



(12) 发明专利

(10) 授权公告号 CN 107636728 B

(45) 授权公告日 2022.03.01

(21) 申请号 201680029471.2

(72) 发明人 C.瓦雷坎普

(22) 申请日 2016.05.06

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

(65) 同一申请的已公布的文献号

申请公布号 CN 107636728 A

(43) 申请公布日 2018.01.26

(51) Int.CI.

G06T 7/50 (2017.01)

G06K 9/62 (2006.01)

(30) 优先权数据

15168577.3 2015.05.21 EP

(85) PCT国际申请进入国家阶段日

2017.11.21

(86) PCT国际申请的申请数据

PCT/EP2016/060221 2016.05.06

(87) PCT国际申请的公布数据

W02016/184700 EN 2016.11.24

(56) 对比文件

US 2013156294 A1, 2013.06.20

US 2014307056 A1, 2014.10.16

US 2014055450 A1, 2014.02.27

JP 2012151851 A, 2012.08.09

CN 104395931 A, 2015.03.04

审查员 冷凝

(73) 专利权人 皇家飞利浦有限公司

权利要求书2页 说明书15页 附图5页

地址 荷兰艾恩德霍芬

(54) 发明名称

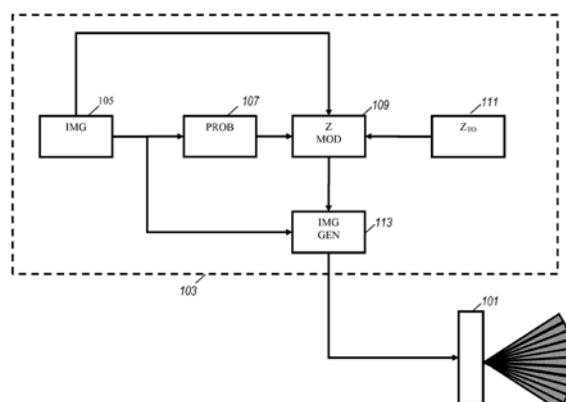
用于确定图像的深度图的方法和装置

(57) 摘要

一种用于确定图像的深度图的装置包括：提供具有相关联的深度图的图像的图像单元(105)，所述深度图包括图像的至少一些像素的深度值。概率单元(107)确定所述图像的概率图，所述概率图包括指示像素属于文本像素对象的概率的概率值。深度单元(109)生成经修改的深度图，其中，所述经修改的深度值被确定为所述输入值与文本图像对象深度值的经加权的组合，所述文本图像对象深度值与文本的优选的深度相对应。所述加权是取决于所述像素的所述概率值的。所述方法提供针对文本对象的更柔性的深度修改，其例如在使用深度图执行视图转换时产生减少的假象和降级。

B

CN 107636728



1. 一种用于确定图像的深度图的装置,所述装置包括:

图像单元(105),用于提供具有相关联的深度图的第一图像,所述深度图包括所述第一图像的至少一些像素的第一深度值;

概率单元(107),用于确定所述第一图像的概率图,所述概率图包括所述图像的至少一些像素的概率值;像素的所述概率值指示所述像素属于文本像素对象的概率;

深度单元(109),用于生成所述第一图像的经修改的深度图,所述深度单元被布置为,将至少一个第一像素的经修改的深度值确定为所述第一像素的所述相关联的深度图的第一深度值与对于多个像素而言相同的文本图像对象深度值的经加权的组合,所述加权是取决于所述第一像素的所述概率值的。

2. 根据权利要求1所述的装置,其中,所述概率单元(107)被布置为,确定所述第一图像的像素组,每个像素组包括多个像素;并且其中,所述概率单元被布置为,确定像素组的概率值。

3. 根据权利要求2所述的装置,其中,所述深度单元(109)被布置为,响应于所述第一像素所属的像素组的第一概率值和相邻的像素组的第二概率值,确定所述经加权的组合的权重;所述第一概率值和所述第二概率值的加权是取决于所述第一像素在所述第一像素所属的所述像素组中的位置的。

4. 根据权利要求1所述的装置,其中,所述概率单元(107)被布置为,确定所述图像的像素组的概率值,每个像素组包括至少一个像素;所述概率单元(107)被布置为,首先确定分类图,所述分类图包括指示所述像素组被指定为属于文本图像对象还是不属于文本图像对象的针对所述像素组的值;以及响应于对所述分类图的过滤生成所述概率图。

5. 根据权利要求4所述的装置,其中,所述分类图包括针对所述像素组的二进制值,每个二进制值指示像素组被指定为属于文本图像对象还是该像素组被指定为不属于文本图像对象。

6. 根据权利要求5所述的装置,其中,所述过滤包括被应用于所述分类图的二进制扩张过滤。

7. 根据权利要求4所述的装置,其中,所述第一图像是图像的时间序列的部分,并且所述过滤包括时间低通过滤。

8. 根据权利要求7所述的装置,其中,所述时间过滤是非对称的,以使得用于增大像素组属于所述文本图像对象的概率的时间常量是与用于降低所述像素组属于所述文本图像对象的概率的时间常量不同的。

9. 根据权利要求4所述的装置,其中,所述过滤包括空间过滤。

10. 根据权利要求9所述的装置,其中,所述空间过滤包括作为具有最大输出值限制的过滤器的softmax过滤器。

11. 根据权利要求9所述的装置,其中,所述空间过滤包括至少两个顺序的空间softmax过滤器。

12. 根据权利要求1所述的装置,其中,所述第一深度值与图像对象在与所述第一图像的场景的不同视图方向相对应的多个图像中的视差相对应。

13. 根据权利要求1所述的装置,其中,为了增大所述概率值的值,增大对所述文本图像对象深度的加权,并且降低对所述第一深度值的加权。

14.一种用于确定图像的深度图的方法,所述方法包括:

提供具有相关联的深度图的第一图像,所述深度图包括所述第一图像的至少一些像素的第一深度值;

确定所述第一图像的概率图,所述概率图包括所述图像的至少一些像素的概率值;像素的所述概率值指示所述像素属于文本像素对象的概率;以及

生成所述第一图像的经修改的深度图,所述生成包括:将至少一个第一像素的经修改的深度值确定为所述第一像素的所述相关联的深度图的第一深度值与对于多个像素而言相同的文本图像对象深度值的经加权的组合,所述加权是取决于所述第一像素的所述概率值的。

15.一种或多种计算机可读存储介质,其存储计算机可读指令,所述计算机可读指令在被执行时执行权利要求14所述的方法。

## 用于确定图像的深度图的方法和装置

### 技术领域

[0001] 本发明涉及用于确定图像的深度图的方法和装置，并且具体地说，涉及用于基于估计的视差值确定深度图的方法和装置。

### 背景技术

[0002] 三维(3D)显示器通过为查看者的两只眼睛提供被观看的场景的不同的视图为查看体验增加了第三维度。这可以通过使用户佩戴用于分隔被显示的两个视图的眼镜来达到。然而，由于这可能被认为是对于用户不方便的，所以在许多场景中优选使用自动立体显示器，自动立体显示器使用显示器处的器件(诸如柱状透镜或者挡板)来分隔视图和在不同的方向上发送它们(它们可以单个地在所述方向上到达用户的眼睛)。对于立体显示器，需要两个视图，而自动立体显示器通常需要更多的视图(诸如例如是九个视图)。

[0003] 然而，实践上的显示器倾向于不具有理想的性能，并且通常不能够呈现完美的三维图像。

[0004] 例如，基于透镜的自动立体3D显示器倾向于面临屏幕外模糊的问题。该效应是与照相机系统中的被称为景深模糊的效应类似的。

[0005] 此外，所呈现的三维图像的质量取决于所接收的图像数据的质量，并且具体地说，三维感知取决于所接收的深度信息的质量。

[0006] 三维图像信息通常是由与场景的不同的视图方向相对应的多个图像提供的。具体地说，诸如电影或者电视节目之类的视频内容日益被生成为包括一些3D信息。这样的信息可以使用专用的3D照相机来捕获，所述专用的3D照相机从稍微偏移的照相机位置捕获两个同时的图像。

[0007] 然而，在许多应用中，所提供的图像不可以直接地与期望的方向相对应，或者需要更多的图像。例如，对于自动立体显示器，需要多于两个图像，并且实际上通常使用9-26个视图图像。

[0008] 为了生成与不同的视图方向相对应的图像，可以使用视图点转换处理。这通常由视图转换算法来执行，所述视图转换算法使用单个视图方向的图像以及相关联的深度信息。然而，为了生成新的视图图像而没有显著的假象，所提供的深度信息必须是足够准确的。

[0009] 不幸的是，在许多应用和使用场景中，深度信息可能不是像期望的那样准确的。实际上，在许多场景中，深度信息是通过经由对不同的视图方向的视图图像进行比较估计和提取深度值来生成的。

[0010] 在许多应用中，使用稍微不同的位置处的两个照相机将三维场景捕获为立体图像。具体的深度值因而可以通过估计两个图像中的对应的图像对象之间的视差来生成。然而，这样的深度提取和估计是有问题的，并且倾向于产生不理想的深度值。这可以再次导致产生假象和降级的三维图像质量。

