



(12) 发明专利申请

(10) 申请公布号 CN 102473307 A

(43) 申请公布日 2012. 05. 23

(21) 申请号 201080028724. 7

(51) Int. Cl.

(22) 申请日 2010. 03. 15

G06T 7/20 (2006. 01)

(85) PCT申请进入国家阶段日

2011. 12. 27

(86) PCT申请的申请数据

PCT/EP2010/001612 2010. 03. 15

(87) PCT申请的公布数据

W02011/113444 EN 2011. 09. 22

(71) 申请人 松下电器产业株式会社

地址 日本大阪府

(72) 发明人 岩崎正宏 登一生 甲本亚矢子

F・加拉索 R・奇波拉

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 张伟 王英

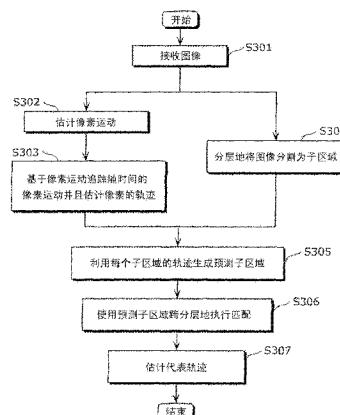
权利要求书 2 页 说明书 23 页 附图 25 页

(54) 发明名称

用于轨迹估计的方法和装置以及用于分割的方法

(57) 摘要

一种轨迹估计装置，包括：图像接收单元(101)，其接收按时间排序的并且包括在视频中的多幅图像；分层子区域生成单元(102)，其通过对由所述图像接收单元(101)接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域，使得在属于彼此不同的分层级的子区域当中，空间上较大的子区域包括空间上较小的子区域；以及代表轨迹估计单元(104)，其通过跨与特定图像不同的图像中的分层级，搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。



1. 一种用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的轨迹估计方法，所述轨迹估计方法包括：

接收包括在所述视频中的多幅图像，所述多幅图像是按时间排序的；

通过对在所述接收中接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域，使得空间上较大的子区域包括空间上较小的子区域，所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级；以及

通过跨与特定图像不同的图像中的分层级，搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

2. 根据权利要求 1 所述的轨迹估计方法，还包括：

通过针对由包括在所述多幅图像中的每一幅图像中的一个或多个像素构成的每个块分析所述多幅图像之间的运动而在所述视频中生成所述块的轨迹，

其中在所述估计中，通过使包括在所述特定图像中的所述子区域与包括在所述子区域中的所述块的所述轨迹一起移动来生成预测子区域，所述预测子区域是在与所述特定图像不同的所述图像中预测的子区域，并且通过跨与所述特定图像不同的所述图像中的所述分层级，搜索与所述预测子区域最相似的所述子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为所述代表轨迹。

3. 根据权利要求 2 所述的轨迹估计方法，

其中在所述估计中，使包括在所述生成中生成的每个子区域中的轨迹对于所述每个子区域是平滑的，并且基于包括在所述每个子区域中的平滑的轨迹来估计所述代表轨迹。

4. 根据权利要求 1 至 3 之一所述的轨迹估计方法，

其中在所述估计中，通过使用相似性测度来确定子区域之间的相似性，从而搜索所述子区域，其中所述相似性测度包括所述子区域的像素值和形状信息中的至少之一。

5. 根据权利要求 1 所述的轨迹估计方法，

其中在所述估计中，通过 (i) 采用曲线图、和 (ii) 将动态编程应用于所述曲线图而生成用于所述曲线图的最优路径来估计所述代表轨迹，在所述曲线图中：包括在所述特定图像中的所述子区域和包括在与所述特定图像不同的所述图像中的所述子区域是节点；使用边来对表示包括在所述特定图像中的所述子区域的一个节点和表示包括在与所述特定图像不同的所述图像中的所述子区域的另一节点进行连接；并且这样的值是边权重，即对于由所述边连接的子区域之间的较大相似性，该值是较小的。

6. 根据权利要求 1 至 5 之一所述的轨迹估计方法，

其中在所述生成中，基于包括轮廓信息的特征，在所述分层级上生成所述子区域。

7. 一种用于对每个移动对象的视频执行分割的分割方法，所述分割方法包括：

包括在根据权利要求 1 至 6 之一所述的轨迹估计方法中的处理；以及

通过将代表轨迹之间具有相似性的子区域进行结合来对所述视频执行分割。

8. 根据权利要求 7 所述的分割方法，

其中，在所述执行中，通过连接小于阈值的距离来将所述代表轨迹之间的距离转换为测地距离、检测在获得的测地距离中的不连续点并且将彼此分开的距离比在所检测的不连续点处的测地距离小的轨迹分到一个类别，从而执行分割。

9. 根据权利要求 7 所述的分割方法，

其中，在所述执行中，基于对应于要被结合的子区域的代表轨迹和邻近所述要被结合的子区域的至少一个子区域的代表轨迹，确定是否结合所述要被结合的子区域。

10. 根据权利要求 9 所述的分割方法，

其中，在所述执行中，基于在根据曲线图计算的最小生成树中表示的连接关系来执行分割，在该曲线图中，将所述代表轨迹假设为节点并且使用边来将所述代表轨迹彼此连接；并且由所述边连接的所述代表轨迹之间的距离被假设为边权重。

11. 根据权利要求 10 所述的分割方法，

其中，在所述执行中，基于用于将具有以下关系的一对代表轨迹与邻近该对代表轨迹的至少一个代表轨迹连接的边权重，通过在最小生成树中确定是否结合该对代表轨迹来执行分割：所述代表轨迹是使用一条边直接彼此连接的。

12. 一种用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的轨迹估计装置，所述轨迹估计装置包括：

图像接收单元，被配置为接收包括在所述视频中的多幅图像，所述多幅图像是按时间排序的；

分层子区域生成单元，被配置为通过对由所述图像接收单元接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域，使得空间上较大的子区域包括空间上较小的子区域，所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级；以及

代表轨迹估计单元，被配置为通过跨与特定图像不同的图像中的分层级，搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

13. 一种用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的程序，并且所述程序使计算机执行：

接收包括在所述视频中的多幅图像，所述多幅图像是按时间排序的；

通过对在所述接收中接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域，使得空间上较大的子区域包括空间上较小的子区域，所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级；以及

通过跨与特定图像不同的图像中的分层级，搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

用于轨迹估计的方法和装置以及用于分割的方法

技术领域

[0001] 本发明涉及轨迹估计方法和轨迹估计装置以及分割方法，并且特别地涉及能够基于两幅或多幅图像而精确地估计诸如移动而改变形状的人等对象在图像上的轨迹的轨迹估计方法和轨迹估计装置以及分割方法。

背景技术

[0002] 常规上，已经广泛地进行了对图像上的轨迹进行估计的研究与开发。具体地，对诸如人等改变形状的对象的轨迹进行估计的技术是普遍用于例如数码摄像机或数码相机中的聚焦控制和图像质量改进处理以及汽车的驾驶安全支持系统或使用机器人的碰撞避免控制或警报的基础技术。

[0003] 用于估计轨迹的最通常的方法是通过在一幅图像中提供块（该块为部分区域）并且在另一幅图像中使用包括在该块中的像素信息（像素值）来执行空间搜索，从而基于像素信息中的相似性来估计运动矢量。通过时间上连接这样的运动矢量可以估计轨迹。

[0004] 然而，根据该方法，运动矢量是基于块内的像素值的相似性来估计的。因此，在包括诸如不具有纹理的均匀区域等区域或具有条纹图案的区域的图像的情况下，不可能精确地估计运动矢量，从而导致误差被包括在轨迹中。

[0005] 另一方面，非专利文献 2 描述了一种用于估计更精确轨迹的方法。利用该方法，在对诸如不具有纹理的均匀区域等区域的轨迹进行估计的情况下，仅通过使用诸如角落或边缘等鲁棒点来进行研究，在鲁棒点处，像素值不太可能随时间变化。这给出了允许更精确估计轨迹的优点。

[0006] 引用列表

[0007] 非专利文献

[0008] [非专利文献 1] 1989 年的 International Journal of Computer Vision, 卷 2, 第 283–310 页, 作者为 P. Anandan 的文章“*A computational Framework and an Algorithm for the Measurement of Visual Motion*”

[0009] [非专利文献 2] 1994 年的 IEEE Conference on Computer Vision and Pattern Recognition, 第 593–600 页, 作者为 Jianbo Shi 和 Carlo Tomasi 的文章“*Good Features to Track*”

发明内容

[0010] 技术问题

[0011] 然而，在常规的方法中，为了处理在诸如人等对象中频繁发生的大小 (scale) 改变或形状改变，需要采取对应于该大小改变或该形状改变的移动模型。因此，如果没有执行移动模型的合适设置，则存在不能够获得精确轨迹的问题。特别是，在不使用任何关于对象的先前知识的情况下，很难针对形状改变设置合适的运动模型。

[0012] 因此，构思了本发明来解决以上的问题，并且本发明的目的是提供一种能够基于

两幅或多幅图像而精确地估计诸如移动而改变形状的人等对象或包括均匀纹理的对象在图像上的轨迹的轨迹估计方法和轨迹估计装置以及分割方法。

[0013] 解决问题的方案

[0014] 为了实现上述目的,根据本发明一方面的轨迹估计方法是用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的轨迹估计方法,并且所述轨迹估计方法包括:接收包括在所述视频中的多幅图像,所述多幅图像是按时间排序的;通过对在所述接收中接收的所述多幅图像中的每一幅图像执行分层分割来在分层级(hierarchical level)上生成子区域,使得空间上较大的子区域包括空间上较小的子区域,所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级;以及通过跨与特定图像不同的图像中的分层级,搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹(representative trajectory)。

[0015] 应该注意,本发明不仅可以被实现为上述轨迹估计方法,而且还可以被实现为轨迹估计装置、程序和计算机可读非易失性记录介质,其中所述轨迹估计装置包括作为组成要素的、包括在所述轨迹估计方法中的特征步骤,所述程序使计算机执行包括在所述轨迹估计方法中的所述特征步骤,以及计算机可读非易失性记录介质例如是光盘只读存储器(CD-ROM),在该计算机可读非易失性记录介质上存储所述程序。

[0016] 发明的有益效果

[0017] 根据本发明,可以精确地对诸如移动而改变形状的人等对象或包括均匀纹理的对象在图像上的轨迹进行估计。

附图说明

[0018] 图1是示出根据本发明第一实施例的轨迹估计装置的基本结构的功能框图。

[0019] 图2是示出根据本发明第一实施例的轨迹估计装置的硬件结构的功能框图。

[0020] 图3是示出根据本发明第一实施例的轨迹估计装置的基本操作的流程图。

[0021] 图4是示出根据本发明第一实施例的用于估计运动矢量的处理的示例的简图。

[0022] 图5是示出根据本发明第一实施例的用于估计轨迹的处理的示例的简图。

[0023] 图6是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的示例的简图。

[0024] 图7是示出由根据本发明第一实施例的分层子区域生成单元输入的输入图像的示例的简图。

[0025] 图8是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0026] 图9是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0027] 图10是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0028] 图11是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0029] 图 12 是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0030] 图 13 是示出由根据本发明第一实施例的分层子区域生成单元执行的处理的结果的示例的简图。

[0031] 图 14 是示出由根据本发明第一实施例的代表轨迹估计单元生成模板的处理的示例的简图。

[0032] 图 15 是示出由根据本发明第一实施例的代表轨迹估计单元执行的匹配处理的示例的简图。

[0033] 图 16 是示出由根据本发明第一实施例的代表轨迹估计单元执行的匹配处理的另一示例的简图。

[0034] 图 17 是示出由根据本发明第一实施例的代表轨迹估计单元估计代表轨迹的处理的示例的简图。

[0035] 图 18 是示出由根据本发明第一实施例的代表轨迹估计单元估计的代表轨迹的示例的简图。

[0036] 图 19 是示出根据本发明第二实施例的分割装置的基本结构的功能框图。

[0037] 图 20 是示出根据本发明第二实施例的分割装置的基本操作的流程图。

[0038] 图 21 是示出由根据本发明第二实施例的分割单元执行的处理的示例的简图。

[0039] 图 22 是示出由根据本发明第二实施例的分割单元执行的处理的优点的简图。

[0040] 图 23 是示出由根据本发明第二实施例的分割单元执行的测地距离 (geodetic distance) 转换的结果的示例的简图。

[0041] 图 24 是示出由根据本发明第二实施例的分割单元执行的处理的结果的示例的简图。

[0042] 图 25 是示出根据本发明第二实施例的分割装置的基本操作的流程图。

[0043] 图 26 是示出由根据本发明第二实施例的变型的分割单元生成的最小生成树 (MST) 的示例的简图。

[0044] 图 27 是示出由根据本发明第二实施例的变型的分割单元生成的当前对的示例的简图。

[0045] 图 28 是示出由根据本发明第二实施例的变型的分割单元执行的用于自动地确定类别数量的示例的简图。

具体实施方式

[0046] 根据本发明的一个实施例是用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的轨迹估计方法，并且所述轨迹估计方法包括：接收包括在所述视频中的多幅图像，所述多幅图像是按时间排序的；通过对在所述接收中接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域，使得空间上较大的子区域包括空间上较小的子区域，所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级；以及通过跨与特定图像不同的图像中的分层级，搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

[0047] 根据该结构,跨分层(hierarchy)地搜索子区域。由于这促进了吸收形状和尺寸的差异,所以可以改进搜索精确度。特别是,可以精确地估计诸如移动而改变形状的人等对象或包括均匀纹理的对象的轨迹。

[0048] 优选地,上述轨迹估计方法还包括通过针对由包括在所述多幅图像中的每一幅图像中的一个或多个像素构成的每个块分析所述多幅图像之间的运动而在所述视频中生成所述块的轨迹,其中在所述估计中,通过使包括在所述特定图像中的所述子区域与包括在所述子区域中的所述块的所述轨迹一起移动来生成预测子区域,所述预测子区域是在与所述特定图像不同的所述图像中预测的子区域,并且通过跨与所述特定图像不同的所述图像中的所述分层级,搜索与所述预测子区域最相似的所述子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为所述代表轨迹。

[0049] 根据该结构,还可以基于属于每个子区域的轨迹,通过在下一时间生成模板而响应于子区域的形状改变来鲁棒地估计轨迹。

[0050] 更优选地,在所述估计中,使包括在所述生成中生成的每个子区域的轨迹对于所述每个子区域是平滑的,并且基于包括在所述每个子区域中的平滑的轨迹来估计所述代表轨迹。

[0051] 根据该结构,可以通过每个子区域地校正所述轨迹来针对每个区域估计连贯一致的轨迹。

[0052] 更优选地,在所述估计中,通过使用相似性测度(similarity measure)来确定子区域之间的相似性,从而搜索所述子区域,其中所述相似性测度包括所述子区域的像素值和形状信息中的至少之一。

[0053] 根据该结构,可以基于包括形状信息和像素值中的至少一个的相似性测度来估计轨迹。

[0054] 此外,在所述估计中,可以通过(i)采用曲线图(graph)、和(ii)将动态编程应用于所述曲线图而生成用于所述曲线图的最优路径来估计所述代表轨迹,在所述曲线图中:包括在所述特定图像中的所述子区域和包括在与所述特定图像不同的所述图像中的所述子区域是节点;使用边来对表示包括在所述特定图像中的所述子区域的一个节点和表示包括在与所述特定图像不同的所述图像中的所述子区域的另一节点进行连接;并且这样的值是边权重,即对于由所述边连接的子区域之间的较大相似性,该值是较小的。

[0055] 根据该结构,可以通过采用上述曲线图并且使用动态编程计算最优路径来从全局优化视角来看在相对长的时间段中精确地估计所述代表轨迹。特别地,在估计用于三个或更多个帧的代表轨迹的情况下,可以利用关于全部帧的信息来估计最优代表轨迹。

[0056] 优选地,在所述生成中,基于包括轮廓信息的特征,在分层级上生成所述子区域。

[0057] 根据该结构,通过使用包括轮廓信息的特征,可以生成所述子区域,所述子区域对于颜色和亮度波动是鲁棒的并且从所述子区域中排除诸如阴影等影响,并且因此可以估计高精确的轨迹。

[0058] 本发明的另一实施例是用于对每个移动对象的视频执行分割的分割方法,并且该分割方法包括:上述轨迹估计方法中包括的处理;以及通过将代表轨迹之间具有相似性的子区域进行结合来对视频执行分割。

[0059] 根据该结构,可以通过使用估计的轨迹和子区域来执行分割。因此,在改变形状的

对象的情况下,也可以更精确地执行分割。

[0060] 优选地,在所述执行中,通过连接小于阈值的距离来将代表轨迹之间的距离转换为测地距离、检测在获得的测地距离中的不连续点并且通过将彼此分开的距离比在所检测的不连续点处的测地距离小的轨迹分为一类来执行分割。

[0061] 根据该结构,将彼此分开的距离比在不连续点处的测地距离小的轨迹分为一类。因此,与使用欧几里得距离(欧几里得距离为线性距离)的聚类(clustering)相比,执行考虑了轨迹之间相似性的时空连续性的聚类。因此,可靠地区分了图片中的每个块属于同一对象(或区域)还是属于另一对象(或区域)。因此,甚至在检测诸如由具有不同运动的区域构成的人等关节式对象的情况下,也可以正确地执行分割,以便可靠地检测移动对象。换言之,可以对诸如移动而改变形状的人等移动对象精确地执行分割,并由此可以检测图像中的移动对象。

[0062] 此外,在所述执行中,基于对应于要被结合的子区域的代表轨迹和邻近所述要被结合的子区域的至少一个子区域的代表轨迹,可以确定是否结合所述要被结合的子区域。

[0063] 根据该结构,可以通过不仅考虑要被结合的子区域而且还考虑邻近的子区域来执行考虑邻近子区域的分割。因此,这产生了以下有益效果,即能够更好地反映图像的结构并且执行不易受诸如图像中的对象尺寸差异等影响的分割。

[0064] 优选地,在所述执行中,基于在根据曲线图计算的最小生成树中表示的连接关系来执行分割,在该曲线图中,将代表轨迹假设为节点并且通过使用边来将代表轨迹彼此连接;并且由边连接的代表轨迹之间的距离被假设为边权重。

[0065] 通过使用最小生成树以有限的连接关系表示所述代表轨迹,可以以最小计算量来执行分割。

[0066] 更优选地,在所述执行中,基于用于将具有以下关系的一对代表轨迹与邻近该对代表轨迹的至少一个代表轨迹连接的边权重,通过在最小生成树中确定是否结合该对代表轨迹来执行分割:代表轨迹是使用一条边直接彼此连接的。

[0067] 根据该结构,可以通过执行这种包括邻域(neighborhood)的成对聚类(pairwise clustering)来考虑包括手和脚之间关系的区域之间的关系。因此,这产生了允许甚至对诸如人等关节式对象进行精确分割的优点。此外,由于还考虑了与空间上远的地方的关系,因此可以执行不易受诸如图像中的对象尺寸变化或差异等影响的分割。

[0068] 本发明的又一实施例是用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的轨迹估计装置,并且所述轨迹估计装置包括:图像接收单元,其接收包括在所述视频中的多幅图像,所述多幅图像是按时间排序的;分层子区域生成单元,其通过对由所述图像接收单元接收的所述多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域,使得空间上较大的子区域包括空间上较小的子区域,所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级;以及代表轨迹估计单元,其通过跨与特定图像不同的图像中的分层级,搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

[0069] 本发明的又一实施例是用于对构成视频中的移动对象的全部或移动对象的部分的子区域的轨迹进行估计的程序,并且所述程序使计算机执行:接收包括在所述视频中的图像,所述图像是按时间排序的;通过对在所述接收中接收的所述多幅图像中的每一幅图

像执行分层分割来在分层级上生成子区域,使得空间上较大的子区域包括空间上较小的子区域,所述空间上较大的子区域和所述空间上较小的子区域属于彼此不同的分层级;以及通过跨与特定图像不同的图像中的分层级,搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。

[0070] 本发明的又一实施例是一种对每个移动对象的视频执行分割的分割装置,并且所述分割装置包括:上述轨迹估计装置;以及分割单元 1101,其通过将代表轨迹之间具有相似性的子区域进行结合来对视频执行分割。

