



## (12) 发明专利

(10) 授权公告号 CN 101960490 B

(45) 授权公告日 2013. 07. 17

(21) 申请号 200880127567. 8

(22) 申请日 2008. 03. 14

(85) PCT申请进入国家阶段日  
2010. 08. 27(86) PCT申请的申请数据  
PCT/GB2008/000918 2008. 03. 14(87) PCT申请的公布数据  
W02009/112790 EN 2009. 09. 17(73) 专利权人 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 A·塔亚南坦 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 陈松涛 蹇炜

(51) Int. Cl.  
G06T 7/20 (2006. 01)(56) 对比文件  
CN 1510915 A, 2004. 07. 07,  
US 2004/0194129 A1, 2004. 09. 30,  
CN 1460380 A, 2003. 12. 03,  
Joshua B. Tenenbaum et al. A Global  
Geometric Framework for Nonlinear  
Dimensionality Reduction. 《Science》. 2000,  
第 290 卷 (第 5500 期), 2319-2323.Di Zhong et al. Long-term moving  
object segmentation and tracking using  
spatio-temporal consistency. 《Image  
Processing, 2001. Proceedings. 2001  
International Conference》. 2001, 第 2 卷  
57-60.Diana Mateus et al. Articulated Shape  
Matching Using Locally Linear Embedding and  
Orthogonal Alignment. 《Computer Vision,  
2007. ICCV 2007. IEEE 11th International  
Conference》. 2007, 1-8.Gabriel J. Brostow et al. Unsupervised  
Bayesian Detection of Independent Motion  
in Crowds. 《Computer Vision and Pattern  
Recognition, 2006 IEEE Computer Society  
Conference》. 2006, 第 1 卷 594-601.Fabio Cuzzolin et al. Robust Spectral  
3D-Bodypart Segmentation Along Time.  
《Lecture Notes in Computer Science,  
Springer-Verlag Berlin Heidelberg》. 2007, 第  
4814 卷 196-211.Odest Chadwicke Jenkins et al. A  
Spatio-temporal Extension to Isomap

(续)

审查员 胡贝贝

权利要求书2页 说明书17页 附图16页

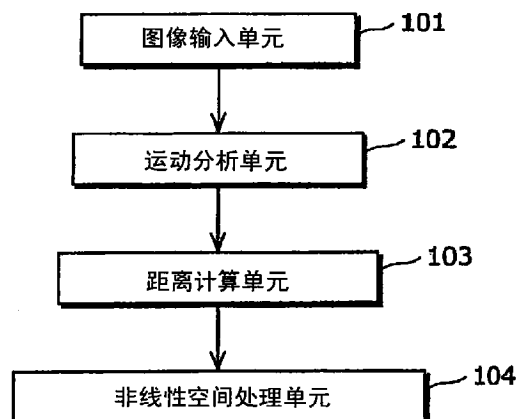
## (54) 发明名称

图像处理方法和图像处理设备

## (57) 摘要

提供了一种对包括具有变形的诸如移动的人的对象的图像鲁棒地执行分割的图像处理设备。所述图像处理设备包括：图像输入单元 (101)，其接收时间上连续的图像；运动分析单元 (102)，其使用至少两个时间上不同的图像计算块的运动，并基于所述块的所述运动，计算所述时间上连续的图像中的所述块的时间运动轨迹；距离计算单元 (103)，其使用由所述运动分析单元计算的块  $i$  的时间运动轨迹和不同于所述块  $i$  的块的时间运动轨迹计算表示所述块的所述运动的相似性的距

离；以及非线性空间处理单元 (104)，其将所述距离计算单元计算的所述距离投影到非线性空间并对所述非线性空间的所述投影的结果执行分割。



[ 转续页 ]

[ 接上页 ]

(56) 对比文件

conference on Machine learning》.2004, 第 69

Nonlinear Dimension Reduction.《Proceedings 卷 56.

of the twenty-first international

1. 一种图像处理方法,所述方法用于基于图像中的块的运动来对包括形状发生变化的具有关节的对象的图像执行分割,每个所述块由一个像素或多个像素构成,所述方法包括:

图像接收步骤,接收时间上连续的图像;

运动分析步骤,使用至少两个时间上不同的图像计算所述块的所述运动,并基于所述块的所述运动,计算所述时间上连续的图像中的所述块的时间运动轨迹;

距离计算步骤,使用在所述运动分析步骤中计算的块  $i$  的时间运动轨迹和不同于所述块  $i$  的块的时间运动轨迹来计算表示所述块的所述运动的相似性的距离;以及

非线性空间处理步骤,通过使用以下可选方式之一将在所述距离计算步骤中计算的所述距离投影到非线性空间并在所述非线性空间对所述投影的结果执行所述分割:

(i)根据至所述块  $i$  的距离按升序选择  $k$  个像素  $j$ ,至所选择的像素  $j$  的距离不发生改变,而至未选择的像素  $j$  的距离改变为无穷大;

(ii)将至所述块  $i$  的距离中大于阈值的距离无限化;以及

(iii)使用指数函数对所述距离进行变换。

2. 如权利要求 1 所述的图像处理方法,

其中,在所述距离计算步骤中,使用至少两个距离函数针对所述距离函数的每一个计算两个或更多距离,分别将所述距离的每一个投影到不同的非线性空间,并且通过整合所述投影的结果来执行所述分割。

3. 如权利要求 1 所述的图像处理方法,

其中,在所述非线性空间处理步骤中,将所述距离计算步骤计算的所述距离投影到非线性空间中,将投影后的距离转换成测地距离来执行所述非线性化。

4. 如权利要求 1 所述的图像处理方法,

其中,在所述非线性空间处理步骤中,通过计算所述非线性化的距离的本征向量来执行维数缩减。

5. 如权利要求 1 所述的图像处理方法,

其中,在所述运动分析步骤中,计算二维运动向量和仿射参数之一作为所述块的所述运动。

6. 如权利要求 1 所述的图像处理方法,还包括

图像显示步骤,在所述图像接收步骤中接收的所述图像上显示在所述非线性空间执行的所述分割的结果。

7. 如权利要求 1 所述的图像处理方法,

其中,在所述非线性空间处理步骤中,基于在所述非线性空间执行的所述分割的结果来探测在所述图像接收步骤中接收的所述图像中包括的至少两个不同的移动物体。

8. 如权利要求 1 所述的图像处理方法,还包括

对象运动预测步骤,基于在所述非线性空间执行的所述分割的结果,根据通过所述分割获得的每个分割中包括的块的时间运动轨迹计算典型轨迹,并根据所述典型轨迹来预测对象的新位置。

9. 如权利要求 1 所述的图像处理方法,还包括

记录和传输步骤,基于在所述非线性空间执行的所述分割的结果来识别在所述图像接

收步骤中接收的所述图像上的分割,以及记录或传输对应于所识别的分割的结果。

10. 一种图像处理设备,用于基于图像中的块的运动来对包括形状发生变化的具有关节的对象的图像执行分割,每个所述块由一个像素或多个像素构成,所述设备包括:

图像接收单元,用于接收时间上连续的图像;

运动分析单元,用于使用至少两个时间上不同的图像计算所述块的所述运动,并基于所述块的所述运动,计算所述时间上连续的图像中的所述块的时间运动轨迹;

距离计算单元,使用由所述运动分析单元计算的块  $i$  的时间运动轨迹和不同于所述块  $i$  的块的时间运动轨迹来计算表示所述块的所述运动的相似性的距离;以及

非线性空间处理单元,用于通过使用以下可选方式之一将所述距离计算单元计算的所述距离投影到非线性空间并在所述非线性空间对所述投影的结果执行所述分割:

(i)根据至所述块  $i$  的距离按升序选择  $k$  个像素  $j$ ,至所选择的像素  $j$  的距离不发生改变,而至未选择的像素  $j$  的距离改变为无穷大;

(ii)将至所述块  $i$  的距离中大于阈值的距离无限化;以及

(iii)使用指数函数对所述距离进行变换。

## 图像处理方法和图像处理设备

### 技术领域

[0001] 本发明涉及图像处理设备,并且具体地,涉及基于图像中的块运动来对包括具有变形的诸如移动的人的对象的图像执行分割以探测对象的设备,每个块由一个或多个像素构成。

### 背景技术

[0002] 通过组合一种用于从图像提取物理对象的候选区域的方法和一种用于将预先准备的物理对象模型应用于物理对象的提取的候选区域的方法,存在一种可用的方法,用于提取物理对象的候选区域的方法是用于探测具有变形的诸如移动的人的对象并对包括对象的图像执行分割的技术。例如,专利文献 1 公开了用于从图像提取诸如人的物理对象的轮廓图像作为物理对象的候选区域并将用于物理对象的模型应用于所提取的轮廓图像,该物理对象的区域被基于有关物理对象的知识预先参数化。以此方式,因为参数化的模型能够被施加于具有变形的诸如移动的人的对象,所以探测对象并对对象的区域执行分割变得可能。

[0003] 此外,非专利文献 1 公开了一种方法,该方法用于利用其中从多个角度捕获了固定对象的图像作为输入来计算每个图像中的像素值数据和其它图像中的像素值数据之间的距离,并利用距离作为输入执行非线性维数缩减以容许从类似角度捕获的图像被投影,使得它们在二维空间中相隔短的距离。这里,与诸如常规主成分分析 (PCA) 的线性维数缩减方法相比,其表明缩减至更低的维数是可能的,并且处理非线性分布的数据也是可能的。

[0004] 专利文献 1:日本特许公开 NO. 8-214289。

[0005] 非专利文献 1:Joshua Tenenbaum, Vin de Silva, John Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction", Science, VOL290, PP. 2319-2322, 2000 年 12 月 22 日。

[0006] 然而,对于上述专利文献 1 的技术,存在如下问题:在诸如人的对象来来往往的场景中,特别是在街上等,鲁棒地探测对象并对对象执行分割是不可能的。如上述,上述专利文献 1 描绘的图像处理方法要求从图像提取每个物理对象的候选区域。当不能正确地提取物理对象的候选区域时,精确地施加其中物理对象被参数化的模型至物理对象的候选区域变得不可能。特别是,难以正确地提取拥挤场景中物理对象的候选区域。此外,甚至在能够正确地提取物理对象的候选区域时,也存在如下问题。在特别是诸如人的有关节的对象是物理对象的情况下,因为物理对象的各种姿势引起的图像上的外观的变化非常大,所以参数化物理对象模型需要大量参数。这不仅引起计算量的增加,而且也会引起应用模型中的错误。从而,难以鲁棒地探测对象并对对象执行分割。

[0007] 另一方面,非专利文献 1 描绘的图像处理方法,以像素之间的距离作为输入执行非线性处理,以便将图像数据投影到有效地缩减的低维数空间中。以此方式,预期具有抗噪性 (noise resistance)。然而,非专利文献 1 的主要目的是将图像投影到低维数空间中,以便可视化图像之间的相似性,而没有公开用于鲁棒地探测对象并对应于具有变形的诸如人

的具有关节的对象的姿势变化对对象执行分割的方法。

[0008] 从而,本发明的目的是提供能够鲁棒地对包括具有其变形的诸如移动的人的对象的图像执行分割的图像处理方法和图像处理设备。

## 发明内容

[0009] 为了解决以上问题,本发明的特征在于一种图像处理方法,所述方法用于基于图像中的块的运动来对包括形状发生变化的对象的图像执行分割,每个所述块由一个或多个像素构成,所述方法包括:图像接收步骤,接收时间上连续的图像;运动分析步骤,使用至少两个时间上不同的图像计算所述块的所述运动,并基于所述块的所述运动,计算所述时间上连续的图像中的所述块的时间运动轨迹;距离计算步骤,使用在所述运动分析步骤中计算的块 i 的时间运动轨迹和不同于所述块 i 的块的时间运动轨迹计算表示所述块的所述运动的相似性的距离;以及非线性空间处理步骤,将在所述距离计算步骤中计算的所述距离投影到非线性空间并在所述非线性空间对所述投影的结果执行所述分割。

[0010] 应当注意,本发明不仅能够实现为上述图像处理方法,而且能够实现为包括上述每一步的图像处理设备、使得计算机执行上述每一步的程序、存储了该程序的诸如 CD-ROM 的计算机可读记录介质等。

[0011] 上述方法和设备能够探测具有变形的诸如移动的人的对象并能够对区域执行分割。此外,使用探测和分割的结果来预测对象的运动等也是可能的。

## 附图说明

[0012] 附图中:

[0013] 图 1 是示出根据本发明的第一实施例的图像处理设备的基本结构的图示;

[0014] 图 2 是示出根据本发明的第一实施例的图像处理设备的基本操作的流程图;

[0015] 图 3 是示出根据本发明的第一实施例的运动分析单元的处理范例的图示;

[0016] 图 4 是示出根据本发明的第一实施例的非线性空间处理单元的效果的概念图示;

[0017] 图 5 是示出根据本发明的第一实施例的非线性空间处理单元的效果的范例的图示;

[0018] 图 6 是示出根据本发明的第一实施例的非线性空间处理单元执行的维数缩减的范例的图示;

[0019] 图 7 是示出根据本发明的第一实施例的非线性空间处理单元执行的探测和分割的范例的图示;

[0020] 图 8 是示出根据本发明的第一实施例的第一修改例的非线性空间处理单元执行的探测和分割的范例的图示;

[0021] 图 9 是示出根据本发明的第一实施例的第二修改例的图像处理设备的基本操作的流程图;

[0022] 图 10 是示出根据本发明的第一实施例的第二修改例的非线性空间处理单元执行的探测和分割的范例的图示;

[0023] 图 11 是示出根据本发明的第一实施例的第三修改例的图像处理设备的结构范例的图示;

