

前景图像获取方法、前景图像获取装置和电子设备

相关申请的交叉引用

本申请要求于 2019 年 7 月 19 日提交中国专利局的申请号为 2019106546426、名称为“前景图像获取方法、前景图像获取装置和电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及图像处理技术领域，具体而言，提供一种前景图像获取方法、前景图像获取装置和电子设备。

背景技术

在图像处理的一些应用中，需要进行前景图像的提取。其中，常见的一些前景图像提取技术包括帧间差分法、背景差分法、ViBe 算法等。经发明人研究发现，上述的前景图像提取技术难以准确、有效地对视频帧进行前景图像提取。

发明内容

本申请的目的在于提供一种前景图像获取方法、前景图像获取装置和电子设备，以提高计算结果的准确性和有效性。

为实现上述目的中的至少一个目的，本申请采用的技术方案如下：

本申请实施例提供了一种前景图像获取方法，包括：

对获得的当前视频帧进行帧间运动检测得到第一掩模图像；

通过神经网络模型对所述当前视频帧进行识别得到第二掩模图像；

基于预设的计算模型、所述第一掩模图像和所述第二掩模图像计算得到所述当前视频帧中的前景图像。

本申请实施例还提供了一种前景图像获取装置，包括：

第一掩模图像获取模块，被配置成对获得的当前视频帧进行帧间运动检测得到第一掩模图像；

第二掩模图像获取模块，被配置成通过神经网络模型对所述当前视频帧进行识别得到第二掩模图像；

前景图像获取模块，被配置成根据预设的计算模型、所述第一掩模图像和所述第二掩模图像计算得到所述当前视频帧中的前景图像。

本申请实施例还提供了一种电子设备，包括存储器、处理器和存储于该存储器并能够在该处理器上运行的计算机程序，该计算机程序在该处理器上运行时实现上述的前景图像获取方法。

本申请实施例还提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被运行时实现上述的前景图像获取方法。

附图说明

图 1 为本申请实施例提供的电子设备的方框示意图。

图 2 为本申请实施例提供的电子设备的应用交互示意图。

图 3 为本申请实施例提供的前景图像获取方法的流程示意图。

图 4 为图 3 中的步骤 110 的流程示意图。

图 5 为本申请实施例提供的神经网络模型的结构框图。

图 6 为本申请实施例提供的第二卷积层的结构框图。

图 7 为本申请实施例提供的第三卷积层的结构框图。

图 8 为本申请实施例提供的第四卷积层的结构框图。

图 9 为本申请实施例提供的前景图像获取方法包括的其它步骤的流程示意图。

图 10 为图 9 中的步骤 140 的流程示意图。

图 11 为本申请实施例提供的计算面积占比的效果示意图。

图 12 为本申请实施例提供的前景图像获取装置包括的功能模块的方框示意图。

图标：300-电子设备；302-存储器；304-处理器；306-前景图像获取装置；306a-第一掩模图像获取模块；306b-第二掩模图像获取模块；306c-前景图像获取模块。

具体实施方式

为使本申请实施例的目的、技术方案和效果更加清楚，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例只是本申请的一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。

因此，以下对在附图中提供的本申请的实施例的示例性描述并非旨在限制要求保护的本申请的范围，而是仅仅表示本申请的部分实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

如图1所示，本申请实施例提供了一种电子设备300，可以包括存储器302、处理器304和前景图像获取装置306。

在一些实施方式中，存储器302和处理器304之间可以直接或间接地电性连接，以实现数据的传输或交互。例如，相互之间可通过一条或多条通讯总线或信号线实现电性连接。

前景图像获取装置306可以包括至少一个可以软件或固件(firmware)的形式存储于存储器302中的软件功能模块。处理器304可以被配置成执行存储器302中存储的可执行的计算机程序，例如，前景图像获取装置306所包括的软件功能模块及计算机程序等，以实现本申请实施例提供的前景图像获取方法。

