

(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
2015 年 12 月 30 日 (30.12.2015) WIPO | PCT

(10) 国际公布号
WO 2015/197029 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2015/082550
- (22) 国际申请日: 2015 年 6 月 26 日 (26.06.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410302816.X 2014 年 6 月 27 日 (27.06.2014) CN
201410306005.7 2014 年 6 月 30 日 (30.06.2014) CN
- (71) 申请人: 北京奇虎科技有限公司 (BEIJING QIHOO TECHNOLOGY COMPANY LIMITED) [CN/CN]; 中国北京市西城区新街口外大街 28 号 D 座 112 室 (德胜园区), Beijing 100088 (CN)。奇智软件 (北京) 有限公司 (QIZHI SOFTWARE (BEIJING) COMPANY LIMITED) [CN/CN]; 中国北京市朝阳区酒仙桥路 6 号院 2 号楼 B 座 2 层、3 层 301-306 室, Beijing 100015 (CN)。
- (72) 发明人: 朱茂清 (ZHU, Maoqing); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。唐雨 (TANG, Yu); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。薛红霞 (XUE, Hongxia); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。胡金辉 (HU, Jinhui); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。李璋 (LI, Zhang); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。韩玉刚 (HAN, Yugang); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。
- (74) 代理人: 北京智汇东方知识产权代理事务所 (普通合伙) (WISEAST INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区花园路 13 号 5 幢 320 房间, Beijing 100088 (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,

[见续页]

(54) Title: HUMAN FACE SIMILARITY RECOGNITION METHOD AND SYSTEM

(54) 发明名称: 一种人脸相似度识别方法和系统

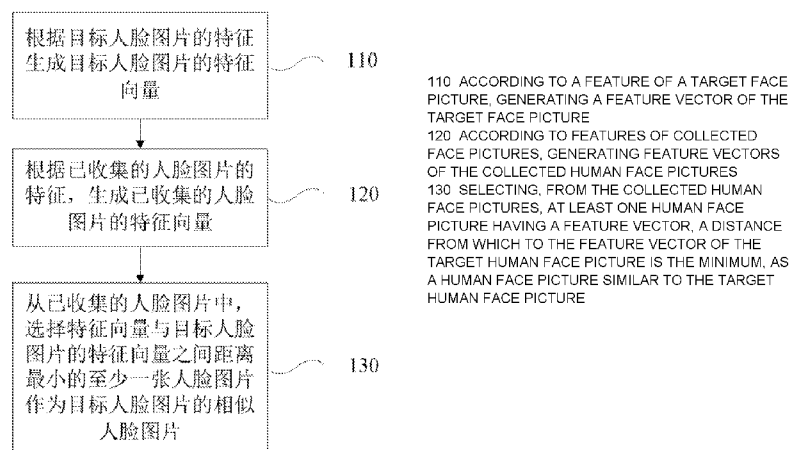


图 1 / FIG. 1

(57) Abstract: Provided are a human face similarity recognition method and system, which relate to the technical field of computers, and are used for accurately recognizing similar human face pictures. The human face similarity recognition method comprises: according to a feature of a target human face picture, generating a feature vector of the target human face picture; according to features of collected human face pictures, generating feature vectors of the collected human face pictures; and selecting, from the collected human face pictures, at least one human face picture having a feature vector, a distance from which to the feature vector of the target human face picture is the minimum, as a human face picture similar to the target human face picture. The present invention is beneficial to recognizing different pictures of the same human face having a difference in the aspects such as expression, make-up, human face angle, etc.

(57) 摘要: 本发明提供了一种人脸相似度识别方法和系统, 涉及计算机技术领域, 用于解决准确识别出相似人脸图片。人脸相似度识别方法包括: 根据目标人脸图片的特征, 生成目标人脸图片的特征向量; 根据已收集的人脸图片的特征, 生成已收集的人脸图片的特征向量; 从已收集的人脸图片中, 选择特征向量与目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为目标人脸图片的相似人脸图片。通过本发明, 有利于识别出表情、化妆、脸角度等方面存在区别的同一人脸的不同图片。

WO 2015/197029 A1



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)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种人脸相似度识别方法和系统

技术领域

本发明涉及计算机技术领域，具体而言，涉及一种人脸相似度识别方法和系统。

5

背景技术

现有技术中的人脸相似度计算，是通过将不同的人脸图片裁剪并转化为单通道图像，获取单通道图像的直方图，以及通过比较不同人脸图片的直方图之间的差异，来计算不同人脸之间的相似度。

10 上述方案的缺陷在于：同一人脸上出现表情、化妆、脸角度等方面发生变化之后，会造成同一人脸的不同图片之间的直方图出现非常大的差异，则基于直方图进行人脸相似度的计算，可能得到同一人脸的不同图片之间的相似度较小的结果，可见计算结果相当不准确。

发明内容

15 鉴于上述问题，提出了本发明以便提供一种克服上述问题或者至少部分地解决上述问题的人脸相似度识别方法和系统。

依据本发明实施例的一个方面，提供了一种人脸相似度识别方法，包括：根据目标人脸图片的特征，生成所述目标人脸图片的特征向量；根据已收集的人脸图片的特征，
20 生成所述已收集的人脸图片的特征向量；从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为所述目标人脸图片的相似人脸图片。

可选地，所述从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张图片作为所述目标人脸图片的相似人脸图片的步骤包
25 括：

将所述已收集的人脸图片聚合为多个分类；

根据所述每个分类中的人脸图片的特征向量，计算所述每个分类中的人脸图片的向量中心点；

30 将与所述目标人脸图片的特征向量之间距离最小的向量中心点对应分类中的人脸图片，作为所述目标人脸图片的相似人脸图片。

可选地，所述方法还包括：

将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分。

35 可选地，所述将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分的步骤包括：

在 $D_x \leq D_{\min}$ 时, 取 $S = S_{\max}$, D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离, $D_{\min}$ 为预设最小距离, S 为所述相似人脸图片与所述目标人脸图片的相似度得分, $S_{\max}$ 为预设最大相似度得分; 和/或

在 $D_i < D_x \leq D_{i+1}$ 时, 取 $S = S_i + K(D_x - D_i)$, 其中 $K = (S_{i+1} - S_i) / (D_{i+1} - D_i)$, D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离, D_i 为预设第一人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离, D_{i+1} 为预设第二人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离, S_i 为所述预设第一人脸图片与所述目标人脸图片的相似度得分, S_{i+1} 为所述预设第二人脸图片与所述目标人脸图片的相似度得分; 和/或

在 $D_x > D_{\max}$ 时, 取 $S = S_{\min}$, D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离, $D_{\max}$ 为预设最大距离, S 为所述相似人脸图片与所述目标人脸图片的相似度得分, $S_{\min}$ 为预设最小相似度得分。

可选地, 所述方法还包括:

在所述相似人脸图片为多张时, 根据与所述目标人脸图片的相似度高低, 对多张所述相似人脸图片进行排序。

可选地, 所述从所述已收集的人脸图片中, 选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张图片作为所述目标人脸图片的相似人脸图片的步骤包括:

将已收集的人脸图片进行聚类得到多个第 1 层分类, 并通过迭代方式继续对至少一个第 i 层分类中的人脸图片进行聚类得到多个第 $i+1$ 层分类, i 从 1 向后进行整数取值;

识别出目标人脸图片所属的第 1 层分类, 并通过迭代方式继续在所述目标人脸图片所属的第 j 层分类中识别出所述目标人脸图片所属的第 $j+1$ 层分类, j 从 1 向后依序进行整数取值;

通过所述迭代方式直至在所述目标人脸图片所属的第 j 层分类中不存在第 $j+1$ 层分类时, 从所述目标人脸图片所属的第 j 层分类中, 识别出所述目标人脸图片的相似人脸图片。

可选地, 所述将已收集的人脸图片进行聚类得到多个第 1 层分类的步骤包括:

设置多个初始中心点, 并根据所述已收集的人脸图片的特征向量与每个所述初始中心点的距离远近, 将所述已收集的人脸图片分为多个第 1 层分类, 并根据每个第 1 层分类的人脸图片的特征向量, 计算所述每个第 1 层分类的向量中心点。

可选地, 所述将已收集的人脸图片进行聚类得到多个第 1 层分类的步骤还包括:

计算所述每个第 1 层分类的初始中心点与向量中心点之间的方差;

如所述方差的大小超过预设阈值, 则重新设置初始中心点, 并重新将所述已收集的人脸图片分为多个第 1 层分类, 并重新计算所述每个第 1 层分类的向量中心点。

可选地, 所述识别出目标人脸图片所属的第 1 层分类的步骤包括:

选择向量中心点与所述目标人脸图片的特征向量之间距离最小的第 1 层分类, 作为

所述目标人脸图片所属的第 1 层分类。

可选地，所述识别出所述目标人脸图片的相似人脸图片的步骤包括：

从所述第 j 层分类的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片，作为所述目标人脸图片的所述相似人脸图片。

5 依据本发明实施例的另一个方面，还提供了一种人脸相似度识别系统，包括：第一特征向量生成模块，用于根据目标人脸图片的特征，生成所述目标人脸图片的特征向量；第二特征向量生成模块，用于根据已收集的人脸图片的特征，生成所述已收集的人脸图片的特征向量；第一相似人脸图片识别模块，用于从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为所述目标
10 人脸图片的相似人脸图片。

可选地，所述系统还包括：

第一分类模块，用于将所述已收集的人脸图片聚合为多个分类；

向量中心点计算模块，用于根据所述每个分类中的人脸图片的特征向量，计算所述每个分类中的人脸图片的向量中心点；

15 所述第一相似人脸图片识别模块用于将与所述目标人脸图片的特征向量之间距离最小的向量中心点对应分类中的人脸图片，作为所述目标人脸图片的相似人脸图片。

可选地，所述系统还包括：

相似度得分计算模块，用于将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分。
20 分。

可选地，所述相似度得分计算模块在 $D_x \leq D_{\min}$ 时，取 $S = S_{\max}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， $D_{\min}$ 为预设最小距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， $S_{\max}$ 为预设最大相似度得分；和/或

