

(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
2023 年 2 月 9 日 (09.02.2023)



(10) 国际公布号
WO 2023/010749 A1

(51) 国际专利分类号:
H04N 21/2343 (2011.01)

南山区深圳大学城学苑大道 1068 号,
Guangdong 518055 (CN)。

(21) 国际申请号: PCT/CN2021/137979

(22) 国际申请日: 2021 年 12 月 14 日 (14.12.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110883118.3 2021年8月2日 (02.08.2021) CN

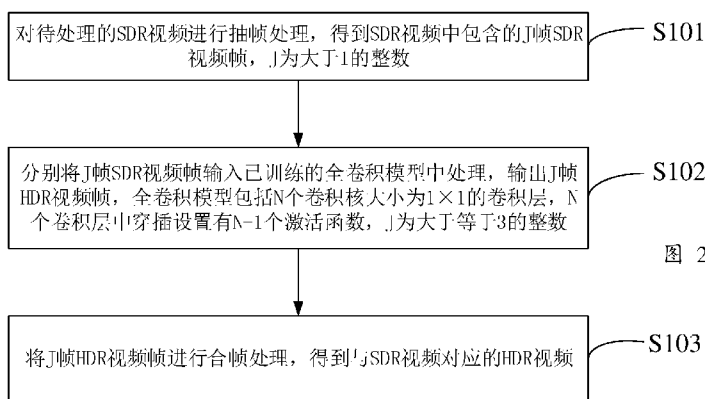
(71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF
ADVANCED TECHNOLOGY CHINESE ACADEMY
OF SCIENCES) [CN/CN]; 中国广东省深圳市

(72) 发明人: 陈翔宇(CHEN, Xiangyu); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。 章政文(ZHANG, Zhengwen); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。 董超(DONG, Chao); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。 乔宇(QIAO, Yu); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。

(74) 代理人: 深圳五邻知识产权代理
事务所(普通合伙) (SHENZHEN WULIN
INTELLECTUAL PROPERTY AGENCY (GENERAL
PARTNERSHIP)); 中国广东省深圳市龙华区

(54) Title: HDR VIDEO CONVERSION METHOD AND APPARATUS, AND DEVICE AND COMPUTER STORAGE MEDIUM

(54) 发明名称: 一种HDR视频转换方法、装置、设备及计算机存储介质



- S101 Perform frame extraction processing on an SDR video to be processed, so as to obtain J SDR video frames included in the SDR video, wherein J is an integer greater than 1
- S102 Input the J SDR video frames into a trained fully convolutional model for processing, and output J HDR video frames, wherein the fully convolutional model comprises N convolutional layers with a convolution kernel size being 1×1 , N-1 activation functions are alternately arranged in the N convolutional layers, and N is an integer greater than or equal to 3
- S103 Perform frame combination processing on the J HDR video frames to obtain an HDR video corresponding to the SDR video

(57) Abstract: The present application relates to the technical field of video processing, and provides an HDR video conversion method and apparatus, and a device and a computer storage medium. The method comprises: performing frame extraction processing on an SDR video to be processed, so as to obtain J SDR video frames included in the SDR video, wherein J is an integer greater than 1; inputting the J SDR video frames into a trained fully convolutional model for processing, and outputting J HDR video frames, wherein the fully convolutional model comprises N convolutional layers with a convolution kernel size being 1×1 , N-1 activation functions are alternately

龙华街道清华社区和平东路幸福城商业大厦A1106, Guangdong 518000 (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, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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))。

arranged in the N convolutional layers, and N is an integer greater than or equal to 3; and performing frame combination processing on the J HDR video frames to obtain an HDR video corresponding to the SDR video. The model used in the present application is simple in structure and has a relatively small amount of parameters, so that the calculation cost of an HDR video conversion task can be effectively reduced, the calculation efficiency is improved, and the speed of video processing is increased.