[0024] 图 12 是示出根据本发明的第一实施例的第四修改例的图像处理设备的结构范例的图示；

[0025] 图 13 是示出根据本发明的第二实施例的图像处理设备的结构范例的图示；

[0026] 图 14 是示出根据本发明的第二实施例的图像处理设备的基本操作的流程图；

[0027] 图 15 是示出根据本发明的第三实施例的图像处理设备的结构的图示；

[0028] 图 16 是示出根据本发明的第三实施例的图像处理设备的基本操作的流程图；

[0029] 图 17 是示出根据本发明的第三实施例的运动预测的范例的图示。

[0030] 101 图像输入单元

[0031] 102 运动分析单元

[0032] 103、103a 距离计算单元

[0033] 104、104a 非线性空间处理单元

[0034] 1001 图像显示单元

[0035] 1101 分割区域记录和传输单元

[0036] 1201 对象运动预测单元

### 具体实施方式

[0037] 本发明的一个实施例是一种图像处理方法，所述方法用于基于图像中的块的运动来对包括形状发生变化的对象的图像执行分割，每个所述块由一个或多个像素构成，所述方法包括：图像接收步骤，接收时间上连续的图像；运动分析步骤，使用至少两个时间上不同的图像计算所述块的所述运动，并基于所述块的所述运动，计算所述时间上连续的图像中的所述块的时间运动轨迹；距离计算步骤，使用在所述运动分析步骤中计算的块 i 的时间运动轨迹和不同于所述块 i 的块的时间运动轨迹计算表示所述块的所述运动的相似性的距离；以及非线性空间处理步骤，将在所述距离计算步骤中计算的所述距离投影到非线性空间并在所述非线性空间对所述投影的结果执行所述分割。基于所述块的所述运动，对具有变形的诸如移动的人的对象执行所述分割，这使得确保探测到图像中的所述对象。

[0038] 本发明的更优选的实施例是，在所述距离计算步骤中，使用至少两个距离函数针对所述距离函数的每一个计算两个或更多距离，将所述距离的每一个投影到分别不同的非线性空间，并且通过整合所述投影的结果来执行所述分割。使用距离函数和非线性空间容许对图像进行更高精度的分割。

[0039] 本发明的更优选的实施例是，在所述距离计算步骤中，针对包括在所述图像中的块计算距离，每个所述距离都表示两个块的运动的相似性，以及在所述非线性空间处理步骤中，将分别的距离非线性化，对所非线性化的距离进行维数缩减，以便将其投影到非线性空间，且通过获得具有关于投影到所述非线性空间的所述距离的相似性的一个或多个分割区域来执行所述分割。这样，鲁棒地探测具有变形的诸如特别是移动的人的对象并同时对象执行分割是可能的。

[0040] 本发明的更优选的实施例是，在所述非线性空间处理步骤中，通过将欧几里德距离转换成测地距离来执行所述非线性化。这样，鲁棒地探测具有变形的诸如特别是移动的人的对象并同时对象执行分割是可能的。

[0041] 本发明的更优选的实施例是，在所述非线性空间处理步骤中，通过计算所非线性

化的距离的本征向量来执行所述维数缩减。容许像素运动的评估的错误使得能够鲁棒地探测具有变形的诸如移动的人的对象并同时对象执行分割。

[0042] 本发明的更优选的实施例是,在所述非线性空间处理步骤中,在通过将除最靠近所述块  $i$  的  $N$  个距离外的距离无限化来执行所述非线性化后,执行所述维数缩减。因为所选择的距离和未选择的距离成非线性关系,所以高效和高精度地缩减诸如人的具有关节的对象的运动变得可能。

[0043] 本发明的更优选的实施例是,在所述非线性空间处理步骤中,在通过将至所述块  $i$  的距离中大于阈值的距离无限化来执行所述非线性化后,执行所述维数缩减。因为所选择的距离和未选择的距离成非线性关系,所以高效和高精度地缩减诸如人的具有关节的对象的运动变得可能。

[0044] 本发明的更优选的实施例是,在所述非线性空间处理步骤中,在通过将至所述块  $i$  的距离乘以包括非线性函数的指数函数来执行所述非线性化后,执行所述维数缩减。因为所选择的距离和未选择的距离是成线性关系,所以高效和高精度地缩减诸如人的具有关节的对象的运动变得可能。

[0045] 本发明的更优选的实施例是,在所述运动分析步骤中,计算二维运动向量和仿射参数之一作为所述块的所述运动。利用运动向量或仿射参数分析块的运动使得能够计算块的时间运动轨迹。

[0046] 本发明的更优选的实施例是,在所述距离计算步骤中,除所述块的所述运动的所述相似性外,还计算包括所述图像上的块之间的距离和所述图像上的块之间的角度的至少之一的距离。这样,有效地捕获形状随绕接合处旋转的旋转运动发生变化的诸如人的移动物体的运动是可能的。

[0047] 本发明的一个实施例是,所述图像处理方法包括图像显示步骤,在所述图像接收步骤中接收的所述图像上显示在所述非线性空间执行的所述分割的结果。图像上的被执行了分割的显示区域能够应用于运动中的式样改革、康复中的行走校正等。

[0048] 本发明的一个实施例是,在所述非线性空间处理步骤中,基于在所述非线性空间执行的所述分割的结果来探测在所述图像接收步骤中接收的所述图像中包括的至少两个不同的移动物体。非线性空间中的探测容许鲁棒地探测移动物体。

[0049] 本发明的一个实施例是,所述图像处理方法还包括对象运动预测步骤,基于在所述非线性空间执行的所述分割的结果,根据通过所述分割获得的每个分割中包括的块的时间运动轨迹计算典型轨迹,并根据所述典型轨迹来预测对象的运动。通过使用块的时间运动轨迹中的典型轨迹来预测对象的运动,高抗噪性的运动预测变得可能。

[0050] 本发明的一个实施例是,所述图像处理方法还包括记录和传输步骤,基于在所述非线性空间执行的所述分割的结果来识别在所述图像接收步骤中接收的所述图像上的分割,以及记录或传输对应于所识别的分割的结果。单独保持包括在图像中的探测的对象的分别的分割使得能够从分割中选择所需的分割、使得所需的分割被记录、并输出至外部。结果,有效地记录和输出分割作为图形元件变得可能。因此,在将分割记录于并传输至诸如移动电话的具有有限处理能力的装置的情况下,其能够有效地操作。

[0051] 以下参照附图描述本发明的实施例。

[0052] (第一实施例)

[0053] 图 1 是示出根据本发明的第一实施例的图像处理设备的结构的图示。图 1 中所示的图像处理设备是能够基于图像中的块的运动对包括具有变形的诸如移动的人的对象的图像鲁棒地执行分割以确切地探测图像中的对象的图像处理设备,其中每个块由一个或多个像素构成,所述图像处理设备包括:图像输入单元 101;运动分析单元 102;距离计算单元 103;以及非线性空间处理单元 104。

[0054] 图像输入单元 101 是接收时间上连续的图像并且例如为摄像机或连接至摄像机的通信接口的处理单元。

[0055] 运动分析单元 102 是使用从图像输入单元 101 接收的图像中选出的至少两个时间上不同的图像来计算块的运动,并基于计算的时间上连续的图像中块的运动来追踪块的运动的处理单元。

[0056] 距离计算单元 103 是使用运动分析单元 102 计算的块  $i$  的时间运动轨迹(temporal motion trajectory)和块  $i$  以外的块的时间运动轨迹来计算包括块的运动的相似性的距离,以捕获移动对象的变形的处理单元。例如,在使用  $N$  个块的时间运动轨迹的情况下,待计算的距离为  $N \times N$  的距离矩阵。这里,通过计算距离而将其运动改变块之间的距离的对象表述为距离矩阵是可能的,该运动尤其是具有变形的如具有关节的对象的诸如移动的人的对象的运动,每一距离都评估块的运动的相似性。如上述,本说明书中的“距离”不仅是二维空间中两点之间的距离,而且包括多维数据之间的数学距离,且为单个值或一组值(距离矩阵)。

[0057] 非线性空间处理单元 104 为首先对距离计算单元 103 计算的距离矩阵执行非线性处理,然后将距离计算单元 103 计算的距离矩阵投影到缩减的空间中(即,非线性空间),并在缩减的非线性空间执行探测和分割的处理单元。特别是,因为通过非线性处理和维数缩减的效果,非线性空间处理单元 104 能够更正确地将具有变形的如具有关节的对象的移动物体的形状信息缩减至低维数空间,所以非线性处理单元 104 能够探测图像中的对象并以低的计算量鲁棒地对图像执行分割。

[0058] 应当注意到,可以通过诸如在计算机中执行的程序的软件或诸如电路的硬件来实现包括在图像处理设备中的每个部件(图像输入单元 101、运动分析单元 102、距离计算单元 103、以及非线性空间处理单元 104),该计算机具有 CPU、RAM、ROM、通信接口、I/O 端口、硬盘、显示器等。以下,这同样适应于根据其它实施例的图像处理设备。

[0059] 关于本发明的图像处理方法,以下使用图 2 中所示的流程图详细描述探测对象并对对象执行分割的范例。

[0060] 首先,第一步骤 S201 中,图像输入单元 101 接收输入图像。

[0061] 接下来,步骤 S202 中,运动分析单元 102 使用至少两个输入图像计算块的运动。应当注意到,这里作为计算块的运动的范例,计算像素的运动。以下,在本实施例中,作为逐块处理的范例,将描述逐像素处理。在执行其中块由像素构成的逐块处理的情况下,通过将对应于多个块的像素的数据求和或通过计算多个块的平均值来获得对应于块的数据(典型值),并且可以以与逐像素处理中相同的方式使用获得的典型值来执行处理。

[0062] 首先,将描述通过光流(optical flow)计算来计算像素的运动的范例。非专利文献 2 或非专利文献 3 能够用于根据光流计算来计算运动向量。

[0063] 非专利文献 2:P. Anandan, "A Computational Framework and an Algorithm for

the Measurement of Visual Motion”, International Journal of Computer Vision, Vol. 2, PP. 283-310, 1989。

[0064] 非专利文献 3: Vladimir Kolmogorov and Ramin Zabih, “Computing Visual Correspondence With Occlusions Via Graph Cuts”, International Conference on Computer Vision, 2001。

[0065] 例如, 当假定在步骤 S201 中输入 T 个图像, 使用在时间 t 和时间 t+1 输入的图像估计像素 i 的运动向量  $(u_i^t, v_i^t)$ 。这里, 输入的图像不必是时间上连续的, 并且例如, 可以使用在时间 t 和时间 t+1 输入的图像计算像素的运动。注意 n 是等于或大于 1 的整数。此外, 除运动向量外, 仿射参数可以估计为像素的运动。此时, 可以获得关于所有像素的运动信息。此外, 在期望以较高速度执行处理的情况下, 将图像划分成网格并且仅可以获得网格上的具有规则间隔的像素的运动信息, 或者如上述, 将图像划分成块, 并且可以以逐块为基础获得运动信息。此外, 在使用非专利文献 2 计算运动向量的情况下, 因为能够计算其可靠程度, 所以仅可以使用具有高可靠程度的运动信息的像素。此外, 在使用非专利文献 3 计算运动向量的情况下, 能够估计遮挡 (occlusion)。从而, 仅可以使用未被遮挡的像素的运动信息。

[0066] 此外, 替代基于块的平移移动的假设来计算运动向量的方法, 可以使用基于块的仿射变换的假设来计算运动向量的方法作为计算像素的运动的运动的方法。非专利文献 4 能够用于基于仿射变换的假设来计算运动向量的方法。

[0067] 非专利文献 4: Jianbo Shi 和 Carlo Tomasi, “Good Features to Track”, IEEE Conference on Computer Vision and Pattern Recognition, PP. 593-600, 1994。

[0068] 以上方法估计仿射参数 A', 其相当于与分别在时间 t 和时间 t+1 输入的图像的像素 i 相邻的运动。关于像素 i, 时间 t 和时间 t+1 的图像上的像素位置  $x_t^i$  和  $x_{t+1}^i$  具有如下关系:

[0069]  $x_{t+1}^i = A_t' x_t^i$  等式 1

[0070] 以上方法能够针对特别是旋转对象以比使用基于平移移动的假设来计算运动向量的方法的情况更高的精度估计像素 i 的运动。

[0071] 接下来, 在步骤 S203 中, 运动分析单元 102 使用在步骤 S202 中计算的运动信息, 针对时间上连续的图像, 根据像素 i 的运动来计算像素 i 的时间运动轨迹。如图 3 中所示, 使用在步骤 S202 中根据时间 t 的输入图像 301 的像素 i 303 的运动计算的运动信息 302 来追踪像素 i 的运动。使用像素 i 通过时的时间 t 时图像上的像素位置  $(x_t^i, y_t^i)$  按以下等式来计算像素 i 的时间运动轨迹:

[0072]  $x^i = (x_1^i, y_1^i, \dots, x_T^i, y_T^i)$  等式 2

[0073] 其中, T 是用于计算时间运动轨迹的图像的数目。

[0074] 接下来, 在步骤 S204 中, 距离计算单元 103 使用按等式 2 计算的像素 i 的时间运动轨迹来计算包括像素的运动的相似性的距离矩阵。能够按以下等式来计算像素 i 和像素 j 之间的线性距离  $f(i, j)$ :

[0075]  $f(i, j) = \text{mean}_{ij} + w \cdot \sqrt{\text{var}_{ij}}$  等式 3

[0076] 其中, w 是加权因子且是设计者设定的参数。此外, 如下表述  $\text{mean}_{ij}$  和  $\text{var}_{ij}$ :

[0077]  $mean_{ij} = \frac{1}{T} \sum_{t=1}^T d'_{ij}$  等式 4

[0078]  $var_{ij} = \frac{1}{T} \sum_{t=1}^T (d'_{ij} - mean_{ij})^2$  等式 5

[0079] 其中

[0080]  $d'_{ij} = \sqrt{(x_t^i - x_t^j)^2 + (y_t^i - y_t^j)^2}$  等式 6

[0081] 如上述,如等式 3 所示,除等式 4 所示的像素之间的距离的时间平均值外,使等式 5 所示的像素之间的距离的时间波动成分为线性距离  $f(i, j)$  的要素。特别是,等式 5 所示的像素之间的距离的波动成分指示像素的运动的相似性,并且这使得不仅能够捕获其中像素之间的距离关系不变的刚性体,而且能够捕获由具有关节的对象等引起的变形。

[0082] 接下来,在步骤 S205 中,非线性空间处理单元 104 对按等式 2 计算的线性距离  $f(i, j)$  执行如下非线性处理:

[0083]  $g(i, j) = \begin{cases} f(i, j) & \text{如果 } i \text{ 为 } K - \text{最近邻} \\ \infty & \text{其它} \end{cases}$  等式 7