25 所述相似度得分计算模块在 $D_i < D_x \leq D_{(i+1)}$ 时，取 $S = S_i + K(D_x - D_i)$ ，其中 $K = (S_{(i+1)} - S_i) / (D_{(i+1)} - D_i)$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， D_i 为预设第一人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， D_{i+1} 为预设第二人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， S_i 为所述预设第一人脸图片与所述目标人脸图片的相似度得分， $S_{(i+1)}$
30 为所述预设第二人脸图片与所述目标人脸图片的相似度得分；和/或

所述相似度得分计算模块在 $D_x > D_{\max}$ 时，取 $S = S_{\min}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， $D_{\max}$ 为预设最大距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， $S_{\min}$ 为预设最小相似度得分。

可选地，所述系统还包括：

35 排序模块，用于在所述相似人脸图片为多张时，根据与所述目标人脸图片的相似度高低，对多张所述相似人脸图片进行排序。

可选地，所述系统还包括：

第二分类模块，用于将已收集的人脸图片进行聚类得到多个第 1 层分类，并通过迭代方式继续对至少一个第 i 层分类中的人脸图片进行聚类得到多个第 $i+1$ 层分类， i 从 1 向后依序进行整数取值；

5 分类迭代识别模块，用于识别出目标人脸图片所属的第 1 层分类，并通过迭代方式继续在所述目标人脸图片所属的第 j 层分类中识别出所述目标人脸图片所属的第 $j+1$ 层分类， j 从 1 向后进行整数取值；

第二相似人脸图片识别模块，用于在所述目标人脸图片所属的第 j 层分类中不存在第 $j+1$ 层分类时，从所述目标人脸图片所属的第 j 层分类中，识别出所述目标人脸图片的相似人脸图片。
10

可选地，所述第二分类模块设置多个初始中心点，并根据所述已收集的人脸图片的特征向量与每个所述初始中心点的距离远近，将所述已收集的人脸图片分为多个第 1 层分类，并根据每个第 1 层分类的人脸图片的特征向量，计算所述每个第 1 层分类的向量中心点。

15 可选地，所述系统还包括：

方差计算模块，计算所述每个第 1 层分类的初始中心点与向量中心点之间的方差；

如所述方差的大小超过预设阈值，则所述分类模块重新设置初始中心点，并将重新将所述已收集的人脸图片分为多个第 1 层分类，并重新计算所述每个第 1 层分类的向量中心点。

20 可选地，所述第二分类模块选择向量中心点与所述目标人脸图片的特征向量之间距离最小的第 1 层分类，作为所述目标人脸图片所属的第 1 层分类。

可选地，所述第二相似人脸图片识别模块从所述第 j 层分类的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片，作为所述目标人脸图片的所述相似人脸图片。

25 根据本发明的又一个方面，提供了一种计算机程序，其包括计算机可读代码，当所述计算机可读代码在计算设备上运行时，导致所述计算设备执行本发明所述的方法。

根据本发明的再一个方面，提供了一种计算机可读介质，其中存储了本发明所述的计算机程序。

本发明的有益效果为：

30 本发明实施例中，将不同人脸图片的特征处理为特征向量并计算特征向量之间的向量距离，以及根据向量距离的大小来识别相似的人脸图片；在同一人脸的不同图片上出现了表情、化妆、脸角度等方面上的变化时，不同图片上的人脸特征却可以保持不变或变化较小，进而不同图片的特征向量之间的距离也必然较小，即说明不同人脸图片之间的相似度较大，这有利于识别出表情、化妆、脸角度等方面存在区别的同一人脸的不同
35 图片。

上述说明仅是本发明技术方案的概述，为了能够更清楚了解本发明的技术手段，而

可依照说明书的内容予以实施，并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂，以下特举本发明的具体实施方式。

附图说明

5 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

图 1 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

图 2 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

10 图 3 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

图 4 示出了根据本发明一个实施例的人脸识别方法的工作示意图；

图 5 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

图 6 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

图 7 示出了根据本发明一个实施例的人脸相似度识别方法的流程图；

15 图 8 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 9 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 10 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 11 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 12 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

20 图 13 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 14 示出了根据本发明的一个实施例的人脸相似度识别系统的框图；

图 15 示意性地示出了用于执行根据本发明方法的计算设备的框图；以及

图 16 示意性地示出了用于保持或者携带实现根据本发明方法的程序代码的存储单元。

25

具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

30

本发明实施例提供了一种人脸相似度识别方法。图 1 示出了根据本发明一个实施例的人脸相似度识别方法的处理流程图。参见图 1，该人脸相似度识别方法至少包括步骤 110 至步骤 130。

步骤 110，根据目标人脸图片的特征，生成目标人脸图片的特征向量。目标人脸图片的特征可以实时提取。依据提取特征的数量，则特征向量可以是多维向量，例如 400 维向量。本实施例的特征包括但不限于脸部器官形状、位置等等。

35

步骤 120, 根据已收集的人脸图片的特征, 生成已收集的人脸图片的特征向量。已收集的人脸图片的特征可以预先提取并存储。依据提取特征的数量, 则特征向量可以是多维向量, 例如 400 维向量。本实施例的特征包括但不限于脸部器官形状、位置等等。

步骤 130, 从已收集的人脸图片中, 选择特征向量与目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为目标人脸图片的相似人脸图片。在本实施例的技术方案中, 如目标人脸图片和某一已收集人脸图片为同一人脸的不同图片, 二者的特征必然相同或差距较小, 则二者的特征向量之间的距离也必然较小, 所以本实施例的技术方案有利于识别出同一人脸的不同图片。

如图 2 所示, 本发明的另一实施例提出一种人脸相似度识别方法, 与上述实施例相比, 本实施例的人脸相似度识别方法, 其中, 步骤 130 包括:

步骤 131, 将已收集的人脸图片聚合为多个分类。例如, 将已收集的人脸图片分为 C1、C2、C3 三个分类。现有的聚类方式较多, 都可以采用到本实施例的技术方案中。

步骤 132, 根据每个分类中的人脸图片的特征向量, 计算每个分类中的人脸图片的向量中心点。例如, 取三个分类的向量中心点分别为 R1、R2、R3。

步骤 133, 将与目标人脸图片的特征向量之间距离最小的向量中心点对应分类中的人脸图片, 作为目标人脸图片的相似人脸图片。例如, 假设目标人脸图片 Q 与 R1、R2、R3 的向量距离值分别为 1.4、1.25、0.2, 其中 Q 与 R3 之间距离的最小, 则取 R3 对应的 C3 分类中的人脸图片为相似人脸图片。

在本实施例的技术方案中, 通过聚类得到多个分类的中心向量点, 并将中心向量点与目标人脸图片的特征向量进行比较, 避免了将已收集的所有人脸图片的特征向量逐个与目标人脸图片的特征向量进行比较, 降低了运算量, 提高了图片识别的效率。

本发明的另一实施例提出一种人脸相似度识别方法, 与上述实施例相比, 本实施例的人脸相似度识别方法, 其中, 还包括:

将相似人脸图片的特征向量与目标人脸图片的特征向量之间的距离, 转换为相似人脸图片与目标人脸图片之间的相似度得分。例如, 结合前述的实施例, 设 C3 分类中的人脸图片与目标人脸图片中的最小向量距离依次为 0.01、0.2、1.2, 则将该三个距离值根据预定的公式转换为 100、91、85 的相似度得分, 则相似度得分的高低可以反映出目标人脸图片与相似人脸图片之间的相似度高低。

本发明的另一实施例提出一种人脸相似度识别方法, 与上述实施例相比, 本实施例的人脸相似度识别方法, 其中, 将相似人脸图片的特征向量与目标人脸图片的特征向量之间的距离, 转换为相似人脸图片与目标人脸图片之间的相似度得分的步骤包括:

在 $D_x \leq D_{min}$ 时, 取 $S = S_{max}$, D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离, D_{min} 为预设最小距离, S 为相似人脸图片与目标人脸图片的相

似度得分， $S_{\max}$ 为预设最大相似度得分。

在 $D_i < D_x \leq D_{i+1}$ 时，取 $S = S_i + K(D_x - D_i)$ ，其中 $K = (S_{i+1} - S_i) / (D_{i+1} - D_i)$ ， D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离， D_i 为预设第一人脸图片的特征向量与目标人脸图片的特征向量之间的距离， D_{i+1} 为预设第二人脸图片的特征向量与目标人脸图片的特征向量之间的距离， S_i 为预设第一人脸图片与目标人脸图片的相似度得分， S_{i+1} 为预设第二人脸图片与目标人脸图片的相似度得分。

在 $D_x > D_{\max}$ 时，取 $S = S_{\min}$ ， D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离， $D_{\max}$ 为预设最大距离， S 为相似人脸图片与目标人脸图片的相似度得分， $S_{\min}$ 为预设最小相似度得分。

10 在本实施例的技术方案中，提供了一种将向量距离转换为相似度得分的技术方案，且相似度得分随向量距离的减小而降低，能够合理反映出目标人脸图片与相似人脸图片的相似程度。

15 本发明的另一实施例提出一种人脸相似度识别方法，与上述实施例相比，本实施例的人脸相似度识别方法，其中，还包括：

在相似人脸图片为多张时，根据与目标人脸图片的相似度高低，对多张相似人脸图片进行排序。

在本实施例的技术方案中，因为相似度最高的人脸图片通常为用户所需图片，通过对多张相似人脸图片进行排序，有利于将用户需求的图片快速提供给用户。

20

如图 3 所示，本发明的另一实施例提出一种人脸相似度识别方法，与上述实施例相比，本实施例的人脸相似度识别方法，其中，步骤 130 包括：

步骤 134，将已收集的人脸图片进行聚类得到多个第 1 层分类，并通过迭代方式继续对至少一个第 i 层分类中的人脸图片进行聚类得到多个第 $i+1$ 层分类， i 从 1 向后进行整数取值。本实施例中，形成的多层分类结构如图 4 所示，例如，其中 C1 分类包括 C11……C1m 等多个分类，C11 分类中又包括 CN1、CN2 等分类。

步骤 135，识别出目标人脸图片所属的第 1 层分类，并通过迭代方式继续在目标人脸图片所属的第 j 层分类中识别出目标人脸图片所属的第 $j+1$ 层分类， j 从 1 向后依序进行整数取值。

