

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 4 月 25 日 (25.04.2019)



(10) 国际公布号
WO 2019/075669 A1

- (51) 国际专利分类号:
G06T 5/00 (2006.01) *H04N 5/213* (2006.01)
G06N 3/04 (2006.01)
- (21) 国际申请号: PCT/CN2017/106735
- (22) 国际申请日: 2017 年 10 月 18 日 (18.10.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 深圳市大疆创新科技有限公司 (SZ DJI TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。
- (72) 发明人: 肖瑾 (XIAO, Jin); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 曹

子晟 (CAO, Zisheng); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 胡攀 (HU, Pan); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。

(74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国北京市海淀区西直门北大街32号枫蓝国际A座8F-6, Beijing 100082 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) Title: VIDEO PROCESSING METHOD AND DEVICE, UNMANNED AERIAL VEHICLE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 视频处理方法、设备、无人机及计算机可读存储介质

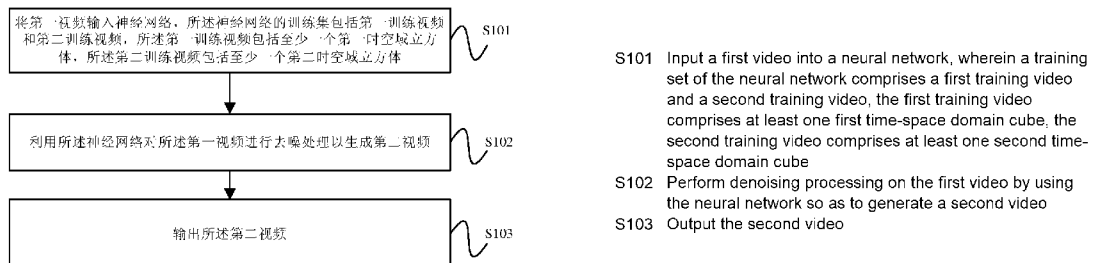


图 1

(57) Abstract: Provided are a video processing method and device, an unmanned aerial vehicle, and a computer-readable storage medium. The method comprises: inputting a first video into a neural network, a training set of the neural network comprising a first training video and a second training video, the first training video comprising at least one first time-space domain cube, the second training video comprising at least one second time-space domain cube; performing denoising processing on the first video by using the neural network so as to generate a second video; and outputting the second video. The embodiments of the present invention improve the computational complexity of video denoising compared with a video denoising method based on motion estimation in the prior art, and improve the effect of video denoising compared with a video denoising method without motion estimation in the prior art.

(57) 摘要: 本发明实施例提供一种视频处理方法、设备、无人机及计算机可读存储介质, 该方法包括: 将第一视频输入神经网络, 所述神经网络的训练集包括第一训练视频和第二训练视频, 所述第一训练视频包括至少一个第一时空域立方体, 所述第二训练视频包括至少一个第二时空域立方体; 利用所述神经网络对所述第一视频进行去噪处理以生成第二视频; 以及输出所述第二视频。本发明实施例相比于现有技术中基于运动估计的视频去噪方法, 提高了视频去噪的计算复杂度, 相比于现有技术中无需运动估计的视频去噪方法, 提高了视频去噪的效果。

LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

视频处理方法、设备、无人机及计算机可读存储介质

技术领域

5 本发明实施例涉及无人机领域，尤其涉及一种视频处理方法、设备、无人机及计算机可读存储介质。

背景技术

10 随着数码产品如相机、摄像头的普及，视频已经广泛的运用于日常生活中,但是噪声在视频的拍摄过程中依旧是不可避免的，噪声直接影响了视频的质量。

15 为了去除视频中的噪声，现有技术中对视频的去噪方法包括：基于运动估计的视频去噪方法和无需运动估计的视频去噪方法。但是，基于运动估计的视频去噪方法的计算复杂度高，无需运动估计的视频去噪方法的去噪效果不理想。

发明内容

本发明实施例提供一种视频处理方法、设备、无人机及计算机可读存储介质，以提高对视频去噪的去噪效果。

20 本发明实施例的第一方面是提供一种视频处理方法，包括：
将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

25 利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及
输出所述第二视频。

本发明实施例的第二方面是提供一种视频处理设备，包括一个或多个处理器，单独或协同工作，所述一个或多个处理器用于：

将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所

述第二训练视频包括至少一个第二时空域立方体；

利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及

输出所述第二视频。

- 5 本发明实施例的第三方面是提供一种无人机，包括：机身；动力系统，安装在所述机身，用于提供飞行动力；以及如第二方面所述的视频处理设备。

本发明实施例的第四方面是提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被一个或多个处理器执行时实现以下步骤：

10 将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

15 利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及

输出所述第二视频。

- 20 本实施例提供的视频处理方法、设备、无人机及计算机可读存储介质，通过将原始的带有噪声的第一视频输入到预先训练成的神经网络中，该神经网络是通过干净的第一训练视频包括的至少一个第一时空域立方体和加噪的第二训练视频包括的至少一个第二时空域立方体训练得到的，通过该神经网络对第一视频进行去噪处理以生成第二视频，相比于现有技术中基于运动估计的视频去噪方法，提高了视频去噪的计算复杂度，相比于现有技术中无需运动估计的视频去噪方法，提高了视频去噪的效果。

25 附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

- 30 图 1 为本发明实施例提供的视频处理方法的流程图；

图 2 为本发明实施例提供的第一训练视频的示意图；

图 3 为本发明实施例提供的第一训练视频中图像帧的分解示意图；

图 4 为本发明实施例提供的一种第一时空域立方体的划分示意图；

图 5 为本发明实施例提供的另一种第一时空域立方体的划分示意图；

5 图 6 为本发明实施例提供的第一训练视频被划分为多个第一时空域立方体的示意图；

图 7 为本发明另一实施例提供的视频处理方法的流程图；

图 8 为本发明另一实施例提供的视频处理方法的流程图；

图 9 为本发明另一实施例提供的第一均值图像的示意图；

10 图 10 为本发明另一实施例提供的对第一时空域立方体进行稀疏处理的示意图；

图 11 为本发明另一实施例提供的视频处理方法的流程图；

图 12 为本发明另一实施例提供的视频处理方法的流程图；

图 13 为本发明实施例提供的视频处理设备的结构图；

15 图 14 为本发明实施例提供的无人机的结构图。

附图标记：

	20-第一训练视频	21-图像帧	22-图像帧
	23-图像帧	24-图像帧	25-图像帧
	2n-图像帧		
	211-子图像	212-子图像	213-子图像
	214-子图像		
20	221-子图像	222-子图像	223-子图像
	224-子图像		
	231-子图像	232-子图像	233-子图像
	234-子图像		
	241-子图像	242-子图像	243-子图像
	244-子图像		
	251-子图像	252-子图像	253-子图像
	254-子图像		
	2n1-子图像	2n2-子图像	2n3-子图像
	2n4-子图像		
25	41-第一时空域立方体	42-第一时空域立方体	
	43-第一时空域立方体	44-第一时空域立方体	
	51-子图像	52-子图像	53-子图像
	54-子图像		
	55-子图像	56-子图像	57-子图像
	58-子图像		
	59-子图像	60-子图像	61-第一时空域立方体
30	62-第一时空域立方体	90-第一均值图像	

510-子图像 530-子图像 550-子图像 570-子图像 590-子图像
130-视频处理设备 131-一个或多个处理器 100-无人机
107-电机 106-螺旋桨 117-电子调速器
118-飞行控制器 108-传感系统 110-通信系统
5 102-支撑设备 104-拍摄设备 112-地面站
114-天线 116-电磁波 109-视频处理设备

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

需要说明的是，当组件被称为“固定于”另一个组件，它可以直接在另一个组件上或者也可以存在居中的组件。当一个组件被认为是“连接”另一个组件，它可以是直接连接到另一个组件或者可能同时存在居中组件。

除非另有定义，本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本发明。本文所使用的术语“及 / 或”包括一个或多个相关的所列项目的任意的和所有的组合。

下面结合附图，对本发明的一些实施方式作详细说明。在不冲突的情况下，下述的实施例及实施例中的特征可以相互组合。

本发明实施例提供一种视频处理方法。图1为本发明实施例提供的视频处理方法的流程图。本实施例的执行主体可以是视频处理设备，该视频处理设备可以设置在无人机，也可以设置在地面站，地面站具体可以是遥控器、智能手机、平板电脑、地面控制站、膝上型电脑、手表、手环等及其组合。在其他实施例中，该视频处理设备还可以直接设置在拍摄设备，例如手持云台、数码相机、摄像机等。具体的，如果视频处理设备设置在无人机，则该视频处理设备可以对无人机搭载的拍摄设备拍摄到的视频进

行处理。如果视频处理设备设置在地面站，地面站可以接收无人机无线传输的视频数据，该视频处理设备对地面站接收到的视频数据进行处理。再或者，用户手持拍摄设备，拍摄设备内的视频处理设备对该拍摄设备拍摄到的视频进行处理。本实施例并不限定具体的应用场景。下面对视频处理方法进行详细介绍。

如图 1 所示，本实施例提供的视频处理方法，可以包括：

步骤 S101、将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体。

10 在本实施例中，第一视频可以是无人机搭载的拍摄设备拍摄的视频，也可以是地面站例如智能手机、平板电脑等拍摄的视频，还可以是用户所持的拍摄设备例如手持云台、数码相机、摄像机等拍摄的视频，其中，第一视频是带有噪声的视频，视频处理设备需要对第一视频进行去噪处理，具体的，视频处理设备将第一视频输入到预先训练成的神经网络中，可以理解的是，视频处理设备将第一视频输入神经网络之前，已根据第一训练视频和第二训练视频训练成所述神经网络。视频处理设备根据第一训练视频和第二训练视频训练所述神经网络的过程将在后续的实施例中详细介绍，下面详细介绍一下所述神经网络的训练集。

20 所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体。

25 可选的，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。也就是说第一训练视频为干净视频，第二训练视频为噪声视频。具体的，第一训练视频可以为无压缩的高清视频，第二训练视频可以是在无压缩的高清视频中添加噪声后的视频。

