

[12] 发 明 专 利 说 明 书

[21] ZL 专利号 97103439.7

[45]授权公告日 2002 年 10 月 23 日

[11]授权公告号 CN 1093290C

[22]申请日 1997.3.10 [21]申请号 97103439.7

[30]优先权

[32]1996.6.6 [33]JP [31]144109

[73]专利权人 富士通株式会社

地址 日本神奈川

[72]发明人 藤本克仁 镰田洋

审查员 齐 霁

[74]专利代理机构 中国国际贸易促进委员会专利商标事务所

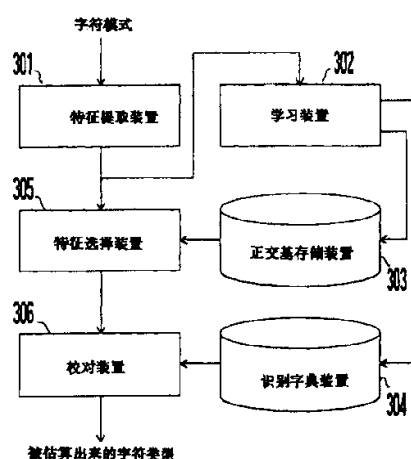
代理人 王以平

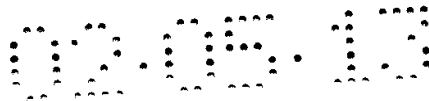
权利要求书 4 页 说明书 15 页 附图 10 页

[54]发明名称 利用部分特征空间上投影特征向量的模式识别装置和方法

[57]摘要

特征提取装置从输入字符模式中提取输入特征向量。学习装置计算并存储原始特征空间的部分本征空间的规范正交基。学习装置计算并存储对应每个目标字符模式在部分本征空间中规定的每个字典选择特征向量。特征选择装置利用上述的规范正交基,通过将由特征提取装置所提取的输入选择特征向量投射在部分本征空间上来计算输入选择特征向量。校对装置校对上述的输入选择特征向量,估算该输入字符模式的类型。





权 利 要 求 书

1.利用代表输入模式特征的特征向量识别某个输入模式的模式识别装置,包括:

一个输入特征向量提取单元,用于提取代表该输入模式的特征的输入特征向量;

一个规范正交基存储单元,用于存放原始特征空间的某个部分特征空间的规范正交基;

一个识别字典单元,存放对应将被识别的一个或多个目标模式的部分特征空间中所规定的每个字典选择特征向量;

一个特征选择单元,利用存放在所述规范正交基存储单元中的规范正交基,将由所述输入特征向量提取单元所提取的输入特征向量投影在部分特征空间上,由此计算出输入选择特征向量;以及

一个校对单元,通过用存放在所述识别字典单元中的每个字典选择特征向量校对由所述特征选择单元计算出来的输入选择特征向量来识别对应输入选择特征向量的输入模式的类型。

2.根据权利要求1的模式识别装置,进一步包括:

一个模式学习单元,用于计算规范正交基和每个字典选择特征向量,其中所述模式学习单元包括:

一个学习特征向量提取单元,用于提取代表学习模式的特征的学习特征向量;

一个规范正交基计算单元,根据由所述学习特征向量提取单元所提取的多个学习特征向量计算部分特征空间的规范正交基;

一个字典特征向量计算单元,根据所述学习特征向量提取单元所提取的多个学习特征向量,计算对应将被识别的每个目标模式的原始特征空间中所规定的每个字典特征向量;以及

一个字典选择特征向量计算单元,利用由所述规范正交基计算单元计算出来的规范正交基,通过将由所述字典特征向量计算单元计算出的每个字典特征向量投射到该部分特征空间上、计算出每个字典选择特征

向量。

3.根据权利要求 1 的模式识别装置，其中，所述的规范正交基是这样计算出来的：

通过对多个学习特征向量执行典型确认分析，从具有最大特征值的一个开始，顺序计算出比原始特征空间的维数少的预定数目的特征向量；并且

通过从具有最大的特征值的一个开始顺序地规范正交化预定的特征向量计算出预定数目的规范正交基向量。

4.根据权利要求 1 的模式识别装置，其中，将输入特征向量或每个字典特征向量投射在部分特征空间上得到的所述输入选择特征向量或每个所述字典选择特征向量是通过将输入特征向量或每个字典特征向量投射在规范正交基向量上得到的。

5.根据权利要求 4 的模式识别装置，其中，将输入特征向量或每个字典特征向量投射在规范正交基向量上得到的所述输入选择特征向量或每个所述字典选择特征向量，是做为计算输入特征向量或每个字典特征向量和规范正交基向量的内积的结果。

6.根据权利要求 1 的模式识别装置，其中，所述字典选择特征向量基于表示每类将被识别的目标模式的特征向量的平均值的平均特征向量。

7.根据权利要求 1 的模式识别装置，其中通过计算输入选择特征向量和每个字典选择特征向量之间的距离并相互比较得到的距离值，用每个字典选择特征向量校对所述输入选择特征向量。

8.根据权利要求 7 的模式识别装置，其中，所述距离基于某个欧几里德距离。

9.根据权利要求 7 的模式识别装置，其中所述距离基于市街区距离。

10.根据权利要求 7 的模式识别装置，其中所述距离基于棋盘距离。

11.根据权利要求 1 的模式识别装置，其中所述模式为字符模式。

12.利用代表输入模式的特征的输入特征向量识别输入模式的模式识别方法，包括步骤：

提取代表输入模式的特征的输入特征向量；
 存储原始特征空间的部分特征空间的规范正交基；
 存储对应将被识别的一个或多个目标模式的部分特征空间中规定的每个字典选择特征向量；
 利用规范正交基将对应输入模式的被提取的输入特征向量投射到部分特征空间计算出输入选择特征向量；并且
 通过用每个字典选择特征向量校对输入选择特征向量，识别对应输入选择特征向量的输入模式的类型。

13.根据权利要求 12 的模式识别方法，进一步包括计算规范正交基和每个字典选择特征向量的步骤；

其中所述步骤包括：

提取代表某个学习模式的特征的学习特征向量；
 根据多个被提取的学习特征向量计算部分特征空间的规范正交基；
 利用多个学习特征向量，计算对应每个将被识别的目标模式在原始特征空间中指定的每个字典特征向量；并且
 利用规范正交基，通过将每个字典特征向量投射在部分特征空间上，计算每个字典选择特征向量。

14.利用代表输入模式的特征的输入特征向量识别输入模式的模式识别装置，包括：

一个特征向量提取单元，用于提取代表输入模式特征的输入特征向量或代表学习模式特征的学习特征向量；

一个规范正交基存储单元，存储原始特征空间的部分特征空间的规范正交基；

一个识别字典单元，存放对应将被识别的一个或多个目标模式在部分特征空间中规定的每个字典选择特征向量；

一个规范正交基计算单元，根据由所述特征向量提取单元所提取的多个学习特征向量计算部分特征空间的规范正交基，并且将计算结果存放在所述规范正交基存储单元中；

一个字典特征向量计算单元，根据由所述特征向量提取单元所提取

的多个学习特征向量，计算对应将被识别的每个目标模式在原始特征空间中规定的每个字典特征向量；

一个字典选择特征向量计算单元，利用由所述规范正交基计算单元计算出来的规范正交基，通过将由所述字典特征向量计算单元计算出来的每个字典特征向量投射到部分特征空间上来计算每个字典选择特征向量，并且将计算结果存放在所述识别字典单元中；

一个特征选择单元，利用存放在所述规范正交基存储单元中规范正交基，将由所述特征向量提取单元提取的输入模式的输入特征向量投射在部分特征空间上，计算出输入选择特征向量；

一个校对单元，通过用由所述识别字典单元存储的每个字典选择特征向量校对由所述特征选择单元计算出来的输入选择特征向量，识别对应输入选择特征向量的输入模式的类型。

利用部分特征空间上投影特征向量的模式识别装置和方法

本发明涉及识别字符等的模式的技术，通过有效地减少指示模式的特征向量的维数来选择特征。

最近，文档识别技术，即采用电子技术手段汇存文档，有效地执行办公室的工作流，并且当需要时对数据编码，这样的一种技术正被认真地提出需求。具体来说，字符识别技术，即文档识别技术的一个方面，对于编码字符串信息来说是十分重要的。在这种技术中，要求采用按所需的识别精度快速估计字符类型的方法，使这种字符识别技术能在各领域中投入实际应用。利用统计技术减少输入字符的字符向量的维数来选择特征的方法对于减少用校对识别字典进行校对的计算量是很有效的。因此，当生产实用的文档识别装置时，使用特征选择法的字符识别技术将发挥重要的作用。对于生产用于识别非字符的各种模式的设备来说，这也是一种十分重要的技术。

下面首先介绍字符识别的一般概念，以字符做为模式的例子。

首先，当某个字符模式被输入时，其尺寸将被规范化。

接着，将规范化得到的矩形字符区域划分为形成该矩形字符区域的多个块。例如，一个矩形字符区域可相等地划分为3行3列的9块或者6行6列的36个块。

接着，存在于每块中的指示字符轮廓的象素（轮廓象素）被提取。对于每个象素，确定包含象素的轮廓的方向。轮廓象素为直接对应某个字符区域的象素，或者也可以通过处理细线处理中的字符区域得到。上述方向可以是8个方向（上、下、左、右和4个对角线）中的一个，或者是进一步细化的36个方向中的一个。接着，对每个块中的每个方向得到轮廓象素的数目。结果，对每个块可得到某个部分特征向量，其维数等于方向数目，其元素的值对应在该元素的方向上的轮廓象素的数

