



(12) 发明专利

(10) 授权公告号 CN 110753239 B

(45) 授权公告日 2022.03.08

(21) 申请号 201810812432.0

(22) 申请日 2018.07.23

(65) 同一申请的已公布的文献号  
申请公布号 CN 110753239 A

(43) 申请公布日 2020.02.04

(73) 专利权人 深圳地平线机器人科技有限公司  
地址 518000 广东省深圳市南山区粤海街  
道粤兴三道6号南京大学产学研基地B  
区601室

(72) 发明人 刘景初

(74) 专利代理机构 北京市正见永申律师事务所  
11497

代理人 黄小临 冯玉清

(51) Int. Cl.

H04N 21/2343 (2011.01)

H04N 21/4402 (2011.01)

(56) 对比文件

CN 101882217 A, 2010.11.10

CN 108267123 A, 2018.07.10

CN 102810161 A, 2012.12.05

CN 105487537 A, 2016.04.13

CN 102881024 A, 2013.01.16

CN 108171324 A, 2018.06.15

CN 104866821 A, 2015.08.26

CN 107577996 A, 2018.01.12

US 2002099594 A1, 2002.07.25

JP 2010250759 A, 2010.11.04

US 2009110236 A1, 2009.04.30

谈东奎. 人机共享的驾驶员横向辅助系统关键技术研究.《中国博士学位论文全文数据库 工程科技 II 辑》. 2018, (第一期),

Heng Wang等. Text Generation Based on Generative Adversarial Nets with Latent Variable.《arXiv》. 2017,

审查员 叶会

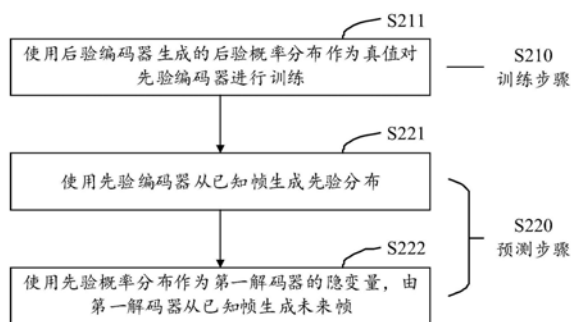
权利要求书2页 说明书9页 附图5页

(54) 发明名称

视频预测方法、视频预测装置、电子设备和车辆

(57) 摘要

公开了一种视频预测方法、视频预测装置、电子设备和车辆。该视频预测方法包括：训练步骤，使用后验编码器生成的后验概率分布作为真值，对先验编码器进行训练；以及预测步骤，使用先验编码器从已知帧生成先验概率分布，使用先验概率分布作为第一解码器的隐变量，由第一解码器从已知帧生成预测帧。由此，可以使用后验编码器生成的后验概率分布训练先验编码器，再以先验编码器的先验分布作为隐变量进行视频预测，从而对于随机视频预测问题，能够获得更鲁棒的隐变量先验估计。



1. 一种视频预测方法,包括:  
训练步骤,包括:  
使用后验编码器基于视频数据的先前帧和后续帧生成的后验概率分布作为真值,对先验编码器进行训练;以及  
预测步骤,包括:  
使用所述先验编码器从输入的视频的已知帧生成先验概率分布;以及  
使用所述先验概率分布作为第一解码器的隐变量,由所述第一解码器从所述已知帧生成未来帧。
2. 如权利要求1所述的视频预测方法,其中,所述后验编码器根据预定的不需要训练的非学习类程序规则从所述先前帧和后续帧生成后验概率分布估计。
3. 如权利要求1所述的视频预测方法,其中,所述训练步骤还包括:  
在对先验编码器进行训练之前,对后验编码器进行训练。
4. 如权利要求3所述的视频预测方法,其中,对后验编码器进行训练包括:  
使用所述后验编码器基于所述先前帧和后续帧生成后验概率分布;  
使用所述后验概率分布作为第二解码器的隐变量,由所述第二解码器从所述先前帧生成预测帧;以及  
以所述预测帧与所述后续帧之间的误差作为损失,调整所述后验编码器的参数。
5. 如权利要求4所述的视频预测方法,其中,所述第一解码器和所述第二解码器彼此相同或不同。
6. 如权利要求4所述的视频预测方法,其中,所述先验编码器、所述后验编码器、所述第一解码器和所述第二解码器包括全卷积网络、循环卷积网络或长短期记忆网络。
7. 如权利要求4所述的视频预测方法,其中,所述先验编码器用于产生先验光流场概率分布,所述后验编码器用于产生后验光流场概率分布,所述第一解码器和所述第二解码器用于通过基于光流场的像素变换来产生预测帧。
8. 如权利要求4所述的视频预测方法,其中,在所述训练步骤中对先验和后验编码器进行训练时,所述第一和第二解码器也分别被训练。
9. 如权利要求8所述的视频预测方法,其中,在所述训练步骤中对后验编码器的训练和对先验编码器的训练交替或同步进行。
10. 一种视频预测装置,包括:先验编码器、后验编码器、第一解码器、训练单元和预测单元,  
所述训练单元配置为使用所述后验编码器基于视频数据的先前帧和后续帧生成的后验概率分布作为真值,对所述先验编码器进行训练,并且  
所述预测单元配置为使用所述先验编码器从输入的视频的已知帧生成先验概率分布,并且使用所述先验概率分布作为所述第一解码器的隐变量,由所述第一解码器从所述已知帧生成未来帧。
11. 如权利要求10所述的视频预测装置,还包括第二解码器,其中,所述训练单元还配置为  
使用所述后验编码器基于所述先前帧和后续帧生成后验概率分布,使用所述后验概率分布作为第二解码器的隐变量,由所述第二解码器从所述先前帧生成预测帧;以及

以所述预测帧与所述后续帧之间的误差作为损失,调整所述后验编码器的参数,从而对所述后验编码器进行训练。

12. 一种电子设备,包括:

处理器;以及

存储器,在所述存储器中存储有计算机程序指令,所述计算机程序指令在被所述处理器运行时使得所述处理器执行如权利要求1至9中的任一项所述的视频预测方法。

13. 一种车辆,包括如权利要求12所述的电子设备。

14. 一种计算机可读介质,其上存储有计算机程序指令,所述计算机程序指令在被处理器运行时使所述处理器执行如权利要求1至9中的任一项所述的视频预测方法。

## 视频预测方法、视频预测装置、电子设备和车辆

### 技术领域

[0001] 本公开总体上涉及自动驾驶的技术领域,并且具体地涉及一种视频预测方法、视频预测装置、电子设备和车辆。

### 背景技术

[0002] 近年来,自动驾驶或者说高级驾驶辅助系统(ADAS)日益受到广泛关注。ADAS系统通常需要使用各种车载传感器来感知车辆的本身和周围环境的各种状态,然后根据所收集的数据对静态和/或动态对象进行辨识、侦测与追踪,并结合地图数据进行运算和分析,从而制定驾驶策略,并最终实现自动驾驶功能。

[0003] 在自动驾驶场景中,通常需要根据通过诸如摄像头等图像采集器件获得的视频或图像序列对环境中的动态对象(例如,行人、车辆等)的运动进行预测,然后将预测结果提供给后续模块使用,从而实现车辆的驾驶控制等功能。

[0004] 对于图像序列或视频的随机预测,通常可以基于像素自回归模型或者自编码器框架提供相应的预测方法。然而,这些方法由于计算效率低而难以应对实时性要求,或者预测效果不佳,或者不具备鲁棒性。

[0005] 因此,期待计算效率高且具有高鲁棒性的视频或图像序列的预测方案。

### 发明内容

[0006] 为了解决或减轻至少上述技术问题,本公开的实施例提供了一种视频预测方法、视频预测装置、电子设备和车辆。通过该视频预测方法,能够使用后验编码器生成的后验概率分布训练先验编码器,然后以先验编码器的先验分布作为隐变量进行视频预测,从而至少对于随机视频预测,能够获得更鲁棒的隐变量先验估计。

[0007] 一方面,本公开提供了一种视频预测方法,该方法可以包括训练步骤和预测步骤。训练步骤可以包括:使用后验编码器生成的后验概率分布作为真值,对先验编码器进行训练。预测步骤可以包括:使用前述先验编码器从已知帧生成先验概率分布;以及使用前述先验概率分布作为第一解码器的隐变量,由前述第一解码器从前述已知帧生成未来帧。

[0008] 在一个实施例中,前述后验编码器可以根据预定的非学习类程序规则从先前帧和后续帧生成后验概率分布估计。

[0009] 在一个实施例中,前述训练步骤还可以包括:在对先验编码器进行训练之前,对后验编码器进行训练。

[0010] 在一个实施例中,对后验编码器进行训练可以包括:使用前述后验编码器基于先前帧和后续帧生成后验概率分布;使用前述后验概率分布作为第二解码器的隐变量,由前述第二解码器从前述先前帧生成预测帧;以及以前述预测帧与前述后续帧之间的误差作为损失,调整前述后验编码器的参数。

[0011] 在一个实施例中,前述第一解码器和前述第二解码器可以彼此相同或不同。

[0012] 在一个实施例中,前述先验编码器、前述后验编码器、前述第一解码器和前述第二

解码器可以包括全卷积网络、循环卷积网络或长短期记忆网络。

[0013] 在一个实施例中,前述先验编码器可以用于产生先验光流场概率分布,前述后验编码器可以用于产生后验光流场概率分布,前述第一解码器和前述第二解码器可以用于通过基于光流场的像素变换来产生预测帧。

[0014] 在一个实施例中,在所述训练步骤中对先验和后验编码器进行训练时,前述第一和第二解码器也可以分别被训练。

[0015] 在一个实施例中,在所述训练步骤中对后验编码器的训练和对先验编码器的训练可以交替或同步进行。

[0016] 另一方面,本公开提供了一种视频预测装置。该视频预测装置可以包括先验编码器、后验编码器、第一解码器、训练单元和预测单元。训练单元可以被配置为使用前述后验编码器生成的后验概率分布作为真值,对前述先验编码器进行训练。预测单元可以被配置为使用前述先验编码器从已知帧生成先验概率分布,并且使用前述先验概率分布作为前述第一解码器的隐变量,由前述第一解码器从前述已知帧生成未来帧。

[0017] 在一个实施例中,后验编码器可以根据预定的非学习类程序规则从先前帧和后续帧生成后验概率分布估计。

[0018] 在一个实施例中,训练单元还可以被配置为在对先验编码器进行训练之前,对后验编码器进行训练。

[0019] 在一个实施例中,上述视频预测装置还可以包括第二解码器。训练单元还可以被配置为使用前述后验编码器基于先前帧和后续帧生成后验概率分布,使用前述后验概率分布作为第二解码器的隐变量,由前述第二解码器从前述先前帧生成预测帧,以前述预测帧与前述后续帧之间的误差作为损失,调整前述后验编码器的参数,从而对前述后验编码器进行训练。

[0020] 在一个实施例中,第一解码器和第二解码器可以彼此相同或不同。

[0021] 在一个实施例中,先验编码器、后验编码器、第一解码器和第二解码器可以包括全卷积网络、循环卷积网络或长短期记忆网络。

[0022] 在一个实施例中,先验编码器可以被配置为产生先验光流场概率分布,后验编码器可以被配置为产生后验光流场概率分布,第一解码器和前述第二解码器可以被配置为通过基于光流场的像素变换来产生预测帧。

[0023] 在一个实施例中,训练单元可以被配置为在对先验和后验编码器进行训练时也分别训练第一和第二解码器。

[0024] 在一个实施例中,训练单元可以被配置为交替或同步地进行对后验编码器的训练和对先验编码器的训练。

[0025] 另一方面,本公开提供了一种电子设备。该电子设备可以包括处理器和存储器。在存储器中可以存储有计算机程序指令,前述计算机程序指令在被处理器运行时使得处理器执行上述视频预测方法。