其中，存储器302可以是，但不限于，随机存取存储器(Random Access Memory, RAM)，只读存储器(Read Only Memory, ROM)，可编程只读存储器(Programmable Read-Only Memory, PROM)，可擦除只读存储器(Erasable Programmable Read-Only Memory, EPROM)，电可擦除只读存储器(Electric Erasable Programmable Read-Only Memory, EEPROM)等。

处理器304可以是一种通用处理器，包括中央处理器(Central Processing Unit, CPU)、网络处理器(Network Processor, NP)、片上系统(System on Chip, SoC)等；还可以是数字信号处理器(Digital Signal Processing, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

可以理解，图1所示的结构仅为示意，电子设备300还可包括比图1中所示更多或者更少的组件，或者具有与图1所示不同的配置，例如，电子设备300还可以包括被配置成与其它设备进行信息交互的通信单元。

其中，本申请对于电子设备300的具体类型不进行限制；例如，在一些实施方式中，电子设备300可以是具有较佳数据处理性能的终端设备，又例如，在一些实施方式中，电子设备300也可以是服务器。

在一种可以替代的示例中，电子设备300可以作为一种直播设备，例如，可以是主播直播时使用的终端设备，也可以是与主播在直播时使用的终端设备通信连接的后台服务器。

在电子设备300作为后台服务器时，如图2所示，图像采集设备可以将采集主播得到的视频帧发送至该主播的终端设备，该终端设备可以将该视频帧发送至后台服务器进行处理。

结合图3，本申请实施例还提供一种可应用于上述电子设备300的前景图像获取方法。其中，该前景图像获取方法有关的流程所定义的方法步骤可以由电子设备300实现。下面将结合图3所示的流程步骤，对本申请提供的前景图像获取方法进行示例性阐述。

步骤110，对获得的当前视频帧进行帧间运动检测得到第一掩模图像。

步骤120，通过神经网络模型对当前视频帧进行识别得到第二掩模图像。

步骤130，基于预设的计算模型、第一掩模图像和第二掩模图像计算得到当前视频帧中的前景图像。

电子设备通过上述方法，基于执行步骤 110 和步骤 120 得到的第一掩模图像和第二掩模图像，可以使得在执行步骤 130 计算前景图像时计算依据增加，从而使得计算结果的准确性和有效性得到提高，进而改善采用一些其他的前景提取方案难以准确、有效地对视频帧进行前景图像获取的情况。

经本申请的发明人研究发现，在一些应用场景中（如在获取视频帧时，如果存在灯光闪烁、镜头抖动、镜头缩放、拍摄对象静止等情况），相较于一些其他的前景图像方案，采用本申请实施例提供的前景图像获取方法可以具有一些较佳的效果。

需要说明的是，本申请对于上述的步骤 110 和步骤 120 的先后执行顺序不进行限制，例如，在一些实施方式中，电子设备可以先执行步骤 110，再执行步骤 120；或者，在另一些实施方式中，电子设备也可以先执行步骤 120，在执行步骤 110；再或者，在其他一些实施方式中，电子设备还可以同时执行步骤 110 和步骤 120。

可选地，在一些实施方式中，对于电子设备执行步骤 110 基于当前视频帧得到第一掩模图像的方式也不进行限制，可以根据实际应用需求进行选择。

例如，在一种可以替代的示例中，可以根据当前视频帧中各个像素点的像素值计算得到第一掩模图像。示例性地，结合图 4，步骤 110 可以采用以下步骤 111 和步骤 113 的方式实现：

步骤 111，根据获得的当前视频帧中各个像素点的像素值计算各个像素点在该当前视频帧中的边界信息。

在一些可能的实施方式中，电子设备在通过图像采集设备获取到采集的当前视频帧或通过连接的终端设备获取到转发的当前视频帧之后，可以对该当前视频帧进行检测，得到各个像素点的像素值。然后，在基于获取到的像素值计算各个像素点在该当前视频帧中的边界信息；其中，每一边界信息可以表征对应的像素点周围其他像素点的像素值水平。