[0084] 其中,在注意力集中于像素  $i$  时,根据至像素  $i$  的线性距离按升序选择  $K$  个像素  $j$ ,至所选择的像素  $j$  的距离不发生改变,而至未选择的像素  $j$  的距离改变为无穷大。这里,虽然按升序选择线性距离,但是也可以按以下关系设定阈值  $TH$ :

[0085]  $g(i, j) = \begin{cases} f(i, j) & \text{如果 } f_x(i, j) < TH \\ \infty & \text{其它} \end{cases}$  等式 8

[0086] 此外,可以使用指数函数按以下等式将线性距离非线性化:

[0087]  $g(i, j) = \exp\left(-\frac{f(i, j)}{\sigma}\right)$  等式 9

[0088] 其中,  $\sigma$  由设计者给定。

[0089] 接下来,在以下等式中使用非线性化距离  $g(i, j)$  将距离更新。

[0090] 关于距离的非线性化,其不限于以上函数,且能够对像素  $i$  和像素  $j$  之间的距离执行非线性转换的任何情况都是可接受的。

[0091] 接下来,如下计算测地距离:

[0092]  $g_{\text{new}}(i, j) = \min(g(i, j), g(i, k) + g(k, j))$  等式 10

[0093] 注意  $\min(x, y)$  是返回值  $x$  和值  $y$  之间的较小值的函数。

[0094] 这里,将参照图 4(a)-(c) 的概念图示描述等式 7 至 10 所示的非线性处理。虽然于此描述了等式 8,但是能够预期执行其它非线性处理带来相同效果。首先,图 4 的左图示 (a) 示出了二维数据分布。这里分别的数据点相当于等式 3 所示的像素  $i$  的时间运动轨迹。在不执行等式 7 至 10 所示的非线性处理的情况下,如图 4 的中心图 (b) 中所示,数据点  $i$  和数据点  $j$  之间的距离变得比数据点  $i$  和数据点  $k$  之间的距离小。然而,如图 4 的右图 (c) 中所示,例如,当执行等式 8 和 10 所示的处理时,数据点  $i$  和数据点  $j$  之间的距离不是欧几里德距离而是称作测地距离的距离,遍及该距离,箭头追踪数据点。结果,与使用欧几里德距离的情况不同,并且数据点  $i$  和数据点  $j$  之间的距离变得比数据点  $i$  和数据点  $k$  之间的距离大。

[0095] 接下来,将参照图 5(a) 和 (b) 所示的概念图示描述执行等式 7 至 10 所示的非线

性处理的有效性。这里,为了描述清楚地描述线性距离  $f(i, j)$  和非线性距离  $g_{\text{new}}(i, j)$  之间的差异,如等式 6 所示,使用时间  $t$  时像素之间的距离作为范例。实际上,除像素之间的距离外,如等式 3 所示,在使用像素之间的距离的波动成分作为像素的运动的相似性时,不仅能够捕获具有关节的对象等的形状,而且能够捕获其变形。图 5(a) 示出了不执行等式 7 至 10 所示的处理的范例。例如,头的像素  $i$  402 和手的像素  $j$  403 之间的距离是线性距离 401 所示的距离。另一方面,当执行等式 7 至 10 所示的非线性处理时,如图 5(b) 中所示,头的像素  $i$  402 和手的像素  $j$  403 之间的距离是经由像素  $k$  404 至像素  $j$  追踪的箭头所示的线性和。因此,诸如人的具有关节的对象的形状不能以线性距离 401 表述,而相反,其使得可以将关节的形状表述为距离。应当注意,计算测地距离的方法不限于等式 10。

[0096] 接下来,在步骤 S206 中,非线性空间处理单元 104 对步骤 S205 中计算的非线性化距离矩阵执行维数缩减。通过在执行 Young-Householder 变换后获得本征系统,能够实现维数缩减。这是将多维空间中分布的数据有效地投影到低维空间中的方法,并且能够针对输入数据噪声表述鲁棒的数据(这里,相当于在使用等式 2 计算像素  $i$  的时间运动轨迹时,像素  $i$  的运动信息中包括误差的情况)。

[0097] 换句话说,非线性距离矩阵表述为:

$$[0098] \quad G = \{g_{\text{new}}(i, j)\} \quad \text{等式 11}$$

[0099] 且如下执行 Young-Householder 变换,其中,等式 11 中的  $G$  被中心矩阵  $H$  在两边相乘。距离矩阵包括点之间的距离,而相反,这是变换至具有重力中心作为原点的距离矩阵。

$$[0100] \quad \tau(G) = \frac{HG^{(2)}H}{2} \quad \text{等式 12}$$

[0101] 注意  $H$  是中心矩阵,且如下表述:

$$[0102] \quad H_{ij} = (I - 1/N) \quad \text{等式 13}$$

[0103]  $I$  是单位矩阵,且  $N$  是等式 2 所示的像素  $i$  的时间运动轨迹的数目。

[0104] 此外,

$$[0105] \quad G^{(2)} = \{g_{\text{new}}(i, j)^2\} \quad \text{等式 14}$$

[0106] 接下来,为执行维数缩减,计算对应于  $\tau(G)$  的  $p$  个本征向量  $e_p$  和对应于  $\tau(G)$  的本征值  $\lambda_p$ 。

[0107] 以此方式,基于等式 10,当进行如下等式 15 的假设时,将  $g_{\text{new}}^i$  投影到维数缩减空间的结果能够按如下等式 16 表述为数据  $z_p^i$ :

$$[0108] \quad g_{\text{new}}^i = (g_{\text{new}}(i, 0), g_{\text{new}}(i, 1), \dots, g_{\text{new}}(i, N)) \quad \text{等式 15}$$

$$[0109] \quad z_p^i = \sqrt{\lambda_p} e_p^i \quad \text{等式 16}$$

[0110] 注意为  $e_p^i$  为第  $p$  个本征向量  $e_p$  的第  $i$  个元。 $p$  为本征向量的数目,其可以根据待使用的场景来实验地确定,或在根据本征值  $\lambda_p$  如下计算贡献比  $a_p$  后基于贡献比来确定:

$$[0111] \quad a_p = \frac{\sum_{p=1}^P \lambda_p}{\sum_{p=1}^N \lambda_p} \quad \text{等式 17}$$

[0112]  $p$  是待使用的本征向量的数目,即缩减的空间的维数的数目。 $N$  是所有本征向量的

数目。从而,在贡献比  $a_p$  等于或大于恒定值时,  $p$  可以是本征向量的数目。

[0113] 如上述,利用等式 2 至 16 所示的处理,能够将等式 2 所示的像素  $i$  的时间运动轨迹与通过本征向量  $e_p$  维数缩减为非线性的空间中的数据关联。

[0114] 图 6 示出了其中人正在行走的图像被输入的情况下,将被维数待缩减为非线性的像素  $i$  的时间运动轨迹投影到空间中的结果。水平轴为本征向量  $e_1$ ,且竖直轴为本征向量  $e_2$ 。二维投影的点  $(z_1^i, z_2^i)$  被  $g_{new}^i$  投影。这里,因为非线性空间中的数据  $z_p^i$  和像素  $i$  的时间运动轨迹  $x^i$  为一一对应关系,所以能够认为  $(z_1^i, z_2^i)$  对应于像素  $i$  的时间运动轨迹  $x^i$ 。应当注意,虽然非线性空间的维数数目为二维以可视化结果,如上述,但是其不必是二维的,并且以更多数目的维数以更高精度投影数据是可能的。

[0115] 接下来,在步骤 S207 中,非线性空间处理单元 104 探测对象并通过将在步骤 S206 中投影到缩减的非线性空间中的数据进行群集来对对象执行分割,如图 6 中所示。

[0116] 首先,分割区域表述为:

[0117]  $\theta = \{\theta_1, \dots, \theta_m, \dots, \theta_M\}$  等式 18

[0118] 其中,  $M$  是分割区域的数目并根据待使用的场景经验地确定。

[0119] 每个分割区域  $\theta_m$  由参数  $\overline{z_m}$ 、 $Z_m$  表述。这里,  $\overline{z_m}$  是缩减的非线性空间中的属于分割  $\theta_m$  的数据的坐标的平均值,且  $Z_m$  是关于属于分割  $\theta_m$  的数据的坐标的方差矩阵。

[0120]  $\overline{z_m}$  的初始值可以随机确定,或者可以将通过将缩减的非线性空间划分成网格而获得的交点的坐标设定为初始值。

[0121] 注意  $\overline{z_m}$ 、 $Z_m$  能够如下表述:

$$[0122] \quad \overline{z_m} = \begin{bmatrix} \overline{z_1^m} \\ \vdots \\ \overline{z_p^m} \end{bmatrix} \quad \text{等式 19}$$

$$[0123] \quad Z_m = \frac{1}{C_m} \sum_{c_m=1}^{C_m} \begin{bmatrix} z_1^{c_m} - \overline{z_1^m} \\ \vdots \\ z_p^{c_m} - \overline{z_p^m} \end{bmatrix} \begin{bmatrix} z_1^{c_m} - \overline{z_1^m} & \dots & z_p^{c_m} - \overline{z_p^m} \end{bmatrix} \quad \text{等式 20}$$

[0124] 其中,  $C_m$  是缩减的非线性空间中属于分割区域  $\theta_m$  的数据的数目。

[0125] 以下描述具体计算方法。首先,使用以下距离函数获得缩减的非线性空间中的数据  $z_i$  所属的分割区域  $\theta_m$ :

[0126]  $\psi_m(z_i) = \varphi_m(z_i) + \ln|Z_m| - \ln p(\omega_m)$  等式 21

[0127] 其中,  $\psi_m(z_i)$  示出了非线性空间中对应于像素  $i$  的时间运动轨迹和每个分割  $\theta_m$  的数据  $z_i$  之间的距离。每个数据属于分割区域  $\theta_m$ , 其中  $\psi_m(z_i)$  是最小值。注意  $\varphi_m(z_i)$  是 Mahalanobis 距离, 且:

[0128]  $\varphi_m(z_i) = (z_i - \overline{z_m})' Z_m^{-1} (z_i - \overline{z_m})$  等式 22

[0129] 此外,  $\varphi_m(z_i)$  可以用于代替  $\psi_m(z_i)$ 。

[0130] 此外,  $p(\omega_m)$  可以是恒定值, 且在对诸如人的常规对象执行分割的情况下, 可以基于躯体区域的形状、面积比等对其进行设定。 $\omega_m$  是用于分割区域  $\theta_m$  的加权因子。

[0131] 接下来, 基于等式 21 计算的结果, 使用属于分割区域  $\theta_m$  的数据  $z_i$  如下更新分割区域  $\theta_m$  的参数  $\bar{z}_m$ 、 $Z_m$  :

$$[0132] \quad \bar{z}_m = \frac{\sum_{c_m=1}^{C_M} \omega_{c_m} z_{c_m}}{\sum_{c_m=1}^{C_m} \omega_{c_m}} \quad \text{等式 23}$$

$$[0133] \quad Z_m = \frac{\sum_{c_m=1}^{C_m} \omega_{c_m}^2 (z_{c_m} - \bar{z}_m)(z_{c_m} - \bar{z}_m)^t}{\sum_{c_m=1}^{C_m} \omega_{c_m}^2} \quad \text{等式 24}$$

[0134] 其中  $z_{c_m}$  是属于缩减的非线性空间中的分割区域  $\theta_m$  的数据。此外, 可以接受  $\omega = 1$  或者可以依赖于与输入数据的平均值的偏离调整  $\omega$ 。如上述, 将使用等式 21 至 24 进行距离计算和参数更新重复固定次数, 容许获得非线性空间中的每个数据所属的分割区域  $\theta_m$ 。应当注意, 除以上外, 可以使用诸如 k-mean 法和竞争学习法的其它群集方法。图 7(a) 和 (b) 示出了对非线性空间中的数据执行分割的范例, 其中, 假定  $M = 9$ , 且使用其中人正在走动的图像作为输入。当在图像上看到缩减的非线性空间中的分割区域  $\theta_1$  至  $\theta_9$  的分别的对应关系时,  $\theta_1$  对应于人的头,  $\theta_2$  对应于上躯干,  $\theta_3$  对应于手臂,  $\theta_4$  对应于下躯干,  $\theta_5$  和  $\theta_7$  对应于大腿, 以及  $\theta_6$  和  $\theta_8$  对应于小腿。注意  $\theta_9$  对应于背景的部分。

[0135] 这里, 如图 7(b) 中的箭头所示, 对应于缩减的非线性空间中的分割区域的不仅是一个图像上的区域, 而且是在时间上连续的图像上对像素进行追踪的结果。即, 执行缩减的非线性空间中的分割, 容许探测图像中的对象 (对象的区域) 和包括对象的图像的分割作为时间上追踪在图像中移动的物体的区域的结果。此外, 因为不必设定人的候选区域作为预处理, 不存在由探测人的候选区域时的错误引起的分割中的失败。如上述, 无需拟合大量参数, 非线性空间中的群集容许对包括具有变形的移动的人的图像的分割, 使得能够鲁棒地探测图像中的对象。

[0136] (第一实施例的第一修改例)

[0137] 接下来, 将描述一个范例作为第一实施例的第一修改例, 其中根据第一实施例执行非线性空间中的分割以探测对象。在第一实施例中, 描述了用于对图像中的人的区域执行分割的方法, 其中图像中的人正在走动。如图 1 中所示, 本图像处理设备包括图像输入单元 101、运动分析单元 102、距离计算单元 103、以及非线性空间处理单元 104。因为其结构与第一实施例中的相同, 因此这里不对其结构进行描述。在第一修改例中, 将描述通过与第一实施例中相同的方法改变参数来探测对象的方法。具体地, 将使用等式 7 时的  $k$  值或使用等式 8 时的  $TH$  值设定为比在使用等式 7 或等式 8 来对区域执行分割时的值大。

[0138] 图 8 示出了缩减的非线性空间中的分割的结果的范例, 其中, 假定  $M = 2$ , 且使用其中人正在走动且自行车正运行的图像作为输入。将描述缩减的非线性空间中分割区域  $\theta_1$  和  $\theta_2$  的对应关系。分割区域  $\theta_1$  对应于图像上自行车的运动, 且分割区域  $\theta_2$  对应于图像上人的走动。如上述, 在低维缩减空间执行分割容许简单并鲁棒地单独探测 (即分割) 自

