



(12) 发明专利

(10) 授权公告号 CN 101488187 B

(45) 授权公告日 2013. 05. 08

(21) 申请号 200810090409. 1

US 6993177 B1, 2006. 01. 31, 全文.

(22) 申请日 2008. 03. 31

US 2005/0018904 A1, 2005. 01. 27, 全文.

(30) 优先权数据

审查员 金晶

08150403. 7 2008. 01. 18 EP

(73) 专利权人 MV 科技软件有限责任公司

地址 德国慕尼黑

(72) 发明人 安德烈亚斯·霍夫豪泽

卡斯滕·施特格

(74) 专利代理机构 北京集佳知识产权代理有限公司

公司 11227

代理人 高少蔚 席兵

(51) Int. Cl.

G06K 9/46 (2006. 01)

(56) 对比文件

US 6005978 A, 1999. 12. 21, 全文.

US 6324299 B1, 2001. 11. 27, 全文.

US 6411734 B1, 2002. 06. 25, 全文.

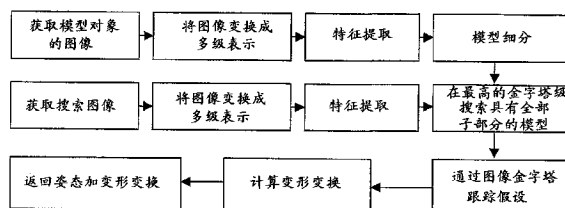
权利要求书2页 说明书10页 附图7页

(54) 发明名称

可变形对象识别系统和方法

(57) 摘要

本发明提供了一种用于甚至在存在部分遮挡、杂波以及非线性照度变化的情况下在图像中检测可变形对象的系统和方法。公开了一种用于可变形对象检测的综合方法,该方法结合了:基于模型点的归一化梯度方向的匹配度量、将模型分解成部分的分解以及将针对所有部分的所有搜索结果同时加以考虑的搜索方法的优点。尽管模型被分解成子部分,但是用来在最高金字塔级搜索的模型的相关尺寸并没有减小。因此,本发明没有现有技术方法中所存在的数目减小的金字塔级的速度限制。



1. 一种用于识别图像中的经历广义非线性变形的对象的方法,该方法包括以下步骤:
 - (a) 在电子存储器中获取模型对象的图像;
 - (b) 将模型对象的图像变换成与搜索空间的递归细分一致的多级表示,所述多级表示至少包括所述对象的图像;
 - (c) 针对搜索空间的每一个离散化级,生成模型对象的至少一个预先计算的模型,所述预先计算的模型包括多个具有对应方向矢量的模型点,所述模型点和方向矢量通过返回每一个模型点的方向矢量的图像处理操作产生;
 - (d) 生成将所述多个模型点细分成多个部分的细分,其中,模型的变形实例通过对所述部分进行变换来表示;
 - (e) 在电子存储器中获取搜索图像;
 - (f) 将搜索图像变换成与搜索空间的递归细分一致的多级表示,所述多级表示至少包括所述搜索图像;
 - (g) 对所述搜索图像的多级表示的每一个变换的图像执行图像处理操作,该图像处理操作返回所述搜索图像内的图像点的子集的方向矢量,该子集对应于搜索所述至少一个预先计算的模型所应该针对的变换范围;
 - (h) 计算将局部度量的结果相结合的全局匹配度量,其中,对于局部度量,在接近于预先计算的模型的受限制的仿射变换范围内搜索模型的所述部分,并将每个部分的最适合者视为该部分对全局匹配度量的贡献;
 - (i) 确定那些其全局匹配度量超过了用户可选择的阈值且其全局匹配度量局部最大的模型姿态,以及,根据这些模型姿态产生处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例的列表;
 - (j) 计算描述所述部分的局部位移的变形变换;
 - (k) 通过搜索空间的递归细分,跟踪处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例,直到达到最精细的离散化级;
 - (l) 在每一级计算相应的变形变换,并将所述变形变换传送到下一级;
 - (m) 在最精细的离散化级提供模型对象实例的变形变换和模型姿态。
2. 根据权利要求1所述的方法,其中,以高于最精细的离散化级的分辨率来确定所述部分的位置。
3. 根据权利要求2所述的方法,其中,除了用户可选择的阈值之外,只有能够达到用户可选择的最大数量的最佳的模型对象实例也被生成到最粗糙的离散化级上的可能匹配的列表中。
4. 根据权利要求3所述的方法,其中,每一个部分的局部匹配度量的得分必须超过用户可选择的局部阈值,否则,认为该部分被遮挡并且将该部分舍弃而不做进一步的处理。
5. 根据权利要求4所述的方法,其中,每一个部分由多个点组成,其中所述细分生成交叠的点集。
6. 根据权利要求4所述的方法,其中,每一个部分由多个点组成,其中所述点集是分离的集合。
7. 根据权利要求6所述的方法,其中,利用k均值聚类或利用归一化分割来进行所述细分。

8. 根据权利要求4所述的方法,其中,在所述步骤(j)和/或(l)中计算的变换是透视变换。

9. 根据权利要求4所述的方法,其中,模型生成另外接收有关成像装置的内部几何参数和模型的度量信息作为输入,在所述步骤(j)和/或(l)中计算的变换是三维姿态。

10. 根据权利要求4所述的方法,其中,在所述步骤(j)和/或(l)中计算的变换是薄板样条、圆柱变换或调和插值。

11. 根据权利要求4所述的方法,其中,所变换的模型部分和搜索图像的方向矢量的归一化点积之和被用于局部匹配度量。

12. 根据权利要求4所述的方法,其中,所变换的模型部分和搜索图像的方向矢量的归一化点积之和的绝对值被用于局部匹配度量。

13. 根据权利要求4所述的方法,其中,所变换的模型部分和搜索图像的方向矢量的归一化点积的绝对值之和被用于局部匹配度量。

14. 一种用于识别图像中的经历广义非线性变形的对象的系统,该系统包括:

(a) 用于在电子存储器中获取模型对象的图像的装置;

(b) 用于将模型对象的图像变换成与搜索空间的递归细分一致的多级表示的装置,所述多级表示至少包括所述对象的图像;

(c) 用于针对搜索空间的每一个离散化级,生成模型对象的至少一个预先计算的模型的装置,所述预先计算的模型包括多个具有对应方向矢量的模型点,所述模型点和方向矢量通过返回每一个模型点的方向矢量的图像处理操作产生;

(d) 用于生成将所述多个模型点细分成多个部分的细分的装置,其中,模型的变形实例通过对所述部分进行变换来表示;

(e) 用于在电子存储器中获取搜索图像的装置;

(f) 用于将搜索图像变换成与搜索空间的递归细分一致的多级表示的装置,所述多级表示至少包括所述搜索图像;

(g) 用于对所述搜索图像的多级表示的每一个变换的图像执行图像处理操作的装置,该图像处理操作返回所述搜索图像内的图像点的子集的方向矢量,该子集对应于搜索所述至少一个预先计算的模型所应该针对的变换范围;

(h) 用于计算将局部度量的结果相结合的全局匹配度量的装置,其中,对于局部度量,在接近于预先计算的模型的受限制的仿射变换范围内搜索模型的所述部分,并将每个部分的最适合者视为该部分对全局匹配度量的贡献;

(i) 用于确定那些其全局匹配度量超过了用户可选择的阈值且其全局匹配度量局部最大的模型姿态以及根据这些模型姿态产生处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例的列表的装置;

(j) 用于计算描述所述部分的局部位移的变形变换的装置;

(k) 用于通过搜索空间的递归细分,跟踪处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例,直到达到最精细的离散化级的装置;

(l) 用于在每一级计算相应的变形变换,并将所述变形变换传送到下一级的装置;

(m) 用于在最精细的离散化级提供模型对象实例的变形变换和模型姿态的装置。

可变形对象识别系统和方法

技术领域

[0001] 本发明总体上涉及机器视觉系统,更具体地,涉及对经历部分遮挡 (partial occlusion)、杂波 (clutter) 或非线性对比度变化的非线性变换对象的视觉识别。

