



(12)发明专利

(10)授权公告号 CN 105096307 B

(45)授权公告日 2018.01.02

(21)申请号 201510232293.0

(22)申请日 2015.05.08

(65)同一申请的已公布的文献号

申请公布号 CN 105096307 A

(43)申请公布日 2015.11.25

(30)优先权数据

14/272,570 2014.05.08 US

(73)专利权人 三菱电机株式会社

地址 日本东京都

(72)发明人 刘洛埏 O·图兹尔

(74)专利代理机构 北京三友知识产权代理有限公司

公司 11127

代理人 吕俊刚 刘久亮

(51)Int.Cl.

G06T 7/00(2017.01)

G06T 7/593(2017.01)

G06K 9/62(2006.01)

(56)对比文件

CN 102026013 A,2011.04.20,

CN 101976455 A,2011.02.16,

CN 103366354 A,2013.10.23,

CN 103226821 A,2013.07.31,

US 2012008857 A1,2012.01.12,

US 6301370 B1,2001.10.09,

Martin Peris 等.Towards a Simulation Driven Stereo System.《21st International Conference on Pattern Recognition》.2012,第1038-1042页.

Sebastian Tuermer 等.Airborne Vehicle Detection in Dense Urban Areas Using HoG Features and Disparity Maps.《IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing》.2013,第6卷(第6期),第2327-2337页.

Asmaa Hosni 等.Fast Cost-Volume Filtering for Visual Correspondence and Beyond.《IEEE Transactions on Pattern Analysis & Machine Intelligence》.2012,第35卷(第2期),第504-511页.

Jonghwan Kim 等.Stereo Vision-Based Improving Cascade Classifier Learning for Vehicle Detection.《Advances in Visual Computing》.2011,第387-397页.

审查员 李秋萍

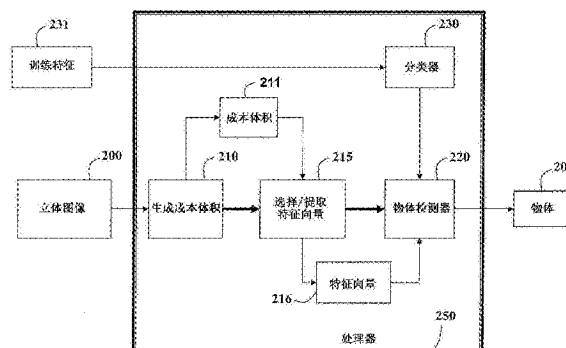
权利要求书1页 说明书4页 附图7页

(54)发明名称

在成对的立体图像中检测物体的方法

(57)摘要

在成对的立体图像中检测物体的方法。一种在从场景获取的成对的立体图像中检测物体的方法,其首先从所述成对的立体图像生成成本体积,其中,所述成本体积包括在成对的立体图像之间针对立体图形的各个像素的视差值范围的匹配成本。根据所述成本体积中的子图像,利用所述子图像内部的区域内具有最小累积成本的视差值的特征函数确定特征向量。然后,将分类器应用于所述特征向量以检测所述子图像是否包括物体。



1. 一种在从场景获取的成对的立体图像中检测物体的方法, 其中各个立体图像包括像素, 该方法包括以下步骤:

从所述成对的立体图像生成成本体积, 其中, 所述成本体积包括在成对的立体图像之间针对各个像素的视差值范围的匹配成本;

根据所述成本体积中的子图像, 利用所述子图像内部的区域内具有最小累积成本的视差值的特征函数确定特征向量; 以及

将分类器应用于所述特征向量以检测所述子图像是否包括物体,

其中, 上述步骤在处理器中执行。

2. 根据权利要求1所述的方法, 该方法还包括:

在所述立体图像内定位所述物体。

3. 根据权利要求1所述的方法, 其中, 从成对的训练立体图像学习所述分类器。

4. 根据权利要求1所述的方法, 该方法还包括:

调整所述成对的立体图像。

5. 根据权利要求1所述的方法, 该方法还包括:

平滑所述成对的立体图像。

6. 根据权利要求1所述的方法, 其中, 所述生成还包括:

利用欧几里德范数来使所述成对的立体图像中的像素的颜色和梯度匹配。

7. 根据权利要求1所述的方法, 其中, 所述特征函数为

$$f_k(J) = \begin{cases} 1 & \text{如果 } d_{\min}(R_k^1) > d_{\min}(R_k^2) \\ 0 & \text{如果 } d_{\min}(R_k^1) = d_{\min}(R_k^2) \\ -1 & \text{其它} \end{cases}, \quad (1)$$

其中J表示所述子图像, k表示所述特征向量的维度, min表示返回最小值的函数, $d_{\min}(R_k^i)$ 表示所述子图像中的矩形区域 R_k^i 中具有最小累积成本的视差值。

8. 根据权利要求7所述的方法, 其中

$$d_{\min}(R_k^i) = \arg \min_d \sum_{(x,y) \in R_k^i} C(x,y,d)$$

其中C(x,y,d)表示所述成本体积。

9. 根据权利要求1所述的方法, 其中, 所述分类器是包括T个决策树分类器的集成分类器。

10. 根据权利要求9所述的方法, 其中, 所述分类器针对子图像J提供检测得分s, 所述检测得分为

$$s(J) = \sum_{t=1}^T \theta_t \delta_t(J),$$

其中 δ_t 是决策树分类器, θ_t 是对应权重。

在成对的立体图像中检测物体的方法

技术领域

[0001] 本发明涉及计算机视觉,更具体地讲,涉及在立体图像中检测物体。

背景技术

[0002] 许多计算机视觉应用使用通过立体相机获取的立体图像来检测物体。立体相机通常具有多个镜头和传感器。通常,镜头之间的轴内距离约与眼睛之间的距离相同以提供交叠视野。

