



# [12] 发明专利说明书

[21] ZL 专利号 98811849.1

[45] 授权公告日 2004 年 10 月 27 日

[11] 授权公告号 CN 1173296C

[22] 申请日 1998.12.3 [21] 申请号 98811849.1

[30] 优先权

[32] 1997.12.5 [33] AU [31] PP0778

[32] 1998.4.8 [33] AU [31] PP2865

[86] 国际申请 PCT/AU1998/001005 1998.12.3

[87] 国际公布 WO1999/030280 英 1999.6.17

[85] 进入国家阶段日期 2000.6.5

[71] 专利权人 动力数字深度研究有限公司

地址 澳大利亚西澳大利亚

[72] 发明人 P·V·哈曼

审查员 孙治国

[74] 专利代理机构 上海专利商标事务所

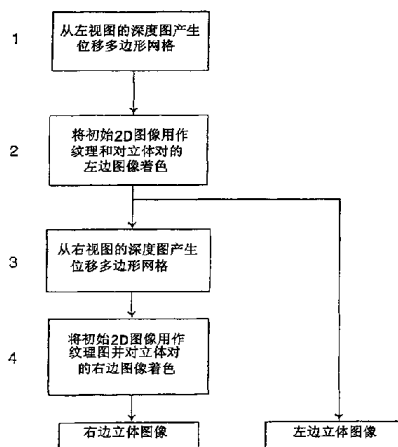
代理人 李 湘

权利要求书 4 页 说明书 18 页 附图 14 页

[54] 发明名称 改进的图像转换和编码技术

[57] 摘要

一种产生用于 2D 图像转换为立体图像的深度图的方法，包括以下步骤：识别 2D 图像内至少一个物体；为所述或每个物体分配识别标签；为所述或每个物体分配深度标签；以及确定并定义每个或所述物体的轮廓。



1. 一种产生用于将 2D 图像转换为立体图像的深度图的方法，其特征在于，它包括以下步骤：

识别 2D 图像内至少一个物体；

为所述或每个物体分配识别标签；

为所述或每个物体分配深度标签；以及

确定并定义每个或所述物体的轮廓。

2. 如权利要求 1 所述的方法，其特征在于，物体轮廓由一系列的坐标、曲线和/或几何形状定义。

3. 如权利要求 1 或 2 所述的方法，其特征在于，所述识别标签是唯一的数字编号。

4. 如权利要求 1 或 2 所述的方法，其特征在于，识别所述至少一个物体包括将 2D 图像与类属场景库比较的步骤。

5. 如权利要求 1 或 2 所述的方法，其特征在于，确定轮廓的步骤进一步包括逐个像素地跟踪物体。

6. 如权利要求 1 或 2 所述的方法，其特征在于，确定轮廓的步骤进一步包括利用线段来逼近物体的轮廓。

7. 如权利要求 1 或 2 所述的方法，其特征在于，确定轮廓的步骤进一步包括利用曲线来逼近物体的轮廓。

8. 如权利要求 1 或 2 所述的方法，其特征在于，确定轮廓的步骤进一步包括利用 Bezier 曲线来逼近物体的轮廓。

9. 如权利要求 1 或 2 所述的方法，其特征在于，确定轮廓的步骤进一步包括将物体与曲线和/或类属或几何形状库比较来逼近物体的轮廓。

10. 如权利要求 9 所述的方法，其特征在于，它还包括对曲线和/或类属或几何形状进行标度，以最佳拟合物体。

11. 如权利要求 1 所述的方法，其特征在于，深度标签包括颜色编码。

12. 如权利要求 11 所述的方法，其特征在于，白色代表靠近观察者的物体，而黑色表示远离观察者的物体。

13. 如权利要求 1 所述的方法，其特征在于，所述深度标签为数值。

14. 如权利要求 13 所述的方法，其特征在于，所述数值的范围为 0—255。

15. 如权利要求 1 所述的方法，其特征在于，所述至少一个物体进一步分割为多个部分，每个部分分配一个深度标签。

16. 如权利要求 15 所述的方法，其特征在于深度变化用斜坡函数定义。

17. 如权利要求 16 所述的方法，其特征在于所述斜坡函数为线性或径向斜坡。

18. 如权利要求 1 所述的方法，其特征在于进一步包括跟踪图像连续帧上该物体或每个物体并且为每帧内物体确定和分配深度标签。

19. 如权利要求 1 所述的方法，其特征在于进一步包括向该物体或每个物体加入纹理起伏图。

20. 如权利要求 19 所述的方法，其特征在于所述纹理起伏图通过将物体分裂为多个单元并为每个单元分配单独的深度标签来定义。

21. 如权利要求 19 所述的方法，其特征在于所述纹理起伏图由物体单个单元的亮度值来定义。

22. 如权利要求 19 所述的方法，其特征在于所述纹理起伏图由物体单个单元的色度、饱和度、颜色分组、反射、阴影、焦距和/或锐度来定义。

23. 如权利要求 1 所述的方法，其特征在于进一步包括产生每个 2D 图像的  $80 \times 60 \times 8$  位分辨率的灰度图像。

24. 一种产生用于将视频序列中 2D 图像转换为立体图像的深度图的方法，其特征在于，它包括以下步骤：

识别每一视频序列帧并对其编号；

识别视频序列内的至少一个物体；

为每个物体分配识别标签；

将所述视频序列分割为多个局部序列；

向多个操作者发送局部序列，每个操作者确定并定义已经分配了所述识别标签的局部序列内的每个物体的轮廓；

从所述多个操作者接收所述局部序列；

整理所述局部序列以重新构成视频序列；以及

为每个物体分配深度标签。

25. 如权利要求 24 所述的方法，其特征在于进一步包括在所述视频序列被分割为多个局部序列之前向序列增加安全措施的步骤。

26. 如权利要求 25 所述的方法，其特征在于所述安全措施包括从视频序列

中去除音频和/或修改视频序列的颜色。

27. 一种对将 2D 图像转换为立体图像时使用的深度图进行编码的方法，其特征在于，它包括以下步骤：

为每个物体分配物体识别符；  
为所述物体分配深度标签；以及  
定义物体轮廓。

28. 如权利要求 27 所述的方法，其特征在于所述物体轮廓由一系列的  $x, y$  坐标定义，每对  $x, y$  坐标由曲线分离。

29. 如权利要求 28 所述的方法，其特征在于每条所述曲线存储在库中并且被分配唯一的编号。

30. 如权利要求 28 或 29 所述的方法，其特征在于所述物体轮廓还包括有关每条曲线取向的数据。

31. 如权利要求 28 或 29 所述的方法，其特征在于每条所述取向为 Bezier 曲线。

32. 如权利要求 27 所述的方法，其特征在于所述物体轮廓由至少一种几何形状定义。

33. 如权利要求 32 所述的方法，其特征在于所述至少一种几何形状是按照形状和所述形状的参数来定义的。

34. 如权利要求 27 所述的方法，其特征在于所述物体深度标签的编码包括：

分配深度类型；以及  
为物体分配深度。

35. 如权利要求 34 所述的方法，其特征在于深度类型包括单值、线性斜坡或径向斜坡。

36. 一种译码深度图数据的方法，其特征在于包括以下步骤：

接收 2D 图像和对应所述 2D 图像的深度图数据；  
确定深度图数据内标识的物体；  
确定所述物体的对应深度；  
根据深度给所述物体着色；以及  
处理图像以形成变形网格，其中失真量取决于深度。

37. 如权利要求 36 所述的方法，其特征在于包括以下步骤：

在形成变形网格之前使深度图模糊，从而在物体之间提供更为平滑的过渡。

38. 一种译码深度图数据的方法，其特征在于包括以下步骤：

从多个多边形中产生未变形网格；

将深度图作用于所述网格，其中网格内多边形的标高取决于深度图图内记录的深度；

将多边形的标高转换为平移位移从而形成变形网格；以及

将变形网格作用于对应深度图数据的 2D 图像。

39. 一种译码深度图数据的译码器，其特征在于包括深度图库，其中输入数据与所述库比较，并且如果所述数据与所述深度图库内的深度图不匹配，则译码器利用如权利要求 38 所述的方法处理所述输入数据。

## 改进的图像转换和编码技术

### 发明领域

本发明通常涉及立体图像合成，特别涉及转换二维(2D)图像的改进方法以用于立体图像显示目的的编码、传输和译码。

### 背景技术

申请人在 PCT/AU96/00820 中描述了一种从初始 2D 图像产生立体显示的左右眼图像的方法，它包括以下步骤：

- a. 识别初始图像中至少一个物体
- b. 勾画每个物体的轮廓
- c. 定义每个物体的深度特性
- d. 使每个物体的选定区域沿横向位移一个作为每个物体深度特性函数的确定量，从而形成拉伸图像供观察者左右眼观看。

这些步骤可以单独或统称为动态深度插入或 DDC。

### 发明内容

本发明进一步改进了申请人先前系统的操作。本发明提供的深度图产生方法用于将 2D 图像转换为立体图像，它包括以下步骤：

识别 2D 图像内至少一个物体；  
为所述或每个物体指定一个识别标签；  
为所述或每个物体指定一个深度标签；以及  
确定和定义每个或所述物体的轮廓。

在进一步方面中，本发明提供一种编码深度图的方法，用于 2D 图像转换为立体图像，方法包括以下步骤：

将物体标识符指定给物体；  
为所述物体指定深度标签；以及  
定义物体轮廓。

物体轮廓可以由一系列的坐标、曲线和/或几何形状定义。通常情况下，

识别标签可以是唯一的编号。

在本发明的另一方面，提供了利用 Bezier 曲线产生 2D-3D 转换过程中物体的轮廓。

在本发明的另一方面，提供了利用曲线定义 2D-3D 转换过程中的物体。

在本发明的另一方面，提供了利用几何形状定义 2D-3D 转换过程中物体的轮廓。

在本发明的另一方面，提供了深度图信息的传输方法，其中信息包含在垂直消隐间隔或 MPEG 数据流内。

在本发明的另一方面，提供了利用类属库辅助 2D-3D 的转换过程。

### 附图的简要说明

为了更好地理解本发明，以下借助附图描述本发明的较佳实施例。

在附图中：

图 1 和 2 示出了将深度图数据转换为变形网格的较佳方法。

图 3、4、5 和 6 示出了如本发明所述确定物体轮廓的各种技术。

图 7 示出了示例性的失真网格。

图 8 示出了用于另一译码器的硬件框图。

图 9 示出了另一译码器译码过程的示意性流程图。

图 10 示出了未变形网格实例。

图 11 示出了圆锥体的示意性深度图。

图 12 示出了经深度图修正的示意性网格。

图 13-16 示出了将深度图 Z 标高转换为 X 位移的方法。

图 17 示出了未失真网格上的初始帧。

图 18 示出了经 X 位移图修正的示意性网格。

图 19 示出了初始帧网格与位移网格的示意性组合。

图 20 示出了另一只眼睛的示意性的最终拉伸图像。

图 21 示出了简化的位移流程图。

### 实施发明的较佳方式

#### 物体识别

被转换的 2D 图像内的物体可以由操作人员通过肉眼识别。操作人员利用计算机

鼠标、光笔、触摸指或其他装置在图像内的每个物体或物体组上作出标签并且为物体指定唯一的编号。编号可以由操作人员人工生成或者由计算机按照特定的顺序自动生成。

物体也可以利用计算机全自动识别或者由操作人员辅助计算机确定物体位置从而半自动识别。

为了自动识别物体，计算机将利用诸如物体尺寸、颜色、运动速度、颜色深浅、纹理、亮度、暗淡、焦距之类的特征以及先前、当前和后续图像之间差异。也可以利用神经网络和专家系统辅助识别物体。

