(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113473124 A
(43)申请公布日 2021.10.01

(21)申请号 202110593819.3

(22)申请日 2021.05.28

(71)申请人 北京达佳互联信息技术有限公司
地址 100085 北京市海淀区上地西路6号1幢1层101D1-7

(72)发明人 章浩 郭晓锋 张德兵

(74)专利代理机构 北京润泽恒知识产权代理有限公司 11319
代理人 李娜

(51)Int.Cl.
*H04N 19/124*(2014.01)
*H04N 19/51*(2014.01)
*H04N 21/44*(2011.01)
*H04N 21/466*(2011.01)

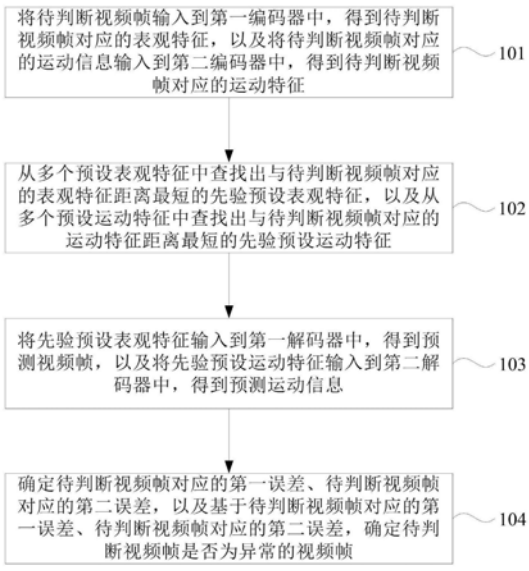
权利要求书4页 说明书11页 附图2页

(54)发明名称

信息获取方法、装置、电子设备及存储介质

(57)摘要

本公开是关于一种信息获取方法、装置、电子设备及存储介质，该方法包括：将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征；从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征；将先验预设表观特征输入到第一解码器中，得到预测视频帧，将先验预设运动特征输入到第二解码器中，得到预测运动信息；基于待判断视频帧对应的第一误差、待判断视频帧对应的第二误差，确定待判断视频帧是否为异常的视频帧。




CN 113473124 A

1.一种信息获取方法,其特征在于,所述方法包括:

将待判断视频帧输入到第一编码器中,得到所述待判断视频帧对应的表观特征,以及将所述待判断视频帧对应的运动信息输入到第二编码器中,得到所述待判断视频帧对应的运动特征,所述第一编码器和所述第二编码器均预先被基于训练集训练,所述训练集中的每一个视频帧均为正常的视频帧,所述正常的视频帧为与异常情况无关联的视频帧;

从多个预设表观特征中查找出与所述待判断视频帧对应的表观特征距离最短的先验预设表观特征,以及从多个预设运动特征中查找出与所述待判断视频帧对应的运动特征距离最短的先验预设运动特征;

将所述先验预设表观特征输入到第一解码器中,得到预测视频帧,以及将所述先验预设运动特征输入到第二解码器中,得到预测运动信息,其中,所述第一解码器、所述第二解码器均预先被基于所述训练集训练;

确定所述待判断视频帧对应的第一误差、所述待判断视频帧对应的第二误差,以及基于所述第一误差、所述第二误差,确定所述待判断视频帧是否为异常的视频帧,其中,所述第一误差指示所述预测视频帧与所述待判断视频帧之间的差异程度,所述第二误差指示所述预测运动信息与所述待判断视频帧对应的运动信息之间的差异程度。

2.根据权利要求1所述的方法,其特征在于,基于所述第一误差、所述第二误差,确定所述待判断视频帧是否为异常的视频帧包括:

计算所述第一误差的归一化值和所述第二误差的归一化值;

基于所述第一误差的预设权重、所述第二误差的预设权重,计算所述第一误差的归一化值和所述第二误差的归一化值的加权和;

将所述加权和确定为所述待判断视频帧的得分;

在所述待判断视频帧的得分大于分数阈值的情况下,确定所述待判断视频帧为异常的视频帧。

3.根据权利要求1所述的方法,其特征在于,所述预设表观特征为第一聚类中心,第一聚类中心为表观特征聚类结果对应的聚类中心,所述表观特征聚类结果通过对预设第一视频帧集合中的视频帧对应的表观特征进行聚类得到,所述预设运动特征为第二聚类中心,第二聚类中心为运动特征聚类结果对应的聚类中心,所述运动特征聚类结果通过对预设第二视频帧集合中的视频帧对应的运动特征进行聚类得到;

从多个预设表观特征中查找出对应于所述待判断视频帧对应的表观特征的先验预设表观特征包括:

计算所述待判断视频帧对应的表观特征与每一个预设表观特征的欧式距离;

将与所述待判断视频帧对应的表观特征的欧式距离最小的预设表观特征确定为所述先验预设表观特征;以及

从多个预设运动特征中查找出对应于所述待判断视频帧对应的运动特征的先验预设运动特征包括:

计算所述待判断视频帧对应的运动特征与每一个预设运动特征的欧式距离;

将与所述待判断视频帧对应的运动特征的欧式距离最小的预设运动特征确定为所述先验预设运动特征。

4.根据权利要求1所述的方法,其特征在于,在将待判断视频帧输入到第一编码器中,

得到所述待判断视频帧对应的表观特征，以及将所述待判断视频帧对应的运动信息输入到第二编码器中，得到所述待判断视频帧对应的运动特征之前，所述方法还包括：

确定所述训练集中的期望输出视频帧、所述训练集中的所述期望输出视频帧的前预设数量个视频帧；

将所述前预设数量个视频帧输入到所述第一编码器中，得到第一编码结果，将所述第一编码结果输入到所述第一解码器中，得到对应于所述期望输出视频帧的预测输出视频帧；

计算所述期望输出视频帧对应的第一损失、所述期望输出视频帧对应的第二损失、所述期望输出视频帧对应的第三损失、所述期望输出视频帧对应的第四损失，其中，所述第一损失为所述预测输出视频帧与所述期望输出视频帧之间的损失，所述第二损失为所述预测输出视频帧对应的运动信息与所述期望输出视频帧对应的运动信息之间的损失，所述预测输出视频帧对应的运动信息为所述预测输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述期望输出视频帧对应的运动信息为所述期望输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述第三损失为所述预测输出视频帧的边缘梯度与所述期望输出视频帧的边缘梯度的差值、所述第四损失为所述预测输出视频帧与所述期望输出视频帧之间的对抗损失；

基于所述第一损失、所述第二损失、所述第三损失、所述第四损失，更新所述第一编码器的网络参数、所述第一解码器的网络参数；

将所述前预设数量个视频帧中的每一个用于预测运动信息的视频帧对应的运动信息输入到所述第二编码器中，得到第二编码结果，将所述第二编码结果输入到所述第二解码器中，得到所述期望输出视频帧对应的预测运动信息，其中，所述用于预测运动信息的视频帧为所述前预设数量个视频帧中的、除了第一个视频帧之外的视频帧；

