

(12) 发明专利

(10) 授权公告号 CN 101334839 B

(45) 授权公告日 2013. 08. 14

(21) 申请号 200810126194. 4

EP 1700567 A1, 2006. 09. 13,

(22) 申请日 2008. 06. 27

CN 1839410 A, 2006. 09. 27,

CN 1940965 A, 2007. 04. 04,

(30) 优先权数据

2007-173112 2007. 06. 29 JP

审查员 刘力

(73) 专利权人 佳能株式会社

地址 日本东京都大田区下丸子 3-30-2

(72) 发明人 金田雄司 森克彦 佐藤博

(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 任之光 郭召道

(51) Int. Cl.

G06K 9/00 (2006. 01)

(56) 对比文件

CN 1830389 A, 2006. 09. 13,  
2006. 09. 13,

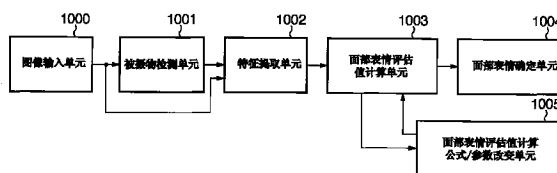
权利要求书2页 说明书11页 附图17页

(54) 发明名称

图像处理装置及图像处理方法

(57) 摘要

本发明提供了一种图像处理装置及图像处理方法。该图像处理装置用于对甚至难以识别面部表情的被摄物执行精确的面部表情识别。从由图像输入单元输入的图像中提取人的面部区域。从所述提取出的面部区域中提取预定的部分区域，所述预定的部分区域在面部表情处于第一状态与第二状态之间变化。通过使用评估值计算公式来计算面部表情评估值。当所述计算出的面部表情评估值大于第一阈值时，确定所述面部表情处于第二状态。如果在预定时间内的所述计算出的面部表情评估值的最大值与最小值之间的差小于预定值，则改变所述评估值计算公式或所述评估值计算公式的参数，以增大该差。



1. 一种图像处理装置,该图像处理装置包括:

面部区域提取单元,用于从输入图像中提取人的面部区域;

部分区域提取单元,用于从所述提取出的面部区域中提取预定的部分区域,在所述预定的部分区域中,面部表情在第一状态与第二状态之间变化;

计算单元,用于通过使用评估公式,来针对所述提取出的部分区域计算面部表情评估值,所述评估公式在所述面部表情指示所述第二状态的情况下生成的值比在所述面部表情指示所述第一状态的情况下生成的值大;

确定单元,在所述计算出的面部表情评估值超过第一阈值的情况下,确定所述面部区域所表达的面部表情指示所述第二状态;

改变单元,用于在预定时间内的所述计算出的面部表情评估值的变化小于预定值的情况下,改变所述评估公式和所述评估公式的参数中的一个,以增大在预定时间内的所述计算出的面部表情评估值的变化。

2. 根据权利要求1所述的图像处理装置,其中,所述改变单元被配置成:改变所述评估公式和所述评估公式的所述参数中的一个,以使所述面部表情评估值的最大值和最小值取预定值。

3. 根据权利要求1所述的图像处理装置,其中,所述计算单元具有如下神经网络,该神经网络的神经元输出值在所述第二状态下变得大并在所述第一状态下变得小;并且所述评估公式表示所述面部表情评估值与所述预定的部分区域的多个像素的神经元输出值中不小于第二阈值的神经元输出值的数量之间的关系。

4. 根据权利要求3所述的图像处理装置,其中,所述面部表情评估值与所述预定的部分区域的所述多个像素的神经元输出值中不小于所述第二阈值的神经元输出值的所述数量之间的所述关系由双曲正切函数表示。

5. 根据权利要求3或权利要求4所述的图像处理装置,其中,所述改变单元被配置成改变作为所述参数的所述第二阈值。

6. 根据权利要求1所述的图像处理装置,其中,

所述预定的部分区域是瞳孔区域;

所述第一状态是眼睁开状态;并且

所述第二状态是眼闭合状态。

7. 一种图像处理方法,所述图像处理方法包括以下步骤:

从输入图像中提取人的面部区域;

从所述提取出的面部区域中提取预定的部分区域,在所述预定的部分区域中,面部表情在第一状态与第二状态之间变化;

通过使用评估公式,来针对所述提取出的部分区域计算面部表情评估值,所述评估公式在所述面部表情指示所述第二状态的情况下生成的值比在所述面部表情指示所述第一状态的情况下生成的值大;

在所述计算出的面部表情评估值超过第一阈值的情况下,确定由所述面部区域所表达的面部表情指示所述第二状态;以及

在预定时间内的所述计算出的面部表情评估值的变化小于预定值的情况下,改变所述评估公式和所述评估公式的参数中的一个来增大在预定时间内的所述计算出的面部表情

评估值的变化。

## 图像处理装置及图像处理方法

### 技术领域

[0001] 本发明涉及图像处理装置及图像处理方法。

### 背景技术

[0002] 已经开发了用于检测作为人类面部表情之一的眼的睁开 / 闭合状态的技术。例如, 参考文献 [1] 中公开的技术对输入图像进行二值化, 并从被二值化的图像中提取对应于瞳孔的黑色区域, 并基于该黑色区域的垂直连续像素的数量来确定眼的睁开 / 闭合状态。在该技术中, 在多个图像中参照黑色 (虹膜) 区域的垂直连续像素的数量的最大值。然后, 基于垂直连续像素的数量的最大值和最小值来设置用于确定是 / 否存在眨眼的阈值。

[0003] 参考文献 [2] 中公开的技术检测上眼睑和下眼睑的边缘, 并基于所述上眼睑边缘与下眼睑边缘之间的距离来确定眼的睁开 / 闭合状态。

[0004] 还开发了用于检测诸如高兴和愤怒的感情的面部表情的技术。例如, 参考文献 [3] 中公开的技术对输入图像执行二维傅立叶变换并生成预定特征向量。根据预先准备的面部表情的隐马尔可夫模型, 来计算特征向量生成的可能性。将与用于计算最大可能性的隐马尔可夫模型相对应的面部表情作为识别结果输出。

[0005] 另一方面, 还开发了基于交互用户、交互量和情境来调整与用户交互的电子秘书的面部表情和该面部表情的程度的技术。例如, 参考文献 [4] 中公开的技术基于交互用户、交互量和情境来从预先存储的表情阈值集合中读取特定表情阈值, 或者通过使用预先定义的几个变换中的特定一个变换来设置表情阈值。利用该方法设置了或改变了电子秘书的面部表情的类型。

[0006] 然而, 包含在面部中的诸如眼和嘴的多个部分的形状和动作因人而异。例如, 对于上眼睑与下眼睑间隔的距离相对远的人, 上眼睑与下眼睑之间的距离的变化量大。但是, 对于上眼睑与下眼睑之间距离近的人, 上眼睑与下眼睑之间的距离的变化量小。

[0007] 在客观地描述面部表情的动作的参考文献 [8] 中, 将作为面部表情之一的“喜悦”描述为 (1) “升高颊部”、(2) “上拉唇端”、……。但是, 颊部或唇端的变化量也因人而异。

[0008] 鉴于以上原因, 如果在确定面部表情的过程中对例如所有人使用相同的参数 (例如, 阈值), 则可能总是将上眼睑与下眼睑距离近的特定人错误地确定为处于眼闭合状态。例如, 则可能总是将少量地移动了诸如眼和嘴之类的多个部分的人错误地确定为无表情。

[0009] 参考文献 [1]: 日本特开平 06-032154 号公报;

[0010] 参考文献 [2]: 日本特开 2000-137792 号公报;

[0011] 参考文献 [3]: 日本特许第 2962549 号公报;

[0012] 参考文献 [4]: 日本特开平 07-104778 号公报;

[0013] 参考文献 [5]: 日本特开 2000-030065 号公报;

[0014] 参考文献 [6]: 日本特开 2003-323622 号公报;

[0015] 参考文献 [7]: 日本特开 2005-056388 号公报;

[0016] 参考文献 [8]: P. Ekman and W. V. Friesen, Facial Action Coding System (FACS);

Manual, Palo Alto :Consulting Psychologists Press, 1978 ;

[0017] 参考文献 [9] :P.Viola and M. Jones, "Rapid object detection using a Boosted Cascade of Simple Features", Proc. of IEEE Conf. CVPR, 1, pp. 511-518, 2001 ;

[0018] 参考文献 [10] :Yann LeCun and Yoshua Bengio, "Convolutinal Networksfor Images, Speech, and Time Series", The Handbook of Brain Theory and Neural Networks, pp. 255-258, 1995 ;

[0019] 参考文献 [11] :Ishii, Ueda, Maeda, and Murase, "Easy-to-Understand Pattern Recognition", Ohmsya, 1998。

## 发明内容

[0020] 鉴于上述问题开发了本发明,并且本发明的目的是甚至对于难以进行面部识别处理的难测的被摄物的面部表情也进行精确的识别。

[0021] 本发明的一个方面提供一种图像处理装置,该图像处理装置包括:面部区域提取单元,用于从输入图像中提取人的面部区域;部分区域提取单元,用于从所述提取出的面部区域中提取预定的部分区域,在所述预定的部分区域中,所述面部表情在第一状态与第二状态之间变化;计算单元,用于利用评估公式,来针对所述提取出的部分区域计算面部表情评估值,所述评估公式在所述面部表情指示所述第二状态的情况下生成大值,并且在所述面部表情指示所述第一状态的情况下生成小值;确定单元,在所述计算出的面部表情评估值超过第一阈值的情况下,确定该面部区域所表达的面部表情指示所述第二状态;改变单元,用于在预定时间内的所计算出的面部表情评估值的变化小于预定值的情况下,改变所述评估公式和所述评估公式的参数中的一个来增大所述变化。

[0022] 本发明的另一方面提供一种图像处理方法,该图像处理方法包括以下步骤:从输入图像中提取人的面部区域;从所述提取出的面部区域中提取预定的部分区域,在所述预定的部分区域中,所述面部表情在第一状态与第二状态之间变化;利用评估公式,来针对所述提取出的部分区域计算面部表情评估值,所述评估公式在所述面部表情指示所述第二状态的情况下生成大值,并且在所述面部表情指示所述第一状态的情况下生成小值;在所述计算出的面部表情评估值超过第一阈值的情况下,确定所述面部区域所表达的面部表情指示所述第二状态;以及在预定时间内的所计算出的面部表情评估值的变化小于预定值的情况下,改变所述评估公式和所述评估公式的参数中的一个来增大所述变化。