[0026] 另一方面,本公开提供了一种车辆,其可以包括上述电子设备。

[0027] 另一方面,本公开提供了一种计算机可读介质,其上可以存储有计算机程序指令,前述计算机程序指令在被处理器运行时使得前述处理器执行上述视频预测方法。

[0028] 通过根据本公开的实施例的视频预测方法、视频预测装置、电子设备和车辆,能够

在训练步骤中使用后验编码器生成的后验概率分布作为真值,对先验编码器进行训练,并且在预测步骤中使用前述先验编码器从已知帧生成先验概率分布,并使用前述先验概率分布作为第一解码器的隐变量,由第一解码器从前述已知帧生成未来帧。因此,能够使用后验编码器生成的后验概率分布训练先验编码器,再以先验编码器的先验分布作为隐变量进行视频预测,从而至少对于随机视频预测,能够获得更鲁棒的隐变量先验估计。

#### 附图说明

- [0029] 图1示出根据本公开的实施例的视频预测方法所应用的系统架构的示例。
- [0030] 图2示出根据本公开的实施例的视频预测方法的示例的流程图。
- [0031] 图3示出根据本公开的实施例的视频预测方法中后验编码器的训练过程的示例。
- [0032] 图4示出根据本公开的实施例的视频预测方法中以后验概率分布对先验编码器的训练过程的示例。
- [0033] 图5示出根据本公开的实施例的视频预测方法中的预测过程的示例。
- [0034] 图6示出根据本公开的实施例的视频预测方法中同时训练先验编码器和第一解码器的过程的示例。
- [0035] 图7示出根据本公开的实施例的视频预测方法中同时训练先验编码器和第一解码器的过程的另一示例。
- [0036] 图8示出根据本公开的实施例的视频预测装置的示例。
- [0037] 图9示出根据本公开的实施例的电子设备的示例。

#### 具体实施方式

[0038] 下面参考附图描述根据本公开的方法、装置和系统。应当理解,所描述的实施例或示例仅仅是本公开的一部分实施例或示例,而不是本公开的全部实施例或示例。本公开局限于所描述的实施例或示例。

[0039] 通常,对于视频预测,需要考虑未来的随机性,并且因此需要建立从历史已知图像到待预测目标图像先验概率分布的估计模型,而非确定性的映射。在本文中,将不知道预测结果的情况称为先验(prior),而将已经知道要预测的结果的情况称为后验(posterior)。

[0040] 对于视频或图像序列的随机预测,要解决的主要问题之一是如何有效地学习和表达待预测目标图像的概率分布。

[0041] 在一个实施例中,可以通过像素自回归模型表达图像像素间的联合概率分布。然而,这样的方式的计算效率低,因此难以应对实时性要求。

[0042] 在另外的示例中,可以使用变分自编码器(Variational Autoencoder)框架,通过变分推断算法学习将概率分布已知的隐变量映射为概率分布未知的待预测目标图像。然而,这样的方式将随机隐变量空间假设为具有固定的概率分布,例如多维标准高斯噪声 $N(0, 1)$ ,而这样的限制约束了预测输出的可能范围,从而可能造成预测效果的恶化。

[0043] 另外,也可以在在变分自编码器的框架中加入可学习的隐变量约束。然,在这种方式中,先验和后验估计器互为学习目标,导致在训练过程中容易出现过早收敛或不能收敛的现象,因此不具有鲁棒性。

[0044] 根据本公开的实施例的视频预测方法可以包括训练步骤和预测步骤。在训练步骤

中,可以使用后验编码器生成的后验概率分布作为真值,对先验编码器进行训练。在预测步骤中,可以使用先验编码器从已知帧生成先验概率分布,并使用先验概率分布作为第一解码器的隐变量,由第一解码器从已知帧生成未来帧。由此,通过使用后验编码器生成的后验概率分布作为监督信号来训练先验概率分布的生成,能够获得随机的视频预测模型,从而至少对于随机视频预测,能够获得更鲁棒的隐变量先验估计。

[0045] 根据本公开的实施例的视频预测方法、视频预测装置、电子设备和车辆可以直接应用于视频预测,并且也可以用于任何可以转化为视频预测的其他预测任务。例如,对于自动驾驶场景中的诸如车辆、行人等动态对象的运动预测,可以转化为全景俯视图中各种动态对象所占据的格点图序列的预测任务。此外,预测的图像不局限于包含单个或三个颜色通道的自然图像,也可以是隐含地表达其他信息(如速度、加速度)的任何形式的张量。

[0046] 图1示出根据本公开的实施例的视频预测方法所应用的系统架构的示例100。如图1所示,系统100可以包括先验编码器110、后验编码器120和解码器130。

[0047] 在一个实施例中,先验编码器110可以被配置为接收视频的当前时刻的前几帧(包括当前帧)的输入并生成先验概率分布,后验编码器120可以被配置为接收视频的当前时刻的前后几帧的输入并生成后验概率分布,并且解码器130可以被配置为使用先验编码器110生成的先验概率分布作为隐变量并根据已知帧生成未来帧。

[0048] 在一个实施例中,先验编码器110、后验编码器120和解码器130可以包括但不限于全卷积网络、循环卷积网络和长短期记忆网络中的一个或多个。

[0049] 如图2所示,根据本公开的实施例的视频预测方法的示例可以包括训练步骤S210和预测步骤S220。训练步骤S210可以包括:步骤S211,使用后验编码器生成的后验概率分布作为真值,对先验编码器进行训练。预测步骤S220可以包括:步骤S221,使用先验编码器从已知帧生成先验概率分布;以及步骤S222,使用先验概率分布作为第一解码器的隐变量,由第一解码器从已知帧生成未来帧。

[0050] 在一个实施例中,根据本公开的实施例的视频预测方法可以开始于训练步骤S210中的步骤S211,以使用后验编码器估计隐变量的后验概率分布,再采用估计获得的后验概率分布监督先验分布的学习。

[0051] 在一个实施例中,该视频预测方法可以适用于解码器较简单,参数确定的情况,例如,可以基于光流的预测模型,通过对像素进行仿射变换(warping)操作来获得预测帧,并且可以将解码器视为是已知或确定的或者经过训练的。

