



(12) 发明专利

(10) 授权公告号 CN 101673346 B

(45) 授权公告日 2013. 06. 05

(21) 申请号 200810215058. 2

CN 1952954 A, 2007. 04. 25, 全文.

(22) 申请日 2008. 09. 09

JP 特开 2008-77536 A, 2008. 04. 03, 全文.

CN 101236599 A, 2008. 08. 06, 全文.

(73) 专利权人 日电(中国)有限公司

地址 100007 北京市东城区东四十条甲 22 号南新仓国际大厦 B 栋 12 层 1222 室

审查员 王平

(72) 发明人 曾炜 张洪明

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

代理人 王波波

(51) Int. Cl.

G06K 9/62(2006. 01)

G06K 9/00(2006. 01)

G06T 5/50(2006. 01)

(56) 对比文件

US 2002/0106114 A1, 2002. 08. 08, 全文.

US 2004/0136574 A1, 2004. 07. 15, 全文.

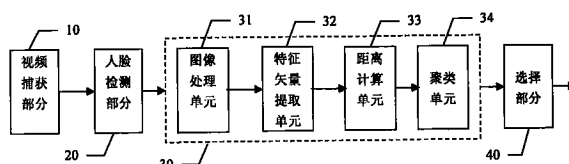
权利要求书2页 说明书5页 附图2页

(54) 发明名称

对图像进行处理的方法、设备和系统

(57) 摘要

提出了一种对图像进行处理的方法, 设备和系统。该图像是由分布在不同位置的多个摄像机对同一场景所捕获的, 所述方法包括如下步骤: 从所捕获的图像中检测人脸图像; 以不同的姿态角度, 用预先确定的人脸姿态模型对检测到的人脸图像中的每一个进行处理, 生成不同姿态角度的合成图像, 作为相应人脸图像的合成图像集; 提取合成图像集中的合成图像的特征矢量; 通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离; 基于不同合成图像集之间的距离对检测的人脸图像进行聚类。该方法、设备和系统可以很容易检测并提取高质量的人脸图像。



1. 一种对图像进行处理的方法,所述图像是由分布在多个位置的多个摄像机对同一场景所捕获的,所述方法包括步骤:

从所捕获的图像中检测人脸图像;

以不同的姿态角度,用预先确定的人脸姿态模型对检测到的人脸图像中的每一个进行处理,生成不同姿态角度的合成图像,作为相应的人脸图像的合成图像集;

提取合成图像集中的合成图像的特征矢量;

通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离;

基于不同合成图像集之间的距离对检测的人脸图像进行聚类;

其中,通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离的步骤包括:

计算不同合成图像集中的合成图像之间的特征矢量距离;

确定所计算的特征矢量距离中最小的特征矢量距离,作为所述不同合成图像集之间的距离。

2. 如权利要求 1 所述的方法,还包括:

从同一聚类的图像中选择两眼之间的距离最大的那幅图像,作为该聚类的代表性图像。

3. 如权利要求 1 所述的方法,还包括:

从同一聚类的图像中选择清晰度最大的那幅图像,作为该聚类的代表性图像。

4. 如权利要求 1 所述的方法,其中所述的特征矢量是 LDA 或 PCA 特征矢量。

5. 如权利要求 1 所述的方法,其中所述基于不同合成图像集之间的距离对检测的人脸图像进行聚类的步骤包括:

将不同合成图像集之间的距离小于预定阈值的人脸图像分为同一聚类。

6. 如权利要求 1 所述的方法,其中,如果不同的人脸图像来自同一摄像机,则将这些人脸图像分为不同类。

7. 如权利要求 1 所述的方法,其中,所述姿态角度的范围是从 -45 度到 +45 度的水平和/或垂直角。

8. 如权利要求 1 所述的方法,其中所述人脸姿态模型是 2D 或 3D 人脸姿态模型。

9. 一种对图像进行处理的设备,所述图像是由分布在多个位置的多个摄像机对同一场景所拍摄的,该设备包括:

检测装置,从所捕获的图像中检测人脸图像;

人脸模型存储装置,存储预先确定的人脸姿态模型;

处理装置,以不同的姿态角度,用人脸模型存储装置中存储的人脸姿态模型对检测到的人脸图像中的每一个进行处理,生成不同姿态角度的合成图像,作为相应人脸图像的合成图像集;

提取装置,提取合成图像集中的合成图像的特征矢量;

距离计算装置,通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离;

聚类装置,基于不同合成图像集之间的距离对检测的人脸图像进行聚类;

其中距离计算装置计算不同合成图像集中的合成图像之间的特征矢量距离,并且确定所计算的特征矢量距离中最小的特征矢量距离,作为所述不同合成图像集之间的距离。

10. 如权利要求 9 所述的设备,还包括:

选择装置,用于从同一聚类的图像中选择两眼之间的距离最大的那幅图像,作为该聚类的代表性图像。

11. 如权利要求 9 所述的设备,还包括:

选择装置,从同一聚类的图像中选择清晰度最大的那幅图像,作为该聚类的代表性图像。

12. 如权利要求 9 所述的设备,其中所述的特征矢量是 LDA 或 PCA 特征矢量。

13. 如权利要求 9 所述的设备,其中所述聚类装置将不同合成图像集之间的距离小于预定阈值的人脸图像分为同一聚类。

14. 如权利要求 9 所述的设备,其中,如果不同的人脸图像来自同一摄像机,则聚类装置将该图像分为不同类。

15. 如权利要求 9 所述的设备,其中,所述姿态角度的范围是从 -45 度到 +45 度的水平和 / 或垂直角。

16. 如权利要求 9 所述的设备,其中所述人脸姿态模型是 2D 或 3D 人脸姿态模型。

对图像进行处理的方法、设备和系统

技术领域

[0001] 本发明涉及人脸图像处理,具体涉及一种对来自多个摄像机的人脸图像进行处理,以便对其进行聚类的方法,设备和系统。

背景技术

[0002] 人脸检测是模式识别研究的热点话题。在过去的几年中,人脸检测技术在检测精度和检测速度方面都得到了长足的发展。人脸检测的目的在于确定图像中是否有包含人脸图像的区域并且自动定位这些人脸图像区域。人脸检测广泛应用于生活的各个领域。例如,近年来,人脸检测已经嵌入到数码相机中,以便使用人脸自动检测技术帮助人们针对人脸获得高质量的图像。另外,在安全应用中,人脸检测技术用来提取人脸图像并且将提取的人脸图像提供给人脸识别器,从而进行人脸的自动分析。

[0003] 另外,基于相机的人脸捕获技术是人脸检测的一个应用。人脸捕获技术首先由相机捕获图像,然后输出带有人脸图像区域的那些图像。这种技术可以用于人脸识别的前端模块或者人脸分析系统的前端模块,以便为后续的分析提供良好的人脸图像数据。

[0004] 人脸图像捕获的目的在于从输入的图像中自动提取人脸图像。通常采用人脸检测技术来定位图像中的人脸区域。由于人脸是 3D 物体,人脸图像实际上是三维人脸物体在二维图像平面上的投影。不同的脸部姿态会从一个 3D 人脸产生多幅人脸图像。因此,人脸捕获技术的任务不但是要检测并且提取正面人脸图像,而且要捕获其他方向的人脸图像,例如侧面图像。

[0005] 目前的人脸捕获技术通常使用单一相机。人脸检测模块首先从相机获得捕获的图像,然后搜索捕获的图像中的各个位置。在每个位置,人脸检测模块首先确定预定尺寸的图像区域,并且判断该区域是否是人脸区域。如果该区域被分类为人脸区域,则该区域被作为候选人脸图像区域来处理。在搜索之后,如果这些人脸区域有重叠部分,则对这些人脸图像进行融合。最终将融合人脸图像区域的位置标记为人脸图像区域。如果仅仅需要人脸图像,则从这些人脸图像区域取出样本区域即可。