背景技术

[0002] 对于许多计算机视觉应用,特别是机器视觉应用来说,在图像中将给定的二维对象模板快速、鲁棒而准确地本地化是自然的先决条件。例如对于拾放应用,对象识别方法必须确定成像对象的位置。当结合成像设备的已知几何参数给定对象的位置时,可以通过本领域众知的方法来计算对象的姿态。当给定这个姿态时,机器人可以从例如传送带抓取对象。在各种检验工作中,提取对象的位置使得能够将图像中的找到区域 (found region) 解扭曲 (un-warp),并且便于光学字符识别或与例如用于对可能的制造误差进行检测的原型图像的比较。

[0003] 在本领域中已经提出了几种用来在图像中确定对象的位置的方法。大多数方法比较一组可能的对象姿态与图像之间的相似性。超过阈值并且就该相似性量度而言局部最大的位置被选作对象的位置。

[0004] 依赖于所使用的相似性量度,实现了相对于不利成像条件的某种不变性。例如,以归一化相关性作为相似性量度,实现了相对于模型图像和搜索图像 (search image) 之间的线性灰度值变化的不变性。与本发明特别相关的是相对于部分遮挡、杂波和非线性对比度变化而具有不变性的相似性量度,通过引用将其结合在本申请中 (US 7,062,093、EP 1193642 和 JP 3776340)。所述度量 (metric) 的总体思想是将图像和模型特征的归一化方向的点积用作模型和图像之间的相似性的量度。

[0005] 典型地,遍及全部姿态参数的穷尽搜索的计算代价很大,并且对于大多数实时应用来说是不容许的。大多数现有技术的方法通过用模型图像和搜索图像建立图像金字塔来克服这种速度限制 (参见例如 Tanimoto (1981) (Steven L. Tanimoto, Template matching in pyramids, Computer Graphics and Image Processing, 16:356-369, 1981), 或者 Brown (1992) (Lisa Gottesfeld Brown, A survey of image registration techniques, ACM Computing Surveys, 24(4):325-376, 1992 年 12 月))。然后,只在最高金字塔级针对整个搜索范围来评估相似性量度。在较低级,只跟踪有希望的匹配候选者,直达到最低的金字塔级。这里,所使用的金字塔级的数目是一个关键的决策,其直接影响对象识别方法的运行时间。典型地,基于最高金字塔图像中的对象的最小尺寸来选择金字塔级的数目。如果该图像中的对象很小,则难以将对象与例如杂波区分开来。这样就必须评估过多的可能匹配候选者。如果选择的金字塔级不足,则在最高金字塔级的搜索将慢到难以进行的程度。

[0006] 加快搜索的另一种方法是假定受检对象的运动参数可以通过线性仿射变换来近似。根据下式,线性仿射变换将输入点 $(x, y)^T$ 映射到输出点 $(x', y')^T$:

$$[0007] \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}.$$

[0008] 可以将这个广义公式进一步分解成更具几何意义的参数表示式：

[0009]

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} 1 & -\sin \theta \\ 0 & \cos \theta \end{pmatrix} \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}.$$

[0010] 这些参数描述了：将原 x 和 y 轴缩放不同的比例因子 s_x 和 s_y ； y 轴相对于 x 轴的倾斜变换 (skew transformation)，即 y 轴旋转角度 θ ，而 x 轴保持固定；两个轴均旋转角度 φ ；最后平移矢量 $(t_x, t_y)^T$ 。典型地，对象识别系统只是针对例如只有平移和旋转的简化子集来评估这些参数。此外，参数被局限于某个固定范围，例如减小的旋转范围。这就减小了对象识别系统在最高的金字塔级所必须检查的可能姿态的空间，从而提高了搜索速度。

[0011] 然而，在多种情况下，必须要找到的对象是根据较之于线性仿射变换或其子集更广义的变换来变换的。一种这样的变换是透视变换，该变换描述了根据下式对从不同相机位置成像的平面对象的映射：

$$[0012] \quad \begin{pmatrix} x' \\ y' \\ t' \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ t \end{pmatrix}$$

[0013] (参见 Hartley 和 Zisserman(2000) [Richard Hartley 和 Andrew Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2000])。这九个参数按比例限定，产生 8 个自由度。

[0014] 我们明确地把两种情况区别开：一种情况是，对象识别系统的最终任务只是修正图像；另一种情况是，必须确定对象的姿态。对于前者，确定透视变换就足够了。这里，反转的透视变换被用来修正图像。

[0015] 对于必须要确定对象的三维姿态，而且提供了相机内部参数的情况，只用 6 个自由度就足以描述姿态 (3 个用于平移，3 个用于旋转)。要指出的重要的一点是，透视变换不能总是被直接变换成姿态，因为为了产生真实的姿态，必须为透视变换的 8 个参数另外施加两个非线性约束 (Berthold K. P. Horn, Projective Geometry considered Harmful, 1999)。一旦得到有效的透视变换，则可以借助于本领域众知的方法将它直接分解成三维姿态 (例如 Oliver Faugeras, Three-dimensional computer vision: a geometric viewpoint, The MIT press, 2001, chapter 7.5)。优选的方法是：直接搜索三维姿态参数，而不是先确定透视变换然后将其分解成姿态。

[0016] 线性变换不能满足需要的另一个实例是当对象的图像发生非线性变形时。这可能是由于未能预先校正相机透镜系统导致的畸变。进一步的实例是当成像是在像热空气那样产生不规则畸变的介质中进行或者图像是在水下所拍摄时。非线性变换的另一个来源是当对象本身可变形时，例如当它被印在弯曲或皱折的表面时。这里，不但要确定姿态，而且还必须同时确定模型的变形。对非刚性变形的数学描述是增加扭曲 $W(x, y)$ ，以便根据下式对点进行变换：

$$[0017] \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = W(x, y) + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}.$$

[0018] 如果 $W(x, y) = \sum_{i=1}^n w_i U(|P_i - (x, y)|)$ 且 $U(r) = r^2 \log r^2$, 则得到薄板样条函数 (Fred L. Bookstein, "Principal Warps: Thin-plate Splines and the Decomposition of Deformations", IEEE Transactions on pattern analysis and machine intelligence, Vol 11, No. 6, 567-585 1989)。这里, 扭曲 (warp) 由锚点 P_i 和系数 w_i 来参数化。所得到的扭曲使锚点之间的曲率最小。

[0019] 用于非线性对象识别的大多数现有技术的方法做这样的假定: 即使整个对象发生变形, 模型的足够小的部分即使在变形之后在图像中仍十分相似。

[0020] 然而, 如何将这个假定结合到对象识别系统的有效搜索方法中仍是悬而未决的问题。一种方法 (参见例如 US 7, 239, 929 或 US 7, 190, 834) 包括以分级的方式来组织模型的分解部分。这里, 一个部分被选作细分 (subdivision) 的根部。从这个根部开始, 其它对象以树状结构组织。应该指出的重要的一点是: 在随后的搜索中对这个根部进行单独检测。一旦检测到这个根部, 后续部分的可能位置的范围基于对象变形的假定而变窄。对其它部分的搜索由此而简化。

[0021] 然而, 这个现有技术的方法存在几个明显的问题。一个问题是: 对部分的搜索典型地不如对整个对象的搜索容易辨别, 因为部分所包含的信息必然较少。这就导致了假匹配并因为必须评估更多的匹配假设 (match hypothesis) 而导致搜索速度减小。进一步的局限是: 部分的尺寸小于整个模型的尺寸, 因此, 在图像中的模型的相对尺寸变得太小而不能被基于特征的方法使用之前, 只能使用较少的金字塔级。

[0022] 本发明的目的是一种用于可变形对象检测的综合方法, 该方法结合了所述不变性匹配度量、将模型分解成部分的方法以及将对所有部分的所有搜索结果同时加以考虑的搜索方法的优点。尽管模型被分解成子部分, 但是用来在最高金字塔级搜索的模型的相关尺寸并没有减小。因此, 本发明没有现有技术方法中所存在的数目减小的金字塔级的速度限制。