[0011] 三维图像降级和假象倾向于对于诸如例如是字幕块之类的文本图像对象是特别

显著的。并非作为场景的部分，文本图像对象倾向于分离的对象，所述分离的对象不被感知为是被集成或者嵌入在场景中的。进一步地，文本图像对象的深度变化倾向于对于观看者更可感知的。此外，在典型的应用中，文本（特别诸如是字幕）被预期通过明确定义的边缘是明锐的和清晰的。相应地，具体地说，以高图像质量呈现诸如字幕块之类的文本图像对象是高度重要的。

[0012] 因此，用于确定文本图像对象的合适深度信息的改进的方法将是有优势的，并且具体地说，允许增加的灵活性、促进的实现、降低的复杂度、改进的3D体验和/或改进的感知的图像质量的方法将是有优势的。

## 发明内容

[0013] 相应地，本发明寻求单一地或者以任何组合优选地减轻、缓和或者消除上面提到的缺点中的一个或多个缺点。

[0014] 根据本发明的一个方面，提供了一种用于确定图像的深度图的装置，所述装置包括：用于提供具有相关联的深度图的第一图像的图像单元，所述深度图包括所述第一图像的至少一些像素的第一深度值；用于确定所述第一图像的概率图的概率单元，所述概率图包括所述图像的至少一些像素的概率值；像素的所述概率值指示所述像素属于文本像素对象的概率；用于生成所述第一图像的经修改的深度图的深度单元，所述深度单元被布置为，将至少一个第一像素的经修改的深度值确定为所述第一像素的相关联的深度图的第一深度值与对于多个像素而言相同的文本图像对象深度值的经加权的组合，所述加权是取决于所述第一像素的所述概率值的。

[0015] 所述方法可以在许多实施例中允许改进的深度图被生成，所述改进的深度图在被用于图像处理时可以提供改进的质量。例如，在许多实施例和场景中，诸如字幕之类的文本图像对象的改进的呈现可以使用所述经修改的深度图来达到。在所述经修改的深度图被用于图像视图转换时和/或在自动立体显示器上呈现三维图像时，所述改进是特别显著的。所述方法可以进一步允许文本图像对象的呈现深度的转换，而减轻或者减少降级、不一致和/或假象。

[0016] 所述图可以是完全的或者部分的图。例如，所述概率图可以包括像素/像素组的仅一个子集的概率值。给定的像素的概率值可以表示对该像素属于文本图像对象（或者是其一部分）的概率或者可能性的估计。例如，概率值可以是对属性的所确定的值与对于文本图像对象所预期的值多么接近地相匹配的指示。例如，处理可以被应用于第一图像以生成值（或者值的集合）。所述像素属于文本图像对象的概率可以由所述值与预期的值多么接近地相匹配来表示。例如，处理可以被应用于所述第一图像，所述处理产生区间 $[a; b]$ 中的数字，其中，例如， $a=0$ 并且 $b=1$ 。对于属于文本图像对象的像素，所确定的属性可以被预期是 $b$ ，而对于不属于文本图像对象的像素，所述值可以被预期是 $a$ 。所确定的值因此可以直接是指示对应的像素属于或者不属于文本图像对象的概率的属性值。例如，对于 $a=0$ 和 $b=1$ ，值越高则更有可能对应的像素属于文本图像对象。取决于单个实施例的偏好和要求，对所述值的精确的处理和确定在不同的实施例中可以是不同的。

[0017] 文本图像对象可以是包括文本的所述图像的区/区域/子集或者片段。具体地说，文本图像对象可以是字幕图像对象。文本图像对象可以与一个或多个字符相对应，或者可

以例如包括所述文本的边界区。

[0018] 具体地说,所述经加权的组合可以是第一深度值和文本图像对象深度值的经加权的和,其中,所述权重是概率值的函数。在一些实施例中,具体地说,经加权的组合可以是第一深度值和文本图像对象深度值的单调函数的经加权的和,其中,所述权重是概率值的函数。

[0019] 文本图像对象深度值可以是预定的值。具体地说,所述文本图像对象深度值可以是时域和/或空间域中的固定的和/或恒定的值。文本图像对象深度值可以是文本图像对象的期望的深度。文本图像对象深度值可以指示文本图像对象的优选的深度,并且可以是固定的和/或预定的值。具体地说,文本图像对象深度值可以指示文本图像对象的优选的深度,并且可以是独立于由第一图像表示的场景的深度属性的。实际上,文本图像对象通常可以是覆盖图形,所述覆盖图形不是由第一图像表示的场景的部分,并且文本图像对象深度值可以表示不是场景的部分的覆盖图形/文本图像对象的优选的深度。文本图像对象深度值是与文本图像对象相关联的,并且可以是对于多个像素或者像素组相同的。在许多实施例中,所述文本图像对象深度值是独立于第一图像的属性和/或相关联的深度图的。

[0020] 所述第一图像可以是作为与不同的查看角度相对应的多个图像的部分的图像,或者可以例如是单个的隔离的并且独立的图像(具有相关联的深度图)。在一些实施例中,第一图像可以是图像的时间序列的图像,诸如例如是来自视频序列的帧。

[0021] 根据本发明的可选的特征,概率单元被布置为,确定所述第一图像的像素组,每个像素组包括多个像素;并且其中,概率单元被布置为,确定像素组的概率值。

[0022] 这可以提供改进的性能和/或促进的操作和/或降低的复杂度和/或资源需求。可以将像素组的所确定的概率值指派给属于该像素组的全部像素。

[0023] 根据本发明的可选的特征,所述深度单元被布置为,响应于所述第一像素所属的像素组的第一概率值和相邻的像素组的第二概率值确定所述经加权的组合的权重;第一概率值和第二概率值的加权是取决于所述第一像素在第一像素所属的像素组中的位置的。

[0024] 这在许多场景中可以允许改进的性能。具体地说,所述方法可以允许在比图像像素分辨率低的分辨率处对概率值的确定,同时允许对于深度水平的组合提高有效的分辨率。

[0025] 在一些实施例中,经加权的组合的权重可以响应于经内插的概率值被确定,所述经内插的概率值是响应于所述第一像素所属的像素组的第一概率值与相邻的像素组的第二概率值之间的空间内插被确定的;所述内插是取决于第一像素在第一像素所属的像素组中的位置的。

[0026] 概率图可以包括多个概率值,并且具体地说可以包括针对多个像素或者像素组中的每个像素或者像素组的多个概率值。类似地,可以对于每个像素执行所述组合,其中,至少一些像素具有不同的概率值并且因此独立的组合。因此,所述概率值和组合可以跨图像改变(其可以是对于不同的像素不同的)。

[0027] 根据本发明的可选的特征,所述概率单元被布置为,确定所述图像的像素组的概率值,每个像素组包括至少一个像素;所述概率单元被布置为,首先确定分类图,所述分类图包括指示所述像素组被指定为属于文本图像对象还是不属于文本图像对象的针对所述像素组的值;以及响应于对所述分类图的过滤生成所述概率图。

- [0028] 这可以提供促进的实现和/或改进的性能/结果。
- [0029] 根据本发明的可选的特征，所述分类图包括针对像素组的二进制值，每个二进制值指示像素组被指定为属于文本图像对象还是该像素组被指定为不属于文本图像对象。
- [0030] 这可以提供促进的实现和/或改进的性能/结果。具体地说，其在许多场景中可以允许对像素组的更鲁棒的并且可靠的初始分类。过滤可以将二进制分类转换成非二进制概率值，所述非二进制概率值也反映像素组的时间和/或空间相邻性的特性。
- [0031] 根据本发明的可选的特征，过滤包括被应用于所述分类图的二进制扩张过滤。
- [0032] 这可以改进性能，并且具体地说可以改进所检测的与文本图像对象相对应的区域的一致性。在许多场景中，其可以减少这样的区域中的孔洞形成。
- [0033] 根据本发明的可选的特征，所述过滤包括时间过滤。
- [0034] 例如在查看由基于所述经修改的深度图的视图转换生成的图像时，这可以例如允许改进的稳定性和一致性，以及提供改进的用户体验。
- [0035] 根据本发明的可选的特征，时间过滤是非对称的。
- [0036] 这可以在许多应用和场景中提供改进的性能。
- [0037] 根据本发明的可选的特征，过滤包括空间过滤。
- [0038] 这可以在许多应用和场景中提供改进的性能。
- [0039] 根据本发明的可选的特征，空间过滤包括作为具有最大输出值限制的过滤器的softmax过滤器。
- [0040] 这可以在许多应用和场景中提供改进的性能。具体地说，其可以允许对于与文本图像对象相对应的区域生成有优势的深度简档。例如，在许多场景中，softmax过滤器可以允许与文本图像对象相对应的一致的深度区的生成，同时减少孔洞的量或者大小和/或同时在文本图像对象的边缘处提供柔性的过渡。
- [0041] softmax过滤器可以是过滤器和将输出限于最大值的限制器的级联。例如，softmax低通过滤器可以执行与低通过滤器和将低通过滤器的输出值限于最大值的限制器的级联相对应的操作。因此，softmax过滤器可以与具有最大输出值限制的过滤器相对应。
- [0042] 根据本发明的可选的特征，空间过滤包括至少两个顺序的空间softmax过滤器。
- [0043] 这可以在许多应用和场景中提供改进的性能。具体地说，其可以允许对于与文本图像对象相对应的区域生成有优势的深度简档。例如，在许多场景中，softmax过滤器可以允许与文本图像对象相对应的一致的深度区的生成。在许多实施例中，过滤器可以减少孔洞的量或者大小，同时还在所述文本图像对象的边缘处提供柔性的过渡。
- [0044] 所述两个顺序的空间softmax过滤器可以被布置为，具有不同的设计/操作参数，并且具体地说，内核尺寸、缩放因子和/或最大值可以是对于所述两个过滤器不同的。在许多实施例中，用于第一过滤器的参数可以是针对减少孔洞形成和提供提高的一致性被优化的，而第二过滤器的参数可以是为了在所述文本图像对象的边缘处提供期望的深度过渡简档被优化的。
- [0045] 根据本发明的可选的特征，第一深度值与图像对象在与第一图像的场景的不同视图方向相对应的多个图像中的视差相对应。
- [0046] 本发明可以在许多实施例中改进由视差估计生成的深度图，并且具体地说可以减轻与文本图像对象相关联的降级和假象。

