

[12] 发明专利说明书

[21] ZL 专利号 95192084.7

[45] 授权公告日 2002 年 7 月 10 日

[11] 授权公告号 CN 1087549C

[22] 申请日 1995.2.27

[21] 申请号 95192084.7

[30] 优先权

[32] 1994.3.14 [33] IL [31] 108,957

[86] 国际申请 PCT/US95/02424 1995.2.27

[87] 国际公布 W095/25399 英 1995.9.21

[85] 进入国家阶段日期 1996.9.13

[73] 专利权人 斯戴尔技术有限公司

地址 以色列赫泽列

[72] 发明人 海姆·克莱特曼 丹·巴-阿尔

约尔·阿米尔 胡德·特劳西

[56] 参考文献

US 5264933 1993.11.23 H04N5/262

US 5353392 1994.10.4 H04N5/262

审查员 张龙哺

[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

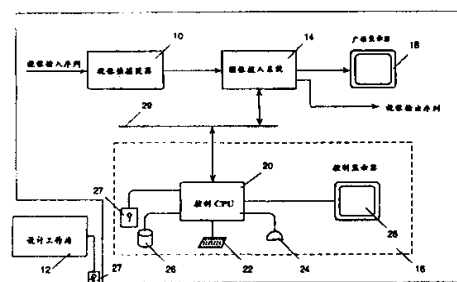
代理人 杨国旭

权利要求书 4 页 说明书 18 页 附图页数 28 页

[54] 发明名称 一种将图像植入视频流的方法和系统

[57] 摘要

公开一种将广告之类的图像与出现在相对不变的空间例如运动场的动作的视频流相混合的系统和方法。系统利用背景空间的模型改变视频流,使在背景空间的某些位置包括有该图像。系统包括一个视频帧捕捉器(10)和一个图像植入系统(14)。帧捕捉器(10)一次捕捉视频信号的一个单帧。图像植入系统(14)一般将图像植入帧里面背景空间的一个预选场面的预定部位(如果该部位被显示在帧中的话)。



ISSN 1008-4274

权 利 要 求 书

1.一种用于一次将一个图像植入表示出现在背景空间的一连串动作的多个视频帧的一个被选帧的方法，背景空间具有一个固定的平面场面并且被至少一架摄像机扫描，方法包括的步骤是：

生成与所述多个视频帧无关的所述固定场面中的一个被选场面的模型，所述模型包括表征所述场面的几何特征的一种图形；以及

利用所述模型将所述图像植入所述帧内，所述利用步骤包括用透视图法改变所述模型形状的步骤。

2.根据权利要求1的方法，其中所述几何特征包括至少一个由直线和弧线组成的线族。

3.根据权利要求2的方法，还包括提供各条所述直线的弧线之间的平面关系的表示法的步骤。

4.根据权利要求1的方法，其中所述图形是一种平面向量图形。

5.用于一次将一个图像植入表示出现在背景空间的一连串动作的多个视像帧中的一个被选帧的装置，背景空间具有固定的平面场面并被至少一架摄像机扫描，装置包括：

用于生成与所述多个视频帧无关的所述固定场面中的一个被选场面的模型的装置，所述模型包括表征所述场面的几何特征一种图形；以及

利用所述模型将所述图像植入所述帧的装置，所述利用装置包括以透视图的方法改变所述模型形状的装置。

6.根据权利要求5的装置，其中所述几何特征包括至少一个由直线和弧线组成的线族。

7.根据权利要求6的装置，还包括用于提供各条所述直线和弧线之间的平面关系的表示法的装置。

8.根据权利要求6的装置，其中所述图形是一种平面向量图形。

9.一种用于一次将一个图像植入表示出现在背景空间的一连串动作的多个视频帧中的一个所选帧的方法，背景空间具有固定的平面场面并且被至少一条摄像机扫描，方法包括：

生成所述固定场面中一个被选场面的模型。所述模型包括表征所述

场面的几何特征的一个图形;

利用所述模型以透视图方式改变所述图像形状,以把它装入总体上同样的所述被选帧的透视图;

产生所述被选帧的一个背景掩罩,所述帧包括至少所述固定场面的一部分,所述背景掩罩确定所述被选帧的可改变区域和它的不可改变区域;以及

将所述以透视图方式变了形的图像共混到所述可改变区域的所述被选部分,从而将所述图像植入所述被选帧中。

10. 根据权利要求 9 的方法,其中所述生成所述模型包括生成与所述多个视频帧无关的所述模型。

11. 根据权利要求 9 的方法,其中所述利用包括利用所述模型和一个被选来用于产生变换矩阵的可改变的区域,所述变换矩阵以透视图方式改变所述图像形状。

12. 根据权利要求 9 的方法还包括在所述利用步骤之前产生一个确定所述图像与所述模型的相对位置掩罩,从而使所述图像位置掩罩随所述图像变换。

13. 根据权利要求 12 的方法,其中所述共混步骤包括将所述变形了的位置掩罩和所述背景掩罩的被选可变区域相加,因此产生一个确定所述帧中象素的允许掩罩,所述图像的对应象素被植入其中。

14. 根据权利要求 13 的方法还包括产生一个以透视图方式变形的共混掩罩,所述共混掩罩确定透明度数值并且其中所述共混步骤还包括将所述图像,所述允许掩罩和所述共混掩罩相混合,从而将所述图像以所要求的色值植入。

15. 根据权利要求 14 的方法,其中所述混合包括向每个象素 (x,y) 分别一个色值,其中

象素 (x,y) 的色值 = $\beta(x,y) * \text{图像}(x,y) + (1 - \beta(x,y)) * \text{帧}(x,y)$

其中 $\beta(x,y) = \alpha(x,y) * P(x,y)$

其中 $\alpha(x,y)$ 和 $P(x,y)$ 为象素 (x,y) 的共混和允许掩罩数值;
以及

其中图像 (x,y) 为象素 (x,y) 的所述图像数值以及所述帧 (x,y) 为象素 (x,y) 的所述被选可改变区域数值。

16. 一种用于一次将一个图像植入表示出现在背景空间的一连串动作的多个视频帧中被选的一个帧的系统，背景空间具有固定的平面场面并且被至少一架摄像机扫描，系统包括：

用于生成所述固定场面中一个被选场面的模型的装置，所述模型包括表征所述场面几何特征的一个图形；

用于利用所述模型以透视图方式改变所述图像形状，以将它装入总体上同样的所述被选帧的透视图中的；

用于产生所述被选帧的背景掩罩的装置，所述帧包括至少一部分所述固定的场面，所述背景掩罩确定所述被选帧的可改变区域和它的不可改变区域；以及

用于将所述以透视图方式变了形的图像共混到所述可改变区域的所述被选部分中，从而将所述图像植入所述被选帧的装置。

17. 根据权利要求 16 的系统，其中所述生成所述模型包括生成所述不依赖于所述多个视频帧的模型。

18. 根据权利要求 16 的系统，其中所述利用装置包括利用所述模型和一个被选的可变区域产生变换矩阵的装置；所述变换矩阵以透视图方式改变所述图像的形状。

19. 根据权利要求 16 的系统还包括用于在所述利用装置工作之前产生确定所述图像与所述模型相对位置的图像位置掩罩，从而使所述图像位置掩罩随所述图像变形的装置。

20. 根据权利要求 19 的系统，其中所述用于共混的装置包括用于将所述变形的图像位置掩罩和所述背景掩罩的所述被选可变区域相加，由此产生一个确定所述帧中象素的允许掩罩（所述图像的对应象素被植入其中）的装置。

21. 根据权利要求 20 的系统还包括用于产生一个以透视图方式变形的共混掩罩的装置，所述共混掩罩确定透明度数值，并且其中所述用于共混的装置还包括用于将所述图像、所述允许掩罩和所述共混掩罩相混合，从而以所要求的色值植入所述图像的装置。

22. 根据权利要求 21 的系统, 其中所述用于混合的装置包括向每个像素 (x,y) 分配一个色值的装置, 其中

像素 (x,y) 的色值 = $\beta(x,y) * \text{图像}(x,y) - (1 - \beta(x,y)) * \text{帧}(x,y)$

其中 $\beta(x,y) = \alpha(x,y) * P(x,y)$

其中 $\alpha(x,y)$ 和 $P(x,y)$ 为像素 (x,y) 的共混和允许掩罩数值;
以及

其中图像 (x,y) 为像素 (x,y) 的所述图像数值以及所述帧 (x,y) 为像素 (x,y) 的所述被选可改变区域数值.

说明书

一种将图像植入视频流的方法和系统

本发明一般地涉及将一个准备好的图象与视频信号并接的技术。

体育竞技场一般包括一个进行比赛的比赛区，一个坐着观众的座席区以及一个将这两个区隔开的某种隔板。一般，隔板至少部分地被赞助比赛的公司的广告所覆盖。当拍摄比赛时，隔板上的广告被作为体育竞技场的一部分拍摄下来。除非电视摄像机专门拍摄，否则广告不能充分地公众显示。

已经知道几种将预先规定的广告并接到体育竞技场的电视场面上的系统。一种系统有一个操作员确定竞技场的一个目标场面。然后系统锁定在目标场面上并将一个预定的广告与相应于该场面的视频流的部分并接。当摄像机不再朝向该场面时，系统便失去目标场面并且操作员必须再一次指定使用哪个场面。

上述系统是实时操作的。其他已知的系统基本执行同样的操作，不过不是实时。

已经知道其他几种用于将数据并接到视频序列的系统。这些系统包括将一个图像插在视频场之间，在电视帧的一个固定位置叠加图像数据（例如电视台徽标的数据）以及甚至电子插入图像数据“替代”专用目标广告牌。后一系统采用彩色键控一类的技术实现。

U.S.5,264,933 叙述了一种改造电视图像的装置和方法，能够使所加入的广告图像成为原显示图像的一部分。操作员选择在所捕获的图像的什么地方植入广告图像。U.S.5, 264, 933 的系统也能根据音频信号例如解说员的标准词句在所选择的主要广播区内植入图像。

PCT 申请 PCT/FR91/00296 叙述了用于在接连的图像中修改一个区域的过程和装置。各图像均显示出一个不变形的目标区（附近有登记标志）。系统搜索出登记标志并利用它确定区域的位置。然后，事前准备的图像就可被叠加到该区域上。登记标志是目标区内或其附近的任何易

于识别的标志（例如十字或其他语义图）。PCT/FR91/00296 的系统以许多个分辨率产生被捕获图像，并在其识别过程中利用这许多分辨率。

本发明的一个目的是提供一种将图像例如广告与出现在相对不变的空间中的活动视频流相混合的系统和方法。这样的空间可以是一个运动场或球场，一个舞台或一个房间，并且其位置一般都是在活动（即比赛或演出）之前就选定的。图像被“植入”到背景空间的被选场面上（名词“植入”在这里意味着图像被混入到显示被选场面的视频流的一部分上面。

具体地说，本发明利用顾及到背景空间对视频流的变化先验信息以包括背景空间里面某些位置的图像。无论在视频流中出现背景空间的哪一个透视图对于系统和方法操作来说都不重要。

根据本发明的一个优选实施例，系统最好包括一个视频帧捕捉器和一个图像植入系统。帧捕捉器一次捕捉一单帧视频信号。图像植入系统一般将广告图像植入到帧中背景空间一个预选场面的预先规定的部位（如果该部位被显示在帧中的话）。为了确定接受植入的部位的位置，图像植入系统包括一个接收单元，用于接收 a)背景空间的固定场面的不透明模型和 b)一个图像掩罩，它指示不透明模型要被混入图像的部位。图像植入系统通过模型识别该部位是否在帧内显示以及显示在什么地方。

此外，根据本发明的一个优选实施例，系统还包括一个设计工作站，图像和表示预选场面的图像掩罩可在工作站设计。

再者，识别过程最好包括 a)审查帧并从中提取固定场面的特征以及 b)确定一个在模型和提取的特征之间的透视图变换。

再又一点，审查和提取过程包括创建一个背景掩罩和一个前景掩罩。背景掩罩表示帧里面背景要素中感兴趣的特征的位置并被用于提取所要求的特征。前景掩罩由帧里面必须保留不变的前景要素组成。

另外，根据本发明的一个优选实施例，植入过程包括的步骤是：a)使用透视图变换变换图像、图像掩罩和共混掩罩（有选择地）；b)将所变换的图像、图像掩罩和可选择共混掩罩与帧和前景掩罩混合。正如上文所述，前景掩罩表示不被所变换的图像遮盖的前景数据的位置。

再者，系统最好包括一个查找表，用于将帧里面的中间多种颜色转换为感兴趣的特征的颜色、背景要素的颜色和表示前景要素的颜色中的一个。查找表最好通过让用户指定相关颜色来创建。如果相关颜色不再表示感兴趣的特征和背景（典型地是由于照明变化），用户可指定表示所要求的要素的新的颜色，从而校正查找表。

再又一点，根据本发明的一个优选实施例，查找表被用来创建表示帧里面感兴趣的特征、背景要素和前景要素的位置的帧的背景和前景掩罩。

根据本发明的一个示例性实施例，特征就是一些直线。在一个实施例中，这些直线通过霍克变换法提取。在另一个实施例中，则通过确定直线线段的角来提取。选出感兴趣的象素并打开一个邻域。再划分邻域并选出具有最大活力的扇区。然后延伸并划分所选出的扇区。根据需要反复进行这一过程。

此外，根据本发明的一个优选实施例，系统将所提取的特征映射到一个渐近函数上，以确定哪些特征是平行直线的透视图形式。

再者，根据本发明的示例性实施例，背景空间是一个其上标有直线的体育竞技场。系统具有一个体育竞技场模型，并且最好有一个模型中各长方形及其拐角的清单。系统最好执行以下操作：

- a) 由提取的特征中选出两条垂直线和两条水平线并定出它们的交点；
- b) 生成一个由模型的每个长方形拐角点至特征交点的变换短阵；
- c) 使用每个变换矩阵变换模型；
- d) 利用背景掩罩的背景要素，使每一个已变换的模型与帧符合；以及
- e) 选出与帧特征符合最佳的变换矩阵。

此外，根据本发明的一个示例性实施例，摄像机参数可被用来减少识别体育场地所需的帧里面的直线数目。对于本实施例，存在有下面的操作：

接收或提取摄像机的坐标；

将当前变换矩阵表示为坐标、俯仰、偏转和焦距矩阵的剩积，然后

确定俯仰、偏转和焦距的数值；以及

识别具有算出的俯仰、偏转和焦距数值的摄像机并将该信息存储起来；以及

每当在视像中出现新的切换，便重复接收、表示和识别步骤。

现在对视像流里面的任一帧均能作为类似于前面的帧，或者作为一架识别过的摄像机拍摄的新切换部分进行处理。

通过下面结合附图所作的详细叙述，本发明将得到更全面的理解和正确评价。在附图中：

图 1 是一个方框图，说明根据本发明的一个优选实施例构造和工作的用于将图像植入视频流的系统；

图 2 是一个网球比赛的概略图，用作一个说明图 1 的系统工作的例子；

图 3 是一个用于理解图 1 系统工作的网球场模型图；

图 4A 是要被植入的图像的一个例图；

图 4B 是用于图 4A 的图像和图 3 的模型的图像区域掩罩的例图；

图 4C 是用于图 4A 的图像和图 3 的模型的共混掩罩的例图；

图 5 是一个方框图，说明图 1 系统的图像植入单元形成部分的诸部件；

图 6 是一个示例性视频帧的例图，图 4A 的图像要被植入其中；

图 7 是一个根据图 6 的视频帧生成的背景掩罩的例图；

图 8 是一个方框图，说明图 5 的图像植入单元的特征识别单元形成部分的工作；

图 9A 是一个流程图，说明特征提取方法；

图 9B 是用于理解图 9A 的方法的背景掩罩的部位例图；

图 9C 是图 9B 的背景掩罩的子扇区的直方图，用来理解图 9A 的方法；

图 10 是一个方框图，说明图 5 的图像植入单元的透视图识别单元形成部分的工作；

图 11A 是由图 7 提取的特征的交汇点的例图；

图 11B 是透视图平行线由于计算的不精确交汇在不同的点的例图；

图 12A 和 12B 是心射投影的例示图，用于理解图 10 透视图识别单元的工作；

图 12C 是用于图 12A 和 12B 的心射投影的示例函数的曲线图；

图 13 是一个详细方框图，说明 10 所示的工作过程；

图 14A 和 14B 用于理解图 13 的工作过程；

图 15 是使用变换矩阵的例图；

图 16 是用于理解四边形与几何模型之间的符合过程的例图，用来理解图 13 的工作过程；

图 17 是一个方框图，说明图 5 的图像植入单元的变换器和混合单元的工作过程；

图 18 是一个方框图，说明用于修正图 5 的图像植入单元中使用的查找表一种校正方法；

图 19 是摄像机参数的简略说明；

图 20 是一个流程图，说明当图 19 的摄像机参数已知或可算出时变换矩阵的工作过程；

图 21 是一个用于图 20 所示过程中的表格的例图；以及

图 22 是一个流程图，说明当摄像机参数已知或可算出时的工作方法。

现在参照图 1 说明一种用于将广告之类的图像和出现在相对不变的背景空间中的活动视频流相混合的系统。图像被植到背景空间的一个被选场面上。系统将被叙述在网球比赛（还参照图 2 说明）的电视播放过程中，不言而喻，本发明工作于活动出现的场面事前已知和可识别的所有情形。