[0003] 图1示出用于基于立体的物体检测的传统系统。立体相机101获取立体图像102。检测方法可包括以下步骤:立体成像100、成本体积(cost volume)确定110、深度/视差图估计120和物体检测130。

[0004] 基于立体的物体检测的大多数传统方法依赖于交叠区域120中的每像素深度信息。此步骤通常被称作深度/范围图估计。此步骤可通过确定视差值(即,两个图像中的对应像素的平移)、确定深度图来实现。然后深度图可用于物体检测130,例如,深度图的梯度方向直方图(HoG)用于物体描述。一种方法估计子图像区域中的主要视差,并且使用相对视差值的共现直方图来进行物体检测。

[0005] 深度/范围/视差图估计是具有挑战的问题。局部方法存在深度确定不准确的问题,而全局方法需要可观的计算资源并且不适合于实时应用。

[0006] 多种方法通过使用关注区域生成的立体提示(stereo cue)来避免深度图确定步骤。例如,一种方法确定标记潜在物体位置的stixel图。各个stixel由相对于相机的3D位置来定义,并且垂直地竖立在地平面上。然后将基于彩色图像内容的检测器应用于所述位置以检测物体。

[0007] 美国公布20130177237使用范围图来确定关注区域,并且使用基于强度直方图的分类器来检测物体。

[0008] 关注区域方法无法直接应用于物体检测。关注区域方法必须与其它物体检测器结合应用。另外,当关注区域没有覆盖物体时,无疑会缺失检测。

发明内容

[0009] 本发明的实施方式提供一种在立体图像中检测物体的方法。从图像计算成本体积。然后,直接对从成本体积获得的特征应用物体检测。所述检测使用从训练特征学习的T个决策树分类器(Adaboost)。

[0010] 本发明避免了现有技术的易于发生错误并且计算上复杂的深度图估计步骤,得到准确且高效的物体检测器。所述方法更加适合于嵌入式系统,因为它不需要为获得良好的深度图所需的复杂的优化模块。另外,所述方法搜索输入图像中的所有子图像以检测物体。这避免了存在于关注区域生成技术中的缺失检测问题。

[0011] 所述检测准确,因为所述方法可借助大量的训练数据并且利用机器学习过程。其在检测准确度方面胜过关注区域生成技术。

附图说明

- [0012] 图1是传统的基于立体的物体检测系统的框图；
- [0013] 图2是根据本发明的实施方式的基于立体的物体检测系统的框图；
- [0014] 图3是图2的基于立体的物体检测系统的物体检测模块的框图；
- [0015] 图4是根据本发明的实施方式的学习基于立体的物体检测器的方法的框图。
- [0016] 图5是根据本发明的实施方式的成本体积确定的示意图；
- [0017] 图6是根据本发明的实施方式的学习的特征的示意图；以及
- [0018] 图7是占据子图像的大部分和小部分的物体的示意图。

具体实施方式

[0019] 图2示出根据本发明的实施方式的在一对立体图像200中检测物体201的方法和系统。从这一对立体图像生成(210)成本体积211。然后选择并提取(215)特征向量216。然后，将物体检测器220应用于特征向量体积以检测物体。物体检测器使用从训练图像特征231学习的分类器230。在检测到物体之后，可对物体进行定位，即，可确定物体在图像中的位置。如本领域中已知的，所述方法可在连接到存储器和输入/输出接口的处理器250中执行。

[0020] 本发明基于这样的认识：深度图中可用的深度信息在成本体积中也可用，因为深度图是从成本体积获得的。

[0021] 本发明的直接使用成本体积的检测器220理论上能够基于深度图使任何检测器的性能匹配。此外，成本体积是比传统深度图更丰富的表示。深度图仅提供各个像素的深度，而成本体积提供立体图像中的各个像素可具有的潜在深度范围(包括真实深度)的匹配成本。因此，检测器使用从成本体积直接获得的特征可访问更多的深度信息，并且实现更好的性能。

[0022] 如图3所示，本发明的一个实施方式包括成本体积生成210、特征提取310、物体检测和定位320、学习的区别特征330以及学习的物体分类模型340。定位确定在哪里检测物体。

[0023] 图4示出用于学习区别特征的机器学习过程以及学习的物体分类模型。从包括成对的训练立体图像的训练数据400选择并学习(410)特征。

[0024] 成本体积生成

[0025] 图5示出成本体积C 211的生成。成本体积 $C: X \times Y \times D$ 是存储在存储器中的三维数据结构，其中X和Y表示图像x和y轴，D表示一组视差值(是两个立体图像 I_L 501和 I_R 502中的对应像素之间的平移)。假设 I_L 和 I_R 被调整，这意味着图像被变换为使得镜头畸变效果被补偿，并且一个图像的一行中的像素被映射至另一图像的同一行中的像素。然后可通过使成对的立体图像 I_L 和 I_R 中的像素外观匹配来确定成本体积。

