



(12) 发明专利

(10) 授权公告号 CN 101017572 B

(45) 授权公告日 2010.07.21

(21) 申请号 200710008088.1

审查员 巴晓艳

(22) 申请日 2007.02.09

(30) 优先权数据

11/352,145 2006.02.09 US

(73) 专利权人 三菱电机株式会社

地址 日本东京

(72) 发明人 法提赫·M·波里克里

昂塞尔·图塞尔

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 马浩

(51) Int. Cl.

G06T 7/20(2006.01)

H04N 5/14(2006.01)

(56) 对比文件

W0 2005/048196 A2, 2005.05.26, 第2页第25-35行, 第6页第4段, 第7页, 第8页第1段至第9页第5段, 第10页第3段.

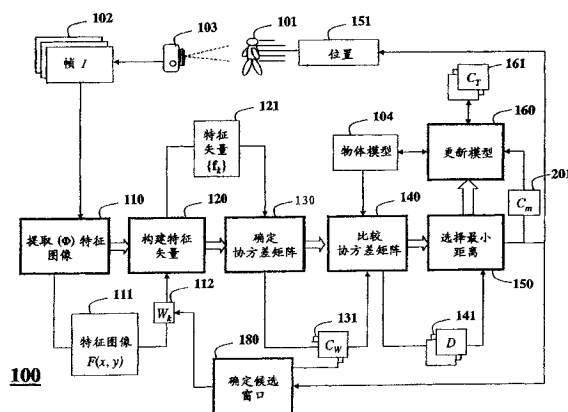
权利要求书 3 页 说明书 8 页 附图 6 页

(54) 发明名称

用于在帧序列中跟踪物体的计算机化的方法

(57) 摘要

提供了一种用于在视频帧序列中跟踪非刚性物体的方法。物体的特征从视频中提取。这些特征包括像素的位置和像素的属性。这些特征用于构建一个协方差矩阵。该协方差矩阵用于跟踪的目的作为物体的描述符。通过基于李代数平均的更新机制来管理物体变形和外表改变。



1. 一种用于在帧序列中跟踪一个物体 (101) 的计算机化的方法, 包括下列步骤:
 从获得的物体的帧序列中的每一帧中提取 (110) 一个特征图像;
 对每个特征图像构建 (120) 一个特征矢量集, 每个特征矢量对应于所述特征图像中的
 诸个特征的不同窗口, 所述窗口是根据物体大小而定义的搜索区域;
 从所述特征矢量集中确定 (130) 一个对应的协方差矩阵集合;
 度量 (140) 所述协方差矩阵集中的每个协方差矩阵和表示所述物体的一个模型协方
 差矩阵之间的距离;

选择 (150) 一个特定窗口作为所述物体在所述帧中的位置, 该特定窗口具有到所述模
 型协方差矩阵有最小距离的对应的协方差矩阵。

2. 如权利要求 1 所述的方法, 还包括:

对每一个帧, 更新 (160) 所述模型协方差矩阵。

3. 如权利要求 1 所述的方法, 其中帧是多谱图像。

4. 如权利要求 1 所述的方法, 其中根据所述物体在所述帧中的估计位置来选择不同的
 窗口。

5. 如权利要求 1 所述的方法, 其中每个特征矢量都包括所述物体的空间属性和外表属
 性。

6. 如权利要求 5 所述的方法, 其中从所述特征图像中的像素坐标中获得所述空间属
 性, 并且所述外表属性包括像素强度。

7. 如权利要求 1 所述的方法, 其中对于特定窗口 W 的协方差矩阵 C_W 为

$$C_W = \frac{1}{MN} \sum_{k=1}^{MN} (f_k - m_\mu)(f_k - m_\mu)^T,$$

其中 m_μ 是对应的特征矢量 f_k 的均值的矢量, M 和 N 是输入图像的维数, T 是转置运算
 符。

8. 如权利要求 1 所述的方法, 其中不同的窗口在所述特征图像中具有不同的大小和位
 置。

9. 如权利要求 1 所述的方法, 其中两个矩阵 C_i 和 C_j 间的距离为

$$D(C_i, C_j) = \sqrt{\sum_{k=1}^d \ln^2 \lambda_k(C_i, C_j)}$$

其中 $\{\lambda_k(C_i, C_j)\}$ 为矩阵 C_i 和 C_j 的归一化的本征值, 并由下式确定

$$\lambda_k C_i x_k - C_j x_k = 0 \quad k = 1 \dots d,$$

对于与 d 个特征相对应的 d 个本征值, 其中 x_k 为归一化本征矢量。

10. 如权利要求 2 所述的方法, 其中所述更新 (160) 使用协方差矩阵集合的子集来确定
 一个固有均值协方差矩阵作为所述模型协方差矩阵。

11. 如权利要求 10 所述的方法, 其中使用李代数确定所述均值协方差矩阵。

12. 如权利要求 10 所述的方法, 还包括:

根据每个协方差矩阵与所述模型协方差矩阵之间的距离, 将每个协方差矩阵距离乘以
 (172) 所述子集中的权重。

13. 如权利要求 10 所述的方法, 还包括:

为所述子集中的每个协方差矩阵指定恒定的权重值。

14. 如权利要求 10 所述的方法,还包括:

将当前模型协方差矩阵包括 (311) 在所述子集中;以及
增加 (312) 所述子集中的协方差矩阵的数目。

15. 如权利要求 10 所述的方法,还包括:

从所述子集中移除 (321) 最老的协方差矩阵;
将当前模型协方差矩阵包括 (323) 在所述子集中以保持矩阵的总数目不变。

16. 如权利要求 10 所述的方法,还包括:

保持 (161) 先前物体位置的一个子集以及在所述窗口中的对应特征;以及
直接由所述特征计算 (140) 所述模型协方差矩阵。

17. 如权利要求 1 所述的方法,还包括:

对于每一个帧,更新模型特征均值矢量。

18. 如权利要求 17 所述的方法,还包括:

度量 (176、177) 所述模型特征均值矢量和窗口特征均值矢量之间的距离以计算一个
特征均值距离;

用协方差融合权重乘以 (172) 每个协方差矩阵与所述模型协方差矩阵之间的距离;

用协方差融合权重乘以 (171) 所述均值距离;

计算所述每个协方差矩阵与所述模型协方差矩阵之间的距离和特征均值距离的加权和;