30 步骤 136，通过迭代方式直至在目标人脸图片所属的第 j 层分类中不存在第 $j+1$ 层分类时，从目标人脸图片所属的第 j 层分类中，识别出目标人脸图片的相似人脸图片。

本实施例的技术方案中，通过对聚类上一层聚类结果进行再次划分聚类通过迭代方式将已收集的人脸图片聚类为多层结构，以及通过迭代方式逐层寻找目标人脸图片的所属分类，直至最终找到目标人脸图片的相似人脸图片；相对于现有的技术方案，本发明的技术方案中的计算量非常小，大大提高了人脸识别效率。

35

如图 5 所示, 本发明的另一个实施例提供了一种人脸相似度识别方法, 其中, 步骤 134 包括:

步骤 1341, 根据已收集的人脸图片的特征, 生成已收集的人脸图片的特征向量。本实施例基于对已收集的人脸图片特征的提取, 已收集人脸图片的特征可以预先提取并存储 5 在特征库中。

步骤 1342, 设置多个初始中心点, 并根据已收集的人脸图片的特征向量与每个初始中心点的距离远近, 将已收集的人脸图片分为多个第 1 层分类, 并根据每个第 1 层分类的人脸图片的特征向量, 计算每个第 1 层分类的向量中心点。

则根据本实施例的技术方案, 都按照已收集的人脸图片特征向量与初始中心点之间的 10 距离, 将其分配到最邻近的分类中, 之后计算向量中心点; 如此反复即可快速形成多层聚类结构。

本发明的另一个实施例提供了一种人脸相似度识别方法, 其中, 步骤 134 还包括:

1343, 计算每个第 1 层分类的初始中心点与向量中心点之间的方差。本实施例基于 15 对目标人脸图片特征的提取, 目标人脸图片的特征可以实时提取。

1344, 如方差的大小超过预设阈值, 则重新设置初始中心点, 并重新将已收集的人脸图片分为多个第 1 层分类, 并重新计算每个第 1 层分类的向量中心点。

在本实施例的技术方案中, 如果方差 <0.000001 (示例, 可以取其他值), 则表示分类中都是特征相近的人脸图片, 否则表示分类中具有特征差距明显不适合置于同一分类 20 的人脸图片, 所以需要重新分类。此时, 本实施例的人脸识别方法的工作流程可以如图 6 所示, 图中, 计算到指定层数的步骤是指, 将已收集的人脸图片聚类为指定层数的聚类结构。

如图 7 所示, 本发明的另一个实施例提供了一种人脸相似度识别方法, 其中, 步骤 25 135 包括:

步骤 1351, 根据目标人脸图片的特征, 生成目标人脸图片的特征向量。

步骤 1352, 选择向量中心点与目标人脸图片的特征向量之间距离最小的第 1 层分类, 作为目标人脸图片所属的第 1 层分类。

则根据本实施例的技术方案, 在各层结构中, 都将目标人脸图片的特征向量与所属 30 上层分类中多个下层分类的向量中心点进行比较, 可以快速找到目标人脸图片所属的最小分类。

本发明的另一个实施例提供了一种人脸相似度识别方法, 其中, 步骤 136 包括:

从第 j 层分类的人脸图片中, 选择特征向量与目标人脸图片的特征向量之间距离最 35 小的至少一张人脸图片, 作为目标人脸图片的相似人脸图片。

结合以上实施例, 假定已收集人脸图片为 1000 万, 需要在其中检索目标人脸图片的

相似人脸图片时：

1、若使用直接比较方式，需要将目标人脸图片的特征向量与已收集人脸图片的特征向量比较 1000 万次。

2、若使用传统聚类方式，将 1000 万数据划分为 10000 个聚类，则需要将目标人脸图片与聚类的向量中心点比较 10000 次，每个聚类平均中有 1000 条数据，则每个分类内部需要比较 1000 次，总体比较 $10000+k \times 1000$ 次，如 k 取 10 则比较次数为： $10000+10 \times 1000=20000$ 次。通过现有的近邻算法在聚类中寻找目标人脸图片的相似人脸图片时， k 表示选取 k 个近邻中心点，10 为常见取值。

3、使用本实施例的技术方案，若划分为 2 层，第一层 100 个聚类，第二层为 200 个聚类，则第二层平均每个聚类中有 500 条数据，比较次数约 $100+m \times 200+n \times 500$ 次，如取 $m=3$ ， $n=10$ ，比较次数为： $100+3 \times 200+10 \times 500=11100$ 次，和 1、2 比较，可以显著减少比较次数。同理， m 表示第一层选取 m 个近邻中心点， n 表示第二层选取 n 个近邻中心点，3、10 为常见取值。

如图 8 所示，本发明的另一实施例提供了一种人脸相似度识别系统，其包括：

第一特征向量生成模块 310，用于根据目标人脸图片的特征，生成目标人脸图片的特征向量。目标人脸图片的特征可以实时提取。依据提取特征的数量，则特征向量可以是多维向量，例如 400 维向量。本实施例的特征包括但不限于脸部器官形状、位置等等。

第二特征向量生成模块 320，用于根据已收集的人脸图片的特征，生成已收集的人脸图片的特征向量。已收集的人脸图片的特征可以预先提取并存储。依据提取特征的数量，则特征向量可以是多维向量，例如 400 维向量。本实施例的特征包括但不限于脸部器官形状、位置等等。

第一相似人脸图片识别模块 330，用于从已收集的人脸图片中，选择特征向量与目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为目标人脸图片的相似人脸图片。在本实施例的技术方案中，如目标人脸图片和某一已收集人脸图片为同一人脸的不同图片，二者的特征必然相同或差距较小，则二者的特征向量之间的距离也必然较小，所以本实施例的技术方案有利于识别出同一人脸的不同图片。

如图 9 所示，本发明的另一实施例提出一种人脸相似度识别系统，与上述实施例相比，本实施例的人脸相似度识别系统，其中，还包括：

第一分类模块 340，用于将已收集的人脸图片聚合为多个分类。例如，将已收集的人脸图片分为 C1、C2、C3 三个分类。现有的聚类方式较多，都可以采用到本实施例的技术方案中。

向量中心点计算模块 350，用于根据每个分类中的人脸图片的特征向量，计算每个分类中的人脸图片的向量中心点。例如，取三个分类的向量中心点分别为 R1、R2、R3。

第一相似人脸图片识别模块 330，用于将与目标人脸图片的特征向量之间距离最小的

向量中心点对应分类中的人脸图片，作为目标人脸图片的相似人脸图片。例如，假设目标人脸图片 Q 与 R1、R2、R3 的向量距离值分别为 1.4、1.25、0.2，其中 Q 与 R3 之间距离的最小，则取 R3 对应的 C3 分类中的人脸图片为相似人脸图片。

在本实施例的技术方案中，通过聚类得到多个分类的中心向量点，并将中心向量点与目标人脸图片的特征向量进行比较，避免了将已收集的所有人脸图片的特征向量逐个与目标人脸图片的特征向量进行比较，降低了运算量，提高了图片识别的效率。

如图 10 所示，本发明的另一实施例提出一种人脸相似度识别系统，与上述实施例相比，本实施例的人脸相似度识别系统，其中，还包括：

相似度得分计算模块 360，用于将相似人脸图片的特征向量与目标人脸图片的特征向量之间的距离，转换为相似人脸图片与目标人脸图片之间的相似度得分。例如，结合前述的实施例，设 C3 分类中的人脸图片与目标人脸图片中的最小向量距离依次为 0.01、0.2、1.2，则将该三个距离值根据预定的公式转换为 100、91、85 的相似度得分，则相似度得分的高低可以反映出目标人脸图片与相似人脸图片之间的相似度高低。

15

本发明的另一实施例提出一种人脸相似度识别系统，与上述实施例相比，本实施例的人脸相似度识别系统，其中，相似度得分计算模块 360 在 $D_x \leq D_{min}$ 时，取 $S = S_{max}$ ， D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离， D_{min} 为预设最小距离， S 为相似人脸图片与目标人脸图片的相似度得分， S_{max} 为预设最大相似度得分。

20

相似度得分计算模块 360 在 $D_i < D_x \leq D_{(i+1)}$ 时，取 $S = S_i + K(D_x - D_i)$ ，其中 $K = (S_{(i+1)} - S_i) / (D_{(i+1)} - D_i)$ ， D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离， D_i 为预设第一人脸图片的特征向量与目标人脸图片的特征向量之间的距离， D_{i+1} 为预设第二人脸图片的特征向量与目标人脸图片的特征向量之间的距离， S_i 为预设第一人脸图片与目标人脸图片的相似度得分， $S_{(i+1)}$ 为预设第二人脸图片与目标人脸图片的相似度得分。

25

相似度得分计算模块 360 在 $D_x > D_{max}$ 时，取 $S = S_{min}$ ， D_x 为目标人脸图片的特征向量与相似人脸图片之间的特征向量之间的距离， D_{max} 为预设最大距离， S 为相似人脸图片与目标人脸图片的相似度得分， S_{min} 为预设最小相似度得分。

30

在本实施例的技术方案中，提供了一种将向量距离转换为相似度得分的技术方案，且相似度得分随向量距离的减小而降低，能够合理反映出目标人脸图片与相似人脸图片的相似程度。

35

如图 11 所示，本发明的另一实施例提出一种人脸相似度识别系统，与上述实施例相比，本实施例的人脸相似度识别系统，其中，还包括：

排序模块 370，用于在相似人脸图片为多张时，根据与目标人脸图片的相似度高低，

对多张相似人脸图片进行排序。

在本实施例的技术方案中，因为相似度最高的人脸图片通常为用户所需图片，通过对多张相似人脸图片进行排序，有利于将用户需求的图片快速提供给用户。

5 如图 12 所示，本发明的另一实施例提出一种人脸相似度识别系统，与上述实施例相比，本实施例的人脸相似度识别系统，其中，还包括：

第二分类模块 380，用于将已收集的人脸图片进行聚类得到多个第 1 层分类，并通过迭代方式继续对至少一个第 i 层分类中的人脸图片进行聚类得到多个第 $i+1$ 层分类， i 从 1 向后依序进行整数取值。本实施例中，形成的多层分类结构如图 2 所示，例如，其中
10 C1 分类包括 C11……C1m 等多个分类，C11 分类中又包括 CN1、CN2 等分类。