[0023] 通过以下参照附图对示例性实施例的详细说明,本发明其他的特征将变得明确。

## 附图说明

[0024] 图 1 是示出根据第一实施例的图像处理装置的结构框图。

[0025] 图 2 是示出输入图像和正规化图像的示例图。

[0026] 图 3 是用于解释用于检测瞳孔区域的卷积神经网络 (CNN) 的图。

[0027] 图 4 是示出该 CNN 的神经元输出值直方图的示例图。

[0028] 图 5 是示出面部表情评估值的特性曲线的曲线图。

[0029] 图 6 是用于解释面部表情评估值的阈值处理的曲线图。

[0030] 图 7 是示出上眼睑与下眼睑之间距离大的人的眼和上眼睑与下眼睑之间距离小

的人的眼的图。

[0031] 图 8 和图 9 是用于解释眼闭合状态检测错误的曲线图。

[0032] 图 10 是示出当改变图 8 中的特性曲线的面部表情评估值计算公式的系数时的面部表情评估值特性曲线的示例的曲线图。

[0033] 图 11 是示出具有根据图 8 中的面部表情评估值特性曲线改变的线性特性曲线的面部表情评估值特性曲线的示例的曲线图。

[0034] 图 12 是例示根据第一实施例的改变面部表情评估值计算公式的系数的处理的流程图。

[0035] 图 13 是示出根据第二实施例的图像处理装置的结构框图。

[0036] 图 14 是用于解释根据第二实施例的神经元输出值的阈值的改变的曲线图。

[0037] 图 15 是例示根据第二实施例的阈值变化处理的流程图。

[0038] 图 16 是示出根据第三实施例的摄像设备的结构框图。

[0039] 图 17A 和 17B 是例示根据第三实施例的摄像设备的操作的流程图。

[0040] 图 18 是用于解释根据第三实施例的面部表情评估值的阈值的改变的曲线图。

### 具体实施方式

[0041] 将根据附图详细描述本发明的优选实施例。本发明并不限于实施例的公开内容，并且对于本发明的解决手段，在实施例中描述的特征的所有组合并不总是必需的。

[0042] (第一实施例)

[0043] 图 1 是示出根据第一实施例的图像处理装置的结构框图。

[0044] 图像输入单元 1000 输入图像信号。被摄物检测单元 1001 从自图像输入单元 1000 输入的图像信号中检测是 / 否存在人脸，并且如果存在人脸，则检测该人脸的位置。特征提取单元 1002 通过使用由被摄物检测单元 1001 检测到的人脸位置和由图像输入单元 1000 生成的图像，来提取面部表情识别所必需的特征。面部表情评估值计算单元 1003 通过使用特征提取单元 1002 提取的面部表情识别所必需的特征，来计算代表各被摄物的面部表情的类型和各个面部表情的程度的面部表情评估值。面部表情确定单元 1004 通过使用由面部表情评估值计算单元 1003 计算的被摄物的面部表情评估值，来确定各被摄物的面部表情。面部表情评估值计算公式 / 参数改变单元 1005 基于由面部表情评估值计算单元 1003 计算的面部表情评估值，来改变用于以后的面部表情评估值的计算公式或该计算公式的系数。

[0045] 以下将使用作为人的面部表情之一的眼闭合作为示例来详细描述本实施例。

[0046] 图像输入单元 1000 输入由摄像设备生成的图像信号，该摄像设备包括镜头、图像捕获元件（例如 CMOS 传感器或 CCD）、A/D 转换器和图像处理电路。例如，输入图 2 所示的图像 2200。

[0047] 被摄物检测单元 1001 从自图像输入单元 1000 输入的图像信号中提取人脸区域的位置。为此，被摄物检测单元 1001 使用例如利用多个矩形滤波器检测面部位置的技术和被称作 Integral Image（参考文献 [9]）的技术。通过使用称作 AdaBoost 的学习方法，学习预先准备的大量的面部图像和非面部图像，来选择用于检测面部位置的矩形滤波器的类型。

[0048] 被摄物检测单元 1001 还利用三个点（即：面部位置、左眼区域的重心位置和右眼

区域的重心位置)来执行图像正规化处理,以获得具有预定大小和预定方向的面部图像。例如,使用以下方法来检测左眼区域的重心位置和右眼区域的重心位置。首先,基于各个人的面部位置,将包括眼白、眼睑和瞳孔的眼搜索区域设置在上侧。类似上述的面部检测,使用参考文献[9]中公开的技术,从所设置的眼搜索区域中检测包括左眼和右眼的眼白、眼睑和瞳孔的眼区域的重心位置。如在面部检测中一样,通过学习预先准备的各种类型的眼图案图像(例如睁大的眼的图像、具有多种视线的图像和闭合的眼的图像),来选择矩形滤波器。

[0049] 因此,能够检测不仅处于眼睁开状态下而且处于眼闭合状态下的包括眼白、眼睑和瞳孔的区域的区域的重心位置。在检测到面部位置和左眼区域的重心位置和右眼区域的重心位置之后,从由图像输入单元 1000 输入的图像 2200 中裁切出各个人的面部区域(如图 2 所示)。执行仿射变换以获得图 2 中的具有预定大小和预定方向的正规化图像 2201 和 2202。例如,从大小为  $720 \times 480$ [像素]的输入图像中提取面部区域。当将左眼区域的重心位置与右眼区域的重心位置之间的距离设置为 40[像素]时,获得从正面观察到的、具有  $100 \times 100$ [像素]大小的正规化图像。

[0050] 当输入图像中的人脸的大小很小(例如,当左眼区域的重心位置与右眼区域的重心位置之间的距离为 20[像素]或更小时),或者面部方向超出预定范围(例如,与被设为  $0^\circ$  的正面成  $30^\circ$  或以上)时,禁止所述处理。

[0051] 特征提取单元 1002 通过使用由被摄物检测单元 1001 检测到的左眼区域的重心位置和右眼区域的重心位置,来提取瞳孔区域作为部分区域。为了检测瞳孔区域,首先,针对左眼区域和右眼区域中的每一个,基于由被摄物检测单元 1001 检测到的它们的重心位置,来设置瞳孔搜索区域。从各个设置的瞳孔搜索区域裁切图像,并提交给已预先完成了学习的卷积神经网络(参见参考文献[10];以下将卷积神经网络称为 CNN)。

[0052] 本实施例中使用的 CNN 具有包括输入层、中间层和输出层的三层结构,如图 3 所示。为了进行学习,准备了大量的表示处于眼睁开状态(第一状态)下的瞳孔周围区域的正解图像(correct image)和表示处于眼闭合状态(第二状态)下的瞳孔周围区域的非正解图像(non-correct image)。

[0053] 接下来,调整神经网络中的多个特征之间的权重系数,使得当如图 3 所示将正解图像 2250 输入输入层时,从输出层输出的对应于瞳孔区域的神经元值变得大,如在图 3 所示的输出值图像 2251 中的那样。更具体地说,使用最速下降法(参见参考文献[11])改变权重系数。相反地,调整所述权重系数,使得当输入非正解图像时,从输出层输出的对应于瞳孔区域的神经元值变得小。

[0054] 重复该学习。然后,当如图 3 所示输入睁开着的眼的图像时,在输出层中只有对应于瞳孔区域的神经元区域的输出值变得大,而其余部分的值较小。注意一个神经元输出值对应于图像的一个像素。

[0055] 在本实施例中,使用了具有三层结构的 CNN(如图 3 所示)。但是,本实施例并不限于此。可以通过使用除 CNN 以外的其他技术来检测对应于瞳孔区域的特征。

[0056] 面部表情评估值计算单元 1003 将特征提取单元 1002 所提取的对应于瞳孔区域的神经元输出值与阈值  $Th1$  进行比较。对等于或大于阈值  $Th1$  的神经元输出值的数量  $c$  进行计数。通过使用数量  $c$  来计算表示眼闭合程度的面部表情评估值  $Exp1$ 。例如,当由特征

提取单元 1002 提取的神经元输出值分布在 0 至 255 之间时,可以将阈值  $Th1$  设置为例如 200(如图 4 所示),并对等于或大于阈值  $Th1$  的神经元输出值的数量  $c$  进行计数。通过使用表示图 5 所示的特征曲线的面部表情评估值计算公式来计算面部表情评估值  $Exp1$  (0 至 255)。面部表情评估值  $Exp1$  越大,眼闭合的可能性越高。图 5 示出了面部表情评估值与瞳孔区域的像素的神经元输出值中等于或大于阈值  $Th1$  的神经元输出值的数量之间的关系。面部表情评估值的特征曲线可以由例如双曲正切函数表示。更具体地说,面部表情评估值计算公式由下式给出:

[0057]  $Exp1 = A0(\tanh(-B0*(c-C0))+D0)$  其中,  $A0$ 、 $B0$ 、 $C0$  和  $D0$  为常数。

[0058] 图 6 示出了面部表情评估值  $Exp1$  在人的眨眼期间的随时间的变化。如图 6 所示,当睁开时,面部表情评估值  $Exp1$  小。当眼闭合时,面部表情评估值增大。当眼再次睁开时,面部表情评估值  $Exp1$  返回到小值。面部表情确定单元 1004 执行二进制确定,以通过将表示眼闭合程度的面部表情评估值  $Exp1$  与阈值  $Th2$  进行比较,来确定眼是闭合的还是睁开的。将阈值  $Th2$  设置为,例如 210。如果面部表情评估值等于或小于阈值  $Th2$ (即  $Exp1 \leq 210$ ),则面部表情确定单元 1004 确定眼是睁开的(第一状态)。另一方面,如果面部表情评估值大于阈值  $Th2$ (即  $Exp1 > 210$ ),则面部表情确定单元 1004 确定眼是闭合的(第二状态)。

[0059] 面部表情评估值计算公式/参数改变单元 1005 改变面部表情评估值计算公式或其系数,使得在预定时间  $t1[s]$  内的所计算出的面部表情评估值  $Exp1$  的最大值变为 255,而最小值变为 0。在本实施例中,最大值是 255,而最小值是 0。但是,本发明并不限于此。

[0060] 特征提取单元 1002 所使用的 CNN 进行学习以检测瞳孔区域。因此,当瞳孔区域变得大时,神经元输出值和 CNN 的输出区域也变得大。对于上眼睑与下眼睑之间具有平均距离(如图 7 所示的 2300)的人,则神经元输出值和 CNN 的输出区域在眼闭合状态与睁开状态之间的变化量相对较大。因此,面部表情评估值  $Exp1$  显示出如图 6 所示的随时间的变化,并且能够基于阈值  $Th2$  正确地确定眼闭合。

[0061] 然而,对于上眼睑与下眼睑之间距离近的天生细缝眼的人(如图 7 中的 2301),则神经元输出值和 CNN 的输出区域在眼闭合状态和睁开状态之间的变化量较小。为此,如图 8 所示,神经元值的数量  $c$  的最大值为  $c_{max1}$ ,而最小值为  $c_{min1}$ 。因此,仅在窄范围 2305 内计算面部表情评估值  $Exp1$ 。面部表情评估值  $Exp1$  显示出如图 9 所示的随时间的变化。结果,实际上没有闭合着眼的人常被错误地检测为处于眼闭合状态。这不仅应用于基于作为检测特征的瞳孔区域面积的确定,还应用于利用上眼睑与下眼睑之间的距离的眼闭合确定。这还不仅应用于细缝眼的人,还应用于由于不同的瞳孔颜色而难以检测到其瞳孔区域的人。

[0062] 如果面部表情评估值  $Exp1$  的最大值与最小值之间的差小于预定值,则面部表情评估值计算公式/参数改变单元 1005 改变面部表情评估值计算公式或其系数,以使该差变大。更具体地说,面部表情评估值计算公式/参数改变单元 1005 改变面部表情评估值计算公式或其系数,使得例如在预定时间  $t1[s]$  内的所计算出的面部表情评估值  $Exp1$  的最大值变为 255,而最小值变为 0。这使得能够对难以确定眼闭合的被摄物进行更精确的检测。注意所述预定时间  $t[s]$  是足够一次眨眼的的时间。通常是,人每 3 到 10 秒眨一次眼。例如,设置预定时间  $t1 = 20[s]$ 。

