(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 107624061 B
(45)授权公告日 2021.01.22

(21)申请号 201680032611.1
(22)申请日 2016.04.20
(65)同一申请的已公布的文献号
申请公布号 CN 107624061 A
(43)申请公布日 2018.01.23
(30)优先权数据
62/150,068 2015.04.20 US
(85)PCT国际申请进入国家阶段日
2017.12.04
(86)PCT国际申请的申请数据
PCT/US2016/028406 2016.04.20
(87)PCT国际申请的公布数据
WO2016/172188 EN 2016.10.27
(73)专利权人 康奈尔大学
地址 美国纽约州
(72)发明人 S·尼伦伯格
(74)专利代理机构 北京市万慧达律师事务所 11111
代理人 白华胜 王蕊
(51)Int.Cl.
A61F 9/08(2006.01)
A61N 5/06(2006.01)
G06K 9/46(2006.01)
G06N 3/04(2006.01)
G06K 9/00(2006.01)
H04N 19/60(2014.01)
H04N 19/85(2014.01)

(56)对比文件
CN 103890781 A,2014.06.25
CN 1985514 A,2007.06.20
US 2003/0235341 A1,2003.12.25
US 2010/0165117 A1,2010.07.01
US 2010/0008568 A1,2010.01.14
CN 1770177 A,2006.05.10

审查员 尹尹

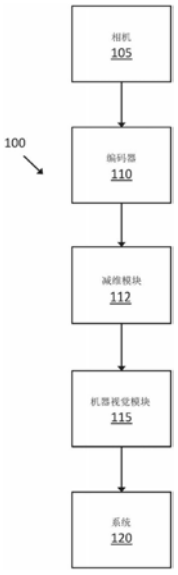
权利要求书2页 说明书21页 附图14页

(54)发明名称

具有维度数据缩减的机器视觉

(57)摘要

描述一种方法,包括接收对应于一系列原始图像的原始图像数据,和用处理装置的编码器处理所述原始图像数据以生成经编码数据。所述编码器的特征在于基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换。所述方法还包括通过对所述经编码数据应用减维算法处理所述经编码数据以生成减维经编码数据。所述减维算法被配置成压缩包含于所述经编码数据中的信息的量。此外描述可与这类方法一起使用的设备和系统。



CN 107624061 B

1.一种通过处理装置处理原始图像数据的方法,包括:

通过处理装置接收对应于一系列原始图像的原始图像数据;

用所述处理装置的编码器处理所述原始图像数据以生成经编码数据,其中所述编码器的特征在于基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换;和

通过处理器处理所述经编码数据以生成减维经编码数据,其中处理所述经编码数据包括对所述经编码数据应用减维算法,其中所述减维算法被配置成压缩包含于所述经编码数据中的信息,并且其中所述减维算法选择用于特定机器视觉任务的所述经编码数据的特征的子集,并忽略用于所述特定机器视觉任务的所述经编码数据的其他特征。

2.根据权利要求1所述的方法,其中所述经编码数据包含一系列经编码视网膜图像,并且其中所述方法还包括从用于特定机器视觉任务的所述经编码数据的特征的子集中生成特征标记数据,其中所述特征标记数据包括具有多个分量的向量,所述多个分量包括与多个视网膜图像区域的各个区域的所述特征的子集的值。

3.根据权利要求2所述的方法,其中所述特征标记数据包含与多个视网膜图像区域相关的信息。

4.根据权利要求3所述的方法,其中所述特征标记数据包含对应于所述多个视网膜图像区域中的每个的运动数据。

5.根据权利要求4所述的方法,其中所述运动数据包含对应于所述多个视网膜图像区域中的每个的速度数据。

6.根据权利要求3所述的方法,其中所述特征标记数据包含对应于所述多个视网膜图像区域中的每个的光流数据。

7.根据权利要求1所述的方法,其中所述处理所述经编码数据包含对所述经编码数据应用经训练的算法。

8.根据权利要求7所述的方法,其中所述经训练的算法包含卷积神经网络。

9.根据权利要求7所述的方法,其中所述经训练的算法已在经编码训练数据的训练数据集上训练,并且其中所述经编码训练数据已使用训练编码器编码,所述训练编码器的特征在于基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换的输入/输出转换。

10.根据权利要求9所述的方法,其中所述经编码训练数据的训练集包含虚拟环境的经编码图像,并且所述原始图像数据包含现实环境的原始图像。

11.根据权利要求9所述的方法,其中所述经编码训练数据的训练集包含在第一组条件下获得的图像,并且其中所述原始图像数据包含在不同于所述第一组条件的第二组条件下获得的原始图像。

12.根据权利要求11所述的方法,其中所述第一组条件和所述第二组条件包含光照条件。

13.根据权利要求1所述的方法,进一步包含对所述减维经编码数据应用机器视觉算法。

14.根据权利要求13所述的方法,其中所述处理所述经编码数据以生成减维经编码数据是在所述处理所述原始图像数据以生成经编码数据之后并且在所述对所述减维经编码

数据应用所述机器视觉算法之前执行。

15.根据权利要求1所述的方法，其中所述处理所述原始图像数据以生成经编码数据包含生成相对于所述原始图像数据在维度上缩减的经编码数据，并且其中所述处理所述经编码数据以生成所述减维经编码数据包含另外压缩相对于所述原始图像数据在维度上已经缩减的所述经编码数据。

16.根据权利要求15所述的方法，其中包含于所述经编码数据中的所述信息的量被至少压缩为所述对应的原始图像数据的约1/2，并且其中所述减维经编码数据被至少压缩为所述对应的经编码数据的约1/2。

17.根据权利要求15所述的方法，其中包含于所述经编码数据中的所述信息的量被至少压缩为所述对应的原始图像数据的约1/10，并且其中所述减维经编码数据被至少压缩为所述对应的经编码数据的约1/10。

18.一种处理原始图像数据的设备，包含：

被配置成存储原始图像数据的至少一个存储器存储装置；和

至少一个处理器，所述至少一个处理器可操作地与所述存储器耦接并且被编程以：

接收对应于一系列原始图像的所述原始图像数据；

使用基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换处理所述原始图像数据以生成经编码数据；和

通过对所述经编码数据应用减维算法，处理所述经编码数据以生成减维经编码数据，其中所述减维算法被配置成压缩包含于所述经编码数据中的信息的量，其中所述减维算法选择用于特定机器视觉任务的所述经编码数据的特征的子集并忽略用于所述特定机器视觉任务的所述经编码数据的其他特征。

19.根据权利要求18所述的设备，进一步包含可操作地耦接到所述至少一个处理器的机器人装置，其中所述机器人装置包含被配置成生成所述原始图像数据的至少一个图像传感器。

20.一种具有计算机可执行指令的非暂时性计算机可读介质，在通过计算装置执行时，使所述计算装置执行包含以下的操作：

接收对应于一系列原始图像的原始图像数据；

使用基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换处理所述原始图像数据以生成经编码数据；和

通过对所述经编码数据应用减维算法处理所述经编码数据以生成减维经编码数据，其中所述减维算法被配置成压缩包含于所述经编码数据中的信息的量，其中所述减维算法选择用于特定机器视觉任务的所述经编码数据的特征的子集并忽略用于所述特定机器视觉任务的所述经编码数据的其他特征。

具有维度数据缩减的机器视觉

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C.§119要求2015年4月20日提交的美国临时专利申请第62/150,068号的优先权,其全部内容以引用的方式并入本文中。本申请此外以引用的方式并入美国临时申请第61/527493号(2011年8月25日提交)、第61/657406号(2012年6月8日提交)、第61/308,681号(2010年2月26日提交)、第61/359,188号(2010年6月28日提交)、第61/378,793号(2010年8月31日提交)、第61/382,280号(2010年9月13日提交)和第13/230,488号(2011年9月12日提交)。本申请此外以引用的方式并入在本文中被称作“假体应用(Prosthesis Applications)”的国际专利申请第PCT/US2011/026526号(2011年2月28日提交)和PCT/US2011/049188(2011年8月25日提交)、在本文中被称作“机器视觉应用(Machine Vision Application)”的国际专利申请第PCT/US2012/052348号(2012年8月24日提交)。前述申请中的每一个的内容均以其对应的全文引用的方式并入。

背景技术

[0003] 提供以下描述来帮助读者理解。所提供信息或所引用参考文献均不承认为现有技术。机器视觉(或计算机视觉)是指允许机器(例如计算机)使用视觉信息例如从图像提取信息以宽的或有限的含义解决一些任务,或“理解”情况的技术。一般来说,机器视觉关注从图像数据提取信息。图像数据可采取许多形式,如单个图像、视频序列、来自多个相机的视图,或更高维数据(例如来自医学扫描仪的三维图像)。已经提出若干方法用于发展基于视觉系统的操作的抽象模型的机器视觉的图像数据预处理方案,但是其效用受到限制。仍存在许多改善机器视觉的空间,使得其可更有效并且实用。

发明内容

[0004] 前述发明概括仅仅是说明性的,且并不旨在以任何方式作为限制。除了上文描述的说明性方面、实施例和特征之外,通过参考以下图式以及详细描述,另外的方面、实施例以及特征将变得显而易见。

[0005] 在说明性中,描述了包括通过处理装置接收对应于一系列原始图像的原始图像数据和用处理装置的编码器处理原始图像数据以生成经编码数据的方法。编码器的特征在于基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换。方法还包括通过对经编码数据应用减维算法,通过处理器处理经编码数据以生成减维经编码数据。此后者减维算法被配置成进一步压缩包含于经编码数据中的信息以使其适用于机器视觉算法。具体地说,此附加压缩允许从经编码数据抽出特征-适用于机器视觉但是在初始原始图像中不可获得的特征。

[0006] 此外,在一些实施方案中,可建立通过附加减维生成的特征以产生特征标记。特征标记包括与多个经编码图像区域相关的信息。特征标记可通过多个经编码图像区域携带比其分量特征更多(即,更具描述性)的信息。举例来说,在一些实施例中,通过对经编码数据应用减维的抽出特征可包括速度信息,所述速度信息包括不与不同光照条件和环境混淆的

特定种类的速度信息。在标记中的多个区域可为对应于人体区域的区域集合。在此实例中，特征标记可通过创建其分量为单个区域中的速度(即，与单个身体部位的运动相关联的速度，如使用经编码数据所测量)的向量来生成。应注意，如上所述，通过使用经编码数据而不是原始图像，可在每个区域中计算速度而不与其它因素(如人在其中移动的光照条件或环境等)混淆。此向量、此标记可随后用于在标记捕获人体移动的许多识别特征(如步态、热量倾斜、身体部位的相对大小等)时识别人，即，区分他或她与其他人。类似地，多个区域或扇区可对应于物体的区域，使得区分人类或其它生物与无生命(即，刚性)物质。

[0007] 在一些实施方案中，经编码数据可表示为一个或一系列经编码图像(通常被称作一系列视网膜图像或经编码视网膜图像)，并且处理经编码数据包括处理经编码的一系列视网膜图像以生成特征或特征标记。如上，特征标记可包括与多个视网膜图像区域相关的信息。

[0008] 在一些实施方案中，处理经编码数据包括对经编码数据应用经训练的算法。经训练的算法可包括卷积神经网络(CNN)。经训练的算法可已在经编码训练数据的训练数据集上训练，并且经编码训练数据可已使用特征在于基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换的输入/输出转换的训练编码器编码。在另一实施方案中，经编码训练数据的训练集包括虚拟环境的经编码图像，并且原始图像数据包括现实环境的原始图像。经编码训练数据的训练集可包括在第一组条件下获得的图像，并且原始图像数据可包括在不同于第一组条件的第二组条件下获得的原始图像。第一组条件和第二组条件可包括不同光照条件。

[0009] 在另一个实施方案中，方法可包括对减维经编码数据应用机器视觉算法。处理经编码数据以生成减维经编码数据可在处理原始图像数据以生成经编码数据之后并且在对减维经编码数据应用机器视觉算法之前执行。处理原始图像数据以生成经编码数据可包括生成相对于原始图像数据在维度上缩减的经编码数据，并且处理经编码数据以生成减维经编码数据可包括另外压缩相对于原始图像数据在维度上已经缩减的经编码数据。包含于经编码数据中的信息的量被至少压缩为所述对应的原始图像数据的约1/2，并且其中所述减维经编码数据被至少压缩为所述对应的经编码数据的约1/2。包含于经编码数据中的信息的量被至少压缩为所述对应的原始图像数据的约1/10，并且其中所述减维经编码数据被至少压缩为所述对应的经编码数据的约1/10。

[0010] 在另一个实施方案中，设备包括被配置成存储原始图像数据的至少一个存储器存储装置和可操作地与存储器耦接的至少一个处理器。处理器被编程以接收对应于一系列原始图像的原始图像数据并且使用基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换处理原始图像数据以生成经编码数据。处理器进一步被编程以通过对经编码数据应用减维算法处理经编码数据以生成减维经编码数据。减维算法可被配置成压缩包含于经编码数据中的信息的量。设备可进一步包括可操作地耦接到至少一个处理器的机器人装置，其中机器人装置包含被配置成生成原始图像数据的至少一个图像传感器。

[0011] 在另一个实施方案中，具有计算机可执行指令的非暂时性计算机可读介质在通过计算装置执行时使计算装置执行包括接收对应于一系列原始图像的原始图像数据，使用基本上模拟脊椎动物视网膜的至少一个视网膜细胞的输入/输出转换的输入/输出转换的处

理原始图像数据以生成经编码数据，和通过对经编码数据应用减维算法处理经编码数据以生成减维经编码数据的操作。减维算法被配置成压缩包含于经编码数据中的信息的量。

附图说明

[0012] 通过以下描述和所附权利要求书结合附图理解，本公开的前述和其它特征将变得更加完全清楚。应理解，这些图式只描绘了根据本公开的若干实施例，并且因此不应被视为限制了本公开的范围，将通过使用这些附图用额外的特性和细节来描述本公开。

