



(12) 发明专利

(10) 授权公告号 CN 101783027 B

(45) 授权公告日 2012. 02. 29

(21) 申请号 201010114994. 1

(22) 申请日 2010. 02. 26

(73) 专利权人 浙江大学

地址 310027 浙江省杭州市西湖区浙大路
38 号

(72) 发明人 刁常宇 朱文峤 鲁东明

(74) 专利代理机构 杭州天勤知识产权代理有限公司 33224

代理人 胡红娟

(51) Int. Cl.

G06T 17/00 (2006. 01)

G06T 7/00 (2006. 01)

H04W 84/18 (2009. 01)

(56) 对比文件

CN 101630406 A, 2010. 01. 20, 全文.

US 4772944 A, 1988. 09. 20, 全文.

CN 101443817 A, 2009. 05. 27, 全文.

CN 101650178 A, 2010. 02. 17, 全文.

WO 99/03068 A1, 1999. 01. 21, 全文.

张伟. 多视图三维重构算法与软件实现. 《中国优秀博硕士学位论文全文数据库(硕士)信息科技辑》. 2006, (第 12 期), I138-1420.

孔晓东等. 基于极约束和边缘点检测的图像密集匹配. 《计算机工程》. 2004, 第 30 卷(第 20 期), 40-41、179.

Pollefeys M, et al.. A Simple and Efficient Rectification Method for General Motion. 《Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on ICCV'99》. 1999, 第 1 卷 496-501.

审查员 胡妮

权利要求书 1 页 说明书 5 页 附图 3 页

(54) 发明名称

基于多台图像传感器的动态场景三维记录方法

(57) 摘要

本发明公开了基于多台图像传感器的动态场景三维记录方法,包括以下步骤:在场景周围放置图像传感器再用 L. Wang 的方法对所有的图像传感器进行标定,算出图像传感器的内参和外参;之后利用图像传感器进行图像采集;利用图像传感器的内参和外参计算需要密集匹配的图像 A 和图像 B 之间的基本矩阵 F;用 Marc Pollefeys 的方法,根据 F 对图像 A, B 进行矫正,使 A, B 的对极线平行;进行密集匹配计算;用三角形法进行结构计算;最后进行点云优化计算;本发明提供一种对基于多台图像传感器的动态场景三维记录的方法,可以方便的对动态场景的三维信息进行记录。

CN 101783027 B

1. 一种基于多台图像传感器的动态场景三维记录方法,其特征在于:包括以下步骤:
 - (1) 在动态场景周围放置图像传感器,至少有两个图像传感器的视域覆盖所述的动态场景中需要进行三维记录的区域;
 - (2) 对所有的图像传感器进行标定,得到图像传感器的内参和外参;
 - (3) 利用图像传感器进行图像采集;
 - (4) 利用图像传感器的内参和外参计算需要密集匹配的图像之间的基本矩阵 F ;
 - (5) 根据步骤 (4) 得到的基本矩阵 F 对需要密集匹配的图像进行矫正,使需要密集匹配的图像中的对极线平行;
 - (6) 对需要密集匹配的图像中的所有点进行密集匹配计算,得到对应点的点对;需要密集匹配的图像为两幅,分别标记为图像 A, 图像 B, 在进行密集匹配计算时,步骤如下:
 - a) 在图像 A 中,以某一需要进行密集匹配计算的点 X 为圆心,以 12 个像素为半径做圆,在圆周上选取均匀分布的 8 个点;
 - b) 分别以所述圆的圆心及圆周上的 8 个点为中心,得到 9 个矩形区域,所述的矩形区域的长均为 21 个像素,宽均为 15 个像素;
 - c) 对每个矩形区域内的所有点做如下计算得到 9 个加和值:
以点的灰度值除以该点到矩形中心的距离,得到灰度值与距离的比值,然后将这些比值求和,得到加和值;
 - d) 将步骤 c) 求得的 9 个加和值组成点 X 的描述向量;
 - e) 根据点 X 得到图像 B 中与其对应的对极线,对该对极线上的所有的点,依照步骤 a) 到步骤 d) 的方法计算其描述向量,找出与点 X 的描述向量距离最小的描述向量所对应的点 Y ,点 Y 就是点 X 在图像 B 中的对应点;
 - f) 遍历图像 A 上的所有点,完成点的密集匹配计算;
 - (7) 利用步骤 (2) 得到的内参和外参,对步骤 (6) 得到的点对用三角形法进行结构计算,得到动态场景中需要进行记录的区域的空间坐标;
 - (8) 对步骤 (7) 得到的点的空间坐标进行优化计算,得到所述点的最终空间坐标。
2. 如权利要求 1 所述的基于多台图像传感器的动态场景三维记录方法,其特征在于:步骤 (3) 中进行图像采集时,所有的图像传感器应同步工作。