指定 (174) 该加权和为模型与当前窗口之间的距离;

选择 (150) 一个特定窗口作为所述物体在搜索区域中的位置,该特定窗口具有最小加权和。

19. 如权利要求 6 所述的方法,其中所述空间属性是像素的空间坐标值。

20. 如权利要求 6 所述的方法,其中所述空间属性是该像素距离所述窗口的原点的距离。

21. 如权利要求 6 所述的方法,其中所述空间属性是像素的空间坐标值相对于窗口中的原点的绝对值。

22. 如权利要求 1 所述的方法,还包括:

由所述物体位置的先前位置预测 (180) 在当前图像中的搜索区域;

在所述搜索区域中构建 (120) 一个特征矢量集;

从所述搜索区域中的所述特征矢量集中确定 (130) 协方差矩阵的一个对应集合;

度量 (140) 该集合中的每个协方差矩阵与在所述搜索区域中表示所述物体的模型协方差矩阵之间的距离;

选择 (150) 一个特定窗口作为所述物体在所述搜索区域中的位置,该特定窗口具有到所述模型协方差矩阵有最小距离的对应的协方差矩阵。

23. 如权利要求 1 所述的方法,还包括:

通过它的协方差矩阵的本征值集合来表示每个窗口。

24. 如权利要求 23 所述的方法,还包括:

从所述特征矢量集中确定对应的本征值集合;

度量该集合中的每个本征值集合与表示所述物体的模型本征值集合之间的距离；

选择一个特定窗口作为所述物体在所述帧中的位置，所述特定窗口具有到所述模型本征值集合有最小距离的对应的本征值集合。

25. 如权利要求 24 所述的方法，还包括：

对于每个帧，更新模型本征值集合。

用于在帧序列中跟踪物体的计算机化的方法

技术领域：

[0001] 本发明一般涉及计算机视觉，尤其涉及在视频中跟踪物体。

背景技术

[0002] 在帧序列或视频序列中，通过一帧帧地确定物体的特性的对应关系，可以跟踪物体。但是，精确地跟踪一个变形的、非刚性的并且快速移动的物体仍然是一个问题。

[0003] 可以通过均值移位操作可以执行跟踪，D.Comaniciu, V.Ramesh, 和 P.Meer 的“Real-time tracking of non-rigid objects using mean shift”, IEEE Conf. on Computer Vision and Pattern Recognition, 第 1 卷第 142-149 页, 2000 年。使用一个非参数的密度梯度估计量来跟踪与给定颜色直方图最近似的物体。该方法提供了精确的定位。但是，该方法要求在连续帧中物体位置要有一些重叠，这不适用于快速移动物体的情况。

[0004] 跟踪也可被看作是物体的状态的估计，给定到一个瞬间为止的所有测量值。这等价于构造所述物体位置的概率密度函数 (pdf)。通过递归贝叶斯滤波器可以提供最优解，其解决了连续预测和更新步骤中的问题。

[0005] 当假定测量噪声为高斯分布时，Kalman 滤波器提供了一种解，它经常用于跟踪刚性物体，Y.Boykov 和 D.Huttenlocher 的“Adaptive Bayesian recognition in tracking rigid objects”, IEEE Conf. on Computer Vision and Pattern Recognition, 第 2 卷第 697-704 页, 2000 年；以及 R.Rosales 和 S.Sclarroff 的“A framework for heading-guided recognition of human activity”, Computer Vision and Image Understanding, 第 91 卷第 335-367 页, 2003 年。Kalman 滤波器限于控制运动特性的“粘性”的预定状态转换参数。

[0006] 当状态空间是离散的，并由有限数目的状态构成时，马尔可夫滤波器可用于物体跟踪。最通用的一类滤波器的代表是基于蒙特卡洛积分法的粒子滤波器。具有相关权重的一组随机样本表示特定状态的当前密度。然后新的密度基于这些权重样本。

[0007] 粒子滤波器可用于恢复条件密度传播以用于视频跟踪和验证。一般说来，粒子滤波器基于随机取样，这是一个由样本退化和贫化引起的难题，尤其对于高维问题。基于内核的贝叶斯滤波器可用于更有效地对状态空间进行取样。多假设滤波器估计移动物体引起特定测量序列的概率。

[0008] 作为一个问题，所有上述基于滤波器的方法可很容易地被局部最优所阻。作为另一个担心，大部分现有技术都缺乏一个既表示统计又表示空间特性的令人满意的类似性准则。大部分现有技术或者仅依赖于颜色分布，或者仅依赖于结构模型。

[0009] 许多不同的表示（从积累统计到外表模型）被用于跟踪物体。柱状图是非常普遍的，因为标准化的柱状图非常类似于模型化数据的概率密度函数。但是，柱状图没有考虑特征值的空间排列。例如，在观察窗口中随机重新排列像素产生相同的柱状图。此外，用小数量的像素构建较高维数的柱状图是主要问题。

[0010] 外表模型映射图像特征（如形状和质地）到统一尺寸的张量窗口上。因为指数的

复杂性,仅可以使用相对较小数目的特征。因此,每个特征都必须是高阶判别式。特征的可靠性严格依赖于物体类型。外表模型趋向于高敏感于比例变形,也是依赖于姿势的。

[0011] 因此,希望提供一种用于在视频中跟踪物体的更好的方法。

发明内容

[0012] 本发明的一个实施方式提供了一种用于在视频帧序列中跟踪非刚性物体的方法。物体的特征从视频中提取。这些特征包括像素的位置和像素的属性。这些特征用于构建一个协方差矩阵。该协方差矩阵用于物体的描述符以用于跟踪的目的。

[0013] 协方差矩阵表示了物体的空间和统计属性,以及相同表示中的相关性。协方差矩阵使得不同类型的特征和模态可以有效地熔合在一起。与现有技术的直方图相比,协方差矩阵的维数较小。

[0014] 不同于现有技术的方法,不用做出对被跟踪物体的运动或位置的任何假定。此外,该方法并不受限于对后验概率函数的估计。相反,该方法提供了一种全局最优解,并显著地简化了对应关系问题。

[0015] 通过一种基于李代数平均 (Lie algebra averaging) 的更新机制来管理物体的变形和外表变化。

[0016] 基于协方差的跟踪能够精确地在非静态摄像序列中检测非刚性的移动物体,而平均来说获得近似完美的检测率,即使对于快速移动物体。

附图说明

[0017] 将基于下面的图对本发明的优选实施方式进行详细描述,其中:

