

[12] 发明专利说明书

[21] ZL 专利号 95104385.4

[45] 授权公告日 2002 年 10 月 9 日

[11] 授权公告号 CN 1092375C

[22] 申请日 1995.4.10 [21] 申请号 95104385.4

[30] 优先权

[32] 1994.4.29 [33] US [31] 235834

[73] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 鲁道夫·M·鲍勒

琼纳森·H·康奈尔 诺曼·哈斯

拉科什·莫汉 加布里尔·塔乌宾

[56] 参考文献

US 5060290A 1991.10.22 G06K9/00

US 5253302A 1993.10.12 G06K9/00

审查员 黄毅斐

[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

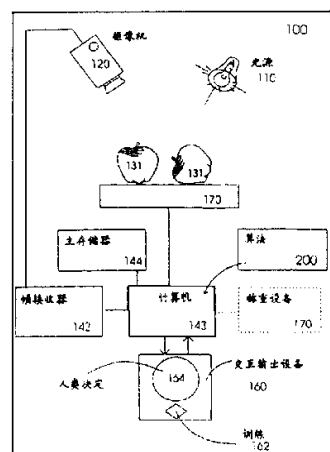
代理人 付建军

权利要求书 6 页 说明书 38 页 附图 16 页

[54] 发明名称 识别对象的系统及方法

[57] 摘要

本系统和装置利用图象处理技术来识别景物中的对象。系统包括用于照明景物的照明源。通过控制照明源,图象处理系统能得到具有以较高水平照明的对象的景物的第一个数字化图象和具有以较低水平照明的对象的第二个数字化的图象。使用某个算法,通过比较这两个数字化的图象,就能把对象图象从景物的背景图象中分割出来。然后把对象的经处理过的图象同被存储的参照图象进行比较。当匹配时,该对象就被识别。



权 利 要 求 书

1. 识别对象的系统，包括：

a. 照明对象的光源，该光源具有在时间周期上不变的非单色光频率分布；

b. 计算机系统，具有能建立包括目标对象图象和背景图象的景物图象的视频输入设备、存储器，所述存储器中存储有多个参考规范化特性，每个特性与一个被分割的参考对象的一个特征相关；该计算机系统还包括用于在景物图象中分割目标对象图象和背景图象，并从被分割目标对象图象的特征获得一个或多个目标规范化特性的装置；以及

把目标规范化特性同一个或多个参考规范化特性比较，并且若目标和参考特性匹配，则目标对象被识别为相关的参考对象的装置。

2. 权利要求 1 的系统，其特征在于特征为色调。

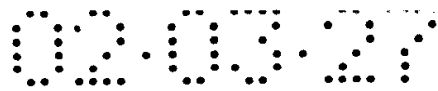
3. 权利要求 1 的系统，其特征在于特性为直方图。

4. 权利要求 1 的系统，其特征在于特性为色调直方图。

5. 识别对象的系统，包括：

a. 照明景物的光源，该光源具有在一段时间内保持不变的非单色光频率分布，光源被控制以第一照明级和低于第一照明级的第二照明级来照明景物，景物包括对象和背景；

b. 计算机系统，具有：存储器，所述存储器中存储有多个



参考规范化直方图，每个直方图与一个被分割的参考对象的一个特征相关；建立包括目标对象图象和背景图象的景物图象的视频输入设备；在景物图象中分割目标对象图象和背景图象的装置，该装置用于通过比较以第一照明级得到的景物图象和以第二照明级得到的景物图象来分割目标对象图象；该计算机系统还具有：

用于计算被分割目标对象图象的每个特征的一个或多个目标规范化直方图的装置，该装置以与计算参考规范化直方图相同的方式计算目标规范化直方图；以及

将目标规范化直方图与一个或多个参考规范化直方图进行比较，并且如果某个特征的目标和参考规范化直方图匹配，则目标对象被识别为相关的参考对象的装置。

6. 权利要求 5 的系统，其特征就在于规范化为面积规范化。

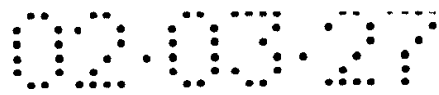
7. 权利要求 5 的系统，其特征就在于规范化为长度规范化。

8. 权利要求 5 的系统，其特征就在于规范化是针对从被分割对象图象中提取的量度。

9. 权利要求 5 的系统，其特征就在于特征为色调，且色调特征被面积规范化。

10. 权利要求 5 的系统，其特征就在于特征为饱和度，而饱和度直方图被面积规范化。

11. 权利要求 9 的系统，其特征就在于还有一个第二特征为饱和度，并且在对象被识别之前，被分割图象的目标规范化饱和度直方图也必须匹配某个参考规范化饱和度直方图。



12. 权利要求 5 的系统，其特征在于特征为质地，并且为了使目标对象能被识别，目标质地直方图必须匹配参考质地直方图。

13. 权利要求 12 的系统，其特征在于，利用区域预测确定质地。

14. 权利要求 12 的系统，其特征在于，利用边缘预测确定质地。

15. 权利要求 7 的系统，其特征在于，特征为形状，并且为了使目标对象能被识别，目标形状直方图必须匹配某个参考形状直方图。

16. 权利要求 5 的系统，其特征在于进一步包括：

a. 称重目标对象的秤；

b. 存在计算机存储器中的对象的参考分割对象平均投射密度；以及

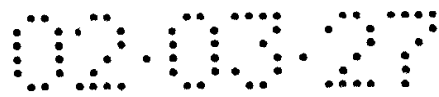
c. 计算机系统被进一步安排用于获取目标分割对象投射密度的装置；

为了辨认目标对象，目标对象投射密度必须匹配参考投射密度。

17. 权利要求 16 的系统，其特征在于，由对象面积划分对象重量，确定平均投射密度。

18. 权利要求 5 的系统，其特征在于，目标和参考对象图象都是经极化过滤器得到的。

19. 权利要求 5 的系统，其特征在于，目标对象具有两个或



更多的目标区域特征，每一个表示目标对象上的一个不同的区域，并且每个目标区域特征直方图和区域特征的相关位置匹配存在计算机存储器中的各自的参考区域特征直方图，以便识别目标对象。

20. 权利要求 19 的系统，其特征不在于，目标对象区域特征就在相关位置中，并且为了使目标对象能被识别，相关位置必须匹配某个被存储的相关位置。

21. 权利要求 5 的系统，其特征不在于，目标对象的面积是这样确定的：通过分别获取没有对象的第一个图象和有对象的第二个图象，并且在第一个图象等于第二个图象的象素位置上从第二个图象中去掉第一个图象而把背景从对象图象中去掉。

22. 权利要求 5 的系统，其特征不在于，获取第一个图象是在光源关闭时，而获取第二个图象则是在光源打开时，并且通过选择在第一个图象中暗的而在第二个图象中较亮的象素来分割目标对象图象。

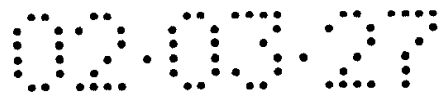
23. 权利要求 5 的系统，其特征不在于，对象为散装货物，而视频输入设备为彩色视频摄象机。

24. 权利要求 5 的系统，其特征不在于，对象为农产品且视频输入设备为彩色视频摄象机。

25. 权利要求 5 的系统，还包括：

确定某个不能识别的目标对象图象是否满足一组存储规则的装置；以及

用于将一个目标规范化直方图与一个或多个参考规范化直



方图相比较，并且如果某个特征的目标和参考直方图不能匹配则目标对象不能被识别为相关的参考对象，并且如果满足存储规则，则该目标规范化直方图被存在存储器中。

26. 权利要求 25 的系统，其特征在于，通过用户界面让用户辨认对象。

27. 权利要求 26 的系统，其特征在于，用户界面为用户提供两个或多个可能的对象供选择。

28. 权利要求 26 的系统，其特征在于，用户界面是一个触摸屏。

29. 权利要求 26 的系统，其特征在于，界面是一种语音识别系统。

30. 权利要求 26 的系统，其特征在于，该界面使用户能浏览对象选择。

31. 权利要求 25 的系统，其特征在于进一步包括：

称重对象的一种秤，使能在识别对象和重量的基础上确定散装货物的价格。

32. 识别对象的方法，包括以下的步骤：

用光源照明对象，该光源具有在一段时间内不变的非单色光频率分布；

用带有视频输入设备的计算机系统建立景物图象，景物图象包括目标对象图象和背景图象，该计算机系统还在存储器中存储一个或多个参考规范化特性；

把目标对象图象从景物的背景图象中分割出来；

从被分割出来的目标对象图象中获得一个或多个目标规范化特性；

将所述一个或多个目标规范化特性与一个或多个在存储器中的参考规范化特性比较，其中每个目标规范化特性对应由所述算法确定的被分割目标对象图象的一个特征，每个参考特性对应与被分割的参考对象相关的一个特征；并

当一个或多个目标规范化特性匹配一个或多个参考规范化特性时，目标对象被识别为参考对象。

说明书

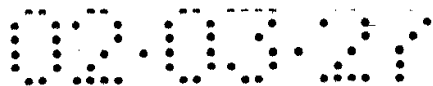
识别对象的系统及方法

本发明涉及利用计算机化的光学扫描设备识别(即,辨认、分类、分级和检验)对象的领域。更具体地,本发明是涉及用图象处理技术识别散装货物的一种可训练系统和方法。

现有技术中存在用于识别对象的图象处理系统,这些系统用直方图来实现这种识别。一种常见的直方图方法是从包含某个对象的(彩色)图象中产生一个灰度直方图或彩色直方图。然后再把这些直方图和参考图象的直方图直接进行比较。或者,抽取这些直方图的特征,并且和从包含参考对象的图象的直方图中的特征进行比较。

参考直方图或这些直方图的特征通常被存在计算机存储器中。现有技术一般用这种方法来验证图象中的目标对象的确是所期望的对象,并且还可能根据相对于参考直方图的其外表的质量来分级/分类该对象。另一个目的可以是通过比较目标图象对象的直方图和多个对象的参考图象的直方图来辨认目标对象。

在本文中,辨认被定义为:给定一组参考对象或类别,决定目标对象是哪个参考对象或目标对象属于哪个参考类别。分类或分级被定义为:确定目标对象被看作是某个对象并且/或者该对象的质量为



某个定量的值。这里,类别中的一个可能是“反”类,意味着对象的质量太差,或者该对象不是已知类别中的成员。另一方面,验证被定义为:确定目标为某个对象或类别,并且简单地验证这是真的或是假的。识别被定义为辨认、分类、分级和/或验证。

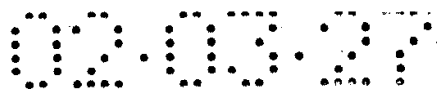
散装货物包括在超级市场、食品杂货店、零售店或五金商店中散装出售的任何货物。例如农产品(水果和蔬菜)、糖、咖啡豆、糖果、铁钉、螺母、螺栓、普通五金器件、零配件和包装货物。

在图象处理技术中,数字图象就是由摄象机得到的模拟图象,通过把该图形划分为称之为图形元素的固定数量的位置、并把在这些图形元素上的图象的值量化为固定数量的值而使之转换为一种离散表示法。所得到的数字图象可由某个计算机算法处理以开发其他的图象。这些图象可以被存在存储器中,并且/或者被用来确定被成象对象的有关信息。象素就是数字图象的图形元素。

图象处理和计算机视觉就是通过对数字图象的计算机处理来修改图象或者从该图象中获取被成象对象的性质,例如对象的特性、位置等等。

景物包括一个或多个感兴趣的对象比及同对象一块被成象的环境。这些环境也被称为背景。背景通常比感兴趣的对象离摄象机更远一些。

图象分割(也被称为图形/背景分离)就是把一个景物图象分割为单独的对象和背景图象。分割指的是辨认与属于背景的图象的象



素相对的那些被包含在对象的图象中的象素。因此,被分割的对象图象就是组成完整景物的原始图象中的对象的象素集。被分割对象图象的面积为对象图象中的象素的个数。

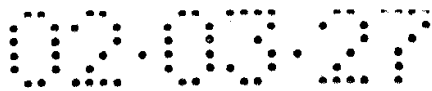
照明度是指照明景物及其中对象的光线。整个景物的照明度直接决定景物中单个对象的照明度,以及由成象装置(如摄象机)所接收到的对象的反射光。

环境照明度是指来自除了专门用于成象某个对象的特定光线之外的任何光源的照明度。例如,环境照明度是由于类似户外太阳光和室内光线这样的环境中产生的光源的照明度。

闪耀或镜子式的反光指的是由光亮的(象镜子那样的、可能是局部的反射特性)对象反射出来的大量光线。闪耀的色彩主要是照明光线的色彩(对比对象的自然色彩)。

图象的特征被定义为该图象的任何特性,能够通过计算来提取。特征通常具有在某个范围中(比如说 R_0-R_1) 的数值。在现有技术中,直方图是在整个图象或图象的窗口(子图)上被计算的。图象的某个特征的直方图是在图象或窗口上特征值分布的数字表示。通过把特征范围 R_0-R_1 划分为 M 个间距(箱)并对每个图象的象素计算特征来产生特征的直方图。只需计数有多少个图象或窗口的象素落在给出特征直方图的每个箱子中。

图象特征包括(但并不局限于)色彩和质地。色彩是一种二维的特性,例如色调和饱和或其他的象素色彩描述(下面将给出说明),但



通常按三维特性处理,即红、绿和蓝(RGB)的数量。用在现有技术中的各种色彩描述包括:(1)RGB空间;(2)对立色彩空间;(3)Munsell(H,V,C)色彩空间;以及(4)色调、饱和、亮度(H,S,I)空间。对于后者来说,类似Munsell空间,色调指的是象素的色彩(从红到绿到蓝),饱和是指色彩的“深度”(例如从带绿色到深饱和的绿色),而亮度则指光亮度,或者在灰度图象中,象素呈现的样子。

另一方面,质地是一种视图特征,更加难以从计算中得到,并且,这种特征不能归属于单个象素,而是一小块图象数据的属性。一个图象块的质地是对该块中的空间亮度变化的描述。这可能是一种质素(texels)的重复模式,如洋葱或菠萝的模式,也可能是更为随机的,如欧芹叶子的模式。这些分别被称为结构质地和统计质地。质地有一个广泛的范围,从二维平面的棋盘形的布置上质素的确定性排列到“盐和胡椒”的白噪声。对图象质地的研究工作已经进行了30多年,并且开发了许多一维的或更高维数的计算方法。然而,在现有的技术中,发明人还不知道有质地特征的直方图。