发明内容

[0023] 本发明提供了一种用于对象识别的系统和方法, 其对部分遮挡、杂波以及非线性照度变化是鲁棒的, 并且, 即使当对象由于透视变形或更广义的变形而变换时, 也能够对对象进行识别。

[0024] 本发明包括一种用于识别图像中的经历广义非线性变形的模型对象的方法, 该方法包括以下步骤:

[0025] (a) 在电子存储器中获取模型对象的图像;

[0026] (b) 将模型对象的图像变换成与搜索空间的递归细分一致的多级表示, 所述多级表示至少包括所述对象的图像;

[0027] (c) 针对搜索空间的每一个离散化级, 生成模型对象的至少一个预先计算的模型, 所述预先计算的模型包括多个具有对应方向矢量的模型点, 所述模型点和方向矢量通过返回每一个模型点的方向矢量的图像处理操作产生;

[0028] (d) 生成将所述多个模型点细分成多个部分的细分, 其中, 模型的变形实例通过对所述部分进行变换来表示;

- [0029] (e) 在电子存储器中获取搜索图像；
- [0030] (f) 将搜索图像变换成与搜索空间的递归细分一致的多级表示，所述多级表示至少包括所述搜索图像；
- [0031] (g) 对所述搜索图像的多级表示的每一个变换的图像执行图像处理操作，该图像处理操作返回所述搜索图像内的图像点的子集的方向矢量，该子集对应于搜索所述至少一个预先计算的模型所应该针对的变换范围；
- [0032] (h) 计算将局部度量的结果相结合的全局匹配度量，其中，对于局部度量，在接近于预先计算的模型的受限制的仿射变换范围内搜索模型的所述部分，并将每个部分的最适合者视为该部分对全局匹配度量的贡献；
- [0033] (i) 确定那些其全局匹配度量超过了用户可选择的阈值且其全局匹配度量局部最大的模型姿态，以及，根据这些模型姿态产生处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例的列表；
- [0034] (j) 计算描述所述部分的局部位移的变形变换；
- [0035] (k) 通过搜索空间的递归细分，跟踪处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例，直到达到最精细的离散化级；
- [0036] (l) 在每一级计算相应的变形变换，并将所述变形变换传送到下一级；
- [0037] (m) 在最精细的离散化级提供模型对象实例的变形变换和模型姿态。
- [0038] 在从属权利要求中说明了根据本发明的优选实施例。
- [0039] 本发明包括一种用于识别图像中的经历广义非线性变形的模型对象的系统，该系统包括：
- [0040] (a) 用于在电子存储器中获取模型对象的图像的装置；
- [0041] (b) 用于将模型对象的图像变换成与搜索空间的递归细分一致的多级表示的装置，所述多级表示至少包括所述对象的图像；
- [0042] (c) 用于针对搜索空间的每一个离散化级，生成模型对象的至少一个预先计算的模型的装置，所述预先计算的模型包括多个具有对应方向矢量的模型点，所述模型点和方向矢量通过返回每一个模型点的方向矢量的图像处理操作产生；
- [0043] (d) 用于生成将所述多个模型点细分成多个部分的细分的装置，其中，模型的变形实例通过对所述部分进行变换来表示；
- [0044] (e) 用于在电子存储器中获取搜索图像的装置；
- [0045] (f) 用于将搜索图像变换成与搜索空间的递归细分一致的多级表示的装置，所述多级表示至少包括所述搜索图像；
- [0046] (g) 用于对所述搜索图像的多级表示的每一个变换的图像执行图像处理操作的装置，该图像处理操作返回所述搜索图像内的图像点的子集的方向矢量，该子集对应于搜索所述至少一个预先计算的模型所应该针对的变换范围；
- [0047] (h) 用于计算将局部度量的结果相结合的全局匹配度量的装置，其中，对于局部度量，在接近于预先计算的模型的受限制的仿射变换范围内搜索模型的所述部分，并将每个部分的最适合者视为该部分对全局匹配度量的贡献；
- [0048] (i) 用于确定那些其全局匹配度量超过了用户可选择的阈值且其全局匹配度量局部最大的模型姿态以及根据这些模型姿态产生处于搜索空间的最粗糙的离散化级的所述

至少一个预先计算的模型的实例的列表的装置；

[0049] (j) 用于计算描述所述部分的局部位移的变形变换的装置；

[0050] (k) 用于通过搜索空间的递归细分,跟踪处于搜索空间的最粗糙的离散化级的所述至少一个预先计算的模型的实例,直到达到最精细的离散化级的装置；

[0051] (l) 用于在每一级计算相应的变形变换,并将所述变形变换传送到下一级的装置；

[0052] (m) 用于在最精细的离散化级提供模型对象实例的变形变换和模型姿态的装置。

[0053] 待识别的对象的模型包括多个具有对应方向矢量的点,这些点可以通过标准图像处理算法来得到,例如通过边缘检测方法或线检测方法。在生成模型时,点集被分成多个部分。这些部分在搜索期间可相对于它们的原始位置移动,从而允许模型灵活地改变它的形状。在一个优选实施例中,模型的每一个部分只包括一个模型点。在另一个优选实施例中,每一个部分包括若干邻近的点,这些点相互之间保持刚性。

[0054] 在搜索期间,例如针对广义仿射姿态范围将原始模型实例化。在每一个位置,模型的实例通过以近距离变换独立地变换每一个部分来变形。针对每一个部分,在这个受限制的范围内的每一个变换处计算匹配度量。在一个优选实施例中,匹配度量是该部分的方向矢量与经预处理的搜索图像的方向矢量的归一化点积。整个模型的匹配度量是最适合部分的变形变换处的最适合部分的归一化总和。在一个优选实施例中,假定关于匹配度量的得分低于阈值的那些部分处于遮挡状态,从而在进一步的处理中舍弃这些部分。匹配度量最大的部分的变换决定该部分相对于原始位置的变形。这个位移被用来计算预先选择的变形模型。在一个优选实施例中,非线性变形的模型是透视变换。在另一个实施例中,它是例如样条函数或另一种本领域公知的用于对点集进行插值或近似的方法。一旦计算出这个变换函数,可以反转找到图像区域的变形来产生修正的图像。

附图说明

[0055] 根据以下详述,并结合附图,可以更加全面地理解本发明,在附图中：

[0056] 图 1 是本发明的一个优选实施例的流程图,示出了所述方法的步骤；

[0057] 图 2 示出了对象的图像以及用于模型生成的对象周围的兴趣区域；

[0058] 图 3A 示出了模型生成,其中对象在平面表面上；

[0059] 图 3B 示出了模型生成,其中对象被校准板遮挡；

[0060] 图 4A 是一组由边缘滤波器产生的模型点；

[0061] 图 4B 示出了将模型点细分成多个部分的示例细分,其中描述了模型中心以及这些部分相对于模型中心的平移；

[0062] 图 4C 示出了典型的由于所述部分在附近的局部运动而产生的模型的变形；

[0063] 图 5 示出了当前图像,该图像包含对象的两个变形的实例以及就模型的两个找到实例而言的对象检测结果；

[0064] 图 6 示出了通过拟合变形函数而产生的刚性模板与变形模板之间的变形映射,其中变形映射的示例点是部分的中心；

[0065] 图 7 示出了实施图 1-6 的方法的对象识别系统的示意性框图；以及

[0066] 图 8 示出了使用计算机可读介质来实施图 1-6 的方法的计算机的示意性框图。

具体实施方式

[0067] 本发明提供了一种用于可变形对象识别的方法,其对遮挡、杂波以及非线性对比度变化是鲁棒的。

[0068] 本发明还提供了一种图 7 所示的对象识别系统,其用于图 8 所示的计算机,以便实施在这里结合图 1-6 所描述的方法。

[0069] 图 1 给出了对所述方法的步骤的概括。所述方法分成产生对象模型的离线阶段和在搜索图像中找到所述对象模型的在线阶段。模型生成的输入是以未变形方式示出对象的示例图像。在图 2 中,示出了房子形状的示例对象的图像 202。兴趣区域 201 限制对象在图像中的位置。典型地,这个区域由对象识别系统的用户在离线训练阶段限定。如果对象识别系统的用户只对随后修正搜索图像中的对象感兴趣,有时只有对象的一小部分对于随后的处理来说是特别令人感兴趣的。例如,在必须从包裹中读取信件的应用中,可以指定另外的兴趣区域 203,其针对随后必须被解扭曲的模型而限定该区域的位置和尺寸。尽管对象检测系统使用来自整个对象的信息,但是必须修正的只是那个对象特别令人感兴趣的部分。这就提高了速度,特别是当对象只有一小部分令人感兴趣时。

