

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200610095675.4

[51] Int. Cl.

G06F 17/30 (2006.01)

G06K 9/52 (2006.01)

[43] 公开日 2007 年 6 月 13 日

[11] 公开号 CN 1979480A

[22] 申请日 2000.4.28

[21] 申请号 200610095675.4

分案原申请号 200410034396.8

[30] 优先权

[32] 1999.4.29 [33] GB [31] 9909963.2

[71] 申请人 三菱电机株式会社

地址 日本东京都

[72] 发明人 M·博贝

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 程天正 梁永

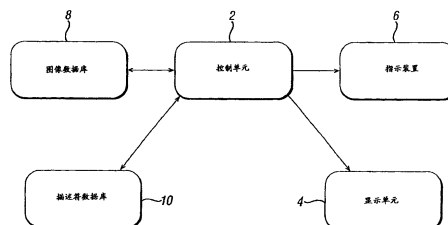
权利要求书 2 页 说明书 6 页 附图 5 页

[54] 发明名称

利用形状进行目标表示与检索的方法和装置

[57] 摘要

本发明涉及供检索使用的对出现在静止或视频图像中的目标进行表示的方法，其中的目标出现在具有第一的二维外形的图像中，通过对相应图像的信号进行处理，包括导出该目标第一外形的视图描述符以及导出至少一个在不同视图中的目标的外形的视图描述符，并且将二个或者更多的视图描述符进行关联以形成一个目标描述符。



1. 对出现在图像中的目标进行表示的一种方法，其中通过处理对应于该图像的信号，出现在图像中的目标具有一个第一的二维外形，该方法包括导出一个该目标的第一外形的视图描述符并且导出至少一个在不同视图中的该目标外形的另外的视图描述符，并且将这二个或者更多的视图描述符进行关联以形成目标描述符。

2. 如权利要求1所述的方法，其中的图像是一序列图像的一部分，并且至少一个另外的对应于出现在该图像序列中的其它地方的目标的视图的视图描述符。

3. 如权利要求1或者权利要求2所述的方法，其中的图像来自视频。

4. 如权利要求1至3中的任何一项权利要求所述的方法，其中通过曲率标度空间表示导出该视图描述符。

5. 如权利要求1至4中的任何一项权利要求所述的方法，还包括导出一个该目标的描述符，该目标描述符与该目标的形状和/或者尺寸相关，而该目标描述符与该目标在该图像中的视图无关。

6. 一种对出现在图像中的目标进行表示的方法以便对用于检索的各个图像或者图像序列标注索引，该方法包括通过处理对应于这些外形的信号，导出对应于该目标在不同视图中的多个二维外形的多个视图描述符，并且将这些视图描述符进行关联以形成一个对各个图像或者图像序列标注索引的目标描述符。

7. 一种对出现在静止或者视频图像中的目标的表示方法，该方法包括导出一个描述符，该描述符与该目标的形状和/或者尺寸相关，而与该目标在该图像中的视图无关。

8. 如权利要求5或者权利要求7所述的方法，其中所述与视图无关的描述符对应于该目标的体积。

9. 通过处理对应于该图像的信号进行图像中目标检索的一种方法，该方法包括以至少一个目标的二维外形的形式输入查询，导出该查询目标的描述符，将所述查询描述符与根据权利要求1至8中的任何一项所述的方法导出的存储的图像中的目标描述符进行比较，并且对至少一个对应于包含目标的图像的结果进行选择 and 显示，各个描述符与该查询描述符之间的比较为该目标指明了该查询与所述目标之间

的相似程度。

10. 如权利要求 9 所述的并根据权利要求 1 至 6 中任何一项的方法，其中以目标的 2 个或者更多的二维外形的形式输入查询，并且其中为每一个所述外形导出一个查询视图描述符，并且其中的比较步骤包括将每一个所述查询视图描述符与每一个在每一个存储的目标描述符中的视图描述符进行比较以导出多个视图相似值。

11. 如权利要求 9 所述的方法，其中对视图相似值进行分析以导出目标相似值。

12. 如权利要求 9 至 11 中任何一项所述的方法，其中至少这些目标描述符中的一些包括根据权利要求 5、权利要求 7 或权利要求 8 所述的方法导出的与视图无关的描述符，并且其中该方法包括输入一个与视图无关的查询值，以及该比较步骤包括将该查询值同该存储的目标描述符的与该视图无关的描述符进行比较。