目。对于每个输入字符模式，可以得到包括对应被包含在对应的矩形字符区域中的每个块的每个部分特征向量的所有元素的特征向量。

如果对于每个这样得到的字符模式的特征向量被分类为目标字符的类型，则对每种字符类型形成一个群集（聚类），即对在所具有的维数对应该特征向量的元素数目的多维空间中相同类字符类型的字符模式的特征向量进行分组。基于这样的特征，对学习字符模式的特征向量进行分类，并且从被包含在该群集中的特征向量中计算出表示对应每个结果群集的字符类型的平均特征向量。平均特征向量是通过计算特征向量的每个元素的平均值得到的。每个字符类型的平均特征向量被存在字典中。

当实际识别某个字符时，输入字符模式的特向向量被计算出来，然后再计算特征向量和进入字典的每个平均特征向量之间的距离。这样，对应指示最短距离的平均特征向量的字符类型被估计为该输入字符的类型。距离可以是欧几里德距离或市街区距离。

当对整个特征空间计算距离时，需要对应该特征空间维数的时间长度。高速执行这样一种计算任务的一种众所周知的方法是在减少特征空间的维数（例如，从 384 维减少到 64 维）后再计算距离。在特征空间中减少维数的方法被称为特征选择。用于特征选择的一种具体方法可以是典型确认分析或主要部分分析。经实验已确认，通过特征选择，仅对原始维数的 1/8，识别率几乎没有什么降低。反之，在特征选择中，通过减少不需要的噪声特征向量，也能够提高识别率。

下面首先结合图 1 所示的配置介绍通过典型确认分析进行特征选择的识别字符的常规技术。

如上所述，特征提取装置 101 从被包含在 9 个类型的每个字符类型之 ($1 \leq i \leq g$) 中的 n_i 个字符模式的示例中提取由下面等式 (1) 所代表的 n_i 组学习特征向量 $x_j^{(i)}$ ($1 \leq j \leq n_i$)。上标 T 表示矩阵（或向量）的转置。

$$x_j^{(i)} = (x_{jk}^{(i)}) = (x_{j1}^{(i)}, \dots, x_{jN}^{(i)})^T \quad \dots (1)$$

下标 K 表示特征向量在范围 $1 \leq K \leq N$ 内的元素数目。

在下面的描述中，带下划线的符号表示向量的数量，具有元素个数

而没有如上所述的下划线的符号则表示向量的一个元素的值。

学习装置 102 计算由下面等式 (2) 表示的平均特征向量 $\underline{m}^{(i)}$, 对应和字符类型之相应的上述特征向量 $\underline{X}_j^{(i)}$ ($1 \leq j \leq n_i$) 中的 g 个类型中的每个字行类型 i ($1 \leq i \leq g$)。

$$\underline{m}^{(i)} = (m_k^{(i)}) = (m_1^{(i)}, \dots, m_N^{(i)})^T \quad \dots\dots (2)$$

学习装置 102 从每个字符类型 i 的 n_i 个示例和上述的平均特征向量 $\underline{m}^{(i)}$ 中为所有的字符类型计算由下式 (3) 表示的平均特征向量 (整体平均特征向量) $\underline{m}$ 。

$$\underline{m} = (m_k) = (m_1, \dots, m_N)^T \quad \dots\dots (3)$$

接着, 学习装置 102 计算字符类型之间的方差矩阵 S_b 以及字符类型内部的方差矩阵 S_w , 根据每个字符类型 i 的特征向量 $\underline{X}_j^{(i)}$, 示例的个数 n_i , 上述的平均特征向量 $\underline{m}^{(i)}$, 以及整体的平均特征向量 $\underline{m}$, 如下式 (4) 至 (7) 所示。下标 p 和 q 表示在范围 $1 \leq p, q \leq N$ 中的特征向量的元素个数。

$$S_b = (b_{pq}) \quad \dots\dots (4)$$

$$b_{pq} = \sum_{i=1}^g n_i (m_p^{(i)} - m_p) (m_q^{(i)} - m_q) \quad \dots\dots (5)$$

$$S_w = (W_{pq}) \quad \dots\dots (6)$$

$$W_{pq} = \sum_{i=1}^g \sum_{j=1}^{n_i} (X_{jp}^{(i)} - m_p^{(i)}) (X_{jq}^{(i)} - m_q^{(i)}) \quad \dots\dots (7)$$

而且, 学习装置 102 利用上述的字符类型之间的方差矩阵 S_b 和字符类型内部的方差矩阵 S_w 来满足下面的等式 (8), 并计算 N 组特征向量 ϕ_k (维数为 N 且其长度为 1) 和一组特征根 λ_k ($1 \leq k \leq N$)。

$$S_b \phi_K = \lambda_K S_w \phi_K \quad (1 \leq K \leq N)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \quad \dots\dots (8)$$

接着, 学习装置 102 从对应被计算的 N 个特征向量 ϕ_K 的特征根 λ_K 中选取较大的 M ($M < N$) 个特征向量 ϕ_h ($1 \leq h \leq M$), 并将其存放在特征向量存储装置 103 中。

学习装置 102 通过计算上述 M 个特征向量 ϕ_h ($1 \leq h \leq M$) 与每个字符类型 i 的 N 维平均特征向量 $\underline{m}^{(i)}$ 的内积, 为每个字符类型之计算 M 维的平均选择特征向量 $\underline{m}^{(i)}$, 由下式 (9) 所示。接着, 将其存放在识别字典装置 104 中。

$$\begin{aligned}\underline{m}^{(i)'} &= (\underline{m}_h^{(i)'}) \\ &= (\phi_1^T \underline{m}^{(i)}, \phi_2^T \underline{m}^{(i)}, \dots, \phi_M^T \underline{m}^{(i)}) \quad \dots\dots (9)\end{aligned}$$

这里的下标 h 表示在范围 $1 \leq h \leq M$ 中选择特征向量的元素个数。这样，通过对被存放在字典中的每个字符类型的平均特征向量，将 N 维减少为 M 维，执行特征选择处理。这种特征选择处理等同于将一个 N 维的平均特征向量投射在用 M 个特征向量 ϕ_h 定义的 M 个座标轴上的处理过程，因此，字符类型之间的方差可以被扩大，而字符类型内部的方差可以被减少，即，使得不同类型的字符可以被区分，而相同类型的字符在经过由 M 个特征向量 ϕ_h 所规定的特征选择之后，能被收集在一个空间中。即，在典型确认分析中，对应所有目标字符类型的原始特征空间中的所有群集都被转换为新的空间。

当实际识别某个字符时，特征提取装置 101 从其字符类型未知的输入字符模式中提取由下式 (10) 所表示的 N 维特征向量 $\underline{X}$ 。

$$\underline{X} = (X_K) = (X_1, \dots, X_N)^T \quad \dots\dots (10)$$

这里的下标 K 指示在特征选择前在 $1 \leq K \leq N$ 的范围内特征向量的元素个数。

特征选择装置 105 计算 M 维选择特征向量 $\underline{Y}$ ，通过计算 M 个特征向量 ϕ_h ($1 \leq h \leq M$) 以及存放在特征向量存储装置 103 中的 N 维特征向量 $\underline{X}$ 的内积，由下式 (11) 表示。

$$\underline{Y} = (y_h) = (\phi_1^T \underline{X}, \phi_2^T \underline{X}, \dots, \phi_M^T \underline{X}) \quad \dots\dots (11)$$

这样，通过对输入的特征向量将 N 维减少为 M 维来执行特征选择处理。

最后，校对装置 106 为每个字型类型 i 计算出 M 维选择特征向量 $\underline{Y}$ 和存放在识别字典 104 中的每个平均选择特征向量 $\underline{m}^{(i)}$ 之间的每个欧几里德距离 $d^{(i)}$ ，利用下面的等式 (12)。

$$d^{(i)} = \left\{ \sum_{h=1}^M (y_h - m_h^{(i)'})^2 \right\}^{1/2} \quad \dots\dots (12)$$

接着，校对装置 106 输出对应具有最短距离 $d^{(i)}$ 的平均选择特征向量 $\underline{m}^{(i)}$ 的字符类型 i 做为被估计的字符类型。

根据上述的利用典型确认分析做为特征选择的常规技术，用来计算距离的元素个数被从 N 项减少为 M 项。因此，通过近似地将 M 设置为 N 的 $1/8$ ，就能大大地提高识别速度。

然而，根据利用典型确认分析做为特征选择的常规技术，不能保证 M 个特征向量 ϕ_h ($1 \leq h \leq M$) 互相正交。因此，如果新的特征空间被定义在这些特征向量 ϕ_h 的基础上，目标字符的特征向量 $\underline{X}$ 被投影在对应上述 M 个特征向量 ϕ_h 的 M 个座标轴上，并且如果在投影结果，即选择特征向量 $\underline{Y}$ 和每个字符类型 i 的平均选择特征向量 $\underline{m}^{(i)}$ 之间的欧几里德距离被计算时，则该距离可能会和在原始的 N 维特征空间中的距离大不一样。

为了容易理解，假定特征选择前的维数为 3，而特征选择后的维数为 2，如图 2 所示。

在特征选择前 3 维特征空间中特征向量 $\underline{X}_2$ 和特征向量 $\underline{X}_1$ 之间的欧几里德距离 d_{org} 由下式 (13) 得到。

$$d_{org} = ||\underline{X}_2 - \underline{X}_1|| \\ \{(\underline{a}_1^T \underline{X}_2 - \underline{a}_1^T \underline{X}_1)^2 + (\underline{a}_2^T \underline{X}_2 - \underline{a}_2^T \underline{X}_1)^2 + (\underline{a}_3^T \underline{X}_2 - \underline{a}_3^T \underline{X}_1)^2\}^{1/2} \dots\dots (13)$$