[0052] 在该实施例中,先验编码器可以用于产生先验光流场概率分布,后验编码器可以用于产生后验光流场概率分布,并且解码器可以用于通过基于光流场的像素变换来产生预测帧。

[0053] 根据不同的实施例,像素变换可以包括仿射变换,也可以包括例如使用稀释卷积等实现的像素移动。

[0054] 在一个实施例中,在步骤S211中,可以使用后验编码器生成的后验概率分布作为真值,对先验编码器进行训练。

[0055] 在一个实施例中,后验编码器可以根据不需要训练的非学习类程序规则从先前帧和后续帧生成后验概率分布估计。例如,后验编码器可以直接或间接地利用已知图像帧和待预测输出帧真值,通过预先确定好的非学习类程序规则(例如,匹配)提取后验估计。

[0056] 在另外的实施例中,后验编码器也可以通过训练获得。例如,可以在对先验编码器进行训练之前,先对后验编码器进行训练。

[0057] 例如,在图3所示的对后验编码器进行训练的示例过程中,可以首先使用后验编码器基于先前帧和后续帧( $X_{t-4}:X_{t+4}$ )生成后验概率分布,然后可以使用后验概率分布作为第二解码器的隐变量,由第二解码器从先前帧生成预测帧( $X_{\sim t+1}$ ),并以预测帧( $X_{\sim t+1}$ )与后续帧( $X_{t+1}$ )之间的误差(例如,均方误差MSE)作为损失,来调整后验编码器的参数。

[0058] 在一个实施例中,在训练后验编码器的过程中,可以使用真实视频数据而不是预测视频数据作为后验编码器的先前帧和后续帧,从而增加训练的稳定性。

[0059] 根据不同的实施例,可以直接将后验概率分布作为第二解码器的隐变量,也可以首先对后验概率分布进行诸如采样、求均值、求方差等处理以获得相应的确定分布,从而将获得的采样值、均值或方差等作为隐变量。

[0060] 另外,针对不同的应用场景,后验编码器生成的后验概率分布可以对应于不同的分布(例如高斯分布),并且相应地,可以对后验概率分布进行不同的处理,并通过不同的方式获得隐变量。

[0061] 然后,可以将该隐变量作为第二解码器的隐变量,并由第二解码器生成预测帧。然后,可以使用预测帧和真值的后续帧之间的误差作为损失来训练后验编码器。

[0062] 在使用训练好的后验编码器获得后验概率分布之后,可以使用后验概率分布作为真值,对先验编码器进行监督训练。

[0063] 另外,该训练过程也可以直接使用概率分布(例如,概率分布之间的交叉熵、KL散度等)作为损失来训练先验编码器。

[0064] 然后,如图4所示,后验编码器可以基于先前帧和后续帧( $X_{t-4}:X_{t+4}$ )生成后验概率分布。先验编码器可以基于先前帧(包括当前帧,例如,如图4所示的 $X_{t-4}:X_t$ )生成先验概率分布。然后,可以使用后验概率分布与先验概率分布之间的KL散度作为损失来训练先验编码器。

[0065] 可以将预测目标帧的先验分布视为后验分布在所有可能的预测目标取值下的平均分布,因此当后验概率估计精确时,上述有监督学习方法可以用于训练先验概率分布的估计。

[0066] 另外,根据本公开的实施例的后验监督先验的过程可以采用以将后验分布边缘化为先验分布作为期望目标的任何监督方式。

[0067] 回到图2,在训练步骤S210之后,根据本公开的方法可以继续到预测步骤S220,以使用先验编码器生成的先验概率分布作为隐变量预测待预测的目标图像帧。

[0068] 在步骤S221中,先验编码器可以从已知帧(例如,图5中的 $X_{t-4}:X_t$ )生成先验概率分布。然后,在步骤S222中,可以使用先验概率分布作为第一解码器的隐变量,并由第一解码器从已知帧生成未来帧(例如,图5中的 $X'_{t+1}$ )。

[0069] 根据不同的实施例,先验概率分布可以直接作为第一解码器的隐变量,也可以首先对先验概率分布进行诸如采样、求均值、求方差等处理以获得确定分布,从而将获得的采样值、均值或者方差等作为隐变量。

[0070] 另外,根据不同的应用场景,先验编码器生成的先验概率分布也可以对应于不同的分布(例如高斯分布),并且相应地,可以对先验概率分布进行不同的处理,并通过不同的



方式获得隐变量。。

[0071] 在一个实施例中,第一解码器和前述的第二解码器可以是相同的解码器,例如长短期记忆网络。在另外的实施例中,第一解码器和第二解码器也可以是不同的解码器。

[0072] 根据第一解码器和第二解码器的类型,在训练步骤S210中对先验编码器和后验编码器进行训练时,可以分别训练第一解码器和第二解码器。

[0073] 在一个实施例中,在训练后验编码器时,可以同时第二解码器进行训练,例如可以使用预测帧与后续帧之间的误差作为损失,同时优化后验编码器和第二解码器的可训练参数。

[0074] 在另外的实施例中,在以后验概率分布训练先验编码器的过程中,可以同时第一解码器进行训练。

[0075] 如图6所示,后验编码器基于先前帧和后续帧(例如,如图6所示的 $X_{t-4}:X_{t+4}$ )生成后验概率分布,且先验编码器基于先前帧(包括当前帧,例如,如图6所示的 $X_{t-4}:X_t$ )生成先验概率分布。

[0076] 进一步地,可以使用先验编码器的先验概率分布作为隐变量,并由第一解码器生成训练预测帧(例如,如图6所示的 $X_{t+1}$ ),然后使用先验概率分布与后验概率分布之间的KL散度以及训练预测帧( $X_{t+1}$ )与后续帧( $X_{t+1}$ )之间的均方误差(MSE)作为损失函数,即 $loss = MSE + KL$ 。然后,可以使用诸如随机最速下降法(SGD)或者其改进方法等方法来训练先验编码器和第一解码器的可训练网络参数,从而优化先验编码器和第一解码器。