[0063] 例如,以下述方式改变面部表情评估值计算公式的系数。首先,获得预定时间

$t1[s]$  内的面部表情评估值  $Exp1$  的最小值  $Exp1min$  和最大值  $Exp1max$ , 以及神经元输出值的最小计数  $cmin$  和最大计数  $cmax$ 。然后, 通过使用这些值和预先准备的表, 来决定面部表情评估值计算公式的系数, 以获得显示出例如如图 10 所示的特性曲线的面部表情评估值计算公式。

[0064] 代替改变面部表情评估值计算公式的系数, 可以通过使用如图 11 所示的线性函数来改变面部表情评估值计算公式本身。

[0065] 图 12 是例示改变面部表情评估值计算公式的系数的处理的流程图。

[0066] 在步骤 S2600 中, 确定是否已经过了预先设置的预定时间  $t1[s]$ 。例如,  $t1 = 20[s]$ 。如果在步骤 S2600 中确定尚未经过预定时间  $t1[s]$ , 则处理进入步骤 S2603。在步骤 S2603 中, 对面部表情评估值  $Exp1$  的最大值和最小值以及等于或大于阈值  $Th1$  的神经元输出值的数量进行存储, 然后处理返回步骤 S2600。

[0067] 如果在步骤 S2600 中确定已经过了预定时间  $t1[s]$ , 则处理进入步骤 S2601。在步骤 S2601 中, 确定预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值是否是 255, 最小值是否是 0。如果在步骤 S2601 中确定面部表情评估值  $Exp1$  的最大值不是 255 或者最小值不是 0, 则处理进入步骤 S2602。

[0068] 在步骤 S2602 中, 使用预先准备的表来改变面部表情评估值计算公式的系数, 然后结束该处理。如果在步骤 S2601 中确定面部表情评估值  $Exp1$  的最大值为 255 且最小值为 0, 则直接结束该处理。

[0069] 优选的是, 针对多人中的每个人执行上述处理。

[0070] 综上所述, 通过改变面部表情评估值计算公式或其系数来使面部表情评估值的变化量变大。这使得能够对难测的被摄物的眼闭合进行更精确的检测。

[0071] (第二实施例)

[0072] 图 13 是示出根据第二实施例的图像处理装置的结构框图。

[0073] 图像输入单元 3000 输入图像信号。被摄物检测单元 3001 从自图像输入单元 3000 输入的图像信号中检测是 / 否存在人脸, 并且如果存在人脸, 则检测该人脸的位置。特征提取单元 3002 通过使用由被摄物检测单元 3001 检测到的人脸位置和由图像输入单元 3000 生成的图像, 来提取面部表情识别所必需的特征。面部表情评估值计算单元 3003 通过使用特征提取单元 3002 提取的面部表情识别所必需的所述特征, 来计算代表各被摄物的面部表情的类型和各个面部表情的程度的面部表情评估值。面部表情确定单元 3004 通过使用由面部表情评估值计算单元 3003 计算的被摄物的所述面部表情评估值, 来确定各被摄物的面部表情。特征提取结果阈值改变单元 3005 基于由面部表情评估值计算单元 3003 计算的面部表情评估值, 来改变要用于特征提取单元 3002 所提取的所述特征的阈值。

[0074] 如第一实施例那样, 以下将使用作为人的面部表情之一的眼闭合作为示例来详细描述本实施例。

[0075] 图像输入单元 3000 至面部表情确定单元 3004 执行与第一实施例的图像输入单元 1000 至面部表情确定单元 1004 相同的处理, 因此不再赘述。

[0076] 特征提取结果阈值改变单元 3005 改变 CNN 的神经元输出值的阈值  $Th1$ , 使得面部表情评估值  $Exp1$  的最大值变为 255, 而最小值变为 0 (如第一实施例中那样)。例如, 如图 14 所示, 将阈值  $Th1$  由 200 变为 150。如在第一实施例中描述的, 对于上眼睑与下眼睑之间

的距离相对近的人,出现的瞳孔区域小,因此,从 CNN 获得的神经元输出值往往也小。通常针对从 CNN 获得的神经元输出值小的人(如上眼睑与下眼睑之间的距离相对近的人),将阈值  $Th1$  设置得小。作为改变阈值  $Th1$  的详细方法,对预定时间  $t1[s]$  内的神经元输出值的最大值  $Numax$  进行计算,并将对应于最大值  $Numax \times 0.8$  的神经元输出值设置为阈值  $Th1$ 。例如,当在预定时间  $t1[s]$  内的神经元输出值的最大值是 200 时,则将阈值  $Th1$  设置为 160。

[0077] 图 15 是例示对来自 CNN 的输出层的神经元输出值的阈值  $Th1$ 、面部表情评估值计算公式或其系数进行改变的处理的流程图。

[0078] 在步骤 S3100 中,确定是否经过了预定时间  $t1[s]$ 。如果在步骤 S3100 中确定尚未经过预定时间  $t1[s]$ ,则处理进入步骤 S3103。在步骤 S3103 中,对面部表情评估值  $Exp1$  和从 CNN 输出的神经元输出值进行存储。如果在步骤 S3100 中确定已经过了预定时间  $t1[s]$ ,则处理进入步骤 S3101。

[0079] 在步骤 S3101 中,确定预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值是否是 255,最小值是否是 0。如果确定最大值不是 255 或最小值不是 0,则处理进入步骤 S3102。在步骤 S3102 中,通过上述方法将阈值  $Th1$  改变为  $Th1'$ 。然后处理进入步骤 S3104。

[0080] 在步骤 S3104 中,通过使用在步骤 S3102 中改变的阈值  $Th1'$ ,来再次计算预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值和最小值,并确定最大值是否是 255,最小值是否是 0。如果在步骤 S3104 中确定在预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值不是 255 或最小值不是 0,则处理进入步骤 S3105。

[0081] 在步骤 S3105 中,通过使用与第一实施例中的方法来改变面部表情评估值计算公式或其系数,使得在预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值变为 255 且最小值变为 0。

[0082] 如果在步骤 S3101 或 S3104 中确定在预定时间  $t1[s]$  内的面部表情评估值  $Exp1$  的最大值是 255 且最小值是 0,则直接结束该处理。

[0083] 综上所述,通过改变 CNN 的神经元输出值的阈值  $Th1$  使面部表情评估值的变化量变大。这使得能够对难测的被摄物的眼闭合状态进行更精确的检测。

[0084] (第三实施例)

[0085] 图 16 是示出根据第三实施例的摄像设备的结构的框图。

[0086] 参照图 16,标号 5000 表示成像镜头组,而 5001 表示具有光圈装置和快门装置的光量调节单元。诸如 CCD 或 CMOS 的图像捕获元件 5002 将透过成像镜头组的与被摄物图像相对应的光束转化为电子信号。

[0087] 模拟信号处理电路 5003 对从图像捕获元件 5002 输出的模拟信号执行钳位处理和增益处理。模拟/数字(以下称作“A/D”)转换器 5004 将来自模拟信号处理电路 5003 的输出转化为数字信号。

[0088] 图像信号处理电路 5005 对来自 A/D 转换器 5004 的数据或来自存储器控制电路 5006 的数据执行预定的像素插值处理和色彩转换处理。图像信号处理电路 5005 还通过使用所摄图像数据来进行预定运算处理。

[0089] 系统控制电路 5009 基于上述运算结果来执行 TTL 型 AF(自动聚焦)处理、AE(自动曝光)处理和 EF(Electronic Flash pre-light emission:电子闪光灯预点亮发光)处理以控制曝光控制电路 5011 和焦点控制电路 5010。

[0090] 图像信号处理电路 5005 还通过使用所摄图像数据来执行预定运算处理,并基于所获得的运算结果来执行 TTL AWB(自动白平衡)处理。图像信号处理电路 5005 还基于所摄图像数据,来执行人脸检测处理、个人身份验证处理和面部表情识别处理。将用于预定图像处理(例如面部检测处理、个人身份验证处理和面部表情识别处理)的数据存储在图像信号处理电路 5005 的存储器中。

[0091] 存储器控制电路 5006 控制模拟信号处理电路 5003、A/D 转换器 5004、图像信号处理电路 5005、存储器 5007 和数字/模拟(以下称作“D/A”)转换器 5012。将由 A/D 转换器 5004A/D 转换的数据经由图像信号处理电路 5005 和存储器控制电路 5006 或直接通过存储器控制电路 5006 写入存储器 5007。

[0092] 存储器 5007 存储要显示在显示设备 5013 上的数据。将记录在存储器 5007 中的数据经由 D/A 转换器 5012 输出给并显示在诸如 TFT 或 LCD 的显示设备 5013 上。存储器 5007 还存储所摄静态图像和动态图像。存储器 5007 具有足够的存储容量以存储预定数量的静态图像和预定时间的动态图像。即使在连续地拍摄多个静态图像的连拍模式或全景拍摄模式中,能高速将大量的数据写入存储器 5007 中。存储器 5007 还作为系统控制电路 5009 的工作区。

[0093] 显示设备 5013 能够通过依次地显示所摄图像数据而实现电子取景器功能。显示设备 5013 能够按照来自系统控制电路 5009 的指示而任意地打开/关闭所述显示。与显示打开状态相比,在显示关闭状态下能够显著地降低摄像设备的功耗。显示设备 5013 还根据系统控制电路 5009 执行的程序利用文本或图像来显示操作状态或消息。

[0094] 标号 5008 表示对存储介质(例如存储卡或硬盘)的接口。通过使用接口 5008,能够从另一计算机或诸如打印机的外围设备传送图像数据和附在该图像数据上的管理信息,或者能够将所述图像数据和附在该图像数据上的管理信息传送给所述另一计算机或外围设备。当接口 5008 符合诸如 PCMCIA 卡或 CF(Compact Flash®)卡的标准时,该接口 5008 能够连接各种类型的通信卡。各种通信卡的示例是:LAN 卡、调制解调器卡、USB 卡、IEEE1394 卡、P1284 卡、SCSI 卡以及 PHS 用通信卡。

[0095] 系统控制电路 5009 控制整个摄像设备。系统控制电路 5009 的存储器存储系统控制电路 5009 的操作使用的常数、变量和程序。

[0096] 曝光控制电路 5011 控制光量调整单元 5001 的光圈装置和快门装置。焦点控制电路 5010 控制成像镜头组 5000 的聚焦和变焦。使用 TTL 方法控制曝光控制电路 5011 和焦点控制电路 5010。系统控制电路 5009 基于图像信号处理电路 5005 对图像数据进行运算处理后获得的运算结果,来控制曝光控制电路 5011 和焦点控制电路 5010。

[0097] 将参照图 17A 和 17B 中的流程图来描述根据本实施例的摄像设备的操作。该处理在接通电源时开始。首先,在步骤 S5500 中,系统控制电路 5009 初始化内部存储器中的各种标志和控制变量。在步骤 S5501 中,系统控制电路 5009 检测摄像设备的模式设置状态。如果设置了除拍摄模式以外的模式,则处理进入步骤 S5502。在步骤 S5502 中,系统控制电路 5009 执行对应于所选择的模式的处理。当该处理完成时,处理返回步骤 S5501。在该情况下,也可以设置应该触发拍摄的面部表情。在本实施例中,从面部表情中选择“笑脸”。

[0098] 如果在步骤 S5501 中设置自动拍摄模式,则处理进入步骤 S5503。系统控制电路 5009 确定电源的剩余电量或操作状态在对摄像设备的操作中是否有问题。如果系统控制电

