(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 106529388 A
(43)申请公布日 2017.03.22

(21)申请号 201610800799.1

(22)申请日 2016.09.01

(30)优先权数据
JP2015-180116 2015.09.11 JP
JP2016-144757 2016.07.22 JP

(71)申请人 佳能株式会社
地址 日本东京都大田区下丸子3丁目30-2

(72)发明人 松下昌弘 椎山弘隆

(74)专利代理机构 北京怡丰知识产权代理有限公司 11293
代理人 迟军

(51)Int.Cl.
G06K 9/00(2006.01)
G06K 9/46(2006.01)
G06T 7/20(2017.01)

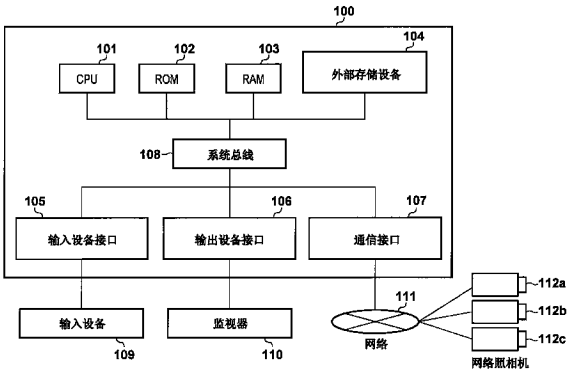
权利要求书2页 说明书11页 附图8页

(54)发明名称

信息处理装置及其控制方法

(57)摘要

本发明提供一种信息处理装置及其控制方法。本发明防止核对数据库变得庞大，并缩短从人检测至在核对数据库中的登记所需的延迟时间。该信息处理装置包括获取视频的获取单元、从获取的视频的至少一帧检测人的全身或部位的检测单元、追踪被检测的人的全身或部位的追踪单元、以及在数据库中登记从在从追踪单元开始追踪人的全身或所述部位的定时至追踪单元追踪结束之前的定时的第一时段期间追踪的人的全身或所述部位所提取的特征量的登记单元。



CN 106529388 A

1.一种信息处理装置,所述信息处理装置包括:

获取单元,其被构造为,获取视频;

检测单元,其被构造为,从所获取的视频的至少一帧检测人的全身或部位;

追踪单元,其被构造为,追踪所检测到的人的全身或所述部位;以及

登记单元,其被构造为,在数据库中登记,从在从所述追踪单元开始追踪人的全身或所述部位的定时至所述追踪单元追踪结束之前的定时的第一时段期间所追踪的人的全身或所述部位提取的特征量。

2.根据权利要求1所述的信息处理装置,其中,所述登记单元在数据库中登记,从在所述第一时段结束之后的第二时段期间由所述追踪单元所追踪的人的全身或所述部位提取的特征量,并且

所述第一时段短于所述第二时段。

3.根据权利要求2所述的信息处理装置,其中,所述登记单元在数据库中登记,从每当在所述第一时段结束之后经过所述第二时段时由所述追踪单元所追踪的人的全身或所述部位提取的特征量。

4.根据权利要求1所述的信息处理装置,所述信息处理装置还包括:

分类单元,其被构造为,将被追踪的人的全身或部位分类成多个预定组;以及

选择单元,其被构造为,针对各个组从被分类成组的人的全身或部位选择数量不大于第一预定数量的人的全身或部位,

其中,所述登记单元在数据库中登记所选择的人的全身或部位的特征量。

5.根据权利要求4所述的信息处理装置,其中,所述组是人的全身或部位的方向。

6.根据权利要求4所述的信息处理装置,其中,所述选择单元基于人的全身或部位的尺寸、模糊、闭合的眼睛和张开的嘴中的至少一者,选择数量不大于所述第一预定数量的人的全身或部位。

7.根据权利要求4所述的装置,其中,所述选择单元基于优先级从多个组选择数量不大于第二预定数量的组,并且,针对数量不大于所述第二预定数量的所选择的组中的各组,选择数量不大于所述第一预定数量的人的全身或部位,并且

预先为所述多个组设置优先级。

8.根据权利要求1所述的信息处理装置,所述信息处理装置还包括:

输入单元,其被构造为,输入查询视频;以及

核对单元,其被构造为,基于所输入的查询视频中的人的整个图像或部分图像的特征量,进行与在数据库中登记的特征量的核对。

9.根据权利要求1所述的信息处理装置,其中,人的所述部位是人的脸部和人的头部中的一者。

10.根据权利要求8所述的信息处理装置,其中,人的所述部位是人的脸部,

所述登记单元在数据库中登记与脸部器官相关联的特征量,并且

所述核对单元核对从所述查询视频提取的关于所述脸部器官的特征量与在数据库中登记的关于所述脸部器官的特征量。

11.一种信息处理装置的控制方法,所述控制方法包括以下步骤:

获取视频;

从所获取的视频的至少一帧检测人的全身或部位；

追踪所检测到的人的全身或所述部位；以及

在数据库中登记，从在从开始追踪人的全身或所述部位的定时至追踪结束之前的定时的第一时段期间所追踪的人的全身或所述部位提取的特征量。

信息处理装置及其控制方法

技术领域

[0001] 本发明涉及一种用于核对图像中的人的信息处理装置及其控制方法。

背景技术

[0002] 传统上已知有这样的装置，其从监视器视频的各帧检测人的全身或部位(part)(例如，脸部)，从该脸部计算图像特征量，并且，将该图像特征量与视频的帧相关联地存储。例如，在日本特开2013-153304号公报(以下被称为文献1)中公开的装置可以用存储的图像特征量核对作为查询的失踪的孩子的脸部，并且显示包含该孩子的视频。

[0003] 然而，在文献1的技术中，从视频的所有的帧检测到的所有的脸部都被登记在数据库中，并且，存储极大数量的图像特征量。所述帧可以被稀疏化(thin out)，以降低处理的帧速率并减少要登记的脸部的图像特征量的数目。在这种情况下，具有优选的方向、尺寸、眼睛拍摄条件等的脸部的特征量可能未被登记，导致核对失败。

发明内容

[0004] 鉴于上述问题而作出了本发明。本发明提供一种用于减少在核对数据库中登记从人的全身或部位提取的特征量为止所需的延迟时间。

[0005] 根据本发明的一个方面，提供一种信息处理装置，所述信息处理装置包括：获取单元，其被构造为，获取视频；检测单元，其被构造为，从所获取的视频的至少一帧检测人的全身或部位；追踪单元，其被构造为，追踪所检测到的人的全身或所述部位；以及登记单元，其被构造为，在数据库中登记，从在从所述追踪单元开始追踪人的全身或所述部位的定时至所述追踪单元追踪结束之前的定时的第一时段期间所追踪的人的全身或所述部位提取的特征量。

[0006] 根据本发明的另一个方面，提供一种信息处理装置的控制方法，所述控制方法包括以下步骤：获取视频；从所获取的视频的至少一帧检测人的全身或部位；追踪所检测到的人的全身或所述部位；以及在数据库中登记，从在从开始追踪人的全身或所述部位的定时至追踪结束之前的定时的第一时段期间所追踪的人的全身或所述部位提取的特征量。

[0007] 根据本发明，可以减少从检测人至在核对数据库中登记从人的全身或部位提取的特征量所需的延迟时间。

[0008] 根据下面(参照附图)对示例性实施例的描述，本发明的另外的特征将变得清楚。

附图说明

[0009] 图1是根据实施例的装置的框图；

[0010] 图2是根据实施例的装置的功能框图；

[0011] 图3A和图3B是用于解释根据实施例的基于脸部方向的分类的视图；

[0012] 图4是示出根据实施例的对象追踪信息的内容的视图；

[0013] 图5是用于解释根据实施例的图像特征存储单元的处理的视图；

[0014] 图6是示出根据实施例的开始和结束对象追踪的定时的示例的视图；

[0015] 图7是示出根据实施例的核对数据库的结构的示例的视图；

[0016] 图8是示出根据第一实施例的视频存储处理的流程图；

[0017] 图9是示出图8中的DB登记处理的流程图；以及

[0018] 图10是示出根据第一实施例的核对处理的流程图。

具体实施方式

[0019] 现在将参照附图描述本发明的实施例。请注意，下述实施例是本发明的具体实现方式的示例或者所附权利要求中所述的布置的具体示例。

[0020] [第一实施例]

[0021] 图1是根据本实施例的图像核对装置100的框图。装置100将被描述为既用作用于图像搜索的服务器装置又用作客户端装置的装置。然而，服务器功能和客户端功能可由独立的装置形成。服务器装置不必由单个装置实现，并且，处理可以被分布给多个计算机装置。在服务器装置由多个计算机装置构成时，它们通过局域网(LAN)连接，以相互通信。各个计算机装置可由诸如个人计算机(PC)或工作站(WS)等的信息处理装置实现。