[0071] 本发明的又一实施例是一种用于对每个移动对象的视频执行分割的程序,并且所述程序使计算机执行:包括在上述轨迹估计方法中的处理;以及通过将代表轨迹之间具有相似性的子区域进行结合来对视频执行分割。

[0072] 在下文中,参考附图来描述本发明的实施例。

[0073] 实施例 1

[0074] 图 1 是示出根据第一实施例的轨迹估计装置的结构的图。图 1 中示出的轨迹估计装置是对诸如移动而改变形状的人等对象或包括均匀纹理的对象的轨迹进行精确估计的装置,并且该轨迹估计装置包括图像接收单元 101、运动分析单元 102、分层子区域生成单元 103、代表轨迹估计单元 104 和输出单元 105。应该注意:本发明中的必要组成元件是图像接收单元 101、分层子区域生成单元 103 和代表轨迹估计单元 104,并且该轨迹估计装置不需要包括运动分析单元 102 和输出单元 105。

[0075] 图像接收单元 101 接收按时间排序的并且包括在视频中的图像的输入。该图像接收单元 101 例如是摄像机或连接到该摄像机的通信接口等。

[0076] 运动分析单元 102 针对由图像接收单元 101 接收的多幅图像中的每一幅图像中包括的一个或多个像素构成的每个块分析所述多幅图像之间的运动而在视频中生成所述块的轨迹。换言之,运动分析单元 102 通过使用图像接收单元 101 接收的多幅图像中的至少两幅时间上不同的图像来估计块运动。基于所估计的块运动,运动分析单元 102 通过追踪块在按时间排序的多幅图像上的运动来估计像素的轨迹。

[0077] 分层子区域生成单元 103 通过对由图像接收单元 101 接收的多幅图像中的每一幅图像执行分层分割来在分层级上生成子区域,使得在属于彼此不同的分层级的子区域中,空间上较大的子区域包括空间上较小的子区域。使用特别是包括轮廓信息的特征,分层子区域生成单元 103 对由图像接收单元 101 获得的图像执行分割,将其分割成子区域。分层子区域生成单元 103 从空间上粗糙的子区域到空间上精细的子区域分层地执行分割。本实施例具有以下特征,即空间上粗糙的子区域(空间上较大尺寸)包括空间上精细的子区域(空间上较小尺寸)。

[0078] 代表轨迹估计单元 104 通过跨与特定图像不同的图像中的分层级,搜索与包括在所述特定图像中的子区域最相似的子区域来将包括在所述特定图像中的所述子区域在所述视频中的轨迹估计为代表轨迹。利用该处理,代表轨迹估计单元 104 估计代表每个子区域的轨迹。更具体地,代表轨迹估计单元 104 通过使包括在所述特定图像中的子区域与包括在所述子区域中的块的轨迹一起移动来生成预测子区域,所述预测子区域是在与所述特定图像不同的图像中预测的子区域,并且代表轨迹估计单元 104 通过跨与所述特定图像不同的所述图像中的分层级,搜索与所述预测子区域最相似的所述子区域来将包括在所述特

定图像中的所述子区域在所述视频中的轨迹估计为所述代表轨迹。这里,产生如下有益效果,即通过使用用于每个子区域的轨迹来生成预测子区域(模板),从而能够处理子区域的形状改变。而且,跨分层执行的空间搜索产生以下有益效果,即能够精确地计算在具有均匀纹理等的区域中的轨迹。应该注意:代表轨迹估计单元104通过使用相似性测度来确定子区域之间的相似性,从而搜索子区域,其中该相似性测度包括子区域的像素值和形状信息中的至少之一。

[0079] 输出单元105将代表每个子区域的轨迹输出到监视器(显示器)等。根据本实施例,也可以基于每个子区域及其轨迹来输出结果,作为时空分割的结果。

[0080] 应该注意:包括在轨迹估计装置中的每个组成元件(图像接收单元101、运动分析单元102、分层子区域生成单元103、代表轨迹估计单元104和输出单元105)可以通过使用诸如在计算机上执行的程序等软件来实现,并且也可以被实现为诸如电子电路等硬件,该计算机包括CPU、随机存取存储器(RAM)、只读存储器(ROM)、通信接口、输入/输出(I/O)端口、硬盘、显示器等。这同样适用于另一实施例中的分割装置。图2是示出通过使用软件实现的根据本发明的轨迹估计装置的硬件结构的简图。在图2中,照相机201捕获并输出图像,并且计算机202获得由照相机201捕获的图像,执行用于估计代表轨迹的处理以及生成用于显示代表轨迹的计算结果的图像。显示器203获得并显示由计算机202生成的图像。计算机202包括:接口(I/F)204、CPU205、ROM206、RAM207、硬盘驱动器(HDD)208和视频卡209。使计算机202运行的程序由ROM206或HDD208预先保存。CPU205从ROM206或HDD208读出该程序,并且该程序由RAM207扩展,其中CPU205是处理器。CPU205执行在由RAM207扩展的程序中的每个编码的命令。用于执行程序的中间处理数据临时地存储在RAM207或HDD208上。根据程序的执行,I/F204收回由照相机201捕获的图像,并将其放到RAM207中。视频卡209根据程序的执行来输出所生成的图像,并且显示器203显示输出图像。

[0081] 应该注意:计算机程序可以存储在例如光盘等(非易失性记录介质)上,不限于为半导体的ROM206或HDD208。此外,计算机程序也可以经由有线或无线网络、广播等发送,并且计算机程序可以被收回并放到计算机的RAM207中。

[0082] 在下文中,使用图3中的流程图来描述根据本发明第一实施例的代表轨迹估计装置的操作。这里描述了以下示例,即接收包括诸如移动而改变形状的人等对象或包括均匀纹理的对象的图像的输入,并且然后估计轨迹。

[0083] 图像接收单元101接收T幅图像($T \geq 2$)的输入(步骤S301)。

[0084] 运动分析单元102估计关于输入的T幅图像之间的运动的信息,并且生成和输出轨迹(步骤S302)。这里,用于估计T幅图像之间的运动的方法参考T幅图像中的特定图像上的点I来搜索包括在其它T-1幅图像中的对应像素。

[0085] 应该注意:替代点I处的像素,用于估计图像之间的运动的方法可以参考I个矩形子区域(块)。例如,如图4所示,通过使用在时间t和时间t+1捕获的输入图像401来估计对应于在时间t的图像上的像素i的像素坐标(x_t^i, y_t^i)($i = 1, \dots, I$)的在时间t+1的图像上的像素坐标(x_{t+1}^i, y_{t+1}^i)。此时,在时间t的矩形子区域402(块)被扫描在时间t+1(虚线箭头)的图像上,并且将具有在时间t的矩形子区域中的像素值与在时间t+1的矩形子区域中的像素值之间差异的最小和的像素坐标(实线箭头)估计为对应的像素。对

于像素坐标而言,可以使用表示块的中心位置的像素坐标。这里,可以预先确定扫描范围。此外,为了计算差异,可以使用任何事物,只要它表示待比较的块的像素值中的差异:可以使用在时间 t 的矩形子区域中的像素值与在时间 t+1 的矩形子区域中的像素值之间的均方误差,并且也可以使用受诸如中值等离群值 (outlier) 影响较小的测度。在三幅或更多幅图像的情况下,通过顺序地估计对应的坐标来估计与在所输入的 I 幅图像中的每一幅图像中的点 I 相对应的点。

[0086] 作为估计如上所述的图像之间的对应点的另一种特定方法,也可以使用在非专利文献 1、3、4 中公开的方法。关于用于估计图像之间的对应点的特定方法,由于通过引用将这些非专利文献并入本文,因此这里不再重复对其的详细说明。

[0087] [非专利文献 3]2001 年的 International Conference on Computer Vision, 作者为 Vladimir Kolmogorov 和 Ramin Zabih 的文章“Computing Visual Correspondence with Occlusions via Graph Cuts”

[0088] [非专利文献 4]2007 年的 International Conference on Medical Image Computing and Computer Assisted Intervention, 作者为 Thomas Pock、Martin Urschler、Christopher Zach、Reinhard Beichel 和 Horst Bischof 的文章“A Duality Based Algorithm for TV-L1-Optimal-Flow Image Registration”

[0089] 在步骤 S302 中,当假设例如在步骤 S301 中输入 T 幅图像时,运动分析单元 102 使用在时间 t 和时间 t+1 捕获的两幅图像来估计像素 i 的运动矢量 (u_t^i, v_t^i)。这里,两幅图像不需要是按时间排序的,并且通过使用例如在时间 t 和时间 t+n 输入的图像,也可以估计像素运动。然而, n 是等于或大于 1 的整数。

[0090] 应该注意:代替使用运动矢量,运动分析单元 102 可以估计和使用仿射参数作为运动信息。这时,运动分析单元 102 可以估计全部像素的运动信息。此外,当旨在执行高速处理时,运动分析单元 102 可以将图像分成网格并且仅估计预定间隔的网格上的像素的运动信息,或者如之前所述,运动分析单元 102 可以将图像分成块并且估计关于每个块的运动信息。此外,当使用在非专利文献 1 中公开的方法来估计运动矢量时,运动分析单元 102 可以估计运动信息的可靠性,并且因此可以仅估计高可靠性的运动信息并且将所估计的信息用作运动信息。此外,当使用在非专利文献 3 中公开的方法来估计运动矢量时,运动分析单元 102 可以估计遮挡 (occlusion)。因此,运动分析单元 102 可以仅估计关于未被遮挡的像素的信息并且将所估计的信息用作运动信息。

[0091] 此外,作为用于估计像素运动的方法,代替使用如前所述的通过采取块的平行移动来估计运动矢量的方法,运动分析单元 102 可以使用通过采取块的仿射变换来估计运动矢量的方法。对于通过采取仿射变换来估计运动矢量的方法,可以使用在非专利文献 2 中公开的方法。非专利文献 2 中公开的方法是估计对应于在时间 t 和时间 t+1 输入的图像中的像素 i 的邻域的运动的仿射参数 A_t^i 。对于像素 i 而言,在时间 t 和时间 t+1 的图像上的像素位置 x_t^i 和 x_{t+1}^i 具有以下的关系 (表达式 1)。

[0092] 数学式 1

$$\mathbf{x}_{t+1}^i = \mathbf{A}_t^i \mathbf{x}_t^i$$

[0094] (表达式 1)

[0095] 根据该方法,特别对于正在旋转移动的移动对象而言,可以估计比在使用通过采

取平行移动来估计运动矢量的方法的情况下更精确地像素 i 的运动。

[0096] 再次,参考图 3,使用在步骤 S302 中估计的运动信息,运动分析单元 102 基于像素 i 的运动针对按时间排序的 T 幅图片估计像素 i 的轨迹(步骤 S303)。如图 5 所示,使用在步骤 S302 中估计的运动信息 502,从在时间 t 输入的输入图像 501 的像素 i 503 开始追踪像素 i 的运动。这时,通过使用像素 i 已经通过的在时间 t 的图像上的像素位置 (x_t^i, y_t^i) ,如下所述地估计像素 i 的轨迹。换言之,表达式 2 表示从第一幅图像到第 T 幅图像的像素位置 (x_t^i, y_t^i) 的轨迹。

[0097] 数学式 2

$$\mathbf{x}^i = (x_1^i, y_1^i, \dots, x_T^i, y_T^i)$$

[0099] (表达式 2)

[0100] 这里, T 是计算轨迹所使用的图像的数量。

[0101] 再次,参考图 3,分层子区域生成单元 103 通过执行多个分割来分层地将由图像接收单元 101 获得的相应图像分割为子区域(步骤 S304),其中所述多个分割中的每一个是使用不同的空间划分数量(粒度)执行的。每个分层级对应于子区域分割的粒度,并且每个子区域的空间大小(尺寸)在具有粗糙粒度的分层级上是相对大的,并且子区域的空间大小(尺寸)在具有精细粒度的分层级上是相对小的。

[0102] 图 6 示出将一幅输入图像 601 分层地分割为子区域 602 的结果的示例。如图 6 所示,在步骤 S304 中,分层子区域生成单元 103 将输入图像 601 分割为子区域 602,从而建立以下关系:较高分层级(具有粗糙粒度)上的子区域 602 包括较低分层级(具有精细粒度)上的子区域 602。应该注意:相反地,也可以建立以下关系:较低分层级上的子区域包括较高分层级上的子区域。

[0103] 分层级的数量可以预先确定,并且分层级的数量也可以开始利用子区域数量的上限来确定。不言自明的是,可以根据时间来改变分层级的数量。在第一实施例中,假设:预先确定分层级的数量。

[0104] 图 7 示出实际输入图像的示例,并且图 8 至图 13 示出按粗糙粒度的次序的分割的结果。换言之,图 8 示出具有最粗糙粒度(处于最高的分层级)的分割结果,并且图 13 示出具有最精细粒度(处于最低的分层级)的分割结果。如由这些分割结果所显示的那样,处于较高分层级的子区域包括处于较低分层级的子区域。

[0105] 由于用于分层地将图像分割为子区域的上述方法是如在非专利文献 5 中公开的广泛已知方法,这里不再重复其详细的描述。利用在非专利文献 5 中公开的方法,使用包括轮廓信息的特征所执行的这种子区域分割产生以下有益效果:可以允许较不易受颜色和亮度的波动影响的分割并且因此允许稳定的分割为子区域。在本实施例中,还可以使用包括轮廓信息的特征来执行子区域分割。

[0106] [非专利文献 5]2009 年的 Computer Vision and Pattern Recognition,作者为 Pablo Arbelaez、Michael Maire、Charles Fowlkes 和 Jitendra Malik 的文章“From Contours to Regions :An Empirical Evaluation”

[0107] 再次,参考图 3,使用子区域和属于该子区域的轨迹,代表轨迹估计单元 104 生成在子区域移动到目的地的时间的预测子区域,以便适应于子区域的时间上的形状改变(步骤 S305)。这里,对于预测子区域而言,像素值包括在子区域中并且在子区域开始从源移

动的时间定位该像素值。应该注意：预测子区域被用于之后要被描述的匹配处理（步骤 S306），但是，可以使用预测子区域的形状信息来执行该匹配或者可以使用像素值信息来执行该匹配。

[0108] 这里，描述了作为模板生成的预测子区域的示例。这里，优选的是，模板是利用预定分层级或利用包括预定数量子区域的分层级生成的。首先，在由分层子区域生成单元 103 执行的分割所产生的子区域和由运动分析单元 102 估计的轨迹中，将通过每个子区域的轨迹表示如下。通过子区域的轨迹是包括在子区域中的像素的轨迹。

[0109] 数学式 3

[0110] $\mathbf{x}^{sc_i} = (x_1^{sc_i}, y_1^{sc_i}, \dots, x_t^{sc_i}, y_t^{sc_i}, \dots, x_{T'}^{sc_i}, y_{T'}^{sc_i})$ $sc_i \in$ 子区域 sc (表达式 3) 这里， sc_i 表示属于子区域 sc 的轨迹 i。

[0111] 使用图 14 来描述用于在步骤 S305 中生成模板的方法。如在图 14(a) 和图 14(b) 中所示，首先，代表轨迹估计单元 104 利用属于子区域 701 的轨迹 702 的目的地的信息，计算在不同于对应于子区域 701 的时间的时间的模板 703 的形状。这里，描述用于生成在下一时间的用于子区域 701 的模板的方法，但是只要关于目的地的信息可以使用，就不需要生成在下一时间的模板，而是可以生成在任何时间的模板。例如，可以生成比对应于子区域 701 的时间早的时间的模板，并且可以以相同的方式执行以下的过程。更具体地，参考图 14(c)，假设：代表轨迹估计单元 104 利用属于在时间 t 的子区域 sc 的轨迹 i 来估计在时间 $t+1$ 的轨迹 705 的一组对应点 706，并且将该组对应点确定为模板 703。应该注意，假设：对于包括在模板 703 中的每个像素的像素值，使用在轨迹 705 的源处的像素值，即时间 t 时的像素值。因此，使用关于轨迹的目的地的信息，代表轨迹估计单元 104 更新每一帧的模板 703。

[0112] 利用这一点，模板的形状随时间而改变。换言之，由于预测子区域随时间而改变，因此具有以下有益效果：能够在响应于对象的形状改变的同时跟踪对象。此外，通过使用具有空间密度的轨迹，也可以甚至对于非线性的形状改变生成鲁棒的模板。出于解释的目的，图 14 已经描述了一个子区域，但是代表轨迹估计单元 104 生成用于多个或全部子区域的模板。

[0113] 再次，参考图 3，通过跨分层地空间搜索不同时间的相似子区域，代表轨迹估计单元 104 估计在子区域中的对应点（步骤 S306）。这里，为了清晰地描述，描述了跨分层在不同时间的两幅图像之间执行搜索的示例。具体地，代表轨迹估计单元 104 从分层地生成的在时间 $t+1$ 的子区域中，跨分层地并且使用在步骤 S305 中生成的预测子区域来搜索相对于预测子区域具有最小误差的子区域。在下文中，作为特定的示例，描述了一个示例，其中通过将预测子区域用作模板，执行与分割为在时间 $t+1$ 生成的子区域的图像或子区域的形状的匹配。

[0114] 参考图 15，代表轨迹估计单元 104 使用在步骤 S305 中生成的模板 801 来执行与在时间 $t+1$ 的多个子区域 802 中的每一个子区域的匹配。这里，可以使用一个模板 801 来执行与在全部分层级上的子区域 802 的匹配，或者仅可以对模板 801 所属的分层级的邻近分层级上的子区域 802 执行该匹配。通过计算模板 801 的像素值与在时间 $t+1$ 的每一个子区域 802 的像素值之间的差异的和来执行该匹配。

[0115] 假设：模板 801 与具有差异的最小和的子区域 802 最匹配。换言之，具有差异的最

小和等同于具有最大的相似性。这时，假设对应点是已经被用于生成模板 801 的在时间 t 的子区域 802 的形心位置 (centroid position) 以及与在时间 t+1 的子区域 802 最匹配的模板 801 的形心位置。通过沿着时间方向重复该计算，可以利用对应点来计算代表轨迹。

[0116] 这里，空间匹配的目标范围可以被预先限定为在时间 t 的模板 801 的位置的外围等，或在不执行空间匹配的情况下，也可以仅沿着分层方向执行匹配，这是因为轨迹信息已经被用于在步骤 S305 中生成模板 801。对于计算差异（相似性）而言，可以使用任何事物，只要它表示模板 801 和子区域 802 中的每一个子区域 802 之间的像素值的差异：可以使用模板 801 的像素值与在时间 t+1 的子区域 802 中的每一个的像素值之间的均方误差，并且也可以使用较不易受诸如中值等离群值影响的测度。此外，也可以使用模板 801 的像素数量与要被匹配的子区域 802 的像素数量之和，对诸如像素值的差异的和和均方误差等测度执行归一化。例如，当假设模板 801 的像素数量是 T_{num} ，要被匹配的子区域 802 的像素数量是 R_{num} ，并且彼此重叠并包括在模板 801 和要被匹配的子区域 802 中的像素的像素数量是 O_{num} ，可以根据以下的表达式 4 来对差异的和归一化。

[0117] 数学式 4

[0118] 归一化的差异的和 = 差异的和 $\times (T_{num} + R_{num}) / O_{num}$ (表达式 4)

[0119] 此外，代替使用像素值，可以将模板 801 的形状与要被匹配的子区域 802 的形状之间的相似性用作相似性测度。例如，可以将通过使用模板 801 和子区域 802 的像素数量对当模板 801 和子区域 802 叠加时重叠的像素的数量进行归一化而得到的值用作相似性。然后，具有最高相似性的子区域 802 可以被认为与模板 801 最匹配。

[0120] 另外，如图 16 所示，当子区域在两幅图像之间的形状改变大时，可能具有如由分层级 A 的匹配结果 807 所示的小相似性，从而导致即使当使用分层级 A 的模板 805 时也不能执行合适匹配的情况。在该情况中，可以执行以下的处理。当不能合适地执行匹配时，相似性小。因此，当使用分层级 A 的模板 805 进行匹配所产生的相似性小于预定值时，使用用于在分层级 B 上的子区域的模板来执行匹配，该分层级 B 包括通过分割生成的、比在分层级 A 上的子区域精细的子区域。这里，由于在步骤 S304 中输入图像 601 被分割为子区域 602 以便建立其中较高分层级（具有粗糙粒度）的子区域 602 包括较低分层级（具有精细粒度）的子区域 602 的关系，因此可以使用分层级 A 的模板和分层级 B 的模板之间的包含关系。具体地，使用两个模板中的每一个（例如分层级 B 的模板 806）来执行匹配，其中该分层级 B 的模板 806 与分层级 A 的模板 805 之间为包含关系并且包括通过分割生成的更精细的子区域。因此，如由分层级 B 的匹配结果 808 所示，甚至当子区域在两幅图像之间改变形状时，也可以执行更精确的匹配。这时，可以将分层级 B 的两个模板的代表轨迹的平均值确定为分层级 A 的模板的代表轨迹。