路 5009 确定有问题,则处理进入步骤 S5504,以使显示设备 5013 利用图像或声音输出预定警报。然后,处理返回步骤 S5501。

[0099] 如果系统控制电路 5009 在步骤 S5503 中确定电源没有问题,则处理进入步骤 S5505。在步骤 S5505 中,系统控制电路 5009 确定存储介质的操作状态在对摄像设备的操作中是否有问题,更具体地说,确定对存储介质的图像数据记录 / 图像数据回放操作是否有问题。如果系统控制电路 5009 确定有问题,则处理进入步骤 S5504,以使显示设备 5013 通过图像或声音输出预定警报。然后,处理返回步骤 S5501。

[0100] 如果系统控制电路 5009 在步骤 S5505 中确定存储介质没有问题,则处理进入步骤 S5506。在步骤 S5506 中,系统控制电路 5009 使显示设备 5013 通过图像或声音来显示摄像设备的各种设置状态的用户界面(下文中称作 UI)。当显示设备 5013 的图像显示是 ON(开启)时,则可能使用显示设备 5013 通过图像或声音来显示摄像设备的各种设置状态的 UI。这样,用户执行各种设置。

[0101] 在步骤 S5507 中,系统控制电路 5009 开启显示设备 5013 的图像显示。在步骤 S5508 中,系统控制电路 5009 设置直通显示(through display)状态以依次地显示所摄图像数据。在该直通显示状态中,将依次地写入存储器 5007 的数据依次地显示在显示设备 5013 上,由此实现电子取景器功能。

[0102] 在步骤 S5509 中,系统控制电路 5009 确定用户(例如拍摄者)是否已按下了快门开关。如果用户尚未按下快门开关,则处理返回步骤 S5501。如果用户已按下了快门开关,则系统控制电路 5009 执行步骤 S5510 中的面部检测处理。所述面部检测处理通过与第一实施例中相同的处理来实现。

[0103] 在步骤 S5511 中,根据步骤 S5510 中检测的面部来进行 AE 和 AF。在步骤 S5512 中,开始对从步骤 S5509 中开关接通起所经过的时间进行测量。

[0104] 在步骤 S5513 中,显示设备 5013 不仅显示所摄图像数据还显示在步骤 S5510 中检测的面部的位置。

[0105] 在步骤 S5514 中,对步骤 S5510 中检测的各个面部执行个人身份验证。该个人身份验证采用例如参考文献 [5] 中的共用子空间方法,或参考文献 [6] 中的将面部区域分割成多个子区域并针对每个子区域进行比较的方法。在本实施例中,使用了参考文献 [6] 中的方法,即将面部区域分割成多个子区域的方法。用于该身份验证的图像不需要总是预先登记,可以在拍摄之前直接登记。个人身份验证方法不限于本实施例中描述的方法。

[0106] 在步骤 S5515 中,只对步骤 S5514 中经身份验证的特定人执行笑脸识别。基于与预先准备的无表情数据的差异来进行笑脸识别(如参考文献 [7])。该面部表情识别方法不限于在本实施例中描述的方法。

[0107] 在步骤 S5516 中,系统控制电路 5009 确定是否满足预定拍摄条件。例如,当在步骤 S5501 中设置了笑脸拍摄模式时,系统控制电路 5009 确定在步骤 S5515 中确定的面部表情是否是“笑脸”。如果在步骤 S5516 中确定满足预定拍摄条件(也就是说,图像中的预定人的面部表情是“笑脸”),则处理进入步骤 S5517。

[0108] 在步骤 S5517 中,进行拍摄。在步骤 S5518 中,显示设备 5013 显示所摄图像。在步骤 S5519 中,将所摄图像通过接口 5008 存储到例如 CF(小型闪存)中。

[0109] 在将所摄图像存储到 CF 中的过程中,除了诸如所摄图像的图像大小等信息外,还

对面部表情识别处理中使用的参数（例如，面部表情确定阈值）、面部表情识别结果和个人身份验证 ID 进行记录。例如，将面部表情识别处理中使用的参数（例如，面部表情确定阈值）或个人身份验证 ID 记录在 Exif 信息中。当将面部表情识别处理中使用的参数、面部表情识别结果和个人身份验证 ID 与所摄图像一起记录时，能够从存储在例如硬盘中的巨量图像中快速地找到做出特定面部表情的人 A 的图像。对于没有面部表情识别结果的图像，优先搜索针对其记录了面部表情识别中使用的参数的同一人的图像。然后，可以通过使用所记录的参数来对没有面部表情识别结果的图像执行面部表情识别。

[0110] 如果在步骤 S5516 中确定不满足预定拍摄条件，则处理进入步骤 S5521。在步骤 S5521 中，系统控制电路 5009 确认从步骤 S5512 中启动计时器操作起是否经过了预定时间  $t_2[s]$ 。如果在步骤 S5521 中确定已经过了预定时间  $t_2[s]$ ，则处理进入步骤 S5522。可以预先针对各个人或各个面部表情设置预定时间  $t_2[s]$ 。在步骤 S5522 中，系统控制电路 5009 确认是否已执行了拍摄操作。

[0111] 如果在步骤 S5522 中完全没有执行拍摄操作，则系统控制电路 5009 确定当前的面部表情确定阈值不是最优值。处理进入步骤 S5523。在步骤 S5523 中，将面部表情确定阈值从  $Th_3$  变为  $Th_3'$ ，并且处理进入步骤 S5520。

[0112] 如果在步骤 S5501 中用户选择了笑脸拍摄模式，则很可能的是，在步骤 S5509 中开关接通之后被摄物会每隔预定时间笑一次。因此，改变面部表情确定阈值  $Th_3$  为  $Th_3'$ （例如如图 18 所示），计算预定时间内的面部表情评估值  $Exp_2$  的最大值  $Exp_{2max}$  和最小值  $Exp_{2min}$ ，并将最大值  $Exp_{2max}$  设置为面部表情确定阈值  $Th_3'$ 。

[0113] 对于“眼闭合”面部表情，人类通常 3 到 10 秒眨眼一次。因此，将阈值  $Th_3$  设置成：当预定时间  $t_2 = 20[s]$  时，至少两次确定出眼闭合状态。

[0114] 为了进行更精确的面部表情识别，如第一和第二实施例所述，可以改变面部表情评估值计算公式、面部表情评估值计算公式的系数或用于特征提取的阈值。

[0115] 如果在步骤 S5521 中确定从步骤 S5512 中启动计时器操作起还未经过预定时间  $t_2[s]$ ，则处理进入步骤 S5520。相似的是，如果在步骤 S5522 中确定拍摄操作已执行预定次数或更多次，或已执行预定次数或更少次，则处理进入步骤 S5520。

[0116] 如果在步骤 S5520 中已完成自动拍摄，则处理结束。否则，处理进入步骤 S5524。在步骤 S5524 中，针对下一帧图像执行面部检测处理。然后，处理返回步骤 S5514 中的个人身份验证处理。

[0117] 综上所述，当改变面部表情确定阈值时，能够检测甚至仅少量变化了其面部部分（如眼和嘴）的形状的人的面部表情。

[0118] （其他实施例）

[0119] 注意本发明能够应用于包含单个装置的设备或由多个装置构成的系统。

[0120] 再者，本发明能够通过以下方式来实现：向系统或设备直接地或间接地提供实现了上述实施例的功能的软件程序，并且通过所述系统或设备的计算机来读取并执行所提供的程序代码。在此情况下，只要该系统或设备具有本程序的功能，则实施方式不必依靠程序。

[0121] 因此，由于本发明的功能能够由计算机实现，所以安装在计算机上的程序代码也实现了本发明。换句话说，本发明的权利要求书涵盖以实现本发明的功能为目的的计算机

程序。

[0122] 在此情况下,只要所述系统或设备具有所述程序的所述功能,则可以以任何形式实现该程序,例如,目标代码、解译机执行的程序或提供给操作系统的脚本数据。

[0123] 能够用来提供所述程序的记录介质的示例是:软盘、硬盘、光盘、磁光盘、CD-ROM、CD-R、CD-RW、磁带、非易失型存储卡、ROM 和 DVD (DVD-ROM 和 DVD-R)。

[0124] 对于提供所述程序的方法,可以通过使用客户端计算机的浏览器来连接到因特网上的网址,并可以将本发明的计算机程序或该程序的可自动安装的压缩文件下载到诸如硬盘的存储介质上。此外,本发明的所述程序还可以通过以下方式提供:将构成所述程序的程序代码分割成多个文件,并从不同的网址下载这些文件。换句话说,本发明的权利要求还涵盖:给多个用户下载实现本发明的功能的程序文件的万维网 (WWW) 服务器。

[0125] 还可以将本发明的程序加密并存储到诸如 CD-ROM 的记录介质上,然后将该记录介质发给多个用户,允许满足一条要求的用户通过互联网从一网址下载解密密钥信息,并允许这些用户通过使用该密钥信息来将加密的程序解密,从而将该程序安装在用户计算机上。

[0126] 除了通过计算机执行所读取的程序来实现根据所述实施例的上述功能外,在该计算机上运行的操作系统等也可以执行实际处理的部分或者全部,使得能够通过该处理实现上述实施例的功能。

[0127] 此外,在将从记录介质读取的程序写入插入计算机的功能扩展板或者设置在连接到计算机的功能扩展单元中的存储器之后,安装在该功能扩展板或者功能扩展单元中的 CPU 等执行实际处理的部分或者全部,使得通过该处理实现上述实施例的功能。

[0128] 虽然参照示例性实施例对本发明进行了描述,但是应当理解的是,本发明并不限于所公开的示例性实施例。应当对以下权利要求的范围给予最宽泛的解释,以包括所有变体、等同结构和功能在内。

