



[12] 发明专利申请公开说明书

[21] 申请号 02141904.3

[43] 公开日 2003 年 2 月 12 日

[11] 公开号 CN 1396775A

[22] 申请日 2002.7.7 [21] 申请号 02141904.3

[30] 优先权

[32] 2001.7.7 [33] GB [31] 0116681.8

[71] 申请人 夏普公司

地址 日本大阪市

[72] 发明人 洪启和 野村敏男

[74] 专利代理机构 中国专利代理(香港)有限公司

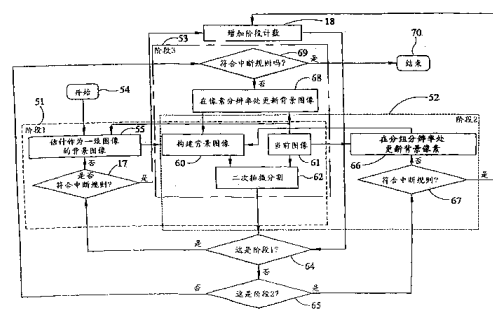
代理人 罗 朋

权利要求书 6 页 说明书 29 页 附图 14 页

[54] 发明名称 分割像素图像的方法和装置

[57] 摘要

为了分割具有前景和背景的像素图像,确定背景一致的视觉特征,例如颜色,并由此产生一非封闭的背景图像(55)。将所产生的背景图像 60 和要分割的图像 61 进行比较,例如使用二次拍摄分割技术 62,且将图像像素指定为前景或背景像素。



1、一种分割像素图像的方法，该图像具有至少一个前景区和至少一个非封闭的背景区，该方法包括：

- 5 (a) 确定至少一个背景区的第一视觉特征；
 (b) 从第一视觉特征中产生第一非封闭的背景图像；
 (c) 对像素图像中的至少一些像素的每个像素同第一背景图像中的相应像素进行比较；和
 (d) 至少部分依赖于步骤 (c) 的结果，将至少一些像素中的每个像素指
10 定为前景或背景像素。
- 2、如权利要求 1 中的方法，其特征在于第一视觉特征基本上是一一的视觉特征。
- 3、如权利要求 2 中的方法，其特征在于第一视觉特征基本上是一一颜色。
- 4、如权利要求 3 中的方法，其特征在于像素图像中的每个像素由多个彩色
15 分量表示；步骤 (a) 包括为每个颜色分量值形成至少部分像素图像的每个颜色分量的像素数的直方图并检测每个直方图的颜色分量值，在该直方图上存在最大的像素数；和步骤 (b) 包括把颜色分量值归于第一背景图像的每个像素。
- 5、如权利要求 4 中的方法，其特征在于将至少部分像素图像进行手工选择并包括至少一个背景区的至少一部分。
- 20 6、如权利要求 4 中的方法，其特征在于像素图像是风景格式且至少部分像素图像包括侧面部分。
- 7、如权利要求 4 中的方法，其特征在于像素图像是肖像格式且至少部分像素图像包括顶面部分。
- 8、如权利要求 4 中的方法，其特征在于像素图像是肖像格式且至少部分像
25 素图像包括顶角部分。
- 9、如权利要求 1 中的方法，其特征在于步骤 (b) 包括将第一视觉特征归于第一背景图像的每个像素。
- 10、如权利要求 1 中的方法，其特征在于包括至少用步骤 (a) 重复步骤 (a) 到(d)一次，步骤 (a) 包括在前述步骤(d)中指定给背景像素的至少一些像素中
30 确定第一视觉特征。

11、如权利要求1中的方法，其特征在于包括下面的步骤：

(e) 产生第二非封闭的背景图像，该图像包括多个像素块，每个像素块具有基本上一致的第二视觉特征；

(f) 将像素图像中至少部分像素的每个像素与第二背景图像的相应像素
5 进行比较；和

(g) 至少部分依赖于步骤(f)的结果，将至少一些像素中的每个像素指定为前景或背景像素。

12、如权利要求11中的方法，其特征在于每个第二视觉特征基本上是同一颜色。

10 13、如权利要求11中的方法，其特征在于通过像素图像的每个像素的相应视觉特征至少部分地确定每个第二视觉特征，该像素相当于块中的像素并将其指定为背景像素。

14、如权利要求11中的方法，其特征在于第二视觉特征至少部分地由第一视觉特征来决定。

15 15、如权利要求9中的方法，其特征在于每个块的第二视觉特征包括第一视觉特征和相应于块中像素并被指定为背景像素的像素图像的像素的相应视觉特征的线性组合。

16、如权利要求14中的方法，其特征在于包括至少用每个第二视觉特征重复步骤(e)到(g)一次，该第二视觉特征至少部分地由相应于块中像素并在前述步骤(g)中指定给背景像素的像素图像的像素的相应视觉特征确定。
20

17、如权利要求11中的方法，其特征在于每个块包括预定数量的像素。

18、如权利要求17中的方法，其特征在于每个块包括像素的方阵列。

19、如权利要求17中的方法，其特征在于包括至少一个循环，该循环或每个循环包括在每个块中用减少数量的像素执行步骤(e)到(g)至少一次。

25 20、如权利要求1中的方法，其特征在于包括下面的步骤：

(h) 产生第三非封闭的背景图像，其中相应于像素图像的所指定背景像素的每个像素具有第三视觉特征，该第三视觉特征至少部分地由所指定的背景像素的相应视觉特征确定；

(i) 将像素图像的至少部分像素的每个像素与第三背景图像的相应像素
30 进行比较；和

(j) 至少部分依赖于步骤(f)的结果, 将至少一些像素中的每个像素指定为前景或背景像素。

21、如权利要求 20 中的方法, 其特征在于第三视觉特征是颜色。

22、如权利要求 20 中的方法, 其特征在于相应于所指定的背景像素的第三背景像素的每个像素的第三视觉特征包括所指定的背景像素的相应视觉特征和第一或第二视觉特征的线性组合。

23、如权利要求 20 中的方法, 其特征在于在步骤(h)中, 相应于像素图像的所指定的前景像素的第三背景图像的每个像素具有第一或第二视觉特征。

24、如权利要求 20 中的方法, 其特征在于包括用在前述步骤(j)中确定
10 的步骤(h)的背景像素重复步骤(h)到(j)至少一次。

25、如权利要求 1 中的方法, 其特征在于步骤(c)和(d)包括步骤:

(k) 选择至少一个背景区中的至少一部分;

(l) 从至少一部分像素值中获得第一阈值以便使第一预定比例的像素的值比第一阈值要低;

15 (m) 形成一差值图像作为在像素图像中的每个像素和第一背景图像的相应像素之间的差值; 和

(n) 如果差值图像像素的值低于第一阈值且大于第一预定数量的相邻差值图像像素的值低于第一阈值, 则指定每个差值图像像素为背景像素。

26、如权利要求 11 中的方法, 其特征在于步骤(f)和(g)包括步骤:

20 (k) 选择至少一个背景中的至少一部分;

(l) 从至少一部分像素值中获得第一阈值以便使第一预定比例的像素的值比第一阈值要低;

(m) 形成一差值图像作为在像素图像中的每个像素和第二背景图像的相应像素之间的差值; 和

25 (n) 如果差值图像像素的值低于第一阈值且大于第一预定数量的相邻差值图像像素的值低于第一阈值, 则指定每个差值图像像素为背景像素。

27、如权利要求 20 中的方法, 其特征在于步骤(i)和(j)包括步骤:

(k) 选择至少一个背景中的至少一部分;

30 (l) 从至少一部分像素值中获得第一阈值以便使第一预定比例的像素的例比第一阈值要低;

(m) 形成一差值图像作为在像素图像中的每个像素和第三背景图像的相应像素之间的差值; 和

(n) 如果差值图像像素的值低于第一阈值且大于第一预定数量的相邻差值图像像素的值低于第一阈值, 则指定每个差值图像像素为背景像素。

5 28、如权利要求 20 中的方法, 其特征在于第一预定比例在 0.5-1 之间。

29、如权利要求 28 中的方法, 其特征在于第一预定比例基本上等于 0.75。

30、如权利要求 28 中的方法, 其特征在于第一预定数量基本上等于相邻差值图像像素数量的一半。

31、如权利要求 25 中的方法, 其特征在于用在前述步骤 (n) 中包括所指定的背景像素的至少一部分步骤 (k) 重复步骤 (k) 到 (n) 至少一次。

32、如权利要求 31 中的方法, 其特征在于步骤 (n) 包括形成指定为背景像素的差值图像值的第一初始直方图, 且步骤 (l) 从第一合成直方图中得到第一阈值, 该合成直方图包括在前述步骤 (n) 中形成的第一初始直方图和在前述步骤 (l) 中的小于 1 的第一预定比值的第一合成直方图的总和。

15 33、如权利要求 32 中的方法, 其特征在于第一预定比值是 1/2。

34、如权利要求 25 中的方法, 其特征在于包括步骤:

(o) 选择至少一个前景区中的至少一部分;

(p) 从至少一部分像素值中获得第二阈值以便使第二预定比例的像素的值超过第二阈值;

20 (q) 指定每个差值图像像素, 该像素未被指定为背景像素, 如果差值图像像素的值超过第二阈值且大于第二预定数量的相邻差值图像像素的值高于第二阈值, 则将其指定为前景像素。

35、如权利要求 34 中的方法, 其特征在于第二预定比例在 0.5-1 之间。

36、如权利要求 35 中的方法, 其特征在于第二预定比例基本上等于 0.75。

25 37、如权利要求 35 中的方法, 其特征在于第二预定数量基本上等于相邻差值图像像素数量的一半。

38、如权利要求 34 中的方法, 其特征在于用包括在前述步骤 (q) 中所指定的前景像素的至少一部分步骤 (o) 重复步骤 (o) 到 (q) 至少一次。

39、如权利要求 38 中的方法, 其特征在于步骤 (q) 包括形成被指定为背景像素的差值图像值的第二初始直方图, 且步骤 (p) 从第二合成直方图中得

到第二阈值，该合成直方图包括在前述步骤（q）中形成的第二初始直方图和在
在前述步骤（p）中小于1的第二预定比值的第二合成直方图的总和。

40、如权利要求39中的方法，其特征在于第二预定比值是一半。

41、如权利要求34中的方法，其特征在于包括指定每个差值图像像素，该
5 像素没有指定为前景或背景像素，如果差值像素的值小于第三阈值则指定为候
选背景像素。

42、如权利要求41中的方法，其特征在于第三阈值在第一和第二阈值之间。

43、如权利要求42中的方法，其特征在于第三阈值是第一和第二阈值的算
术平均。

10 44、如权利要求41中的方法，其特征在于包括指定每个差值图像像素，该
像素没有指定为前景或背景像素，如果将大于相邻像素的第三预定数量指定为
背景像素或候选像素，则将其指定为背景像素。

45、如权利要求44中的方法，其特征在于第三预定数量基本上等于相邻差
值图像像素数量的一半。

15 46、如权利要求44中的方法，其特征在于将每个没有指定为前景或背景像
素的差值图像像素指定为前景像素。

47、如权利要求25中的方法，其特征在于包括形成二进制掩码，其元素相
应于差值图像像素，如果将相应的差值图像像素指定为背景像素，则每个元素
具有第一值，且将如果相应的差值图像像素指定为前景像素，则每个元素具有
20 不同于第一值的第二值。

48、一种用于分割像素图像的装置，该图像具有至少部分封闭背景区的一个
前景区，和至少一个非封闭的背景区，其特征在于包括确定至少一个背景的
视觉特征的装置，从视觉特征中产生非封闭的背景图像的装置，比较像素图
像的至少一些像素的每个像素与背景图像的相应像素的装置，和至少部分取决
25 于比较装置来指定至少一些像素的每个像素为前景像素或背景像素的装置。

49、一种指定像素图像的装置，其特征在于包括可编程数据处理器和控制
数据处理器来执行权利要求1所述方法的程序。

50、一种存储介质，其特征在于包括用于控制数据处理器来执行权利要求
1所述方法的程序。

30 51、一种程序，用于控制数据处理器来执行权利要求1所述的方法。

52、一种图像捕获装置，其特征在于包括权利要求 48 中的装置。

分割像素图像的方法和装置

5

技术领域

本发明涉及一种将像素图像分割为至少一前景区和至少一背景区的方法和装置。为了简化信息率和/或提高前景区的压缩质量,这种技术可以用在视频压缩领域中。这种技术还可以用在合成新图像顺序,通过用另一个背景图像或另一个背景的顺序来重复分割的背景。此外,可能的应用包括视频通讯,视频会议,电视播送,因特网多媒体应用, MPEG-4 应用例如图像压缩,人脸检测应用和实时视频跟踪系统例如观测员跟踪具有立体效果画面的 3D 显示。

10

发明背景

15

这种技术的一个具体应用是在数字摄影机和其它在多媒体应用中的数字图像捕获和记录装置。这种装置的一个例子是夏普因®特网 ViewCam VN-EZ 系列。

许多已有的图像处理和分析应用包括具有前景目标的图像顺序,通常是临时被激活,且背景是相对静止的。当前景目标移动和/或改变形状时,部分背景图幅可以被覆盖和/或未被覆盖。对这些应用来说,具有将图像分割成前景区和背景区的能力是很有用的。

20

夏普公司因®特网 ViewCam VN-EZ 是一制作多媒体应用的 MPEG-4 数字记录装置。该记录装置使计算机用户将活动图像合并到它们的多媒体应用中,例如主页,因特网传播,和 e-mail 通信。这种记录装置使用 MPEG-4 数字移动图像压缩标准和微软高级流动格式(advanced streaming format)来产生活动图像文件,该文件较小,因此,对因特网分布来说更具实际意义。这种视频数据记录在智能多媒体存储卡中,大约提供一个小时的记录时间。

25

例如,图像捕获可以在具有平面墙壁的家庭中作为静止的背景来使用。例如,成功的分割将使不同的压缩技术在前景区和背景区中使用。然后可以实现一更高的压缩率,具有较长的记录时间,在前景区具有高质量。此外,背景可以由其他场景来重复以产生一特殊的效果来对用户增强吸引力。新的背景图像可以是自然图像和使用因特网连接的合成图像或计算机生成图像。用户能存

30

储, 浏览, 打印和交换背景图像和用因特网连接和数据供给网站提供的合成图像, 例如夏普® Spacetown。生成具有特殊效果的新图像的能力增强了这种装置对用户的吸引力。

这种特征可以用来增强具有图像处理能力的其它移动装置的吸引力。这种例子是夏普®移动电话 J-SH 系列。这些可移动电话用多媒体性能来装备, 用于作为 e-mail 接收和发送图像。图像由内置式薄膜晶体管 (TFT) 液晶显示器 (LCD) 来显示。但是, 可移动电话经常设计为处理单个图像而不是视频顺序, 因此不同于既能用于单个图像又能够处理视频顺序的因特网 ViewCam,。

限于处理单个图像的可移动电话和其他装置, 期望具有分割单个图像的能力。使用分割视频顺序的技术分割单个图像是可能的。但是视频顺序包括多个前景和背景的信息。例如, 前景目标能从一个图像移动到另一个图像, 而背景保持静止。移动信息可被摘取来促进前景和背景的分割。实际上, 使用分割单个图像的技术可以是用于分割视频顺序的最有效的技术。在有些情况中, 分割单个图像比分割视频顺序是更困难的。

早期的系统使用精细控制的背景诸如同一的彩色屏幕或者在前景对象之后照度背托执行分割。例如, US5808682 公开了从特殊的背景中分割前景对象的数据压缩系统, 它通过已知的颜色统一被照亮。可以利用任意颜色, 不过蓝色更为通用。因此这种类型的颜色背托经常被称为蓝色背托。然后前景对象可利用著名的色度键技术来分割。在大型的彩色背托上, 实现同一的照度不是非常简单的事情。US5424781 公开了一线性图像合成系统, 该系统对非同一照度和/或没有引起边发光, 边发暗, 边细节损失和其它变形的颜色背托的颜色进行校正。

对黑白图像, 公认的是利用一控制的背景以致于试图分离前景对象和背景为两个不同的灰度级别范围。典型地通过在灰度级的直方图中发现波谷值来实现分割。Nobuyuki Otsu 的“从灰度级别直方图中的 A 阈值选择方法”, IEEE 关于系统、人和控制论的学报 1979 年 1 月的第 SME-9 卷第 1 期的 62-66 页 (Nobuyuki Otsu “A threshold selection method from grey-level histograms” IEEE Trans on systems, Man and Cybernetics Vol. SME-9, No.1, January 1979 pp.62-66) 公开了这种寻找一个从它们的背景中分割前景对象的最优阈值的方法。附图 1 示出了这种类型的直方图, 其中 $h(t)$ 代表像素的数量, t 代表像素值的振幅。控制