[0018] 图 1 是一种根据本发明的一个实施方式的用于在视频中跟踪移动物体的方法的方框图;

[0019] 图 2 是根据本发明的一个实施方式的距离测定的方框图;

[0020] 图 3 是根据本发明的一个实施方式的候选窗口的选择的方框图;

[0021] 图 4 是根据本发明的一个实施方式的物体模型的更新的方框图;

[0022] 图 5 是根据本发明的一个实施方式向物体模型更新分配权重的方框图;

[0023] 图 6 是根据本发明的一个实施方式的用于将对应于最小距离的矩阵与先前协方差矩阵的数据集相加的两种方法的方框图。

具体实施方式

[0024] 基于协方差的物体跟踪

[0025] 图 1 示出了根据本发明的一个实施方式的用于在摄像头 103 所捕获的帧 102 或视频的序列中跟踪移动物体 101 的方法 100 的步骤。所示步骤对每一个帧 102 重复。我们假定我们具有物体 101 的模型 104。该模型的形式为如本文所述而构建的协方差矩阵。例如可以通过使用手动或自动物体定位方法而在第一个帧中定位物体而获得初始模型。此后,对每一帧 102 该模型被更新 160,如这里所要描述的。

[0026] 对于视频序列中的每一帧 102,我们构建 110 一个特征图像 111。对于特征图像 111 中的各个区域或窗口 W 112,我们构建 120 一组特征矢量 121。对于该组特征矢量 121

确定 130 对应的协方差矩阵 131。我们比较 140 该协方差矩阵 131 与当前模型 104 的协方差矩阵,以确定距离 141。在当前帧中,我们定位出具有与模型 104 的协方差矩阵有最小协方差距离 201 的区域(窗口),并选择 150 对应的区域作为所述运动物体 101 的位置 151。为了动态适应移动物体的位置及外表的变化,我们保持一组先前的协方差矩阵 161,并使用李代数提取出一个固有的“平均”协方差矩阵作为更新的模型。在下面的帧中,我们使用所述物体的先前位置 151 以确定 180 一组候选区域(窗口)以构建特征矢量。

[0027] 现在我们更详细描述我们的发明的具体实施方式。

[0028] 特征提取

[0029] 我们以 I 表示一个两维像素图像或帧 102。每个像素都具有位置 (x, y) 和强度。对于彩色像素,每个像素具有三个强度(RGB)。根据本发明的一个实施方式的方法可以被推广为其它类型的图像,例如多谱图像。

[0030] 从图像 I 102 中,我们根据下式得到一个 $M \times N \times d$ 维的特征图像 111,

[0031] $F(x, y) = \Phi(I, x, y)$

[0032] 其中, M 和 N 是输入图像的维数,函数 Φ 可以是任何特征映射,如颜色、图像梯度、边的量值、边的方位、滤波响应等。可以扩展这些特征以包括高阶导数、质地评价(texture score)、和时域帧差,例如运动矢量。

[0033] 下面,我们检查特征图像 111 中的一组窗口 112。可以根据物体的当前位置的最佳估计来选择该组窗口,考虑物体的外形、尺寸、方位、位置、速度和方向,这些由前面的帧中确定。

[0034] 对于该组中的 k 个窗口 W_k 112 或区域中的每一个,我们构建 120 一个 d 维特征矢量 $\{f_k\}$ 121, $k = 1, \dots, n$ 。我们使用两种属性映射:空间和外形,来构建 120 特征矢量 $\{f_k\}$ 121。空间属性从像素坐标 (x, y) 中获得,外形属性可以包括颜色、梯度、边、滤波响应等。特征 121 可以被直接与像素坐标相联系,

[0035] $f_k = [x \ y \ I(x, y) \ I_x(x, y) \dots]_k$ 。

[0036] 或者,所述特征可以被认为是对称的:

[0037] $f_k^s = \begin{bmatrix} |x'| & |y'| & I(x, y) & I_x(x, y) & \dots \end{bmatrix}_k,$

[0038] 其中 $(x', y') = (x - x_0, y - y_0)$ 是相对坐标, (x_0, y_0) 是窗口 112 的中心的坐标。也可以用对称径向关系来指定特征矢量,

[0039] $f_k^r = \begin{bmatrix} \|(x', y')\| & I(x, y) & I_x(x, y) & \dots \end{bmatrix}_k,$

[0040] 其中

[0041] $\|(x', y')\| = \sqrt{(x'^2 + y'^2)}。$

[0042] 空间属性和外形属性间的不同关系使得可以施加不同的混合则。例如,对于相对于窗口原点 (x', y') 的物体旋转,特征矢量 f_k 为外形模型而占优,而径向特征矢量 f_k^r 提供了特征的姿态和旋转不变的空间构造。

[0043] 协方差矩阵的构建

[0044] 对于每个窗口 112 中的每个特征矢量 121,我们确定 130 对应的协方差矩阵 C_w 131。协方差矩阵是对两个变量怎样一起变化的量度。对于沿相同方向不同于它们的均值

的每对变量,协方差变成更大的正值,并且对于沿相反方向不同于它们的均值的每对变量,协方差变成更大的负值。

[0045] 因此,我们将特征的每个窗口 W_{112} 表示为像素特征的 $d \times d$ 的协方差矩阵 C_w 131:

$$[0046] \quad C_w = \frac{1}{MN} \sum_{k=1}^{MN} (f_k - m_\mu)(f_k - m_\mu)^T,$$

[0047] 其中 m_μ 是区域或窗口 W_{112} 中的像素对应的特征的均值矢量, M 和 N 是输入图像维数, T 是转置运算符。

[0048] 根据上述定义,协方差矩阵是对称阵。该矩阵的对角线元素表示每个特征的方差,非对角线元素表示特征各个相关。矩阵 C_w 131 表示单窗口或区域内的多个特征的协方差。

[0049] 使用协方差矩阵作为物体的模型有多个好处。协方差矩阵提供了一种融合多个特征的自然的方法,而不用标准化特征或使用混合权重。协方差矩阵包含了嵌入在现有直方图中的相同信息,以及可以被从外表模型中得到的信息。