行车的运动和人的运动。

[0139] 应当注意,如等式 2 所示,因为这里不使用无运动像素的时间运动轨迹作为输入,仅两个移动的具有关节的实体,即走动的人和自行车,被投影到非线性空间中。理所当然,无运动像素的时间运动轨迹可以用作输入。这里,对应于如第一实施例中的缩减的非线性空间中的分割区域的不仅是一个图像上的区域,而且是在时间上连续的图像上对像素进行追踪的结果。即,执行缩减的非线性空间中的分割,容许探测每个物体作为时间上追踪在图像中移动的物体的区域的结果。此外,因为不必设定人的候选区域作为预处理,不存在由探测人的候选区域时的错误引起的分割中的失败。如上述,部分改变参数容许移动对象的鲁棒的探测和分割。

[0140] (第一实施例的第二修改例)

[0141] 接下来,将描述在不同于第一实施例的非线性空间中实现分割的方法作为第一实施例的第二修改例。如图 1 中所示,本图像处理设备包括图像输入单元 101、运动分析单元 102、距离计算单元 103、以及非线性空间处理单元 104。因为其结构与第一实施例中的相同,所以这里不对其结构进行描述。在第二修改例中,非线性空间处理单元 104 以不同于第一实施例的方法执行非线性空间中的分割。这里将描述一种方法,利用该方法,非线性空间处理单元基于非专利文献 5 中公开的方法执行非线性空间中的处理。

[0142] 非专利文献 5:Arik Azran, Zoubin Ghahramani, "A new Approach to Data Driven Clustering", International Conference on Machine Learning, 2006。

[0143] 图 9 是示出根据本发明的第一实施例的第二修改例的处理的流程图。因为步骤 S201 至 S206 与第一实施例中相同,所以这里不对它们进行描述。

[0144] 在步骤 S801 中,非线性空间处理单元 104 对非线性化距离  $g(i, j)$  执行以下归一化 (normalization) 处理,其中,已经对该非线性化距离执行了等式 7 至 10 所示的任何处理。

$$G_{ij} = \frac{g(i, j)}{\sum_{j=1}^N g(i, j)} \quad \text{等式 25}$$

[0146] 其中,  $N$  是等式 2 所示的像素  $i$  的时间运动轨迹的数目。且其中  $G$  描绘数据之间的相似性。接下来,对归一化  $G$  执行以下计算:

$$\Omega = G^t \quad \text{等式 26}$$

[0148] 作为执行此计算的结果,在  $G_{ij}$  是小值的情况下,即数据之间的相似性小,其变为更小的值,而在  $G_{ij}$  为大值的情况下,其变为更大值。换句话说,具有高相似性的数据之间的距离变得占主要地位。这容许鲁棒地执行以下处理。

[0149] 关于  $t$ ,例如,准备矩阵  $\Omega^r$ ,诸如  $\Omega^1 = G^{t1}, \dots, \Omega^r = G^{tr}, \dots, \Omega^R = G^{tR}$  (注意  $t1 < tr < tR$ ),并且期望从步骤 S802 至步骤 S803 单独执行每个  $\Omega^r$  的计算。此外,因为非专利文献 4 中描述了高速执行等式 26 的计算的方法,所以可以使用该方法。

[0150] 接下来,在步骤 S802 中,非线性空间处理单元 104 设定群集  $Q_m$  的初始值以执行用于执行分割的群集。这里群集数目为  $M$ 。这里群集数目  $M$  设定为使得,  $R$  设定得越大,则值  $M$  变得越小。

[0151] 首先,如下确定第一群集的初始值:

$$[0152] \quad Q_i = \frac{1}{N} \sum_{i=1}^N \Omega_i \text{ 等式 27}$$

[0153] 其中,  $\Omega_i$  是矩阵  $\Omega$  中的第  $i$  行向量。这里如等式 15 中,  $\Omega_i$  对应于等式 2 所示的像素  $i$  的时间运动轨迹。接下来, 如等式 28 所示, 从用于  $N$  个  $\Omega_i$  的每一个的群集  $Q_1$  至  $Q_{m-1}$  中选择最小距离的群集, 于是选择群集为最小距离时为  $\Omega_i$  中的最大距离的  $\Omega_{\arg \max_{\min}}$ , 且如下地将第一群集后的群集的初始值设定为新群集的初始值。对从  $m = 2$  至  $M$  重复此操作。

$$[0154] \quad Q_m = \Omega_{\arg \max_{\max}} \quad \text{等式 28}$$

[0155] 此处理是设定为属于群集的数据中的最大距离的数据作为新群集的初始值的处理, 并容许群集的初始值更广泛地分布, 使得能够防止归因于群集在一个点的集中而不能正确执行群集数据分布的问题。

[0156] 注意, 使用 KL 散度如下计算群集  $Q_m$  和  $\Omega_i$  之间的距离, 该距离用于计算群集  $Q_1$  至  $Q_{m-1}$  中的最小距离。

$$[0157] \quad J_{mn} = KL(\Omega_n \| Q_m) \quad \text{等式 29}$$

[0158] 如上述, 能够确定群集  $Q_1$  至  $Q_m$  的初始值。

[0159] 接下来, 在步骤 S803 中, 非线性空间处理单元 104 对  $N$  个  $\Omega_i$  的每一个执行群集, 使用如等式 29 中的 KL 散度确定  $N$  个  $\Omega_i$  所属的群集  $Q_1$  至  $Q_m$ 。这里, 在 KL 散度为最小值时, 每  $\Omega_i$  个属于群集  $Q_m$ 。

[0160] 接下来, 非线性空间处理单元 104 使用对  $N$  个  $\Omega_i$  进行群集的结果按以下等式更新群集  $Q_m$  :

$$[0161] \quad Q_m^{new} = \frac{1}{C_m} \sum_{c_m=1}^{C_m} \Omega_{c_m} \text{ 等式 30}$$

[0162] 其中,  $\Omega_{c_m}$  是属于群集  $Q_m$  的  $\Omega_i$ , 且  $C_m$  是属于群集  $Q_m$  的  $\Omega_i$  的数目。将步骤 S803 重复固定次数, 以确定  $\Omega_i$  属于哪个群集  $Q_m$ 。应当注意  $\Omega_i$  和像素  $i$  的时间运动轨迹  $x^i$  是一一对应关系。也就是说, 因为在确定  $\Omega_i$  所属的群集  $Q_m$  时,  $\Omega_i$  对应于等式 2 所示的像素  $i$  的时间运动轨迹, 所以应当对像素  $i$  的时间运动轨迹  $x^i$  执行分割。

[0163] 将参照图 10(a) 至 (c) 描述范例, 其中, 假定群集号  $M = 9$ 、 $M = 5$ 、以及  $M = 2$ , 在非线性空间中分别针对  $r = t1$ 、 $t2$ 、 $t3$  (注意  $t1 < t2 < t3$ ) 产生  $\Omega^r$  和对的数据  $\Omega_i$  执行分割。在  $r = t3$  且  $M = 2$  的情况下分割的结果是数据被分割成人的不包括一边的手臂区域的区域和其它区域。此外, 在  $r = t2$  且  $M = 5$  的情况下, 数据被分割成作为一个区域的头区域和躯干区域以及分别的腿区。此外, 在  $r = t3$  且  $M = 9$  的情况下, 获得了与图 7(a) 和 (b) 中相同的探测和分割结果。以此方式, 本修改例的优点是以复数准备  $\Omega^r$  容许在不同水平执行分割。

[0164] 应当注意, 数据没有分割到一边的手臂区域中的原因是, 由于追踪像素时发生了遮挡, 包括在手臂区域中的像素  $i$  的时间运动轨迹 (等式 2) 没有用作输入。此外, 对应于如第一实施例中的每个分割区域的不仅是一个图像上的区域而且是在时间上连续的图像上对像素追踪的结果。也就是说, 作为对在图像中移动的具有关节的实体的区域时间上追踪的结果, 在缩减的非线性空间执行分割容许进行探测和分割。此外, 因为不必设定人的候选区域作为预处理, 不存在探测人的候选区域时的错误引起的分割中的失败。

[0165] 此外,虽然与第一实施例不同,在本实施例中,由维数缩减引起的加速的效果小,但是因为相同方法能够用于如图 10(a) 至 (c) 中所示的粗分割和细分割,所以存在的优点是,能够如第一实施例和第一实施例的第一修改例所示,同时执行对人的探测和对人的区域的分割。如上述,探测具有变形的诸如移动的人的对象并对区域执行分割是可能的。

[0166] (第一实施例的第三修改例)

[0167] 接下来,将描述根据本发明的第一实施例的第三修改例的图像处理设备。这里,将描述范例,其中,增加了用于显示第一实施例中探测对象并对区域执行分割的结果作为图像的功能。如图 11 中所示,将图像显示单元 1001 增加至根据第三修改例的该图像处理设备,该图像处理设备包括图像输入单元 101、运动分析单元 102、以及非线性空间处理单元 104,并且因此,图像处理设备能够在监视器等上显示分割的区域作为图像。

[0168] 如图 7(b) 中所示,图像显示单元 1001 通过色彩将图像上的对应于分割区域  $\theta_m$  的分割区域分开,容许分别区别分割的区域,并在监视器等上显示它们。理所当然,可以不同地显示每个区域。这里,等式 2 所示的像素  $i$  的时间运动轨迹对应于任何分割区域  $\theta_m$ 。因此,在基于属于分割区域  $\theta_m$  的像素  $i$  的时间运动轨迹选择图像上的像素并基于分割区域数目  $m$  执行色彩编码时,容易显示图像上的分割区域。

[0169] (第一实施例的第四修改例)

[0170] 接下来,将描述根据本发明的第一实施例的第四修改例的图像处理设备。这里,将描述范例,其中,增加了用于针对每个分割区域记录和传输第一实施例中探测对象并对区域执行分割的结果的功能。如图 12 中所示,通过增加分割区域记录和传输单元 1101,能够实现根据第四修改例的该图像处理设备,该图像处理设备包括图像输入单元 101、运动分析单元 102、以及非线性空间处理单元 104。

[0171] 类似于图像显示的情况,作为分割的结果,分割区域记录和传输单元 1101 根据分割区域数目  $m$  单独记录并传输图像。此外,通过如下对属于每个分割的运动信息进行平均能够缩减运动信息。虽然,通常有必要给每个像素保存运动信息,但是当执行以下处理时,仅需要给一个分割保存一条运动信息。例如,当像素的运动向量  $(u_i^t, v_i^t)$  用作运动信息时,能够如下计算对每个分割进行了平均的运动信息。

$$[0172] \quad u_i^m = \frac{1}{C_m} \sum_{c_m=1}^{C_m} u_i^{c_m} \text{等式 31}$$

$$[0173] \quad v_i^m = \frac{1}{C_m} \sum_{c_m=1}^{C_m} v_i^{c_m} \text{等式 32}$$

[0174] 其中,  $C_m$  是属于分割区域  $\theta_m$  的像素的数目或像素的时间运动轨迹的数目。 $u_i^{c_m}$ 、 $v_i^{c_m}$  分别是属于分割区域  $\theta_m$  的像素的运动向量。

[0175] 应当注意,在使用仿射运动缩减运动信息的情况下,代替等式 31 和 32,可以计算像素位置的平均值,其中,基于仿射运动将像素移动至该像素位置。如上述,记录并传输分割区域同时缩减像素的运动信息是可能的。

[0176] (第二实施例)

[0177] 接下来,将描述使用两个或更多距离函数实现更高精度的探测和分割的方法作为第二实施例。如图 13 中所示,图像处理设备包括图像输入单元 101、运动分析单元 102、距离计算单元 103a、以及非线性空间处理单元 104a。图像输入单元 101 和运动分析单元 102

与第一实施例中的相同,这里不对它们进行描述。

[0178] 虽然距离计算单元 103a 和非线性空间处理单元 104a 各基本与第一实施例的距离计算单元 103 和非线性空间处理单元 104 具有相同功能,但是有一点与第一实施例不同,即距离计算单元 103a 和非线性空间处理单元 104a 使用了附加参数以执行更高精度的处理。以下主要描述与第一实施例不同的点。

[0179] 距离计算单元 103a 使用用于图像上的像素间的距离和距离波动的距离函数和用于像素间的角度和角度波动的距离函数来计算每个距离函数的距离矩阵。

[0180] 在将距离计算单元 103a 计算的两个距离矩阵投影到分别不同的非线性空间后,非线性空间处理单元 104a 通过整合 (integrating) 该两个不同的非线性空间来执行群集。

[0181] 这里,按图 14 中所示的处理的流程图进行描述。首先,因为步骤 S201 至 S203 与第一实施例中相同,所以这里不对它们进行描述。

[0182] 接下来,在步骤 S204,距离计算单元 103a 使用按等式 2 计算的像素 i 的时间运动轨迹来计算包括像素的运动的相似性的距离矩阵。这里,除等式 3 所示的距离  $f(i, j)$  外,将描述范例,其中,基于像素 i 和像素 j 之间的角度的距离  $f_2(i, j)$  用作第二距离函数。

[0183]  $f_2(i, j) = a\_mean_{ij} + w_a \cdot \sqrt{a\_var_{ij}}$  等式 33

[0184] 其中,  $w_a$  是加权因子和设计者设定的参数。此外,  $a\_mean_{ij}$  和  $a\_var_{ij}$  如下表述:

[0185]  $a\_mean_{ij} = \frac{1}{T} \sum_{t=1}^T a'_{ij}$  等式 34

[0186]  $a\_var_{ij} = \frac{1}{T} \sum_{t=1}^T (a'_{ij} - a\_mean_{ij})^2$  等式 35

[0187] 其中,

[0188]  $a_{ij} = \arctan\left(\frac{y'_i - y'_j}{x'_i - x'_j}\right)$  等式 36

[0189] 如上述,使用像素 i 和像素 j 之间的角度及角度的波动成分能够表述像素的运动的相似性。这不仅容许捕获刚体而且容许捕获具有关节的对象等的变形,特别是包括旋转运动。