计算第五损失，基于所述第五损失，更新所述第二编码器的网络参数、所述第二解码器的网络参数，其中，所述第五损失为所述期望输出视频帧对应的预测运动信息与所述期望输出视频帧对应的运动信息之间的损失。

5.一种信息获取装置，其特征在于，所述装置包括：

获取模块，被配置为将待判断视频帧输入到第一编码器中，得到所述待判断视频帧对应的表观特征，以及将所述待判断视频帧对应的运动信息输入到第二编码器中，得到所述待判断视频帧对应的运动特征，所述第一编码器和所述第二编码器均预先被基于训练集训练，所述训练集中的每一个视频帧均为正常的视频帧，所述正常的视频帧为与异常情况无关联的视频帧；

查找模块，被配置为从多个预设表观特征中查找出与所述待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与所述待判断视频帧对应的运动特征距离最短的先验预设运动特征；

解码模块，被配置为将所述先验预设表观特征输入到第一解码器中，得到预测视频帧，以及将所述先验预设运动特征输入到第二解码器中，得到预测运动信息；

确定模块，被配置为确定所述待判断视频帧对应的第一误差、所述待判断视频帧对应的第二误差，以及基于所述第一误差、所述第二误差，确定所述待判断视频帧是否为异常的视频帧，其中，所述第一误差指示所述预测视频帧与所述待判断视频帧之间的差异程度，所

述第二误差指示所述预测运动信息与所述待判断视频帧对应的运动信息之间的差异程度。

6.根据权利要求5所述的装置，其特征在于，所述确定模块进一步被配置为计算所述第一误差的归一化值和所述第二误差的归一化值；基于所述第一误差的预设权重、所述第二误差的预设权重，计算所述第一误差的归一化值和所述第二误差的归一化值的加权和；将所述加权和确定为所述待判断视频帧的得分；在所述待判断视频帧的得分大于分数阈值的情况下，确定所述待判断视频帧为异常的视频帧。

7.根据权利要求5所述的装置，其特征在于，所述预设表观特征为第一聚类中心，第一聚类中心为表观特征聚类结果对应的聚类中心，所述表观特征聚类结果通过对预设第一视频帧集合中的视频帧对应的表观特征进行聚类得到，所述预设运动特征为第二聚类中心，第二聚类中心为运动特征聚类结果对应的聚类中心，所述运动特征聚类结果通过对预设第二视频帧集合中的视频帧对应的运动特征进行聚类得到；查找模块进一步被配置为计算所述待判断视频帧对应的表观特征与每一个预设表观特征的欧式距离；将与所述待判断视频帧对应的表观特征的欧式距离最小的预设表观特征确定为所述先验预设表观特征；计算所述待判断视频帧对应的运动特征与每一个预设运动特征的欧式距离；将与所述待判断视频帧对应的运动特征的欧式距离最小的预设运动特征确定为所述先验预设运动特征。

8.根据权利要求5所述的装置，其特征在于，所述装置还包括：

训练模块，被配置为在将待判断视频帧输入到第一编码器中，得到所述待判断视频帧对应的表观特征，以及将所述待判断视频帧对应的运动信息输入到第二编码器中，得到所述待判断视频帧对应的运动特征之前，确定所述训练集中的期望输出视频帧、所述训练集中的所述期望输出视频帧的前预设数量个视频帧；将所述前预设数量个视频帧输入到所述第一编码器中，得到第一编码结果，将所述第一编码结果输入到所述第一解码器中，得到对应于所述期望输出视频帧的预测输出视频帧；计算所述期望输出视频帧对应的第一损失、所述期望输出视频帧对应的第二损失、所述期望输出视频帧对应的第三损失、所述期望输出视频帧对应的第四损失，其中，所述第一损失为所述预测输出视频帧与所述期望输出视频帧之间的损失，所述第二损失为所述预测输出视频帧对应的运动信息与所述期望输出视频帧对应的运动信息之间的损失，所述预测输出视频帧对应的运动信息为所述预测输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述期望输出视频帧对应的运动信息为所述期望输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述第三损失为所述预测输出视频帧的边缘梯度与所述期望输出视频帧的边缘梯度的差值、所述第四损失为所述预测输出视频帧与所述期望输出视频帧之间的对抗损失；基于所述第一损失、所述第二损失、所述第三损失、所述第四损失，更新所述第一编码器的网络参数、所述第一解码器的网络参数；将所述前预设数量个视频帧中的每一个用于预测运动信息的视频帧对应的运动信息输入到所述第二编码器中，得到第二编码结果，将所述第二编码结果输入到所述第二解码器中，得到所述期望输出视频帧对应的预测运动信息，其中，所述用于预测运动信息的视频帧为所述前预设数量个视频帧中的、除了第一个视频帧之外的视频帧；计算第五损失，基于所述第五损失，更新所述第二编码器的网络参数、所述第二解码器的网络参数，其中，所述第五损失为所述期望输出视频帧对应的预测运动信息与所述期望输出视频帧对应的运动信息之间的损失。

9.一种电子设备，其特征在于，包括：

处理器；

用于存储所述处理器可执行指令的存储器；

其中，所述处理器被配置为执行所述指令，以实现如权利要求1至4中任一项所述的方法。

10.一种计算机可读存储介质，当所述计算机可读存储介质中存储的指令由电子设备的处理器执行时，使得所述电子设备能够执行如权利要求1至4中任一项所述的方法。

# 信息获取方法、装置、电子设备及存储介质

## 技术领域

[0001] 本公开涉及视频领域，尤其涉及信息获取方法、装置、电子设备及存储介质。

## 背景技术

[0002] 在审核视频是否合规时，可以确定视频中的一部分帧即视频帧是否为异常的视频帧，以确定视频是否与异常事件相关。在相关技术中，以有监督方式训练一个用于识别异常的视频帧的神经网络。为了使得用于识别异常的视频帧的神经网络达到较高的精度，需要收集大量的正样本即正常的视频帧和大量的负样本即异常的视频帧，同时，需要由相关人员对大量的正常的视频帧和大量的异常的视频帧进行标注，从而，导致实现确定视频帧是否为异常的视频帧的成本高。此外，确定视频帧是否为异常的视频帧的准确性受到收集到的负样本的影响，例如，对于某一个类型的异常事件，收集到的与该类型的异常事件相关的异常的视频帧的数量较少或者未收集到与该类型的异常事件相关的异常的视频帧，则会导致完成训练之后用于识别异常的视频帧的神经网络识别与该类型的异常事件相关的异常的视频帧的准确性较低。