需要说明的是，在对当前视频帧进行检测得到像素值之前，电子设备还可以先将该当前视频帧转化为灰度图。在一种可以替代的示例中，还可以根据需要对当前视频帧进行大小调整，例如，可以将当前视频帧的尺寸缩放至 256*256。

步骤 113，根据每个像素点的边界信息判断该像素点是否属于前景边界点，并根据属于前景边界点的各像素点的掩模值得到第一掩模图像。

在一些实施方式中，电子设备在通过步骤 111 得到各个像素点在当前视频帧中的边界信息之后，可以根据得到的边界信息判断各像素点是否属于前景边界点。然后，获取属于前景边界点的各个像素点的掩模值，从而基于得到的各个掩模值得到第一掩模图像。

可选地，在一些实施方式中，本申请对于电子设备执行步骤 111 计算边界信息的方式不进行限制，可以根据实际应用需求进行选择。

例如，在一种可以替代的示例中，针对每一个像素点，电子设备可以基于与该像素点相邻的多个像素点的像素值计算得到该像素点的边界信息。

示例性地，电子设备可以通过以下计算公式计算各个像素点的边界信息：

$$Gx = (fr_BW(i+1, j-1) + 2 * fr_BW(i+1, j) + fr_BW(i+1, j+1)) \\ - (fr_BW(i-1, j-1) + 2 * fr_BW(i-1, j) + fr_BW(i-1, j+1))$$

$$Gy = (fr_BW(i-1, j+1) + 2 * fr_BW(i, j+1) + fr_BW(i+1, j+1)) \\ - (fr_BW(i-1, j-1) + 2 * fr_BW(i, j-1) + fr_BW(i+1, j-1))$$

$$fr_gray(i, j) = \sqrt{Gx^2 + Gy^2}$$

其中，fr_BW（）是指像素值，fr_gray（）是指边界信息，Gx 是指横向边界差异，Gy

是指纵向边界差异, i 是指横向的第 i 个像素, j 是指纵向的第 j 个像素。

可选地, 在一些实施方式中, 本申请对于电子设备执行步骤 113 以根据边界信息得到第一掩模图像的方式也不进行限制, 可以根据实际应用需求进行选择。

例如, 在一种可以替代的示例中, 电子设备可以将当前视频帧与之前获取的视频帧进行对比, 以得到第一掩模图像。

示例性地, 电子设备可以通过以下步骤执行步骤 113:

首先, 针对每一个像素点, 电子设备可以根据该像素点在当前视频帧中的边界信息、在前 N 帧视频帧中的边界信息和在前 M 帧视频帧中的边界信息, 确定该像素点的当前掩模值和当前频数值。

然后, 针对每一个像素点, 电子设备可以根据该当前掩模值和该当前频数值判断该像素点是否属于前景边界点, 并根据属于前景边界点的各像素点的当前掩模值得到第一掩模图像。

其中, 在一种可以替代的示例中, 电子设备可以按照以下方式确定像素点的当前掩模值和当前频数值:

首先, 若一个像素点的边界信息满足第一条条件, 电子设备可以将该像素点的当前掩模值更新为 255、当前频数值加 1。其中, 在一些实施方式中, 该第一条条件可以包括: 像素点在当前视频帧中边界信息大于 $A1$, 且该像素点在当前视频帧中的边界信息与在前 N 帧视频帧中的边界信息的差值、或与在前 M 帧视频帧中的边界信息的差值大于 $B1$;

其次, 若一个像素点的边界信息不满足上述第一条条件, 但满足第二条条件, 电子设备可以将该像素点的当前掩模值更新为 180、当前频数值加 1。其中, 在一些实施方式中, 该第二条条件可以包括: 像素点在当前视频帧中边界信息大于 $A2$, 且该像素点在当前视频帧中的边界信息与在前 N 帧视频帧中的边界信息的差值、或与在前 M 帧视频帧中的边界信息的差值大于 $B2$;

然后, 若一个像素点的边界信息不满足上述的第一条条件或者是第二条条件, 但满足第三条条件, 电子设备可以将该像素点的当前掩模值更新为 0、当前频数值加 1。其中, 在一些实施方式中, 该第三条条件可以包括: 像素点在当前视频帧中的边界信息大于 $A2$;