[0047] 在一些实施例中,图像单元被布置为,估计图像对象在多个图像中的视差。

[0048] 根据本发明的可选的特征,为了增大概率值的值,增大对文本图像对象深度的加权,并且降低对第一深度值的加权。

[0049] 这可以在许多应用和场景中提供改进的性能。

[0050] 根据本发明的一个方面,提供了一种确定图像的深度图的方法,所述方法包括:提供具有相关联的深度图的第一图像,所述深度图包括所述第一图像的至少一些像素的第一深度值;确定所述第一图像的概率图,所述概率图包括所述图像的至少一些像素的概率值;像素的所述概率值指示所述像素属于文本像素对象的概率;以及生成所述第一图像的经修改的深度图,所述生成包括:将至少一个第一像素的经修改的深度值确定为所述第一像素的所述相关联的深度图的第一深度值与对于多个像素而言相同的文本图像对象深度值的经加权的组合,所述加权是取决于所述第一像素的所述概率值的。

[0051] 本发明的这些和其它的方面、特征和优点将是从下文中描述的(一个或多个)实施例中显而易见的,并且将参考下文中描述的(一个或多个)实施例被阐述。

## 附图说明

[0052] 将仅作为示例参考附图描述本发明的实施例,其中

[0053] 图1是对包括根据现有技术的一些实施例的装置的显示器系统的示例的图示;

[0054] 图2图示了从自动立体显示器被投影的视图图像的示例;

[0055] 图3图示了图像和相关联的所检测的文本区域的示例;

[0056] 图4图示了字幕块在深度图中的定位的示例;以及

[0057] 图5图示了输入图像和因此在根据本发明的一些实施例的装置中被生成的概率图的示例。

## 具体实施方式

[0058] 以下说明聚焦于本发明的实施例,所述实施例适用于用于确定用于在生成针对场景的不同的视图方向的图像时使用的经修改的深度图的系统,诸如例如是用于生成用于输入立体图像在自动立体显示器上的呈现的额外的图像的方法。然而,将认识到,本发明不限于该应用,而可以被应用于许多其它的应用和系统。

[0059] 图1图示了根据窗口的一些实施例的系统的示例。在该具体的示例中,与自动立体显示器101的不同的视图相对应的图像是根据输入三维图像被生成的。输入三维图像可以例如由具有相关联的深度图的单个图像表示,或者可以例如由立体图像表示,其中,可以从所述立体图像中提取相关联的深度图。在一些实施例中,图像可以是来自图像的时间序列的图像,诸如来自视频序列/信号的帧。

[0060] 通常,自动立体显示器产生视图的“圆锥体”,其中,每个圆锥体包含与场景的不同的查看角度相对应的多个视图。邻近的(或者在一些情况下被较远地放置的)视图之间的查看角度差被生成为与用户的右和左眼之间的查看角度差相对应。相应地,其左和右眼看到两个合适的视图的查看者将感知三维效果。在图2中图示了自动立体显示器在每个查看圆锥体中生成九个不同的视图的示例。

[0061] 自动立体显示器倾向于使用诸如柱状透镜或者位差栅栏/栅栏掩体之类的器件来

分隔视图和在不同的方向上发送它们,以使得它们单个地到达用户的眼睛。对于立体显示器,需要两个视图,但多数自动立体显示器通常使用更多的视图。实际上,在一些显示器中,在图像上执行视图方向的渐进的过渡,以使得图像的不同的部分可以在不同的查看方向上被投影。因此,在一些更新近的自动立体显示器中,图像区在视图方向上的更渐进的并且连续的分布可以被应用,而非自动立体显示器渲染固定数量的完整的视图。这样的自动立体显示器通常被称为提供部分视图而非完全视图。关于部分视图的更多信息可以例如在W0 2006/117707中找到。

[0062] 然而,对于多数自动立体显示器来说,通常它们需要针对相对大量的不同的视图方向生成图像信息。然而,通常,三维图像数据是作为立体图像或者作为具有深度图的图像被提供的。为了生成所需的视图方向,图像视图转换算法通常被应用,以生成用于渲染的合适的像素值。然而,这样的算法通常是欠优化的,并且可能引入假象或者失真。

[0063] 发明人已经认识到,结合特别诸如是字幕图像对象之类的文本图像对象,这样的假象、降级和质量降低可以是特别普遍、可感知和/或显著的。因此,在包含文本的图像被接收和呈现时(例如,通过自动立体显示器),可以经常在字幕块和类似的文本对象周围感知假象。

[0064] 为了解决这样的问题,可以应用寻求标识这样的图像对象的算法。因而可以将固定深度水平应用于所标识的图像对象,以便将全部文本定位在固定的深度处。然而,这样的检测通常是非常困难的,并且向文本图像对象(和非文本图像对象)中的准确的分段通常是不可行的。因此,这样的方法通常还导致可感知的假象。

[0065] 例如,可以标识字幕框,并且可以将该框的深度水平设置为屏幕深度。图3图示了可以如何完成此的一个示例。

[0066] 字幕框被用于在屏幕深度处放置字幕。图3图示了可以如何完成这一点。在该示例中,在具有字幕文本(图3a)的照度图像中检测字幕框。在该示例中,检测可以每8x8像素块地、使用从每个8x8块内提取的特征被执行。每个黑块可以首先被分类为是(在图3b中由白色示出)或者不是(在图3b中由黑色示出)字幕的部分。最后,首先通过水平地集成检测结果并且找到沿y坐标的开始和停止位置以及然后通过水平地集成检测结果并且找到x坐标的开始和停止位置来确定边界字幕框。具体地说,字幕框的边界可以被设置为包括全部被分类为属于字幕框的块的最小矩形。

[0067] 字幕框内的全部像素的深度可以然后被设置为具体的深度值,并且具体地说,可以被设置为屏幕或者显示器深度。这可以减少假象,并且可以减少由自动立体显示器向不位于屏幕水平处的对象引入的模糊。方法可以由图4来图示,图4示出了具有位于屏幕深度之后的部分401和位于屏幕深度之前的部分403的图像。所描述的方法可以标识字幕块405,并且将该字幕块定位在屏幕深度处。

[0068] 然而,尽管该解决方案将产生明锐的(具体地说,没有屏幕外模糊的)并且几何上正确的字幕对象,但深度图中的新被引入的陡峭边缘(由于文本框)可能经常导致产生文本框的边界附近的非常可见的文本失真(断裂)(诸如在图4中由407指示的)。此外,在文本框自身分裂成多个较小的文本框时,渲染错误将也变成可见的。例如,句子中的一些字符可能是明锐的,而其它字符可能是模糊的。此外,该行为可以随时间迅速地改变,这在实践中倾向于对查看者是相当可见的。

[0069] 在实践中已经发现,寻求标识字幕区并且将该区的深度设置为(具体地说)屏幕深度通常导致产生各种问题。首先,检测算法中的噪声可以频繁地使所检测的框分裂成两个或更多个单独的框。这可以由于多个开始和停止位置可能沿x和/或y坐标存在而发生。此外,所检测的框的开始和停止位置倾向于是对分类错误敏感的。因此,产生的深度图可能变得在时间上不稳定,这导致产生被自动立体显示器显示的图像中的可见的时间错误。

[0070] 图1的装置可以减轻或者减少已知的方法中已知的缺点中的一些缺点。装置包括驱动自动立体显示器101的显示器驱动器103。显示器驱动器103包括用于为自动立体显示器101生成多个视图图像以及向其馈送这些视图图像的功能。使用视图转换算法针对不同的视图方向生成视图图像,视图转换算法具有输入图像和深度图像作为输入。

[0071] 然而,并非只是标识字幕区并且将字幕区设置为屏幕深度,系统被布置为,对输入深度图(从外部源或者内部源被接收,并且通常是由显示器驱动器103根据与不同的视图方向相对应的图像生成的,具体地说,诸如立体图像)进行处理以生成经修改的深度图,其中在经修改的深度图中,使用字幕深度的更巧妙的重新布置。这将倾向于提供大致上改进的视觉体验,包括围绕字幕或者其它文本图像对象的可视的假象的减少或者减轻。

[0072] 显示器驱动器103的方法涉及首先生成也被称为 $\alpha$ 图的概率图,概率图包括指示对应的像素是否属于文本图像对象的概率值。因此,并非简单地二进制指定确定给定的像素或者像素组是否是文本图像对象,概率图提供非二进制的并且通常大致上连续的值,所述值指示像素或者像素组属于或者不属于文本图像对象的估计的概率。概率图的概率值也被称为 $\alpha$ 值,并且通常由离散的值表示,并且在多数实施例中,每个值可以由具有至少4个但通常至少8、16、32、64、128、256或者甚至更多个离散的水平的值表示。在许多实施例中,每个概率值可以由至少3、4、6、8、10或者16比特的二进制值表示。