基于多台图像传感器的动态场景三维记录方法

技术领域

[0001] 本发明涉及到无线传感网络, 计算机视觉领域, 特别涉及到计算机视觉中的三维重建技术。

背景技术

[0002] 基于图像序列的场景三维重建是计算机视觉中的重要领域, 它包括: 相机标定: 相机标定的过程可以理解为求解相机矩阵 P 的过程, 矩阵 $P = K[R \ t]$, 其中矩阵 K 表示相机的内参, 它包括 x, y 方向的焦距, 主点坐标和切向畸变。 R, t 表示相机的外部参数, 它描述世界坐标系在相机坐标系中的偏转和平移。通过矩阵 P , 世界坐标系中的点 X 和图像坐标系中的点 x 可以通过下面的式子进行关联:

$$[0003] \quad s_x = PX$$

[0004] 其中 s 是一个常数, 对上面的式子进行展开, 我们有:

$$[0005] \quad s[u \ v \ 1]^T = P[x \ y \ z \ 1]^T$$

[0006] 其中 u, v 是点在图像中的坐标, x, y, z 是点在世界坐标系中的坐标。对极几何: 空间中的一个点, 在两幅图像上的像 x, x' 有如下的性质,

$$[0007] \quad x'^T F x = 0$$

[0008] 其中 F 是一个 3×3 的并且秩为 2 的矩阵。另外, x 在另外一副图像上的点在它的对极线上, 这条线可以用方程 $l' = Fx$ 表示。对应两张图像的 F 和它们的矩阵 P, P' 有如下关系:

$$[0009] \quad F = [e' \]_x P' P^+$$

[0010] 其中 P^+ 是 P 的伪逆矩阵, $e' = P' C$, 而 C 通过求解 $PC = 0$ 解出。

[0011] 结构计算: 结构计算中最常用的是线性三角形法, 给定一个点在两个图像中的像 x, x' 我们有 $x = PX, x' = P' X$, 通过 $x \times (PX) = 0$ 消去齐次因子, 两个式子展开, 我们组成形如 $AX = 0$ 的方程:

[0012] 其中 A 可以表示成:

$$[0013] \quad A = [A1 \ A2 \ A3 \ A4]^T$$

$$[0014] \quad A1 = [xP^{3T} - P^{1T}]$$

$$[0015] \quad A2 = [yP^{3T} - P^{2T}]$$

$$[0016] \quad A3 = [x' \ P' \]^{3T} - P' \]^{1T}]$$

$$[0017] \quad A4 = [y' \ P' \]^{3T} - P' \]^{2T}]$$

[0018] P^{iT} 表示矩阵 P 的第 i 行, 通过求解上述方程, 我们可以求出点的三维齐次坐标。因为实际的测量数据会带有误差, 因此 $x = PX$ 并不能完全满足, 可以通过用 Levenberg-Marquardt 算法来最小化重投影误差, 来优化计算出的点云数据。

发明内容

[0019] 本发明提供一种基于多台图像传感器的动态场景三维记录方法, 一种基于多台图

像传感器的动态场景三维记录方法,它包括以下的步骤:

[0020] (1) 在动态场景周围放置图像传感器;

[0021] 所述的动态场景是指场景中物体的三维结构或者说几何形状会随着时间的改变而发生变化。我们所指的记录动态场景的三维信息,是指记录任一时间点的场景三维信息,这些任一时间点的场景三维信息集合就组成了一个时间段内的场景的三维信息及其变化的过程。