的背景是大多数背景像素具有相当低的水平而前景像素具有趋向占据更高范围的水平。Otsu 尝试定义在两个范围之间的波谷的阈值。

- 这种技术存在几个问题。例如，尽管图 1 指出了在背景和前景灰度级别范围之间存在明确的波谷，这只是在非常仔细控制的背景和一些可能的但不完全确定为前景的情况下。

如果该技术不限于非常精细的控制条件，那么问题就变得严重。特别地，如果不是对所有的图像进行分割，则前景像素的显著数量将具有在阈值以下的级别而背景像素的显著数量将具有在阈值以上的级别。因此，任何被选择的阈值将引起不正确的分割。

- 10 分割图像的另一个技术在 T Fugimoto 等的“从移动图像中去掉背景区域的方法” SPIE 卷, 1606 可视化通讯和图像处理, 599-606 页(“A method for removing back-ground regions from moving images”, SPIE vol. 1606 Visual communications and image processing 1991, imaging processing, pp. 599-606)。为了照亮强度波动，这个技术利用像素值和极性。

- 15 附图 2 为与图 1 具有相同轴的直方图，但示出了照亮强度波动的影响。在缺少这种波动的情况下，在直方图中示出的干扰在与斜坡边对称垂直的轴中心具有窄的峰值。当照亮强度波动发生，这个峰值在水平上偏移。Fugimoto 等的技术是通过匹配峰值的实际位置的高斯分布和模拟剩余的曲线而获得不对称的正的和负的阈值 T_1 和 T_2 ，它假定代表恒定函数的前景像素水平。在高斯分布和常函数之间的相交为要处理的图像给定阈值 T_1 和 T_2 。然后它被假设为在
- 20 阈值之间的所有像素值代表噪声。

这个技术具有与 Otsu 相同的问题。尽管它可抵抗发光强度波动，当时阈值的选择不能在每个图像被分割之前完成。

- US 5878163 公开了成像目标追踪器和决定用来从其背景中优化辨别目标的阈值的方法。目标被分别假定为占据与目标的内部和外部区域相关的两个直
- 25 方图中识别的灰度级区域。直方图都被递归地平滑化且然后将实际观测像素值的检索表计算出来。两个最优的阈值被选择并设置在各自的直方图分割端。相似的图在时间上适合目标特征。目标的灰度级分布被用来选择它们所属于的目标是高的灰度级带的阈值。对这种类型的应用的精确分割不是必要的。

- 30 JP10/63855 公开了提取期望的目标区域的方法。图像分为彩色区域，每个

包含相同颜色的像素。通过用户选择所要区域，且如果彩色区域具有与其所要区域基本上重叠，它作为期望目标的一部分进行选择。然后将包含所有选择的颜色区域的整个区域定义为期望目标区域。这个技术可应用到静止图像和顺序图像。然而，它需要一预设的注意区域且包含落在这个区域的背景颜色区域。

- 5 尽管这些方法可取得对期望应用的合理分割结果且通常有计算效率，不过具有从在强度或颜色中的目标中识别仔细控制的背景的需求严格限制了应用的范围。

- 于是更大的挑战任务是如何从一般情况下的背景中分割前景对象。这些方法经常要求当前帧与预设帧之间的差异为特征的差值图像的计算。预定帧可以是背景的预记录图像，或者为以前帧，或者为从以前帧中生成的图像。
- 10

US5914748 公开了一将对象插入到差异背景的电子合成系统。这种方法减去背景的以前记录图像的每个图像以生成差值图像。屏蔽图像然后通过对差值图像定限。屏蔽图像被用来从其背景分割前景对象。方法执行简单但需要通过人工校正，从而在分割的前景区和背景区域中去除膺象。

- 15 JP5-89243 公开了使用三个连续的帧来提取对象的背景图像可替换装置。水平扫描每个线以检测对象的完整边。这一技术不能在静止图像中使用且只能提取形状相当简单的对象。

- JP6-52311 公开了包含三个步骤的提取移动对象的方法。第一步检测每个图像的静止区域。在这一步，作为两个连续图像的直接差异的绝对值而获得差值图像。然后该差值图像与预定的阈值相比较以确定静止图像区域。第二步从所有差值图像中提取的所有静止区域中建立背景图像。第三步通过对由第二步中形成的背景图像中导出的差值图像和输入图像来识别移动对象。这一技术仅能用于移动对象且不能用于静止图像。
- 20

- JP2000-251079 公开了动态图像对象提取技术，它使用时间方向的照度直方图。对每个像素位置建立直方图以描述在时间方向上的照度分布。直方图被用来定义背景像素和移动对象像素。然而，每个像素需要自己的直方图，从而这个技术需要非常大的计算能力和内存。
- 25

- 这些已知技术计算昂贵且不适合快速或者甚至实时应用，诸如在此提到的摄像机和移动电话，它们具有有限的计算能力和内存容量。进一步，很难或不可能在没有牺牲性能的情况下，有效执行这样的方法。
- 30

US5848183 公开了从扫描的文档图形中生成和利用直方图数据的系统和
方法。利用在样本窗口内的图形像素生成该直方图数据，窗口位置和尺寸需要
预先设定，例如，通过用户使用用户界面，或使用特殊的记号或者描述样本窗
口边界的记号集。然后分析确定与白背景相关的具有峰值频率的第一灰度值，
5 与黑阈值相关所谓具有峰值频率的第二灰度值和与图形直方图相关的具有最小
频率的第三灰度值，和与第一和第二灰度值相关的具有等于四分之一和/或八分
之五的峰值频率的其它灰度值。这些不同灰度值可用来增强复印机的图形对比
度。

US6043900 改进了检测扫描系统扫描文档的背景类型和确定基于从扫描
10 图形中发展铅制直方图对图形着色的最优方法的方法。

US5280367 公开了将复杂文档的扫描图像转变为文本被保留并从背景分
离出图像的系统。该系统首次将扫描图像再划分为块，然后逐个检查每个块的
像素以建立该像素的灰级值的直方图。该直方图被划分为第一，中间和最后区
域。如果一个或更多峰值出现在第一和最后区域中且单个直方图峰值在中间区
15 域内出现，则重新检查像素以确定在附近的具有第一区域峰值的像素中间的峰
值灰级的像素的出现频率。如果这个频率高，则中间峰值被认为是背景信息。
在确定阈值后，系统重新扫描应用该阈值从块内背景信息分离文本的块。

US5831748 公开了数字复印机的图像处理器，特别对快速移动到底色和
文档图像的背景。底色是没有记录信息的区域，诸如白纸上的白色基础区。背
20 景是位于主要对象后面的重复或者涉及主要对象的文档不重要的部分的区域。
例如，黑色字符可在形成背景的灰色区域上印制。在文档图像或者每线的更多
优选中生成许多区域的直方图。底色趋向在白色级别或者在白色级别附近产生
峰值，在文档中的黑色字符是稀疏的且低级别地广泛散布在直方图中，而背景
趋向于占据低于底色范围的范围。直方图的三个窗口可被定义以提取底色，背
25 景和分离特征的特征量。根据移动通过使用底色峰值和三个频率确定的底色的
底色峰值和底色范围，底色数量的窗口被确定。然后图像数据根据底色范围被
校正，根据校正的数据在纸张上形成图像以致于这或每个底色在没有人工干预
的情况下被移除。

1999 IEEE 计算机协会关于计算机视觉和模式识别论文集的会议，1999 年
30 6 月 23-25 日第 459-464 页第 2 卷，Gordon 等的“基于将范围和颜色的背景估

计和去除”(1.999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings, 23-25 June 1999, pp 459-464, v6i. 2, Gordon et al, “Background estimation and removal based on range and colour”)公开了基于组合范围分割和颜色分割技术的分割图像为前景和背景区域的技术。尽管未能很清晰或者详细地描述,但基于分析包含移动前景对象和
5 在一些方式从这一非封闭背景图像衍生的帧顺序的颜色分割技术出现了。

EP1115254, 这里合并参考一些内容,公开了分割像素图像为前景和背景区域的高效技术。作为在被分割的图像和非封闭背景的图像之间的差异,形成差值图像。前边图像的背景和前景区被分析以致提供于背景区域有关的噪声直
10 方图和与前景区有关的信号直方图。第一阈值作为最低级别从噪声直方图中获得以致背景像素的预定部分具有较低噪声级别。第二阈值作为最高级别从信号直方图中获得以致前景像素的预定部分具有更高的信号级别。

作为在被分割的图像和非封闭背景图像的相关像素之间的差异,形成差值图像。如果差值图像像素值小于第一阈值和像素被大多数指定到背景区域的
15 相邻差值图像的像素所包围,则每个差值图像的像素被作为背景像素指定。如果差值图像像素值大于第二阈值和如果像素被大多数指定到前景区的相邻像素所包围,则每个剩余的像素被作为前景像素指定。然后通过对比在第一和第二阈值之间的第三阈值,剩下的像素作为候选前景或背景像素被指定。如果大多数与其相邻的像素已经指定到背景区域或者为候选背景像素,每个候选像素然
20 后作为背景像素被指定。余下的像素被指定到前景。

尽管这个技术对分割静止背景的图像次序非常有效,但它需要一个非封闭背景的预先记录图像。例如,这可以通过在前景对象进入摄像机的视觉领域前捕获一个或更多背景的初始图像来实现。对比被分割的图像和背景图像的这种类型的方法和
25 其它方法这里称为“两次摄像”方法。

背景在整个图像捕获过程期间实际上应当保持一致从而图像捕获装置应当被安装在固定支撑上。这对许多应用来说不方便或不可能。例如,一个用户可能希望在对另一个人对焦时手持相机。一旦实现了正确的对焦,用户按下开始按钮便开始图像捕获。为了方便起见,在允许相机移动实现分割应当是可能的,即使仅在允许平缓移动时。两次摄像方法对这样的应用不适合,因此至少
30 为了在没有可利用的背景图像的情况下开始分割,希望有一次摄像技术。同样,

对一些应用诸如从因特网上下载图像，可能没有可利用的背景图像。一次摄像方法是必须的。

- 有许多用于分割的一个图像而非可用的图像次序情况。这种情况需要在没有预先记录背景图像时分割单个图像的“一帧”方法。如前所述，用于分割
- 5 单个图像的技术对分割录像次序不十分有效。因此术语“一次摄像”和“一帧”被用来辨别不同的情况。

发明内容

根据本发明的第一方面，提供一种分割像素图像的方法，该图像具有至少部分封闭背景区的一个前景区和至少一个非封闭的背景区，该方法包括步骤：

- 10 (a) 确定至少一个背景区的第一视觉特征；
- (b) 从第一视觉特征中产生第一非封闭的背景图像；
- (c) 将像素图像中的至少一些像素的每个像素同第一背景中的相应像素进行比较；和
- (d) 至少部分依赖于步骤 (c) 的结果，将至少一些像素中的每一个像素
- 15 定为前景或背景像素。

第一视觉特征可以基本上是同一的视觉特征。第一视觉特征可以基本上是单色。

- 像素图像中的每个像素可以由多个彩色分量表示；步骤 (a) 可以包括为每个颜色分量值的至少部分像素图像的像素数的每个颜色分量的形成直方图并计
- 20 算每个直方图的颜色分量值，在该直方图上存在最大的像素数；和步骤 (b) 可以包括把颜色分量值归于第一背景图像的每个像素。

可以将至少部分像素图像进行手工选择并包括至少背景的至少一部分。

像素图像可以是风景格式且至少部分像素图像包括侧面部分。

像素图像可以是肖像格式且至少部分像素图像包括顶面部分。

- 25 像素图像可以是肖像格式且至少部分像素图像包括顶角部分。

步骤 (b) 包括将第一视觉特征归于第一背景图像的每个像素。

该方法可以包括至少用步骤 (a) 重复步骤 (a) 到(d)一次，包括在前述步骤(d)中指定给背景像素的至少一些像素中确定第一视觉特征。

该方法可以包括下面的步骤：

- 30 (e) 产生第二非封闭的背景图像，该图像包括多个像素块，每个像素的

块具有基本上一致的第二视觉特征;

(f) 将像素图像的至少部分像素的每个与第二背景图像的相应像素进行比较; 和

(g) 至少部分依赖于步骤 (f) 的结果, 将至少一些像素中的每一个指定
5 为前景或背景像素。

第二视觉特征可以基本上是单色。

可以通过像素图像的每个像素的相应视觉特征至少部分地确定每个第二视觉特征, 该像素相当于块中的像素并将其指定为背景像素。

每个第二视觉特征可以至少部分地由第一视觉特征来决定。

10 每个块的第二视觉特征可以包括第一视觉特征和相当于块中像素的像素图像的像素并指定为背景像素的相应视觉特征的线性组合。

该方法可以包括至少用每个第二视觉特征重复步骤 (e) 到(g)一次, 该第二视觉特征由相当块中像素的像素图像的像素的相应视觉特征部分地确定, 并在前述步骤(g)中指定给背景像素。

15 每个块可以包括一预定数量的像素。每个块可以包括像素的方阵列。该方法包括至少一个循环, 该循环包括在每个块中用像素的简化数执行步骤 (e) 到(g)至少一次。

该方法可以包括下面的步骤:

(h) 产生第三非封闭的背景图像, 其中相应于像素图像所指定的背景像素的每个像素具有第三视觉特征, 该第三视觉特征由所指定的背景像素的相应视觉特征至少部分地确定;
20

(i) 将像素图像的至少部分像素的每个与第三背景图像的相应像素进行比较; 和

(j) 至少部分依赖于步骤 (f) 的结果, 将至少一些像素中的每一个指定
25 为前景或背景像素。

第三视觉特征可以是颜色。

相当于所指定的背景像素的第三背景像素的每个像素的第三视觉特征可以包括所指定的背景像素的相应视觉特征和第一或第二视觉特征的线性组合。

在步骤 (h), 相当于像素图像的所指定的前景像素的第三背景图像的每个
30 像素可以具有第一或第二视觉特征。

该方法可以包括用在前述步骤 (j) 中确定的步骤 (h) 中背景像素重复步骤 (h) 到 (j) 至少一次。

步骤 (c) 和 (d) 可以包括步骤:

(k) 选择至少一个背景区中的至少一部分;

- 5 (l) 从至少一部分像素值中获得第一阈值以便使像素的第一预定比值比第一阈值要低;

(m) 形成一差值图像作为在像素图像中的每个像素和第一背景图像的相应像素之间的差值; 和

- 10 (n) 如果差值图像的值低于第一阈值且大于相邻差值图像像素的第一预定数量的值低于第一阈值, 指定每个差值图像像素为背景像素。

步骤 (f) 和 (g) 可以包括步骤:

(k) 选择至少一个背景中的至少一部分;

(l) 从至少一部分像素值中获得第一阈值以便使像素的第一预定比值比第一阈值要低;

- 15 (m) 形成一差值图像作为在像素图像中的每个像素和第二背景图像的相应像素之间的差值; 和

(n) 如果差值图像的值低于第一阈值且大于相邻差值图像像素的第一预定数量的值低于第一阈值, 指定每个差值图像像素为背景像素。

步骤 (i) 和 (j) 可以包括步骤:

- 20 (k) 选择至少一个背景中的至少一部分;

(l) 从至少一部分像素值中获得第一阈值以便使像素的第一预定比值比第一阈值要低;

(m) 形成一差值图像作为在像素图像中的每个像素和第三背景图像的相应像素之间的差值; 和

- 25 (n) 如果差值图像的值低于第一阈值且大于相邻差值图像像素的第一预定数量的值低于第一阈值, 指定每个差值图像像素为背景像素。

第一预定比值可以在 0.5-1 之间, 第一预定比值可以基本上等于 0.75。

第一预定数量可以基本上等于相邻差值图像像素数量的一半。

用在前述步骤 (n) 中包括背景像素的至少一部分步骤 (k) 可以重复步骤

- 30 (k) 到 (n) 至少一次。

每个步骤 (n) 可以包括形成指定为背景像素的差值图像值的第一初始直方图, 且步骤 (l) 从第一合成直方图中得到第一阈值, 该合成直方图包括在前述步骤 (n) 中形成的第一初始直方图和小于前述步骤 (l) 中的第一合成直方图的第一预定比值的总和。第一预定比值是一半。

5 该方法包括步骤:

(o) 选择至少一个前景中的至少一部分;