[0190] 应当注意,只要  $f_2(i, j)$  是容许评估像素的运动的相似性的距离值,则任何值都能够用于  $f_2(i, j)$ ,且例如可以以使用等式 1 的仿射参数  $A'_t$  的等式 37 或等式 38 来代替等式 34 至 39。等式 37 能够使用像素 i 和像素 j 的运动向量之间的角度差和角度差的波动成分来表述像素的运动的相似性。这容许捕获包括旋转的运动改变。等式 38 能够使用像素 i 和像素 j 的相邻块的仿射参数的差和该差的波动成分来表述像素的运动的相似性。这容许捕获包括旋转、平移移动、以及阶梯跃迁 (scale transition) 的运动改变。

[0191]  $f_2(i, j) = a\_mean_{ij} + w_a \cdot \sqrt{a\_var_{ij}}$  等式 37

[0192]  $a\_mean_{ij} = \frac{1}{T-1} \sum_{t=1}^{T-1} a'_{ij}$

[0193]  $a\_var_{ij} = \frac{1}{T-1} \sum_{t=1}^{T-1} (a'_{ij} - a\_mean_{ij})^2$

$$[0194] \quad a'_{ij} = \arctan\left(\frac{y'_{i+1} - y'_i}{x'_{i+1} - x'_i}\right) - \arctan\left(\frac{y'_{i+1} - y'_i}{x'_{i+1} - x'_i}\right)$$

$$[0195] \quad f_2(i, j) = a\_mean_{ij} + w_a \sqrt{a\_var_{ij}} \text{ 等式 38}$$

$$[0196] \quad a\_mean_{ij} = \frac{1}{T-1} \sum_{i=1}^{T-1} a'_{ij}$$

$$[0197] \quad a\_var_{ij} = \frac{1}{T-1} \sum_{i=1}^{T-1} (a'_{ij} - a\_mean_{ij})^2$$

$$[0198] \quad a'_{ij} = \|A'_i - A'_j\|$$

[0199] 接下来,在步骤 S205a 中,非线性空间处理单元 104a 使用等式 7 至 10 所示的处理计算对应于等式 3 的距离  $f(i, j)$  的  $g_{\text{new}}(i, j)$ 。此外,除以上计算外,计算对应于基于像素  $i$  和像素  $j$  之间的角度的距离  $f_2(i, j)$  的  $g_{2\_new}(i, j)$ 。也就是说,单独计算两个非线性化距离。虽然于此描述了像素之间的距离和像素之间的角度,但是可以使用表述其它像素的运动的相似性的距离来计算三个或更多个距离。

[0200] 接下来,在步骤 S206a 中,非线性空间处理单元 104a 以与等式 11 相同的方式根据单独非线性化的距离  $g_{2\_new}(i, j)$  来生成  $G_2$ 。此外,类似地对  $G_2$  执行等式 12 至 16 所示的处理,  $G_2$  被投影到维数缩减空间,如等式 16,且计算  $z'_{2\_p2}$ 。

[0201] 接下来,在步骤 S207 中,非线性空间处理单元 104a 对  $z'_{2\_p2}$  和按等式 16 计算的  $z'_p$  进行整合,并执行群集,以探测对象并同时对象执行分割。具体地,通过如下重写等式 19 和 20 能够将其实现。

$$[0202] \quad \overline{z_m} = \begin{bmatrix} \overline{z_1^m} \\ \vdots \\ \overline{z_p^m} \\ \overline{z_{2\_1}^m} \\ \vdots \\ \overline{z_{2\_P}^m} \end{bmatrix} \text{ 等式 39}$$

$$[0203] \quad Z_m = \frac{1}{C_m} \sum_{c_m=1}^{C_m} \begin{bmatrix} z_1^{c_m} & - & \overline{z_1^m} \\ & \vdots & \\ z_p^{c_m} & - & \overline{z_p^m} \\ z_{2\_1}^{c_m} & - & \overline{z_{2\_1}^m} \\ & \vdots & \\ z_{2\_P}^{c_m} & - & \overline{z_{2\_P}^m} \end{bmatrix} \left[ z_1^{c_m} - \overline{z_1^m} \cdots z_p^{c_m} - \overline{z_p^m} \quad z_{2\_1}^{c_m} - \overline{z_{2\_1}^m} \cdots z_{2\_P}^{c_m} - \overline{z_{2\_P}^m} \right] \text{ 等式}$$

40

[0204] 以下,在步骤 S207a 中,等式 21 至 24 与以前的相同,不对它们进行描述。

[0205] 如上述,根据本实施例的图像处理设备使用针对图像上的像素间的距离与距离的波动的距离函数和针对像素间的角度与角度的波动的距离函数来计算每个距离函数的距离矩阵、将两个不同的非线性空间进行整合、并在将每个距离矩阵投影到分别的非线性空

间后执行群集,以便以更高精度探测并执行分割作为时间上追踪图像中移动的物体的区域的结果。此外,使用图像上像素间的角度和角度的波动作为距离函数容许更正确地捕获旋转运动,特别是由结合点运动引起的旋转运动。此外,因为不必设定人的候选区域作为预处理,不存在探测人的候选区域时的错误引起的分割中的失败。如上述,探测具有变形的诸如移动的人的对象并以更高精度对区域执行分割是可能的。

[0206] (第三实施例)

[0207] 在本实施例中,将描述图像处理设备,其中,增加了用于基于使用第一实施例中所述的图像处理方法对对象进行探测并执行分割的结果来预测对象的运动的功能。如图 15 中所示,对象运动预测单元 1201 增加至图像处理设备,该图像处理设备包括图像输入单元 101、运动分析单元 102、距离计算单元 103、以及非线性空间处理单元 104,并且因此,图像处理设备具有用于预测对象的运动的功能。

[0208] 对象运动预测单元 1201 基于分割的结果根据每个分割中包括的像素的时间运动轨迹来计算典型轨迹,并根据典型轨迹来预测对象的运动。图 16 示出了根据第三实施例的处理的流程图。因为步骤 S201 至 S206 与第一实施例中相同,这里不对它们进行描述。

[0209] 接下来,在步骤 S1301 中,对象运动预测单元 1201 使用步骤 S207 中执行的分割的结果如下获得群集的典型点和典型轨迹。

[0210] 这里,对应于非线性空间中的属于图 7(a) 和 (b) 以及图 8 中所示的分割区域  $\theta_m$  的数据  $z_{c_m}$  的像素的时间运动轨迹表述为  $x^{c_m}$ 。如前述,因为非线性空间中的数据  $z_p^i$  和像素  $i$  的时间运动轨迹  $x^i$  是一一对应关系,所以当选择属于群集区域  $\theta_m$  的数据  $z_{c_m}$  时,自动确定对应于像素的时间运动轨迹  $x^{c_m}$ 。此外,如以下等式,对每个群集区域  $\theta_m$  获得典型时间运动轨迹。虽然这里的范例计算平均时间运动轨迹作为典型时间运动轨迹,但是可以执行对像素的时间运动轨迹  $x^{c_m}$  的每一个的加权等,或者可以将对应于图像上群集的重力中心的像素的时间运动轨迹用作典型时间运动轨迹。

$$[0211] \quad \overline{x^m} = \frac{1}{C_m} \sum_{c=1}^{C_m} x^{c_m} \text{等式 41}$$

[0212] 图 17 示出了一个范例,其中,基于等式 41 对每个群集  $\theta_m$  获得了典型时间运动轨迹。应当注意,考虑到可视性,仅示出了关于对应于头的群集区域  $\theta_1$  的和关于对应于小腿的群集区域  $\theta_8$  的典型时间运动轨迹。图中的十字形是分别对应于时间  $t$  的  $\overline{x^m}$  的元并且每一个表示像素位子。此外,如等式 3 和 33 所示,因为考虑像素的运动的相似性来执行非线性空间中通过群集进行的分割,所以仅使用具有相似运动的像素的时间运动轨迹来进行计算是可能的。因此,与用于简单地获得相邻像素的运动轨迹的平均的方法相比,能够以更高精度获得典型时间运动轨迹。以此方式,给每个群集获得典型运动轨迹容许精确并容易地表述每个区域的运动。

[0213] 接下来,在步骤 S1302 中,对象运动预测单元 1201 基于步骤 S1301 中计算的典型时间运动轨迹来预测时间  $T$  后的时间处对象的位置。首先,将描述范例,其中,根据典型时间运动轨迹计算加速度,并预测  $T+1$  后对象的位置。在输入三个或更多个时间系列图像的情况下,能够对每个典型时间运动轨迹  $\overline{x^m}$  获得加速度向量  $s^m$ ,如以下等式:

$$[0214] \quad s_{t+1}^m = \{u_{t+1}^{m1} - u_t^m\} - \{u_t^m - u_{t-1}^n\} \text{等式 42}$$

[0215] 其中,  $u_t^m$  是运动向量, 并能够按以下等式表述:

[0216]  $u_t^m = (u_t^m, v_t^m)$  等式 43

[0217] 如图 17 中的虚箭头和圆所示, 能够使用等式 43 的加速度向量针对对象的每个区域如下预测时间  $T+t'$  时对象的区域的位置  $pos_m(T+t')$ 。虽然这里将对象的区域用作范例, 但是预测诸如图 8 中所述的自行车和步行者的每个对象的位置是可能的。

[0218]  $pos_m(T+t') = pos_m(T) + t' u_t^m + \frac{1}{2} t'^2 s_t^m$  等式 44

[0219] 这容许考虑加速度的预测。在运动突然变得迅速或突然停止的情况下, 利用该加速度预测对象的位置是可能的。此外, 可以使用仿射参数代替运动向量。仿射参数能够表述包括旋转运动的运动并适于表述手臂和腿的旋转运动。

[0220] 此外, 代替上述运动向量和加速度, 直接拟合 N 次函数为典型时间运动轨迹  $x^m$  是可能的。在输入 T 时间系列图像的情况下, 拟合 N 次函数为包括在 T 图像上的  $\overline{x^m}$  中的位置信息是可能的。以此方式, 能够评估时间  $(T+t')$  时图像上的位置  $pos_m(T+t')$ , 使得其符合被执行了拟合的函数的值。具体地, 因为利用函数拟合容许表述较平滑的运动, 所以更高精度的预测变得可能。理所当然, 对这些图像上的位置的预测能够用于时间插值。

[0221] 如上述, 根据本实施例, 因为能够作为相同群集计算具有相似运动的像素的时间运动轨迹, 所以能够高精度地获得典型时间运动轨迹。能够表述特别是具有关节的对象等的每个区域的运动, 并且高精度地预测对象的区域的位置, 而无需设定人的候选区域作为预处理是可能的。

[0222] 虽然已经基于实施例和修改例描述了根据本发明的图像处理方法和图像处理设备, 但是本发明不限于这些实施例和修改例。本发明包括通过在本发明的要旨的范围内给每个实施例施加本领域技术人员可设想的各种修改而实现的实施例。

[0223] 此外, 本发明包括通过任意组合每个实施例的特征部件而实现的实施例。

[0224] 本发明能够用于图像处理设备, 该图像处理设备基于图像中具有一个或多个像素的块的运动来执行对包括具有变形的诸如移动的人的对象的图像的分割, 以便探测对象, 例如, 可以是包括在运动分析装置、监控装置、以及诸如摄像机和 TV 的 AV 装备中的图像处理设备。