[0022] 在动态场景的周围放置图像传感器,使得需要进行三维记录的区域至少在两个图像传感器的视域内,并且使得对着同一视域的图像传感器中的任意两个图像传感器的主轴的夹角在 10 度到 30 度之间。通过图像传感器连接的同步模块,使多个图像传感器可以同步工作。图像传感器间的同步通过图像传感器链接的专有同步模块进行设置的。同步的工作方式可以通过各个图像传感器间发送时间同步包来进行同步,或是通过激光脉冲方法进行同步,激光脉冲方法进行同步可以达到较高的时间同步精度。同步工作的实现,可以根据当时的环境选择有线网络,或无线网络接入,也可以是通过接入大容量的存储设备来实现。

[0023] 有时候我们并不需要对整个动态场景进行记录,而仅仅是需要记录整个动态场景中的局部区域,那么至少保证有两个图像传感器的视域可以覆盖需要进行三维记录的局部区域,对该局部区域进行图像记录的至少两个图像传感器中,任意两个图像传感器的主轴的夹角在 10 度到 30 度之间。

[0024] 需要说明的是,有可能还有其他图像传感器的视域也覆盖了需要进行三维记录的局部区域,但由于不是专为该局部区域设置的,所以并不严格要求这些非专设的图像传感器与专设的图像传感器的主轴的夹角。

[0025] (2) 用 L. Wang 的方法对所有的图像传感器进行标定,求出图像传感器的内参和外参;

[0026] 通过使用 L. Wang 的方法,用一个带至少 3 个标志点的直杆对多个图像传感器(相机)进行标定。具体可以参照 L. Wang 的论文(参考文献 1)“L. Wang F. C. Wu and Z. Y. HU Multi-Camera Calibration with One-Dimensional Object under General Motions, ICCV2007”。

[0027] 两个图像传感器同步地拍摄 10 到 15 张图片,这些图片中要包含上面所述的带至少有 3 个标志点的直杆。将两个图像传感器中的图像的中的直杆的标记点的图像坐标提取出来,把这些标志点的图像坐标以及标志点间的实际距离带入到 L. Wang 的相机标定算法中,计算出所有图像传感器的相机矩阵。如果标定成功,进行下一步工作,否则重复上述过程重新对两个图像传感器进行标定。

[0028] L. Wang 的方法在大范围内散乱分布的相机阵列的标定时可以得到很好的效果,而其他的方法只在范围较小或单个相机的标定中比较试用。

[0029] (3) 利用图像传感器进行图像采集;

[0030] 在进行图像采集的过程中,所有的图像传感器应该是同步工作。因为我们记录的场景可以是动态的,因此需要记录一个时间段内的场景的几何结构及其变化的过程。通过图像传感器同步的方式,每间隔一定时间通过图像传感器对场景进行一次同步拍摄,通过拍摄的图片,计算出这一个时间点的场景的三维信息。可以根据场景机构变化的速度来调整同步拍摄的时间间隔。通过这些时间点的场景的三维信息,我们可以得出一个时间段内

的场景的三维信息及其变化过程。

[0031] (4) 利用图像传感器的内参和外参计算需要密集匹配的图像 A 和图像 B 之间的基本矩阵 F；

[0032] 通过前面的标定过程求出的图像 A 对应的相机矩阵 P 和图像 B 所对应的相机矩阵 P'，我们可以求出矩阵 F，具体公式如下：

$$[0033] \quad F = [e']_x P' P^+$$

[0034] 其中 P⁺ 是 P 的伪逆矩阵， $e' = P' C$ ，C 通过求解 $PC = 0$ 解出。

[0035] (5) 用 Marc Pollefeys 的方法，根据 F 对图像 A, B 进行矫正，使 A, B 的对极线平行；

[0036] 有了上一步求出的 F 我们对两张图像进行矫正，矫正后的图像满足对应的对极线是平行的。具体可以参考 Marc Pollefeys 的论文（参考文献 2）“Marc Pollefeys and Reinhard Koch and Luc Van Gool A simple and efficient rectification method for general motion, 1999”。