[0070] 在一些应用中,对象识别系统的用户对对象的姿态感兴趣,例如在机器人操作装置的拾放应用中。然而,图像中的兴趣区域 201 只是指定了对象在图像中的位置和尺寸。要确定对象的度量姿态,必须将成像装置的内部几何参数提供给系统。成像装置 300(见图 3A)的内部几何参数典型地由以下几个方面来描述:它的焦距、主点(principal point)在图像中的位置、像素元件在行和列方向的尺寸以及模拟由透镜引起的枕形畸变或桶形畸变的畸变系数。要利用对象识别系统来控制机器人,还需要进一步的信息,如相机在世界坐标系中的姿态。具有相机和机器人在世界坐标系中的位置允许将姿态从相机的局部坐标系转换成机器人的坐标框架以及相反过程。因此,相机坐标系中的姿态信息可以转换成机器人能够直接执行的控制任务。相机的内部和外部参数可以通过本领域已知的方法事先确定(参见例如 MVTec Software GmbH, HALCON 8.0 Documentation, Solution Guide II-F, 3d Machine Vision, 2007)。

[0071] 一旦确定了这些参数,则需要模型对象 301 的兴趣区域在相机坐标系中的相对姿态来对对象进行相对姿态估算(见图 3 和 4)。这是重要的,因为通常没有成像对象的先验度量信息可供利用,因而不能说对象是例如小且接近相机的或者是大且远离相机的。这里的两种情况会得到相同的图像。提供这种度量信息的典型方法是人工指定例如世界点和图像点之间的点对应,并且测量它们在图像和世界两者中的位置,从而消除了成像过程的比例模糊(scale ambiguity)。

[0072] 然而,这种方法容易出错,并且使对象识别系统使用起来麻烦,因为以人工方式提供点对应是一项费时的任务。在一个优选实施例中,将已经测量的平面校准板 303 覆在兴趣对象上,并且获取显示校准板的图像(见图 3B)。在这个示意图中,校准板 303 包含限定点对应的暗圆形。因为校准板的尺寸和点的精确度量位置是已知的,所以可以在相机坐标系中确定校准平面的相对姿态。然后将校准板从对象移走,并获取显示处在与校准板相同的位置的对象的第二幅图像。因为校准板的姿态和用于模型生成的对象的姿态在世界和图像中是相同的,对象的对应姿态得以自动确定。校准板 303 的区域直接与用于模型生成的对

象的图像结合使用。这个过程使得能够容易地使用系统,因为用户不再需要在图像中指定对象的区域。代之以他一次性地将校准网格放置在对象上、拍摄图像、然后去掉校准板并拍摄对象的第二幅图像。在另一优选实施例中,由校准板限定的兴趣区域的矩形形状可以以人工方式调节以适合于任意形状的平面对象,而不是校准网格的矩形形状。这里,只要限定模型的兴趣区域与校准板处在同一平面内,就仍然可以使用由校准板所给出的点对应。此外,本领域的任何技术人员都能够直接将这个过程扩展为甚至可以用于非平面对象。这里,可以例如以循环方式收集平面子结构的兴趣区域和度量信息,并将这些部分组合成对对象的三维描述。

[0073] 对象识别系统将模型生成的图像变换成包含原始图像的经平滑和子采样的版本的递归细分。在以下表述中,递归细分、多级表示和图像金字塔在使用时具有相同的含义。在一个优选实施例中,递归细分是均值图像金字塔。在另一个优选实施例中,应用高斯图像金字塔。从限定模型位置的兴趣区域产生同一个多级表示。针对每一个多级表示,模型生成从图像的所述区域提取边缘点。边缘检测结果在图 4A 中示出。这里,边缘检测不仅提取位置,还提取强对比度变化的方向。所使用的边缘检测是例如 Sobel 滤波或 Canny 边缘检测滤波,或任何本领域已知的从图像提取有向特征点的其它滤波。本发明不限于边缘特征,本领域的技术人员能够容易地将其扩展到线特征或兴趣点特征。为清楚起见,在进一步的讨论中我们只限于边缘点。图 4A 的小箭号 400 表示边缘点的位置和方向。针对每一个模型点,所提取的边缘点被变换成模型坐标框架(由圆圈 401 表示)并被保存到存储器中。因此,系统得到对被成像对象的几何描述。

[0074] 限定模型原点的模型坐标框架 401 典型地通过获取点集的重心来计算。坐标框架的取向与图像相同。因此,将模型坐标框架映射成模板图像坐标框架的变换是简单的平移。在另一优选实施例中,用户针对模型原点提供不同的点。以人工方式设置原点的典型应用是:应该找到在该应用中特别令人感兴趣的对象上的位置,例如金属部件的钻孔中心。显然,这个位置典型地不对应于模型中心。这个中心模型点特别令人感兴趣,因为:由于对象识别系统认为它表示对象位置,所以返回这个点的姿态。

[0075] 通过施加从模型坐标框架到图像坐标框架的广义仿射变换映射,可以将模型的不同实例投射到图像中。然而,到目前为止,模型实例是刚性的。

[0076] 为考虑相继的非线性模型变形,将所述多个边缘点组织成若干组的次多个(sub-plurality)点。通过局部地变换所述若干组的次多个点,组与组相互之间的空间关系发生变化,从而导致整个对象的非线性形状变化。这里,施加于每一个组的局部变换是足够小的仿射变换,或者是其子集,如刚性变换或平移。示例的模型细分在图 4B 中示出。部分生成(part generation)的输入是事先通过特征提取产生的边缘点 400 的集合。

[0077] 一旦提取了边缘点,部分生成的任务是将这些点分组成空间相关的结构 403。这里,本发明假定空间相关的结构甚至在变换之后都保持相同。本发明的一个方面是以人工方式进行这种聚类。这里,用户将他认为会保持相似的部分选择成为一组。本发明的另一个实施例通过自动方法执行聚类。一种直接的方法是对模型设置固定的细分并将在一个细分单元内的点归属于一个部分。另一种方法是计算模型点的邻域曲线图并将固定数量的最近的点选在一个部分内。另一种方法是对点集进行谱分解。这些谱分解方法的实例在本领域内被称为例如 k 均值聚类或归一化分割(Jianbo Shi 和 Jitendra Malik, Normalized

cuts and image segmentation, In IEEE Conference on Computer Vision and Pattern Recognition, 第 731-737 页, 1997)。要指出的重要的一点是, 本发明不限于这样的情况: 即不同组的所述次多个点是分离的集合。在一个优选实施例中, 针对每一个点及其最近的邻近点生成一组次多个点。与所使用的细分方法无关, 模型点被分成 n 个部分, 每个部分包含 k_i 个模型点。为了加速随后的计算, 使用一种数据结构, 对于每一个部分, 该数据结构包含它所包含的模型点的索引 n_{ij} 。这里, 索引 i 的范围是从 1 到 n , 并限定哪一个部分被选择; j 从 1 到 k_i , 并限定该部分的点。如果例如每一个部分具有相同数量的模型点, 则使用矩阵表示, 其中每一行限定部分, 每一列限定该部分中的索引。

[0078] 限定这样一种细分之后, 例如通过获取相应点集的重心来计算每一个部分 403 的中心 402。部分的中心与模型 401 的原点之间的变换 404 保存在模型中。因此, 部分的中心的相对位置被转换成将模型的坐标框架变为部分的坐标框架的变换, 如欧几里德变换式。这些变换允许将模型点的位置和方向从部分的坐标框架转换成模型的坐标框架的转换以及相反的转变。例如通过沿 x 和 y 轴的小的移动或绕部分的中心的旋转来改变模型与部分之间的相对变换 404, 允许将模型的变形版本实例化。由于沿 x 和 y 方向的小的平移而产生的一些示例的变形在图 4C 中说明。