分类迭代识别模块 390，用于识别出目标人脸图片所属的第 1 层分类，并通过迭代方式继续在目标人脸图片所属的第 j 层分类中识别出目标人脸图片所属的第 $j+1$ 层分类， j 从 1 向后进行整数取值。

第二相似人脸图片识别模块 3100，用于在目标人脸图片所属的第 j 层分类中不存在
15 第 $j+1$ 层分类时，从目标人脸图片所属的第 j 层分类中，识别出目标人脸图片的相似人脸图片。

本实施例的技术方案中，通过对聚类上一层聚类结果进行再次划分聚类通过迭代方式将已收集的人脸图片聚类为多层结构，以及通过迭代方式逐层寻找目标人脸图片的所属分类，直至最终找到目标人脸图片的相似人脸图片；相对于现有的技术方案，本发明
20 的技术方案中的计算量非常小，大大提高了人脸识别效率。

如图 13 所示，本发明的另一个实施例提供了一种人脸相似度识别系统，其中，还包括：

第三特征向量生成模块 3110，用于根据已收集的人脸图片的特征，生成已收集的人
25 脸图片的特征向量。本实施例基于对已收集的人脸图片特征的提取，已收集人脸图片的特征可以预先提取并存储在特征库中。

第二分类模块 380 设置多个初始中心点，并根据已收集的人脸图片的特征向量与每个初始中心点的距离远近，将已收集的人脸图片分为多个第 1 层分类，并根据每个第 1 层分类的人脸图片的特征向量，计算每个第 1 层分类的向量中心点。

30 则根据本实施例的技术方案，都按照已收集的人脸图片特征向量与初始中心点之间的距离，将其分配到最邻近的分类中，之后计算向量中心点；如此反复即可快速形成多层聚类结构。

本发明的另一个实施例提供了一种人脸相似度识别系统，其中，还包括：

35 方差计算模块 3120，计算每个第 1 层分类的初始中心点与向量中心点之间的方差。本实施例基于对目标人脸图片特征的提取，目标人脸图片的特征可以实时提取。

如方差的大小超过预设阈值，则第二分类模块 380 重新设置初始中心点，并重新将已收集的人脸图片分为多个第 1 层分类，并重新计算每个第 1 层分类的向量中心点。

在本实施例的技术方案中，如果方差 <0.000001 （示例，可以取其他值），则表示分类中都是特征相近的人脸图片，否则表示分类中具有特征差距明显不适合置于同一分类的人脸图片，所以需要重新分类。此时，本实施例的人脸识别方法的工作流程可以如图 6 所示，图中，计算到指定层数的步骤是指，将已收集的人脸图片聚类为指定层数的聚类结构。

如图 14 所示，本发明的另一个实施例提供了一种人脸相似度识别系统，其中，还包括：

第四特征向量生成模块 3130，用于根据目标人脸图片的特征，生成目标人脸图片的特征向量。

第二分类模块 380 选择向量中心点与目标人脸图片的特征向量之间距离最小的第 1 层分类，作为目标人脸图片所属的第 1 层分类。

则根据本实施例的技术方案，在各层结构中，都将目标人脸图片的特征向量与所属上层分类中多个下层分类的向量中心点进行比较，可以快速找到目标人脸图片所属的最小分类。

本发明的另一个实施例提供了一种人脸相似度识别系统，其中，第二相似人脸图片识别模块 3100 从第 j 层分类的人脸图片中，选择特征向量与目标人脸图片的特征向量之间距离最小的至少一张人脸图片，作为目标人脸图片的相似人脸图片。

结合以上实施例，假定已收集人脸图片为 1000 万，需要在其中检索目标人脸图片的相似人脸图片时：

1、若使用直接比较方式，需要将目标人脸图片的特征向量与已收集人脸图片的特征向量比较 1000 万次。

2、若使用传统聚类方式，将 1000 万数据划分为 10000 个聚类，则需要将目标人脸图片与聚类的向量中心点比较 10000 次，每个聚类平均中有 1000 条数据，则每个分类内部需要比较 1000 次，总体比较 $10000+k \times 1000$ 次，如 k 取 10 则比较次数为： $10000+10 \times 1000=20000$ 次。通过现有的近邻算法在聚类中寻找目标人脸图片的相似人脸图片时， k 表示选取 k 个近邻中心点，10 为常见取值。

3、使用本实施例的技术方案，若划分为 2 层，第一层 100 个聚类，第二层为 200 个聚类，则第二层平均每个聚类中有 500 条数据，比较次数约 $100+m \times 200+n \times 500$ 次，

如取 $m=3$ ， $n=10$ ，比较次数为： $100+3 \times 200+10 \times 500=11100$ 次，和 1、2 比较，可以显著减少比较次数。同理， m 表示第一层选取 m 个近邻中心点， n 表示第二层选取 n 个近邻中心点，3、10 为常见取值。

在此处所提供的说明书中，说明了大量具体细节。然而，能够理解，本发明的实施

例可以在没有这些具体细节的情况下实践。在一些实例中，并未详细示出公知的方法、结构和技术，以便不模糊对本说明书的理解。

类似地，应当理解，为了精简本公开并帮助理解各个发明方面中的一个或多个，在上面对本发明的示例性实施例的描述中，本发明的各个特征有时被一起分组到单个实施
5 例、图、或者对其的描述中。然而，并不应将该公开的方法解释成反映如下意图：即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说，如下面的权利要求书所反映的那样，发明方面在于少于前面公开的单个实施例的所有特征。因此，遵循具体实施方式的权利要求书由此明确地并入该具体实施方式，其中每个权利要求本身都作为本发明的单独实施例。

10 本领域那些技术人员可以理解，可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件，以及此外可以把它分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外，可以采用任何组合对本说明书（包括伴随的权利要求、摘要和附图）中公开的所有特征以及如此
15 公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述，本说明书（包括伴随的权利要求、摘要和附图）中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

此外，本领域的技术人员能够理解，尽管在此所述的一些实施例包括其它实施例中
20 所包括的某些特征而不是其它特征，但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如，在下面的权利要求书中，所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

本发明的各个部件实施例可以以硬件实现，或者以在一个或者多个处理器上运行的软件模块实现，或者以它们的组合实现。本领域的技术人员应当理解，可以在实践中使用微处理器或者数字信号处理器（DSP）来实现根据本发明实施例的人脸相似度识别系
25 统中的一些或者全部部件的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的设备或者装置程序（例如，计算机程序和计算机程序产品）。这样的实现本发明的程序可以存储在计算机可读介质上，或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到，或者在载体信号上提供，或者以任何其他形式提供。

30 例如，图 15 示出了可以实现根据本发明的方法的计算设备。该计算设备传统上包括处理器 1510 和以存储器 1520 形式的计算机程序产品或者计算机可读介质。存储器 1520 可以是诸如闪存、EEPROM（电可擦除可编程只读存储器）、EPROM、硬盘或者 ROM 之类的电子存储器。存储器 1520 具有用于执行上述方法中的任何方法步骤的程序代码
1531 的存储空间 1530。例如，用于程序代码的存储空间 1530 可以包括分别用于实现上
35 面的方法中的各种步骤的各个程序代码 1531。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。这些计算机程序产品包

括诸如硬盘，紧致盘（CD）、存储卡或者软盘之类的程序代码载体。这样的计算机程序产品通常为如参考图 16 所述的便携式或者固定存储单元。该存储单元可以具有与图 15 的计算设备中的存储器 1520 类似布置的存储段、存储空间等。程序代码可以例如以适当形式进行压缩。通常，存储单元包括计算机可读代码 1531’，即可以由例如诸如 1510 之

5 类的处理器读取的代码，这些代码当由计算设备运行时，导致该计算设备执行上面所描述的方法中的各个步骤。

本文中所称的“一个实施例”、“实施例”或者“一个或者多个实施例”意味着，结合实施例描述的特定特征、结构或者特性包括在本发明的至少一个实施例中。此外，请注意，这里“在一个实施例中”的词语例子不一定全指同一个实施例。

10 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制，并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中，不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机

15 来实现。在列举了若干装置的单元权利要求中，这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

此外，还应当注意，本说明书中使用的语言主要是为了可读性和教导的目的而选择的，而不是为了解释或者限定本发明的主题而选择的。因此，在不偏离所附权利要求书

20 的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。对于本发明的范围，对本发明所做的公开是说明性的，而非限制性的，本发明的范围由所附权利要求书限定。

权 利 要 求

1、一种人脸相似度识别方法，包括：

根据目标人脸图片的特征，生成所述目标人脸图片的特征向量；

5 根据已收集的人脸图片的特征，生成所述已收集的人脸图片的特征向量；

从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为所述目标人脸图片的相似人脸图片。

2、根据权利要求 1 所述的方法，其中，所述从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张图片作为所述目标人脸
10 图片的相似人脸图片的步骤包括：

将所述已收集的人脸图片聚合为多个分类；

根据所述每个分类中的人脸图片的特征向量，计算所述每个分类中的人脸图片的向量中心点；

15 将与所述目标人脸图片的特征向量之间距离最小的向量中心点对应分类中的人脸图片，作为所述目标人脸图片的相似人脸图片。

3、根据权利要求 1-2 任一项所述的方法，其中，还包括：

将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分。

4、根据权利要求 3 所述的方法，其中，所述将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分的步骤包括：

在 $D_x \leq D_{min}$ 时，取 $S = S_{max}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， D_{min} 为预设最小距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， S_{max} 为预设最大相似度得分；和/或

25 在 $D_i < D_x \leq D_{(i+1)}$ 时，取 $S = S_i + K(D_x - D_i)$ ，其中 $K = (S_{(i+1)} - S_i) / (D_{(i+1)} - D_i)$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， D_i 为预设第一人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， D_{i+1} 为预设第二人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， S_i 为所述预设第一人脸图片与所述目标人脸图片的相似度得分， $S_{(i+1)}$ 为所述预设第二人脸图片与所述目标人脸图片的相似度得分；和/或