13. 如权利要求 9 至 12 中的任何一项所述的方法，其中使用该查询的目标外形的曲率标度空间表示导出该查询描述符。

14. 通过处理对应于所述图像的信号来表示出现在图像中目标的一种方法，该方法包括导出对应于该目标的不同二维视图的表示。

15. 一种适合实施权利要求 1 至 14 中任何一项所述的方法的图像检索和恢复系统。

16. 一种用于实施权利要求 1 至 14 中任何一项所述的方法的计算机程序。

17. 一种被编程的根据权利要求 1 至 14 中任何一项所述的方法运行的计算机系统。

18. 一种计算机可读的存储介质，存储用于实施权利要求 1 至 14 中任何一项所述的方法的计算机可执行过程步骤。

19. 一种用于检索在静止或视频图像中的目标的方法，该方法基本上如前文参照附图所述。

20. 一种基本上如前文参照附图所述的计算机系统。

利用形状进行目标表示与检索的方法和装置

本申请为分案申请。母案是 2004 年 4 月 15 日提交的分案申请 200410034396.8，其发明名称同样是“利用形状进行目标表示与检索的方法和装置”，申请日为 2000 年 4 月 28 日。

本发明涉及利用形状对图像中的目标进行表示的方法和装置，特别在检索中使用。本发明还涉及利用形状表示对图像中的目标进行检索的方法和装置。

众所周知，利用出现在图像中的目标的表示来存储静止或者视频图像，例如在图像库中。在检索方法中使用这些表示使得图像包含供以后要恢复的令人感兴趣的目标。这些表示可以基于目标的各种不同的特征包括：色彩、纹理和形状。

已知有许多方法用于表示图像中目标的形状。已知的方法包括链式编码(chain coding)、四树法(quad-tree)和曲率标度表示(curvature scale space representation)法。

为了在图像检索系统中进行检索，用户通过给系统提供被搜寻目标的草图或图像，或者选择一幅存储在系统中目标的视图来输入查询。然后，该系统导出或获得该查询目标的一个表示并且通过适当的匹配算法将该查询表示与存储在数据库中的表示进行比较。将最接近的匹配表示在显示单元上。

在大多数情况下，出现在视频图像中的目标是三维真实目标在二维图像平面上的投影。所以，出现在图像中目标的二维形状或外形依赖于诸如观察位置、观察角度以及拍摄装置和光学系统参数等因素。因此，目标将具有与不同的观察视图相关的不同的外形。

已知的图像数据库系统的一个缺点是具有不同外形的一个目标的不同视图被处理成了不同的目标。作为结果，例如，如果用户输入一个基于目标 A 的前视图的查询，并且该目标 A 只出现在来自后视图和侧视图的视频序列中，结果无法成功匹配并且无法恢复该目标。

本发明提供了一种对出现在数字图像中的目标进行表示的方法，该方法包括导出对应于该目标的多个不同的二维视图的表示。

本发明还提供了一种匹配目标的方法，该方法包括通过处理对应

于目标图像的信号，输入查询和将该查询与目标的不同视图的多个表示进行比较，以发现最接近的一个匹配或一些匹配。

下面将参照以下附图描述本发明的实施方案：

图 1 是根据本发明的实施方案的系统的方框图；

图 2 是展示如何获得针对一个目标的描述符的示意图；

图 3 是展示如何获得针对目标的描述符的流程图；

图 4 是展示用于检索的设备和方法示意图；

图 5 是展示检索方法的流程图。

图 1 展示了一个根据本发明的实施方案的系统。该系统包括控制单元 2 例如用于控制系统运行的计算机、与控制单元 2 相连接用于显示包括图像和文字的输出的显示单元 4 例如监视器、以及给控制单元 2 输入指令的指示装置 6 例如鼠标。该系统还包括存储多个视频序列的数字版本的图像数据库 8 以及存储描述符信息的描述符数据库 10，这将在下文中针对出现在存储图像数据库中的视频序列中的每一个的目标进行更详细的描述。图像数据库 8 和描述符数据库 10 中的每一个数据库都与控制单元 2 相连接。

在该实施方案中，系统元在一个单一位置例如图像库被提供，在该位置上系统部件被永久链接。

参照图 2 和图 3 将对根据本发明的实施方案的目标的描述符的导出方法进行描述。将针对一个具有相对简单形状的的目标，目前情况下是一个圆柱形目标对该方法进行描述。

在该实例中，该圆柱形目标在存储在图像数据库 8 中的视频序列中出现二次。参照图 2，第一次出现时该目标外形对应视图 1，该视图是一个侧视图，第二次出现时该目标的外形对应于视图 3，该视图是侧视图和俯视图。

现在参照图 3 对该方法的步骤进行描述。

数据库索引器(database indexer)认为二个外形属于同一个目标以及该三维目标对应的外形是一个圆柱形(步骤 10)。对于视图 1 和视图 3 中的每一个，采用曲率标度空间(CSS)表示来导出形状描述符(步骤 20)。另外，该数据库索引器选择任意另外的被认为能代表目标的视图，即使它们不在被考虑的视频序列中出现(步骤 30)。在该实例中，选择视图 2 即该圆柱体的俯视图作为代表。也获得了该视图