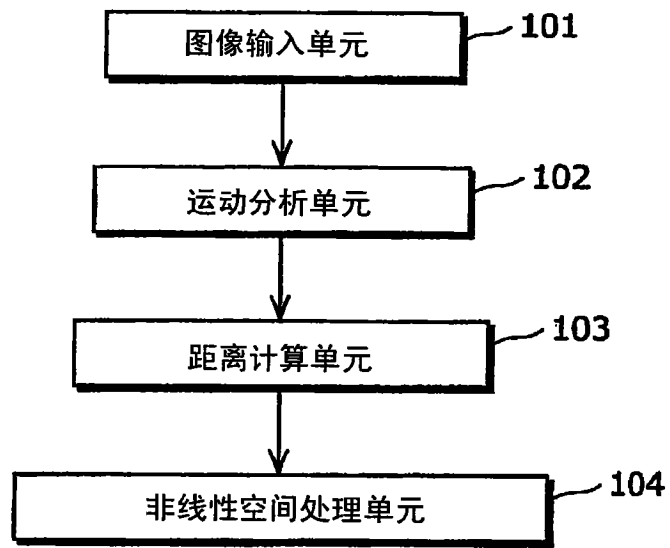


图 1

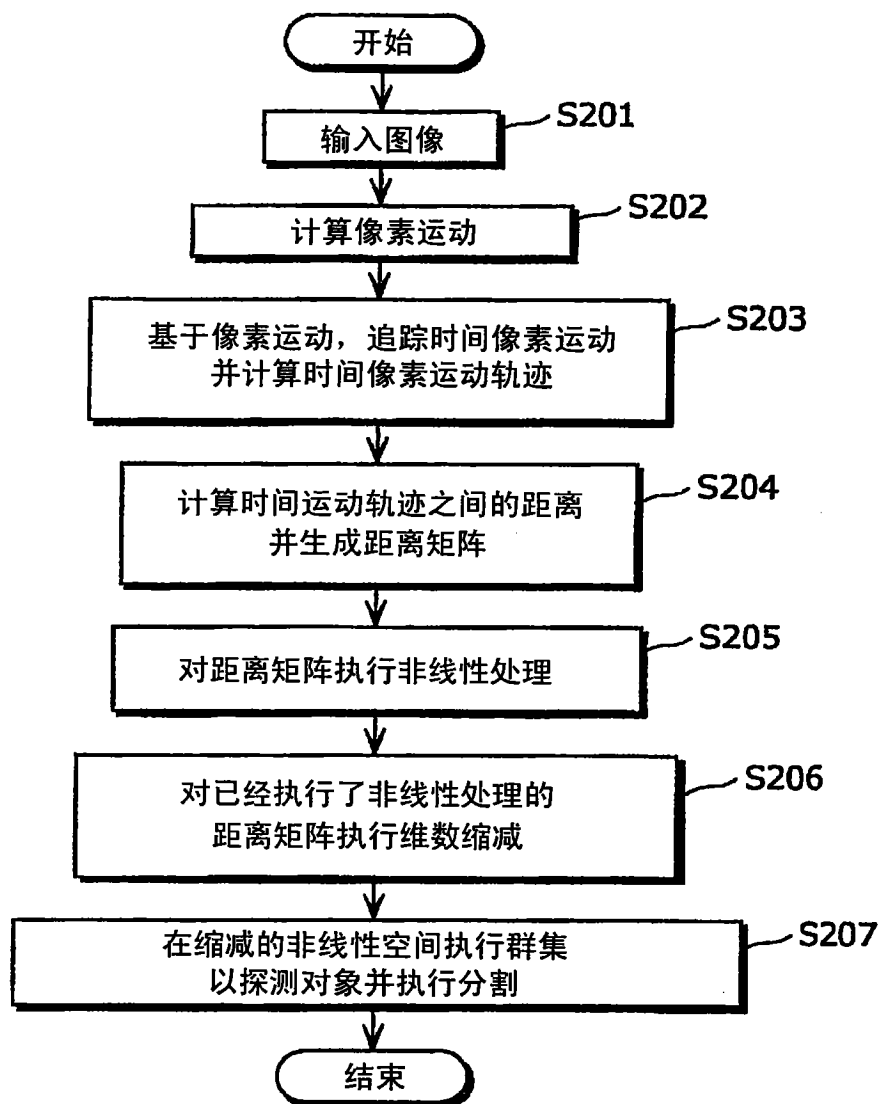


图 2

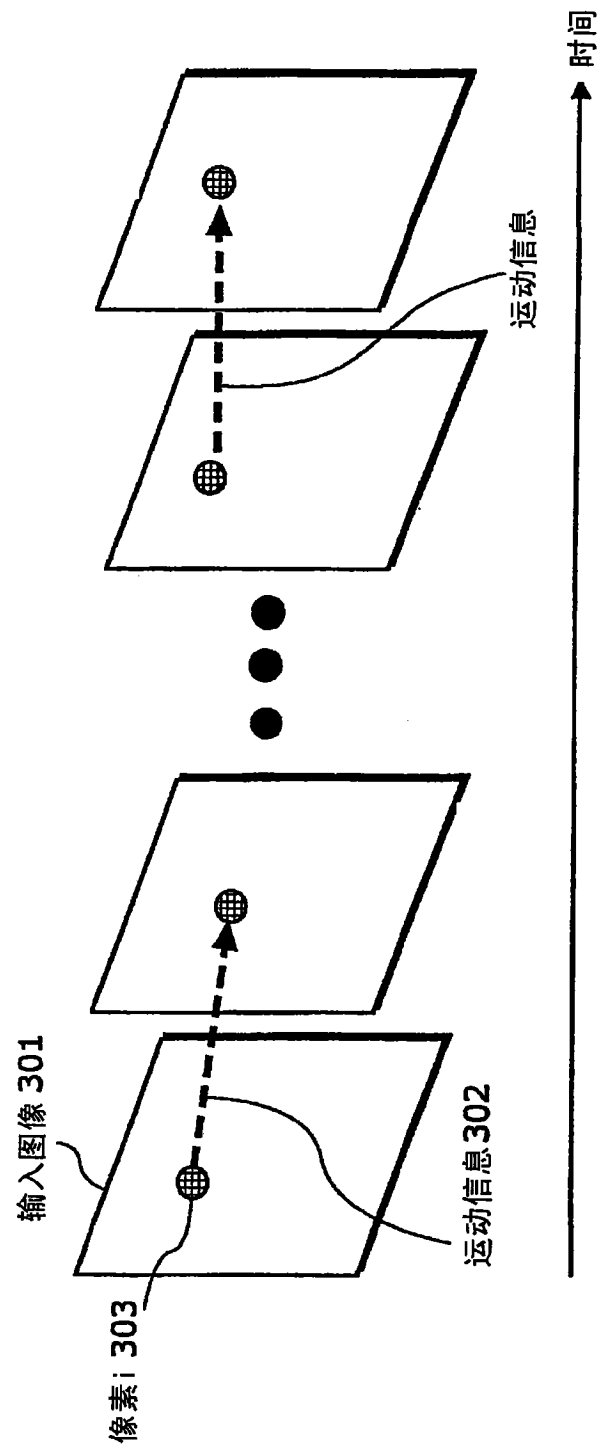


图 3

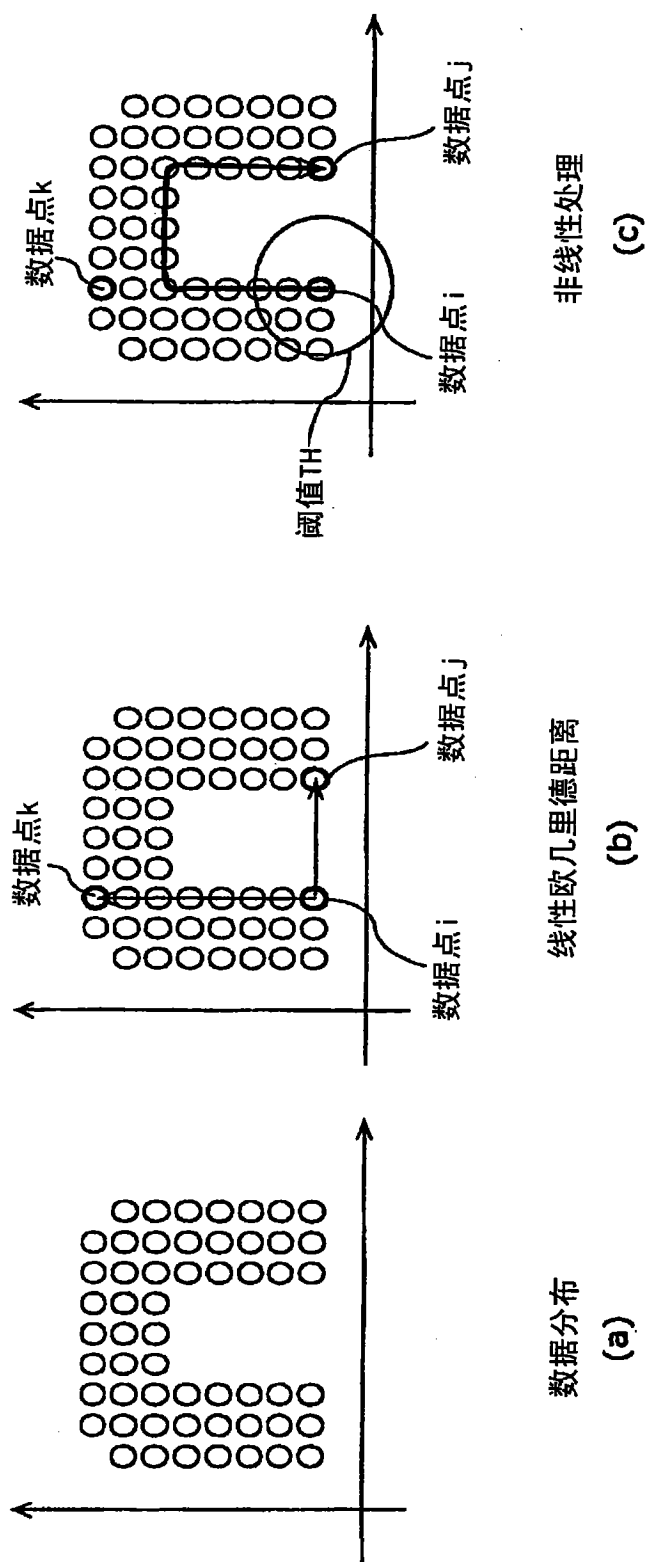


图 4

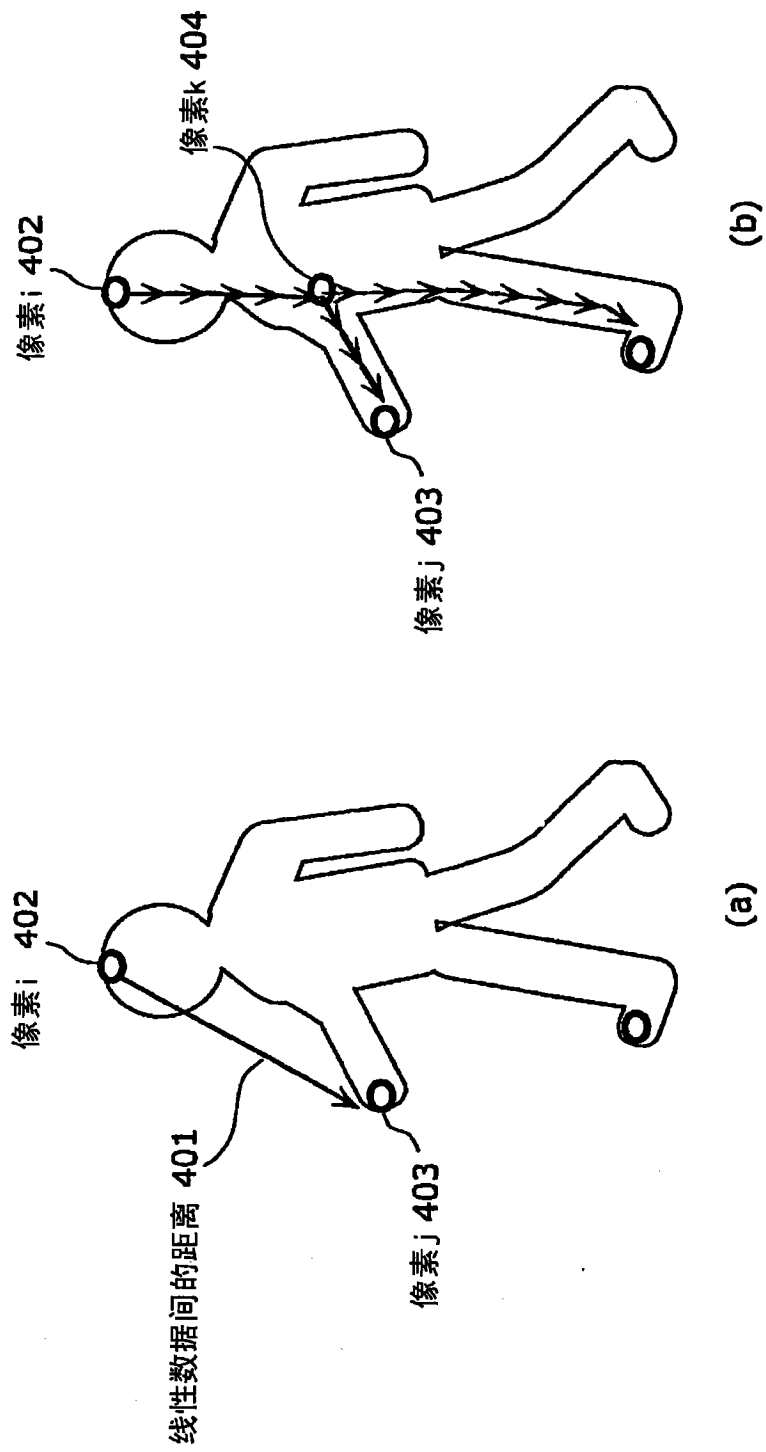


图 5

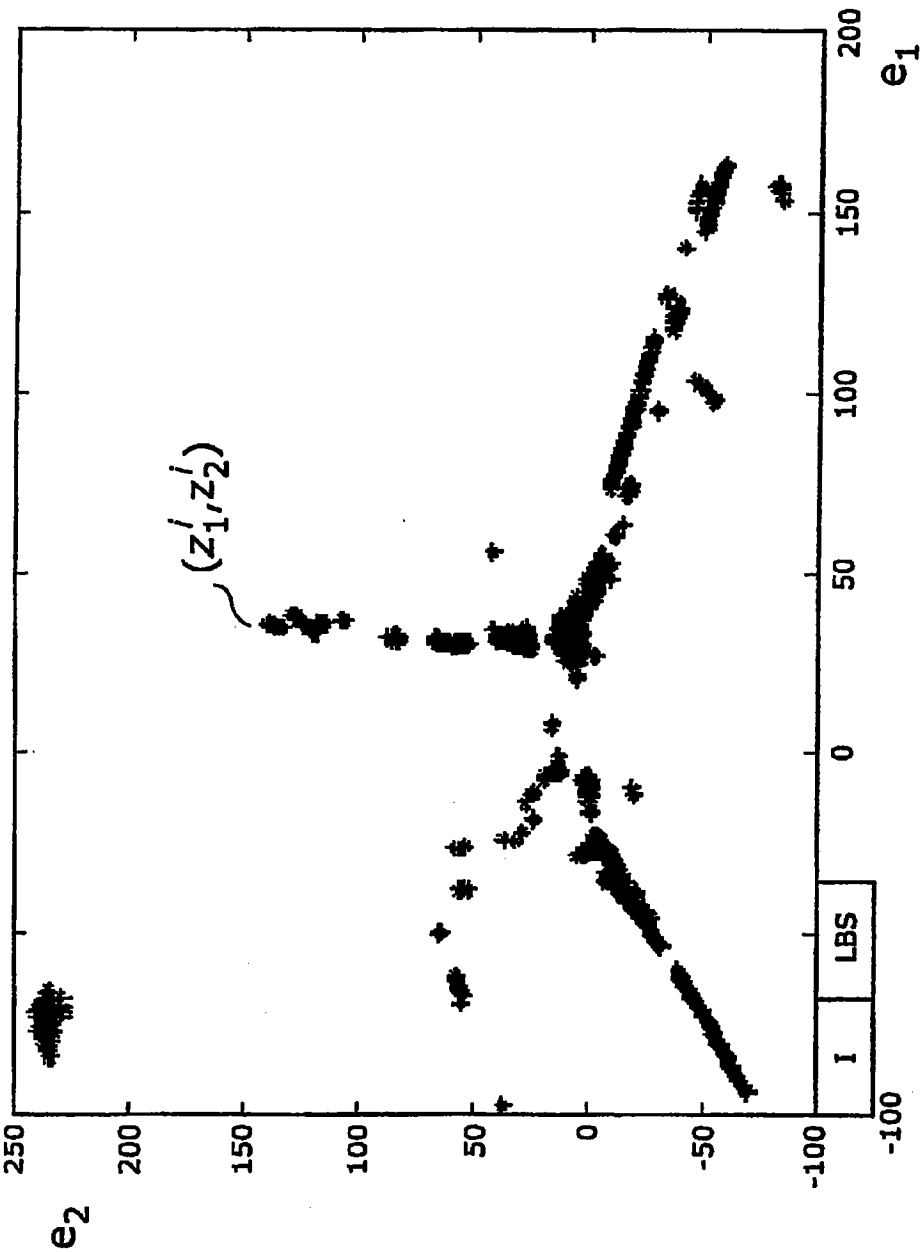