[0022] 参照图1，CPU 101是控制整个图像核对装置100的中央处理单元。ROM 102是存储不需要改变的程序(例如，BIOS)或参数的只读存储器。RAM 103是临时地存储从外部装置等供给的程序或数据的随机存取存储器。外部存储设备104是永久地置于计算机装置100中的诸如硬盘等的大容量存储设备。外部存储设备104存储装置的操作系统(OS)、与图像核对相关联的应用和在核对时要参考的核对数据库(以下被称为核对DB)。如稍后将描述的，由摄像机拍摄的视频数据也被存储在外部存储设备104中。

[0023] 输入设备接口105是用于接收用户操作和输入数据的诸如指向设备或键盘等的输入设备109的接口。输出设备接口106是被构造为显示由计算机装置100保持的显示数据或供给的数据的监视器110的接口。通信接口107是被构造为将该装置连接到诸如因特网等的网络111的通信接口。网络照相机112a至网络照相机112c与网络111连接。请注意，给网络照相机112a至网络照相机112c中的各个分配照相机ID。这里假设分别给网络照相机112a、网络照相机112b和网络照相机112c分配照相机ID 1、2和3来进行描述。至少一个网络照相机就足够了，并且可以连接任意数量的网络照相机。系统总线108是可通信地连接上述单元101至107的传输线。

[0024] 在上述布置中，当该装置被通电时，CPU 101执行在ROM 102中存储的BIOS程序，从外部存储设备104加载OS(操作系统)到RAM 103，并且，执行OS。结果是，该装置用作信息处理装置。在OS的控制下，从外部存储设备104执行与图像核对相关联的客户端程序和服务器程序。该装置由此用作图像核对装置。

[0025] 图2是根据实施例的视频核对装置的功能框图。图2是在CPU 101执行搜索服务器和客户端的两种功能的程序的情况中的功能框图。

[0026] 在本实施例中，将描述这样的示例，其中，从脸部获得的脸部特征用作图像特征，并且，使用该图像特征进行图像特征核对。

[0027] 控制处理单元211进行用于控制由附图标记201至207、209和210表示的处理单元的主要处理。

[0028] 视频输入单元201经由通信接口107从网络照相机112a至网络照相机112c接收视频帧，并且将它们存储在外部存储设备104中。输入帧速率被假设是，例如，30帧/秒。结果是，在外部存储设备104中创建网络照相机112a至网络照相机112c的运动图像文件。视频输入单元201也将接收到的视频帧供给到脸部检测单元206。

[0029] 脸部检测单元206使用脸部方向检测单元207和脸部特征提取单元209进行脸部检测。脸部特征提取单元209检测视频帧中的脸部区域，提取特征量，并且将其返回到脸部检测单元206。在本实施例中，作为脸部特征量，获得脸部器官点的SIFT(Scale Invariant Feature Transform，尺度不变特征变换)特征量。在提取脸部特征量时，脸部特征提取单元209也将脸部区域的外接矩形的左上角的坐标和该外接矩形的尺寸(宽度W和高度H)返回到脸部检测单元206。请注意，可以替代尺寸返回外接矩形的右下角的坐标。

[0030] 脸部方向检测单元207检测由脸部检测单元206检测到的脸部的方向，并且，将该方向分类到预设方向图案中的一个。脸部方向检测单元207将脸部方向图案编号作为分类结果返回到脸部检测单元206。请注意，检测图像中的人脸的方向的技术是已知的，将省略其详细描述。请注意，关于脸部方向检测的文献是：

[0031] Erik Murphy-Chutorian，"Head pose estimation for driver assistance systems:A robust algorithm and experimental evaluation"，in Proc.IEEE Conf.Intelligent Transportation Systems，2007，pp.709-714。

[0032] 根据本实施例的脸部方向检测单元207确定图3A所示的5×5图案中的哪个对应于关注的脸部的方向。请注意，尽管在图3A中示出5×5图案，然而，图案的数量并不受特别的限制。如果各个网络照相机具有足够高的分辨率，则特征量或方向的精确度变高。因此，可以增加图案的数量。在所示的图案中，向前看的脸部的图案位于中心。垂直和水平方向上的偏差角越大，离中心的距离就越长。换句话说，脸部方向图案被布置为使得脸部向前看，因为其位置变得靠近5×5图案的中心。在脸部核对中，在使用向前看的脸部时，获得高的精确度。也就是说，对于在5×5图案的中心的脸部方向，优先级最高，并且，随着离中心的距离增大，优先级下降。图3B示出脸部方向图案编号的示例。如图3B所示，向前看的脸部的脸部方向图案编号是“1”。在其周围分配脸部方向图案编号2至9，并且，在这些编号的外面分配脸部方向图案编号10至25。请注意，脸部方向图案编号2至9不必总是如图所示的那样。这也适用于脸部方向图案编号10至25。

[0033] 如上所述，脸部检测单元206使用脸部方向检测单元207和脸部特征提取单元209从由网络照相机112a至网络照相机112c接收到的各个视频帧获得人脸的位置、尺寸、特征量和脸部方向图案编号。

[0034] 在从控制处理单元211接收追踪开始指令信息时，追踪处理单元205设置当前帧中的指定脸部作为追踪开始目标脸部，并且在后续的帧图像中进行对应脸部的追踪处理。追踪开始指令信息包括表示作为帧图像的源的网络照相机的信息(照相机ID)、表示追踪目标脸部的初始位置的信息、以及用来指定要追踪的脸部的对象ID。尽管细节从稍后的描述将变得清楚，但追踪开始指令信息还包括追踪计数和表示结束帧的定时的信息(在本实施例中，Tp1和Tp2中的一个)。对象ID的初始值是1。每当找到要追踪的脸部，该值就被递增“1”。请注意，使用日本特开2002-373332号公报中公开的技术来进行追踪图像中的人的处理。

[0035] 在接收追踪开始指令信息时，追踪处理单元205在RAM 103上生成如图4所示的对

象追踪信息。一个对象追踪信息由照相机ID和对象ID指定。针对一个对象追踪信息设置追踪计数和最后一帧编号的数据。稍后将对这些数据进行详细的描述，然而，这里将省略其详细描述。

[0036] 在对象追踪信息中存储的信息包括在开始一天的摄像记录之后指示装置开始追踪时的帧编号(或时间)。另外，包括对应视频帧中的追踪目标脸部的外接矩形的左上角的位置、以及尺寸(宽度和高度)、脸部方向图案编号和脸部的特征量。作为这些信息，使用来自脸部方向检测单元207和脸部特征提取单元209的信息。

[0037] 只要视频帧被顺序地输入，并且，正继续追踪对应的脸部，追踪处理单元205就将上述的帧编号、位置、尺寸、脸部方向图案编号和特征量添加至对应对象追踪信息。如果脸部变成不可追踪，则追踪处理单元205请求控制处理单元211在核对DB中登记对应的对象ID。请注意，在由于处于追踪的人(的脸部)移到照相机的视场之外而无法识别脸部或者人向后看的情况下，脸部变成不可追踪。

[0038] 在接收到登记请求时，控制处理单元211将对应的对象追踪信息传输到图像特征存储单元202并使其在核对DB中进行存储处理(登记处理)。稍后将描述图像特征存储单元202的细节。当登记处理结束时，控制处理单元211从RAM 103擦除登记的对象追踪信息。

[0039] 请注意，移到照相机的视场之外的人可以再次移入照相机的视场中。本实施例的装置不能确定移到照相机的视场之外的人和再次移入照相机的视场中的人是否为同一人。由于这种原因，对于再次移入照相机的视场中的人的脸部，分发新的对象ID，并且，追踪开始。因此，追踪处理单元205最终对照相机的摄像场中的被识别为脸部的所有的对象进行追踪处理。因此，生成与实际上正被追踪的脸部的数量一样多的图4所示的多个对象追踪信息。在一个网络照相机的视场中至多存在约40至约50个脸部区域，各个脸部区域具有可识别为脸部的尺寸。即使当考虑存在三个网络照相机时，追踪对象脸部的数量也最多为150。当前计算机可以充分地应对该数量的脸部。

[0040] 假设从网络照相机接收到视频帧，并且，在图6中的时刻t0第一次检测到人A的脸部。在这种情况下，为人A分发对象ID，并且，生成人A的对象追踪信息。对人A进行追踪处理，直到该人A变成不可追踪的时刻t3。在这段时间，如果在时刻t1第一次在同一摄像场中检测到另一个人B的脸部，则为该人B分发对象ID，并且，生成该人B的对象追踪信息。对人B进行追踪处理，直到该人B变成不可追踪的时刻t2。

[0041] 接下来，将参照图5描述图像特征存储单元202的存储处理。如已描述的，在对某人脸部的追踪处理结束时进行存储处理。

[0042] 图5中的附图标记500表示在对象追踪信息中存储的脸部方向图案编号的时间转变。请注意，为了描述方便，图5示出在对象追踪信息中存储共八个脸部方向图案的示例。

[0043] 首先，图像特征存储单元202对在对象追踪信息中存储的各个脸部方向图案编号的出现计数进行总计处理。图5中的附图标记501表示总计结果。各个脸部图案的右下侧的编号表示出现计数(累积数)。