[0006] 基于单一相机的人脸检测的缺点在于人脸检测模块对于正面人脸图像能够取得较高的检测精度,而对于非正面人脸图像,则精度较差。另一个问题在于,只有当人们正对着相机时才能够捕获正面图像,而实际的情况并不能总满足这一点。因此需要人脸检测系统尽可能地获得人脸的正面图像。换言之,在基于单一相机的人脸检测技术中,难以避免由于被拍摄者的人脸姿态所造成的问题,原因在于人们并不总是正对着相机镜头。如果人们没有正对着相机的镜头,则捕获的人脸图像必然是非正面图像或者侧面图像。

[0007] 在基于多相机的人脸捕获技术中,非常容易获得人脸的正面图像,这是由于有多个相机对着人们,使得捕获人脸的正面图像的机会增大。在基于多相机的人脸检测技术中,多个相机在同一时刻会捕获相同人脸的不同方向的图像,这就产生了需要将来自不同摄像机的人脸图像进行聚类的问题。

发明内容

[0008] 本发明的图像处理技术基于多相机。由于多个相机在同一时刻捕获了相同人脸的不同图像,本发明的技术意欲提供一种聚类方法,以便将这些图像对应于不同的人而聚类。在本发明的实施例中,利用图像之间的距离来测量不同姿态的图像之间的相似度,从而免去了要进行人脸姿态估计所引起的复杂计算量问题。换言之,本发明的图像处理技术可以有效地捕获人脸图像,而不存在姿态估计问题,并且能够输出高质量的人脸图像。

[0009] 在本发明的一个方面,提出了一种对图像进行处理的方法,所述图像是由分布在不同位置的多个摄像机对同一场景所捕获的,所述方法包括步骤:从所捕获的图像中检测人脸图像;以不同的姿态角度,用预先确定的人脸姿态模型对检测到的人脸图像中的每一个进行处理,生成不同姿态角度的合成图像,作为相应人脸图像的合成图像集;提取合成图像集中的合成图像的特征矢量;通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离;基于不同合成图像集之间的距离对检测的人脸图像进行聚类。

[0010] 在本发明的另一方面,提出了一种对图像进行处理的设备,所述图像是由分布在不同位置的多个摄像机对同一场景所拍摄的该设备包括:检测装置,从所捕获的图像中检测人脸图像;人脸模型存储装置,存储预先确定的人脸姿态模型;处理装置,以不同的姿态角度,用人脸模型存储装置中存储的人脸姿态模型对检测到的人脸图像中的每一个进行处理,生成不同姿态角度的合成图像,作为相应人脸图像的合成图像集;提取装置,提取合成图像集中的合成图像的特征矢量;距离计算装置,通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离;聚类装置,基于不同合成图像集之间的距离对检测的人脸图像进行聚类。

[0011] 在本发明的又一方面,提出了一种图像处理系统,包括:分布在不同位置的摄像机,用于针对同一场景捕获图像;检测装置,从所捕获的图像中检测人脸图像;人脸模型存储装置,存储预先确定的人脸姿态模型;处理装置,以不同的姿态角度,用预先确定的人脸姿态模型对检测到的人脸图像中的每一个进行处理,生成不同姿态角度的合成图像,作为相应人脸图像的合成图像集;提取装置,提取合成图像集中的合成图像的特征矢量;距离计算装置,通过计算不同合成图像集中的合成图像之间的特征矢量距离来计算不同合成图像集之间的距离;聚类装置,基于不同合成图像集之间的距离对检测的人脸图像进行聚类。