图象中某些边界的形状是多个边界象素的特征。边界形状指的是局部特征,例如曲率。苹果具有基本上不变的曲率边界,而黄瓜则有一条低的曲率,一条低的负曲率和两条高曲率(端点)。还可以使用其他的边界形状量度。

某些现有的技术利用色彩直方图来辨认对象。给出目标对象的(R,G,B)彩色图象,用于该直方图的色彩表示为“对立色彩”:rg=

$R-G$, $by=2*B-R-G$, 并且 $wb=R+G+B$ 。 wb 轴被划分为 8 段, 而 rg 和 by 轴被分为 16 段。这就产生了 2048 个箱的一个三维直方图。该系统把目标图象直方图同 66 个预先存起来的参考图象直方图进行匹配。这组 66 个预先存放的参考图象直方图是固定的, 因此不是一个可训练的系统, 即: 在某个示例中不可识别的目标图象在以后的示例中也是不能识别的。

Kelly 和 Klein 的美国专利 5060290 提出了基于灰度直方图的杏仁分级。落下的杏仁被一致的光线照射并由一个线性摄象机成象。产生杏仁图象的灰度直方图(被量化为 16 级)。通过让所有的箱计数被 1700 除而实现对直方图的规范化,这里的 1700 个象素是预期最大杏仁的尺寸。从该直方图中抽取 5 个特征:(1)峰值的灰度值;(2)直方图的范围;(3)在峰值上象素的数目;(4)在峰值右边的箱子中象素的个数;以及(5)在箱 4 中象素的个数。通过查找表,产生一个 8 位数字的代码,如果该代码就在库里,则该杏仁就被接受。该系统是不可训练的。可接受质量的杏仁的外形被硬编码在算法中,因此,系统不能被训练来对由杏仁的新示例的表示的杏仁进行不同的分级。

$$\text{级(level)} = (\text{全部反射光的亮度}) / (\text{被桔子反射的绿光的亮度})$$

确定该直方图的中值 N , 并把它作为桔子色彩的表示。在 N 的基础上, 可以把桔子的色彩分为 4 个等级: “优”、“好”、“较好”和“差”, 还可以分得更细。该系统是不可训练的, 其中, 不同等级的桔子的外形被硬编码在算法中。

使用灰度和彩色直方图对于分级或验证图象中的对象是一种非常有效的方法。其主要理由是: 直方图是参照对象的非常紧凑的表示, 不依赖于图象中对象的位置或方位。

然而, 为了有效地使用基于直方图的图象识别方法, 必须满足某些条件。要求 (1) 图象中对象的大小基本上是已知的, (2) 对象几乎不受到遮断 (即对象的大部分都在图象中, 并且不被其他对象所遮蔽), (3) 产生参照对象直方图和目标对象直方图的图象 (参照和目标图象) 的景物照明度几乎没有什么不同, 以及 (4) 能够方便地把对象从背景中分离出来, 或者在背景中几乎没有什么干扰。在这些条件下, 现有的技术中有许多方法能对目标对象的图象直方图和参照对象的图象直方图进行比较。

某些现有技术的匹配系统和方法, 声称不受背景、视点变化、遮蔽和图象分辨率变化的影响。然而, 在某些现有的技术中, 照明条件得不到控制。当获取参照对象直方图的照明度色彩不同于获取目标对象的图象直方图的照明色彩时, 系统就会失败。图象中图象点的 RGB 值在很大程度上取决于照明的色彩 (即使人们能毫无困难地命名给出整个图象的色彩)。因此, 当照明的色彩 (光线频率分布) 变

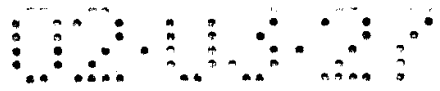


化时,图象的彩色直方图可能发生显著的变化。而且,在这些现有技术的系统中,对象没有从背景中分割出来,因此,图象的直方图不是规范化的区域。这意味着:为了准确识别,目标图象中的对象必须具有和参照图象中的对象一样的尺寸,因为相对于象素大小的对象大小的变化可能显著地改变彩色直方图。这还意味着:图象中对应背景的部分必须是非彩色的(如黑色的),或者至少是对象中所没有的颜色,否则,这将严重地干扰所得到的图象彩色直方图。

如果图象中杏仁的尺寸明显地不同于所预期的杏仁的尺寸,则美国专利 5060290 中所公开的现有技术就会失败。另外,这也是由于该系统不能明确地把对象和背景分开。这个系统只能用于对杏仁分级,而不能区分杏仁和(比如说)花生。

同样,例如在美国专利 4735323 中公开的这类现有技术只能识别不同级别的桔子。一个变红的葡萄柚可能都要被识别为一个非常大的桔子。该系统的设计不是同时针对不同类别的水果,因此没法应付例如绿色的比例对白色反射率这样的弱特征。

总之,在农业领域中的许多现有技术,由美国专利 4735323 和 5060290 所代表的,都涉及到分类/分级农产品货物。如果在扫描仪前一次只通过一个对象,这种现有技术也只能分类/辨认对象/产品/农产品。而且还要求对象/产品/农产品的尺寸范围事先知道(从最小到最大可能的对象尺寸)。如果同时扫描多个货物,或者更准确地说,如果在同一时间上多个对象出现在扫描位置上,则这些系统将会



失败。

此外,现有的技术通常要求细心地工程化和昂贵的机械环境来小心地控制灯光条件,把货物传送到预定的空间位置。这些装置被设计来专门用于一种固定形状的对象(圆的、椭圆的等等),并且不可能或至少不能方便地被修改以处理其他的对象类型。对象的形状决定对象传送的装置,因此对于这些传送装置来说,不可能或难以传送不同的对象类型。对于奇形怪状的对象(如花椰菜或姜)来说更是这样。由于这种原因以及使用为特定对象选择的特征,不允许现有的技术区分不同类型的农产品。

另外,没有一种现有的技术是可训练的系统,在这种系统中,通过人类或计算机的介入,学习新的项目或放弃旧的项目。即:这类系统能被示教,以识别那些最初没有被程序设计在系统中的对象,或停止识别那些最初被程序设计在系统中的对象。

现有技术不能有效发挥作用的一个领域是农产品结帐,目前用于农产品结帐的装置和方法还存在着许多问题。顾客和农产品零售商/批发商都不喜欢在新鲜的农产品上贴上(PLU—价格查找)标签。也不喜观预先包装好的农产品,因为增加了包装和处置(固体垃圾)的费用,并且无法在预包装的形式中检查农产品的质量。

自从食品杂货店第一次出现以来,农产品销售结帐的处理并没有发生多大的变化。在销售点(POS)上,出纳员必须识别农产品的项目、重量或者计数货物的个数,并确定其价格。目前,在大多数商店

中,后者是通过手工输入和农产品连系在一块的非助记符 *PLU* 代码来实现的。在 *POS* 上,这些代码以打印清单或带有图片的小册子的形式出现。

这种农产品结帐处理导致以下的几个问题:

(1)商店遭受损失(减少收入)。首先,出纳员可能由于疏忽而输入错误的代码号。如果这对顾客有利,顾客将很少会主动地提醒出纳员的注意。其次,对于朋友和亲戚,出纳员可能故意输入较低价格的农产品货物的代码(*Sweethearting*)。

(2)由于农产品的辨认问题,农产品的结帐倾向于降低了结帐的处理过程。

(3)每个新的出纳员都必须接受对农产品的品名、外形和 *PIU* 码的培训。

本发明的一个目的是用于识别象农产品这样的对象的一种改进的装置和方法。

本发明的一个目的是一种用于识别象农产品这样的对象的改进的可训练的装置和方法。

本发明的另一个目的是一种用于在销售点或农产品商店中识别和标价象农产品这样的对象的改进的装置和方法。

本发明的进一步目的是用于自动化产品(如农产品)的辨认的用户界面的改进装置和方法。

本发明是利用图象处理技术来识别景物中的对象的一种系统和

装置。该系统包括照明景物的照明源。通过控制照明源,图象处理系统就能得到具有按较高水平照明的对象的第一个数字化的影物图象,以及具有按较低水平照明的对象的第二个数字化图象。使用某个算法,通过比较所得到的这两个数字化的图象,就能新颖地把对象图象从景物的背景图象中分割出来。然后把对象的被处理过的图象(可以用来表示特征的特性)同被存储的参考图象进行比较。当匹配成立时,对象就被识别。

不被识别的对象的被处理过的图象可以用对象的个性加以标号并存在存储器中,建立在某些准则的基础上,因此,当该对象在将来被成象时,这个不被识别的对象就能被识别。用这种新颖的方式,本发明就能被示教来识别先前未知的对象。

对象的识别不依赖对象的尺寸或数目,因为对象的图象在同参照图象比较之前就被新颖地规范化。

可选择确定对象的其他特征(如重量)的用户界面和装置用在系统中。

图 1 是本系统的一个最佳实施例的简图。

图 2 是一个流程图,表示用来识别对象的本方法的一个最佳实施例。

图 3 说明把一个景物分割为对象图象和背景图象。

图 4 是用于分割图象并识别图象中的对象的装置的最佳实施例的简图。

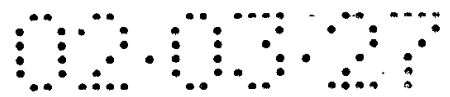


图 5 是分割目标对象图象的一种最佳方法的流程图。

图 6 是一个流程图,说明表示引用目标对象特征的特性的最佳方法。

图 7 是一个流程图,表示对象特征特性(面积/长度)规范化的一种最佳方法。

图 8 表示面积/长度规范化目标对象特性同一个或多个面积规范化参照对象特性的比较。

图 9 是一个流程图,说明训练本装置识别新图象的一种最佳(算法的)方法。

图 10 是一个简图,表示被提取的对象的多个特征。

图 11 是一个流程图,表示质地特征的直方图化和规范化。

图 12 是一个流程图,表示边界形状特征的直方图化和规范化。

图 13 是一个称重设备的简图。

图 14 给出一个图象,其中,通过分割对象图象确定该被分割的对象具有两个不同的区域并且把这些区域加入到识别算法中。

图 15 表示本发明装置的一个人类界面,提供了被成象农产品的最可能的个性的有序等级。

图 16 给出了一种装置,用于由人类通过浏览所有先前被存储的图符图象的子集来确定对象的个性,以及选择子集的装置。

图 17 是利用对象的重量来为对象定价的本发明的最佳实施例。

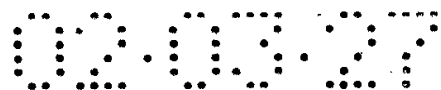
图 1 所示的装置 100 是本发明的一个最佳实施例,利用图象处

理技术自动识别一个或多个对象 131。

具有不随时间而变化的光频率分布的光源 110 照明对象 131。光线是非单色的，并且可能包括红外线或紫外线频率。非单色的并且具有固定频率分布的光线能保证对象 131 的色彩显现不会因所得到的不同图象之间光线的变化而改变并保证给定对象的被存储图象能匹配后来由该对象得到的图象。最佳的灯光为闪光灯管 *Mouser U-4425* 或两个 *GE* 的冷—白荧光灯泡(22 瓦和 30 瓦)，分别为 *GE FC8T9-CW* 和 *GE FC12T9-CW*。这种光源是众所周知的。

视频输入设备 120 被用来把被反射的光线转换为图象。这种图象一般是二维的。最佳的视频输入设备是彩色摄象机，但也可以使用其他任何能把光线转换为图象的设备。这类摄象机应该包括 *CCD* 摄象机和 *CID* 摄象机。彩色摄象机的输出可以是 *12GB*、*HSI*、*YC* 或任何其他色彩表示。一种较好的摄象机是 *Sony* 杆卡式的摄象机 *CCB-C35YC* 或 *Sony XC-999*。象 120 这类的视频输入设备是众所周知的。

在本发明中，彩色图象是最佳的传感模态，然而，其他的传感器模态也是可以的，例如，红外线和紫外线图象，嗅觉/气味(可量度的，例如用质量分光计)，热衰减特性，超声波和磁共振图象，*DNA*，基频，刚度和硬度。这些模态可以用已知的照明和量度方法使能，或者对对象 131 抽样并用一种可兼容的成象设备 120 来产生图象。

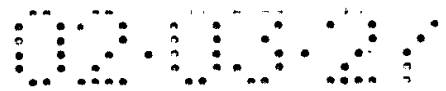


对象 131 是由系统 100 成象和识别的对象。对象 131 可能包括一个或多个项目。虽然最好对象 131 是一种类型的,例如一个或多个苹果,但这些项目可以属于不同的类型,例如谷类食品盒(对象 A)和苹果(对象 B)。系统 100 将识别对象为(1)对象 A,(2)对象 B,(3)对象 A 和对象 B,或(4)认为对象是不可识别的。这些对象实际上可以是由系统 100 成象的任何东西,但最佳的对象 131 是散装货物,包括农产品(水果和蔬菜),五金,盒装货物等。

计算设备 140,一般为计算机 140,用来处理由视频输入设备 120 产生并由帧接收器 142 数字化(和计算机 140 兼容)的图象。

处理过程由算法 200 实现。其他类型的计算设备 140 包括:个人计算机和工作站。计算设备 140 也可以是一个或多个数字信号处理器,或独立的,或被装在计算机中。也可以是能执行算法 200 的任何特定的硬件。一种最佳实施例是和插在 *IBM Valuepoint* 计算机或 *IBM4690* 系列的 *POS* 现金出纳机中的与收据转换 *DT2871* 帧接收器板连接的数据转换 *DSP* 板 *DT2878*。帧接收器 142 是数字化来自摄象机 120 的图象信号的一种设备。如果摄象机 120 是一种数字式的摄象机,则不要求单独的帧接收器 142。数字化仪可以和计算机 140 分离或者作为其组成部分。图象可以存在标准的存储器设备 144 中。给出本公开,熟悉这一技术的人就能开发其他同等的计算设备 140 和帧接收器 142。

一个可选的交互输出设备 160 可以连接到计算设备 140 上,作



为与用户(如出纳员)的界面。输出设备 160 可以包括辅助用户决策 164 的屏幕,也可以提供能训练(162)系统 100 识别新对象的机制。一种可选的称重设备 170 也能把对象 131 的重量(或密度)输入到计算设备 140。见下面(图 13)的描述。