[0044] 如图3B所示，脸部方向图案编号是1至25。因此，脸部方向ID的总计结果被表达为变量F(脸部方向图案编号)。在本实施例中，与追踪对象脸部对应地在核对DB中最多登记四个特征量。当然，该数量仅仅是一个示例，并且，可以登记更多的特征量。

[0045] 在本实施例中，按编号的顺序设置脸部方向图案编号的优先级(图3B)。图像特征

存储单元202按此顺序布置总计变量F(1)至F(25)，并且找到出现计数为非零(1或更大)的前四个图案。在图5中，如附图标记502所示，阴影图案F(1)、F(2)、F(3)和F(6)对应于这些图案。也就是说，决定对象追踪信息中的脸部方向图案编号＝1，2，3和6的特征量作为核对DB的登记候选。

[0046] 这里，脸部方向图案编号“1”的出现计数F(1)是“1”。也就是说，在对象追踪信息中只存在一个脸部方向图案编号“1”的特征量。由于这种原因，图像特征存储单元202在核对DB中登记对象追踪信息中的脸部方向图案编号“1”的特征量。

[0047] 由于脸部方向图案编号“2”的出现计数F(2)是“2”，所以决定要登记所述特征量中的一个。在这种实施方式下，登记两个脸部中的具有较大尺寸的一个脸部的特征量，作为表示对应方向的特征量。如果脸部尺寸大，则在拍摄人的脸部时离照相机的距离短，并且，可以预期高的精确度。这也适用于脸部方向图案编号＝3和6。

[0048] 结果是，在本实施例中，在核对DB中与一个对象ID对应地登记四个脸部特征量。另外，如果脸部方向不重复，并且，脸部向前看的程度高，则可以使登记脸部特征量的可能性高。因此，可以防止核对DB变得庞大。另外，随着脸部尺寸变大，要登记的特征量的优先级也变高，并且，能够预期达到精确的匹配处理。

[0049] 仍存在一个问题。此问题是由于在核对DB中登记脸部特征量的定时被设置为脸部(人)变成不可追踪的定时而发生。例如，考虑当前时刻是图6中的时刻t4，并且，作为查询图像输入作为失踪的孩子的人A的脸部的情况。在这种情况下，网络照相机中的任意一个在当前时刻(时刻t4)正拍摄人A的脸部。然而，该脸部仍处于追踪中，并且，人A的脸部的特征量仍未被登记在核对DB中。也就是说，即使作为查询图像输入人A的脸部的特征量，并且，在时刻t4开始核对处理，在从时刻t0至t4的最新时段内的人A的脸部也不能是搜索目标。当搜索例如失踪的孩子等的人时，指定最靠近最新时刻的位置是重要的。因此，不能忽视此问题。

[0050] 考虑到上述问题，作为本实施例的特性特征，不仅脸部变成不可追踪的定时、还有时限，都被设置为在核对DB中登记脸部特征量的定时。例如，在图6中，在从开始追踪人A的时刻t0起经过预设时间Tp1之后的时刻t5，即使仍可以检测人A的脸部，也在核对DB中进行对人A的脸部的特征量的登记处理。结果是，由于经由追踪处理在早期阶段在核对DB中开始对人A的脸部的特征量的登记处理，所以当在时刻t4开始核对处理时可以将人A设置为搜索目标。时刻t4是一个示例。如果在时刻t5之后在核对DB中登记人A的脸部的特征量，则可以将人A设置为搜索目标。

[0051] 在本实施例中，要在核对DB中登记的一个对象的脸部特征量的数量是4。然而，在决定要登记的特征量之前的候选的数量优选地更大。这是因为对各个脸部方向上的出现脸部的最大尺寸进行更新的可能性变高。因此，当在第一次的时间Tp1进行了核对DB中的登记之后，在第二次或后续次开始核对DB中的登记处理之前的时间优选地长于时间Tp1。将参照图6对此进行描述。在从对该人A的追踪开始时刻t0起经过了时间Tp1之后的时刻t5，开始核对DB中的人A的第一登记处理。令Tp2是对同一人A进行核对DB中的第二次或后续次的登记处理之前的时间。在本实施例中，保持由Tp2≥Tp1给出的关系。更具体地，Tp1是约10秒，并且，Tp2是约20秒。这些值可以被适当地设置，或者，可以设置为同一值。因为需要缩短第一次在DB中登记对象的脸部特征量为止的时间，以提早将该人设置为搜索目标，所以使Tp1更短以保持Tp2≥Tp1。

[0052] 请注意，各个对象ID必须是唯一的。在本实施例中，即使同一人A正被追踪，从时刻t0至t5(第一次追踪)，从时刻t5至t6(第二次连续追踪)和从时刻t6至t7(第三次连续追踪)，也分配不同的对象ID。

[0053] 从追踪开始时刻起经过时间Tp1或Tp1+Tp2×n(n＝1，2...)之后的时刻在对象之间变化。需要可以识别追踪处理是第一次追踪还是第二次或后续次追踪。

[0054] 由于上述原因，根据实施例的追踪处理单元205设置“1”给图4所示的对象追踪信息中的由附图标记402表示的追踪计数，以指示它是人检测之后的第一次追踪。此时刻可以在新创建对象追踪信息时由第一帧编号指定。例如，在图4的示例中，当创建照相机ID＝1和对象ID＝00001的对象追踪信息时的第一帧编号是“12345”。假设网络照相机112a至网络照相机112c中的各个以30帧/秒的帧速率拍摄图像。如果Tp1＝10秒，则在此时段内拍摄300个帧。也就是说，核对DB中的登记的时刻t5是“12645”(＝12345+300)。因此，当新创建对象追踪信息时，追踪处理单元205设置“12645”给由附图标记401表示的结束帧编号。上述处理由追踪处理单元205根据由控制处理单元211设置的追踪开始指令信息进行。

[0055] 如上所述，追踪处理单元205追踪目标帧，并且，将包括帧编号、位置和尺寸、脸部方向图案编号和特征量的记录添加至对象追踪信息。如果正追踪的脸部由于该人移到视场之外或向后看而变成不可追踪，则追踪处理单元205将包括作为引数(argument)的对应对象ID的对核对DB的登记请求发送(发出)到控制处理单元211。

[0056] 如果在追踪处理期间当前帧编号到达了结束帧编号401，则追踪处理单元205将包括对应对象ID和表示当前帧编号已经到达结束帧编号的信息的登记请求发送到控制处理单元211。

[0057] 在后一种情况下，也就是说，如果当前帧编号到达了结束帧编号401，则控制处理单元211可以检测正在进行连续追踪，因此，向追踪处理单元205通知包括新对象ID和用来定义结束帧的定时的“Tp2”(由于连续追踪)的追踪开始指令信息。结果是，追踪处理单元205创建用于连续追踪的对象追踪信息并继续追踪对应的脸部。

[0058] 图7示出根据本实施例的核对DB的结构的示例。在外部存储设备104中分配核对DB。如图7所示，核对DB的一个记录，由存储对象ID、特征量(在本实施例中，最多四个特征量)、照相机ID、检测到登记的脸部特征量的脸部的时段的第一帧编号(时刻)、坐标和脸部尺寸的字段而形成。在这些数据当中，特征量用作搜索中的密钥。

[0059] 基于上述解释，接下来将参照图8的流程图描述根据本实施例的视频记录的主要处理。本处理由控制处理单元211进行。

[0060] 在步骤S81中，控制处理单元211经由通信接口107从网络照相机112a至112c接收视频帧，并且，将它们作为核对目标的运动图像文件而记录在外部存储设备104中。通过添加例如照相机ID来创建此时创建的三个运动图像文件的文件名。

[0061] 接下来，控制处理单元211将当前视频帧传输到脸部检测单元206。在步骤S82中，控制处理单元211使脸部检测单元206执行脸部检测处理。结果是，可以获得在视频帧中存在的各个脸部的检测结果。在步骤S83中，控制处理单元211基于当前视频帧的检测结果和紧邻之前的视频帧的检测结果来确定在当前视频帧中是否存在未追踪的脸部。如果存在未追踪的脸部，则处理前进到步骤S84。在步骤S84中，控制处理单元211将包括存在要新追踪的脸部的位置、脸部的尺寸、新对象ID、照相机ID、帧编号、追踪计数“1”和用来定义结束帧

作为引数的时间“Tp1”的追踪开始指令信息，发送到追踪处理单元205。结果是，追踪处理单元205创建如图4所示的对象追踪信息，并对脸部进行追踪处理。

[0062] 另一方面，在步骤S85中，控制处理单元211确定是否从追踪处理单元205接收到登记请求。如果接收到登记请求，则控制处理单元211使处理前进到步骤S86，将登记请求的对象追踪信息传输到图像特征存储单元202，并且使其在核对DB中进行登记处理。在步骤S87中，控制处理单元211确定在步骤S86中接收到的登记请求是否包括表示当前帧编号已经到达结束帧编号401的信息，换句话说，确定是否由于该人不可追踪而发出登记请求。在确定由于该人不可追踪而发出登记请求时，控制处理单元211在步骤S89中删除在核对DB中登记的对象追踪信息并进行从步骤S81起的处理。