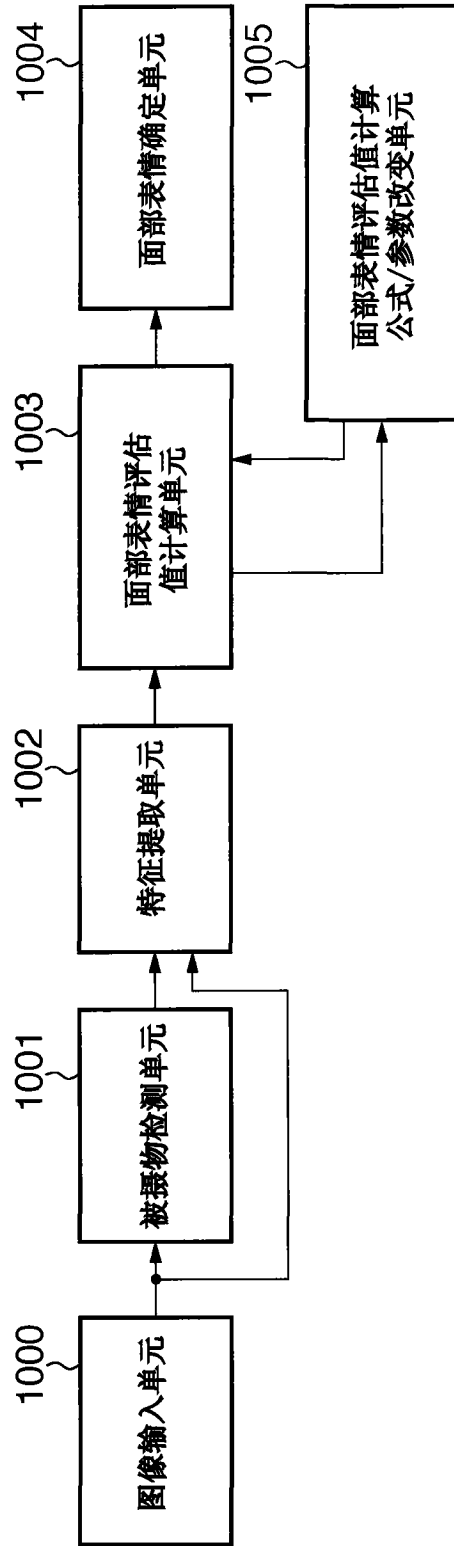


图1

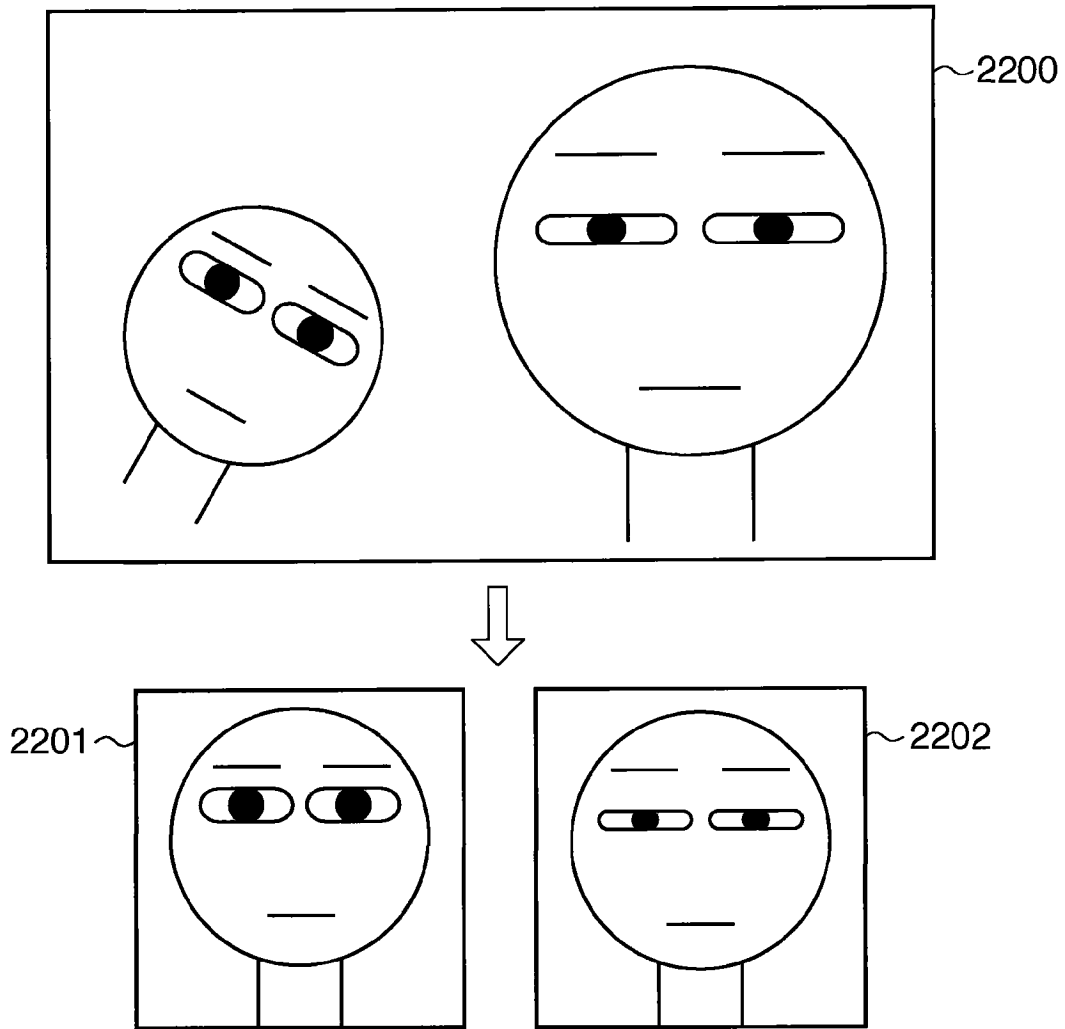


图 2

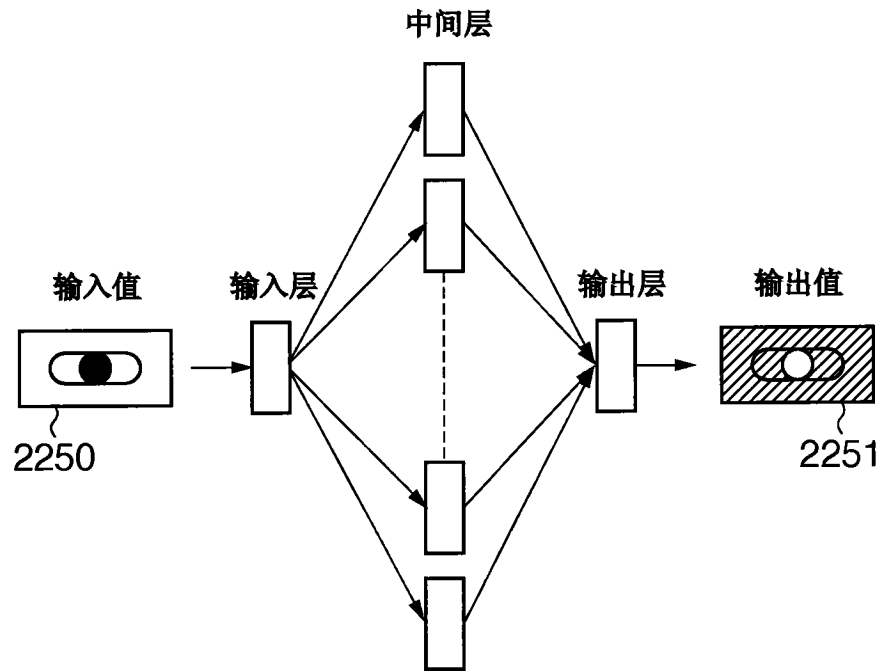


图 3

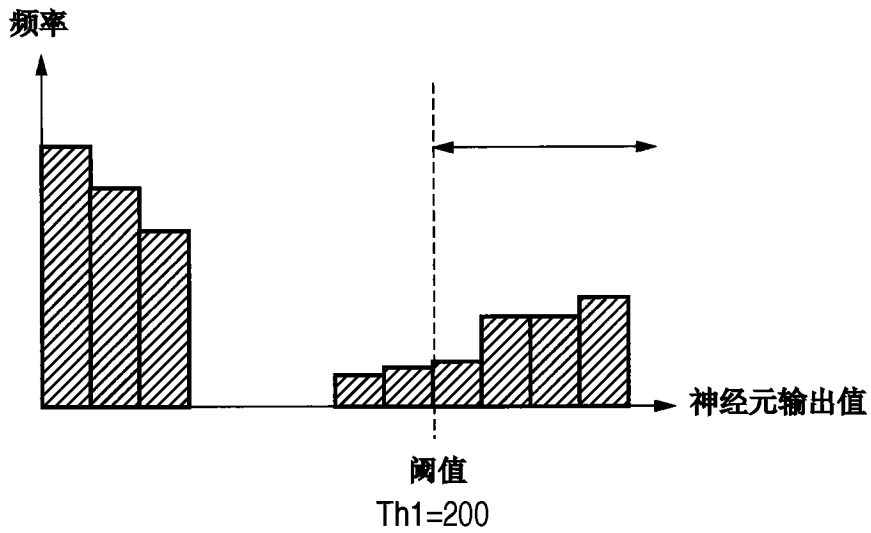


图 4

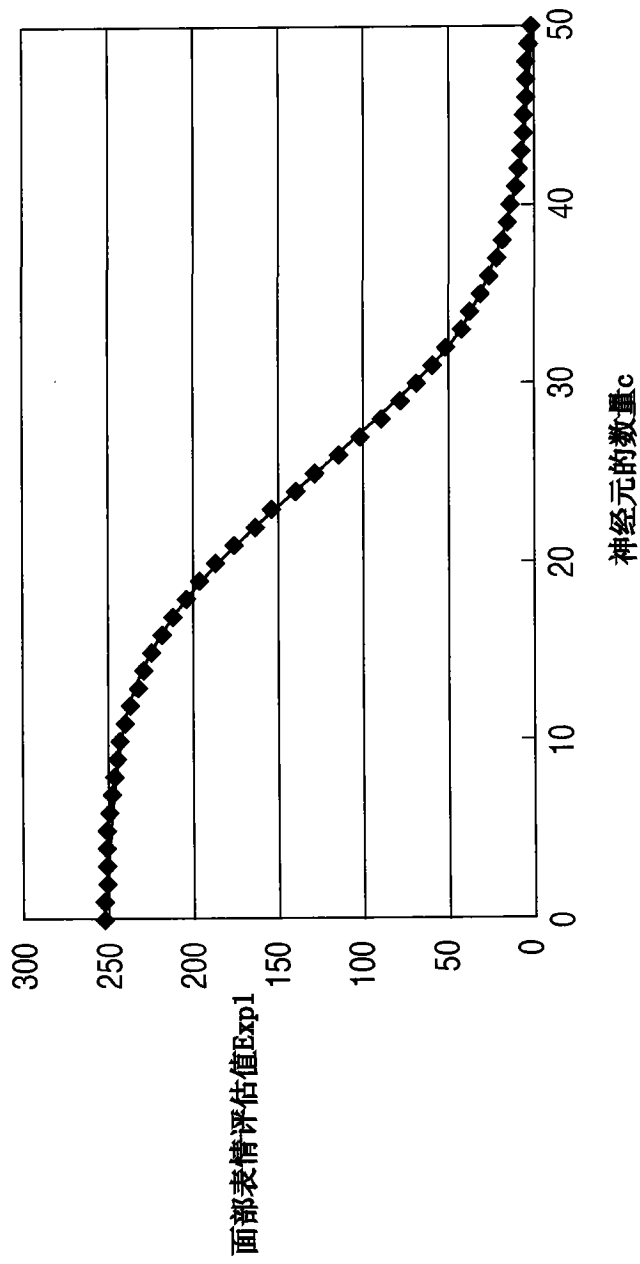