## 发明内容

[0003] 为克服相关技术中存在的问题，本公开提供一种信息获取方法、装置、电子设备及存储介质，以至少解决相关技术中的实现确定视频帧是否为异常的视频帧的成本高和确定视频帧是否的异常的视频帧的准确性受到收集到的负样本的影响的问题。本公开的技术方案如下：

[0004] 根据本公开实施例的第一方面，提供一种信息获取方法，包括：

[0005] 将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征，所述第一编码器和所述第二编码器均预先被基于训练集训练，训练集中的每一个视频帧均为正常的视频帧，所述正常的视频帧为与异常情况无关联的视频帧；

[0006] 从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征；

[0007] 将所述先验预设表观特征输入到第一解码器中，得到预测视频帧，以及将所述先验预设运动特征输入到第二解码器中，得到预测运动信息，其中，所述第一解码器、所述第二解码器均预先被基于所述训练集训练；

[0008] 确定所述待判断视频帧对应的第一误差、所述待判断视频帧对应的第二误差，以及基于所述第一误差、所述第二误差，确定所述待判断视频帧是否为异常的视频帧，其中，所述第一误差指示预测视频帧与所述待判断视频帧之间的差异程度，所述第二误差指示预测运动信息与所述待判断视频帧对应的运动信息之间的差异程度。

[0009] 根据本公开实施例的第二方面，提供一种信息获取装置，包括：

[0010] 获取模块，被配置为将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征，所述第一编码器和所述第二编码器均预先被基于训练集训练，训练集中的每一个视频帧均为正常的视频帧，所述正常的视频帧为与异常情况无关联的视频帧；

[0011] 查找模块，被配置为从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征；

[0012] 解码模块，被配置为将所述先验预设表观特征输入到第一解码器中，得到预测视频帧，以及将所述先验预设运动特征输入到第二解码器中，得到预测运动信息，其中，所述第一解码器、所述第二解码器均预先被基于所述训练集训练；

[0013] 确定模块，被配置为确定所述待判断视频帧对应的第一误差、所述待判断视频帧对应的第二误差，以及基于所述第一误差、所述第二误差，确定所述待判断视频帧是否为异常的视频帧，其中，所述第一误差指示预测视频帧与所述待判断视频帧之间的差异程度，所述第二误差指示预测运动信息与待判断视频帧对应的运动信息之间的差异程度。

[0014] 本公开的实施例提供的技术方案可以包括以下有益效果：

[0015] 利用第一编码器、第二编码器、第一解码器、第二解码器确定待判断视频对应的第一误差和待判断视频对应的第二误差，基于待判断视频对应的第一误差和待判断视频对应的第二误差，确定待判断视频帧是否为异常的视频帧。第一编码器、第二编码器、第一解码器、第二解码器被预先训练时利用的训练集仅包括一种类型的视频帧即正常的视频帧，因此，无需由相关人员对训练集中的视频帧进行标注，避免了需要由相关人员对大量的正常的视频帧和大量的异常的视频帧进行标注导致成本高的情况，从而，低成本地实现确定视频帧是否为异常的视频帧。同时，确定视频帧是否的异常的视频帧的准确性不会受到负样本的影响，适用于确定与任意一个类型的异常事件相关的异常的视频帧。

## 附图说明

[0016] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

[0017] 图1是根据一示例性实施例示出的一种信息获取方法的一个实施例的流程图；

[0018] 图2是得到待判断视频的得分的一个流程示意图；

[0019] 图3是根据一示例性实施例示出的一种信息获取装置的结构框图；

[0020] 图4是根据一示例性实施例示出的一种电子设备的结构框图。

## 具体实施方式

[0021] 为了使本领域普通人员更好地理解本公开的技术方案，下面将结合附图，对本公开实施例中的技术方案进行清楚、完整地描述。

[0022] 需要说明的是，本公开的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本公开的实施例能够以除了在这里图示或

描述的那些以外的顺序实施。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0023] 图1是根据一示例性实施例示出的信息获取方法的一个实施例的流程图。该方法包括以下步骤：

[0024] 步骤101，将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征。

[0025] 在本公开中，待判断视频帧并不特指待判断视频帧所属的视频中的某一视频帧。在审核一个视频是否与异常事件相关时，可以将该视频中的至少部分视频帧中的每一个视频帧分别作为待判断视频帧，分别对该每一个视频帧执行步骤101-104。

[0026] 正常的视频帧是指与异常情况无关联的视频帧。异常情况由于表观特征异常和/或运动特征异常造成。异常的视频帧是指与异常情况相关联的视频帧，异常的视频帧通常出现在违反法律规定的视频、与违反法律规定的事件例如暴力事件相关的视频。待判断视频帧对应的表观特征可以包括待判断视频帧中的物体的颜色特征、待判断视频帧中的物体的轮廓特征。待判断视频帧对应的运动特征可以包括：待判断视频帧中的物体的运动方向、待判断视频帧中的物体的运动速度。

[0027] 在本公开中，第一编码器、第二编码器可以为编码器-解码器(Encoder-Decoder)类型的神经网络中的Encoder编码器。第一解码器、第二解码器可以为编码器-解码器类型的神经网络中的Decoder解码器。

[0028] 在本公开中，第一编码器、第二编码器、第一解码器、第二解码器均预先被基于训练集训练，训练集中的每一个视频帧均为正常的视频帧。可以基于训练集，以自监督的方式预先训练第一编码器、第二编码器、第一解码器、第二解码器。

[0029] 例如，每一次训练，可以将一个正常的视频帧输入到第一编码器中，得到第一编码器输出的编码结果，将第一编码器输出的编码结果输入到第一解码器中，得到对应于该正常的视频帧的重建视频帧。可以计算该重建视频帧与该正常的视频帧之间的损失，基于该重建视频帧与该正常的视频帧之间的损失，更新第一编码器的网络参数、第一解码器的网络参数。每一次训练，在将该正常的视频帧输入到第一编码器中的同时，可以将该正常的视频帧对应的运动信息输入到第二编码器中，得到第二编码器输出的编码结果，该正常的视频帧对应的运动信息可以为该正常的视频帧与训练集中的该正常的视频帧的关联视频帧之间的运动信息。该正常的视频帧的关联视频帧可以为该正常的视频帧所属的视频中的、该正常的视频帧的前一个视频帧。可以将第二编码器输出的编码结果输入到第二解码器中，得到对应于该正常的视频帧的重建运动信息。可以计算该重建运动信息与该正常的视频帧对应的运动信息之间的损失，根据该重建运动信息与该正常的视频帧对应的运动信息之间的损失，更新第二编码器的网络参数、第二解码器的网络参数。

[0030] 在一些实施例中，在将待判断视频帧输入到第一编码器中，得到该待判断视频帧对应的表观特征，以及将该待判断视频帧对应的运动信息输入到第二编码器中，得到该待判断视频帧对应的运动特征之前，还包括：确定训练集中的期望输出视频帧、该训练集中的该期望输出视频帧的前预设数量个视频帧；将该前预设数量个视频帧输入到该第一编码器