[0050] 单个协方差矩阵与物体 102 的不同视图和姿态下其它窗口 121 的协方差矩阵 131 的集合相比较 140,通常已经是足够的。当协方差矩阵的集合被确定 130 后,通过平均,恶化各个像素样本的噪声被大部分滤掉了。所有的协方差矩阵具有相同的大小 ($d \times d$)。

[0051] 因此,不同大小的不同区域的矩阵可以被直接比较,而不用被限制为固定大小的区域。事实上,在特征(例如图像梯度和方位)是在不同尺度下提取的情况下,在不同图像的区域中,协方差矩阵具有一个固有的尺度恒定属性。如上所述,协方差矩阵可以是旋转不变的。然而,当与像素的方位相关的信息被嵌入到特征矢量 121 中时,可以检测到旋转差异。

[0052] 因此,我们认为协方差矩阵是所述运动物体的很好的模型。

[0053] 协方差矩阵的构建的完整描述在下面的美国专利申请中给出,申请号为 11/305,427,“Method for Constructing Covariance Matrices from Data Features”,由 Porikli 等在 2005 年 12 月 14 日申请。

[0054] 除了协方差矩阵,我们还计算一个 ($d \times 1$) 的特征均值矢量 μ_w 175,参见图 3。该矢量 μ_w 的每一个元素是窗口中对应特征的均值。

[0055] 比较区域

[0056] 为了在帧 102 中跟踪移动物体 101,我们将每一帧 102 中的各个窗口 112 的协方差矩阵集与物体 102 的‘当前’模型 104 的协方差矩阵进行比较 140,以获得距离 141。我们选择 150 一个最小距离作为物体 101 的当前位置 151。当跟踪物体时,对每一帧,模型的协方差矩阵 104 可被更新 160,如下所述。

[0057] 为了在当前帧中为所述运动物体 101 获得一个最近似的区域,我们度量 140 当前帧的协方差矩阵集 131 与物体模型的协方差矩阵 104 间的‘距离’ 141。

[0058] 但是,协方差矩阵并不位于欧几里得空间。例如,该空间在与负定标器(negative scaler)相乘的情况下是不闭合的。也即,两个协方差矩阵的简单的算术减法并不能度量相应区域内的差异。

[0059] 协方差矩阵或者是正定的或者是半正定的。如果特征矢量中的特征是精确相同

的,也就是说协方差矩阵是对称正定的,那么就可以使用下面的文章中所述的距离度量,Forstner等的“A metric for covariance matrices”,Technical report,Dept. of Geodesy and Geoinformatics, Stuttgart University, 1999,在此引入作为参考。该度量使用特征值的平方对数之和来确定协方差矩阵间的差异或‘距离’ 141,公式为:

$$[0060] \quad D(C_i, C_j) = \sqrt{\sum_{k=1}^d \ln^2 \lambda_k(C_i, C_j)}$$

[0061] 其中 $\{\lambda_k(C_i, C_j)\}$ 为矩阵 C_i 和 C_j 的归一化的本征值 (eigenvalue) 并由下式确定

$$[0062] \quad \lambda_k C_i x_k - C_j x_k = 0 \quad k = 1 \dots d,$$

[0063] 其中 x_k 是归一化本征矢量 (eigenvector)。所述对数保证了将偏差度量为因子,所述平方保证了偏差因子被平等地求值。

[0064] 对正定对称矩阵 C_i 和 C_j , 该距离度量 D 141 满足度量公理。另外,协方差对均值变化是不变的,例如颜色值的相同缩放。当在变化照明的条件下跟踪物体时,这成为一个优势属性。由承载照明所引起的外表变化,除了协方差距离外,还可通过包括平均距离来检测。

[0065] 图 2 示出了根据本方法的一个实施方式的用于比较 140 协方差矩阵以确定距离度量 D 的步骤。除了协方差矩阵比较外,我们还计算 176 模型的特征均值矢量与候选区域间的距离。该距离被定义为模型 104 和当前窗口间的平均特征矢量差值 $\|\mu_w - \mu^*\|$ 的模。

[0066] 所述模型和当前窗口间的协方差距离 177 和平均距离 176 被用融合权重来缩放,以在距离函数上施加应用特定的限制。这些权重 171 和 172 调节协方差的重要性,以及距离函数对平均差值的敏感性,如颜色和照明变化。在我们用协方差融合权重 172 乘以协方差距离,以及用均值融合权重 171 乘以特征均值距离之后,我们将它们的和指定为所述模型和当前窗口 174 之间的距离。该距离得分还考虑到当前窗口与物体的先前位置的空间距离 173。对于空间距离 173,所述模型对当前窗口距离得分 174 按比例增加。这使得我们可以在估计的位置上施加运动特定的限制。在例如不要限制物体的运动的应用中,在距离项 174 中不包含空间距离。

[0067] 额外的基于运动的限制被集成在候选窗口选择 180 中,如图 3 中所示。使用物体的运动历史 181,也即物体的先前位置集,以及物体 182 的大小,限制搜索范围。对于物体的运动被限制以适合线性或非线性运动模型的特定应用,我们使用基于 Kalman 滤波的方法来预测 183 物体在当前帧中的位置。我们根据物体的大小定义 184 在此预测位置附近的一个搜索区域。例如,该搜索区域可以是物体两倍大的窗口。然后,对仅在该搜索区域内的候选窗口 112 使用特征提取、协方差比较和距离计算。

[0068] 对于运动不受限的应用,搜索区域是整个图像。

[0069] 模型更新策略

[0070] 因为非刚性和移动的物体会在时间上经受形状、大小和外表的变形,所以可以动态地适应 160 这些变化。但是,应该理解,如果物体(例如一个球或车辆)的外表保持相对稳定,更新 160 步骤并不总是必须的。

[0071] 我们构建并更新与最近估计的物体区域 $w_1, \dots, w_T$ 相对应的协方差矩阵 161 的子集。我们保留 T 个先前的协方差矩阵 $[C_1, \dots, C_T]$ 的子集,其中 C_1 表示当前的协方差矩阵。从这个子集,我们确定一个抽样均值协方差矩阵。该抽样均值协方差矩阵‘混合’所有的先前矩阵以便类似于所述物体 101 的属性。

[0072] 然后,所有先前检测到的区域及对应的特征测量都是可获得的,可以通过下式获得聚合的协方差矩阵 104,

$$[0073] \quad \hat{C} = \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,2}^2 & \cdots \\ \sigma_{1,2}^2 & \sigma_{2,2}^2 & \\ \vdots & & \ddots \end{bmatrix}_{d \times d}$$