(p) 从至少一部分像素值中获得第二阈值以便使像素的第二预定比值超过第二阈值;

(q) 指定每个差值图像不指定为背景像素, 如果差值图像的值超过第二
10 阈值且大于相邻差值图像像素的第二预定数量的值高于第一阈值, 则将其指定为前景像素。

第二预定比值可以在 0.5-1 之间, 第二预定比值可以基本上等于 0.75。

第二预定数量可以基本上等于相邻差值图像像素数量的一半。

用包括在前述步骤 (q) 中指定的前景像素的至少一部分步骤 (o) 可以重
15 复步骤 (o) 到 (q) 至少一次。

每个步骤 (q) 可以包括形成指定为背景像素的差值图像值的第二初始直方图, 且步骤 (p) 可以从第二合成直方图中得到第一阈值, 该合成直方图包括在前述步骤 (q) 中形成的第二初始直方图和小于前述步骤 (p) 中的第二合成直方图的第一预定比值的总和。

20 第二预定比值可以是一半。

该方法可以包括指定每个差值图像像素, 该像素没有指定为前景或背景像素, 如果差值像素的值小于第三阈值则指定为候选背景像素。第三阈值可以在第一和第二阈值之间, 第三阈值可以是第一和第二阈值的算术平均。

该方法可以包括指定每个差值图像像素, 该像素没有指定为前景或背景像
25 素, 如果大于相邻像素的第三预定数量指定为背景像素或候选像素, 则将其指定为背景像素。第三预定数量可以基本上等于相邻差值图像像素数量的一半。
该方法可以包括指定每个没有指定为前景或背景像素的差值图像像素为前景像素。

该方法可以包括形成二进制掩码, 其元素相当于差值图像像素, 如果相应
30 的差值图像像素被指定为背景像素则每个元素具有第一值, 且如果相应的差值

图像像素指定为前景像素，则第二值不同于第一值。

根据本发明的第二个方面，提供一种用于分割像素图像的装置，该图像具有至少部分封闭背景的一个前景，和至少一个非封闭的背景，其特征在于包括确定至少一个背景的视觉特征的装置，从视觉特征中产生非封闭的背景图像
5 的装置，比较像素图像的至少一些像素的每个像素与背景图像的相应像素的装置，和至少部分取决于比较装置来指定至少一些像素的每个像素为前景像素或背景像素的装置。

根据本发明的第三个方面，提供一种分割像素图像的装置，包括一可编程数据处理器和控制数据处理来执行根据本发明第一方面的方法的程序。

10 根据本发明的第四个方面，提供一种存储介质，包括用于控制数据处理来执行根据本发明第一方面的方法的程序。

根据本发明的第五个方面，提供一种程序，用于控制数据处理来执行根据本发明第一方面的方法。

15 根据本发明的第六方面，提供一种图像捕获装置，包括根据本发明第二或第三方面的装置。

该技术可以用于分割一个像素图像成一个或多个前景和一个或多个相对同一的背景且不需要非封闭背景的预记录图像。

为了分割单一图像，用户可以直接使用手提图像捕获装置进行捕获并不需要固定支承例如三脚架。为了捕获图像的顺序，图像捕获装置也可以是手提的
20 只要是逐渐移动就可以。

至少本发明的一些实施例可以从双镜头拍摄方法中得到，例如在 EP1115254 中公开的类型，具有稳定性和易于实现的优点。例如，双镜头拍摄方法在分割视频顺序时使用。在一些实施例中，分割的质量可以在分割过程中自我改进。在这种实施例中，分割的结果提高了噪音的信号直方图的评价，也提高了形成
25 连续改进环的下一分割。

该技术可以在按照功率和仅包括运算的存储需求以有效计算的方式来实施，该运算可以只使用整数来运行。因此该技术特别适合快速和实时应用，例如在上面所描述的类型。

附图说明

30 下面将参照附图，通过例子对本发明进行进一步的描述，其中：

- 附图 1 示出了分割图像的第一已知技术的直方图；
附图 2 示出了分割图像的第二已知技术的直方图；。
附图 3 示出了构成本发明第一实施例的图像分割方法的流程图；
附图 4 和 5 示出了图 3 中方法的部分初始步骤图；
5 附图 6, 7 和 8 示出了用于计算噪声频率图的自动选择区域图；
附图 9 示出了图 3 中的部分方法的流程图；
附图 10 图解示出了不同图像的生成。
附图 11 示出了在图 3 方法中阈值选择的两个直方图；
附图 12 示出了在图 9 中所示的技术改进的流程图；
10 附图 13 示出了像素的四个邻域和八个邻域图；
附图 14 示出了构成本发明的另一个实施例的方法的流程图；
附图 15 到 17 示出了图 3 所示的方法改进的流程图；和
附图 18 示出了构成本发明的另外一个实施例的分割图像的装置的框图。

具体实施方式

- 15 图 3 示出了利用三个步骤 51, 52, 53 进行分割图像的方法。每个步骤执行循环过程，该过程根据在 EP1115254 中公开的技术基础上利用双摄像方法通过对该图像和估计的背景图像进行比较，重复估计和更新背景图像并分割像素图像。在第一步骤 51 中，背景图像作为一个完全一致的图像而被估计。在第二步骤 52 中，背景作为一个分组图像而被更新。在第三和最后步骤中，背景
20 图像以单个像素分辨率被更新。

该方法开始于 54 且在第一循环期间，在 55 中利用下面的技术估计作为同一图像的背景图像。

- 彩色图像可用大量的彩色分量来表示，诸如红色 (R)，绿色 (G) 和蓝色 (B) 分量，亮度 (Y) 和两个彩色的不同分量 (U,V) 和色调 (H)，饱和度
25 (S) 和值 (V)。该技术可用于任何彩色格式，但为了方便期间且不失普遍性，该技术对 RGB 分量的应用将在下文详细描述。同样，不同于色彩的视觉特征可用于估计和代表背景图像，且这些特征包括图像平均信息量和纹理（例如 Law's 的纹理参数）。

- 图像 $I(x,y)$ 的每个像素可用 $R(x,y)$, $G(x,y)$, $B(x,y)$ 来表示，其中 x 和 y 为
30 像素的离散坐标。如果像素代表前景对象或区域，则称为前景像素。如果像素

代表不被前景对象或区域阻挡的背景区域，则被称为是背景像素。

应用本技术的图像，其背景基本上是同一的且包含所有基本上相同颜色中的一个或多个区域。因此，所有的背景像素有相同的不变的颜色分量 R_0 , G_0 , B_0 ，因此它们不依赖于位置。

- 5 例如，为了确定从包括前景像素的图像的所有像素中的红色分量中生成红色分量 R_0 ，直方图 $h_r(r)$ 。通常，图像像素有遍布一个范围的红色分量，但是因为背景像素都有相同的红色分量值，它们在如图 4 所示的直方图上的 R_0 位置处形成一个峰值。如果图像上足够数量的背景像素是可视的，则一般该峰值基本上大于在直方图上且在前景区与普遍红色分量有联系的其它任何峰值。这种峰值可容易地被辨认出来且其值 R_0 代表同一背景像素的红色分量。
- 10

对其它每个颜色分量来说，则执行相同的直方图过程，以便确定背景像素三种颜色的分量值。然后同一背景图像通过将三种背景颜色分量归于背景图像的所有像素中而被估计或产生。这相当于将同一背景颜色延伸到被阻挡的背景区域或者在前景或区域后面的区域。

- 15 实际上，图像的背景不是完全一致的，且即使实际的背景高度一致，由于背景平面的不平坦，背景亮度的不一致而出现差异，在图像捕获装置和图像捕获装置的电子噪声中视觉系统色差通过前景对象，阴影被投影在背景上。因而对背景区域的彩色分量来说，占据小范围而不是具有单个值更有用。这在图 5 中示出，它示出了更宽的峰值出现在与最普遍的颜色分量相应的最大值的直方图上。再者，直方图最大值的红色分量值 R_0 出现且通过将此值和相应的绿色和蓝色分量值赋给背景图像的所有像素来生成图像。
- 20

- 在前景区具有相当大的同一颜色的区域是可能的，该颜色在直方图上产生一比与相应背景区更大的最大值。为了避免这种情况，图像的背景区域可进行人工或自动判别。例如，如图 6 所示，目标前景区 56 基本上位于“风景格式”图像的中心，该“风景格式”图像具有仅包含背景的左侧和右侧区域 57，在这种情况下，为了形成直方图，可选择区域 57。这种区域的选择可以是自动的，因为只要在这些区域的大多数像素是背景像素，则不需要进行特别计算。因此这种技术对风景格式的图像来说可能最有用。
- 25

- 对肖像格式的图像，更适合选择其它区域以形成直方图。图 7 示出了肖像格式图像的一般构成，因而下半部一般包含在图像顶部具有相对小的背景区
- 30

域的前景区。在这种情况下，可选择在图像顶部的区域 58 以形成直方图。可选择的是，图 8 示出了选择用于形成直方图的顶部角落区域。

在一些应用中，例如在分割是脱机执行时，则可允许用户手工选择背景区域或所有区域。例如在显示器上显示图像且用户可用鼠标指针选择区域以在 5 形成直方图时使用。

在这种情况下，通过用户的操作而画出边界（例如鼠标或者输入笔的操作）使得选择一边界区域。可替换地，默认边界显示在显示器上，然后用户可改变默认边界。

在诸如风景格式图像的情况下，如图 6 所示清晰地显示器显示了边界 10 区域和前景区之间的边界，从而用户可检查该边界。当用户确定在显示器上显示的边界不合适时，通过用户的操作（通过操作触板，拖动鼠标，或者操作遥控或键盘的按钮），向右或向左改变边界。当用户确定在显示器上显示的边界合适时，通过用户的操作把边界确定下来，从而终止背景区域的选择。

在肖像格式图像的情况下，过程与上述相似，使得如图 7 所示的边界通过 15 用户在显示器上检查，并且如有必要，通过用户的操作，向上或向下改变边界。然后，把边界确定下来从而终止背景区域的选择。

可以使用如图 8 所示的边界而不是图 7 所示的边界。在这种情况下，通过用户在显示器上检查，并且如有必要，通过用户的操作，向上或向下，向左或向右改变边界。然后，把边界确定下来从而终止背景区域的选择。

20 背景不限于直线，也可以是之字形线或者曲线。例如，可事先预备形成一代表人体上半部的大体剖面的椭圆形或者曲线的边界。边界的移动不限于各种变化，但也涉及旋转和/或变形（放大或减小边界的部分或整体）。

可替换地，可以事先准备边界的多种模式，用户可选择模式之一来对背景区域进行选择。将选择的模式显示在显示器上。当用户确定显示的模式是不 25 合适时，可以移动，旋转和/或变形边界直到如上所示的用户操作将边界确定下来。

如上所述，描述了用户对背景区域的选择。可选择地，在移动照相机时定位边界可获得背景区域。更特别的是，用户检查在摄影期间照相机照的图像，并在观察显示器中显示的边界时，调整照相机使得目标在前景区内。照相机可 30 通过诸如直接移动照相机或利用照相机的变焦功能来调整。

如图3所示,估计的或生成的背景图像作为构建背景图像60与该图像(仅有一个图像用来分割)或当前图像61(图像顺序)一起被提供到两次拍摄的分割过程62中,这基于EP1115254中公开的技术,如图9所示,它是更详细地以整个第一阶段51所示的“扩大”流程图的形式的块功能图。该图示出了数据的供给和步骤或操作的顺序。

当前图像61的像素和每个与构建的背景图像60相应的像素之间的差值在5处计算并确定差值图像。这在图10中图解,它示出了背景图像60和当前图像61。步骤5计算如图8中在步骤8处示出的差值图像。差值图像具有大部分或全部代表噪声被像素估值的噪声区域。图像8也具有与前景图像对应的信号区域10,前景图像像素代表在图像像素和与背景图像的背景像素相应的像素之间的差值。

步骤3确定差值图像是第一个这样的差值图像且步骤1c从差值图像8的背景和前景区9,10的统计特征中计算噪声和信号直方图1d。

在步骤1c之后,或者如果步骤3确定差值图像不是第一差值图像,那么控制进入到步骤11,其在差值图像的背景像素中识别有力的候选者。特别地,每个像素依次与第一阈值进行对比,该阈值根据在步骤1c确定的噪声的统计特征,在步骤24中被确定。值低于第一阈值的像素和其值低于第一阈值的大多数相邻像素被指定到背景区域(或者多于一个这样的背景区域之一)。

步骤12通过将步骤11中不指定给至少一个背景区域的每个差值图像像素与第二阈值进行比较以辨别前景像素的强候选者,根据在步骤1c中确定的信号区域10的信号强度直方图,第二阈值在步骤25被确定。这些其值大于第二阈值且大多数值高于第二阈值的相邻像素的像素被指定给该或每个前景区。

然后步骤13识别是否没有指定到背景和前景区的像素应当被指定给至少一个背景区域和至少一个前景区。每个这样的像素与第三阈值相比较,该第三阈值在步骤26中被确定,并在第一和第二阈值之间具有一个级别。那些级别低于第三阈值的像素被作为候选背景像素识别。如果大多数相邻像素已经被作为比较像素或者作为候选像素被识别出来,然后每个候选背景像素被指定到至少一个背景区域。然后余下的像素作为前景像素被指定。

在每个步骤11,12和13中,对整个差值图像来说,通过“空间过滤”来决定像素是如何指定的。然而,对接近每个图像的边缘和角落的像素,每个

像素邻域受到限制。尽管空间过滤在图像边界总是以相同的方式进行，但是，至少在一些应用中，这更合适将空间过滤限制到每个图像的中间部分，该图像中，每个像素具有对相邻像素的完全补充。留下围绕边界的像素未被过滤。

- 例如，如果图像尺寸是 $M \times N$ 个像素和用于空间过滤的窗口尺寸为 $m \times n$ 个像素，然后包含 $(M-m+1) \times (N-n+1)$ 个像素的中心区域受到空间过滤且可以通过利用合适的更小窗口尺寸将周围边界的像素进行空间过滤或留下来未进行空间过滤。对分割不必如此精确的应用，周围边界的像素的指定可利用阈值和没有过滤过程来单独确定。

- 步骤 14 形成二进制掩码。特别地，该掩码包含与差值图像相关的元素。
- 10 如果相关的差值图像像素被指定到至少一个背景区域，则把每个二进制掩码的像素归于第一个值，或者如果相关的差值图像像素被指定到至少一个前景区，则把每个二进制掩码的像素归于第二个值。因此二进制掩码代表期望的分割且可以用作图像的进一步处理。

- 步骤 16 在分析系列当前图像的基础上更新信号和噪声直方图且将更新的直方图返回到步骤 1。步骤 17 确定是否满足终止条件。如果没有，控制返回到步骤 55。否则，在步骤 18 阶段计数器加 1。

任何合适的终止条件可用于步骤 18 中。例如，当重复的次数达到被设为 2 或 3 的预定数字时，可终止阶段 1。

- 在每个顺序重复期间，根据二进制掩码 14 分割图像，并在步骤 55 中在作为背景像素辨别的像素基础上估计背景图像。相似地，噪声和信号直方图根据被二进制掩码 14 所定义的背景和前景像素的统计特征被更新，其改进分割的下一个循环。

将更详细地描述用于图 9 所示方法中的单个技术。

步骤 5 定义差值图像，给出 $D(x, y)$ 如下：

- 25 $D(x, y) = F[I(x, y), B(x, y)]$

其中 F 为表征在 $I(x, y)$ 和 $B(x, y)$ 之间差值的用户定义度量法。

在 $I(x, y)$ 和 $B(x, y)$ 为灰度级图像的最简单的情况下，例如， $D(x, y)$ 也可定义为它们的直接差值，也就是：

$$D(x, y) = I(x, y) - B(x, y)$$

- 30 在没有噪声的理想情况下， $D(x, y)$ 的值则给定为：

$$D(x, y) = \begin{cases} s(x, y) & \text{如果像素} \in \text{前景对象} \\ 0 & \text{如果像素} \in \text{背景} \end{cases} \quad 16$$

其中 $s(x, y)$ 为信号的测量, $\in$ 意指属于。 $S(x, y)$ 项代表目标像素和在背景图像同样位置的像素之间的差值测量。如果所有的前景像素生成非零信号值, 则它们从给定差值测量为 0 的背景像素中识别和分离出来。

- 5 这个简单的方法在实际中是不可行的, 因为在图像中总是有噪声且可能有一些像素与背景图像相关的那些有同一值的情况。为了包括噪声的影响, 最

$$D(x, y) = \begin{cases} s(x, y) + n(x, y) & \text{如果像素} \in \text{前景对象} \\ n(x, y) & \text{如果像素} \in \text{背景} \end{cases}$$