中，得到第一编码器输出的第一编码结果，将该第一编码结果输入到第一解码器中，得到对应于该期望输出视频帧的预测输出视频帧；计算该期望输出视频帧对应的第一损失、该期望输出视频帧对应的第二损失、该期望输出视频帧对应的第三损失、该期望输出视频帧对应的第四损失，其中，该第一损失为该预测输出视频帧与该期望输出视频帧之间的损失，该第二损失为该预测输出视频帧对应的运动信息与该期望输出视频帧对应的运动信息之间的损失，该预测输出视频帧对应的运动信息为该预测输出视频帧与该期望输出视频帧的前一个视频帧之间的运动信息，该期望输出视频帧对应的运动信息为该期望输出视频帧与该期望输出视频帧的前一个视频帧之间的运动信息，该第三损失为该预测输出视频帧的边缘梯度与该期望输出视频帧的边缘梯度的差值、该第四损失为该预测输出视频帧与该期望输出视频帧之间的对抗损失。基于该第一损失、该第二损失、该第三损失、该第四损失，更新第一编码器的网络参数、第一解码器的网络参数；将该前预设数量个视频帧中的每一个用于预测运动信息的视频帧对应的运动信息输入到第二编码器中，得到第二编码结果，将该第二编码结果输入到第二解码器中，得到该期望输出视频帧对应的预测运动信息，其中，用于预测运动信息的视频帧为该前预设数量个视频帧中的、除了第一个视频帧之外的视频帧；计算该期望输出视频帧对应的第五损失，基于该第五损失，更新第二编码器的网络参数、第二解码器的网络参数，其中，该第五损失为该期望输出视频帧对应的预测运动信息与该期望输出视频帧对应的运动信息之间的损失。

[0031] 在本公开中，每一次训练，可以确定训练集中的期望输出视频帧、该训练集中的该期望输出视频帧的前预设数量个视频帧，该期望输出视频帧和该期望输出视频帧的前预设数量个视频帧来自同一个视频，该期望输出视频帧的前预设数量个视频帧为连续的多个视频帧。该期望输出视频帧的前预设数量个视频帧中的最后一个视频帧为该期望输出视频帧的前一个视频帧。

[0032] 例如，在一次训练过程中，确定的期望输出视频帧为一个视频中的第5个视频帧，预设数量为4，该期望输出视频帧的前4个视频帧由该视频中的、该第5个视频帧的前4个视频帧即该视频中的第1-4个视频帧组成。

[0033] 对于期望输出视频帧的前预设数量个视频帧中的每一个用于预测运动信息的视频帧，该用于预测运动信息的视频帧对应的运动信息为该用于预测运动信息的视频帧与该前预设数量个视频帧中的、该用于预测运动信息的视频帧的前一个视频帧之间的运动信息。

[0034] 以下举例说明一次训练第一编码器、第一解码器、第二编码器、第二解码器的过程：

[0035] 假设在一次训练过程中，确定的期望输出视频帧为一个视频中的第5个视频帧，该第5个视频帧简称为第5个视频帧。预设数量为4，该期望输出视频帧的前预设数量个视频帧为第5个视频帧的前4个视频帧。

[0036] 可以将该期望输出视频帧的前4个视频帧简称为前4个视频帧。

[0037] 在该次训练过程中，对于第一编码器、第一解码器，将前4个视频帧输入到第一编码器中，第一编码器输出第一编码结果，将第一编码器输出的第一编码结果输入到第一解码器中，第一解码器输出对应于第5个视频帧的预测输出视频帧。

[0038] 可以将对应于第5个视频帧的预测输出视频帧简称为预测第5视频帧。

[0039] 计算第5个视频帧对应的第一损失、第5个视频帧对应的第二损失、第5个视频帧对应的第三损失、第5个视频帧对应的第四损失。

[0040] 该第一损失为预测第5视频帧与第5个视频帧之间的损失。

[0041] 第5个视频帧的前一个视频帧为前4个视频帧中的第4个视频帧。该第二损失为预测第5视频帧对应的运动信息与第5个视频帧对应的运动信息之间的损失，换言之，该第二损失为预测第5视频帧对应的光流图与第5个视频帧对应的光流图之间的损失。

[0042] 该预测第5视频帧对应的运动信息为该预测第5视频帧与该第4个视频帧之间的运动信息。该第5个视频帧对应的运动信息为该第5个视频帧与该第4个视频帧之间的运动信息。

[0043] 该第三损失为预测第5视频帧的边缘梯度与第5个视频帧的边缘梯度的差值。该第四损失为预测第5视频帧与第5个视频帧之间的对抗损失。

[0044] 在该次训练过程中，对于第二编码器、第二解码器，将前4个视频帧中的每一个用于预测运动信息的视频帧对应的运动信息输入到第二编码器中，第二编码器输出第二编码结果，将第二编码器输出的第二编码结果输入到第二解码器中，第二解码器输出第5个视频帧对应的预测运动信息。

[0045] 用于预测运动信息的视频帧为前4个视频帧中的、除了第一个视频帧之外的视频帧，前4个视频帧中的第2个视频帧、第3个视频帧、第4个视频帧均为用于预测运动信息的视频帧。第2个视频帧对应的运动信息为第2个视频帧与第1个视频帧之间的运动信息、第3个视频帧对应的运动信息为第3个视频帧与第2个视频帧之间的运动信息，第4个视频帧对应的运动信息为第4个视频帧与第3个视频帧之间的运动信息。

[0046] 计算第5个视频帧对应的对应的第五损失，基于该第五损失，更新第二编码器的网络参数、第二解码器的网络参数。

[0047] 该第五损失为第5个视频帧对应的预测运动信息与第5个视频帧对应的运动信息之间的损失。

[0048] 在本公开中，在以自监督的方式预先训练第一编码器、第二编码器、第一解码器、第二解码器时，每一次训练第一编码器、第二编码器、第一解码器、第二解码器，均可以多个视频帧参与第一编码器、第一解码器的训练，多个运动信息参与第二编码器、第二解码器的训练。相对于每一次单个视频帧参与训练，多个视频帧具有更加丰富的表观特征，使得第一编码器、第一解码器可以学习到更加丰富的表观特征，进一步提升训练完成之后第一编码器的精度、第一解码器的精度。相对于每一次单个运动信息参与训练，多个运动信息具有更加丰富的运动特征，使得第二编码器、第二解码器可以学习到更加丰富的运动特征，进一步提升训练完成之后第二编码器的精度、第二解码器的精度。