[0013] 图1为根据说明性实施例具有机器视觉的装置的框图。

[0014] 图2为根据说明性实施例示出编码器模块的操作的流程图。

[0015] 图3A描绘根据说明性实施例通过由实例视网膜导航仪控制的现实世界环境的机器人导航的轨迹。

[0016] 图3B描绘根据说明性实施例通过由实例像素导航仪控制的现实世界环境的机器人导航的轨迹。

[0017] 图描绘4根据各种说明性实施例通过不同环境的视网膜导航仪和像素导航仪的结果的直方图。

[0018] 图5描绘根据说明性实施例表明在具有各种光照条件的各种环境中视网膜导航仪和像素导航仪表现如何的图。

[0019] 图6A到6E描绘根据说明性实施例通过像素导航仪和视网膜导航仪具有各种光照条件的各种环境做出的正确决策的直方图。

[0020] 图7描绘根据说明性实施例在视频中人类辨识的阶段。

[0021] 图8描绘根据说明性实施例使用经编码图像(此处被称作虚拟视网膜图像(VR))和原始图像数据识别人的卷积神经网络的效用的直方图。

[0022] 图9描绘根据说明性实施例经编码图像(此处被称作虚拟视网膜图像(VR))以创建人脸的运动标志的实例使用。

[0023] 图10描绘根据说明性实施例机器视觉的装置的框图。

具体实施方式

[0024] 在以下详细描述中参考附图，附图形成本文的一部分。在图式中，相似的符号通常识别相似的组件，除非上下文另外规定。在详细描述、图式以及权利要求书中所描述的说明性实施例并不旨在为限制性的。在不脱离本文呈现的主题的精神或范围的情况下，可利用其它实施例，并且可作出其它改动。将容易了解的是，如本文中所一般描述且在图中说明的本公开的各方面可以广泛多种不同配置来布置、取代、组合和设计，所有广泛多种不同配置明确被预期并构成本公开的部分。

[0025] 机器视觉具有在相对简单的任务(如用于计数在生产线上通过的物体的工业系统)到更复杂任务(如面部辨识)和感知任务(例如允许机器人导航复杂环境)的范围内的大量应用。机器视觉应用的实例的非限制性列表包括用于控制过程(例如工业机器人或自控车辆)、检测事件(例如用于视觉监视或人数计数)、组织信息(例如用于为图像和图像序列的数据库编索引)、建模物体或环境(例如工业检测、医学图像分析或表面形状建模)和交互(例如输入到计算机-人类交互的装置)的系统。

[0026] 在一些应用中，机器视觉可涉及高度计算昂贵任务。单一颜色数字图像可由数百万的像素或更多构成。每个像素可具有相关联值，如在颜色空间(例如熟悉的RGB颜色空间、YCbCr空间、HSV空间等)中限定像素的坐标的多个(例如8个或24个)比特值。视频流可包括对应于每秒数百兆比特或更多的比特率的在例如每秒几十帧的帧率下这类图像的序列。许多机器视觉应用需要这类图像或视频流的快速处理(例如跟踪物体的运动并且对其做出反应，以在物体沿装配线移动时识别物体或对物体进行分类，以允许机器人对其环境做出实时反应等)从而使机器视觉实用。

[0027] 在这类时间限制下处理这类大量的数据可极其具有挑战性。在一些实施例中，使用减少信息的原始量的技术来处理图像数据。在这类实施例中，技术还可保持(或甚至增强)对于在手边的机器视觉任务突出的图像数据的特征。经预处理的图像数据(而不是原始数据)可输入到包括机器学习算法的机器视觉系统，由此使得算法学习之前难以或不可能学习的图像的特征。这降低在系统上的处理负荷，使学习易控制并且允许快速响应和可能改善的性能。

[0028] 脊椎动物眼睛的视网膜提供此种类的图像处理。即，脊椎动物眼睛的视网膜可接受视觉刺激并且将刺激转换成可被大脑理解的形式。此系统(经数百万年进化的过程发展)高效并且有效，如通过在哺乳动物(尤其猴和人类)中的高等级的复杂视觉感知证明。

[0029] 用于基于脊椎动物视觉系统的操作的抽象模型发展用于机器视觉的图像数据预处理方案的一些方法可基于视网膜的实际性能的粗略近似，但是这些尚未高效。

[0030] 视觉世界极其复杂。人类大脑连续浸没视觉信号，然而信号被快速解析和利用。对于这类速度的一个原因为通过视网膜执行的减维。数百万年的进化已将视网膜形成强有力的减维系统。脊椎动物视觉系统获取视觉输入，抽出所需的并且丢弃剩余的。如果利用此减维，那么机器视觉的能力将大大提高。

[0031] 如在机器视觉应用中所论述，我们已开发执行此减维的视网膜输入/输出关系的模型。它与其它视网膜模型不同，因为它推广到任意复杂度的刺激(面部、景观、人行走等，参见机器视觉应用和假体应用)。模型所做的是收缩(在本文中也被称作压缩)存在于视觉世界中的信息的量成易控制形式。收缩形式可随后用作包括机器学习和机器视觉算法(两者在本领域中已知并且在本公开中描述)的其它算法的输入。收缩形式使使得算法可以学习视觉环境和借助其以比当算法直接依赖于正常、原始、高维视觉输入时可能的惊人地更有效的方式执行任务在本公开中，我们描述此方法对一系列任务的应用，所述任务包括导航、面部辨识、人辨识、物体辨识、情绪/表达辨识、经济/地质学/气候等数据的趋势辨识、疾病检测(例如使用医学图像，如MRI图像、CT扫描图像、病理学切片等)和通过动物或人类很好执行的其它辨识/检测任务。

[0032] 在机器视觉应用中，我们展示方法在包括导航任务的应用中的效用，因为这尤其具有挑战性(需要在空间和时间两者上处理)。我们使用使用开放源Panda3D图形程序包构造的虚拟现实3D环境作为环境。使用的学习算法来自“深度学习”家族，具体地说，卷积神经网络(CNN)。我们在两个条件下比较导航仪的性能：当导航仪使用正常原始图像作为其输入时和当导航仪使用通过视网膜模型(即编码器)过滤的图像时。我们将第一个称为“像素导航仪”并且将第二个称为“视网膜导航仪”。如举例来说在国际申请第PCT/US2012/052348号(参见例如图10和11)中所示，当学习算法使用后者-从视觉环境提取相关特征的算法时，性

能基本上较好(好许多倍),并且非常重要的是,学习推广到到其它环境(乡村、郊区、不同光照条件等)、对于人工系统已经高度难以解决的问题。

[0033] 在本公开中,我们描述在一些实施例中呈现以下优点中的一些或全部的技术:1)将减维方法的功率转换成现实世界环境,和2)添加附加减维方法,生物启发和人工两者,以创建用于面部辨识、人辨识、物体辨识、情绪/表达辨识、经济/地质学/气候等数据的趋势辨识和动物或人类表现良好的其它辨识/检测任务的新算法。

[0034] 在一个方面,公开一种方法,包括:接收对应于一系列原始图像的原始图像数据;用编码器处理原始图像数据以生成经编码数据,其中编码器的特征在于输入/输出转换基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换;和通过应用被配置成压缩包含于经编码数据中的信息的量的减维算法,处理经编码数据以生成减维经编码数据。

[0035] 各种实施方案可单独或以任何合适的组合包括以上描述的装置、技术和其元件中的任一个。

[0036] 下文是涉及用于教示目标语言的方法和系统的各种概念的更详细描述和用于教示目标语言的方法和系统的实施方案。上文引入并且下文更详细地论述的各种概念可以多种方式中的任一种实施,因为描述的概念不限于实施方案的任何特定方式。主要出于说明性目的提供了特定的实施方案和应用的实例。

[0037] 在一些实施例中,可使用提供通过视网膜执行的操作的几乎完全复制的编码器。如详细描述于上文以引用的方式并入的各种国际专利申请中(此后“假体应用”),这类编码器可用于开发高效视网膜假体。在一些实施例中,编码器应用于机器视觉。

[0038] 当用作预处理步骤(具体来说,减维步骤或维度偏移)时,编码器可基本上增强机器视觉算法的性能。在一些实施例中,编码器可允许机器视觉算法在广泛范围的环境和光照条件中有效地提取信息。在其中现有机器视觉算法部分有效的情况下,减维可充当强增强子。编码器可允许更有效(产生例如较高性能)以及更快和更高效地进行提取。

[0039] 由编码器输出的数据可通过使用另一种减维算法进一步处理。附加处理可在应用机器视觉算法之前(或作为其一部分)完成。此附加减维算法可为生物启发或部分或全部人工的。如在机器视觉应用中所论述,编码器将原始图像转换成新表示,例如视网膜编码的图像。这些新表示为原始图像中不存在的表示。表示可有利地用于进一步减维处理,其允许初始原始图像数据的更有效和高效处理。因此,附加减维算法可应用于视网膜编码的图像(包括其中包括的附加特征或表示)。应注意,视网膜编码的图像根本上不同于原始图像,并且因此对原始图像执行减维产生之前未生成并且不可通过直接对原始图像应用减维方法生成的用于机器学习的特征(或参数)(例如编码器创建随时间进化的新颖模式;因为模式通过对图像的神经元响应而不是图像自身生成,所以选择捕获模式的参数不同于可用于捕获原始图像的特征的那些)。

[0040] 在一些实施例中,学习算法(例如机器视觉算法)可在已使用视网膜编码器和一个或多个附加减维方法处理的数据(例如图像)上训练。如例如在本文所提供的实例中所描述,发明人已发现,在广泛多种应用中,在这类经处理数据集上训练的算法(与例如原始图像数据相反)呈现增强的性能。另外,发明人已发现增强的性能为高度可泛化的。即,甚至当呈现有在与在训练图像中发现的那些显著不同的条件(例如光照条件、环境调色板等)下生

成的输入(例如图像)时,各种算法也表现良好。在一些情况下,发明人已发现,甚至当呈现有显著不同于虚拟训练环境的现实世界环境时,在已使用本文所述的技术处理的虚拟图像(例如基于环境的虚拟绘制生成的图像)上训练的算法也表现良好。

[0041] 如在本文所提供的实例中所示,使用本文所述的技术处理(例如特征在于视网膜编码随后一个或多个附加减维方法)的训练数据集可提供增强的训练环境用于包括用于导航、面部识别、身体识别等的机器视觉算法的广泛范围的学习算法。

[0042] 一些实施例可采用详细地描述于假体应用中的类型的视网膜编码过程。这类实施例可包括接收刺激并且用一组编码器将刺激转换成一组代码的装置。装置还可使用接口将代码转换成信号。信号可用高分辨率变换器激活多个视网膜细胞。多个视网膜细胞的活化可引起对广泛范围的刺激的视网膜神经节细胞响应,这可基本上类似于来自对相同刺激作出响应的哺乳动物视网膜的视网膜神经节细胞的时间依赖性响应。用于这类装置的编码器可适于处理用于机器视觉应用中的图像数据并且不必限于视网膜假体。

[0043] 在假体应用中描述的视网膜假体(像正常视网膜)可为图像处理器。视网膜假体可从接收的刺激提取必需的信息并且可将信息重新格式化成大脑可理解的动作电位的模式。通过哺乳动物视网膜生成的动作电位的模式在视网膜的代码或神经节细胞的代码中。视网膜假体可将视觉刺激转换成此相同代码或其紧密代理,使得受损或退化的视网膜可生成正常或几乎正常的输出。因为视网膜假体使用与正常视网膜或相同的代码其紧密代理,所以在受损或退化的视网膜中的神经节细胞的放电模式与正常神经节细胞的放电模式相同(或基本上类似)。因此,这类假体可允许视网膜与指示视觉世界的正常视网膜信号相同(或基本上类似)发送到大脑。

[0044] 如在假体应用中所论述,编码器可使用用于视网膜细胞的输入/输出模型。输入/输出模型可使用从对多种刺激(例如白噪声(WN)和自然场景(NS)影片)的哺乳动物视网膜细胞的输入/输出响应的研究获得的数据生成。在一些实施例中,编码器可基于包括通过多个参数来表征的时空转换的线性非线性级联模型。这类参数可基于使用哺乳动物视网膜通过实验获得的数据优化。使用参数可产生接近模拟哺乳动物细胞对广泛范围的刺激的响应的人工过程。结果可为捕获用于自然图像(静态或空间变化),如面部、景观、人行走、儿童玩耍等的输入/输出关系的模型。模型还可用高斯统计学捕获用于白噪声刺激或刺激的输入/输出关系。

[0045] 一些实施例可充分利用通过实验获得的数据。因此,生成的编码器可精确模拟视网膜处理而不需要视网膜的基础处理方案的详细抽象理解。举例来说,据相信,在灵长类动物和人类中的视网膜处理突出显示在适用于图案辨识任务(例如面部辨识)的视觉刺激中的特征同时不强调或消除其它特征(例如冗余信息或噪声)以允许在大脑中高效处理。迄今为止,没有对此处理方案的细节的完全抽象理解,其由于在万古的过程中自然选择而发展。尽管如此,即使如此缺乏抽象理解,但是本文所述的装置和技术可通过精确模拟视网膜响应捕获此处理的益处。

[0046] 换句话说,在本文所述的各种实施例中,模型或模块可为数据驱动的。模型或模块可为视网膜输入/输出关系的数据驱动的模型并且提供真实图像预处理。模型或模块可将已预处理的信号输出到下游机器视觉算法。预处理步骤可完成与生物视网膜相同(或基本上类似)种类和相同(或基本上类似)幅度的减维。因此,这类模型或模块可提供与哺乳动物