[0063] 如果由于当前帧编号到达了结束帧编号401而发出登记请求，则在步骤S88中，控制处理单元211将包括新的对象ID和Tp2的追踪开始指令信息发送到追踪处理单元205。结果是，追踪处理单元205进行连续的追踪处理。下一个最长的追踪时间是由Tp2设置的定时。然后，控制处理单元211进行步骤S89的处理。

[0064] 接下来将参照图9的流程图描述步骤S86的DB登记处理(图像特征存储单元202的处理)。

[0065] 在步骤S91中，图像特征存储单元202从控制处理单元211接收对象追踪信息。图像特征存储单元202分析接收到的对象追踪信息，并且，进行各个脸部方向图案编号的总计处理。也就是说，图像特征存储单元202获得脸部方向图案编号1至25的出现计数F(1)至F(25)。

[0066] 在步骤S93中，图像特征存储单元202将变量i初始化为1。图像特征存储单元202使处理前进到步骤S94，并且，确定变量i是否为25或更小，即，是否完成对所有出现计数的检查。如果变量i是25或更小(检查是不完整的)，则图像特征存储单元202在步骤S93中将变量i初始化为1并将变量j初始化为0。变量i是用来指定出现计数F()的变量，并且，变量j是用来计数要登记的特征量的数量的变量。

[0067] 在步骤S94中，图像特征存储单元202将变量i与25进行比较。如果变量i是25或更小，则未检查所有的出现计数F()。因此，图像特征存储单元202使处理前进到步骤S95，以将变量j与在核对DB中要登记的特征量的预设的上限数N(在本实施例中，N＝4)进行比较。如果j＜N，则要在核对DB中登记的特征量的数量尚未达到上限数N。因此，图像特征存储单元202使处理前进到步骤S96。在步骤S96中，图像特征存储单元202确定脸部方向图案编号i的出现计数F(i)是否为非零(1或更大)。如果出现计数F(i)是非零，则在步骤S97中，图像特征存储单元202决定脸部方向图案编号＝i中的最大脸部尺寸的特征量作为用于核对DB的登记目标。由于决定了一个登记目标，所以图像特征存储单元202在步骤S98中进行使变量j递增1的处理。然后，图像特征存储单元202在步骤S100中进行使变量i递增1的处理，并且，使处理返回到步骤S94。

[0068] 如果步骤S94和S95的确定结果之一是“否”，则图像特征存储单元202使处理前进到步骤S99。在步骤S99中，图像特征存储单元202将被决定作为登记目标的特征量与对象ID、帧编号、照相机ID、位置和尺寸信息一起登记在核对DB中，并结束处理。请注意，在变量j到达N之前，变量i可以是25。在这种情况下，要登记的特征量的数量小于4。

[0069] 上面描述了根据本实施例的视频数据和特征量登记处理。接下来将描述根据本实

施例的核对处理。核对处理由查询视频输入单元203、脸部检测单元206、脸部特征核对单元210、核对结果显示单元204和控制这些单元的控制处理单元211实现。下面将参照图10的流程图描述核对处理。

[0070] 在步骤S601中，控制处理单元211指示查询视频输入单元203输入查询视频。查询视频的输入源并不受特别的限制。例如，由搜索客户端保持的便携式终端可以与该装置连接，并且，拍摄要搜索的人的脸部的图像可作为查询视频被传输到该装置。请注意，如果该装置与因特网连接，则查询视频可以作为邮件被接收。无论如何，当输入查询视频时，进行指定查询人的脸部的处理。例如，脸部方向检测单元207检测在查询视频中存在的所有的脸部，将矩形框架添加至检测到的脸部，并且在显示设备上显示脸部，以使搜索客户端选择它们中的一个作为搜索目标脸部图像(查询脸部图像)。请注意，如果只有一个脸部被检测到，则该脸部可以被处理为无条件选择的脸部。当决定查询脸部图像时，将在检测到该查询脸部图像时提取的特征量决定为检测密钥。

[0071] 在步骤S602中，控制处理单元211获得在核对DB中登记的对象ID的总数M(记录数)。控制处理单元211请求脸部特征核对单元210使用获取的数M和查询脸部图像的特征量作为引数来进行核对处理。

[0072] 在步骤S603中，图像特征核对单元210将变量i初始化为0。在步骤S604中，确定变量i是否等于或小于M。如果i＜M，则对所有记录的核对是不完整的。因此，处理前进到步骤S606。在步骤S606中，脸部特征核对单元210读出核对DB的第i记录，并且，读出在该记录中存储的Ni特征量。在本实施例中，Ni最大为4，如已经描述的。计算查询脸部图像的特征量与Ni特征量中的各个之间的距离，并且，决定最短距离作为查询脸部图像和第i记录的对象ID的人脸之间的距离。脸部特征核对单元210将这对对象ID和决定距离存储在RAM 103的预定区域中。此后，为了用下一个记录进行核对处理，变量i被递增“1”，并且，处理返回到步骤S604。

[0073] 重复上述处理。在与所有记录的比较处理结束时(在确定变量i＝M时)，在RAM 103中存储与所有的登记的对象ID的距离。在步骤S605中，脸部特征核对单元210按升序排列在RAM 103中存储的距离，并结束处理。

[0074] 作为此后的处理，控制处理单元211将核对结果传输到核对结果显示单元204，从而进行核对结果显示处理。关于如何显示，例如，显示按距离的升序布置的预定数量的脸部图像的列表。当选择脸部图像中的一个时，基于帧编号和对应于脸部的照相机ID，进行例如从对应的位置再现运动图像的处理。

[0075] 请注意，在上述实施例中，如果在一个对象追踪信息中存在多个相同的脸部方向图案编号，则决定最大脸部尺寸的特征量作为登记目标。然而，可以在考虑除了尺寸以外的条件或者作为尺寸的补充条件作为特征量的情况下决定登记目标。例如，模糊程度小的视频帧中的特征量可被优选地登记，或者，具有闭合的眼睛或张开的嘴的脸部可以以低优先级被登记。

[0076] 如用于静止图像的照相机中一样，甚至在用于拍摄运动图像的照相机中快门速度也可以取决于场地的亮度改变。因此，脸部图像可能会在暗地方或者由于对象的移动速度而模糊。这样直接导致图像特征量或属性信息降级。为了估计模糊，获得脸部图像区域中的频率分量，并且，获得低频分量和高频分量之比。如果低频分量的比例超过预定值，则可以

确定已经发生模糊。另外，如果存在闭合的眼睛或张开的嘴，则该器官的图像特征量劣化。在一些情况下，也在属性信息中发生错误。

[0077] 此外，当使用诸如尺寸、闭合的眼睛和模糊的多个参数等决定登记目标的特征量时，假设具有对应于参数类型的坐标轴的空间，并且，获得该坐标空间中的距离。

[0078] 脸部方向图案编号2至9被布置为包围中心位置，因此，可以被视为同一级别。因而，如果在脸部方向图案编号2至9当中存在其总计结果是非零的多个图案，则重复决定最大尺寸的特征量，直到该数量达到要登记的特征量的数量(在本实施例中，为4)。然而，如上所述，对于一个脸部方向图案编号，最多登记一个特征量。这也适用于脸部方向模式编号10至25。

[0079] 如上所述，根据本实施例，当从检测到要追踪的脸部起经过预定时间(在本实施例中，Tp1)时，在核对DB中登记此时段检测到的脸部特征量。因此，当Tp1被设置为几秒至几十秒时，接近当前时刻的人实质上可以被设置为核对目标。

[0080] 根据本实施例，其方向相互不同的脸部的特征量可以与一个对象ID对应地被登记在核对DB中。也就是说，由于同一方向上的脸部的特征量没有被反复地登记，所以可以防止核对DB不必要地变得庞大。另外，在本实施例中，给向前看的脸部赋予高优先级，并且，按优先级的顺序在核对DB中登记预定数量的特征量。这样使得可以进一步防止核对DB变得庞大，并且可以登记精确的特征量。即使脸部方向相同，也登记可以预期对于诸如脸部尺寸等的条件具有高精确度的特征量。因此，也可以使核对精确度变高。

[0081] 请注意，在本实施例中，追踪处理单元205对追踪目标脸部进行追踪，并且，在此时段将特征量和脸部的方向添加至对象追踪信息。然而，在脸部检测中，只需要确定对象是否为人脸。因此，可以进行简单的脸部检测，并且，在结束追踪之后，可以计算要在核对DB中登记的更精确的特征量或脸部方向。

[0082] 在本实施例中，与一个脸部方向对应地登记的特征量的数量是1。如果允许DB稍微庞大，则该数量可以被设置为预定数量或更小。在这种情况下，可以由用户设置预定数量(管理员适当地设置该数量)。

[0083] 请注意，在本实施例中，已经描述这样的示例，其中，从脸部获得的脸部特征用作图像特征，并且，使用该图像特征进行图像特征核对。然而，可以使用从比脸部宽的头部区域或者直到脚趾的全身区域获得的图像特征。例如，从头部区域或全身区域提取局部二值模式(Local Binary Pattern，LBP)特征量或方向梯度直方图(Histogram of Oriented Gradient，HOG)特征量并使用该特征量。