具体的，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

30 如图 2 所示，20 表示第一训练视频，第一训练视频 20 中包括多帧图像，本实施例不限定第一训练视频 20 包括的图像的帧数，如图 2 所示，

图像帧 21、图像帧 22、图像帧 23 只是第一训练视频 20 中任意的相邻的三帧图像。

如图 3 所示, 假设将图像帧 21 分成 4 个子图像, 例如子图像 211、子图像 212、子图像 213、子图像 214; 将图像帧 22 分成 4 个子图像, 例如子图像 221、子图像 222、子图像 223、子图像 224; 将图像帧 23 分成 4 个子图像, 例如子图像 231、子图像 232、子图像 233、子图像 234, 不失一般性, 第一训练视频 20 包括 n 帧图像, 最后一帧图像表示为 $2n$ 。依次类推可以将第一训练视频 20 中的每个图像帧分解成 4 个子图像, 直至将图像帧 $2n$ 分成 4 个子图像, 例如子图像 $2n1$ 、子图像 $2n2$ 、子图像 $2n3$ 、子图像 $2n4$ 。此处只是示意性说明, 并不限定每个图像帧可以分解成的子图像的个数。

根据图 3 可知, 子图像 211 在图像帧 21 中的位置、子图像 221 在图像帧 22 中的位置以及子图像 231 在图像帧 23 中的位置相同, 可选的, 将第一训练视频 20 中相邻的若干图像帧中位置相同的子图像构成一个集合, 该集合记为第一时空域立方体, 此处的第一时空域立方体是为了和后续的第二训练视频包括的第二时空域立方体进行区分。例如将第一训练视频 20 中每相邻 5 帧图像中位置相同的子图像构成一个集合, 如图 4 所示, 图像帧 21-25 是相邻的 5 帧图像, 来自图像帧 21-25 相同位置的子图像 211、子图像 221、子图像 231、子图像 241、子图像 251 构成一个第一时空域立方体 41; 来自图像帧 21-25 相同位置的子图像 212、子图像 222、子图像 232、子图像 242、子图像 252 构成一个第一时空域立方体 42; 来自图像帧 21-25 相同位置的子图像 213、子图像 223、子图像 233、子图像 243、子图像 253 构成一个第一时空域立方体 43; 来自图像帧 21-25 相同位置的子图像 214、子图像 224、子图像 234、子图像 244、子图像 254 构成一个第一时空域立方体 44。此处只是示意性说明, 并不限定一个第一时空域立方体中包括的子图像的个数。

在其他实施例中, 第一训练视频 20 中的每个图像帧可以不完全被划分成多个子图像, 如图 5 所示, 图像帧 21-25 是相邻的 5 帧图像, 只在每个图像帧中分别截取两个二维矩形块, 例如只在图像帧 21 上截取两个二维矩形块作为子图像 51 和子图像 52, 并没有像图 3 或图 4 所示, 将整个

图像帧 21 划分成 4 个子图像。此处只是示意性说明，并不限定从一个图像帧中截取出的二维矩形块的个数。同理，在图像帧 22 上截取两个二维矩形块作为子图像 53 和子图像 54；在图像帧 23 上截取两个二维矩形块作为子图像 55 和子图像 56；在图像帧 24 上截取两个二维矩形块作为子图像 57 和子图像 58；在图像帧 25 上截取两个二维矩形块作为子图像 59 和子图像 60。来自图像帧 21-25 相同位置的子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 构成一个第一时空域立方体 61；来自图像帧 21-25 相同位置的子图像 52、子图像 54、子图像 56、子图像 58、子图像 60 构成一个第一时空域立方体 62。此处只是示意性说明，并不限定一个第一时空域立方体中包括的子图像的个数。

同理于图 4 或图 5 所示的第一时空域立方体的划分方法，可以从如图 2 所示的第一训练视频 20 中划分出多个第一时空域立方体，如图 6 所示，第一时空域立方体 A 只是从第一训练视频 20 中划分出的多个第一时空域立方体中的一个。本实施例不限定第一训练视频 20 中包括的第一时空域立方体的个数，也不限定每个第一时空域立方体包括的子图像的个数，也不限定从图像帧中截取或划分子图像的方法。

不失一般性，假设第一训练视频 20 表示为 x ， x_t 表示第一训练视频 20 中的第 t 帧图像， $1 \leq t \leq n$ ， $x_t(i, j)$ 表示第 t 帧图像中的一个子图像， (i, j) 表示该子图像在第 t 帧图像中的位置，也就是说 $x_t(i, j)$ 表示从干净的第一训练视频 20 中截取的一个二维矩形块， (i, j) 表示二维矩形块的空域索引， t 表示二维矩形块的时域索引。将第一训练视频 20 中相邻的若干图像帧中位置相同、大小相同的子图像构成一个集合，该集合记为第一时空域立方体，第一时空域立方体 V_x 表示为如下公式 (1)：

$$V_x = \{x_{t_0-h}^{\Delta}(i, j), K, x_{t_0}(i, j), K, x_{t_0+h}(i, j)\} = \{x_{t_0+s}^{\Delta}(i, j)\}_{s=-h}^h \quad (1)$$

根据公式 (1) 可知，第一时空域立方体 V_x 中包括 $2h+1$ 个子图像。也就是说将第一训练视频 20 中相邻的 $2h+1$ 个图像帧中位置相同、大小相同的子图像构成一个集合，时域索引 $t_0-h, \dots, t_0, \dots, t_0+h$ 和空域索引 (i, j) 决定了第一时空域立方体 V_x 在第一训练视频 20 中的位置，根据时域索引和/或空域索引的不同，可从第一训练视频 20 中划分出多个不同的第一时空域立方体。

所述第二时空域立方体包括多个第二子图像，所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。假设第二训练视频表示为 Y ， Y_t 表示第二训练视频中的第 t 帧图像， $y_t(i, j)$ 表示第二训练视频中第 t 帧图像中的一个子图像， (i, j) 表示该子图像在第 t 帧图像中的位置，也就是说 $y_t(i, j)$ 表示从添加有噪声的第二训练视频中截取的一个二维矩形块， (i, j) 表示二维矩形块的空域索引， t 表示二维矩形块的时域索引。将第二训练视频中相邻的若干图像帧中位置相同、大小相同的子图像构成一个集合，该集合记为第二时空域立方体，第二时空域立方体的划分原理和过程和第一时空域立方体的划分原理和过程一致，此处不再赘述。

具体的，视频处理设备根据第一训练视频包括的至少一个第一时空域立方体，以及第二训练视频包括的至少一个第二时空域立方体训练所述神经网络，训练所述神经网络的过程将在后续的实施例中详细介绍。

步骤 S102、利用所述神经网络对所述第一视频进行去噪处理以生成第二视频。

视频处理设备将第一视频也就是有噪声的原始视频输入到预先训练成的神经网络中，利用该神经网络对第一视频进行去噪处理，也就是说通过该神经网络去除第一视频中的噪声得到干净的第二视频。

步骤 S103、输出所述第二视频。

视频处理设备进一步输出干净的第二视频。例如，第一视频是无人机搭载的拍摄设备拍摄的视频，视频处理设备设置在无人机，则第一视频经过视频处理设备的处理可转换为干净的第二视频。无人机可进一步通过通信系统将干净的第二视频发送给地面站，以供用户观赏。

本实施例通过将原始的带有噪声的第一视频输入到预先训练成的神经网络中，该神经网络是通过干净的第一训练视频包括的至少一个第一时空域立方体和加噪的第二训练视频包括的至少一个第二时空域立方体训练得到的，通过该神经网络对第一视频进行去噪处理以生成第二视频，相比于现有技术中基于运动估计的视频去噪方法，提高了视频去噪的计算复杂度，相比于现有技术中无需运动估计的视频去噪方法，提高了视频去噪的效果。

本发明实施例提供一种视频处理方法。图7为本发明另一实施例提供的视频处理方法的流程图。如图7所示，在图1所示实施例的基础上，步骤S101将第一视频输入神经网络之前，还包括：根据第一训练视频和第二训练视频训练所述神经网络。具体的，根据第一训练视频和第二训练视频训练所述神经网络包括如下步骤：

步骤S701、根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型。

具体的，步骤S701根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，包括如图8所示的步骤S7011和步骤S7012：

步骤S7011、对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理。

具体的，所述对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理，包括：根据所述第一时空域立方体包括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

如图5所示，来自图像帧21-25相同位置的子图像51、子图像53、子图像55、子图像57、子图像59构成一个第一时空域立方体61。以第一时空域立方体61为例，第一时空域立方体61包括子图像51、子图像53、子图像55、子图像57、子图像59，由于子图像51、子图像53、子图像55、子图像57、子图像59大小相同，假设均为 2×2 ，此处只是示意性说明，并限定各子图像的大小。也就是说，子图像51、子图像53、子图像55、子图像57、子图像59分别为2行2列的二维矩形块，如图9所示，假设子图像51的4个像素点的像素值分别为h11、h12、h13、h14；子图像53的4个像素点的像素值分别为h31、h32、h33、h34；子图像55的4个像素点的像素值分别为h51、h52、h53、h54；子图像57的4个像素点的像素值分别为h71、h72、h73、h74；子图像59的4个像素点的像素值

分别为 h_{91} 、 h_{92} 、 h_{93} 、 h_{94} 。进一步的，计算子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 的第 1 行第 1 列的像素值的平均值得到 $H1$ ，即 $H1$ 等于 h_{11} 、 h_{31} 、 h_{51} 、 h_{71} 、 h_{91} 的平均值，同理，计算子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 的第 1 行第 2 列的像素值的平均值得到 $H2$ ，即 $H2$ 等于 h_{12} 、 h_{32} 、 h_{52} 、 h_{72} 、 h_{92} 的平均值；计算子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 的第 2 行第 1 列的像素值的平均值得到 $H3$ ，即 $H3$ 等于 h_{13} 、 h_{33} 、 h_{53} 、 h_{73} 、 h_{93} 的平均值；计算子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 的第 2 行第 2 列的像素值的平均值得到 $H4$ ，即 $H4$ 等于 h_{14} 、 h_{34} 、 h_{54} 、 h_{74} 、 h_{94} 的平均值。 $H1$ 、 $H2$ 、 $H3$ 、 $H4$ 构成第一均值图像 90，即第一均值图像 90 中每个位置的像素值是子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 中相同位置的像素值的平均值。