在半自动物体识别中，操作人员可以通过向计算机提出何处可发现物体的图像性质的建议来帮助计算机。例如，操作人员可以建议计算机场景是“新闻播音员”类属格式，这样计算机将试图定位行为阅读者的头部和肩部、桌子和背景等。操作人员可以从可能的类属场景菜单中选择。操作人员可以手动地取消和/或纠正和调整计算机作出的物体选择。计算机程序例如可以利用神经网络或专家系统从这些纠正中学习，从而持续改进物体识别和编号的精度。

一旦识别出物体并进行编号，则随着物体在连续帧图像之间的运动，可以对其作人工、自动或半自动地跟踪。

操作人员也可以利用其他操作人员在处理同一序列或类似场景转换中获得的物体识别信息。

### 物体轮廓勾画

物体的轮廓可以人工、自动或半自动确定。

在人工勾画中，操作人员利用计算机鼠标、光笔、触摸指或其他装置跟踪物体的轮廓。操作人员可以沿被选物体的他/她认定的轮廓逐个选择轮廓像素，然后，利用曲线或类属形状库内的直线或曲线逼近、Bezier 曲线或最佳拟合。操作人员也可以从形状和尺度已经得到校正的类属库中选择或动态调整形状。例如操作人员希望选择人体轮廓，此时可以从库中搜索人体类属轮廓并且手工、自动或半自动地作些调整。操作人员还可以从诸如圆、椭圆、三角形、正方形等几何形状中选择。

在自动勾画中，计算机可利用诸如物体尺寸、颜色运动速度、颜色深浅、纹理、亮度、暗淡、焦距之类的特征以及先前、当前和后续图像之间差异。也可以利用神经网络和专家系统确定物体的轮廓。



在半自动勾画中，操作人员可以通过向计算机提出何处可发现物体的图像性质的建议来帮助计算机。例如，操作人员可以建议计算机场景是“新闻播音员”类属格式，这样计算机将试图定位行为阅读者的头部和肩部、桌子和背景等。操作人员可以从可能的类属场景菜单中选择。操作人员可以手动地取消和/或纠正和调整计算机作出的物体勾画。计算机程序例如可以利用神经网络或专家系统从这些校正中学习，从而持续改进物体勾画的精度。

一旦勾画出物体，则随着物体在连续帧图像之间的运动，可以对其作人工、自动或半自动地跟踪。

操作人员也可以利用其他操作人员在处理同一序列或类似场景转换中获得的物体轮廓信息。操作人员也可以从包括诸如圆、椭圆、三角形、正方形等几何形状的预定义库中选择，并且人工、半自动或自动地调整库轮廓以拟合选定的物体。库可以用单独的轮廓(例如新闻播音员)或根据特定物体族(例如赛马、晚间新闻)索引。

### 定义深度

物体的深度可以人工、自动或半自动确定。可以利用字母、视觉、听觉或触觉信息为物体指定深度。在较佳实施例中，物体的深度通过用特定的颜色深浅表示。对于一旦转换在 3D 位置上最靠近观看者的物体用白色，而离开观看者 3D 距离最远的物体用黑色。显然，这种惯例是可以改变的，例如采用相反的深浅或颜色表示相对或绝对深度。

在另一实施例中，可以为物体深度指定一个数值。该值可以是线性或非线性序列中的正数或负数并包含一个或多个位数。在较佳实施例中，该值介于 0—255 之间，从而使数值编码为一个字节，这里 255 表示一旦转换在 3D 位置上最靠近观看者的物体而 0 表示离开观看者 3D 距离最远的物体。显然，这种惯例是可以改变的，例如采用相反的代表或者采用其他范围的数值。

在人工深度定义中，操作人员可以利用计算机鼠标、光笔、触摸指或其他装置为物体指定深度。操作人员可以通过将定位装置放置在物体轮廓内部并输入深度值为物体指定深度。深度可以由操作人员键入作为数值、字符或图像值，并且可以由操作人员指定或者由计算机从允许值的预设范围内自动指定。操作人员也可以从允许深度库或菜单中选择物体深度。

操作人员也可以在物体范围内指定一个深度范围或者随时间、物体位置或运动或这些因素组合而变化的深度范围。例如物体可以是一张桌子，它的最近

边缘朝向观看者而最远的边缘和离开观看者。当转换为 3D 时，桌子的表观深度必需沿长度方向变化。为此，操作人员可以将桌子分割为多个段并且为每个段指定单独的深度。操作人员可以通过使物体呈深浅颜色变化而在物体内部指定一个连续变化的深度，深浅的程度表示桌子特定位置上的深度。在该实例中，浅色表示较近的物体而深色表示较远的物体。对于桌子实例，最近的边缘颜色很浅，然后逐渐加深，直到最远的边缘。

物体内的深度变化可以是线性或非线性的并且可以随时间、物体位置或运动或这些因素的组合而变化。

物体内深度变化可以是斜坡形。线性斜坡具有始点(A)和结束点(B)。定义点 A 和 B 上的颜色。在垂直线上应用从点 A 至点 B 的梯度。

虽然径向斜坡采用的是从中心点(A)到半径(B)的距离，但是与线性斜坡是相似的。

对径向斜坡的简单推广是把边缘外部形成锥形或者允许尺寸可变的中心点。

线性扩展是从线段的距离相对于到垂线的距离。在该实例中，为线段定义颜色，并且为“外部”定义颜色。沿线段定义颜色，并且逐渐变为“外部”颜色。

各种斜坡可以容易地编码。斜坡也可以基于更加复杂的曲线、方程、可变透明度等。

在另一实例中，物体可以经过一个帧周期从图像前面运动到后面。操作人员可以为第一帧内物体指定一个深度并且为最后或后续场景内物体指定一个深度。计算机随后在连续帧内以线性或预定方式对物体深度径向插值。该过程也可以是全自动的，从而使计算机根据物体相对时间运动时尺寸的变化指定变化的物体深度。

在自动深度定义中，计算机可以利用诸如物体尺寸、颜色、运动速度、颜色深浅、亮度、暗淡、焦距之类的特征以及先前、当前和后续图像之间差异。也可以利用神经网络和专家系统确定物体的深度。

在半自动深度定义中，操作人员可以通过向计算机提出何处可发现物体的图像性质的建议来帮助计算机。例如，操作人员可以建议计算机场景是“新闻阅读者”类属格式，这样计算机将试图定位行为阅读者的头部和肩部、桌子和背景等并且将这些放入逻辑深度序列内。操作人员可以从可能的类属物体和深

度中选择。操作人员可以手动地取消和/或纠正和调整计算机确定的物体深度。计算机程序例如可以利用神经网络或专家系统从这些校正中学习，从而持续改进物体深度指定的精度。

一旦为物体指定具体的深度，则随着物体在连续帧之间图像内的运动，可以对其作人工、自动或半自动地跟踪。

操作人员也可以利用其他操作人员在处理同一序列或类似场景转换中获得的物体深度定义。

### 多操作人员

为了及时地转换视频序列，需要多个操作人员处理 2D 源材料。尽管在同一前提下，但是利用在线计算机(例如因特网)，操作人员可以位于世界上不同的地方。在这种布局下，为了保证源材料的安全性，可能需要去除声音并且修改图像的颜色。这对操作人员确定物体轮廓的能力没有影响但是防止了源材料的偷盗。由于物体轮廓的实际选择相对简单，所以在低劳动力成本的国家完成是最合算的。在利用这种布局中，转换程序通常具有如下特征：

1. 监督员识别转换为 3D 的视频序列并且给每个序列帧赋予编号。
2. 如有需要监督员采用必要的安全程序。
3. 监督员识别场景内需要勾画轮廓的物体并且如上所述作出唯一的标签。
4. 视频序列随后被转换为合适的数字格式并且经在线服务发送至远处的目的地。对于较长的视频序列，这样做可能不合算，比较好的是复制在 CD-ROM 或其他复制介质上。
5. 序列由远端接收，那里有操作人员承担物体操作工作。
6. 由于操作结果是数据随后被压缩的识别物体的轮廓，所以文件尺寸大大小于初始图像。这是利用在线电子邮件服务可以方便地将物体信息返回监督员。
7. 监督员承担对接收的物体轮廓的质量控制工作并且使帧编号与初始视频源材料匹配。
8. 监督员随后将物体轮廓和初始源材料送至后续为每个物体赋予所需深度信息的操作人员。

由于深度信息的赋予是艺术性和创造性的过程，所以虽然不是必要，可考

虑由一小群操作人员在核心位置内进行。这还将保证长序列范围内物体深度的一致性。

### 定义复杂深度

为了产生更逼真的 3D 图像，有时候需要比简单斜坡或线性变化更为复杂的深度定义。对于具有复杂内部结构的深度变化很多的物体(例如树)特别需要。这种物体的深度图可以通过向物体加入纹理起伏图形成。例如，如果我们考虑一棵树，第一步是跟踪树的轮廓并为其指定深度。随后可以加入纹理起伏图从而为树上每片叶子赋予单独的深度。对于本发明来说，为了增加相对简单物体的细节，这种纹理图是有用的。

但是对于精细的细节(例如树上的叶子或其他复杂物体)，由于树随风而动或者随摄像机角度而逐帧变化，这种方法将更加复杂，所以它不是优选的方法。较好的方法是利用原始物体的照度(或黑色和白色分量)来形成所需的起伏图。一般而言，较为靠近观看者的物体元素较亮而较远的较暗。因此通过为近距离元素赋予较亮的照度而为远距离元素赋予较暗的照度可以自动形成起伏图。这种技术的优点是物体本身可以用来形成自己的起伏图并且自动跟踪物体逐帧的运动。物体的其他属性也可以用来形成起伏图，这些包括但是并不局限于色度、饱和度、颜色分组、反射、阴影、焦距、锐度等。

从物体属性获得的起伏图比较好的是进行标度，从而使物体内部深度变化的范围与整个图像总的深度范围一致。

### 深度图

检测物体、确定其轮廓并指定深度的过程被称为深度图的形成。在较佳实施例中，深度图由  $80 \times 60 \times 8$  位分辨率的灰度图像组成，从而可以 256 种深度定义相关的 2D 图像内的物体。

曲线形状可以定义为序列  $xy$  坐标之间的距离与这些点之间曲线离直线的位移的比率。 $x_1y_1$  和  $x_2y_2$  位于直线 A 上并且由曲线连接。这些点之间的曲线具有从直线 A 到曲线中点的最大位移 B。曲线因此定义如下：

$$\text{曲线} = B/A$$

它比较好的是取值在  $-128 \sim +128$  之间，0 表示两点之间为直线。值得指出的是，由于为曲线指定的值是两个测量值的比率，所以同一曲线值可以指定给

具有相同 B/A 比率的其他曲线。

### 深度图的编码

深度图可以有多种编码方式。在较佳实施例中，物体编号、深度和物体轮廓编码如下。考虑如图 3 所示人体轮廓。该人被分配的物体编号为，深度为 20。物体的轮廓已经如上所述确定并且位于具体的 x、y 位置。在物体轮廓方向改变的地方作上特殊的标记。该标记可以是字符、形状、颜色或其他形式的视觉指示。每个这样的标记具有特定的 x、y 位置。在较佳实施例中，其范围为 0~255。在每对 x、y 位置之间存在一条曲线。每条曲线可以通过从所有可能曲线形状的库中选择确定。在较佳实施例中，每条曲线被赋予-127~+128 范围内的数值，从而可以用 1 个字节定义曲线。从 x、y 位置顺时针行进至下一 x、y 位置的曲线可以被赋予正数而逆时针行进的曲线可以被赋予负数。也可以采用其他指定方式。

### 深度阈值

将深度阈值加入转换算法以确保阈值前的物体不变形。在前景物体与背景物体相交时，这避免了前景物体边缘的某些细小变形。

在较佳的转换算法中，深度图被用来生成构成最终场景的 3D 轮廓的连续深度图。当阈值被用于该处理时，深度图被处理为检测阈值转变，并且转变之上和之下的深度被单独处理。