[0073] 在所述方法中,至少一些像素的深度水平不是简单地通过选择原始深度值或者将其设置为预定的深度水平(诸如屏幕水平)来生成的。相反,经修改的深度图的至少一些深度值是通过执行原始深度值和文本图像对象深度值的经加权的组合来生成,文本图像对象深度值具体地说是文本图像对象的优选的深度水平(诸如屏幕深度)。给定像素的深度水平的加权是基于该像素的概率值被确定的。

[0074] 具体地说,如果 $\alpha[i, j]$ 指代像素位置 $[i, j]$ 处的概率图的值,经修改的深度图的经修改的深度水平可以被确定为

$$D_{\text{mod}}[i, j] = \alpha[i, j]D_{\text{text}} + (1 - \alpha[i, j])D[i, j]$$

[0076] 其中,  $D_{\text{text}}$  指代文本图像对象深度,以及  $D[i, j]$  是原始深度水平。

[0077] 因此,在所述方法中,生成经修改的深度图以提供更连续地跨越原始深度水平与对文本图像对象所期望的深度水平之间的差的深度水平。这样,文本图像对象可以有效地在原始深度呈现与期望深度之间在空间上(在深度方向上)被“弯曲”。

[0078] 在实践中已经发现这样的方法提供显著更吸引人的视觉体验,所述视觉体验具有对文本图像对象(诸如具体地说字幕块)周围的假象和降级的显著减少的感知。

[0079] 所述方法的另一个优点在于,通过使用文本图像对象检测结果来生成概率图,这可以更易进行空间和时间过滤以改进产生的渲染图像。此外,这样的过滤可以被执行而不影响原始深度图。

[0080] 显示器驱动器103包括图像单元105,图像单元105被布置为,提供具有相关联的深度图的图像,所述深度图包括图像的至少一些像素的深度值。图像将被称为输入图像,并且深度图被称为输入深度图(并且像素和深度值被称为输入像素和深度值)。

[0081] 图像单元105在一些实施例中可以被布置为简单地从任何合适的内部的或者外部的源接收输入图像和输入深度图。例如,可以从网络(诸如互联网)、广播信号、媒体载波等接收包括由图像和相关联的深度图表示的三维图像的视频信号。

[0082] 在一些实施例中,图像单元105可以被布置为,生成输入图像和/或输入深度图。具体地说,在许多实施例中,图像单元105可以被布置为,接收由与不同的视图方向相对应的多个图像形成三维图像(诸如,具体地说,立体图像),并且其可以被布置为,执行视差估计以生成输入深度图。

[0083] 被包括在深度图中的深度值可以是深度的任何合适的表示,诸如具体地说,深度坐标(z)或者表示不同的视图方向的图像之间的转换的视差值。

[0084] 在该示例中,输入深度值是所观察或者估计的深度值。具体地说,它们可以是通过在外部源处或者由图像单元105自身执行的视差估计被生成的值。相应地,深度值将通常是相对不准确的,并且可以包含一些错误和误差。

[0085] 此外,在输入信号中,诸如字幕之类的文本图像对象在被渲染在自动立体显示器上时可能不被定位在优选的深度处。这可能例如是由于技术上最优的深度(例如,就明锐度而言)随不同显示器而改变或者由于主观的偏好在应用之间改变而导致的。

[0086] 例如,对于电视节目和运动图片,字幕通常被放置在屏幕之前,即,使字幕出现在屏幕之前的深度水平处,并且通常是作为最前面的对象的。然而,对于许多自动立体显示器,对于这样的深度水平引入了大量增加的模糊,并且因此,将字幕放置在屏幕深度处可能是优选的。如所描述的,显示器驱动器103可以相应地执行对所估计的文本图像对象的深度水平的渐进的重新调整。

[0087] 具体地说,图像单元105被耦合到概率单元107,概率单元107被布置为,生成包括概率值的概率图,概率值指示对应的像素属于文本图像对象的概率或者可能性。

[0088] 给定的像素的概率值可以表示对其属于文本图像对象的概率或者可能性的估计。概率可以例如是对包括像素的像素组与文本图像对象的预期特性的集合多么接近地相匹配的指示。匹配越接近,像素属于文本图像对象的概率越高。

[0089] 像素属于文本图像对象的概率可以是基于关于文本图像对象仅包括作为文本字符的部分的像素的考虑的。然而,在许多实施例中,可以通常认为文本图像对象包括边界框。例如,可以认为文本图像对象包括包围的字幕框。将认识到,被认为表示文本图像对象的精确的特性和属性将取决于具体的偏好和单个实施例的要求,并且具体地说,对概率值的确定可以是适于反映被认为是文本框的优选的特性和属性的。

[0090] 例如,在一些实施例中,字幕可以例如被提供为灰色或者黑色框中的白色字符。在这样的示例中,认为文本图像对象包括整个字幕框可能是高度合期望的,即,所述概率可以被确定为指示像素属于字幕框的概率产生字幕块在期望的深度处的渐进的定位。在其它的应用中,字幕可以被简单地提供为底层图像上的白色字符。在这样的示例中,不存在任何边界框,并且概率可以反映像素属于文本字符的概率。

[0091] 在许多实施例中,文本图像对象具体地说可以是图像覆盖图像对象,并且具体地

说可以是字幕或者标题图像对象。

[0092] 将认识到,可以在不同的实施例中使用各种用于确定概率图的方法,并且将在稍后描述提供具有高性能的高度高效的操作的具体的示例。

[0093] 概率单元107被耦合到深度单元109,深度单元109被布置为,生成经修改的深度图。深度单元109被进一步耦合到图像单元105和存储器/存储装置111,文本图像对象深度值被存储在所述存储器/存储装置111中。深度单元109相应地接收输入深度图、文本图像对象深度值和概率图,并且其继续执行对输入深度值和文本图像对象深度值的经加权的组合,其中,加权是取决于概率值的。

[0094] 具体地说,如之前描述的,所述组合可以是线性的组合,诸如具体地说,输入深度值和文本图像对象深度值的经加权的和,其中,权重是取决于概率值的。然而,将认识到,可以使用其它的经加权的组合,其中,来自输入深度水平和文本图像对象深度水平的相对贡献是取决于概率值的。例如,在一些实施例中,可以使用非线性的组合。

[0095] 因此,所述组合(对于给定的像素)根据输入深度值、文本图像对象深度值和概率值(对于该像素)提供输出深度值,其中,分别来自输入深度值和文本图像对象深度值的对输出深度值的贡献取决于概率值。

[0096] 经加权的组合是使得为了增大概率值的值而增大对文本图像对象深度的加权并且降低对第一深度值的加权。因此,越有可能估计给定的像素属于文本图像对象,则输出深度值将越接近文本图像对象深度值,并且越不可能估计给定的像素属于文本图像对象,则输出深度值将越接近输入深度值。精确的关系将取决于偏好和单个实施例的要求。

[0097] 在许多实施例中,对于给定的像素,输出深度值可以是概率值的函数,其中,该函数是对于零的概率值输出像素的输入深度值并且对于一的概率输出文本图像对象深度值的函数。概率值的函数具体地说可以是单调函数。概率值的函数对于给定的像素可以将范围从0到1的概率映射到范围从像素的输入深度值到文本图像对象深度值的深度值。

[0098] 在多数实施例中,文本图像对象深度值可以是预定的值,并且具体地说可以对于全部图像和/或对于整个图像是恒定的值。然而,在一些实施例中,文本图像对象深度值可以在图像的不同的区之间改变,例如,图像的较低部分的优选的文本图像对象深度值可以与图像的较高部分的优选的文本图像对象深度值不同。

[0099] 深度单元109在该示例中被耦合到图像生成器113,图像生成器113被布置为,为自动立体显示器101生成视图图像。图像生成器113从深度单元109接收经修改的深度图,并且被进一步耦合到图像单元105(其从图像单元105接收输入图像)。图像生成器113被布置为,为自动立体显示器101生成视图图像,这是通过执行视图转换以为与由自动立体显示器101产生的不同的视图相关联的具体的视图方向生成视图图像进行的。图像生成器113被布置为,由视图转换算法基于输入图像和经修改的深度图生成这些图像。因此,视图图像将被生成为在渐进地反映它们是否被认为有可能属于文本图像对象的深度处呈现像素。因此,被查看者感知的经投影的三维图像将倾向于具有对文本图像对象的更一致的渲染(具有文本图像对象与其它图像区之间的更渐进的过渡)。这通常将大量减少3D呈现的被感知的不完美。

[0100] 因此,在该示例中,输入图像和经修改的深度图被图像生成器113用于生成属于输入图像的场景但具有与输入图像的视图方向不同的视图方向的图像。

[0101] 应当认识到,技术人员将知道许多不同的视图转换算法,并且可以使用任何合适的算法,而不减损本发明。

[0102] 在一些实施例中,图像单元105可以直接地随输入图像一起接收深度图。在一些情况下,深度图可以在与文本图像对象被包括或者添加到例如所捕获的图像中所在的相同的时间和位置处被生成。在一些情况下,深度图因此可以被生成为具有位于具体的优选深度处的针对诸如字幕之类的文本图像对象的深度水平。因此,对于文本图像对象来说,深度图可以具有与文本图像对象的区域相对应的完整的并且一致的区,并且该区中的深度值可以全部是相同的。这可以允许几乎没有误差和假象地放置文本图像对象。然而,甚至在这种情况下,文本图像对象的深度水平可能不是具体的使用场景的优选深度水平。例如,适于使用基于眼镜的方法进行的查看的深度水平对于使用自动立体显示器进行的呈现可能不是理想的。进一步地,在执行视图转换时,明锐的深度过渡可以仍然导致产生假象。