图 2 是由计算设备或计算机 140 运行的算法 200 的流程图。在步 210 中,将被识别的目标对象由摄像机 120 成象。象这类的成象技术是众所周知的。然后把目标对象 131 的图象从其背景中分割出来(步 220)。步 220 的目的是把目标对象 131 和背景分开,使得系统 100 能独立于景物的背景计算被分离对象 131 的图象象素的特性。在步 230 中,从被分离的目标对象图象,最好是逐个象素地计算出对象 131 的一个或多个特征。在步 240 中,产生这些逐个象素被计算特征集的特性。在步 250,规范化能保证这些特性不依赖实际的面积、长度、大小、或者和对象 131 在图象中所占的面积/长度/大小有关的特性,使得对象 131 的一个或多个示例被识别为同一个对象类型。规范化由对象 131 在图象中所占的段的特性化的最佳手段是对被计算的特征特性的次数计数。(这将在图 7 中进一步描述,最佳的规范化手段是通过面积或长度)。在步 260 中,目标对象的计数—规范化特性和被存在内存 270 中的参考对象的计数—规范化特征进行比较。存储器 270 可能位于存储设备 144 或计算机 140 中。(见图 8 中的描述)。在步 251 中,根据计算机存储器 270 中的某准则 255 存放面积—规范化特性。这一步使得系统 100 能被训练,因为存储标



准 255 可能允许存储(步 251)新的参照图象,以后可能被用来比较目标 131 图象(见图 15 中的描述)。

步 220 把对象的图象和背景图象分割或分离出来。执行这一步,使得目标对象的特征能免受景物背景的影响和干扰而被处理。图 3 给出了分割对象图象和背景图象的两种最佳的方法(分别为图 3a 和图 3b)。

图 3a 给出了两个景物。第一个被成象的景物 310 表示没有任何其他对象出现在摄像机 120 的视域中的背景 311 的图象。第二个成象的景物 320 同时包括景物背景 311 的图象和具有一个或多个对象 131 的图象 130。这里,成象对象 130 的象素在对象 131 存在的景物图象 320 那些区域中代替了背景图象 311 中的象素。所以,这是具有对象 131 的示例存在于景物中的背景 311 的图象。

景物 310 和 320 比较,最好是逐个象素比较,允许对象图象 130 从景物的背景图象 311 中被分到(分离)出来。如果对于图象 320 中的某个给定的象素,其亮度不同于(如大于)景物 310 中相同象素的图象亮度,则该象素属于对象图象 130。如果对于图象 320 中的某个给定象素,其亮度等于景物 310 中的相同象素,则该象素属于背景图象 311(见图 5 的描述)。

图 3b 给出了具有由能够分割对象图象的本发明的最佳实施例产生的背景 和一个或多个对象的一个景物的两个图象。图 3b 中的图象 330 是关闭了光源 110 的景物(包括对象 131 和背景 311)的图

象。即,景物图象 330 包括由环境光线照明的背景 311 的图象。另外,在景物图象 330 中还有遮蔽了背景的对象图象 135。由于光源 110 是关闭的,在景物图象 330 中对象图象 135 显得黑,因为没有被光源 110 照明。

图 3b 中的图象 340 是打开光源 110 的景物图象。在这种情况下,光源 110 用大于环境光线的光量照明在摄象机视域中的对象 131。这导致了景物图象 340 中的对象图象 130 更亮(比在图象 330 中)。

图 4 是一个简图,表示最佳系统 400 成象景物(如图 3 中所述的那些景物)、把对象图象 130 从其物理背景 312 的背景图象 311 中分割出来、并且识别对象 131。对这个系统 400 的组成部分 401 的一般描述可以在 1994 年 4 月申请的、题目为“对象成象系统”的美国专利申请号 235064 中找到(授权后的专利号是 5631976),在这里将被引用。

最佳系统 400 把对象 131 放在灯光 110 和摄象机 120 上面,因此提供了从下向上看的对象 131 的图象。系统 400 为对象 131 提供一个支撑物 405,并且还保证对象是固定的且离摄象机 120 有一个可重复的距离。此外,系统 400 允许闪光对象的成象(如塑料包),通过提供由 410 和 420 组成的过滤系统,减少闪光对象图象中的闪耀(镜子式的反射)。

系统 400 包括对灯光 110 和摄象机 120 的一个不透光的外壳 401,该外壳只有一个面向对象 131 的开口 403。开口 403 有足够的

尺寸允许对象 131 被摄象机 120 成象并由灯光 110 照明。开口可以是方的、圆的或任何其他形状。一个透明的表面 405 复盖着外壳 401 中的开口 403。这个表面 405 可以是一块玻璃。透明表面 405 提供一个支撑,可以把被成象的对象 131 放置其上。通过把对象 131 放在表面 405 之上,使得摄象机 120 和对象 131 之间的距离 407 保持固定,因此提供了可重复成象的手段。

为了从对象 131 的图象中消除闪光(从对象 131 本身并且可能从外层的塑料包),在摄象机 120 的透镜中加入一个偏振过滤器 420,或者就把这个过滤器 420 放在摄象机 120 的透镜之上。第二个偏振过滤器 410 放在灯光 110 和开口 403 之间。这保证到达对象 131 的灯光被偏振。另外,灯光也可以被偏振器完全包围起来。如果灯光只是部分地被放在一个盒子中(如摄象机的闪光灯)或借助一个折射器(如照相的聚光灯),则偏振器只需被放在允许灯光通过的灯光装置中的开口上。在第一个过滤器 410 中偏振(极化)的方向垂直于第二过滤器 420 偏振的方向。现有技术公认:从某个对象(如对象 131)反射出来的镜子式反射光被极化为与从该对象反射出来的漫射(无光)反射光相反。因此,具有偏振过滤器 420 的成象对象 131 减少了图象中的闪光。进而,其灯光 110 被 410 极化的照明装置 131 减少了对对象 131 上的闪光量。410 也保证了从对象 131 反射出来的反射光的偏振角度垂直于偏振器 420。因此,成象的对象 131 通过垂直于偏振器 410 的偏振器 420 进一步减少了对象图象 130 中

图 2 的分割步骤 220 将在图 5 中进一步描述, 给出从景物中分割对象图象的一种最佳方法。

在步 520 中,景物的第二个图象是在灯光 110 被关闭或置于步 510 中的亮度级之下的情况下产生的。灯光 110 应该这样被设置,使得对象 131 在第二图象中显得比在第一个图象中较暗。通过执行这些新颖的步骤,对象图象 130 就能在下面的步骤中从背景图象 311 中被分离或分割出来。

- 18 -

象中的被相同编号的象素完全一致。保证正确的定位可通过快速接续来获得第一和第二个图象,或者相对于静止的背景 312 成象静止的对象 131。

获得第一和第二个图象的次序可以颠倒,即,步 520 可以在步 510 之前执行。

在算法 220 的步 530 中,第一和第二图象在帧接收器 142 中被数字化。在计算机 140 中,第一个数字化图象中的每个象素同第二个数字化图象中的相同位置上的对应象素相比较。象这样的逐个象素比较的方式在图象处理技术中是已为人们所了解的。例如,虽然每一对被比较的象素必须互相对应(即在每个图象中各自相同的位置上),但图象中的对应象素对可以按任何次序比较。此外,也可以采用其他的比较方式,例如,常常比较第二或第三象素。

在步 540,在逐个象素的基础上执行检查,确定第一个图象中的象素是否比第二个图象中的对应象素更亮且大于某个值 T 。在任何象素比较中,如果第一个图象中的象素比第二个图象中的对应象素更亮且大于值 T ,则算法 220 执行分枝 542 并指定该象素对应到对象 131。同理,若象素的比较表明:第一个图象中的象素并不比第二个图象中的对应象素的亮度大于值 T ,则算法 220 执行分枝 544,并指定该象素对应物理背景 312 的图象 311。

容差 T 的值可以是一个常量。最佳的容差 T 为最大图象亮度的 5%。另外, T 的值也可以变化,取决于象素在图象中的位置,或取决

图象处理技术进一步处理。

同样,对应背景 312 的象素被集中在被分割的背景图象 311 中。尤其是,来自第一个图象的并不显著地亮于第二个图象中对应的象素的所有象素在其第一个图象的位置上被集中在分割的背景图象中。(在最佳实施例中,“显著地亮于”意味着在对应的象素对中的象素之间的亮度上的差别大于容差 T)。被分割的背景图象对应其对象 130 被去掉的背景 311 的图象。如果需要的话,被分割背景图象中留下的象素(即那些对应被去掉的对象象素位置的象素)可以被赋以任何值,并且/或者用已知的图象处理技术对其作进一步的处理。

如果只需要对象 130 的图象,则不用执行获取 311 的步骤 544。同样,若只需要背景 312 的图象,则不必执行步 542 以获取 130。

在替换的最佳实施例中,对象 131 的半透明部分(如塑料盖)可以通过增加步 552、554 和 556 使之和对象 131 的不透明部分分开。

在该实施例中,分枝 542 执行步 552 而不是步 540。在步 552 之前,已经确定:第一个图象中的象素比其在第二个图象中对应的象素更亮。步 552 确定:第二个图象的对象 130 的象素(在低照明度下的对象 131)是否亮于第二个容差值 V 。若是,则执行分枝 553,并使该对象象素属于对象 130 的半透明部分 554。(该对象在这一象素位置上 是半透明的,因为某些环境的光线通过对象 130 并且当灯光 110 被开到低照明度时在该位置上被成象的)。如果不是,则执行分

枝 555, 并且象素属于对象 130 的不透明部分 556。(没有环境光线或者在容差 V 下的光线数量通过对象 130 的不透明部分被测量到)。对于第二个图象中的每个象素, V 值可以是常量, 也可以是变量, 例如, 取决于该象素在第二个图象中的位置。注意, V 值可以象上述那样进一步被计算, 单独从背景 135 的图象中, 通过选择这样的 V , 使得背景图象的 95% 到 85% 比 V 更亮。 V 的最佳值是最亮的图象亮度的 20%。

在步 554 中, 建立半透明的对象图象。在这一步中, 第一个图象中对应第二个图象中亮于 V 值的象素的每个象素(属于该对象)对应对象 130 的半透明部分, 并且被存在半透明的对象图象中。在对第一个和第二个图象中的所有象素进行如此的处理之后, 半透明的对象图象只包括对象 130 的半透明部分的图象。如果需要的话, 半透明对象图象的剩余象素可以被赋以任意值并且/或者被进一步处理。

在步 556 中建立不透明的对象图象。在该步中, 对应第二个图象中等于或暗于 V 值的象素的第一图象中的每个象素(属于该对象)对应对象图象 130 的不透明部分, 并且被存在不透明对象图象 556 中。在第一和第二个图象中的所有象素都被如此处理之后, 不透明对象图象将只包含对象 130 的不透明部分的图象。如果需要的话, 不透明对象图象的剩余象素可以被赋以任何所需的值, 并且/或者被进一步处理。

如果只需对象 130 不透明部分的图象, 不必执行步 554。同样。

若只要对象 130 的半透明部分的图象,则不需要执行步 556。

在另一个最佳实施例中,步 552 结合步 540 和步 542 去掉对象 130。这样做的结果是产生半透明的对象图象或不透明的对象图象(或两者都产生),但不产生完整的被分割的对象图象 130。

步 552、554 和 556 和前面步骤的其他组合也在本发明的考虑之中。

在图象被分割(步 220)之后,计算一个或多个的目标对象特征。参看图 2 的步 230。计算机 140 执行计算步骤 230,并确定目标对象的特征。执行步 230 仅仅确定在步 220 中得到的目标对象的被分离出来(被分割)的图象 130。对于被分割对象图象中的每个象素,确定特征。例如,利用单个象素的色彩,或利用象素的(色彩)值及其周围象素的(色彩)值来计算这样的特征。特征包括,但不局限于,目标对象被分割图象的色彩、形状、质地和密度。通常,特征由一个或多个特征值表示。

一旦确定了一个或多个特征(步 230),该特征或该组特征被特性化(步 240)。直方图化是实现特性化(步 240)的一种最佳方法。见下面图 6 的描述。然而,也可以使用其他表示特性的方法。例如,中值特征值、计算化特征值的一阶(平均值)和/或高阶统计量、或者能从特征值的计算化集中导出的任何统计量都可以使用。给出本公开,熟悉本领域技术的人就能开发出表示特征的特性的其他等同的方法。

算法 200 的规范化步骤 250 刻划对象特征的特性,而不考虑被成象的实际对象 131 的尺寸。该步骤还使对象 131 的一个或多个示例能被装置 100 识别,而不必考虑对象 131 在景物中的个数或大小。规范化步骤 250 执行一个或多个计算化特征的特性化过程。执行规范化的最佳方法是相对于(例如)通过计数被分割对象图象中象素的个数(见下面图 7 的描述)或计数被分割对象图象边界上象素的个数(见下面图 12 的描述)所得到的面积或长度。

其他的规范化方法,例如,相地于从被分割对象图象中得到的任何其他特性的规范化也在本发明的考虑之中。

本发明的另一新颖特征使系统 100 能被训练。如果对象 131 的规范化特性不能被识别,即,不能和参考信息匹配(步 260),则检查该规范化的特性(步 251)是否满足某些存储规则 255。若不被识别的对象的规范化特性满足存储规则 255,则和其他的参考信息存在一块(步 270)。因此,下次该对象 131 再被系统 100 成象时,就能被和某个参考图象匹配并被识别。见下面图 9 的描述。训练使得系统 100 能识别那些没有为了被识别而被“硬连线”(预先被程序设计好)的对象,这样使得系统更加灵活。被存储特性的规范化使得用于参考特性的对象 131 的个数可以不同于用来开发目标特性的对象 131 的个数。

存储规则可以包括由系统 100 设计所建立的任何准则。选择并且/或者建立存储规则 255 的能力也可以通过界面 160 给用户。简单

的存储规则 可以把关于不被识别对象所提供的信息存在参考数据库 270 中。其他的存储规则可以包括,但不局限于:(1)图象 210 的质量是好的;(2)目标对象占了图象 210 的大百分比;(3)特性对数据库中目标对象的引用应该是足够封闭的(图 8 中的描述 840)。

在算法 200 的步 260 中,目标对象 131 的规范化特性和一个或多个规范化的参考对象特性进行比较(步 260)。这种比较(步 260)在很大程度上取决于特性化对象特征的方法,上述的步骤 240 中给出了这样的例子。一种最佳的比较方法(260)是用面积或长度的规范化直方图来实现的。

一个或多个参考对象特性被存在存储设备中(步 270)。这种设备可能位于计算机 140 上的存储器中,或者是一种独立的外存设备 144。存储参考对象特性(步 270)的最佳方法是利用刻划了对象特征的一系列面积规范化特征直方图。这些面积规范化直方图系列中的每一个都有一个唯一的对象类型标识符。存储面积规范化特征直方图的一种最佳方法是使用规范化特征直方图的一个向量,即不同特征值出现的规范化频率。