30 在 $D_x > D_{max}$ 时，取 $S = S_{min}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， D_{max} 为预设最大距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， S_{min} 为预设最小相似度得分。

5、根据权利要求 1-4 任一项所述的方法，其中，还包括：

35 在所述相似人脸图片为多张时，根据与所述目标人脸图片的相似度高低，对多张所述相似人脸图片进行排序。

6、根据权利要求1所述的方法，其中，所述从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张图片作为所述目标人脸图片的相似人脸图片的步骤包括：

5 将已收集的人脸图片进行聚类得到多个第1层分类，并通过迭代方式继续对至少一个第*i*层分类中的人脸图片进行聚类得到多个第*i*+1层分类，*i*从1向后进行整数取值；

识别出目标人脸图片所属的第1层分类，并通过迭代方式继续在所述目标人脸图片所属的第*j*层分类中识别出所述目标人脸图片所属的第*j*+1层分类，*j*从1向后依序进行整数取值；

10 通过所述迭代方式直至在所述目标人脸图片所属的第*j*层分类中不存在第*j*+1层分类时，从所述目标人脸图片所属的第*j*层分类中，识别出所述目标人脸图片的相似人脸图片。

7、根据权利要求6所述的方法，其中，所述将已收集的人脸图片进行聚类得到多个第1层分类的步骤包括：

15 设置多个初始中心点，并根据所述已收集的人脸图片的特征向量与每个所述初始中心点的距离远近，将所述已收集的人脸图片分为多个第1层分类，并根据每个第1层分类的人脸图片的特征向量，计算所述每个第1层分类的向量中心点。

8、根据权利要求6-7任一项所述的方法，其中，所述将已收集的人脸图片进行聚类得到多个第1层分类的步骤还包括：

计算所述每个第1层分类的初始中心点与向量中心点之间的方差；

20 如所述方差的大小超过预设阈值，则重新设置初始中心点，并重新将所述已收集的人脸图片分为多个第1层分类，并重新计算所述每个第1层分类的向量中心点。

9、根据权利要求6-8任一项所述的方法，其中，所述识别出目标人脸图片所属的第1层分类的步骤包括：

25 选择向量中心点与所述目标人脸图片的特征向量之间距离最小的第1层分类，作为所述目标人脸图片所属的第1层分类。

10、根据权利要求6-9任一项所述的方法，其中，所述识别出所述目标人脸图片的相似人脸图片的步骤包括：

从所述第*j*层分类的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片，作为所述目标人脸图片的所述相似人脸图片。

30 11、一种人脸相似度识别系统，包括：

第一特征向量生成模块，用于根据目标人脸图片的特征，生成所述目标人脸图片的特征向量；

第二特征向量生成模块，用于根据已收集的人脸图片的特征，生成所述已收集的人脸图片的特征向量；

35 第一相似人脸图片识别模块，用于从所述已收集的人脸图片中，选择特征向量与所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片作为所述目标人脸图片

的相似人脸图片。

12、根据权利要求 11 所述的系统，其中，还包括：

第一分类模块，用于将所述已收集的人脸图片聚合为多个分类；

5 向量中心点计算模块，用于根据所述每个分类中的人脸图片的特征向量，计算所述每个分类中的人脸图片的向量中心点；

所述第一相似人脸图片识别模块用于将与所述目标人脸图片的特征向量之间距离最小的向量中心点对应分类中的人脸图片，作为所述目标人脸图片的相似人脸图片。

13、根据权利要求 11-12 任一项所述的系统，其中，还包括：

10 相似度得分计算模块，用于将所述相似人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离，转换为所述相似人脸图片与所述目标人脸图片之间的相似度得分。

14、根据权利要求 13 所述的系统，其中，所述相似度得分计算模块在 $D_x \leq D_{\min}$ 时，取 $S = S_{\max}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， $D_{\min}$ 为预设最小距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， $S_{\max}$ 为预设最大相似度得分；和/或

15 所述相似度得分计算模块在 $D_i < D_x \leq D_{i+1}$ 时，取 $S = S_i + K(D_x - D_i)$ ，其中 $K = (S_{i+1} - S_i) / (D_{i+1} - D_i)$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， D_i 为预设第一人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， D_{i+1} 为预设第二人脸图片的特征向量与所述目标人脸图片的特征向量之间的距离， S_i 为所述预设第一人脸图片与所述目标人脸图片的相似度得分， S_{i+1} 为所述预设第二人脸图片与所述目标人脸图片的相似度得分；和/或

所述相似度得分计算模块在 $D_x > D_{\max}$ 时，取 $S = S_{\min}$ ， D_x 为所述目标人脸图片的特征向量与所述相似人脸图片之间的特征向量之间的距离， $D_{\max}$ 为预设最大距离， S 为所述相似人脸图片与所述目标人脸图片的相似度得分， $S_{\min}$ 为预设最小相似度得分。

25 15、根据权利要求 11-14 中任一项所述的系统，其中，还包括：

排序模块，用于在所述相似人脸图片为多张时，根据与所述目标人脸图片的相似度高低，对多张所述相似人脸图片进行排序。

16、根据权利要求 11 所述的系统，其中，还包括：

30 第二分类模块，用于将已收集的人脸图片进行聚类得到多个第 1 层分类，并通过迭代方式继续对至少一个第 i 层分类中的人脸图片进行聚类得到多个第 $i+1$ 层分类， i 从 1 向后依序进行整数取值；

分类迭代识别模块，用于识别出目标人脸图片所属的第 1 层分类，并通过迭代方式继续在所述目标人脸图片所属的第 j 层分类中识别出所述目标人脸图片所属的第 $j+1$ 层分类， j 从 1 向后进行整数取值；

35 第二相似人脸图片识别模块，用于在所述目标人脸图片所属的第 j 层分类中不存在第 $j+1$ 层分类时，从所述目标人脸图片所属的第 j 层分类中，识别出所述目标人脸图片的

相似人脸图片。

17、根据权利要求 16 所述的系统，其中，

所述第二分类模块设置多个初始中心点，并根据所述已收集的人脸图片的特征向量与每个所述初始中心点的距离远近，将所述已收集的人脸图片分为多个第 1 层分类，

5 并根据每个第 1 层分类的人脸图片的特征向量，计算所述每个第 1 层分类的向量中心点。

18、根据权利要求 16-17 任一项所述的系统，其中，还包括：

方差计算模块，计算所述每个第 1 层分类的初始中心点与向量中心点之间的方差；

如所述方差的大小超过预设阈值，则所述分类模块重新设置初始中心点，并重新
10 将所述已收集的人脸图片分为多个第 1 层分类，并重新计算所述每个第 1 层分类的向量中心点。

19、根据权利要求 16-18 任一项所述的系统，其中，所述第二分类模块选择向量中心点与所述目标人脸图片的特征向量之间距离最小的第 1 层分类，作为所述目标人脸图片所属的第 1 层分类。

15 20、根据权利要求 16-19 任一项所述的系统，其中，

所述第二相似人脸图片识别模块从所述第 j 层分类的人脸图片中，选择特征向量与
所述目标人脸图片的特征向量之间距离最小的至少一张人脸图片，作为所述目标人脸图
片的所述相似人脸图片。