[0103] 此外,在许多应用中,深度图可以不与包括文本图像对象(诸如字幕)的相同的时间处被生成和被放置为包括文本图像对象。例如,在许多场景中,三维图像可以由立体图像表示,其中,字幕被包括在这两者图像中,并且字幕的深度是由这两者立体图像之间的视差控制的。这样的三维立体表示不包括任何深度图,并且因此,如果需要,这样的图可以作为处理后操作被生成。实际上,在许多场景中,深度图可以在终端用户设备中被生成。

[0104] 例如,图像单元105可以被布置为,接收与同一个场景的不同的视图方向相对应的多个图像。具体地说,可以接收包括左眼图像和右眼图像的立体图像。立体图像可以包括例如字幕之类的文本图像对象,其中,该文本图像对象的深度由这两者图像之间的视差反映。

[0105] 图像单元105可以然后被布置为,响应于视图方向图像之间的视差检测生成深度图。因此,图像单元105可以继续找到图像中的对应的图像对象,确定这些图像对象之间的相对转换/视差,以及为图像对象分配对应的深度水平。将认识到,可以使用任何合适的用于基于视差估计确定深度的算法。

[0106] 这样的视差估计可以导致产生相对准确的深度图。然而,深度图将仍然通常包括相对大量的错误,并且将通常不是完全一致的。具体地说,假象和不一致在大的并且明锐的深度过渡周围可能是普遍的,其中,大的并且明锐的深度过渡具体地说对于文本图像对象可能经常出现。

[0107] 因此,例如在执行视图转换时,直接地使用根据对不同方向的图像的视差估计生成的深度图将倾向于导致产生被感知的质量降级和假象的引入。

[0108] 然而,基于概率检测和融合文本图像对象的所确定的深度值与期望的深度值的视差的对经修改的深度图的生成倾向于导致产生大量改进的图像,并且具体地说,导致产生具有提高的一致性和减少的假象的文本图像对象周围的被大量改进的感知的质量。具体地说,在执行视差估计时通常可能在文本图像对象周围出现的错误、不一致或者假象的出现可以被所描述的方法大量减轻。

[0109] 在下面,将描述用于确定概率图的一种具体的方法。

[0110] 在该方法中,可以对可以包括多个像素的像素组执行对概率值的确定。因此,尽管以下内容中描述的原理可以被应用于单个像素,但它们在该具体的示例中是基于像素组(以及具体地说,基于矩形块)被执行的。在该具体的示例中,每个像素组是8x8像素的块。

[0111] 相应地,在该示例中,概率单元107包括用于确定输入图像的像素组的功能,其中,

每个像素组包括多个像素。对概率值的确定因而是基于这些像素组的。

[0112] 在该方法中,概率单元107被布置为,首先确定分类图,分类图包括针对像素组的值,其中,每个值指示对应的像素组被指定为属于文本图像对象还是不属于文本图像对象。

[0113] 在该具体的所描述的实施例中,分类值是二进制值,并且每个值相应地指示对应的像素组被指定为属于文本图像对象或者该像素组被指定为不属于文本图像对象。该方法在许多实施例中可以促进分类过程,并且可以生成鲁棒的决策。此外,其可以促进处理,并且已经发现其导致产生高度适于下面的不同深度水平的混合或者组合的概率图。实际上,已经发现其导致被感知具有高质量的三维图像的生成。

[0114] 然而,将认识到,在其它实施例中,分类图的值可以是非二进制的值。例如,对象分类过程可以生成用于指示像素组是否属于文本图像对象的软决策值。例如,所述值可以是区间[0;1]中的离散值的集合,其中,1表示指定为属于文本图像对象,以及0表示指定为不属于文本图像对象,并且其中,其间的值反映像素组被认为多么接近地与被指定为属于文本图像对象的要求相匹配。实际上,所述值在一些实施例中可以被看作初始概率值,并且分类图被看作初始概率图。

[0115] 概率单元107可以对输入图像应用文本图像对象分类过程,所述文本图像对象分类过程提供指示每个像素组是否属于文本图像对象的二进制决策。

[0116] 将认识到,可以使用不同的算法来将像素组指定为属于或者不属于文本图像对象。作为一个简单的示例,像素组内的颜色分布可以被评估,并且如果其主导地包括与那些被用于字幕框的像素颜色(例如,白色和黑色)相对应的像素颜色,则像素组可以被指定为是文本图像对象,以及否则其可以被指定为不是文本图像对象。

[0117] 将认识到,在多数实施例中,可以使用复杂得多的分类或者指定。例如,分类可以是基于针对每个框被计算的多个特征的。特征可以例如是平均水平像素梯度和像素强度直方图的特定箱体内的像素的数量。诸如AdaBoost (<http://cseweb.ucsd.edu/~yfreund/paper/IntroToBoosting.pdf>; 日期:2015-05-20) 之类的机器学习方法因而可以被用于(自动地)通过线性地组合例如50个“弱”分类规则来训练所谓的“强分类器”。应当指出,通常,候选特征的大得多的集合(例如,>300)被用于从其中进行选择,以使得可以作出合适的选择。为获得候选特征的这样的大的集合,诸如平均水平像素梯度这样的基础特征各自被用作具有不同的内核大小的多个空间卷积(过滤器)操作的输入,并且因此可以各自生成例如10个新的特征候选项。对分类器的训练是使用训练图像的给定的集合在8x8块级别进行的。产生的算法然后可以被用于将每个块分类为属于或者不属于文本图像对象。

[0118] 如已提到的,所生成的分类值在所描述的示例中是指示像素组被指定为属于文本图像对象还是其被指定为不属于文本图像对象的二进制值。然而,在一些实施例中,分类过程可以生成软决策值,并且这些软决策值可以取代二进制值而被使用。例如,AdaBoost分类器可以在内部使用并且生成软决策指示,将所述软决策指示与阈值进行比较以指定像素组。然而,在一些实施例中,这些软决策值可以作为代替被用于分类图,而没有用于生成二进制值的与阈值进行的任何比较。

[0119] 用于生成概率图的第二阶段将是继续对所生成的分类图应用过滤。过滤可以对分类图强加时间和空间约束和平滑化。

[0120] 具体地说,在分类过程返回二进制应决策分类结果时,这些初始二进制值可以被

转换成指示单个像素组(或者像素)属于或者不属于文本图像对象的概率的非二进制值。因此,初始二进制分类/指定可以通过过滤被应用于分类图而被转换成渐进的非二进制概率值。该转换不仅允许具有相对低的复杂度和资源要求的高效的处理,而还已经被发现为反映不仅组的分类自身而还反映该分类与其它分类在像素组的时间和/或空间相邻性上如何相关的概率值提供非常好的结果。实际上,所述方法倾向于导致生成具有高的一致性程度并且高度适于对文本图像对象的深度的操纵的概率图。

[0121] 过滤可以包括多个过滤操作。在许多实施例中,过滤可以包括多个顺序的过滤操作,其中,通常随后的过滤操作根据前一个过滤操作的结果被执行。

[0122] 在下面,将描述使多个顺序的过滤操作被执行的具体的方法的示例。然而,将认识到,这仅是一个示例,并且在不同的实施例中,可以使用所描述的过滤器操作的仅一个子集,并且过滤器操作的次序在其它的实施例中可以是不同的。例如,在一些实施例中,时间过滤可以在空间过滤之前被应用,而在其它的实施例中,空间过滤可以在时间过滤之前被应用。

[0123] 在该示例中,过滤器操作以像素组分辨率被执行,其中,每个像素组包括多个像素。具体地说,全部过滤器以8x8的块分辨率操作。在说明中,索引[m,n]将相应地指块索引。产生的概率图将也是处在块分辨率处的。在生成经修改的深度图期间,可以对概率值进行内插以提供更高的分辨率,并且具体地说以提供像素级分辨率。作为一个具体的示例,可以对 $\alpha[m,n]$ 进行(线性地)内插以计算处在像素级处的 $\alpha[i,j]$ 。

[0124] 在一些实施例中,过滤可以包括被应用于分类图的二进制扩张过滤。

[0125] 例如, $\alpha_{init}[m,n]$ 可以指代块索引[m,n]处的原始检测结果,即,分类图的二进制值。原始检测结果取1(属于文本图像对象)或者0(不属于文本图像对象)的值。为了提高结果的密度(即,增大被检测的区域的大小),可以首先使用(通常是二进制的)扩张过滤器(具体地说,诸如具有“十字形”的形态扩张过滤器)对 $\alpha_{init}[m,n]$ 进行过滤:

$$[0126] \quad \alpha_{dil}[m,n] = \max(\alpha_{init}[m,n], \alpha_{init}[m+1,n], \alpha_{init}[m,n+1], \alpha_{init}[m-1,n], \alpha_{init}[m,n-1])$$