本发明的系统一般包括一个视频帧捕捉器 10，用于将输入的视频序列（例如网球比赛）转换成视频帧；一个设计工作站 12，用于设计要被植入到视频帧内看得见的被选场面（例如网球场）上的图像（例如广告）；一个图像植入系统 14，用于将视频帧与所设计的图像并接；一个计算机控制系统 16，用于控制操作员的操作并为操作员提供向图像植入系统 14 的输入；以及一个广播监视器 18。

计算机控制系统 16 一般包括一个中央处理单元（CPU）20，一个

键盘 22，一个鼠标 24，一个磁盘 26，一个可更换的介质驱动器（例如软盘 27）以及一个显示器 28。显示器 28 一般用构成 CPU 20 的部件的图形适配器驱动。设计工作站 12 一般还包括诸如软盘 27 一类的可更换介质驱动器。

计算机控制系统 16 和图像植入系统 14 一般通过总线系统 29 通信。设计工作站和图像植入系统 16 一般通过可更换的介质通信。

视频序列可从任一个信号源例如从录像带，通过卫星、微波或任何其他形式视频通信等等的远地发射台接收。如视频序列由卫星提供，则系统不控制视频速率。因而图像植入系统 14 必须在视频速率内完成其操作，卫星视像流的各帧之间的视像速率一般为 30ms。如果序列来自录像带，则系统能够控制视频速率并以任何要求的速度工作。

视频序列原本产生于比赛场。正如由图 2 所能见到的，对于网球比赛一般有两架电视摄像机 30 拍摄网球场 32 上的活动。电视摄像机 30 的位置一般不动。

球场 32 被网 34 分成两个半场。每个半场具有多个一般漆成第一绿色色调的区域 36，各区域 36 被一般漆地白色的线条分开。外场区 40 一般被漆成第二绿色色调。

在现实中，直线 38 是一些平行线和垂直线。由于摄像机 30 是从一个角度而不是从上方推向活动体，它所接收到的活动图像都是透视图像。于是，在摄像机 30 的视像输出中，平行线 38 显得好像会聚于有限远处。视频输出的透视图的角度随摄像机 30 的角度和焦距量改变。

本发明将在各个透视图角度和焦距量的情形下把一个图像 42（例如文字“IMAGE”）植在被选背景场面上的一个要求的位置。对于网球场来说，可能的位置是由四条直线 38 确定的网球场 32 的一个半中的任何长方形。如图 2 所示，图像 42 不影响选手 44 的活动；图像 42 显得好像是漆在球场场面上的。

在现实中，由于球场 32 的色调和球场 32 内的直线 38 的位置均不变，如果图像植入系统有一个包括有要被植入图像位置的比赛空间的模型，并至少能够识别拍摄角度和焦距量，就能够将图像并接到视频序列中，使得看起来好像图像被植在所要求的位置。为了做到这一点，图像

植入系统还需知道摄像机所看到的球场的颜色。这些颜色能够随照明（日光或人造光）的变化而变化。

现在附带参照附图 3 说明网球场的几何模型 50，以及参照图 4A、4B 和 4C 说明植入设计师准备的数据。

植入设计师工作在设计工作站 12（例如由以色列赫兹利亚的西太克斯公司制造的 BLAZE 工作站），一般有一个顶视图形式的网球场几何模型 50。模型 50 一般为球场 32 的一个按比例缩小形式，它指出其要被植入系统 14 识别的要素（例如直线 38）。其他的运动场可以包括圆弧或别的明确定义的曲线。其他的识别要素包括直线 38 的交点 54。

植入设计师设计要被植入的图像 42（示于图 4A）并决定把它放在模型 50 上的什么地方。若干可能的位置 52 示于图 3。然后，设计师准备一个图像位置掩罩 56（图 4B）以识别图像 42 要被放在模型 50 中的什么地方。掩罩 56 在模型 50 中要被放置图像 42 的位置处透光而在其他各处则不透光。

由于图像 42 可能是色彩鲜明的，因而可能要求的不是植入图像本身而是其经过柔和处理的形式，以便不致明显干扰球场 32 上的活动。因而，植入设计师可以有选择地准备一个共混掩罩 58（图 4C），表示图像 42 在位置掩罩 56 所指出的植入位置处应当怎样地与球场 32 的颜色共混。共混掩罩 58 可以是这一技术中已知的任何适合掩罩。在图 4C 中，示出掩罩 58 具有四个区 59，每一区表示球场颜色的最终不同数量，其中外区 59 一般较内区掺有更多的球场颜色。

现在回过头来参照图 1 和 2。由几何模型 50、图像 42、图像位置掩罩 56 以及可选择的共混掩罩 58 组成的植入数据一般都在相应的网球比赛之前准备好，并且一般都通过可更换的存储介质提供给图像植入系统 14 用于在比赛发生时植入到所输入的视频序列中。

大多数电视直播比赛的视频序列都用一个初始化的操作序列开头，使本地电视台的操作人员能够将自己的系统与输入的序列同步。这一点对录像带数据也是正确的。

在本发明中，初始化的视频数据被帧捕捉器 10 捕捉，并被首先送到计算机控制系统 16。电视台操作员选出一个具有清楚的赛场景像的帧，

并如下文所述，用它提供校准信息。校准信息被图像植入系统 14 用来识别球场 32 及其特征（例如直线 38）。在下文叙述的实施例中，校准信息包括背景中所感兴趣的特征例如场地线、运动场（球场 32）以及运动场以外的地面（外球场区 40）的颜色。可以接收到的其余颜色均被定为前景颜色。别的运动场可以要求或多或少的特征作定义，并因而要求或多或少的颜色。

电视台的操作员利用鼠标 24 和键盘 22 交互定义校准信息。这一点可用许多种方法达到，下面将叙述其中的一种。一个四色层被叠加在控制显示器 28 当前显示的帧上。开始时，这个四色层仅包括一种颜色 - 一种透明的颜色。于是，这一当前帧在开始时是可见的。

操作员指定描述三个特征即直线 38、内运动场 36 以及外运动场 40 之一的象素。当他选出一个象素时，叠加层中对应于当前帧中具有被选颜色的象素的那些象素被着以单一的已调换颜色，因而覆盖当前帧中它们所对应的象素。所选的颜色被存储。对所有三个区重复这一过程。所有未选中的颜色被指定为第四调换色。

如果操作员认可这最后的四色层，则由当前帧选出的颜色之间的查找表（LUT）和调换色便产生出来了。

如果需要，计算机控制系统 16 能够将操作员为以后在 LUT 校正周期使用选出的象素存储起来，下文中将参照图 18 叙述。

计算机控制系统 16 向图像植入系统 14 提供由 LUT 构成的帧数据和用于产生 LUT 的象素。系统 14 利用上述帧数据识别输入视像信号的每一帧里面的所要求的特征。

现在参照图 5 说明图像植入系统 14 的通用部件。也参照用于理解系统 14 的工作的图 6 和 7。

系统 14 一般包括一个用于识别在每个输入视频帧中存在球场 32 的哪些特征的特征识别单元 60（图 5），以及用于识别活动摄像机 30 的拍摄角度和焦距并确定模型 50 和输入视频帧之间的适当透视图变换的透视图识别单元 62。系统 14 还包括一个用于将植入数据由模型平面变到图像观察平面的变换器 64 以及一个用于将透视图植入数据与当前视频相混合从而将图像 42 植到球 32 上的混合器 66。

正如下文中更详细叙述的一样，特征识别单元 60 利用 LUT 创建出输入帧的一个背景掩罩，指出帧的哪一部分具有可能感兴趣的背景特征以及哪一部分是前景，因而在后面的操作中是不会被改变的。图 6 和 7 分别提供一个示例性输入帧 68 及其相应的背景掩罩 70。

图 6 的输入帧 68 在球场 32 上有两个选手 44。图 7 的背景掩罩 70 示出各个四色区。标为 1 - 4 的区域分别为直线颜色、内场颜色、外场颜色以及保留颜色的区域。注意选手 44 的区域被标为背景颜色 4 并遮盖别的重要区域（例如白色直线 1 的那些区域）。

单元 60（图 5）由背景掩罩 70 提取运动场的特征。对于网球场来说，所感兴趣的特征是直线 38。适视图识别单元 62 将所提取的特征与模型 50 的那些相比较，并从中产生出一个变换矩阵。

通过使用变换矩阵，变换器 64 将图像植入数据（即要被植入的图像 42、图像位置掩罩 56 以及共混掩罩 58）转换为输入视频帧的透视图。

最后，通过使用已变换的图像位置掩罩 56 和背景掩罩 70，混合器 66 将透视图形式的图像 42 植入输入视频帧的所要求的部位。于是，如果选手在球场 32 植入图像 42 的部分走动，他们将好像就在被植入的图像“上方”行走。如果需要，已变换的共混掩罩 58 可被用来将图像 42 与植入图像 42 的场地的颜色共混。

现在参照图 8 详细说明特征识别单元 60 的工作过程。在步骤 72，单元 60 使用 LUT 将来自许多彩色帧的输入视频转换为称作背景掩罩 70 的四色图。具体地说，对于网球场 32，LUT 提供一个第一数值至具有直线 38 的颜色的象素，一个第二数值至具有内球场 36 的颜色的象素，一个第三数值至具有外球场 40 的颜色的象素以及一个第四数值（表示前景象素）至其余的象素。这一点示于图 7。LUT 可以这一技术中已知的许多种方法中的任何一种适当方法实现。

背景掩罩 70 不仅确定哪些象素属于感兴趣的背景，而且包括例如直线 38 一类的感兴趣的特征。于是，在步骤 74，特征识别单元 60 对背景掩罩 70 进行处理以提取感兴趣的特征。一般尽管不必要，LUT 还是设计得提供具有单色数值的特征。

对于网球比赛的例子，提取过程包括审查背景掩罩 70 具有第一数值

的那些象素并从中提取直线线段。例如，步骤 74 可通过在背景掩罩 70 上的一个霍克变换操作来实现。霍克变换叙述在 1982 年科学院出版社出版的阿兹利尔·罗逊费尔德与艾维纳什·C 凯克所著的书“数字图像处理”第二版的第二卷第 121 - 126 页。本文中引用该书作为参考。

结果为一组直线参数 - 各描述背景掩罩 70 中的一个直线线段。每个线段的直线参数包括描述线段的直线方程的系数以及表示包含在线段中的象素数目的权重数值。

一种另外的提取方法示于现在就来简单参考的图 9A、9B 和 9C，正如 9A 总体示出的，该方法在背景掩罩 70 中具有感兴趣的颜色（本例中为白色）的一个第一象素 69（图 9B）处开始，并考察其邻域 75 以确定哪里存在更多的白色象素（用阴影标出）。为了做到这一点，将邻域 75 分成一个预定大小的子扇区 71 - 74，并完成各子扇区中白色象素分布的直方图。图 9C 示出图 9B 的扇区 71 - 74 的直方图。具有强极大值的一个（子扇区 73）被选出作为继续查找的下一个扇区。

在下一个步骤，确定一个由被选的子扇区 73 及其延伸部分组成的新邻域 78。整个邻域 78 是邻域 75 的两倍长。这一新的邻域 78 再划分为四个子扇区 76 并重复上面的过程。

继续这一过程直到满足下面的判断标准：

1. 子扇区窄到足以被确定为一根直线。
2. 在直方图中得不出强极大值。

如果达到条件 1，则直线的系数被存储起来并且组成直线的象素随即被“着色”列具有“保留颜色”因而从查找中删去。

特征提取过程产生出一组可能的特征，它包括真正的特征和杂散的直线。

现在参照图 10 一般说明图 5 的透视图识别单元 62 的工作过程。还要参照用于一般理解单元 62 的工作过程的图 11A 和 11B，参照以网球场 32 作例子详述单元 62 工作的图 13 以及参照用于理解图 13 详述的工作过程的图 12A、12B、12C、14A 和 14B。

通过使用一个先验信息，单元 62 在步骤 80 处理可能的特征组并决定哪一个最可能是感兴趣的特征。在步骤 82，单元 62 从最后的真正特

征中选出特征的一个最小集合并试图将它们与模型 50 的特征符合。每当需要时就重复这一过程直到找出符合为止。在步骤 84，符合的特征被用来生成一个变换矩阵 M ，将模型变换为输入视频帧中的特征。

在网球场 32 的例子中，步骤 80 利用模型 50 的直线 38 在两个方向（垂直和水平）平行并且在透视图（例如在输入视像帧中）平行线实际上相交于一有限点的事实。这一点示于图 11A，图中所有被提取的线段（用实线表示）均用虚线延长。对应于实际平行线的透视图直线（即伪平行线 90）相交于一个远离帧的外边沿 92 的点 91。一切别的交点（标为 94）均出现在边沿 92 之内或者紧靠其边界。

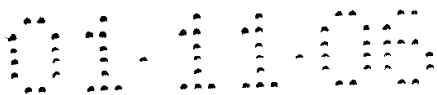
但是，如图 11B 所示，由于数字化误差的关系，有可能确定出三个伪平行线的延长线不在一个单一的点相交。事实上，这些线可能在隔开很远的三个点 96 相交。

申请人已经实现了，由于透视图平行线在有限远处相交，被提取的直线向一个渐近函数的映射将使得交点靠在一起出现。因而，根据本发明的一个优选实施例，被提取的直线段映射到一个二维渐近函数上。一个这样的映射被称作心射图法，在马萨诸塞州坎布里奇市的 MIT 出版社 1986 年出版的伯索尔德·克劳斯·保尔·霍恩所著的“机器人视觉”一书的第 258、259 和 275 页有叙述，本文中引用这几页作为参考。心射图法的例子示于图 12A 和 12B。

在心射图法中， XY 平面 102 上的点 100 被映射到半球面 104 上的点 100'。 XY 平面上的直线 106 被映射为半球面 104 的大弧 106'（即一个球面的大圆周的弧）。原点用南极点 109 表示，无限远点用中纬线 108 表示。于是，靠近中纬线 108 的任何点集 110（图 12B）均表示伪平行线的交点，由此，通过点集 110 的点的直线均为平行线。

图 12B 示出对应于一些任意提取的直线段（未示出）的多条大弧（标为 120a - 120f）。三条弧线 120a - 120c 具有的交点 122 在中纬线 108 附近形成一个点集 110a。大弧 120d - 120f 也在中纬线附近相交不过是相交于点集 110b。各条大弧互相相交，不过它们的另一交点均在比较靠近南极点 109 的位置，而不是靠近中纬线 108。

在步骤 130（图 13），心射图法被用来根据由特征提取过程（图



8 的步骤 74) 产生的一组直线线段阵列产生一组大弧。

在步骤 132 , 围绕中纬线 108 的区域均被查找以求出所有交点 122 . 对每个交点赋以一个数值 V_K . 数值 V_K 是各相交直线段的权重 W_i 和相交点 122 的 z 坐标的函数. 方程 (1) 提供了函数 V_K 的一个例子:

$$V_K = W_{\text{直线}1} \cdot W_{\text{直线}2} \cdot f(z_{\text{交点}}) \quad (1)$$

式中 $f(z_{\text{交点}})$ 为具有类似于图 12C 的曲线 134 的曲线的任何函数, 其中大多数的点均获得一个低的数值, 只有靠近中纬线 108 的那些点 ($z=1$) 获得接近于 1 的数值. 例如, $f(z_{\text{交点}})$ 可以是 z^5 .

在步骤 136 , 围绕各个交点 122 的一个小的邻域被用来查找其他交点. 若找出任何一交点, 则该交点与已找出的交点被作为点集 110 (图 12B) 存储起来. 点集 110 也被定义为一个 $f(z_{\text{交点}})$ 的数值高于预定的门限值的点集. 这样, 点集 10 可以只包含一个交点. 在图 12B 中有三个点集 110a - 110c , 其中之一即 110c 就只包含一个交点 122 .

一旦所有的点均被找遍, 每一点集 110 的位置就通过找出点集中各点的“重心”来确定. 点集 110 的权重为点集 110 中各点的数值 V_K 之和.

在步骤 138 , 选出具有最高权重的两个点集. 对于图 12B 的例子, 选出 3 点集 110a 和 110b .

在步骤 140 , 假定一个点集表示“垂直”线, 而另一个则表示“水平线”. 而且, 在步骤 140 中, 对应于两个选出的点集的直线段被分别标上“垂直”和“水平”字样.

在步骤 142 , 审查“垂直”和“水平”线, 并选出两根最重的垂直线和两根最重的水平线, 这里的“最重”由 W_i 的数值确定. 所选出的垂直线 (标为 146) 示于图 14A 并用作图 11A 的直线. 在步骤 144 , 确定四条选出的直线的交点 (标为 A、B、C 和 D) 并存储起来. 如图 14A 所示, 选出的直线可以相交在帧外.

步骤 130 - 144 是识别视频帧里面的真正特征 (图 10 的步骤 80) 所必须的操作过程. 步骤 144 的输出便是要与模型符合的特征. 其余的步骤将特征与模型符合和将变换 (图 10 的步骤 82 和 84) 确定为整套的操作过程.