该物体的深度图数据因此定义如下：

<物体编号><物体深度><x1, y1, 曲线 1, x2, y2, 曲线 2……x1, y1>

物体深度信息包含了产生当前物体深度所需的数据。如上所述，该深度数据可以是单个值、斜坡(线性、径向或其他)或者其他描述单个物体深度的方法。下列方法示出了编码单个物体深度数据的可能手段。

单个深度值深度数据可以编码如下：

<深度标志 1><深度值>

对于深度值具有线性斜坡变化的物体，深度数据编码如下：

<深度标志 2><x1, y1, 深度值 1, x2, y2, 深度值 2>

这里物体的深度从 x1、y1 处的值 1 线性变化至 x2、y2 处的值 2。

对于深度值呈非线性斜坡变化的物体，深度数据编码如下：

<深度标志 3><x1, y1, 深度值 1, x2, y2, 深度值 2, gamma>

这里 gamma 为描述深度在 x1, y1~x2, y2 之间范围内非线性变化的数值。

对于深度值呈径向斜坡变化的物体，深度数据编码如下：

<深度标志 4><x1, y1, 深度值 1, 半径, 深度值 2>

这里物体在 x1、y1 的深度值为 1 并且深度呈线性变化至离开 x1、y1 的所有半径像素处的深度值 2。

应该理解的是，一旦发送物体深度值，则无需再发送深度图，直到物体运动或形状改变。如果只是物体位置变化，则可以通过如下为物体位置分配偏移而发送物体新的位置：

<物体编号><x 偏移, y 偏移>

同样，物体深度变化而位置或尺寸不变时可以是发送下述

<物体编号><深度>

将会理解的是，邻近的接触的物体将共享 x, y 坐标并且因此在需要发送以唯一定义场景内每个物体深度图的 x, y 坐标内存在冗余。

为了使发送或存储的附加数据量最小，需要压缩包含深度图的数据。压缩可以利用任何形式的压缩算法并且对于本领域内技术人员来说许多都是已知的。压缩的实例包括但不局限于游程长度编码和 Huffman 编码。由于物体可能在帧之间没有运动，所以仅需发送帧之间深度图的差异。测量并处理帧之间差

异的技术对于本领域内技术人员来说也是已知的。

将会看到，如上变形网格传输中所述，深度图信息可以包含在模拟电视信号的垂直消隐间隔(VBI)或数字电视信号的 MPEG 或其他数字传输流内。同样，深度图数据可以加入 DVD 的 VOB 文件内。

数据如何包含在 VBI 和 MPEG 数据流中是已知的并且较佳实施例是当前用于将封闭的标题和图文包含在标准电视图像内的技术。在另一较佳实施例中，数据可以包含在 MPEG 数据流中的用户数据区内。

对于将数据包含在 VBI 或 MPEG2 流的情况，以下计算表明了可能需要的数据大小。

假定：

VBI 规定允许 32 字节/视频线

每幅图像的最多物体数 = 20

每个物体的最大 X, Y 坐标 = 20

物体#、物体深度、X, Y 和形状数据每种占用 1 个字节

则字节/物体 =  $1+1+3(20) = 62$  个字节

因此对于 20 个物体，VBI 数据 =  $20 \times 62 = 1240$  个字节/帧

应该理解的是，这是最差的情况，实践中一般的场景需要 200 字节/帧。如果采用合适的数据压缩并且考虑到冗余，该数值将明显减小。

对于将信息包含在 MPEG 数据流的情况，MPEG 标准允许向接收位置发送数据流。提供 MPEG 流内数据发送的技术可以用于向接收译码器发送深度图数据。也可以将该信息包含在 MPEG 信号的其中一个声道内。在记录 MPEG 信号的介质(例如 CD-ROM 或 DVD)上，信息可以包含在数字音频文件内作为分离的数字或模拟文件，或者记录在其他装置的盘片上。对于本领域内技术人员来说，其他技术也是显而易见的。

也可以发送初始深度图作为 MPEG 数据流的一部分。在较佳实施例中，深度图的分辨率可以从  $640 \times 480 \times 8$  压缩至  $80 \times 60 \times 8$  而在最终 3D 图像内物体深度的误差不会很明显。这种分辨率与 MPEG 编码视频信号内 DCT 块尺寸相同。因此，通过向转换为 3D 时定义每个块深度的 DCT 块加入附加信息，深度图信息可以包含在 MPEG 信号内。如上所述，深度图也可以包含在 MPEG 数据流内，例如音频信道或其他本领域内技术人员熟悉的方法。分辨率减小的深度图也可以在包含入 MPEG 流之前利用标准的压缩技术(包括但不限于 JPEG、MJPEG、

MPEG 等)进行压缩。

在进一步的较佳实施例中，物体轮廓利用 Bezier 曲线定义。考虑如图 4 所示人的轮廓。Bezier 曲线应用至轮廓，从而得当所示的  $x, y$  坐标。物体的深度图可以定义为

<物体编号><物体深度>< $x_1, y_1, x_{1a}, y_{1a}, x_{2b}, y_{2b}, x_2, y_2, \dots, x_{1b}, y_{1b}$ >

也可以生成如图 5 所示仅需 3 个  $x, y$  坐标的 Bezier 曲线并且定义如下

<物体编号><物体深度>< $x_1, y_1, x_{1a}, y_{1a}, x_2, y_2, \dots, x_{8a}, y_{8a}$ >

由于需要较少的元素数来定义曲线，所以这种方法比较好。

在进一步的较佳实施例中，物体轮廓利用几何形状来定义。考虑如图 5 所示的人的轮廓。将几何形状用于轮廓从而得到所示构造。构成头部的圆的圆心为  $x_1, y_1$ ，而半径为  $r_1$ 。三角形可以描述为  $x_{2a}, y_{2a}, x_{2b}, y_{2b}, x_{2c}, y_{2c}$ ，并且对于其他多边形也是如此。每种几何形状具有一般的形式

<形状><参数>

物体的深度图因此定义为

<物体编号><物体深度><形状 1><参数>……<形状 n><参数>

显而易见的是，利用这些方法生成的轮廓和/或深度图，不管是压缩还是未压缩的，都可以合适的模拟或数字格式和介质带有或不带相关的 2D 图像存储。存储介质包括但不限于软盘、硬盘、CD-ROM、激光盘、DVD、RAM、ROM、磁带、录像带、录像机等。存储的轮廓和/或深度图可以在后面调用并且/或回放以重建深度图，用于产生生成 3D 图像所用的变形网格或者作进一步的调整和精细调整用。

## 译码器



如上所述，变形网格可以用来将 2D 图像转换为 3D。

现在可以从深度图产生所需的变形网格。深度图本身从 2D 视频内发送的附加信息中产生。可以实时、半实时或离线地从深度图产生变形网格并且可以在本地完成，也可以经合适的传输介质在远端完成。可以软件或硬件方式完成生成。

因此与发送变形网格的子像素点作为 2D 图像部分不同，可以发送重新生成深度图所需的信息。深度图随后可以在译码器处重建并且转换为变形网格。这些转换可以在接收处实时、半实时或离线地完成并且可以软件或硬件方式完成。图 1 的软件流程图和图 2 的硬件示出了从深度图数据转换为深度图。随后又转换为变形网格的较佳方法。软件转换过程的单个单元的功能如下：

图像序列源—2D 影片或视频或其他图像序列源。

区域和深度源—该信息与图像序列一起发送并且在较佳实施例中包含在 VBI 或 MPEG 数据流内。它包含每个物体的位置、形状和深度信息。

将带深度的区域作用至深度图—为了表示物体，物体内部的“区域”被按照深度信息填充/涂上阴影。阴影区域外部的所有区域是未接触的。该过程重建初始的深度图。

使深度图模糊—随后模糊(高斯、快速或其他方法)硬的深度图以去除锐利的边缘。模糊化在物体之间提供了平滑的过渡以消除图像重叠。模糊化在水平方向上权重略大。垂直模糊通过渗透入图像之上和之下防止图像撕裂，从而使远近物体的过渡更为平滑。

利用深度图处理图像—模糊化的深度图随后被用作变形网格位移的源，白色为最大位移，黑色为无位移。沿水平轴的变形量按照给定轴位上深度图的深度标度。在较佳实施例中，左边图像的位移向右，右边图像的位移向左。可以对图像阴影整体强迫视差，从而白色(前景)位移物体汇聚到屏幕水平。黑色(背景)区域随后具有与未偏移图像相等的强迫视差。位移的方向和强迫视差可以变化以适合显示被转换图像的 3D 显示系统特殊的要求。

一旦生成变形网格，则如上所述将 2D 图像变换为 3D 图像。

图 2 示出了从变形网格产生分离的左和右图像的硬件转换器的较佳实施例，它可以是全数字的。图 2a 示出了实现该方法并且操作如下。

系统采用两线存储器，它是多端口的以允许同时存取。视频线被写入其中一个线存储器，而另一线存储器被读取以产生输出视频信号。在当前线的结束

处，交换线存储器。

深度信息从视频信号中提取以重新生成当前图像的深度图。对于每个输出像素，深度图被平移一个(变形网格的)像素偏移。当视频线从线存储器中读取时，像素偏移被加入像素计数器。像素偏移是一个分数值，所以需要所需像素每侧的像素值并且内插中间值。来自视频译码器的奇/偶场信号被用来控制场序列的视频输出并且使观众快门镜头与输出视频信号同步。基本的电路可以复制以产生需要该视频格式的 3D 显示的左右视频信号。

图 2b 示出了 DDC 译码器的功能框图。第一个过程是从输入的可插入 VBI 或 MPEG 数据流的视频中提取物体数据。提取的数据将是压缩格式并且随后利用微处理器解压缩。来自微处理器的输出是原始物体轮廓的信息并且经过再次处理以产生每个物体的深度信息。该数据被送至一组三个旋转场缓冲器内，缓冲器由微处理器控制。第一缓冲器重新创建原始深度图。深度图随后被送至下一缓冲器，在那里经过水平和垂直模糊化。一旦完成模糊化，最终的数据被施加在最后缓冲器，在那里数据被送至图 2a 所示的深度-像素偏移转换器内。一旦数据被转送至偏移转换器，最后缓冲器就被清除并且准备接收下一深度图。

图 2c 示出了 DDC 译码器过程。这示出了该过程的时序图并且假定当前微处理器的速度无法同时承担所有的译码处理。因此译码过程以流水线过程顺序执行。随着微处理器性能的提高，预期如果不是全部也是多个这样的过程将被同时执行。在图 2c(1)中，示出了 4 幅视频帧，每帧包含奇场和偶场。在(2)中，产生帧 4 的物体列表，而在(3)中产生帧 4 的深度图。在(4)中，应用水平和垂直模糊化，并且在(5)中输出帧 4 的深度图并清除缓冲器以准备下一物体列表。在(5)中，帧 4 的深度图和 2D 图像可以同时利用以转换为 3D。值得注意的是，图 2c 示出了单个帧的处理，实际上，在任一时刻，硬件的不同部分产生 4 幅不同帧的深度图。

### 替换的译码器

如上所述，当前可用的微处理器的速度无法同时承担所有的译码处理。因此以下描述无需使用快速微处理器的译码器的另一较佳实施例。该替换译码器利用处理 2D 和 3D 计算机图形的集成电路。这种专用图形处理器能够每秒提供 500,000 个多边形。由于这些集成电路的产量很大，所以价格便宜，并且可制造出低成本的 DDC 译码器。译码器采用图形处理器的最简单的多边形生成能