最后, 针对不满足上述第一条条件、第二条条件或者是第三条条件的像素点, 电子设备可以将该像素点的当前掩模值更新为 0。

需要说明的是, 在一些实施方式中, 上述的当前频数值可以是指, 像素点在各视频帧中被判定为属于前景边界点的次数。例如, 针对像素点 (i, j) , 若在第一帧视频帧中被判定为属于前景边界点, 则当前频数值为 1; 若在第二帧视频帧中也被认为属于前景边界点, 则当前频数值为 2; 若在第三帧视频帧中也被认为属于前景边界点, 则当前频数值为 3。

其中, 在一些实施方式中, N 和 M 的范围可以在 1-10, 本申请对于 N 和 M 的具体取值不进行限制, 只要满足 N 不等于 M 即可。例如, 在一种可以替代的示例中, N 可以为 1, M 可以为 3。也就是说, 针对每一个像素点, 电子设备可以根据该像素点在当前视频帧中的边界信息、在前 1 帧视频帧中的边界信息和在前 3 帧视频帧中的边界信息, 确定该像素点的当前掩模值和当前频数值。

另外, 在一些实施方式中, 本申请对于上述的 $A1$ 、 $A2$ 、 $B1$ 和 $B2$ 的具体数值也不受限制, 例如, 在一种可以替代的示例中, $A1$ 可以为 30, $A2$ 可以为 20, $B1$ 可以为 12, $B2$ 可以为 8。

在一些实施方式中, 电子设备在通过上述方式得到像素点的当前掩模值和当前频数值之后, 可以将当前掩模值大于 0 的像素点确定为前景边界点, 将当前掩模值等于 0 的像素点确定为背景边界点。

并且, 为了提高确定前景边界点和背景边界点的准确性, 电子设备还可以基于以下方式判断像素点是否属于前景边界点, 该方式可以包括:

首先, 针对当前掩模值大于 0 的像素点, 若该像素点的当前频数值与当前帧数的比值

大于 0.6, 且在当前视频帧中的边界信息与在前一帧视频帧中的边界信息的差值、与前三帧视频帧中的边界信息的差值都小于 10, 电子设备可以将该像素点重新确定为背景边界点;

其次, 针对当前掩模值等于 0 的像素点, 若该像素点的当前频数值与当前帧数的比值小于 0.5, 且在当前视频帧中的边界信息大于 60, 电子设备可以将该像素点重新确定为前景边界点, 并将该像素点的当前掩模值更新为 180;

最后, 为了提高对之后的视频帧的前景图像提取的准确性, 针对不满足上述两个条件的像素点, 可以将该像素点的当前频数值减 1。

可选地, 在一些实施方式中, 本申请对于电子设备执行步骤 120 以基于当前视频帧得到第二掩模图像的方式也不进行限制, 可以根据实际应用需求进行选择。

例如, 在一种可以替代的示例中, 神经网络模型模型可以包括多个网络子模型以进行不同的处理, 从而得到第二掩模图像。

示例性地, 结合图 5, 在一种可能的实施方式中, 神经网络模型可以包括第一网络子模型、第二网络子模型和第三网络子模型。电子设备可以通过以下步骤执行步骤 120:

首先, 通过第一网络子模型对当前视频帧进行语义信息提取处理得到第一输出值。

其次, 通过第二网络子模型对第一输出值进行尺寸调整处理得到第二输出值。

然后, 通过第三网络子模型对第二输出值进行掩模图像提取处理得到第二掩模图像。

其中, 在一些实施方式中, 第一网络子模型可以通过第一卷积层、多个第二卷积层和多个第三卷积层构建。第二网络子模型可以通过第一卷积层和多个第四卷积层构建。第三网络子模型可以通过多个第四卷积层和多个上采样层构建。