而特征选择后在 2 维特征空间中的特征向量 $\underline{X}_2$ 和特征向量 $\underline{X}_1$ 之间的近似的欧几里德距离 d_{new} 由下式 (14) 得到。

$$d_{new} = \{(\phi_1^T \underline{X}_2 - \phi_1^T \underline{X}_1)^2 + (\phi_2^T \underline{X}_2 - \phi_2^T \underline{X}_1)^2\}^{1/2} \dots\dots (14)$$

在上面的等式 (14) 中，右手边上每一项中的物理量由图 2 所示。如图 2 所示，等式 (14) 并不基于勾股定理。在特征选择后的 2 维特征空间中的欧几里德距离 d_{new} 完全不同于特征选择前 3 维特征空间中的欧几里德距离 d_{org} 。

根据使用典型确认分析做为特征选择的常规技术，特征向量 $\underline{X}$ 的投影和由典型确认分析法得到的 M 个特征向量 ϕ_h ($1 \leq h \leq M$) 的组合并不能看作原始特征空间的部分特征空间上的特征向量 $\underline{X}$ 的投影，因此在识别字符中将导致精度的降低。所以，实现具有实用识别精度的字符识别装置是很困难的。

利用主要部分分析法，即特征选择的另一种方法，需要对每种字符类型计算出一组能使每个字符类型互相区分的主要部分向量 (每种字符

类型对应一组主要部分向量)。使用这种分析方法,不是为了确定多类字符而对字符分类,而是当特征空间相邻区域中存在类似的字符类型时,通过将某个目标字符的特征向量投影到对应每种字符类型的主要部分向量以获得正确的识别结果。即,在主要部分分析法中,为每个字符类型生成指定某个主要部分向量做为座标轴的新的独立空间。因为在目标字符的特征向量被投影在每个字符类型的主要部分向量上之后才计算目标字符和每个字符类型之间的距离,因此,当涉及大量的字符类型时,就需要大量的计算。所以,该分析方法主要用于,例如,确定诸如数字字符、少量类型的字型等这样的字符。

特征向量距离的定义涉及 Mahalanobis 距离和贝叶斯 (Bayes) 距离。当某种识别装置利用这些距离设计时,在做了这样的限制后再计算距离,即,用来计算距离的座标轴应互相正交。因此,典型确认分析法中的问题不会出现在该方法中,但存在的问题是:计算被复杂化且需要进行大量的计算。

上述的问题不仅涉及字符识别装置,而且对于通过特征向量识别其它的各种模式(例如图像模式、声音模式等)的技术也同样存在。

本发明是基于上述的背景而被开发的,目的是利用某种特征选择方法减少特征向量的维数,实现能在高速执行识别过程的同时保持识别精度的一种模式识别装置,以便识别各种模式。

本发明的第一方面是基于一种模式识别装置,该装置利用代表例如某个输入字符模式等一类的输入模式的特征的输入特征向量来识别某个输入模式。

首先,本发明包括输入特征向量提取装置,用于提取代表输入模式的特征的输入特征向量。

还包括规范正交基存储装置,用于保存原始特征空间的部分特征空间的规范正交基。

还包括识别字典装置,存放对应一个或多个目标模式在部分特征空间中被规定的每个字典选择特征向量。

进一步包括特征选择装置,利用被存放在规范正交基存储装置中的规范正交基,计算被投影在由输入特征向量提取装置所提取的输入特征

向量的部分特征空间上的输入选择特征向量。

还包括校对装置，通过校对由字符选择装置计算出来的输入选择特征向量和存放在识别字典装置中的每个字典选择特征向量，识别对应该输入选择特征向量的输入模式的类型。这种校对是通过例如计算输入选择特征向量和每个字典选择特征向量之间的距离并互相比对所得到的距离的值来实现的。距离可以是，例如，基于欧几里德距离，市街区距离或棋盘距离等的距离。

本发明的第二方面是基于模式学习装置，用于计算规范正交基和由根据本发明第一方面的模式识别装置得到的每个字典选择特征向量。

首先，本发明包括学习特征向量提取装置，用于提取代表学习模式的特征的每个学习特征向量。

还包括规范正交基计算装置，根据由学习特征向量提取装置所提取的学习特征向量计算部分特征空间的规范正交基。

还包括字典特征向量计算装置，用于计算在对应基于由学习特征向量提取装置所提取的学习特征向量的每个目标模式的原始特征空间中指定的每个字典特征向量。

进一步包括字典选择特征向量计算装置，利用规范正交基计算装置所计算的规范正交基计算被投影在由字典特征向量计算装置所计算出来的每个字典特征向量的部分特征空间上的每个字典选择特征向量。

根据本发明第一或第二方面的配置，可以计算正交基，例如，在学习模式的学习特征向量上执行典型确认分析之后，从具有最大特征根的一个开始，顺序计算少于原始特征空间的维数的预定数目的特征向量，并且，从具有最大特征根的一个开始，顺序地规范正交化预定数目的特征向量。

根据本发明的第一或第二方面的配置，将输入特征向量或字典特征向量投影在部分特征空间上得到的输入选择特征向量或字典选择特征向量可以被配置得似乎是将输入特征向量或字典特征一投射在正交基向量上得到的一样。此外，做为投影结果得到的输入选择特征向量或每个字典选择特征向量可以被配置得似乎是做为计算输入特征向量或字典特征向量与每个正交基向量的内积后所得到的结果。

根据上述的本发明第一或第二方面，字典选择特征向量可以被配置得似乎基于做为每类目标模式的特征向量的平均向量得到的平均特征向量。

本发明可以被配置到具有根据上述的本发明第一或第二方面的功能。

本发明也可以被配置成为计算机可读的存储介值，被用来指导计算机执行根据上述的本发明第一或第二方面的功能。

正如上述配置所指出的那样，本发明通过减少在常规技术中那样的距离计算中的元素个数来实现高速识别。而且，本发明的唯一特征是通过将特征向量投影在原始特征空间的部分特征空间上来存储距离的，由此实现具有实用识别精度的字符识别装置。

在这种情况下，在用常规的典型确认分析法等方法计算特征向量之后，利用规范正交化的一个简单递归方程计算正交基向量，就能用简单的处理计算得到正交基。

可以预料在字符类型内部方差可以被最小化而字符类型之间的方差可以被最大化的典型确认分析法中，通过规范正交化上述的特征向量，其特征可能会在某种程度上受到损失。然而，从实验中知道，通过提高由正交化距离计算的规范化得到的识别率，就足以补偿识别率的这种损失，由此确切地显示出本发明的意义。

对于一般熟悉本技术的人来说，结合附图和本发明的最佳实施例说明，就能很容易地理解本发明的其他目的和特征。

图 1 表示常规技术的配置；

图 2 表示常规技术的问题；

图 3 表示根据该最佳实施例的配置；

图 4 表示本发明的最佳实施例的原理；

图 5 是一个流程图，表示在本发明最佳实施例中学习过程的操作；

图 6 是一个流程图，表示计算字符类型之间方差矩阵 S_b 的过程；

图 7 是一个流程图，表示计算字符类型内部方差矩阵 S_w 的过程；

图 8 是一个流程图，表示计算正交基向量的过程；以及

图 9 表示计算 ψ_h 的过程。

图 10 表示实施最佳实施例的系统结构。

结合附图顺序地介绍本发明的最佳实施例。

根据本发明的最佳实施例，在根据典型确认分析法从具有最大特征值的一个开始顺序计算 M 个特征向量 ϕ_h ($1 \leq h \leq M$) 之后，再利用施密特 (schmidt) 规范正交化方法规范正交化特征向量 ϕ_h 计算 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$)。接着，通过将 N 维特征向量投射在上述的 M 个规范正交基向量 ψ_h 上来计算选择特征向量。利用选择特征向量执行识别处理，并通过特征选择方法减少特征一的维数。因此，根据本发明，既能高速执行识别处理，又能保持高识别精度。

为了便于理解，假定在特征选择前的维数为 3，而特征选择后的维数为 2，如图 4 所示。

首先，在根据典型确认分析法从具有最大特征值的一个开始顺序计算两个特征向量 ϕ_1 和 ϕ_2 之后，再利用施密特规范正交化方法规范正交化特征向量来计算两个规范正交基向量 ψ_1 ($=\phi_1$) 和 ψ_2 。

在特征选择之前，在 3 维特征空间中从特征向量 X_2 到特征向量 X_1 的欧几里德距离 d_{org} 由上述的等式 (13) 得到。

特征选择后，在 2 维特征空间中从特征向量 X_2 到特征向量 X_1 的近似欧几里德距离 d_{new} 由下式 (15) 得到。

$$d_{new} = \{ (\phi_1^T x_2 - \phi_1^T x_1)^2 + (\phi_2^T x_2 - \phi_2^T x_1)^2 \}^{1/2} \dots (15)$$

在上述的等式 (15) 中，右手边每一项的物理量如图 4 所示。正如图 4 中所清楚表示的那样，等式 (15) 基于勾股定理，而且特征选择后的在 2 维特征空间中的欧几里德距离 d_{new} 近似于特征选择前在 3 维特征空间中的欧几里德距离 d_{org} 。