一个标准的网球场具有五条垂直线和四条水平线. 由于不可能把球



场的两个半场区分开，因而仅有三条水平线是重要的。由三条中选出两条水平线（有三种可能的组合）和由五条中选出两条垂直线（有10种可能的组合）能够组成的不同四边形的数目为30。这30个四边形可以在四个不同的方向形成总共120个矩形。

在步骤150，通过选择其四个角（图14中标为A'、B'、C'和D'）选出几何模型50中120个矩形里面的一个来。正如所能看到的，这不是正确的符合。

在步骤152，将模型的四个点A'、B'、C和D'（图14B）变为视帧的四个点A、B、C和D（图14A）的矩阵M被确定。正如参照图15所说明的，矩阵M可表示为一连串的变化叠加。

图15示出三个四边形180、182和184。四边形180是示于XY平面上的模型四边形ABCD，四边形182是在TS平面内具有点(0, 1)、(1, 1)、(0, 0)和(1, 0)的单位正方形，而四边形184则是UV平面上的透视图四边形。

由模型四边形180至透视图四边形184的变换M可以由两个变换即一个由四边形180至单位正方形182的变换并改变比例的矩阵T和一个由单位正方形182至四边形184的透视图矩阵P的叠加来表示。矩阵T（采用齐次坐标）具有形式：

$$T = \begin{vmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ T_x & T_y & 1 \end{vmatrix} \quad (2)$$

式中 S_x 和 S_y 分别为X和Y方向的比例因子， T_x 和 T_y 为X和Y变换因子。 S_x 、 S_y 、 T_x 和 T_y 由方程式

$$(X, Y, 1) * T = (S, t, 1) \quad (3)$$

对四边形180的四个坐标(X, Y, 1)和单位正方形182的四个坐标(S, t, 1)求出。

矩阵P（采用齐次坐标）具有形式：

$$P = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (4)$$

矩阵 P 的元素通过解下面的方程求出:

$$(S, t, 1) * P = (U, V, W) \quad (5)$$

式中 (U, V, W) 表示图 15 所示的四边形 184 的点 A、B、C 和 D 的四个已知坐标, 并且 W 总是归一化的。

假设 $a_{33}=1$, 则可按以下方法计算出 P。

由 $(S, t, 1) = (0, 0, 1)$, 我们求出

$$a_{31} = U_0$$

$$a_{32} = V_{00} \quad (6)$$

由 $(S, t, 1) = (1, 0, 1)$, 我们求出:

$$a_{11} + a_{31} = U_{10}(a_{13} + 1) \Rightarrow a_{11} = U_{10}(a_{13} + 1) - U_{00} \quad (7)$$

$$a_{12} + a_{32} = V_{10}(a_{13} + 1) \Rightarrow a_{12} = V_{10}(a_{13} + 1) - V_{00}$$

由 $(S, t, 1) = (0, 1, 1)$, 我们求出:

$$a_{21} + a_{13} = U_{01}(a_{13} + 1) \Rightarrow a_{21} = U_{01}(a_{23} + 1) - U_{00} \quad (8)$$

$$a_{22} + a_{32} = V_{01}(a_{23} + 1) \Rightarrow a_{22} = V_{01}(a_{23} + 1) - V_{00}$$

由 $(S, t, 1) = (1, 1, 1)$, 我们求出:

$$a_{11} + a_{21} + a_{31} = U_{11}(a_{13} + a_{23} + 1) \quad (9)$$

$$a_{12} + a_{22} + a_{32} = V_{11}(a_{13} + a_{23} + 1)$$

由方程 (7) - (9), 产生出含两个未知数 a_{13} 和 a_{23} 的两个方程

如下:

$$a_{13}(U_{10} - U_{11}) + a_{23}(U_{01} - U_{11}) = U_{11} + U_{00} - U_{10} - U_{01} \quad (10)$$

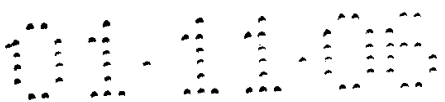
$$a_{13}(V_{10} - V_{11}) + a_{23}(V_{01} - V_{11}) = V_{11} + V_{00} - V_{10} - V_{01}$$

一旦求出 a_{13} 和 a_{23} , 其余的元素就能由方程 (7) 和 (8) 求出。

变换或映射矩阵 M 等于矩阵 T 和 P 的矩阵乘积, 如下式:

$$M = T * P$$

在步骤 154, 模型 50 的直线 38 被用映射矩阵 M 映射到视频帧上。结果得到一个在存在模型的已转换象素的地方具有 1 在别的地方则具有 0 的变形帧 156 (图 16)。正如所能看到的, 点 A'、B'、C' 和 D' 分别



与点 A、B、C 和 D 符合。但是，几何模型 50 的其余地方则不相符合。

在步骤 158，求变形帧 156 与背景掩罩（图 7）的“异或”值。“异或”步骤在两种情形输出一个 0：a) 变形帧 156 的象素具有 1 值和视频帧的象素具有场地线颜色；b) 变形帧 156 的象素具有 0 值和视频帧的象素具有“非场地线”颜色。其余的情形均得到 1 值。

在步骤 160 和 161，对具有 1 值的象素数目计数并将计数数值与变换矩阵 M 相联系。

所有的矩阵 M 都已经确定之后，在步骤 162，选出具有最小权重的矩阵。由于存在不能达成任何符合的可能性（即电视上正在放映商业广告，电视摄像机正在对着观众，等等），在步骤 164，用一个门限值检验被选矩阵的权重，若高于那个值，则提供一个零值变换矩阵。否则，将所选矩阵确定为变换矩阵 M。在前面任何步骤的试验条件出现故障时，也提供零变换矩阵。

现在参照图 17 说明图 5 的变换器 64 和混合器 66 的工作过程。变换器 64 利用变换矩阵 M 使每个图像 42、图像区域掩罩 56 以及共混掩罩 58 变形到视频帧的平面（步骤 170）。它也对变形的图像区域掩罩和背景掩罩进行逻辑乘（AND），从中产生一个允许掩罩。允许掩罩表示那些既是背景象素又处在图像区域之中的视频帧的象素。图像将被植到这些象素上。

混合器 66 根据共混和允许掩罩将变了形的图像与视频帧组合。每一象素（x,y）的实现公式一般为：

$$\text{输出}(x,y) = \beta(x,y) * \text{图像}(x,y) + (1 - \beta(x,y)) * \text{视像}(x,y) \quad (12)$$

$$\beta(x,y) = \alpha(x,y) * P(x,y) \quad (13)$$

式中输入（x,y）为输出帧的象素数值，图像（x,y）和视频（x,y）分别为植入图像 42 和视频帧的数值， $\alpha(x,y)$ 为共混掩罩 58 中的数值以及 $P(x,y)$ 为允许掩罩中的数值。

上文的叙述假定了产生背景掩罩 70 的 LUT 在整个比赛期间保持正确。如果照明变化（一般出现在室外比赛中），则视频序列的色彩会变化，其结果是，背景掩罩 70 不再正确表示背景要素。因此可以周期性地执行一个校正过程。校正过程详述在图 18 中，现在就来参考它。

注意，在校准过程中，表示背景中感兴趣的特征（例如场地线、内球场和外球场）的测试点由操作员选择。点的位置及它们的彩色数值被储存起来。

一旦用于校准视频帧的矩阵被确定，测试点的位置便由视频帧平面转换到几何模型平面（即使用矩阵 M 的逆变换）。在以后某些时候需要校准时，测试点再转换到当前的视频帧平面。选出当前视频帧里面变形的测试点并对其邻域取样。计算出每个领域的颜色特性（例如用直方图），并将结果与存储的点的特性相比较。如果在颜色中有任何显著的变化，则校正 LUT 并将相关的点转换为几何模型且存储起来。

将正确评价在上文中就网球比赛以及就其他背景信息固定不变且已知的情形叙述的本发明所包括的操作程序。如下文所述，上文叙述的操作程序可通过跟踪和通过知道摄像机参数用许多方法改进。

当提供有关摄像机位置、转动角和焦距量的信息（由外部或者由系统确定）时，由于透视图矩阵 P 的自由度数目的减少的关系，上文叙述的操作过程可以缩短。

具体地说，透视图矩阵 P 包括有考虑所用摄像机的位置、转动角以及焦距的信息。这个信息可以提取出来，并且透视图矩阵 P （或者同样地，变换矩阵 M ）可以被重新定义为各个摄像机参数的函数。

图 19 示出一架摄像机及其参数。其位置用具有离 X ， Y ， Z 坐标系 172 的原点 O 坐标 (x,y,z) 的向量 171 表示。摄像机关于其基轴 U 、 V 和 W 的相应滚动、俯仰、偏转和扫调分别如箭头 173、174 和 175 所示。此外，摄像机镜头可以沿 V 轴调焦距，如箭头 176 所示。

假设摄像机不滚动并且摄像机的纵横比（摄像机生成的图像中象素的宽度与高度之间的比例）规定为方形象素，则透视图矩阵 P 可以参数化为摄像机的位置及其俯仰、偏转和焦距的函数。假设摄像机从前帧至后帧不改变位置而仅仅改变其俯仰、偏转、角度或焦距。

图 20、21 和 22 表示确定和利用摄像机参数的方法。在图 20 中，当在视频流中识别出一个新的切换时，则在新切换的第一帧上执行如图 10 所示的整个透视图识别操作过程（步骤 180）。步骤 180 产生透视图距阵 P 的元素 $a(i,j)$ 。过程在两个方向继续进行：

a)由图 13 的步骤 154 开始, 确定变换矩阵 T ; 以及

b)如 MIT 出版社 1993 年出版的奥利弗、福格拉斯所著“三维计算机视觉: 一种几何观点”一书的第 3.4 节所说的那样由矩阵 P 提取出摄像机坐标 (x, y, z) (步骤 184)。本文中引用该书作为参考。

一旦提取出摄像机坐标 (x, y, z) , 便执行如下两个检查 (步骤 186 和 188):

条件 186: 摄像机在方向 174 不滚动 (旋转)。若元素 a_{13} 不等于零时, 则存在滚动。

条件 188: 摄像机的纵横比 (AR) 定义一个正正方形象素 (即 $AR = 1$)。

如果任何一个条件不真实, 则缩短操作过程的剩余部分立即停止。

如果两个条件均被满足, 则像在“三维计算机视觉: 一种几何观点”一书中所说的那样, 矩阵 P 可以被重新表示 (步骤 190) 成下面各矩阵的乘积:

a)焦距 (f): 摄像机焦平面的投影的矩阵;

b)变换: 由坐标系原点至计算出的摄像机位置 (x, y, z) 的变换的矩阵;

c)俯仰 (α): 围绕 U 轴转一角度 α 的旋转矩阵; 以及

d)偏转 (θ): 围绕 W 轴转一角度 θ 的旋转矩阵。

使用焦距、俯仰、偏转和变换的数值, 将第一摄像机彻底校准 (步骤 192) 并将其参数登载到识别过的摄像机的表格 194 (示于图 21) 中。别的摄像机将接下文叙述的方法识别并登记在表格 194 中。

然后, 在所有的帧上面执行根据图 22 所叙述的缩短计算过程。使用 α 、 θ 和 f 考查一个帧 (步骤 196) 以确定它与前面的帧的相似性。相似性则通过一个符合系数 (即使用计算出的矩阵相继映射到模型上的帧中感兴趣的象素的百分率) 来度量。如果得到的相似性良好, 则所计算得的矩阵可用于插入操作 (根据图 17 所述)。如果符合系数很小, 则可能该帧是用表格 194 里的另一架摄像机来拍摄的。

为了找出另外的摄像机, 必须审查当前帧并识别其中的一根直线。而且, 还必须识别被识别直线上的一个点例如和另一直线的交点 (步骤 198)。通常被识别的直线都是“最粗状的”直线。

在步骤 200，按下法确定列在表格 194 中的各架摄像机的符合值：

将被识别的直线和点与几何模型中的一根直线和点联系起来，并且确定出这一联系的透视图矩阵 P （它将模型的直线和点交换到被识别的直线和点）。由于每一个透视图矩阵都是当前摄像机（已知）的坐标 (x,y,z) 和俯仰角 α 、偏转角 θ 以及焦距 f （未知数）的一个函数，因而，在假定被识别的直线和点很好地符合模型的直线和点的情况下，最终的透视图矩阵 P 可通过能够算得出来的俯仰、偏转及焦距的数值确定。

如在图 10 的方法中一样，变换矩阵 M 根据透视图矩阵 P 确定，并且几何模型通过矩阵 M 变换为帧内图像的平面。模型的直线符合图像里的直线并产生一个符合数值。

对几何模型里的直线和点的每一组合反复进行将模型的直线和点与被识别的直线和点联系，根据已知的摄像机坐标及直线和点的联系产生一个透视图矩阵 P ，并将一个符合数值确定为的过程。如果符合数值大大小于 1，就表示符合很差，则对坐标 (x,y,z) 已知的另一架摄像机重复进行上文所述的与被识别直线和点的符合过程。

将对每一架摄像机计算出的最大符合系数登入表格 194（图 21）的一个标为 202 的栏中。在步骤 204，选出具有系数 202 的最高值的摄像机，如果系数大于一个预定的门限值，则它的透视图矩阵 P 被用于图 17 的图像插入过程。如果栏 202 的最高系数有一个低于门限的数值，则有一个未知的摄像机被用来拍摄了当前帧。执行图 10 的过程必须继之以图 20 的摄像机识别过程。