进一步的，如图 10 所示，将子图像 51 中各位置的像素值减去第一均值图像 90 中相同位置的像素值得到新的子图像 510，即将子图像 51 的 h_{11} 减去第一均值图像 90 的 $H1$ 得到 $H11$ ，将子图像 51 的 h_{12} 减去第一均值图像 90 的 $H2$ 得到 $H12$ ，将子图像 51 的 h_{13} 减去第一均值图像 90 的 $H3$ 得到 $H13$ ，子图像 51 的 h_{14} 减去第一均值图像 90 的 $H4$ 得到 $H14$ 。 $H11$ 、 $H12$ 、 $H13$ 、 $H14$ 构成新的子图像 510。同理，将子图像 53 中各位置的像素值减去第一均值图像 90 中相同位置的像素值得到新的子图像 530，子图像 530 包括像素值 $H31$ 、 $H32$ 、 $H33$ 、 $H34$ 。将子图像 55 中各位置的像素值减去第一均值图像 90 中相同位置的像素值得到新的子图像 550，子图像 550 包括像素值 $H51$ 、 $H52$ 、 $H53$ 、 $H54$ 。将子图像 57 中各位置的像素值减去第一均值图像 90 中相同位置的像素值得到新的子图像 570，子图像 570 包括像素值 $H71$ 、 $H72$ 、 $H73$ 、 $H74$ 。将子图像 59 中各位置的像素值减去第一均值图像 90 中相同位置的像素值得到新的子图像 590，子图像 590 包括像素值 $H91$ 、 $H92$ 、 $H93$ 、 $H94$ 。

如图 5 所示，子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 分别来自相邻的图像帧 21-25，而相邻的图像帧之间的关联性 or 相似性较强。如图 9 所示，根据子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 计算出第一均值图像 90。如图 10 所示，再将子图像 51、子图

像 53、子图像 55、子图像 57、子图像 59 中的每一个子图像分别减去第一均值图像 90 后得到子图像 510、子图像 530、子图像 550、子图像 570、子图像 590。子图像 510、子图像 530、子图像 550、子图像 570、子图像 590 之间的关联性或相似性较低，因此由子图像 510、子图像 530、子图像 550、子图像 570、子图像 590 构成的时空域立方体要比子图像 51、子图像 53、子图像 55、子图像 57、子图像 59 构成的第一时空域立方体 61 具有更强的稀疏性，即由子图像 510、子图像 530、子图像 550、子图像 570、子图像 590 构成的时空域立方体是对第一时空域立方体 61 进行稀疏处理后的第一时空域立方体。

如图 6 所示，第一训练视频 20 中包括多个第一时空域立方体，需要对多个第一时空域立方体中的每个第一时空域立方体进行稀疏处理，具体的，对多个第一时空域立方体中的每个第一时空域立方体进行稀疏处理的原理和过程与对第一时空域立方体 61 进行稀疏处理的原理和过程一致，此处不再赘述。

不失一般性，公式 (1) 表示的第一时空域立方体 v_x 中包括 $2h+1$ 个子图像，根据第一时空域立方体 v_x 中包括的 $2h+1$ 个子图像确定的第一均值图像表示为 $\mu(i, j)$ ， $\mu(i, j)$ 的计算公式如下公式 (2) 所示：

$$\mu(i, j) = \frac{1}{2h+1} \sum_{s=-h}^h \{x_{t0+s}(i, j)\} \quad (2)$$

对第一时空域立方体 v_x 进行稀疏处理后得到的时空域立方体表示为

$\bar{v}_x$ ， $\bar{v}_x$ 可表示为公式 (3)：

$$\bar{v}_x = \{\bar{x}_{t0+s}(i, j)\}_{s=-h}^h = \{x_{t0+s}(i, j) - \mu(i, j)\}_{s=-h}^h \quad (3)$$

步骤 S7012、根据每个稀疏处理后的第一时空域立方体训练局部先验模型。

由于 $\bar{v}_x$ 比 v_x 具有更强的稀疏性，因此，根据第一训练视频 20 中每个稀疏处理后的第一时空域立方体建模更容易，具体的，将第一训练视频 20 中每个稀疏处理后的第一时空域立方体中的每个二维矩形块构成一个列向量，例如，子图像 510、子图像 530、子图像 550、子图像 570、子图像 590 构成的时空域立方体是第一训练视频 20 中的一个稀疏处理后的第一时空域立方体，分别将子图像 510、子图像 530、子图像 550、子图像 570、

子图像 590 各自的 4 个像素值构成一个 4*1 的列向量，得到 5 个 4*1 的列向量。同理，将第一训练视频 20 中其他的稀疏处理后的第一时空域立方体中的每个二维矩形块构成一个列向量，进一步采用混合高斯模型 (Gaussian Mixture Model, 简称 GMM) 对第一训练视频 20 中每个稀疏处理后的第一时空域立方体对应的列向量建模得到局部先验模型，该局部先验模型具体为局部时空先验 (Local Volumetric Prior, 简称 LVP) 模型，同时约束同一个稀疏处理后的第一时空域立方体中的所有二维矩形块属于同一个高斯类。从而得到如下公式 (4) 所示的似然函数 $P(\bar{V}_x)$ ：

$$P(\bar{V}_x) = \sum_{k=1}^K \pi_k \prod_{s=-h}^h N(\bar{x}_{t0+s}(i, j) \setminus \mu_k, \Sigma_k) \quad (4)$$

其中， K 表示高斯类的个数， k 表示第 k 个高斯类， π_k 表示第 k 个高斯类的权重， μ_k 表示第 k 个高斯类的均值， Σ_k 表示第 k 个高斯类的协方差矩阵， N 表示概率密度函数。

进一步的，对每个高斯类的协方差矩阵 Σ_k 进行奇异值分解，得到正交字典 D_k ，正交字典 D_k 和协方差矩阵 Σ_k 之间的关系如公式 (5) 所示：

$$\Sigma_k = D_k \Lambda_k D_k^T \quad (5)$$

其中，正交字典 D_k 是由协方差矩阵 Σ_k 的特征向量组成的， Λ_k 表示特征值矩阵。

步骤 S702、根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频。

具体的，步骤 S702 根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，包括如图 11 所示的步骤 S7021 和步骤 S7022：

步骤 S7021、对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理。

具体的，所述对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理，包括：根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个第二子图像在所述位置的像素值的平均值；将所述第二时空域立方体包括的多个第二子图像中的每

个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

假设第二训练视频表示为 Y ， Y_t 表示第二训练视频中的第 t 帧图像， $y_t(i, j)$ 表示第二训练视频中第 t 帧图像中的一个子图像， (i, j) 表示该子图像在第 t 帧图像中的位置，也就是说 $y_t(i, j)$ 表示从添加有噪声的第二训练视频中截取的一个二维矩形块， (i, j) 表示二维矩形块的空域索引， t 表示二维矩形块的时域索引。

将第二训练视频中相邻的若干图像帧中位置相同、大小相同的子图像构成一个集合，该集合记为第二时空域立方体 V_y ，第二训练视频 Y 可以被划分出多个第二时空域立方体 V_y 。第二时空域立方体的划分原理和过程和第一时空域立方体的划分原理和过程一致，此处不再赘述。不失一般性，一个第二时空域立方体 V_y 可以表示为如下公式 (6)：

$$V_y = \{y_{t-l}(i, j), K, y_t(i, j), K, y_{t+l}(i, j)\} = \{y_{t+s}(i, j)\}_{s=-l}^l \quad (6)$$

第二时空域立方体 V_y 中包括 $2l+1$ 个子图像，该 $2l+1$ 个子图像的第二均值图像表示为 $\eta(i, j)$ ， $\eta(i, j)$ 的计算公式如下公式 (7) 所示：

$$\eta(i, j) = \frac{1}{2l+1} \sum_{s=-l}^l \{y_{t+s}(i, j)\} \quad (7)$$

进一步对第二时空域立方体 V_y 进行稀疏处理，稀疏处理后得到的第二时空域立方体表示为 $\overline{V}_y$ ， $\overline{V}_y$ 可表示为公式 (8)：

$$\overline{V}_y = \{\overline{y}_{t+s}(i, j)\}_{s=-l}^l = \{y_{t+s}(i, j) - \eta(i, j)\}_{s=-l}^l \quad (8)$$

稀疏处理后得到的第二时空域立方体 $\overline{V}_y$ 比第二时空域立方体 V_y 具有更强的稀疏性。由于第二训练视频 Y 可以被划分出多个第二时空域立方体 V_y ，对每个第二时空域立方体 V_y 的稀疏处理过程均可以采用公式 (7)、公式 (8) 的方法。

步骤 S7022、根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。

具体的，根据步骤 S7012 确定出的局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理得到初始去噪后的第二训练视频。

步骤 S703、根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

具体的，所述根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络，包括：将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。可选的，以初始去噪后的第二训练视频作为训练数据、以第一训练视频作为标签训练出的神经网络是一个深度的神经网络。

本实施例通过干净的第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，再根据训练出的局部先验模型对带有噪声的第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频，最后将初始去噪后的第二训练视频作为训练数据，将干净的第一训练视频作为标签训练神经网络，该神经网络是一个深度的神经网络，深度的神经网络可提高对噪声视频的去噪效果。

本发明实施例提供一种视频处理方法。图 12 为本发明另一实施例提供的视频处理方法的流程图。如图 12 所示，在图 7 所示实施例的基础上，步骤 S7022 根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理，可以包括如下步骤：

步骤 S1201、根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类。

具体的，根据公式（4）所示的似然函数 $P(\bar{V}_x)$ 确定稀疏处理后得到的第二时空域立方体 $\bar{V}_y$ 属于混合高斯模型中的哪一个高斯类。由于稀疏处理后得到的第二时空域立方体 $\bar{V}_y$ 可以是多个，因此，根据公式（4）所示的似然函数 $P(\bar{V}_x)$ 确定每一个 $\bar{V}_y$ 所属的高斯类。

步骤 S1202、根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

具体的，根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始

去噪处理，包括如下步骤 S12021 和步骤 S12022：

步骤 S12021、根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵。

所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵，包括：对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

假设稀疏处理后得到的第二时空域立方体 $\bar{V}_y$ 属于混合高斯模型中的第 k 个高斯类，根据上述公式（5）所述的对第 k 个高斯类的协方差矩阵 Σ_k 进行奇异值分解可确定出第 k 个高斯类的正交字典 D_k 和特征值矩阵 Λ_k 。