图5

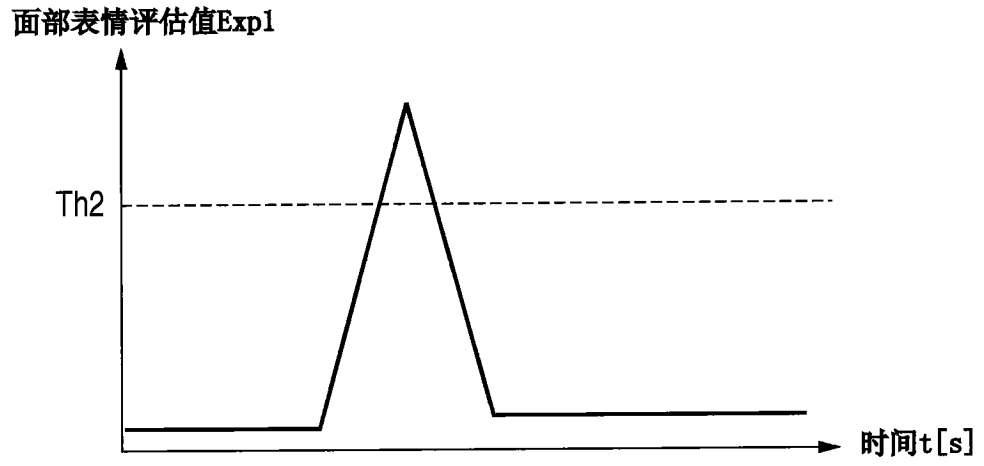


图 6

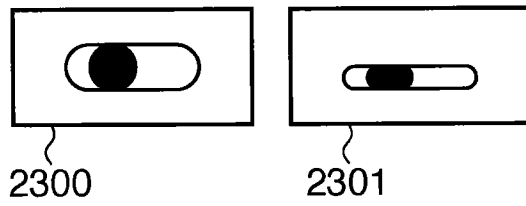


图 7

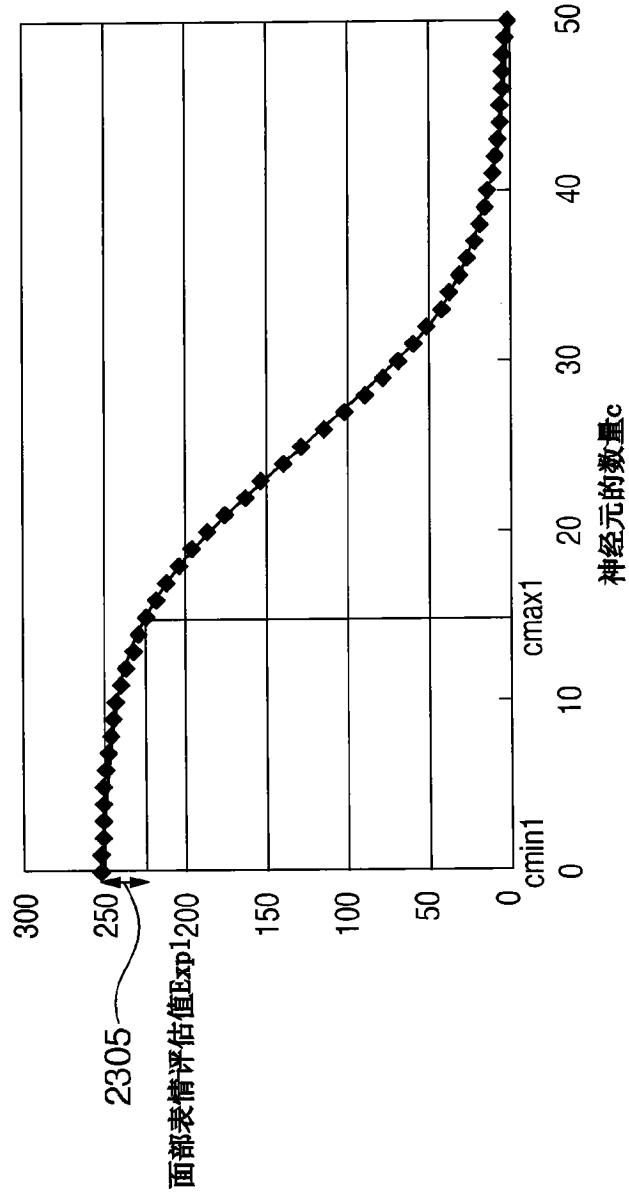


图8

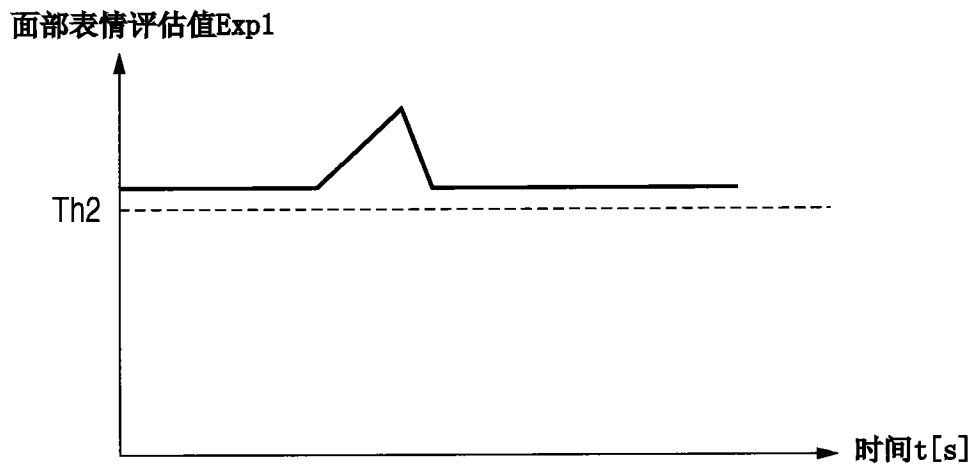


图 9

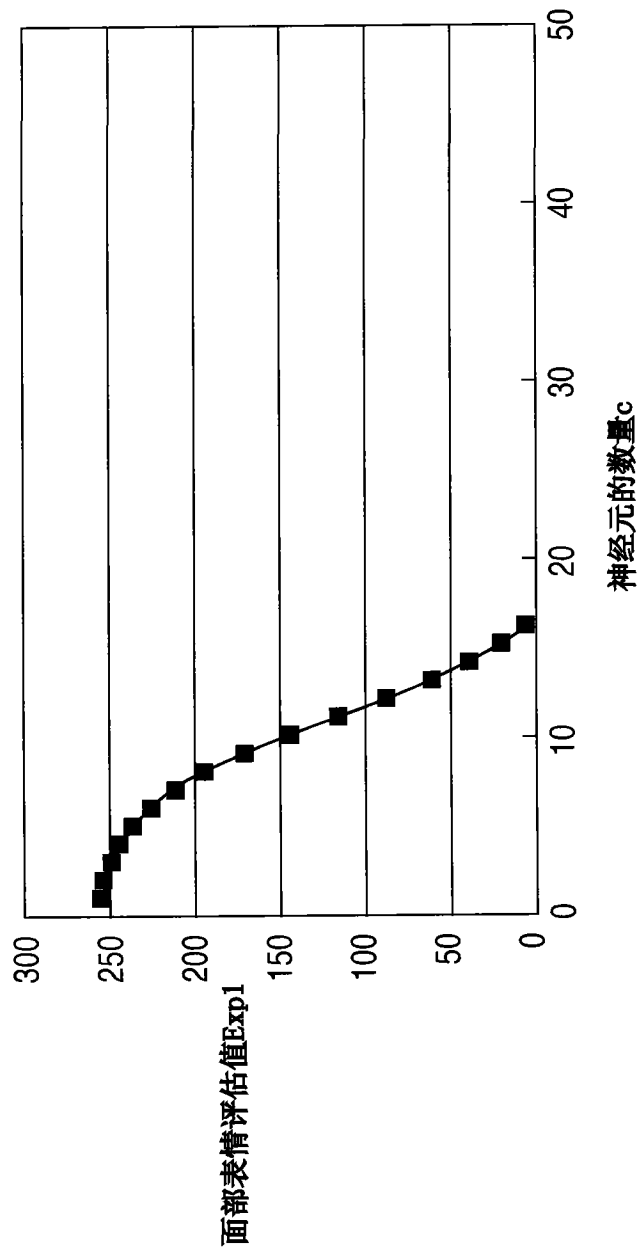