上述的原理适用于维数很大的情况。在下面实际介绍的最佳实施例中，在正交基向量 ψ_h 上特征向量 X 的每个投影的组合等于在原始特征空间的部分特征空间上特征向量 X 的投影。结果，即使在特征选择之后，也能保证足够的字符识别精度。

图 3 表示根据上述原理的最佳实施例的配置。

首先，在学习过程中，特征提取装置 301 和学习装置 302 执行图 5

中的流程图所表示的学习过程。

特征提取装置 301 在步 503 中从被包含在每个字符类型之中的 n_i 个字符模式的示例中提取由上述等式 (1) 指示的 n_i 组学习特征向量 $\underline{X}_j^{(i)}$ ($1 \leq j \leq n_i$)。在示例数目字在步 502 被置为 1 之后, 由步 504 对示例数目字增 1, 重复执行步 503 中的处理直到在步 505 中判定示例数字已经超出了字符类型 i 的示例数 n_i 。因此, 字符类型 i 的 n_i 组学习特征向量 $\underline{X}_j^{(i)}$ 被提取。

接着, 学习装置 302 利用上述步 502 至 505 中得到的对应字符类型 i 的 n_i 组学习特征向量 $\underline{X}_j^{(i)}$, 通过对向量的每个元素的位置计算向量元素的平均值, 对一个字符类型 i 计算由对应字符类型 i 的等式 (2) 所代表的平均特征向量 $\underline{m}^{(i)}$ 。

在将字符类型数 i 在步 501 置为 1 之后, 由步 507 对字符类型数 i 增 1, 执行步 502 至步 506 中的处理, 直到在步 508 中判定字符类型数 i 已经超过了字符类型的总数 g 。因此, 对字符类型 i 的 g 次变化 ($1 \leq i \leq g$), 在 n_i 组学习特征向量 $\underline{X}_j^{(i)}$ 和一个平均特征向量 $\underline{m}^{(i)}$ 上顺序执行计算处理。

接着, 在步 509 中, 学习装置 302 从每个字型类型 i 的 n_i 个示例中计算出由上述等式 (3) 所示的所有字型类型的平均特征向量 (总体平均特征向量) $\underline{m}$ 的每个元素 m_k , 上述平均特征向量 $\underline{m}^{(i)}$ 的每个元素 $\underline{m}_k^{(i)}$ 由下式 (16) 所示

$$m_k = \left(\sum_{i=1}^g m_k^{(i)} \cdot n_i \right) / \sum_{i=1}^g n_i \quad \dots (16)$$

在步 510 中, 学习装置 302 根据每个字符类型 i 的特征向量 $\underline{X}_j^{(i)}$ 、示例的个数 n_i 、平均特征向量 $\underline{m}^{(i)}$ 以及总体平均特征向量 $\underline{m}$ 计算字符类型之间的方差矩阵 S_b 。图 6 的流程图表示这一过程的细节。

在第一元素个数 P 在步 601 被设置为 1 之后, 在步 610 中增 1 直到在步 611 中确定元素个数 P 已经超出个数 N 。在第二元素数 q 在步 602 被置为 1 之后, 在步 608 中被增 1, 直到在步 609 中被确定该元素数 q 已经超出个数 N 。在步 603 至步 607 中对由这些控制循环确定的每组第一元素数 P 和第二元素数 q 。通过上述的等式 (5) 执行计算。

即, 在步 604 中, 被包含在由第一元素数 P 和第二元素数 q 确定的

字符类型之间方差矩阵 S_b 中的元素值 b_{pq} 被清零，并且在字符类型数 i 在步 603 中被置为 1 后，在步 606 中增 1 直到在步 607 中确定字符类型数 i 已经超过了字符类型的总数 g 。接着，在步 605 中，利用下式 (17) 重复执行计算得到元素值 b_{pq} 。

$$b_{pq} = b_{pq} + n_i (m_p^{(i)} - m_p) (m_q^{(i)} - m_q) \dots (17)$$

被包含在字符类型之间的方差矩阵 S_b 中的所有元素值 b_{pq} 都是由第一元素数 P 和第二元素数 q 的组合确定的，利用步 601、602 和 608 至 611 中的控制循环，在步 603 至 607 中重复执行这些处理过程。

接着，在步 612 中，字符类型之间的方差矩阵 S_b 被输出，做为由上述等式 (4) 所计算的一组所有的元素值 b_{pq} 。

在步 510 中，学习装置 302 得到字符类型之间的方差矩阵 S_b ，如图 5 所示，并且在步 511 中，根据特征向量 $\underline{X}_j^{(i)}$ 、示例的个数 n_i 以及平均特征向量 $\underline{m}^{(i)}$ 计算字符类型内部的方差矩阵 S_w 。图 7 为这些处理过程细节的流程图。

在步 701 中，第一元素数 P 被置为 1，接着在步 713 中增 1 直到在步 714 中确定第一元素数 P 已经超出了个数 N 。第二元素数 q 在步 702 中被置为 1 并在步 711 中增 1 直到在步 712 中被判定第二元素数 q 已经超过个数 N 。在步 703 至 710 中，对由控制循环确定的每组第一元素数 P 和第二元素数 q ，由上述的等式 (7) 执行计算过程。

即，在步 705 中，被包含在由第一元素数 P 和第二元素数 q 确定的字符类型内部的方差矩阵 S_w 中的元素值 W_{pq} 被清零，并且在字符类型数 i 在步 703 被置为 1 之后，在步 709 中对字符类型数 i 增 1，直到在步 710 中确认字符类型数 i 已经超过字符类型的总数 g 。接着，重复执行步 706 至 708 中的循环。而且在该循环中，示例数字在步 704 中被置为 1，并且在步 707 中增 1 直到在步 708 中被确认示例数字已经超出示例的总数 n_i 。接着，在步 706 中，由下式 (18) 重复执行计算过程得到元素值 W_{pq} 。

$$W_{pq} = W_{pq} + (x_{jp}^{(i)} - m_p^{(i)}) (x_{jq}^{(i)} - m_q^{(i)}) \dots (18)$$

所有被包含在字符类型内部的方差矩阵 S_W 中的元素值 W_{pq} 都是由第一元素数 P 和第二元素数 q 的组合确定的, 利用步 701、702 以及 711 至 714 中的控制循环, 在步 703 至 710 中重复执行这些处理。

接着, 在步 715 中, 字符类型内部的方差矩阵 S_W 被输出做为由上述等式 (6) 计算出来的一组所有的元素值 W_{pq} 。

学习装置 302 利用图 5 所示的步 510 中计算的字符类型之间的方差矩阵 S_b , 以及在步 511 中计算的字符类型内部的方差矩阵 S_W , 在步 512 中满足上述的等式 (8), 然后对 N 组特征向量 ϕ_k 和特征值 λ_k ($1 \leq k \leq N$) 执行计算过程。由于该计算过程是众所周知的, 因此在这里忽略了详细的解释。学习装置 302 在步 512 中从 N 个字符向量 ϕ_k 中选择具有较大的特征值 λ_k 的 M ($M < N$) 个特征向量 ϕ_h ($1 \leq h \leq M$)。

在步 513 中, 学习装置 302 通过规范正交化上述的特征向量 ϕ_h ($1 \leq h \leq M$) 计算出 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$)。所用的规范正交方法可以是施密特规范正交化方法。在这种规范正交化中, 从具有较大的特征值 λ_h 的特征向量 ϕ_h 开始顺序地计算 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$), 通过执行由下面的等式 (19) 和 (20) 所代表的递归方程以及由式 (21) 表示的规范化方程实现。

$$\phi_1 = \phi_1 \quad \dots(19)$$

$$t_h = \phi_h - \sum_{i=1}^{h-1} (\phi_i^T \phi_h) \phi_i \quad (2 \leq h \leq M) \quad \dots(20)$$

$$\phi_h = t_h / \|t_h\| \quad \dots(21)$$

这里, 如图 9 和等式 (20) 所示的那样, 向量 t_h 和基于具有大于当前被处理的特征向量 ϕ_h 特征值的特征向量 ϕ_i ($1 \leq i \leq h-1$) 得到的所有规范正交基向量 ψ_i ($1 \leq i \leq h-1$) 正交, 并且顺序地对应从具有较大的特征值的特征向量开始的特征向量 ϕ_h 。接着, 利用式 (21) 将向量 t_h 的长度规范化为 1, 计算对应特征向量 ϕ_h 的规范正交基向量 ψ_h 。

图 8 为图 5 中所示的步 513 中的详细处理过程的流程图。

首先，学习装置 302 在步 801 中利用方程 (19) 执行计算过程。

学习装置 302 在步 802 中置维数 h 为 2，接着重复执行步 803 中对应等式 (20) 的计算过程和步 804 中对应等式 (21) 的计算过程，维数 h 增 1 直到在步 806 中判定维数 h 已经超过维数 M 。

在根据本实施例如上所述利用常规的典型确认分析法计算特征向量 ϕ_h 之后，利用简单的递推方程就能计算规范正交基向量 ψ_h 。

在图 5 所示的步 513 中，学习装置 302 将上述计算得到的 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$) 存储在部分本征空间的规范正交基存储装置 303 中。

接着，在图 5 所示的步 514 中，学习装置 302 计算每个字符类型 i 的 M 维平均选择特征向量 $\underline{m}^{(i)}$ ，通过对每个字符类型 i 用下面的等式 (22) 计算上述的 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$) 和在步 506 中得到的 N 维平均特征向量 $\underline{m}^{(i)}$ 的内积。接着，将结果存在识别字典装置 304 中。

$$\begin{aligned} \underline{m}^{(i)'} &= (m_h^{(i)'}) \\ &= (\psi_1^T \underline{m}^{(i)}, \psi_2^T \underline{m}^{(i)}, \dots, \psi_M^T \underline{m}^{(i)}) \quad \dots(22) \end{aligned}$$