后的公式可改变为:

其中噪声项 $n(x, y)$ 被假定为均值为零, 方差为 σ 的随机变量。

- 10 为了能从背景像素中分割前景像素, 信号的强度比噪声的强度要显著的强。可将述公式则重写为:

$$D(x, y) = \begin{cases} s(x, y) & \text{如果像素} \in \text{前景对象} \\ n(x, y) & \text{如果像素} \in \text{背景} \end{cases}$$

其中上述公式中的 $s(x, y)$ 可认为是包含噪声元素, 实际上它不可以从实际信号本身区别开来。

- 15 从噪声项 $n(x, y)$ 中分离信号项 $s(x, y)$ 的最简单的方式是找到阈值 T , 从而所有的前景像素具有超过 T 的信号值, 且所有背景像素具有低于 T 的噪声。二进制掩码功能则可推导为:

$$M(x, y) = \begin{cases} 1 & \text{如果 } |D(x, y)| > T \\ 0 & \text{其它} \end{cases}$$

其中 1 表明前景像素且 0 为背景像素。

- 20 可以有一些限定条件, 例如, 当背景为同一颜色诸如深蓝色。差值图像 $D(x, y)$ 可利用下文表述的公式 (3) 计算出来。如果前景对象不包含同一深蓝的区域且如果成像系统的噪声级别低, 则差值图像应当在背景区域具有强的信号值。然后寻找阈值去分开这些信号和噪声值是可能的。

实际上, 这种简单的方法对背景的一般情况不能令人满意地进行。例如, 一些前景像素可与相应的背景像素相似, 并且生成低于任何给定阈值的信号值。同样地, 一些背景像素可产生高于 T 的噪声值。导致后边问题的可能原因

有许多, 包括在成像系统中的电子噪声的存在, 背景照明的变化和/或成像系统的小干扰。该方法的多个阈值实际上克服了此后描述的这个问题。

在两个图像 $I(x, y)$ 和 $B(x, y)$ 之间普遍使用的差值测量是颜色位距:

$$F\{I, B\} = \sqrt{\alpha(R_1 - R_0)^2 + \beta(G_1 - G_0)^2 + \gamma(B_1 - B_0)^2}$$

- 5 其中 α , β 和 γ 是加权因子且 $\{R_1, G_1, B_1\}$ 和 $\{R_B, G_B, B_B\}$ 分别是图像 $I(x, y)$ 和 $B(x, y)$ 的 RGB 分量。

在上述公式中使颜色分量标准化是有吸引力的。通过对颜色分量求和, 并且除以每个图像分量而达到标准化。例如 R_I 分量可被标准化为:

$$R'_I = \frac{R_I}{R_I + G_I + B_I}$$

- 10 利用这样标准化给出颜色格式的 F 例子:

$$F\{I, B\} = \sqrt{\alpha(R'_I - R'_B)^2 + \beta(G'_I - G'_B)^2 + \gamma|Y_I - Y_B|}$$

其中 $Y_I = R_I + G_I + B_I$ 和 $Y_B = R_B + G_B + B_B$ 。当在减少由于不同光照密度的影响情况下的这种标准化是有效的时候, 在 Y_I 或 Y_B 值非常小时存在一缺点。被非常小的值的除将导致对每个颜色分量中噪声有非常大的放大作用, 因此使得任

- 15 何分割办法都很困难。因而这种标准化应当小心应用。

为了降低计算成本, 在公式 (1) 中描述的色距测量大约为:

$$F\{I, B\} = \alpha|R'_I - R'_B| + \beta|G'_I - G'_B| + \gamma|B_I - B_B|$$

加权因子可预先设定以反映每个颜色分量的重要性。一般 $\alpha = \beta = \gamma = 1$ 的值设置已经在许多应用中充分使用。这导致:

- 20
$$F\{I, B\} = |R'_I - R'_B| + |G'_I - G'_B| + |B_I - B_B|$$

定义差值测量函数 F 有其它方法。例如, 在信号和噪声的统计特征的基础上进行熵测量。这一方法可生成更好的差异特征, 但趋向于更大的计算强度。

基于公式 (3) 描述的差值测量函数 F , 给出差值图像 $D(x, y)$ 为:

$$D(x, y) = \begin{cases} |s_r(x, y)| + |s_g(x, y)| + |s_b(x, y)| & \text{对前景像素} \\ |n_r(x, y)| + |n_g(x, y)| + |n_b(x, y)| & \text{对背景像素} \end{cases}$$

- 25 其中每个颜色分量的噪声假定均值为零, 方差为 σ 的随机变量。这些同样也假定为从一个颜色通道到另一个颜色通道和从一个像素到另一个像素的随机变量是独立的。

理论上, 噪声的值 n_r , n_g 和 n_b 可从负无穷到正无穷变化。这些值可由在无穷范围内整数重复, 例如从 $-N$ 到 $+N$, 其中 N 为整数且一般等于 255。 $|n_r|$,

$|n_g|$ 和 $|n_b|$ 的值则从 0 到 N 变化。这就是用绝对值操作 $||$ 表示噪声术语为 $|n_r|$, $|n_g|$ 和 $|n_b|$ 的原因。在后面对在利用绝对值之前的成像平均降低噪声的讨论中, 这变得更清晰。噪声的统计特征可以利用直方图方法来初始估计。例如, 如果图像 $I(x, y)$ 仅变化背景且在从当 $B(x, y)$ 被捕获时的不同时间被捕获, 则差值图像 $D(x, y)$ 代表每个像素的噪声变量。直方图, $h_n(t)$, 可通过对一特定的噪声值 t 出现的总
5 数目进行计数而建立起来, 其中 $D(x, y) = |n_r| + |n_g| + |n_b| = t$ 。图 5 的上部图形示出了这一噪声直方图的例子。

理论上, $|n_r| + |n_g| + |n_b|$ 从 0 到 $3N$ 变化。这意味着 $h_n(t)$ 应当具有 $3N$ 元件或接收器, 每个为整数计数器。然而, 如果和只有当所有三项达到它们的最大值
10 N 时, $|n_r| + |n_g| + |n_b|$ 的值等于 $3N$ 。实际上, $|n_r| + |n_g| + |n_b|$ 的值可能远小于这个最大值 $3N$ 。经常使用仅有 N 个接收器的直方图就足够了。对超过 N 个值的那些像素, 它们的值可被截取为 N 。当计算能力和内存存储有限时这非常有用。

在分割开始之前必须计算噪声直方图。另一方面, 如果背景像素已知则
15 可仅计算噪声直方图。这些矛盾需求的一个可能解决方法是在图像中寻找一些可能的背景区域。

例如, 在图 10 中所示的仅包含背景的左边和右边边界上的一些区域 9 的
图像的中心区域, 该目标或每个目标被初始定位。在左边和右边边界上的区域
9 可估计噪声直方图。

如果在一些应用中脱机进行计算, 则同样允许用户手工选择估计噪声特
20 征的背景区域。如果应用在个人电脑 (PC) 上执行, 例如, 图像可显示在显示器上且用户可使用鼠标指针选择这种目标区域。

一旦计算出差值图像, 利用第一阈值 T_n 第一次确定背景像素的强候选者。
这个阈值如此确定以致于最多的前景像素具有高于它的信号值且背景像素具有
低于它的噪声值。

25 这个阈值从噪声直方图 $h_n(t)$ 中确定。理想噪声值是有界的以致直方图的第一接收器不为零。然后 T_n 阈值为第一个 t 值, 余下的接收器为零。即:

$$h_n(t) = 0 \quad \text{对任何 } t > T_n$$

不幸的是, 这样一个接近理想的情况一般不存在。

当直方图超过 t 的更高范围不会都为零时, 大多数噪声值通常在图 11 所
30 示的第一接收器周围群集。剩下的接收器趋向于形成一当 t 增加时通常减少的

长尾。然后阈值 T_n 作为大于给定的背景像素百分比的 t 的最小值被定义。

该阈值 T_n 涉及一预设 η_n 值的百分比, 其中 $0.5 < \eta_n \leq 1$ 。它被确定为满足下述不等式的最小 T_n :

$$\sum_0^{T_n} h_n(t) \geq \eta_n \cdot K_n$$

- 5 其中 K_n 为背景像素的总数目且被指定为:

$$K_n = \sum_0^N h_n(t)$$

然后利用第一阈值判别背景像素的强候选者定限差值图像。二进制掩码为:

$$M_n(x, y) = \begin{cases} 1 & \text{如果 } |D(x, y)| > T \\ 0 & \text{其它} \end{cases}$$

- 10 其中 0 代表背景像素的强候选者且 1 代表为通过定限的前景像素和背景像素的那些像素。这里只对强候选者才感兴趣。那些像素在此后的描述的过程中将被进一步分级。对包含 K_n 背景像素的图像 $I(x, y)$, 大约 $\eta_n K_n$ 像素将具有低于阈值 T_n 的噪声值。余下的 $(1 - \eta_n) K_n$ 像素将具有超过 T_n 的值并在此过程中将作为背景像素而未被检测到。

- 15 这里隐含假定前景像素的绝大多数具有超过第一阈值的信号强度。然而, 小数量的前景像素可具有低于此阈值的信号值。这个简单的定限方法将使它们作为背景像素而被错误的分级。

- 20 如果 η_n 的值被设置得更高, 则 T_n 的值也同样更高且很少背景像素将未能被检测。另一方面, 更大的阈值 T_n 意味着对一些前景像素来说很容易作为背景像素来错误分级。既然分割的目的是从背景中分离移动对象, 理想分割应当即使在信号值小的时候得知前景对象和背景之间的差值。这意味着, 对分割目的来说, 阈值 T_n 越小越好。因此折中是必要的且一般地 η_n 被设置为 0.75。

- 25 一个真正的背景像素不会单独存在而总是与其它背景像素连接。如果像素实际上是背景像素, 那么其大多数相邻的像素同样是背景像素。同时, 作为背景像素未被检测到的那些前景像素在背景中经常作为小的臆象。这种观察导致后面移去这些错误检测到的像素的过滤方法。

对背景像素的当前候选者, 小的邻域, 通常选择在当前像素集中的方形

窗口。如果窗口中的像素的总数目为 m ，那么期望大约 $\eta_n \cdot m$ 个像素被作为背景像素正确分级，其中 $\eta_n > 0.5$ 。如果 m 值充分大时，这则是正确的。

- 如果背景像素在前景目标的相邻区域中，小窗口可包含一些前景像素。直观地，即使当前背景像素与前景目标相邻，则大约一半其相邻区域为背景像素且大约一半为前景像素。很明显，只有在窗口尺寸足够小时，这才是正确的。
- 5 这与上述 m 应当“足够大”的需求矛盾。

因此折中是必要的。试探性的结果表明 m 的典型值为 100，因此，给定窗口尺寸为 9×9 或 11×11 。通常窗口尺寸为单数，因此当前像素被安排在窗口中心。

- 10 然后基于相邻区域中背景像素的 ψ_n 百分比接收或拒绝背景像素的候选者。如果 ψ_n 大于或等于 0.5，则它被接收；否则被拒绝：

$$M_n(x, y) = \begin{cases} 0 & \text{如果 } M_n(x, y) = 0 \text{ 且 } \psi_n \geq 0.5 \\ 1 & \text{其它} \end{cases}$$

通过计算，作为应用于二进制图像 $M_n(x, y)$ 的移动窗口平均操作执行该过程，图像 $M_n(x, y)$ 后面是阈值为 0.5 的定限方法。

- 15 同样地，前景图像的强候选者可利用阈值 T_s 被检测， T_s 可从信号直方图 $h_s(t)$ 中计算出来。如前所述，信号直方图在初始不能利用直到第一图像被分割。对第一图像，这个阈值可被从 T_n 中简单估计且给定为：

$$T_s = \mu T_n$$

其中 μ 为大于 1 的实数。典型地， μ 被设定为 1.5 到 2。

- 20 上述公式仅在第一图像被完全分割之前信号强度的初始估计中使用。如果有关于前景目标的初始位置的另外信息，则不能使用该公式。如果前景目标开始集中在图 10 所示的图像的中心处，例如，图像的中心区域可用于建立初始信号直方图。

- 一般，信号直方图 $h_n(t)$ 从分割的前景目标中建立，只要第一图像一被分割，
- 25 它就可以利用。第二阈值 T_s 被定义为 t 的最大值，其满足下面不等式：

$$\sum_{T_s}^N h_s(t) \geq \eta_s \cdot K_s$$

其中 $0.5 < \eta_s \leq 1$ ， K_s 为前景像素的总数目。典型的 η_s 值为 0.75。

通常阈值 T_s 大于 T_n 。如果不是这种情况，则它意味着移动目标的强度和颜色与背景太相似，以致于它们的差异小于噪声的影响。分割将变得非常困难。

一旦获得第二阈值，它被用于定限差值图像以检测前景像素的候选者。

在定限后获得二进制掩码为：

$$M_s(x, y) = \begin{cases} 0 & \text{如果 } M_n(x, y) = 1 \text{ 且 } |D(x, y)| \geq T_s \\ 1 & \text{其它} \end{cases}$$

已经作为强背景像素被辨认的那些像素不需要再处理，但是它们将贡献

5 到下面描述的过滤过程。

对包含 K_s 前景像素的图像 $I(x, y)$ ，大约 $\eta_s \cdot K_s$ 个像素将有超过阈值的信号值。剩下的 $(1 - \eta_s) \cdot K_s$ 像素将有低于该阈值 T_s 的值，且此过程中将不能作为前景像素被检测到。

10 当在背景像素的强候选者的检测中，如果噪声值通过第二阈值 T_s ，则存在在作为前景像素的背景像素的错误检测。这些错误检测的像素一般单个随机分布且可利用此前描述的相似移动窗口操作被去处掉。该定位也是相似的。如果初始辨别候选者前景图像基于至少一般相邻像素作为前景像素的候选者，则其被作为一个前景像素而确认；否则被拒绝。

$$M_s(x, y) = \begin{cases} 0 & \text{如果 } M_s(x, y) = 1 \text{ 且 } \psi_s \geq 0.5 \\ 1 & \text{其它} \end{cases}$$

15 其中 ψ_s 是当前像素的小相邻区域中前景像素的强度的百分比。邻域通常作为在一个在当前像素中心的方形窗口中选择。其尺寸一般设置为 9×9 和 11×11 ，用作检测背景像素的强候选者。有时，如果前景目标小，则窗口尺寸可以减少。

20 那些既不被接受为背景像素的强候选者，也不被接受为前景像素的像素在进一步定限和过滤过程中被分级。第三阈值 T_m 被引入且具有阈值 T_n 和 T_s 之间的值：

$$T_m = \alpha T_s + (1 - \alpha) T_n$$

其中 $0 < \alpha < 1$ 。一般 α 值可被设置为 0.5。

25 利用这第三阈值 T_m 再次定限差值图像。这次，然而只有那些剩余像素需要处理。新二进制掩码函数可给定为：

$$M_m(x, y) = \begin{cases} 1 & \text{如果 } M_s(x, y) = 1 \\ 0 & \text{如果 } M_n(x, y) = 0 \\ 1 & \text{如果 } M_s(x, y) = 0 \text{ 且 } M_n(x, y) = 1 \text{ 且 } |D(x, y)| \geq T_m \\ 0 & \text{其它} \end{cases}$$

在以前的过程中，对“未定”的像素应用过滤操作。如果超过在窗口中一半像素为背景像素，则当前像素被分级为背景像素，否则为前景像素。二进制掩码最后给定为：

$$M(x, y) = \begin{cases} 1 & \text{如果 } M_s(x, y) = 1 \\ 0 & \text{如果 } M_n(x, y) = 0 \\ 1 & \text{如果 } M_m(x, y) = 1 \text{ 且 } \psi_m \geq 0.5 \\ 0 & \text{其它} \end{cases}$$

- 5 其中 $M(x, y)$ 是最终的二进制掩码， ψ_m 为需要验证在当前像素居中的方形窗口中前景像素的百分比。在以前，窗口尺寸一般设置为 9×9 和 11×11 但可被调整以适合不同尺寸的对象。

- 一旦利用二进制掩码 $M(x, y)$ 分割了图像，则信号和噪声直方图可分别利用前景和背景区域被再计算。同样，图像与二进制掩码一起被存储。然后获得
10 下一个图像并通过重复上述过程进行分割。