熟悉这一技术的人员将会意识到本发明并不限于上文中特别示出和叙述的那些东西。更确切地说，本发明的范围系由后面的权利要求书所规定。

说明书附图

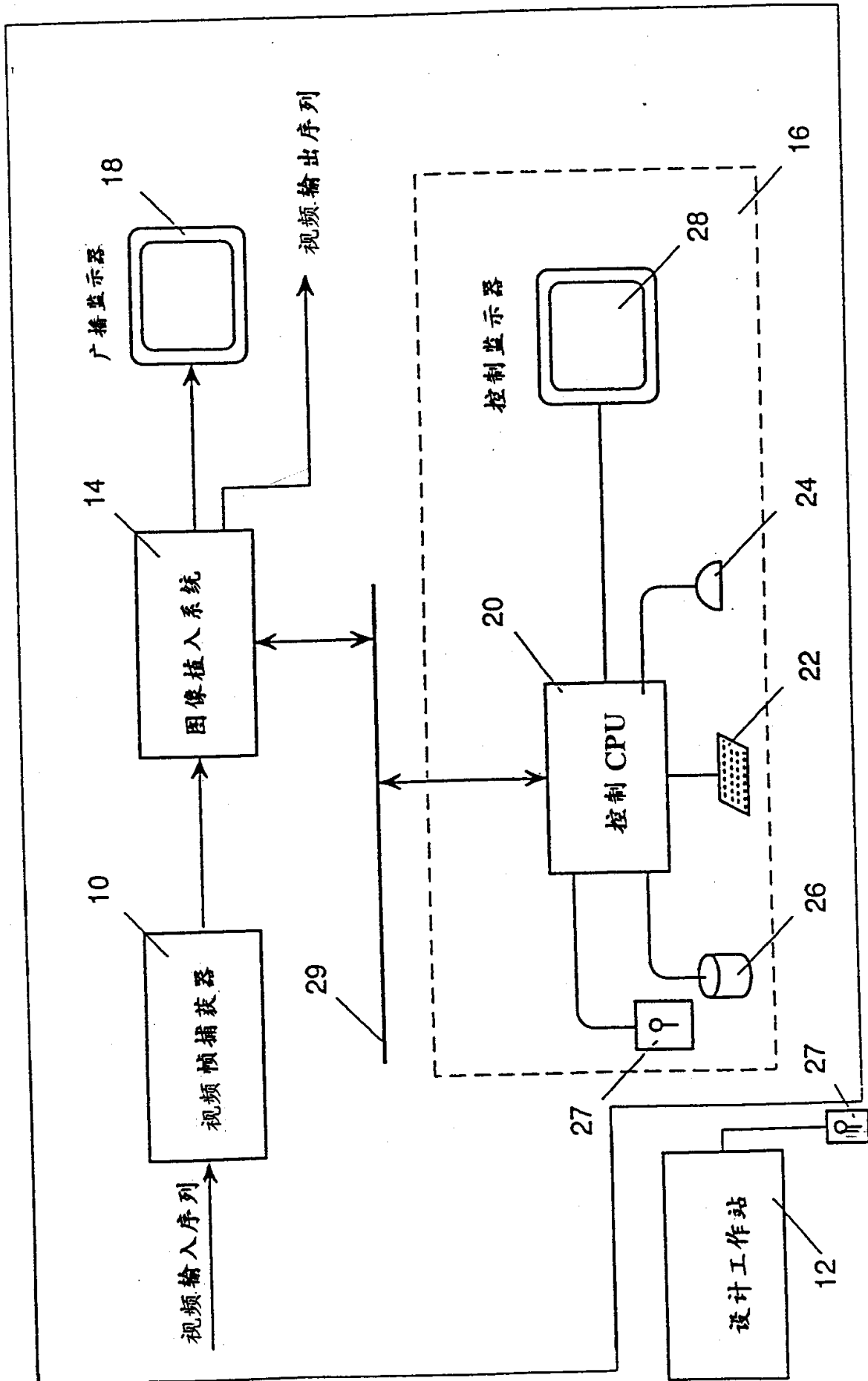


图 1

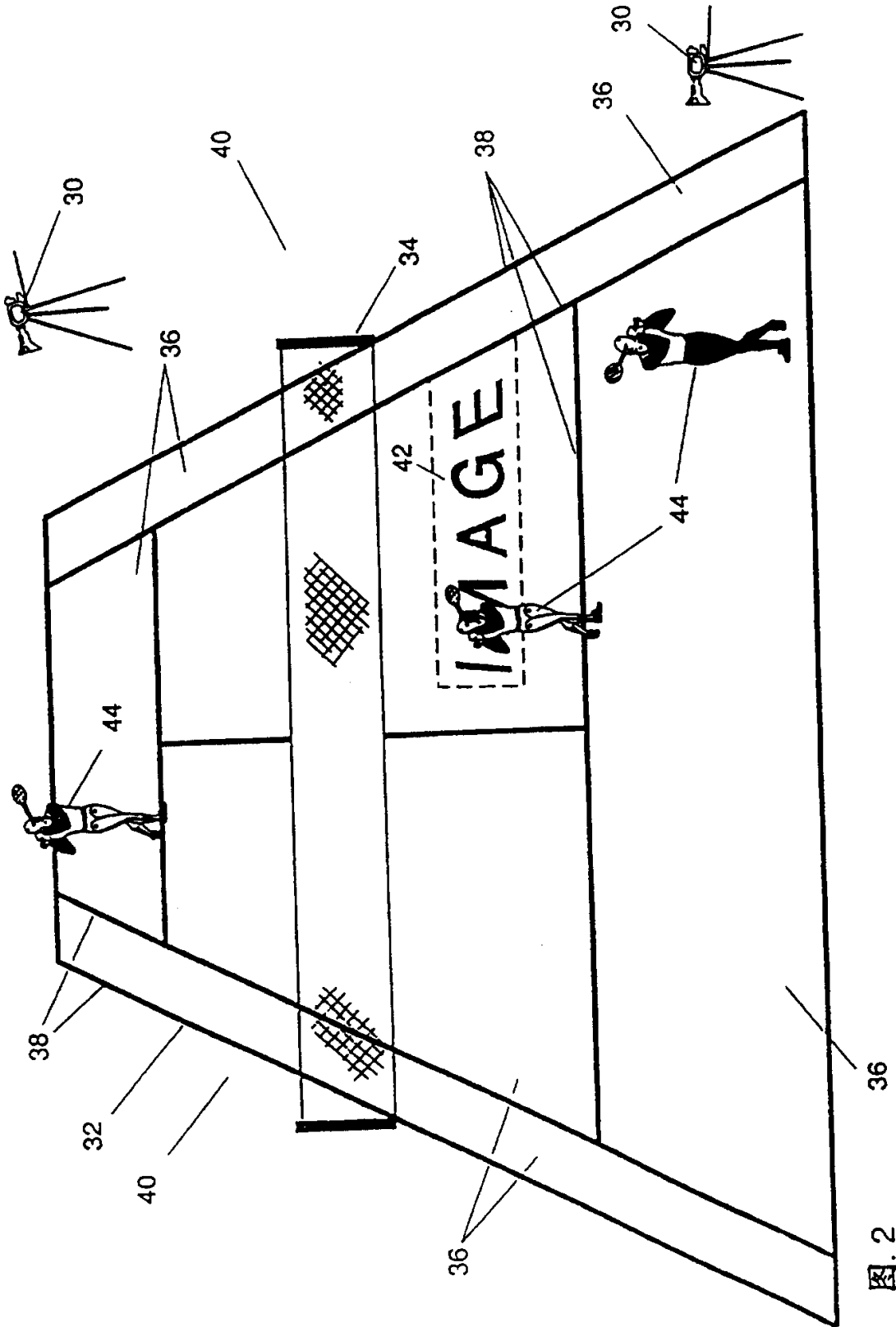


图. 2

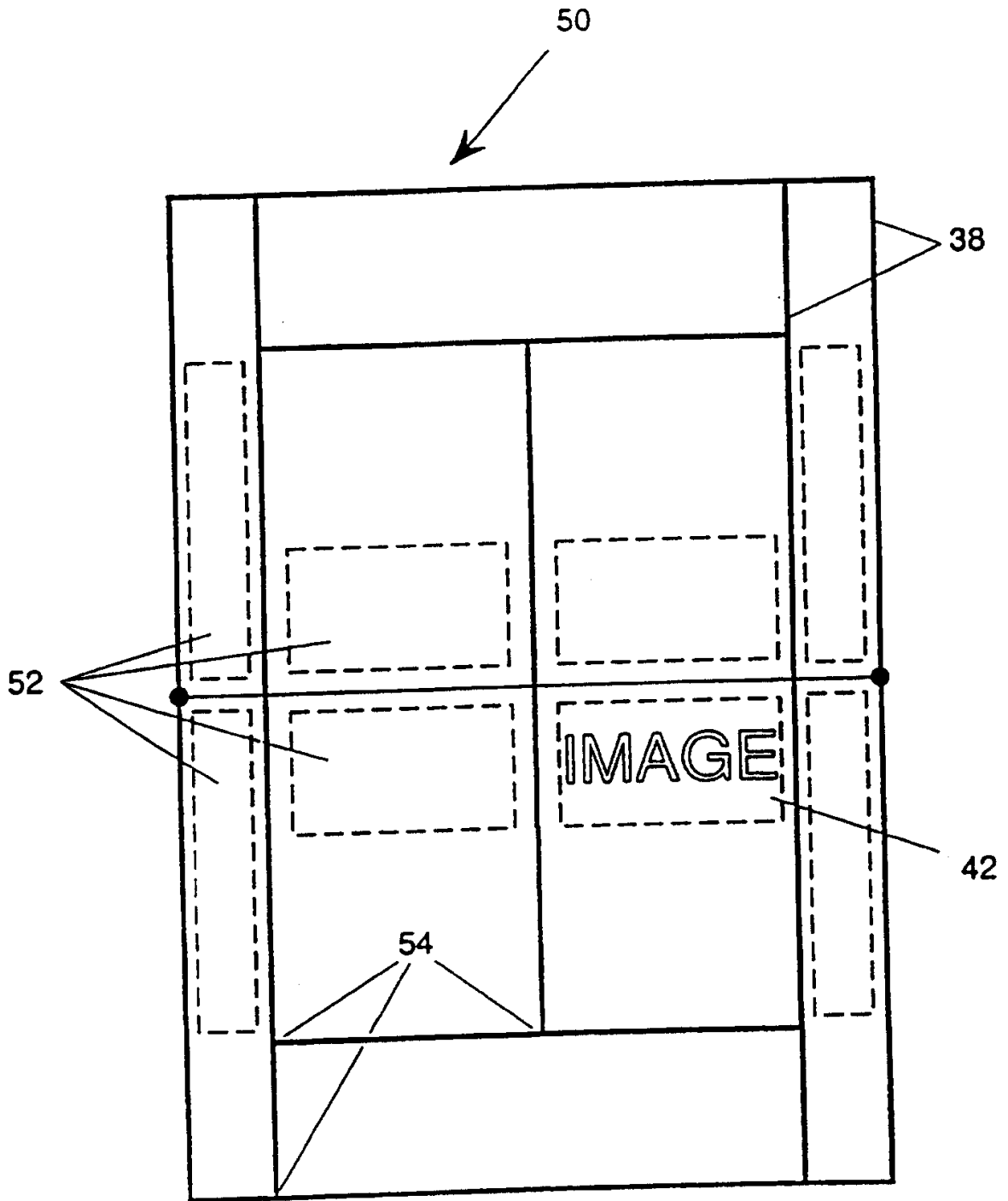


图 3



图 4A

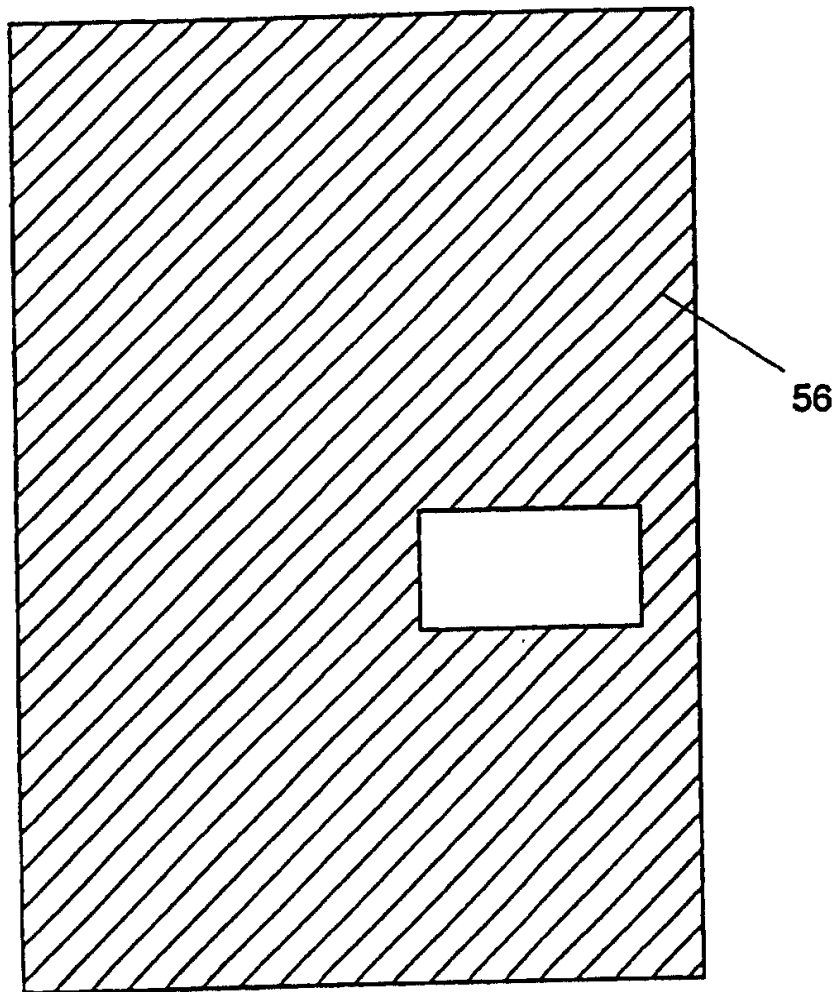


图 4B

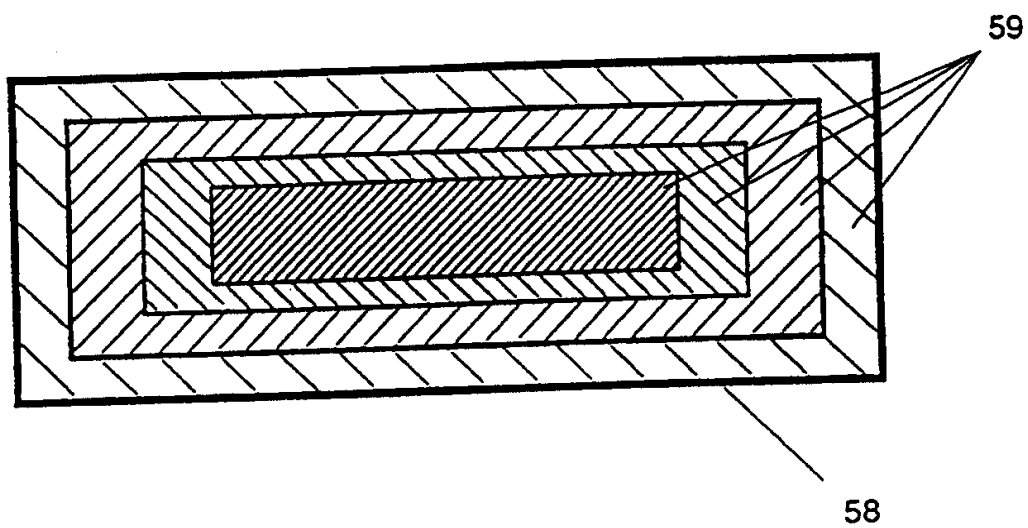