[0037] (6) 对图像 A 和图像 B 进行密集匹配计算，得到对应点的点对；

[0038] 首先对图像 A 中的每个像素点 p1，按先下列方式设置描述向量，以 p1 为圆心，12 个像素为半径的圆上，选取均匀分布的 8 个点，c1, c2, c3, c4, c5, c6, c7, c8。以 p1, c1, c2, c3, c4, c5, c6, c7, c8 这 9 个点为中心，选取 9 个长为 21 个像素宽为 15 个像素的矩形区域。对每一个矩形区域内的所有点做如下计算，以点的灰度值除以该点到矩形中心的距离，得到灰度值与距离的比值，然后将这些比值求和。将上面求得的 9 个加和值组成向量，同时计算 p1 在图像 B 上的对极线上的所有点的描述向量，并求与 p1 距离最小的描述向量所对应的点。遍历图像 A 上的所有点，找到它在图像 B 上的对应点，如果不存在对应点，则将 A 上的点忽略。

[0039] (7) 用三角形法对所求出的点对进行结构计算，得到动态场景中需要进行记录的区域的空间坐标。

[0040] 对每一组对应点 x, x' 我们有 $x = PX, x' = P'X$ 通过 $x \times (PX) = 0$ 消去齐次因子，两个式子展开，我们组成形如 $AX = 0$ 的方程：

[0041] 其中 A 可以表示成：

$$[0042] \quad A = [A1 \ A2 \ A3 \ A4]^T$$

$$[0043] \quad A1 = [xP^{3T} - P^{1T}]$$

$$[0044] \quad A2 = [yP^{3T} - P^{2T}]$$

$$[0045] \quad A3 = [x' \ P' \ ^{3T} - P' \ ^{1T}]$$

$$[0046] \quad A4 = [y' \ P' \ ^{3T} - P' \ ^{2T}]$$

[0047] PiT 表示矩阵 P 的第 i 行，通过求解上述方程，我们可以求出点的三维齐次坐标。对所有的对应点进行上面的计算，最后求出三维场景的需要进行记录的区域的空间坐标数据。

[0048] (8) 对所求得的空间坐标数据进行优化计算；

[0049] 因为实际的测量数据带有噪音，因此我们通过 Levenberg-Marquardt 算法来最小化重投影误差，来优化计算出的点云数据。最终得到需要记录的场景的三维信息。

[0050] 本发明基于多台图像传感器的动态场景三维记录方法，具有以下优点：

[0051] 可以对动态的场景进行建模。通过引入同步工作的图像传感器,我们可以采集同一时刻的场景的各个视角的图像。通过这些图像,采用计算机视觉中的三维重构方法,我们可以对这一个时间点的场景进行建模。

[0052] 密集匹配准确度提高。通过引入我们设计的图像点的描述向量,图像密集匹配的精度可以得到提高。特别是在对低纹理的场景的重构中,可以具有很好的准确度。

附图说明

[0053] 图 1 是本发明基于多台图像传感器的动态场景三维记录方法流程图;

[0054] 图 2 是图像传感器的逻辑结构图;

[0055] 图 3 是构造像素点 p_1 的描述向量的示意图;

[0056] 图 4 是未矫正之前的对极线示意图;

[0057] 图 5 是矫正之后的对极线示意图。

具体实施方式:

[0058] 本发明基于多台图像传感器的动态场景三维记录方法实施过程如下:

[0059] (1) 在场景周围放置图像传感器;

[0060] 在场景的周围放置图像传感器,使得需要进行三维记录的点的区域在两个图像传感器的视域内,这两个图像传感器的主轴的夹角为 20 度。

[0061] (2) 用 L. Wang 的方法对所有的图像传感器进行标定,求出图像传感器的内参和外参;

[0062] 通过使用 L. Wang 的方法,用一个带 3 个标志点的直杆对两个图像传感器(相机)进行标定。“L. Wang F. C. Wu and Z. Y. HU Multi-Camera Calibration with One-Dimensional Object under General Motions, ICCV2007”。L. Wang 的方法通过平行线段的长度比在射影变换下的不变性以及标杆的长度对相机的内参进行约束,同时利用无穷远单应对相机矩阵进行约束,最后求出相机的矩阵。为了减少噪声对结果的影响,该方法还通过使用捆集调整对结果进行优化。