的 CSS 表示 (步骤 40)。

Farzin Mokhtarian, Sadegh Abbassi, Josef Kittler 等人的“通过曲率标度空间进行加强和有效的形状指标化”不列颠机器版本大会会议录, 53 页 - 62 页, 爱丁堡, 英国, 1964 年 (“Robust and Efficient Shape Indexing through Curvature Scale Space” Farzin Mokhtarian, Sadegh Abbassi, Josef Kittler, Proc. British Machine Vision Conference, pp. 53-62, Edinburgh, UK, 1964) 一文对曲率标度空间表示进行了描述并且结合在此作为参考。简要说来, 形状外形的曲率表示是采用对曲线平滑来进行的。该曲线被考虑成许多的不同的演化阶段。更具体而言, 在演化的每一个阶段的曲率函数中确认曲率的零交叉。通过这些来自所有演化步骤的零交叉组合得到一个曲率零交叉图。该图的一个轴对应 σ , 它代表曲线的演化参数并且另一个轴对应曲线的弧长参数 u 。那么形状由图形中的轮廓线最大值的位置表示。

除了该目标的代表视图的 CSS 形状表示, 还获得了一个独立于该目标视图的全局三维形状参数 (步骤 50)。在该实例中, 该全局参数是目标在现实生活中的体积。该体积可能已知, 或者参照出现在视频序列中的其它目标可以大致估计, 例如可以大致估计人的外形尺寸。

将全局参数与该视图形状描述符组合形成三维目标形状描述符 (步骤 60)。

对所有令人感兴趣的出现在图像数据库 8 中的图像中的目标可以通过上述方式获得三维目标形状描述符。根据任何给定目标的复杂程度来使用该目标的视图的数目以及这些视图。有些目标在三维目标描述符中可能没有全局参数, 例如如果不知道或者不容易导出体积值。每一视图有一个参考指针指明它在视频序列中的哪一帧中出现, 以及是该帧中的哪一个目标, 例如, 目标的视图可以具有一个指针指明它出现在第 1000 帧中并且是第 3 号目标。

在采用不同形状表示方法中的其它实施方案中, 该视图的数目和特征依赖于所使用的形状表示。例如, 对由于视图几何变化引起的形状变形不敏感的方法需要较小的视图数目。

该描述符数据库 10 为存储在图像数据库 8 中的视频序列中的目标存储三维目标形状描述符。

现在参照图 4 和图 5 对在视频序列中进行目标检索的方法进行描述。

用户通过输入查询来启动检索。该查询通过指示装置 6 在显示单元 4 上画出一个或者更多的形状外形（步骤 62）。在该实例中，输入了二个查询外形 100。这些是查询视图。该用户还输入了代表他正在检索的目标的体积参数（步骤 64）。

然后，控制单元 2 对每一个查询视图导出 CSS 视图描述符（步骤 66）。

在另一个实施方案中，用户通过选择一个查询形状或者从由控制单元 2 在显示单元 4 上显示的形状菜单中选择形状来输入查询。在该实施方案中，有可能在描述符数据库 10 中已经能够获得该视图描述符。

该查询体积参数与该查询视图描述符被组合形成一个三维的目标描述符。

然后系统为确定该查询目标描述符与存储在描述符数据库 10 中的目标描述符之间的相似性进行匹配操作，这在下文中被描述为模型目标描述符（model object descriptor）。在描述符数据库中有选择器 205 按照顺序选择每一个模型目标描述符（步骤 68）并且对每一个目标描述符按照其顺序进行以下步骤。

首先，由比较器 200 获得一个全局相似测度（global similarity measure）GS（步骤 70），对查询目标描述符和模型目标描述符使用该全局参数。在该实施方案中，通过将查询体积参数与模型体积参数之比作为从数据库中采用的描述符来导出 GS。如果该比位于间隔 $(1/c, c)$ 中其中 $c > 1$ ，那么认为目标相似并且 GS 取值为 0。不然的话，GS 取值为无穷大。C 的值根据应用而定。例如，对电影片， $c = 5$ 。

如果至少查询描述符和模型目标描述符中的一个不具有全局参数值，那么 $GS = 0$ 。

如果 $GS \neq 0$ 那么从该描述符数据库中选一个新的模型目标描述符；如果 $GS = 0$ ，那么进行如下的视图描述符比较（步骤 72）。

利用匹配功能在比较器 810 中将每一个查询视图描述符与每一个针对被考虑的模型目标描述符的视图描述符进行比较以导出视图相似测度（步骤 74）。第 i 个查询视图描述符与第 j 模型视图描述符的比