[0077] 在另外的实施例中,在训练步骤中对后验编码器的训练和对先验编码器的训练可以交替或同步地进行,从而使用对抗训练的方式来训练先验编码器和后验编码器。

[0078] 在同时训练先验编码器和第一解码器的情况下,如图7所示,也可以使用先验编码器的先验概率分布和后验编码器的后验概率分布两者作为隐变量,并由第一解码器生成训练预测帧,然后使用先验概率分布与后验概率分布之间的KL散度以及训练预测帧与后续帧之间的均方误差(MSE)作为损失函数来训练先验编码器和第一解码器。

[0079] 由此,可以通过基于先验/后验分布两者来训练第一解码器,并且可以使用对抗训练的方式进行训练,使得第一解码器的预测效果更好。最终,通过训练,先验分布与后验分布收敛为趋于一致。

[0080] 图8示出根据本公开的实施例的视频预测装置的示例300。如图8所示,示例装置300可以包括先验编码器310、后验编码器320、第一解码器330、训练单元340和预测单元350。

[0081] 在一个实施例中,训练单元340可以被配置为使用后验编码器320生成的后验概率分布作为真值,对先验编码器310进行训练。预测单元350可以被配置为使用先验编码器310从已知帧生成先验概率分布,并且使用先验概率分布作为第一解码器330的隐变量,由第一解码器330从已知帧生成未来帧。

[0082] 在一个实施例中,后验编码器320可以根据预定的非学习类程序规则从先前帧和后续帧生成后验概率分布估计。

[0083] 在一个实施例中,训练单元340还可以被配置为在对先验编码器310进行训练之前,对后验编码器320进行训练。

[0084] 如图8所示,在一个实施例中,示例装置300还可以包括第二解码器360。

[0085] 在该实施例中,第二解码器360与第一解码器330可以是相同的解码器或同一解码器。

[0086] 在该实施例中,训练单元340还可以被配置为使用后验编码器320基于先前帧和后续帧生成后验概率分布,使用后验概率分布作为第二解码器360的隐变量,由第二解码器360从先前帧生成预测帧。另外,训练单元340还可以被配置为使用预测帧与后续帧之间的误差作为损失,调整后验编码器320的参数,从而对后验编码器320进行训练。

[0087] 根据不同的实施例,第一解码器300和第二解码器360可以彼此相同或不同。

[0088] 根据不同的实施例,先验编码器310、后验编码器320、第一解码器330和/或第二解码器360可以包括但不限于全卷积网络、循环卷积网络和长短期记忆网络中的一个或多个。

[0089] 在一个实施例中,先验编码器310可以被配置为产生先验光流场概率分布,后验编码器320可以被配置为产生后验光流场概率分布,第一解码器330和/或第二解码器360可以被配置为通过基于光流场的像素变换来产生预测帧。

[0090] 在一个实施例中,训练单元340可以被配置为在对先验编码器310和后验编码器320进行训练时分别训练第一解码器330和第二解码器360。

[0091] 在一个实施例中,训练单元340还可以被配置为交替或同步地进行对后验编码器320的训练和对先验编码器310的训练。

[0092] 关于示例装置300中的各个单元和/或模块的功能和操作,还可以参考上文参考图2到图7描述的视频预测方法中的步骤,重复之处不再赘述。

[0093] 根据本公开的实施例的视频预测装置(例如,图8所示的示例装置300)可以实现在各种终端设备(例如辅助驾驶的车载设备)中。

[0094] 在一个实施例中,根据本公开的实施例的视频预测装置可以作为一个软件模块和/或硬件模块而集成到所述终端设备中。例如,图8所示的示例装置300可以是该终端设备的操作系统中的一个软件模块,或者可以是针对于该终端设备所开发的一个应用程序,其运行在CPU(中央处理单元)和/或GPU(图形处理单元)上,或者运行在专用的硬件加速芯片,例如适于运行深度神经网络的专用芯片上。另外,示例装置300也可以作为该终端设备的众多硬件模块之一。

[0095] 在另外的实施例中,根据本公开的实施例的视频预测装置与该终端设备可以是分立的设备,并且可以通过有线和/或无线网络连接到该终端设备,并且按照约定的数据格式来传输交互信息。

[0096] 图9示出根据本公开的实施例的电子设备的示例10。如图9所示,电子设备10可以包括一个或多个处理器11和存储器12。

[0097] 处理器11可以是诸如CPU这样的通用处理装置或计算装置,也可以是具有数据处理能力和/或指令执行能力的其他形式的处理单元,并且可以控制电子设备10中的其他组件以执行期望的功能。

[0098] 存储器12可以包括一个或多个计算机程序产品。这样的计算机程序产品可以包括各种形式的计算机可读存储介质,例如易失性存储器和/或非易失性存储器。易失性存储器例如可以包括随机存取存储器(RAM)和/或高速缓冲存储器(cache)等。非易失性存储器例如可以包括只读存储器(ROM)、硬盘、闪存存储器等。在该计算机可读存储介质上可以存储一个或多个计算机程序指令。处理器11至少可以运行该一个或多个程序指令,以执行

根据本公开的各个实施例的视频预测方法。在计算机可读存储介质中还可以存储诸如先前帧、后续帧、已知帧等各种数据。

[0099] 在一个实施例中,电子设备10还可以包括输入装置13和输出装置14,这些组件通过总线系统和/或其他形式的连接机构(未示出)互连。根据不同的实施例,输入装置13可以包括但不限于键盘、鼠标等。输出装置14可以向外部输出各种信息(包括作为预测结果的视频帧等),并且可以包括但不限于显示器、扬声器、打印机、通信网络及其所连接的远程输出设备等。

[0100] 为了简洁,图9的示例中仅示出电子设备10中与本公开有关的一些组件,而省略了诸如总线、输入/输出接口等其他组件。根据需要,电子设备10还可以包括任何其他适当的组件。

[0101] 另外,本公开的实施例还包括一种计算机程序产品,其包括计算机程序指令,该计算机程序指令在被处理器运行时使处理器执行根据本公开的实施例的视频预测方法中的步骤。