这里的下标 h 表示在 $1 \leq h \leq M$ 的范围内选择特征向量的元素个数。因此，特征选择处理过程，即，将形成字典一部分的每个字符类型的平均特征向量的维数从 N 维减少为 M 维的处理过程被执行。根据本实施例的特征选择处理过程相当于将 N 维平均特征向量投射在由 M 个规范正交基向量 ψ_h 指定的 M 个座标轴上，因此字符类型之间的方差可以被增加，而字符类型内部的方差则可以被减少，正如在常规的技术中一样。由于 M 个规范正交基向量 ψ_h 互相正交，因此，上述的投影过程等价于将 N 维平均特征向量投射在原始空间的部分特征空间上。结果，在原始特征空间中的向量之间的距离对应在由 M 个规范正交基向量 ψ_h 指定的部分特征空间中的向量之间的距离，如图 4 所示。

如上所述，学习装置 302 执行根据图 5 所示流程图的学习过程。

当实际识别某个字符时，特征提取装置 301 从具有未知字符类型的

输入字符模式中提取由上述等式 (10) 所指示的 N 维特征向量 $\underline{X}$ 。

接着, 特征选择装置 305 通过获取存放在部分特征空间中的规范正交基存储装置 303 中的 M 个规范正交基向量 ψ_h ($1 \leq h \leq M$) 和 N 维特征向量 $\underline{X}$ 的内积来计算 M 维选择特征向量 $\underline{Z}$, 如下式 (23) 所示。

$$\underline{Z} = (z_h) = (\psi_1^T \underline{x}, \psi_2^T \underline{x}, \dots, \psi_M^T \underline{x})^T \quad \dots(23)$$

接着, 输入特征向量被投射 (特征选择) 在 M 维部分特征空间上。

最后, 校对装置 306 对每个字符类型之用下式 (24) 计算 (例如) M 维选择特征向量 $\underline{Z}$ 和存放在识别字典装置 304 中的平均选择特征向量 $\underline{m}^{(i)}$ 之间的欧几里德距离。

$$d^{(i)} = \left\{ \sum_{h=1}^M (z_h - m_h^{(i)})^2 \right\}^{1/2} \quad \dots(24)$$

校对装置 306 输出对应具有最短距离 $d^{(i)}$ 的平均选择特征向量 $\underline{m}^{(i)}$ 的字符类型 i , 做为对应输入字符模式的估算字符类型。

根据上述的最佳实施例, 识别过程可以被大大加速, 因为距离计算中元素的个数从 N 项被减少到 M 项, 正如在常规的技术中一样。而且, 本发明通过将特征向量投射到原始特征空间的部分特征空间上来存储距离的值, 由此实现具有实用识别精度的字符识别装置。

可以预料在字符类型内部的方差被最小化而字符类型之间的方差被最大化的典型确认分析法中, 其特征将稍微受到损失。从实验结果看, 识别率的这种降低可通过正交化的距离计算规范化提高识别率来得到更多的补偿。因此, 本发明的有效性已被实践证明。

根据上述的最佳实施例, 欧几里德距离被用作向量之间的距离。然而, 本发明并不局限于这种应用, 基于欧几里德距离的其他距离, 市街区距离、棋盘距离等, 也可以被使用。向量 $\underline{X}_1$ 和向量 $\underline{X}_2$ 之间的市街区距离 d 是由下面的等式定义的。然而, X_{1h} 和 X_{2h} 分别为向量 $\underline{X}_1$ 和 $\underline{X}_2$ 的元素。距离的计算是沿着座标轴的垂直或水平的向进行的。

$$d = \sum_{h=1}^M |x_{1h} - x_{2h}| \quad \dots(25)$$

而向量 $\underline{X}_1$ 和 $\underline{X}_2$ 之间的棋盘距离 d 是由下式定义的。沿着座标轴的垂直、水平或对角线方向中的一种方向计算出最短的距离。

$$d = \max_h |x_{1h} - x_{2h}| \quad \dots(26)$$

本发明并不局限于在上述最佳实施例中所介绍的字符识别装置，也可以被应用到各种模式识别装置中。

此外，本发明也包括例如软盘、CD - ROM（只读光盘），光盘和可装卸式硬盘等一类的存储介质，存放计算机程序以实现上述的模式识别功能。

图 10 表示实现上述最佳实施例的系统结构。

扫描仪装置 1001、中央处理器（CPU）1003、随机访问存储器（RAM）1004、硬盘驱动器 1005 以及 CD - ROM 驱动器 1006 被连接到公共总线 1001 上。

CD - ROM 介质存放计算机程序以实现上述的模式识别功能。该程序通过 CD - ROM 驱动器 1006 和公共总线 1001 被装入到硬盘驱动器 1005 上。

CPU1003 将程序从硬盘驱动器 1005 中装入随机存取存储器 1004 并执行这些程序。结果，对通过扫描仪装置 1002 读入的字符模式执行上述的模式识别过程。

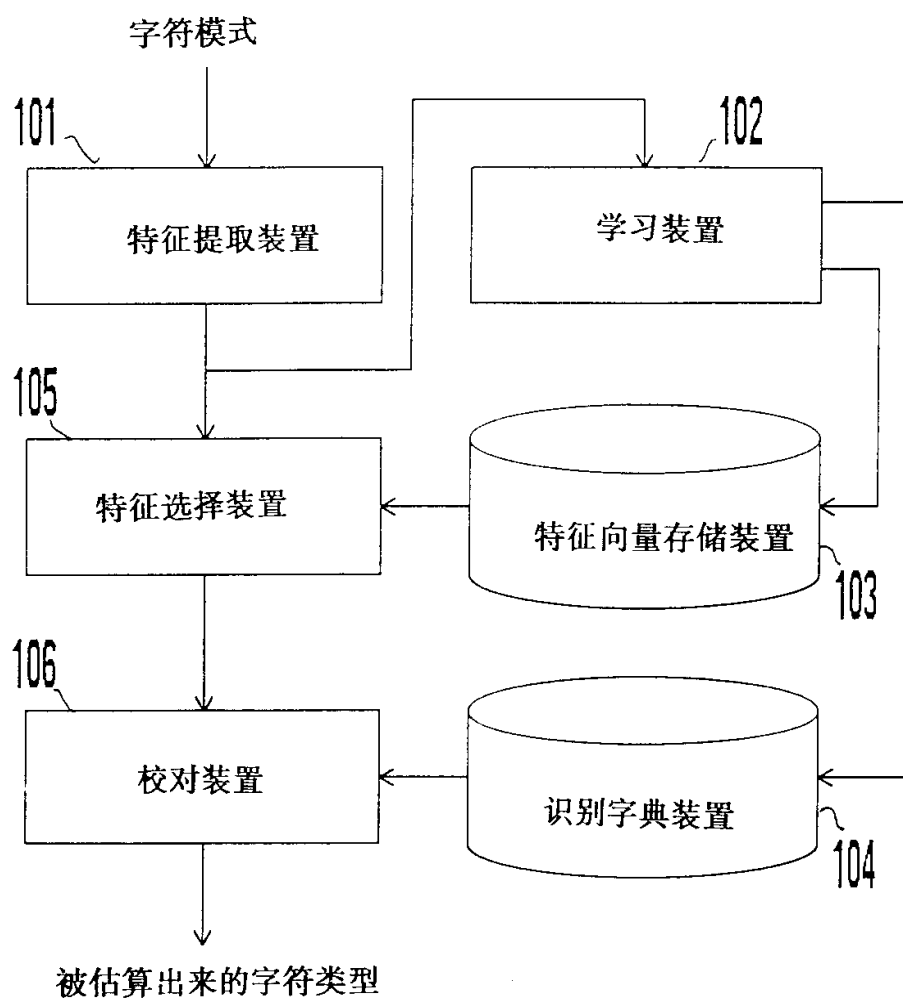
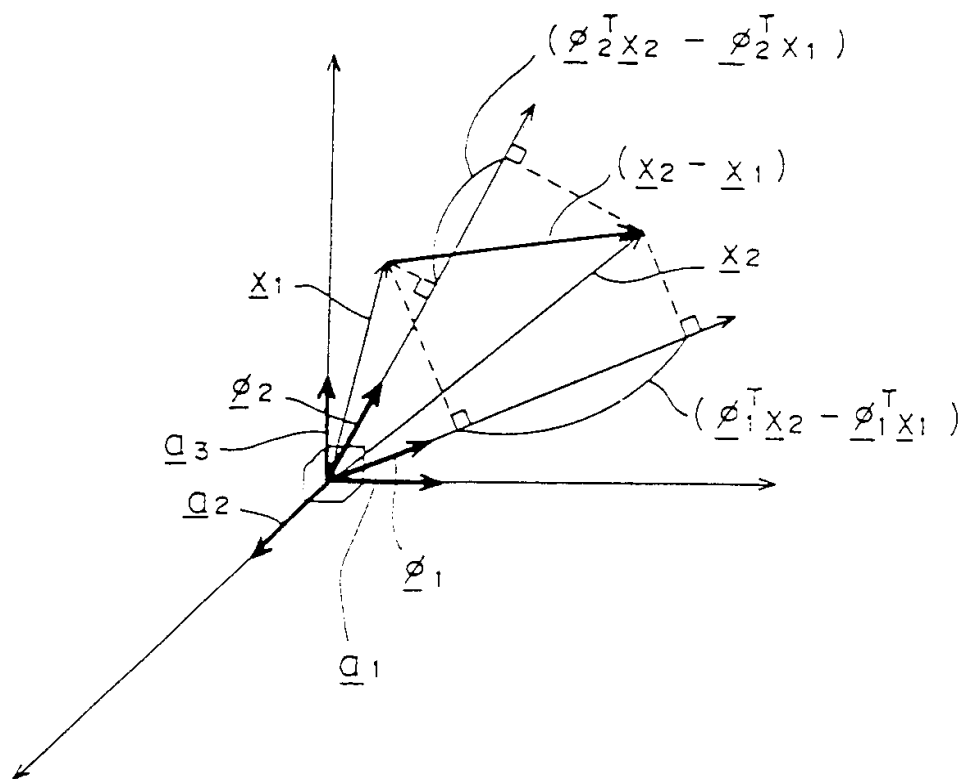