图 6 是一个流程图,说明产生特征的直方图的一种最佳方法。在这个没有限制的例子中,使用了特征 F1 色调。然而,可以使用能从被分割图象中提取的任何特征。注意,本发明只从被分割的对象图象 130 中新颖地提取特征。直方图化的特征也可以从关于被分割对象的其他信息中获取。例如,色调可以从彩色图中的其他某些特征中得

到。

为了产生直方图 650,首先对景物成象(图 2 的步 210)并且分割对象图象 130(步 220)。然后利用现在的技术计算或确定被直方图化的特征(步 230),最好在逐个象素的基础上(但也可以在每隔一个象素或任何预定的象素子集的基础上)执行这一步。然后利用现有的技术来形成(步 640)特征 $F1$ 的直方图 650。

例如, M 个区间(箱)的直方图数组首先被初始化为零。然后,在逐个象素的基本上,计算象素的色调。对应某个具体象素的这个被计算出来的色调值被量化,使之落入该直方图的 M 个箱子中的某一个,比如说 $Bin(x)$ 。然后对 $Bin(x)$ 的内容增 1,即新的 $Bin(x) = \text{原来的 } Bin(x) + 1$ 。对被分割的对象图象 130 中的所有象素或这些象素的经选择过的子集(如每隔一个象素)执行这一步骤。

色调直方图 650 表示色彩(色调)是如何分布在被分割的对象 130 的图象中的。换句话说,每个箱子的内容描述了在对象图象 130 中多少象素具有该箱子所表示的色彩。如果 $F1$ 是某些其他特征,则这表示该特征是如何被分布在对象 130 的图象中的。每个箱子的内容描述了对象 130 中多少象素具有该箱子所代表的特征值。

图 7 是一个流程图,表示规范化直方图特征的步骤,以及这些规范化的特征直方图如何不受被成象的对象 131 的尺寸或个数的影响的。

图象 320 是一个被分割的彩色图象,展示了一个被分割的对象

131 的图象 130。图象 720 是对象 131 的三个示例的被分割彩色图象,展示了三个被分割的对象图象 130。象图 6 中所描述的那样计算一个或多个特征 $F1$,并产生两个直方图,分别为直方图 745 和 740。在步 750 中,利用相同的规范化方法规范化每个直方图(745,740)。因为本发明只对每个图象(320,720)中对象的被分割图象规范化(步 750),因此,所得的规范化直方图(分别为 770 和 760)是一致的。即使具有较大数目的对象 131 的图象 720 将分布到直方图 740 的每个箱子上的象素个数比具有较少数目的对象 131 的图象 320 将分布到其相应的直方图 745 的每个箱子上的象素个数多,也会出现这种结果。(注意,如果对象 131 的尺寸在一个图象 720 中大于在另一个图象 320 中的尺寸,也会产生相同的效果)。例如,面积规范化建立了近似相等的规范化直方图(760,770),因为被分割图象对其直方图的分布由其相应的图象面积所划分。

通过增加直方图 745 的所有箱子的内容,计算彩色图象 320 中被分割对象图象面积 $130Area1$,用同样的方式计算图象 720 的所有对象 130(或较大尺寸的对象)的被分割面积 $Area2$ 。为了得到面积规范化直方图 760,直方图 745 被值 $Area1$ 划分(逐个箱子地)。再逐个箱子地由 $Area2$ 划分直方图 740 来计算面积规范化直方图 770。经过这种操作这后,面积规范化直方图 760 约等于面积规范化直方图 770,并且准备好象图 8 的描述中那样被比较(步 260)。

可以对任何能从被分割对象图象 130 中提出的特性执行规范

化。例如面积、长度和大小。还可以使用描述形状的其他量度,这样的量度包括(但不局限于)对象图象 130 的第二和较高阶(形状)的矩量、边界矩形的大小以及凸壳的面积。

图 8 说明算法 200 的步骤 260, 把被分割目标图象 130 的规范化特性(760, 770)同一个或多个被存储起来的规范化参考特征 270 进行比较(步 840)。特性 810 代表包含目标对象的某些被分割图象的规范化特性。对该特性的获取如图 7 所示。块 820 是如图 9 所述那样得到的面积规范化参考特性的一个表(数据库)。这些代表系统将能够识别的那些对象。这些规范化特性表示中的每一个一般都被标号为 831, ..., 837。只给出 6 个, 但面积规范化直方图表示的数目可以是很大的, 例如 100 个, 甚至 1000 个。每个被识别的对象至少应该由一个规范化的特性表示, 但也可以由多个规范化的特征表示。820 中的每个面积规范化特性都有一个从中产生规范化特性的对象的描述性标识符。特性 810 和参考特性 820 不限于一个特性, 可以用多个特性表示。在那种情况下, 从目标对象的图象中产生多个特性, 而多个特性表示每一个参考对象。另外, 每个这样的特性集都有一个唯一性的对象标识符。见图 10 的描述。

块 840 表示目标特性与参考特性的比较/匹配过程。匹配/比较特性的一种最佳方法是确定目标直方图和参考直方图之间的距离量度 $L1$ 。例如, 让目标直方图 810 被表示为编号的矢量 T , 而参考直方图 820 则表示为通过某个 RN 的矢量 $R1$ 。对于本公开, 目标直方

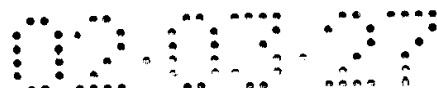
图 T 的最好匹配被定义为这样的参考直方图 RJ , 满足 T 和 $R1 \cdots R1 \cdots RN$ 之间的距离 $L1$ (有时称为曼哈坦距离) 是最小的。即, RJ 将给出距离

$$Dist(T - RJ), \dots, J=1, 2, \dots, N$$

的最小距离 $L1$ 。象最邻近分类法这样的匹配算法是众所周知的。展示距离量度的一般特性的任何距离量度(现有技术)都能用在这里。此外, 不能展示距离特性的其他量度(如直方图相交)也可以使用。目标直方图 T 和参考直方图 $R1, \dots, RN$ 的组成部分可以使用权值, 这就产生了部分方式加权的距离量度。

如果目标对象和参考对象是由多个直方图表示的, 最佳表示可以被看作包含多个被并置的直方图 T' 和 $R1' \cdots RN'$ 的较高维矢量。定义这样一个被并置的目标直方图到 T' 的最好匹配的一种最佳方法是被定义为满足 T' 和 $R1' \cdots RN'$ 之间的距离 $L1$ 是最小的并置参考直方图 RJ' 。这里, 不同的权值可以被赋给距离 $L1$ 中不同的子向量(代表不同的特征直方图)。另外, 可以使用任何距离量度, 并且还可以使用不展示距离特性的量度方法, 如直方图相交法。权值可以被加到目标直方图 T' 和参考直方图 $R1' \cdots RN'$ 的组成部分上, 导致了部分方式的加权距离量度方法。

在本发明中, 故意让所示的对象 131 都属于同一类型。若提供多种对象, 对象 A 和对象 B , 则可能导致预想不到的结果。最可能出现的结果是: 对象被标识为不可识别的。然而, 由于所用的距离量度



方法,也可能发生这样的识别结果:(1)对象为对象 A;(2)对象为对象 B;(3)对象为对象 A 或对象 B,在图 5 的用户界面中供选择。当混合对象具有类似的外形时,如 *Granny Smith* 苹果和 *Golden Delicious* 苹果,则会出现后一种情况。对象被识别为存在 820 中的某个其他对象 C 的这种情况是不可能出现的。

图 9 是一个流程图,表示方法 910 中通过把满足一定存储规则 255 的(被并置的)参考直方图加到存储器 270 中来训练系统的步骤。训练方法 910 使装置 100 能识别新的对象/项目,即那些最初没有被存在系统存储器 270 中的对象/项目。训练方法 910 首先提供具有对象/项目的图象 920 的装置。该图被分割(步 220),然后为如上所述的直方图化 640 确定特征。象上面那样,规范化(被并置)的直方图 750 被比较(步 260)。若目标规范化直方图匹配了存储器中的某个参考规范化直方图,则该目标图象被识别。否则,方法 910 根据一定的存储规则 255 继续检查目标规范化图象。若存储规则没有被满足,方法终止(940)。若该目标规范化图象满足存储规则 255,则把该目标规范化图象存在存储器设备 270 中,可用于以后匹配其他的目标图象。

采用以后将被用来识别所述对象 131 的不同示例的相同方式通过设备操作获取图象 920 是很重要的。图 4 中描述了这种设备的最佳实施例,使被极化的灯光通过过滤器 410 和摄象机上的极化过滤器 420。极化(偏振)具有特殊的重要性,因为在训练和识别期间,



对象 131 可能具有非常不可预期的镜式反射(闪耀)效应。

训练也可以通过交互的输入/输出设备 160 来实现;可以通过人类的干预来实现(通过前台的出纳员或后面屋子中的农产品管理员)。图 15 中将作进一步的描述。

图 10 是一个流程图,表示从图象中提取多个特征并利用多个特征来辨认对象的步骤。该方法首先从对象 130 和背景 311 的图象开始。象前面那样,把对象图象 130 从背景 311 中分割出来。然后按照如上所述的抽取色调 $F1(230)$ 同样的方法从被分割的图象 320 中提取多个特征。块 1010...1020 通常指的是其他被提取的多个特征。这些包括,但不限于,饱和度、亮度、质地(如图 11 中所述)、边界形状(图 12 中所示)和密度(图 13 中所示)。至于色彩,HSI 表示法是本发明表示色彩的最佳手段。其他的色彩表示法也可以使用,包括(但不局限于)RGB、Munsell、以及对立色彩。

一旦特征 $F1-FN$ 被提取,就如上所述那样被直方图化和规范化。虽然许多特征(如色彩)可以被面积规范化,但其他的规范化(如长度和边界)可能更适合于某个特征。例如,见下面图 12 中的形状直方图。

在步骤 840 中,每个抽取的特征($F1-FN$)被比较。图 8 中已经描述了这种比较。被提取的 N 个直方图(特征 $F1-FN$ 的直方图)基本上被并置为一个长的直方图,并且比较是基于目标并置直方图和参考并置直方图之间的某种距离量度的。在这种距离量度方法中,



单个特征值 $F1-FN$ 的直方图可以用不同的权值 $W1-WN$ 加权。另外,在距离量度中,并置直方图的每个单个的组成部分也可以有自己的权值。如上所述,特征 $F1-FN$ 包括,但不局限于,色调、饱和度、亮度、质地和形状。不是所有的这些特征都具有同等的识别能力,因此,可以使用权值 $W1-WN$ 。然而,不是单个特征(比如说 $F1$)的每个部分都具有相等的识别能力。因此,单个特征直方图也可以按不同组成部分的方式进行加权。

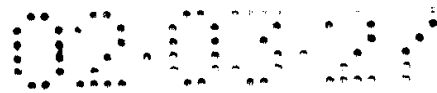
图 11 说明质地是如何被用作规范化识别特征的。块 210 是展示质地 1120 特征的对象/项目的对象图象。象前面一样,对象图象被从其背景中分割出来(步 220)。在步 1140 中,从被分割的图象中计算质地特征。在现有技术中已知的任何质地量度法都可以用在步 1140 的这种计算中。然而,有两种新颖的计算是最佳的。

第一种最佳的质地计算方法是一种新颖的质地量度 Λ :

利用现有技术中的方法选择阈值 Tb ,把被分割的图象转换为一个二进制图象。若图象的亮度大于 Tb ,二进制图象被置 1;若图象的亮度小于 Tb ,二进制图象被置 0。也可以使用为熟悉这一技术的人所了解的、用于二进制化图象的其他方法。结果为小斑点似的黑白图象。每个小斑点可以用宽度和长度特性化;与小斑点有关的质地量度(IVE)由

宽度

偏心率 = 宽度/长度



给出。这是使用“区域预测”确定的一种质地量度。

第二种最佳的新颖质地量度 B 如下所述：利用现有技术中的方法，用 $\begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$ 掩码对图象卷积，这种卷积是在垂直和水平方向上同时进行的，分别用 $Vconv$ 和 $Hconv$ 表示。在卷积结果超过某个阈值 $T2$ 的每个象素上，包括大小和方向的向量

$$\text{大小} = \text{Sqrt}(Vconv(x) * * 2 * Hconv(x) * * 2)$$

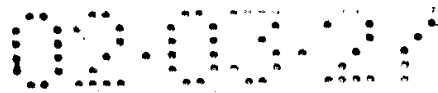
$$\text{方向} = \arctan(Vconv(x)/Hconv(x))$$

被定义为质地量度。这是一种用“边缘预测”确定的质地量度。

质地特征也可以用一种新颖的方法 1150 使之直方图化，即，只对被分割的对象图象 1120。如上所述的那样在被分割的图象上对质地量度直方图化。对质地量度 A ，这就产生了(宽度—偏心率)直方图，其中的宽度和偏心率已经在前面被定义。质地量度 B 给出了(大小—方向)直方图，其中的大小和方向如上所定义。对于方向直方图、方向直方图中最大方向被计算并且该直方图周期性地移动，将其峰值带到中心。这将使得方向直方图不受质地成象时转动的影响。

质地直方图通过计数规范化。其中计数可以是被分割对象图象 1120 中的每个象素，或者也可以是真正正展示质地的被分割对象图象 1120 中的那些象素。熟悉这一技术的人还可以设想从质地化区域中提取的其他形状量度。一种得到的规范化质地直方图被表示为 1170。使用质地直方图来识别对象被认为是一种新颖的方法。

图 12 是一个流程图，说明用形状作为识别特征的步骤。象前面



那样分割图象 210(步 220)。接着,确定对象图象 130 中哪些象素为边界象素(1210)。如果象素 P 的一个或多个相邻象素属于背景图象 311,则象素 P 为边界象素。接着,确定每个边界象素 P 的边界形状特性(1220)。本发明所用的一种最佳形状量度是象素 P 上的局部边界曲率。计算机 140 首先计算适合中心象素 P 和周围边界象素的个数的圆的半径。曲率 $1/R$ 描述象素 P 变化的局部程度—零曲率对应直线边界,高曲率对应局部“波形的”边界。苹果具有基本不变的曲率边界,而黄瓜则有一个低曲率、一个低的负曲率和两个高曲率(端点)。其他的形状量度是可能的。