步骤 S12022、根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

所述根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：根据所述特征值矩阵确定权值矩阵；根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

进一步的，根据特征值矩阵 Λ_k 确定权值矩阵 W 。以稀疏处理后的第二时空域立方体 $\bar{V}_y$ 中的一个子图像 $\bar{y}_t(i, j)$ 为例，根据第 k 个高斯类的正交字典 D_k 和权值矩阵 W ，采用带权稀疏编码的方法对 $\bar{y}_t(i, j)$ 进行初始去噪处理的方法如公式（9）和公式（10）所示：

$$\hat{x} = \arg \min_x \left\| \bar{y}_t(i, j) - \bar{x} \right\|_2^2 + \|W^T \alpha\| \quad (9)$$

$$s.t. \bar{x} = D_k \alpha \quad (10)$$

其中， $\bar{x}$ 表示需要求的对 $\bar{y}_t(i, j)$ 进行初始去噪处理后的子图像， $\hat{x}$ 表示 $\bar{x}$ 的估计值。进一步的，在 $\hat{x}$ 的基础上再加上第二均值图像 $\eta(i, j)$ 可得到对 $y_t(i, j)$ 进行初始去噪处理后的子图像。 $y_t(i, j)$ 是第二时空域立方体 V_y 中的一个子图像， $\bar{y}_t(i, j)$ 是对第二时空域立方体 V_y 进行稀疏处理后 $y_t(i, j)$ 对应的子图像，即 $y_t(i, j)$ 减去 $\eta(i, j)$ 可得到 $\bar{y}_t(i, j)$ ，因此，当计算出对 $\bar{y}_t(i, j)$ 进行初始去噪处理后的子图像的估计值 $\hat{x}$ 时，在 $\hat{x}$ 的基础上再加上第二均值图像 $\eta(i, j)$ 即可得到对 $y_t(i, j)$ 进行初始去噪处理后的子图像。同理可计算出对第二时空域立方体 V_y 中每个子图像进行初始去噪处理后的子图像。由于第二

训练视频 V 可以被划分出多个第二时空域立方体 v_y ，因此采用前述方法可以对多个第二时空域立方体 v_y 中每一个第二时空域立方体 v_y 中的每个子图像进行初始去噪处理，从而得到初始去噪后的第二训练视频 $\hat{X}_t$ ，在初始去噪后的第二训练视频 $\hat{X}_t$ 中，大量的噪声被抑制。

- 5 在本实施例中，为了能够学习视频全局的时空结构信息，设计一个感受野大小为 35×35 的神经网络，神经网络的输入为初始去噪后的第二训练视频 $\hat{X}_t$ 的相邻帧 $\{\hat{X}_{t0+s}\}_{s=-h}^h$ ，恢复其最中间的一帧 X_{t0} ，由于大小 3×3 的卷积核在神经网络中得到了广泛运动，本实施例可以采用 3×3 的卷积核，并设计了 17 层的网络结构。在网络的第一层，由于输入是多帧，可以采用 64
- 10 个 $3 \times 3 \times (2h+1)$ 的卷积核，在网络的最后一层，为了重构出一张图像，可以采用 $3 \times 3 \times 64$ 的卷积层。网络的中间 15 层，可以采用 64 个 $3 \times 3 \times 64$ 的卷积层，网络的损失函数如下公式 (11) 所示：

$$l(\Theta) = \frac{1}{2} \sum_{t0} \left\| F(\{\hat{X}_{t0+s}\}_{s=-h}^h : \Theta) - (\hat{X}_{t0} - X_{t0}) \right\|_F^2 \quad (11)$$

- 其中， F 表示神经网络，最小化损失函数可计算出参数 Θ ，从而确定
- 15 出神经网络 F 。

可选的，本发明采用线性整流函数 (ReLU) 作为非线性层，并在卷积层和非线性层之间添加了归一化层。

- 本实施例通过局部先验模型确定稀疏处理后的第二时空域立方体所属的高斯类，根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对稀疏处理后的第二时空域立方体进行初始去噪处理，实现了无需运动估计的局部时空先验辅助的深度神经网络视频去噪方法。
- 20 用带权稀疏编码的方法对稀疏处理后的第二时空域立方体进行初始去噪处理，实现了无需运动估计的局部时空先验辅助的深度神经网络视频去噪方法。

- 本发明实施例提供一种视频处理设备。图 13 为本发明实施例提供的
- 25 视频处理设备的结构图，如图 13 所示，视频处理设备 130 包括一个或多个处理器 131，单独或协同工作，一个或多个处理器 131，用于：将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；利用所述神经网络对所述第一视频

进行去噪处理以生成第二视频；以及输出所述第二视频。

可选的，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。

本发明实施例提供的视频处理设备的具体原理和实现方式均与图 1 所示实施例类似，此处不再赘述。

5 本实施例通过将原始的带有噪声的第一视频输入到预先训练成的神经网络中，该神经网络是通过干净的第一训练视频包括的至少一个第一时空域立方体和加噪的第二训练视频包括的至少一个第二时空域立方体训练得到的，通过该神经网络对第一视频进行去噪处理以生成第二视频，相比于现有技术中基于运动估计的视频去噪方法，提高了视频去噪的计算复杂度，相比于现有技术中无需运动估计的视频去噪方法，提高了视频去噪的效果。

15 本发明实施例提供一种视频处理设备。在图 13 所示实施例提供的技术方案的基础上，一个或多个处理器 131 将第一视频输入神经网络之前，还用于：根据第一训练视频和第二训练视频训练所述神经网络。

20 具体的，一个或多个处理器 131 根据第一训练视频和第二训练视频训练所述神经网络时，具体用于：根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型；根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频；根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

 可选的，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

25 一个或多个处理器 131 根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型时，具体用于：对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理；根据每个稀疏处理后的第一时空域立方体训练局部先验模型。一个或多个处理器 131 对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理时，具体用于：根据所述第一时空域立方体包

括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

- 5 可选的，所述第二时空域立方体包括多个第二子图像，所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。

一个或多个处理器 131 根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理时，具体用于：对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理；根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。一个或多个处理器 131 对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理时，具体用于：根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个第二子图像在所述位置的像素值的平均值；将所述第二时空域立方体包括的多个第二子图像中的每个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

- 20 本发明实施例提供的视频处理设备的具体原理和实现方式均与图 7、图 8、图 11 所示实施例类似，此处不再赘述。

本实施例通过干净的第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，再根据训练出的局部先验模型对带有噪声的第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频，最后将初始去噪后的第二训练视频作为训练数据，将干净的第一训练视频作为标签训练神经网络，该神经网络是一个深度的神经网络，深度的神经网络可提高对噪声视频的去噪效果。

- 30 本发明实施例提供一种视频处理设备。在图 7、图 8、图 11 所示实施

例提供的技术方案的基础上，一个或多个处理器 131 根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理时，具体用于：根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类；根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

具体的，一个或多个处理器 131 根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理时，具体用于：根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵；根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

一个或多个处理器 131 根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵时，具体用于：对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

一个或多个处理器 131 根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理时，具体用于：根据所述特征值矩阵确定权值矩阵；根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

可选的，一个或多个处理器 131 根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络时，具体用于：将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。

本发明实施例提供的视频处理设备的具体原理和实现方式均与图 12 所示实施例类似，此处不再赘述。

本实施例通过局部先验模型确定稀疏处理后的第二时空域立方体所属的高斯类，根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对稀疏处理后的第二时空域立方体进行初始去噪

处理，实现了无需运动估计的局部时空先验辅助的深度神经网络视频去噪方法。

本发明实施例提供一种无人机。图 14 为本发明实施例提供的无人机的结构图，如图 14 所示，无人机 100 包括：机身、动力系统、飞行控制器 118 和视频处理设备 109，所述动力系统包括如下至少一种：电机 107、螺旋桨 106 和电子调速器 117，动力系统安装在所述机身，用于提供飞行动力；飞行控制器 118 与所述动力系统通讯连接，用于控制所述无人机飞行。

另外，如图 8 所示，无人机 100 还包括：传感系统 108、通信系统 110、支撑设备 102、拍摄设备 104，其中，支撑设备 102 具体可以是云台，通信系统 110 具体可以包括接收机，接收机用于接收地面站 112 的天线 114 发送的无线信号，116 表示接收机和天线 114 通信过程中产生的电磁波。

视频处理设备 109 可以对拍摄设备 104 拍摄到的视频进行视频处理，视频处理的方法和上述方法实施例类似，视频处理设备 109 的具体原理和实现方式均与上述实施例类似，此处不再赘述。

本实施例通过将原始的带有噪声的第一视频输入到预先训练成的神经网络中，该神经网络是通过干净的第一训练视频包括的至少一个第一时空域立方体和加噪的第二训练视频包括的至少一个第二时空域立方体训练得到的，通过该神经网络对第一视频进行去噪处理以生成第二视频，相比于现有技术中基于运动估计的视频去噪方法，提高了视频去噪的计算复杂度，相比于现有技术中无需运动估计的视频去噪方法，提高了视频去噪的效果。

本发明实施例提供一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被一个或多个处理器执行时实现以下步骤：将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以

及

输出所述第二视频。

可选的，所述将第一视频输入神经网络之前，还包括：

根据第一训练视频和第二训练视频训练所述神经网络。

- 5 可选的，所述根据第一训练视频和第二训练视频训练所述神经网络，包括：

根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型；

- 10 根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频；

根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

可选的，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。

- 15 可选的，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

可选的，所述根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，包括：

- 20 对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理；

根据每个稀疏处理后的第一时空域立方体训练局部先验模型。

可选的，所述对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理，包括：

- 25 根据所述第一时空域立方体包括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；

将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

- 30 可选的，其特征在于，所述第二时空域立方体包括多个第二子图像，

所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。

可选的，所述根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，包括：对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理；

根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。

可选的，所述对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理，包括：

根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个第二子图像在所述位置的像素值的平均值；

将所述第二时空域立方体包括的多个第二子图像中的每个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

可选的，所述根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类；

根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

可选的，所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵；

根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

可选的，所述根据所述稀疏处理后的第二时空域立方体所属的高斯

类，确定所述高斯类的字典和特征值矩阵，包括：

对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

可选的，所述根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述特征值矩阵确定权值矩阵；

根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

可选的，所述根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络，包括：

将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。

在本发明所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计

计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）或处理器（processor）执行本发明各个实施例所述方法的部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、随机存取存储器（Random Access Memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