图 6

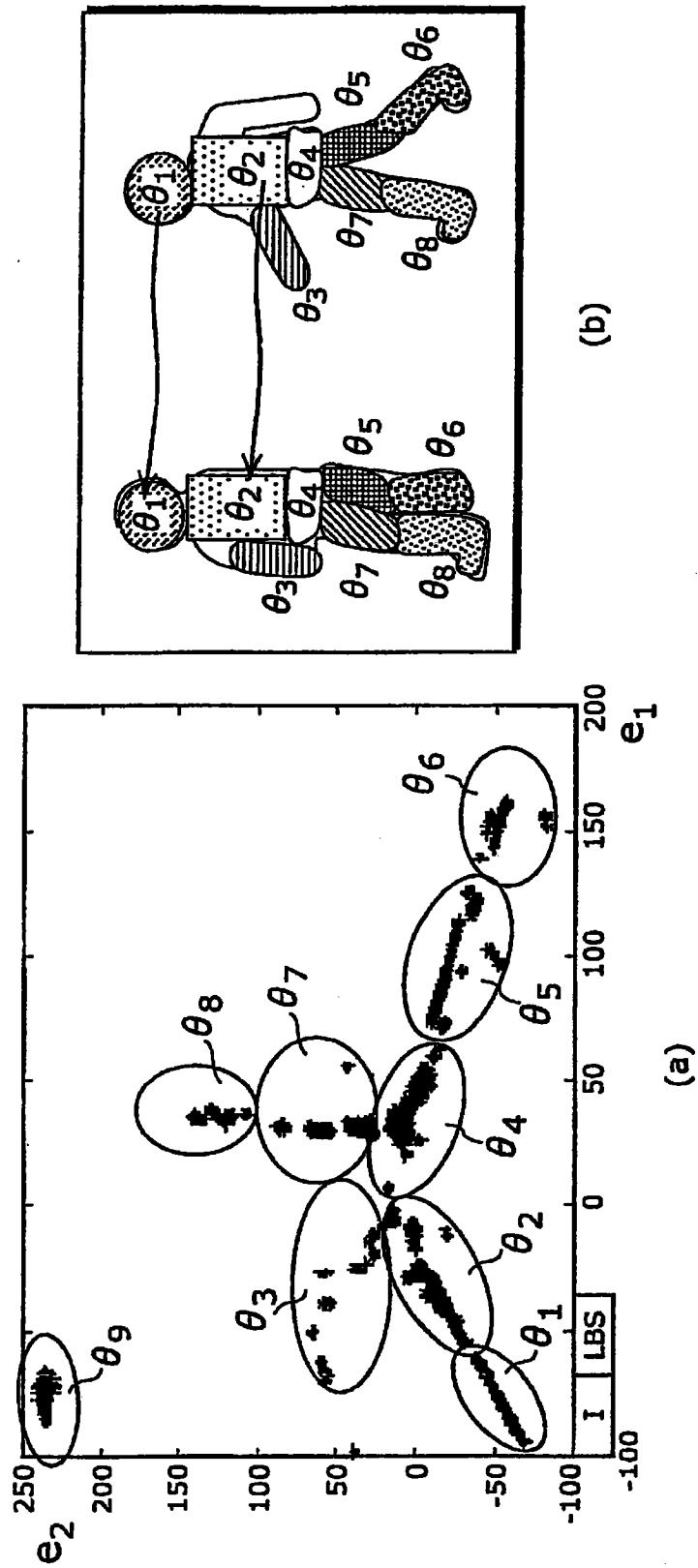


图 7

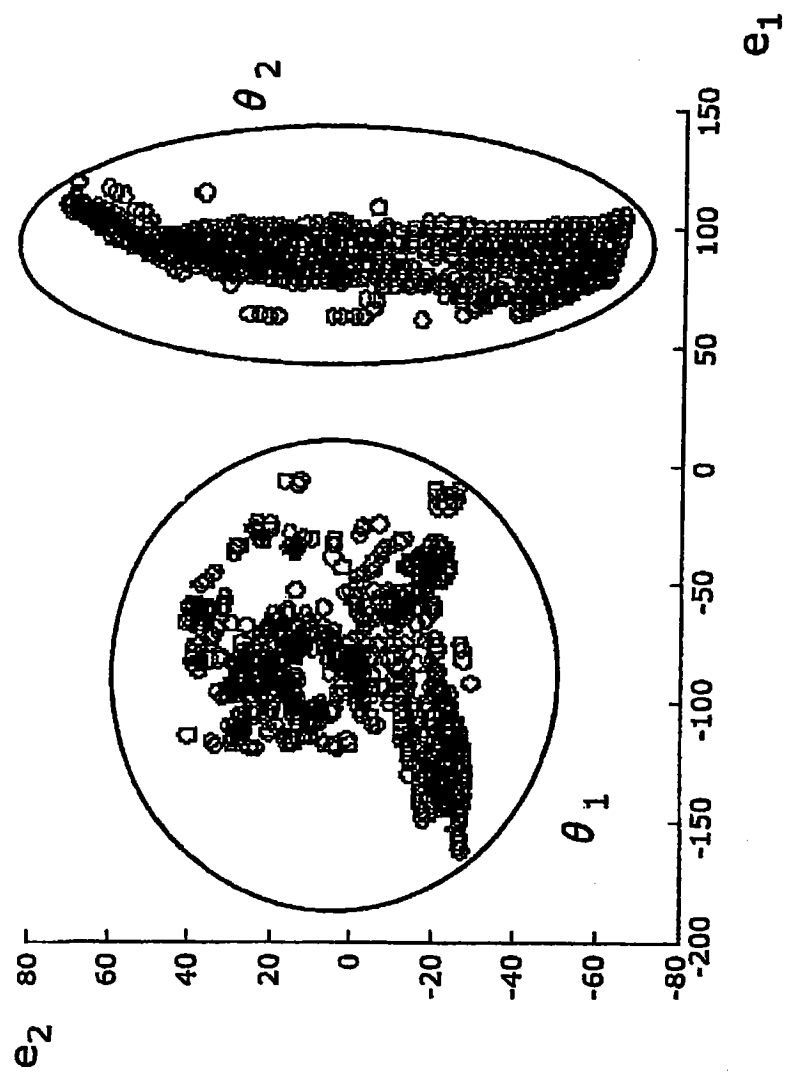


图 8

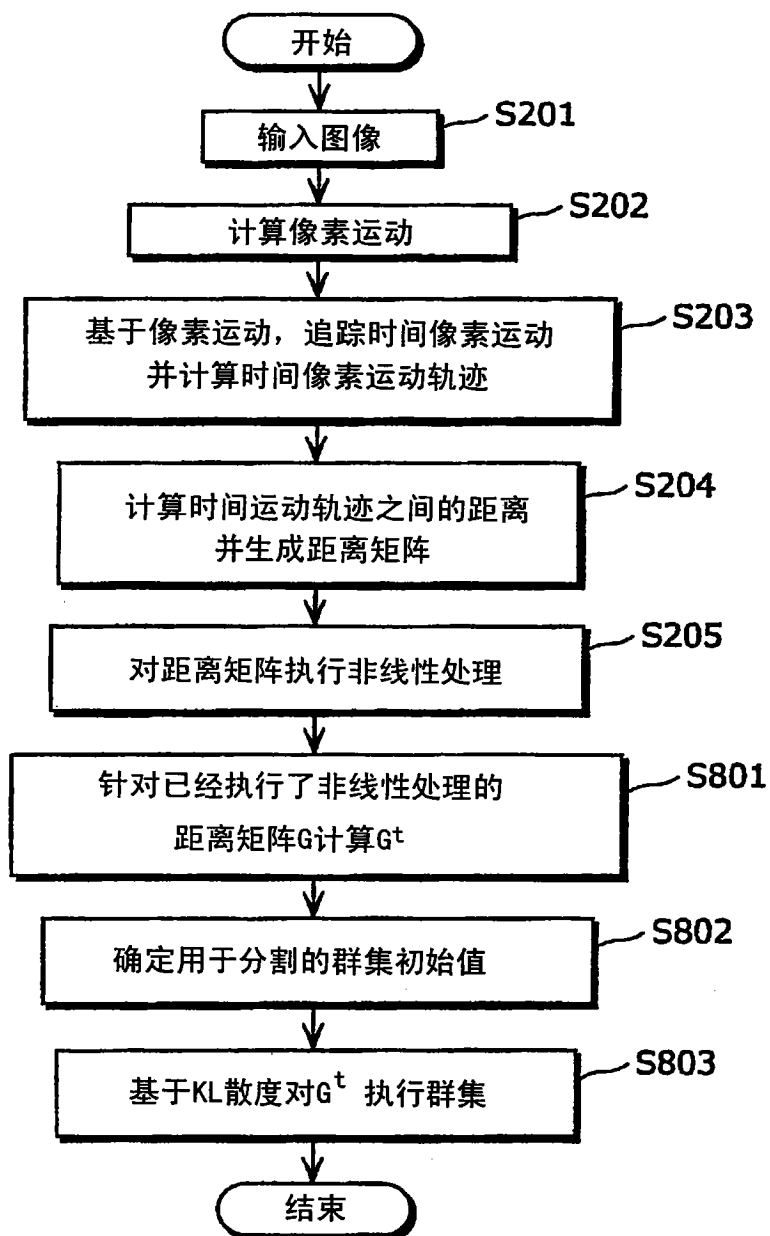


图 9

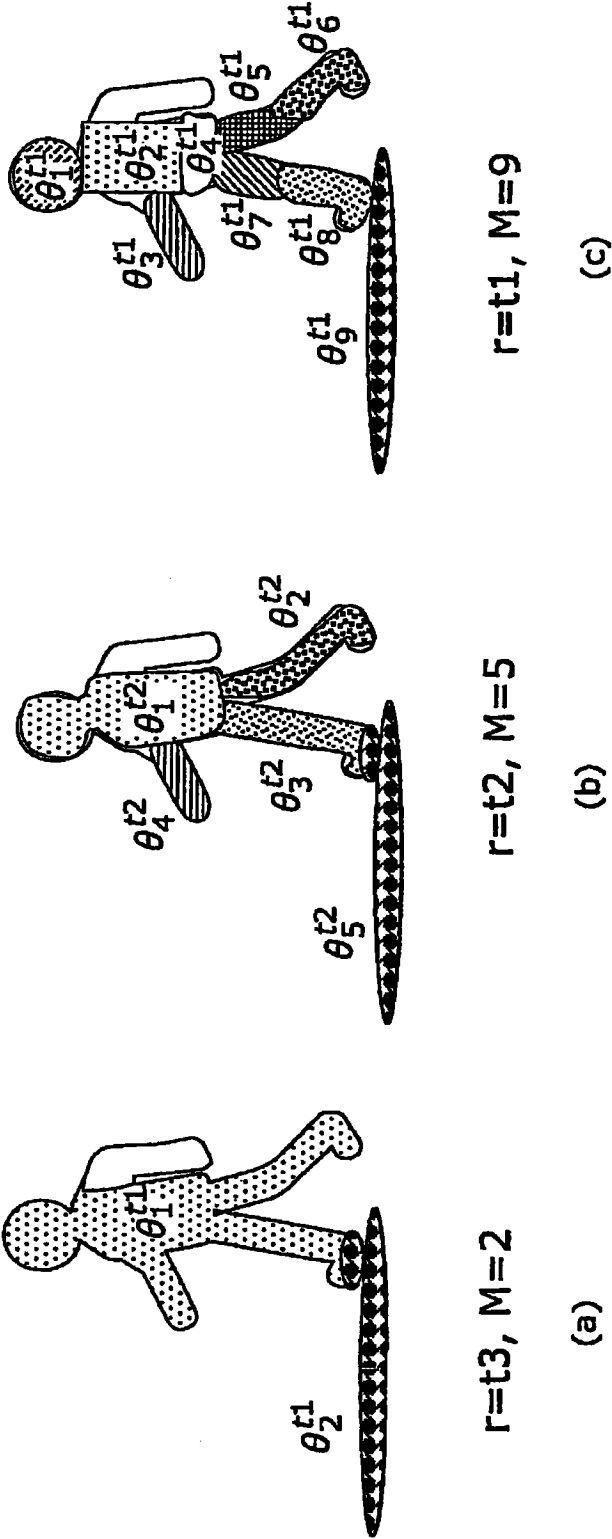


图 10

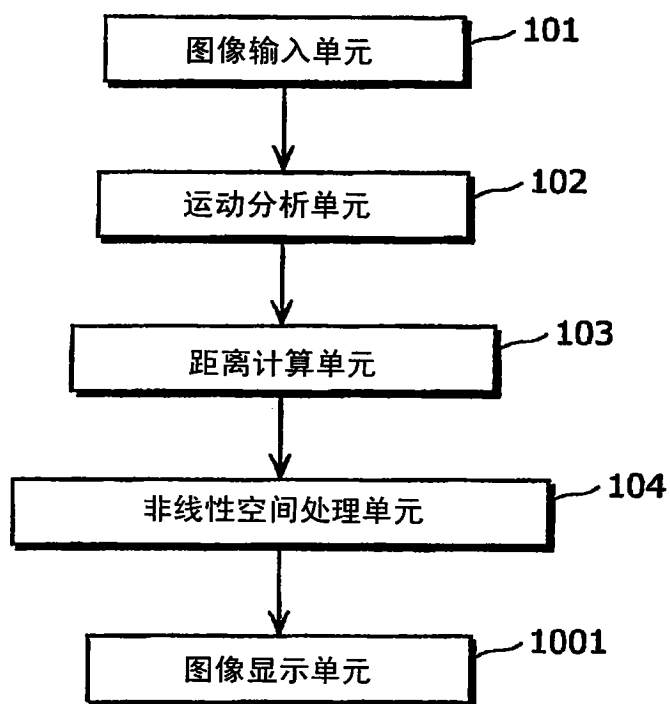


图 11