接着对边界形状特征直方图化(1230)。这是通过计算边界象素 P 的形状特性来实现的。在这里,直方图是从包括对象图象 130 的边界的图象象素 P 的某个集合中、而不是在某个面积上产生的。

步 1235 中实现的规范化是一种形状直方图的长度规范化。一箱接一箱,1230 的直方图被边界象素 P 的总数划分。其结果是:一个对象图象 130 的长度规范化形状直方图等于多个对象图象 130 的长度规范化形状直方图。长度规范化对象图象边界形状直方图是本发明的一个新颖特征。其他关于对象图象边界长度的规范化是可能的。

密度可能是一个重要的识别特征。一磅白洋葱的重量和一磅蘑菇一样,但蘑菇的体积要比白洋葱的体积大得多。因此,重量和体积之间的关系是很重要的。这个关系是由下式决定的对象密度:

$$\text{密度} = \text{重量(对象 131)} / \text{体积(对象 131)}$$

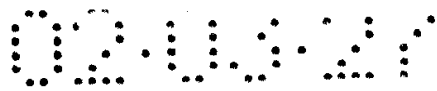


图 3 是一个示意图,表示计算机 140 连接确定对象 131 重量的称重设备 170。为了把重量用作识别特征,设备 170 把对象 131 的重量报告给计算设备 140。在最佳实施例中,系统 100 用重量把“密度”特征定义为:

“密度”=重量(对象 130)/面积(被分割对象 131)

这种量度并没有体现引用密度的常规方法,而是压力的量度。称为“平均投影密度”。

对象 131 的真正密度只能很粗略地计算。为了得到对象 131 的体积的概念,对象 130 的边界轮廓近似椭圆,而其体积则近似于从该近似的椭圆中旋转产生的椭圆球的体积。然后,由重量/体积给出密度。

从投影分割对象图象 130 中估算体积的其他手段也在本发明的范围中。

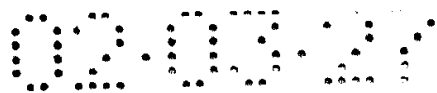
代表同一特征 $F1$ (如色调)的多个参考直方图也能用来识别一个给定的对象。图 14 表示图象 1405,其中,分割对象图象 130 有两个不同的区域,即叶子 1410 和葡萄 1420。图象 1405 包括对象 130 和背景 311。对象图象 130 和它的第一对象区域 1410 及其第二对象区域 1420 一块被分割。利用某个分割算法来识别和定义这些对象区域。一种最佳的算法是使用面积规范化色调直方图来检测是否存在两个或多个不同的峰值。

这些区域分别被直方图化和面积规范化。面积规范化直方图

1450 和 1455 分别对应第一区域 1410 和第二区域 1420, 并且如上所述比较参考直方图。此外, 在匹配期间(图 8), 区域 1410 和 1420 的相对位置 1430 可以被考虑。本发明的这一部分说明这样的项目, 其中的一种特征(如色彩, 但不局限于色彩)在对象 131 的表面上不是一致的, 因此, 在被分割的对象图象 130 上不一致。典型的例子是带有叶子部分的胡萝卜。

图 15 给出了一个可选的人类界面 160。它包括显示由装置 100 确定的对象 131 的各种可能的个性的图象(否则被解释)描述 1510、1520、1530 和 1540 的一个最佳装置。在一般情况下, 通过和参考数据库 270 的比较 260 能唯一地辨认对象 131。然而, 在某种情况下, 可能匹配多个的参考直方图, 即, 目标对象规范化直方图可能近似相同于多个的参考直方图。在这些情况下, 可以通过界面 160 让人来作出最终的识别决定。界面 160 的一个最佳实施例提供了 4 种或更少的选择—1510, 1520, 1530, 1540。象上面所解释的那样, 还可以有更多的选择。人类可以通过各种方式, 如触摸、声音、鼠标、键盘把决定告诉计算机 140。此外, 在界面上可以提供—个装置(按钮)162, 使用户能决定何时并是否把一个直方图加到存储器 270 中的参考对象数据库 820 上, 即, 系统是否用该数据来训练, 使之能在某个将来的时候, 当对象 131 的示例被提供给系统 100 时能识别(或更好地识别)。

图 16 给出了具有类似浏览键功能的界面 160, 使用户能浏览对



象的个性。浏览键指的是一个关键字或键特征,由此能缩小人类对数据库 820 中的对象个性的指导性搜索。这种键的例子如(但不局限于)红色 1612、绿色 1613、黄色 1614,褐色 1615,圆形 1616,直线 1617,叶子 1618,苹果 1619,柑桔类水果 1620,胡椒 1621,以及土豆 1622,都象 1610 中那样被显示。用户可以通过触摸、声音、鼠标、键盘等进行通讯。键 1600 将作出这样的响应:或者用 1610 的另一个示例,其中所提供的选择 1612—1622 更为具体,或者用屏幕 1630 作出最后的决定。若 1619,如苹果,被选择,1600 将为人类提供屏幕 1630,给出对象 131 的个性描述(草图,图片,语句)1631—1641。用户可以利用各种已知的输入设备在屏幕上进行选择。也可以使用其他对人类友好的方法或装置。

图 17 是一个示意图,给出了系统 100 用来对对象定价的可选装置。称重设备 170 被用来确定对象 131 的重量。装置 100 如上所述那样识别该对象。一旦该对象被识别出来,就确定该对象的价格。在定价过程中,如果需要的话,可以使用对象的重量 170 和/或个数(现有项目的个数)。对象的价格被存在存储器 144 中。

定价设备 1710 连接装置 100,以便把价格告诉用户。定价设备 1710 可以是打印设备、显示设备或其他任何能传送对象价格的装置。价格也可以被显示在交互的输出设备 160 上。

如果价格以磅为单位标识、计算机 140 计算价格按:

价格=重量*(对象 131 每磅的价格)



如果以个数为单位,计算机按下式计算价格:

价格=个数*(对象 131 的单价)

项目的个数可通过人类的干预或估计得到。

为了由人类干预而输入个数,如果项目 131 在计算机存储器 140 中被指定为按个数出售(如棕檬,酸橙),则系统 100 只需提示用户输入个数。有两种自动估算个数的方法:

a)装置 100 在存储器 144 中存有对象 131 的平均重量,则按下式计数:

个数=重量/平均重量(对象 131)