[0012] 利用本发明实施例的技术,可以很容易检测并且提取高质量的人脸图像。由于使用了布置在不同位置的多个相机,本发明实施例的技术可以解决人脸姿态问题。人脸姿态空间被多个相机划分成多个子人脸姿态空间。每个摄像机之间的人脸姿态改变程度很小。

[0013] 本发明实施例还采用了有效的人脸距离来进行人脸图像聚类,从而更为鲁棒并且计算量更小。

附图说明

[0014] 从下面结合附图的详细描述中,本发明的上述特征和优点将更加明显,其中:

[0015] 图 1 示出了根据本发明实施例的图像处理系统的结构示意图;

[0016] 图 2 示出了根据本发明实施例的图像处理方法的流程图;

[0017] 图 3 是对人脸图像进行处理所用的姿态角的示意图;以及

[0018] 图 4 是进行距离计算所用的距离矩阵的示意图。

具体实施方式

[0019] 下面,参考附图详细说明本发明的优选实施方式。为了清楚和简明,包含在这里的已知的功能和结构的详细描述将被省略,以防止它们使本发明的主题不清楚。

[0020] 图 1 示出了根据本发明实施例的图像处理系统的结构示意图。如图 1 所示,根据本发明实施例的图像处理系统包括视频捕获部分 10,人脸检测部分 20,人脸聚类部分 30 和选择部分 40。根据本发明的实施例,人脸聚类部分 30 包括图像处理单元 31,特征矢量提取单元 32,距离计算单元 33 和聚类单元 34。

[0021] 视频捕获部分 10 例如是设置在不同位置的多个摄像机,针对同一场景,例如大厦的门口进行拍摄,并且将捕获的视频信号转换成数字图像数据。然后捕获的图像被送入人脸检测部分 20。在人脸检测部分 20,定位捕获的图像中包含人脸的区域的位置,并且基于这些位置从图像中提取人脸图像。然后,在人脸聚类部分 30 中对来自不同摄像机的各个人脸图像进行聚类,形成针对不同人的图像组。最后,在选择部分 40 中,基于预定的准则,例如清晰度或者两眼之间的距离,从聚类后的图像组中选为每一类选择相应的代表性图像,作为输出。

[0022] 在人脸聚类部分 30 中,图像处理单元 31 以事先存储在人脸模型存储器(未示出)中的 3D 姿态模型或者 2D 姿态模型对各个人脸图像进行处理,生成各个姿态角度的合成人脸图像,作为针对个人脸图像的合成图像集。然后,特征矢量提取单元 32 提取各个合成图像的 LDA 或者 PCA 矢量。由距离计算单元 33 计算不同合成图像集的合成人脸图像之间的距离,并且将最小的距离作为两个集合之间的距离。接下来,聚类单元 34 基于合成图像集之间的距离对人脸图像进行聚类,产生针对不同人的图像组。

[0023] 接下来,如上所述,选择部分 40 基于预定的准则,例如清晰度或者两眼之间的距离,从聚类后的图像组中选为每一类选择相应的代表性图像,作为输出。

[0024] 下面结合如图 2~4 详细说明上述各个部分的详细构成和操作过程。图 2 示出了根据本发明实施例的图像处理方法的流程图。

[0025] 在本发明的实施例中,采用多个相机一起协作的形式来捕获图像,这些相机布置成针对同一目标进行拍摄,例如大厦的门口。换言之,这些相机中的至少部分具有公共的视场。

[0026] 在步骤 S11,包括摄像头和视频采集卡的视频捕获部分 10 针对同一场景产生视频信号,对视频信号进行采样并且将其转换成数字视频图像。产生的数字视频图像存储在系统的缓冲存储器(未示出)中。根据本发明的实施例,图像的格式可以是 PAL 或者 NTSC 或者按照用户的需要而确定。同样,图像的尺寸同样可以是事先确定的或者是按照用户的需求而确定的。

[0027] 在步骤 S12,人脸检测部分采用各种检测器,例如非专利文献 1(Ming-Hsuan Yan, David J. Kriegman, and Narendra Ahuja. Detecting Faces in Images :A Survey. IEEE Transactions On Pattern analysis and Machine Intelligence, Vol. 24, No. 1, pp. 34-58, 2002) 中描述的各种人脸检测器,或者非专利文献 2(Paul A. Viola, Michael J. Jones :Rapid Object Detection using a Boosted Cascade of Simple Features.