力，无阴影纹理映射的多边形。

通过解释貌似手工完成的处理可以更为容易地理解译码过程。这可由图 9 和以后各图的流程说明。过程开始时产生未变形网格作为  $xy$  平面内完成相对平滑变形所需的多个许多多边形。在较佳实施例中，每场一般采用 10,000 多边形。图 10 示出了未变形网格部分的实例。被转换为 3D 的物体(在该实例中是图 11 所示顶部朝向观众的圆锥)的深度图被作用于修改为网格多边形的  $z$  轴标高依赖于深度图内对应像素值的网格。它示于图 12 中。该过程的下一步是将每个多边形的  $z$  轴标高转换为相等的  $x$  位移量。这示于图 13—16。在图 13 中，示出了  $x$  轴部分— $z$  标高网格。在图 14 中，沿  $x$  轴选择一行点并且围绕点  $y=0$  旋转 90 度。图 15 示出了在 45 度点旋转的效果而图 16 是 90 度旋转后的效果。对于有效地将深度图  $z$  轴标高转换为  $x$  位移量的所有  $x$  行重复该过程。

该过程的下一步骤是如图 17 所示将初始视频帧映射到未变形网格上。未变形网格随后如图 18 所示形态化为前述产生的  $x$  位移图。最终的视频图像随后如图 19 所示根据网格位移扩张。这与在前面申请 PCT/AU96/00820 中所述的拉伸图像具有相同的效果。拉伸图像可以用来构成一对立体图的一个视角，另一个通过将图 13 中的点旋转-90 度(将产生网格并对应图 20 所示的图像)形成。

当利用 2D/3D 图形处理器在硬件中实现该过程时，可以省略将  $z$  轴标高转换为相等的  $x$  位移量的步骤。由于已知靠近观众的多边形需要比远离观众的多边形横向偏移更多，所以图 18 网格的位移可以直接从图 11 的深度图中产生。由于深度图的灰度值与每个对应多边形的偏移具有直接的关系，所以可以这样做。图 21 的流程图示出了该简化过程。

### 替换的硬件译码器

图 8 示出了基于 2D/3D 图形处理器的硬件 DDC 译码器的框图。从 DDC 数据中提取和生成深度图如上所述并且示于图 2b 中。译码器的操作如下。输入的视频送至 DDC 数据译码器，从视频流中提取 DDC 信息并且回复每个视频场的深度图。视频也可以转换为 RGB、YUV 或其他标准的视频格式并且放入双场存储器。这使得视频场在加载新的场时被同时读取入 2D/3D 图形处理器。从 DDC 数据译码器输出的深度图被送至定义被 2D/3D 图形处理器处理的多边形形状的深度图—多边形网格转换器。其他向图形处理器的输入是初始 2D 视频图像，它

被用作作用于多边形的纹理图。图像处理器的输出被送至使视频以交错格式读取的场存储器。随后被送至 PAL/NTSC 编码器，其输出将是标准的场顺序 3D 视频信号。

### 深度图的再使用

显然，由于同样的深度图在再次显示同一或类似场景时将被复用，所以无需向接收机发送整个深度图。因此在存储器中译码器保存先前发送的深度图序列供复用而无需重新处理先前发送的深度图。在译码器存储器(可以是易失或非易失存储器并包括但是不局限于 RAM、EEPROM、快闪存储器、磁性或光学存储器等)内可以保存深度图或最终的变形网格。也可以在译码器内存储类属深度图和/或变形网格。这常常发生在无需发生或转换深度图的图像转换中。通过将数据包含在对译码器唯一识别哪个缺省的深度图要应用的视频信号内，可以选择正确的深度图。译码器应该能接收新的或改变的深度图，从而维护译码器内的深度图和/或变形网格库。该库可以但是不局限于保存在下列介质内，例如 RAM、EEPROM、快闪存储器、磁性或光学存储器等。可以通过发送包含在视频信号内的特定深度图或变形网格更新库。库也可以借助外部或内部插入模块(包含这样的深度图或变形网格)并经视频信号、调制解调器或因特网下载至译码器来维护。对于本领域内技术人员来说，维护库的其他手段是显而易见的。

在较佳实施例中，包含在视频信号内的 DDC 数据的通用格式可以包括向译码器指示随后数据性质的头部标志。可以有多种已有标准用于这种格式，该格式通常为：

<标志#><译码器处理的数据>

标志的实例包括但不局限于：

标志 1—下列数据为深度图

标志 2—下列数据涉及已有物体的重新定位

标志 3—下列数据涉及物体深度的变化

标志 4—下列数据涉及先前发送的深度图的复用

- 标志 5—下列数据涉及库内深度图的使用
- 标志 6—下列数据涉及库内深度图的修改
- 标志 7—下列数据涉及库内新深度图的增加
- 标志 8—下列数据涉及已有库深度图的删除
- 标志 9—下列数据涉及运动视差延迟的利用
- 标志 10—下列数据涉及强迫视差的利用
- 标志 11—下列数据涉及数学算法的使用
- 标志 12—下列数据涉及数学算法库的使用

每个数据包的长度可以不同，它唯一定义每个包并且减少了标志的需要。在前面描述中，相同的处理可作用于变形网格上。

译码器应该能通过从库内标称范围内自动选择确定作用于相关 3D 图像的最合适的深度图。例如，DDC 数据可以指导译码器搜索特定索引点之间或通过类属(即晚间新闻、马赛)分类的深度图。译码器随后根据物体尺寸、形状、速度、方向、颜色、深浅、浓淡选择合适的图。

初始深度图作为译码处理的副产品，在译码过程中生成，它可以合适的格式与需要 2D 图像和物体深度信息的 3D 显示系统使用。这些显示在性质上可以是自动立体的和/或有容量的。

### 替换的途径

网格变形处理可以用数学算法定义。该算法可以存储在译码器内并且 DDC 数据包包含算法应用的参数。例如考虑通用公式

$$f(x, y) = [1 - \exp(-(|x| - rx) \cdot dx)] \cdot \sin(((PI \cdot x) / rx) + PI/2) \cdot [1 - \exp(-(|y| - ry) \cdot dy)] \cdot \sin(((PI \cdot y) / ry) + PI/2)$$

这里

PI—常数 3.14159...

|x|—x 的绝对值

rx—x 的范围， $-rx \leq x \leq rx$

ry—y 的范围， $-ry \leq y \leq ry$

dx—x 的斜率因子

## dy-y 的斜率因子

如果下列值通过 DDC 数据被传递至方程，则产生图 7 的变形网格

$$rx=ry=50$$

$$dx=dy=0.1$$

就 DDC 数据而言，发送下列数据

<标志 11><50, 50, 0.1, 0.1>

此外，这些参数可以库的形式存储在译码器内部的存储器内并且通过在 DDC 数据内发送库索引调用。

就 DDC 数据而言，发送下列数据：

<标志 12><库索引>

以下考虑利用标志 9 运动视差的实例。现有技术表明，水平方向运动的 2D 图像可以利用运动视差转换为 3D。图像需要因照相机的水平运动(即拍摄全景)而运动。在这种技术中，观众的其中一只眼睛接收当前的视频场，而另一只眼睛接收前一场，即，提供给每只眼睛的图像之间存在延迟。哪一只眼睛接收延迟的图像和延迟量的选择取决于 2D 图像水平运动的方向和速度。延迟一般在 1~4 场范围内。可以通过考虑 2D 图像内的总体运动矢量并根据矢量的尺寸、方向和稳定性选择这些参数来选择方向和延迟。在现有技术中，需要在观看位置实时完成这些计算。较佳的方法是计算运动矢量和发送位置的场延迟的方向和量，随后将这些值作为视频信号的一部分发送。因此在较佳实施例中，发送的数据如下：

<标志 9><方向和延迟>

这里<方向和延迟>一般在-4~+4 的范围内。

DDC 译码器随后恢复该数据并利用其将正确的场延迟量和方向插入处理图像。

通过将摄像机加入已有的 2D 视频或影片摄像机(它利用变焦透镜和锐度检测算法确定摄像机观察的图像内的物体深度)也可以实时获得变形网格。物体深度可以从一对立体摄像机中获得,每幅图像内像素之间的相关度指示物体深度。这种结构的输出在处理前提供了变形网格,可以用来生成深度图。为此处理初始 2D 图像并且采用阴影或其他指示来表示这里描述中所述的物体深度。每个物体的轮廓可以从诸如物体尺寸、颜色、运动速度、阴影、纹理、亮度、浓淡以及先前、当前和后续帧之间差异之类的物体特性中获得。神经网络和专家系统也可以用来辅助识别物体。也可以在摄像机内移动图像从而在摄像机图像传感器上获得后续图像的物理偏移。这种偏移可以本领域内技术人员熟知的光学、电光、机械、电机、电子或其他方法产生。偏移可以是单向的,即 x 方向,或者多方向的,即,顺序地或随机地。在摄像机传感器上,靠近摄像机的物体的偏移将更大。通过使连续图像内的像素相关可以确定每个物体的深度。也可以采用多个摄像机。

也可以采用其他技术确定场景内的物体深度。这些包括但不限于根据光学、激光、超声或微波原理利用范围寻找器或者将网格投影在场景内的物体上并根据最终的网格变形确定物体深度。

大量的计算机辅助绘图(CAD)软件包可以产生图像的线帧模型。这些线帧模型是物体侧面的投影,可以用来确定场景内物体的位置。

同样,来自 3D Studio 之类软件包的 3D 非立体图像的着色处理部分允许输出摄像机与每个像素之间的距离。这种着色可以产生灰度图像,最靠近的物体为白色,离开摄像机最远的点为黑色。这种灰度图可以用作兼容的深度图,用于转换为立体 3D。