较导致产生一个视图相似测度 s_{ij} 。

更详细的描述如下，利用选择器 600 选择该查询视图描述符并且利用选择器 700 选择该数据库视图描述符。首先，将针对第一查询视图的视图描述符与来自数据库的模型目标描述符中的每一个视图描述符进行比较。为每一对利用合适的匹配算法计算视图相似值 s 。在本实施方案中，利用在上述 Mokhtarian, Abbasi 和 Kittler 一文中所描述的匹配算法计算视图相似值 s 。在使用该具体的相似测度时，该视图相似值越小，该匹配就越接近。这导致对该第一查询视图产生一组 k 个视图相似测度，这里 k 是被考虑模型目标描述符中的视图描述符的数目，并且该 k 个测度被存储在全局和局部相似组合器 (combiner) 820 中。

然后计算视图相似值并将其存储以便采用相似的方式对第二查询视图描述符和来自数据库的模型视图描述符进行处理，由此得到另外 k 个视图相似测度。

对于每一个查询视图，选择最小的视图相似值给被考虑数据描述符 (步骤 76)。该最小值是各自查询视图与被考虑的目标描述符中的视图中的一个之间匹配最为接近的测度。这导致产生 p 个最小视图相似值，这里 p 是查询视图的数目。在该实例中， $p = 2$ 。

针对该查询描述符和该被考虑模型目标描述符采用一个总相似测度 S 作为 p 个相似值的媒介 (步骤 78)。它表示该查询描述符与考虑了所有视图的模型目标描述符之间的匹配接近程度。因此，如果一个查询视图与该数据库描述符中的一个视图紧密地匹配而其它查询视图不能与数据库描述符中的任何视图紧密地匹配，那么这在 S 中以一个中间值反映出来。

对描述符数据库 8 中的每一个目标描述符重复上述步骤，结果产生 n 个相似测度 S ，其中 n 是该描述符数据库中的三维目标描述符的数目 (步骤 80)。然后从指明最接近的最低值开始对这 n 个相似测度进行排序 (步骤 82)。然后选定 m 个最低值，其中 m 是由用户选定或者控制单元装置确定的一个值，并且在显示单元 4 上显示来自对应于包括该目标的 m 个视频序列中的每一个的图像 (步骤 84)。

根据本发明，将一个单一目标的多个视图进行存储以形成一个完整的或者半完整的外形形状的描述。如果仅存储被认为是对恢复重要

的视图，那么该表示可以半完整的。例如，对典型的存储正片的数据库，只将汽车的前、后和俯视图作为统一的表示来存储，但是没有将汽车的底视图存储，因为不可能采用该视图作为查询。

根据本发明的系统可以提供给，例如图像库。另一种情况，该数据库与系统控制单元的距离可以很远，该数据库通过临时连接如电话线或者互联网与控制单元相连接。可以在永久存储或者便携式数据存储介质诸如，CD-ROM 或者 DVD 中提供图像和描述符数据库。

所述系统的部件诸如选择器和比较器可以以软件或者硬件的形式提供。尽管对本发明以计算机系统的形式进行了描述，但是它可以以其它的形式实施，例如采用专用芯片。

虽然给出了表示二维形状目标的方法以及计算表示 2 个形状之间相似值的方法的具体实例，但是可以采用任何适合的这种方法。

构成完整或者半完整的形状描述的各种目标视图可以由例如制片商提供。例如，在影片中包括一辆汽车，制片商可以安排拍摄 20 张不同的汽车视图供在数据库中对影片标注索引使用。另一种情况，可以在目标的所有不同的视图之间按照顺序提供链接使得可以找到该视图，并且当包括了该目标的第一幅视图的数据时，例如，可以提供不出现在该序列中的任何有用视图的形状。