In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2001), Vol. 1, pp. 511-518, Kauai, HI, USA, 8-14 December 2001) 提出的分类器来检测捕获图像中的人脸区域。例如首先对用带人脸的图像和不带人脸的图像对分类器进行训练, 然后将其应用于所感兴趣的区域的检测中。如果某个区域包含了人脸图像, 则分类器输出 '1', 否则输出 '0'。这样, 分类器以不同的尺度搜索图像中的所有位置后找到图像中的人脸区域。

[0028] 在步骤 S13, 以不同的姿态角度对人脸图像进行处理, 得到合成图像集。如上所述, 在基于多相机的应用中, 一个人被多个相机捕获得到不同姿态角度的人脸图像, 这是由于相机处于不同的位置和方位。因此聚类的过程可以看出是将同一人的不同姿态的人脸图像进行分类的过程。

[0029] 通常, 来自同一人的两幅相同姿态的人脸图像之间的相似度要大于不同姿态的图像之间的相似度。来自不同人的两幅相同姿态的人脸图像之间的相似度要小于来自同一人的两幅相同姿态的人脸图像之间的相似度。因此相同姿态的人脸图像将很容易被聚为一类。根据本发明的实施例, 使用非专利文献 3 (W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys, Vol. 35, Issue 4, pp. 399-458, December 2003) 所述的诸如 LDA 或者 PCA 之类的特征矢量之间的距离来表示人脸图像之间的相似度。根据本发明的实施例, 以不同的姿态角度对人脸图像进行处理, 例如渲染, 产生针对不同姿态角度, 例如水平和 / 或垂直 -45 度到 +45 度, 的合成人脸图像, 作为该人脸图像的合成图像集。图 3 是各个人脸姿态角度的例子。

[0030] 在步骤 S14, 提取合成图像集中的各个图像的特征矢量。在步骤 S15, 通过计算特征矢量之间的距离来确定合成图像集之间的距离。

[0031] 例如, 计算不同图像集的各个合成图像之间的 LDA 或 PCA 特征矢量距离, 然后计算这些距离中的最小距离, 作为合成图像集之间的距离。下面详细描述该过程。

[0032] 给定两幅人脸图像 f_i 和 f_j , 通过用不同的姿态角度对两幅人脸图像进行渲染来得到相应的合成人脸图像集 F_i 和 F_j 。合成人脸图像集表示如下:

[0033] $F_i = \{f_i(-n\theta), \dots, f_i(-\theta), f_i(0), f_i(\theta), \dots, f_i(n\theta)\}$,

[0034] $F_j = \{f_j(-n\theta), \dots, f_j(-\theta), f_j(0), f_j(\theta), \dots, f_j(n\theta)\}$ 。

[0035] $f_i(k\theta)$ 和 $f_j(k\theta)$ 是使用某个姿态角度 $k\theta$ 的模型渲染后得到的合成人脸图像, θ 是预定的单位人脸姿态角度, k 是 $-n$ 到 n 的整数变量。对于每对人脸图像 $f_i(k\theta)$ 和 $f_j(k\theta)$, 可以通过 LDA 或 PCA 特征矢量的特征矢量距离获得距离矩阵。图 4 示出了这样的一个距离矩阵的例子。在计算得到了特征矩阵之后, 将人脸图像 f_i 和 f_j 之间的最小人脸姿态距离 (MFPD) 定义为距离矩阵之间的最小距离。MFPD 表示如下:

[0036] $MFPD(f_i, f_j) = \min_{k_i, k_j \in [-n, n]} d(f_i(k_i\theta), f_j(k_j\theta))$ 。

[0037] 在步骤 S16, 基于图像集之间的距离对人脸图像进行聚类。一旦得到了两幅人脸图像之间的距离, 则可以使用基于约束的分级聚类方法可以被用来对人脸图像进行聚类。这里的约束是如果不同的人脸图像来自同一摄像机, 则将这些人脸图像分为不同类。例如, 将距离小于预定的阈值的两类合并成同一类, 直到没有可合并的类别。两类 C_i 和 C_j 之间的距离定义如下:

[0038]
$$D(C_i, C_j) = \min_{\substack{f \in C_i \\ f' \in C_j}} MFPD(f, f').$$

[0039] 两类之间的距离还可以采用最大或者平均 MFPD 距离。

[0040] 在步骤 S17, 在聚类之后, 根据预定的准则从人脸图像中选择一幅图像作为该类图像的代表性图像。例如选择两眼之间的距离最大的那幅图像或者清晰度最大的那幅图像作为该类的代表性图像

[0041] 上面的描述仅用于实现本发明的实施方式, 本领域的技术人员应该理解, 在不脱离本发明的范围的任何修改或局部替换, 均应该属于本发明的权利要求来限定的范围, 因此, 本发明的保护范围应该以权利要求书的保护范围为准。