需要说明的是, 在一些实施方式中, 第一卷积层可以被配置成执行一次卷积操作 (卷积核大小为 3×3)。第二卷积层可以被配置成执行两次卷积操作、一次深度可分离卷积操作和两次激活操作 (如图 6 所示)。第三卷积层可以被配置成执行两次卷积操作、一次深度可分离卷积操作和两次激活操作, 并将进行操作得到的值与输入值一并输出 (如图 7 所示)。第四卷积层可以被配置成执行一次卷积操作、一次深度可分离卷积操作和两次激活操作 (如图 8 所示)。上采样层可以被配置成执行双线性差值上采样操作 (如上采样 4 倍的操作)。

其中, 为了便于神经网络模型对当前视频帧进行识别处理, 还可以预先将该当前视频帧缩放为 $256 \times 256 \times 3$ 的数组 P , 然后, 通过归一化计算公式 (如 $(P/128) - 1$) 进行归一化处理 (得到属于 -1 到 1 的值), 并将处理得到的结果输入至神经网络模型进行识别处理。

可选地, 作为一种可能的实施方式, 本申请对于电子设备执行步骤 130 以基于预设的计算模型计算前景图像的方式也不进行限制, 可以根据实际应用需求进行选择。

例如, 在一种可以替代的示例中, 电子设备可以采用以下步骤执行步骤 130:

首先, 根据预设的第一加权系数和第二加权系数对第一掩模图像和第二掩模图像进行加权求和处理。

然后, 将进行加权求和处理得到的结果与预先确定的参数进行求和处理, 得到当前视频帧中的前景图像。

比如, 作为一种可能的实施方式, 该计算模型可以表示如下:

$$M_fi = a1 * M_fg + a2 * M_c + b$$

其中, $a1$ 为第一加权系数, $a2$ 为第二加权系数, b 为预先确定的参数, M_fg 为第一掩模图像, M_c 为第二掩模图像, M_fi 为前景图像。

需要说明的是, 上述的 $a1$ 、 $a2$ 和 b , 可以根据具体的前景图像的类型进行确定。例如, 在前景图像为人像时, 可以通过采集多张样本人像进行拟合得到。

另外, 在一些实施方式中, 上述确定出的前景图像可以被配置成进行一些特定的显示或播放控制。例如, 在直播场景中, 为了避免显示或播放的弹幕对主播人像的遮挡, 可以先确定主播人像在视频帧中的位置, 并在弹幕播放至该位置时, 对弹幕进行透明或隐藏处理。

也就是说，在一些可能的场景中，电子设备还可以对上述前景图像进行显示或播放处理。另外，为了避免在显示或播放时，出现人像抖动的情形，电子设备还可以进行抖动消除处理。

示例性地，在一种可以替代的示例中，结合图 9，电子设备在执行步骤 130 之前，该前景图像获取方法还可以包括下述的步骤 140 和步骤 150。

步骤 140，计算当前视频帧的第一掩模图像与前一帧视频帧的第一掩模图像之间的第一差值，并计算该当前视频帧的第二掩模图像与该前一帧视频帧的第二掩模图像之间的第二差值。

步骤 150，若第一差值小于预设差值，则将当前视频帧的第一掩模图像更新为前一帧视频帧的第一掩模图像；若第二差值小于预设差值，则将当前视频帧的第二掩模图像更新为前一帧视频帧的第二掩模图像。

在一些实施方式中，电子设备可以通过计算第一掩模图像和第二掩模图像在当前视频帧和前一帧视频帧之间的变化量，来判断前景图像是否有发生较大的变化。并且，电子设备可以在判断出前景图像在相邻两帧（当前帧和前一帧）之间未发生较大变化时，用前一帧的前景图像来代替当前帧的前景图像（即用前一帧的第一掩模图像来代替当前帧的第一掩模图像，用前一帧的第二掩模图像来代替当前帧的第二掩模图像），从而避免帧间抖动的问题。

如此，在前景图像（如人像）发生的变化较小时，可以使得当前帧获取到的前景图像与前一帧获取到的前景图像相同，从而实现了帧间的稳定，避免了帧间抖动而导致用户体验较差的问题。