[0074] 其中互协方差定义为

$$[0075] \quad \sigma_{u,v}^2 = \frac{1}{MNT} \sum_{t=1}^T \sum_{k=1}^{MN} [f'_k(u) - m_\mu(u)][f'_k(v) - m_\mu(v)], \text{ 并且 } f'_k \in W_T.$$

[0076] 在所有区域 $W_1, \dots, W_T$ 内确定均值 m_μ 。此公式假定所有的窗口具有相同的大小并且有相同的影响。

[0077] 但是,希望在不被限制于恒定窗口大小并保持所有先前测量的情况下获得聚合协方差矩阵。

[0078] 我们想要确定均值协方差矩阵,即固有平均。但是,协方差矩阵并不符合欧几里得几何,结果,取矩阵的平均是不准确的。但是,我们可以使用黎曼几何来确定多个协方差矩阵的均值,因为正定协方差矩阵具有李群结构。

[0079] 我们提供一个李群和代数学的简要介绍。群是一个非空集合 G ,其具有的群操作具有下列属性,对于所有 $A, B, C \in G$:

[0080] 闭合性: $AB \in G$

[0081] 结合性: $(AB)C = A(BC)$

[0082] 幺元: $e \in G$, 使得 $Ae = eA = A$

[0083] 逆元: $A^{-1} \in G$, 使得 $AA^{-1} = A^{-1}A = e$ 。

[0084] 李群是一个解析流形,其也是一个群,使得群操作的乘法 $(A, B) \rightarrow AB : G \times G \rightarrow G$, 并且逆运算 $A \rightarrow A^{-1} : G \rightarrow G$ 是可微映射。李群可以被局部看作拓扑等价于矢量空间, $\mathbb{R}^d$ 。因此,任意群元素 G 的局部邻居关系可以被通过它的切空间充分描述。群的单位元素 e 处的切空间形成一个李代数 $\mathfrak{g}$ 。李代数 $\mathfrak{g}$ 是一个在李括号下闭合的矢量空间 $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, 它是满足下面的恒等式的双线性操作:

[0085] 反对称性: $[a, b] = -[b, a]$

[0086] Jacobi 恒等式: $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$ 。

[0087] 指数映射, $\exp : \mathfrak{g} \rightarrow G$, 将李代数中的矢量映射为李群。为了我们的目的,我们注意力集中在矩阵李群。矩阵的指数映射以及其逆被定义为:

$$[0088] \quad \exp(a) = \sum_{n=0}^{\infty} \frac{1}{n!} a^n ; \log(A) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (A - e)^n.$$

[0089] 对于交换群,指数映射满足恒等式 $\exp(a)\exp(b) = \exp(a+b)$ 。该恒等式对非交换李群不成立,例如协方差矩阵。等价的映射被定义为 $\mu : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, 即 $\exp(a)\exp(b) = \exp(\mu(a, b))$ 。映射 $\mu(\dots, \dots)$ 由 Baker-Campbell-Hausdorff 公式定义为

$$[0090] \quad \mu(a, b) = a + b + \frac{1}{2}[a, b] + O(|(a, b)|^3). \quad (1)$$

[0091] 我们使用李代数来获得协方差矩阵 161 的集合的固有均值。令 c 为李代数中的一

个元素,并且 $C = \exp(c)$ 为元素到李群的映射,即,对应的协方差矩阵。流形上的距离被定义为该流形上的点之间最小长度的曲线。具有最小长度的曲线被称为最短连线,该曲线的长度是固有距离。

[0092] 点 C 到群的单位元 e 的固有距离为 $\|\log(C)\|$ 。由群元素的逆元 C^{-1} 左乘将点 C 映射到 e ,以及 C 处的切空间到李代数。该映射是一个同构 (isomorphism)。使用此映射,我们可以度量任何两个元素间的固有距离为:

$$[0093] \quad \rho(C_i, C_j) = \|\log(C_i^{-1}C_j)\|。 \quad (2)$$

[0094] 给定协方差矩阵 161, $\{c_t\}_{t=1 \dots T}$, 作为群上的数据点, 映射

$$[0095] \quad c_t = \log(C^{-1}C_t) \quad (3)$$

[0096] 将数据点转换为李代数,以及 C 到零。

[0097] 因为李代数是矢量空间,我们确定第一阶近似值为真或该空间上的固有均值。开始于一个初始矩阵 C_1 , 迭代地确定第一阶近似值至固有均值,我们收敛于该群上的一个固定点。此过程为

[0098] 初始化 $C=C_1$

[0099] 重复

[0100] for $t=1$ to T

[0101] 计算 $c_t = \log(\hat{C}^{-1}C_t)$

[0102] 计算 $\Delta \hat{C} = \exp\left(\frac{1}{T} \sum_{t=1}^T c_t\right)$

[0103] 赋值 $\hat{C} = \hat{C} \Delta \hat{C}$

[0104] 直到 $\|\log(\Delta \hat{C})\| < \varepsilon$

[0105] 在迭代的最后,我们找到固有均值 $\hat{C}$ 。我们使用固有价值协方差矩阵 $\hat{C}$ 作为当前模型 104。

[0106] 当我们估计物体在后面的帧中的位置时,我们比较各个区域的协方差矩阵 131 与模型 104 的协方差矩阵。

[0107] 在上面的公式中,我们认为先前矩阵 $C_1, \dots, C_T$ 的子集对结果具有相同的影响,而不管结果正确不正确。

[0108] 可以根据类似于先前模型的李群矩阵,按比例地为李代数的数据点加权重。然后,上述过程的第二个‘计算’步骤变成

$$[0109] \quad \Delta \hat{C} = \exp\left(\frac{1}{\rho^*} \sum_{t=1}^T \rho(C_t, C^*) c_t\right) \quad (4)$$

[0110] 其中 $\rho^* = \sum_{t=1}^T \rho(C_t, C^*)$ 并且 C^* 是从先前帧中确定的模型。

[0111] 换言之,我们使用与最小距离 C_m 201 相对应的给定协方差矩阵来更新物体的模型 C^* 104。参考图 4,我们首先将矩阵 C_m 201 加入 300 到先前的协方差矩阵 $[C_1, \dots, C_T]$ 161 的数据集中。然后,我们指定 301 数据集中的第一个矩阵为临时的固有均值协方差矩阵 $\hat{C}$ 。我们计算它的逆 302。我们通过被乘以矩阵的逆 (如上述过程所示) 的对应的协方差矩阵的对数计算 303 映射 c_t , 其中 $t=1, \dots, T$ 。我们使用权重系数 400 计算 305 这些映射的加权和。我们计算 306 一个误差项,并将它与误差阈值 ε 进行比较 307。如果误差项大于阈