[0049] 在本公开中，可以将待判断视频帧输入到第一编码器中，得到第一编码器输出的待判断视频帧对应的表观特征。

[0050] 在本公开中，待判断视频帧对应的运动特征为对应于待判断视频帧对应的运动信息的运动特征。待判断视频帧对应的运动信息为待判断视频帧与待判断视频帧的关联视频帧之间的运动信息。待判断视频帧的关联视频帧可以为该待判断视频帧所属的视频中的、位于待判断视频帧之前的用于光流估计的视频帧，例如，待判断视频帧的关联视频帧为待判断视频帧的前一个视频帧。待判断视频帧与待判断视频帧的关联视频帧之间的运动信息

可以通过利用预设光流估计算法对待判断视频帧与待判断视频帧的关联视频帧进行光流估计得到。

[0051] 在本公开中，可以将待判断视频帧对应的运动信息输入到第二编码器中，得到第二编码器输出的待判断视频帧对应的运动特征。

[0052] 步骤102，从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征。

[0053] 在本公开中，可以预先获取用于得到多个预设表观特征的预设第一视频帧集合。预设第一视频帧集合中的每一个视频帧均为正常的视频帧。例如，预先获取多个与异常事件无关联的视频，与异常时间无关的视频中的每一个视频帧均为正常的视频帧。对于获取到的每一个视频，从该视频中抽取出一部分视频帧。从多个视频抽取出的所有视频帧组成预设第一视频帧集合。

[0054] 可以利用第一编码器获取预设第一视频帧集合中的每一个视频帧对应的表观特征。可以采用预设聚类算法例如k-means聚类算法对预设第一视频帧集合中的每一个视频帧对应的表观特征进行聚类，得到多个表观特征聚类结果和每一个表观特征聚类结果对应的聚类中心。

[0055] 对于每一个表观特征聚类结果，可以将该表观特征聚类结果中的、与该表观特征聚类结果对应的聚类中心相似度最高的表观特征确定为一个预设表观特征。

[0056] 在查找与待判断视频帧对应的表观特征距离最短的先验预设表观特征时，可以将多个预设表观特征中的、与待判断视频帧对应的表观特征的相似度最高的预设表观特征作为与待判断视频帧对应的表观特征距离最短的先验预设表观特征。

[0057] 在本公开中，可以预先获取用于得到多个预设运动特征的预设第二视频帧集合。预设第二视频帧集合中的每一个视频帧均为正常的视频帧。

[0058] 可以利用第二编码器获取预设第二视频帧集合中的每一个视频帧对应的运动特征。可以采用预设聚类算法对预设第二视频帧集合中的每一个视频帧对应的运动特征进行聚类，得到多个运动特征聚类结果和每一个运动特征聚类结果对应的聚类中心。

[0059] 对于每一个运动特征聚类结果，可以将该运动特征聚类结果中的、与该运动特征聚类结果对应的聚类中心相似度最高的运动特征确定为一个预设运动特征。

[0060] 在查找与待判断视频帧对应的运动特征距离最短的先验预设运动特征时，可以将多个预设运动特征中的、与待判断视频帧对应的运动特征的相似度最高的预设运动特征作为与待判断视频帧对应的运动特征距离最短的先验预设运动特征。

[0061] 在一些实施例中，预设表观特征为第一聚类中心，第一聚类中心为表观特征聚类结果对应的聚类中心，表观特征聚类结果通过对预设第一视频帧集合中的视频帧对应的表观特征进行聚类得到，预设运动特征为第二聚类中心，第二聚类中心为运动特征聚类结果对应的聚类中心，运动特征聚类结果通过对预设第二视频帧集合中的视频帧对应的运动特征进行聚类得到；从多个预设表观特征中查找出对应于待判断视频帧对应的表观特征的先验预设表观特征包括：计算待判断视频帧对应的表观特征与每一个预设表观特征的欧式距离；将与待判断视频帧对应的表观特征的欧式距离最小的预设表观特征确定为先验预设表观特征；从多个预设运动特征中查找出对应于待判断视频帧对应的运动特征的先验预设运

动特征包括:计算待判断视频帧对应的运动特征与每一个预设运动特征的欧式距离;将与待判断视频帧对应的运动特征的欧式距离最小的预设运动特征确定为先验预设运动特征。

[0062] 在本公开中,可以将对预设第一视频帧集合中的每一个视频帧对应的表观特征进行聚类而得到的表观特征聚类结果对应的聚类中心称之为第一聚类中心。可以将对预设第二视频帧集合中的每一个视频帧对应的运动特征进行聚类而得到的运动特征聚类结果对应的聚类中心称之为第二聚类中心。

[0063] 在本公开中,可以将每一个第一聚类中心分别作为一个预设表观特征。可以将每一个第二聚类中心分别作为一个预设运动特征。可以将多个预设表观特征中的、与待判断视频帧对应的表观特征的欧式距离最小的预设表观特征确定为先验预设表观特征。可以将多个预设运动特征中的、与待判断视频帧对应的运动特征的欧式距离最小的预设运动特征确定为先验预设运动特征。

[0064] 在本公开中,先验预设表观特征与待判断视频帧对应的表观特征的欧式距离最小,利用先验预设表观特征可以对待判断视频帧对应的表观特征进行较为精确地二次表达,先验预设运动特征与待判断视频帧对应的运动特征的欧式距离最小,利用先验预设运动特征可以对待判断视频帧对应的运动特征进行较为精确地二次表达,进而可以利用先验预设表观特征、先验预设运动特征,较为精确地确定待判断视频帧是否为异常的视频帧。

[0065] 步骤103,将先验预设表观特征输入到第一解码器中,得到预测视频帧,以及将先验预设运动特征输入到第二解码器中,得到预测运动信息。

[0066] 在本公开中,可以将先验预设表观特征输入到第一解码器中,第一解码器输出预测视频帧。可以将先验预设运动特征输入到第二解码器中,第二解码器输出预测运动信息。

[0067] 步骤104,确定待判断视频帧对应的第一误差、待判断视频帧对应的第二误差,以及基于待判断视频帧对应的第一误差、待判断视频帧对应的第二误差,确定待判断视频帧是否为异常的视频帧。

[0068] 在本公开中,待判断视频帧对应的第一误差指示预测视频帧与待判断视频帧之间的差异程度,待判断视频帧对应的第二误差指示预测运动信息与待判断视频帧对应的运动信息之间的差异程度。可以利用预设损失函数计算预测视频帧与待判断视频帧之间的损失例如L1损失、L2损失,将预测视频帧与待判断视频帧之间的损失作为待判断视频帧对应的第一误差。可以利用预设损失函数计算预测运动信息与待判断视频帧对应的运动信息之间的损失,将预测运动信息与待判断视频帧对应的运动信息之间的损失作为待判断视频帧对应的第二误差。

[0069] 在本公开中,在确定待判断视频帧对应的第一误差、待判断视频帧对应的第二误差之后,可以基于待判断视频帧对应的第一误差、待判断视频帧对应的第二误差,确定待判断视频帧是否为异常的视频帧。