也就是说，在一些实施方式中，电子设备在执行步骤 150 对当前视频帧的第一掩模图像和第二掩模图像进行更新之后，在执行步骤 130 时，电子设备可以基于该更新之后的一掩模图像和第二掩模图像计算前景图像。

其中，若第一差值大于或等于预设差值，第二差值大于或等于预设差值，表明前景图像变化较大。为了使得直播观众可以有效地看到主播的动作，电子设备在执行步骤 130 时，可以根据执行步骤 110 得到第一掩模图像和执行步骤 120 得到的第二掩模图像计算前景图像，使得该前景图像与前一帧的前景图像不同，从而在前景图像的播放时反映出主播的动作。

其中，本申请对于电子设备执行步骤 140 以计算第一差值和第二差值的方式不进行限制，可以根据实际应用需求进行选择。

经本申请发明人的研究发现，由于通过步骤 150 会使得主播的较小动作被消除，从而导致前景图像在播放时出现跳跃。

例如，第一帧视频帧中主播的眼睛闭合，第二帧视频帧中主播的眼睛睁开 0.1cm，第三帧视频帧中主播的眼睛睁开 0.3cm。由于从第一帧视频帧到第二帧视频帧，主播的眼睛变化较小，为了避免帧间抖动，获取到第二帧视频帧的前景图像和第一帧视频帧的前景图像会保持一致，使得在获得的第二帧视频帧的前景图像中的主播的眼睛也是闭合。

然而，由于从第二帧视频帧到第三帧视频帧，主播的眼睛变化较大，此时在获取到的第三帧视频帧的前景图像中主播的眼睛会睁开 0.3cm。如此，会使得观众观看到主播的眼睛从闭合直接变到睁开 0.3cm，也就是出现帧间（第二帧和第三帧之间）的跳跃。

考虑到部分观众可能会不适应上述的帧间跳跃的情形，因而，为了避免该情形的出现，在一种可以替代的示例中，结合图 10，电子设备可以通过下述的步骤 141 和步骤 143 执行步骤 140，以进行第一差值和第二差值的计算。

步骤 141，对当前视频帧的第一掩模图像进行帧间平滑处理得到新的第一掩模图像，并对该当前视频帧的第二掩模图像进行帧间平滑处理得到新的第二掩模图像。

步骤 143，计算新的第一掩模图像与前一帧视频帧的第一掩模图像之间的第一差值，并计算新的第二掩模图像与该前一帧视频帧的第二掩模图像之间的第二差值。

在一些实施方式中,若第一差值大于或等于预设差值,电子设备则可以将当前视频帧的第一掩模图像更新为新的第一掩模图像,使得电子设备在执行步骤 150 时,可以基于该新的第一掩模图像进行计算。

若第二差值大于或等于预设差值,电子设备则可以将当前视频帧的第二掩模图像更新为新的第二掩模图像,使得电子设备在执行步骤 150 时,可以基于该新的第二掩模图像进行计算。

其中,本申请对于电子设备执行步骤 141 进行帧间平滑处理的方式不进行限制,例如,在一种可以替代的示例中,电子设备可以通过以下步骤执行步骤 141:

首先,计算当前视频帧之前的所有视频帧的第一掩模图像的第一均值,并计算该所有视频帧的第二掩模图像的第二均值。

然后,根据第一均值和当前视频帧的第一掩模图像计算得到新的第一掩模图像,并根据第二均值和当前视频帧的第二掩模图像计算得到新的第二掩模图像。

其中,可以理解的是,在电子设备根据第一均值和第二均值计算新的第一掩模图像和新的第二掩模图像时,本申请对于具体的计算方式不进行限制。

比如,在一种可以替代的示例中,电子设备可以基于加权求和的方式计算新的第一掩模图像。例如,电子设备可以根据以下公式计算新的第一掩模图像:

$$\begin{aligned} M_{-k1} &= \alpha 1 * M_{-k2} + \beta 1 * A_{-k-1} \\ A_{-k-1} &= \alpha 2 * A_{-k-2} + \beta 2 * M_{-k2-1} \\ \alpha 1 + \beta 1 &= 1, \quad \alpha 2 + \beta 2 = 1 \end{aligned}$$