[0121] 如上所述，在由于子区域的形状改变所致而不能执行匹配的情况下，由此使用具有包含关系的不同分层级的模板所执行的匹配产生了能够计算精确的轨迹的有益效果。当使用如在非专利文献 5 中公开的用于分割为子区域的方法，发生这样的情况，其中由于对象随时间移动或者在不同时间的照相机运动的原因，分割为子区域的结果在时间 t 和时间 t+1 之间不同。另外，也存在这样的情况，其中分层中的子区域 802 的粗糙度取决于时间而变化。如果是这种情况，具有大尺寸差异和大形状差异的子区域 802 与模板 801 之间的匹配将要被执行，从而引起不能够执行合适匹配的问题。反之，如通过图 15 中的匹配结果 803

所示,跨分层的匹配很可能吸收形状和尺寸的差异,从而产生改进匹配精确度的有益效果。

[0122] 再次,参考图3,代表轨迹估计单元104使用在步骤S306中执行的匹配的结果来估计每个子区域802的代表轨迹(步骤S307)。如图17所示,代表轨迹估计单元104通过在时间上连接在步骤S306中获得的对应点904来将每个模板801的轨迹估计为代表轨迹。代表轨迹可以由以下的表达式5表示。

[0123] 数学式5

$$[0124] \mathbf{x}^r = (x_1^r, y_1^r, \dots, x_t^r, y_t^r, \dots, x_T^r, y_T^r)$$

[0125] (表达式5)

[0126] 这里, x_i^t, y_t^t 是在时间i的图像中的对应点904的像素位置,并且代表轨迹由按照时间排列的对应点904构成。

[0127] 如图18所示,轨迹估计装置可以通过执行上述步骤S301到步骤S307的处理来获得用于每一个子区域802的代表轨迹1002,其中该代表轨迹1002按照时间连接对应点1001。因此,子区域802具有空间连接性,并且对应点1001具有子区域802的时间连接性。因此,可以使作为代表轨迹1002的时空图像信息集中。当执行分割或识别图像时,这产生了能够进一步减少计算量的有益效果。另外,还可以使用这里提取的子区域802和按照时间连接子区域802的代表轨迹1002作为对时空图像执行分割的结果。

[0128] 如上所述,在本发明的第一实施例中,图像被分层地分割为区域,并且子区域之间的对应性通过执行跨分层的匹配而建立。据此,这种基于轮廓信息来分割为子区域的方案允许每个子区域保持形状信息。基于形状信息的匹配产生了以下有益效果,即该匹配对于由阴影和亮度改变引起的像素值改变是鲁棒的。另一方面,为了执行基于形状信息的匹配,需要每个匹配目标保持形状信息。即,目标应该被分割为子区域。然而,很难将具有相同粒度但时间上不同的图像分割为子区域。因此,通过将图像分割为具有在分层上不同粒度的子区域并且执行跨分层的搜索,可以对包括在这种彼此不同的图像中的子区域执行匹配。

[0129] 此外,这有助于建立子区域之间的对应性同时吸收形状和尺寸的差异,从而产生改进匹配精确度的有益效果。具体地,这产生了以下有益效果,即能够精确地建立子区域之间的对应性,甚至对于诸如移动而改变形状的人或包括均匀纹理的人等对象也是如此。另外,基于子区域之间的对应性来估计轨迹产生了能够计算精确轨迹的另一有益效果。此外,使用根据本发明实施方法估计的代表轨迹还产生了减少计算量以及改进分割和图像识别的精确度的有益效果。

[0130] 应该注意:在第一实施例中,已经描述了通过执行在时间t和时间t+1输入的图像之间的匹配来估计代表轨迹的示例,但是代表轨迹也可以通过使用在时间t和时间t+n输入的图像执行匹配来估计。然而,n是大于等于1的整数。

[0131] 应该注意:在第一实施例的步骤S305中,如表达式3所示,已经从由运动分析单元102估计的轨迹中提取了经过由分层子区域生成单元103分割所得到的每一个子区域的轨迹,但是这并没有将轨迹限制用在步骤S305中,并且也可以使用采用另一方法获得的轨迹。例如,可以使用属于子区域sc的轨迹i来重新校正轨迹。换言之,代表轨迹估计单元104每个子区域地使包括在子区域中的轨迹平滑,并且基于子区域和包括在子区域中的平滑轨迹来估计代表轨迹。具体地,如以下的表达式6所示,通过将平滑滤波型的双边滤波应用到包括在由表达式3表示的子区域sc中的轨迹i,可以减少离群值的影响并且估计具有

空间平滑性的轨迹。这时,通过不使用不属于子区域 sc 的轨迹来执行双边滤波,可以为每个子区域计算连贯一致的轨迹。

[0132] 此外,代替执行双边滤波,可以通过执行处理,例如排除以等于或大于预定阈值的值 (level) 偏离属于子区域 sc 的轨迹的平均值的轨迹来排除离群值。

[0133] 数学式 6

$$[0134] \hat{\mathbf{x}}^{sc_i} = \frac{\sum w^{sc_j} \mathbf{x}^{sc_j}}{\sum_{\mathbf{x}^{sc_j} \in sc} w^{sc_j}}$$

[0135] (表达式 6)

[0136] 这里,数学式 7 : $\hat{\mathbf{x}}^{sc_i}$ 是由双边滤波校正的轨迹。应该注意 :w 是根据以下的表达式 7 被计算为加权参数。

[0137] 数学式 8

$$[0138] w^{sc_j} = N(\sqrt{(p^{sc_i} - p^{sc_j})^2 + (q^{sc_i} - q^{sc_j})^2}, \sigma)$$

$$[0139] \cdot N(\sqrt{(u^{sc_i} - u^{sc_j})^2 + (v^{sc_i} - v^{sc_j})^2}, \sigma)$$

$$[0140] \cdot N(|I^{sc_i} - I^{sc_j}|, \sigma)$$

[0141] (表达式 7)

[0142] 这里, N 表示高斯分布,而权重 w 是基于像素位置、运动矢量和像素值的从左侧的差异来确定的,并且根据表达式 6 来校正轨迹。

[0143] 因此,通过使用属于相同子区域的相邻轨迹来计算加权平均值产生了以下有益效果,即允许减少离群值的影响以使得邻域中的相似性增加,并且允许获得一致的轨迹。这里,高斯分布被用在表达式 7 中,但是使用高斯分布不是必需的。而且,也可以使用关于运动矢量、像素值和像素位置之一的信息来计算权重。因此,只要方法可以减少轨迹的离群值的影响,就没有关于使用什么方法的限制。

[0144] 应该注意 :在第一实施例中,已经描述了以下示例,其中 :在步骤 S305 中,通过使用子区域和属于子区域的轨迹,预测子区域被生成为模板,而在步骤 S306 中,通过使用生成的模板跨分层地执行在时间 t+1 生成的被分割为子区域的图像或子区域的形状之间的匹配。然而,该方法不限于上述内容。即,可以使用任何其它方法,只要该方法跨分层地建立了在不同时间的子区域之间的对应性。

[0145] 这里描述了通过在三个或更多个图像上执行搜索并且生成时间上的最优路径来估计代表轨迹的示例。代表轨迹估计单元 104 通过采取曲线图并且将动态编程应用到曲线图来生成用于该曲线图的最优路径来估计代表轨迹,在曲线图中,包括在特定图像中的子区域和包括在与该特定图像不同的图像中的子区域是节点,使用边来连接表示包括在该特定图像中的子区域的节点之一和表示包括在与该特定图像不同的图像中的子区域的另一节点,并且这样的值是边权重,即对于由所述边连接的子区域之间的较大相似性,该值是较小的。换言之,代替获得在最匹配模板的子区域中的对应点,也可以在步骤 S306 中,通过假设在时间 t 的子区域为节点并且假设用于匹配的值(例如在模板和子区域之间的像素值的差异)是边权重来生成曲线图。利用该曲线图,使用边来连接在不同时间的子区域(节点)。应该注意 :在将相似性用于匹配的情况下,期望的是边权重对于较高相似性具有较小

值（在下文中，称为路径值），例如其倒数。曲线图的使用允许处理被执行，同时保持用于候选对应点的路径值，而不会将最佳匹配对应点仅限于一个点。也可以通过将动态编程（DP）匹配应用到该曲线图而在曲线图中生成最优路径，从而生成代表轨迹。该方法的使用还允许估计与在表达式 5 中相同的代表轨迹。使用根据子区域和模板的路径值的候选对应点来生成曲线图从而最终计算最优路径产生了以下有益效果，即能够从全局优化的视角在相对长的时间段中精确地估计代表轨迹。应该注意：在非专利文献 6 中详述了 DP 匹配。对于 DP 匹配而言，由于通过引用将非专利文献 6 并入本文，这里不再重复对其的详细描述。

[0146] [非专利文献 6] 1993 年的 IEEE Press, 第 624–633 页，作者为 John R. Deller、Jr. John H. L. Hansen 和 John G. Proakis 的文章“Discrete-Time Processing of Speech Signals”

[0147] 应该注意：本发明的第一实施例已经描述了：在步骤 S307 中，估计用于每个区域的一条代表轨迹，但是在步骤 S307 中要被估计的代表轨迹不限于此，并且还可以使用另一方法来估计代表轨迹。例如，在估计具有空间密度的轨迹的情况下，可以使用在步骤 S306 中估计的对应点来校正在步骤 S302 中估计的运动矢量和轨迹。

[0148] 更具体地，可以通过以下操作来获得轨迹 $x^{\text{new_i}}$ ：从在步骤 S302 中估计的运动矢量中，基于最接近模板的形心位置 (xm, ym) 的像素的运动矢量 $(u1, v1)$ 和在步骤 S306 中估计的对应点 (xc, yc) 来例如按照表达式 8 计算校正矢量 $(\Delta u, \Delta v)$ ，并且将该校正矢量应用到属于模板的所有轨迹。

[0149] 数学式 9

$$\Delta u = u1 - (xc - xm)$$

[0150] $\Delta v = v1 - (yc - ym)$

[0151] (表达式 8)

[0152] 数学式 10

$$x^{\text{new_i}} = x^{\text{sc_i}} + \Delta u^{\text{sc}} \quad (\text{表达式 9})$$

[0154] 这里，

[0155] 数学式 11

$$\Delta u = (\Delta u, \Delta v) \quad (\text{表达式 10})$$

[0157] 如上所述，使用校正的轨迹允许获得具有密度的轨迹，从而产生允许使用轨迹来进行对象的更详细运动分析等有益效果。

[0158] 应该注意：在本发明第一实施例的步骤 S306 中，使用预测子区域来跨分层地执行匹配，但是可以在没有由运动分析单元 102 执行的处理的情况下实现匹配。在该情况下，将在不使用预测子区域的情况下执行匹配。换言之，代表轨迹估计单元 104 可以使用子区域本身而非预测子区域来在两幅图像之间执行跨分层的匹配。这时，在使用预测子区域的情况下，具有以下有益效果，即能够在执行子区域匹配时基于运动分析的结果来缩小搜索范围；然而，在不使用预测子区域的情况下，足以预先确定空间搜索范围并执行范围内的搜索。

[0159] 实施例 2

[0160] 接下来，作为第二实施例，描述了一种示例，其中使用在第一实施例中估计的子区域和对应于该子区域的轨迹，通过进一步对轨迹执行聚类来执行分割。虽然可以通过使用

K 均值方法来聚类代表轨迹，从而执行分割，但是这里描述了甚至在改变形状的对象的情况下也可以执行更精确分割的示例。

[0161] 图 19 是示出根据第二实施例的分割装置的结构的简图。图 18 所示的分割装置是通过精确地估计诸如移动而改变形状的人等对象或包括均匀纹理的对象的代表轨迹，并且通过使用估计的代表轨迹来执行聚类而对图像执行分割的装置。分割装置包括图像接收单元 101、运动分析单元 102、分层子区域生成单元 103、代表轨迹估计单元 104、分割单元 1101 和输出单元 105a。

[0162] 由于图像接收单元 101、运动分析单元 102、分层子区域生成单元 103 和代表轨迹估计单元 104 与在第一实施例中的那些单元相同，因此不重复其描述。

[0163] 分割单元 1101 通过使用由代表轨迹估计单元 104 估计的代表轨迹并且基于代表轨迹之间的相似性来结合具有相似代表轨迹的子区域，从而对代表轨迹执行聚类，并且使用由分层子区域生成单元 103 生成的子区域和聚类结果来对图像执行分割。换言之，分割单元 1101 通过连接小于阈值的距离将代表轨迹之间的距离转换为测地距离，检测在获得的测地距离中的不连续点，并且将彼此分开的距离比在所检测的不连续点处的测地距离小的轨迹分为一类来执行分割。之后将描述使用测地距离的分割。

[0164] 输出单元 105a 从由分割单元 1101 执行的聚类结果将结合的轨迹输出到监视器等，或者将时空分割的结果作为图像输出到监视器等。

[0165] 在下文中，使用图 20 中的流程图来描述根据本发明第二实施例的分割装置的操作。这里详细描述了使用在第一实施例中估计的子区域和对应于该子区域的轨迹，通过进一步对子区域执行聚类来执行分割的示例。

[0166] 由于步骤 S301 到步骤 S307 与在第一实施例中的那些步骤相同，因此不重复其描述。

[0167] 接下来，分割单元 1101 通过对由表达式 5、表达式 6 或表达式 9 表示的代表轨迹执行聚类来将代表轨迹分类（步骤 S1201）。因此，可以通过将被分为相同类别的代表轨迹所归属的子区域假设为相同的段（segment）来执行分割。在下文中，描述了对由表达式 5 表示的代表轨迹执行聚类的示例，但是在不修改的情况下它甚至也适用于表达式 6 和表达式 9 的情况。

[0168] 第一，分割单元 1101 使用由表达式 5 表示的代表轨迹 r 来计算表示代表轨迹之间的相似性的距离矩阵。可以根据下面的表达式 11 来计算代表轨迹 r 和代表轨迹 s 之间的线性距离 $f(r, s)$ 。

[0169] 数学式 12

$$[0170] f(r, s) = \text{mean}_{rs} + w \cdot \sqrt{\text{var}_{rs}}$$

[0171] （表达式 11）

[0172] 这里， w 是加权因子，并且 w 是由设计者设置的参数。此外， mean_{rs} 和 var_{rs} 分别由下面的表达式 12 和表达式 13 来表示。

[0173] 数学式 13

$$[0174] \text{mean}_{rs} = \frac{1}{T} \sum_{t=1}^T d_{rs}^t$$

[0175] （表达式 12）

[0176] 数学式 14

$$[0177] \quad \text{var}_{rs} = \frac{1}{T} \sum_{t=1}^T (d_{rs}^t - \text{mean}_{rs})^2$$

[0178] (表达式 13)

[0179] 这里,

[0180] 数学式 15

$$[0181] \quad d_{rs}^t = \sqrt{(x_t^r - x_t^s)^2 + (y_t^r - y_t^s)^2} \quad (\text{表达式 14})$$

[0182] 如表达式 11 所示,除了由表达式 12 表示的代表轨迹之间的时间均值之外,由表达式 13 表示的代表轨迹之间的距离的时间波动分量是线性距离 $f(r, s)$ 的元素。特别地,由表达式 13 表示的代表轨迹之间的距离的波动分量表示运动的相似性,并且据此,可以不仅捕获代表轨迹之间的距离不改变的刚性对象,而且还可以捕获关节式对象等的形状改变。

[0183] 接下来,分割单元 1101 对根据表达式 11 计算的线性距离 $f(r, s)$ 执行由下面的表达式 15 示出的非线性化。

[0184] 数学式 16

[0185]

$$f'(r, s) = \begin{cases} f(r, s) & \text{if } r \text{ 是 } K-\text{NearestNeighbor} \\ \infty & \text{否则} \end{cases}$$

[0186] (表达式 15)

[0187] 换言之,当代表轨迹 r 是当前代表轨迹时,分割单元 1101 以与代表轨迹 s 的线性距离的上升次序来选择 K 条代表轨迹 s ,并且分割单元 1101 不改变与选择的代表轨迹 s 的距离,而无限地改变与未选择的代表轨迹 r 的距离。这里, K 条轨迹是以线性距离的上升次序而被选择的,但是可以如在下面的表达式 16 中所示的那样设置阈值 R 。换言之,当代表轨迹 r 是当前轨迹时,在与代表轨迹 s 的线性距离低于阈值 R 的情况下,分割单元 1101 不改变距离;而在线性距离等于或高于阈值 R 的情况下,分割单元 1101 将与代表轨迹 r 的距离改变为无穷大值。

[0188] 数学式 17

[0189]

$$f'(r, s) = \begin{cases} f(r, s) & \text{如果 } f(r, s) < R \\ \infty & \text{否则} \end{cases}$$

[0190] (表达式 16)

[0191] 接下来,分割单元 1101 通过使用代表轨迹 $f'(r, s)$ 之间的非线性化距离,如下面的表达式 17 表示的那样更新距离,从而计算测地距离 $g(r, s)$ 。

[0192] 数学式 18

[0193] $g(r, s) = \min(f'(r, s), f'(r, k) + f'(k, s))$ (表达式 17) 应该注意: $\min(x, y)$ 是将值 x 和值 y 中较小的一个值返回的函数。此外, k 是代表轨迹,并且 k 是将经过的中继点以便从代表轨迹 r 到达代表轨迹 s 。这里, $f'(r, k) + f'(k, s)$ 中的中继点 s 不限于一个点。该方法是称为 Dijkstra 方法的最短路径搜索方法并且在以下的非专利文献 7 中描述了该方法。

[0194] [非专利文献 7] 1959 年的 Numerische Mathematik, 第 269–271 页, 作者为

E. W. Dijkstra 的文章 “A note on two problems in connexion with graphs”

[0195] 这里, 使用在图 21(a) 到图 21(c) 中的概念图来描述由表达式 15 到表达式 17 所表示的非线性化。这里描述使用表达式 16 的非线性化, 但是利用其它非线性化处理可以预期同样的有益效果。首先, 图 21(a) 示出二维数据分布。这里, 每个数据点对应于由表达式 5 表示的代表轨迹。在不根据表达式 15 到表达式 17 执行非线性化的情况下, 如图 21(b) 所示, 数据点 i 和数据点 j 之间的距离小于数据点 i 和数据点 k 之间的距离。然而, 如图 21(c) 所示, 作为执行例如表达式 16 中的处理和表达式 17 中的处理的结果, 数据点 i 和数据点 j 之间的距离变为不是欧几里得距离而是被称为测地距离, 该测地距离是描绘如由箭头指示的数据点的距离。因此, 不同于使用欧几里得距离的情况, 数据点 i 和数据点 j 之间的距离变得大于数据点 i 和数据点 k 之间的距离。

[0196] 接下来, 使用在图 22(a) 和图 22(b) 中的概念图来描述执行这种如表达式 15 到表达式 17 所示的非线性化的有效性。这里, 为了清楚地描述线性距离 $f(r, s)$ 和距离 $g_{\text{new}}(r, s)$ 之间的差异, 代表轨迹之间的距离被用作示例, 如表达式 14 所示。实践中, 通过使用除了代表轨迹之间距离之外的、如表达式 11 表示的作为运动相似性的代表轨迹之间的距离的波动分量, 不仅可以捕获关节式对象等的形状而且还可以捕获关节式对象等的形状的改变。图 22(a) 是不执行表达式 15 到表达式 17 中的处理的情况示例。例如, 头部区域的代表轨迹 r 1402 和手部区域的代表轨迹 s 1403 之间的距离是由线性距离 1401 表示的距离。另一方面, 如图 22(b) 所示, 作为执行如在表达式 15 到表达式 17 中表达的非线性化的结果, 头部区域的代表轨迹 r 1402 和手部区域的代表轨迹 s 1403 之间的距离变为以下距离, 即该距离是由箭头所示的线性和, 该距离经过代表轨迹 k 1404 引导到代表轨迹 s 1403。因此, 尽管线性距离 1401 不可以表示诸如人等用关节连接的对象的形状, 但是测地距离可以将用关节连接的形状表示为距离。应该注意: 计算测地距离的方法不限于表达式 17。