[0102] 根据不同的实施例,上述计算机程序产品可以是使用一种或多种程序设计语言的任意组合来编写用于执行本公开实施例操作的程序代码。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务上执行。

[0103] 此外,本公开的实施例还可以包括一种计算机可读存储介质,其上存储有计算机程序指令,该计算机程序指令在被处理器运行时使处理器执行根据本公开的实施例的视频预测方法中的步骤。

[0104] 根据不同的实施例,上述计算机可读存储介质可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以包括但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者以上的任意组合。可读存储介质可以包括但不限于:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪速存储器)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0105] 贯穿说明书和权利要求书,除非上下文清楚地另有要求,否则措词“包括”、“包含”等应当以与排他性或穷尽性的意义相反的包括性的意义来解释,也就是说,应当以“包括但不限于”的意义来解释。另外,措词“在本文中”、“上文”、“下文”以及相似含义的措词在本申请中使用时应指作为整体的本申请,而不是本申请的任何具体部分。在上下文允许时,在使用单数或复数的以上描述中的措词也可以分别包括复数或单数。例如,前文所提到的“第一障碍物”或“第二障碍物”可以指一个或多个的“第一障碍物”或“第二障碍物”,并且“另一”也可以表示另外的一个或多个。关于在提及两个或多个项目的列表时的措词“或”,该措词涵盖该措词的以下解释中的全部:列表中的任何项目,列表中的所有项目,以及列表中的项目的任何组合。另外,措词“第一”、“第二”等旨在用于区分,而不是用于强调次序或重要程度。

[0106] 虽然已经描述了本公开的一些实施例,但是这些实施例仅作为示例而呈现,而不打算限制本公开的范围。实际上,在本文中所描述的方法和系统可以采用多种其他形式来实施。另外,可以在不脱离本公开的范围的情况下,在本文中所描述的方法和装置的形式上