[0084] 在本实施例中，由脸部检测单元206检测到的脸部由追踪处理单元205追踪。然而，如日本特开2002-373332号公报中所述，可以从运动向量检测对象，可以估计下一帧中的搜索位置，并且，可以通过模板匹配来进行人追踪。在这种情况下，从被追踪的人的位置检测脸部，并且，如果检测到脸部，则进行通过脸部检测对其分类的处理。如果没有检测到脸部，则可以将没有检测到脸部的图像分类为一组。如上所述，本实施例可以被广泛地应用于这样的形式，其中，当检测人的全身或部位并追踪检测到的人全身或部位时，登记从该人的全身或部位提取的特征量。在本实施例中，描述了脸部通过脸部方向被分类并被登记在核对DB中的示例。然而，可以通过例如脸部表情(微笑的脸部、含泪的脸部等)对脸部分类。也就是说，本实施例可以被广泛地应用于将从输入视频提取的脸部分类到预定组(脸部方向或

脸部表情)并登记的布置。

[0085] [第二实施例]

[0086] 在第一实施例中,如图6所示,即使追踪处理单元205在从时刻t0到t3的时段追踪人A,独立的对象ID也从时刻t0至t5、从时刻t5至t6和从时刻t6至t7被登记在DB中。在第二实施例中,将描述在从时刻t0至t3的区间中使用一个公共对象ID的示例。更具体地,在时刻t5和t6,连续地使用由追踪处理单元205在时刻t0创建的对象追踪信息。

[0087] 控制处理单元211和图像特征存储单元202的处理不同于第一实施例。下面将描述与第一实施例的不同之处。

[0088] 只有当正在进行连续追踪时,控制处理单元211不执行步骤S89。换句话说,只有当人不可追踪时,控制处理单元211删除登记的对象追踪信息。另外,替代步骤S88,控制处理单元211进行如下处理,将对应的对象追踪信息中的追踪计数402递增"1",并且,将结束帧编号更新为"当前组编号+F×Tp2"(F是由摄像机摄像的帧速率)。

[0089] 在接收到具有与在核对DB中已经登记的对象ID相同的ID的对象追踪信息的登记请求时,图像特征存储单元202通过将该请求视为更新对应的对象ID的登记的特征量的请求来进行处理。

[0090] 结果是,例如,在图6所示的时刻t5,在核对DB中登记在"t0-t5"区间中通过追踪处理获得的人A的特征量。因此,如第一实施例中一样,在开始追踪人A时的早期阶段(对应于时间Tp1的时段)的特征量可以被设置为核对目标。在时刻t6,更新在区间"t0-t6"中通过追踪处理获得的具有相同的对象ID的人A的特征量。也就是说,只要进行连续的追踪,要登记的特征量的数量就决不增加。另外,随着时间流逝,追踪间隔变长。因此,可以提高要登记的特征量的可靠性。

[0091] 上面描述了第一实施例和第二实施例。在这些实施例中,在核对DB中与一个脸部方向对应地登记一个特征量。然而,可以设置与一个脸部方向对应的可登记的特征量的数量。在这些实施例中,在连续追踪中使用两种时间Tp1和Tp2。然而,可以根据追踪计数改变这些时间。

[0092] 其他实施例

[0093] 还可以通过读出并执行记录在存储介质(也可更完整地称为"非暂时性计算机可读存储介质")上的计算机可执行指令(例如,一个或更多个程序)以执行上述实施例中的一个或更多个的功能、并且/或者包括用于执行上述实施例中的一个或更多个的功能的一个或更多个电路(例如,专用集成电路(ASIC))的系统或装置的计算机,来实现本发明的实施例,并且,可以利用通过由系统或装置的计算机例如读出并执行来自存储介质的计算机可执行指令以执行上述实施例中的一个或更多个的功能、并且/或者控制一个或更多个电路以执行上述实施例中的一个或更多个的功能的方法,来实现本发明的实施例。计算机可以包括一个或更多个处理器(例如,中央处理单元(CPU),微处理单元(MPU)),并且可以包括分开的计算机或分开的处理器的网络,以读出并执行计算机可执行指令。计算机可执行指令可以例如从网络或存储介质被提供给计算机。存储介质可以包括例如硬盘、随机存取存储器(RAM)、只读存储器(ROM)、分布式计算系统的存储器、光盘(诸如压缩光盘(CD)、数字通用光盘(DVD)或蓝光光盘(BD)™)、闪存装置以及存储卡等中的一个或更多个。

[0094] 本发明的实施例还可以通过如下的方法来实现,即,通过网络或者各种存储介质

将执行上述实施例的功能的软件(程序)提供给系统或装置，该系统或装置的计算机或是中央处理单元(CPU)、微处理单元(MPU)读出并执行程序的方法。

[0095] 虽然参照示例性实施例对本发明进行了描述，但是应当理解，本发明不限于所公开的示例性实施例。应当对所附权利要求的范围给予最宽的解释，以使其涵盖所有这些变型例以及等同的结构和功能。

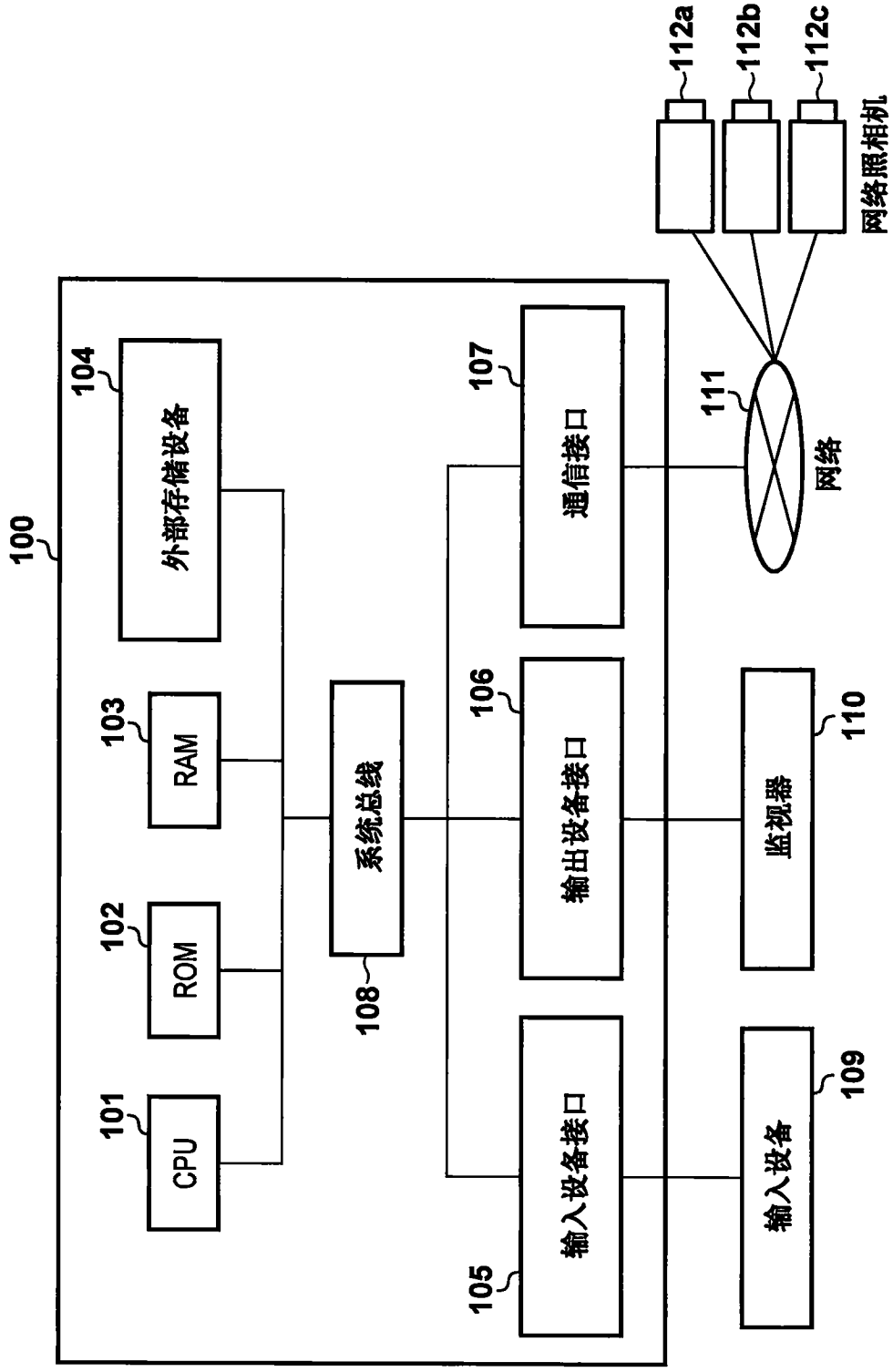

112a
112b
112c
网络照相机
111
网络
100
104
外部存储设备
107
通信接口
103
RAM
106
输出设备接口
110
监视器
系统总线
102
ROM
108
101
CPU
105
输入设备接口
109
输入设备