[0197] 接下来, 通过使用已经被转换为测地距离的 $g(r, s)$ 来检测测地距离中的不连续点, 从而执行聚类。这里, 测地距离中的不连续点落入代表轨迹 r 和代表轨迹 s 之间, 其中 $g(r, s)$ 是无穷大的。参考图 23 描述测地距离转换的结果, 其中该测地距离转换是相对于由表达式 16 表示的阈值 R 获得的。这里, 图 23(a) 是示出在步骤 S306 中估计的轨迹 a 到 h 的简图。图 23(a) 中示出的子区域 1502 是代表轨迹所归属的子区域。图 23(b) 是包括图 23(a) 所示的轨迹 a 到 h 的较高维度空间的概念图。这里, 代表轨迹的数量是八个, 但是在实践中, 可以使用对应于每个子区域的代表轨迹。这里, 包括代表轨迹的较高维度空间 1503 中的一个点对应于由表达式 5 表示的代表轨迹中的每一条。此外, 在较高维度空间 1503 上, 点到点的距离不对应于矢量之间的欧几里得距离, 而是对应于如由表达式 17 表示的测地距离。

[0198] 图 23(c) 是示出聚类结果的简图。这里, 在图 23(c) 中, $f(e, f) > f(c, d) > f(f, g) > f(b, c)$, 其中 $f(a, b)$ 是如由表达式 11 表示的代表轨迹 a 和代表轨迹 b 之间的距离。这里, 假设: 当阈值被设置为 R 时, 距离 $f(f, g)$ 具有大于阈值 R 的值。在该情况下, 即使当根据表达式 17 来计算测地距离时, $g(e, f)$ 、 $g(c, d)$ 和 $g(f, g)$ 中的每一个变为无穷大。因此, 分割单元 1101 确定代表轨迹 c 和 d 之间、代表轨迹 e 和 f 之间以及代表轨迹 f 和 g 之间的每个点是测地距离中的不连续点。因此, 代表轨迹 a 、 b 和 c 之间的测地距离不具有无穷大值, 这是因为代表轨迹 a 、 b 和 c 可以在不通过彼此的测地距离中的不连续点的情况下

被追踪；反之，例如，从这些代表轨迹 a、b 和 c 中的每一条到另一条代表轨迹，代表轨迹 a、b 和 c 通过测地距离中的不连续点 g(c, d)，使得代表轨迹 a、b 和 c 中的每一条与另一条代表轨迹之间的测地距离具有无穷大值。因此，不具有无穷大测地距离的一对代表轨迹 i 和 j 被分为相同的类别，并且具有无穷大测地距离的那些代表轨迹被分为不同的类别。然后，通过将对分为具有无穷大测地距离的对和不具有无穷大测地距离的对，总共可以将轨迹分为四个类别 θ_1 、 θ_2 、 θ_3 和 θ_4 。这里，分割单元 1101 将分为每个类别 θ_i 的代表轨迹所归属的子区域结合为一个段。例如，图 23 的示例示出分为类别 θ_1 的代表轨迹所归属的子区域对应于两个头部子区域和一个胸部子区域，并且可以通过时空上将这三个子区域结合为一个段来执行分割。这同样适用于类别 θ_2 到 θ_4 。

[0199] 利用上述处理，分割单元 1101 可以确定这种代表轨迹属于相同的类别，从而判断不具有无穷大测地距离的代表轨迹是连续的。此外，分割单元 1101 可以通过确定具有无穷大测地距离的代表轨迹是不连续的，基于测地距离中的不连续点来将代表轨迹分类。因此，通过结合分为一个类别的代表轨迹所归属的子区域，可以执行分割。

[0200] 输出单元 105b 将分割结果输出到监视器等。

[0201] 应该注意：根据本实施例，通过为代表轨迹计算测地距离，甚至还可以对改变形状的对象执行分割。因此，在图 23 中示出的示例中，反映了头部区域运动和手臂运动之间的差异和大腿运动和小腿运动之间的差异，使得可以将头部区域、手臂、大腿和小腿分割为单独的类别。此外，根据本方法，可以不仅对一个移动对象执行分割以便将该移动对象分割为多个身体部分，而且还可以对包括在如图 24(a) 所示的不止存在一个人的场景中的背景执行分割，并且还可以对如图 24(b) 所示的每个移动对象执行详细的分割。例如，通过为图像设置较大的阈值 R，还可以在监视器上显示图像，使得可以通过将不同颜色分配给图像上的区域并且每个区域对应于类别 θ_m 来将由分割获得的区域彼此区分开来。如图 24(a) 所示，每个移动对象可以以不同的颜色显示，或者如图 24(b) 所示，每个区域可以被单独显示。换言之，如图 24 所示，根据本实施例的分割装置可以通过执行上述分割（图 24(a)）来检测移动对象，并且也可以检测包括在一个移动对象中的区域，即，根据本实施例的分割装置可以对一个移动对象执行分割（图 24(b)）。

[0202] 如上所述，通过基于代表轨迹之间的相似性来执行聚类，从而对相似的代表轨迹聚类，使得分割被执行。因此，作为将具有小距离和具有相似运动的区域分为一个类别并且时间上追踪在视频中移动的对象的区域的结果，可以对包括移动对象的时空图像执行分割。此外，基于轨迹的测地距离中的相似性来执行聚类。因此，甚至当在图像中存在不同尺寸的移动对象时或当移动对象的一部分被隐藏时，也可以鲁棒地对对象中的区域执行分割。而且，基于测地距离中的不连续点的聚类允许聚类不受类别尺寸的影响。因此，甚至当大的对象和小的对象同时存在时或当具有大运动的对象和具有小运动的对象同时存在时，也可以可靠地执行分割。

[0203] 如上所述，根据第二实施例，将彼此分开的距离小于在不连续点处的测地距离的轨迹分为一个类别。因此，与使用欧几里得距离（为线性距离）的聚类相比，执行考虑了轨迹之间的相似性中的时空连续性的聚类。因此，可靠地区分了图片中的每个块属于同一对象（或同一区域）还是属于另一对象（或区域）。因此，甚至在检测诸如由在运动方面不同的区域构成的人等关节式对象的情况下，也可以精确地执行分割，以便可靠地检测移动对

象。换言之,可以对诸如移动而改变形状的人等移动对象正确地执行分割并由此检测图像中的移动对象。

[0204] 实施例 2 的变型

[0205] 接下来描述已经在实施例 2 中描述的示例的变型,在该变形中,将执行时空分割。通过在使用估计的代表轨迹执行聚类时执行包括邻近代表轨迹的成对聚类,根据本变型的分割装置对图像执行分割,同时自动地确定类别数量。分割装置包括图像接收单元 101、运动分析单元 102、分层子区域生成单元 103、代表轨迹估计单元 104、分割单元 1101 和输出单元 105a。

[0206] 由于图像接收单元 101、运动分析单元 102、分层子区域生成单元 103、代表轨迹估计单元 104、分割单元 1101 和输出单元 105a 与在第二实施例中的那些单元相同,因此不重复其描述。分割单元 1101 执行的处理不同于第二实施例 2 中的处理。因此,在本变型中,描述将集中在由分割单元 1101 执行的处理上。

[0207] 分割单元 1101 基于对应于要被结合的子区域的代表轨迹和与该要被结合的子区域邻近的至少一个子区域的代表轨迹来确定是否结合该要被结合的子区域。更详细地,分割单元 1101 利用曲线图来计算最小生成树,在曲线图中,代表轨迹被假设为节点并且使用边来连接代表轨迹,并且由边连接的轨迹之间的距离被假设为边权重,以及分割单元 1101 基于在最小生成树中表示的连接关系来执行分割。优选地,分割单元 1101 通过基于连接具有以下关系的一对代表轨迹的边权重和邻近该对代表轨迹的至少一条代表轨迹来确定是否结合该对代表轨迹,从而执行分割,其中该对代表轨迹具有的以下关系是使用一条边将代表轨迹相互连接。

[0208] 在下文中,使用图 25 中的流程图描述根据本变型的分割装置的操作。这里,详细地描述以下示例,其中通过使用在第一实施例中估计的子区域和对应于子区域的轨迹来执行包括邻近代表轨迹的代表轨迹的成对聚类,从而执行分割。

[0209] 可以通过使用 K 均值方法来聚类代表轨迹,从而执行分割,在该情况下,代表轨迹之间的距离不限于欧几里得距离。然而,诸如树状图之类的成对聚类具有能够自由地设置两条代表轨迹之间的距离的优点。例如,如表达式 11 所示,也可以包括在距离中的运动差异。

[0210] 由于步骤 S301 到步骤 S307 与在第一实施例和第二实施例中的步骤相同,所以不重复其描述。

[0211] 首先,分割单元 1101 基于根据表达式 5、表达式 6 和表达式 9 计算的代表轨迹来计算代表轨迹之间的距离,并且基于计算的代表轨迹之间的距离来生成最小生成树(在下文中,被称为“MST”),其中该最小生成树是表示代表轨迹之间的相关性的信息(步骤 S1701)。

[0212] 在下文中,详细地描述了步骤 S1701 中的处理。具体地,为了计算代表轨迹之间的距离以便执行成对聚类,可以使用在第二实施例中描述的表达式 11 中的线性距离或表达式 17 中的测地距离。代表轨迹之间的距离可以不仅包括像素位置而且还可以包括以距离表示运动相似性的分量。因此,可以不仅捕获代表轨迹之间的距离不改变的刚性体而且还可以捕获关节式对象的形状变化等。

[0213] 接下来,分割单元 1101 使用计算的代表轨迹之间的距离,基于作为连接关系的代表轨迹之间的相关性来生成 MST。图 26(a) 示出在生成 MST 之前由表达式 11 表示的代表

轨迹之间的连接关系,将其作为曲线图。在曲线图中,每条代表轨迹 1801 被显示为节点,并且各条代表轨迹 1801 之间的关系和轨迹之间的全部其它距离 1802 被表示为由虚线所示的边。应该注意:每条边(虚线)拥有代表轨迹之间的距离,并且由将加权作为距离的边来连接全部代表轨迹。另一方面,由于边被生成使得代表轨迹之间距离的和为最小,所以可以利用如图 26(b) 中所示的代表轨迹之间的有限距离构成 MST。在图 26(b) 中,没有连接关系的代表轨迹等同于具有表示为无穷大的代表轨迹的那些代表轨迹。因此,在这种有限连接关系内表示代表轨迹产生了能够减少存储量或者当执行聚类时进一步减少计算量的有益效果。利用 MST,可以使用在非专利文献 8 中描述的方法,因此,这里不重复其详细的描述。

[0214] [非专利文献 8] MIT Press, 第 561–579 页, 作者为 Thomas H. Cormen、Charles E. Leiserson、Ronald L. Rivest 和 Clifford Stein 的文章“Introduction to Algorithms”

[0215] 接下来,通过从包括如图 26(b) 所示的代表轨迹之间的这种有限距离的 MST 中,选择一对代表轨迹,并且执行用于确定组成该对的代表轨迹是否属于同一类别的处理,分割单元 1101 对代表轨迹执行聚类,其中该对代表轨迹具有以下关系,即使用一条边来直接将该对代表轨迹彼此连接。因此,再次,参考图 25,分割单元 1101 通过将被分类为相同的类别的代表轨迹所归属的子区域假设为相同的段来执行分割(步骤 S1702)。

[0216] 在下文中,描述步骤 S1702 的细节。

[0217] 首先,分割单元 1101 从 MST 中选择具有直接连接关系的一对代表轨迹。例如,如图 27(a) 所示,选择当前对 1901。应该注意:这里,描述了示例,虽然在实践中,足以选择具有直接连接关系的所有代表轨迹并且对它们中的每一个执行随后的处理。

[0218] 接下来,分割单元 1101 根据由 MST 表示的连接关系,选择可以从当前对经过 N 条路径能够到达的代表轨迹(节点)。这确定了当执行聚类时应该考虑的邻近关系。图 27(b) 示出其中 N = 2 的邻近关系。实线所表示的部分是在执行聚类时应当被考虑的代表轨迹。图中所示出的诸如“1”和“2”等数字对应于从所选择的节点开始的路径的数量,其中所选择的节点为当前对。这里,N 可以被预先确定,或者可以预先确定要被考虑的代表轨迹的数量。与诸如树状图之类的成对聚类相比较,可以通过考虑在周围的代表轨迹,执行不仅考虑最接近的邻域而且还考虑周围的信息的聚类,在树状图中,通过考虑仅邻域中的最接近的两个点来执行聚类。因此,这产生了以下有益效果,即能够更好地反映图像结构并且执行不易受诸如图像中的对象的尺寸差异等影响的分割。

[0219] 接下来,分割单元 1101 使用当前对和邻域中的代表轨迹来执行聚类。

[0220] 这里,在诸如树状图等成对聚类中,确定是否仅基于关于当前对的信息来执行到类别的结合。因此,对于诸如人等关节式对象,不可能考虑诸如手和脚等空间远离区域之间的关系。然而,通过执行包括邻域的成对聚类,可以考虑包括手和脚之间的关系的区域之间的关系。因此,具有能够精确地甚至对诸如人等关节式对象执行分割的优点。作为另一有益效果,由于还考虑了与空间远离地方的关系,所以可以执行不易受诸如图像中的对象的尺寸改变或差异等影响的聚类。

[0221] 这里,可以通过将用于确定当前对是否属于相同类别的处理应用到所有当前对来自动地确定最终的类别数量。应该注意:可以分两个阶段执行聚类,即预聚类和结合。首先,使用具体的示例来描述预聚类。这里,为了清楚地描述,如替代图 27(c) 的图 28 所示,

假设 A 和 B 是当前对的节点（代表轨迹），而 C 和 D 是除该对外以外的邻近节点（代表轨迹）。应该注意：这不限制要被用于预聚类的节点数量。

[0222] 首先，在预聚类中，在模型选择的框架内确定代表轨迹 A 到 D 的预类别数量，并且基于该预类别数量对代表轨迹 A 到 D 执行预聚类。具体地，可以通过以下的表达式 18 来表示预聚类。

[0223] 数学式 19

[0224]

$$P(y | m) = \int P(y | \vartheta_m, m) P(\vartheta_m | m) d\vartheta_m$$

[0225] （表达式 18）

[0226] 这里， m 表示预类别数量。如图 28(b) 所示，当对代表轨迹 A 到 D 执行预聚类时，具有四种类型的预类别数量 1(A 到 D 全部在同一类别中) 到 4(A 到 D 全部在不同的类别中)。因此，在该情况下， m 采用从 1 到 4 的四种类型的值。

[0227] 然后，数学式 20： $P(y | \vartheta_m, m)$ 表示在每个预聚类中的可能性，其中预类别数量是 m 。这里，例如，在预类别数量 $m = 2$ 的情况下，分割单元 1101 将下面的 7 个预聚类考虑为用于将代表轨迹 A 到 D 分为两个预类别的可能的候选。在实践中，如图 28(b) 所示，当确定预类别数量为 $m = 1$ 到 4 时，分割单元 1101 计算全部可能的预聚类候选。

[0228] 数学式 21

[0229]

$$(\vartheta_m, m = 2) = \left\{ \begin{array}{l} (A - B - C, D) \\ (A - B, C - D) \\ (A - C, B - D) \\ (A - D, B - C) \\ (A - C - D, B) \\ (A - B - D, C) \\ (A, B - C - D) \end{array} \right\}$$

[0230] （表达式 19）

[0231] 这里，“A-B-C”等中的“-”表示相同的预类别关系，并且“，”表示属于不同的预类别。例如在预聚类候选 (A-B-C, D) 的情况下，代表轨迹 A、B 和 C 属于同一类别，而仅 D 属于不同的类别。在该情况下，

[0232] 数学式 22： $P(y | \vartheta_m, m)$ 由如下表达式 20 来表示。

[0233] 数学式 23

[0234] $P(y | \vartheta_m, m) = P_{AB} \cdot P_{BC} \cdot P_{AC} \cdot (1 - P_{AD}) \cdot (1 - P_{BD}) \cdot (1 - P_{CD})$ （表达式 20）

[0235] 其中 $m = 2$ 。

[0236] 因此，分割单元 1101 使用在相同预类别情况下的 P 并且使用在不同类别情况下的 $(1-P)$ 来计算预聚类的可能性。同样地，在 $m = 2$ 的情况下，分割单元 1101 针对由表达式 19 表示的所有预聚类候选来计算如表达式 20 表示的可能性。此外，在 $m = 1, 3$ 和 4 的情况下，分割单元 1101 针对在图 28(b) 中所示的所有预聚类候选来计算如由表达式 20 表示的预聚类的可能性。

[0237] 然后，数学式 24： $P(\vartheta_m | m)$ 可以根据在预类别数量为 m 情况下的预聚类候选的数量

Z_m 而被表示如下。例如,在代表轨迹 A 到 D 中, $Z_m = 7$, 其中 $m = 2$ 。

[0238] 数学式 25

[0239] $P(\theta_m | m) = 1/z_m$ (表达式 21)

[0240] 然后,可以利用正态分布和表达式 11 中的代表轨迹之间的距离 f 来如下所示的那样计算表达式 20 中的 P_{RS} 。

[0241] 数学式 26

$$[0242] P_{RS} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(f(R,S))^2}{2\sigma^2}\right)$$

[0243] (表达式 22)

[0244] 不言自明的是,可以使用表达式 17 中表示的测地距离 g 来代替代表轨迹之间的距离 f 。如上所述,可以使用表达式 18 到表达式 22 中的信息来根据表达式 18 执行积分计算。

[0245] 然后,分割单元 1101 通过使用针对每个预类别数量 m 而计算的 $P(y|m)$ 中的最大预类别数量,从由表达式 19 表示的预聚类候选中选择预聚类候选,使得数学式 27: $P(y|\theta_m, m)$ 变为最大。这里,如图 28(b) 中所示,可以从全部预聚类候选中选择预聚类候选,使得数学式 28: $P(y|\theta_m, m)$ 变为最大,或如之前描述地,可以计算 $P(y|m)$ 变为最大时所用的预类别数量 m 以便在预类别数量 m 的情况下选择预聚类候选,使得数学式 29: $P(y|\theta_m, m)$ 变为最大。

[0246] 在聚类时,通常,存在由称为过拟合引起的概括性能损失的问题,该过拟合是作为通过增加类别数量而增加数据表达能力的结果所致的对数据的过度拟合 (excessive fitting)。通过使用上述方法,可以在不过度地增加预类别数量的情况下自动地确定预类别数量 m ,由此产生了获得高度概括的有益效果。

[0247] 接下来,描述结合。这里,在当前对 2001 被确定为属于相同的预类别时,作为当前对的代表轨迹 A 和代表轨迹 B 被结合。

[0248] 作为具体的示例,假设在 $m = 2$ 的情况下,根据表达式 18, $P(y|m)$ 被计算为最大值。接下来,假设:在 $m = 2$ 的情况下并且当 (A-B-C, D) 是预聚类候选时,数学式 30: $P(y|\theta_m, m)$ 根据表达式 20 被计算为最大值。在该情况下,由于当前对 A 和 B 属于相同的预类别,所以轨迹 A 和 B 被确定为属于同一类别,并且因此轨迹 A 和 B 被结合。例如,同样地假设:在 $m = 2$ 的情况下并且当 (A, B-C-D) 是预聚类候选时,数学式 31: $P(y|\theta_m, m)$ 根据表达式 20 被计算为最大值。在该情况下,由于属于当前对 2001 的代表轨迹 A 和 B 属于不同的预类别,所以代表轨迹 A 和 B 不被结合。

[0249] 如上所述,分割单元 1101 确定是否对属于当前对 2001 的代表轨迹进行结合。分割单元 1101 通过对由 MST 彼此直接连接的所有对执行这种确定来结合代表轨迹。据此,可以最终获得如在图 28(c) 中所示的输出。图 28(c) 示出获得的确定结果,即连接人和背景的当前对被确定为属于不同的预类别,并且人中的当前对和背景中的当前对中的每一对属于相同的预类别。然后,通过结合被确定为属于相同预类别的代表轨迹和代表轨迹所归属的子区域,可以执行如在第二实施例中的时空分割。