图10

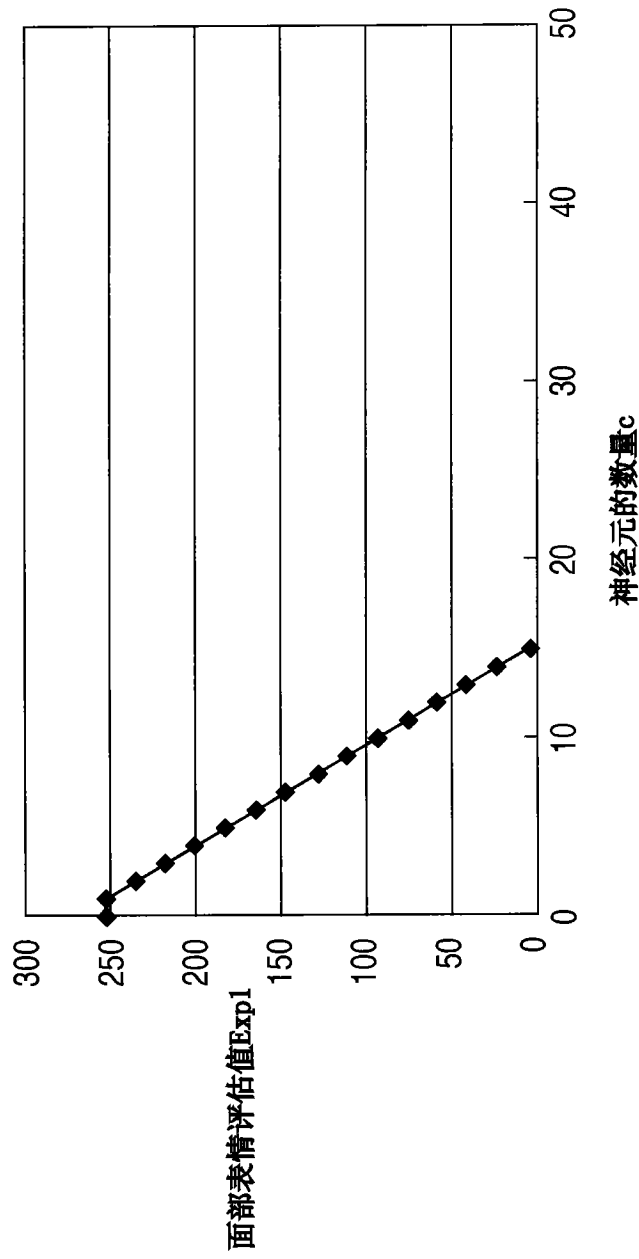


图11

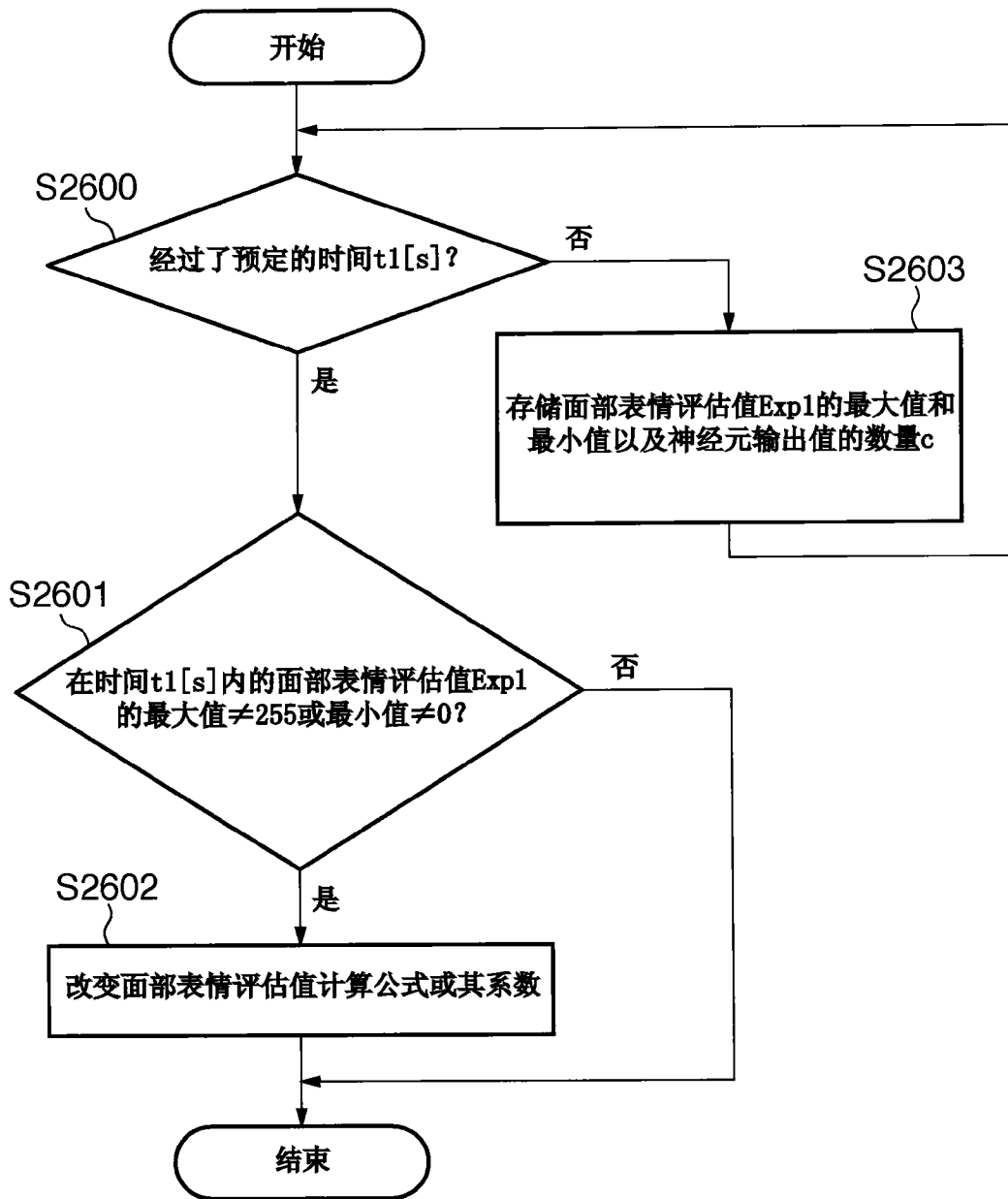


图 12

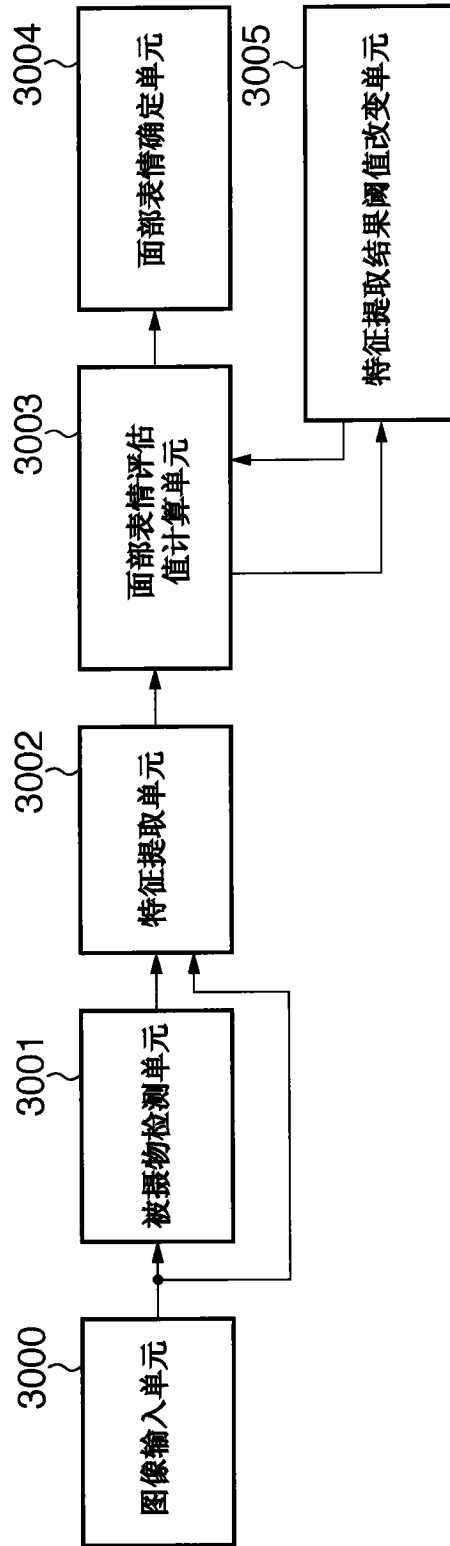


图13

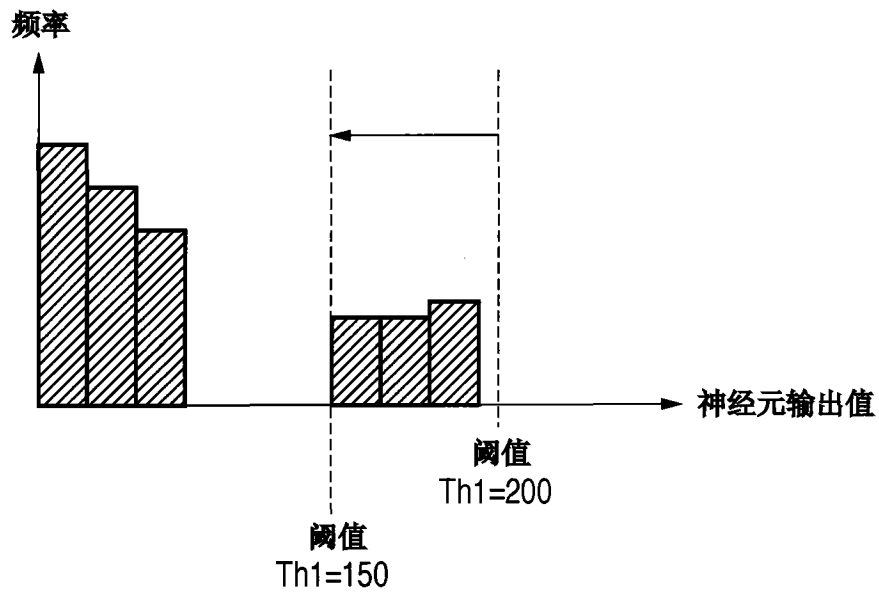


图 14