其中, M_{-k1} 为新的第一掩模图像, M_{-k2} 为通过步骤 110 得到的第一掩模图像, A_{-k-1} 为当前视频帧之前所有视频帧计算得到的第一均值, A_{-k-2} 为前一帧视频帧之前所有视频帧计算得到的第一均值, M_{-k2-1} 为前一视频帧对应的第一掩模图像, $\alpha 1$ 和 $\alpha 2$ 均可以为预先设置的值, $\alpha 1$ 的取值范围可以为[0.1, 0.9], $\alpha 2$ 的取值范围可以为[0.125, 0.875]。

可以理解的是,电子设备也可以基于加权求和的方式计算新的第二掩模图像,具体的计算公式可以参照上述计算新的第一掩模图像的公式,本申请在此不再一一赘述。

需要说明的是,电子设备在通过上述方法进行帧间平滑处理,得到新的第一掩模图像和新的第二掩模图像之后,电子设备还可以对该新的第一掩模图像和该新的第二掩模图像进行二值化处理,并在后续的步骤中基于该二值化处理的结果进行相应的计算。

其中,本申请对于电子设备进行二值化处理的方式不进行限制,例如,在一种可以替代的示例中,电子设备可以采用大津算法进行二值化处理。

需要说明的是,本申请对于电子设备执行步骤 143 进计算第一差值和第二差值的方式不进行限制,例如,在一种可以替代的示例中,电子设备可以通过以下步骤执行步骤 143:

首先,根据新的第一掩模图像中各个连通区域的面积判断该连通区域是否属于第一目标区域,并根据新的第二掩模图像中各个连通区域的面积判断该连通区域是否属于第二目标区域。

其次,计算属于第一目标区域的连通区域的第一重心坐标,并将新的第一掩模图像的重心坐标更新为第一重心坐标;计算属于第二目标区域的连通区域的第二重心坐标,并将新的第二掩模图像的重心坐标更新为第二重心坐标。

然后,计算第一重心坐标与前一帧视频帧的第一掩模图像的重心坐标之间的第一差值,并计算第二重心坐标与该前一帧视频帧的第二掩模图像的重心坐标之间的第二差值。

需要说明的是,在一种可以替代的示例中,电子设备可以基于以下方式判断新的第一掩模图像中每个连通区域是否属于第一目标区域:

首先,可以先计算新的第一掩模图像中每个连通区域的面积,并确定最大面积的目标连通区域。

其次，针对新的第一掩模图像中的每个连通区域，判断该连通区域的面积是否大于目标连通区域的三分之一（也可以是其它比例，根据实际应用需求进行确定即可）。

然后，将面积大于目标连通区域的三分之一的连通区域确定为第一目标区域。

可以理解的是，电子设备判断新的第二掩模图像中每个连通区域是否属于第二目标区域的方式，可以参照上述判断新的第一掩模图像中每个连通区域是否属于第一目标区域的方式，本申请在此不再一一赘述。

需要说明的是，在一种可以替代的示例中，电子设备可以基于以下方式计算属于第一目标区域的连通区域的第一重心坐标：

首先，判断属于第一目标区域的连通区域的数量是否大于设定的数量阈值（比如可以将该设定的数量阈值设置为 2；当然，在本申请一些其他的实施方式中，该设定的数量阈值也可以是其它的数值，根据实际应用需求进行确定即可）。

其次，若该数量大于该设定的数量阈值，则根据其中面积最大的两个属于第一目标区域的连通区域的重心坐标，计算第一重心坐标。若该数量不大于该设定的数量阈值，则直接基于属于第一目标区域的连通区域的重心坐标计算第一重心坐标。

其中，电子设备计算属于第二目标区域的连通区域的第二重心坐标的方式可以参照上述计算第一重心坐标的方式，本申请在此不再一一赘述。

需要说明的是，在通过第一均值和第二均值计算得到新的第一掩模图像和新的第二掩模图像之后，电子设备可以将通过步骤 110 得到的第一掩模图像更新为根据该新的第一掩模图像，并将通过步骤 120 得到的第二掩模图像更新为该新的第二掩模图像。