(57) 摘要: 本申请提供一种HDR视频转换方法、装置、设备及计算机存储介质。涉及视频处理技术领域, 方法包括: 对待处理的SDR视频进行抽帧处理, 得到SDR视频中包含的J帧SDR视频帧, J为大于1的整数; 分别将J帧SDR视频帧输入已训练的全卷积模型中处理, 输出J帧HDR视频帧, 全卷积模型包括N个卷积核大小为 1×1 的卷积层, N个卷积层中穿插设置有N-1个激活函数, N为大于等于3的整数; 将J帧HDR视频帧进行合帧处理, 得到与SDR视频对应的HDR视频。本申请使用的模型结构简单且参数量相对较少, 能够有效降低HDR视频转换任务的计算成本, 提高计算效率, 加快视频处理的速度。

一种 HDR 视频转换方法、装置、设备及计算机存储介质

技术领域

本申请属于视频处理技术领域，尤其涉及一种 HDR 视频转换方法、装置、设备及计算机存储介质。

背景技术

随着高动态范围（High Dynamic Range, HDR）拍摄和视频显示技术的成熟，HDR 视频也逐渐发展起来。相较于传统的标准动态范围（Standard Dynamic Range, SDR）视频，HDR 视频具有更大的动态范围、更宽广的色域，能够展现出更高的对比度和更加丰富的色彩，因此，在许多消费级超高清（Ultra-high definition, UHD）电视上设置了播放 HDR 视频的功能用于 HDR 视频的播放。但是，目前能够拍摄真正符合 HDR 视频标准的拍摄设备尚未普及，因此，如何将拍摄的大量 SDR 视频转换为 HDR 视频，成为目前的热门问题。

目前，一般采用基于神经网络的方法或 HDR 图片转换算法将 SDR 视频转换为 HDR 视频。但已提出的神经网络模型的复杂度较高，计算量较大。HDR 图片转换方法需要先把视频帧从视频编码格式转换为图像编码格式，然后对图像编码格式的数据进行 HDR 转换后重新转回视频编码格式，得到 HDR 视频，处理方法也较为复杂。

发明内容

本申请实施例提供一种 HDR 视频转换方法、装置、终端设备及存储介质，可以解决 HDR 视频转换过程中方法复杂、计算成本高的问题。

第一方面，本申请实施例提供了一种 HDR 视频转换方法，该方法包括：

对待处理的 SDR 视频进行抽帧处理,得到所述 SDR 视频中包含的 J 帧 SDR 视频帧, J 为大于 1 的整数;

分别将 J 帧所述 SDR 视频帧输入已训练的全卷积模型中处理, 输出 J 帧 HDR 视频帧, 所述全卷积模型包括 N 个卷积核大小为 1×1 的卷积层, N 个所述卷积层中穿插设置有 N-1 个激活函数, N 为大于等于 3 的整数;

将所述 J 帧 HDR 视频帧进行合帧处理, 得到与所述 SDR 视频对应的 HDR 视频。

在本申请提供的 HDR 视频转换方法中, 使用本申请提供的全卷积模型实现 HDR 视频转换任务, 由于全卷积模型由卷积核大小为 1×1 的 N 个卷积层及穿插设置的 N-1 个激活函数构成, 模型结构简单且使用的参数量相对较少, 能够有效降低 HDR 视频转换任务的计算成本, 提高计算效率, 加快视频处理的速度。

可选地, $3 \leq N \leq 10$ 。

可选地, 所述激活函数为非线性激活函数。

可选地, 对待处理的 SDR 视频进行抽帧处理, 得到所述 SDR 视频中包含的 J 帧 SDR 视频帧, 包括: 采用抽帧工具对所述 SDR 视频帧进行抽帧处理, 得到所述 J 帧 SDR 视频帧;

所述将所述 J 帧 HDR 视频帧进行合帧处理, 得到与所述 SDR 视频对应的 HDR 视频, 包括:

采用所述抽帧工具对所述 J 帧 HDR 视频帧进行合帧处理, 得到与所述 SDR 视频对应的 HDR 视频。

可选地, 所述全卷积模型的训练方式包括:

利用预设的训练集和预设的损失函数对初始的全卷积模型进行迭代训练, 得到所述全卷积模型;