[0079] 本发明的一个方面是对已知的用于在存在部分遮挡、杂波以及非线性照度变化的情况下检测图像中的刚性对象的方法和系统 (参见 US7, 062, 093) 加以扩展。

[0080] 将模型的有向点集与搜索图像的稠密梯度方向场相比较。即使传送到梯度幅度的非线性照度变化相当大, 梯度方向仍保持相同。此外, 在搜索图像中完全避免了滞后阈值 (hysteresis threshold) 或非最大抑制 (non-maximum suppression), 从而实现了相对于任意照度变化的真正的不变性。部分遮挡、噪声以及杂波导致搜索图像中的随机梯度方向。这些效应降低了关于这个度量的得分的最大值, 但不改变它的位置。得分值 (score value) 的语义是匹配模型点的分数。

[0081] 有效搜索的思想是对象识别系统仅对广义仿射变换或它的子集进行全局实例化。通过允许部分的局部移动以及将最大响应作为最适合者, 搜索隐含地对更高级的非线性变换进行了评估。在图 5 中对此进行了说明, 其中示出了具有两个变形模型实例的搜索图像。在左边示出了模型的透射变换实例 500。在右边描述了更复杂的任意变形 501。如图所示, 局部适配的部分 403 在搜索图像中对对象进行近似。改变部分的刚性位置与局部适配的位置之间的局部变换, 允许表示很多种类的模型外观。

[0082] 一项重要的观测结论是: 通过将图像变换成金字塔表示, 在每一级只需补偿小的变形。例如, 即使对象在最低的金字塔级具有复杂的变形, 最高的金字塔级处的外观也不发生大的变化。另一方面, 如果对象具有一个大的变形, 则可以在最高级对其进行补偿。在本发明中, 变形以递归的方式沿着金字塔传送。如果所有较高级别的变形都在较高的金字塔级得到补偿, 则对象的外观在每一级只发生相对小的变化。

[0083] 因此, 通过将搜索度量分成全局部分 s_g 和局部部分 s_l , 本发明扩展了用于可变形对象检测的所述度量。为清楚起见, 我们只给出用于平移的公式, 这意味着只针对每一行 r 和列 c 计算得分。可以直接将其扩展为用于广义仿射参数。如上所述, 模型被分成 n 部分, 每一部分包含 k_i 个模型点。

[0084] 全局度量定义为:

$$[0085] \quad s_g(r, c) = \frac{1}{n} \sum_{i=1}^n s_i(r, c, i),$$

[0086] 意思是：它是针对索引 i 所限定的每一个部分计算的局部匹配的得分值的组合。

[0087] 局部匹配度量定义为：

$$[0088] \quad s_i(r, c, i) = \max_{T_i} \frac{1}{k_i} \sum_{j=1}^{k_i} \frac{\langle T_i(d_{ij}^m), d_{(r+T_i(r_{ij}), c+T_i(c_{ij}))}^s \rangle}{\|d_{ij}^m\| \cdot \|d_{(r+T_i(r_{ij}), c+T_i(c_{ij}))}^s\|}$$

[0089] 这里，ij 对限定指示哪一个模型点在哪一部分中的索引，其中每一部分具有 k_i 个模型点。 r_{ij} 和 c_{ij} 是相应的模型点在模型坐标系中的行和列位移。局部变换 T_i 用来改变模型的形状。典型地，这些是具有例如在每一个方向上 1 个像素平移的小作用的欧几里德变换。上标 m 和 s 限定 d 是模型的方向矢量或搜索图像中的相应位置的方向矢量。

[0090] 在每一个可能姿态位置，每一个部分具有如独立的得分值，该得分值由类似于 US 7,062,093 中所述的度量给出。针对每一个部分，在其原始仿射位置附近的范围内评估所述度量。局部邻域内的最大得分作为该部分的最适合者。通过计算由每一个部分中的模型点的数量归一化的局部度量的结果的总和来获取全局度量。在不偏离本发明的范围的情况下，可以得出各种不同的度量。本发明的一个变体是可以针对每一个部分设置一个该部分所必须超过的阈值。否则，认为所述部分被遮挡，因此将其舍弃而不做进一步的处理。

$$[0091] \quad s_g(r, c) = \frac{1}{n} \sum_{i=1}^n \mathcal{G}(s_i(r, c, i)), \text{ 其中 } \mathcal{G}(x) = \begin{cases} x: x \geq \text{阈值} \\ 0: x < \text{阈值} \end{cases}$$

[0092] 另一个优选实施例是当部分的尺寸不同时。此时，人们通过每一个部分所包含的模型点的数量来对其影响进行加权。

[0093] 甚至在不知道确切的变形时，一组广义仿射变换的全局得分值也允许确定对象近似地位于何处。另一个变体是：为了实现相对于对比度反转 (contrast reversal) 的不变性，从局部得分度量中舍弃极性信息。这是通过在局部度量中使用模型点和图像点的方向矢量的归一化点积的值的绝对值或绝对值的和来完成的。

[0094] 通过获取每一个部分的最佳匹配，不仅获得了得分值，还获得了对变形的估算。这些是限定最大局部得分的局部变换 T_i 。在具有了每一个部分的局部位移之后，相应的非线性模型被拟合。即使对于没有模型点的位置，也可以计算平滑变形。图 6 示出了一个示例变形。部分 402 的中心移动到附近的位置 603。针对这些点来拟合非线性变换，将原始刚性空间（示意性地描述为网格 601）变换成变形空间 602。在本领域这是一个众知的问题，已经提出了根据函数插值和近似来实现的各种解决方案。这里，本发明的一个方面是只将每一个部分的局部位移用作函数点并为每一个点拟合例如透视变换。这可以例如通过直接线性变换方法 (Hartley 和 Zisserman 2000) 来进行。如果假定模型是刚性的，则可以通过诸如 Levenberg-Marquardt 算法的鲁棒迭代非线性最小化来直接拟合三维姿态。对于可变形模型的情况，样条函数由位移来定义。这个样条函数是例如 B 样条函数或薄板样条函数。这些函数的系数通过直接方法计算。然而，如果使用了例如薄板样条函数，则必须反转很大的线性系统 (linear system) 来获取扭曲系数。因此，在另一个优选实施例中使用由模型点的

变形限定的调和插值方法。此时,将模型点的位移插入描述沿行和列方向扭曲的两个图像中。然后,通过被称为调和修补 (harmonic inpainting) 的方法为没有模型点的区域修补变形 (参见 Aubert,G 和 Kornprobst,P.,Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations(second edition), volume 147of Applied Mathematical Sciences, Springer-Verlag,2006)。为了使扭曲很平滑,将变形传回到模型点的原始区域。因此,不仅获得了插值函数,还获得了近似函数。这个方法的优点是:运行时间只是线性相关于对象的尺寸,而不是例如像薄板样条那样三次方地相关于锚点的数量。

[0095] 通常,特别是对于严重变形,不可能一步提取到变形。当给定变形映射时,对全部模型点和相应的方向进行变换。利用这个被变换的模型,现在针对局部位移再一次独立地搜索模型的每一组次多个点。这给出了确定小位移并拟合被评估的模型的循环,直到达到收敛。典型地,通过检查位移是否变得小于预定阈值来检验收敛。对于所限定的超过阈值并且局部最大的全局实例的范围,将具有位置、得分和变形信息的对象假设 (object hypothesis) 放入列表,以便在较低的金字塔级对它们作进一步的检查。在一个优选实施例中,不仅设置关于全局得分值的阈值,而且还设置产生于最高金字塔级的假设的最大数目。此时,全部假设根据它们的得分值而分类,而且只有固定数量的最佳匹配的候选者才被放入进一步处理的假设的列表中。

[0096] 一旦确定了模型在特定金字塔级的确切位置和变形,必须沿着金字塔将该变形传送到下一个金字塔级。这样做是重要的,以便在较低的级仅仅必须评估小搜索范围的局部变形。在一个优选实施例中,来自较低级的原始仿射模型通过递归细分而变换成较高的金字塔级。已经提取的较高级的变形被施加到模型上,而现在已变换的来自较低级的模型被变回到它的原始金字塔级。在这个级上的搜索从根据较高金字塔级的变形变换的模型的实例开始。