[0127] 这样,增加了被检测为属于文本图像对象的区域,因此提高文本图像对象像素组被包括的可能性,并且提供这些文本图像对象像素组之间的增加的重叠度等。

[0128] 概率单元107还可以应用低通时间过滤。如果扩张过滤被执行,则可以对该过滤的结果执行时间过滤。否则,其可以例如直接对分类图应用。

[0129] 在许多实施例中,时间过滤可以是非对称的,即,低通效果可以是在一个方向上被在另一个方向上更显著的。具体地说,用于改变以提高像素组属于文本图像对象的可能性的时间常量是低于用于改变以降低像素组属于文本图像对象的可能性的时间常量的。

[0130] 作为一个具体的示例,可以例如对扩张过滤器的输出(或者例如直接对分类图)应用时间的、非对称的递归过滤器。适于许多实施例的过滤器的一个具体的示例可以是以下这样的:

$$[0131] \quad \alpha_{\text{tmp}}^{(t)} = \begin{cases} \frac{1}{4}\alpha_{\text{tmp}}^{(t-1)} + \frac{3}{4}\alpha_{\text{dil}}^{(t)} & \text{if } \alpha_{\text{dil}}^{(t)} > \alpha_{\text{tmp}}^{(t-1)} \\ \frac{7}{8}\alpha_{\text{tmp}}^{(t-1)} + \frac{1}{8}\alpha_{\text{dil}}^{(t)} & \text{otherwise} \end{cases}$$

[0132] 其中,上标t指帧号。

[0133] 该过滤器可以确保,在文本第一次被检测时,深度将迅速被修改。然而,文本对深度图的效果将在文本被移除时仅渐进地减少。净效果是更好的时间稳定性。应当指出,  $\alpha_{\text{dil}}^{(t)}$  通常对于全部  $t = 0$  的块被设置为零。

[0134] 在一些实施例中,过滤包括空间过滤。可以直接对分类图执行(例如,在扩张过滤之后(如果这样的过滤被包括的话))或者可以例如在时间过滤之后执行该过滤。

[0135] 在一些实施例中,空间过滤可以是线性低通过滤器。然而,在许多实施例中,空间过滤具体地说可以包括至少一个softmax过滤器。

[0136] softmax过滤器可以是过滤器和将输出限于最大值的限制器的级联。例如, softmax低通过滤器可以执行与低通过滤器和将低通过滤器的输出值限于最大值的限制器的级联相对应的操作。因此,softmax过滤器可以与具有最大输出值限制的过滤器相对应。最大值可以是值1,但将认识到,这在不同的实施例之间可以不同。

[0137] 具体地说,softmax过滤器可以是这样的过滤器:其执行诸如低通过滤器之类的标准过滤器操作,但将过滤器输出与大于1的值相乘,以使得过滤器输出被偏置为较高的值,在此之后,从因此所获得的输出和预定义的最大值中取最小值。这样所定义的softmax过滤器因此由三个部件组成:(标准的)过滤器操作、乘法因子和预设的最大值。因此,softmax过滤可以包括应用具有将过滤器输出限于最大值的限制的空间(低通)过滤。通常,低通过滤可以是之后跟随将过滤器输出值缩放通常被预定的缩放因子的使用(通常是归一化的)低通过滤器的第一过滤的两阶段过程。产生的经缩放的输出因而被限于最大值。然而,缩放可以被看作总体过滤的部分,并且例如可以例如通过缩放空间过滤器的(内核)系数来实现。最大值通常是预定的。

[0138] 对softmax过滤器的使用可以提供改进的性能。具体地说,其可能倾向于生成这样的区,在所述区中,内部元素被指示为(几乎确定地)是文本图像对象,同时提供向区的边界的软的并且渐进的过渡。这样的特性是特别适于使用概率值标识例如字幕框的。

[0139] 在一些实施例中,空间过滤可以包括至少两个顺序的空间soft-max过滤器。不同的softmax过滤器可以具有不同的特性。具体地说,空间过滤可以提供两个功能。一项功能是填充与文本图像对象相对应的区,即,使所检测的区域重叠,以使得所检测的区内优选地不存在任何间隙。另一项功能是提供空间平滑化,以使得围绕所检测的区域的过渡是平滑的并且渐进的。这两者功能导致产生改进的视觉体验。

[0140] 因此,在该具体的示例中,可以应用第一softmax过滤器以“填充”所检测的与文本图像对象相对应的区。然后可以对第一softmax过滤器的结果应用第二softmax过滤器以生成平滑的过渡。

[0141] 作为一个具体的示例,在时间过滤之后,一些文本对象可以导致产生对区域属于文本图像对象的指示的非常稀疏填充的图。因此,为填充例如字幕框内的孔洞,可以应用

“soft-max”过滤器。例如,可以使用具有21个块的内核高度和35个块的内核宽度的softmax过滤器:

$$[0142] \quad \alpha_{\text{softmax}}[m, n] = \min \left( \frac{\sum_{p=m-10}^{m+10} \sum_{q=n-17}^{n+17} \alpha_{\text{tmp}}[p, q]}{K_1}, 1 \right)$$

[0143] 其中

$$[0144] \quad K_1 < \sum_{p=m-10}^{m+10} \sum_{q=n-17}^{n+17} \alpha_{\text{tmp}}[p, q]$$

[0145] 作为一个具体的示例,α可以由8比特数字表示,以使得1的值由255表示。在该示例中,通常可以使用值  $K_1 = 32$ 。这样的softmax过滤器可以在许多实施例中填充字符与字之间的孔洞。

[0146] 该第一softmax过滤器之后可以跟随第二softmax过滤器,第二softmax过滤器寻求移除与文本图像对象相对应的深度图的区与深度图的周围区域之间的硬空间边界,并且因此在文本与三维图像的剩余部分之间(在该三维图像被呈现时)提供软过渡。

[0147] 第二softmax过滤器可以与第一softmax过滤器相对应,但具有不同的参数,并且具体地说,具有更大的分母,以使得该过滤器更像低通过滤器而较不像最大值过滤器那样作出反应。在该具体的示例中,可以使用高度为11并且宽度为25像素的内核大小:

$$[0148] \quad \alpha[m, n] = \min \left( \frac{\sum_{p=m-5}^{m+5} \sum_{q=n-12}^{n+12} \alpha_{\text{softmax}}[p, q]}{K_2}, 1 \right)$$

[0149] 再次地,α可以由8比特数字表示,并且分母的合适的值可以是  $K_2 = 256$ 。

[0150] 对两个顺序的softmax过滤器的使用不仅可以导致产生改进的性能,而还可以允许促进的实现。具体地说,可以使用相同的算法或者函数,而仅参数值被改变。实际上,在许多实施例中,甚至图像数据可以被更高效地(再次)使用(诸如例如通过使用积分图像方法)。

[0151] 图5图示了所描述的用于确定概率图的示例性方法的结果的示例。图5示出了输入图像501和二进制扩张过滤和时间过滤之后的分类图503。如可以看到的,该阶段处的概率图可以是被非常稀疏地填充的(白色指示具有属于文本图像对象的高的概率的区域,以及黑色指示具有属于文本图像对象的低的概率的区域)。示出了第一softmax过滤器的输出505。如可以看到的,该过滤结果导致显著增大了的区域被认为很可能属于文本图像对象。然而,过渡是相当突然的。还示出了第二softmax过滤器的输出507。如可以看到的,该输出与第一softmax过滤器的输出相对接近地相对应,但具有大大被平滑化的过渡,导致被自动立体显示器101渲染的图像中的渐进得多的深度过渡。

[0152] 在所描述的示例中,分类图和对其的过滤/处理是使用包括多个像素的像素组(以

及具体地说,像素块)被执行的。实际上,在该示例中,使用了8x8像素的块。这将分辨率降低因子64,并且相应地可以允许显著更高效的并且较少需求的操作。此外,发明人已经认识到,可以应用该分辨率降低,同时仍然达到期望的性能和被感知的图像质量。

[0153] 然而,为了改进被感知的图像质量,经加权的组合的权重可以以比块级分辨率高的分辨率被生成,并且具体地说,可以以像素分辨率被生成。

[0154] 这可以通过使用基于块的深度值之间的内插来达到。例如,可以使用相邻的块的概率值之间的内插来生成块的单个像素的深度值。内插的权重可以基于像素在像素组中的位置来确定。

[0155] 例如,可以执行第一与第二深度值之间的内插。对于大致上位于针对第一和第二深度值的像素组之间的边界上的像素,可以通过大致上相等地对第一和第二深度值进行加权来确定深度值。例如,可以执行简单的求平均。然而,对于位于一个像素组的中心处的像素,深度值可以被简单地确定为已为该像素组确定的深度值。

[0156] 应当认识到,以上为了澄清所作出的说明已经参考不同的功能电路、单元和处理器描述了本发明的实施例。然而,应当显而易见,在不同功能电路、单元或者处理器之间的功能的任何适当分布可以在不偏离本发明的情况下使用。例如,被图示为由单独处理器或者控制器执行的功能可以由相同处理器或者控制器执行。因此,对具体功能单元或者电路的引用仅仅看做是对用于提供所描述的功能的适当器件的引用,而不只是严格的逻辑或者物理结构或组织。