21、一种计算机程序，包括计算机可读代码，当所述计算机可读代码在计算设备上
20 运行时，导致所述计算设备执行根据权利要求 1-10 中的任一个所述的方法。

22、一种计算机可读介质，其中存储了如权利要求 21 所述的计算机程序。

1/10

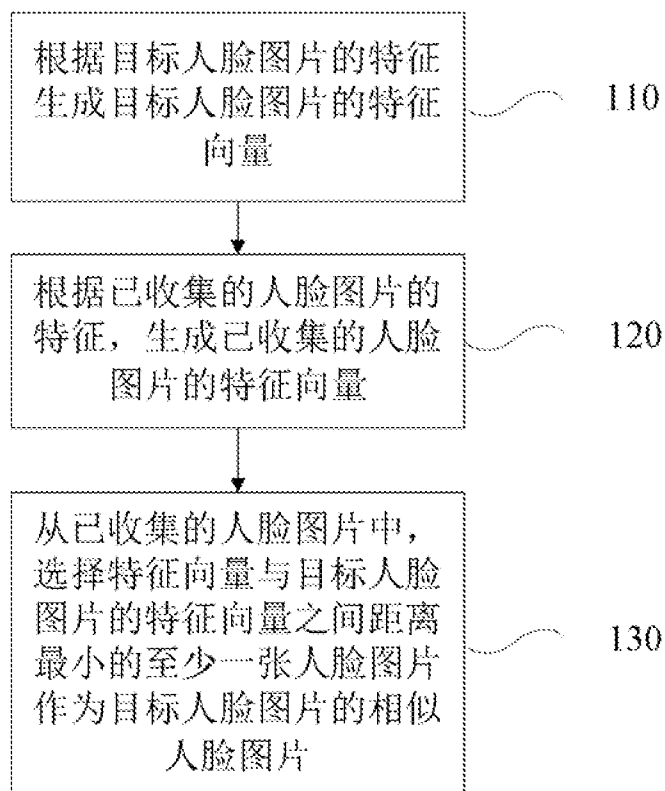


图 1

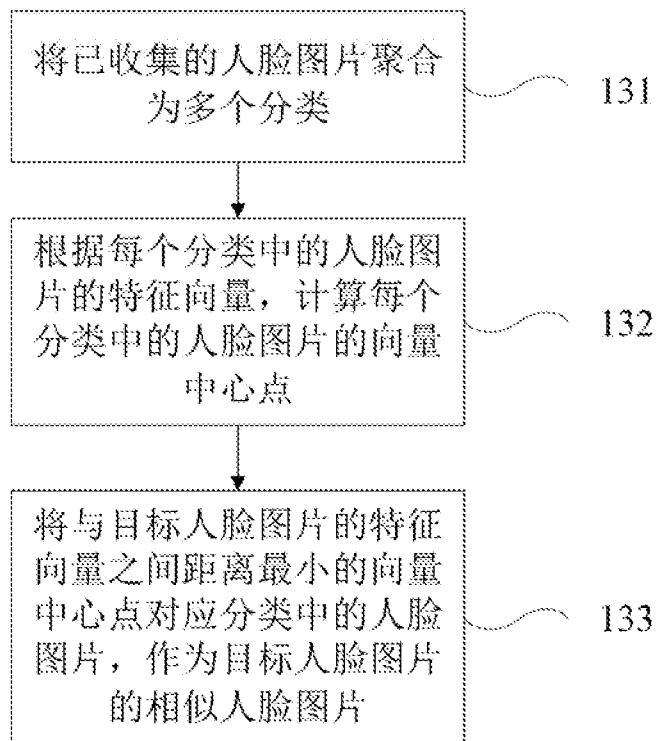


图 2

2/10

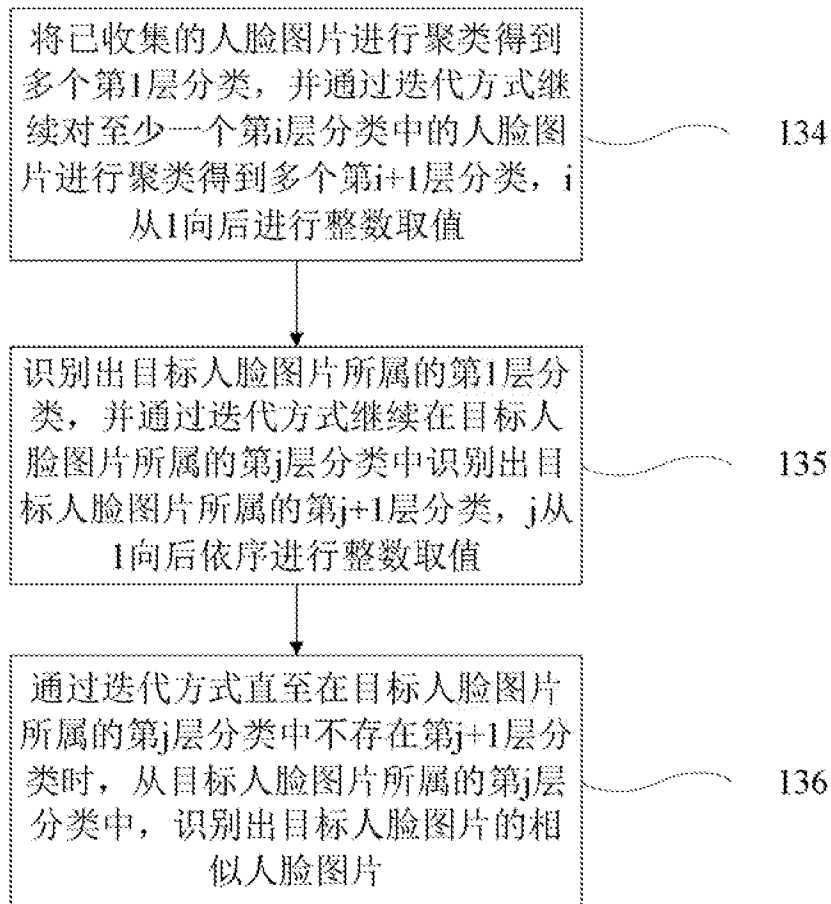


图 3

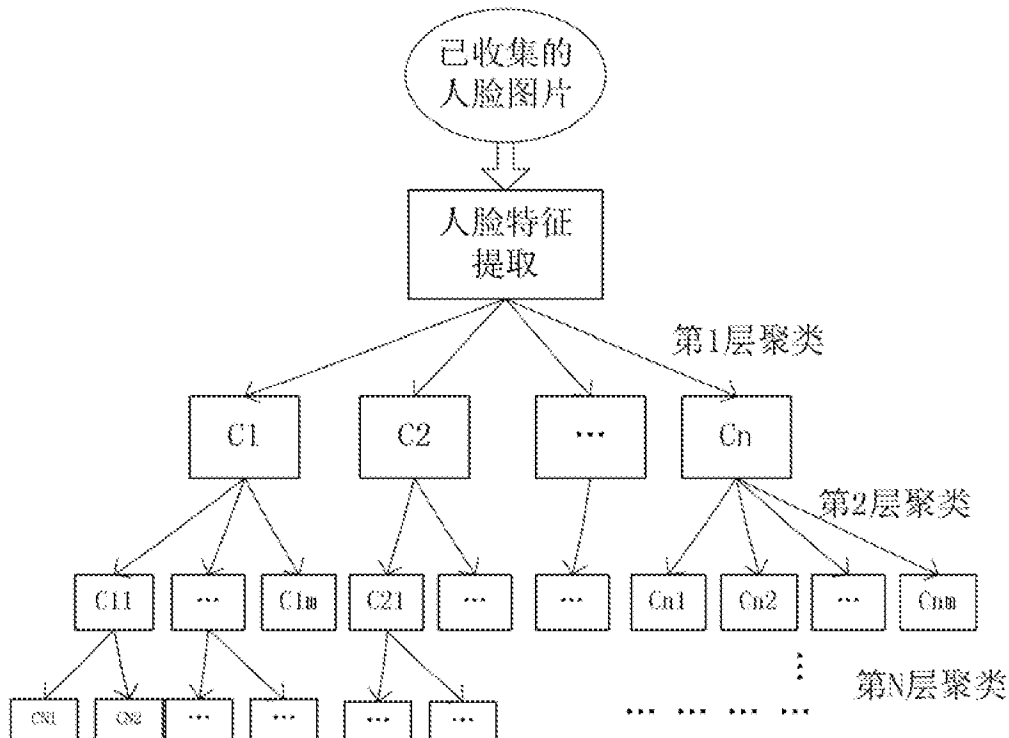


图 4

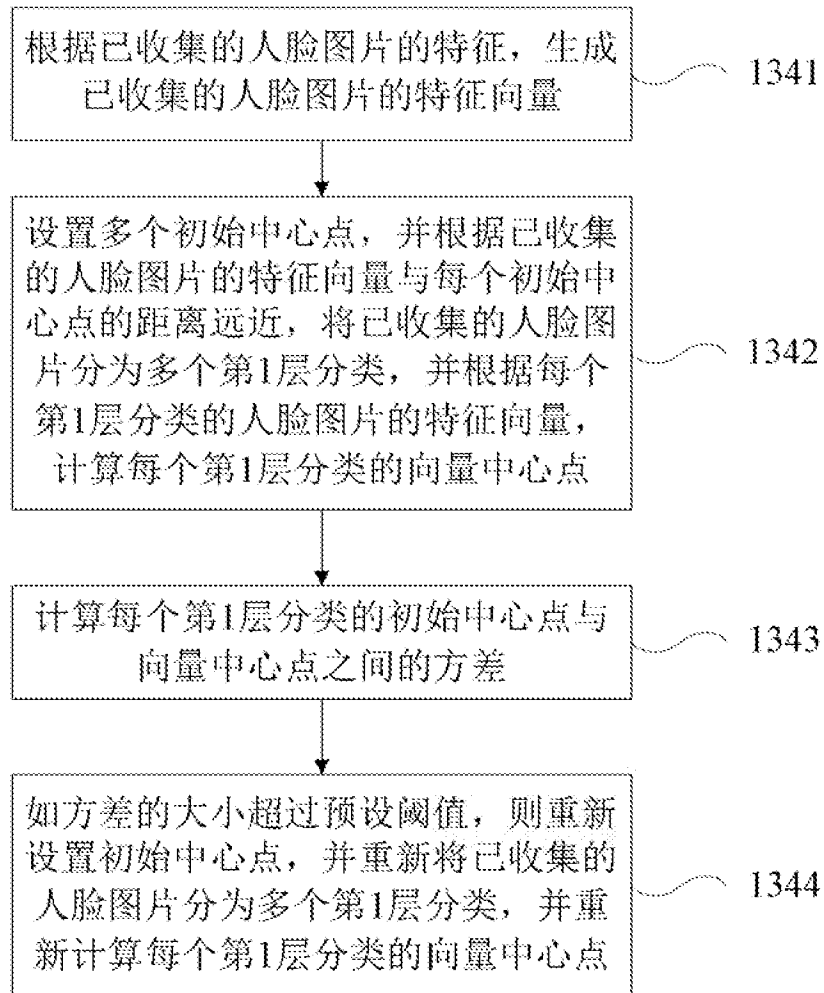


图 5

4/10

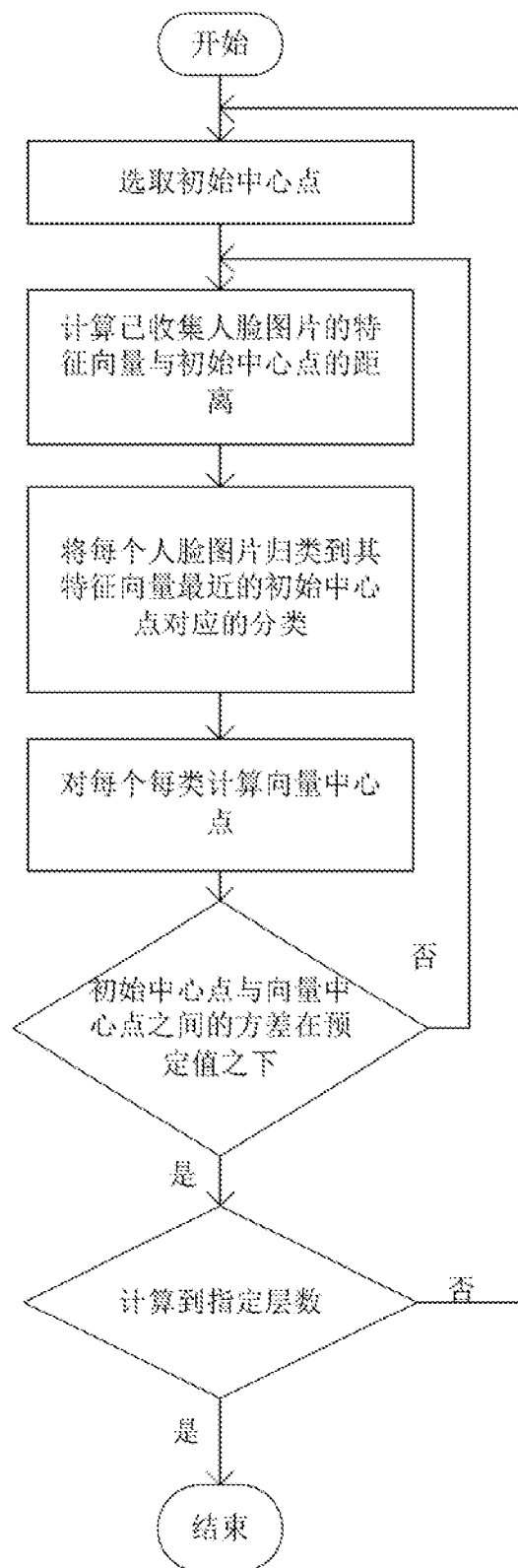


图 6

5/10

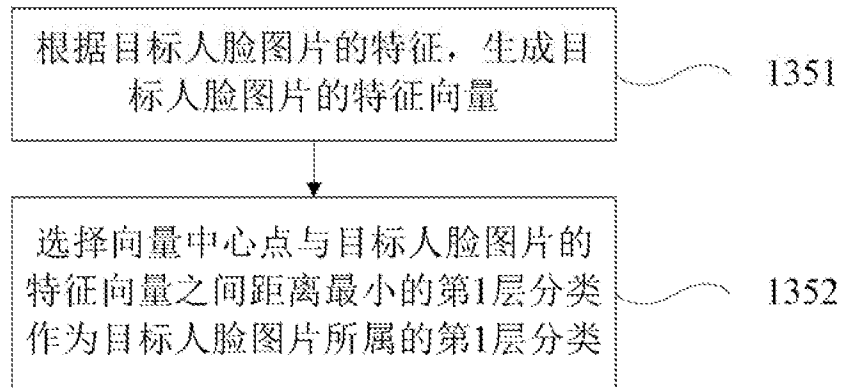


图 7

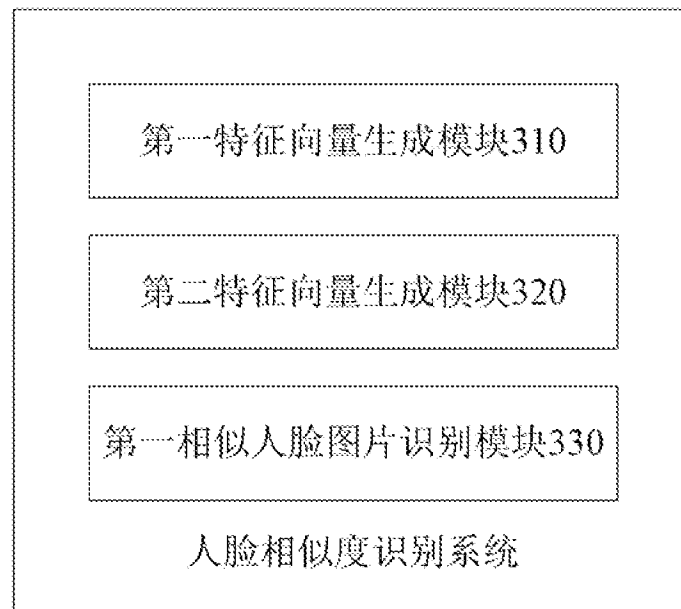


图 8

6/10

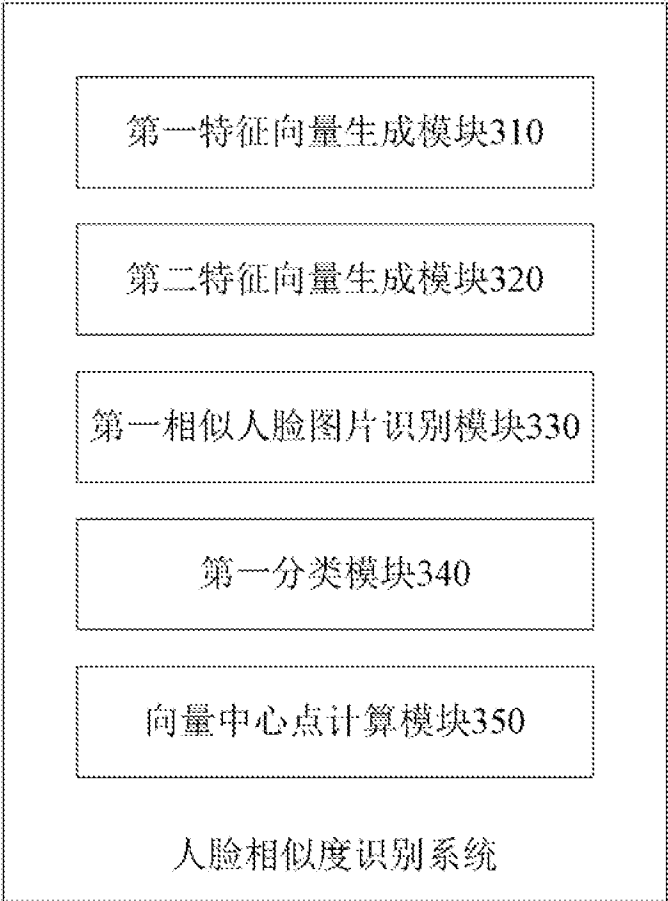


图 9

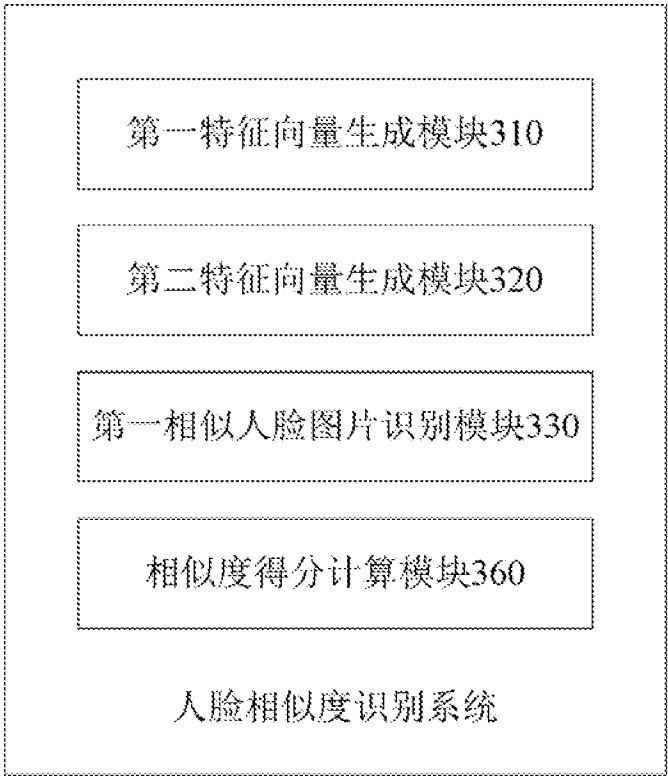


图 10

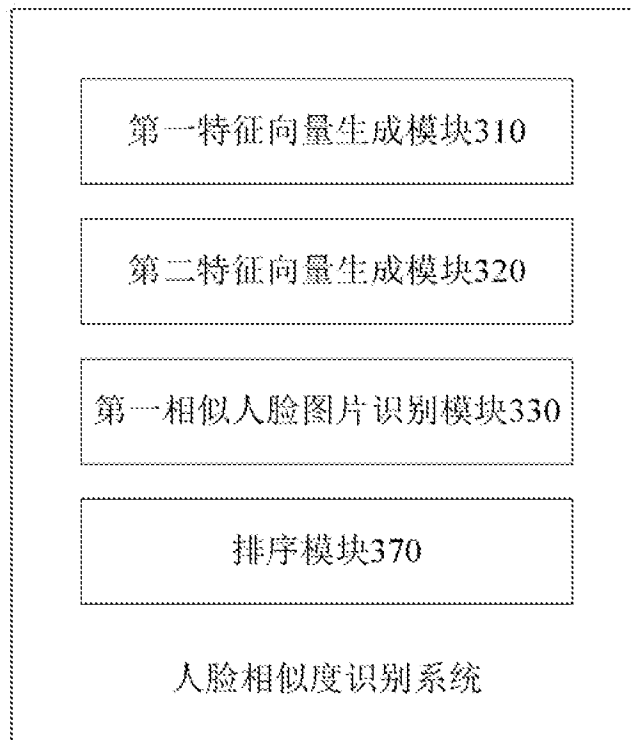


图 11

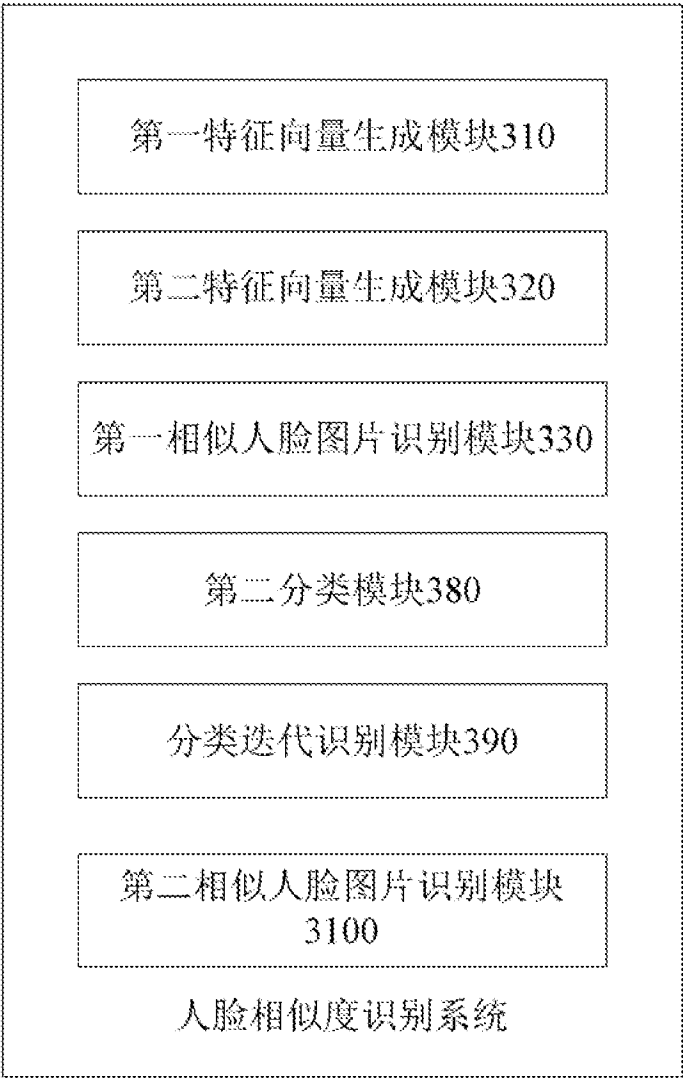


图 12

9/10

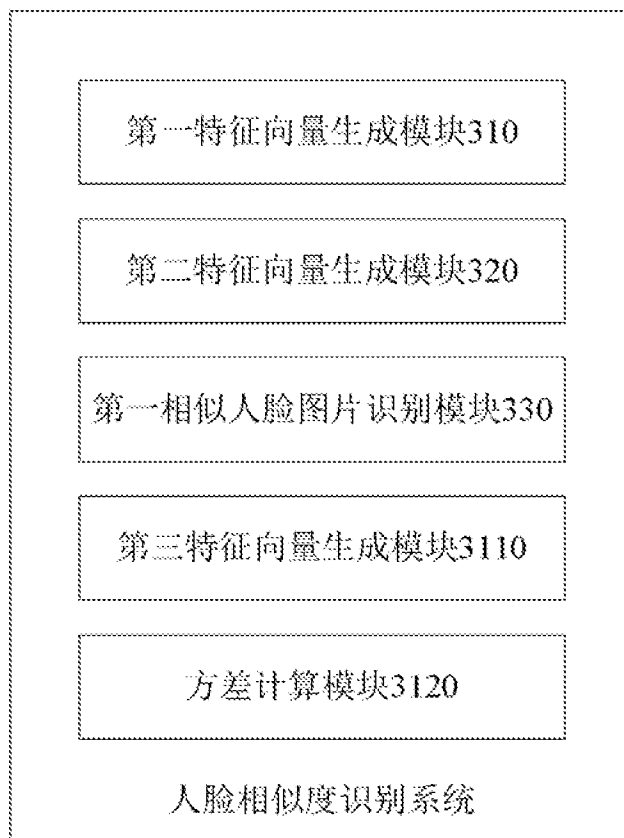


图 13



图 14

10/10

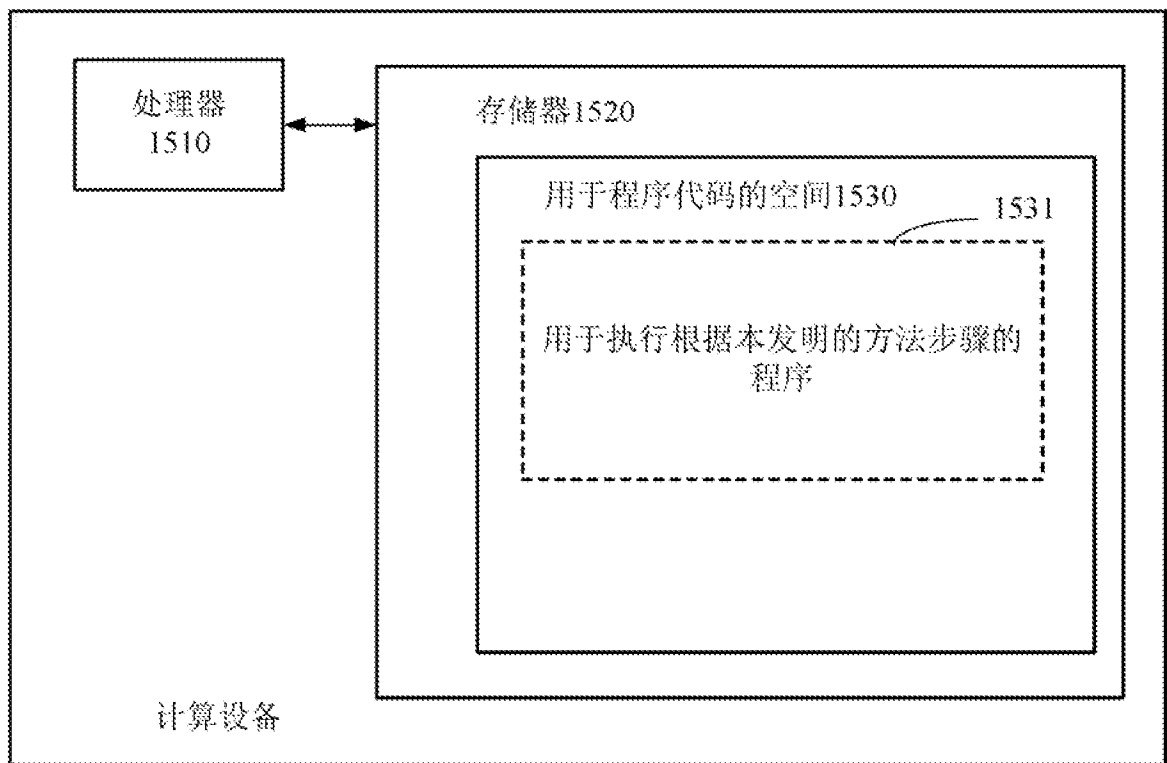


图 15

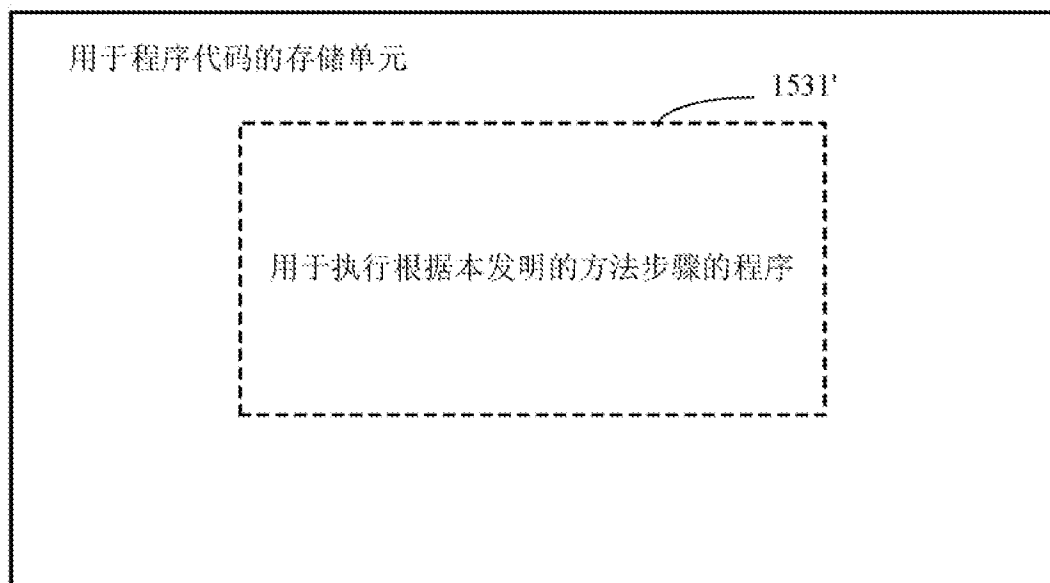


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/082550

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00 (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, G06F

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)