视网膜相同的优点。

[0047] 在各种实施例中，操作视网膜处理以减少来自原始图像数据的信息的总量同时保持对于给定应用的突出特征。以类似于哺乳动物视网膜实现功效的方式的方式减少信息的总量可为高效的。这类过程在本文中被称作“减维”或“生物减维(BDR)”。举例来说，在一些实施例中，即使减少在视网膜经编码数据中的信息的总量，当视网膜经编码数据减少时，机器视觉算法可呈现比如果机器视觉算法接收原始图像数据更好的性能。更好的性能基本上是向机器视觉算法提供通过视网膜接收的最重要信息并且减少或丢弃对机器视觉算法的特定任务不太重要的一些信息的结果。

[0048] 在各种实施例中，视网膜经编码数据可为原始视觉数据(例如通过视网膜接收)的至少压缩为1/1.5、至少压缩为1/2、至少压缩为1/3、至少压缩为1/4、至少压缩为1/5或至少压缩为更多分之一(如分母在1到100范围内或其任何子范围内)的压缩形式。压缩可对应于通过编码器生成的减维。举例来说，在一些实施例中，视网膜编码器的比特率可定量并且可与被编码器用作刺激的原始图像数据的熵(也以比特每单位时间为单位测量)相比，并且采用比率以确定压缩比。举例来说，在假体应用中描述的一些情况下，编码器用与4.9bit/s的输入原始数据比特速率相比的2.13bit/s秒的比特速率描述。因此，通过编码器生成的数据压缩在此实例中接近7倍。

[0049] 如本文所述，一些实施例可包括对经编码数据应用附加减维算法(生物启发的、人工的或其组合)。在各种实施例中，减维视网膜经编码数据可以被进一步至少压缩到1/1.5、1/2、1/3、1/4、1/5或更多分之一(如分母在1到100的范围内或其任何子范围内)。在一些实施例中，此压缩对应于通过经编码数据的额外处理生成的减维。举例来说，在一些实施例中，后处理经编码数据的比特速率可定量，并且可与预处理编码器输出的比特速率相比，并且采用比率以确定压缩比。

[0050] 图1为根据说明性实施例具有机器视觉的装置100的框图。在替代实施例中，装置100可具有附加、更少或不同元件。另外，在图1中使用箭头不意指关于元件顺序的限制。装置100包括相机105、编码器110、机器视觉模块115和系统120。在一些实施例中，系统120可至少部分通过机器视觉模块控制。相机105接收视觉刺激并且将其转换成数字图像数据。在一些实施例中，数字图像数据可为数字图像流。此数字图像数据在本文中可被称作“原始”图像数据。原始图像数据可包括在通过视网膜编码器处理之前的任何图像数据。

[0051] 在一些实施例中，编码器110可接收图像数据并且使用一个或多个视网膜编码器处理图像数据。在一些实施例中，视网膜编码器可具有本文和/或在假体应用中所述的类型。编码器110可将所接收的图像数据转换成与由哺乳动物视网膜输出的格式相同(或基本上类似)的格式。编码器110的输出(被称作经编码视网膜图像数据)传递到减维模块112。减维模块112处理来自编码器110的输出，以提供超出通过编码器110执行的任何减维的附加减维。举例来说，减维模块112可从编码器110接收经编码视网膜图像流并且生成减维经编码数据。换句话说，编码器110通过编码来自原始图像的信息创建经编码视网膜图像数据。减维模块112使用来自经编码视网膜图像数据的特征用于多种机器视觉任务。减维模块112根据对于特定机器视觉任务的需要仅使用由经编码视网膜图像数据表示的特征的子集并且忽略对于特定机器视觉任务不必要的特征。

[0052] 减维模块112的输出传递到机器视觉模块115。机器视觉模块115可通过使用本领

域中已知和/或本文所述的一种或多种机器视觉技术处理数据。机器视觉模块115可生成可用于任何合适目的的输出。如图1所示，机器视觉模块115的输出可发送到一个或多个系统120。在一些实施例中，系统120可为机器人系统。在一些实施例中，图像处理和/或控制可实时或近实时执行。

[0053] 在图1中示出的系统意指仅为示例性的，并且可使用各种其它类型的机器视觉系统。举例来说，在一些实施例中，可不包括受控制的系统120。举例来说，可存储机器视觉模块115的输出用于进一步处理而不是直接用于控制。在一些实施例中，相机105可被存储的图像数据的来源代替。举例来说，相机105可被存储数字视频的存储装置代替。在一些实施例中，附加元件可包括在装置100中，如各种处理器、控制器、用户控件、输入或输出装置等。

[0054] 在一些实施例中，还可建立减维经编码数据以生成特征标记。特征标记包括与多个经编码视网膜图像区域或扇区相关的信息。特征标记可通过多个图像区域携带比其分量特征更多并且更具描述性的信息。在一些实施方案中，特征标记可使用从经编码数据提取的相对简单的特征(如横向运动、横向运动速率、竖直运动、竖直运动速率、密度、开关速率、在经编码数据中的其它特征)允许例如人、面部和物体的识别或分类。可创建向量，其多个分量为在每个区域中的特征的值(例如在每个区域中横向运动的速率，如使用经编码数据所测量)。此向量、此标记可随后用于执行机器视觉任务。举例来说，对于面部辨识，在标记中的区域(在向量中的分量)对应于在面部中的区域(例如覆盖左眉毛的区域、覆盖左眼的区域等)。如果特征为例如横向运动的速率(即，横向速度)，那么向量捕获单个面部的若干识别特征，如眉毛和眼睛的相对移动。这类向量可用于区分面部和由面部生成的情绪。另外的细节在下文提供，其中说明性实施例(图7到9)在“实例2-视觉任务”中描述。类似地，多个区域或扇区可对应于身体的区域(例如头部、上臂、下臂等)。如果特征再次为横向运动的速率，那么向量可捕获人体的若干识别特征，如步态、头部倾斜、身体部位的相对大小等。这类向量可用于区分人。同样，另外的细节在下文在“实例2-视觉任务”中的图7到9的详细描述中提供。同样地，多个区域还可对应于物体的区域，使得人类或其它生物区别于无生命(例如刚性的)物质。多个区域还可对应于移动场景的区域，使得标记捕获整体特性，例如光流。应注意，因为经编码数据可含有一系列图像(经编码数据的影片)，所以标记可为时间序列数据的向量，如图中所示。

[0055] 在各种实施例中，相机105可为能够将视觉刺激转换成数字形式(例如数字图像流)的任何装置。各种实施例可包括基于电荷耦接装置(CCD)的相机105；有源像素传感器(APS)，如互补金属氧化物半导体(CMOS)传感器、薄膜晶体管(TFT)，或光电二极管的阵列；和其组合。

[0056] 通过相机105生成的数字图像中的每个可包括至少0.01百万像素、至少0.1百万像素、至少1百万像素、至少2百万像素或更多，例如在的0.01到1000百万像素范围或其任何子范围内。数字图像流的特征可在于至少10Hz、至少50Hz、至少100Hz或更多，例如在1到1000Hz的范围或其任何子范围内的帧率(例如每秒图像帧的数量)。数字图像可为彩色、灰阶、黑白，或其它合适类型的图像。

[0057] 在一些实施例中，相机105可基于电荷耦接装置(CCD)。在一个实施例中，相机105可为Point Grey牌型号Firefly MV图像捕获装置(能够在60帧/秒下以8比特/像素用752×480像素捕获图像)。在另一个实施例中，相机105可为e-con Systems牌型号e-CAM50 OMAP

GSTIX图像捕获装置。相机105可集成OmniVision牌型号OV5642相机模块，其能够在30帧/秒下以8比特/像素用1280×720像素捕获图像。

[0058] 在一些实施例中，图像可通过相机105获得并且以足以允许装置100运行而没有不期望滞后时间的速度传输到编码器110。在一些实施例中，可在相机105和编码器110之间提供高带宽连接。举例来说，大于20MB/sec的数据传送可在相机105和编码器110之间使用通用串行总线(USB)2.0接口实现。在其它实施例中，可在相机105和编码器110之间使用并行接口，如在Texas Instruments牌OMAP 3530型号处理器上集成到相机图像信号处理器中的并行接口。在各种实施例中，可使用其它合适的连接，包括有线或无线连接。相机105可使用能够高速数据传送的任何连接与编码器110接口，包括(但不限于)串行接口，如电气电子工程师学会(Institute of Electrical and Electronics Engineers，IEEE)标准1394，USB 2.0、USB 3.0、USB 3.1等；并行接口；模拟接口，如国家电视系统委员会(National Television System Committee，NTSC)标准或逐行倒相(PAL)标准；或无线接口。在一些实施例中，相机105可集成到与编码器110相同的电路板上。

[0059] 编码器110可使用本文所述的技术(包括例如实施将图像转换成代码的编码器，模拟视网膜回路的操作)处理图像流。即，编码器110可接收图像并且输出与由给予相同图像的哺乳动物视网膜输出的数据流相同(或基本上类似)的数据流。通过编码器实施的转换可应用到一系列输入图像，生成经编码输出。举例来说，经编码输出可呈类似于如果哺乳动物视网膜已接收图像而应已产生的视网膜细胞的放电速率的值的形式。在一些实施例中，编码器110的输出还可为例如模拟视网膜细胞的发生器电位的信息。发生器电位可为视网膜模型的线性分量的输出(借助线性滤波器的输入图像的卷积的输出)。如果哺乳动物视网膜细胞用相同输入图像刺激，那么编码器110的经编码输出可为将通过哺乳动物视网膜细胞产生的尖峰的脉冲串或脉冲。

[0060] 在一些实施例中，不同编码器110的集合可用于模拟通过哺乳动物视网膜的处理。在哺乳动物视网膜中存在不同类型的视网膜输出细胞，并且各种编码器110可用于复制不同类型的哺乳动物视网膜输出细胞。差异可对应于特定细胞类型(例如“打开”细胞或“关闭”细胞)或对应于在视网膜上的细胞位置(例如在中心视网膜中的“打开”细胞和在视网膜周边中“关闭”细胞)。在其中装置100具有多于一个编码器110的实施例中，编码器110可独立地或通过至少一个或多个耦接机构并行操作。

[0061] 图2为根据说明性实施例示出编码器模块110的操作的流程图。在替代实施例中，可执行附加、更少和/或不同操作。另外，使用流程图不意指关于执行操作的顺序进行限制。在操作205中，编码器模块110可接收一个或多个图像。在一些实施例中，可从相机105(或一些其它合适的来源)接收一个或多个图像。在操作210中，一个或多个图像可经预处理。举例来说，一个或多个图像可经预处理以重新按比例缩放一个或多个图像的对比度和/或强度，以对一个或多个图像应用静噪滤波器以裁剪一个或多个图像等。在一些实施例中，操作210可不执行。

[0062] 在操作215中，一个或多个图像可经编码。一个或多个图像可以模拟哺乳动物视网膜细胞对图像的响应的方式处理。举例来说，在一个实施例中，对于在图像场中的各种位置，可输入一个或多个图像，并且可输出对应于如果一个或多个图像通过哺乳动物视网膜处理那么将已通过哺乳动物视网膜细胞(或细胞群)产生的放电速率的时间依赖性值。在一

个实施例中，放电速率输出如下格式化：对于给定时间t，输出为比特矩阵，其中在位置(x,y)处的元素对应于在位置(x,y)处的视网膜细胞的放电速率。

[0063] 在一些实施例中，操作215可包括使用除放电速率之外的度量产生模拟哺乳动物视网膜细胞的响应的信息。举例来说，编码器的输出可对应于细胞的活化状态、胞内电位、上述发生器电位等。

[0064] 在操作220中，来自操作215的经编码信息可用于生成一个或多个经编码视网膜图像。在一些实施例中，经编码视网膜图像可适合于通过机器视觉模块115处理。举例来说，其中如上所述经编码信息输出为放电速率的矩阵，可生成放电速率视网膜图像，其中在经编码视网膜图像中每个像素的强度通过在矩阵中的对应元素的放电速率值确定(对于实例参见图3)。可使用在放电速率和像素强度之间的任何合适的关系，包括线性关系、非线性关系、多项式关系、对数关系等。操作220可包括使用包括使用查询表的任何合适的技术将像素强度转换成放电速率。在一些实施例中，放电速率可使用除强度之外的图像特征在经编码视网膜图像中表示。举例来说，在其中经编码视网膜图像为彩色图像的实施例中，每个像素的颜色空间坐标可对应于放电速率。

[0065] 在操作225中，经编码视网膜图像进行附加减维处理，其可包括处理一个或多个经编码图像以生成减维经编码数据。可建立减维经编码数据以产生特征标记数据。，如在上文所论述和另外下文在图7到9的描述中进一步详细论述。

[0066] 附加处理还可应用于经编码视网膜图像，包括重新按比例调整、过滤、裁剪、平滑等。

[0067] 在操作230中，一个或多个视网膜图像可输出到机器视觉模块115。一个或多个视网膜图像可使用任何合适的方法(其可为有线或无线)输出。在一些实施例中，视网膜图像可类似于哺乳动物视网膜如何输出视网膜图像来输出。

[0068] 实例

[0069] 实例1-导航处理器

[0070] 机器视觉应用描述在虚拟现实环境中有效的导航仪。然而，根据本公开的一个实例，导航仪还可在现实世界环境中有效。因此，本公开的各种实施例可用于多个环境和多个应用中。