图 4C

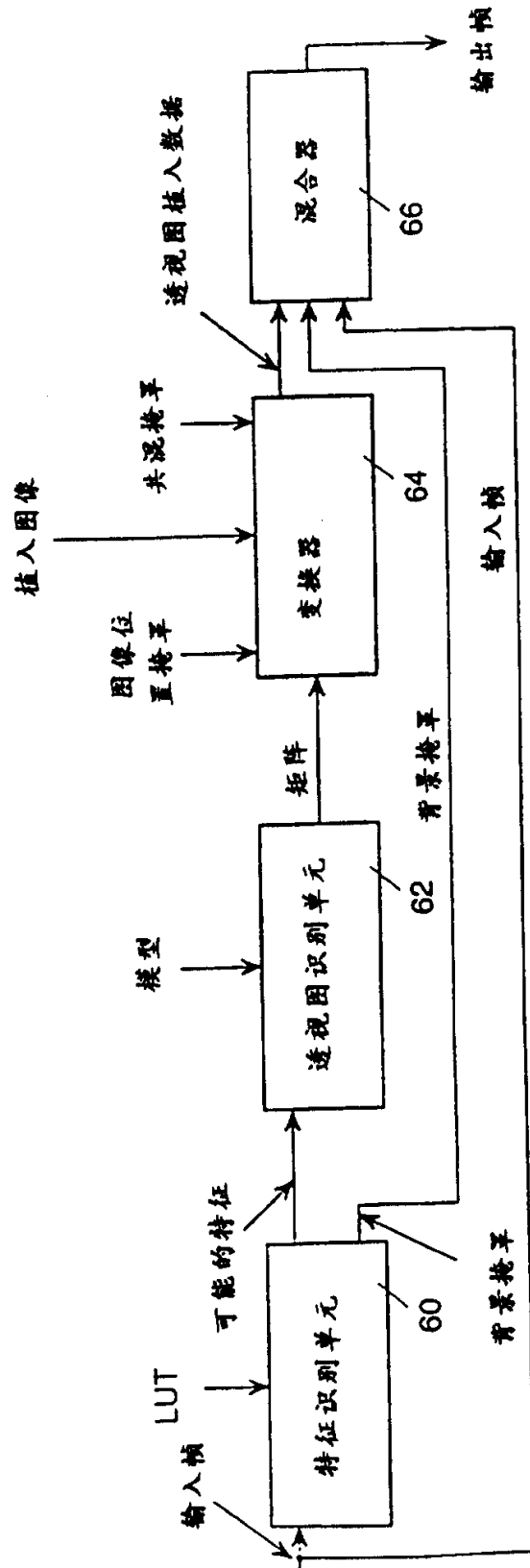


图 5

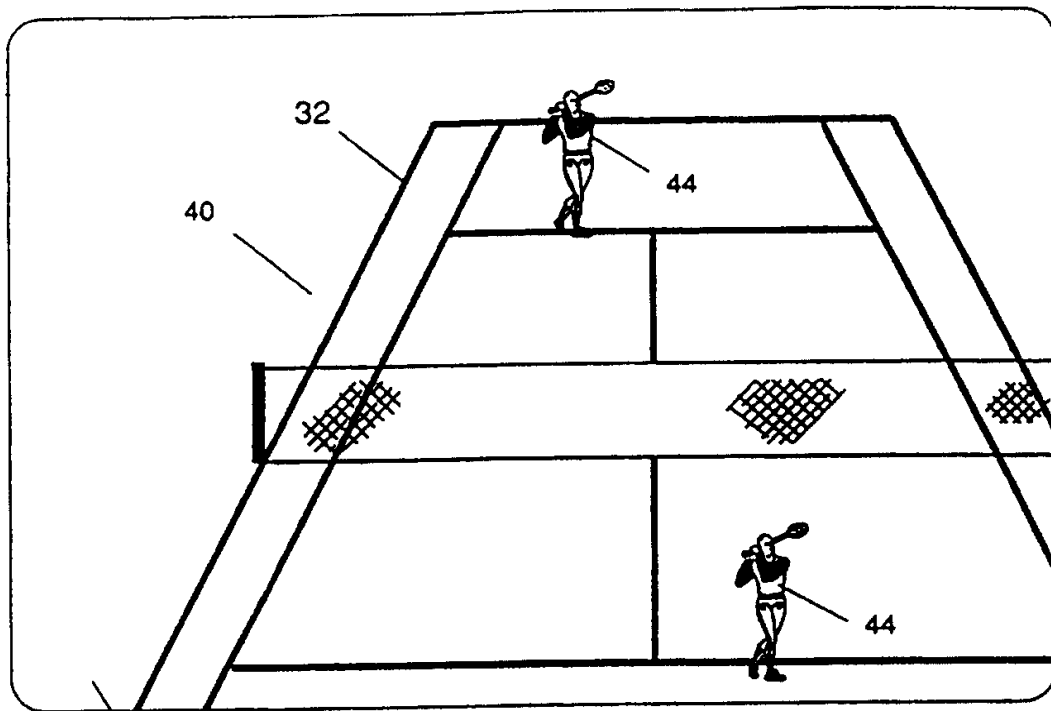


图 6

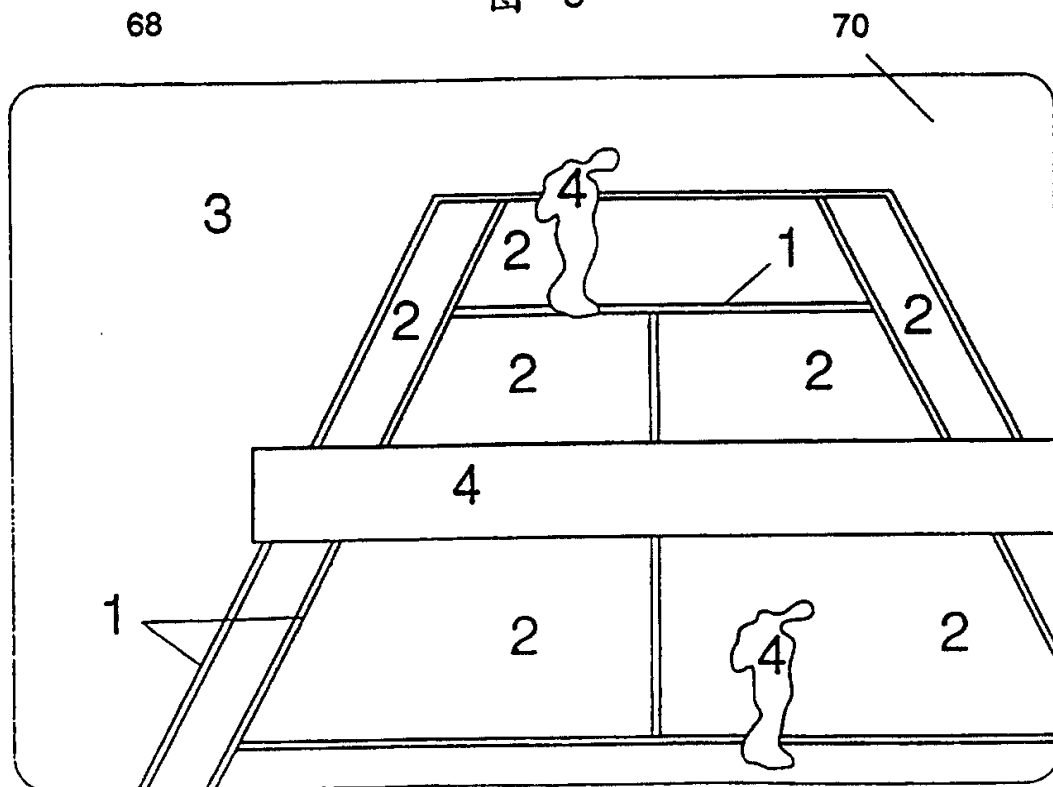


图 7

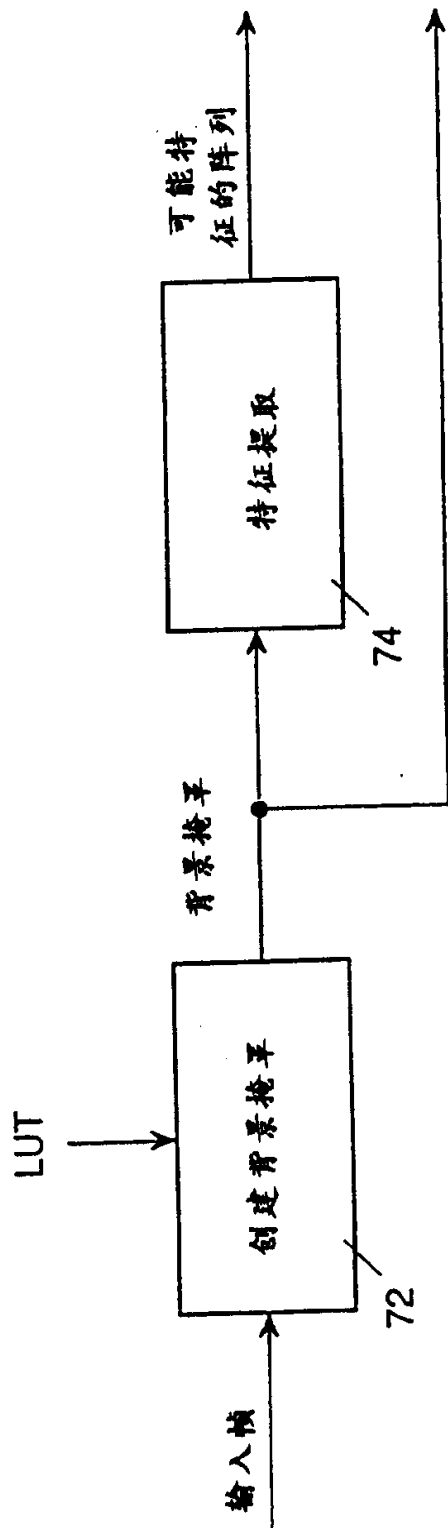


图 8

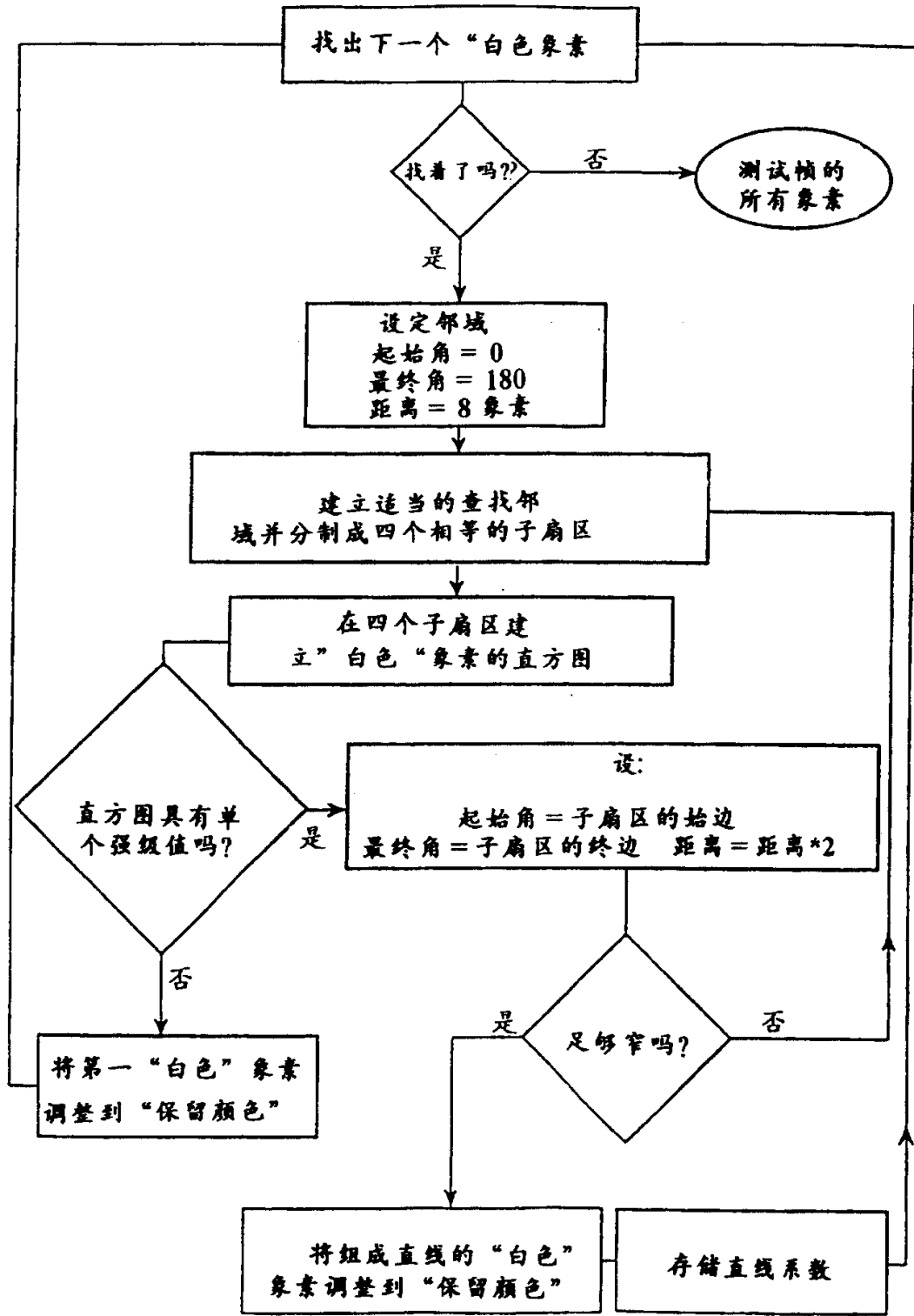


图 9A

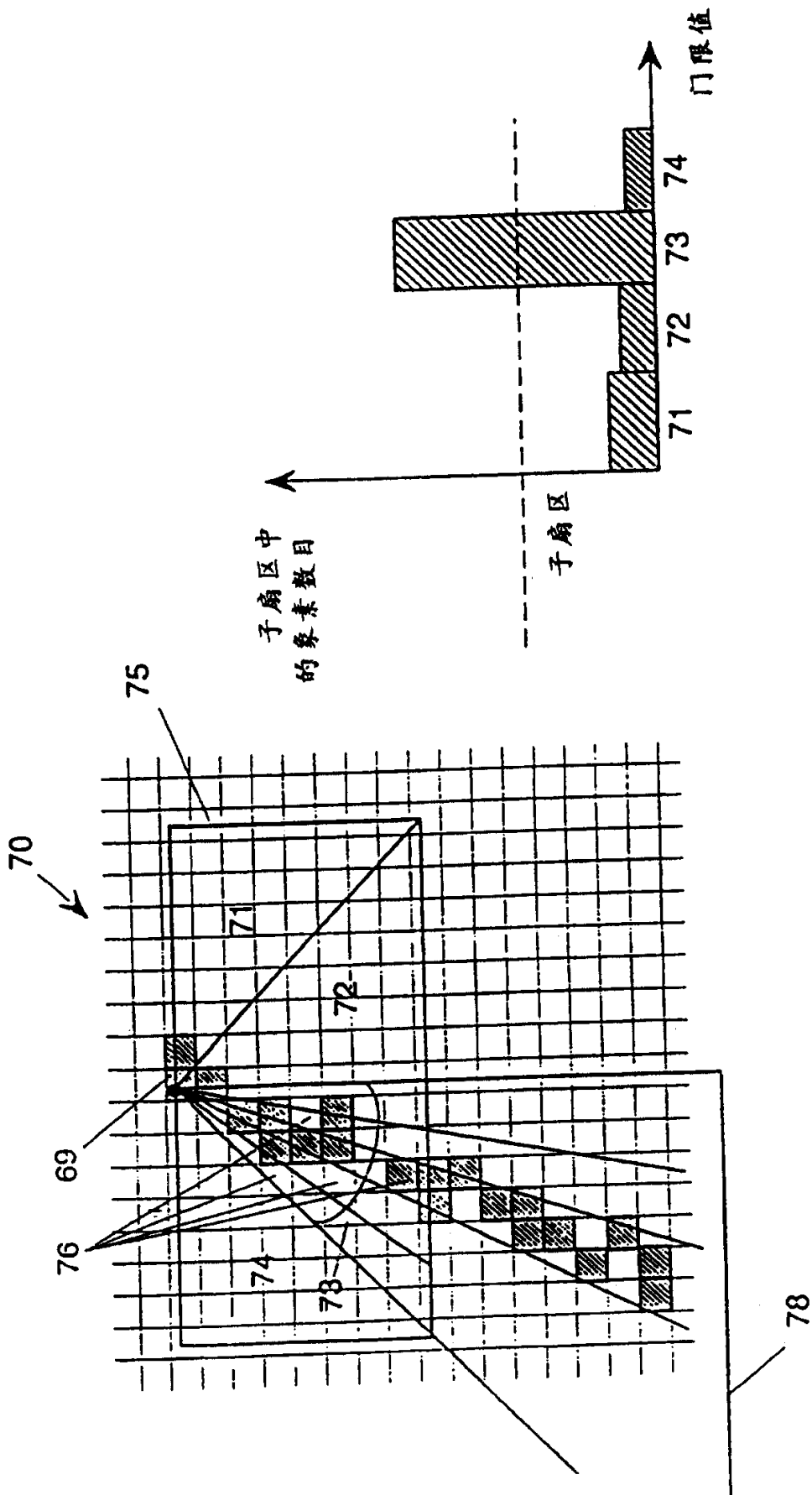