[0063] 两个图像传感器(相机)采用激光脉冲方法进行同步。两个图像传感器同步地拍摄 15 张图片,这些图片中要包含上面所述的带有 3 个标志点的直杆。将两个图像传感器中的图像的中的直杆的标记点的图像坐标提取出来,把这些标志点的图像坐标以及标志点间的实际距离带入到 L. Wang 的相机标定算法中,计算出所有图像传感器的相机矩阵。如果标定成功,进行下一步工作,否则重复上述过程重新对两个图像传感器进行标定。

[0064] (3) 利用同步工作的图像传感器进行图像采集;

[0065] 每间隔 1 秒钟通过图像传感器对场景进行一次同步拍摄。

[0066] (4) 利用图像传感器的内参和外参计算需要密集匹配的图像 A 和图像 B 之间的基本矩阵 F ;

[0067] 通过前面的标定过程求出的 P, P' 我们可以求出需要密集匹配的图像 A 和图像 B 之间的基本矩阵 F 。具体公式如下:

$$[0068] F = [e']_x P' P^+$$

[0069] 其中 P^+ 是 P 的伪逆矩阵, $e' = P' C$, C 通过求解 $PC = 0$ 解出。

[0070] (5) 用 Marc Pollefeys 的方法, 根据 F 对图像 A, B 进行矫正, 使 A, B 的对极线平行;

[0071] 有了上一步求出的 F 我们对两张图像进行矫正, 矫正后的图像满足对应的对极线是平行的。“Marc Pollefeys and Reinhard Koch and Luc Van Gool A simple and efficient rectification method for general motion, 1999”。Marc Pollefeys 的方法在对极点附近对极图像进行极坐标参数化, 并通过图像间的基本矩阵将两幅图像矫正成对极线平行。通过 Marc Pollefeys 的方法我们可以简化后面所述的密集匹配计算过程, 提高了程序的运算效率。

[0072] (6) 需要密集匹配的图像中的所有点进行密集匹配计算, 得到对应点的点对;

[0073] 首先对图像 A 中的每个像素点 p_1 , 按先下列方式设置描述向量, 以 p_1 为圆心, 12 个像素为半径的圆上, 选取均匀分布的 8 个点, $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ 。以 $p_1, c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ 这 9 个点为中心, 选取 9 个长为 21 个像素宽为 15 个像素的矩形区域。对每个区域做如下计算, 对区域内的所有点的像素值除以其到中心距离的半径, 然后求和。将上面求得的 9 个值组成向量, 同时计算图像 B 中与 p_1 对应的对极线上的点的描述向量。所有的向量中与 p_1 的描述向量距离最小的点, 置为 p_1 在图像 B 中的对应点。遍历图像 A 上的所有点, 找到它在图像 B 上的对应点, 如果不存在对应点, 则将 A 上的点忽略。在具体的实施过程中, 可以根据场景的复杂度以及图像传感器采集图像的大小来设置半径的大小和矩形区域的长度和宽度。

[0074] (7) 用三角形法进行结构计算;

[0075] 对每一组对应点 x, x' 我们有 $x = PX, x' = P'X$ 通过 $x \times (PX) = 0$ 消去齐次因子, 两个式子展开, 我们组成形如 $AX = 0$ 的方程:

[0076] 其中 A 可以表示成:

$$[0077] \quad A = [A_1 \ A_2 \ A_3 \ A_4]^T$$

$$[0078] \quad A_1 = [xP^{3T} - P^{1T}]$$

$$[0079] \quad A_2 = [yP^{3T} - P^{2T}]$$

$$[0080] \quad A_3 = [x' \ P' \ ^{3T} - P' \ ^{1T}]$$

$$[0081] \quad A_4 = [y' \ P' \ ^{3T} - P' \ ^{2T}]$$

[0082] 通过求解上述方程, 我们可以求出点的三维齐次坐标。对所有的对应点进行上面的计算, 最后求出三维场景的点云数据。

[0083] (8) 进行点云优化计算;