所述训练集包括多个 SDR 视频帧样本以及与所述 SDR 视频帧样本对应的 HDR 视频帧样本。

所述预设的损失函数用于描述预测的 HDR 视频帧和所述 HDR 视频帧

样本之间的 L2 损失,所述预测的 HDR 视频帧为所述全卷积模型对所述 SDR 视频帧样本进行处理得到的。

第二方面,本申请实施例提供了一种 HDR 视频转换装置,该装置包括:
抽帧单元,用于对待处理的 SDR 视频进行抽帧处理,得到所述 SDR 视频中
5 包含的 J 帧 SDR 视频帧, J 为大于 1 的整数;

处理单元,用于分别将 J 帧所述 SDR 视频帧输入已训练的全卷积模型
中处理,输出 J 帧 HDR 视频帧,所述全卷积模型包括 N 个卷积核大小为 1×1
的卷积层, N 个所述卷积层中穿插设置有 N-1 个激活函数, N 为大于或者
等于 3 的正整数;

10 合帧单元,用于将所述 J 帧 HDR 视频帧进行合帧处理,得到与所述 SDR
视频对应的 HDR 视频。

可选地, $3 \leq N \leq 10$ 。

可选地,所述激活函数为非线性激活函数。

第三方面,本申请实施例提供了一种终端设备,包括存储器、处理器
15 以及存储在存储器中并可在处理器上运行的计算机程序,处理器执行计算
机程序时实现如上述第一方面中任一项的方法。

第四方面,本申请实施例提供了一种计算机可读存储介质,计算机可
读存储介质存储有计算机程序,计算机程序被处理器执行时实现如上述第
一方面中任一项的方法。

20 第五方面,本申请实施例提供了一种计算机程序产品,当计算机程序
产品在终端设备上运行时,使得终端设备执行上述第一方面中任一项的方
法。

可以理解的是,上述第二方面至第五方面的有益效果可以参见上述第
一方面和第一方面的各可能的实施方式所带来的有益效果的相关描述,在
25 此不再赘述。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，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。

5 图 1 是本申请一实施例提供的 HDR 和 SDR 色域表示范围的示意图；

图 2 是本申请一实施例提供的一种 HDR 视频转换方法的一个实施例的流程图；

图 3 是本申请一实施例提供的 HDR 视频转换的全卷积模型的架构图；

图 4 是本申请一实施例提供的一种 HDR 视频转换装置的示意图；

10 图 5 是本申请一实施例提供的一种终端设备的结构示意图。

具体实施方式

为使本申请实施例的目的、技术方案和优点更加清楚，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

相较于标准动态范围（Standard Dynamic Range, SDR），高动态范围（High Dynamic Range, HDR）的动态范围更大且色域范围具有更宽广，由于 HDR 的色域广、动态范围大，因此，HDR 视频能够展现出对比度更高和色彩更加丰富的视频。

如图 1 所示为 HDR 和 SDR 色域表示范围的示意图，其中，BT.709 和 BT.2020 都是 ITU（国际电信联盟）发布的电视参数标准，DCI-P3 是美国电影工业为数字电影院所制定的色域标准，多用来测试投影机所能覆盖的色彩范围。就色域范围来讲，在图 1 所示的 DCI-P3、BT.709 和 BT.2020 中范围最大的是 BT.2020，DCI-P3 的色域范围次之，BT.709 所表示的色域范

围最小。

通常 SDR 视频采用 BT.709 色域，而 HDR 视频采用的是色域更为宽广的 BT.2020 色域，在实际应用 HDR 视频也会采用 DCI-P3 色域。就同一视频而言，无论 HDR 视频采用 BT.2020 色域还是 DCI-P3 色域，HDR 视频展
5 现出对比度和色彩都要优于 SDR 视频。

除此之外，在编码格式上，SDR 视频通常采用的是 8 比特编码，而 HDR 视频采用的是 16 比特编码或者 10 比特编码的编码格式，编码格式所采用的比特位越大，视频呈现出来的对比度越高、色域范围广。