做出各种省略、替换和改变。

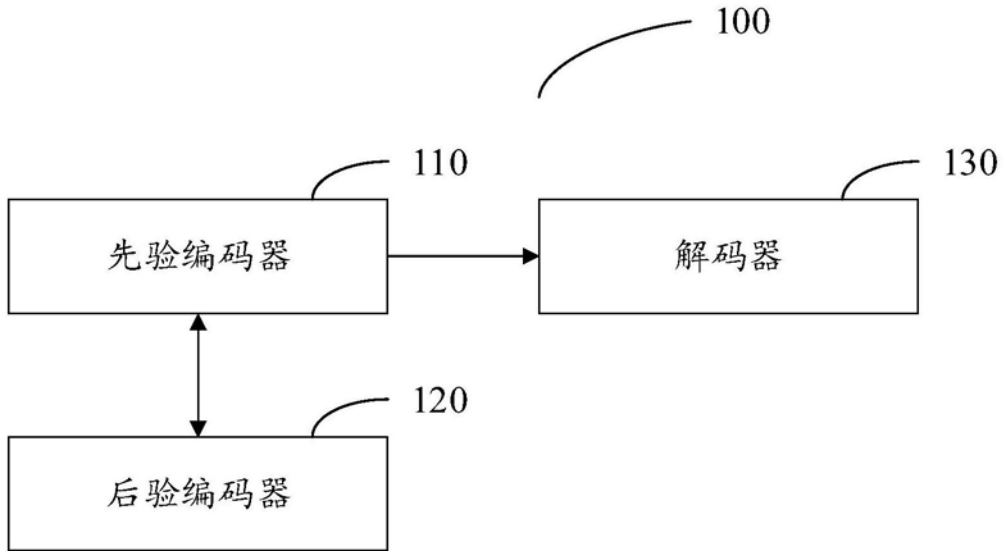


图1

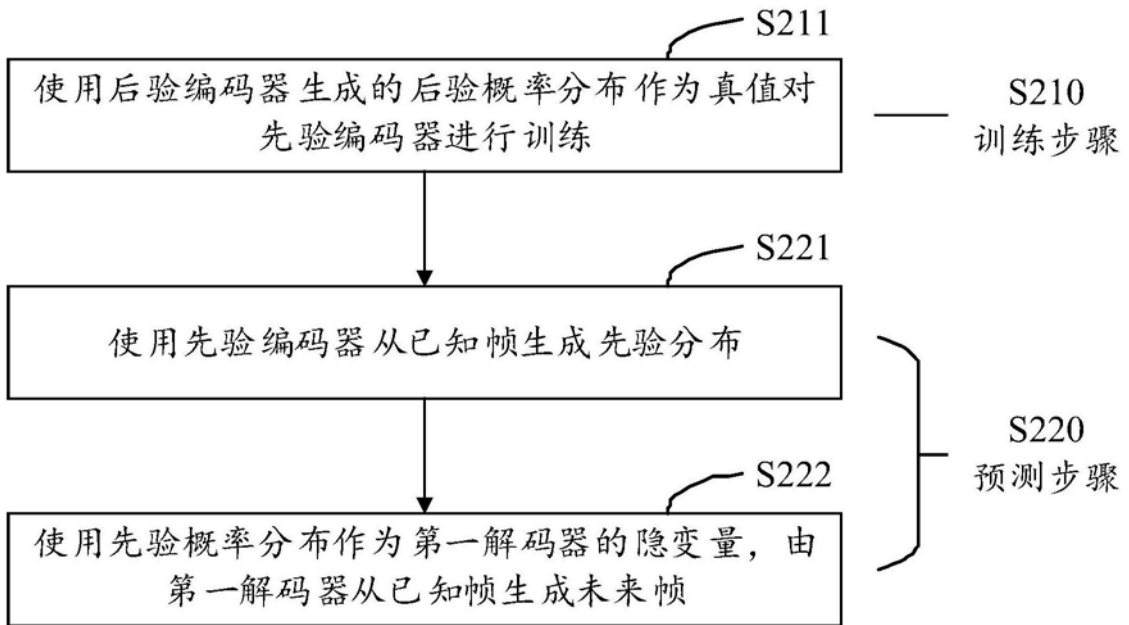


图2

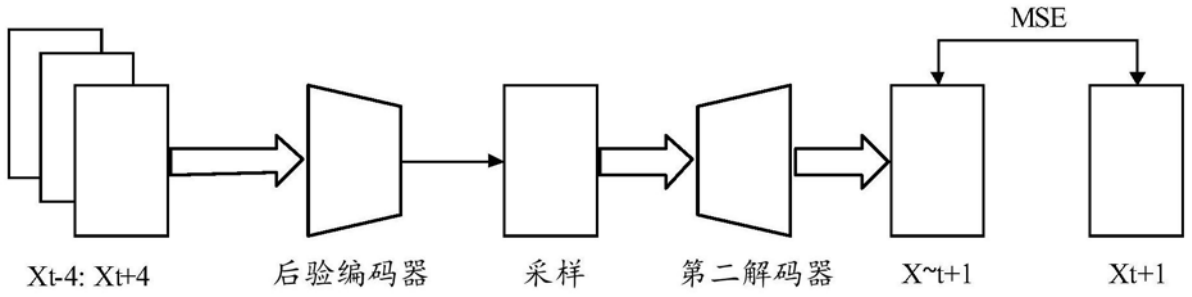


图3

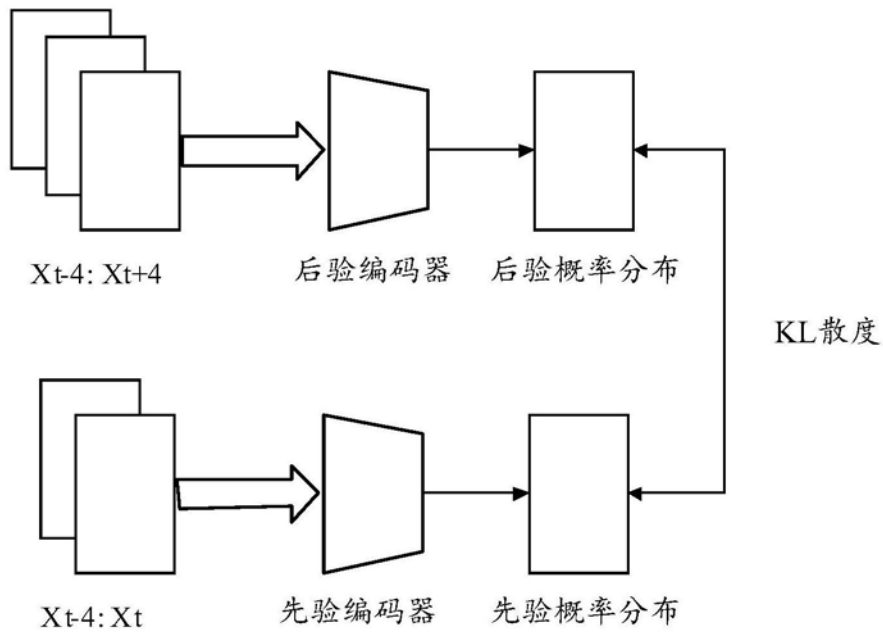


图4