[0084] 因为实际的测量数据带有噪音, 因此我们通过 Levenberg-Marquardt 算法来最小化重投影误差, 来优化计算出的点云数据, 即前面算出的所有对应点的三维坐标的集合, 最终得到需要记录的场景的三维信息。

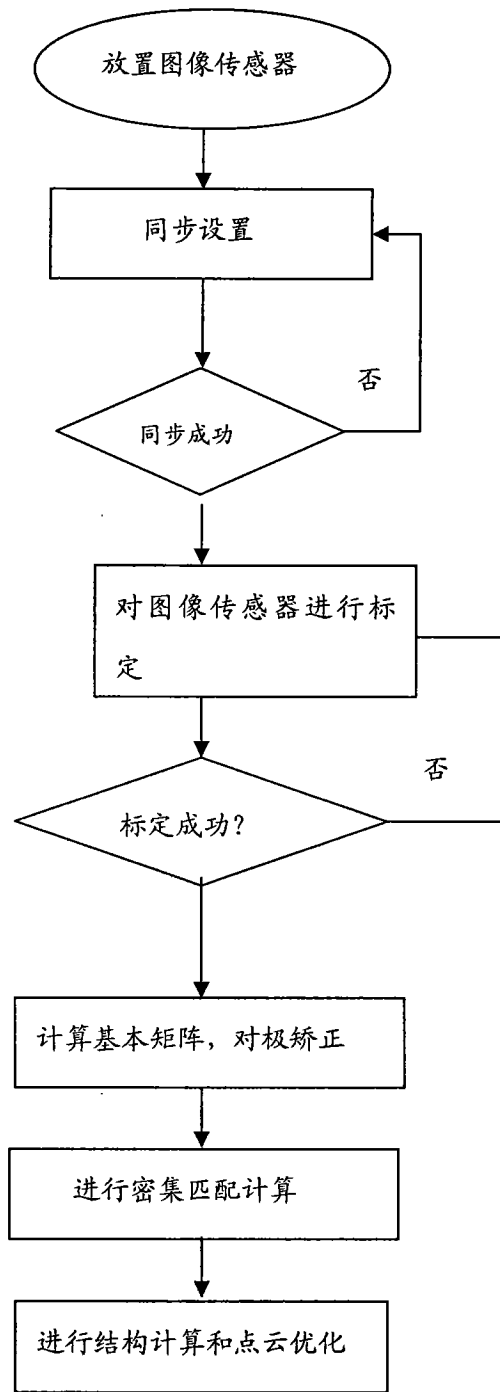


图 1

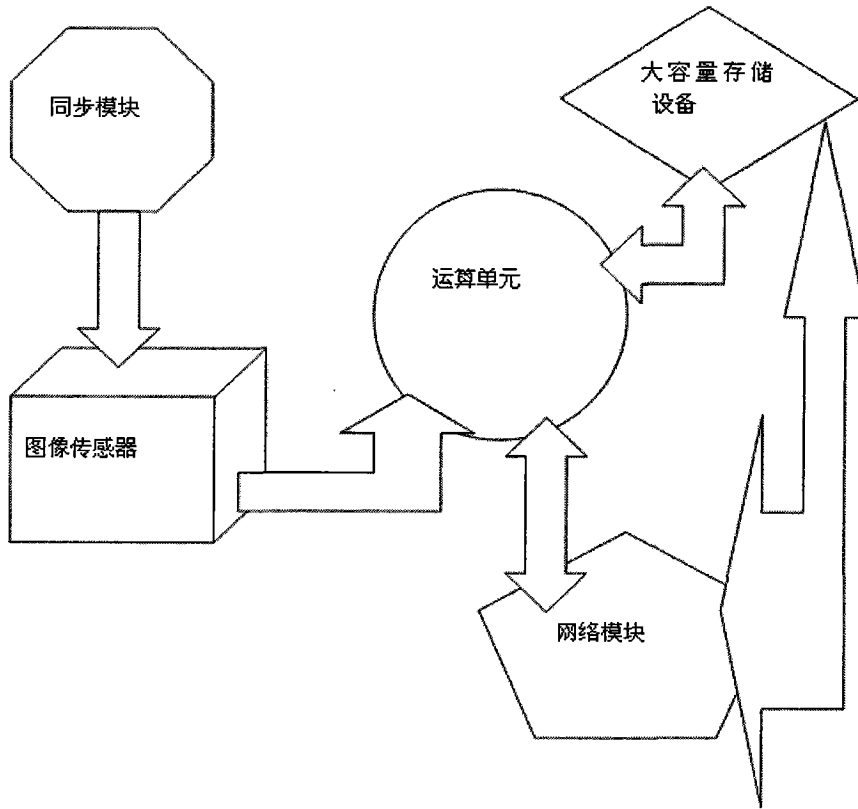


图 2