值,那么我们通过从临时固有均值中指定逆元以继续到 302。如果误差很小,我们将临时固有均值协方差矩阵 $\hat{C}$ 指定为新的更新过的模型 $C * 104$ 。

[0112] 为了如图 5 中所示指定上述权重 400,我们定义一个固有距离函数 401,并计算各个距离项 402 和一个总距离 403。然后我们根据 404 中所示的公式来指定权重。

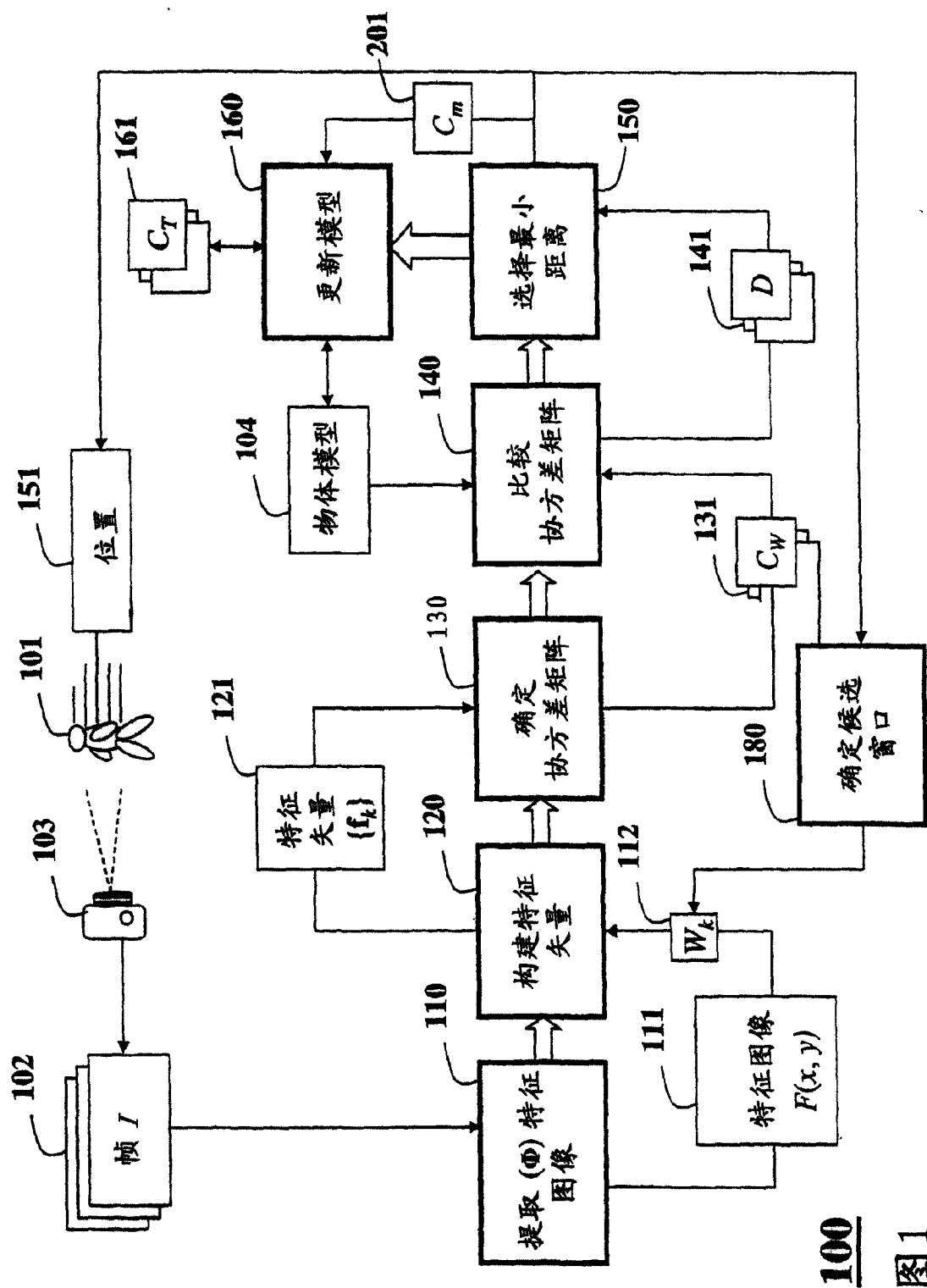
[0113] 如图 6 中所示,有不同的方法来将对应于最小距离的当前矩阵加入 300 到先前协方差矩阵的数据集中。我们直接将当前矩阵作为一个新矩阵加入到数据集 311 中,并增加矩阵 312 的数目。可替换地,我们移除 321 加入到数据集中的最老的矩阵。我们更新索引 322。这些索引指示了矩阵被加入到数据集中的时间。然后作为最新的组分,我们加入当前矩阵 323。