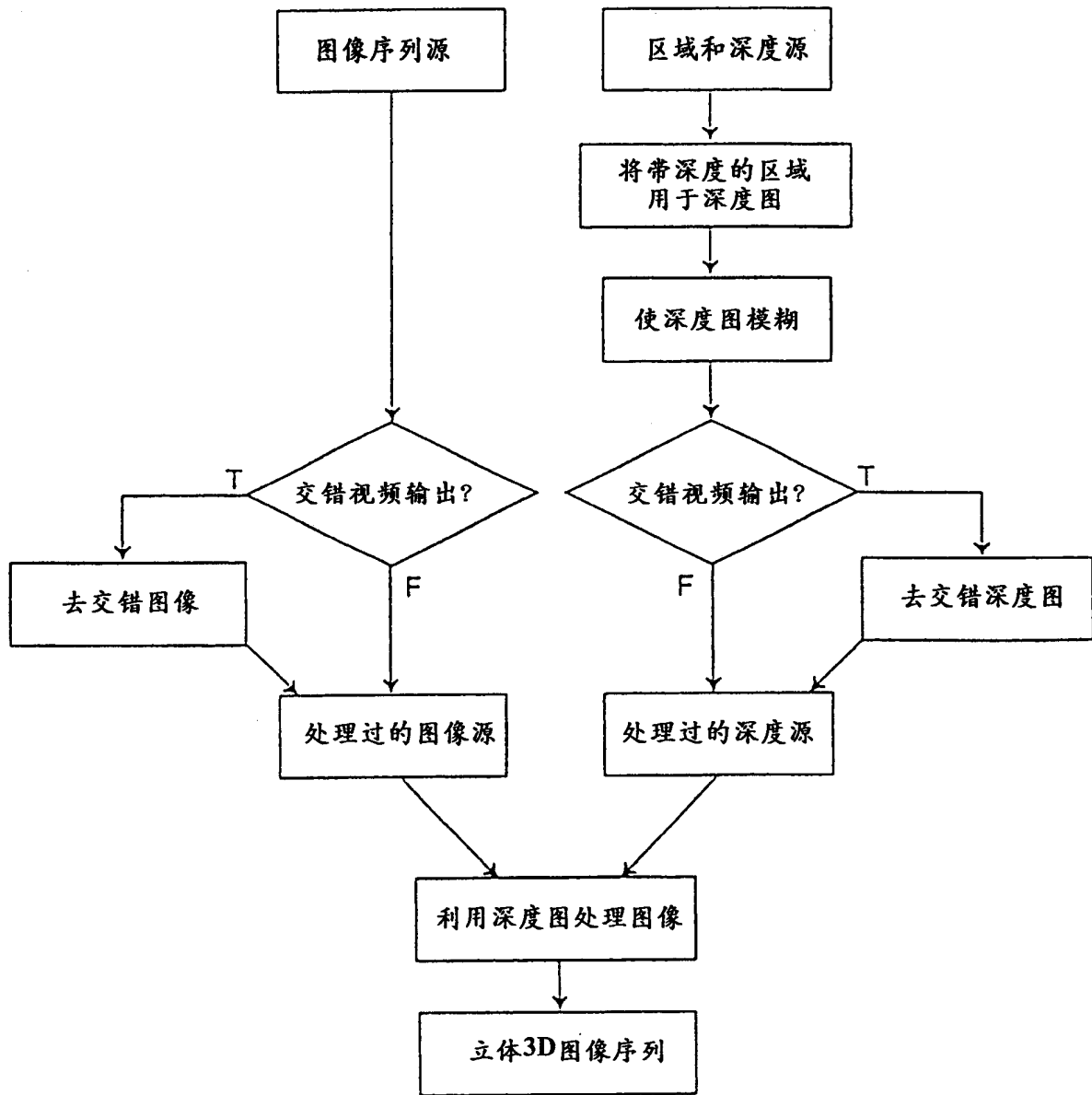


图 1



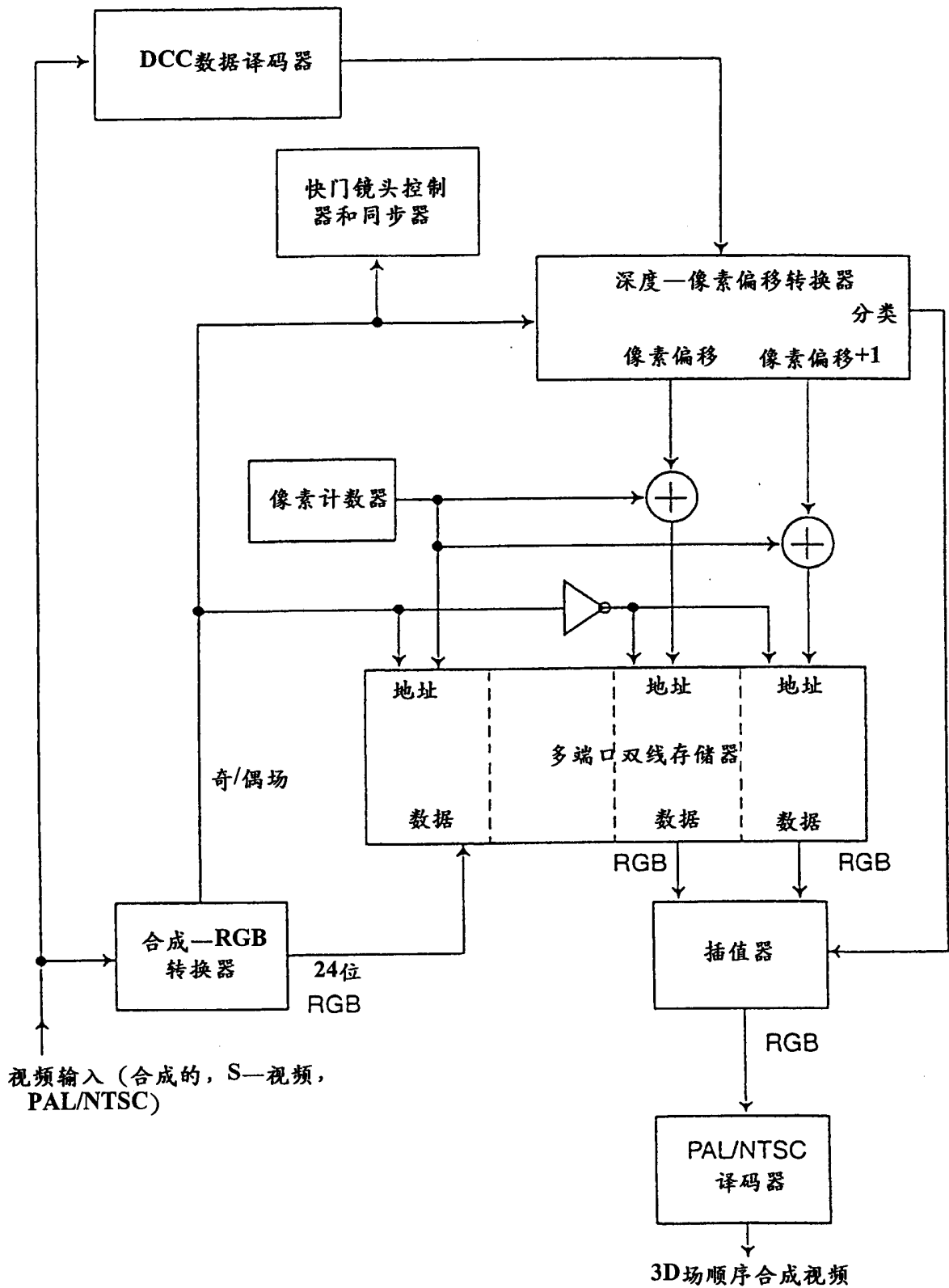


图 2a

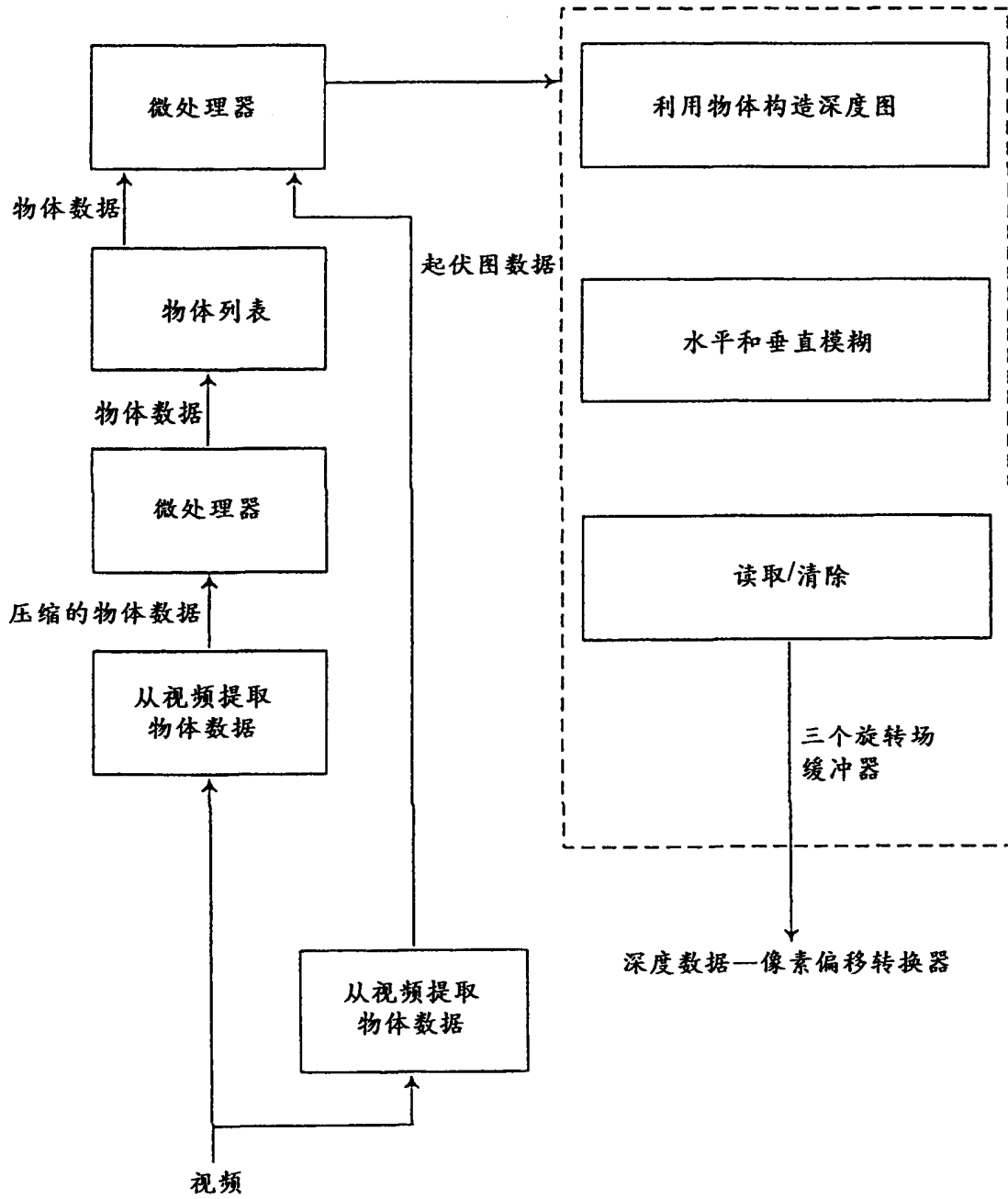
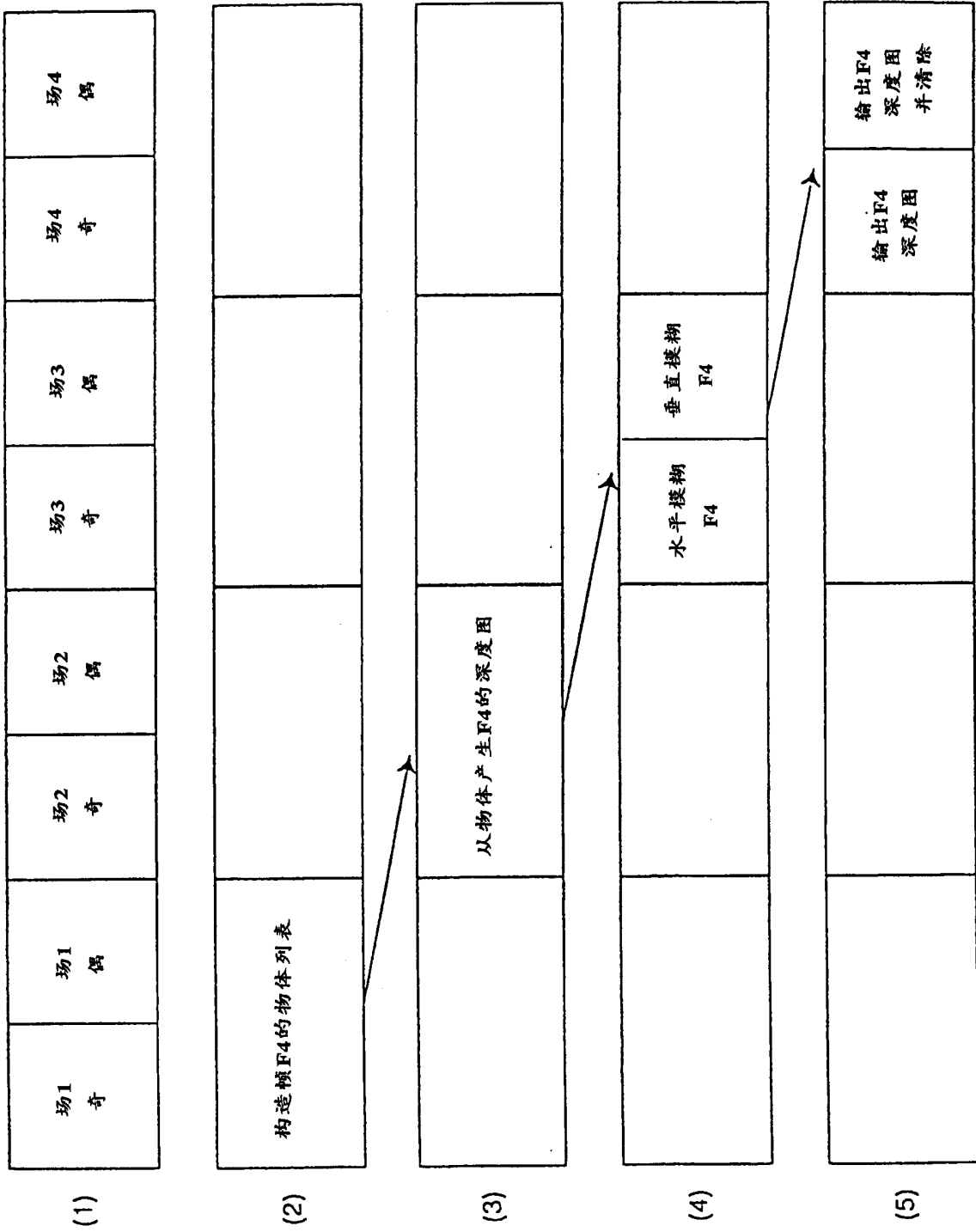
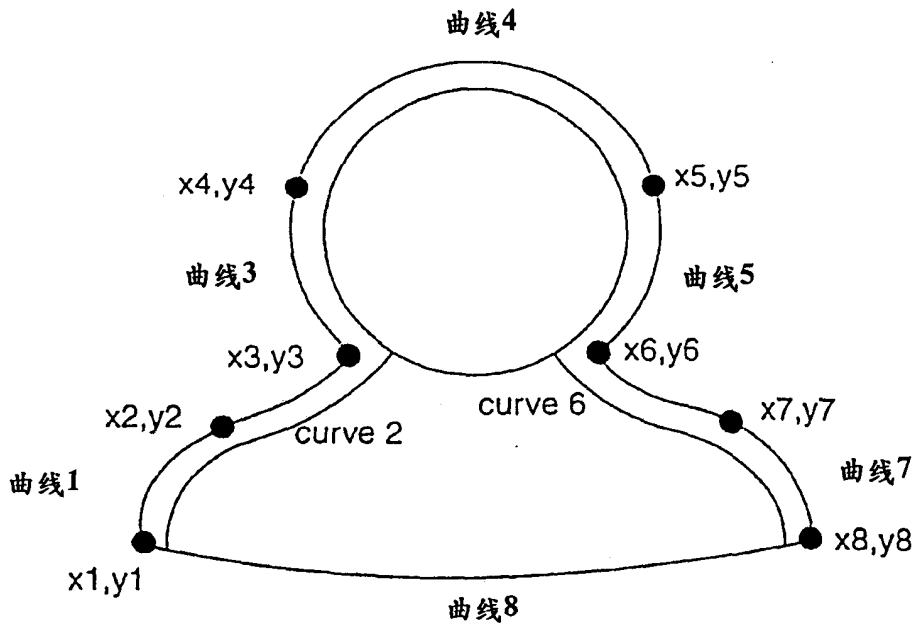