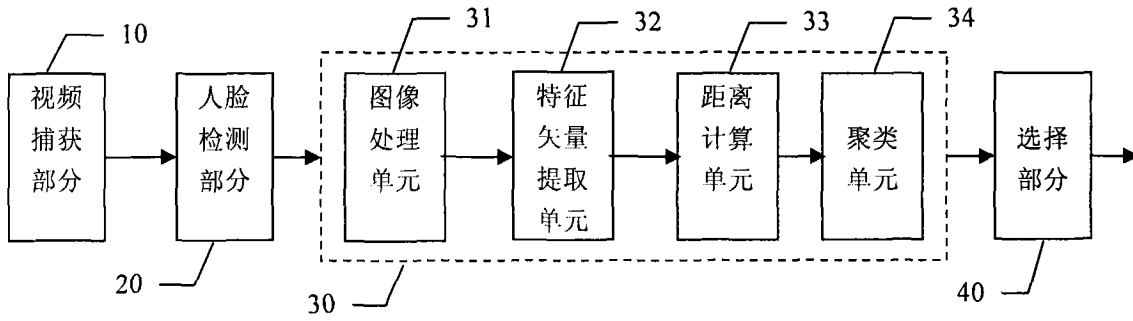


图 1

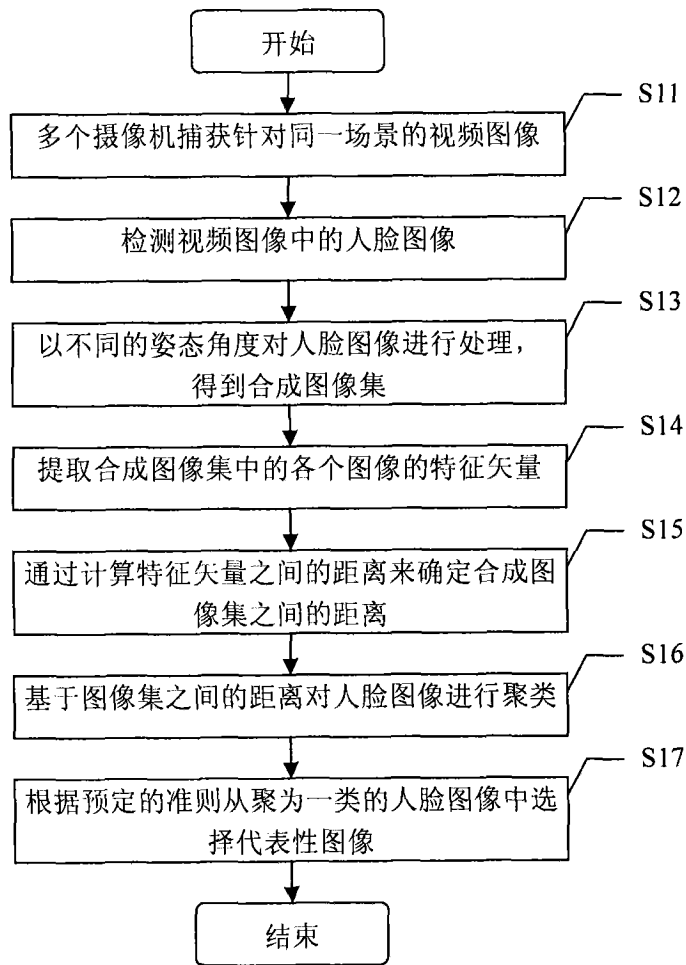


图 2

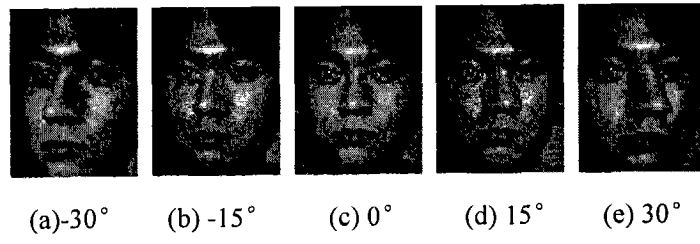


图 3

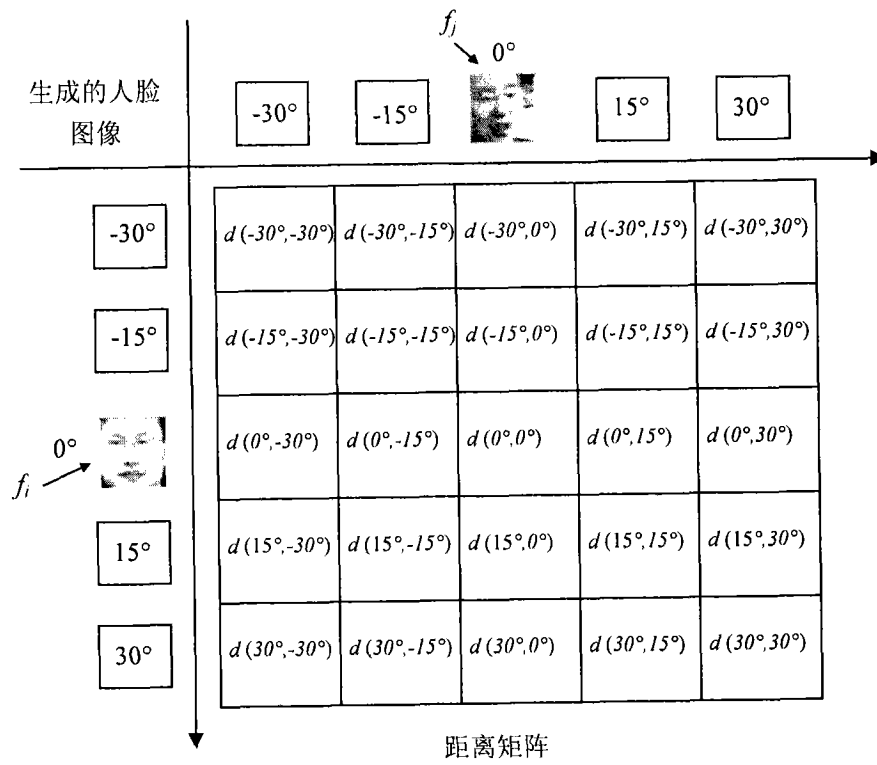


图 4