[0070] 例如,可以在待判断视频帧对应的第一误差大于第一误差阈值和/或待判断视频帧对应的第二误差大于第二误差阈值的情况下,确定待判断视频帧为异常的视频帧。

[0071] 在本公开中,利用第一编码器、第二编码器、第一解码器、第二解码器确定待判断视频对应的第一误差、待判断视频对应的第二误差,基于待判断视频对应的第一误差和待判断视频对应的第二误差,确定待判断视频帧是否为异常的视频帧。第一编码器、第二编码器、第一解码器、第二解码器被预先训练时利用的训练集仅包括一种类型的视频帧即正常

的视频帧，因此，无需由相关人员对训练集中的视频帧进行标注，避免了需要由相关人员对大量的正常的视频帧和大量的异常的视频帧进行标注导致成本高的情况，从而，低成本地实现确定视频帧是否为异常的视频帧。同时，确定视频帧是否的异常的视频帧的准确性不会受到负样本的影响，适用于确定与任意一个类型的异常事件相关的异常的视频帧。

[0072] 在一些实施例中，基于待判断视频帧对应的第一误差、待判断视频帧对应的第二误差，确定该待判断视频帧是否为异常的视频帧包括：计算该第一误差的归一化值和该第二误差的归一化值；基于该第一误差的预设权重、该第二误差的预设权重，计算该第一误差的归一化值和该第二误差的归一化值的加权和；将该加权和确定为该待判断视频帧的得分；在该待判断视频帧的得分大于分数阈值的情况下，确定该待判断视频帧为异常的视频帧。

[0073] 在本公开中，若针对待判断视频帧所属的视频中的、包括待判断视频帧的多个视频帧中的每一个视频帧，均通过步骤101-104判断视频帧是否为异常的视频帧。可以分别确定该多个视频帧中的每一个视频帧对应的第一误差和每一个视频帧对应的第二误差。对该多个视频帧中的每一个视频帧对应的第一误差进行归一化处理，得到每一个视频帧对应的第一误差的归一化值。对该多个视频帧中的每一个视频帧对应的第二误差进行归一化处理，得到每一个视频帧对应的第二误差的归一化值。可以基于判断视频帧对应的第一误差的预设权重、判断视频帧对应的第二误差的预设权重，计算待判断视频帧对应的第一误差的归一化值和待判断视频帧对应的第二误差的归一化值的加权和，将待判断视频帧对应的第一误差的归一化值和待判断视频帧对应的第二误差的归一化值的加权和确定为待判断视频帧的得分。将待判断视频帧的得分与分数阈值进行比较，在待判断视频帧的得分大于分数阈值的情况下，确定待判断视频帧为异常的视频帧。

[0074] 请参考图2，其示出了得到待判断视频的得分的一个流程示意图。

[0075] 将待判断视频帧输入到第一编码器中，第一编码器输出待判断视频帧对应的表观特征。通过光流提取得到待判断视频帧对应的运动信息。将待判断视频帧对应的运动信息输入到第二编码器中，第二编码器输出待判断视频帧对应的运动特征。从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征。将先验预设表观特征输入到第一解码器中，第一解码器输出预测视频帧，将先验预设运动特征输入到第二解码器中，第二解码器输出预测运动信息。计算预测视频帧与待判断视频帧之间的损失，将预测视频帧与待判断视频帧之间的损失作为待判断视频帧对应的第一误差。计算预测运动信息与待判断视频帧对应的运动信息之间的损失，将预测运动信息与待判断视频帧对应的运动信息之间的损失作为待判断视频帧对应的第二误差。计算该第一误差的归一化值和该第二误差的归一化值，基于该第一误差的预设权重、该第二误差的预设权重，计算该第一误差的归一化值和该第二误差的归一化值的加权和，将该加权和确定为该待判断视频帧的得分。

[0076] 在本公开中，可以将待判断视频帧与正常的视频帧之间的差异程度量化为待判断视频帧的得分，待判断视频帧的得分可以较为精确地表示待判断视频帧与正常的视频帧之间的差异程度，利用较为精确地表示待判断视频帧与正常的视频帧之间的差异程度的待判断视频帧的得分与分数阈值进行比较，确定待判断视频帧是否为异常的视频帧，从而，可以

较为精确地确定待判断视频帧为异常的视频帧。对相关数据归一化处理，利用归一化后的数值进行计算可以提高计算精度，在本公开中，在计算待判断视频帧的得分时，可以利用待判断视频帧对应的第一误差的归一化值和待判断视频帧对应的第二误差的归一化值，计算出待判断视频帧的得分，可以进一步提高待判断视频帧的得分的准确性。同时，在计算待待判断视频的得分时，还考虑了待判断视频帧对应的第一误差、待判断视频帧对应的第二误差与表示待判断视频帧与正常的视频帧之间的差异程度的待判断视频帧的得分的关联程度，待判断视频帧对应的第一误差的预设权重、待判断视频帧对应的第二误差的预设权重参与计算，使得计算出的待判断视频帧的得分可以更加精确地表示待判断视频帧与正常的视频帧之间的差异程度。

[0077] 图2是根据一示例性实施例示出的一种信息获取装置的结构框图。参照图3，信息获取装置包括：获取模块301，查找模块302，解码模块303，确定模块304。

[0078] 获取模块301被配置为将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征，所述第一编码器和所述第二编码器均预先被基于训练集训练，所述训练集中的每一个视频帧均为正常的视频帧，所述正常的视频帧为与异常情况无关联的视频帧；

[0079] 查找模块302被配置为从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征；

[0080] 解码模块303被配置为将先验预设表观特征输入到第一解码器中，得到预测视频帧，以及将先验预设运动特征输入到第二解码器中，得到预测运动信息，其中，所述第一解码器、所述第二解码器均预先被基于所述训练集训练；

[0081] 确定模块304被配置为确定所述待判断视频帧对应的第一误差、所述待判断视频帧对应的第二误差，以及基于所述第一误差、所述第二误差，确定所述待判断视频帧是否为异常的视频帧，其中，所述第一误差指示预测视频帧与所述待判断视频帧之间的差异程度，所述第二误差指示预测运动信息与待判断视频帧对应的运动信息之间的差异程度。

[0082] 在一些实施例中，所述确定模块304进一步被配置为计算所述第一误差的归一化值和所述第二误差的归一化值；基于所述第一误差的预设权重、所述第二误差的预设权重，计算所述第一误差的归一化值和所述第二误差的归一化值的加权和；将所述加权和确定为所述待判断视频帧的得分；在所述待判断视频帧的得分大于分数阈值的情况下，确定所述待判断视频帧为异常的视频帧。