可以重复这些步骤以形成一循环过程。信号和噪声直方图可被再计算因为前景和背景区域现在被开始分割。这些再计算的信号和噪声直方图导致下一个循环的噪声和信号阈值 T_n 和 T_s 的更好估计。

- 可替换地，可利用累积过程更新信号和噪声直方图。例如，信号直方图
15 可被更新为：

$$H_s^t(x) = H_s(x) + \lambda H_s^{t-1}(x)$$

其中 $H_s^{t-1}(x)$ 为用于分割上次帧的以前信号直方图， $H_s(x)$ 为利用分割上次帧的结果获得新的直方图和 λ 为典型具有 $1/2$ 值的常量。

- 这相当于从那些以前的具有加权直方图，最靠近帧给定一更大的权重，
20 即

$$H_s^t(x) = H_s(x) + \frac{1}{2} H_s^{t-1}(x) + \frac{1}{4} H_s^{t-2}(x) + \dots$$

噪声直方图同样地可利用这种累积方法计算出来。

- 如图 3 所示，每次阶段计数器在 18 中增加，步骤 64 和 65 确定应当执行的阶段。因此，一旦已经满足 17 则终止规则，阶段计数器 18 增加阶段计数到
25 2 且步骤 64 和 65 促使执行第二阶段 52。第二阶段 52 与第一阶段 51 相似并利用相同的双摄像分割 62。然而，第二阶段不同于在 66 中按组分辨率更新的背景图像。

尽管背景图像一般不完全一致，在背景区域中的色变一般是渐变的。即，颜色在小面积上几乎为常数。因而背景图像改进的估计图像作为小块的镶嵌被建立，每个具有恒定的颜色。

例如，在给定一 $mk \times nl$ 个像素的图像尺寸，图像尺寸背景图像可分为 $m \times n$ 像素块和包含 $k \times l$ 块的图像。给定以 8×8 像素或 16×16 像素的方块， m 和 n 典型地被设置为 8 或 16。这涉及到基于块的处理，例如，MPEG 应用中每个块在 Y 分量上包含 16×16 像素，在 U 和 V 分量上包含 8×8 像素。

对在 i th 水平位置和 j th 垂直位置上的块，所有像素被归结一恒定的颜色值 R_{ij} , G_{ij} , B_{ij} ，其中 $0 \leq i < k$ 和 $0 \leq j < l$ 。 R_{ij} , G_{ij} 和 B_{ij} 值被认为是块的“平均”颜色且从块到块不同。初始分割背景和前景区对这个改进提供有用的信息。例如， R_{ij} 可被定义为

$$R_{i,j} = \alpha R_0 + (1 - \alpha) \Delta_{i,j}$$

其中 $0 < \alpha \leq 1$ 且为

$$\Delta_{i,j} = \frac{R_0 + \sum_{(x,y) \in \text{当前块}} R(x,y) \cdot [1 - M(x,y)]}{1 + \sum_{(x,y) \in \text{当前块}} [1 - M(x,y)]}$$

其中 $R(x, y)$ 为在当前块的每个像素的红色分量， $M(x, y)$ 是二进制掩码。背景像素在当前的块中越多，它们对 Δ_{ij} 值贡献越大。

如果在当前块中没有背景像素， Δ_{ij} 值等于 R_0 且也同样等于 R_{ij} 的值。

如果在当前块中所有的像素为背景像素，那么背景像素的总数量等于 mn 。然后 Δ_{ij} 值基本上代表该块的平均颜色值。 α 典型地被设置为 0.5，以致 R_{ij} 大约为 R_0 和 Δ_{ij} 的平均数。

同样 G_{ij} 和 B_{ij} 的值可被计算出来。然后构建由这些块的镶嵌组成的背景，并给出真正背景值的较优近似。

重复执行第 2 阶段，直至步骤 67 可以确定已满足终止规则。例如，终止规则可能是一个预先定义好的数字，如 5 到 10 之间，来进行循环。当满足终止规则时，该阶段的计数器 18 递增，而且步骤 64 和 65 将导致执行第三阶段的 53 步。

背景图像的按组更新是对第 1 阶段使用的完全一致的图像的改进。然而，它可以进一步被改进，特别在前景对象的边界区域，或在背景有十分明显的颜

色突变。在 68 步通过改进基于像素分辨率，以使得背景图像尽可能真实地反映实际背景。

如将果一个像素分类为背景像素，则它的值可直接应用来更新背景图像。在将像素分类为前景像素的位置，背景图像相应的像素就由第二阶段 52 相应的背景像素值所重复。背景图像和当前图像可以各自描述成 $\{R_{bg}(x,y), G_{bg}(x,y), B_{bg}(x,y)\}$ 和 $\{R_l(x,y), G_l(x,y), B_l(x,y)\}$ 。

下面就是更新背景图像中红色分量的公式：

$$R_{bg}(x,y) = \begin{cases} R_{l,i} & \text{如果 } M(x,y) = 1 \\ \alpha R_{l,i} + (1-\alpha)R_{bg}(x,y) & \text{如果 } M(x,y) = 0 \end{cases}$$

其中 $0 < \alpha < 1$ ， $R_{l,i}$ 是第二阶段 52 的值。一般将 α 设置为 1/2。

10 背景图像中的绿色和蓝色分量以相同的方式进行更新。

69 步决定何时满足第三阶段 53 的终止规则。例如，第三阶段当总循环次数超出预先定义的数字时（一般为 10 到 20）终止。图像分割完成，该流程终止于步骤 70。

如果可以降低噪声水平，可进一步改进分割结果。正如前面提到的，噪声项 n_r, n_g, n_b 是像素间的独立随机变量。在此，在应用绝对值操作前，噪声项是直接差值。另一方面，前景像素的信号值倾向于在一个小的邻域互相关。对红色分量而言，例如，直接差值图像 D_r 可由下式获得：

$$D_r(x,y) = R_l(x,y) - R_b(x,y)$$

如果在绝对值操作前，应用移动窗口平均操作，将会降低噪声变动而平均信号将仍保持不变，前提是窗口尺寸较小。为简化期间，假定每一个噪声项 n_r, n_g, n_b 有个幅度值 σ 。对一个窗口尺寸为 $n \times n$ 像素的窗口，平滑图像的噪声变动被从 σ 降低为 σ/n 。典型地窗口尺寸被设置为 3×3 ，以至于各个颜色分量的噪声变动降低为 1/3。在绝对值操作时应用来产生在公式 3 定义的差值图像前，可对所有的三色分量执行该操作。

25 图 12 示出了一种不同于图 9 所示的方法，因为它作了包括降低噪声在内的修改。步骤 5 由 5a 到 5c 几个子步骤组成。步骤 5a 为对全部颜色分量或渠道直接计算差值图像，步骤 5b 执行移动窗口平均操作以降低每个单一颜色分量的差值图像噪声。步骤 5c 生成最终差值图像。

执行运算和录像图像的背景区域的一致性密切相关。背景区域的颜色变

化使得背景图像的估计很困难，由此导致需手动进行背景分割。

背景区域的一致性可以通过使用颜色转换方法进行改善。例如，可将
对数函数应用整个图像，以至于颜色变化在背景中不很重要。不幸的是，对数函数
同时也降低了整体的颜色范围，因此使得从背景区域中分离前景区更加困
5 难。

既然背景大概一致，颜色则围绕着最平常的颜色而聚集，背景的颜色范
围与前景相当小。只有这些背景颜色的范围需要进一步降低。这点直接导致
了新的基于分段线性函数的颜色转换方法。

该方法的第一步是确定背景和它们所在范围中最通常的颜色。这可以红
10 色分量为例。在第一个图像被分割后，利用背景中的所有像素构造一个直方图
 $H(R)$ 。在位置 R_0 处找到最大顶点，它代表了背景中最通常的红色分量。一旦
找到顶点后，就执行从顶部向两边的搜索，以确定颜色范围 $[R_1, R_2]$ ，在此如果
整个颜色范围是 $[0, 255]$ ，则 $0 < R_1 < R_0$ 、 $R_0 < R_2 < 255$ 。 R_1 和 R_2 的值由下列流程
确定：

15 1. 设 $R_1 = R_0$ 、 $R_2 = R_0$ 、 $S = H(R_0)$
 2. 如果 $S > \alpha$ ，（在此 $0 < \alpha < 1$ ），则停止搜索；否则执行下一步
 3. 如果 $H(R_2 + 1) > H(R_1 - 1)$ ， S 加 $H(R_2)$ 之前 R_1 加 1；否则 S 加 $H(R_1)$ 之前 R_2
 减 1，

 4. 转到第 2 步

20 一般 α 为 50%，正如在范围 $[R_1, R_2]$ 内 50%的背景像素是红色。一个分段
线性函数可被用于控制使整个函数范围为 $[0, 255]$ 。通常的形式如下：

$$R' = \begin{cases} f_1(R) & 0 < R < R_1 \\ f_0(R) & R_1 < R < R_2 \\ f_2(R) & R_2 < R < 255 \end{cases}$$

25 其中 $f_1(R)$ 、 $f_0(R)$ 和 $f_2(R)$ 是线性的、递增的函数，满足以下条件：

$$\begin{cases} f_1(0) > 0 \text{ 且 } f_1(R_1) = f_0(R_1) \\ f_0(R_0) = R_0 \\ f_2(R_2) < 255 \text{ 且 } f_2(R_2) = f_0(R_2) \end{cases}$$

下面是满足上面条件的一系列函数的例子：

30

$$\begin{cases} f_1(R) = \alpha_1(R-R_1) + [R_0 + \alpha_0(R_1-R_0)] \\ f_0(R) = R_0 + \alpha_0(R-R_0) \\ f_2(R) = \alpha_2(R-R_2) + [R_0 + \alpha_0(R_2-R_0)] \end{cases}$$

其中:

$$\begin{cases} \alpha_0 + (1-\alpha_0)R_0/R_1 < \alpha_1 < 1 \\ 0 < \alpha_0 < 1 \\ 1 < \alpha_2 < [255 - R_0 - \alpha_0(R_2 - R_0)] / [255 - R_2] \end{cases}$$

一般 $\alpha_0 = 1/2$, $\alpha_1 = \alpha_2 = 1$ 。

可将类似的尺度函数推导应用于绿色和蓝色中。这些颜色转换可在计算
10 差值图像前应用。

一旦执行完分割, 可移除静态背景并用新的背景图像替代。这个新的背景图像可能与前景对象有较大的反差, 因此在前景对象周围边界有清晰的图案。也可通过对边界像素应用平滑操作进行“柔化”。

如果一个对象像素的全部相邻像素都是前景像素, 则将它看为在对象
15 内部。可将这些毗邻像素定义为如图 13 中在 71 所示的 4 个相联系的邻接点或者是如图 13 中在 72 所示的 8 个相联系的邻接点。在对象内部的像素不需要操作。

如果有至少毗邻像素之一为背景像素, 则认为对象像素是在边界上。这个像素可由下式给出的新像素重复:

$$20 \quad I'(x, y) = \epsilon I(x, y) + (1 - \epsilon) B'(x, y)$$

其中 $B'(x, y)$ 是用户选择的新背景图像而 ϵ 是个常量, 一般等于 $1/2$

在此前描述的实施例, 估计背景图像并以 3 个不同的分辨率更新。这可能延伸到多于 3 个阶段, 每一个比前面在图 14 中所示的有较好的分辨率。这可以通过改变块尺寸来实现, 由整个图像开始, 终止于每个像素的分辨率。

25 块尺寸可能通过固定的步长减少或除以 2 以使得每一阶段的块尺寸是前一阶段的一半。如果图像有 $N \times N$ 个像素且 N 是一个由 2 的指数构成的整数, 则阶段的总数将是 $N+1$ 。

图 14 所示的方法开始于被认为有相同颜色的整个背景区域或区域的图 3 中在 51 所示的第一阶段。步骤 75 检查是否每一个阶段都执行了期望的循环次
30 数及何时完成, 步骤 76 决定最后一阶段是否被执行。如果否, 该阶段计数器

在 18 增加, 而下一阶段的背景分辨率在 74 步被决定。背景分辨率于是逐阶段的增加直至像素级分辨率在最后阶段被完成。因此, 每个中间阶段和具有一个或更多分段分辨率的图 3 中的第二阶段 52 相符合, 最后阶段和图 3 中的第三阶段 53 相符合。

- 5 至此描述的方法是针对分割单一图像的一个帧方法。然而, 该方法可以推广到去构造一次拍摄的方法。该一帧方法可应用于第一个帧, 背景图像可用于分割下一个图像。然而这些下面图像的分割可能只包含通过在像素分辨率中更新背景图像的一个阶段。第二阶段的循环次数可能被置成固定的数 L , 一般为 10。下面图像的循环次数被减为更小的数 k , 其中 $k < L$ 。例如, 循环次数 $\ln(i)$
- 10 可被赋值:

$$\ln(i) = \max\{k, L - i + 1\}$$

其中 i 为从第一个图像从 0 开始的图像数。

该技术在图 15 中说明, 它不同于通过加入步骤 80、81 和 82 的图 3 中的技术。下一图像按顺序在 80 中获得, 而且在这个图像中进行三阶段的分割。

- 15 特别地, 步骤 81 决定第一个图像, 步骤 64 和 65 确保第二个和第三个阶段 51 和 53 在阶段 51 后执行。

对后面所有的图像, 步骤 81 导致只有第三阶段 53 被执行。该步骤将继续直至检测到该序列的最后一个图像在步骤 82 中, 则终止该步。

- 在此前描述的实施例, 各阶段的终止规则是执行预先定义的循环次数。
- 20 然而, 也可使用其他的终止规则, 图 16 示出了一个不同于图 3 中表示的方法, 该方法在一个预定义的时间后终止, 即使这三阶段可能还没完成。特别地, 在每个分割 62 后, 步骤 84 检测是否预定义时间超时, 如果是, 该方法立即在 70 步终止。

- 图 17 示出了一个不同于图 16 的方法, 因为其允许用户交互。特别地,
- 25 执行完第三阶段 53 步后, 在 85 步形成了一个图像, 并在 86 步检查结果。如果结果被接受, 在 70 终止过程。然而, 如果该结果不被接受, 对该图像重复执行该方法, 可能需要进一步的用户交互以尽可能地得到一个可接受的结果。

- 该方法可在诸如图 18 所示的任何一个合适的装置中执行。程序化数据处理器包含连接到 CPU 总线 31 的中央处理器 (CPU)。把系统内存 32 连到总线
- 30 31, 并包含了所有操作数据处理器的系统软件或程序。

图像通过图像设备 33 获得，如夏普®网络图像设备，它被连到和数据总线 35 相连的数字化转换器机 34（如果图像信号不以数字格式存放）。也可将一个可选的显示设备 36 连到数据总线 35 上，以使得如果输出图像也由另一个窗口显示，输入全屏或窗口显示式。

- 5 可替换地，输入图像可来自其他的资源，并且也以计算机文件的格式存放在硬盘上或其他连到数据总线 35 上的存储设备 37 中。

这个系统包含一个帧缓冲区 38 用于存储输入图像，帧缓冲区 39 用于存储背景图像，帧缓冲区 40 用于存储二进制掩码和作为必需品的其它内存缓冲区 44 用于存储在分割过程中生成的暂时数据。