[0071] 举例来说，可使用被称作Turtlebot的移动机器人。Turtlebot为被设计用于机器人技术开发和测试的开放源个人机器人。Turtlebot在机器人操作系统(ROS)上运行，这有助于硬件和通信机制，并且将来自在机器人上的传感器和硬件组件的数据一起带入单个软件框架。机器人包括来自被称作iClebo Kobuki的Yujin机器人的3轮圆形机车基座、包括相机的微软Kinect传感器，和用于机器人的较好传感器输入和状态估计的机载工厂校准陀螺。膝上型计算机机载放置机器人，并且动态处理所有传感器数据以作出在线实时决策。膝上型计算机运行Linux操作系统(Ubuntu)，并且ROS在下方一层以在膝上型计算机上控制和处理所有机器人数据。

[0072] 然而，这类实例仅为本公开的具体实施方案。其它实施例可包括附加、更少或不同元件。举例来说，一般实施例可包括机车元件、视觉感测元件(例如相机)和被配置成从视觉感测元件接收数据并且为机车元件提供指令的处理元件。

[0073] 在一个测试条件中，忽略Kinect的三维相机系统并且仅使用二维彩色相机。因此，

通过装置的软件作出的决策基于二维图像。装置的软件被配置成从相机接收输入图像。输入图像通过装置分成7个区域。软件确定每个区域的速度和每个区域适配到十二类别中的哪些。软件随后用最慢的速度选择区域并且指导装置在最慢区域的方向上行进。

[0074] 上文所述过程允许装置导航环境而几乎没有或没有碰撞。在实例1中，类似的前端软件如用于在机器视觉应用中描述的虚拟导航任务使用。即，允许虚拟导航仪确定在虚拟世界中的速度的类似训练集用于确定在实例1中现实环境中的速度。在实例1中，软件的改变包括在速度确定之后的步骤的修改和与在有限、现实世界空间中的装置形状的框架相关的导航规则的小改变。具体地说，软件通过在决策之间的时间期间积累所有对帧的数据(例如在1秒时间段内29对帧)和确定七个区域中的哪些在所有的帧对上具有最慢的速度来决定是否向前移动或转弯。

[0075] 即，装置每秒作出一次是否转弯或继续向前的决策。每秒，相机捕获二十九个图像。每个图像被分成七个区域。每个区域的速度通过比较最近捕获的图像的区域与先前捕获的图像的对应区域来确定。装置随后确定每个区域的速度属于十二个类别中的哪个(例如A-L，其中A为最慢的并且L为最快的)。装置记录对应于每个区域的速度类别，并且确定从最后一个决策开始七个区域中的哪个在最低类别中具有最多记录。举例来说，对于区域1到7和类别A到L，如果最左区域1在类别A中具有一个记录并且剩余的记录在类别B到L中，并且在类别A中没有其它区域具有任何记录，那么装置确定在区域1的方向上转弯。在另一个实例中，如果最左区域1在类别B中具有两个记录，并且剩余的记录在类别C到L中，并且最右区域7在类别B中具有一个记录，并且剩余的记录在类别C到L中，那么装置确定面向区域1。通过从最后一个决策开始对所有对帧求平均值，作出更稳定的决策。

[0076] 图3A和3B示出根据说明性实施例通过现实世界环境导航的实例1的装置的性能。图3A示出使用“视网膜导航仪”的装置的性能。图3B示出使用“像素导航仪”的装置的性能。装置的围长通过在图中的圆的直径指示。如图3A和3B所示，装置的围长相对大，并且装置必须很好地操纵以避免与壁和障碍物碰撞，其中Xs指示椅子并且黑色正方形指示废纸篓。在两种情况下，机器人以0.4米每秒(m/s)向前移动。

[0077] 图3A示出当装置通过“视网膜导航仪”控制时装置的轨迹。视网膜导航仪可为在机器视觉应用中描述的导航仪，其从通过视网膜模型过滤的影片(例如减维视觉世界)学习在其视觉环境中的速度。减维视网膜模型可为与关于上文图1和2描述的过程一致的模型。

[0078] 图3B示出当装置通过“像素导航仪”控制时装置的轨迹。像素导航仪可为在机器视觉应用中描述的导航仪，其从未过滤的影片(例如未通过视网膜模型减维的影片)学习在其视觉环境中的速度。在两种情况下(在图3A和3B中示出)，导航仪在Panda-3虚拟环境中训练。在其它实施例中，不同虚拟环境可用于训练导航仪。

[0079] 如图3A所示，当装置通过视网膜导航仪控制时，装置操纵通过在现实环境中的障碍物而不运行到壁或障碍物中。相比之下，如图3B所示，当装置通过像素导航仪控制时，装置撞到壁和障碍物中。装置在任何情况下都不能够使其通过航线。

[0080] 在图3A和3B中示出的实例的不同版本中，装置使用上文所述视网膜导航仪。装置随后在朝向站立的人的开放空间中的路径上启动。在装置接近站立的人时，装置检测将在当前轨迹上出现碰撞并且转向以避开站立的人。当人移动到装置的轨迹中时，装置再次转向以避开与人碰撞。此过程重复若干次，并且每次装置转向以避开与人碰撞。

[0081] 另一个实例包括具有如上文所解释的软件的装置。在一个版本中，使用视网膜导航仪，如上文参考图1和2所论述，过滤输入图像。在另一个版本中，使用像素导航仪，其中未执行输入图像的过滤。每个导航仪使用相同图像训练。每个导航仪随后在不同于用于训练导航仪的环境的各种虚拟环境中运行。举例来说，在训练期间使用的虚拟环境为乡村环境的复制品。每个导航仪随后放置于乡村虚拟环境中，并且导航仪表现类似，很好地导航通过环境并且很少运行到障碍物中。视网膜导航仪使其约95％的时间通过障碍航线，而像素导航仪约使其90％的时间通过障碍航线。

[0082] 然而，当导航仪在没有重新训练的情况下放置于具有不同光照条件不同环境中时，不同导航仪不同地执行。视网膜导航仪具有相对直的轨迹，并且很少运行到障碍物中，并且具有与在乡村环境中的导航仪类似的性能。然而，像素导航仪具有杂乱的轨迹(看起来无规的)并且通常运行到障碍物中。在三种不同环境中，视网膜导航仪使其约90％到95％的时间通过障碍航线。然而，像素导航仪使其约0％到25％的时间通过障碍航线。

[0083] 图4示出视网膜导航仪和像素导航仪通过不同环境的结果的直方图。图4的左边的图说明如何确定最慢的区域的正确决策。如上所述，每个图像被分成七个区域，每个区域为图像的竖条。在图4的左边的图中示出的实例中，中间区域为具有最慢移动的区域。因此，如果导航仪确定例如中间区域右边的第二区域为最慢的区域，那么决策在图4的右边的图上接收+2的评分。因此，在此情形下，当应保持直线时导航仪将面向右侧方向移动。

[0084] 图4的最左边的图示出在乡村环境(类似于训练环境)中通过像素导航仪(在上部)和视网膜导航仪(在下部)作出的正确决策的直方图。如图所示，视网膜导航仪比像素导航仪作出更多正确决策(评分为0的决策)和更少不正确决策(不同于0的决策)。但是，与图4右边的图相比，视网膜导航仪和像素导航仪表现类似。在图4右侧的图示出示出在不同于乡村环境并且具有与乡村环境不同光照条件的三种环境中通过像素导航仪(在上部)和视网膜导航仪(在下部)作出正确决策的数量的直方图。如图4所示，视网膜导航仪当放置于不同环境中时作出与当其放置于乡村环境中时类似数量的正确决策。然而，像素导航仪在不同环境中比其在乡村环境中作出更多和更大错误。因此，相对于以上图1和2描述的视网膜图像处理在没有图像处理的情况下大大改善装置的性能。

[0085] 图5示出根据说明性实施例表明在具有各种光照条件的各种环境中视网膜导航仪和像素导航仪表现如何的图。图5的最左边图示出像素导航仪(左条形图)和视网膜导航仪(右条形图)成功使其在具有两种导航仪训练的光照条件的乡村环境中通过障碍航线的时间分数。如图所示，两种导航仪使其大于90％的时间通过航线。各种其它图示出像素导航仪(左条形图)和视网膜导航仪(右条形图)在各种光照条件(即，太阳在位置A到F处，其中太阳的位置为从左边30°、40°、50°、90°(空中竖直)、150°和160°)下成功使其通过各种环境(具有训练物体的乡村、具有新颖物体的乡村、郊区和游乐场)的时间分数。如图所示，像素导航仪很少使其通过到达航线的终点。实际上，像素导航仪的最好表现为略微高于40％的时间，并且那是使用测试乡村环境，但是具有不同障碍物。在其它环境中，像素导航仪表现差。然而，视网膜导航仪在所有环境和所有光照条件中使其高于90％的时间成功地通过航线。

[0086] 类似于在图4中示出的图，图6A到6E的图示出视网膜导航仪和像素导航仪通过具有各种光照条件的各种环境的结果的直方图。图6A示出在具有与训练环境相同光照条件的乡村环境中通过像素导航仪(在上部)和视网膜导航仪(在下部)作出的正确决策的直方图。

图6B到6E示出通过在具有各种光照条件的各种环境中像素导航仪和视网膜导航仪作出正确决策的直方图。如在各种直方图中所示，在不同于训练环境的环境中，像素导航仪比视网膜导航仪作出更多和更大错误(非零评分)。

[0087] 总之，图3A、3B、4、5和6A到E示出视网膜(或在本申请中，编码器)的减维可允许机器学习算法学习可普遍适用于视觉任务(如导航)的特征。结果示出在一个训练集中学习的特征允许导航仪在包括现实世界环境的许多环境中非常有效地操纵通过空间。

[0088] 视网膜或编码器的减维还可允许机器学习算法学习看起来普遍适用于其它任务的特征，如面部辨识、人辨识、物体辨识、情绪/表达辨识、经济、地质、气候等数据的趋势辨识、疾病检测(例如使用医学图像如MRI图像、CT扫描图像、病理学切片等)和动物或人类表现良好的其它辨识/检测任务。为清楚起见，编码器执行初始减维，并且随后一个或多个附加减维步骤随后执行(对经编码数据)，作为单独的步骤或作为并入到机器学习过程中的步骤，以抽出特征。在一些实施例中，给定特征或特征集合可从在经编码图像中的多个区域建立；建立的特征可生成特征标记。特征标记可识别面部、物体、情绪等，如上文所述并且在通篇的各个图中示出。

[0089] 实例2-视觉任务

[0090] 如上所述，在机器视觉应用中所描述的方法允许机器学习算法高效并且以一般化方式学习视觉世界的特征。这类方法通过减少视觉输入的维度(例如使用视网膜编码)实现此。本申请关注对经编码数据应用一个或多个附加减维方法，其方式为使得机器学习算法(例如卷积神经网络，或CNN)当搜索参数空间(例如在CNN中发现权重)时，找到一般解决方案，而不是落入本地解决方案(例如由于在参数空间中本地最小值)。

[0091] 举例来说，在各种实施例中，用于给定训练数据集的解决方案可为捕获转换(或计算或映射)的权重集合。减维训练集可允许算法找到推广到多个问题的解决方案。即，经训练的算法可从视觉信息的各种来源提取有用信息(或执行任务)。当呈现有来自单个样本的图像集合(例如训练集)时和当呈现有样本外图像集合(例如来自在不同光照条件下的不同环境的图像集合等)时，算法可表现良好。

[0092] 举例来说，在虚拟导航仪的情况下，视觉输入传递通过编码器，也被称作虚拟视网膜。经编码数据随后被分成7个区域，并且每个区域被传递到CNN中。随后训练CNN以提取在每个区域中的特征、平均速度。每个区域的速度被分类为十二个类别中的一个。训练CNN以通过最小化在其输出和真实速度之间的误差确定速度，这可由Panda 3D虚拟环境直接确认。由于此训练在CNN中出现的权重为一般的，并且在完全新的环境(例如此外已以相同方式通过编码器和附加减维步骤减维的完全新的视觉输入)中具有精确的速度估计值。

[0093] 在一些例子中，如描述于上实例1中，这类训练的结果可一般化以适用于包括与虚拟训练环境显著不同的现实世界环境的图像的不同环境(例如完全不同物体和光照条件等)。

[0094] 在各种实施例中，上文所述的原理可用于为机器视觉提供工具。

[0095] 在一些实施例中，通过编码器提供的减维可为起点，并且其它减维(生物启发或以其它方式)可与它组合。

[0096] 涉及横向或竖直运动、速度、平均放电速率、闪烁速率的经编码数据的减维为可以此方式使用的实例。这些特征对于捕获人、物体、数据等可能有价值。因此，在各种实施例

中，使用减维方法(通过编码器的原始图像中的第一个并且随后对经编码数据的另外减维，以提取特征并且生成特征标记)可允许执行大量机器视觉任务，如导航、人辨识、面部辨识、物体辨识、情绪/表达辨识、经济、地质、气候等数据的趋势辨识、疾病检测(例如使用医学图像如MRI图像、CT扫描图像、病理学切片等)和动物或人类表现良好的其它辨识和/或检测和/或分类任务。

[0097] 图7示出根据说明性实施例在视频中人辨识的阶段。在一个实施例中，区域或区段速度可用作用于人辨识的工具。在实例实施例中，通过编码器也被称作虚拟视网膜(VR)预处理人行走的3,500个视频的训练集，如上所述。编码器的输出转换回视频，分成栅格(如图7所示)，并且输入到CNN中。CNN对于每个栅格区域作出速度估计，如上文所论述。

[0098] 图7的速度标记包括具有速度分量的向量。在图7的实例中，人的身体转换成48×N分量的向量，其中48为在栅格中的区域的数量，并且N为根据其测量速度的帧对的数量。在各种实施例中，可使用其它合适的数量的栅格元件和对或帧。