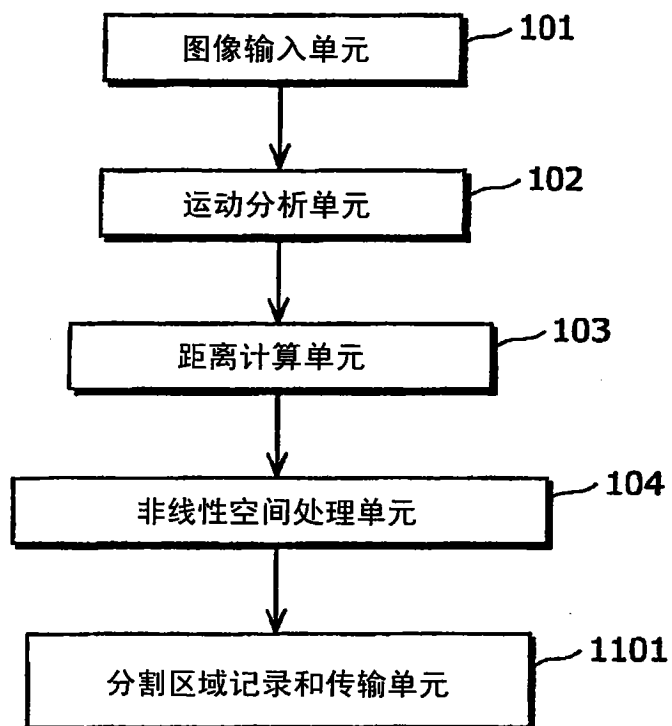


图 12

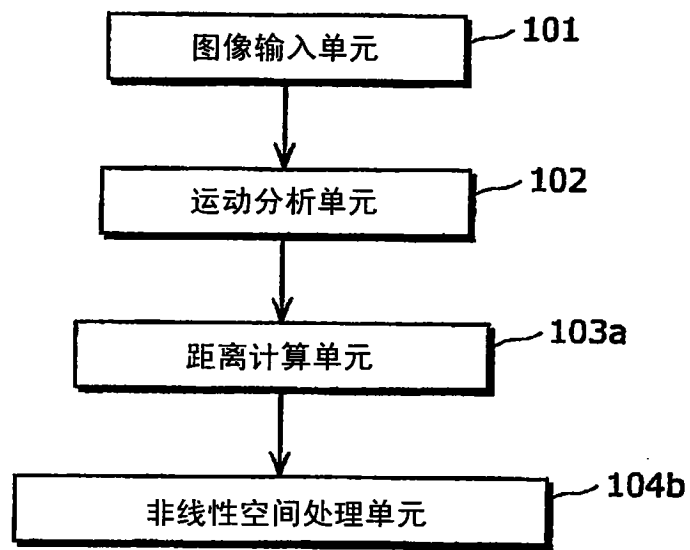


图 13

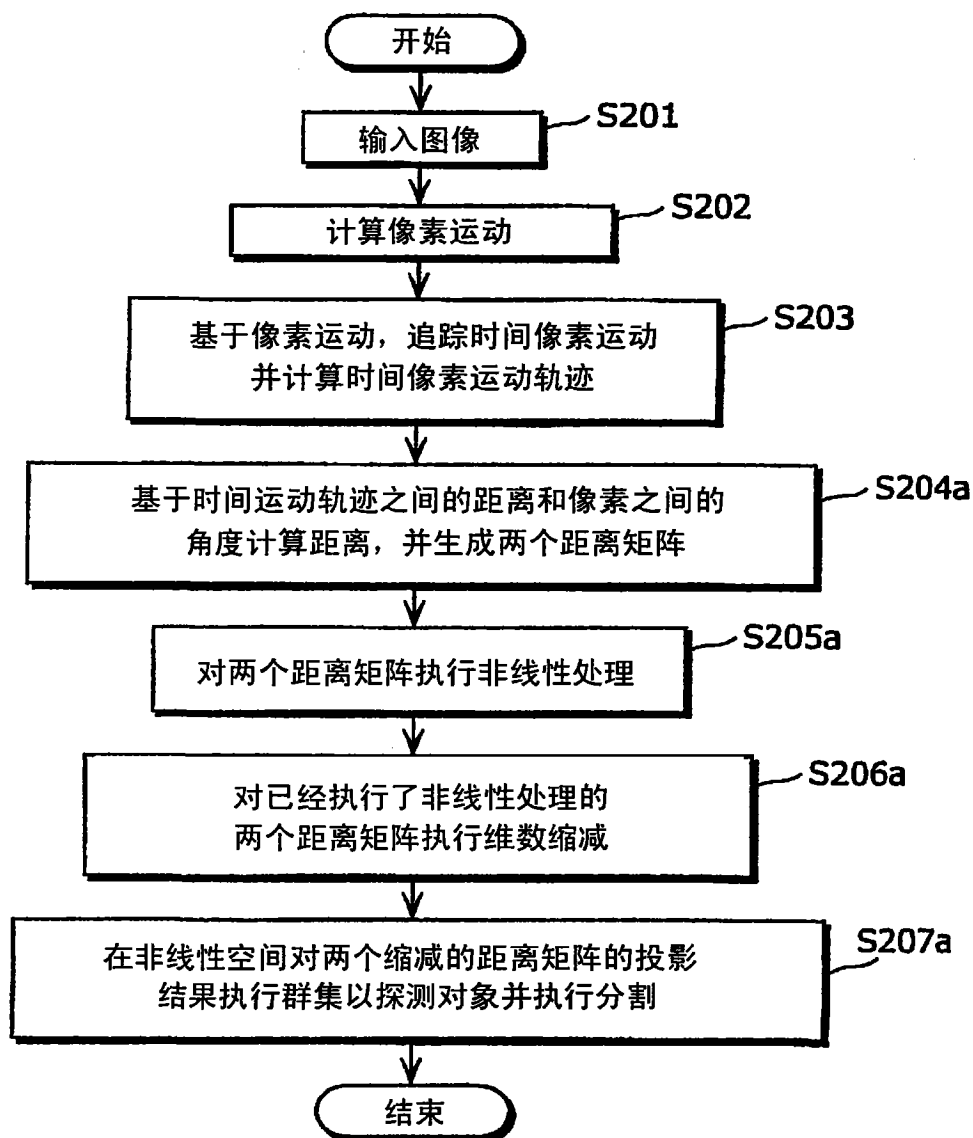


图 14

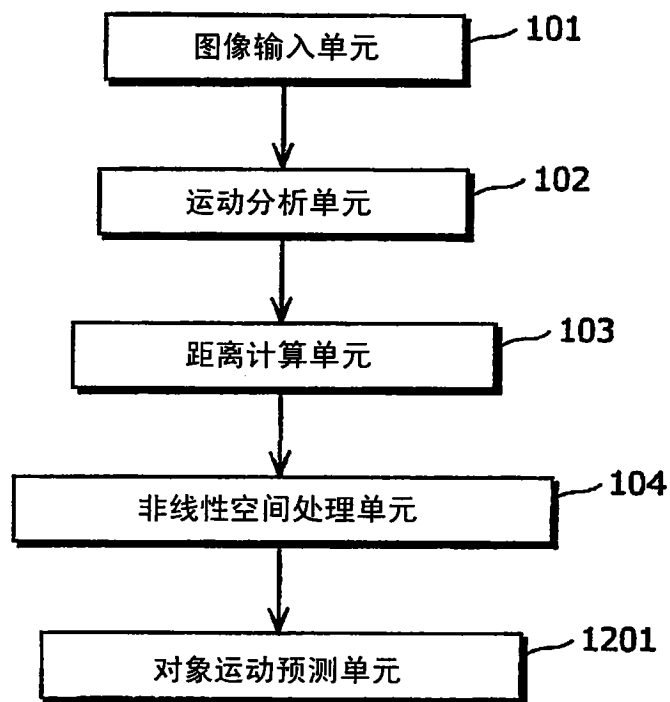


图 15

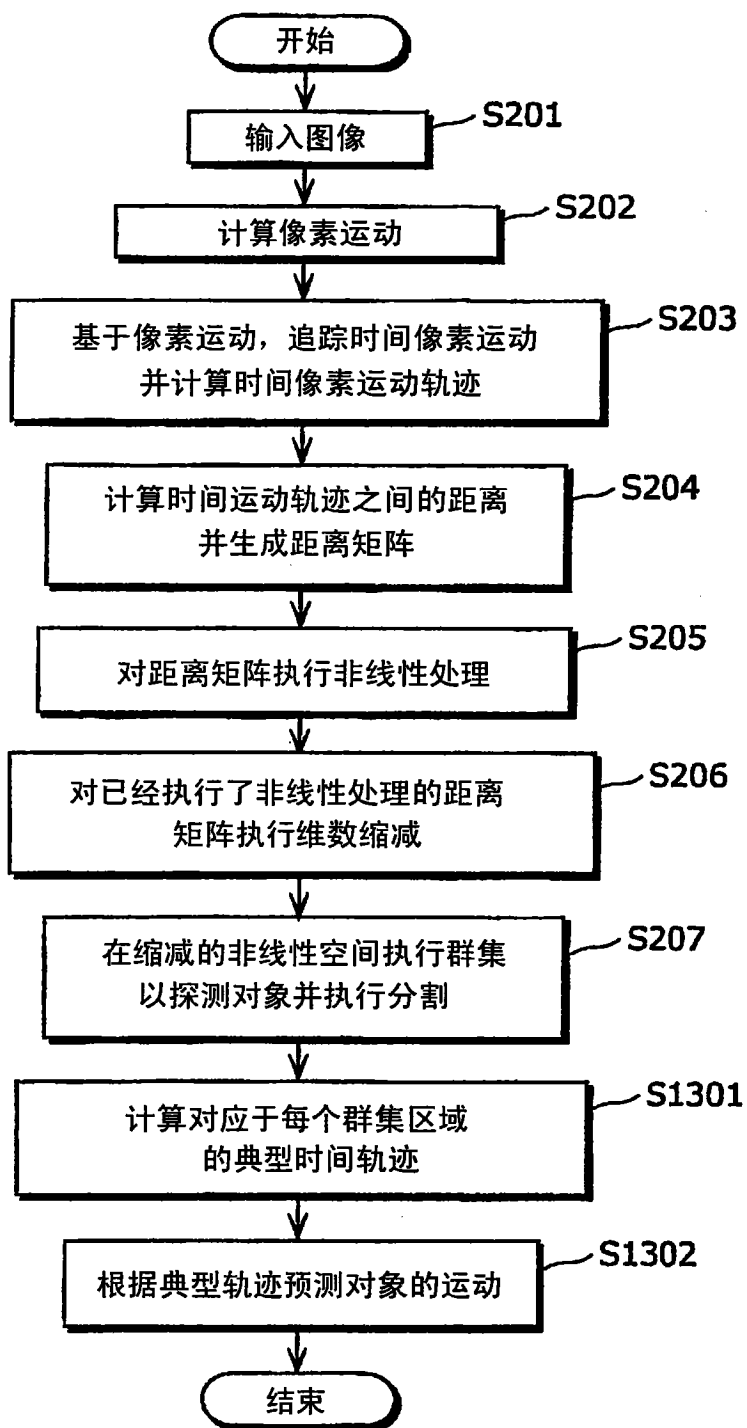


图 16

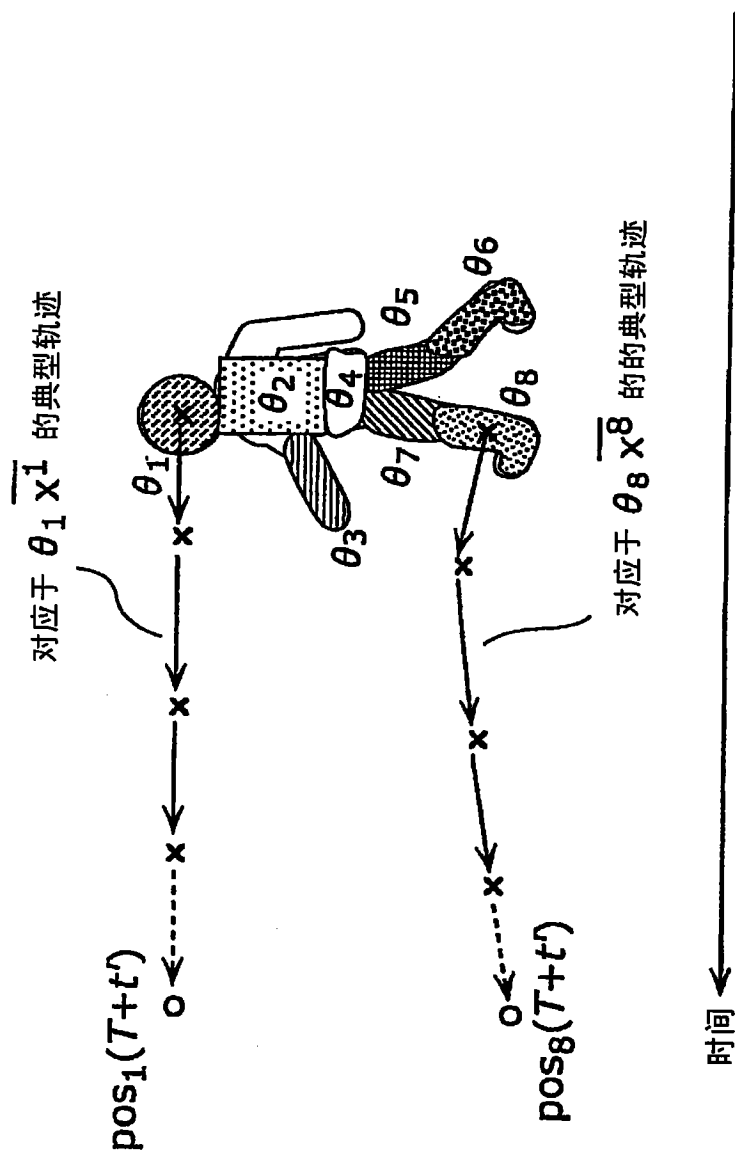


图 17