随着 HDR 视频的拍摄及显示技术的逐渐成熟，越来越多的播放设备支持 HDR 视频的播放。因此，如何将 SDR 视频转换为 HDR 视频，成为当前的热门问题。

目前，采用基于神经网络的方法或 HDR 图片转换算法将 SDR 视频转换为 HDR 视频。但已提出的神经网络模型的复杂度较高，计算量较大。例如，常使用生成对抗网络实现 SDR 视频转换 HDR 视频，其中生成对抗网络使用的网络参数达到了 1.06M（其中，M 是数量级“百万”的缩写符号），
15 有的网络参数甚至达到了 2.87M，网络参数越多，设计的网络越复杂，计算量越大。而 HDR 图片转换方法需要先把视频帧从视频编码格式转换为图像编码格式，然后对图像编码格式的数据进行 HDR 转换后重新转回视频编码格式，得到 HDR 视频，处理方法也较为复杂。

针对现有的 HDR 视频转换方法复杂、计算成本高的问题，本申请提供一种全卷积模型，能够实现 HDR 视频转换任务。其中，全卷积模型是由卷积核大小为 1×1 的 N 个卷积层及穿插设置的 N-1 个激活函数构成，模型结构简单且使用的参数量相对较少，能够有效降低 HDR 视频转换任务的计算成本，提高计算效率，加快视频处理的速度。

一般来说，卷积核大小为 1×1 的卷积层一般在实现特定功能的复杂的神经网络模型中用作升/降维功能，即增加或者减少特征图的通道数，从而提升该神经网络模型的计算效率。而在本申请中，经试验发现，通过将卷

积核大小为 1×1 的卷积层与激活函数进行简单的叠加，构建的全卷积模型，能够实现 HDR 视频转换任务，并且取得良好的转换效果。

下面以具体地实施例对本申请的技术方案进行详细说明。下面这几个具体的实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例中不再赘述。

参见图 2，为本申请提供的一种 HDR 视频转换方法的一个实施例的流程图。该方法的执行主体可以是视频处理设备。其中，视频处理设备可以是智能手机、平板电脑、摄像机等移动终端设备，还可以是台式电脑、机器人、服务器等能够处理视频数据的终端设备。如图 2 所示，该方法包括：

10 S101，对待处理的 SDR 视频进行抽帧处理，得到 SDR 视频中包含的 J 帧 SDR 视频帧，J 为大于 3 的整数。

其中，待处理的 SDR 视频可以是拍摄、下载或者是从本地存储区域中读取的完整的视频，也可以是从完成的视频中截取的 SDR 视频片段。

示例性的，可以采用抽帧工具对待处理的 SDR 视频进行抽帧。例如，
15 采用 FFmpeg (Fast Forward mpeg) 工具对待处理的 SDR 视频进行抽帧。

S102，分别将 J 帧 SDR 视频帧输入已训练的全卷积模型中处理，输出 J 帧 HDR 视频帧，全卷积模型包括 N 个卷积核大小为 1×1 的卷积层，N 个卷积层中穿插设置有 N-1 个激活函数，N 为大于等于 3 的整数。

其中，激活函数可以是非线性激活函数，能够增加全卷积模型的非线
20 拟合能力，提高全卷积模型的灵活度。示例性的，该激活函数可以是 ReLU 激活函数。

N 的设置可以根据实际精度要求来设置。

示例性的，设置 $N=3$ ，如图 3 所示，全卷积模型包括 3 层卷积核大小为 1×1 的卷积层及穿插设置在 3 层卷积层中的 2 个 ReLU 激活函数。分别
25 将 K 帧 SDR 视频输入如图 3 所示的全卷积模型进行处理，即可得到对应的 K 帧 HDR 视频帧。

下面结合如下表 1，对图 3 所示全卷积模型的性能进行说明：

表 1

模型	Params	PSNR	SSIM	SR-SIM	ΔE_{ITP}	HDR-VDP3
ResNet	1.37M	37.32	0.9720	0.9950	9.02	8.391
Pixel2Pixel	11.38M	25.80	0.8777	0.9871	44.25	7.136
CycleGAN	11.38M	21.33	0.8496	0.9595	77.74	6.941
HDRNet	482K	35.73	0.9664	0.9957	11.52	8.462
CSRNet	36K	35.04	0.9625	0.9955	14.28	8.400
Ada-3DLUT	594K	36.22	0.9658	0.9967	10.89	8.423
Deep SR-ITM	2.87M	37.10	0.9686	0.9950	9.24	8.233
JSI-GAN	1.06M	37.01	0.9694	0.9928	9.36	8.169
全卷积模型	5K	36.14	0.9643	0.9961	10.43	8.035