例如，本发明还可以用于出于检验或者过滤的目的而进行的图像匹配。

本发明可以应用于单一图像，以及按照图像顺序的图像，诸如来自影片或者视频图像的图像，或者以某种方式相关联的图像集，诸如位于相同网页上的图像。

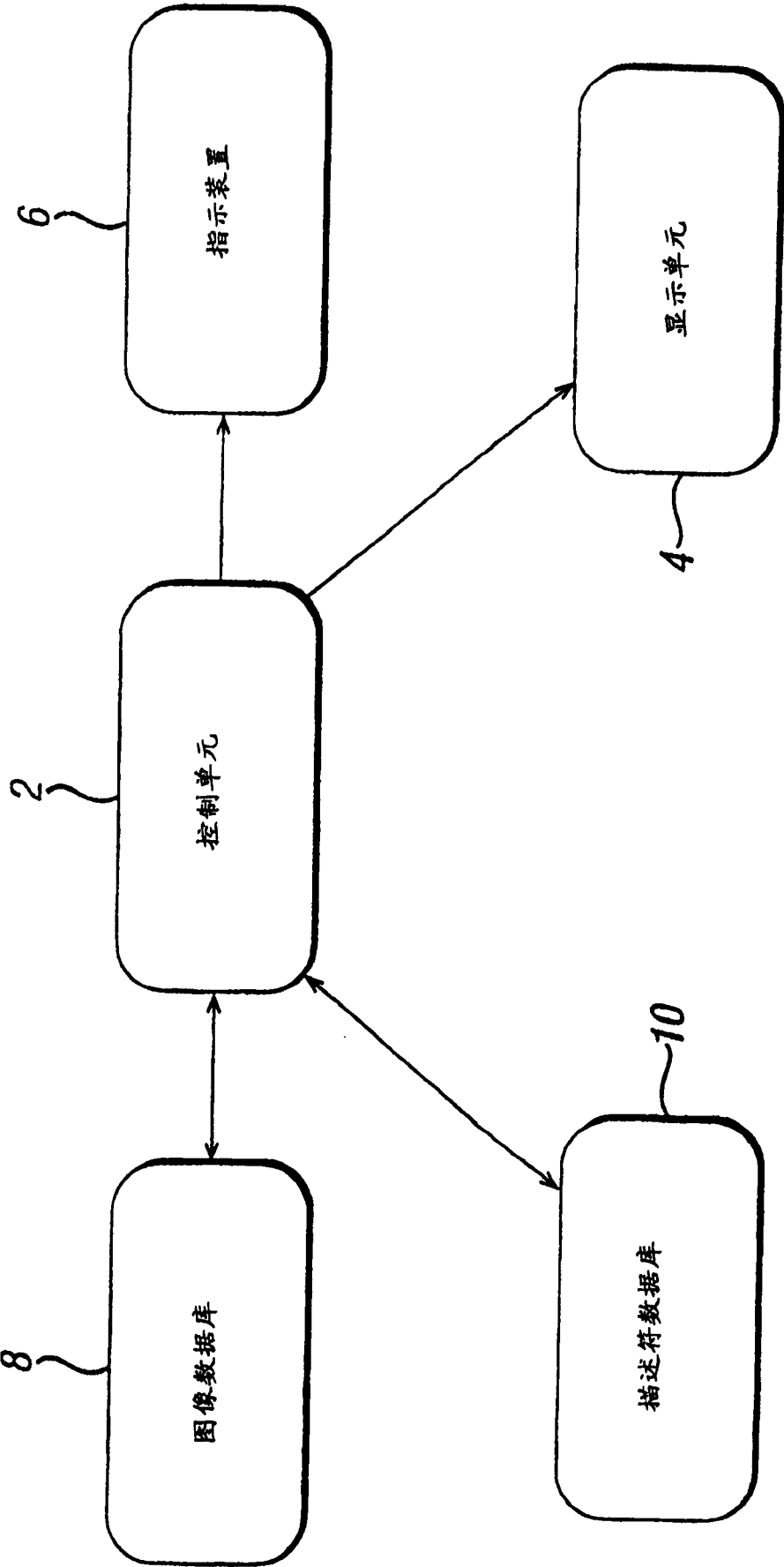


图 1