- 10 分割的前景和背景区域和二进制掩码一起被送到其他设备以用作进一步处理或被存储在诸如硬盘的其他存储设备上。

合成图像可利用分割前景区和其它用户定义的新背景图像而生成。这个新形成的图像可在设备 36 上以全屏或以窗口的形式显示，以和在单个窗口中与原始图像进行比较。

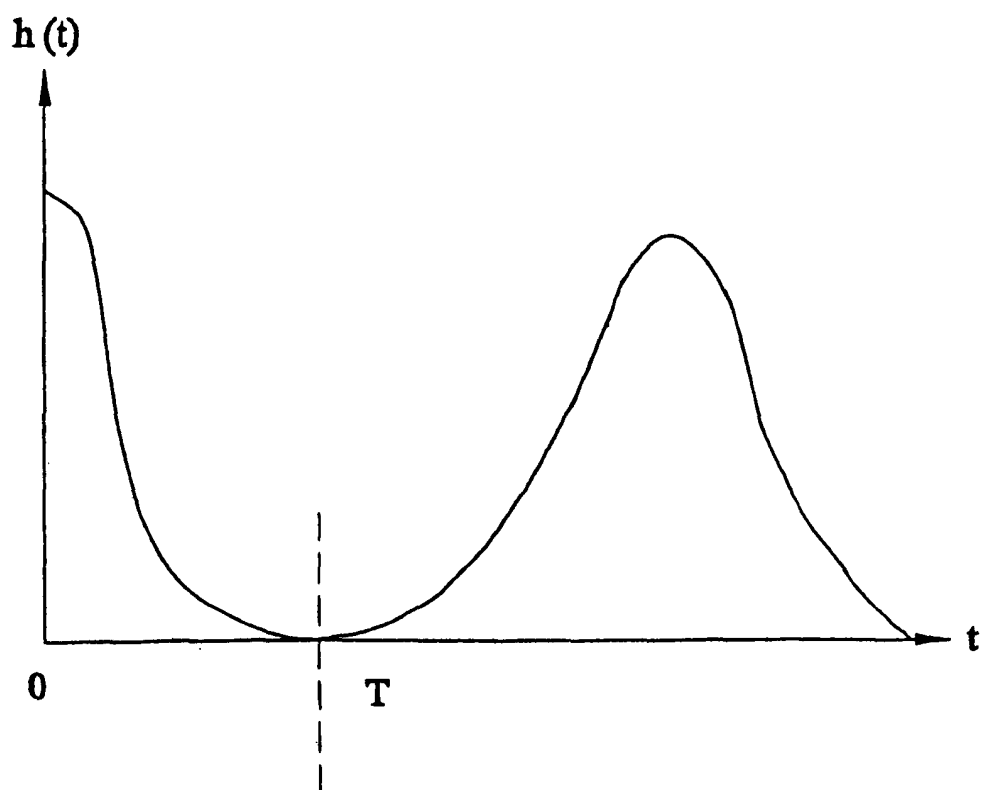


图 1

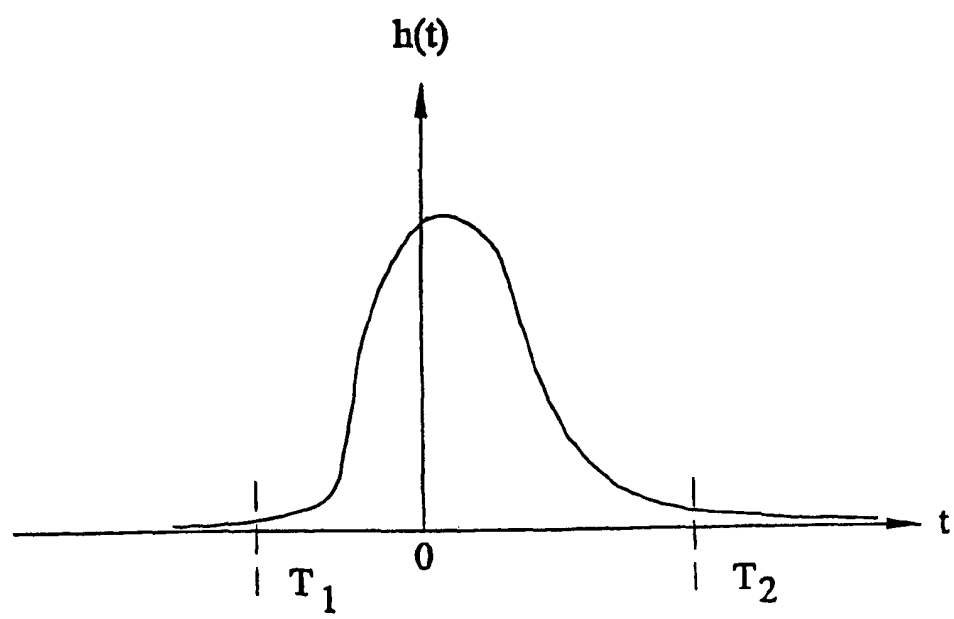


图 2

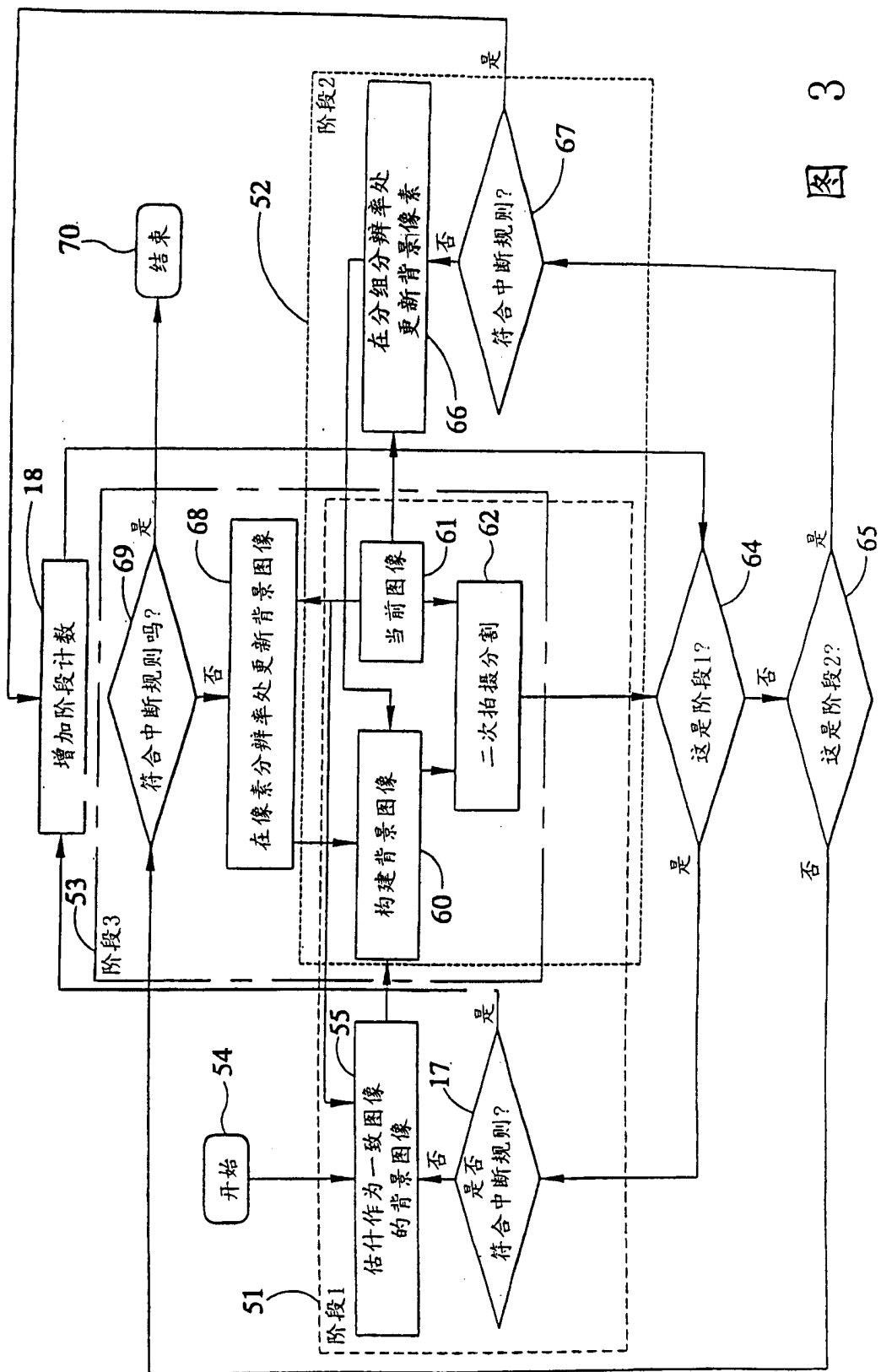


图 3

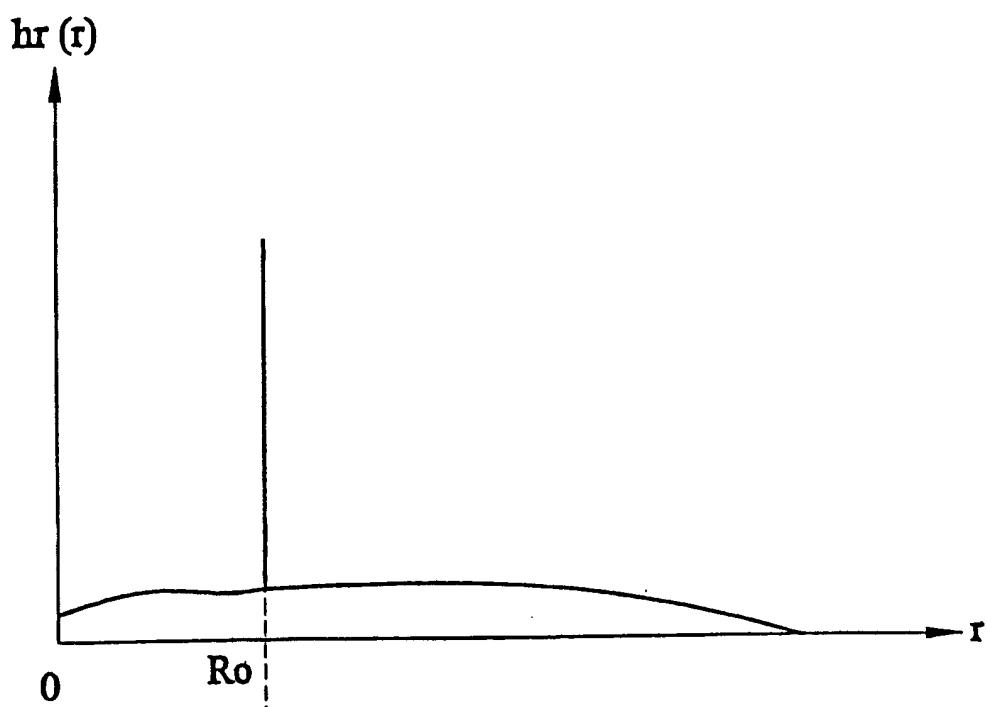


图 4

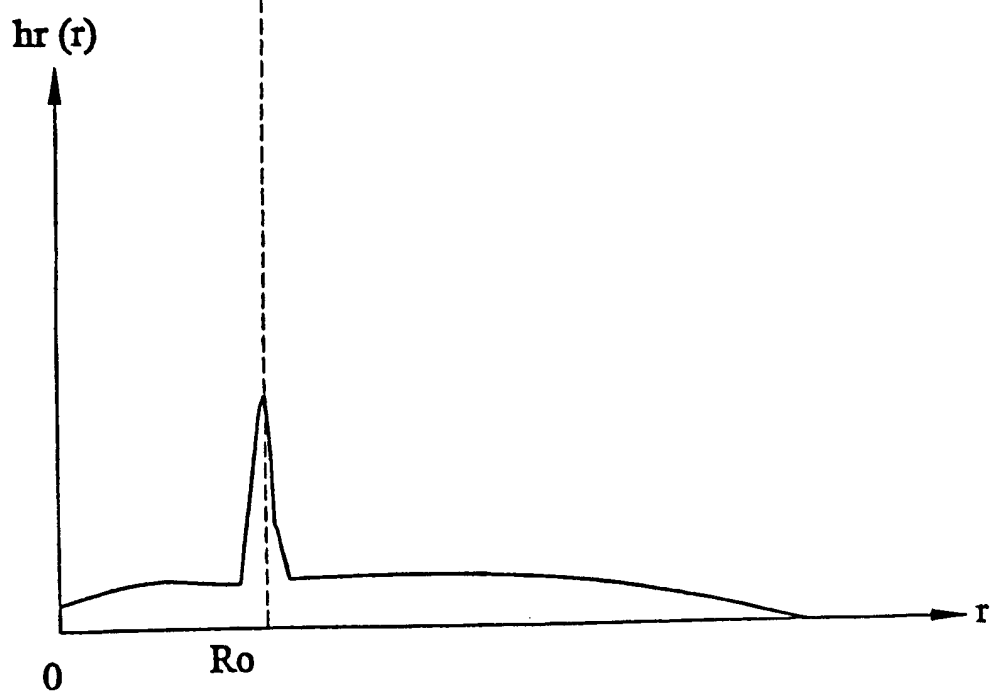


图 5