[0250] 而且,基于由分层子区域生成单元生成的子区域的包含关系,可以对表达式 22 中的代表轨迹之间的距离 f 执行加权。更具体地,在由分层子区域生成单元生成的特定两个区域包括在处于较高分层级的相同子区域(属于相同的子区域)的情况下,可以通过

将两个子区域的代表轨迹之间的距离 f 乘以小于 1 的系数来获得相对小的距离。另一方面，在特定两个子区域没有包括在处于较高分层级的相同子区域（属于不同的子区域）的情况下，可以通过将两个子区域的代表轨迹之间的距离 f 乘以等于或大于 1 的系数来获得相对大的距离。因此，即使在代表轨迹包括误差的情况下，也可以通过反映基于图像轮廓的空间结构来执行更鲁棒的分割。

[0251] 应该注意：使用本变型中的方法产生了以下有益效果，即能够在既不需要预先设置类别数量的情况下也不限制本方法所应用的场景的情况下执行分割。此外，由于预聚类不仅考虑最接近的邻域，而且还考虑周围的代表轨迹，所以还可以执行不受类别尺寸的大小影响的聚类。因此，即使当大对象和小对象同时出现时或者当具有大运动的对象和具有小运动的对象同时出现时，也可以可靠地执行分割。如上所述，可以以高精确度执行分割。

[0252] 至此，已经基于实施例及其变型描述了根据本发明实施方式的轨迹估计方法和轨迹估计装置，以及分割方法，但是本发明不限于这些实施例和变型。本领域的技术人员将容易地理解，可以在示例性实施例中进行许多修改，而没有脱离本发明的新颖性教导和优点。因此，所有这样的修改将被包括在本发明的范围内。

[0253] 此外，本发明还包括由各个实施例中的特征组成元素的任意组合所实现的实施例。

[0254] 工业应用性

[0255] 本发明特别地适用于估计诸如移动而改变形状的人等对象的轨迹以及执行时空分割的图像处理装置。特别地，本发明被应用为能够并入到运动分析装置、监视装置或摄像机、或诸如 TV 装置等音频视频设备中的图像处理设备等。

