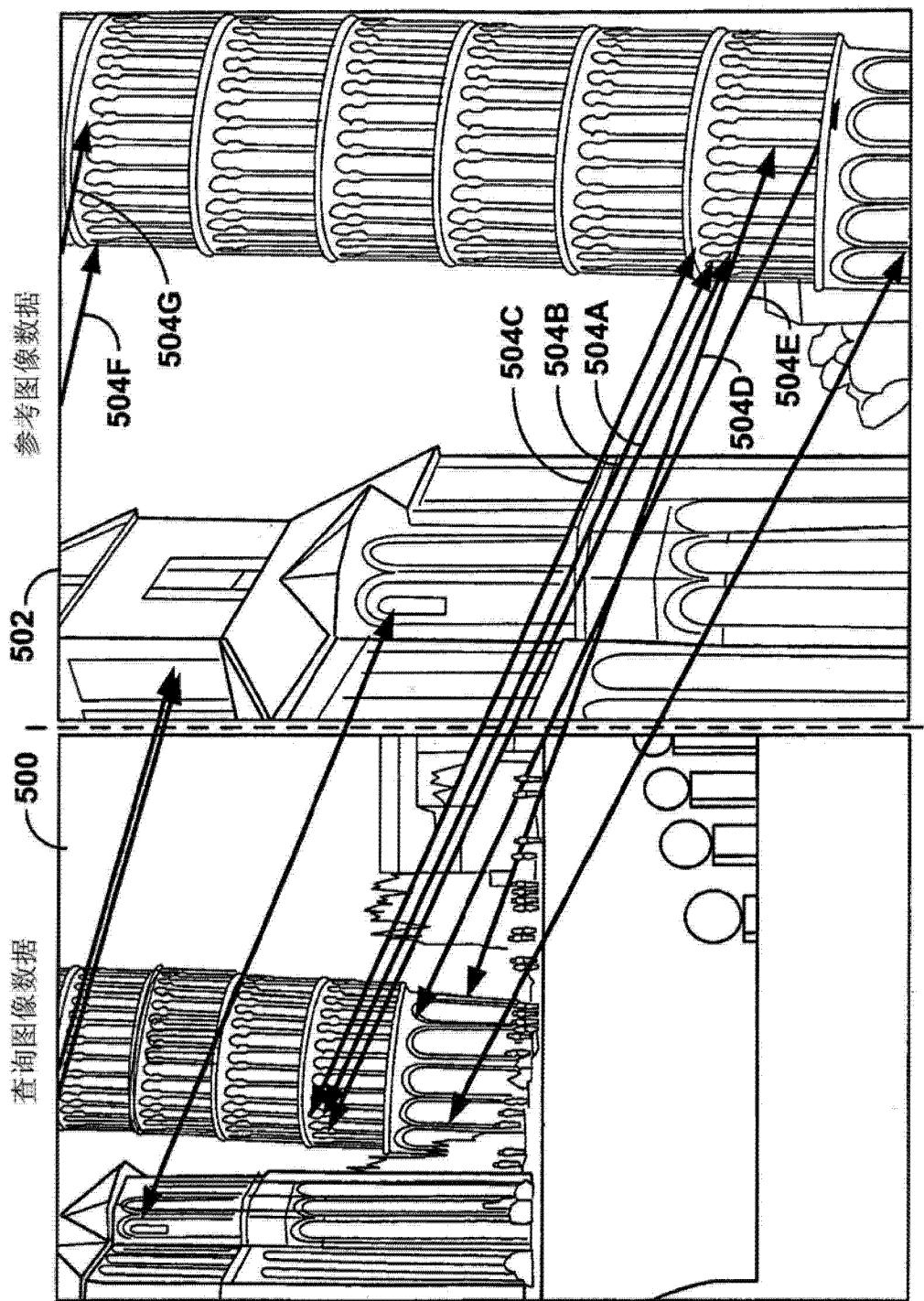


图 7