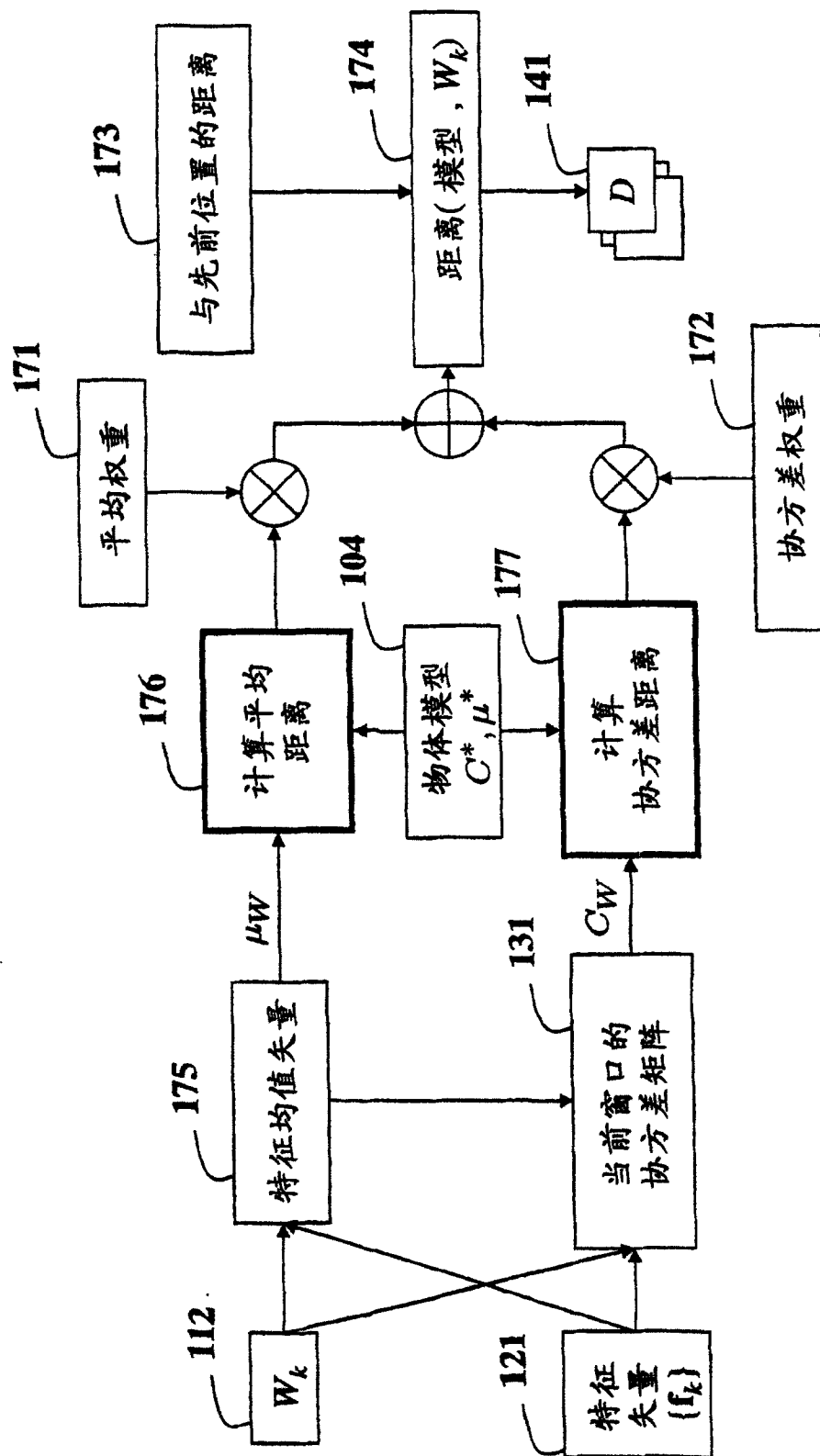
[0114] 本发明的效果

[0115] 根据本发明的具体实施方式的用于在视频中跟踪物体的方法,既使用物体的空间属性又使用物体的外表属性,并提供了一种融合多个物体特征的很好的解决方法。

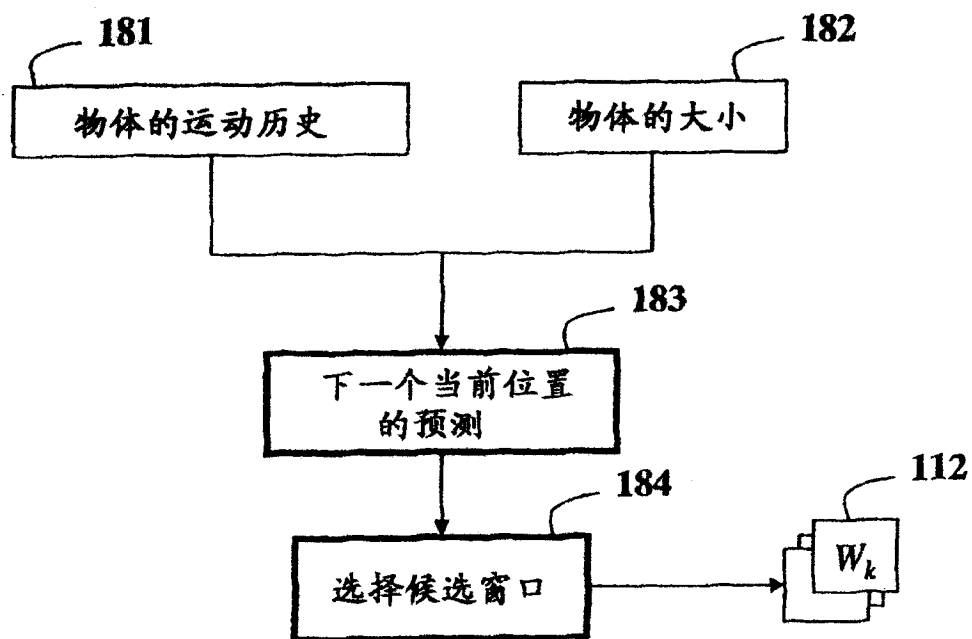
[0116] 该方法并不对物体的外表和运动做出假定。该方法并不限于最大速度。这意味着该方法可以跟踪物体,即使它们的运动是无规律的并且快速的。与现有的均值移位方法相比,该方法并不假定连续帧间的物体区域间的重叠。

[0117] 该方法找出一个全局最优解,不像现有的局部取样方法,例如粒子滤波。该方法可以有效地适应于临时模型变化,并且具有一个非常低的 $(d^2+d)/2$ 维数。该方法还可以比较不同大小的物体区域,如果物体的感觉到的大小是变化的,如由摄像头所观察到的那样,该方法是有益的。