图1

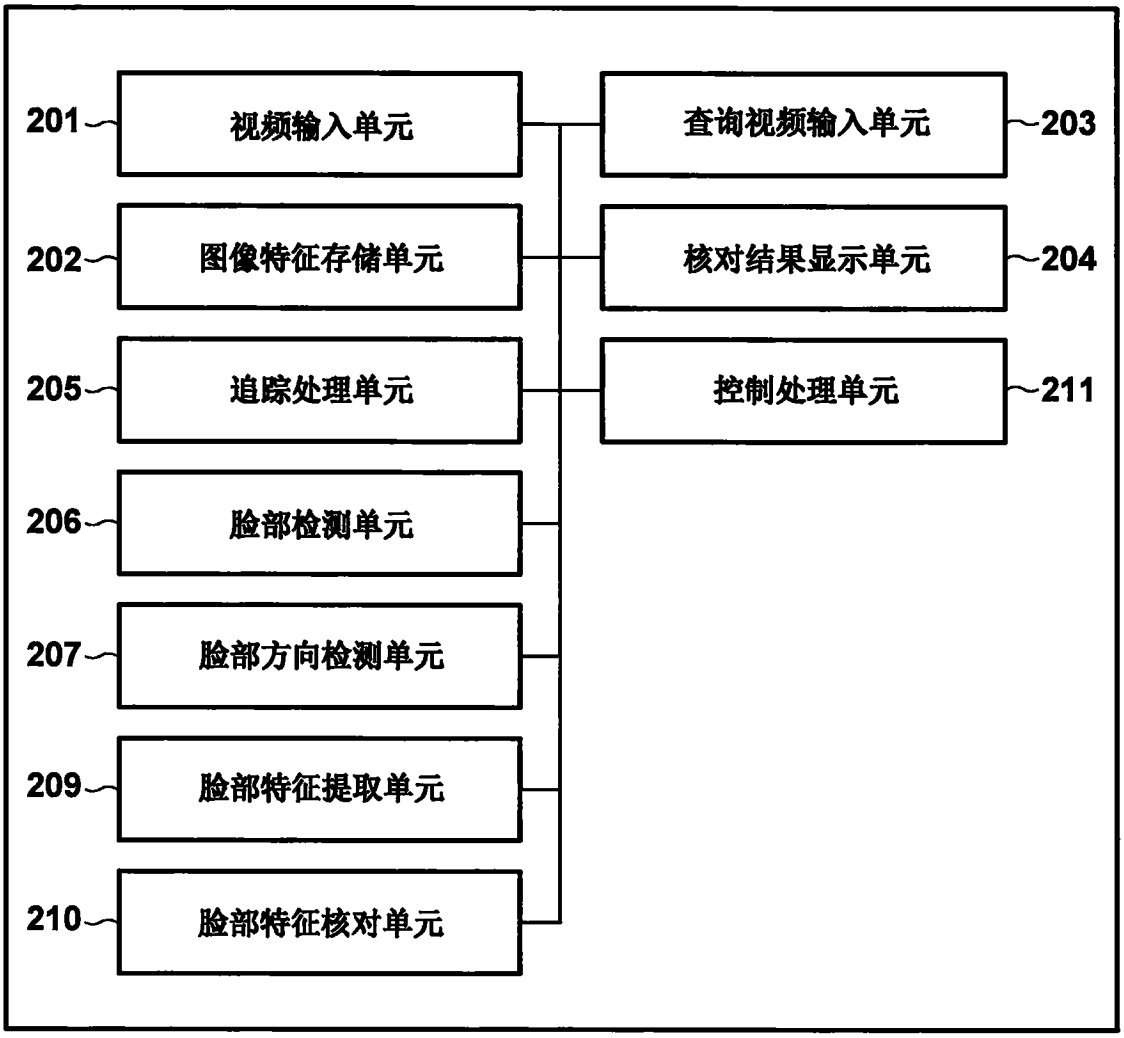


图2

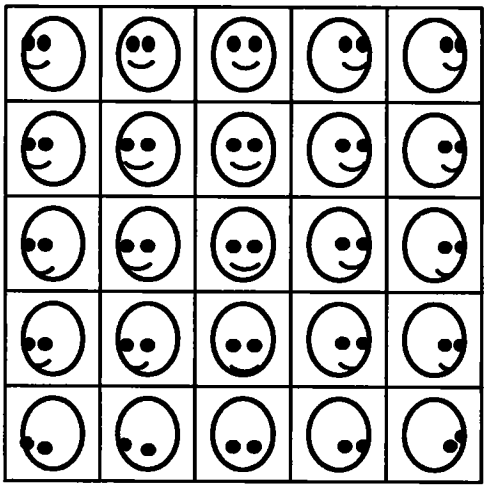


图3A

脸部方向图案编号

24	25	10	11	12
23	9	2	3	13
22	8	1	4	14
21	7	6	5	15
20	19	18	17	16

图3B

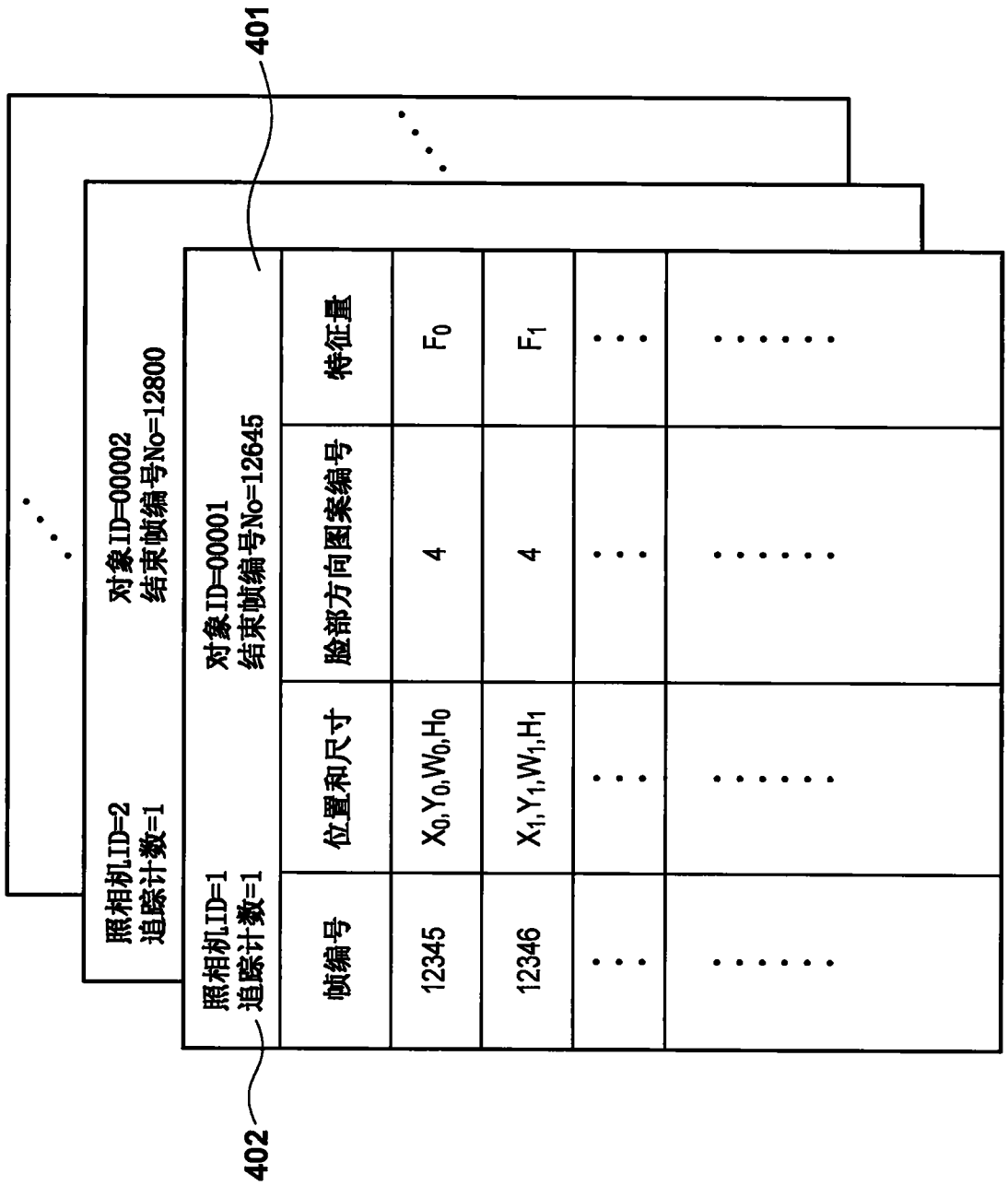

401
照相机ID=1
追踪计数=1
对象ID=00001
结束帧编号No=12645
帧编号
位置和尺寸
脸部方向图案编号
特征量
12345
X_0,Y_0,W_0,H_0
4
F_0
12346
X_1,Y_1,W_1,H_1
4
F_1
照相机ID=2
追踪计数=1
对象ID=00002
结束帧编号No=12800
402