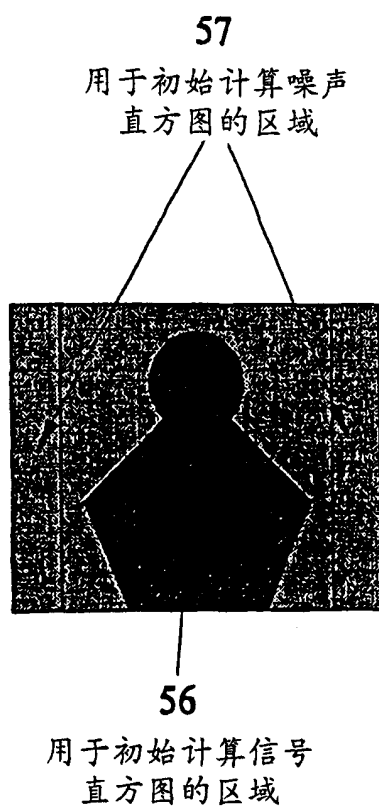


图 6

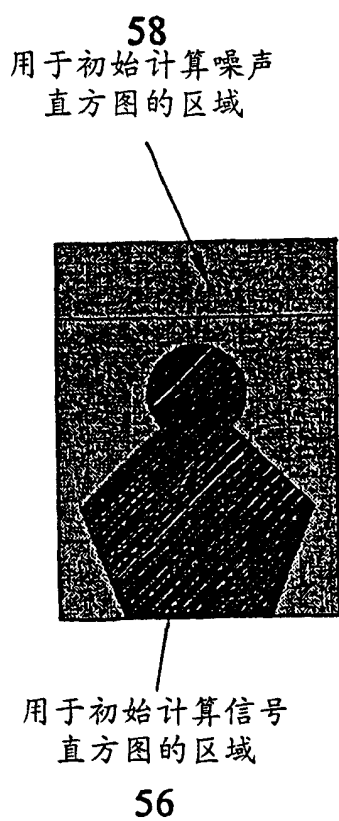


图 7



图 8

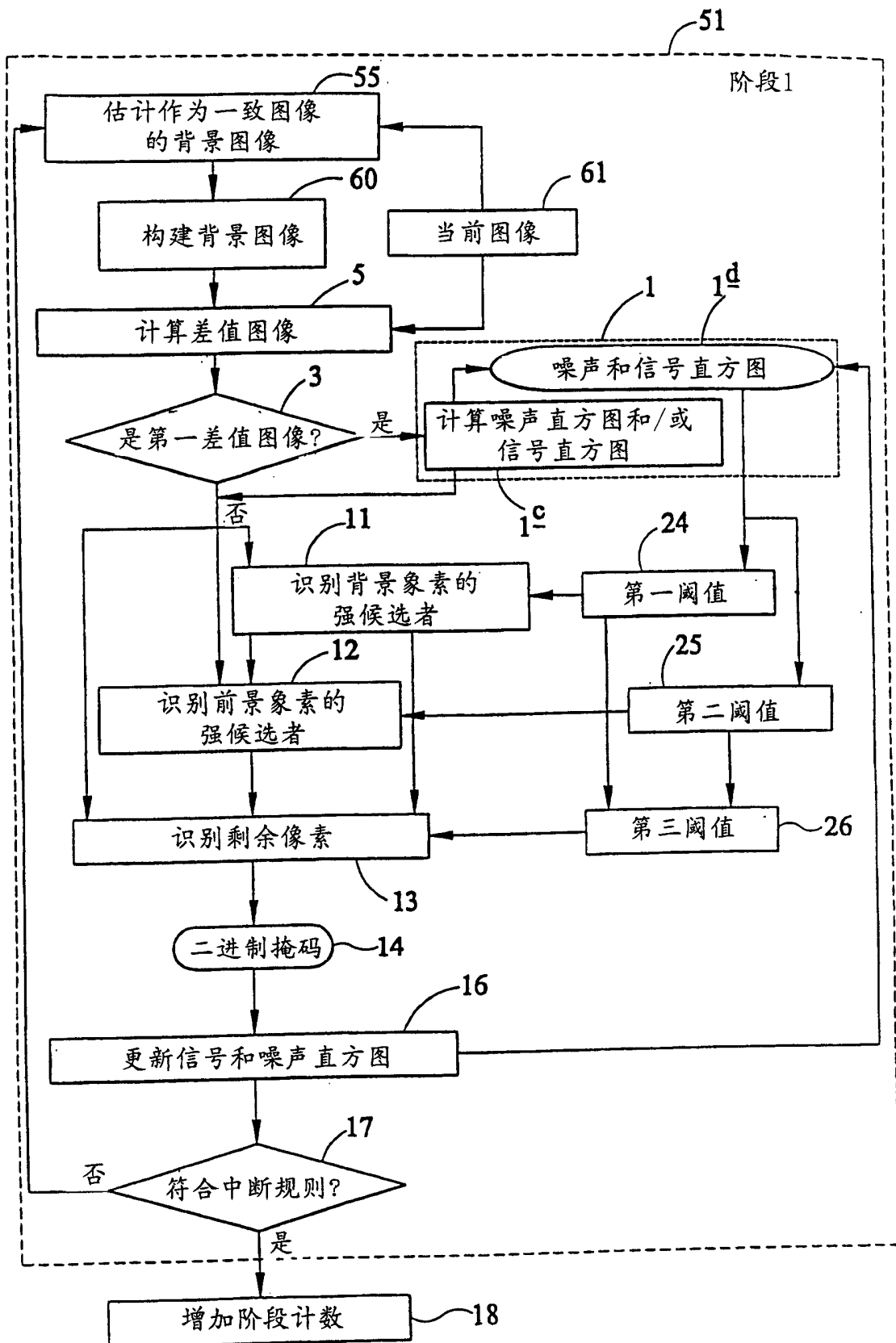


图 9

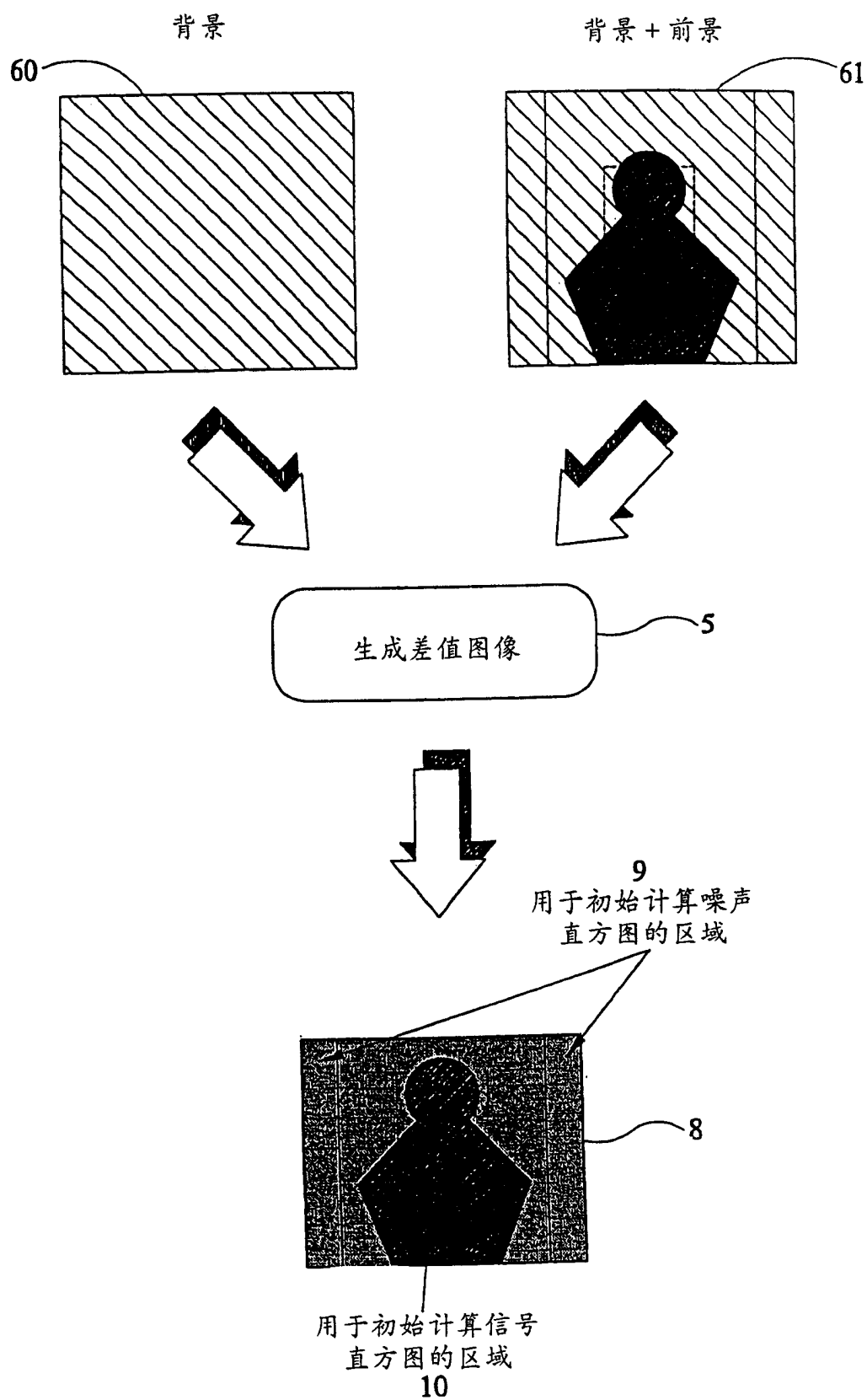


图 10

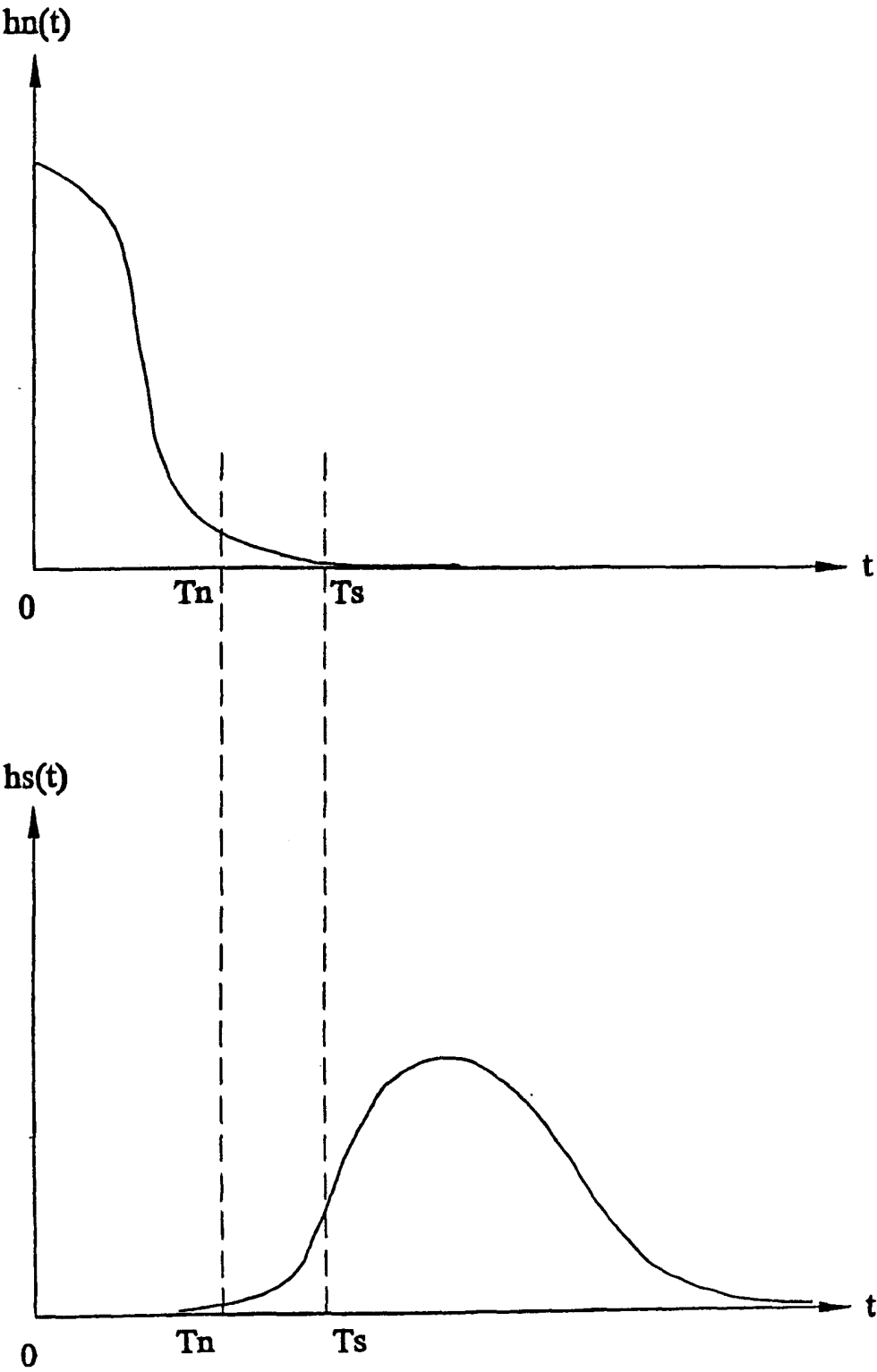


图 11

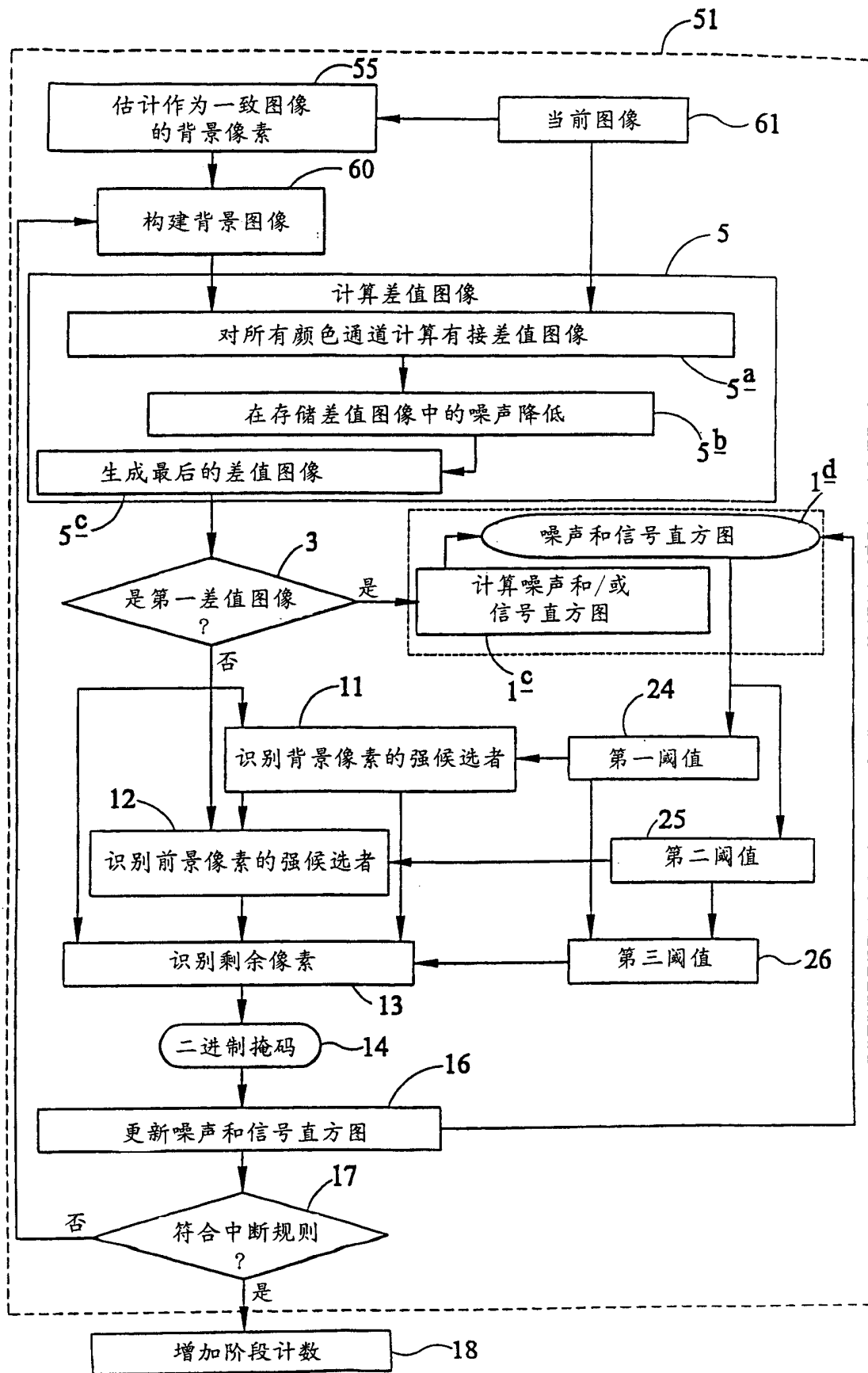
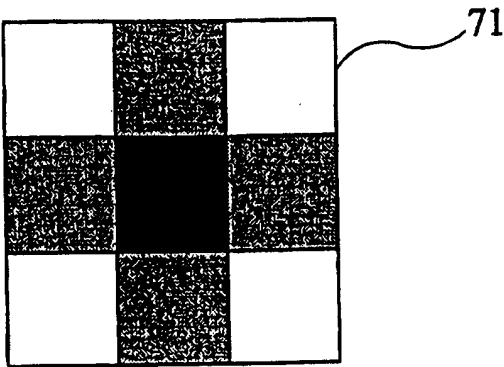
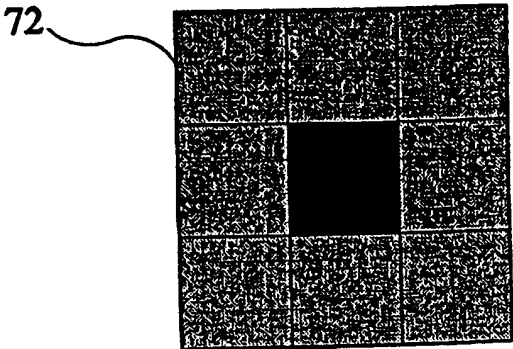