[0099] 在一些实施例中，速度标记可用于识别个体。图8示出使用经编码数据(虚拟视网膜数据)和使用原始图像数据识别人的CNN的效用的直方图。使用速度标记之间的欧几里德(Euclidean)(例如二维)距离，有可能正确地将在视频剪辑中的个体与在另一个剪辑中的同一个体匹配。这类匹配可为“人辨识”。

[0100] 图8示出使用含有十一个不同人的3,500个视频剪辑的速度标记的人辨识的实例的结果。速度标记具有行走的人的全身。速度标记通过CNN评估，并且CNN负责基于速度标记识别人。用于每个试验，选择3,500个剪辑的参考剪辑，并且CNN将参考剪辑的速度标记与其它3,499个剪辑的速度标记比较。最类似于参考剪辑的速度标记的其它3,499个剪辑的速度标记通过CNN选择。选择的剪辑随后与参考剪辑相比，以确定选择的剪辑和参考剪辑是否包含同一行走的人。图8示出对于十一个人中的每个的正确百分比。对于每个人，使用经编码数据(通过虚拟视网膜生成的数据)的正确百分比示出为左侧条形图，并且使用原始图像数据的正确百分比示出为右侧条形图。使用虚拟视网膜，约90％的时间CNN将参考剪辑与包含同一人的剪辑正确地匹配。然而，如图8所示，当使用原始图像数据时，CNN具有低得多的成功率。

[0101] 身体、物体、面部、动画数据(气象图、经济数据等)具有移动部分，并且可利用移动以捕获在空间和时间两者中的移动。图7示出使用行走的人的全身图像的实例。图9示出使用虚拟视网膜创建人脸的运动标记的实例。可选择栅格的大小使得在栅格中的每个区域主要在一个方向上移动。举例来说，对于全身图像(如在图7中的那些)，上臂可在一个方向上移动，下臂可在另一个方向上移动等。对于面部图像，眼眉可在一个方向上移动，下嘴唇可在另一个方向上移动等。通过选择捕获主要在单个方向上移动的栅格大小，CNN可容易地捕获在每个区域中的速度。运动方向在速度标记中隐含地捕获，因为身体部位或面部部位实时移动(例如上臂在相对于其它身体部位的某一方向上移动，并且这包含于向量分量中)。因此，大量关于人的物理对象可包含于这类标记中。对于物理(通常)固体物体，组成部分的移动方向、组成部分的相对移动和运动的自然流动在时间和空间上从栅格的一个区域到另一个。因此，使用这类标记允许对于识别所需要的信息用相对少数目的参数(例如构成速度标记的参数)捕获。

[0102] 在各种实施例中，这类方法可以多种形式实施。举例来说，各种区域的速度可使用

CNN提取。在其它实施例中，可使用其它学习网络如支持向量机。在另外其它实施例中，可使用其它固定方法如光流算法(例如Fahrenback算法)。在各种实施例中，可修改各种参数，如栅格区域的数量、栅格区域的大小和形状、在视频剪辑中的帧的数量、用于估计速度的帧的数量、估计的速度的数量等。还可修改用于比较速度标记的度量。在上文所论述的一些实例中，测量欧几里德距离。然而，测量欧几里德距离的一些替代方案包括通过深度学习分类器、贝叶斯(Bayesian)分类器、支持向量机、Lebesgue空间(L^p)距离、马氏(Mahalanobis)距离确定的距离等。

[0103] 图10为根据说明性实施例的机器视觉的装置1000的框图。在替代实施例中，可使用附加、更少或不同元件。机器视觉装置1000可包括处理器1005、存储器1010、收发器1015、图像捕获装置1020、电源1025、用户接口1030和定位装置1035。

[0104] 在一些实施例中，机器视觉装置1000可包括处理器1005。处理器1005可被配置成进行和/或导致进行本文所述的一个或多个操作。处理器1005可执行本领域的技术人员已知的指令。指令可由专用计算机、逻辑电路或硬件电路执行。因此，处理器1005可在硬件、固件、软件中实施，或这些方法的任何组合。术语“执行”为运行应用程序或进行由指令要求的操作的过程。可使用一种或多种编程语言、脚本语言、汇编语言等编写指令。处理器1005执行指令，意指其执行由所述指令要求的操作。处理器1005可操作地与存储器1010、收发器1015、图像捕获装置1030、电源1025、用户接口1030等耦接，以接收、发送和处理信息并且控制机器视觉装置1000的操作。处理器1005可从永久存储器装置(如只读存储器(ROM)装置)检索一组指令并且将指令以可执行形式复制到通常某一形式的随机存取存储器(RAM)的暂时性存储器装置。机器视觉装置1000可包括使用相同或不同处理技术的多个处理器。在说明性实施例中，指令可存储在存储器1010中。

[0105] 在一些实施例中，机器视觉装置1000可包括存储器1010。存储器1010可为用于信息的电子保持位置或储存器，使得信息可由如本领域的技术人员已知的处理器1005访问。存储器1010可包括(但不限于)任何类型的随机存取存储器(RAM)、任何类型的只读存储器(ROM)、任何类型的闪速存储器等，如磁性存储装置(例如，硬盘、软盘、磁带等)、光盘(例如，紧密光碟(CD)、数字通用光盘(DVD)等)、智能卡、闪存装置等。机器视觉装置1000可具有使用相同或不同存储器媒体技术的一个或多个计算机可读介质。机器视觉装置1000可具有支持存储器媒体的加载的一个或多个驱动器，如CD、DVD、闪存卡等。在一些实施例中，存储器1010可用于存储视频。因此，在一些实施例中，可使用存储器1010代替图像捕获装置1020，以为机器视觉装置1000提供视频。

[0106] 在一些实施例中，机器视觉装置1000可包括收发器1015。收发器1015可被配置成接收和/或传输信息。在一些实施例中，收发器1015可经由有线连接，如乙太网连接、一个或多个双绞线电线、同轴电缆、光缆等传送信息。在一些实施例中，收发器1015可经由使用微波、红外波、无线电波、扩频技术、卫星等的无线连接传送信息。收发器1015可被配置成使用蜂窝式网络、局域网、广域网、因特网等与另一个装置通信。在一些实施例中，机器视觉装置1000的一个或多个元件经由有线或无线通信来通信。

[0107] 在一些实施例中，机器视觉装置1000可包括图像捕获装置1020。在其它实施例中，图像捕获装置1020可为独立装置并且不集成到机器视觉装置1000中。图像捕获装置1020可被配置成捕获现实世界的图像。在一些实施例中，图像捕获装置1020可捕获二维图像。在其

它实施例中，图像捕获装置1020可捕获三维图像。图像捕获装置1020可为静态图像相机、摄像机等。图像捕获装置1020可被配置成捕获彩色图像、黑白图像、过滤的图像(例如深棕过滤器、彩色过滤器、模糊过滤器等)、通过一个或多个镜头捕获的图像(例如放大镜头、广角镜头等)等。在一些实施例中，图像捕获装置1020(和/或处理器1005)可修改一个或多个图像设定或特征，如颜色、对比度、亮度、白色比例尺、饱和度、清晰度等。

[0108] 在一些实施例中，机器视觉装置1000可包括电源1025。电源1025可被配置成为机器视觉装置1000的一个或多个元件提供电力。在一些实施例中，电源1025可包括交流电源，如可用线路电压(例如，在美国的60赫兹的120伏交流电)。电源1025可包括将电力转换成可由机器视觉装置1000的一个或多个元件使用的电力(如1.5伏、8伏、12伏、24伏等)的一个或多个变压器、整流器等。电源1025可包括一个或多个电池。

[0109] 在一些实施例中，机器视觉装置1000可包括用户接口1030。用户接口1030可被配置成从用户接收信息和/或向用户提供信息。用户接口1030可为本领域中已知的任何用户接口。用户接口1030可为如本领域的技术人员已知的用于接收用户输入和/或用于输入机器视觉装置1000的机器指令的接口。用户接口1030可使用各种输入技术，包括(但不限于)键盘、触笔和/或触摸屏、鼠标、轨迹球、小键盘、麦克风、语音辨识、运动辨识、磁盘驱动器、远程控制器、输入端口、一个或多个按钮、拨号盘、操纵杆等，以允许外部来源，如用户，将信息输入机器视觉装置1000。用户接口1030可用于导航菜单、调节选项、调节设定、调节显示器等。

[0110] 用户接口1030可被配置成提供用于将来自机器视觉装置1000的信息呈现到外部系统、用户或存储器等的接口。举例来说，用户接口1030可包括用于显示器、打印机、扬声器、警报灯/指示灯、网络接口、磁盘驱动器、计算机存储器装置等的接口。用户接口1030可包括彩色显示器、负极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器等。

[0111] 在一些实施例中，定位装置1035可用于识别定位，例如机器视觉装置1000的定位。在一些实施例中，定位装置1035可包括全球定位服务(GPS)装置、陀螺仪等。举例来说，机器视觉装置1000可包括机车机器(例如机器人、机动车等)并且接收指令以到达目的地。机器视觉装置1000可使用定位装置1035以识别目的地和/或机器视觉装置1000当前所在的定位。机器视觉装置1000可使用定位装置1035以导航到目的地。

[0112] 如上文所论述，机器视觉装置1000可为单独单元或可集成到另一个系统中。举例来说，机器视觉装置1000可与机车机器结合使用以导航航线。

[0113] 虽然上文实例处理导航、面部辨识和人辨识，但是应理解本文所述的技术可用于多种应用，包括物体辨识、情绪/表达辨识、经济/地质学/气候等数据的趋势辨识、疾病检测(例如使用医学图像如MRI图像、CT扫描图像、病理学切片等)和动物或人类表现良好的其它辨识/检测任务。

[0114] 本文所使用的技术可用于处理任何类型的图像数据。举例来说，图像数据可响应于可见光而生成，但是还可通过其它类型的电磁辐射(如红外线、紫外线或在电磁波谱上的其它波长)来生成。在一些实施例中，图像数据可为人工或虚拟形象数据(例如基于虚拟环境的模型生成的)。在一些实施例中，人工图像数据可与任何种类的合适数据的可视化相关，包括例如医学成像数据(磁共振成像数据、计算机辅助断层扫描数据)、地震成像数据、

LIDAR数据、金融数据等。

[0115] 在各种实施例中，图像数据可为单个图像或多个图像；此外，图像可为静止的或可以时空形式变化。可使用简单形状如图，或相对复杂刺激如自然场景。此外，图像可为灰阶或彩色或灰色和彩色的组合。

[0116] 上文所论述的实施例的各种实施方案涉及一种方法：包括接收对应于一系列原始图像的原始图像数据，用编码器处理原始图像数据以生成经编码数据，其中编码器的特征在于输入/输出转换基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换，和通过应用被配置成压缩包含于经编码数据中的信息的量的减维算法处理经编码数据以生成减维经编码数据。经编码数据可包括一系列视网膜图像。处理经编码数据的步骤可包括处理一系列视网膜图像以基于视网膜图像生成特征标记数据。特征标记数据可包括与多个视网膜图像区域相关的信息。特征标记数据可包括对应于多个视网膜图像区域中的每个的运动数据。运动数据可包括对应于多个视网膜图像区域中的每个的速度数据。特征标记数据可包括对应于多个视网膜图像区域中的每个的光流数据。处理经编码数据的步骤可包括对经编码数据应用经训练的算法。经训练的算法可包括卷积神经网络(CNN)。

[0117] 经训练的算法可已在经编码训练数据的训练数据集上训练，并且经编码训练数据可已使用特征在于基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换的输入/输出转换的训练编码器编码。经编码训练数据的训练集可包括虚拟环境的经编码图像，并且原始图像数据可包括现实环境的原始。经编码培训数据的培训集可包括在第一组条件下获得的图像，并且原始图像数据可包括在不同于第一组条件的第二组条件下获得的原始图像。第一组条件和第二组条件可包括光照条件。

[0118] 方法可进一步包括对减维经编码数据应用机器视觉算法。机器视觉算法可包括选自以下组成的列表的至少一个：物体辨识算法、图像分类算法、面部辨识算法、光学字符辨识算法、基于内容的图像检索算法、姿势估计算法、运动分析算法、自运动(egomotion)确定算法、移动追踪算法、光流确定算法、场景重构算法、3D体积辨识算法和导航算法。

[0119] 机器视觉算法当应用于减维经编码数据时比当应用于尚未使用编码器或减维算法处理的对应的原始图像集合时呈现更好性能。机器视觉算法当应用于减维经编码数据时比当应用于已仅使用减维算法处理的对应的经处理原始图像集合时呈现更好性能。机器视觉算法当应用于减维经编码数据时比当应用于尚未使用减维算法处理的对应的原始图像集合时呈现更好性能。机器视觉算法包括用于在一系列图像内检测或识别人的算法，并且机器视觉算法当应用于基于包括人的图像的减维经编码数据时比当应用于尚未编码器或减维算法处理的对应的数据集合时呈现更好检测或识别精确性。

[0120] 在上文所描述的方法的实施方案中，包含于经编码数据中的信息的量被至少压缩为所述对应的原始图像数据的约1/2，并且减维经编码数据被至少压缩为所述对应的经编码数据的约1/2。包含于经编码数据中的信息的量被至少压缩为所述对应的原始图像数据的约1/5，并且减维经编码数据被至少压缩为所述对应的经编码数据的约1/5。包含于经编码数据中的信息的量被至少压缩为所述对应的原始图像数据的约1/10，并且减维经编码数据被至少压缩为所述对应的经编码数据的约1/10。

[0121] 在上文所描述的方法的实施方案中，脊椎动物可包括选自以下组成的列表中的至少一个：鼠和猴。视网膜细胞可包括神经节细胞。视网膜细胞可包括一个或多个类别的细