图 1



$$\begin{cases}
 d_{org} = |x_2 - x_1| \\
 = \{ (q_1^T x_2 - q_1^T x_1)^2 + (q_2^T x_2 - q_2^T x_1)^2 + (q_3^T x_2 - q_3^T x_1)^2 \}^{1/2} \\
 d_{new} = \{ (\phi_1^T x_2 - \phi_1^T x_1)^2 + (\phi_2^T x_2 - \phi_2^T x_1)^2 \}^{1/2}
 \end{cases}$$

图. 2

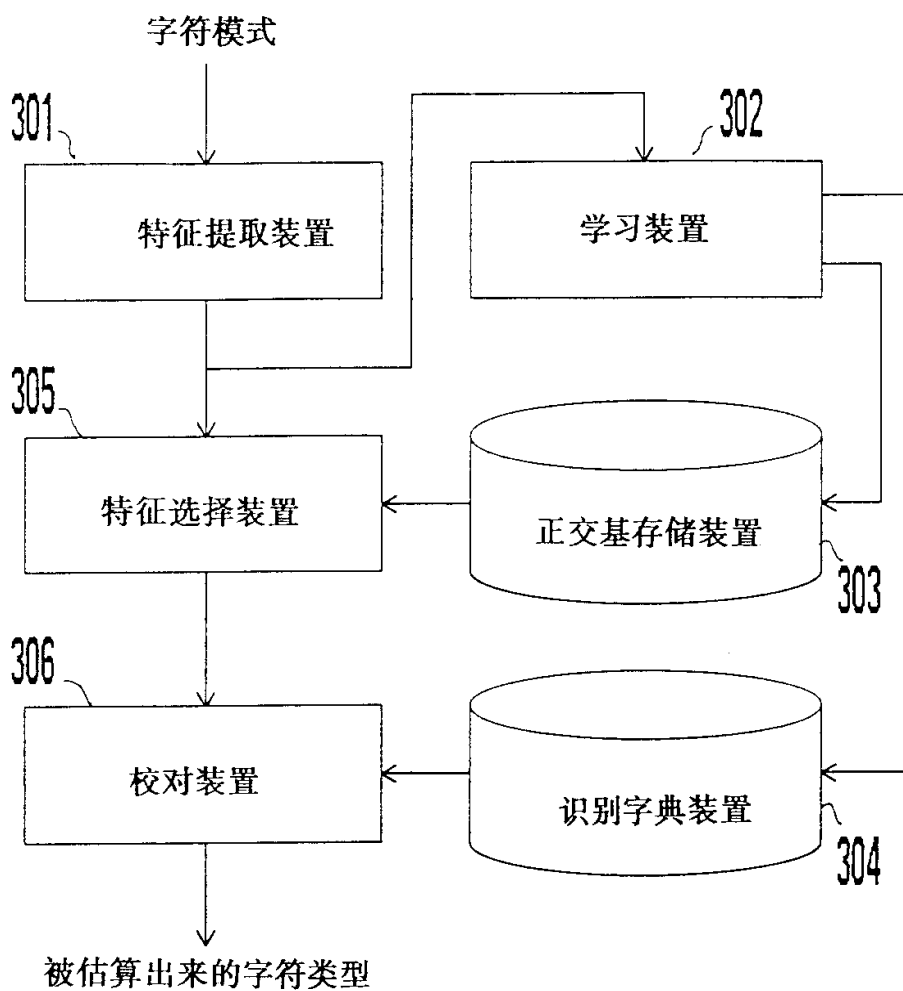
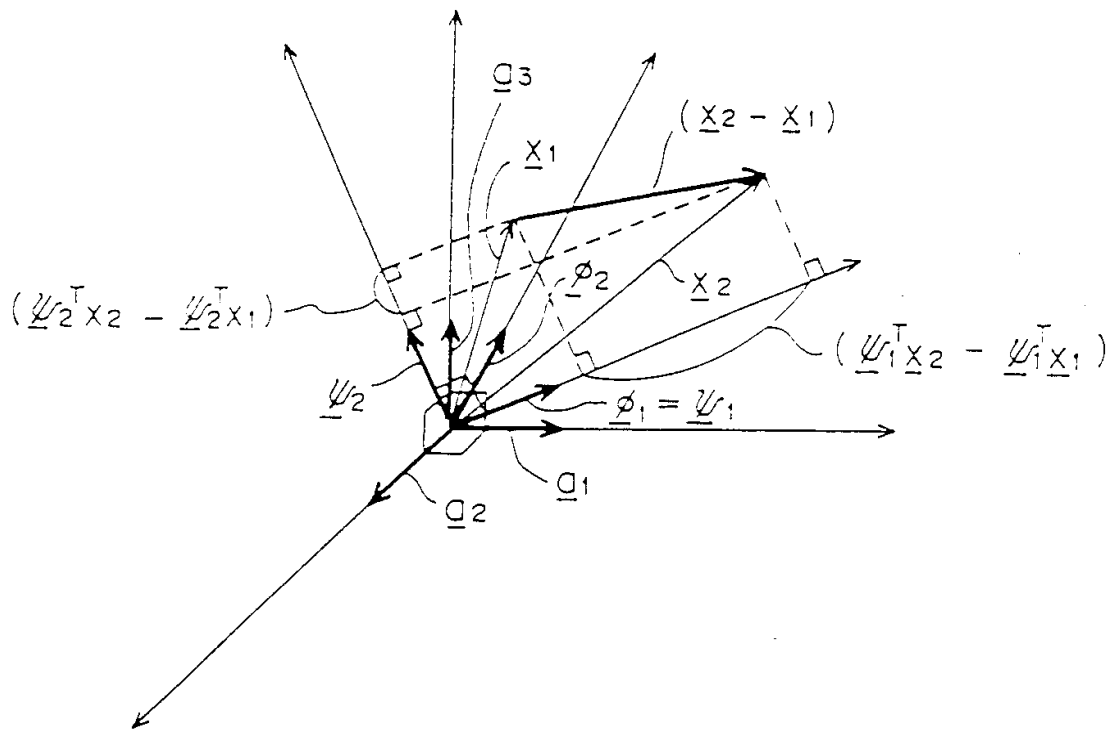


图 3



$$\begin{aligned}
 d_{org} &= | \underline{x}_2 - \underline{x}_1 | \\
 &= \{ (\underline{q}_1^T \underline{x}_2 - \underline{q}_1^T \underline{x}_1)^2 + (\underline{q}_2^T \underline{x}_2 - \underline{q}_2^T \underline{x}_1)^2 + (\underline{q}_3^T \underline{x}_2 - \underline{q}_3^T \underline{x}_1)^2 \}^{1/2} \\
 d_{new} &= \{ (\underline{\psi}_1^T \underline{x}_2 - \underline{\psi}_1^T \underline{x}_1)^2 + (\underline{\psi}_2^T \underline{x}_2 - \underline{\psi}_2^T \underline{x}_1)^2 \}^{1/2}
 \end{aligned}$$