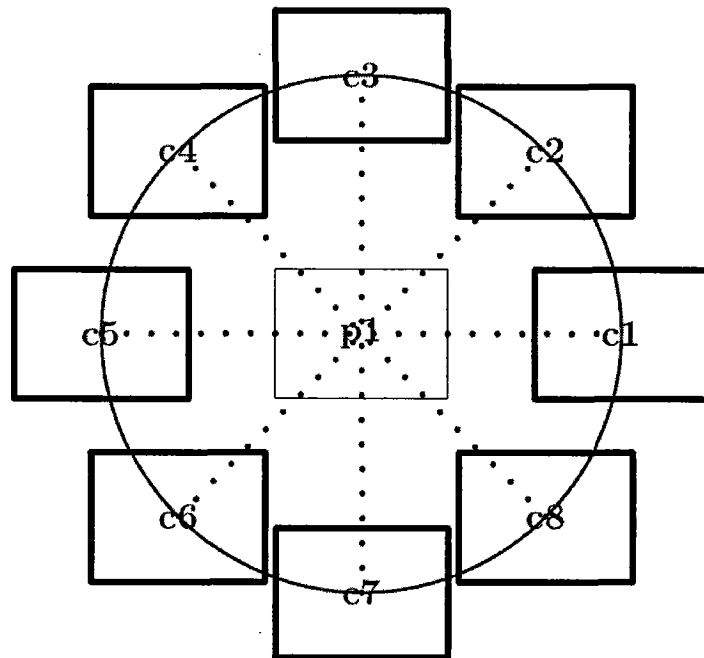


图 3

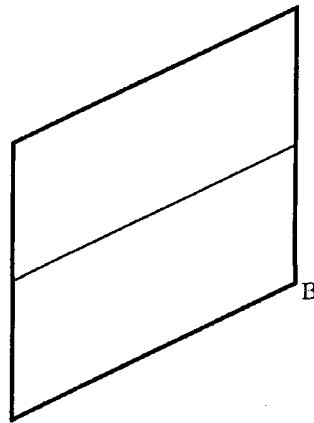
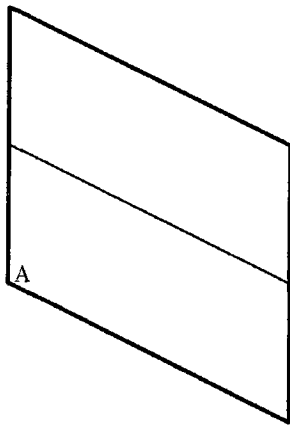


图 4

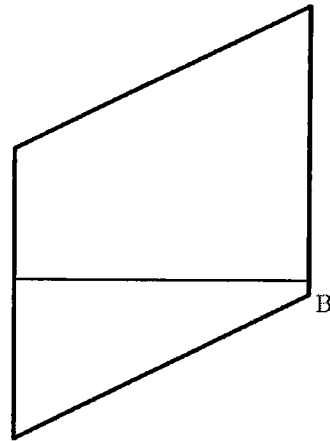
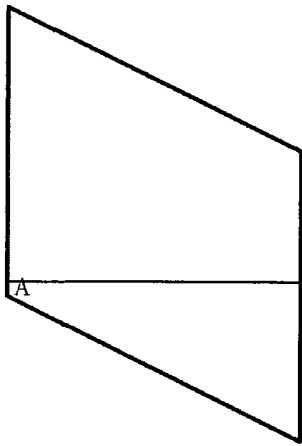


图 5