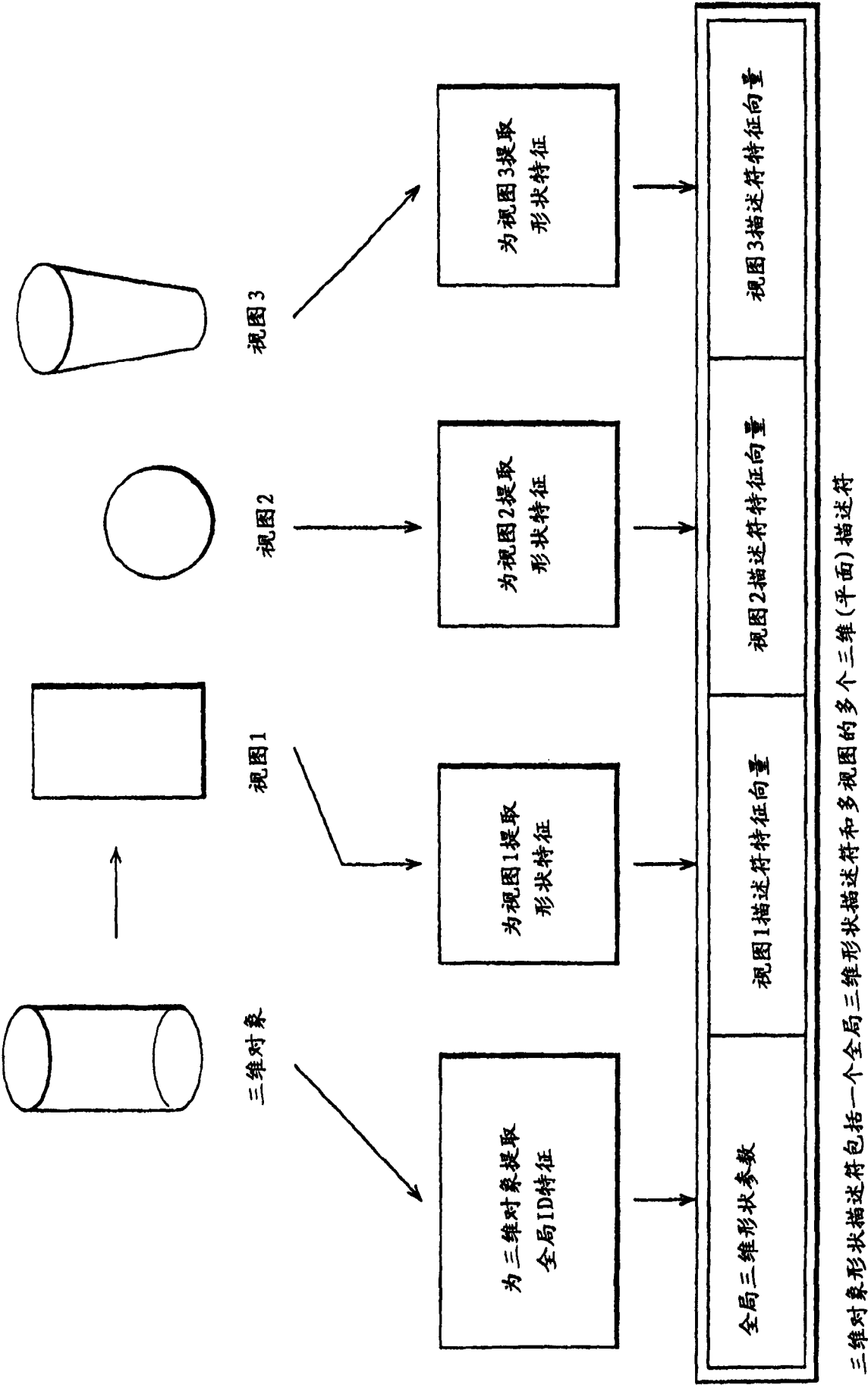


图 2

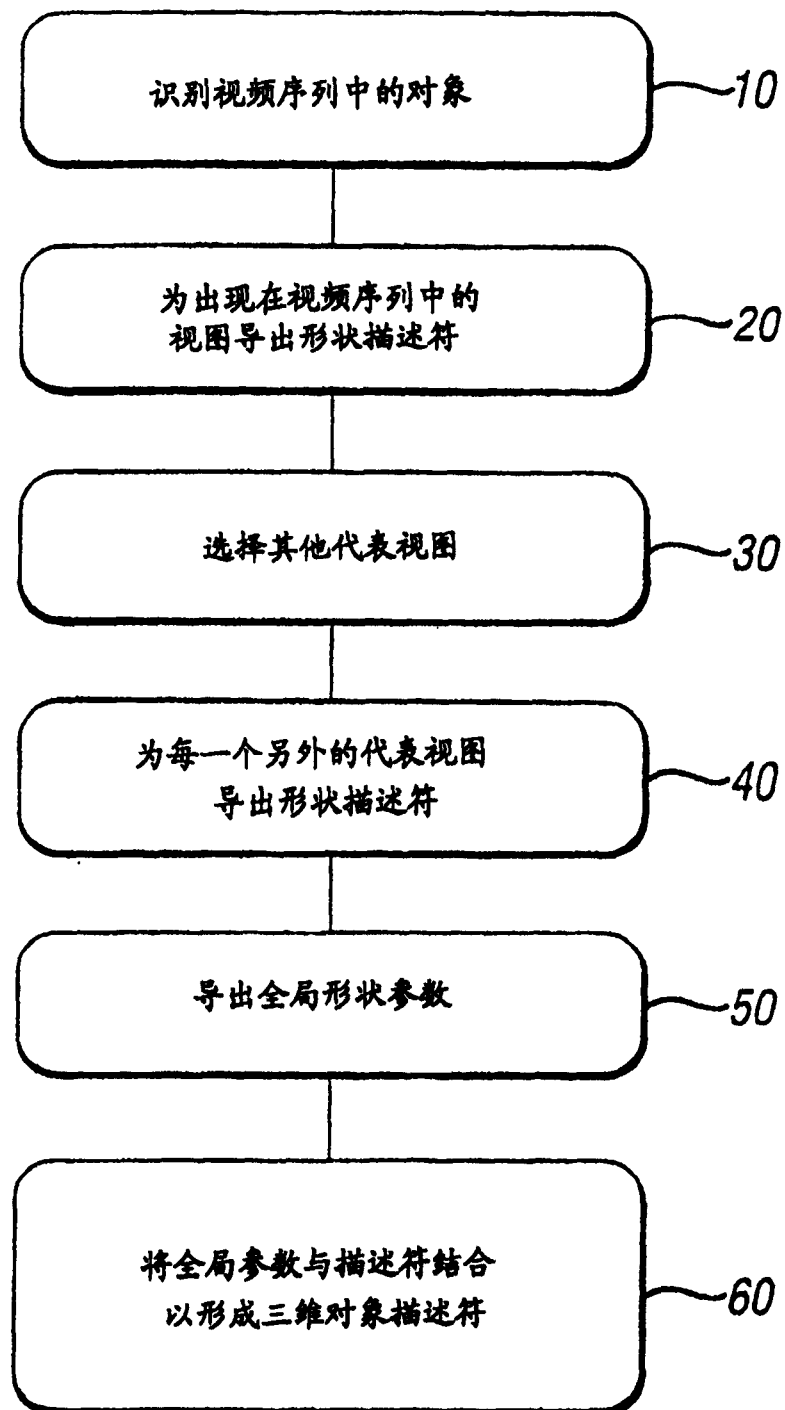


图 3