[0157] 本发明可以任何适当的形式实现,所述形式包括硬件、软件、固件或者这些的任何组合。本发明可选地可以至少部分地实现为在一个或者多个数据处理器和/或数字信号处理器上运行的计算机软件。本发明的实施例的元件和组件可以物理地、功能地、和逻辑地以任何适当方式实现。实际上,功能可以实现在单个单元中、多个单元中、或者作为其他功能单元的一部分。这样,本发明可以被实现在单个单元中,或者可以在不同单元、电路和处理器之间物理地和功能地分布。

[0158] 虽然本发明已经连同一些实施例描述,但是其不打算被限制为本文阐述的具体形式。相反,本发明的范围仅由所附权利要求限制。此外,虽然特征可能看起来是连同特定实施例描述的,但是本领域技术人员将认识到,所描述的实施例的各种特征可以按照本发明组合。在权利要求中,术语包括不排除其他元件或者步骤的存在。

[0159] 此外,虽然单独列出,但是多个器件、元件、电路或者方法步骤可以由例如单个电路、单元或者处理器实现。此外,虽然单个特征可以被包括在不同权利要求中,但是这些可能有可能有利地组合,并且包括在不同权利要求中不暗示特征的组合是不可行的和/或有利的。将特征包括在权利要求的一个类别中也不暗示对该类别的限制,而是指示该特征酌情等同适用于其他权利要求。此外,特征在权利要求中的次序不暗示特征必须起作用的任何具体次序,并且具体的说,在方法权利要求中的单个步骤的次序不暗示步骤必须以该次序执行。相反,步骤可以以任何适当次序执行。此外,单个引用不排除多个。因此,对“一”、“一个”、“第一”、“第二”的引用不排除多个。在权利要求中的参考符号仅仅提供为对示例的阐述,不应被解读为以任何方式限制权利要求的范围。