图 9C

图 9B

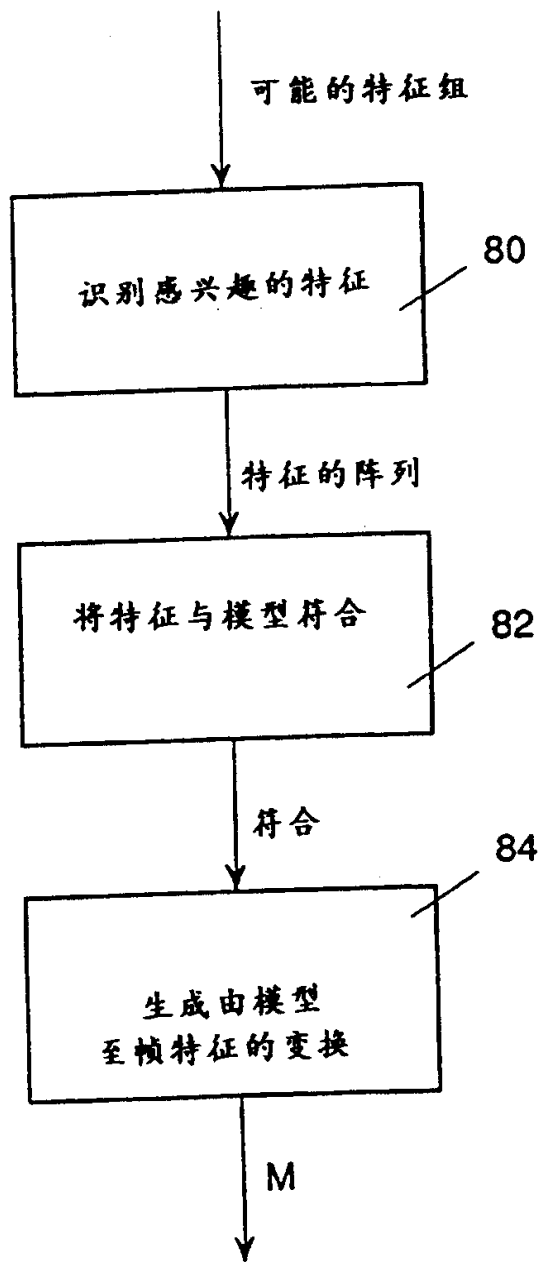


图 10

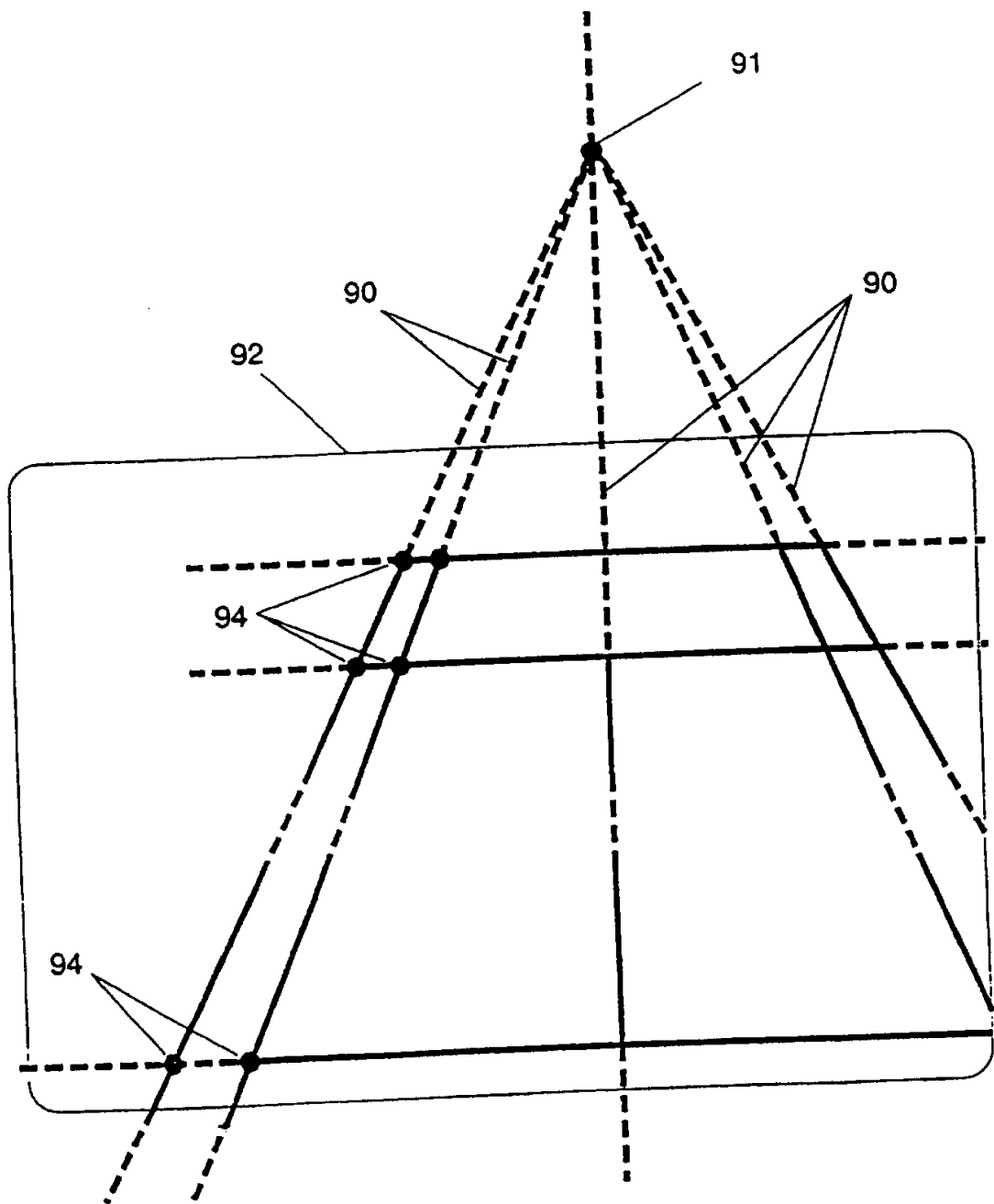


图 11A

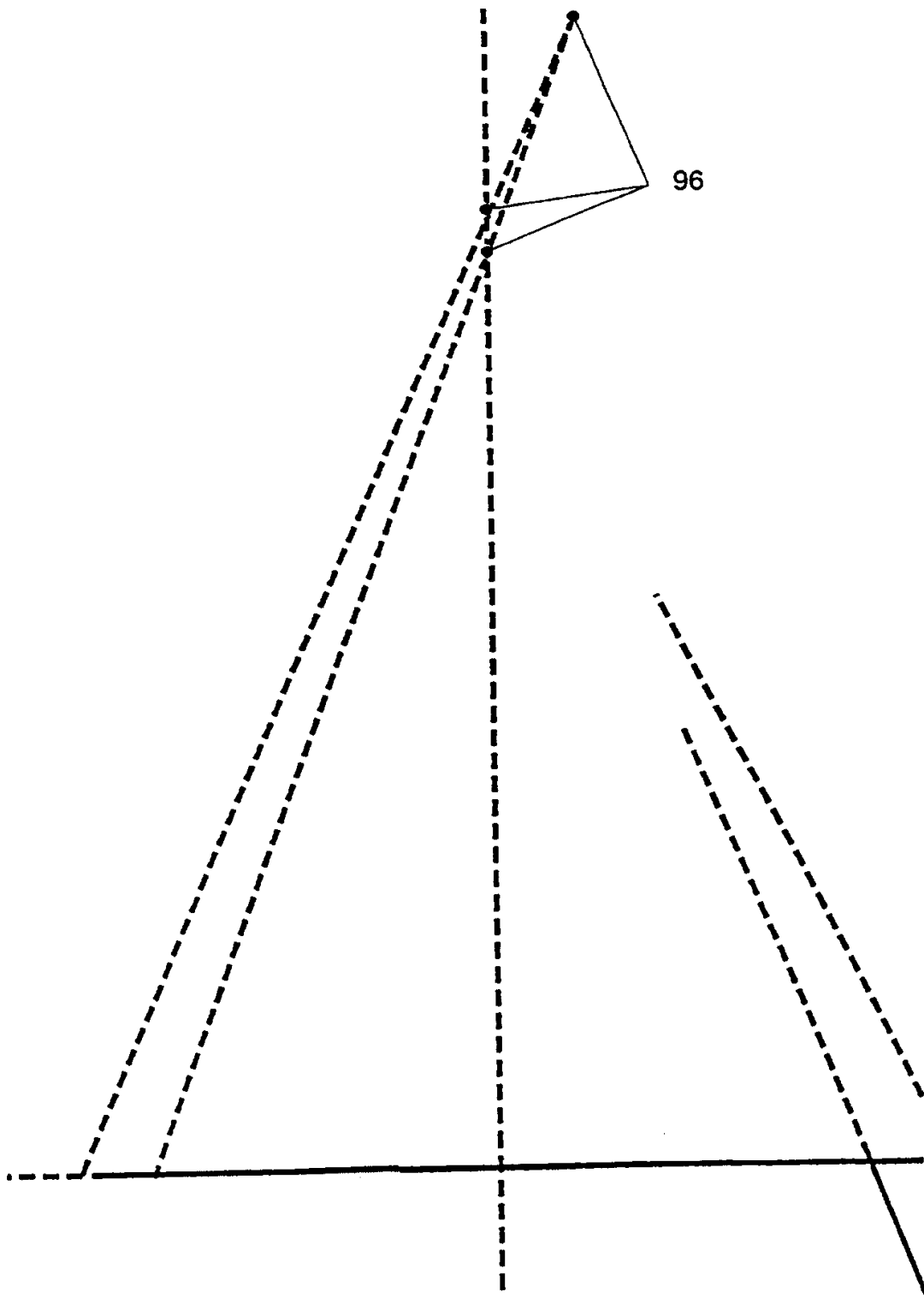


图 11B

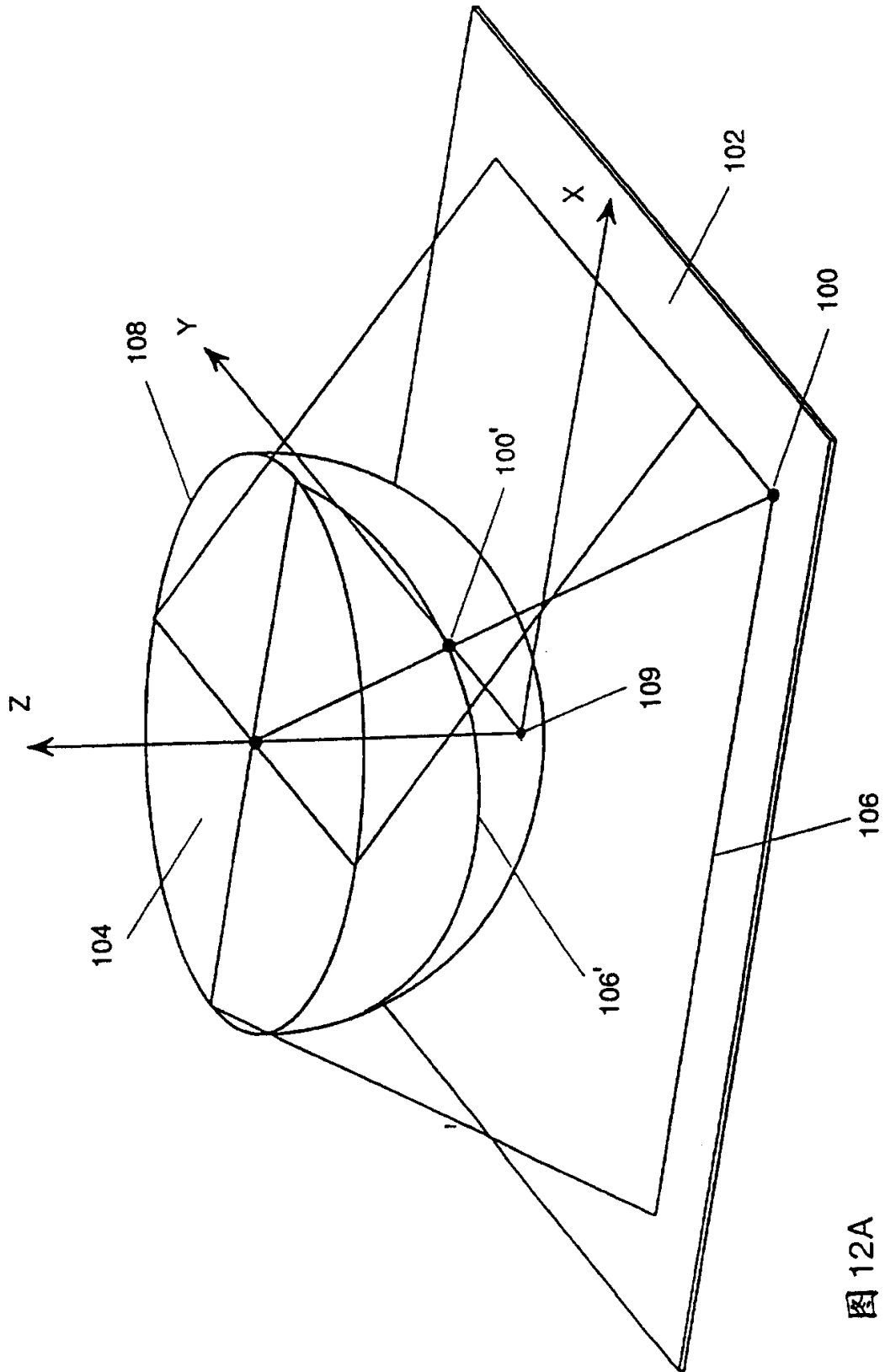


图 12A

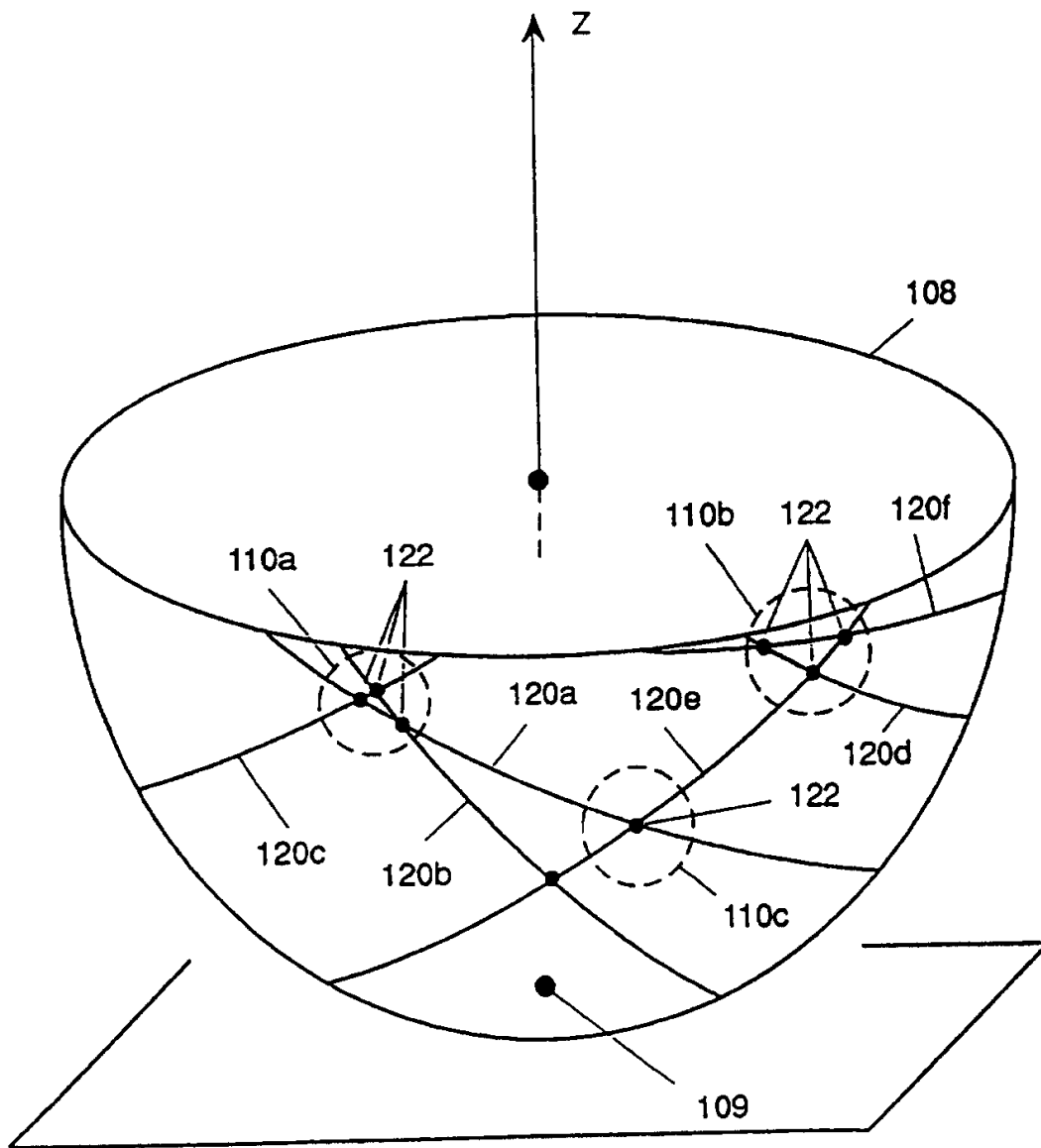


图 12B

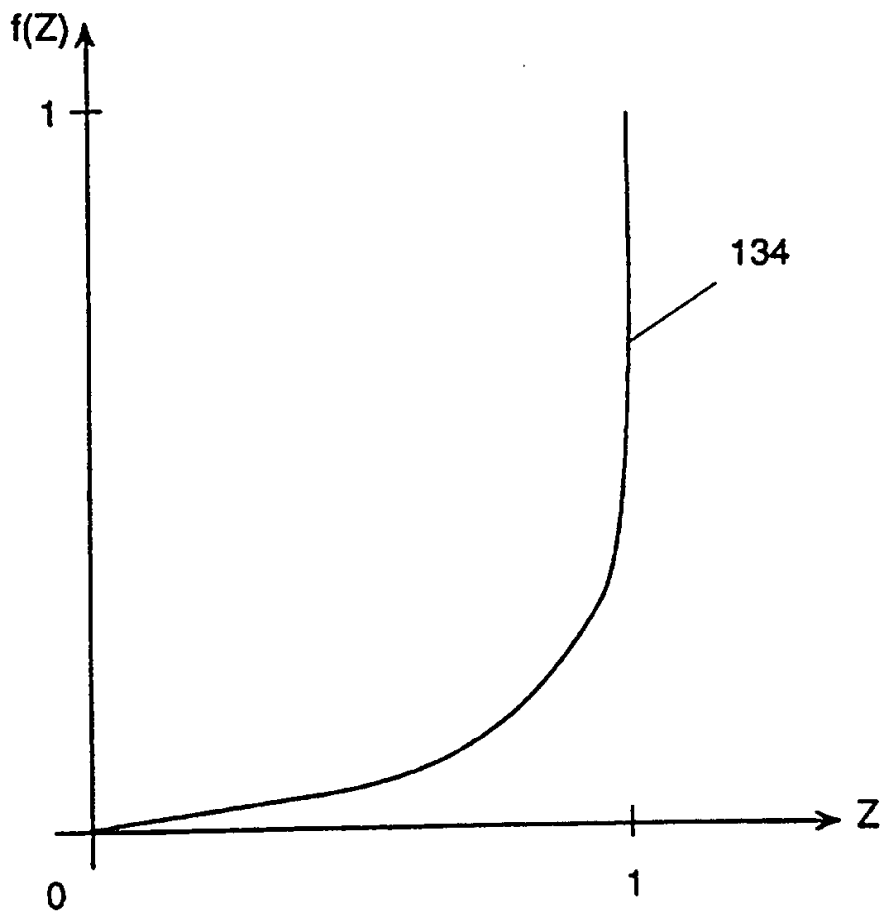