本领域技术人员可以清楚地了解到，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。上述描述的装置的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权 利 要 求 书

1、一种视频处理方法，其特征在于，包括：

将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及

输出所述第二视频。

2、根据权利要求 1 所述的方法，其特征在于，所述将第一视频输入神经网络之前，还包括：

根据第一训练视频和第二训练视频训练所述神经网络。

3、根据权利要求 2 所述的方法，其特征在于，所述根据第一训练视频和第二训练视频训练所述神经网络，包括：

根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型；

根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频；

根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

4、根据权利要求 3 所述的方法，其特征在于，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。

5、根据权利要求 3 或 4 所述的方法，其特征在于，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

6、根据权利要求 5 所述的方法，其特征在于，所述根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，包括：

对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理；

根据每个稀疏处理后的第一时空域立方体训练局部先验模型。

7、根据权利要求 6 所述的方法，其特征在于，所述对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理，包括：

- 5 根据所述第一时空域立方体包括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；

将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

- 10 8、根据权利要求 3-7 任一项所述的方法，其特征在于，所述第二时空域立方体包括多个第二子图像，所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。

- 15 9、根据权利要求 8 所述的方法，其特征在于，所述根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，包括：对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理；

根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。

- 20 10、根据权利要求 9 所述的方法，其特征在于，所述对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理，包括：

- 25 根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个第二子图像在所述位置的像素值的平均值；

将所述第二时空域立方体包括的多个第二子图像中的每个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

- 30 11、根据权利要求 9 或 10 所述的方法，其特征在于，所述根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类；

根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

- 5 12、根据权利要求 11 所述的方法，其特征在于，所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵；

- 10 根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

13、根据权利要求 12 所述的方法，其特征在于，所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵，包括：

- 15 对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

14、根据权利要求 12 所述的方法，其特征在于，所述根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

- 20 根据所述特征值矩阵确定权值矩阵；

根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

- 25 15、根据权利要求 3-14 任一项所述的方法，其特征在于，所述根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络，包括：

将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。

16、一种视频处理设备，其特征在于，包括一个或多个处理器，单独或协同工作，所述一个或多个处理器用于：

- 30 将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频

和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及

5 输出所述第二视频。

17、根据权利要求 16 所述的视频处理设备，其特征在于，所述一个或多个处理器将第一视频输入神经网络之前，还用于：

根据第一训练视频和第二训练视频训练所述神经网络。

10 18、根据权利要求 17 所述的视频处理设备，其特征在于，所述一个或多个处理器根据第一训练视频和第二训练视频训练所述神经网络时，具体用于：

根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型；

15 根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频；

根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

20 19、根据权利要求 18 所述的视频处理设备，其特征在于，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。

20、根据权利要求 18 或 19 所述的视频处理设备，其特征在于，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

25 21、根据权利要求 20 所述的视频处理设备，其特征在于，所述一个或多个处理器根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型时，具体用于：

对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理；

30 根据每个稀疏处理后的第一时空域立方体训练局部先验模型。

22、根据权利要求 21 所述的视频处理设备，其特征在于，所述一个或多个处理器对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理时，具体用于：

5 根据所述第一时空域立方体包括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；

将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

23、根据权利要求 18-22 任一项所述的视频处理设备，其特征在于，
10 所述第二时空域立方体包括多个第二子图像，所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。

24、根据权利要求 23 所述的视频处理设备，其特征在于，所述一个或多个处理器根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理时，具体用于：
15

对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理；

20 根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。

25、根据权利要求 24 所述的视频处理设备，其特征在于，所述一个或多个处理器对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理时，具体用于：

25 根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个第二子图像在所述位置的像素值的平均值；

将所述第二时空域立方体包括的多个第二子图像中的每个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

26、根据权利要求 24 或 25 所述的视频处理设备，其特征在于，所述
30 一个或多个处理器根据所述局部先验模型对每个稀疏处理后的第二时空

域立方体进行初始去噪处理时，具体用于：

根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类；

根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

27、根据权利要求 26 所述的视频处理设备，其特征在于，所述一个或多个处理器根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理时，具体用于：

10 根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵；

根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

15 28、根据权利要求 27 所述的视频处理设备，其特征在于，所述一个或多个处理器根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵时，具体用于：

对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

20 29、根据权利要求 27 所述的视频处理设备，其特征在于，所述一个或多个处理器根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理时，具体用于：

根据所述特征值矩阵确定权值矩阵；

25 根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

30、根据权利要求 18-29 任一项所述的视频处理设备，其特征在于，所述一个或多个处理器根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络时，具体用于：

30 将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。

31、一种无人机，其特征在于，包括：

机身；

动力系统，安装在所述机身，用于提供飞行动力；

以及如权利要求 16-30 任一项所述的视频处理设备。

5 32、一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被一个或多个处理器执行时实现以下步骤：

将第一视频输入神经网络，所述神经网络的训练集包括第一训练视频和第二训练视频，所述第一训练视频包括至少一个第一时空域立方体，所述第二训练视频包括至少一个第二时空域立方体；

10 利用所述神经网络对所述第一视频进行去噪处理以生成第二视频；以及

输出所述第二视频。

33、根据权利要求 32 所述的计算机可读存储介质，其特征在于，所述将第一视频输入神经网络之前，还包括：

15 根据第一训练视频和第二训练视频训练所述神经网络。

34、根据权利要求 33 所述的计算机可读存储介质，其特征在于，所述根据第一训练视频和第二训练视频训练所述神经网络，包括：

根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型；

20 根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，得到初始去噪后的第二训练视频；

根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络。

25 35、根据权利要求 34 所述的计算机可读存储介质，其特征在于，第一训练视频为无噪声视频，所述第二训练视频为噪声视频。

36、根据权利要求 34 或 35 所述的计算机可读存储介质，其特征在于，所述第一时空域立方体包括多个第一子图像，所述多个第一子图像来自所述第一训练视频中相邻的多个第一视频帧，一个第一子图像来自一个第一视频帧，每个第一子图像在第一视频帧中的位置相同。

30

37、根据权利要求 36 所述的计算机可读存储介质，其特征在于，所述根据第一训练视频包括的至少一个第一时空域立方体训练局部先验模型，包括：

对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理；

根据每个稀疏处理后的第一时空域立方体训练局部先验模型。

38、根据权利要求 37 所述的计算机可读存储介质，其特征在于，所述对第一训练视频包括的至少一个第一时空域立方体中的每个第一时空域立方体分别进行稀疏处理，包括：

根据所述第一时空域立方体包括的多个第一子图像，确定第一均值图像，所述第一均值图像中每个位置的像素值是所述多个第一子图像中每个第一子图像在所述位置的像素值的平均值；

将所述第一时空域立方体包括的多个第一子图像中的每个第一子图像在所述位置的像素值减去所述第一均值图像中所述位置的像素值。

39、根据权利要求 34-38 任一项所述的计算机可读存储介质，其特征在于，所述第二时空域立方体包括多个第二子图像，所述多个第二子图像来自所述第二训练视频中相邻的多个第二视频帧，一个第二子图像来自一个第二视频帧，每个第二子图像在第二视频帧中的位置相同。

40、根据权利要求 39 所述的计算机可读存储介质，其特征在于，所述根据所述局部先验模型对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行初始去噪处理，包括：对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理；

根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理。

41、根据权利要求 40 所述的计算机可读存储介质，其特征在于，所述对第二训练视频包括的至少一个第二时空域立方体中的每个第二时空域立方体分别进行稀疏处理，包括：

根据所述第二时空域立方体包括的多个第二子图像，确定第二均值图像，所述第二均值图像中每个位置的像素值是所述多个第二子图像中每个

第二子图像在所述位置的像素值的平均值；

将所述第二时空域立方体包括的多个第二子图像中的每个第二子图像在所述位置的像素值减去所述第二均值图像中所述位置的像素值。

- 42、根据权利要求 40 或 41 所述的计算机可读存储介质，其特征在于，
5 所述根据所述局部先验模型对每个稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述局部先验模型确定所述稀疏处理后的第二时空域立方体所属的高斯类；

- 10 根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

43、根据权利要求 42 所述的计算机可读存储介质，其特征在于，所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

- 15 根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵；

根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

- 20 44、根据权利要求 43 所述的计算机可读存储介质，其特征在于，所述根据所述稀疏处理后的第二时空域立方体所属的高斯类，确定所述高斯类的字典和特征值矩阵，包括：

对所述高斯类的协方差矩阵进行奇异值分解，得到所述高斯类的字典和特征值矩阵。

- 25 45、根据权利要求 43 所述的计算机可读存储介质，其特征在于，所述根据所述高斯类的字典和特征值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理，包括：