100
图1

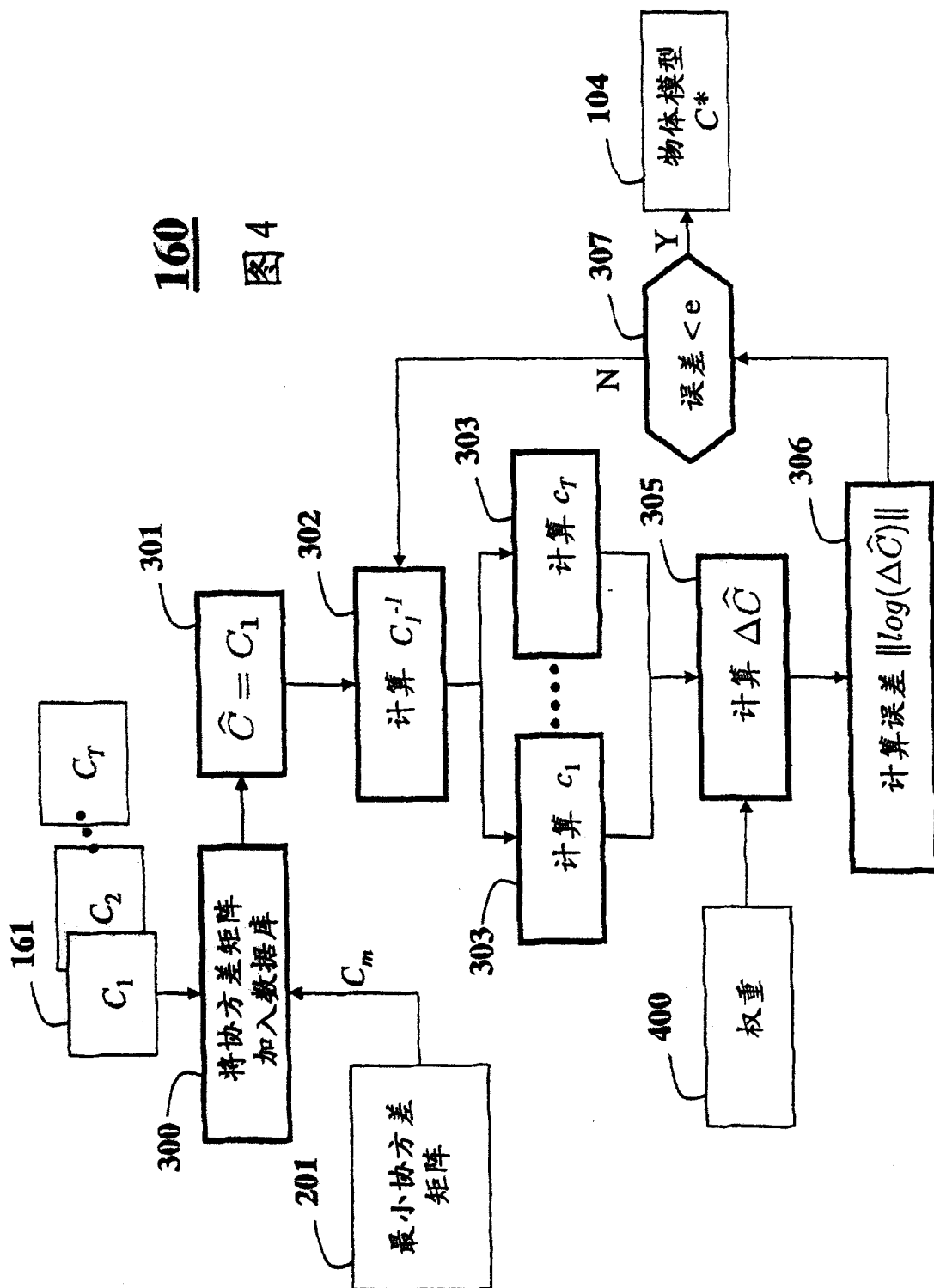


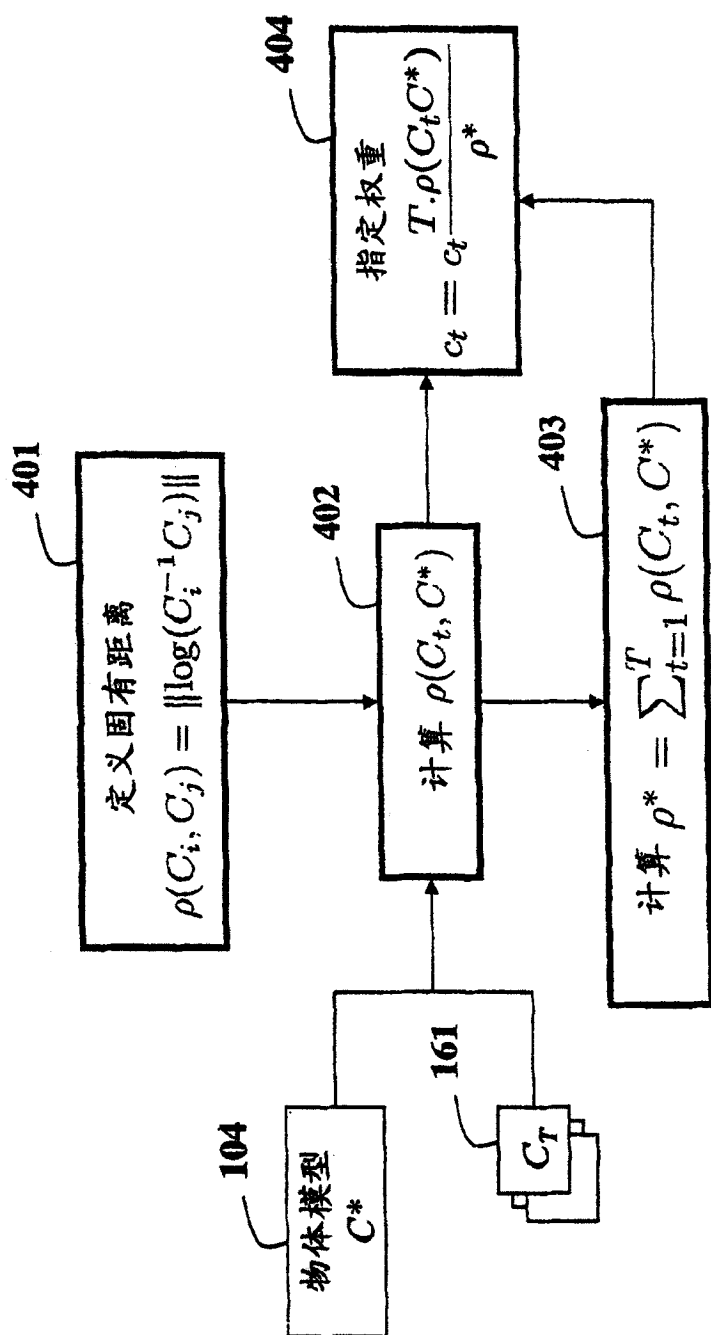
140 图2



180

图 3





400

图 5