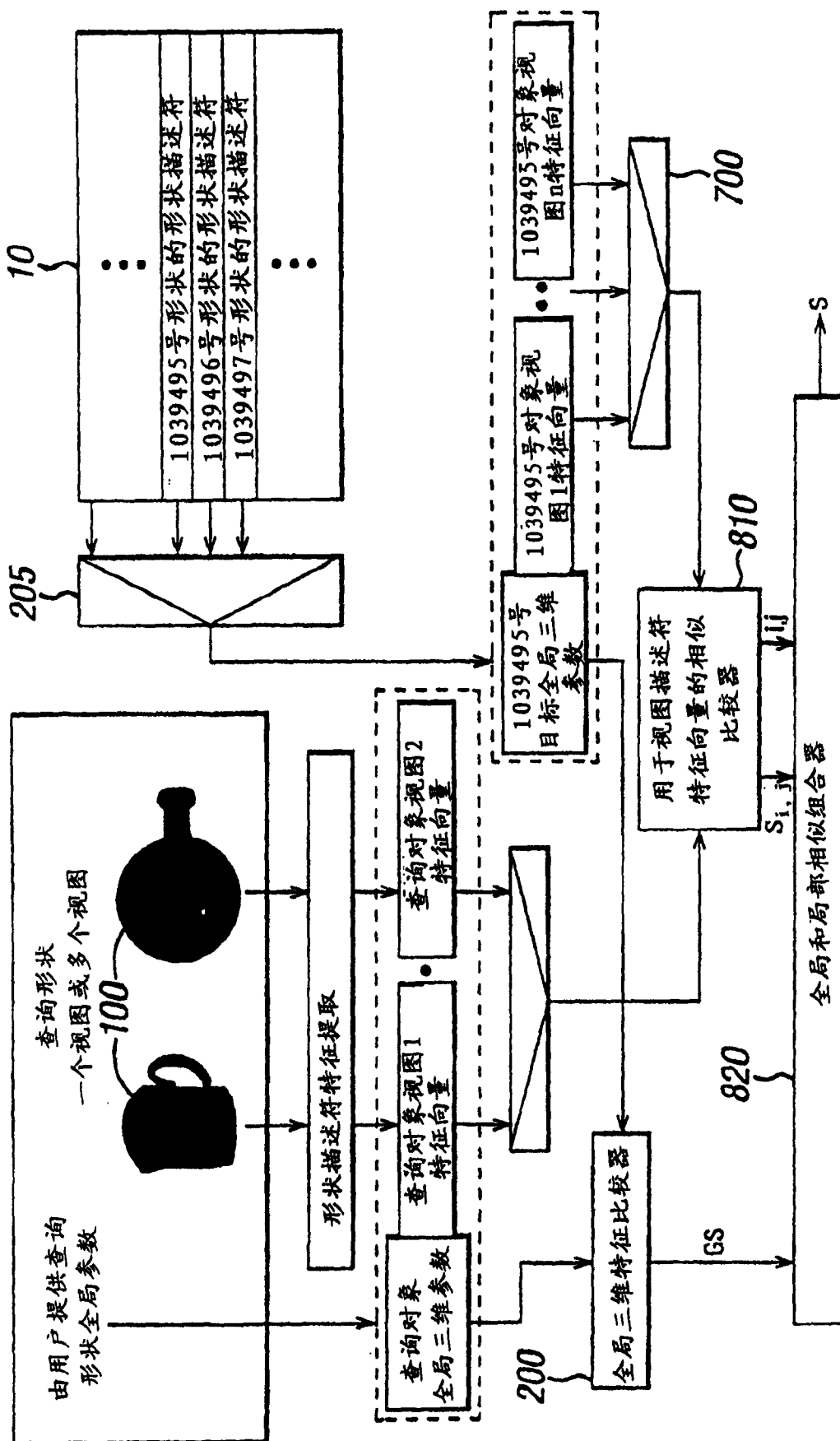


图 4

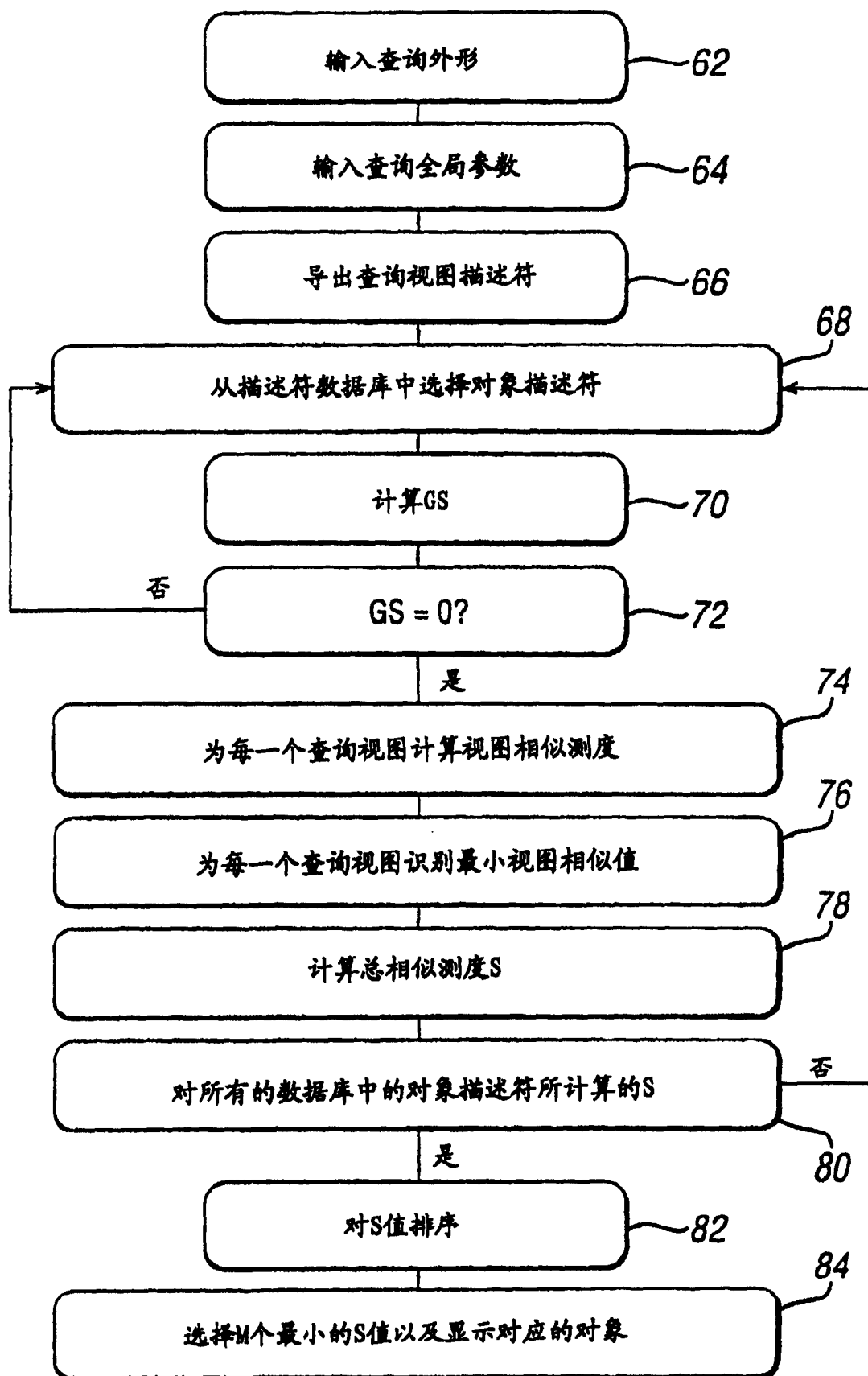


图 5