图 12



4个相邻区



8个相邻区

图 13

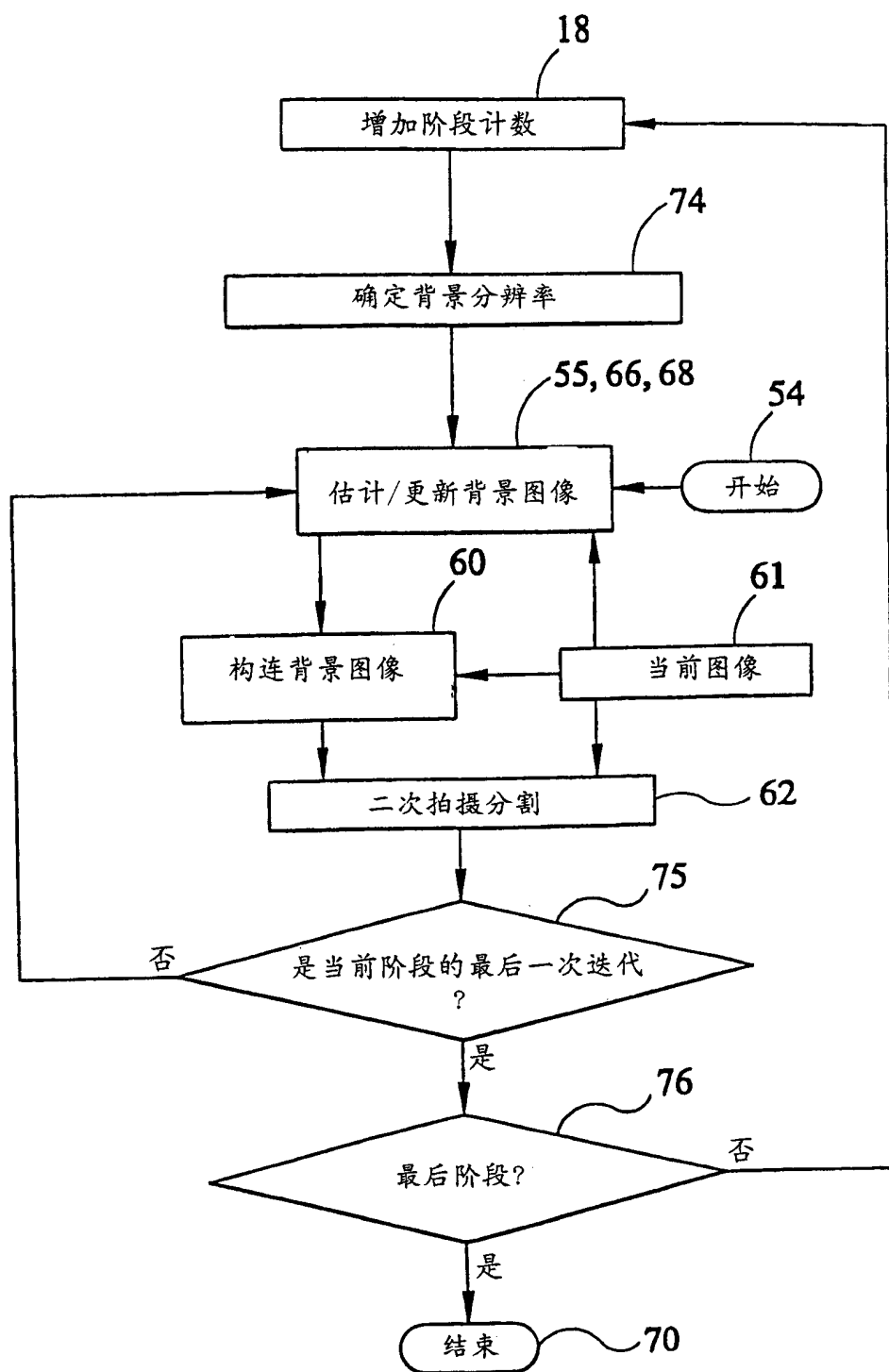


图 14

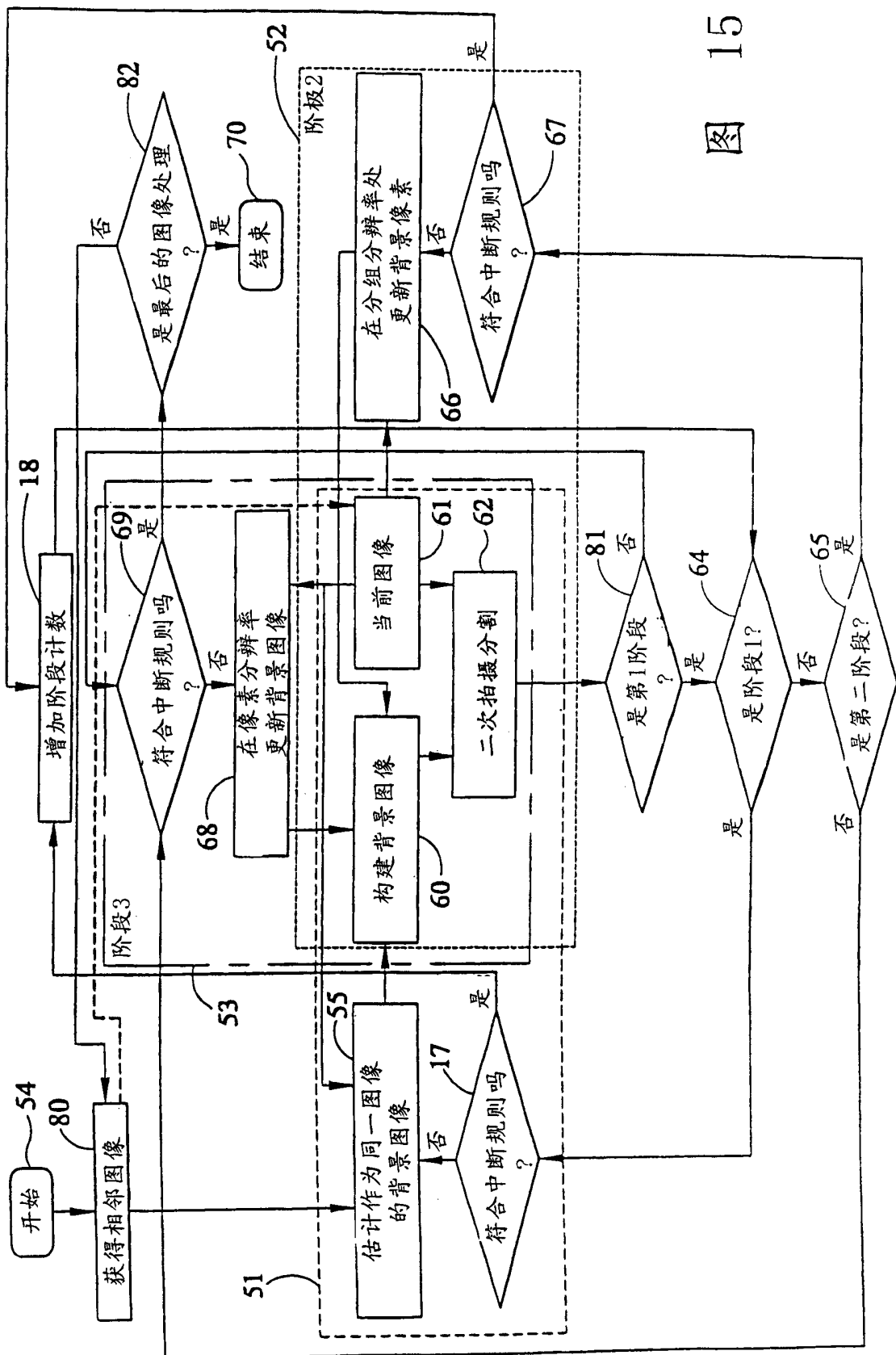


图 15

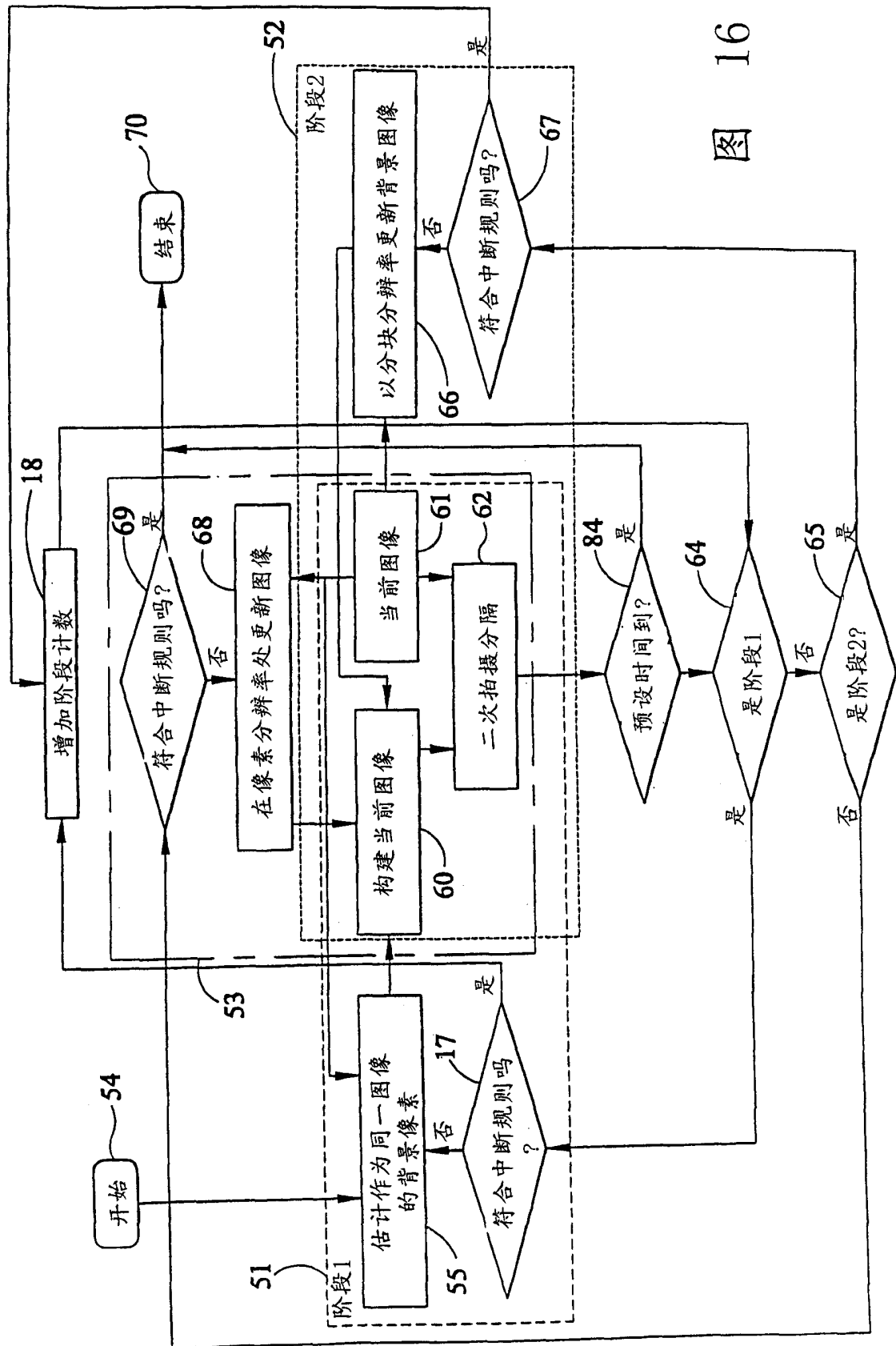
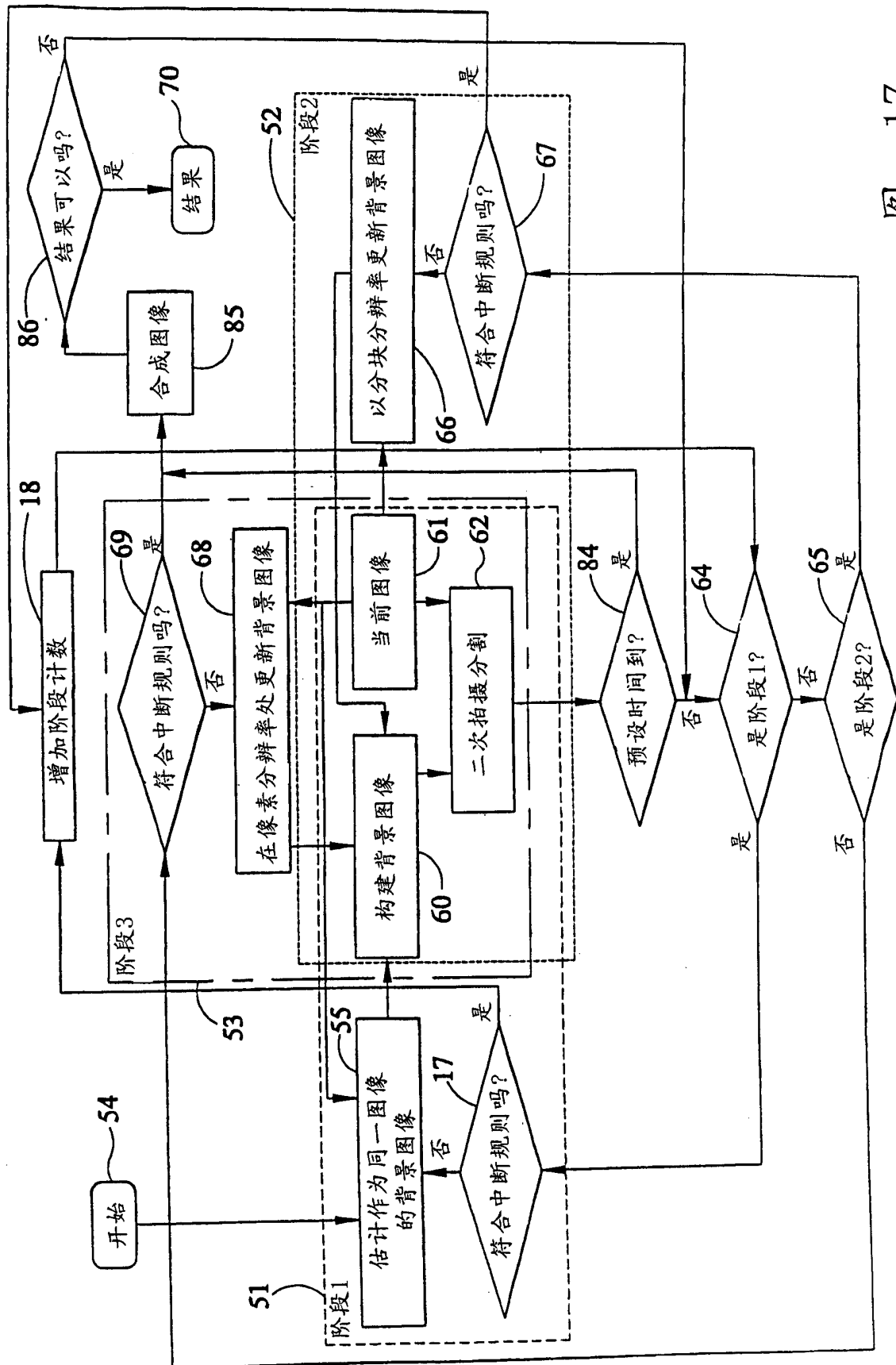


图 16



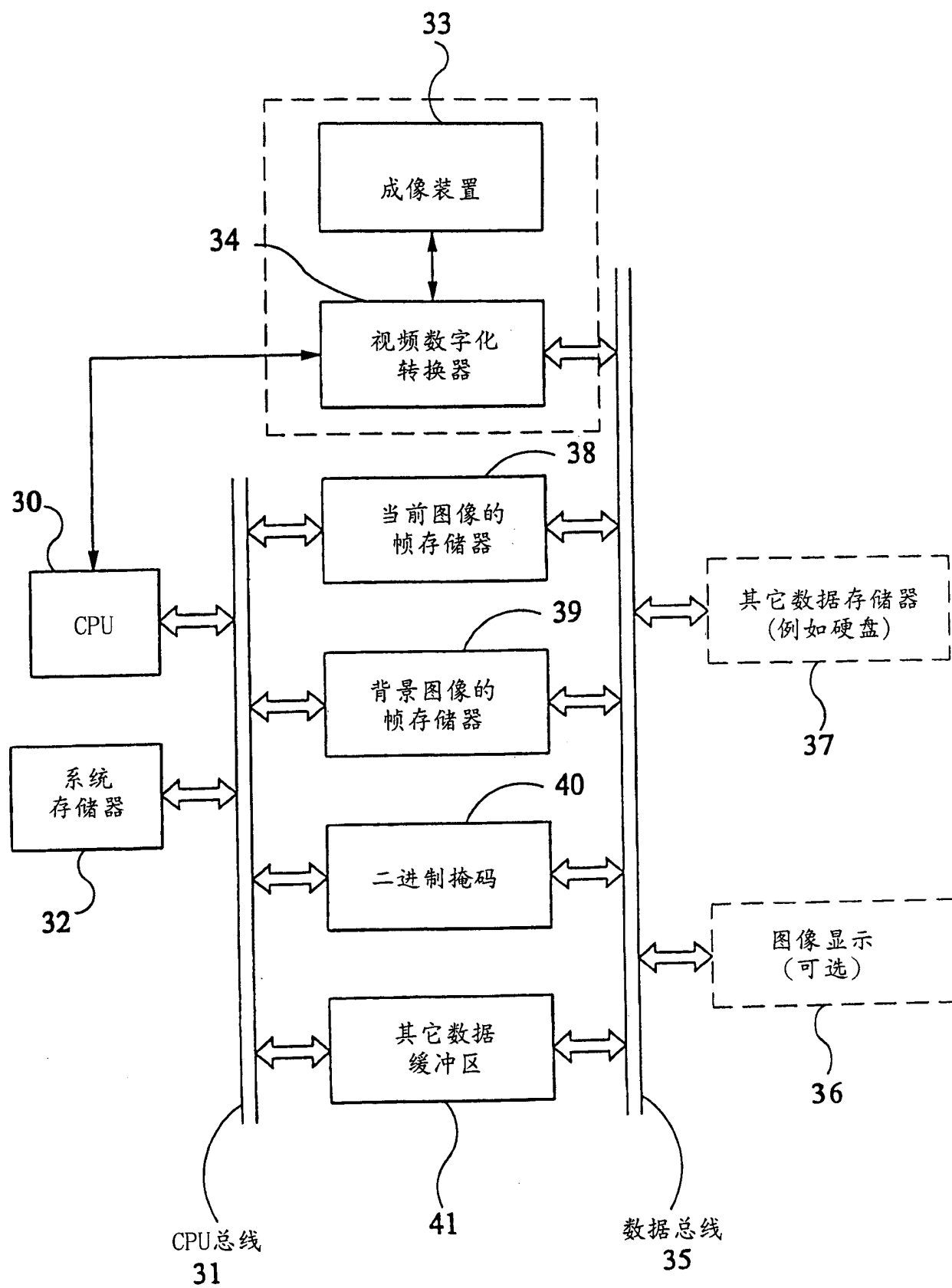


图 18