其中，在上述各个步骤中，存在着对第一掩模图像和第二掩模图像的更新处理，因此，电子设备在执行各步骤时，若在该步骤之前有进行更新处理，那么，电子设备在执行该步骤时则可以根据最近一次的更新处理后的第一掩模图像和第二掩模图像进行处理。

另外，在一些实施方式中，为了避免电子设备 300 处理器 304 的计算资源的浪费，电子设备在执行步骤 140 之前，还可以对通过步骤 110 得到的第一掩模图像和通过步骤 120 得到的第二掩模图像进行区域特征计算处理。

其中，电子设备可以通过计算在第一掩模图像中有效区域的面积占比和在第二掩模图像中有效区域的面积占比，并在该面积占比都未达到预设占比时，判定当前视频帧中不存在前景图像，因而，电子设备可以选择执行后续的步骤，从而降低电子设备 300 的处理器 304 的数据计算量，以节约电子设备 300 的计算资源。

结合图 11，在一种可以替代的示例中，可以先计算由各个前景边界点围合形成的连通区域的面积。其次，将面积最大的连通区域作为有效区域。然后，可以计算该有效区域的面积与涵盖该有效区域的最小方框的面积比值，从而得到面积占比。

结合图 12，本申请实施例还提供一种前景图像获取装置 306，前景图像获取装置 306 可以包括第一掩模图像获取模块 306a、第二掩模图像获取模块 306b 和前景图像获取模块 306c。

第一掩模图像获取模块 306a，被配置成对获得的当前视频帧进行帧间运动检测得到第一掩模图像。在一些实施方式中，第一掩模图像获取模块 306a 可以被配置成执行图 3 所示的步骤 110，关于第一掩模图像获取模块 306a 的相关内容可以参照前文对步骤 110 的描述，在本申请在此不再一一赘述。

第二掩模图像获取模块 306b，被配置成通过神经网络模型对当前视频帧进行识别得到第二掩模图像。在一些实施方式中，第二掩模图像获取模块 306b 可以被配置成执行图 3 所示的步骤 120，关于第二掩模图像获取模块 306b 的相关内容可以参照前文对步骤 120 的描述，本申请在此不再一一赘述。

前景图像获取模块 306c，被配置成根据预设的计算模型、第一掩模图像和第二掩模图像计算得到当前视频帧中的前景图像。在一些实施方式中，前景图像获取模块 306c 可以被配置成执行图 3 所示的步骤 130，关于前景图像获取模块 306c 的相关内容可以参照前文对

步骤 130 的描述，本申请在此不再一一赘述。

在本申请实施例中，对应于上述的前景图像获取方法，还提供了一种计算机可读存储介质，该计算机可读存储介质中存储有计算机程序，该计算机程序运行时执行上述前景图像获取方法的各个步骤。

其中，前述计算机程序运行时执行的各步骤，在此不再一一赘述，可参考前文对本申请提供的前景图像获取方法的解释说明。

综上，本申请提供的前景图像获取方法、前景图像获取装置和电子设备，通过对同一视频帧分别进行帧间运动检测和神经网络识别，并根据得到的第一掩模图像和第二掩模图像计算得到该视频帧中的前景图像。如此，使得计算前景图像时依据增加，从而提高计算结果的准确性和有效性，进而改善一些其他的前景提取技术方案难以准确、有效地对视频帧进行前景图像提取的问题。

以上所述仅为本申请的部分实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

工业实用性

由于本申请实施例提供的技术方案，可以通过对同一视频帧分别进行帧间运动检测和神经网络识别，并根据得到的第一掩模图像和第二掩模图像计算得到该视频帧中的前景图像。如此，使得计算前景图像时依据增加，从而提高计算结果的准确性和有效性，进而改善一些其他的前景提取技术方案难以准确、有效地对视频帧进行前景图像提取的问题。