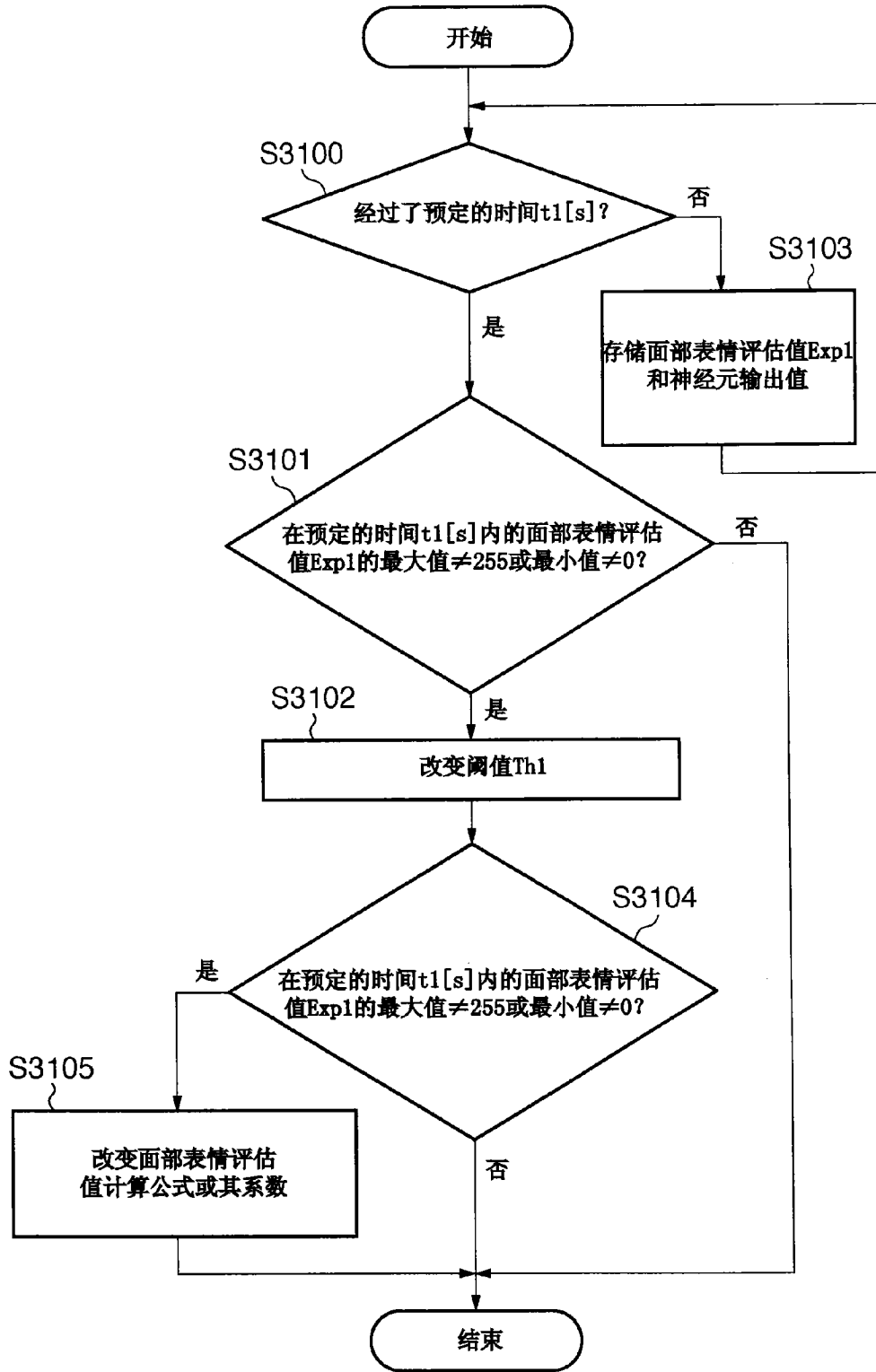


图 15

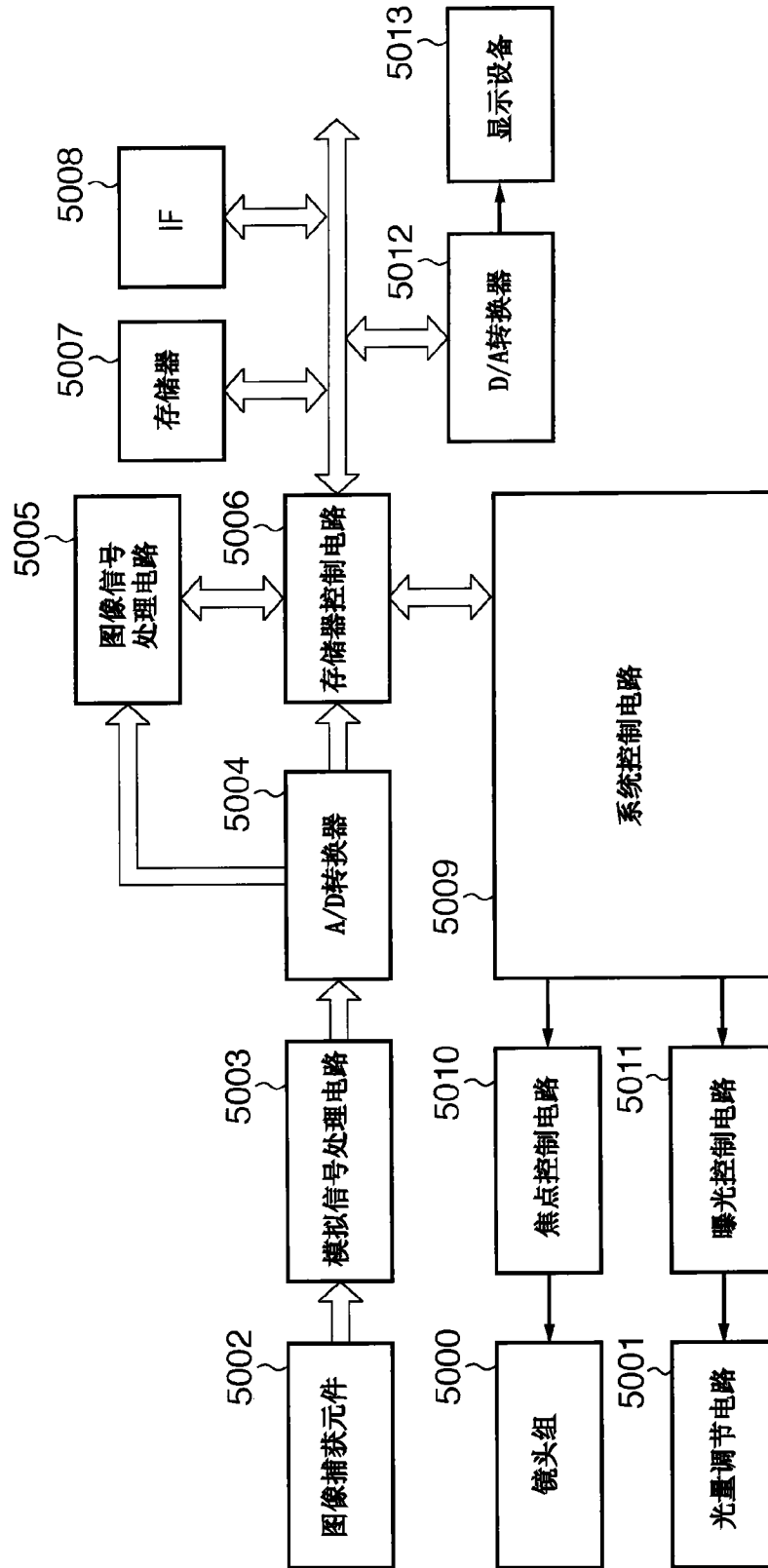


图16

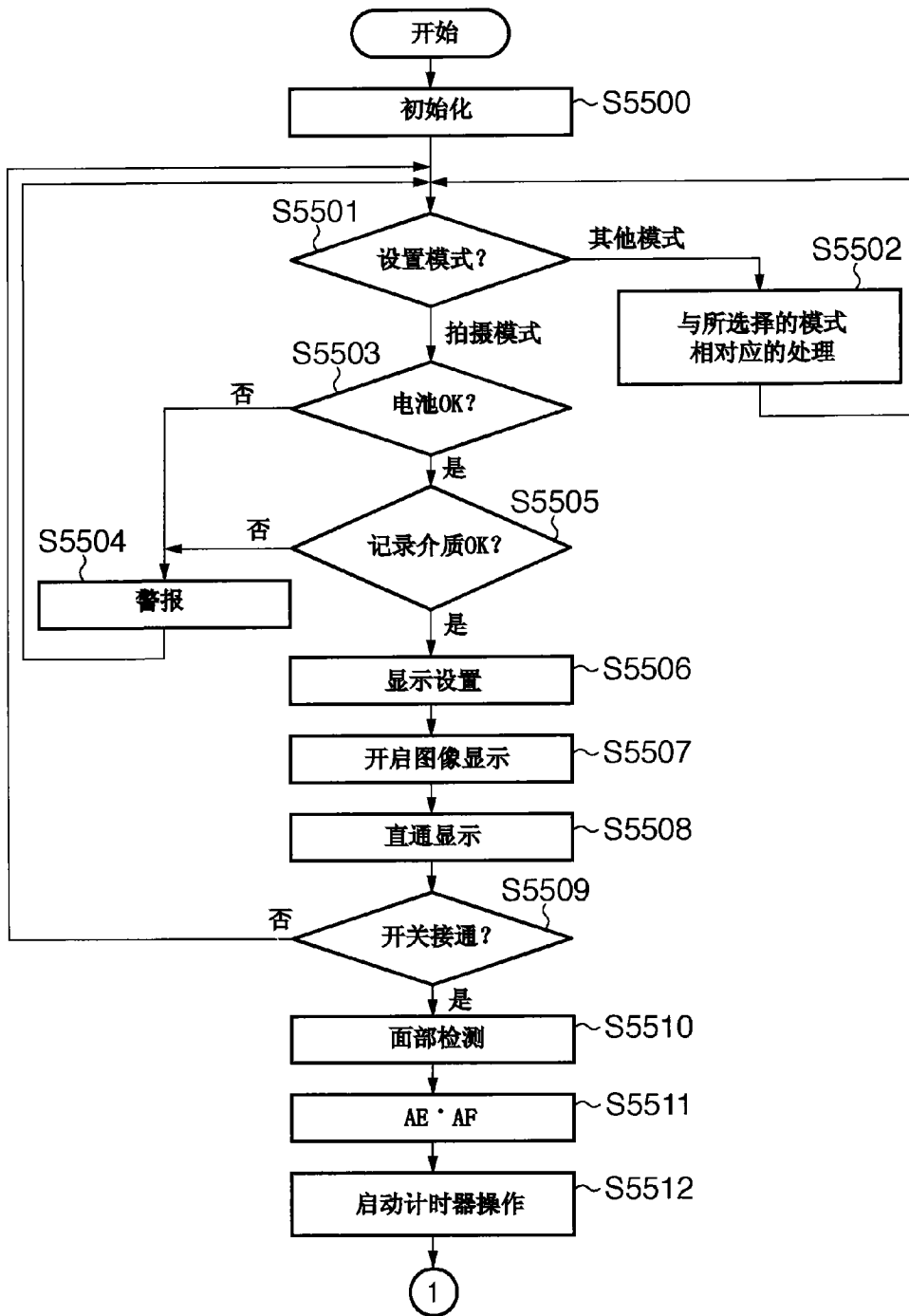


图 17A

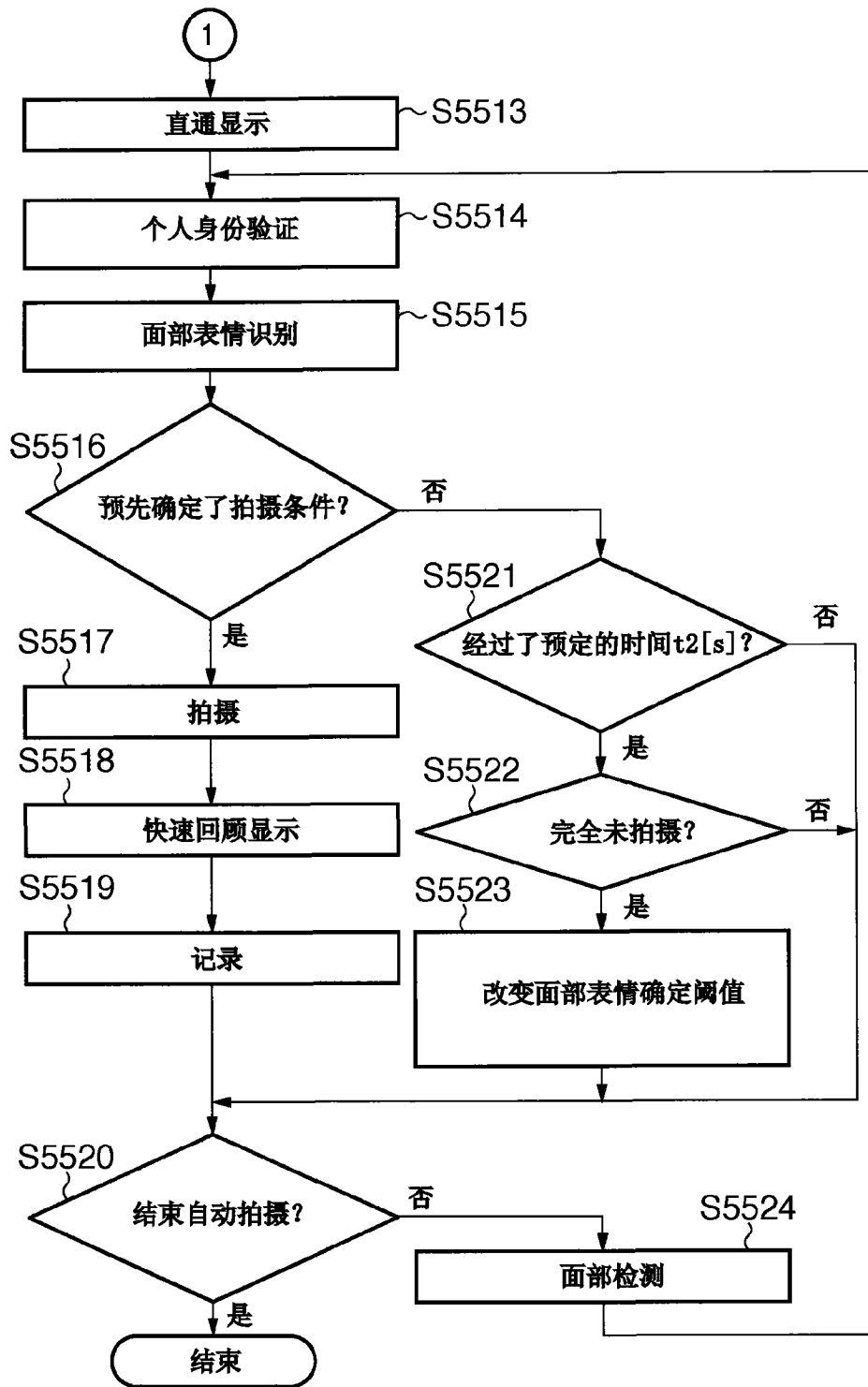


图 17B

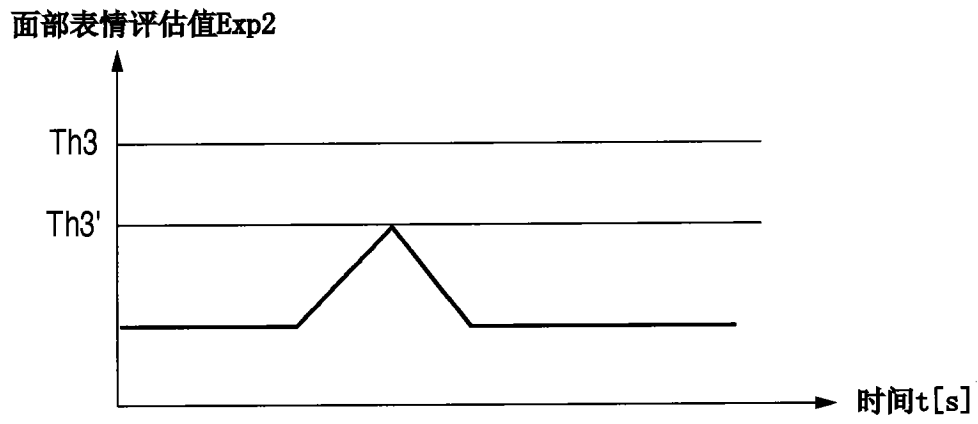


图 18