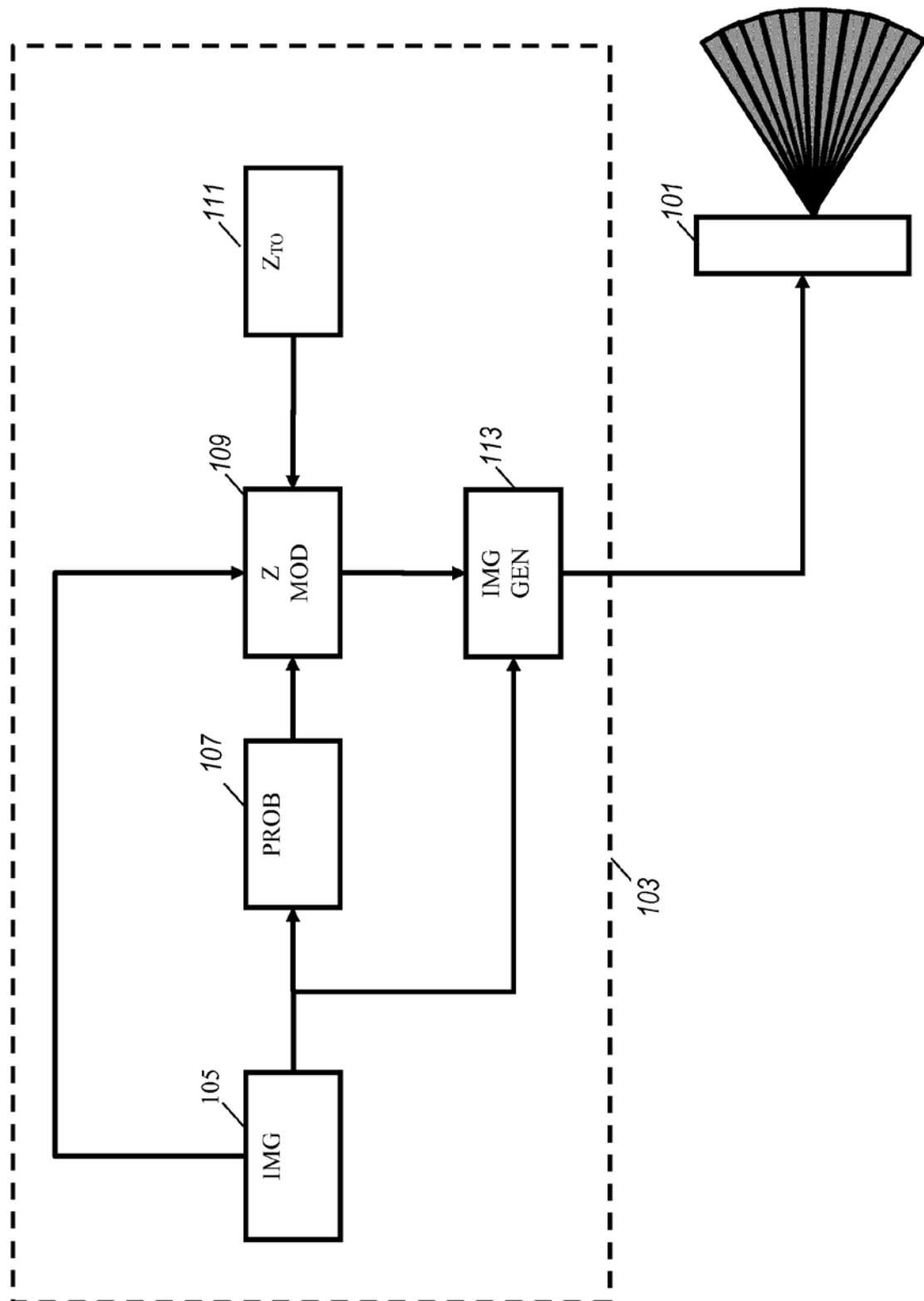


图 1

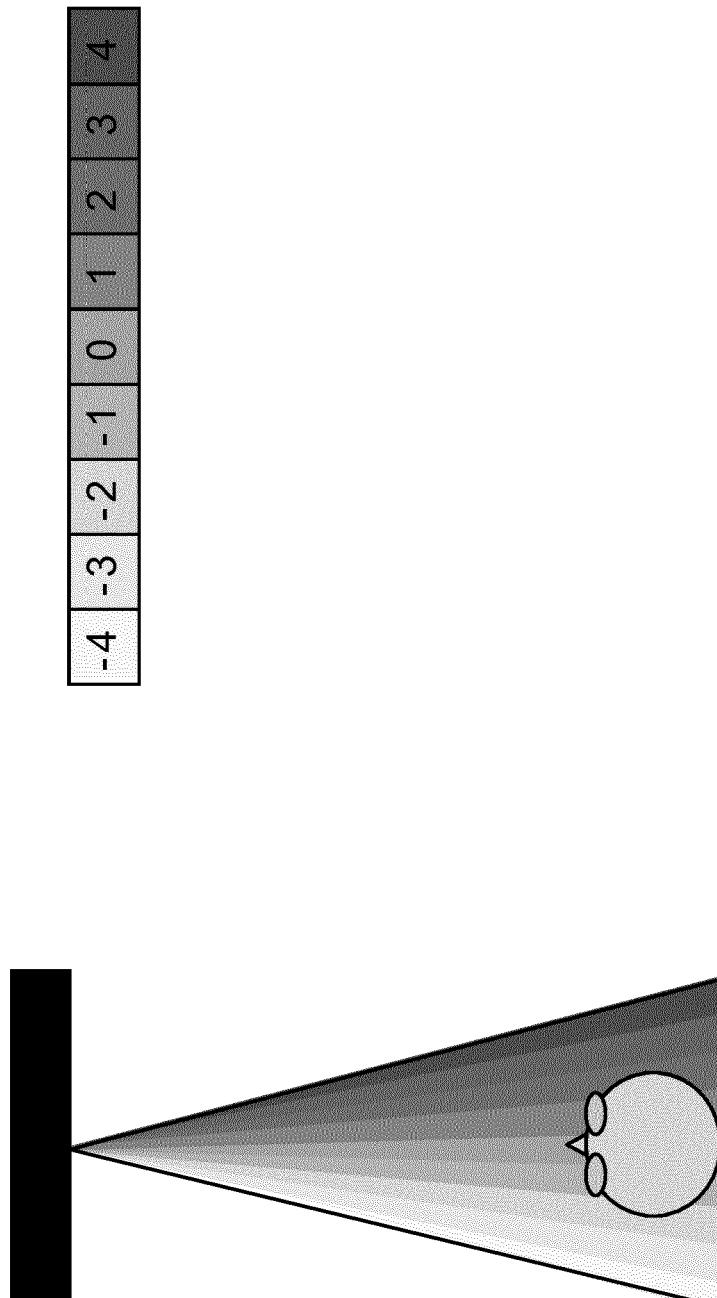


图 2

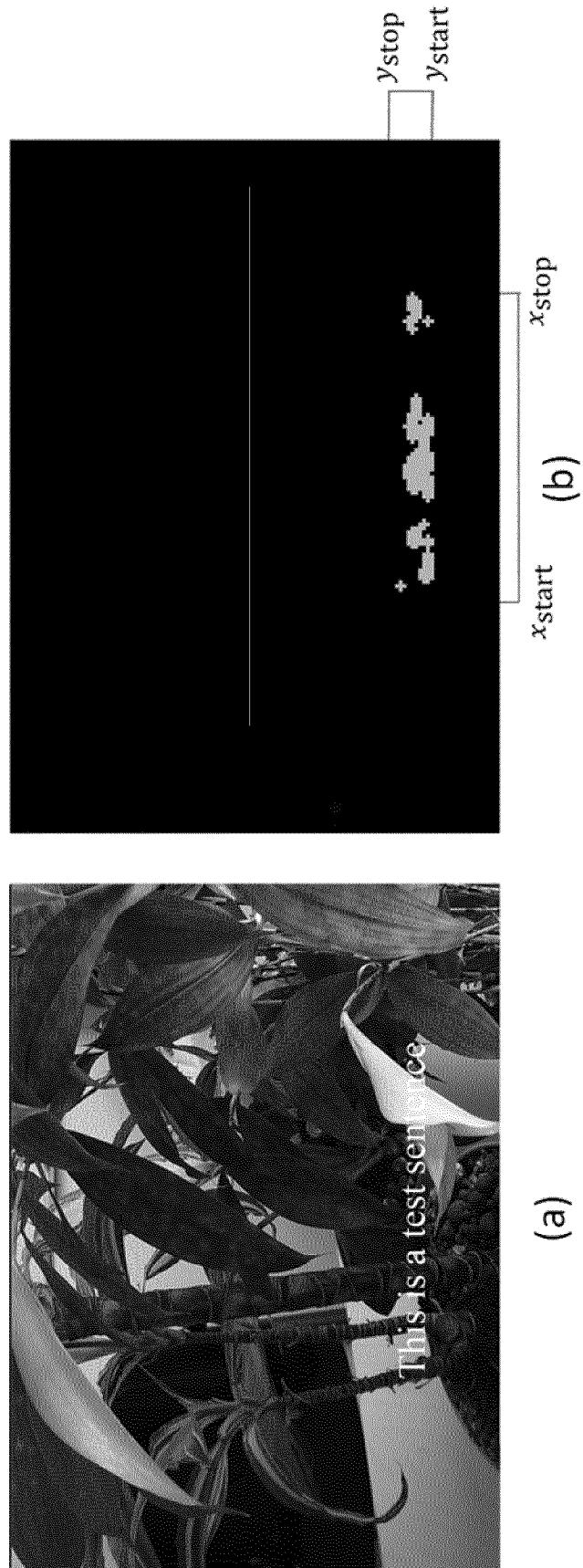


图 3

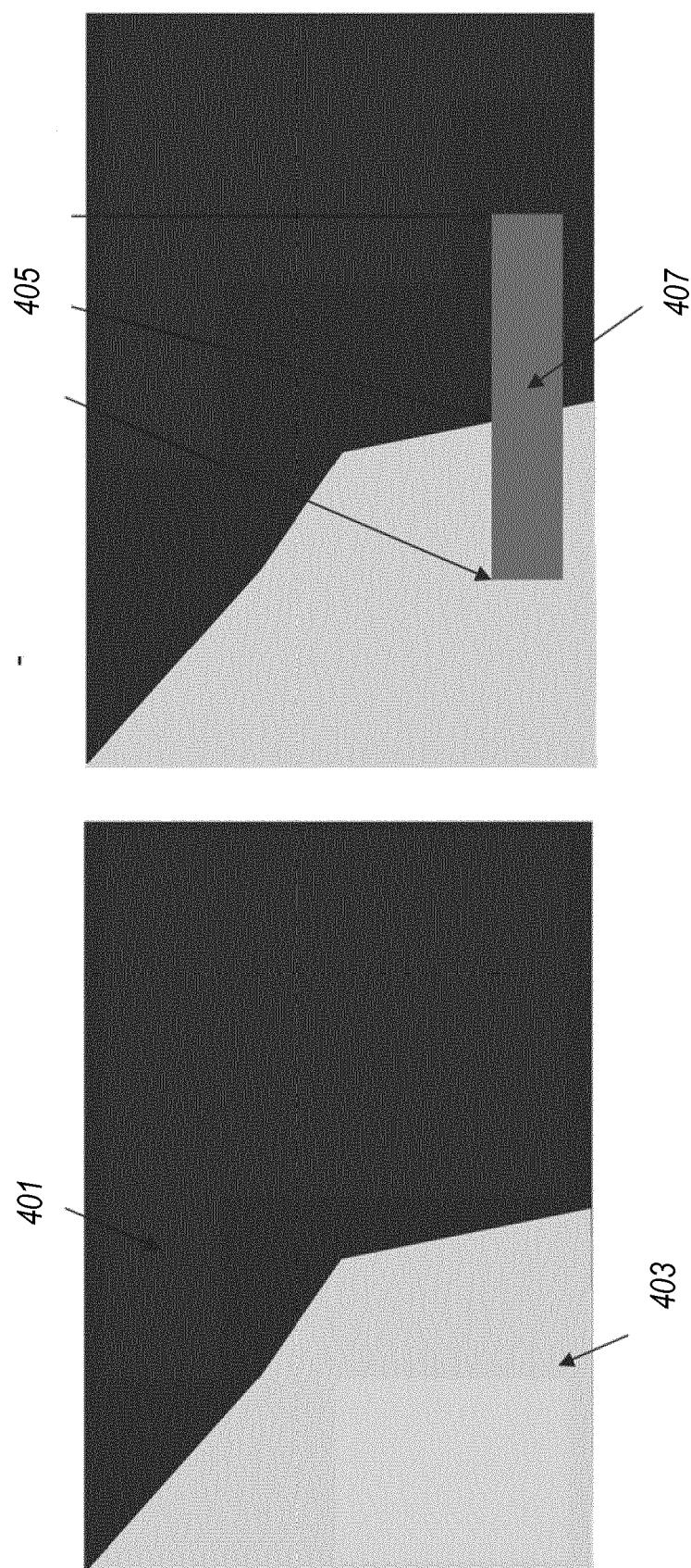


图 4

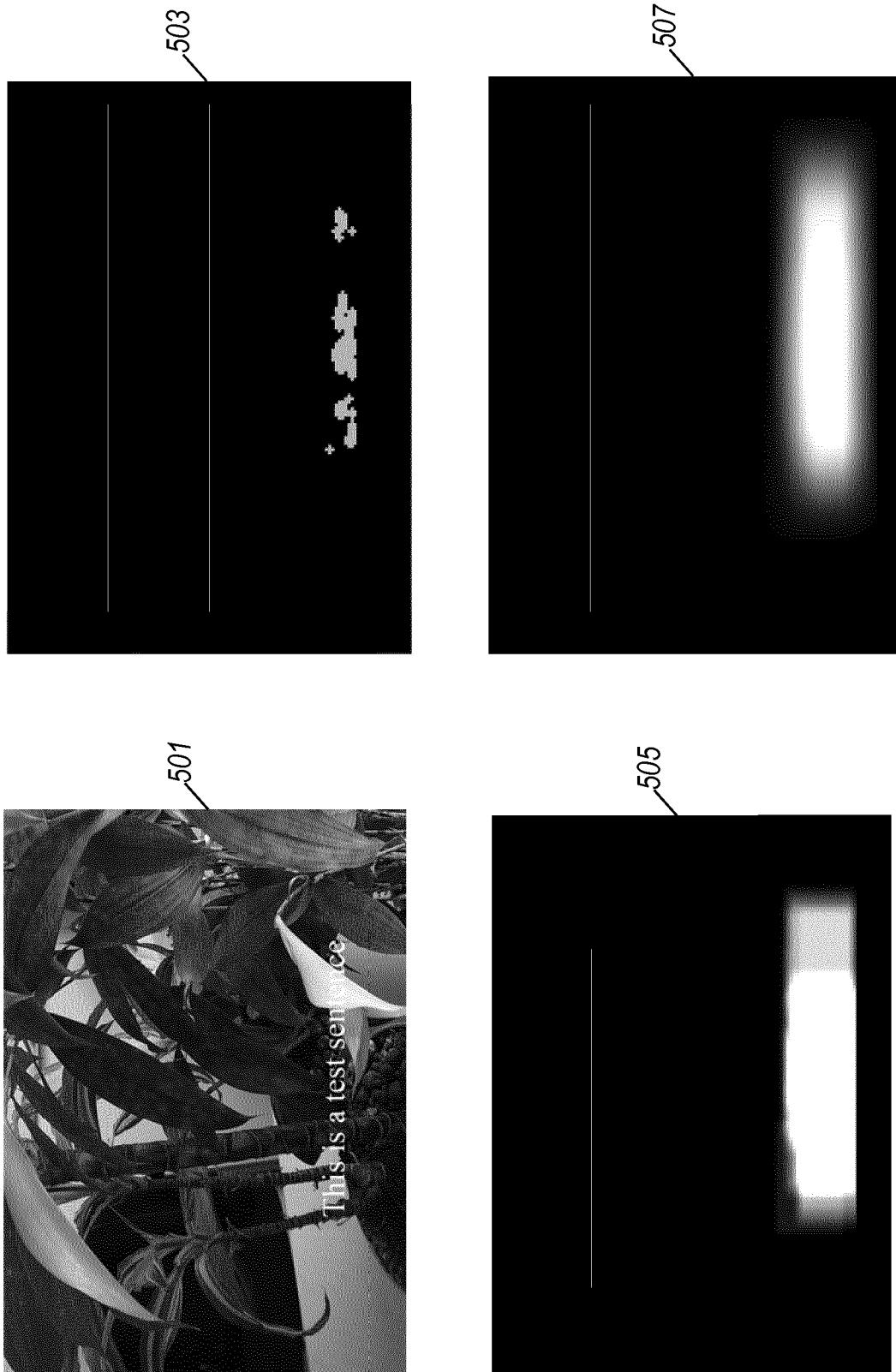


图 5