[0083] 在一些实施例中，所述预设表观特征为第一聚类中心，第一聚类中心为表观特征聚类结果对应的聚类中心，所述表观特征聚类结果通过对预设第一视频帧集合中的视频帧对应的表观特征进行聚类得到，所述预设运动特征为第二聚类中心，第二聚类中心为运动特征聚类结果对应的聚类中心，所述运动特征聚类结果通过对预设第二视频帧集合中的视频帧对应的运动特征进行聚类得到；查找模块202进一步被配置为计算待判断视频帧对应的表观特征与每一个预设表观特征的欧式距离；将与待判断视频帧对应的表观特征的欧式距离最小的预设表观特征确定为所述先验预设表观特征；计算待判断视频帧对应的运动特征与每一个预设运动特征的欧式距离；将与待判断视频帧对应的运动特征的欧式距离最小

的预设运动特征确定为所述先验预设运动特征。

[0084] 在一些实施例中，信息获取装置还包括：

[0085] 训练模块，被配置为在将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征之前，确定所述训练集中的期望输出视频帧、所述训练集中的所述期望输出视频帧的前预设数量个视频帧；将所述前预设数量个视频帧输入到所述第一编码器中，得到第一编码结果，将所述第一编码结果输入到所述第一解码器中，得到对应于所述期望输出视频帧的预测输出视频帧；计算所述期望输出视频帧对应的第一损失、所述期望输出视频帧对应的第二损失、所述期望输出视频帧对应的第三损失、所述期望输出视频帧对应的第四损失，其中，所述第一损失为所述预测输出视频帧与所述期望输出视频帧之间的损失，所述第二损失为所述预测输出视频帧对应的运动信息与所述期望输出视频帧对应的运动信息之间的损失，所述预测输出视频帧对应的运动信息为所述预测输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述期望输出视频帧对应的运动信息为所述期望输出视频帧与所述期望输出视频帧的前一个视频帧之间的运动信息，所述第三损失为所述预测输出视频帧的边缘梯度与所述期望输出视频帧的边缘梯度的差值、所述第四损失为所述预测输出视频帧与所述期望输出视频帧之间的对抗损失；基于所述第一损失、所述第二损失、所述第三损失、所述第四损失，更新所述第一编码器的网络参数、所述第一解码器的网络参数；将所述前预设数量个视频帧中的每一个用于预测运动信息的视频帧对应的运动信息输入到所述第二编码器中，得到第二编码结果，将所述第二编码结果输入到所述第二解码器中，得到所述期望输出视频帧对应的预测运动信息，其中，所述用于预测运动信息的视频帧为所述前预设数量个视频帧中的、除了第一个视频帧之外的视频帧；计算第五损失，基于所述第五损失，更新所述第二编码器的网络参数、所述第二解码器的网络参数，其中，所述第五损失为所述期望输出视频帧对应的预测运动信息与所述期望输出视频帧对应的运动信息之间的损失。

[0086] 关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

[0087] 图4是根据一示例性实施例示出的一种电子设备的结构框图。参照图4，电子设备包括处理组件422，其进一步包括一个或多个处理器，以及由存储器432所代表的存储器资源，用于存储可由处理组件422执行的指令，例如应用程序。存储器432中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件422被配置为执行指令，以执行上述方法。

[0088] 电子设备还可以包括一个电源组件426被配置为执行电子设备的电源管理，一个有线或无线网络接口450被配置为将电子设备连接到网络，和一个输入输出(I/O)接口458。电子设备可以操作基于存储在存储器432的操作系统，例如Windows ServerTM，MacOS XTM，UnixTM，LinuxTM，FreeBSDTM或类似。

[0089] 在示例性实施例中，还提供了一种包括指令的存储介质，例如包括指令的存储器，上述指令可由电子设备执行以完成上述方法。可选地，计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0090] 在示例性实施例中，本申请还提供一种计算机程序产品，包括计算机可读代码，当

计算机可读代码在电子设备上运行时，使得电子设备执行上述信息获取方法。

[0091] 本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本发明的其它实施方案。本申请旨在涵盖本发明的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的，本发明的真正范围和精神由下面的权利要求指出。

[0092] 应当理解的是，本公开并不局限于上面已经描述并在附图中示出的精确结构，并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制。

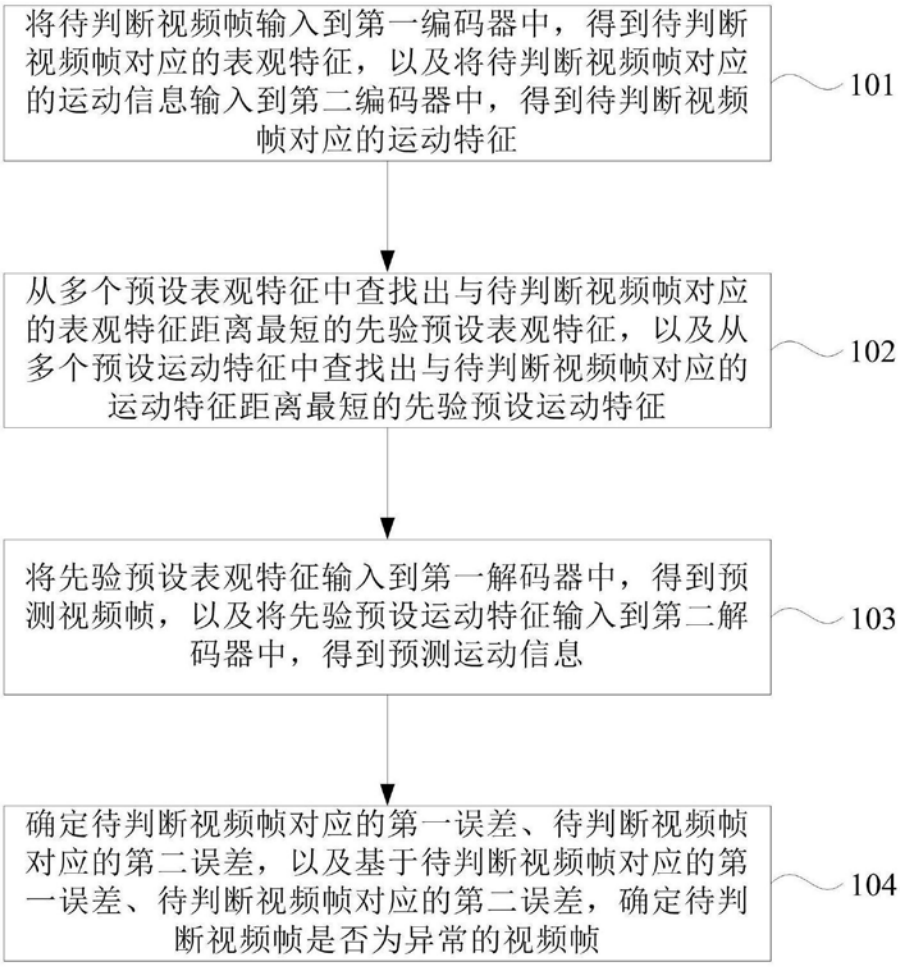

将待判断视频帧输入到第一编码器中，得到待判断视频帧对应的表观特征，以及将待判断视频帧对应的运动信息输入到第二编码器中，得到待判断视频帧对应的运动特征
101
从多个预设表观特征中查找出与待判断视频帧对应的表观特征距离最短的先验预设表观特征，以及从多个预设运动特征中查找出与待判断视频帧对应的运动特征距离最短的先验预设运动特征
102
将先验预设表观特征输入到第一解码器中，得到预测视频帧，以及将先验预设运动特征输入到第二解码器中，得到预测运动信息
103
确定待判断视频帧对应的第一误差、待判断视频帧对应的第二误差，以及基于待判断视频帧对应的第一误差、待判断视频帧对应的第二误差，确定待判断视频帧是否为异常的视频帧
104