图 12C
- 17 -

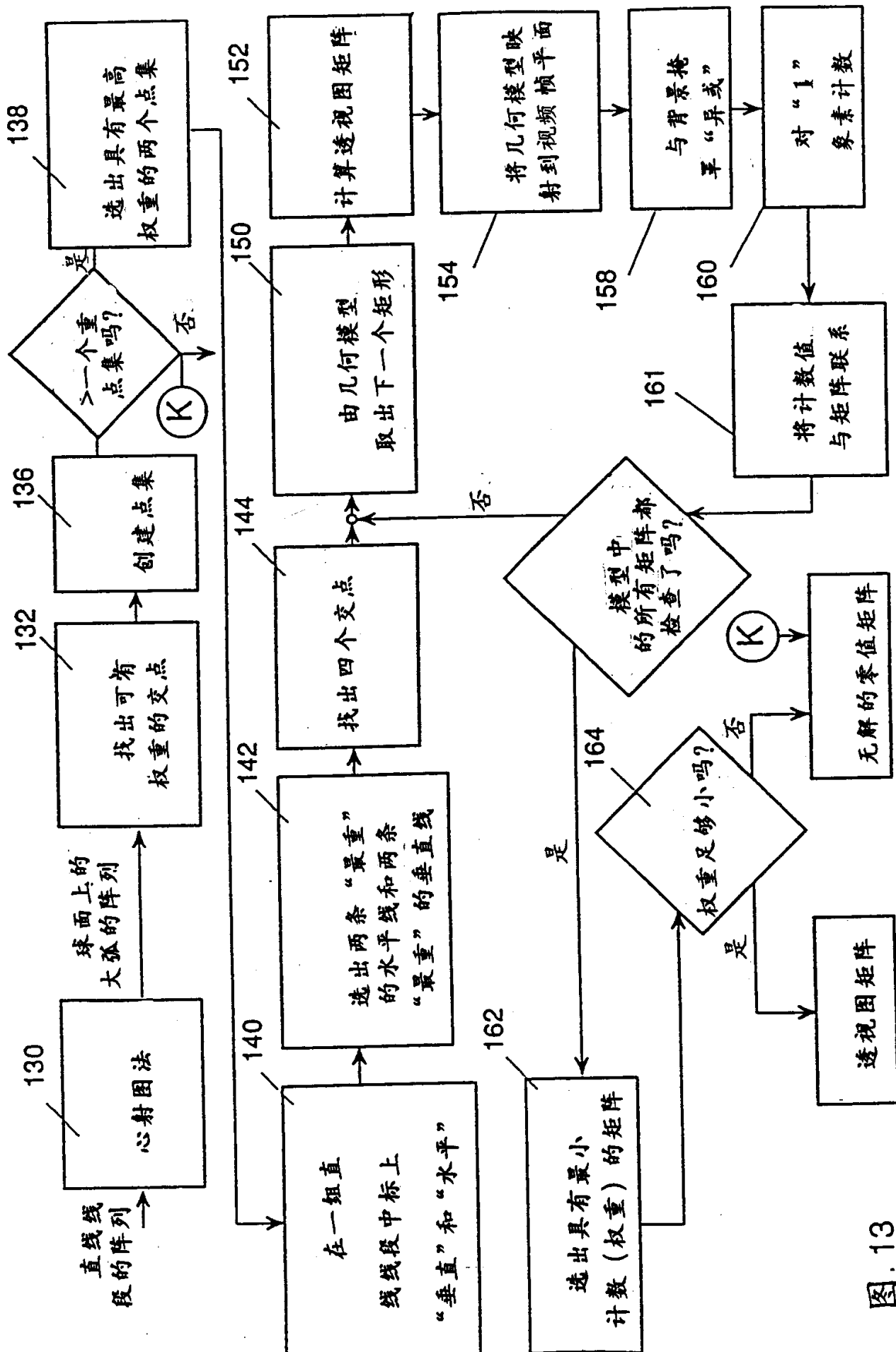


图. 13

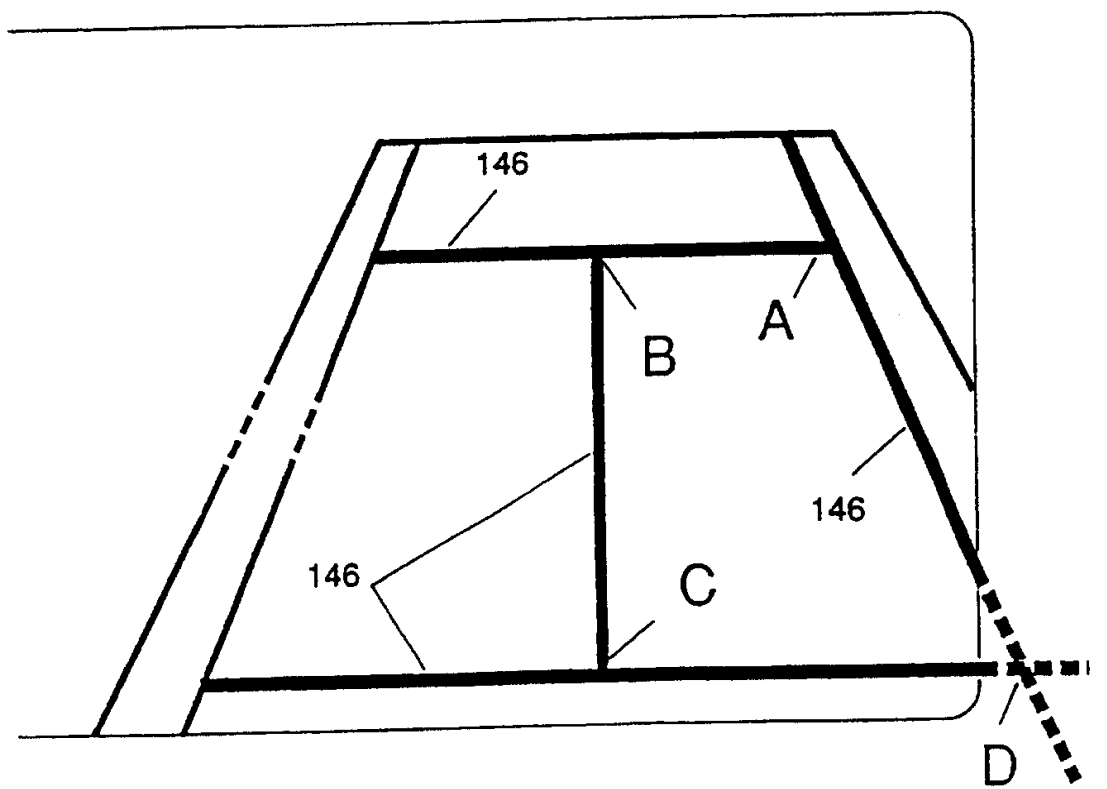


图 14A

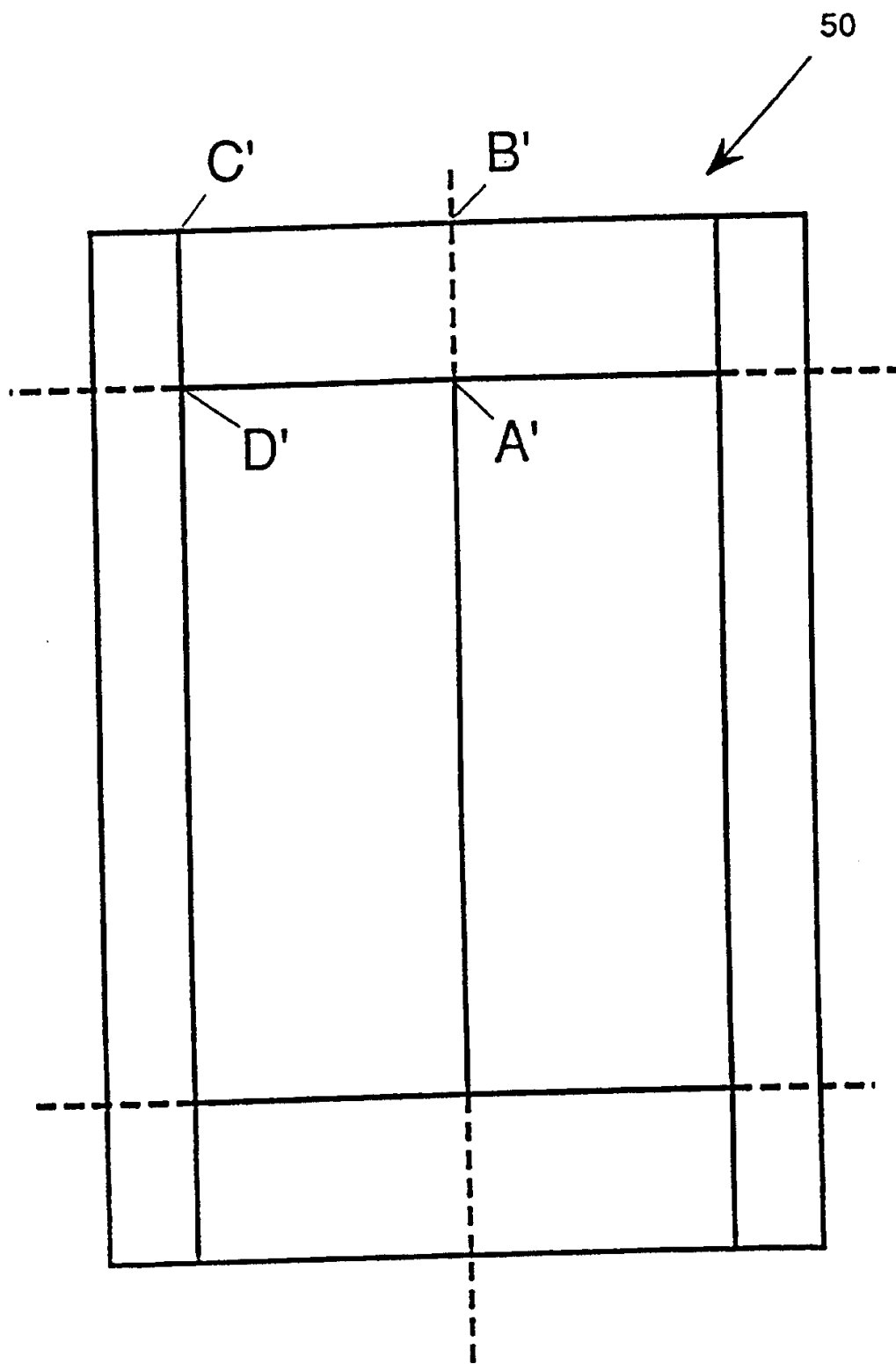


图 14B

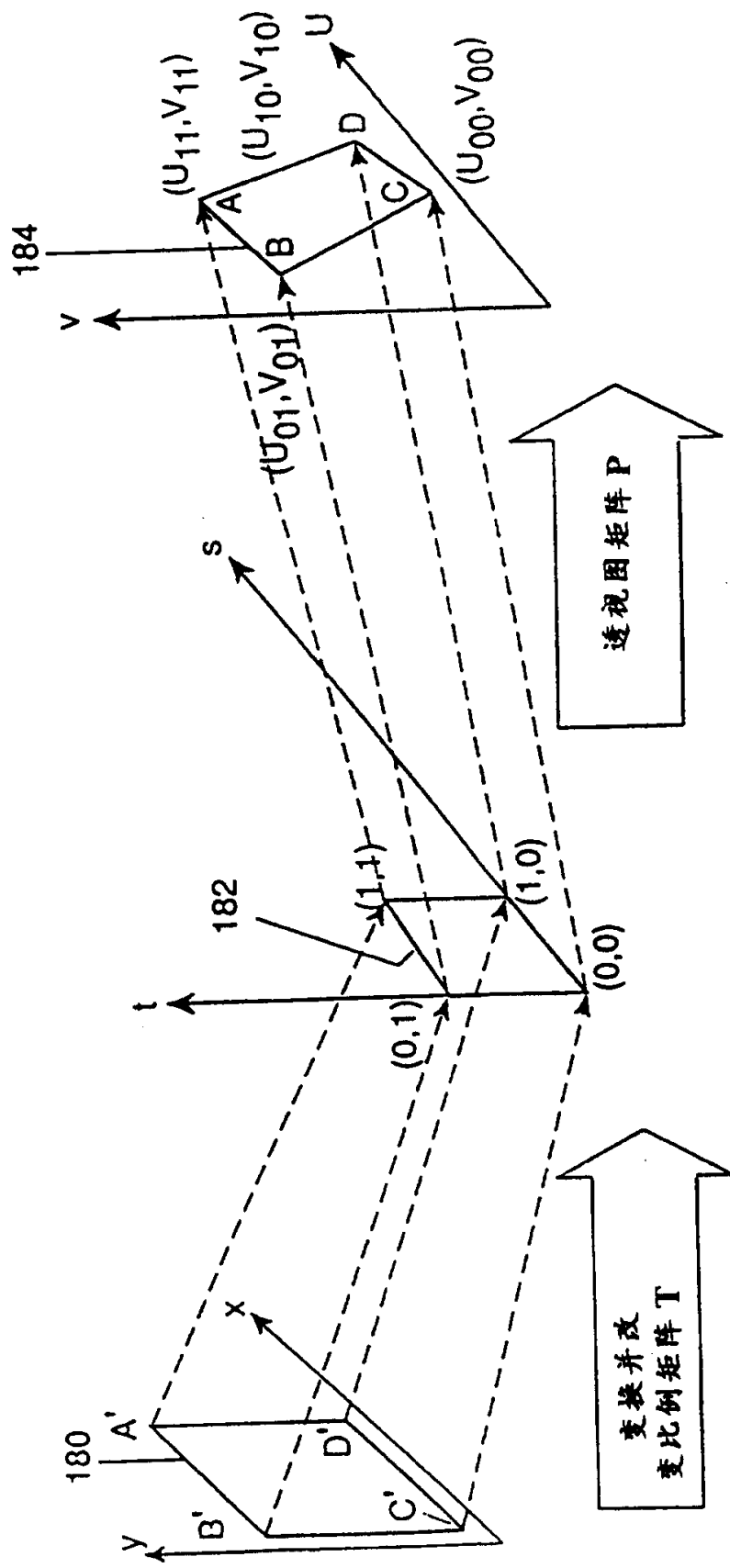


图. 15

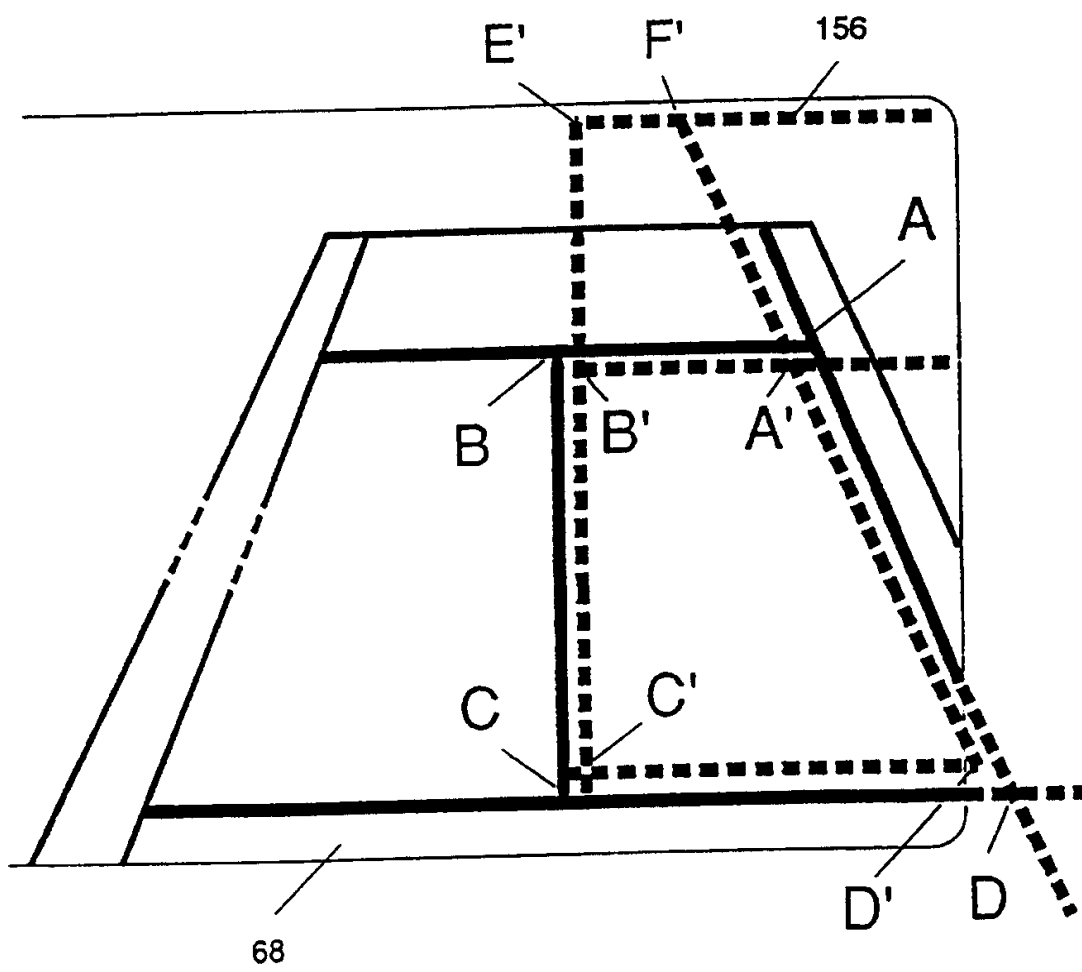


图 16

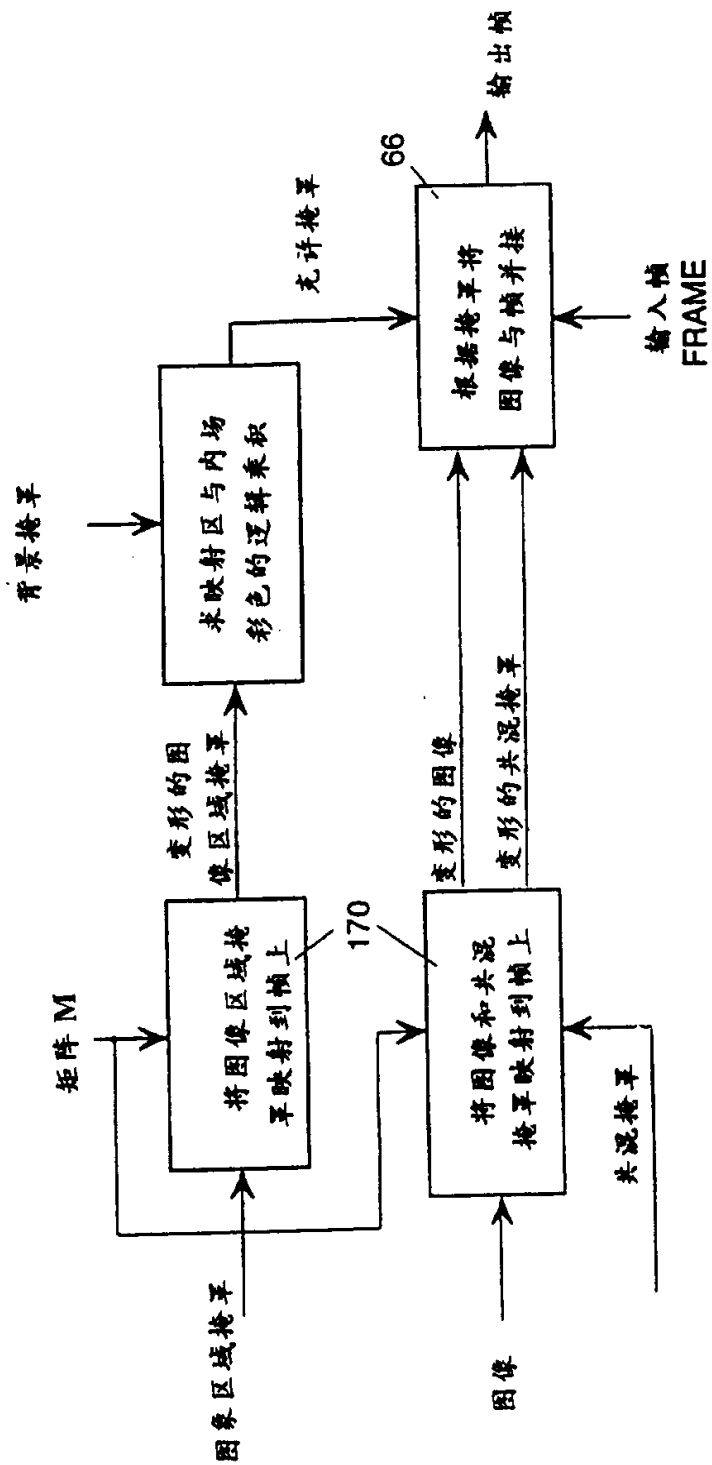