当从被分割对象图象 13 中建立了对象 131 的个性之后。

b)若装置 100 已经估算了被分割对象图象 130 的个数,则

个数=被分割对象图象 130 的个数。

图 1

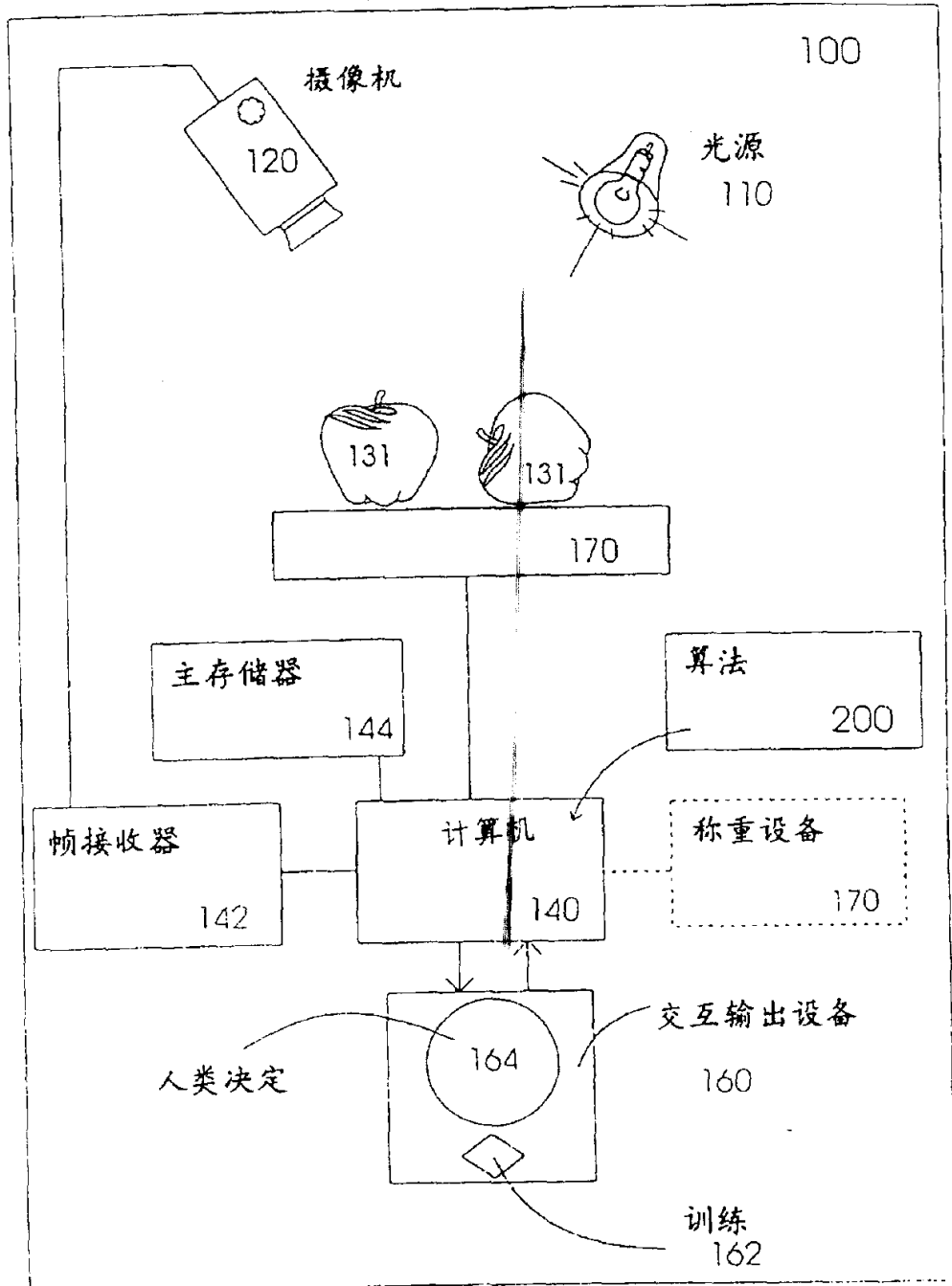


图 2

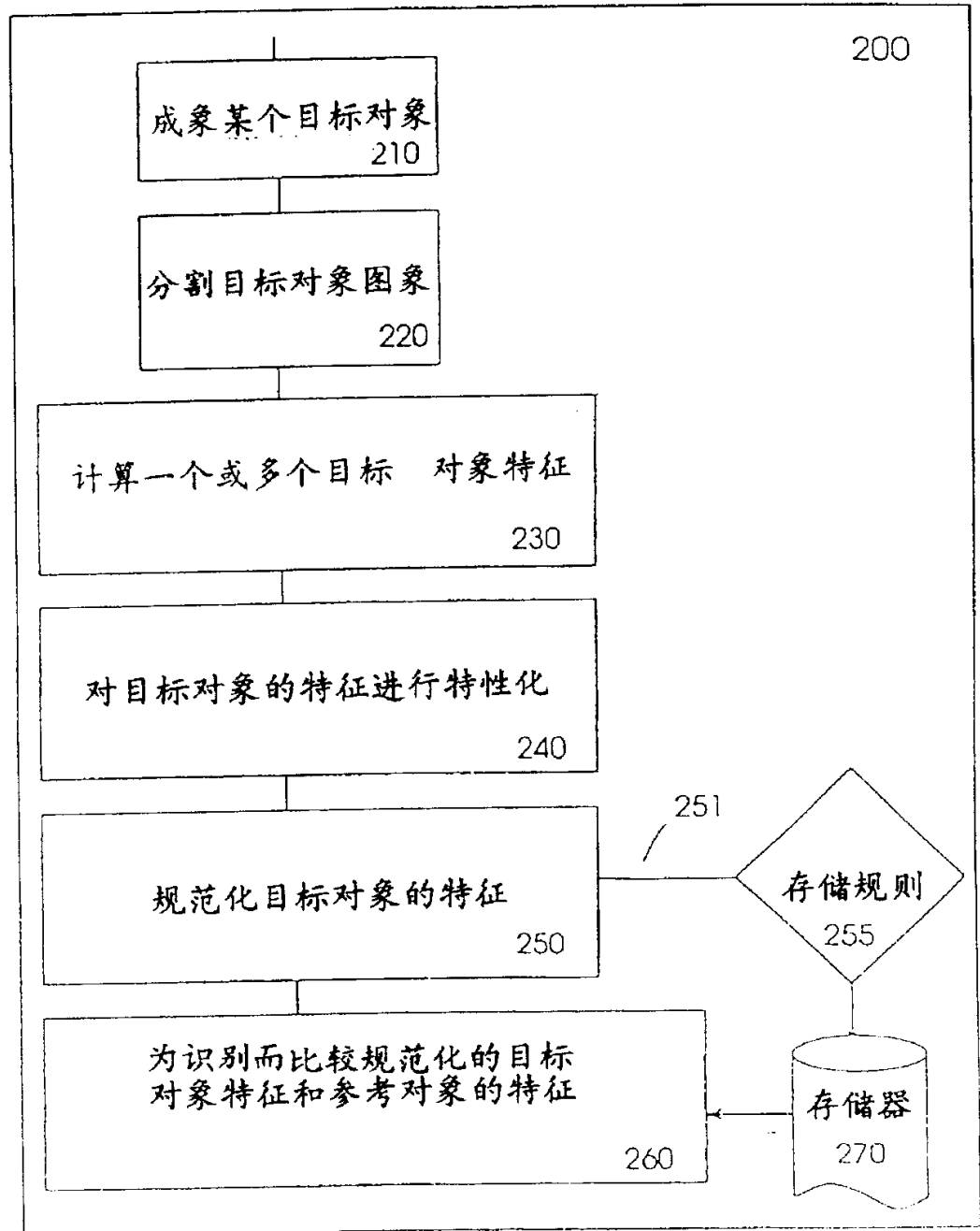


图 3a

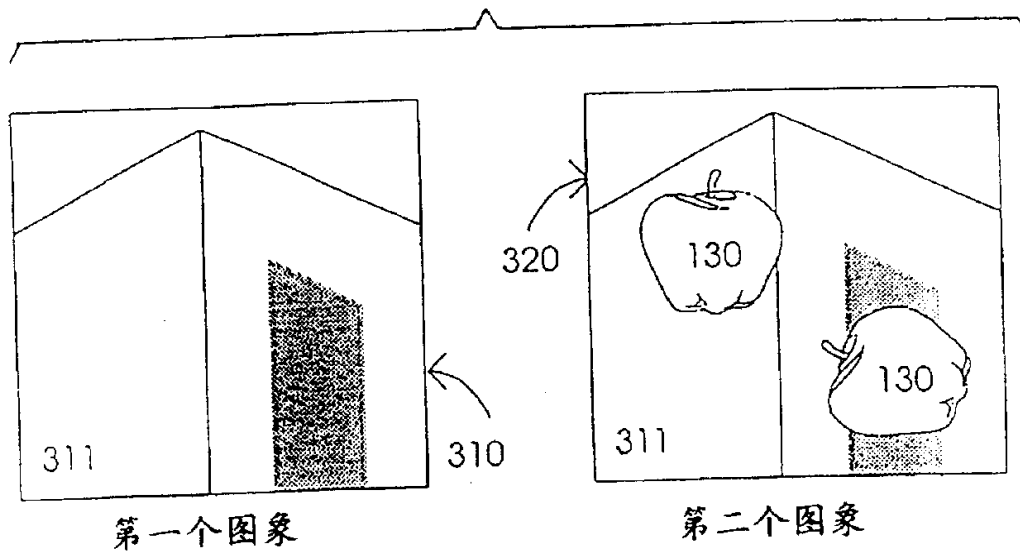


图 3b

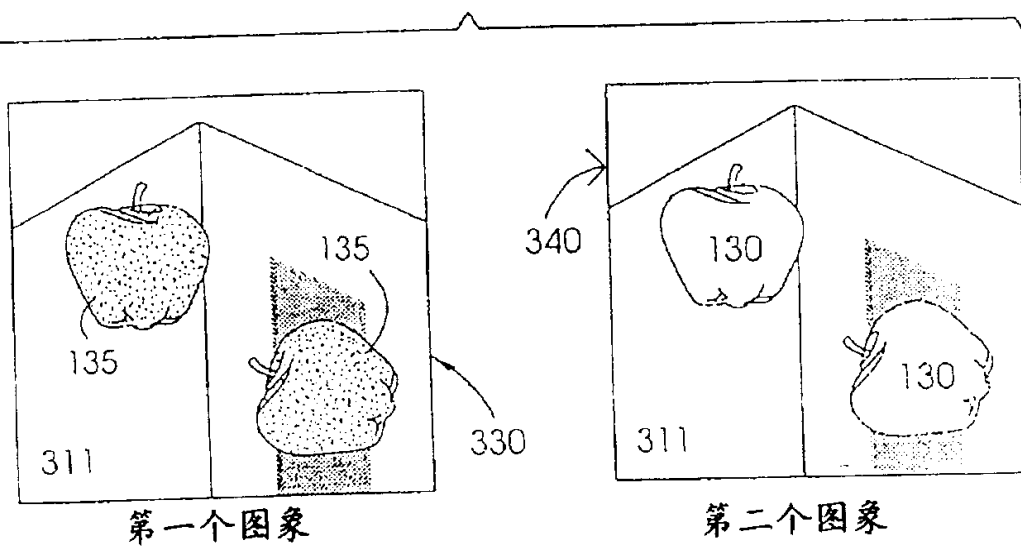


图 4

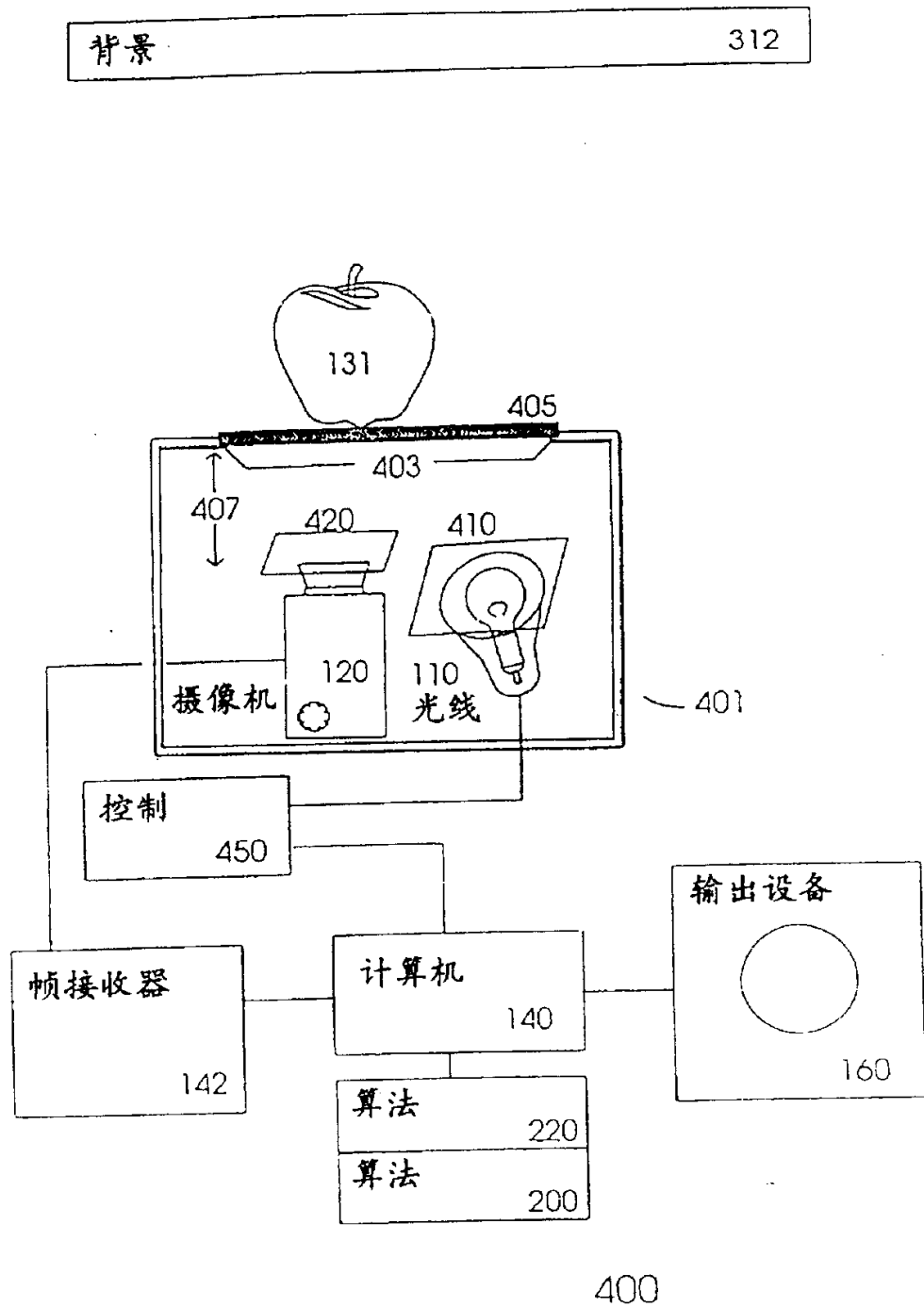
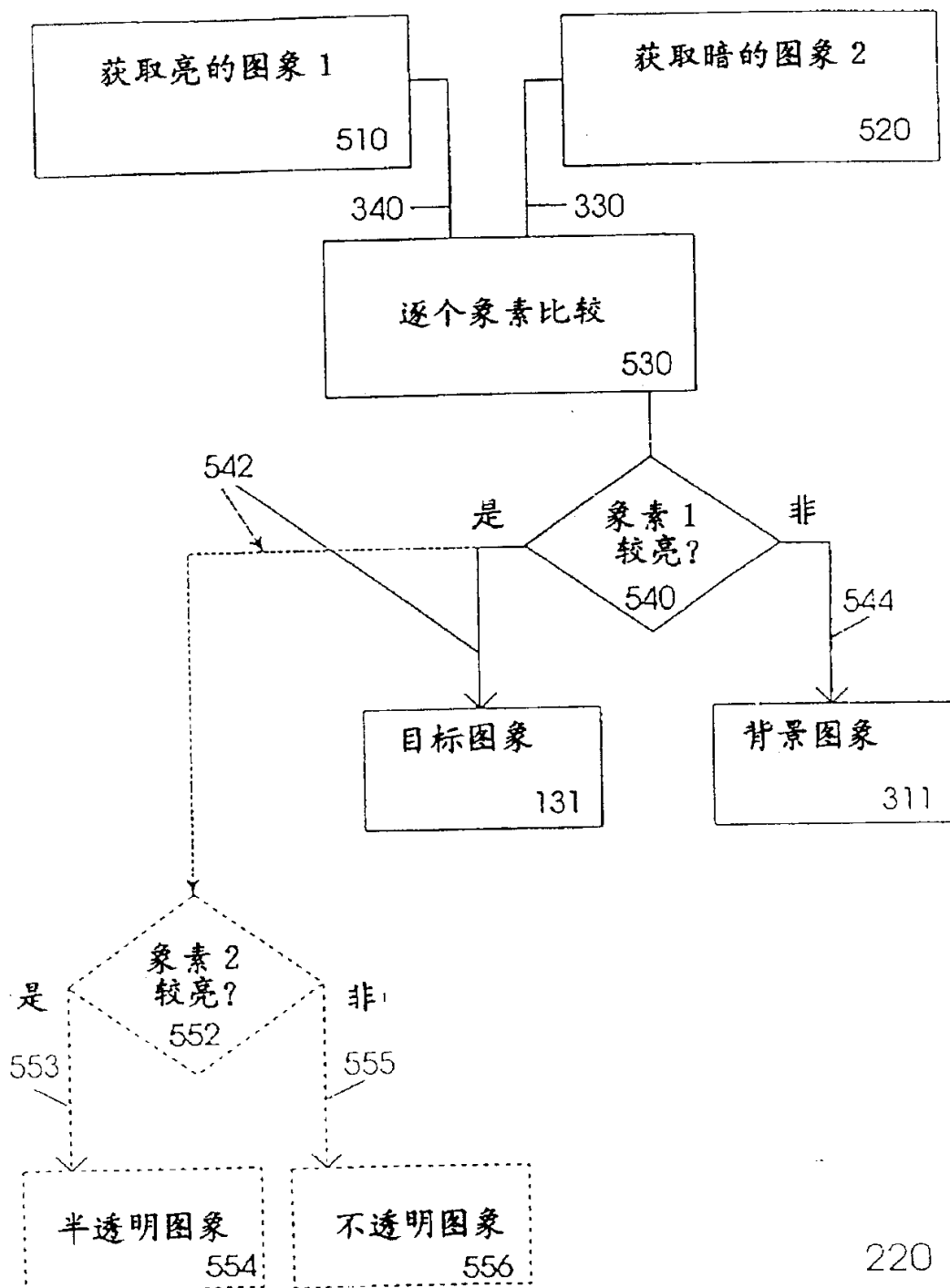