[0097] 沿着图像金字塔对假设的跟踪一直进行到达到最低的金字塔级为止。在最低的金字塔级,以甚至高于原始图像的分辨率来确定位移。因此,在亚像素精度的位置对部分进行实例化,并且确定图像中的对应的最大边缘幅度。此时,部分的位移不再由梯度方向限定,而由梯度幅度限定。按照以上方法,使用小位移以很高的精度来拟合变形函数。一旦在最低级找到对象,则返回位置、姿态和变形函数。此外,返回全局得分函数的值,以便为用户提供以多好的程度找到对象的量度。

[0098] 尽管上面详细描述了本发明的几个具体实施例,然而在不偏离本发明的范围的情况下可以对优选实施例进行各种修改。因此,除非在所附权利要求中指出,以上描述不应作为对本发明的限制。

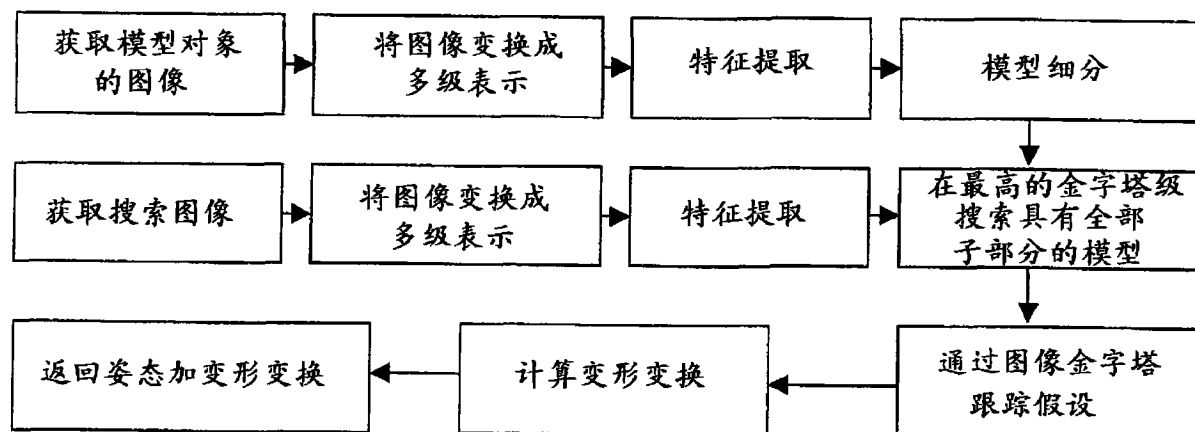


图 1

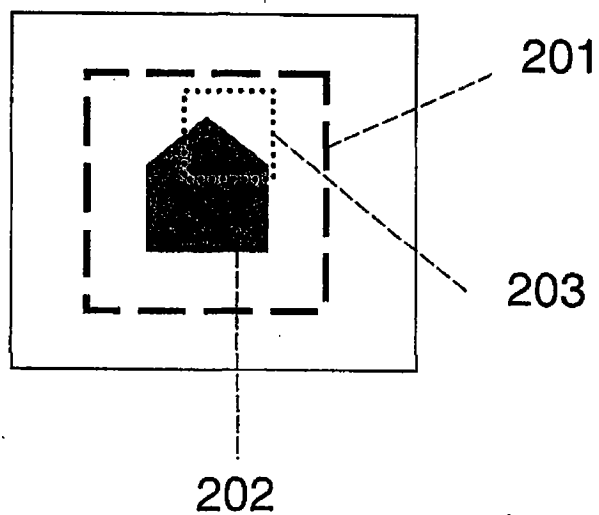


图 2

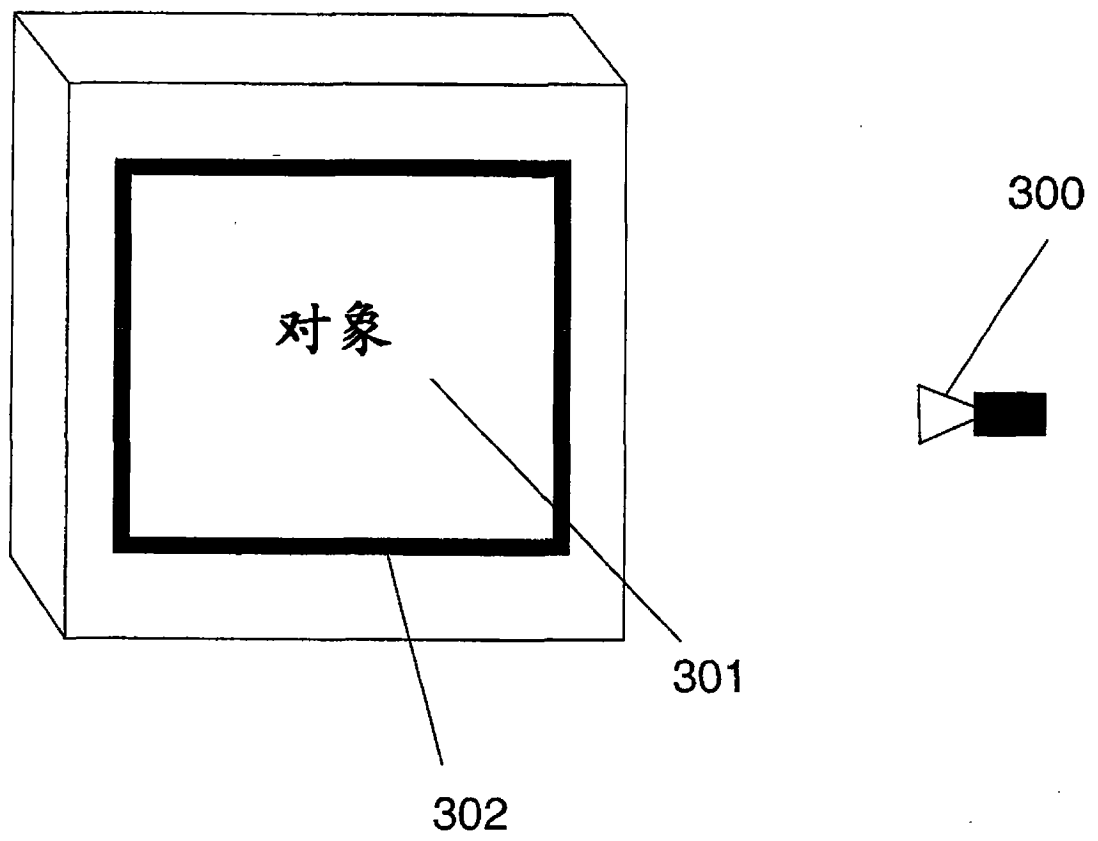


图 3A

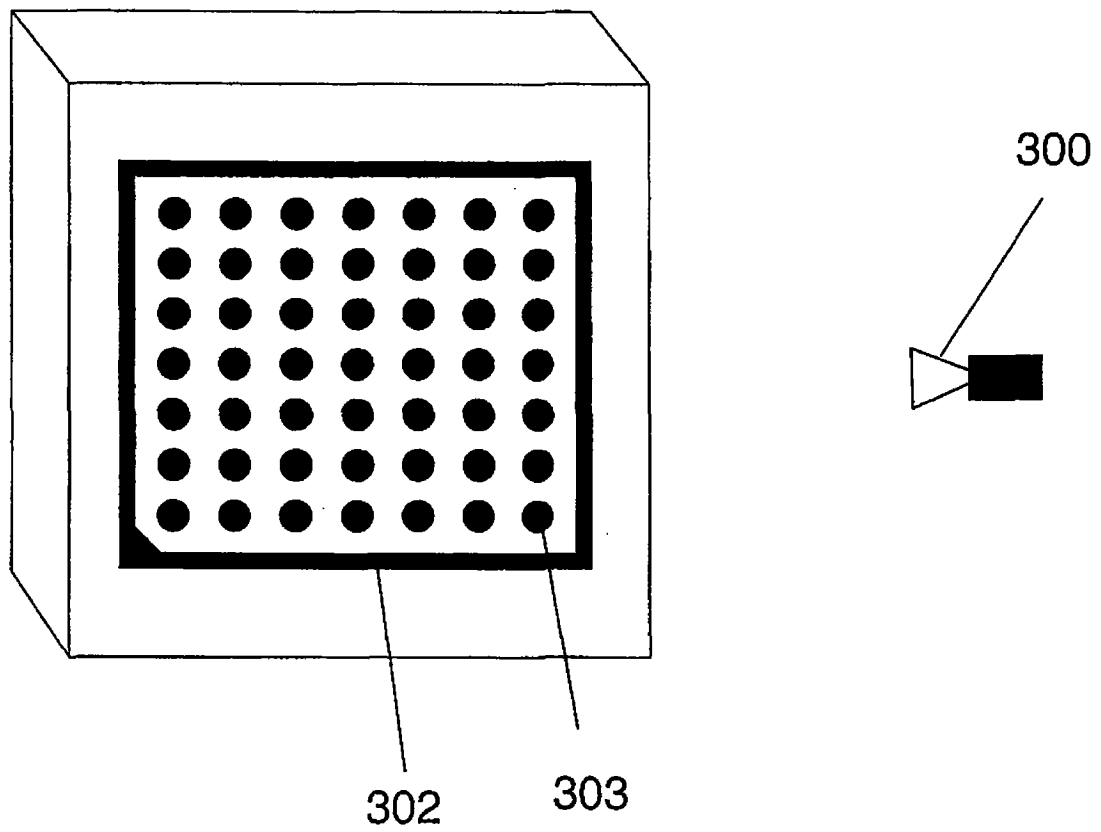


图 3B

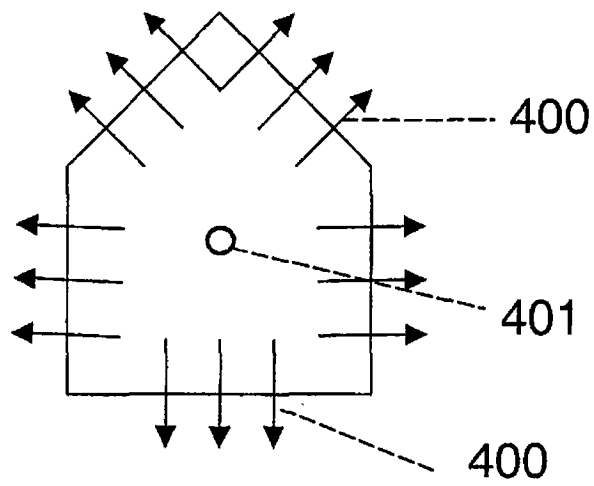


图 4A