图 2b

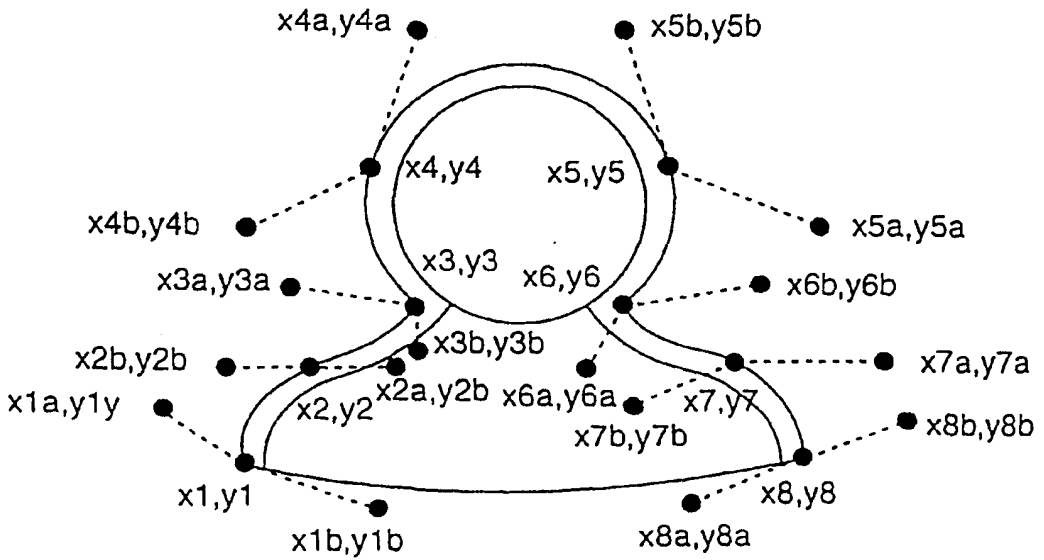


2c



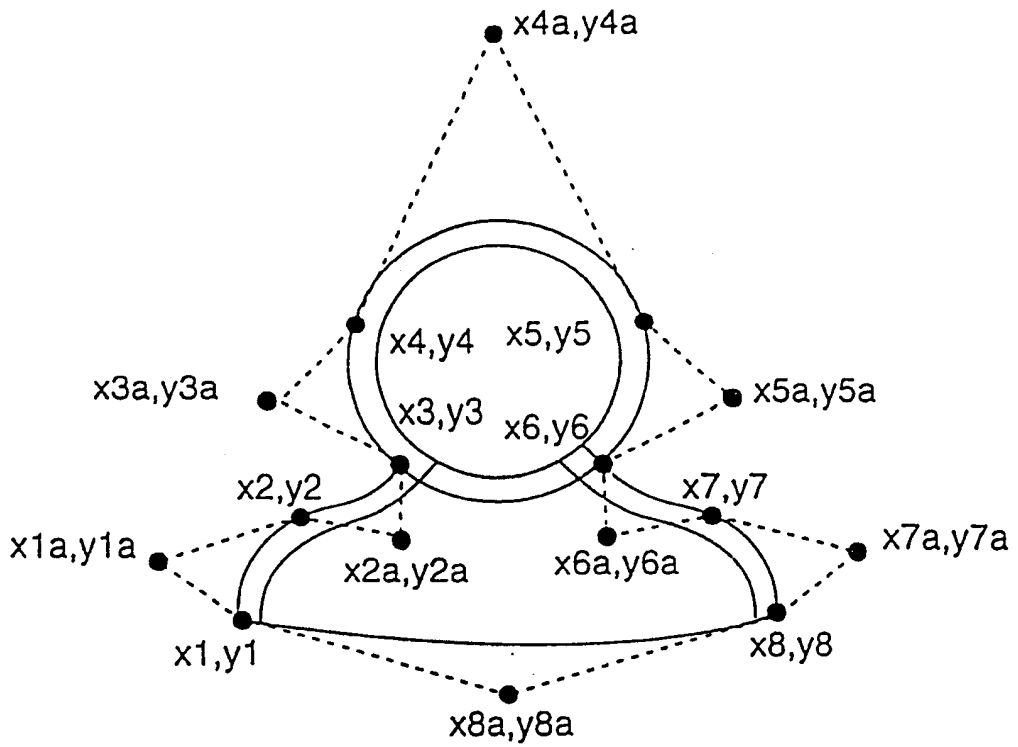
物体编号1, 深度20

图 3



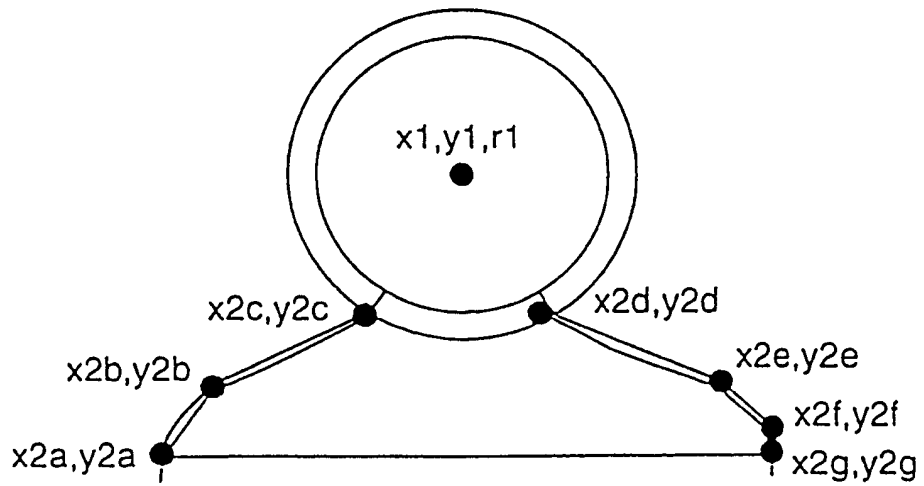
物体编号1, 深度20

图 4



物体编号1, 深度20

图 5



物体编号1, 深度20

图 6

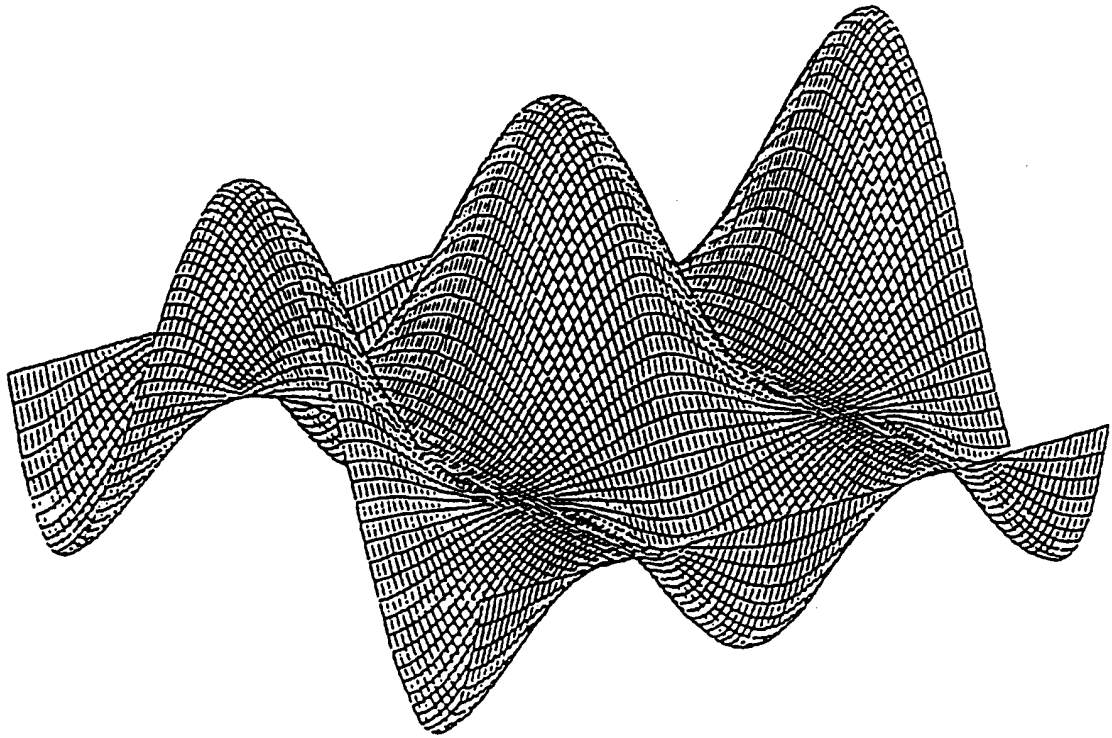


图 7

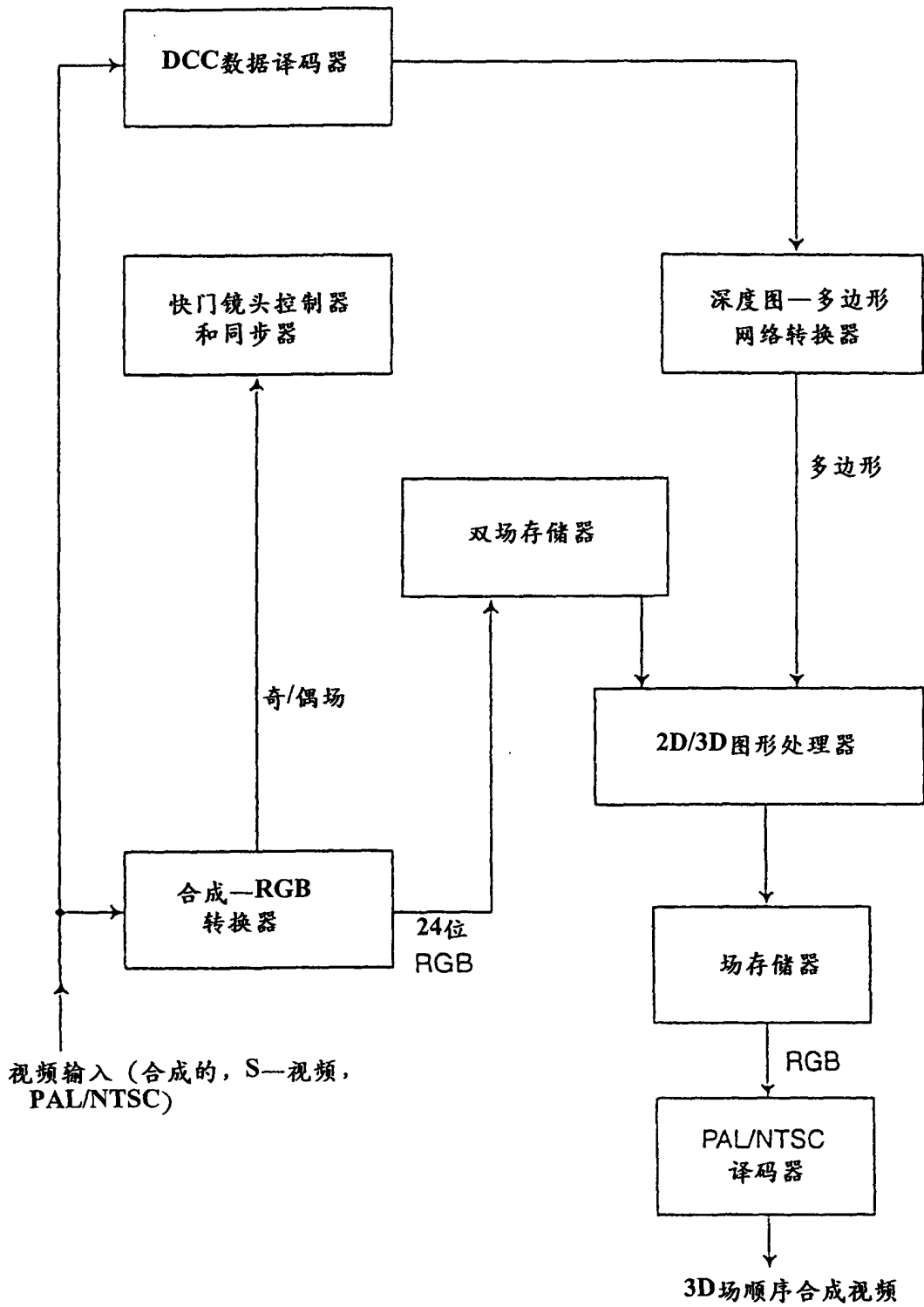


图 8