[0256] 附图标记

[0257] 101 图像接收单元

[0258] 102 运动分析单元

[0259] 103 分层子区域生成单元

[0260] 104 代表轨迹估计单元

[0261] 105、105a 输出单元

[0262] 1101 分割单元

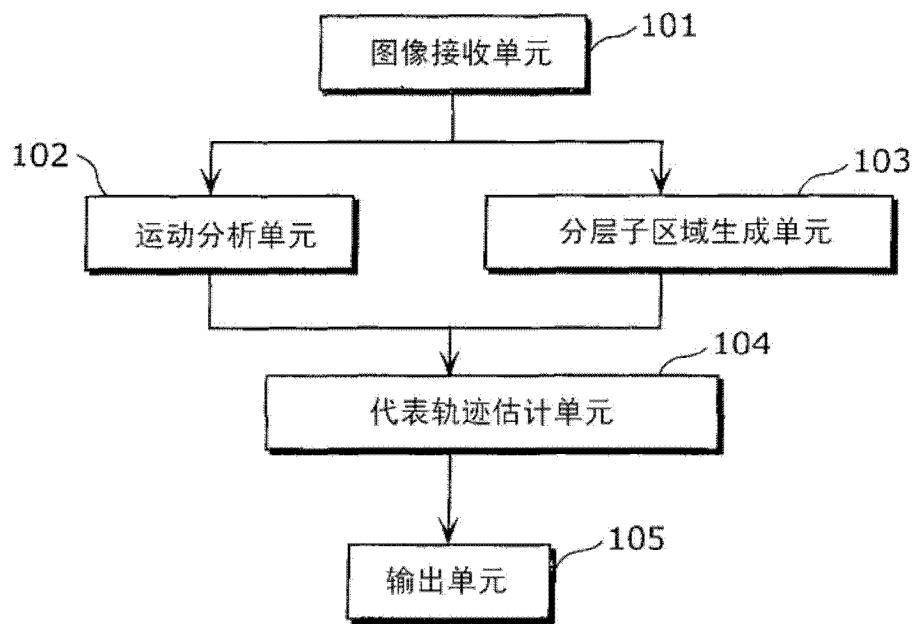


图 1

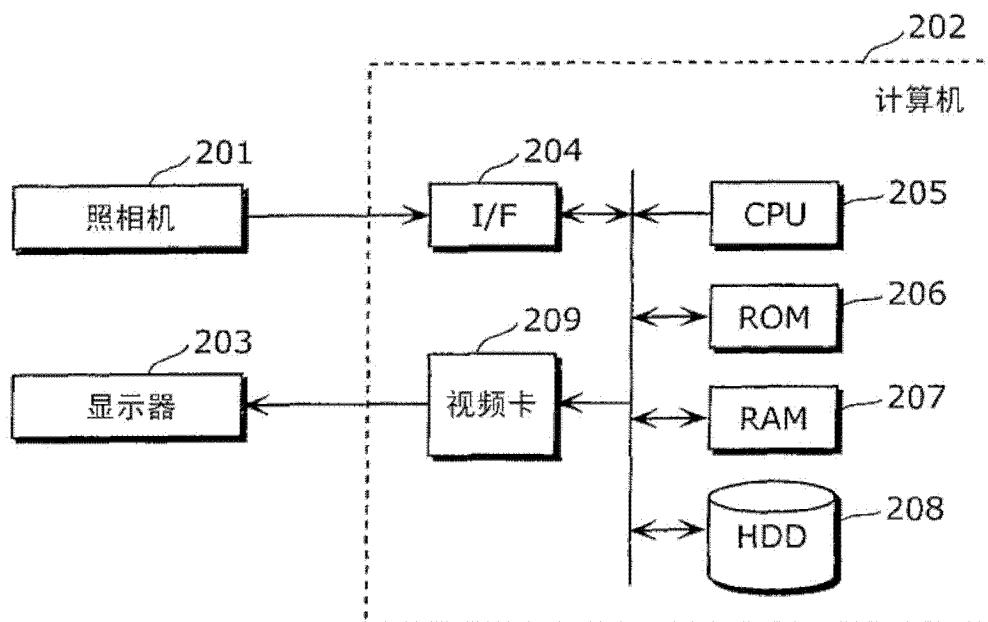


图 2

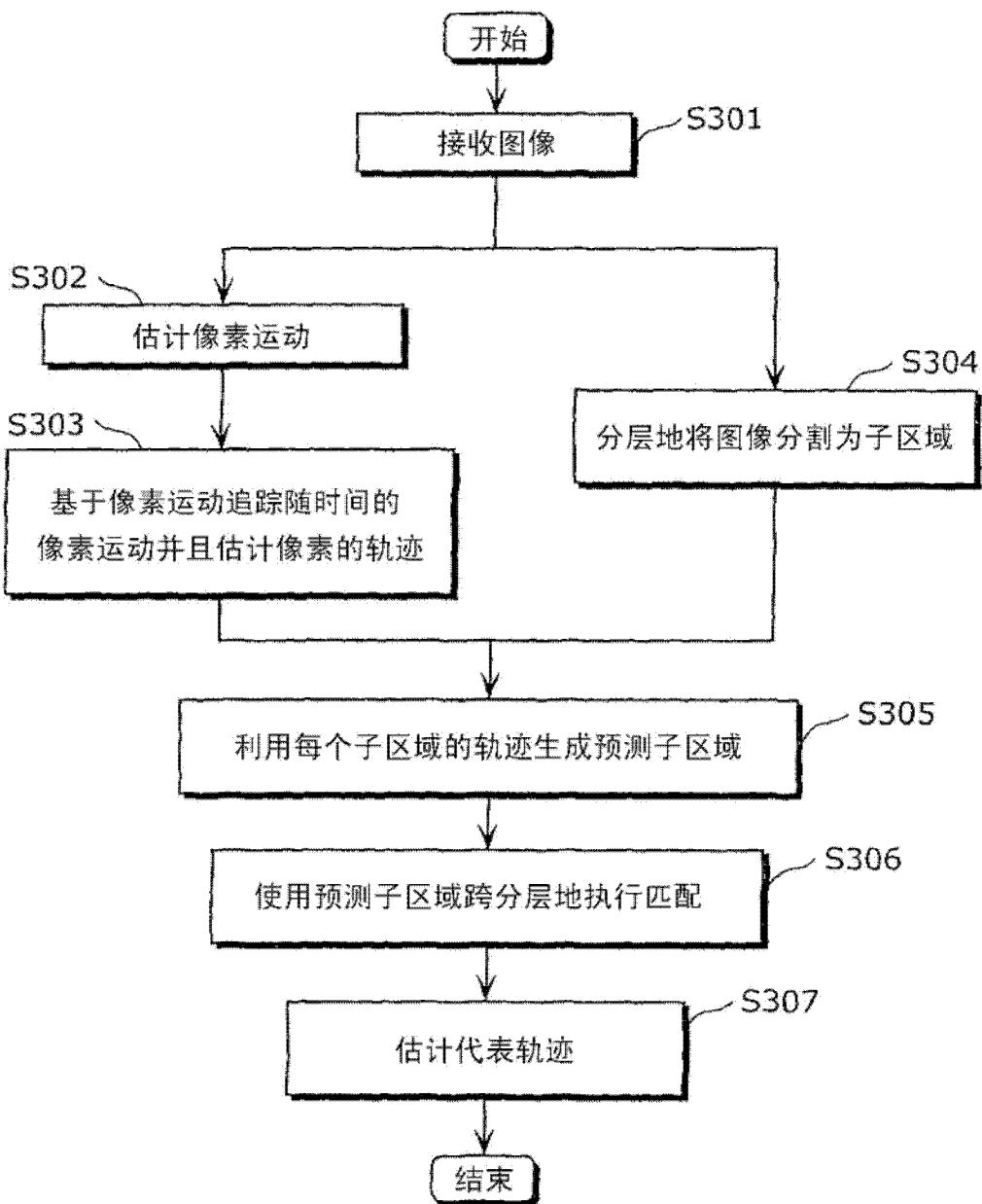


图 3

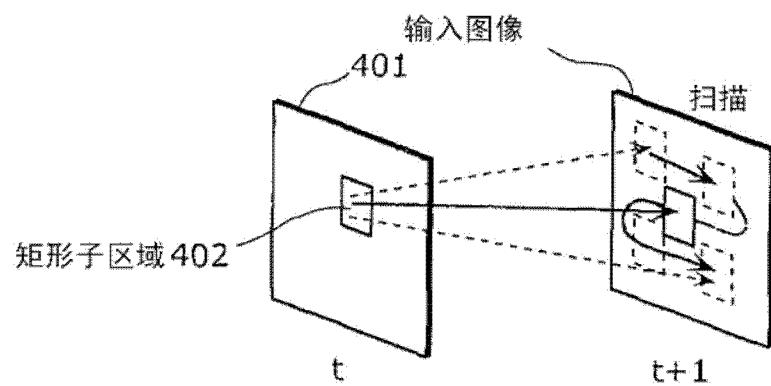


图 4

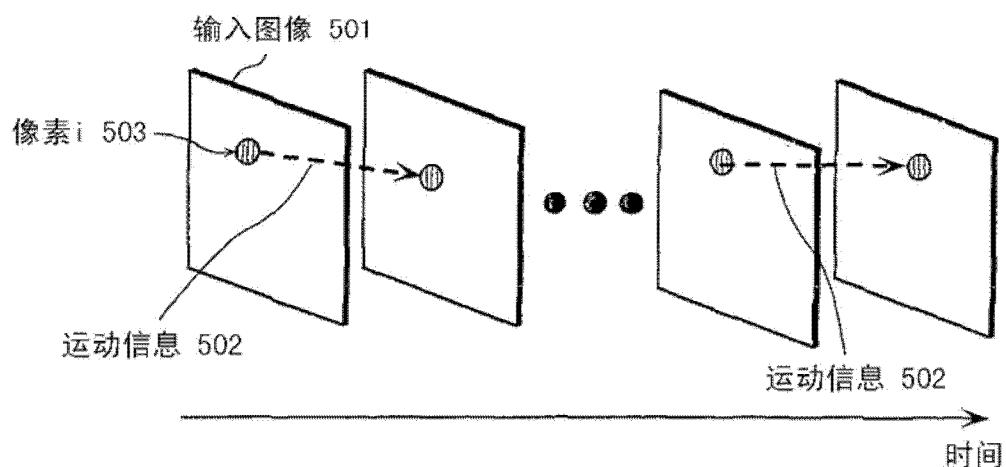


图 5

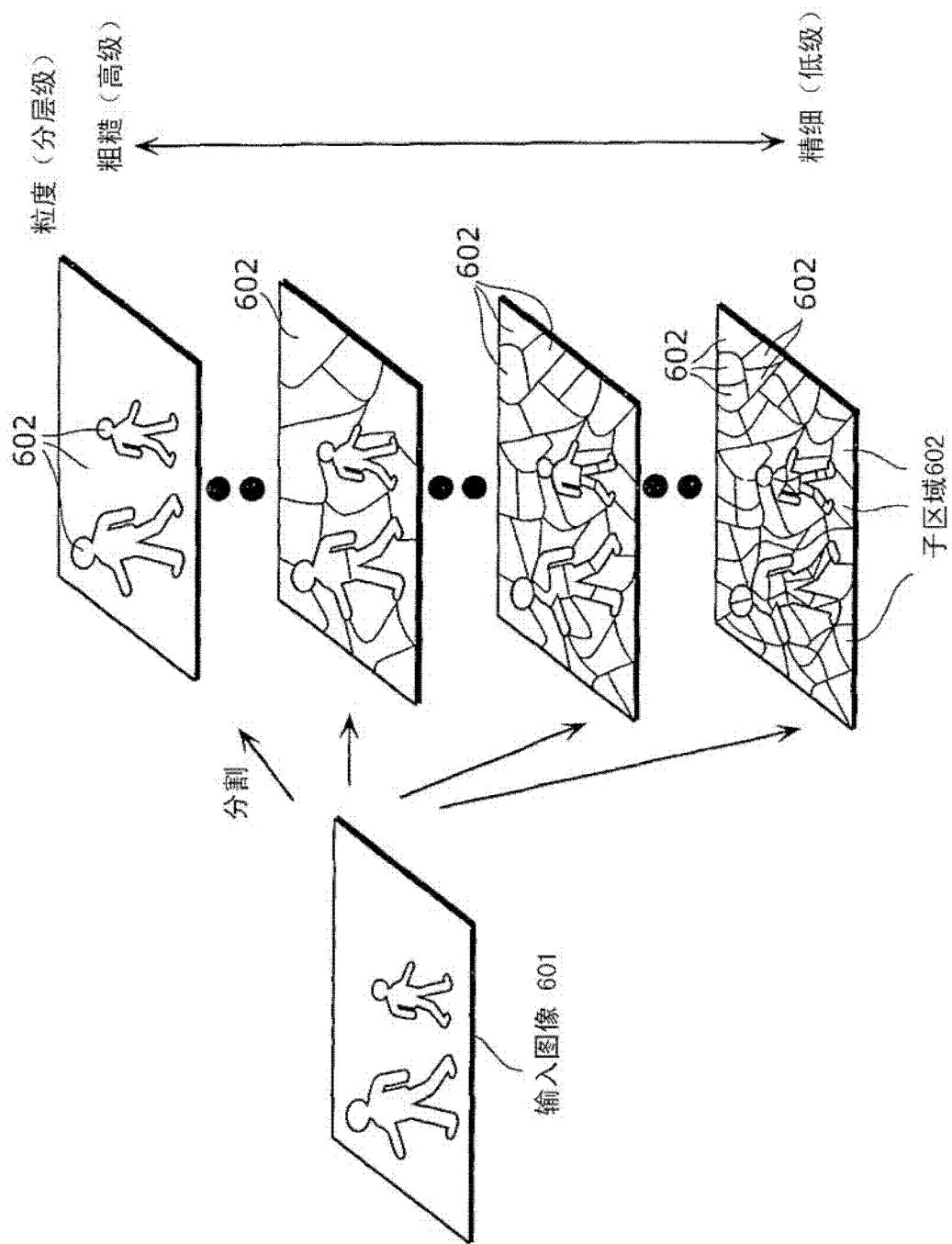


图 6

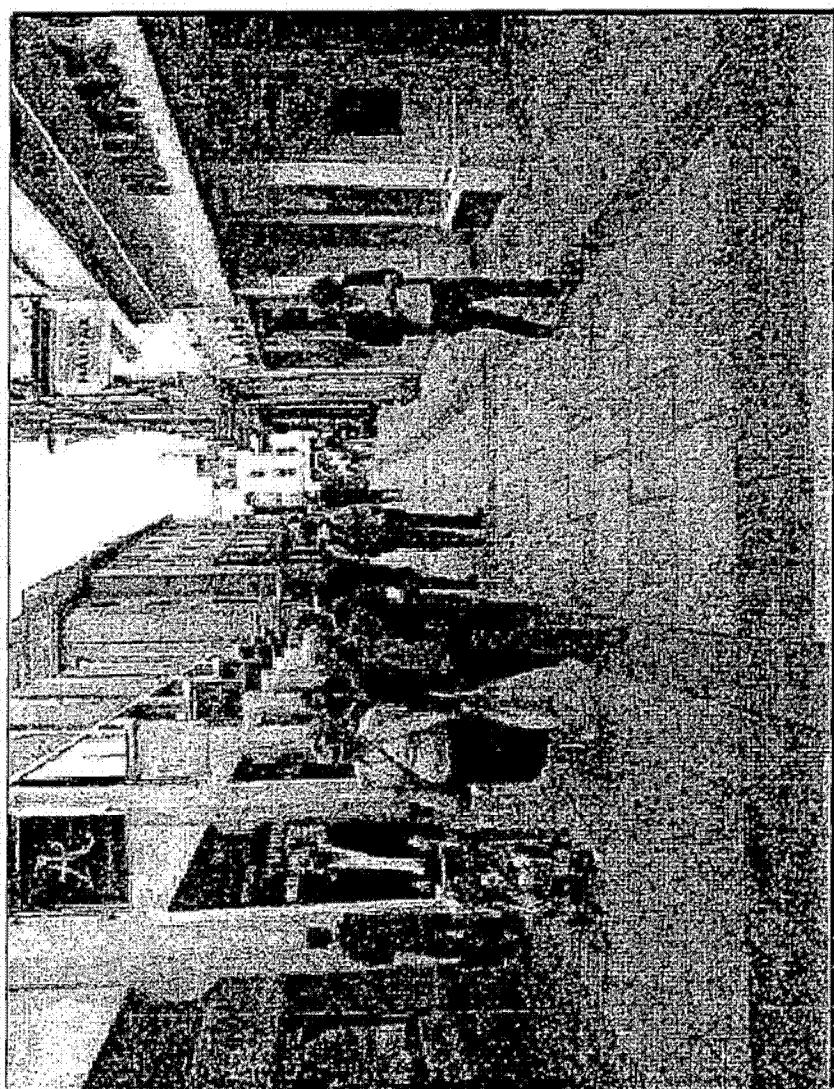


图 7