图 4

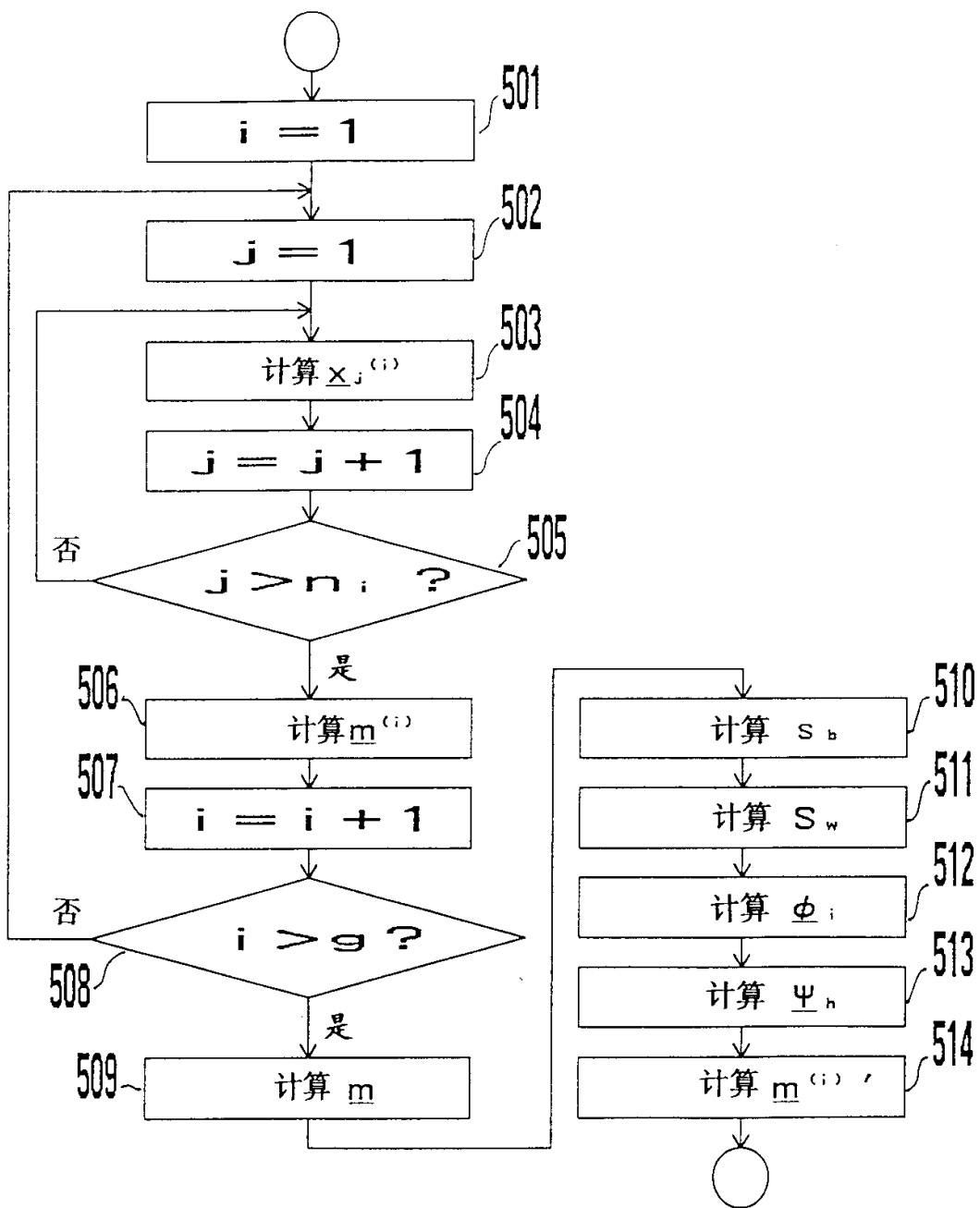


图 5

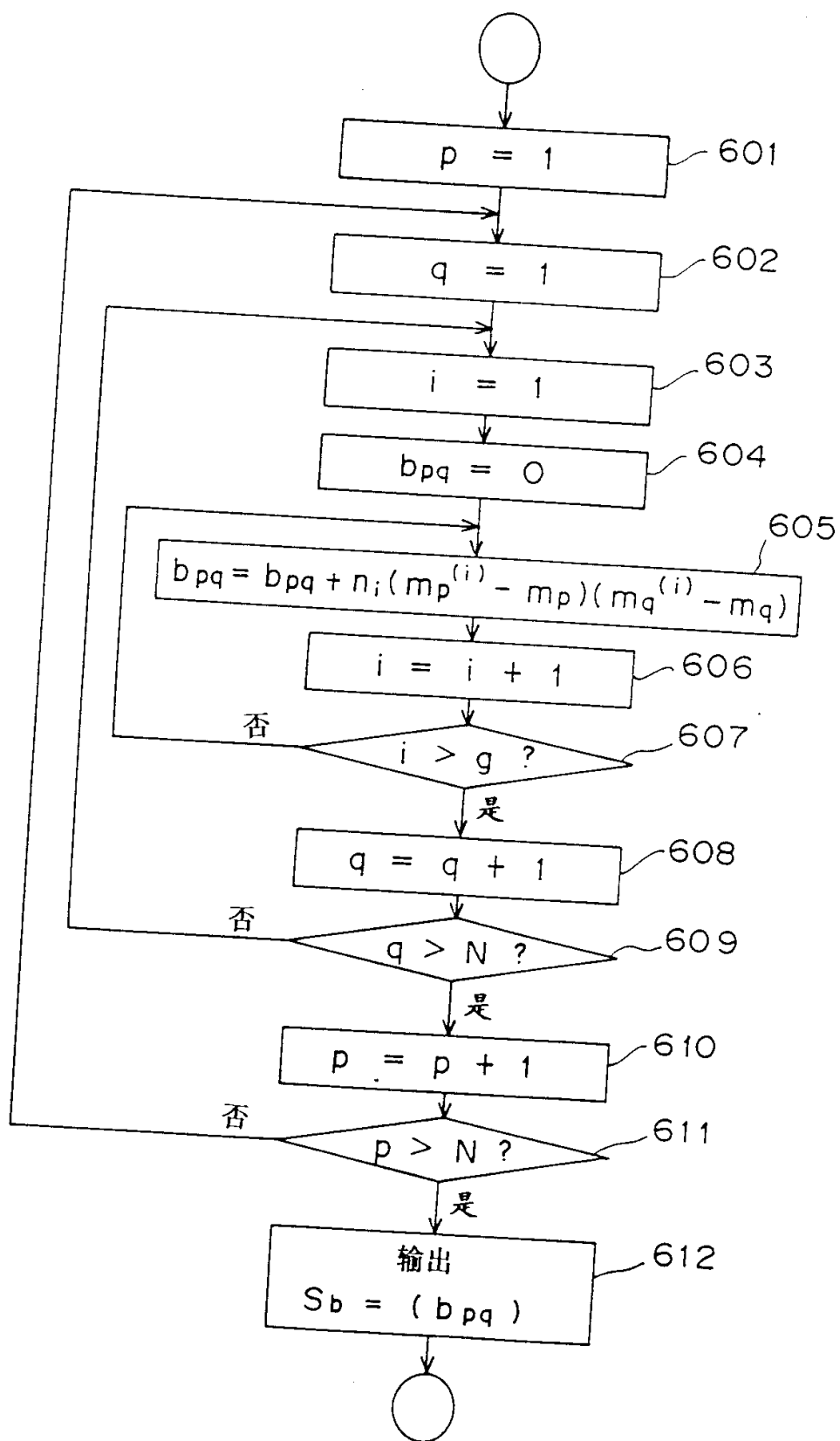


图 6

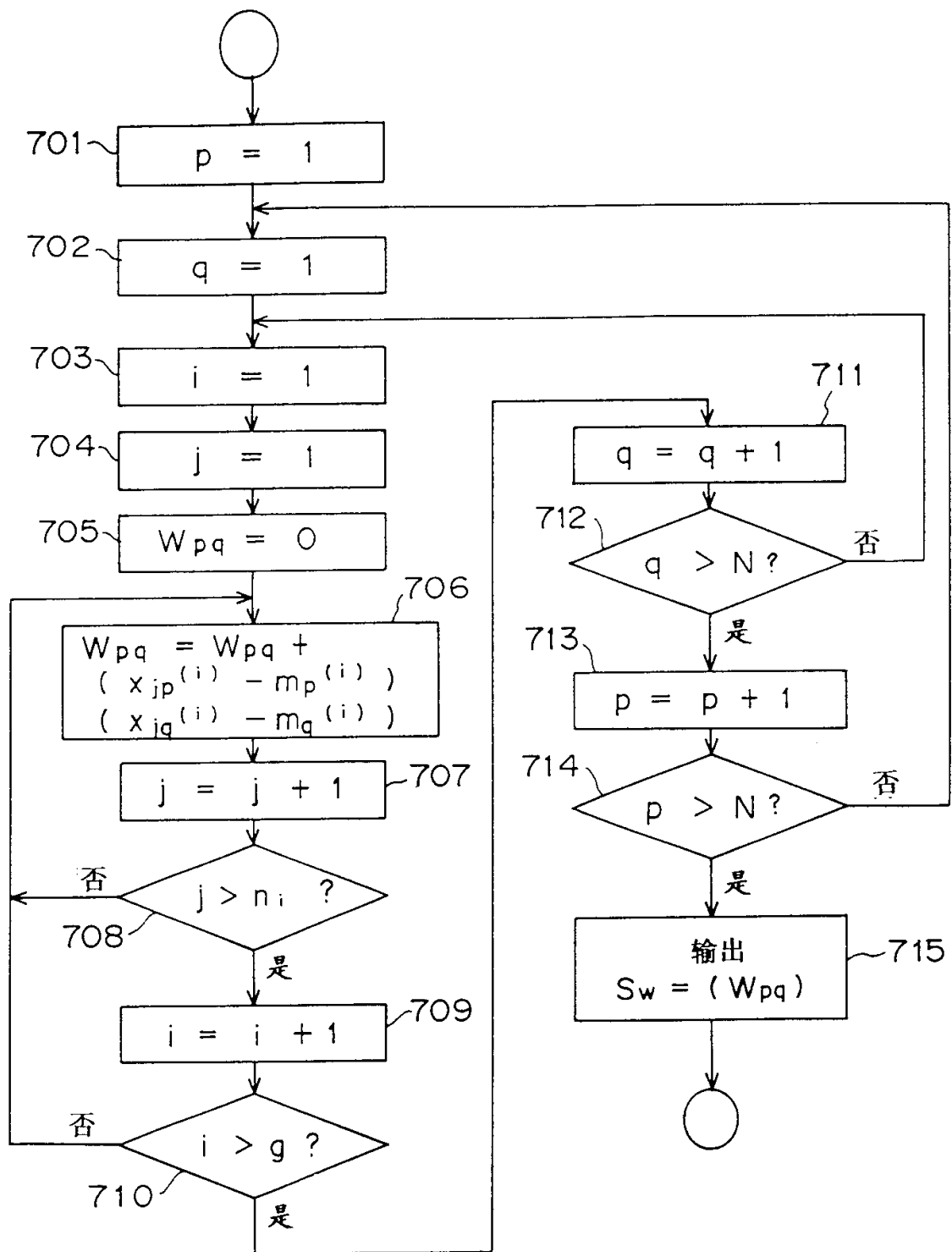