胞，并且两个类别的细胞可包括打开细胞和关闭细胞。编码器的特征可在于在包括包括时空变化图像的自然场景图像的输入范围内基本上模拟脊椎动物视网膜的一个或多个视网膜细胞的输入/输出转换的输入/输出转换。

[0122] 在另一个实施方案中，设备包括被配置成存储原始图像数据的至少一个存储器存储装置；和可操作地与存储器耦接并且被编程以执行前述段落中任一个的方法的至少一个处理器。设备可进一步包括可操作地耦接到至少一个处理器的至少一个机器人装置，并且机器人装置可包括被配置成生成原始图像数据的至少一个图像传感器。另一个实施方式包括具有用于实施上文所描述的方法任一种的步骤的计算机可执行指令的非暂时性计算机可读介质。

[0123] 如本文所用，术语“光”和相关术语(例如“光学”、“视觉”)应理解为包括在可见光谱内和在可见光谱外两者的电磁辐射，包括例如紫外线和红外辐射。

[0124] 虽然本文中已经描述和说明各种发明性实施例，但是本领域的技术人员将容易设想多种其它装置和/或结构用于执行本文所描述的功能和/或获得本文所描述的结果和/或一个或多个优点，并且这类变化和/或修改中的每一个被认为在本文所描述的发明性实施例的范围内。更一般来说，本领域的技术人员将容易了解，本文所描述的所有参数、尺寸、材料以及配置都意指示例性的，并且实际参数、尺寸、材料和/或配置将取决于具体应用或使用发明性教示的应用。本领域的技术人员顶多使用常规实验即可认识到或能够确定本文所描述的具体发明性实施例的许多等效物。因此，应理解，前述实施例仅借助于实例呈现并且在所附权利要求书以及其等效物的范围内，发明性实施例可用与具体描述和要求不同的方式实践。本公开的发明性实施例涉及本文所述的每一单独特征、系统、制品、材料、套件和/或方法。此外，两个或更多个这类特征、系统、制品、材料、套件和/或方法(如果这类特征、系统、制品、材料、套件和/或方法并非互不一致)的任何组合包括在本公开的发明性范围内。

[0125] 上述实施例可以大量方式中的任一种实施。举例来说，可使用硬件、软件或其组合来实施实施例。当在软件中实施时，软件代码可在任何合适的处理器或处理器的集合上执行，无论在单个计算机中提供或分配在多个计算机中。

[0126] 另外，计算机可具有一个或多个输入和输出装置。可使用这些装置，以及其它物质，呈现用户接口。可用于提供用户接口的输出装置的实例包括用于视觉呈现输出的打印机或显示屏和用于音频呈现输出的扬声器或其它声音产生装置。可用于用户接口的输入装置的实例包括键盘和指示装置，如鼠标、触摸板和数字化平板计算机。作为另一个实例，计算机可通过语音辨识或以其它音频格式接收输入信息。

[0127] 这类计算机可通过以任何合适的形式的一个或多个网络互连，包括局域网或广域网，如企业网和智能网(IN)或因特网。这类网络可基于任何合适的技术，并且可根据任何合适的协议运行，并且可包括无线网络、有线网络或光纤网络。

[0128] 用于实施本文所述的功能性的至少一部分的计算机可包含存储器、一个或多个处理单元(在本文中也简单地被称作“处理器”)、一个或多个通信接口、一个或多个显示单元和一个或多个用户输入装置。存储器可包含任何计算机可读介质，并且可存储用于实施本文所述的各种功能性的计算机指令(在本文中也被称作“处理器可执行指令”)。(一个或多个)处理单元可用于执行指令。(一个或多个)通信接口可耦接到有线或无线网络、总线或其它通信装置，并且可因此允许计算机将通信传输到其它装置和/或从其它装置接收通信。可

提供(一个或多个)显示单元,例如以允许用户查看与指令的执行有关的各种信息。可提供(一个或多个)用户输入装置,例如以允许用户在指令的执行期间做出手动调节、做出选择、输入数据或多种其它信息和/或以多种方式中的任何一种与处理器交互。

[0129] 在本文中概述的多种方法可编码为在采用多种操作系统或平台中的任一种的一个或多个处理器上可执行的软件。此外,这类软件可使用多种合适的编程语言和/或编程或脚本处理工具中的任一种编写,并且还可汇编为可执行机器语言代码或在框架或虚拟机上执行的中代编码。

[0130] 在这方面,各种发明性概念可以计算机可读存储媒体(或多个计算机可读存储媒体)(例如,计算机存储器、一个或多个软盘、压缩光盘、光盘、磁带、快闪存储器、现场可编程门阵列或其它半导体装置中的电路配置或其它非暂时性媒体或有形计算机存储媒体),所述计算机可读存储媒体编码有一个或多个程序,所述程序当在一个或多个计算机或其它处理器上执行时,执行实施上文所论述的本发明的各种实施例的方法。一种或多种计算机可读介质可是可运输的,使得存储在上面的一个或多个程序可加载到一个或多个不同计算机或其它处理器上,以便实施如上文所论述的本发明的各种方面。

[0131] 本文中在通用意义上使用术语“程序”或“软件”,以指代任何类型的计算机代码或计算机可执行指令集,其可用于编程计算机或其它处理器以实施如上文所论述的实施例的各种方面。此外,应了解,根据一个方面,当被执行时执行本发明的方法的一个或多个计算机程序不必驻存在单个计算机或处理器上,但是可以模块形式分配在多种不同计算机或处理器之中以实施本发明的各种方面。

[0132] 计算机可执行指令可呈许多形式,例如通过一个或多个计算机或其它装置执行的程序模块。一般来说,程序模块包括执行特定任务或实施特定抽象数据类型的例程、程序、对象、组件、数据结构等。通常在各种实施例中程序模块的功能性可按照期望组合或分配。

[0133] 另外,数据结构可以任何合适的形式存储在计算机可读介质中。为简单说明起见,数据结构可示出具有通过在数据结构中的定位相关的字段。这类关系可同样地通过在计算机可读介质中为字段的存储指派传达字段之间的关系的位置来实现。然而,任何合适的机制可用于建立在数据结构的字段中的信息之间的关系,包括通过使用指针、标记或建立在数据元素之间的关系的其它机制。

[0134] 另外,各种发明性概念可实施为一个或多个方法,其中已提供实例。作为方法的一部分执行的动作可以用任何合适的方式排序。因此,可构造按不同于所说明的次序执行动作的实施例,其可包括同时执行一些动作,即使在说明性实施例中展示为依序的动作也是如此。

[0135] 应理解如本文中限定和使用的所有定义优先于字典定义、以引用的方式并入的文档中的定义和/或限定的术语的普通含义。

[0136] 如本文在说明书和权利要求中所使用,不定冠词“一个(a/an)”除非明确相反指示,否则应理解为意指“至少一个”。

[0137] 如本文在说明书中和在权利要求书中所使用,词组“和/或”应理解为意指如此结合的元件中的“任一者或两者”,即,元件在一些情况中结合地存在并且在其它情况中分离地存在。用“和/或”列出的多个元件应以相同方式解释,即,“一个或多个”的元件如此结合。除通过“和/或”条目具体识别的元件之外,可任选地存在其它元件,无论与具体识别的那些

元件是否相关。因此，作为非限制性实例，提及“A和/或B”，当与开放性语言如“包含”结合使用时，可在一个实施例中，是指仅A(任选地包括除B之外的元件)；在另一个实施例中，是指仅B(任选地包括除A之外的元件)；在又一个实施例中，是指A和B两者(任选地包括其它元件)等。

[0138] 如本文在说明书和权利要求书中所使用，“或”应理解为具有与如上文所定义的“和/或”相同的含义。举例来说，当在列表中分隔多个项目时，“或”或“和/或”将解释为包含性的，即，包含至少一个，但也包含多个元件或元件列表中的多于一个和任选地额外的未列出的项目。只有明确相反指示的术语(例如“仅仅……中的一个”或“恰好……中的一个”或当在权利要求书时使用时“由……组成”将是指包含多个元件或元件列表中的恰好一个元件。一般来说，如本文所用的术语“或”当前面是例如“任一”、“……中的一个”、“仅……中的一个”或“恰好……中的一个”的排他性术语时，仅应解释为指示排他性替代方案(即“一个或另一个但并非两者”)。“基本上由……组成”当在权利要求书中使用时，应具有其如专利法领域中所用的普通含义。

[0139] 如本文在说明书和权利要求书中所使用，参照一个或多个元件的列表的词组“至少一个”应理解为意指选自元件列表中的元件中的任何一个或多个的至少一个元件，但未必包含元件列表内具体列举的每一和每个元件中的至少一个，并且不排除元件列表中的元件的任何组合。这个定义还允许除了元件列表内具体识别的词组“至少一个”所指的元件之外可任选地存在元件，无论其是否与具体识别的那些元件相关。因此，作为非限制性实例，“A和B中的至少一个”(或等效地，“A或B中的至少一个”，或等效地，“A和/或B中的至少一个”)在一个实施例中可是指至少一个，任选地包括多于一个A，不存在B(并且任选地包括除了B之外的元件)；在另一个实施例中可是指至少一个，任选地包括多于一个B，不存在A(并且任选地包括除了A之外的元件)；在又一个实施例中可是指至少一个，任选地包括多于一个A，以及至少一个，任选地包括多于一个B(并且任选地包括其它元件)；等。

[0140] 在权利要求书中以及在上述说明书中，如“包含”、“包括”、“带有”、“具有”、“含有”、“涉及”、“容纳”、“由……组成”等所有连接词应理解为是开放性的，即，意指包括但不限于。仅连接词“由……组成”以及“主要由……组成”将分别是封闭或半封闭的连接词，如美国专利局专利审查程序手册(United States Patent Office Manual of Patent Examining Procedures)2111.03章所阐述。

[0141] 关于本文中实质上任何复数和/或单数术语的使用，本领域的技术人员可在适于上下文和/或应用的情况下将复数转换成单数和/或将单数转换成复数。为清晰起见，本文中可明确地阐述各种单数/复数排列。

[0142] 本领域的技术人员将理解，一般来说，本文中并且尤其所附权利要求书(例如所附权利要求书主体)中所使用的术语通常意思指“开放性”术语(例如术语“包括”应解释为“包括(但不限于)”，术语“具有”应解释为“至少具有”，术语“包含”应解释为“包含(但不限于)”等)。本领域的技术人员将进一步理解，如果打算引入特定数目的权利要求叙述，那么此打算将明确叙述于所述权利要求中，且在不存在此叙述的情况下，不存在此打算。举例来说，出于辅助理解，以下所附权利要求书可含有介绍性词组“至少一个”和“一个或多个”的使用，以介绍权利要求叙述。然而，这些词组的使用不应解释为暗示通过不定冠词“一个(a/an)”对一个权利要求陈述的介绍将含有这类所介绍的权利要求陈述的任何特定权利要求

都限于仅含有一个这类陈述的发明，即使当相同权利要求包括介绍性词组“一个或多个”或“至少一个”以及不定冠词如“一个(a/an)”时也是如此(例如，“一个(a/an)”通常应解释为意指“至少一个”或“一个或多个”)；对于用于介绍权利要求陈述的定冠词的使用，情况也是如此。另外，即使明确叙述特定数目的所介绍权利要求叙述，本领域的技术人员将认识到这类叙述通常应被解释为意指至少所叙述数目(例如无其它修饰语的明确叙述“两种叙述”通常意指至少两种叙述或两种或更多种叙述)。此外，在其中使用类似于“A、B和C等中的至少一个”的约定的那些情形中，一般来说，在本领域的技术人员将理解约定的意义上预期这类构造(例如“具有A、B和C至少一个的系统”将包括(但不限于)仅具有A、仅具有B、仅具有C、具有A和B、具有A和C、具有B和C和/或具有A、B和C等的系统)。在其中使用类似于“A、B或C等中的至少一个”的约定的那些情形中，一般来说在本领域的技术人员将理解约定的意义上预期这类构造(例如“具有A、B或C等中的至少一个的系统”将包括(但不限于)仅具有A、仅具有B、仅具有C、具有A和B、具有A和C、具有B和C和/或具有A、B和C等的系统)。本领域内的技术人员将进一步理解，无论在说明书、权利要求还是图式中，应将呈现两种或更多种替代术语的实际上任何分离性词语及/或词组理解为涵盖包括所述术语中的一种、所述术语中的任一种或两种术语的可能性。举例来说，词组“A或B”理解为包括“A”或“B”或“A和B”的可能性。另外，除非另外指出，否则使用词“大致”、“约”、“大约”等意指加或减百分之十。

[0143] 已出于说明和描述的目的呈现说明性实施例的前述描述。并非旨在为穷举的或限制所公开的确切形式，并且根据以上教示，修改和变化是可能的或可自所公开的实施例的实践获得修改和变化。旨在本发明的范围由所附权利要求书和其等效物限定。