[0026] 确定成本体积的一个方式是对于任何 $(x, y, d) \in X \times Y \times D$ ，应用由 $C(x, y, d) = || I_L(x, y) - I_R(x-d, y) ||_2 + \lambda || \text{grad}(I_L(x, y)) - \text{grad}(I_R(x-d, y)) ||_2$ 给出的映射。

[0027] 其中 $|| \cdot ||_2$ 表示欧几里德范数， $I_L(x, y)$ 是指 I_L 图像的 (x, y) 位置的像素颜色值， $I_R(x, y)$ 是指 I_R 图像的 (x, y) 位置的像素颜色值， $\text{grad}(I_L(x, y))$ 是指 I_L 图像的 (x, y) 位置的梯度， $\text{grad}(I_R(x-d, y))$ 是指 I_R 图像的 (x, y) 位置的梯度， λ 是控制梯度信息的重要性的权重。需

要注意,可应用诸如双边滤波或引导滤波的图像平滑技术,以增强成本体积。

[0028] 特征提取

[0029] 图6示出图2的特征选择和提取215。从各个子图像600提取K维特征向量以用于确定子图像中是否存在物体。子图像可被视为例如按照多个比例针对各个像素以光栅扫描顺序在图像上通过的移动窗口。

[0030] 需要注意,实施方式仅直接使用成本体积来确定特征。没有执行现有技术中的深度图估计。

[0031] 特征向量的各个维度对应于子图像600中的两个(例如,矩形)区域 R_k^1 601和 R_k^2 602的图成本视差值之间的数值比较结果。将子图像表示为J,特征向量的第k维度表示为 $f_k(J)$ 。 $f_k(J)$ 的值为

$$[0032] \quad f_k(J) = \begin{cases} 1 & \text{如果 } d_{\min}(R_k^1) > d_{\min}(R_k^2) \\ 0 & \text{如果 } d_{\min}(R_k^1) = d_{\min}(R_k^2) \\ -1 & \text{其它} \end{cases}, \quad (1)$$

[0033] 其中 $d_{\min}(R_k^i)$ 表示子图像的区域 R_k^i 中具有最小(min)累积成本的视差值。即,

$$[0034] \quad d_{\min}(R_k^i) = \arg \min_d \sum_{(x,y) \in R_k^i} C(x,y,d). \quad (2)$$

[0035] 需要注意,确定区域中最小成本视差值相对简单,因为如本领域已知的,可有效地利用积分图像技术来获得累积成本。利用下述机器学习过程来学习区域的位置和大小。

[0036] 物体检测和定位

[0037] 与子图像关联的K维特征向量被传递给集成分类器以用于确定检测得分。集成分类器包括T个决策树分类器。各个决策树分类器以K维特征的少量维度作为输入,并且将子图像分类为正(包含物体)或负(不包含物体)。从分类器获得的子图像J的检测得分s由下式给出

$$[0038] \quad s(J) = \sum_{t=1}^T \theta_t \delta_t(J), \quad (3)$$

[0039] 其中 δ_t 是决策树分类器, θ_t 是对应权重。如果得分大于预设阈值,则系统宣布在子图像中检测到。

[0040] 如图7所示,分类器可被训练,以在物体占据子图像701的较大部分时给出较高得分,在物体仅占据子图像702的较小部分时给出较低得分,这是因为与较小的物体相比,较大的物体提供物体在图像内位于何处的更好估计。

[0041] 特征选择和分类器学习过程

[0042] 使用离散AdaBoost过程以用于选择区域

$$[0043] \quad \{(R_k^1, R_k^2) | \forall k = 1, 2, \dots, K\}, \quad (4)$$

[0044] 并且用于学习决策树分类器权重

$$[0045] \quad \{\theta_t | \forall t = 1, 2, \dots, T\} \quad (5)。$$

[0046] 收集用于学习任务的一组数据,所述一组数据包括一组立体训练图像。包含物体的子图像被标记为正实例,而其它被标记为负实例。使正和负子图像对准以使得它们的中心重合。子图像还被缩放为具有相同的高度。经对准和缩放的子图像被表示为

$$[0047] \quad D = \{ (J_i, l_i), i=1, 2, \dots, V \}, \quad (6)$$

[0048] 其中 J_i 表示第 i 子图像, l_i 是标签, V 是子图像的总数。

[0049] 采样一组 N 个区域作为特征池 $\{R_i, i=1, 2, \dots, N\}$,其具有不同的位置和大小并且被经对准的子图像覆盖。随机将两个区域组对并且比较它们的最小成本的视差值。这被执行 K 次以构造 K 维特征向量。

[0050] 使用离散AdaBoost过程来学习 T 个决策树分类器及其权重。所述过程开始于将一致的权重指派给训练样本。然后基于当前训练样本权重学习决策树。不正确分类的样本的权重被增加,以使得权重在下一轮决策树分类器学习期间具有更多影响。基于加权错误率将权重指派给决策树分类器。此处理被重复 T 次以构造 T 个决策树分类器的集成分类器。下面描述该过程的伪代码。