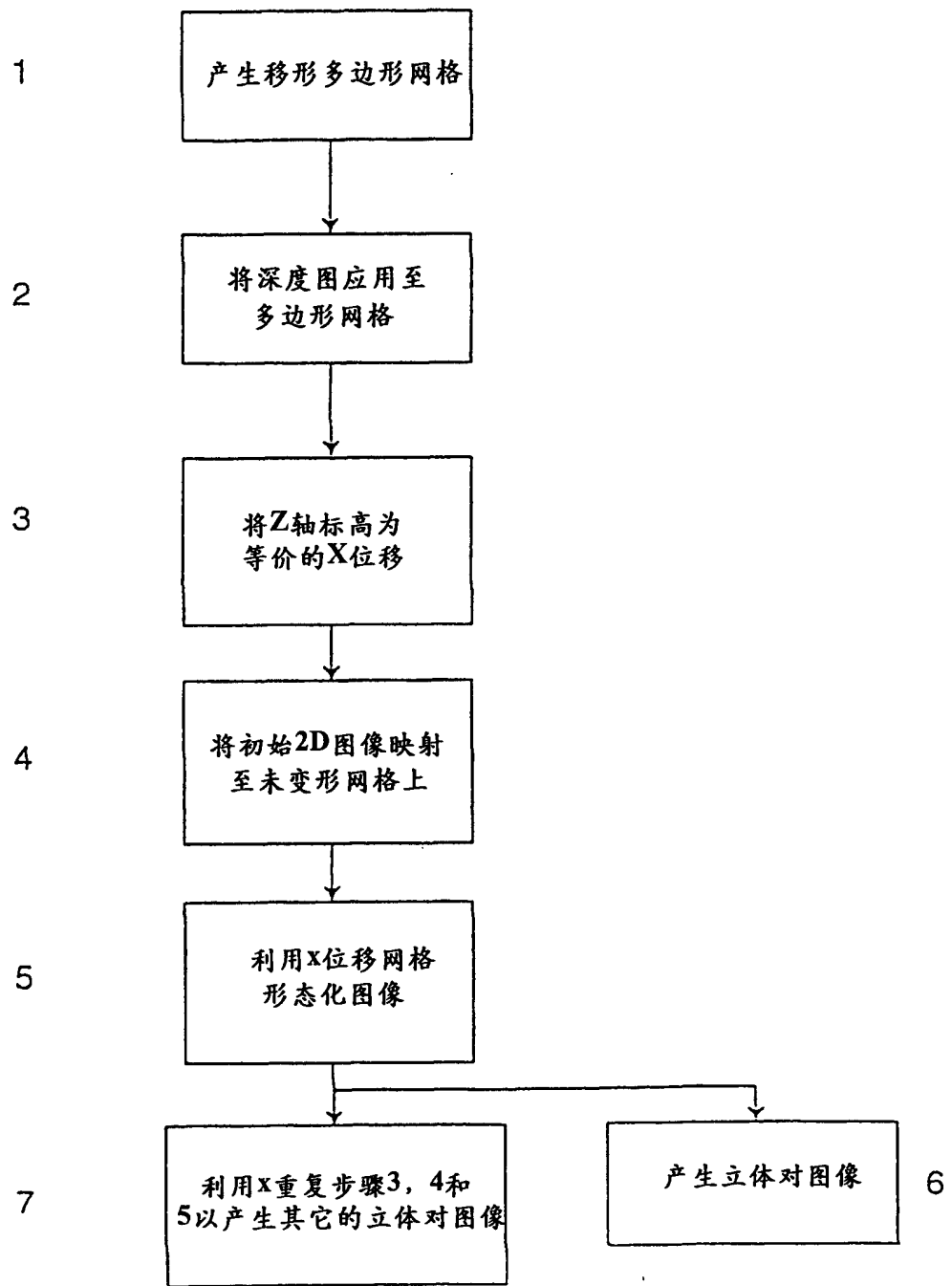


图 9



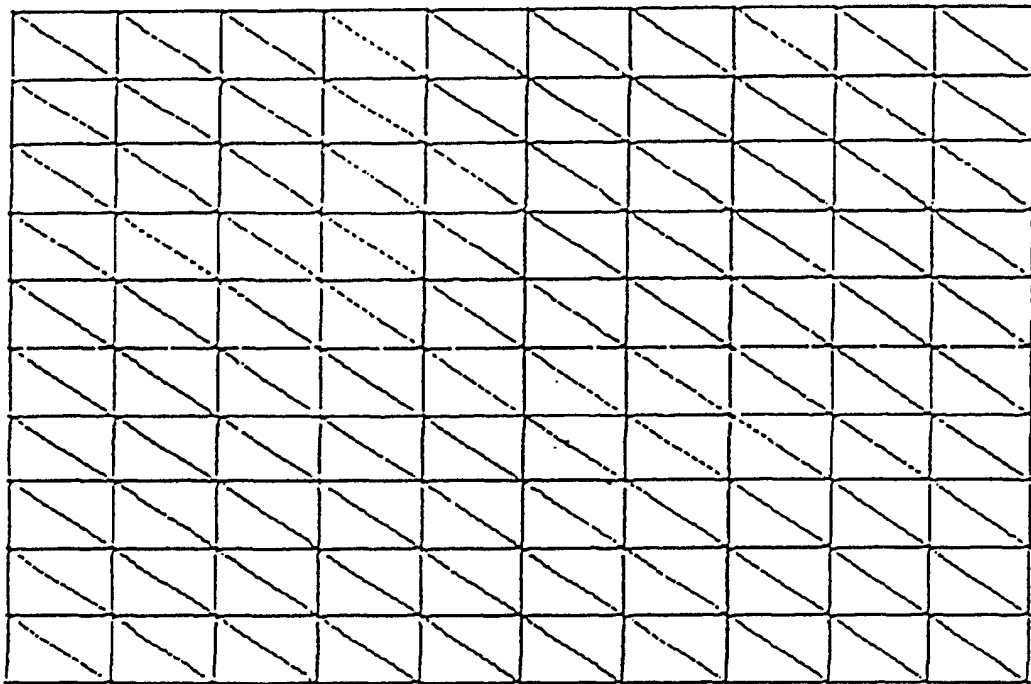


图 10

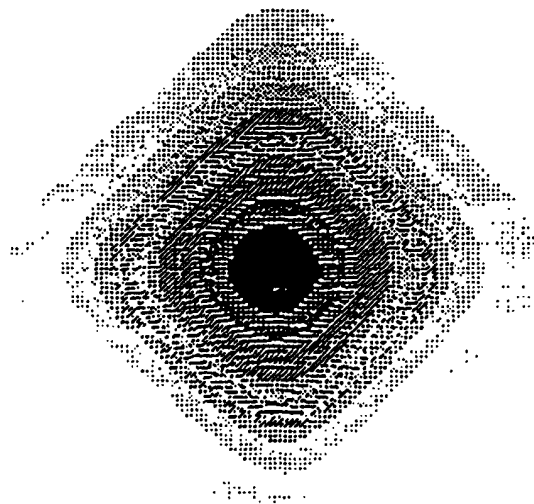


图 11

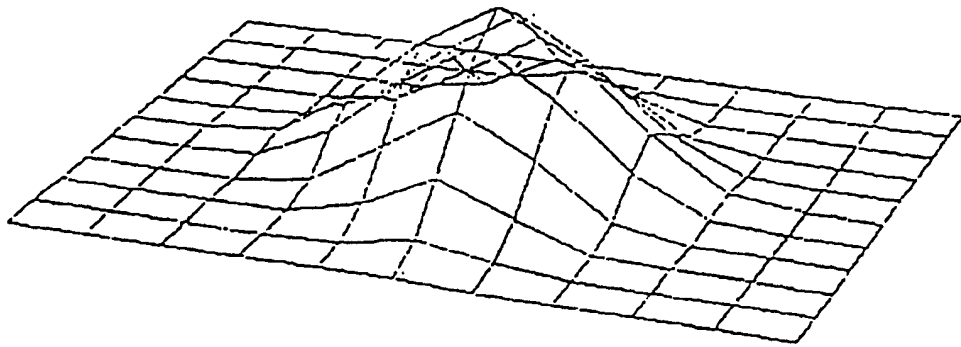


图 12

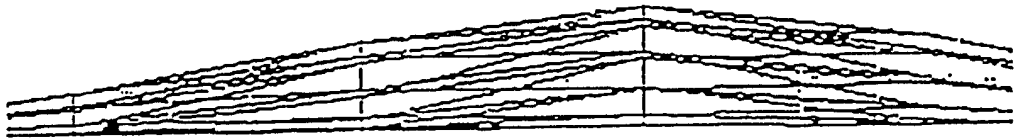


图 13

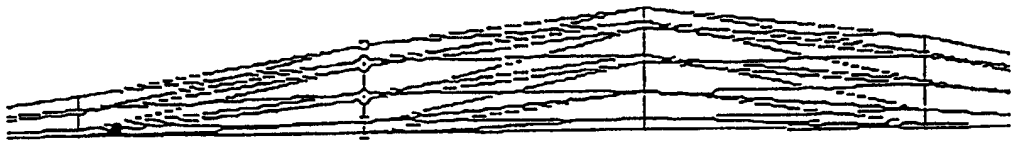


图 14

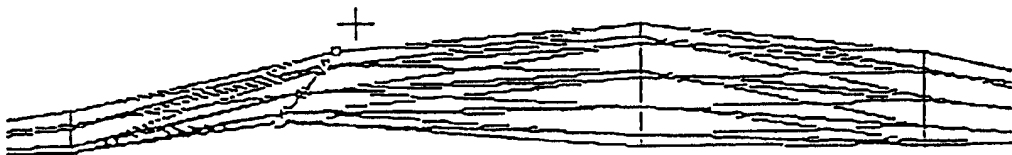


图 15