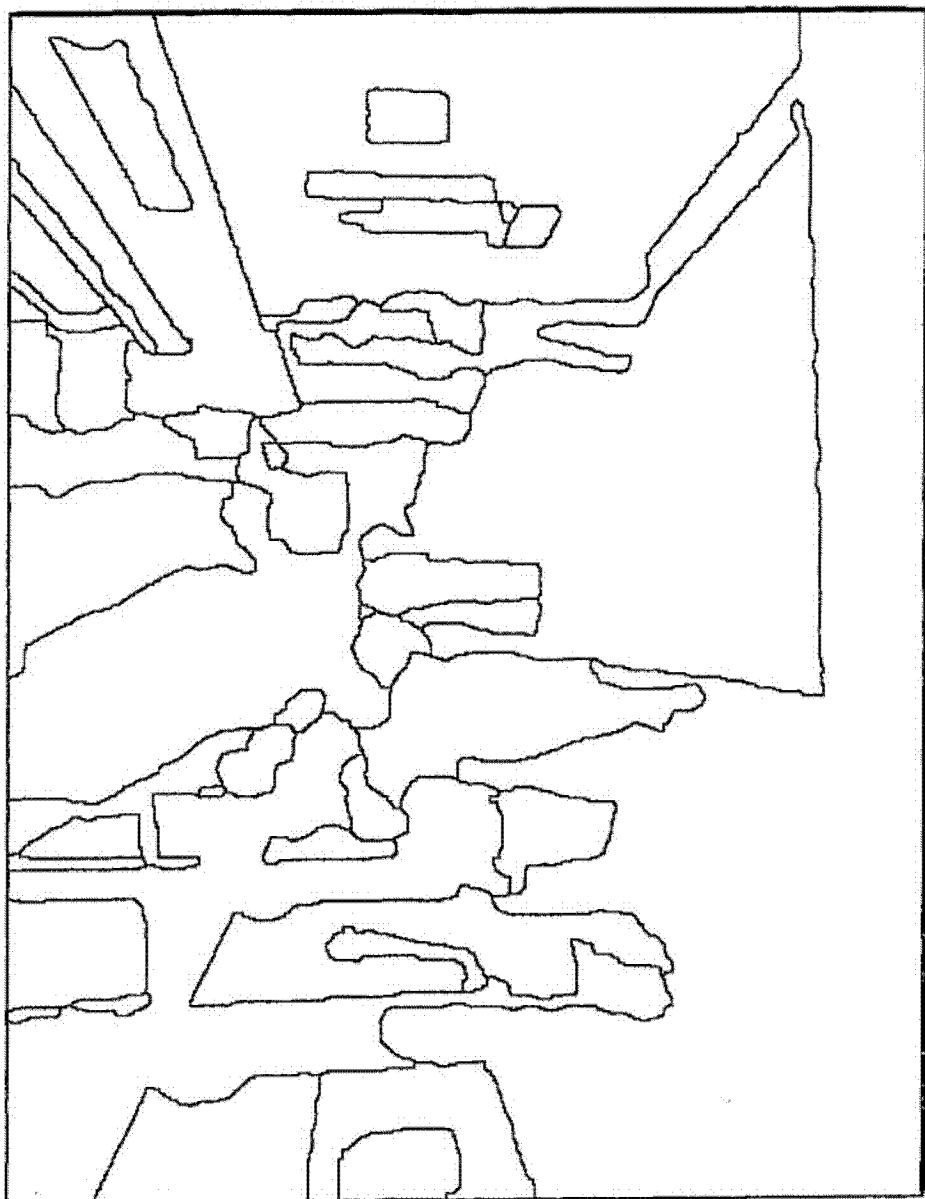


图 8

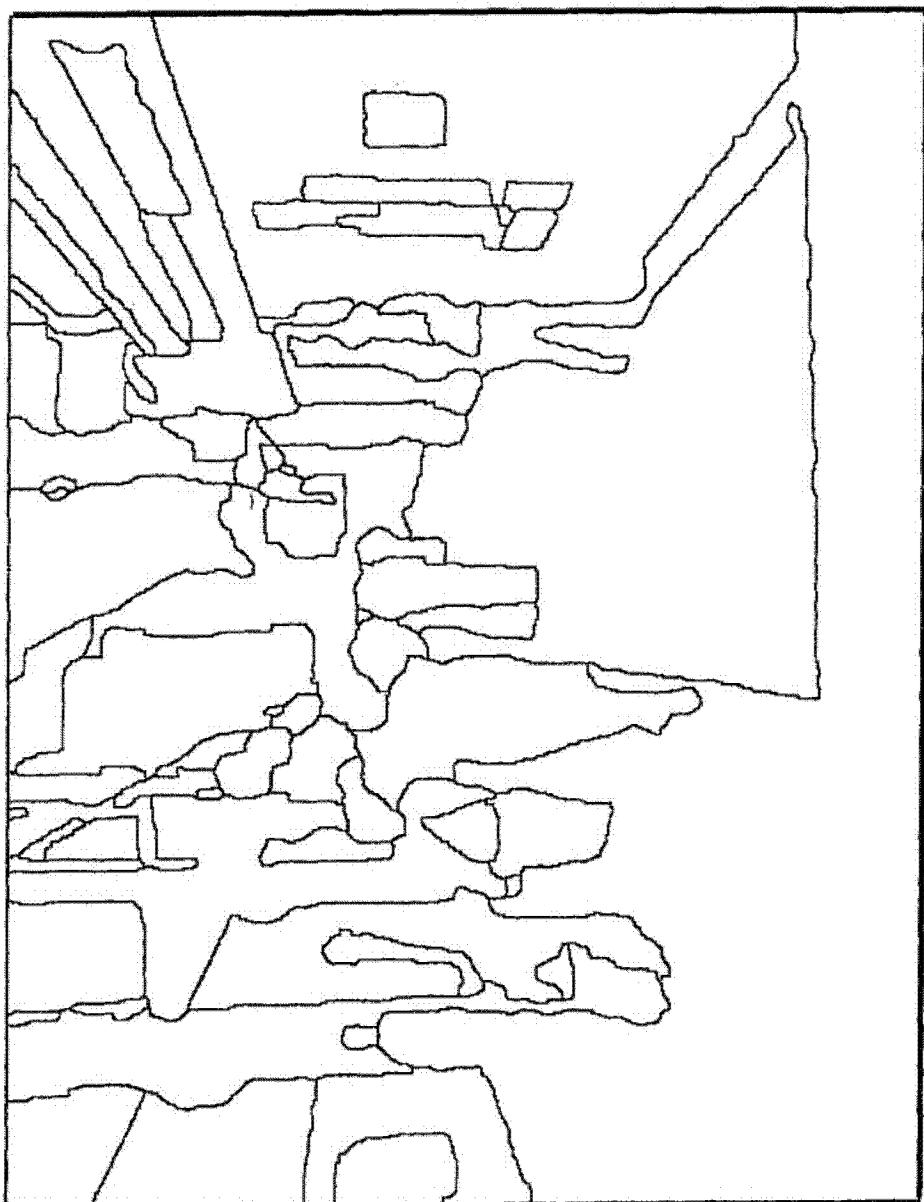


图 9

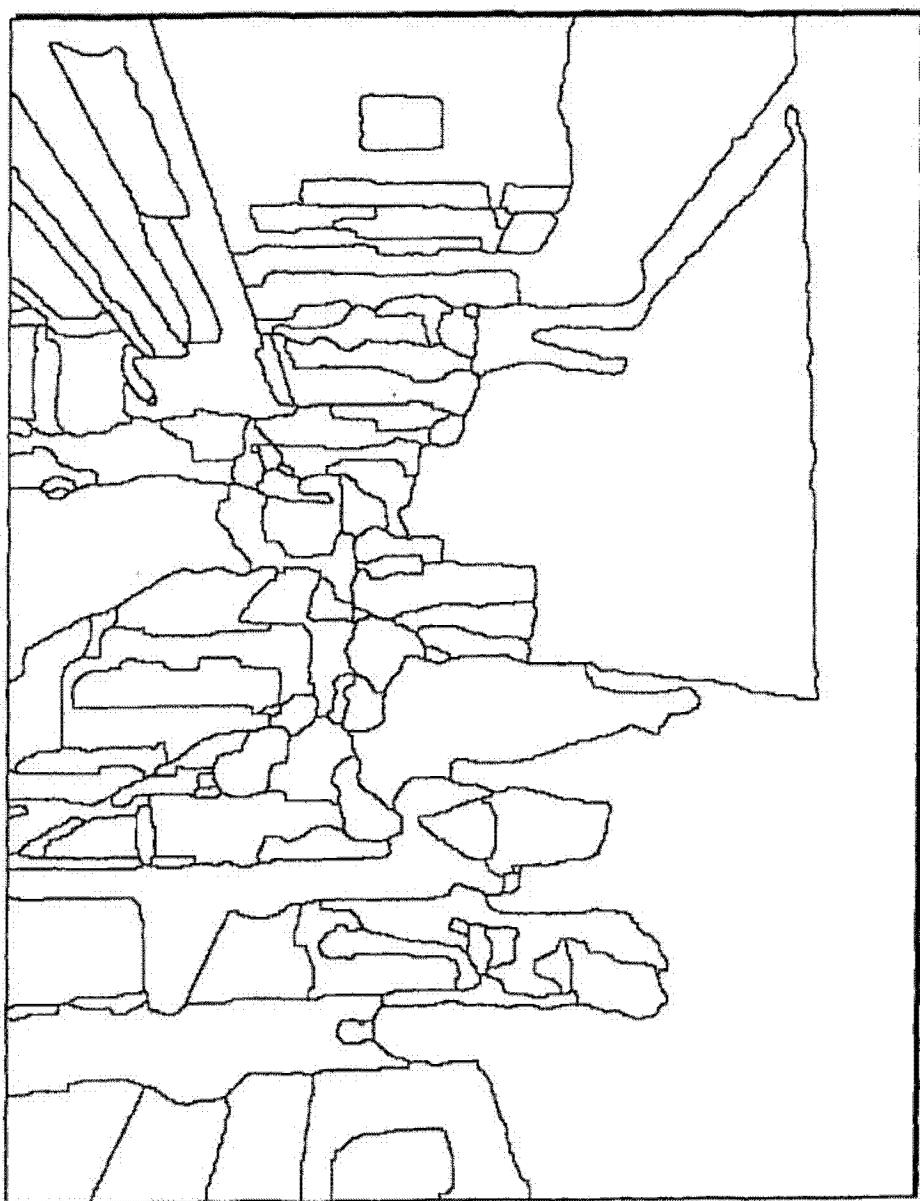


图 10

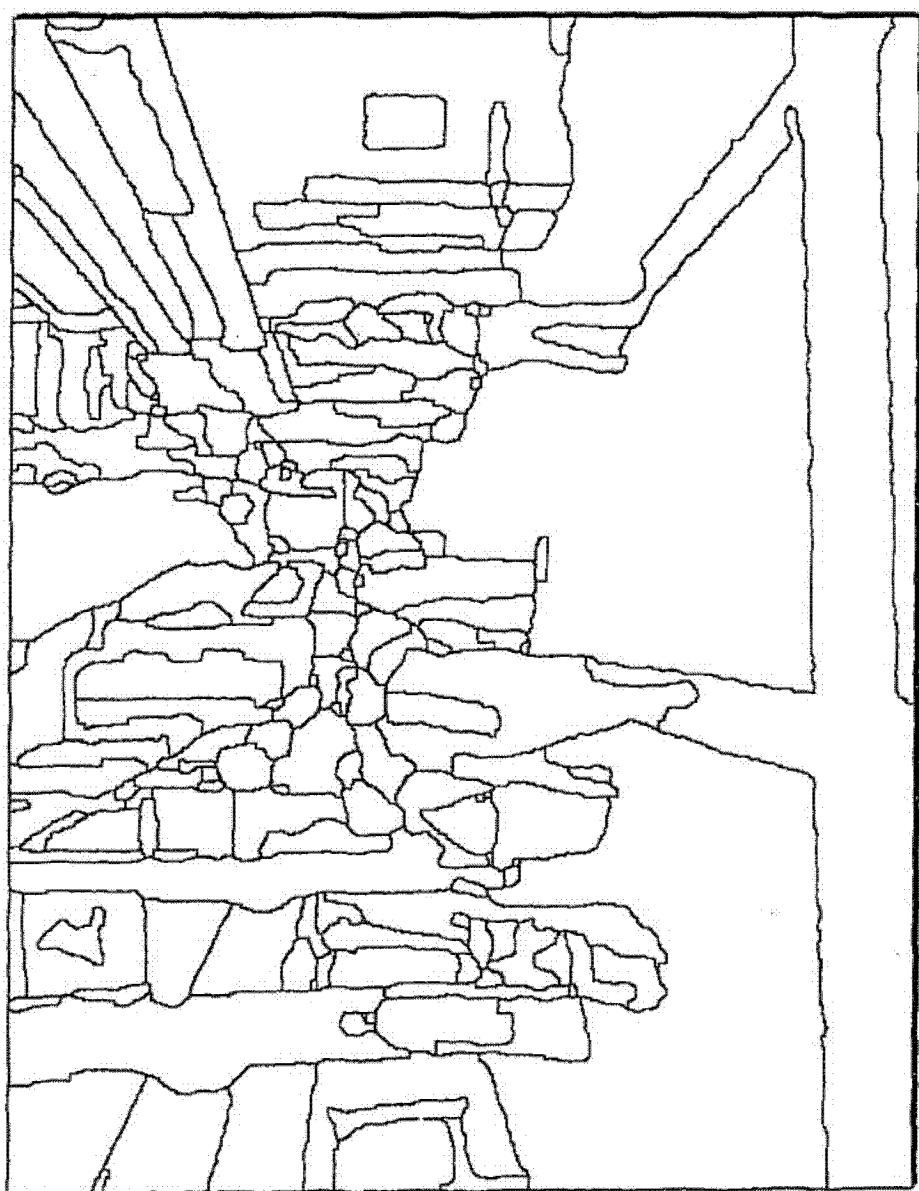


图 11

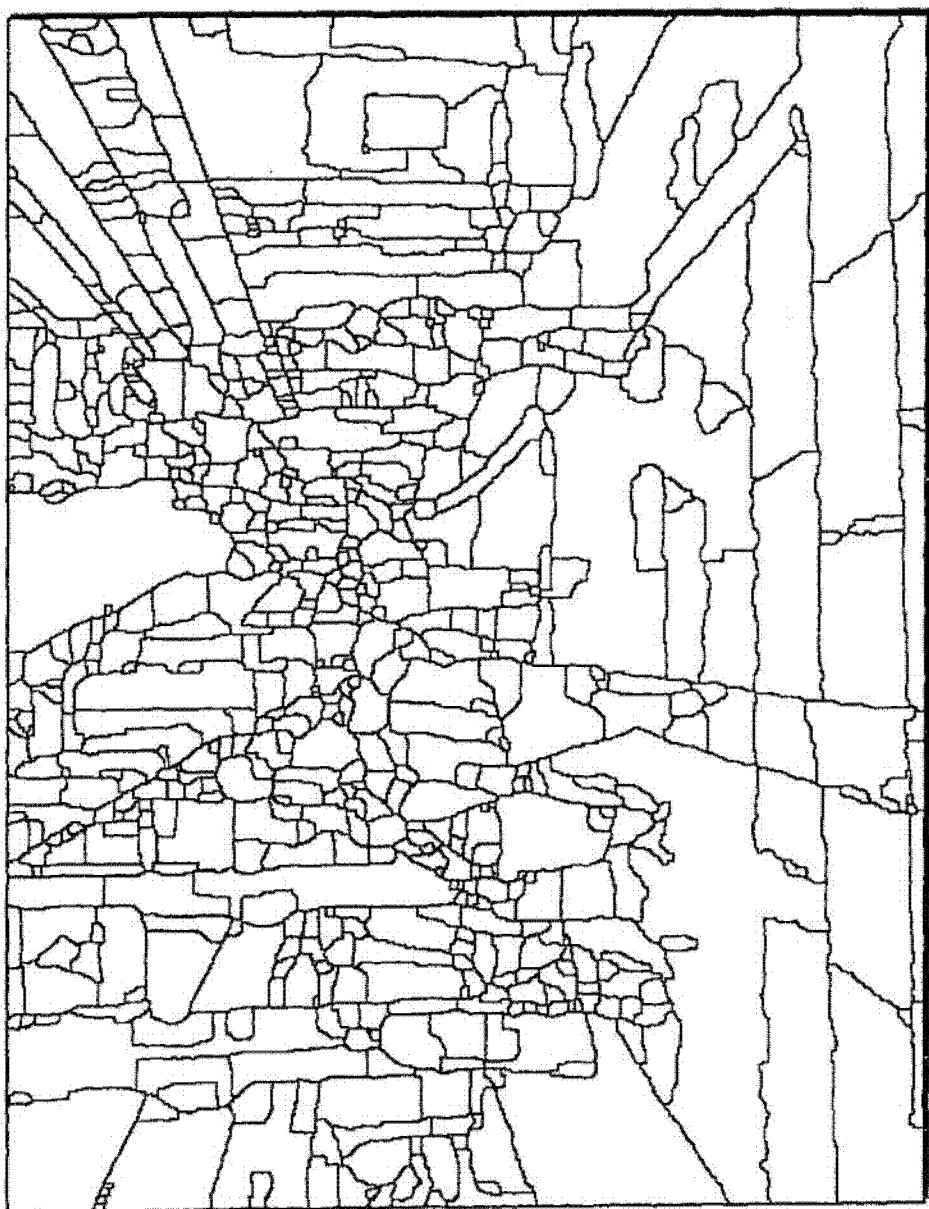


图 12



图 13

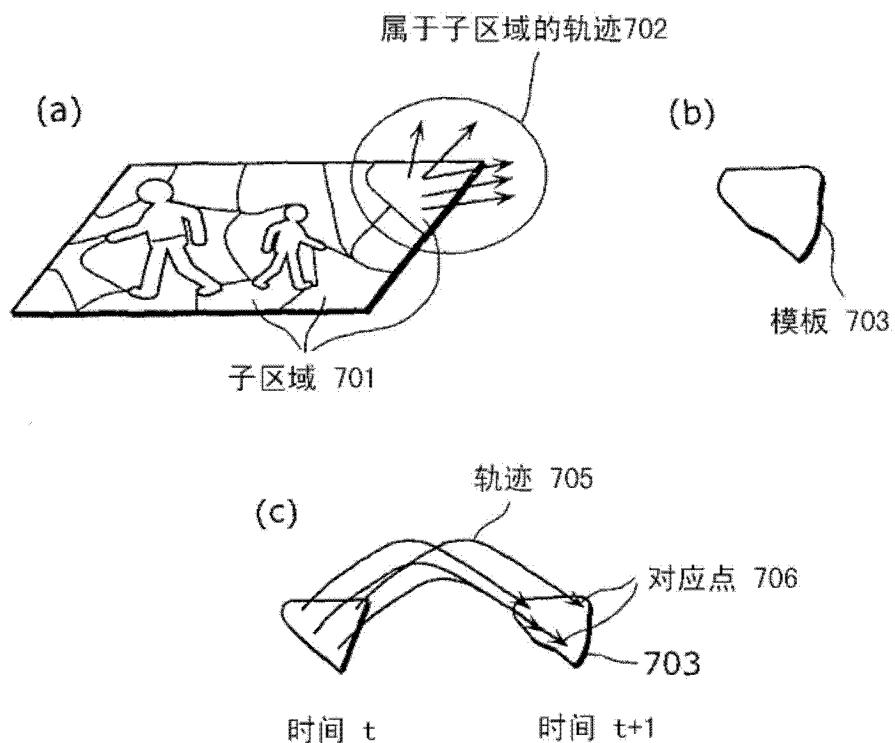


图 14

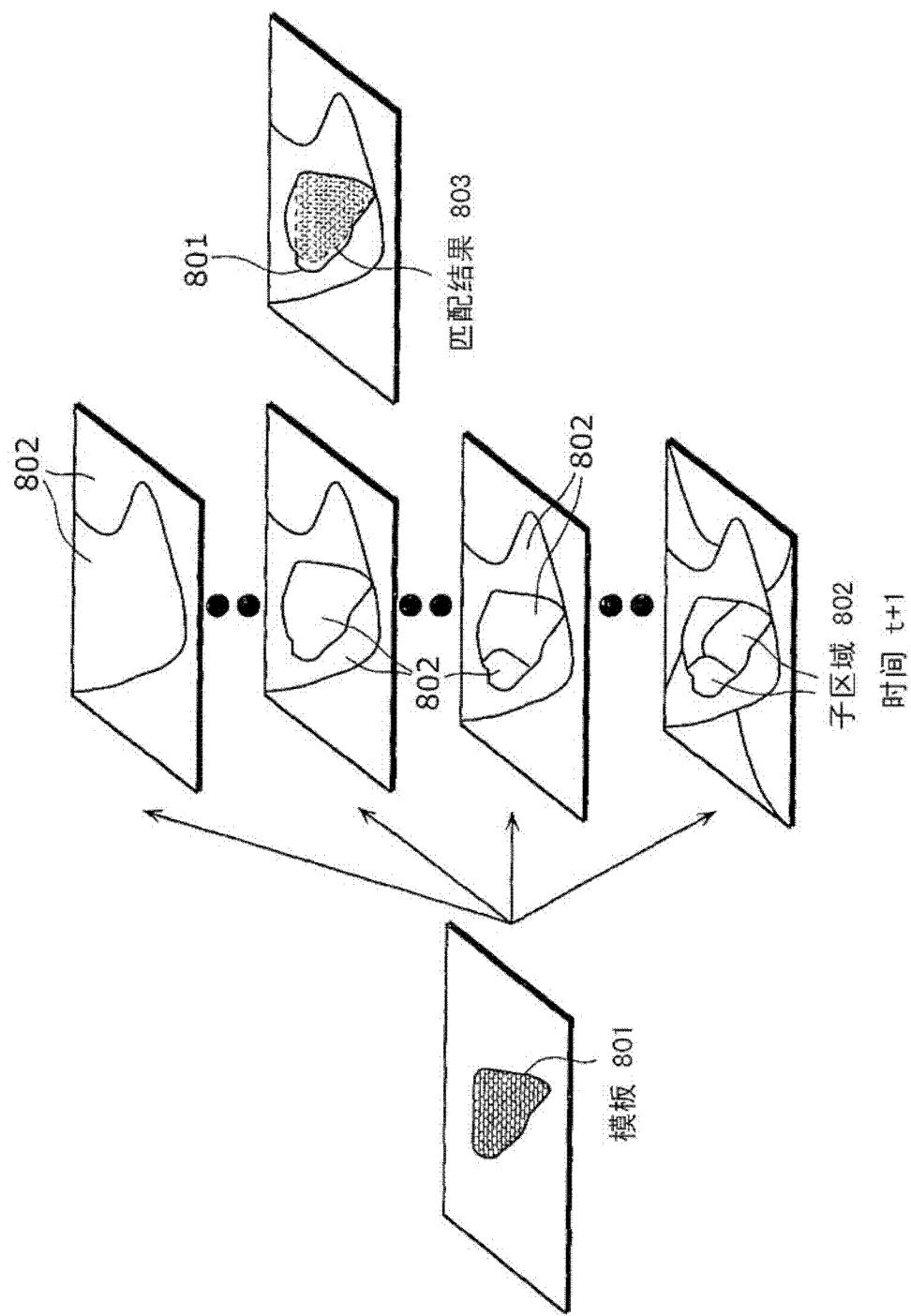


图 15

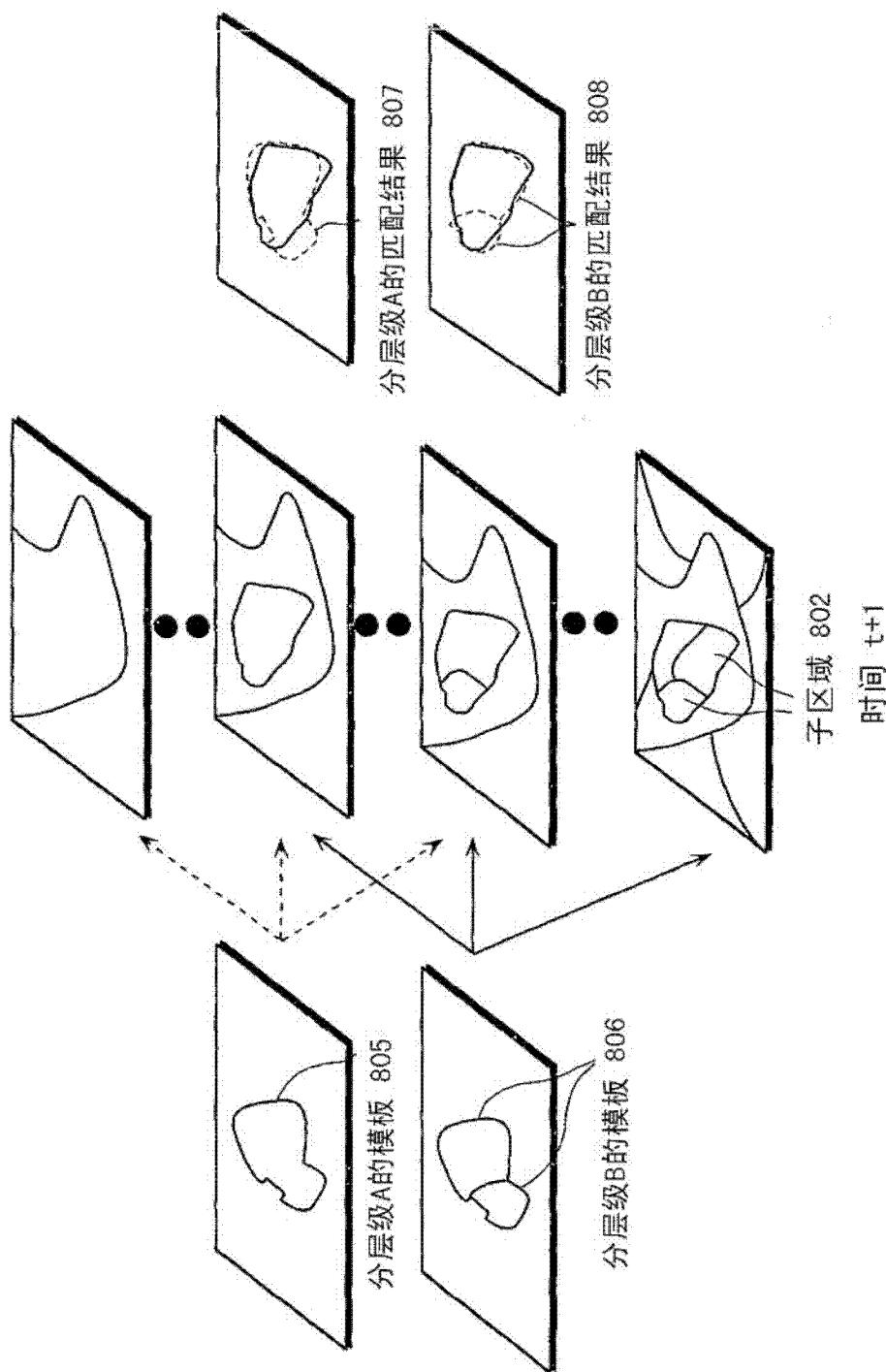


图 16

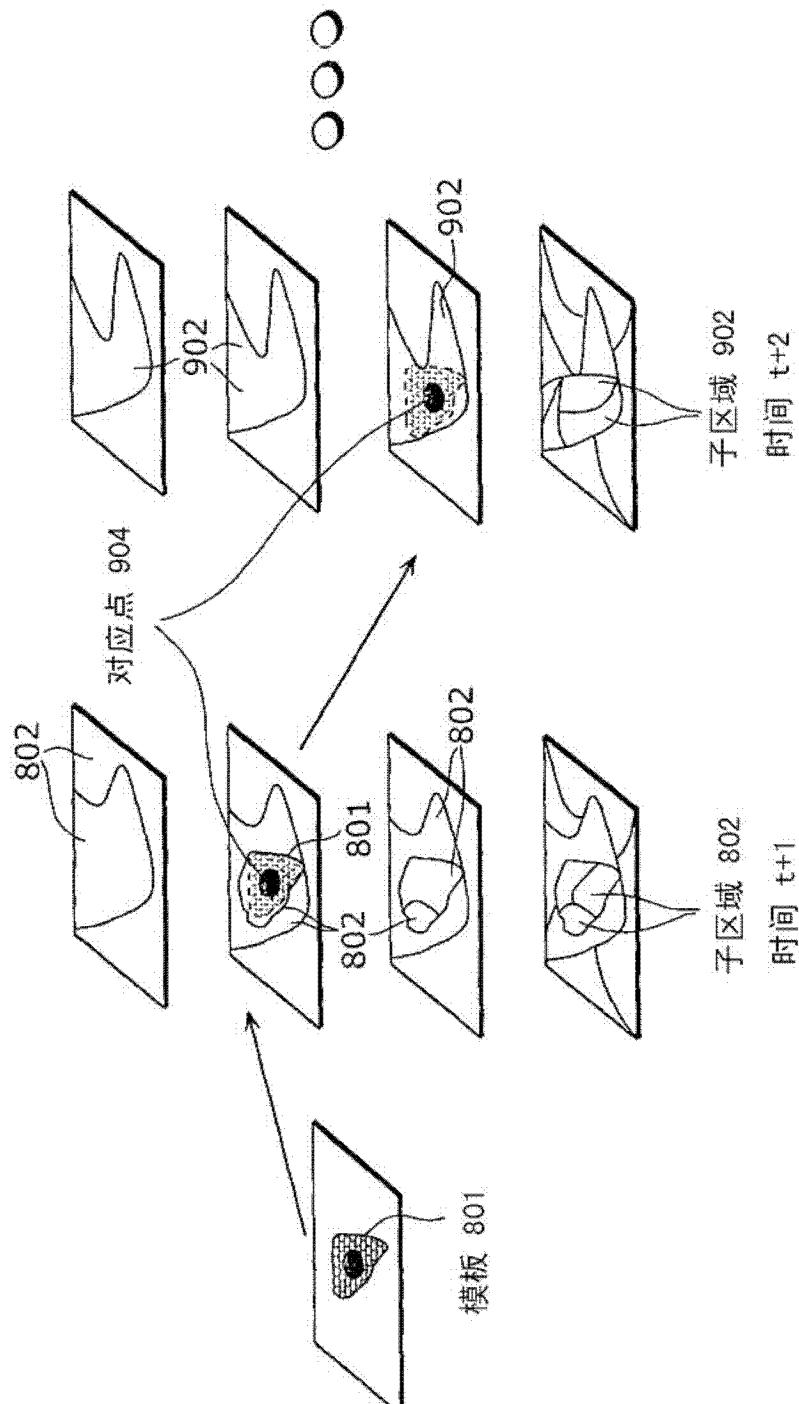


图 17

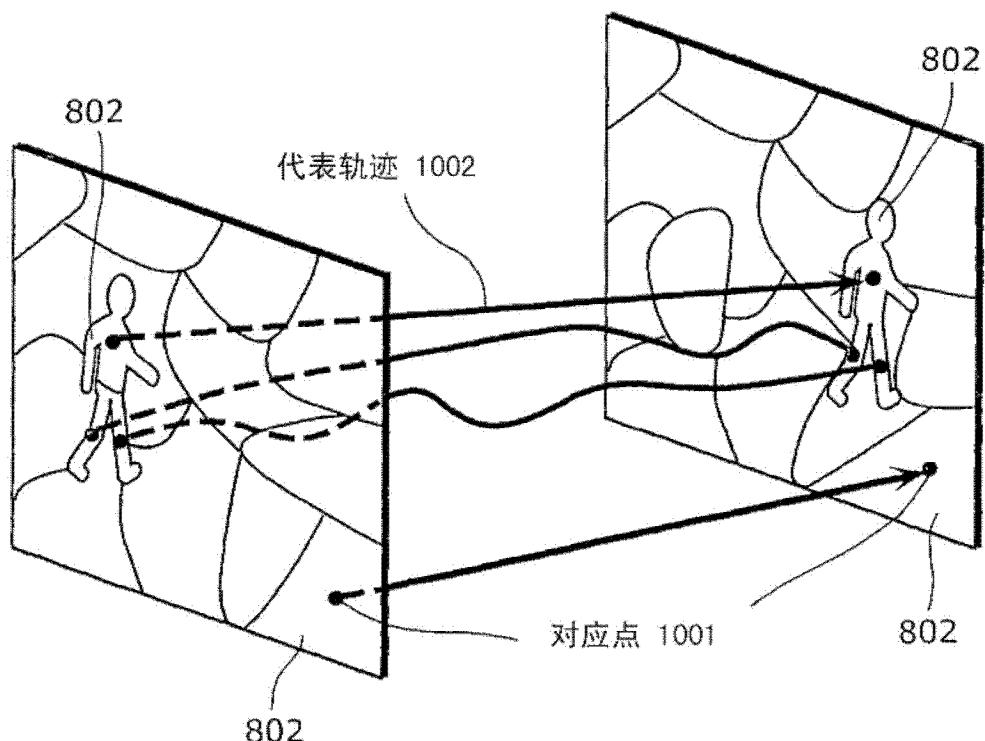


图 18