[0051] 输入:特征向量和类标签 $D = \{ (f(J_i), l_i), i=1, 2, \dots, V \}$

[0052] 输出:集成分类器

[0053] 开始于一致的权重

$$[0054] \quad w_i = \frac{1}{V}, \quad i=1, 2, \dots, V$$

[0055] 对于 $t=1, 2, \dots, T$

[0056] 1. 利用权重 w_i 学习决策树分类器 $\delta_t(J) \in \{-1, 1\}$;

[0057] 2. 确定错误率 $\varepsilon = \sum_i w_i |(\delta_t(J_i) \neq l_i)|$;

[0058] 3. 确定决策树分类器权重 $\theta_t = \log \frac{1-\varepsilon}{\varepsilon}$;

[0059] 4. 对于 $i=1, 2, \dots, V$, 设定 $w_i \leftarrow w_i \exp(\theta_t |(\delta_t(J_i) \neq l_i)|)$; 并且

[0060] 5. 将样本权重 $w_i \leftarrow \frac{w_i}{\sum_i w_i}$ 正规化。

[0061] 步骤2和4中使用的函数 δ_t 表示指示函数,其在括号内的陈述为真的情况下返回一,否则返回零。

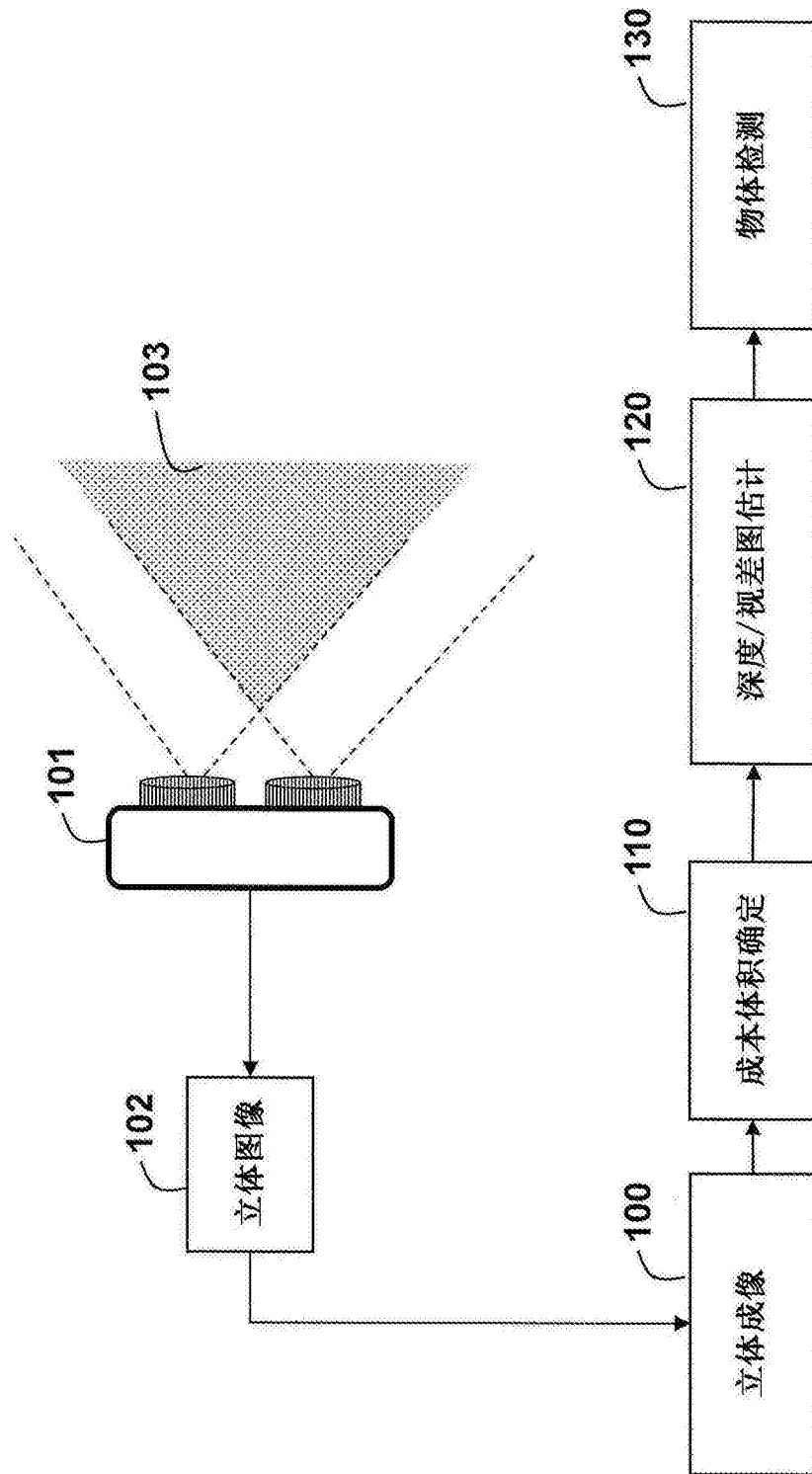


图1

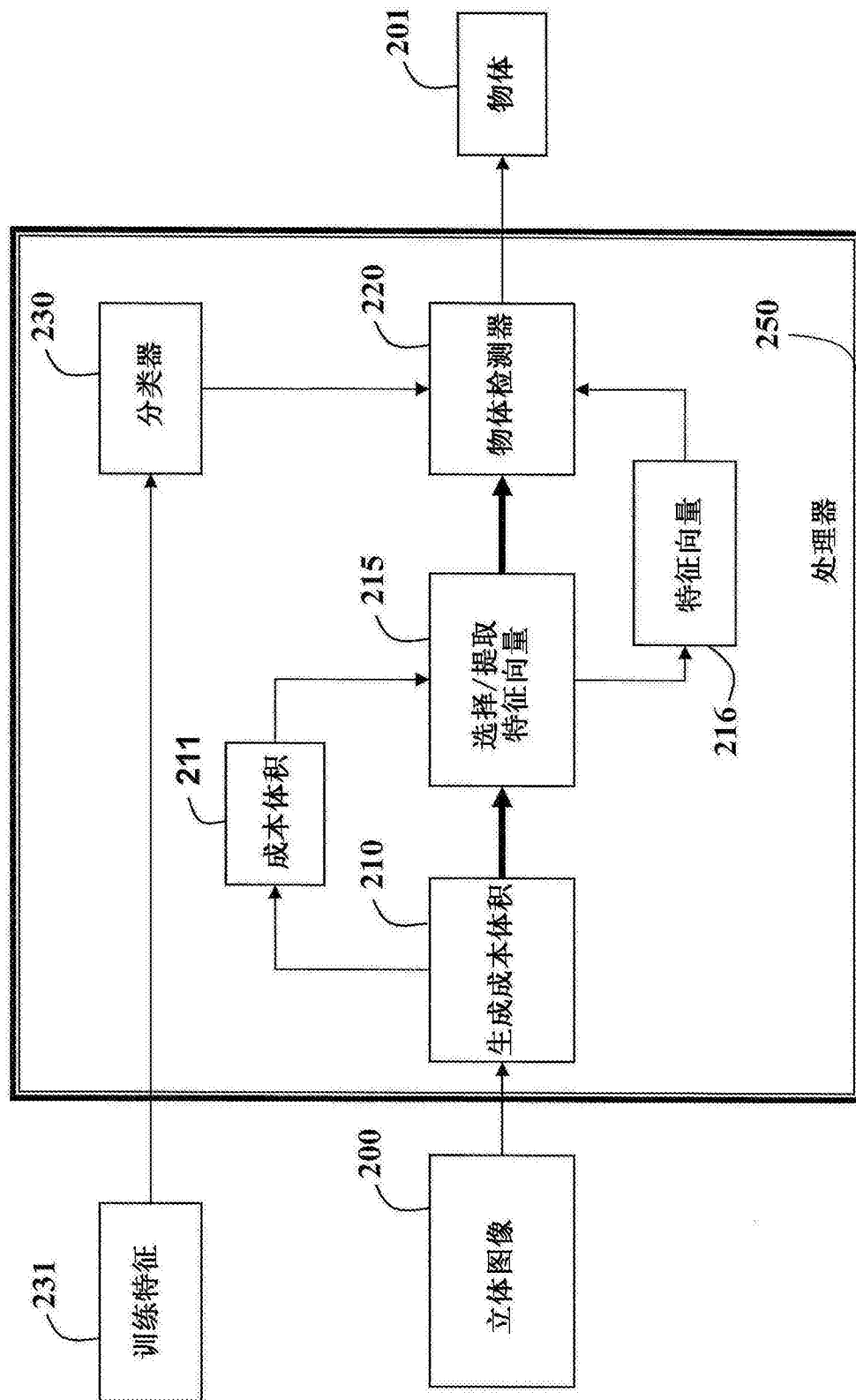


图2

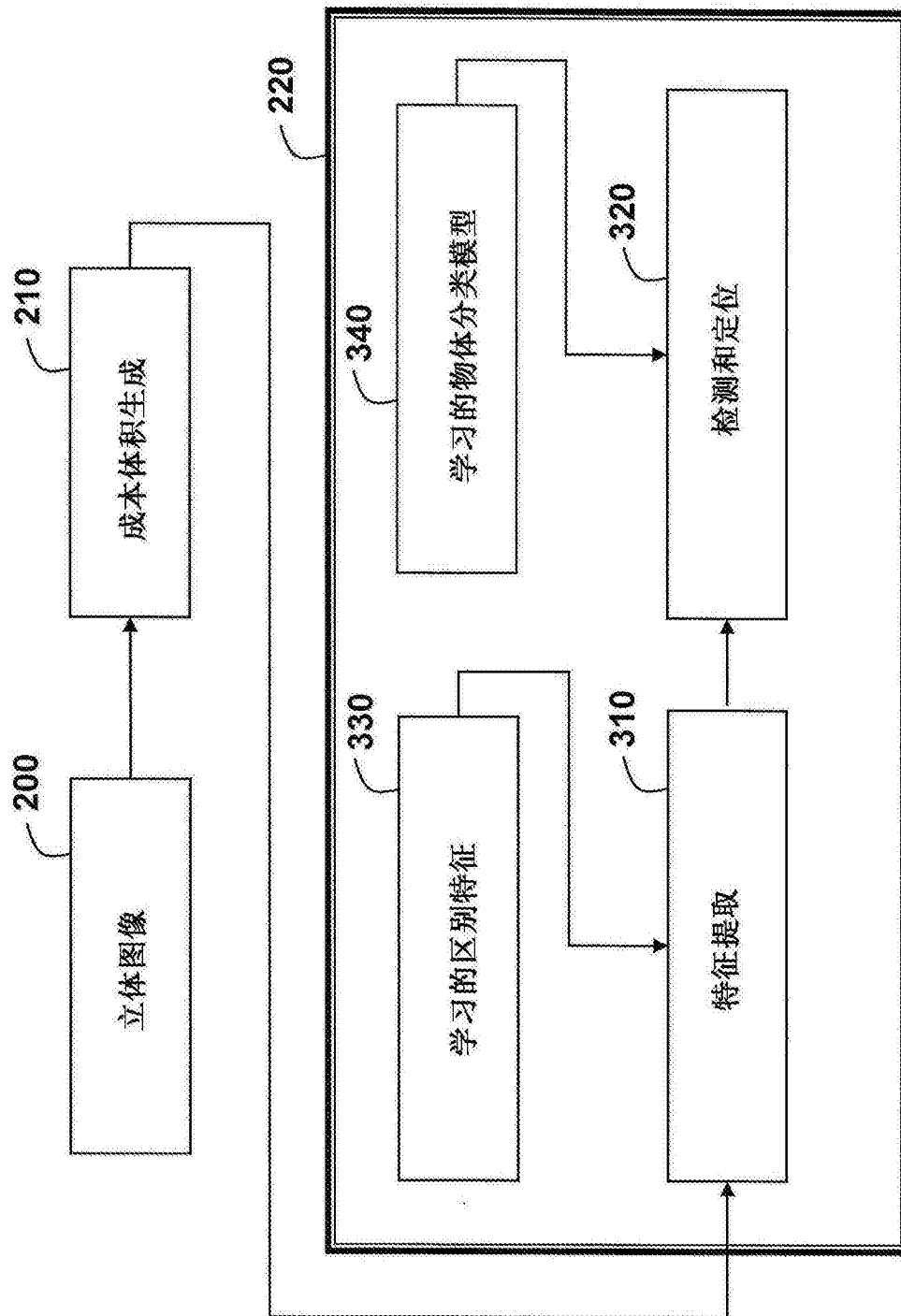