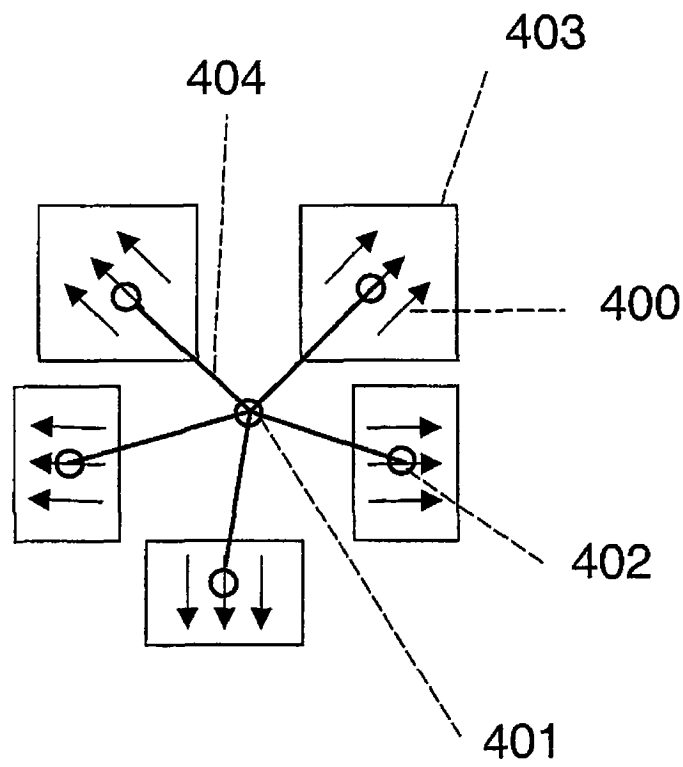


图 4B

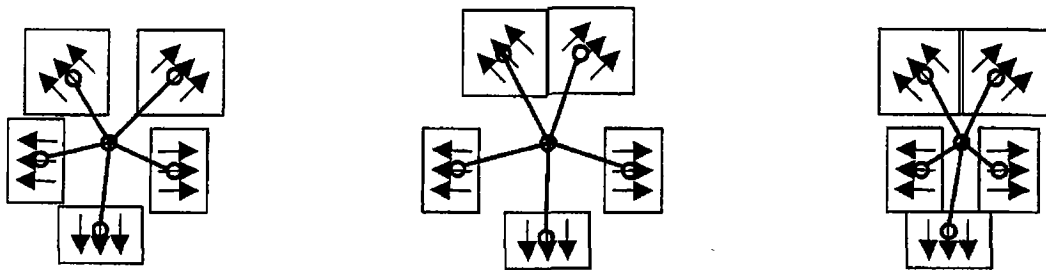


图 4C

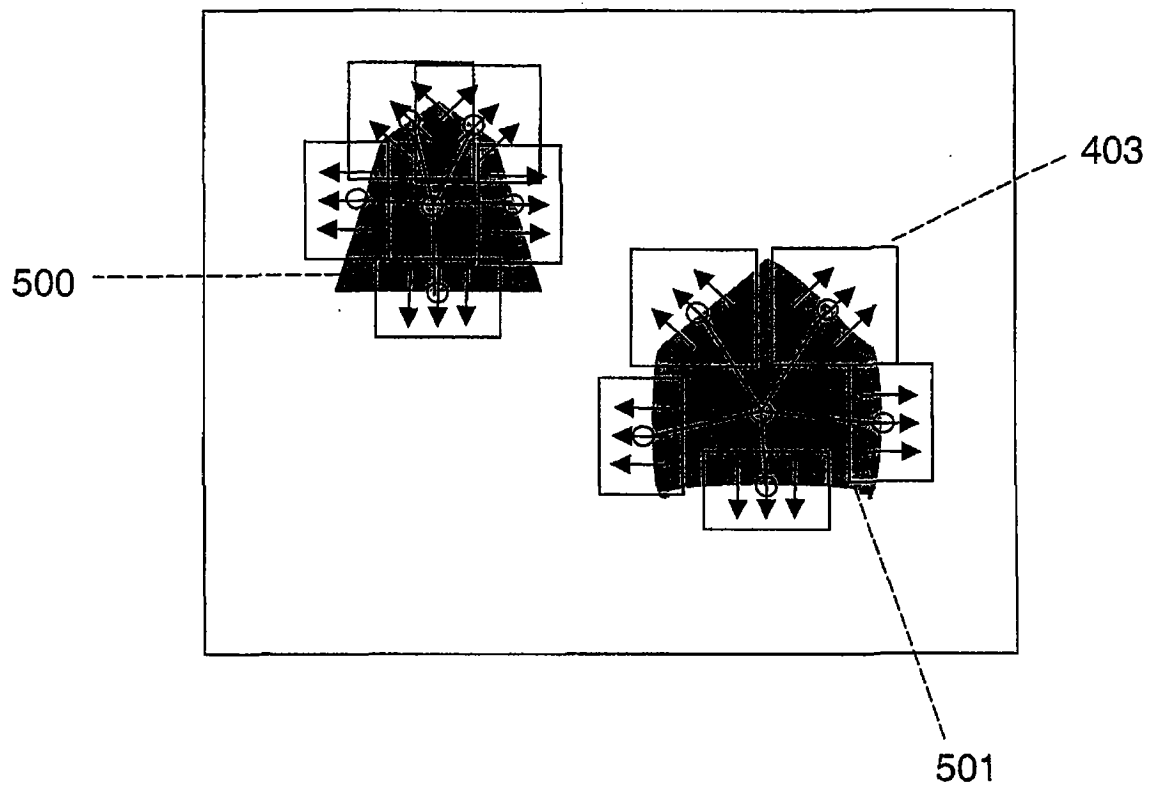


图 5

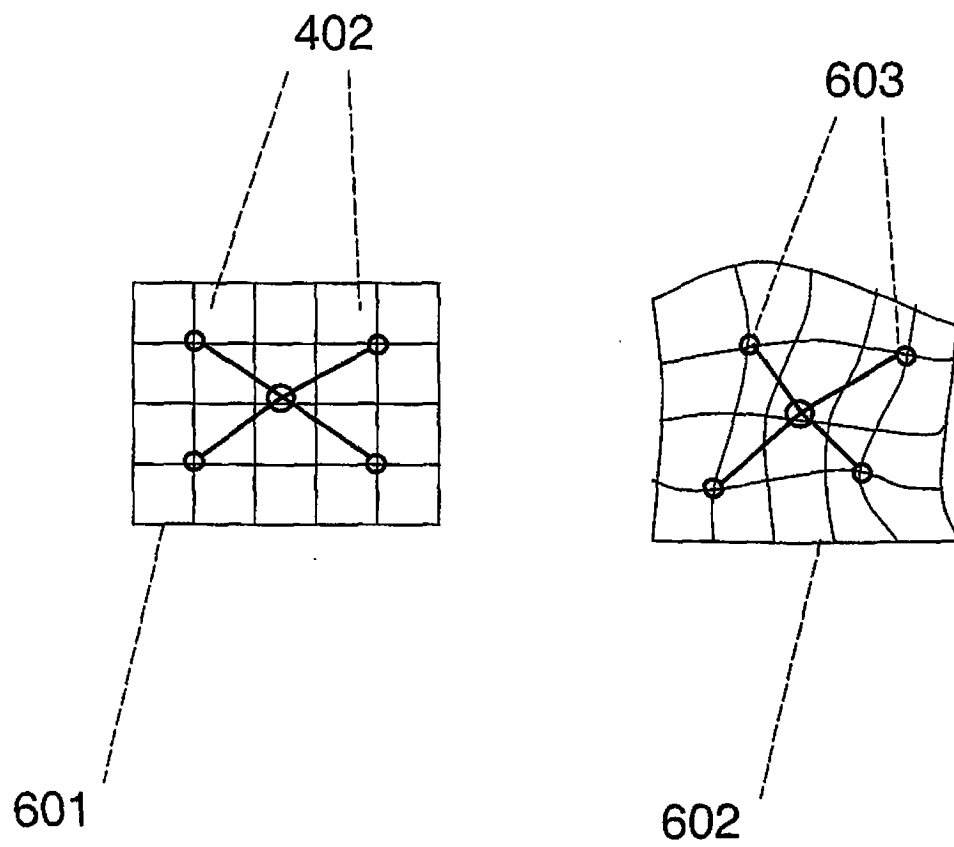


图 6

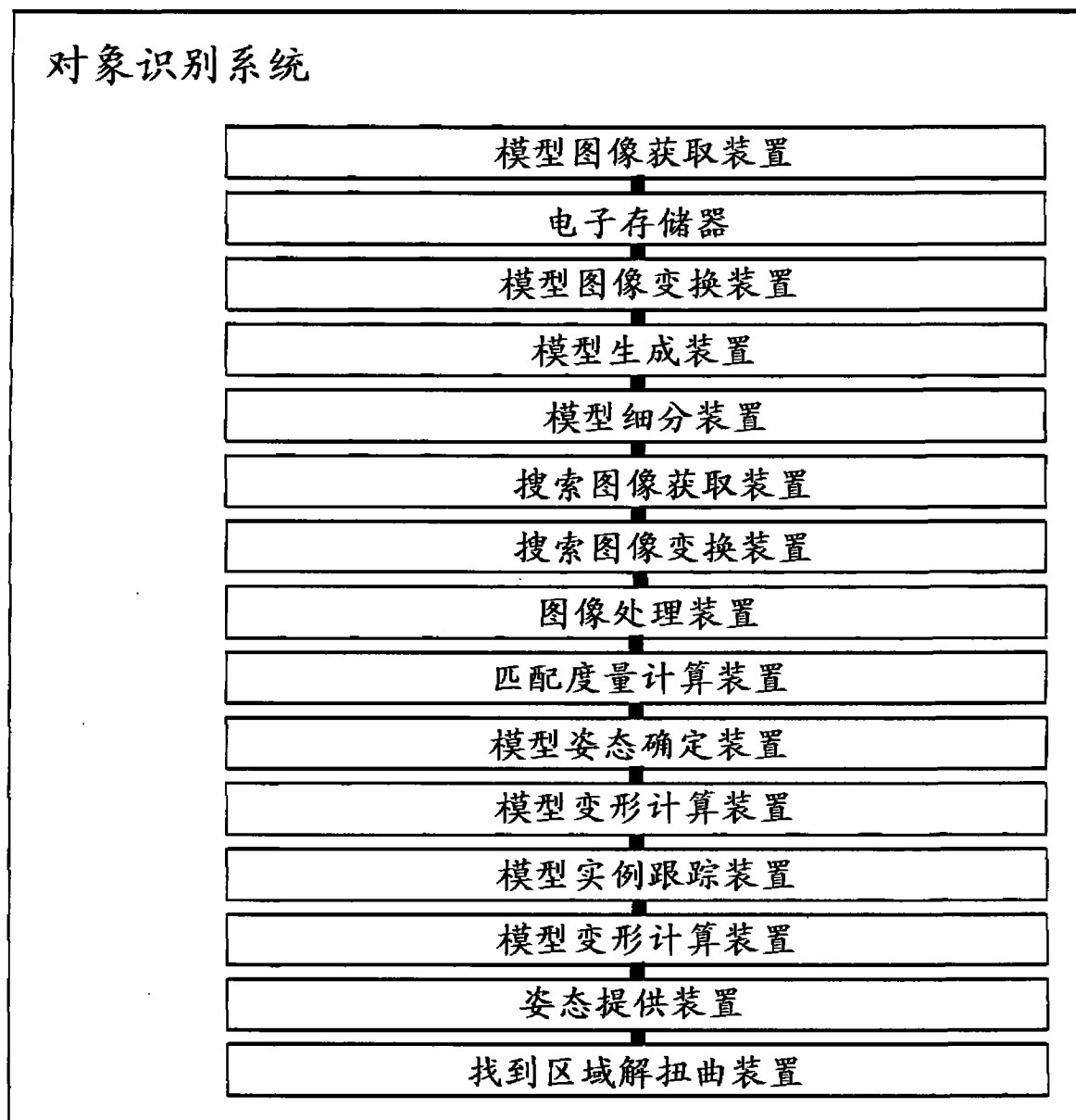


图 7

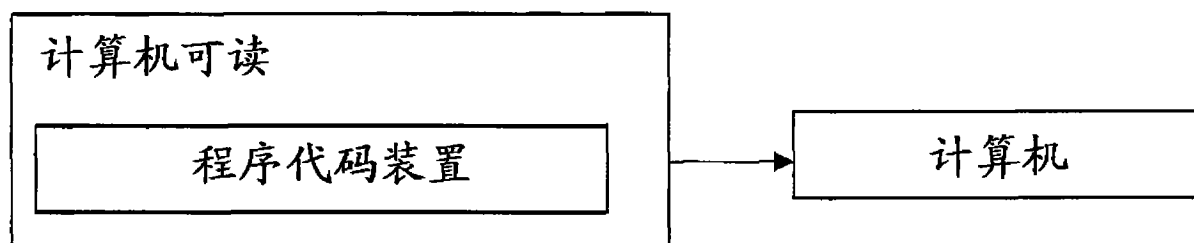


图 8