图1

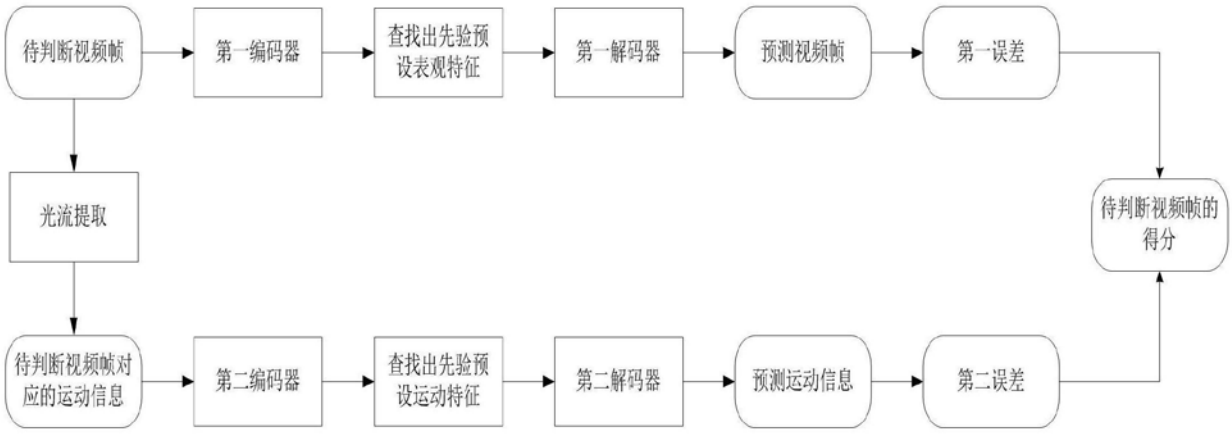

待判断视频帧
第一编码器
查找出先验预设表观特征
第一解码器
预测视频帧
第一误差
光流提取
待判断视频帧的得分
待判断视频帧对应的运动信息
第二编码器
查找出先验预设运动特征
第二解码器
预测运动信息
第二误差

图2

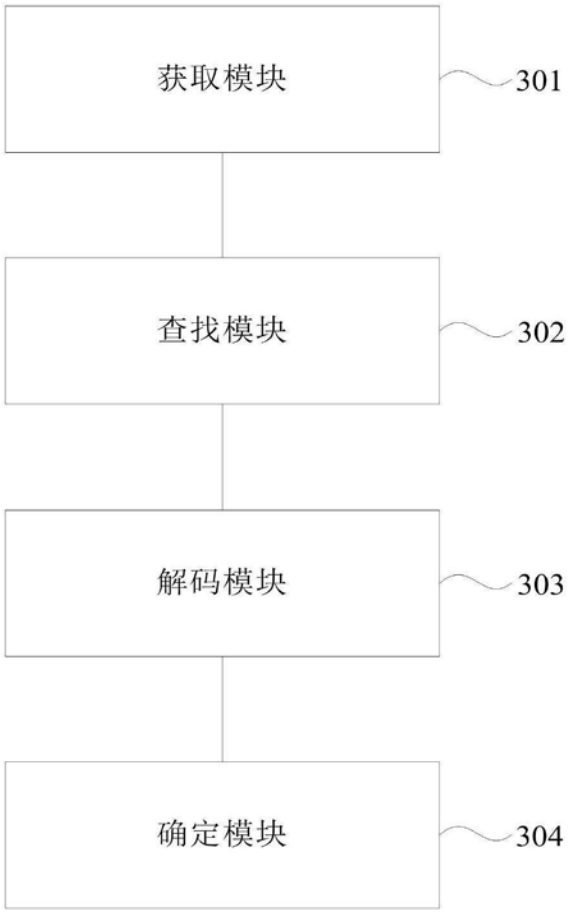

获取模块
301
查找模块
302
解码模块
303
确定模块
304

图3

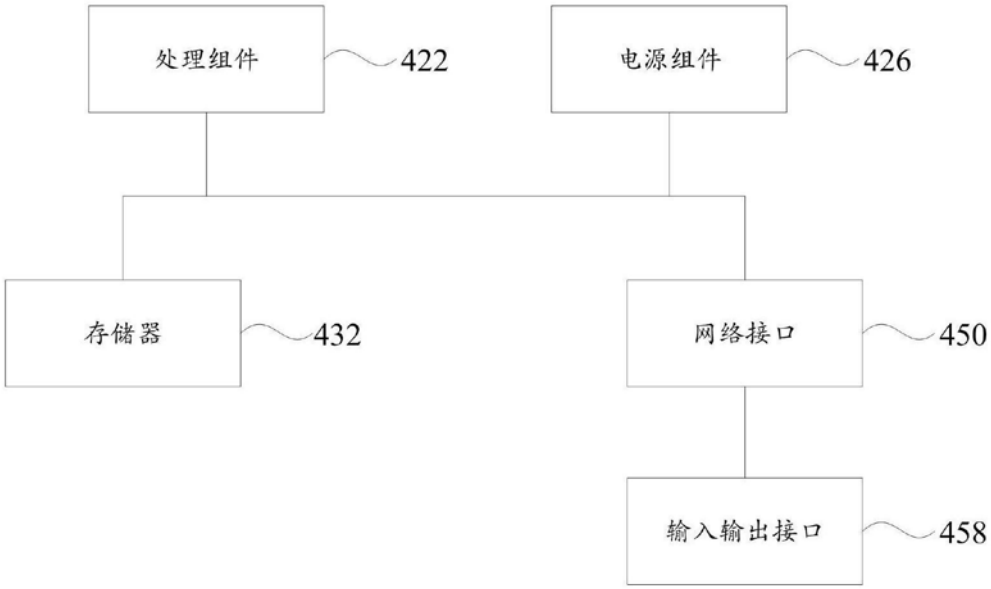

处理组件
422
电源组件
426
存储器
432
网络接口
450
输入输出接口
458

图4