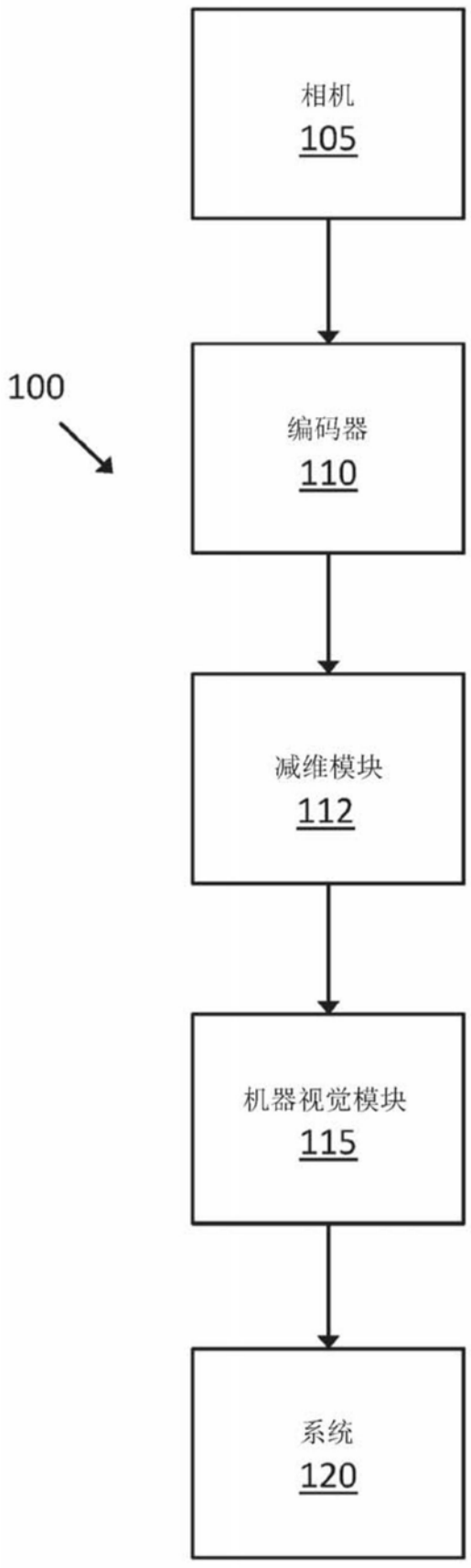


图1

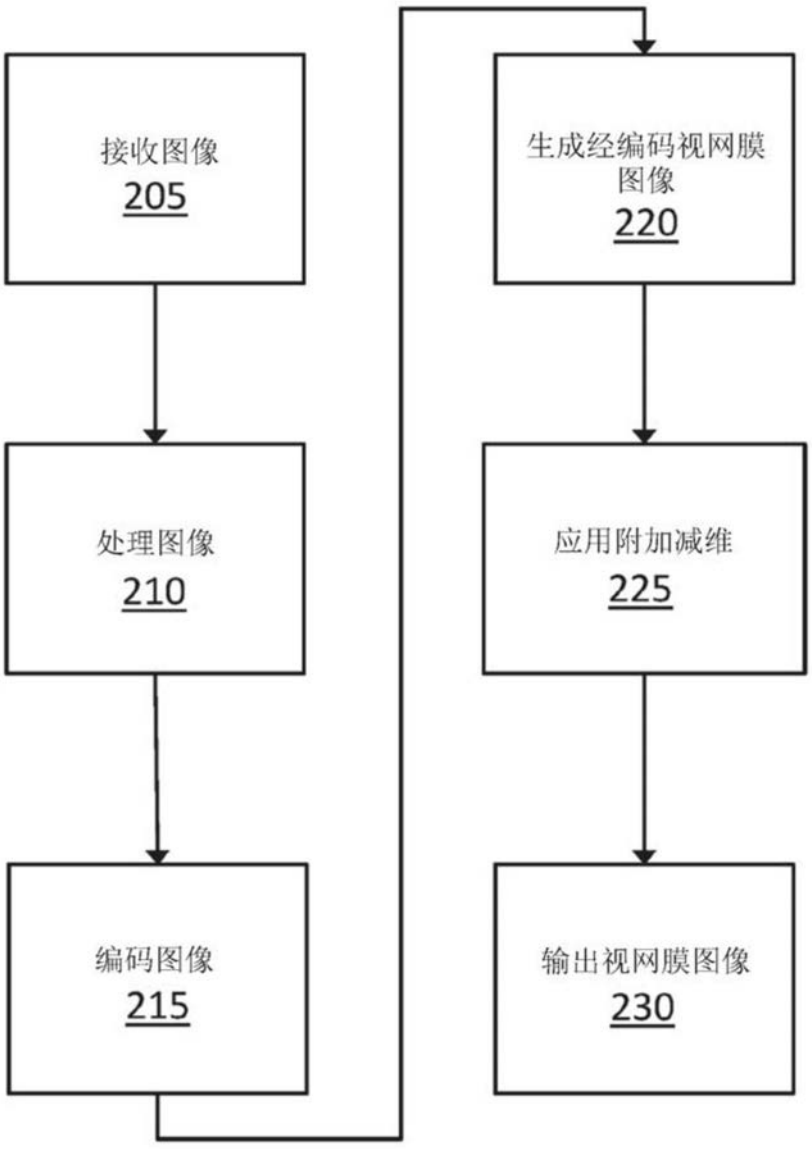

接收图像
205
处理图像
210
编码图像
215
生成经编码视网膜图像
220
应用附加减维
225
输出视网膜图像
230

图2

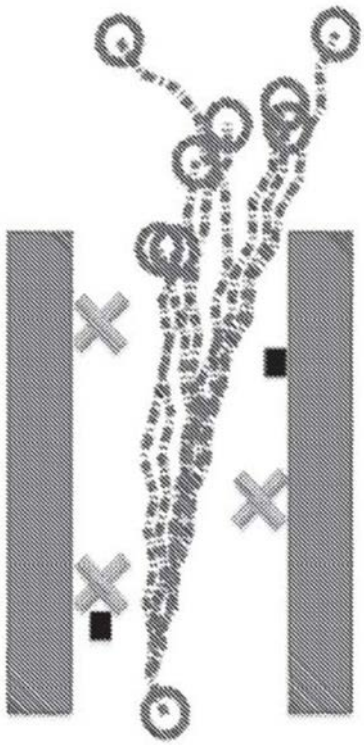


图3A

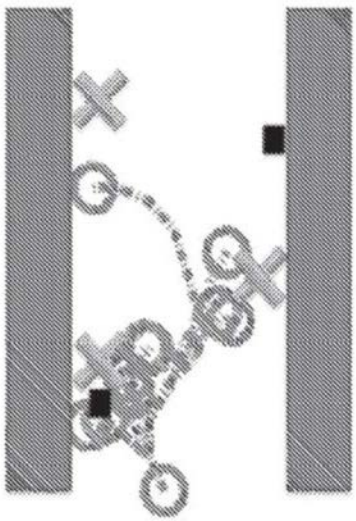


图3B

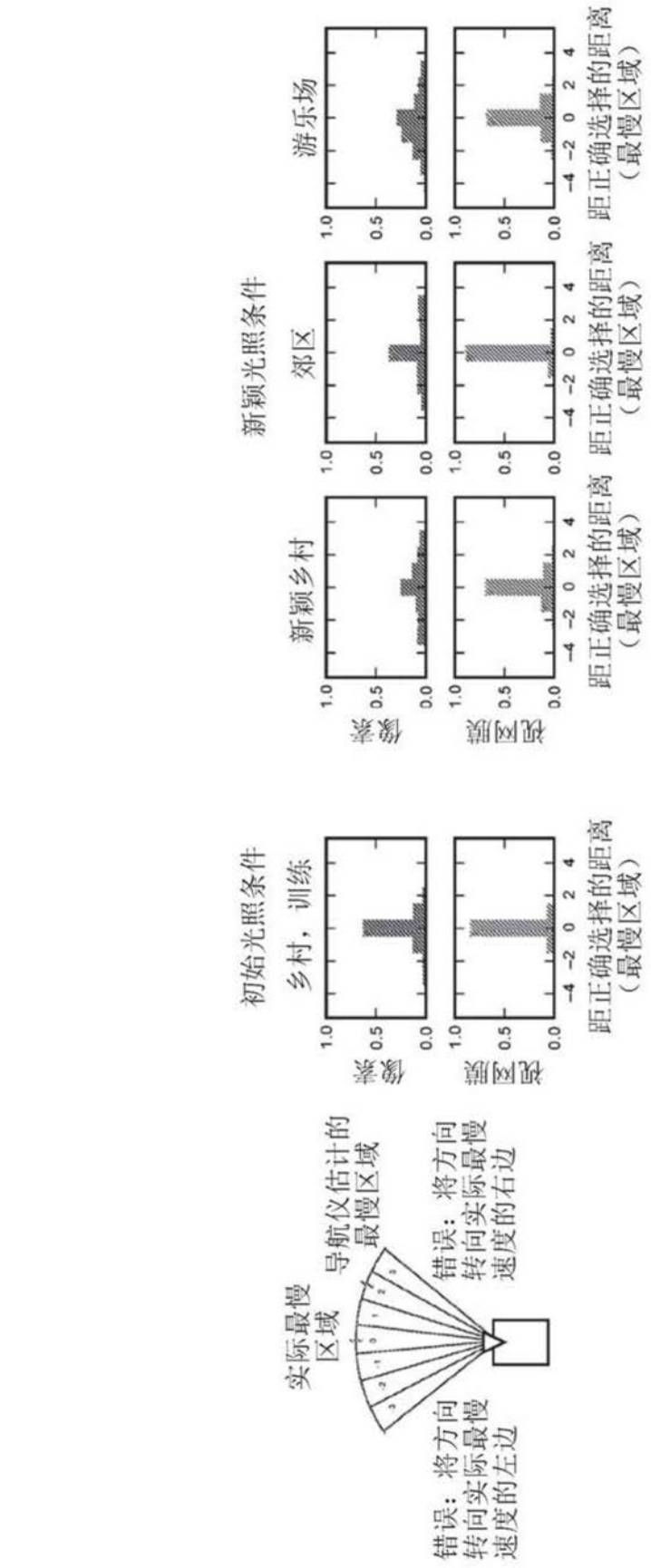

初始光照条件
乡村，训练
新颖光照条件
新颖乡村
郊区
游乐场
像素
距正确选择的距离
（最慢区域）
实际最慢区域
导航仪估计的最慢区域
错误：将方向转向实际最慢速度的左边
错误：将方向转向实际最慢速度的右边

图4

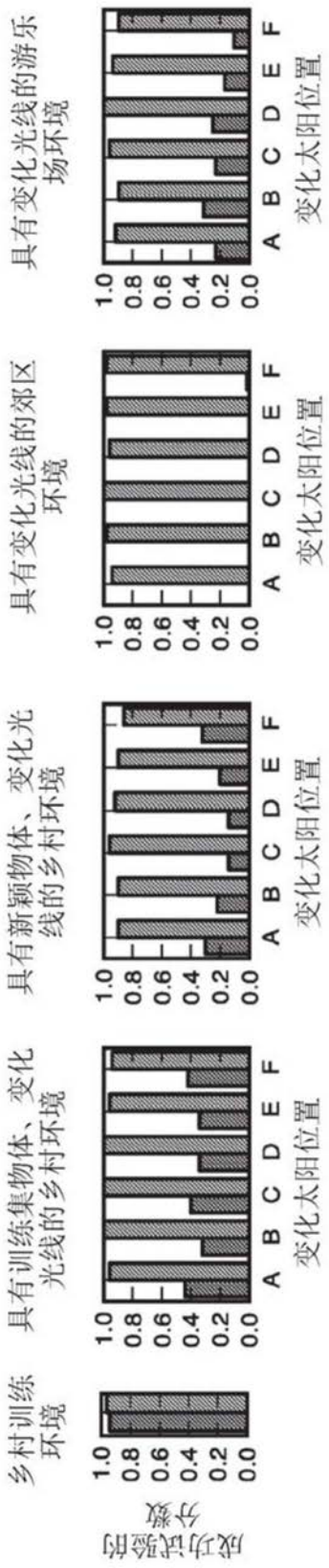

成功试验的分数
1.0
0.8
0.6
0.4
0.2
0.0
乡村训练环境
具有训练集物体、变化光线的乡村环境
具有新颖物体、变化光线的乡村环境
具有变化光线的郊区环境
具有变化光线的游乐场环境
A B C D E F
变化太阳位置