根据所述特征值矩阵确定权值矩阵；

根据高斯类的字典和所述权值矩阵，采用带权稀疏编码的方法对所述稀疏处理后的第二时空域立方体进行初始去噪处理。

- 30 46、根据权利要求 34-45 任一项所述的计算机可读存储介质，其特征

在于，所述根据所述初始去噪后的第二训练视频和所述第一训练视频训练所述神经网络，包括：

将所述初始去噪后的第二训练视频作为训练数据，将所述第一训练视频作为标签训练所述神经网络。

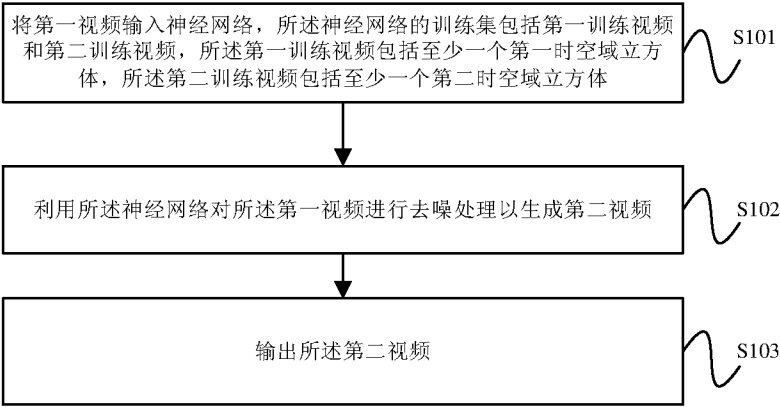


图 1

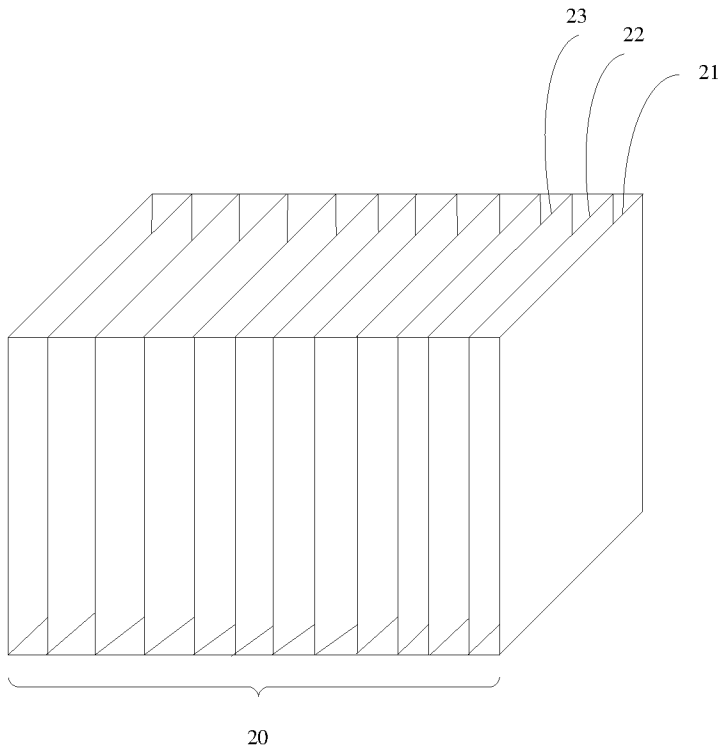


图 2

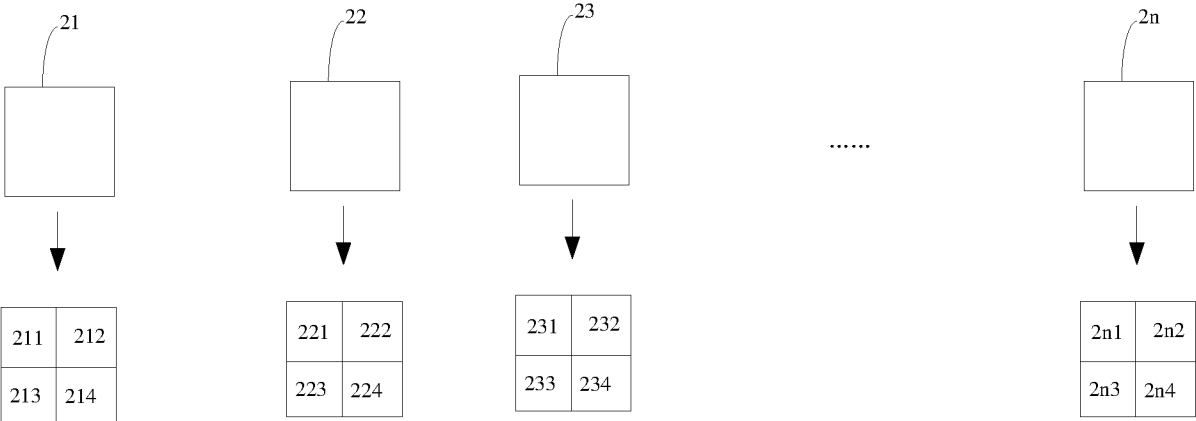


图 3

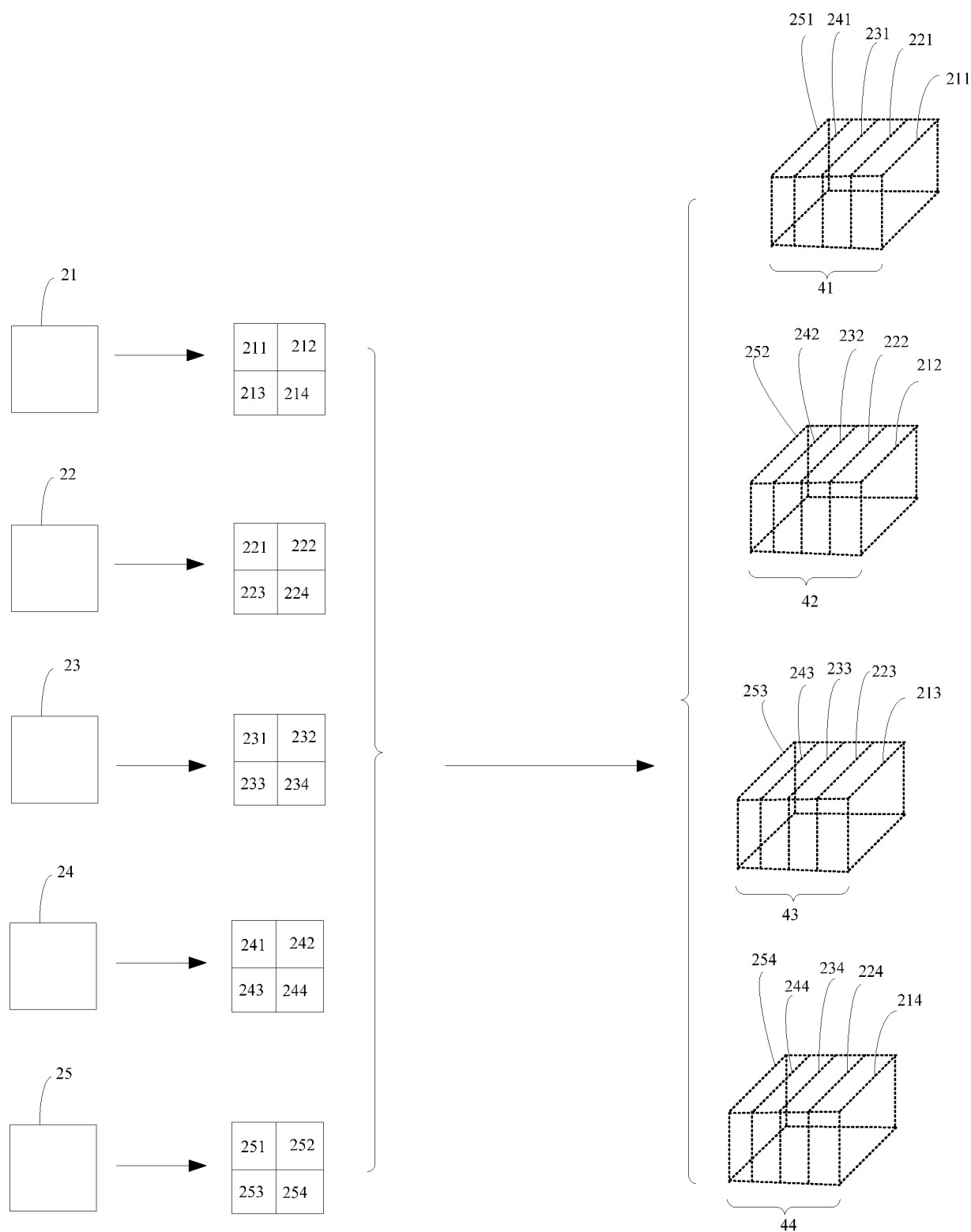


图 4

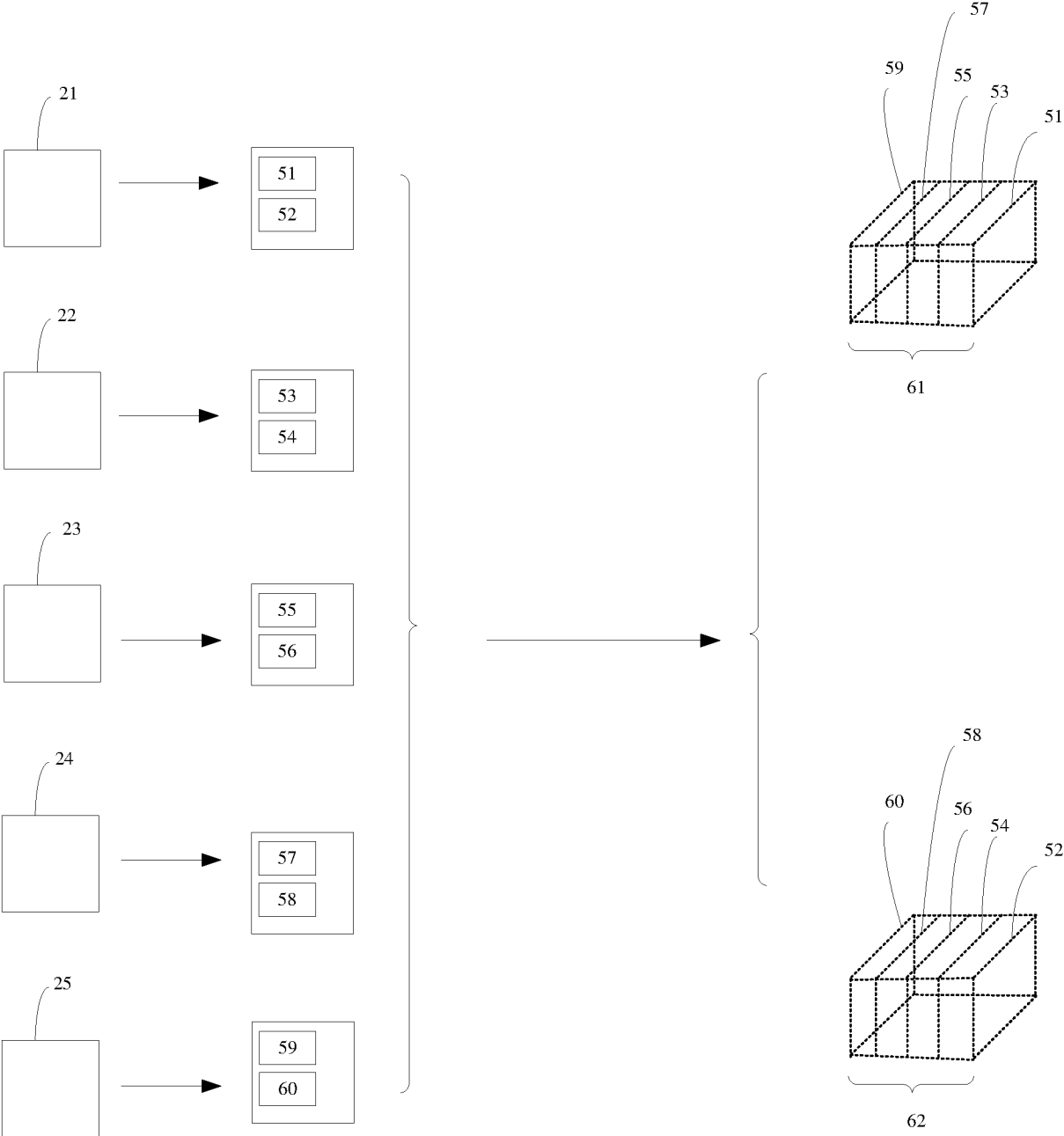


图 5

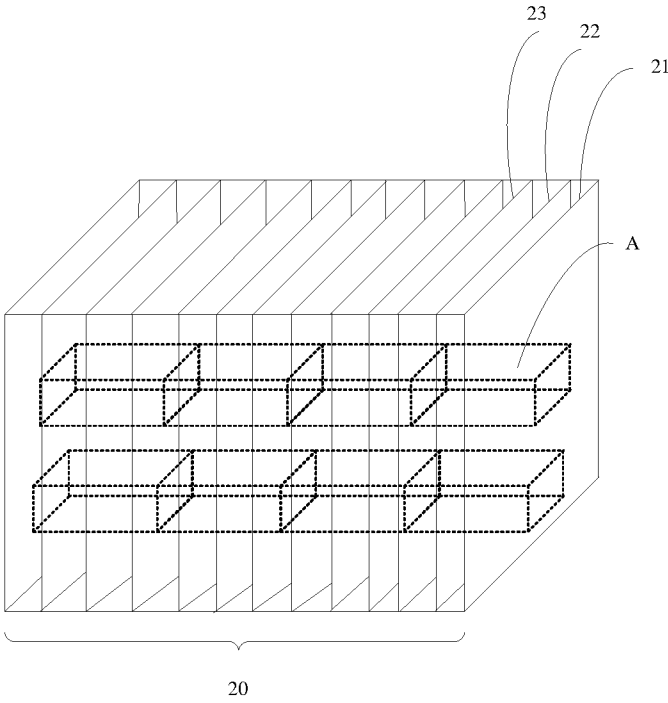


图 6

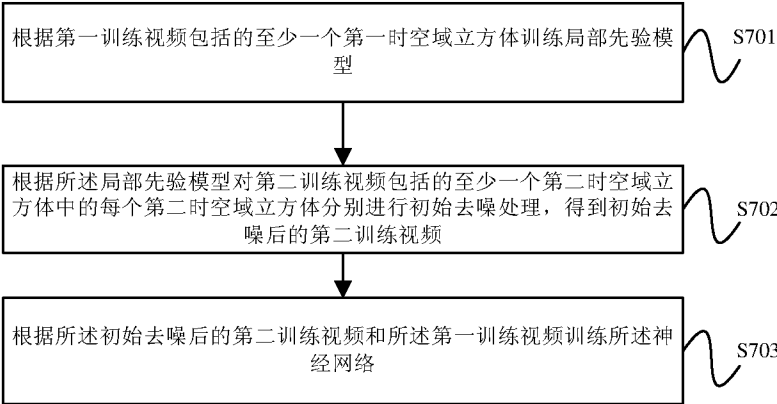


图 7

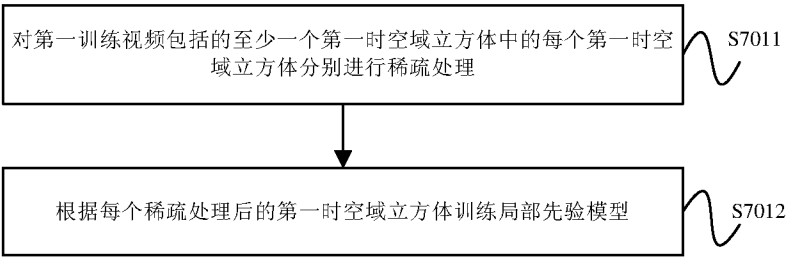


图 8

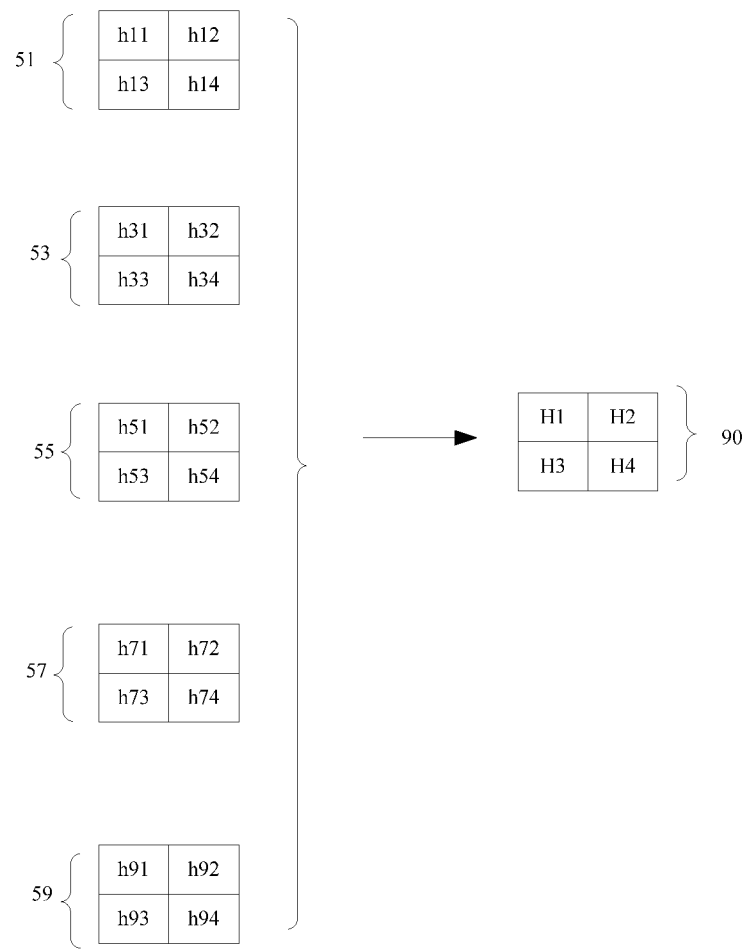


图 9

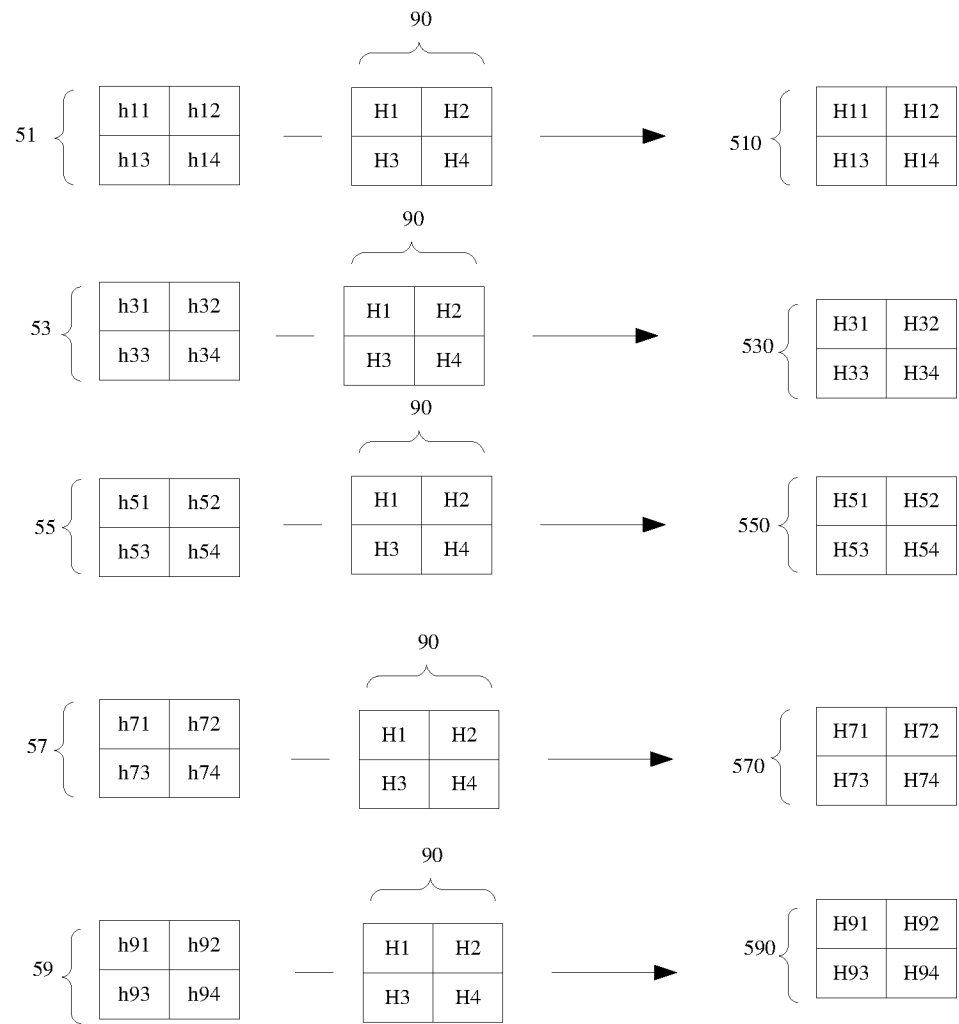


图 10

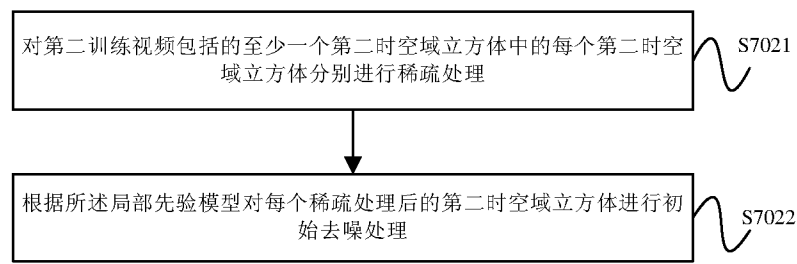


图 11

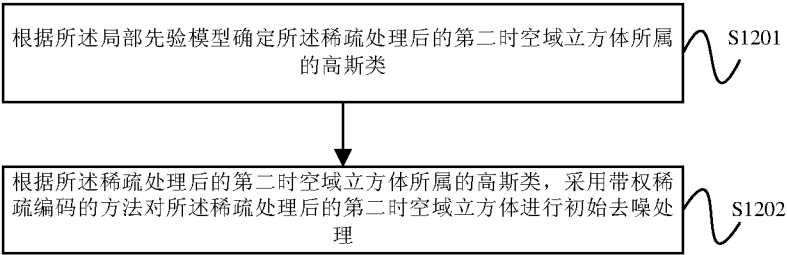


图 12

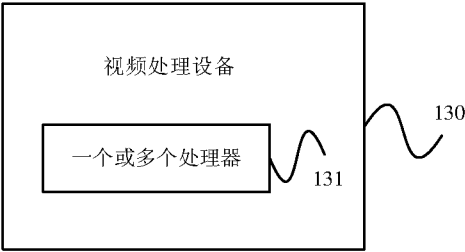


图 13

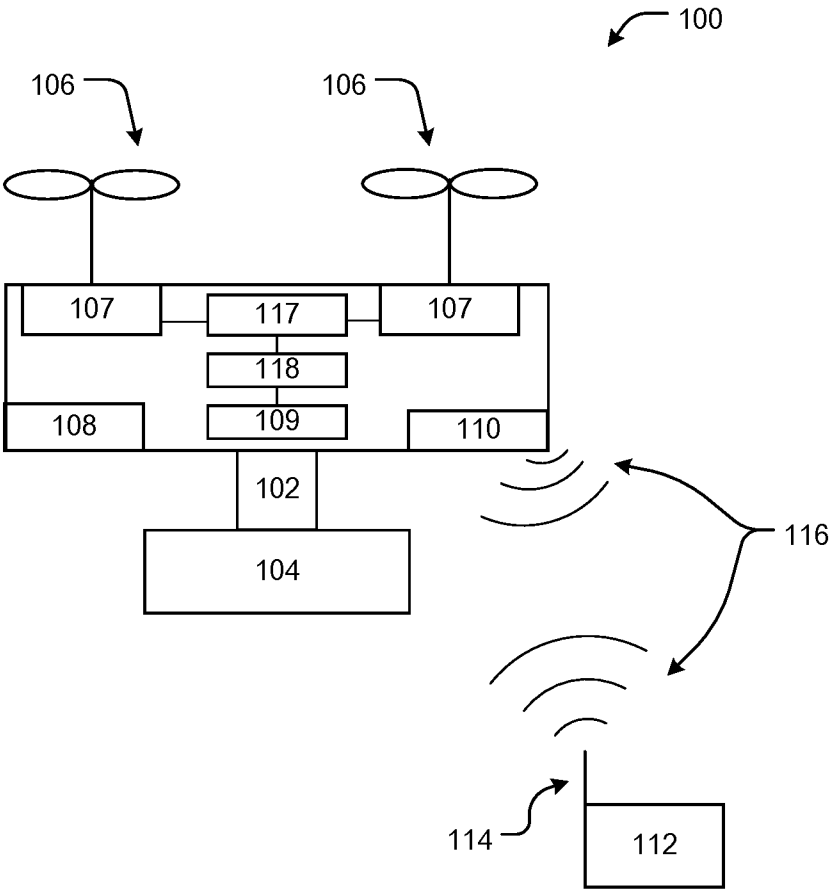


图 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/106735

A. CLASSIFICATION OF SUBJECT MATTER

G06T 5/00 (2006.01) i; G06N 3/04 (2006.01) i; H04N 5/213 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T; G06N; H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, EPODOC, WPI, IEEE: 视频, 图像, 去噪, 神经网络, 训练, 集, 时间, 空间, 时空, 立方体, 稀疏, video, image, denoising, neural network, train, set, time, space, cube, sparse

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106331433 A (SHANGHAI JIAOTONG UNIVERSITY) 11 January 2017 (11.01.2017), description, paragraphs [0008], [0021] and [0034]-[0061], and figures 1-4	1, 2, 16, 17, 32, 33
Y	CN 106331433 A (SHANGHAI JIAOTONG UNIVERSITY) 11 January 2017 (11.01.2017), description, paragraphs [0008], [0021] and [0034]-[0061], and figures 1-4	31
Y	CN 105791702 A (WANG, Zhengzuo) 20 July 2016 (20.07.2016), description, paragraphs [0026]-[0030], and figure 1	31
A	CN 107248144 A (SOUTHEAST UNIVERSITY) 13 October 2017 (13.10.2017), entire document	1-46
A	CN 106204467 A (SHENZHEN RESEARCH INSTITUTE OF FUTURE MEDIA COMPUTING et al.) 07 December 2016 (07.12.2016), entire document	1-46

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 19 June 2018	Date of mailing of the international search report 06 July 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer YU, Lina Telephone No. (86-10) 53961756