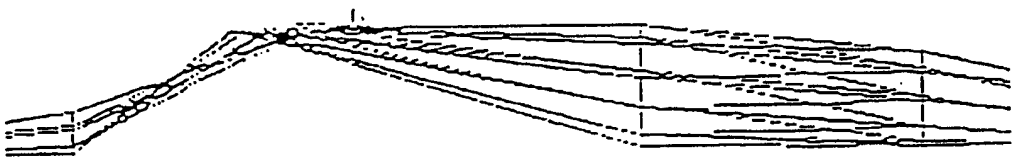


图 16

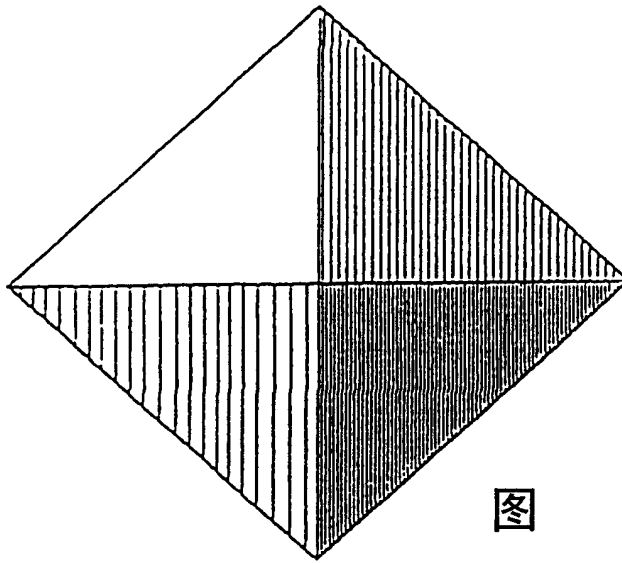


图 17

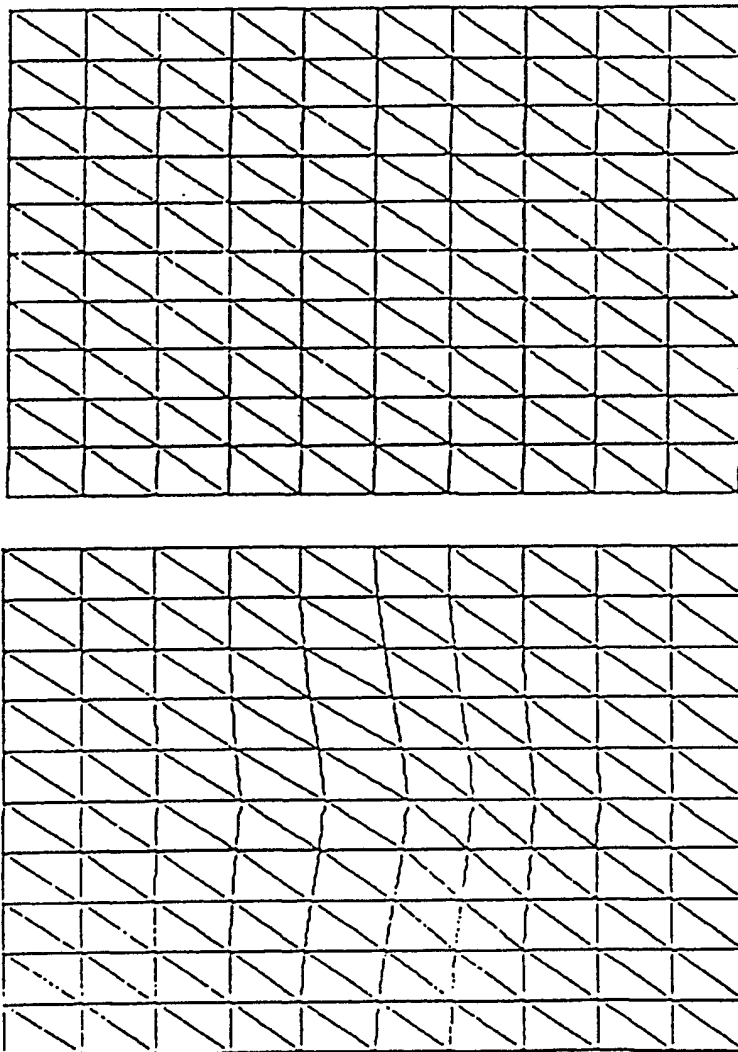


图 18

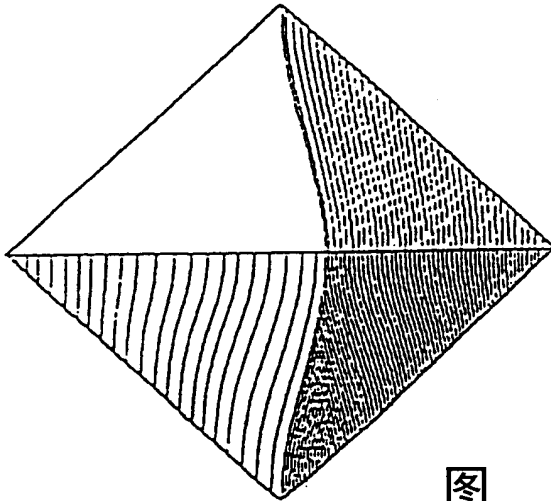


图 19

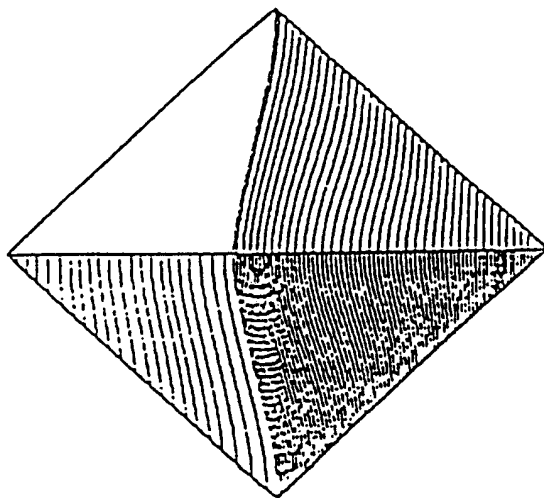
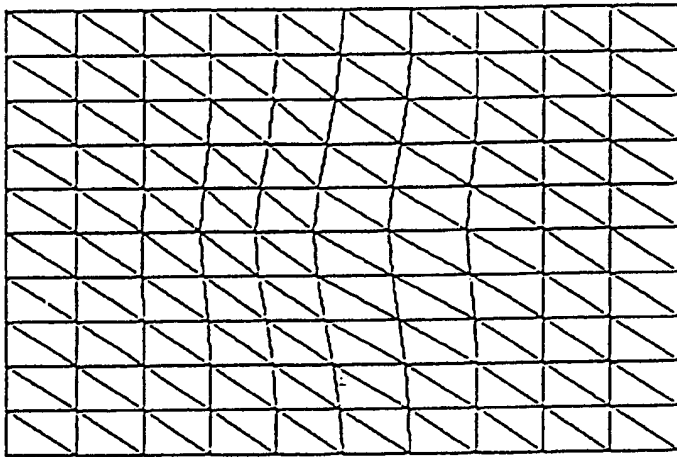
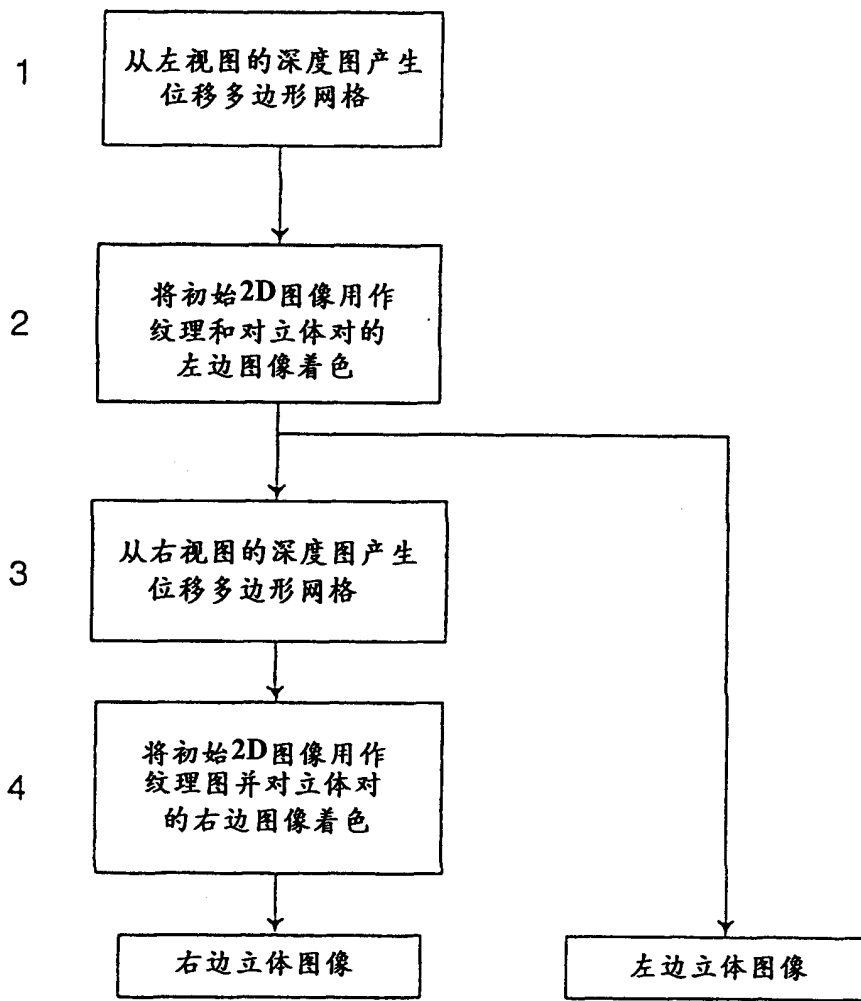


图 20



图

21