图5

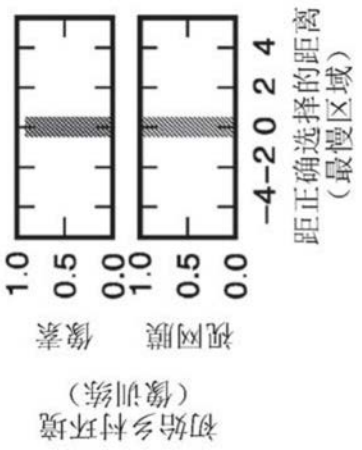


图6A

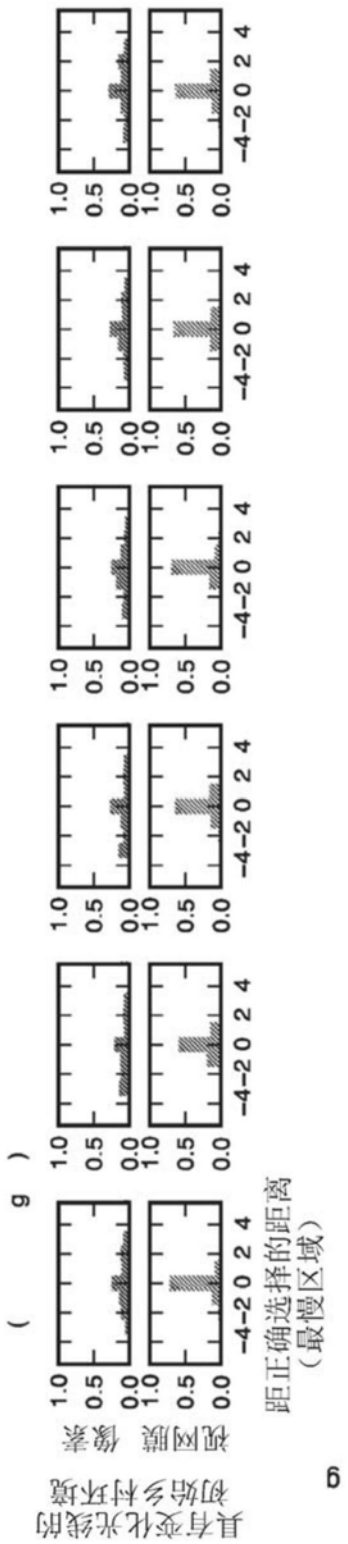


图6B

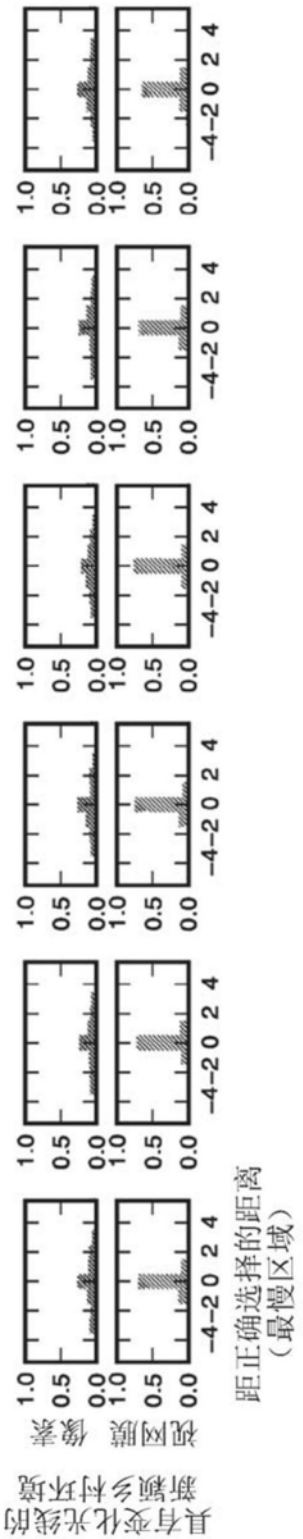


图6C

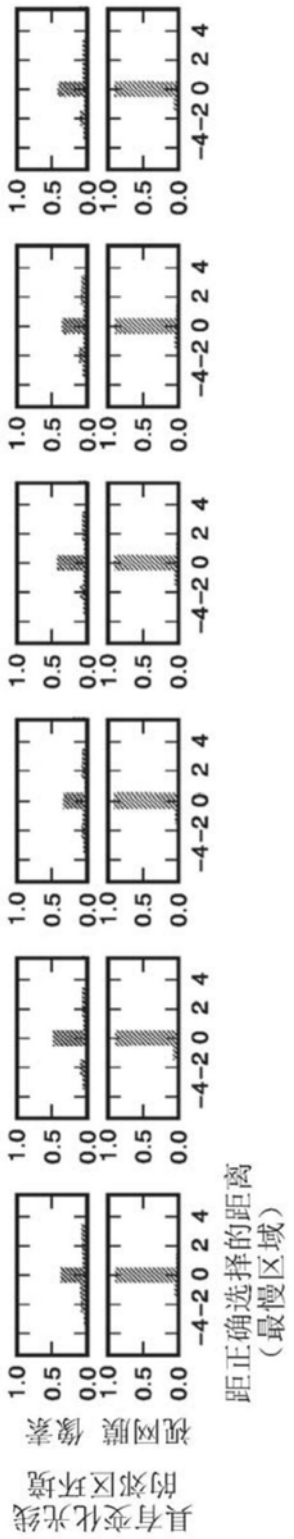


图6D

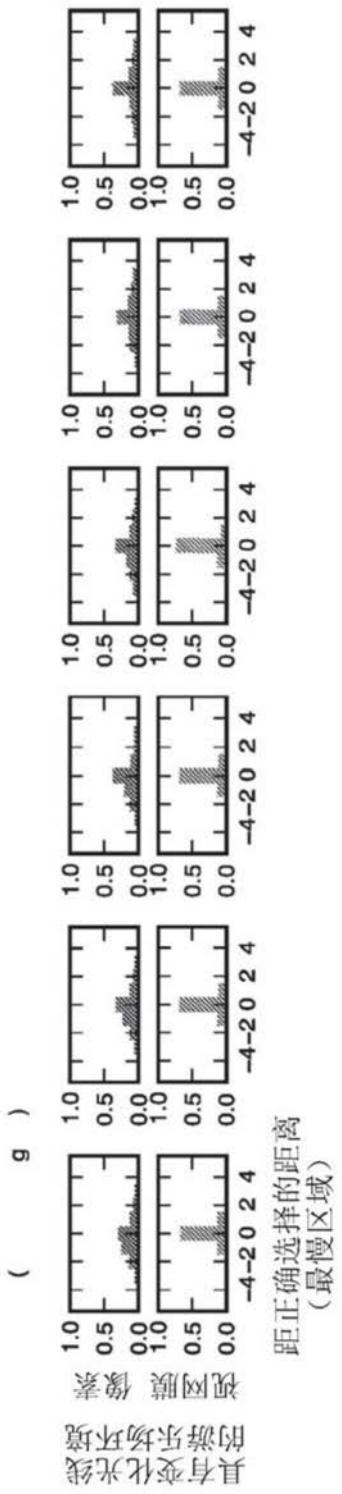


图6E

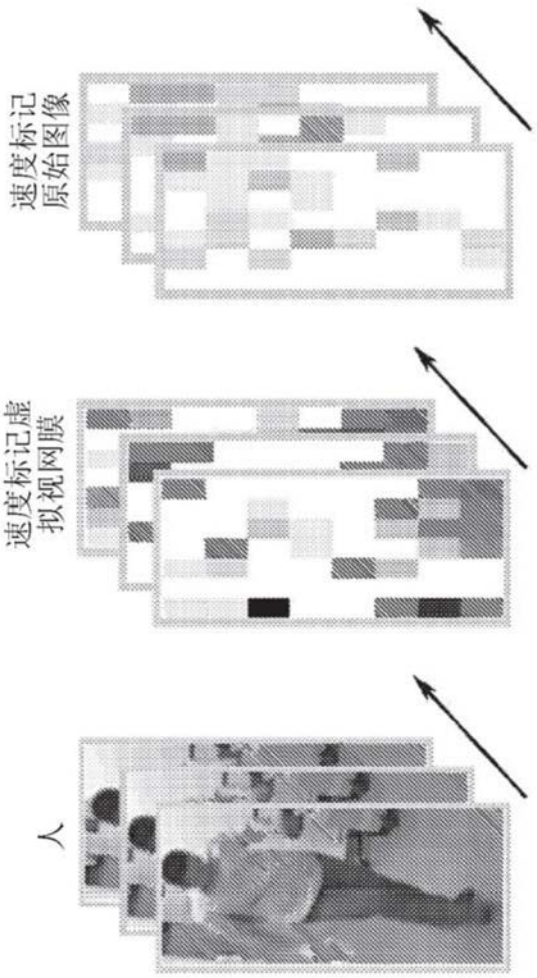


图7

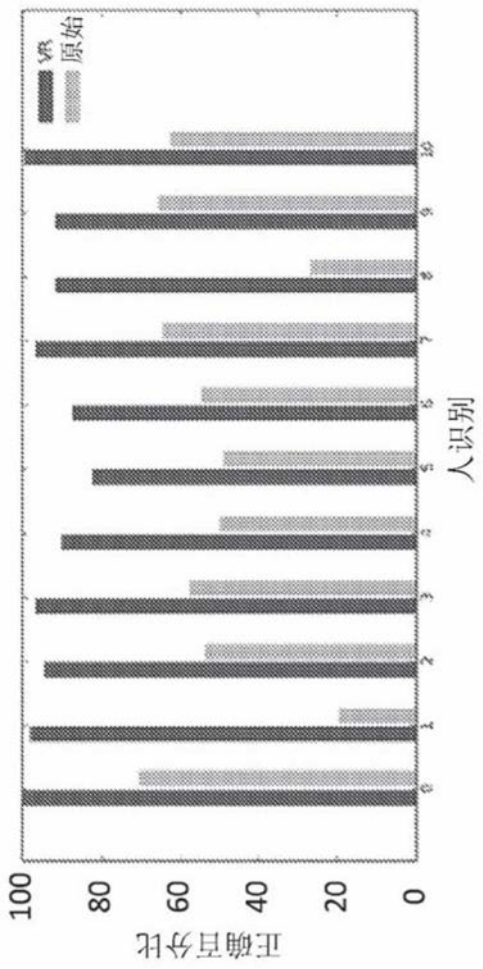


图8

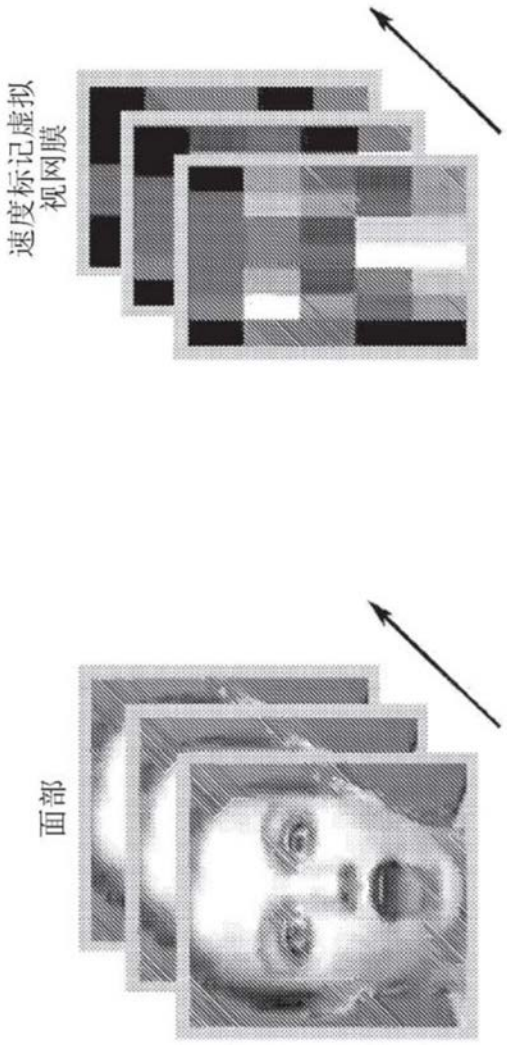


图9

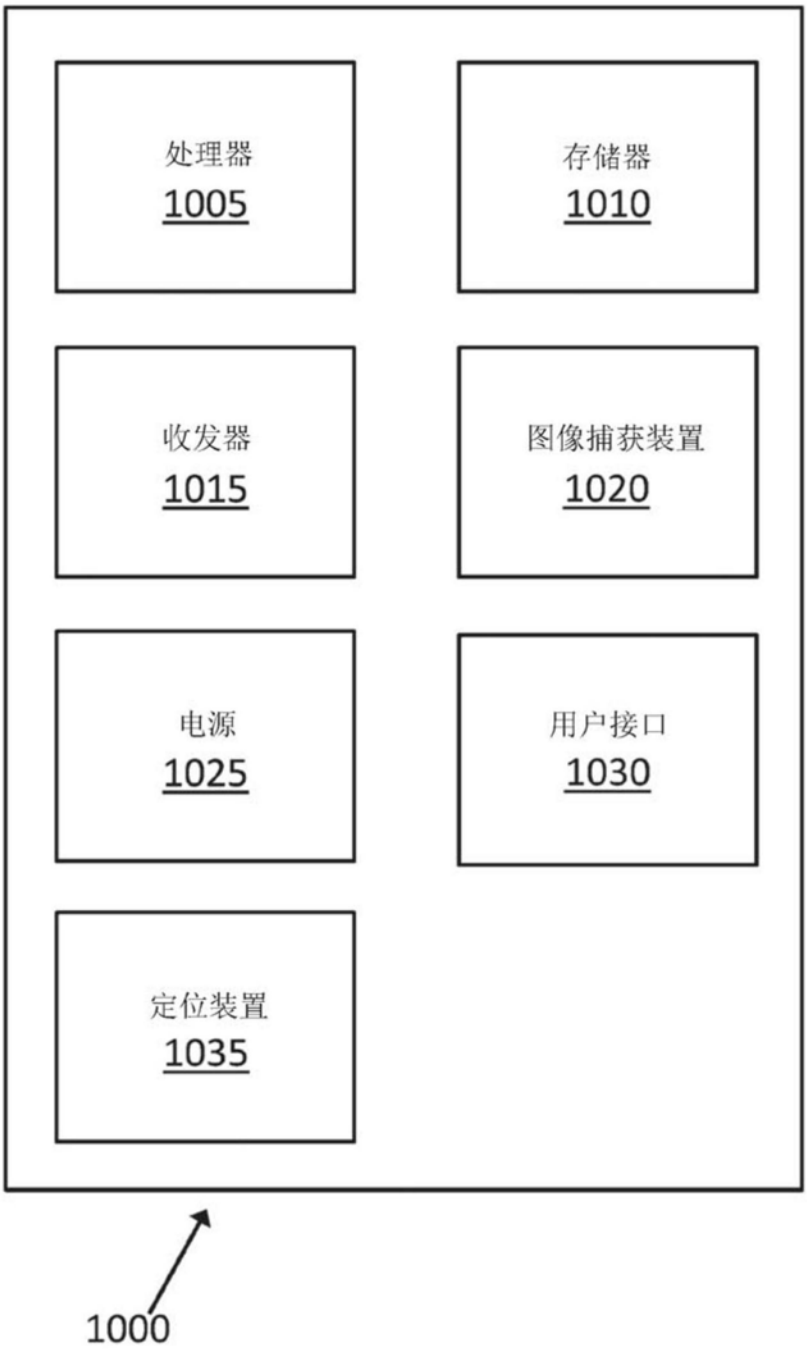

处理器
1005
存储器
1010
收发器
1015
图像捕获装置
1020
电源
1025
用户接口
1030
定位装置
1035
1000

图10