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CN2017/106735

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107133948 A (UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA) 05 September 2017 (05.09.2017), entire document	1-46
A	CN 104820974 A (SCIENTIFIC AND TECHNOLOGICAL INSTITUTES OF ZHEJIANG) 05 August 2015 (05.08.2015), entire document	1-46
A	US 2017084007 A1 (WRNCH, INC.) 23 March 2017 (23.03.2017), entire document	1-46
A	US 9449371 B1 (PIXELWORKS, INC.) 20 September 2016 (20.09.2016), entire document	1-46

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/106735

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 106331433 A	11 January 2017	None	
CN 105791702 A	20 July 2016	None	
CN 107248144 A	13 October 2017	None	
CN 106204467 A	07 December 2016	None	
CN 107133948 A	05 September 2017	None	
CN 104820974 A	05 August 2015	None	
US 2017084007 A1	23 March 2017	WO 2015172235 A1	19 November 2015
US 9449371 B1	20 September 2016	None	

国际检索报告

国际申请号

PCT/CN2017/106735

A. 主题的分类 G06T 5/00(2006.01)i; G06N 3/04(2006.01)i; H04N 5/213(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类																													
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06T; G06N; H04N 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNPAT, CNKI, EPODOC, WPI, IEEE: 视频, 图像, 去噪, 神经网络, 训练, 集, 时间, 空间, 时空, 立方体, 稀疏, video, image, denoising, neural network, train, set, time, space, cube, sparse																													
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4</td> <td>1, 2, 16, 17, 32, 33</td> </tr> <tr> <td>Y</td> <td>CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4</td> <td>31</td> </tr> <tr> <td>Y</td> <td>CN 105791702 A (王正作) 2016年 7月 20日 (2016 - 07 - 20) 说明书第26-30段, 附图1</td> <td>31</td> </tr> <tr> <td>A</td> <td>CN 107248144 A (东南大学) 2017年 10月 13日 (2017 - 10 - 13) 全文</td> <td>1-46</td> </tr> <tr> <td>A</td> <td>CN 106204467 A (深圳市未来媒体技术研究院 等) 2016年 12月 7日 (2016 - 12 - 07) 全文</td> <td>1-46</td> </tr> <tr> <td>A</td> <td>CN 107133948 A (电子科技大学) 2017年 9月 5日 (2017 - 09 - 05) 全文</td> <td>1-46</td> </tr> <tr> <td>A</td> <td>CN 104820974 A (浙江科技学院) 2015年 8月 5日 (2015 - 08 - 05) 全文</td> <td>1-46</td> </tr> <tr> <td>A</td> <td>US 2017084007 A1 (WRNCH, INC.) 2017年 3月 23日 (2017 - 03 - 23) 全文</td> <td>1-46</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4	