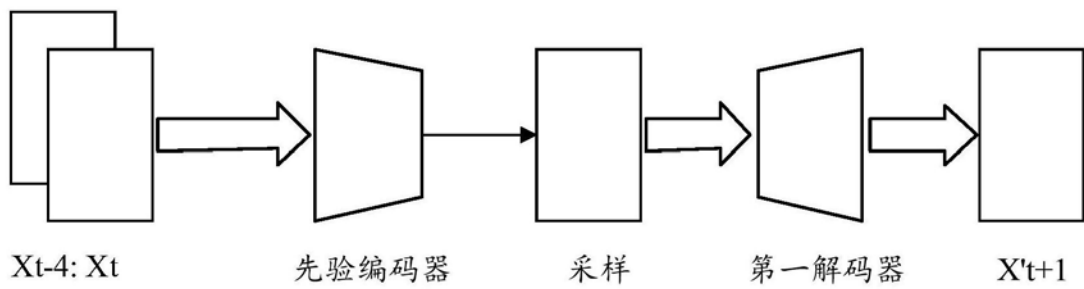


图5

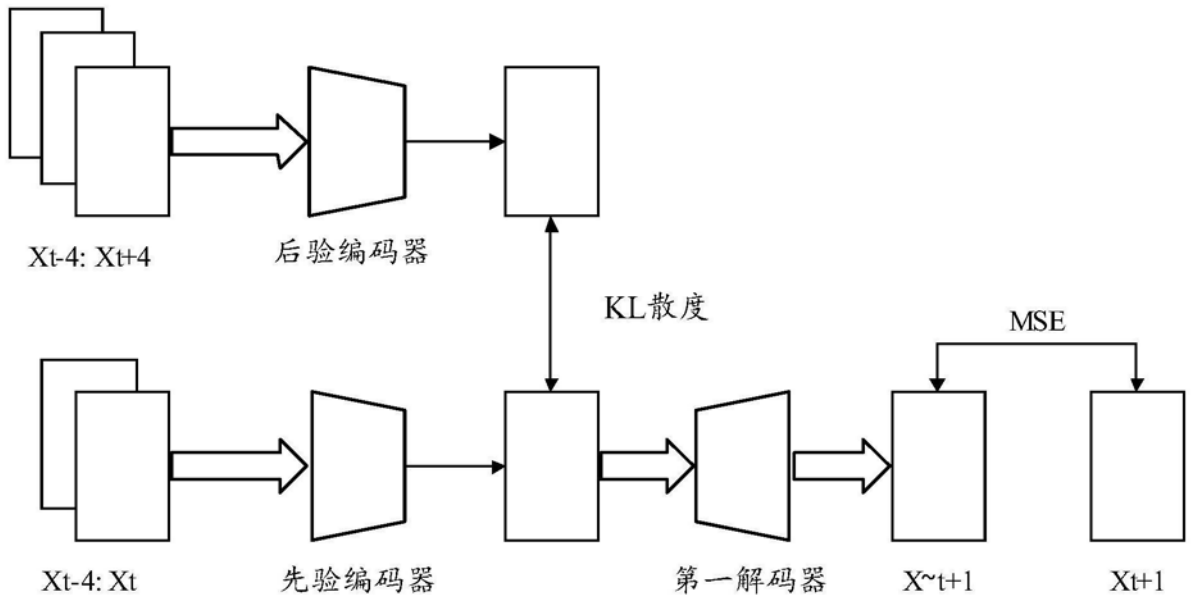


图6

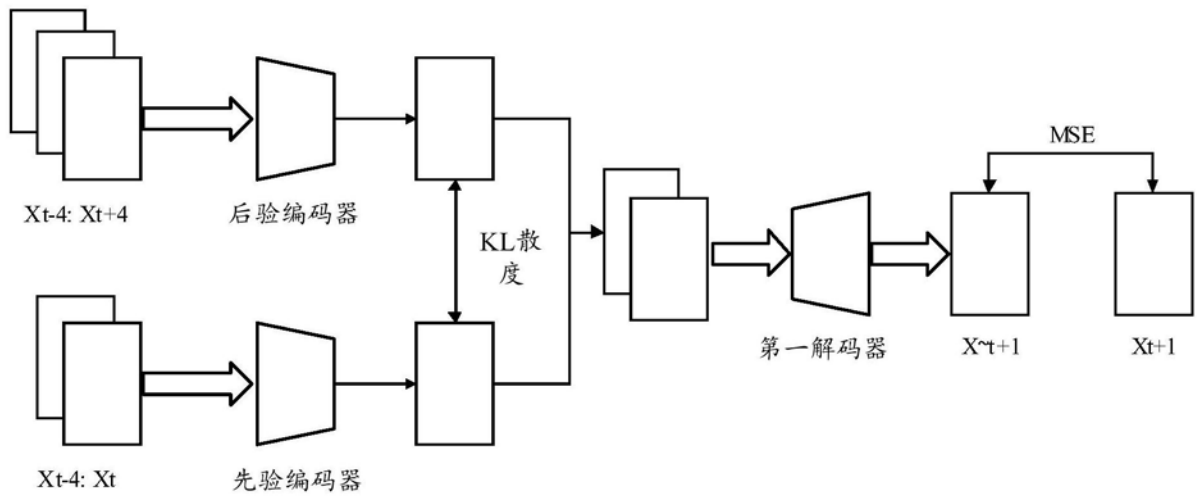


图7

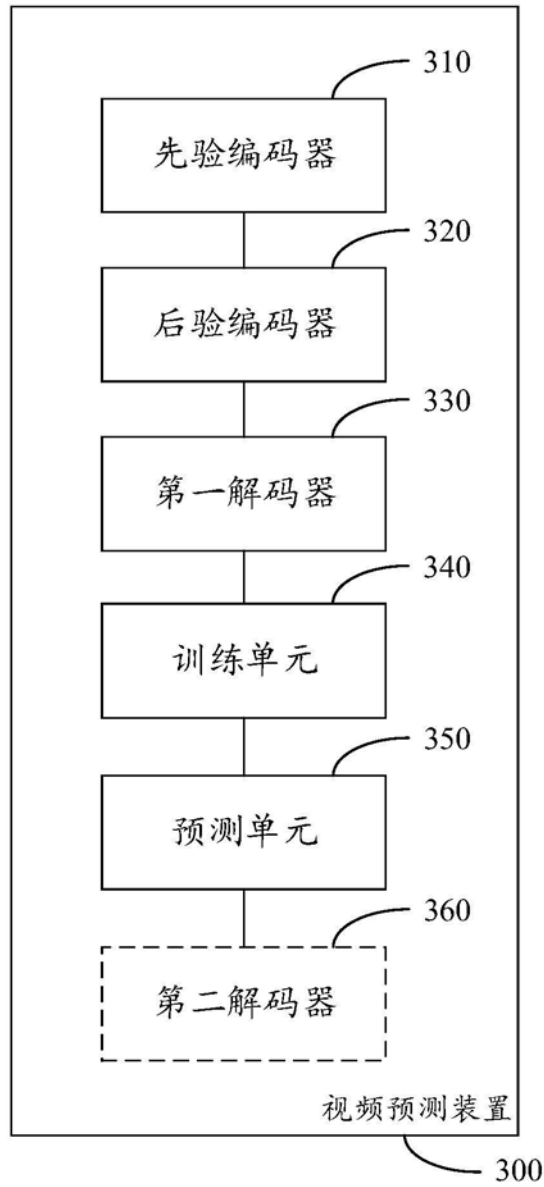


图8



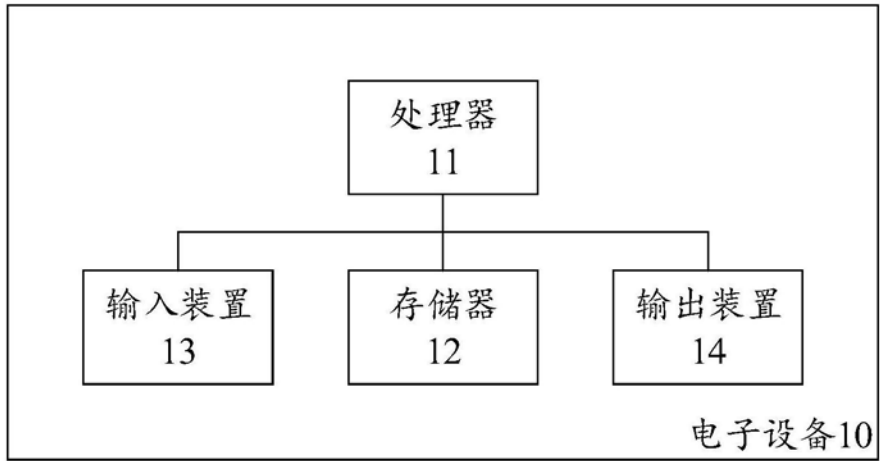


图9