CNABS, CNTXT, VEN: feature vector, similarity, identify, feature, vector, face, iteration, distance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103678984 A (HUBEI MICROPATTERN TECHNOLOGY DEVELOPMENT CO., LTD.), 26 March 2014 (26.03.2014), description, paragraphs [0037]-[0069]	1-3, 5, 11-13, 15, 21, 22
A	CN 101281598 A (TSINGHUA UNIVERSITY), 08 October 2008 (08.10.2008), the whole document	1-22
PX	CN 104036261 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 10 September 2014 (10.09.2014), claims 1-10	6-10, 16-20, 21, 22
PX	CN 104036259 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 10 September 2014 (10.09.2014), claims 1-10	1-5, 11-15, 21, 22
A	CN 103678984 A (HUBEI MICROPATTERN TECHNOLOGY DEVELOPMENT CO., LTD.), 26 March 2014 (26.03.2014), the whole document	4, 6-10, 14, 16-20

☐ 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
22 September 2015 (22.09.2015)

Date of mailing of the international search report
09 October 2015 (09.10.2015)

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
LIU, Fang
Telephone No.: (86-10) **62089108**

International application No.
PCT/CN2015/082550

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2015/082550

A. 主题的分类

G06K 9/00(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, G06F

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))

CNABS, CNTXT, VEN: 相似度, 识别, 特征向量, 人脸, 迭代, 叠代, 距离, similarity, identify, feature, vector, face, iteration, distance

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
X	CN 103678984 A (湖北微模式科技发展有限公司) 2014年 3月 26日 (2014 - 03 - 26) 说明书第[0037]-[0069]段	1-3, 5, 11-13, 15, 21, 22
A	CN 101281598 A (清华大学) 2008年 10月 8日 (2008 - 10 - 08) 全文	1-22
PX	CN 104036261 A (北京奇虎科技有限公司等) 2014年 9月 10日 (2014 - 09 - 10) 权利要求1-10	6-10, 16-20, 21, 22
PX	CN 104036259 A (北京奇虎科技有限公司等) 2014年 9月 10日 (2014 - 09 - 10) 权利要求1-10	1-5, 11-15, 21, 22
A	CN 103678984 A (湖北微模式科技发展有限公司) 2014年 3月 26日 (2014 - 03 - 26) 全文	4, 6-10, 14, 16-20

☐ 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

2015年 9月 22日

国际检索报告邮寄日期

2015年 10月 9日

ISA/CN的名称和邮寄地址

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

传真号 (86-10)62019451

授权官员

刘芳

电话号码 (86-10)62089108

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2015/082550

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	103678984	A	2014年 3月 26日	无			
CN	101281598	A	2008年 10月 8日	CN	100557624	C	2009年 11月 4日
CN	104036261	A	2014年 9月 10日	无			
CN	104036259	A	2014年 9月 10日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)