1, 2, 16, 17, 32, 33	Y	CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4	31	Y	CN 105791702 A (王正作) 2016年 7月 20日 (2016 - 07 - 20) 说明书第26-30段, 附图1	31	A	CN 107248144 A (东南大学) 2017年 10月 13日 (2017 - 10 - 13) 全文	1-46	A	CN 106204467 A (深圳市未来媒体技术研究院 等) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-46	A	CN 107133948 A (电子科技大学) 2017年 9月 5日 (2017 - 09 - 05) 全文	1-46	A	CN 104820974 A (浙江科技学院) 2015年 8月 5日 (2015 - 08 - 05) 全文	1-46	A	US 2017084007 A1 (WRNCH, INC.) 2017年 3月 23日 (2017 - 03 - 23) 全文	1-46
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																											
X	CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4	1, 2, 16, 17, 32, 33																											
Y	CN 106331433 A (上海交通大学) 2017年 1月 11日 (2017 - 01 - 11) 说明书第8、21、34-61段, 附图1-4	31																											
Y	CN 105791702 A (王正作) 2016年 7月 20日 (2016 - 07 - 20) 说明书第26-30段, 附图1	31																											
A	CN 107248144 A (东南大学) 2017年 10月 13日 (2017 - 10 - 13) 全文	1-46																											
A	CN 106204467 A (深圳市未来媒体技术研究院 等) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-46																											
A	CN 107133948 A (电子科技大学) 2017年 9月 5日 (2017 - 09 - 05) 全文	1-46																											
A	CN 104820974 A (浙江科技学院) 2015年 8月 5日 (2015 - 08 - 05) 全文	1-46																											
A	US 2017084007 A1 (WRNCH, INC.) 2017年 3月 23日 (2017 - 03 - 23) 全文	1-46																											
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																													
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																													
国际检索实际完成的日期 2018年 6月 19日		国际检索报告邮寄日期 2018年 7月 6日																											
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 于利娜 电话号码 53961756																											

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	US 9449371 B1 (PIXELWORKS, INC.) 2016年 9月 20日 (2016 - 09 - 20) 全文	1-46

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/106735

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106331433	A	2017年 1月 11日	无	
CN	105791702	A	2016年 7月 20日	无	
CN	107248144	A	2017年 10月 13日	无	
CN	106204467	A	2016年 12月 7日	无	
CN	107133948	A	2017年 9月 5日	无	
CN	104820974	A	2015年 8月 5日	无	
US	2017084007	A1	2017年 3月 23日	WO 2015172235 A1	2015年 11月 19日
US	9449371	B1	2016年 9月 20日	无	