图4

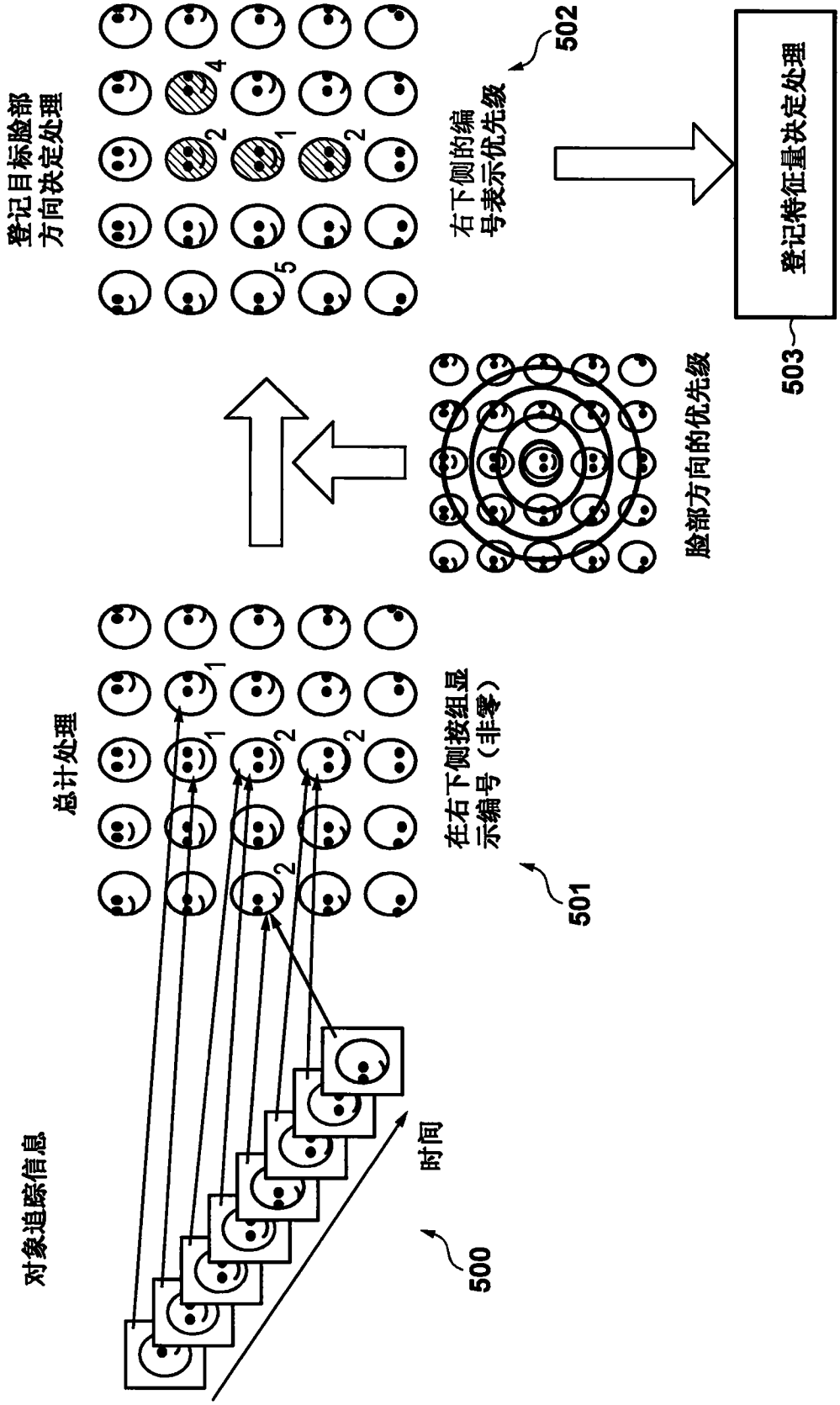

对象追踪信息
时间
500
总计处理
在右下侧按组显示编号（非零）
501
脸部方向的优先级
右下侧的编号表示优先级
502
登记目标脸部方向决定处理
登记特征量决定处理
503

图5

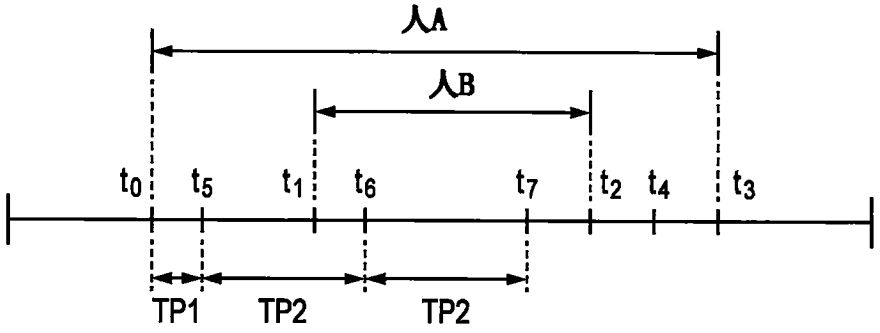


图6

对象ID	特征量（四个）	照相机ID	帧编号、位置、尺寸
00001	XXXXXXX	1	XXXXXXXXX
00002	XXXXXXX	1	XXXXXXXXX
00003	XXXXXXX	2	XXXXXXXXX
⋮	⋮	⋮	⋮

图7

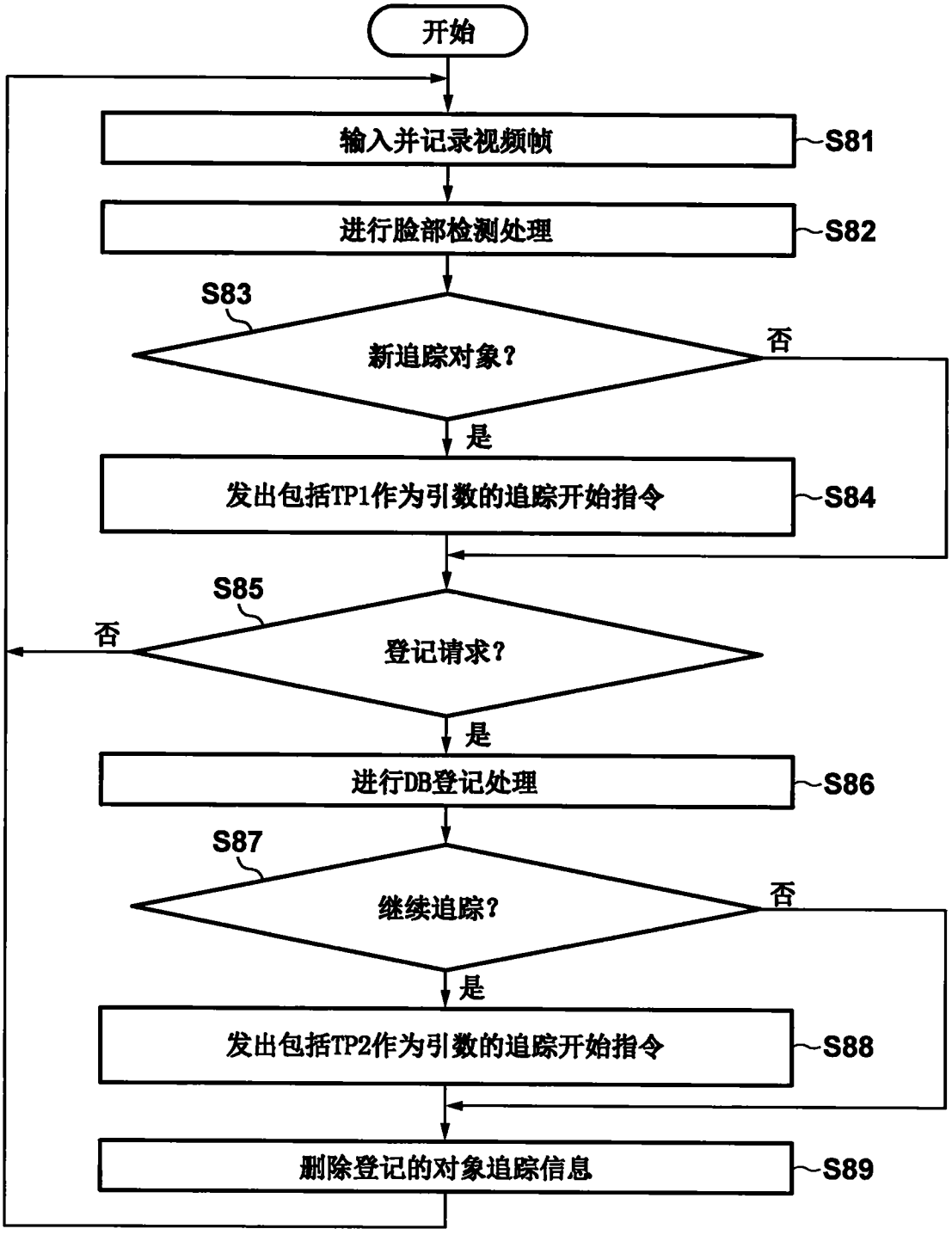

开始
输入并记录视频帧
S81
进行脸部检测处理
S82
S83
新追踪对象？
否
是
发出包括TP1作为引数的追踪开始指令
S84
S85
否
登记请求？
是
进行DB登记处理
S86
S87
继续追踪？
否
是
发出包括TP2作为引数的追踪开始指令
S88
删除登记的对象追踪信息
S89