图 5



220

图 6

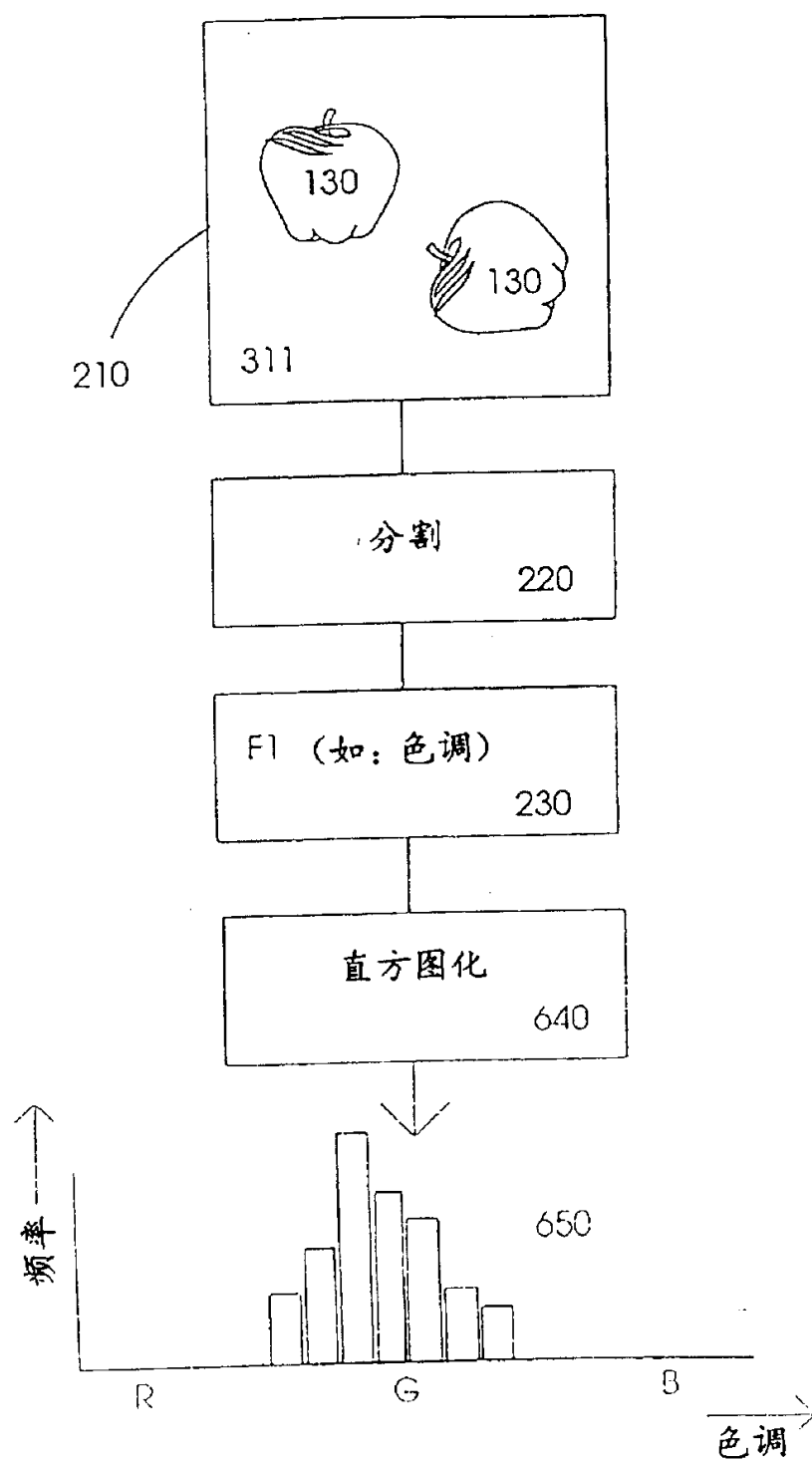


图 7

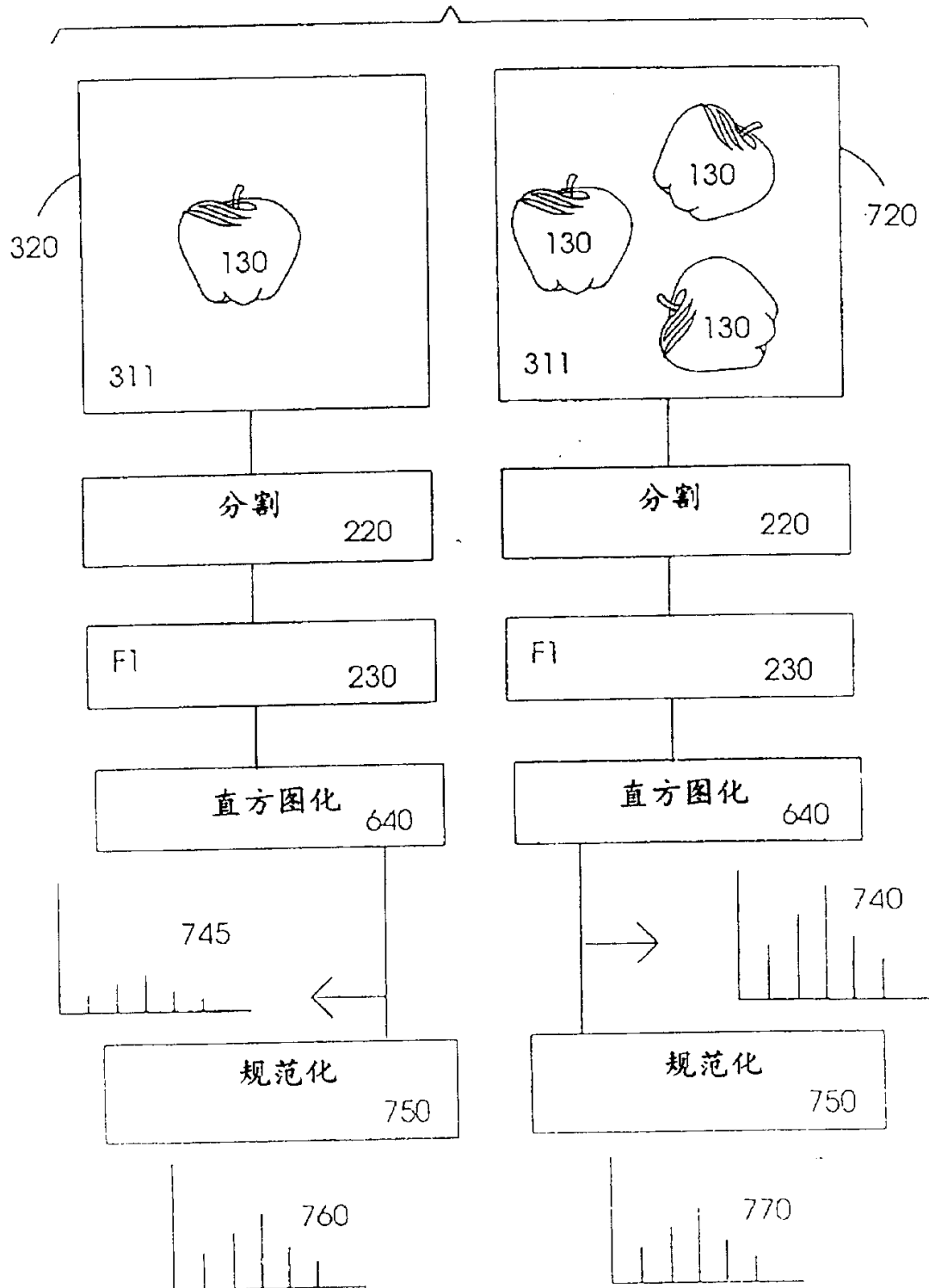


图 8

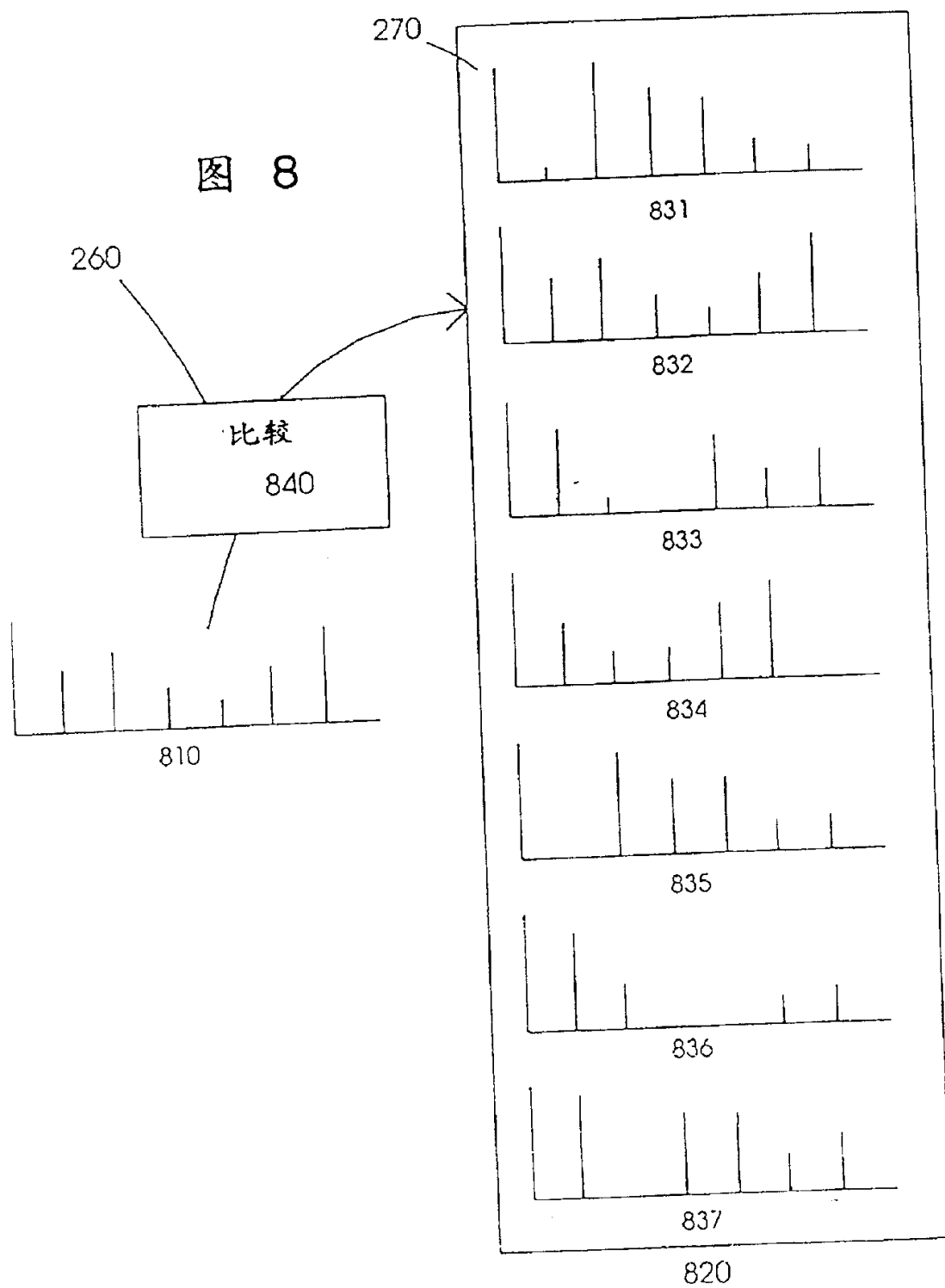


图 9

训练 910

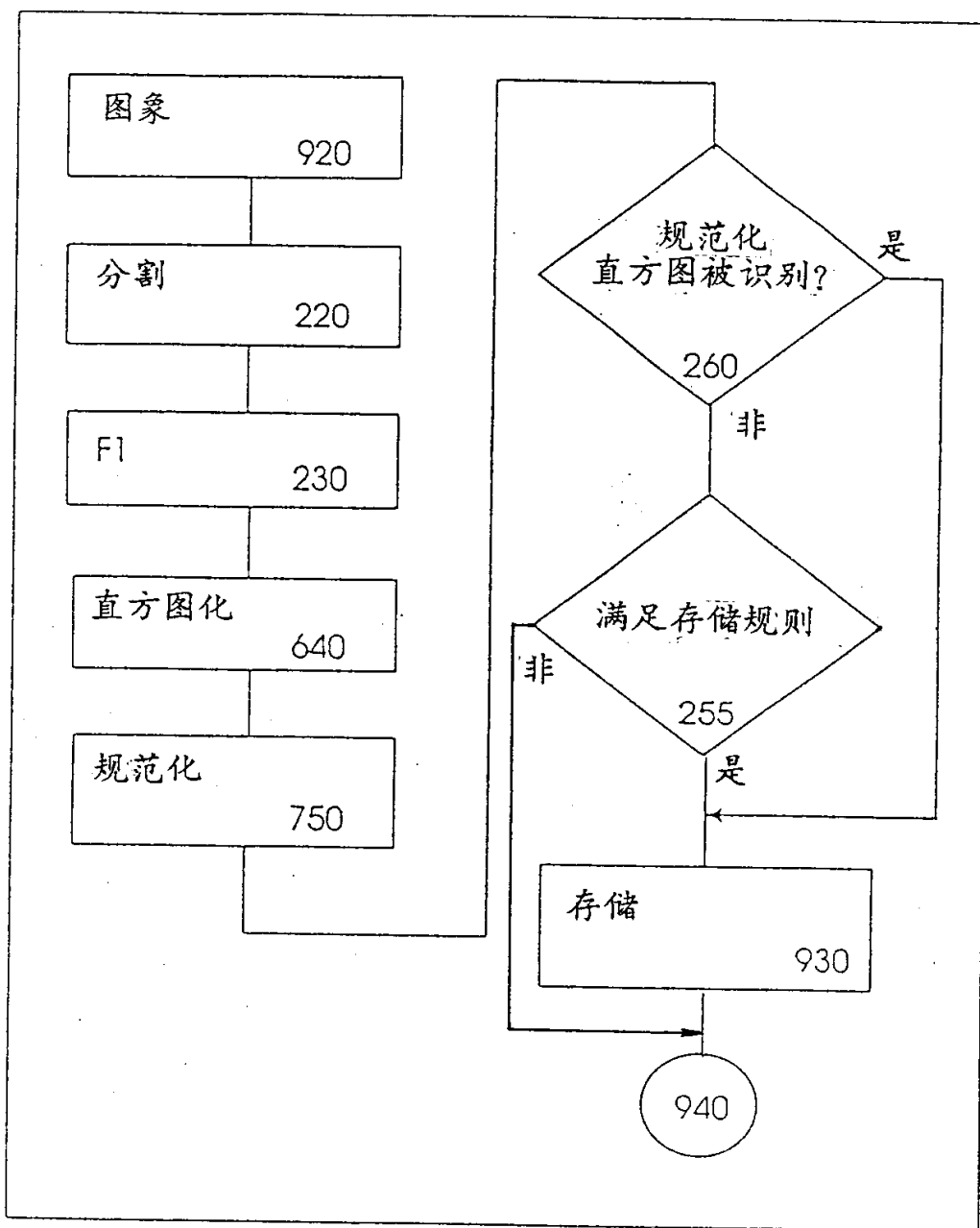


图 10

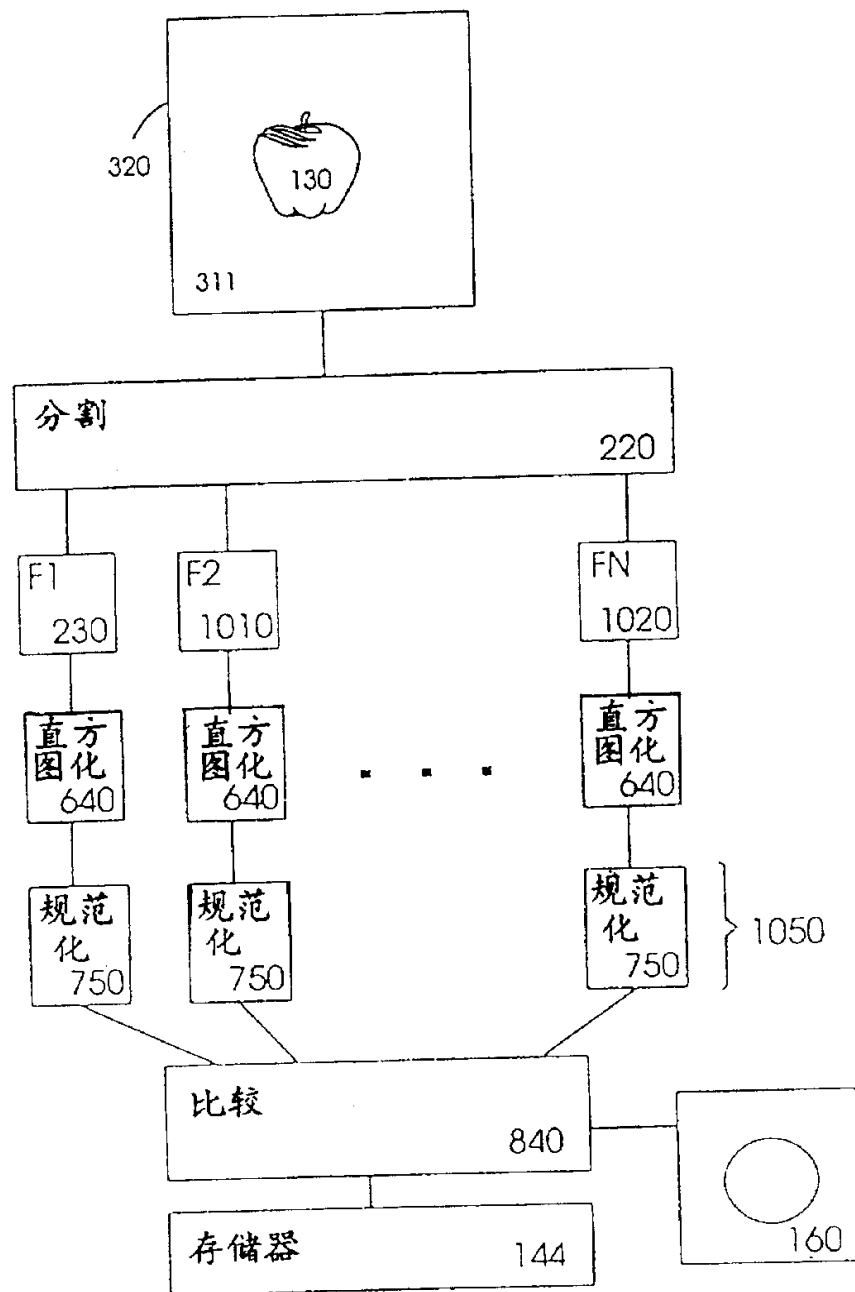


图 11

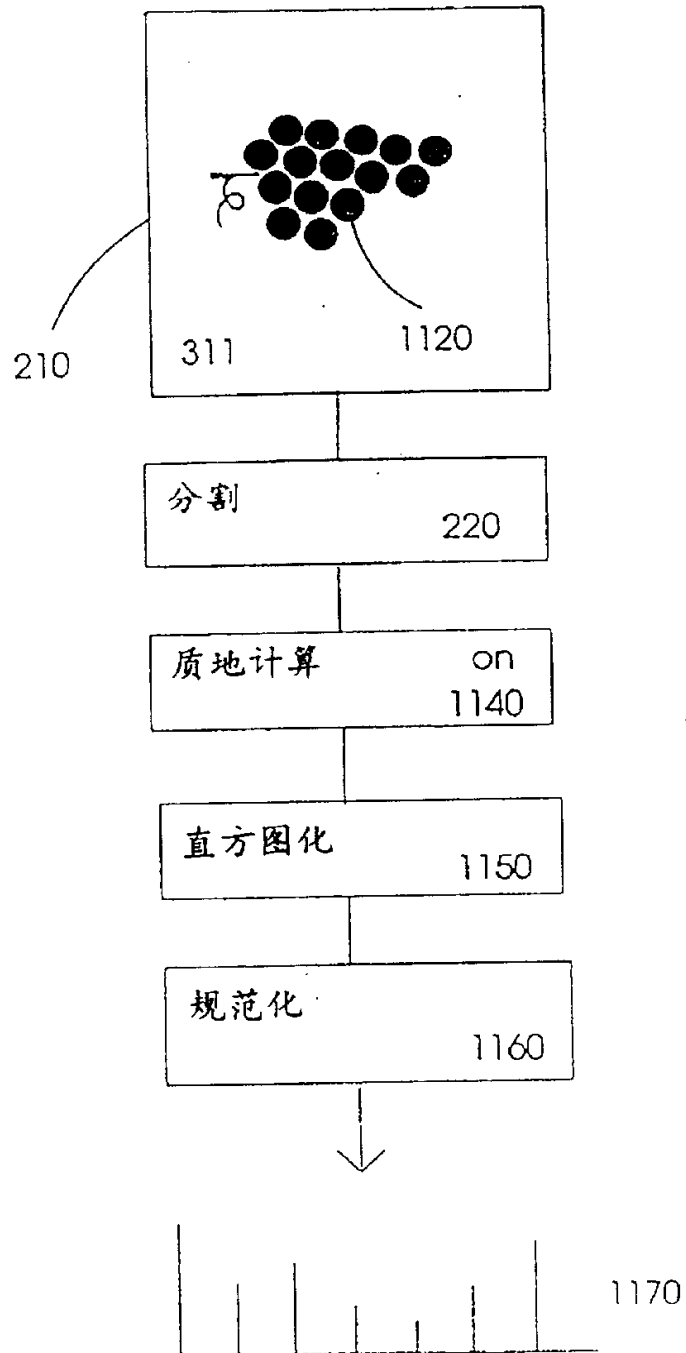


图 12

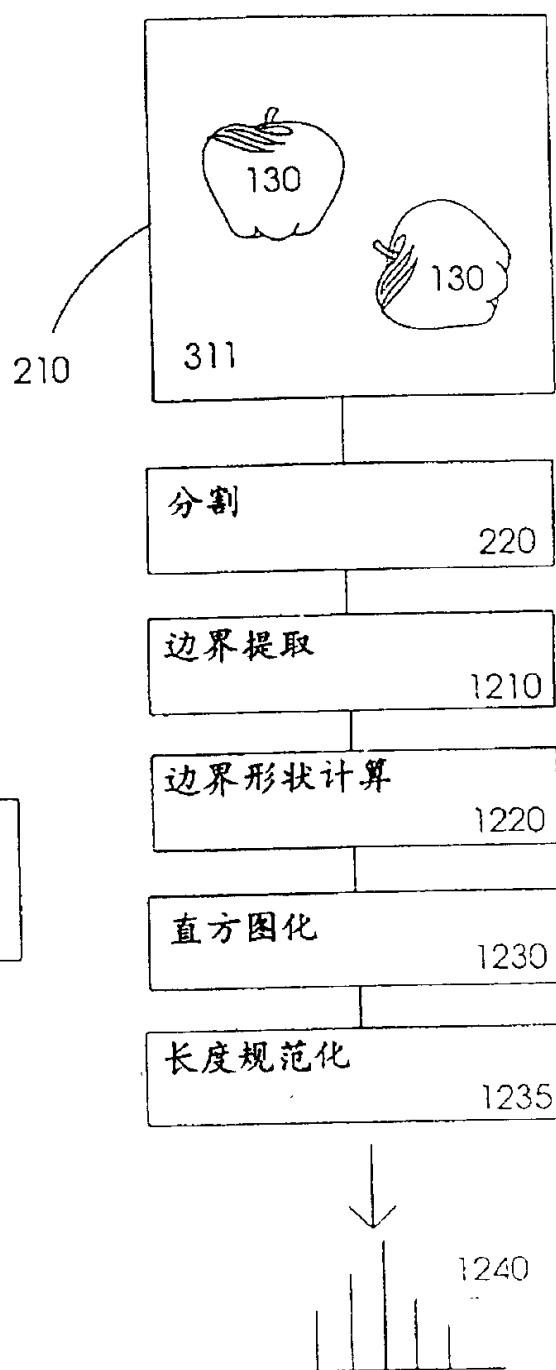