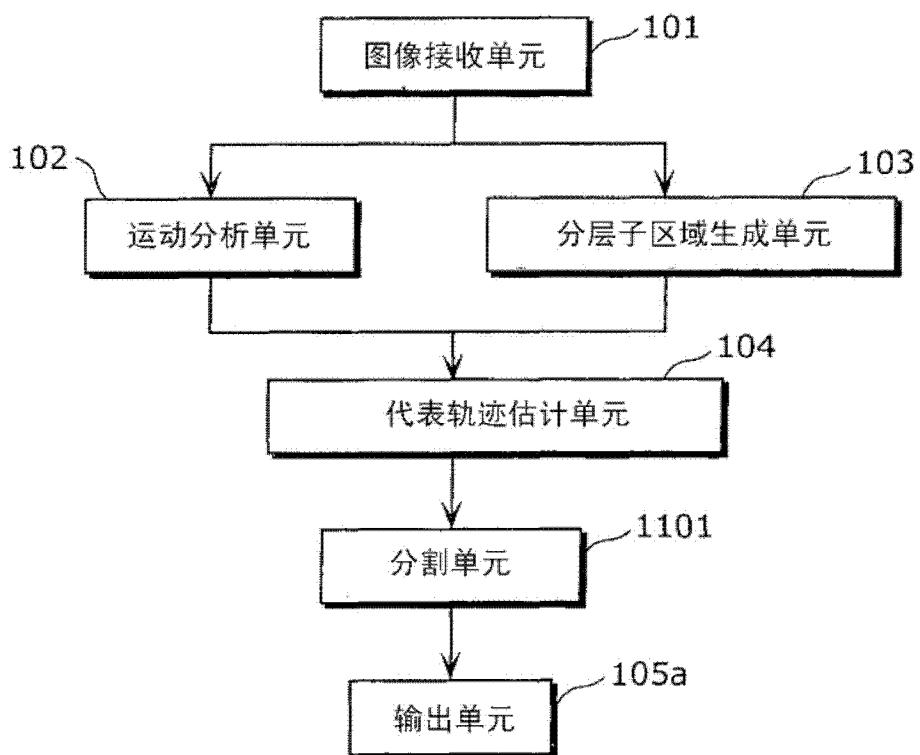


图 19

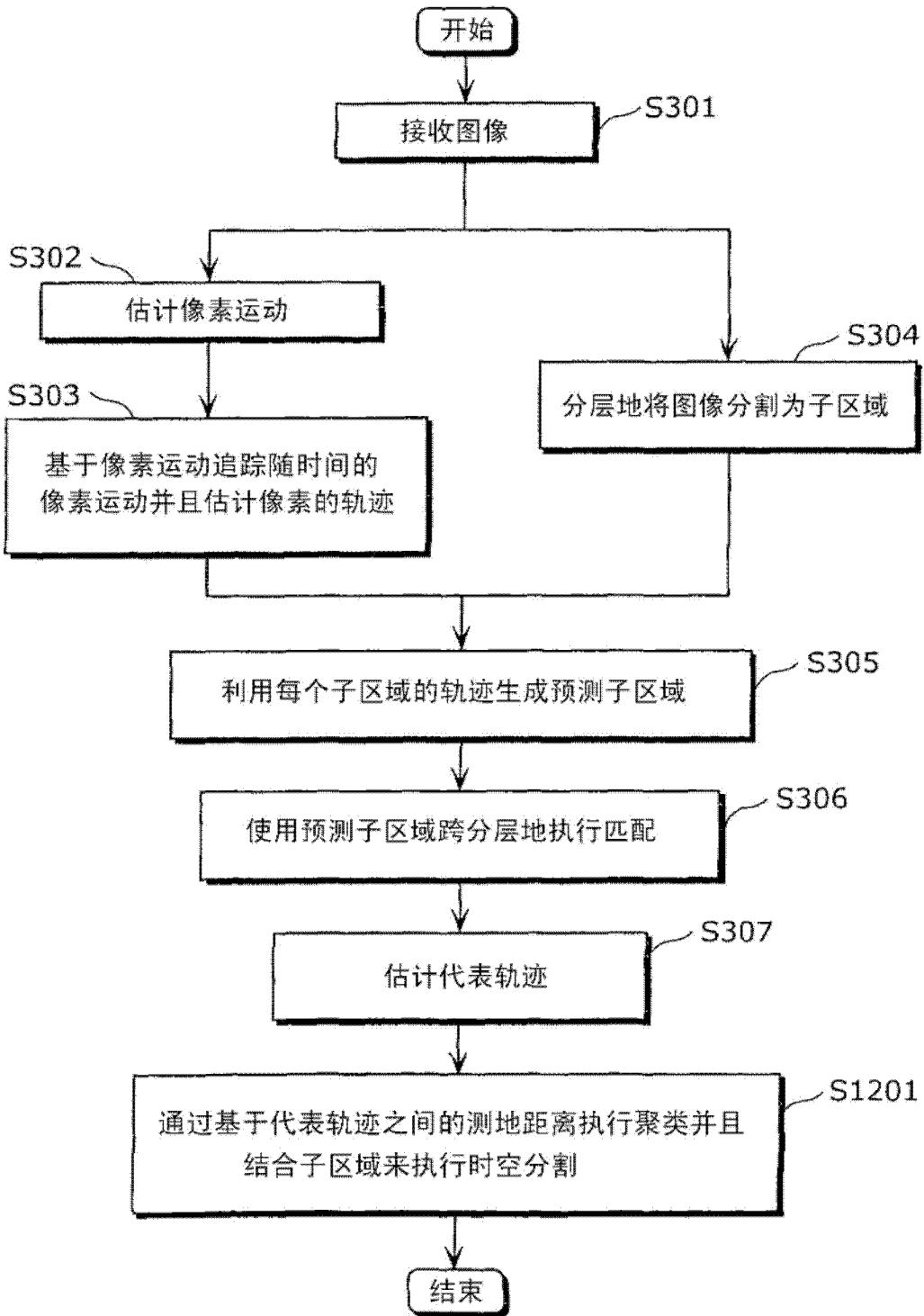


图 20

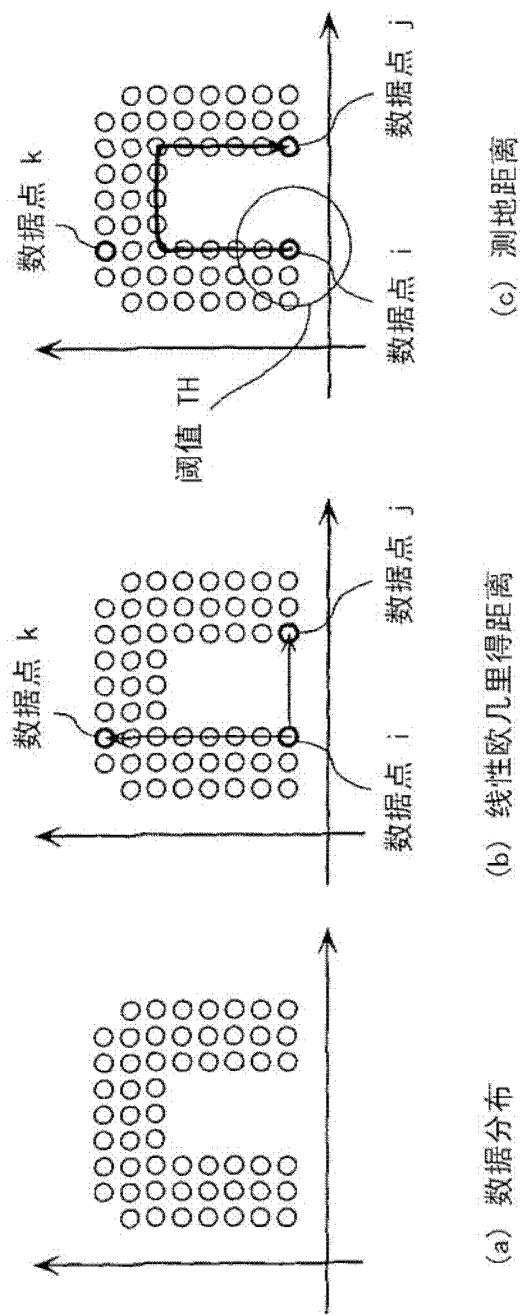


图 21

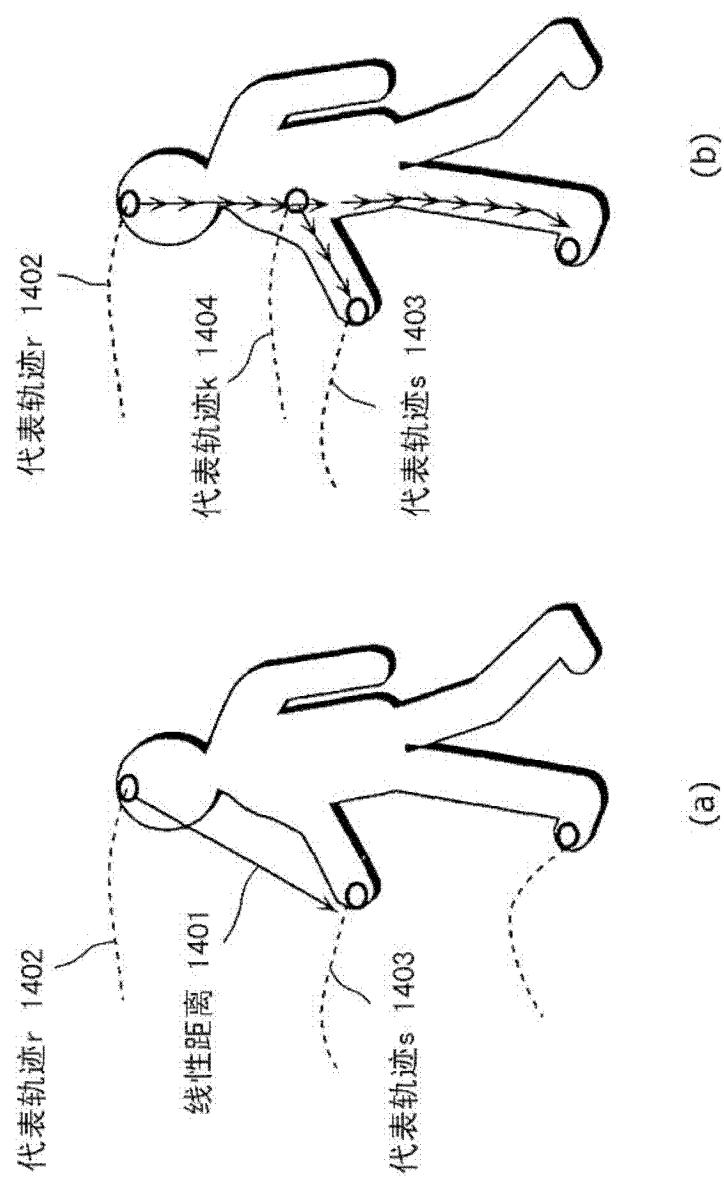


图 22

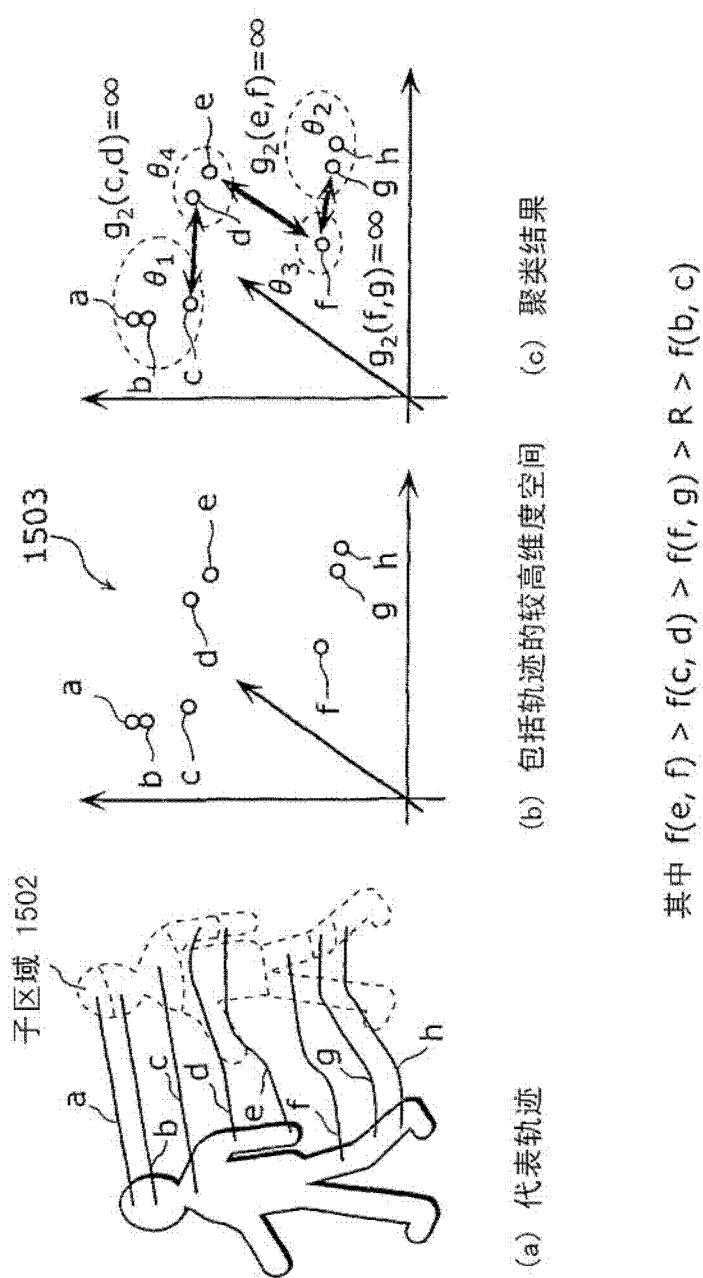


图 23

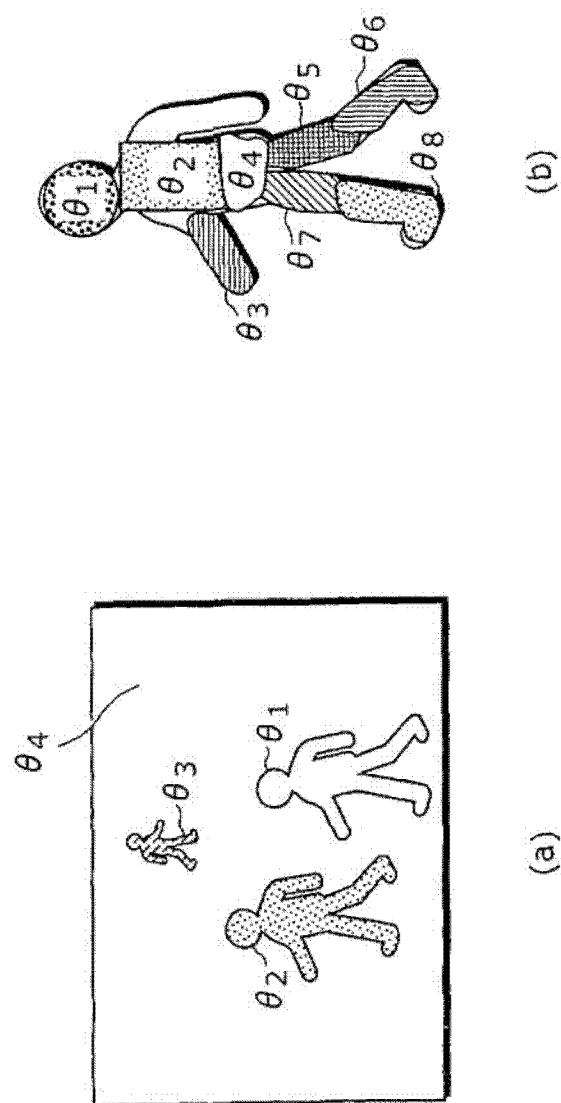


图 24

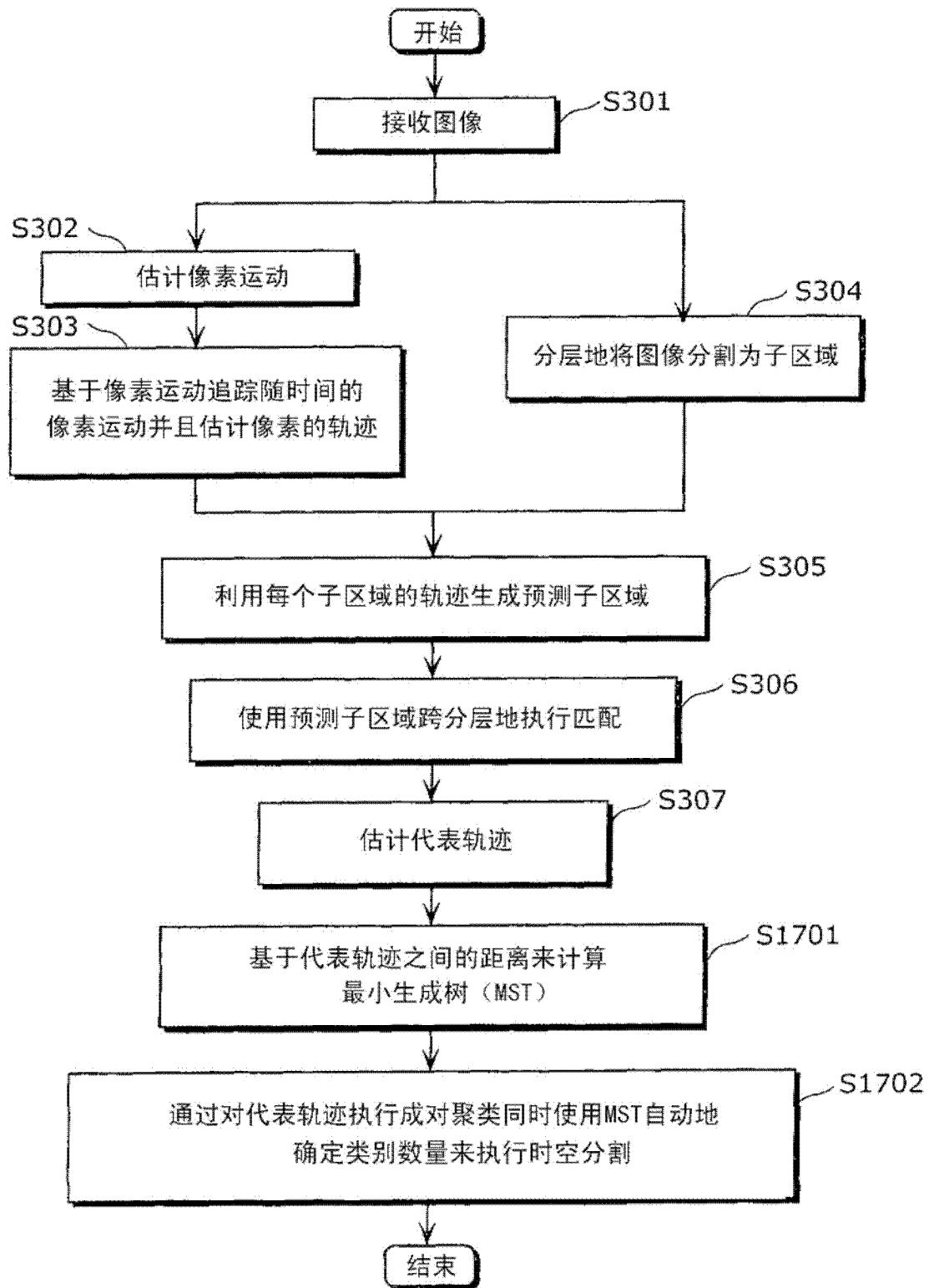


图 25

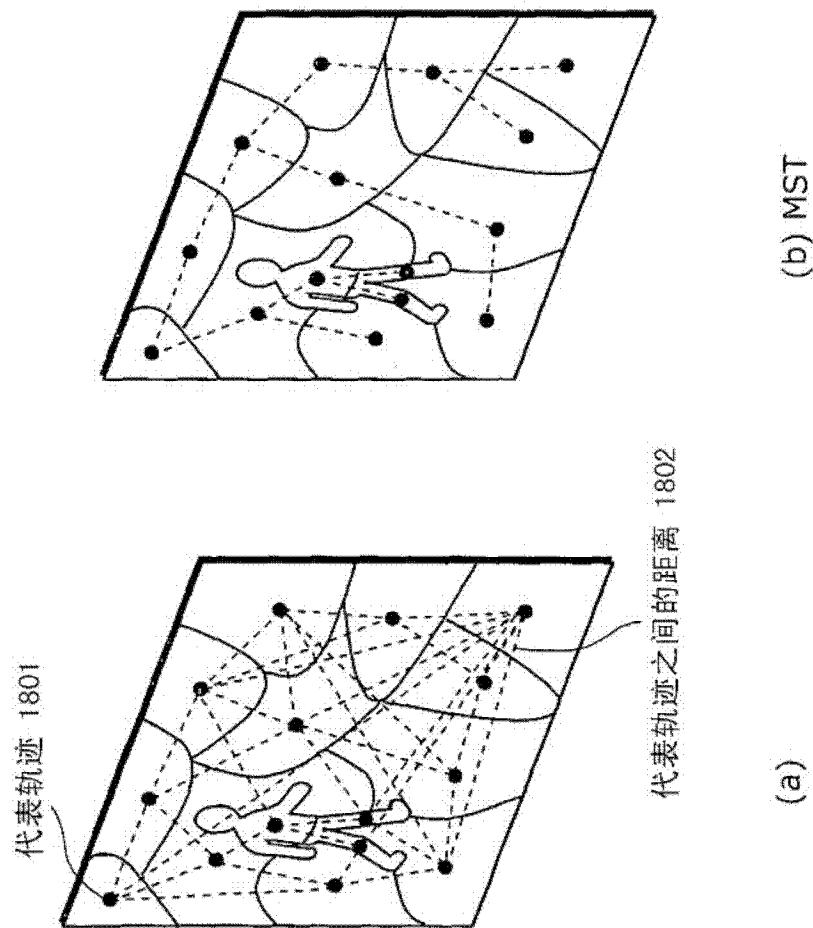


图 26

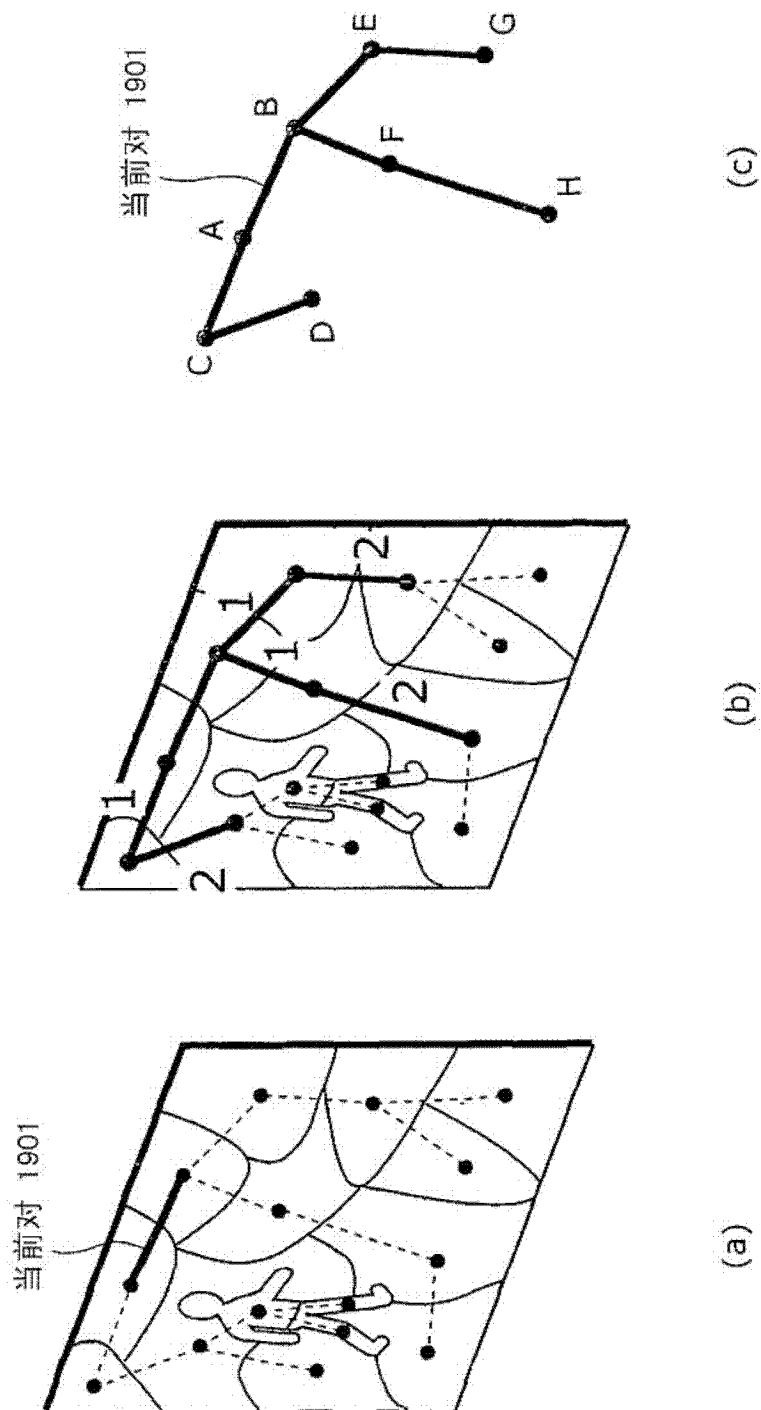


图 27

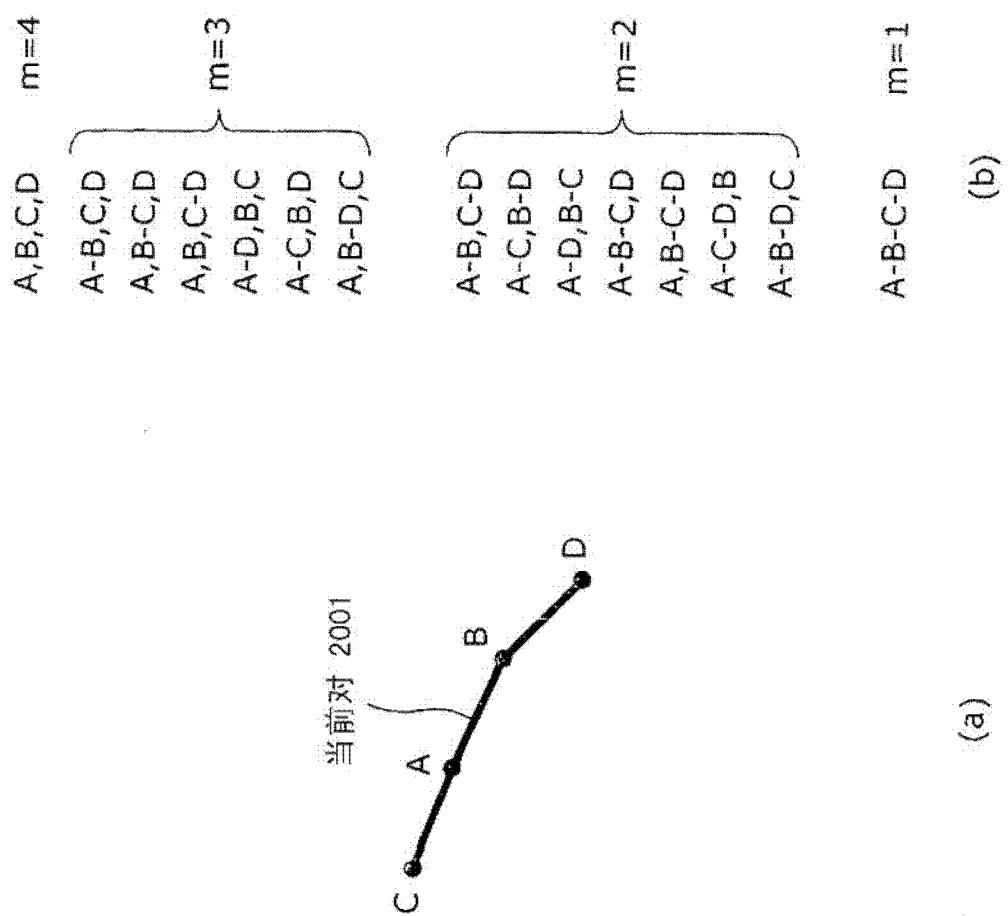


图 28