图8

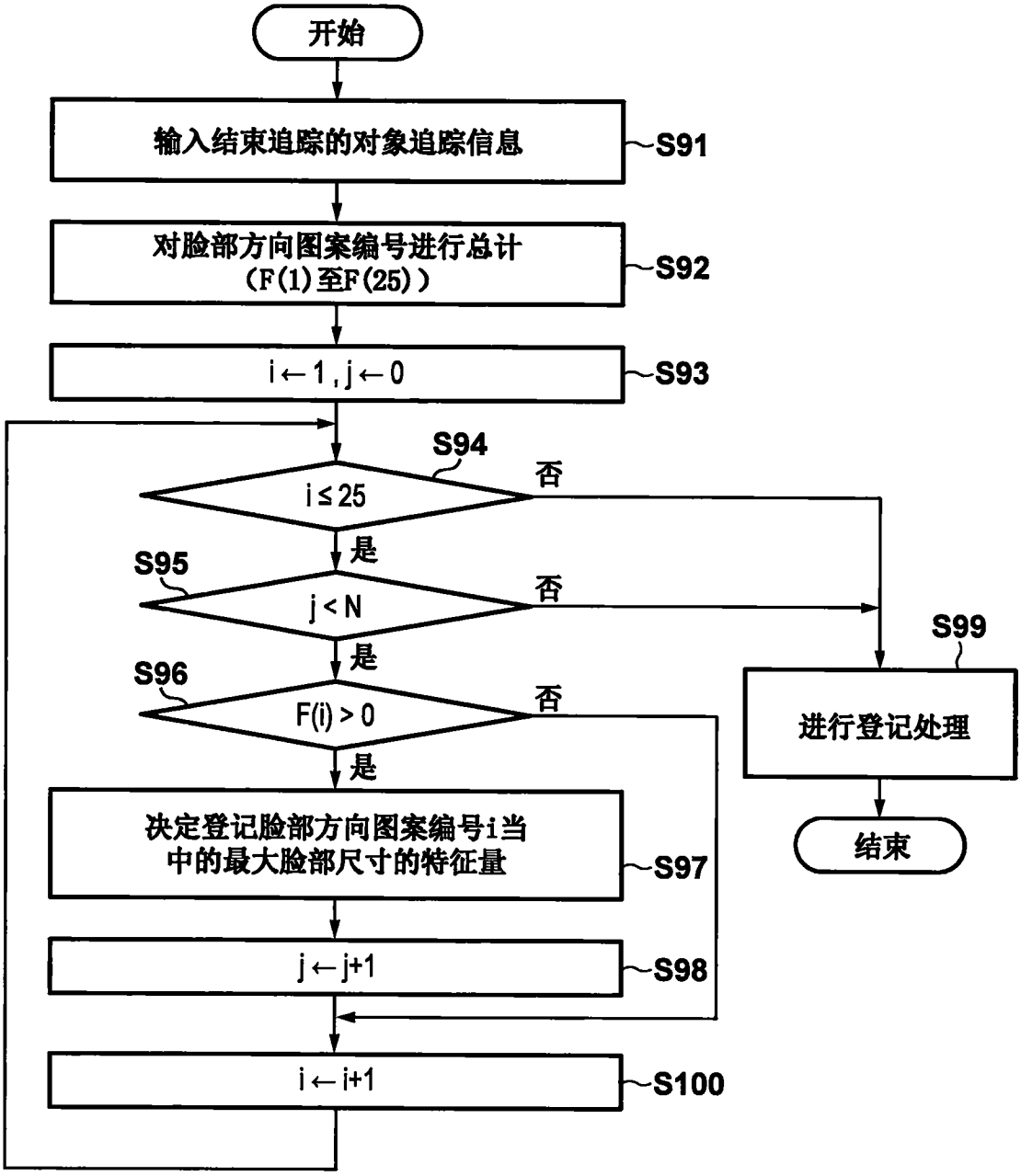

开始
输入结束追踪的对象追踪信息
S91
对脸部方向图案编号进行总计
(F(1)至F(25))
S92
i←1,j←0
S93
S94
i≤25
否
是
S95
j<N
否
是
S96
F(i)>0
否
是
决定登记脸部方向图案编号i当
中的最大脸部尺寸的特征量
S97
j←j+1
S98
i←i+1
S100
S99
进行登记处理
结束

图9

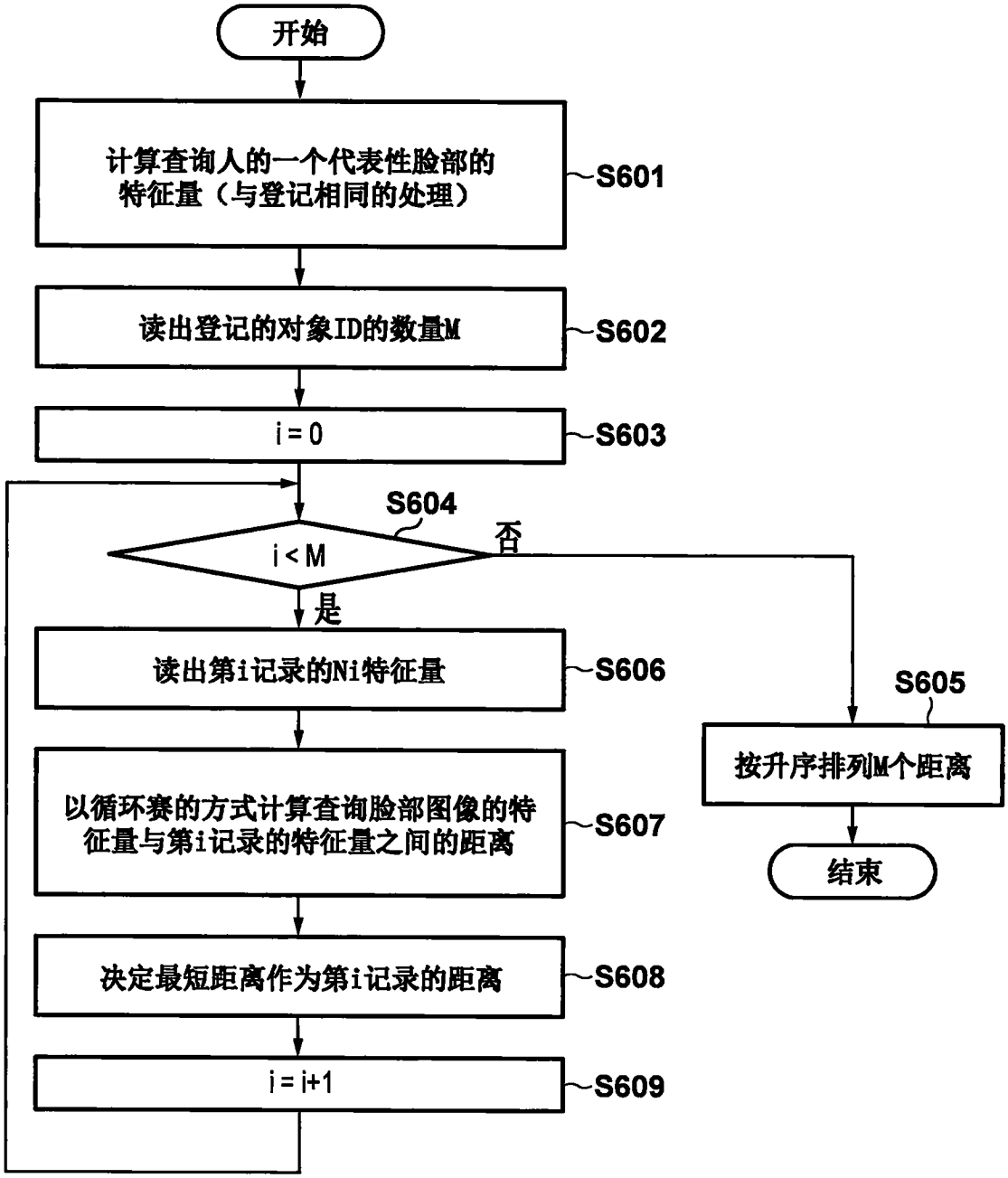

开始
计算查询人的一个代表性脸部的特征量（与登记相同的处理）
S601
读出登记的对象ID的数量M
S602
i = 0
S603
S604
i < M
否
是
读出第i记录的Ni特征量
S606
以循环赛的方式计算查询脸部图像的特征量与第i记录的特征量之间的距离
S607
决定最短距离作为第i记录的距离
S608
i = i+1
S609
S605
按升序排列M个距离
结束

图10