图 7

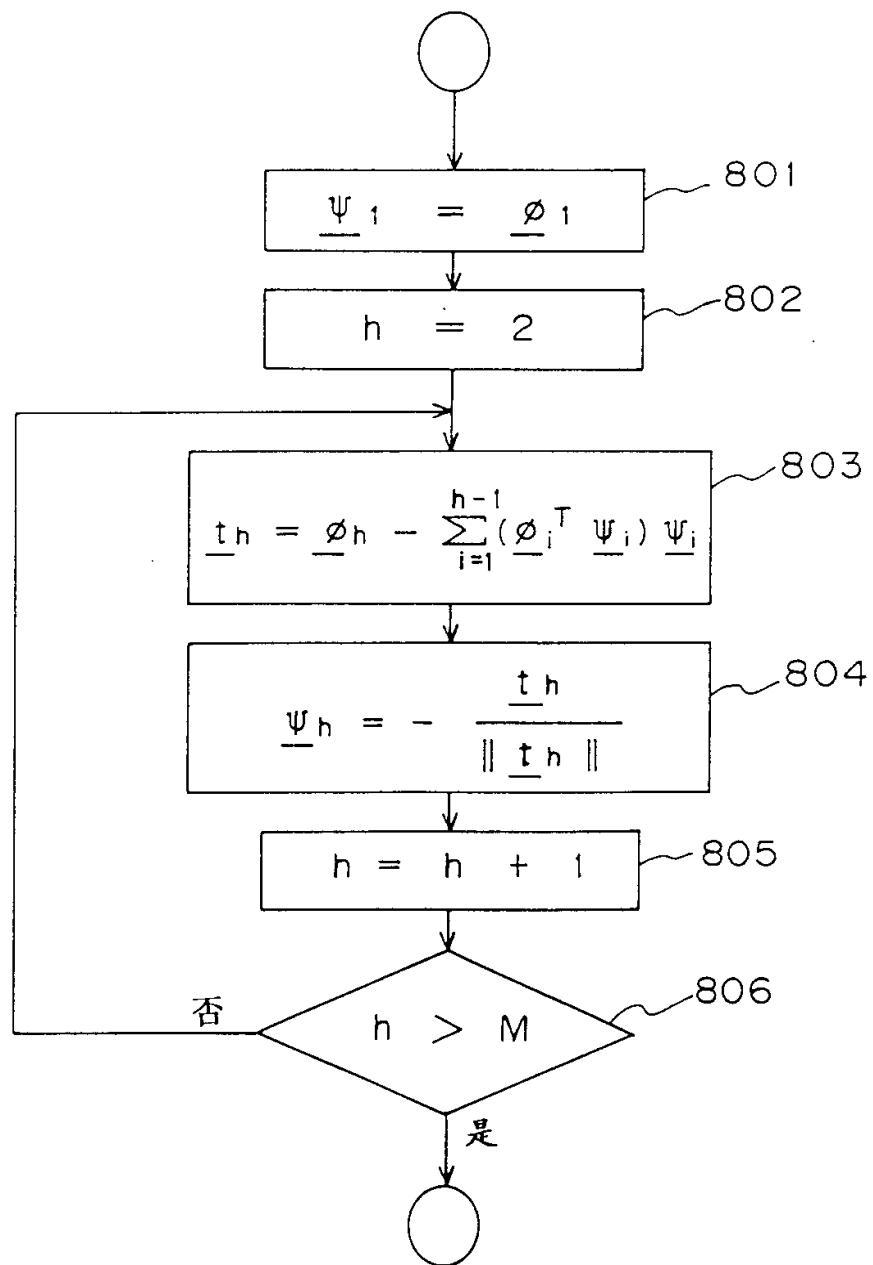


图 8

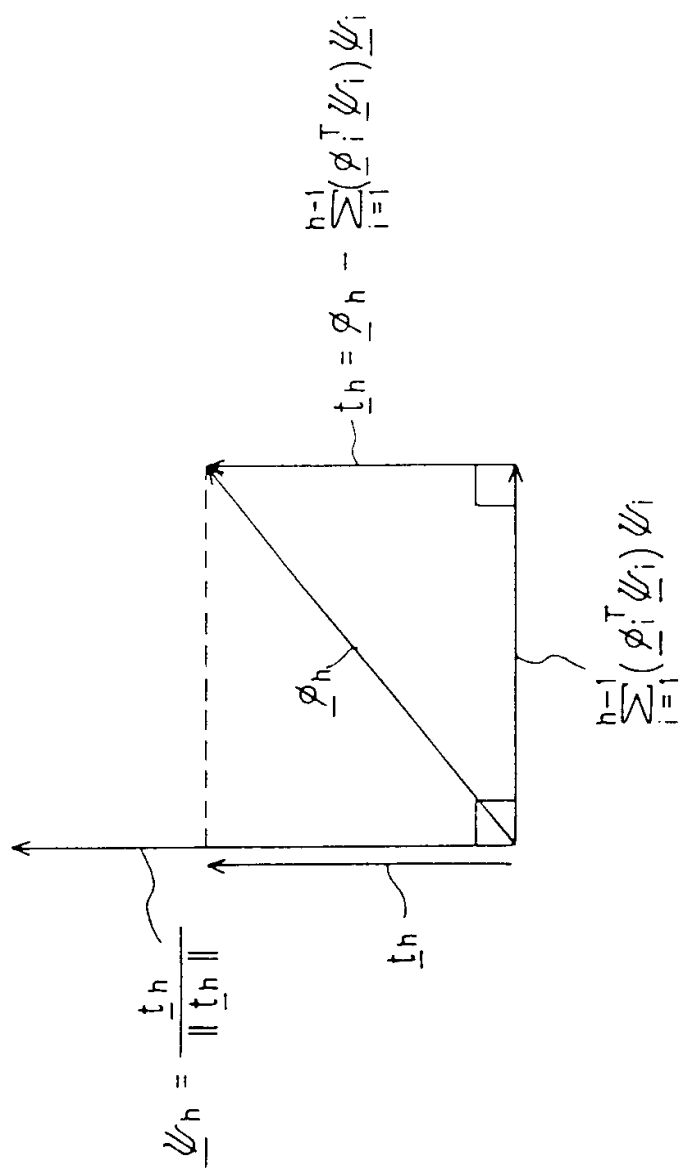


图 9

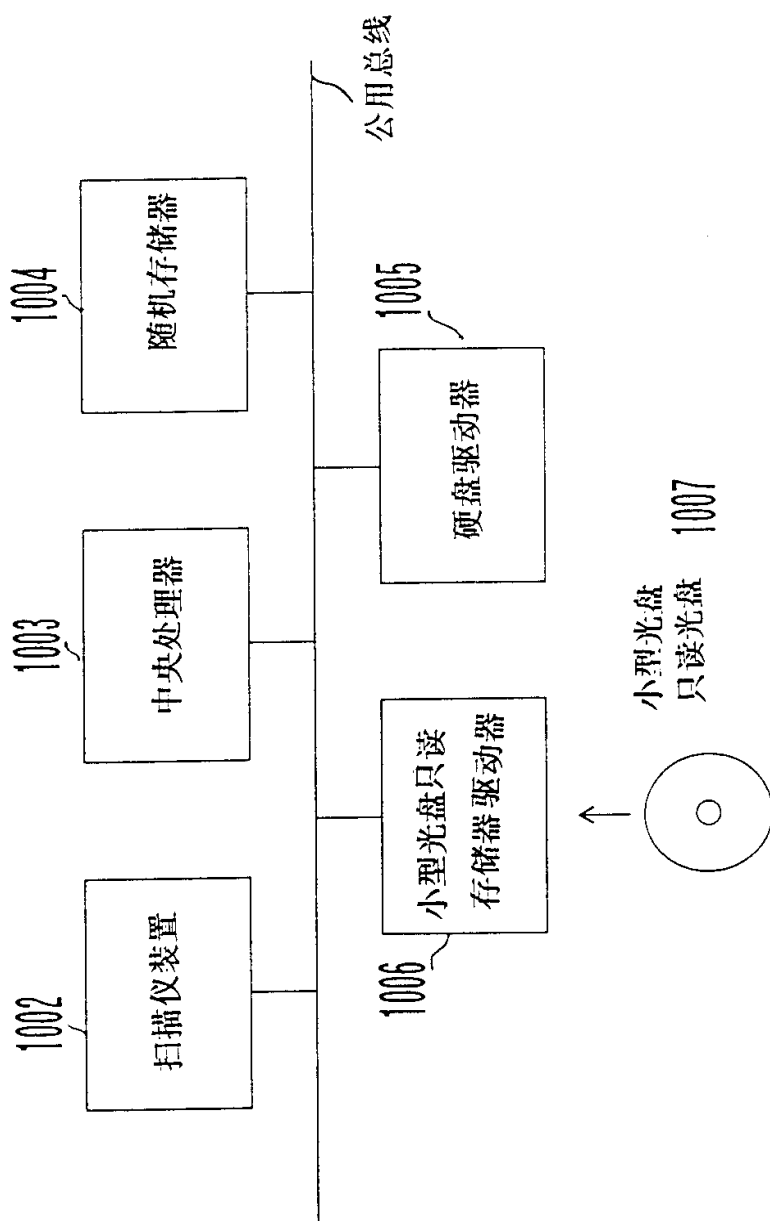


图 10