图3

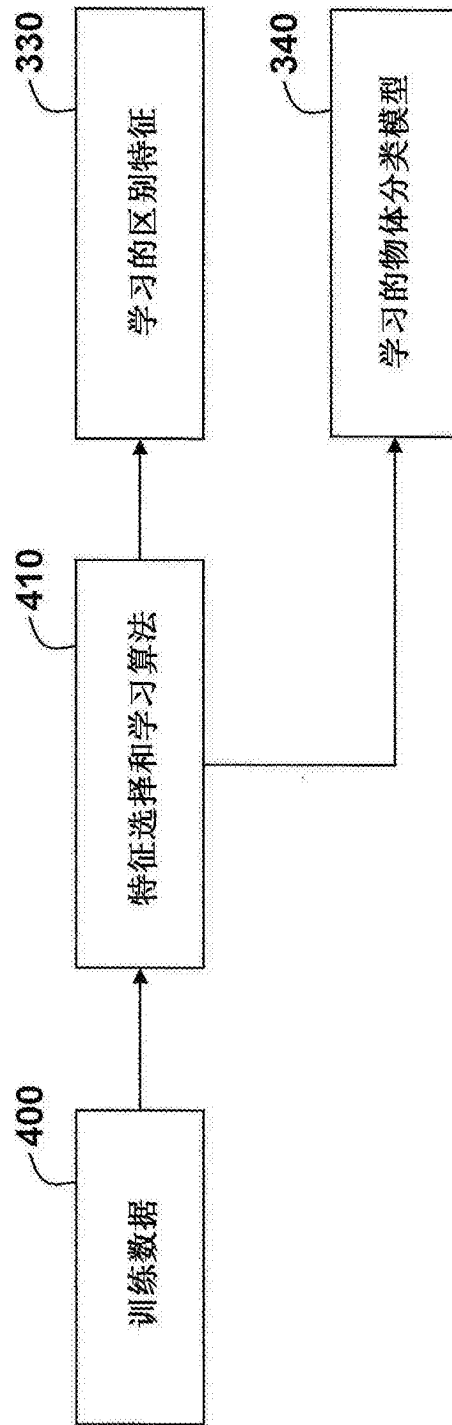


图4

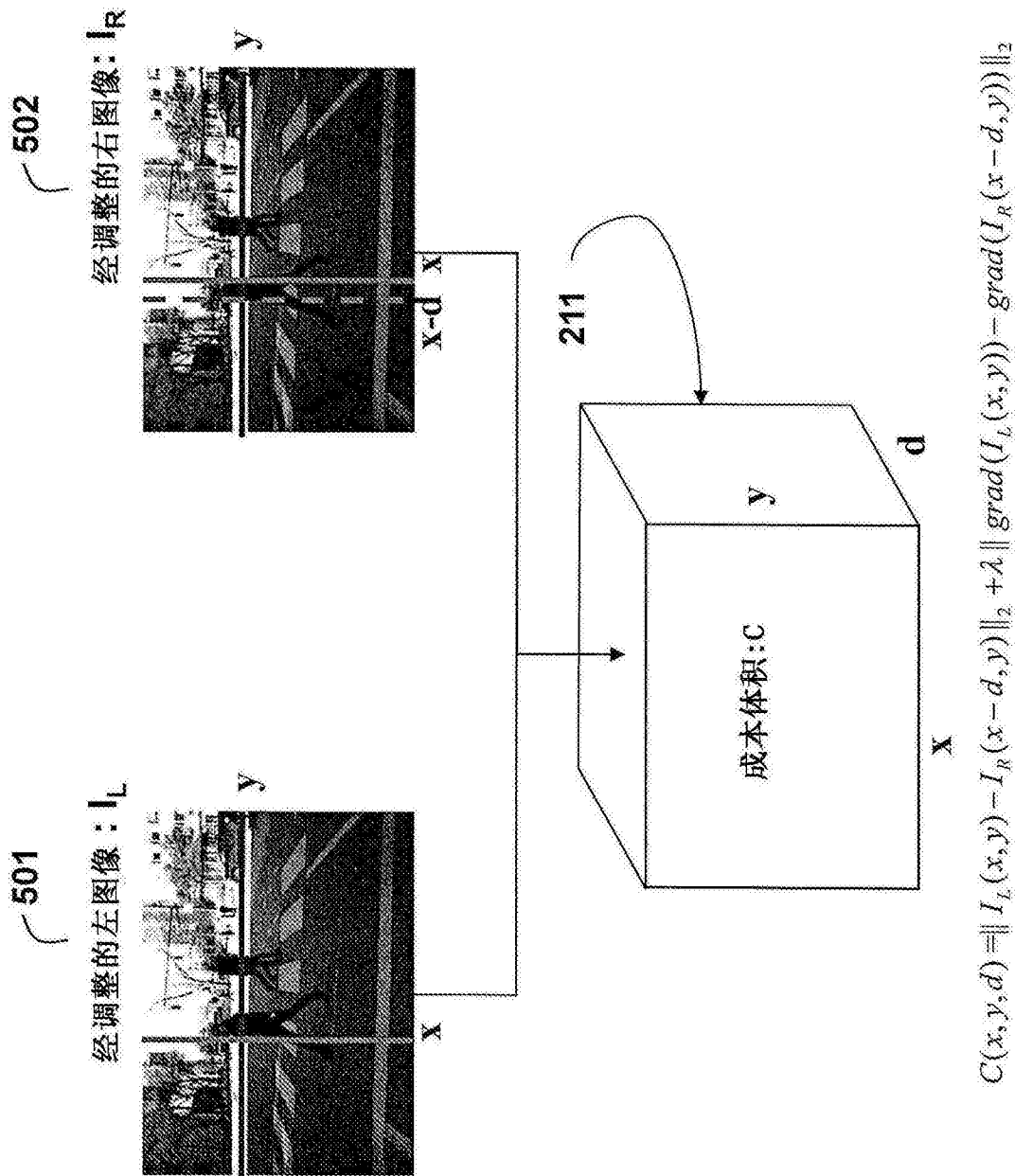


图5

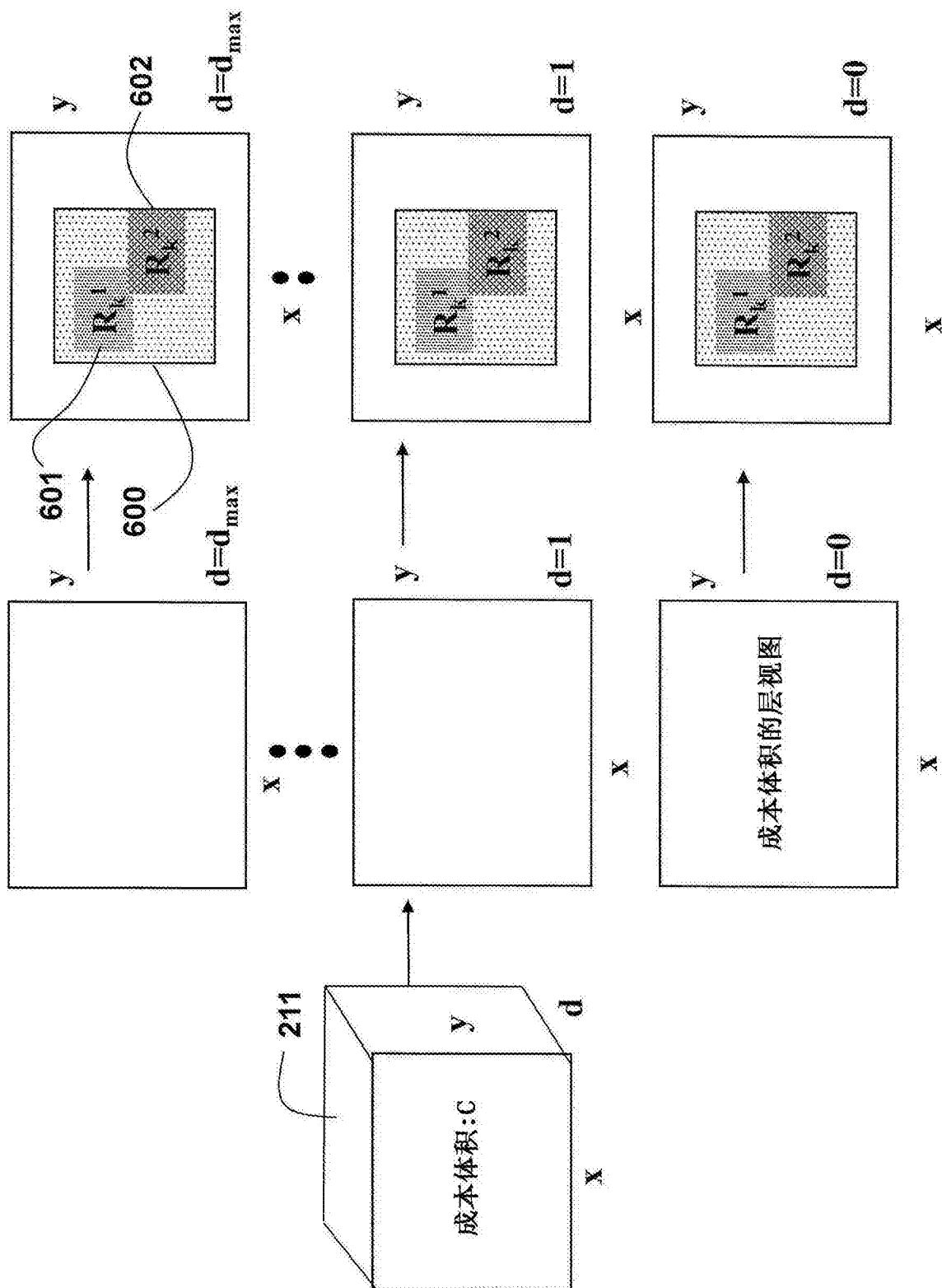


图6

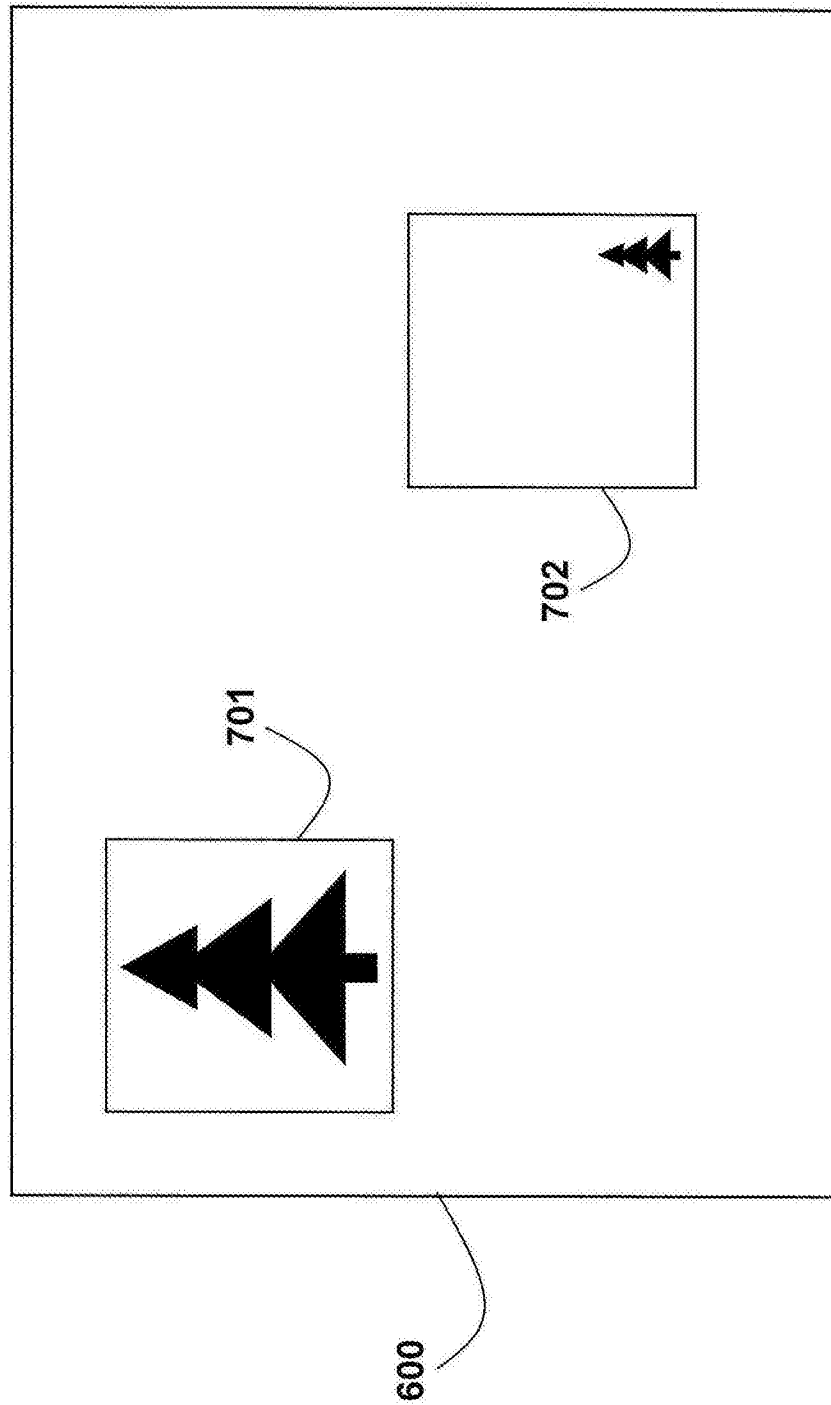


图7