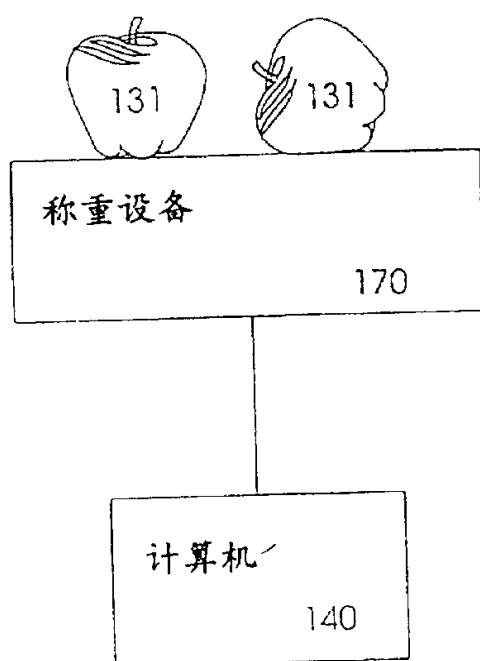


图 13



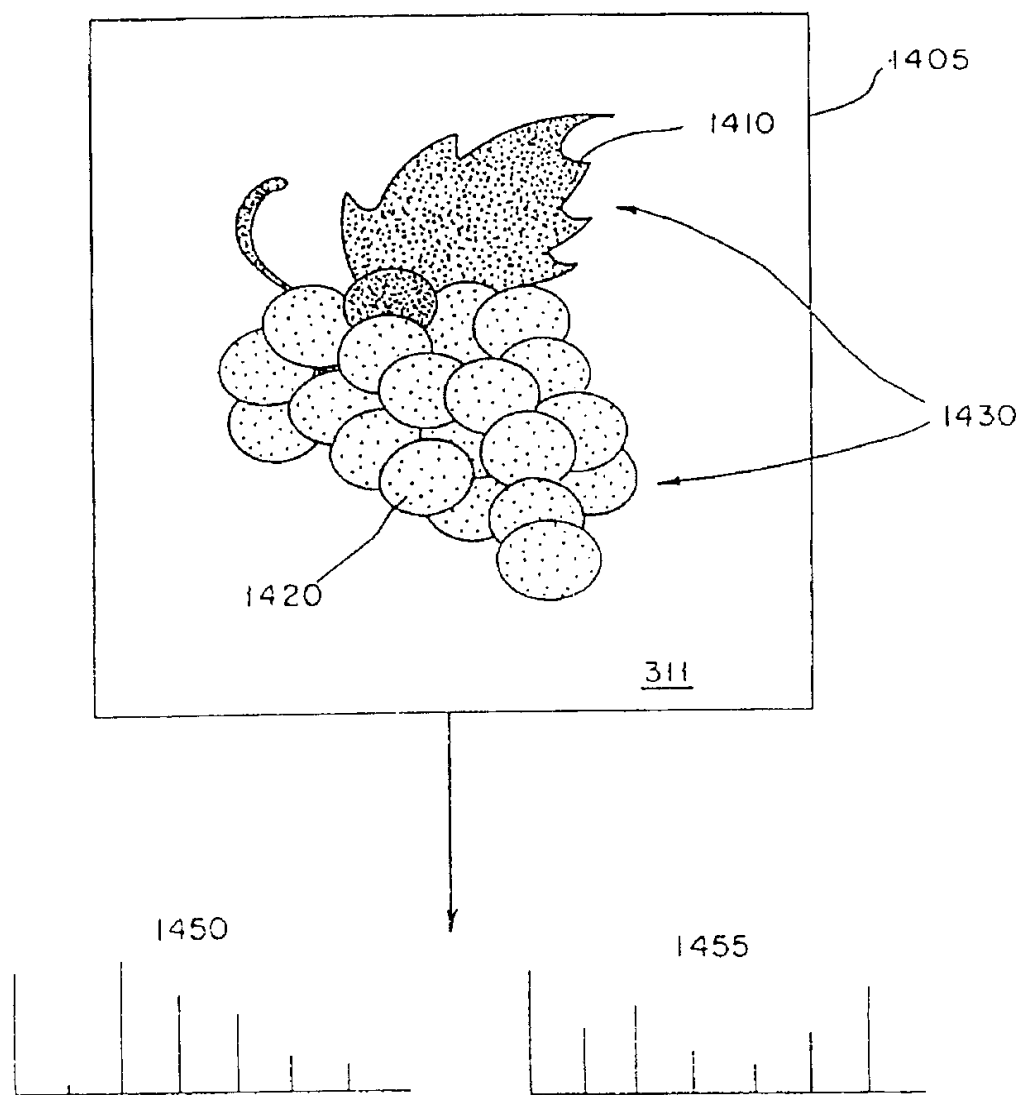
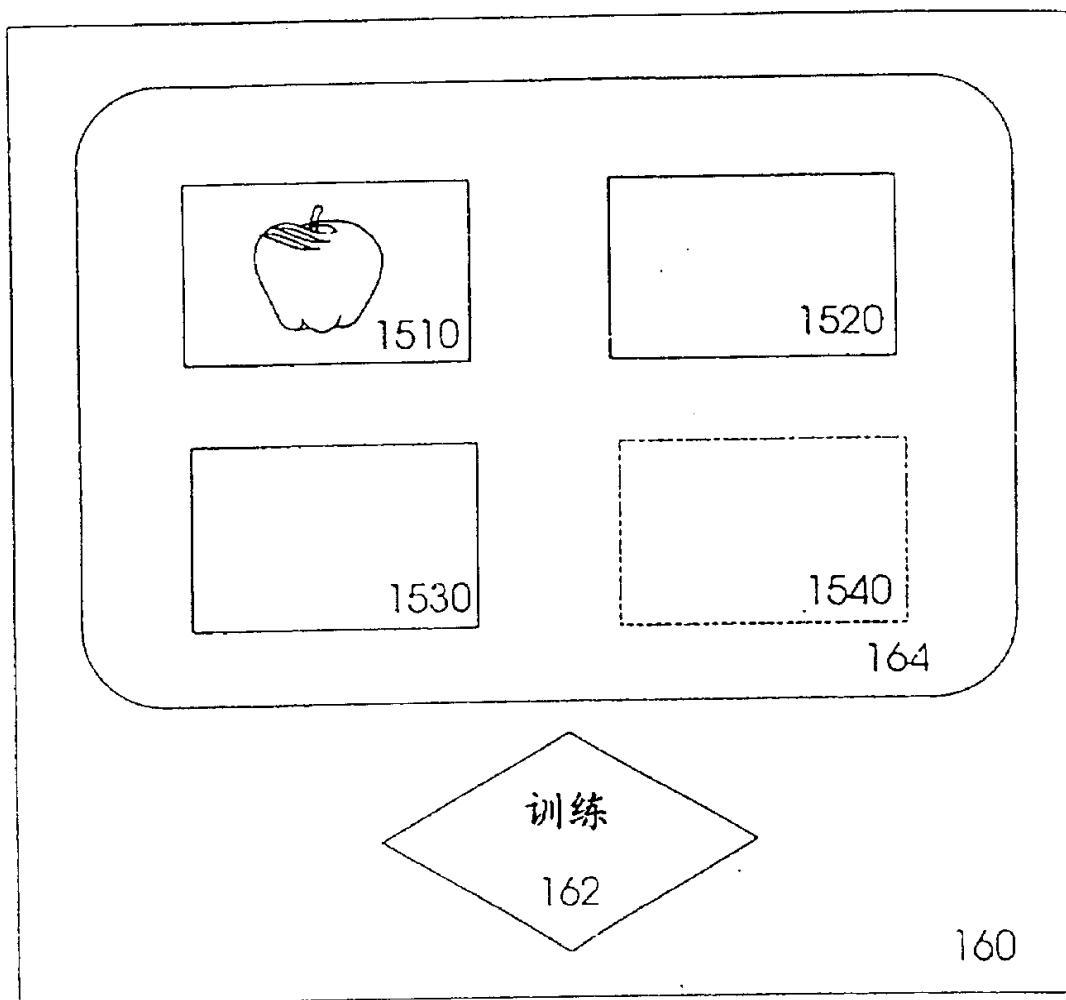


图 14

图 15



1600

图 16

红 1612	绿 1613	黄 1614	褐 1615
圆 1616	直 1617	叶 1618	苹果 1619
柑桔类 水果 1620	胡椒 1621	土豆 1622	1610

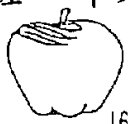
红 DEL 苹果  1631	GALA 苹果 1632	1633	1634
1635	1636	1637	1638
1639	1640	1641	1630

图 17