图 17

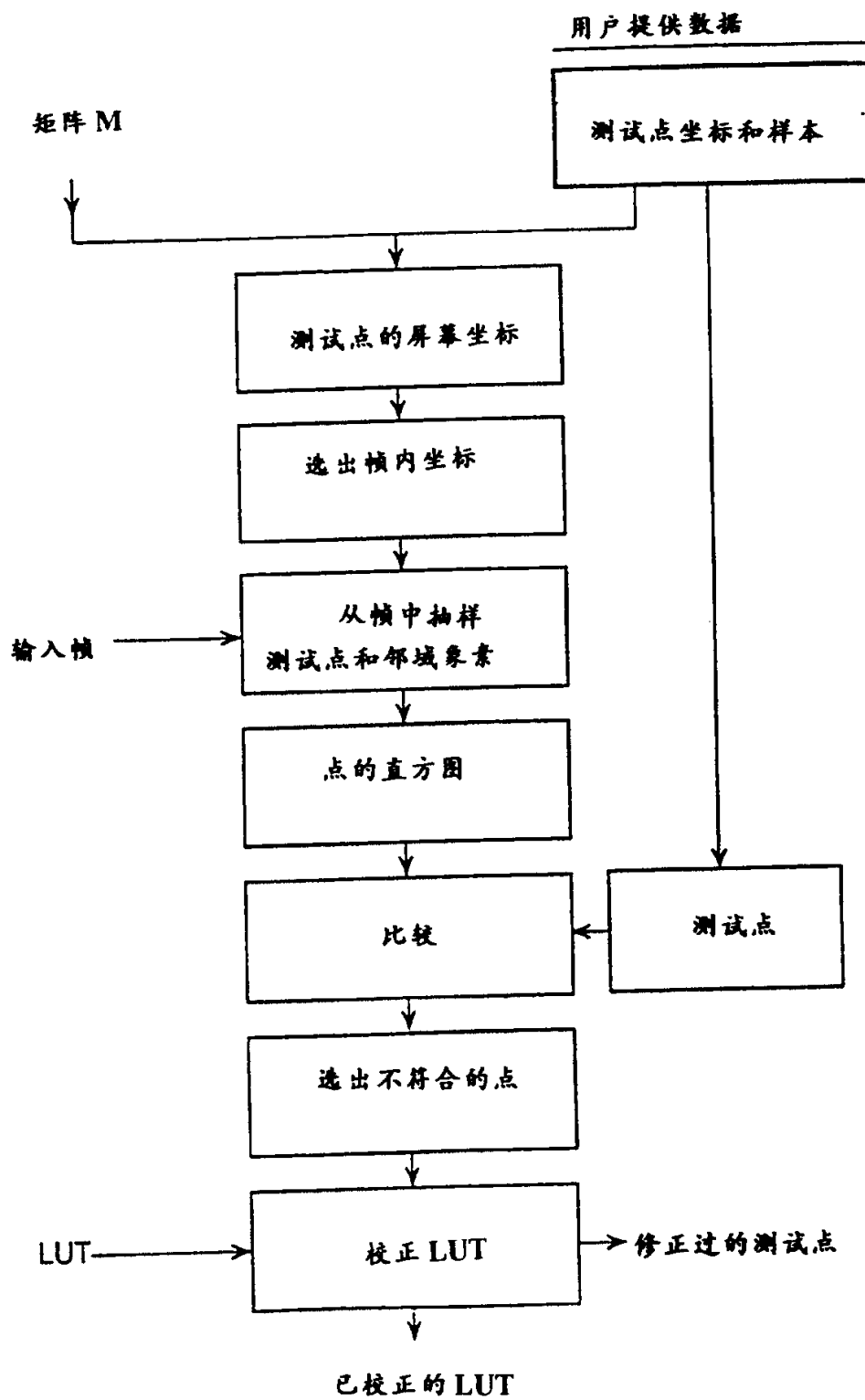


图 18

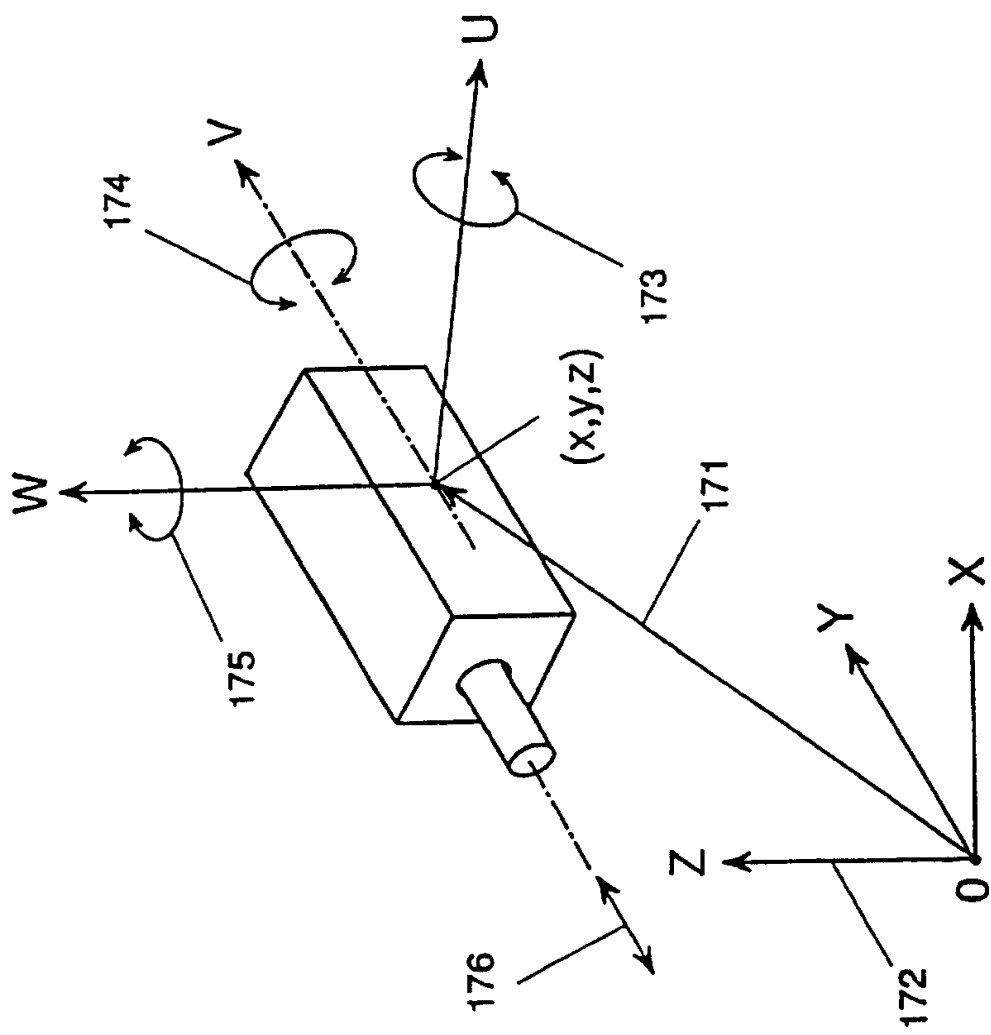


图 19

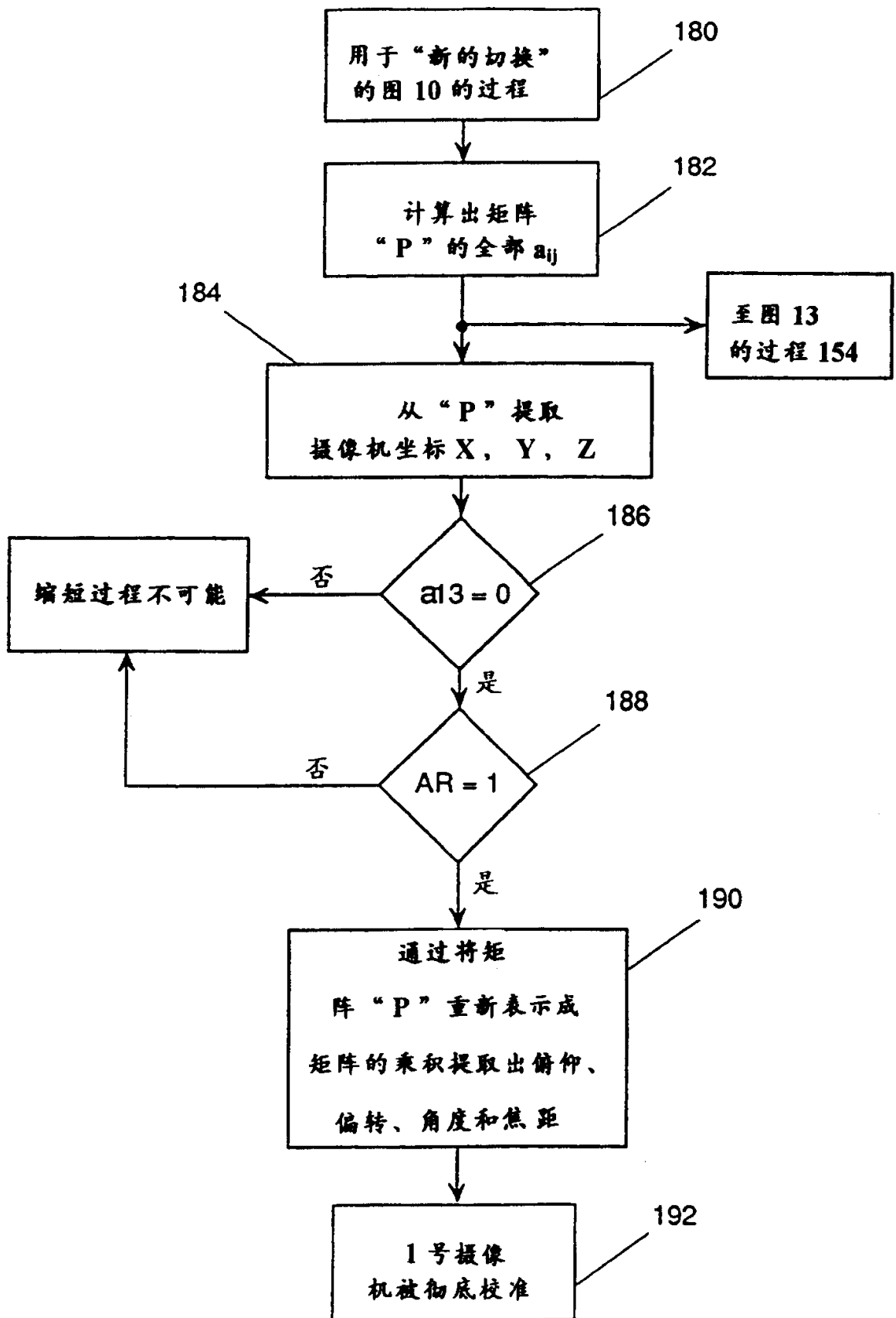


图 20

摄像机号	不变的摄像机坐标	可变偏 转、俯仰、 角度和焦距	符合系数 (PCT)
1	X,Y,Z	α_1, θ_1, f_1	
2			
3			
N			

202

194

图 21

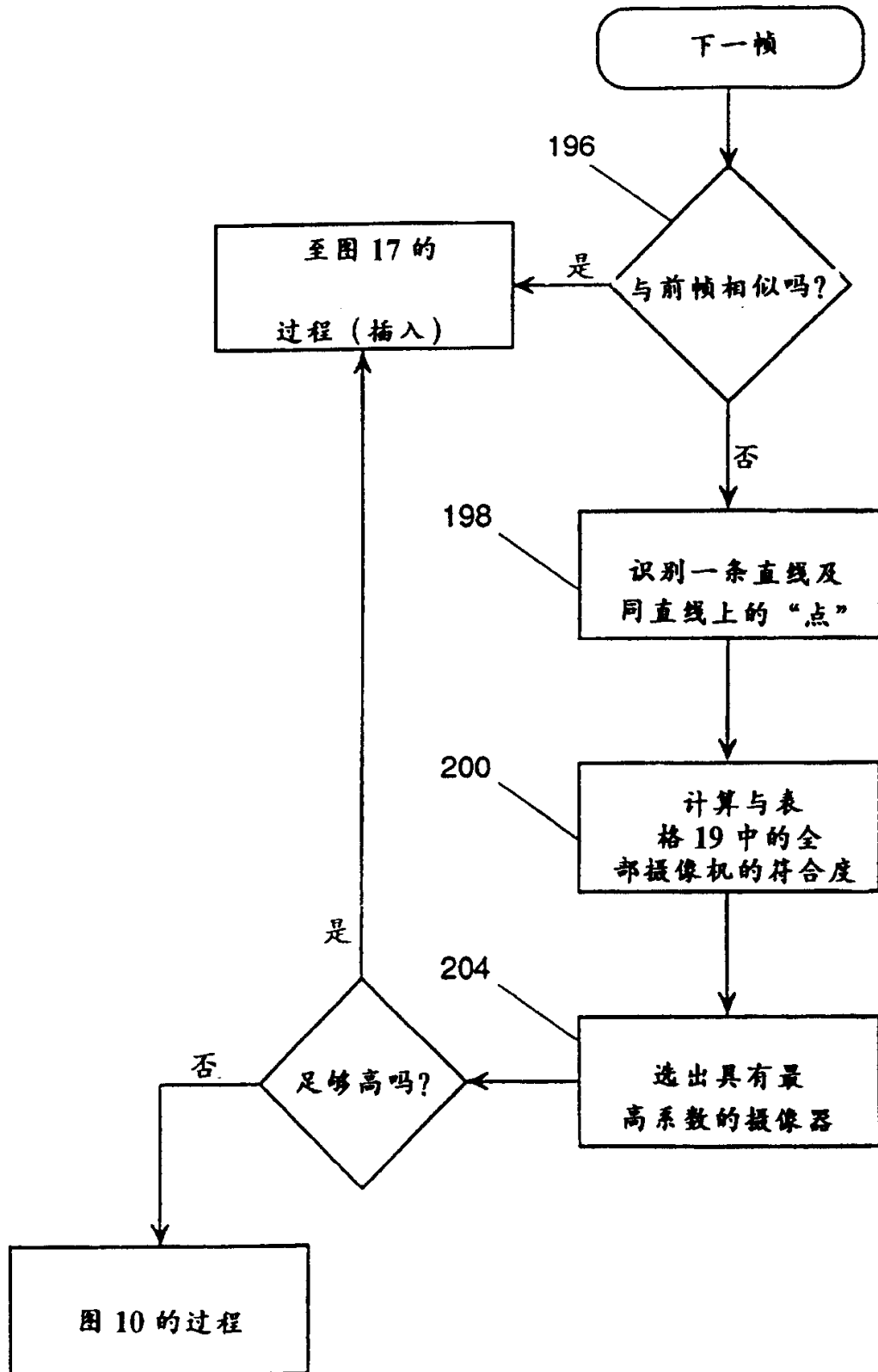


图 22