表 1 中，残差网络（ResNet）、环形生成对抗网络（CycleGAN）和像素到像素生成网络（Pixel 2 Pixel）是用于图像到图像的转换（image-to-image translation）的算法模型。高动态范围网络（High Dynamic Range Net, HDRNet）、条件序列图像修饰网络（Conditional Sequential Retouching Network, CSRNet）和自适应 3D 查找表（Adaptive 3D lookup table, Ada-3DLUT）网络是用于图像修饰（photo retouching）的算法模型。深度超分辨联合逆色调映射方法（Deep super-resolution inverse tone-mapping, Deep SR-ITM）和超分辨联合逆色调映射生成对抗网络（GAN-Based Joint Super-Resolution and Inverse Tone-Mapping, JSI-GAN）是用于 SDR 视频到 HDR 视频转换的算法模型。

从表 1 可以看出，当 N=3 时，激活函数采用 ReLU 激活函数时，本申请提供的全卷积网络模型的参数量不到 5K（其中，K 是数量级“千”的缩写符号）个，在参数量上远小于表 1 中列举的其他算法模型。因此，具备高效的处理效率。

且全卷积网络模型在峰值信噪比（Peak Signal to Noise Ratio, PSNR）、

结构相似性度量指标 (structural similarity index measure, SSIM)、基于光谱残差的相似度度量指标 (spectral residual based similarity index measure, SR-SIM)、色彩保真度 ΔE_{ITP} 、高动态范围可见差异预测 (High Dynamic Range Visible Difference Predictor, HDR-VDP3) 等性能指标上也具备良好的实验效果。

当然, N 具体可以取值在 3 到 10 之间。例如, 当 N=5 时, 全卷积模型的参数量不到 13K, 在表 1 所列举的算法模型中, 在参数量上通样具备优势, 属于高效的模型。且当 N=5 时, 全卷积模型的 PSNR = 36.15, SSIM = 0.9642, SR_SIM = 0.9963, ΔE_{ITP} = 10.43, HDR-VDP = 38.032。可见, 在其他性能, 也具备良好的实验效果。

值得说明的是, 本实施例中全卷积模型的训练方式包括: 利用预设的训练集和预设的损失函数对初始的全卷积模型进行迭代训练, 得到上述全卷积模型; 其中, 训练集包括多个 SDR 视频帧样本以及与上述 SDR 视频帧样本对应的 HDR 视频帧样本。示例性的, SDR 视频样本及其对应的 HDR 视频样本, 可以从公开的视频网站中获取 SDR 视频样本及对应的 HDR 视频样本。也可以对同一 RAW 数据格式的视频分别进行 SDR 和 HDR 处理, 得到 SDR 视频样本及其对应的 HDR 视频样本。还可以分别利用 SDR 相机和 HDR 相机在同一场景下, 分别拍摄对应的 SDR 视频样本和 HDR 视频样本。

在获取到 SDR 视频样本及其对应的 HDR 视频样本之后, 分别对 SDR 视频样本及其对应的 HDR 视频样本进行抽帧处理, 得到多个 SDR 视频帧样本以及在时序上和空间上与多个 SDR 视频帧样本一一对应的 HDR 视频帧样本。

本实施例中采用 L2 作为全卷积模型训练的预设损失函数。预设的损失函数用于描述预测的 HDR 视频帧和 HDR 视频帧样本之间的损失, 其中预测的 HDR 视频帧为全卷积模型对 SDR 视频帧样本进行处理得到的。

S103, 将 J 帧 HDR 视频帧进行合帧处理, 得到与 SDR 视频对应的 HDR

视频。

值得说明的是，经上述全卷积模型处理后得到的 J 帧 HDR 视频帧是与从待处理的 SDR 视频中提取的 J 帧 SDR 格式的图像数据对应的 HDR 格式的图像数据，由于 HDR 视频为了能够适配更高对比度和色域范围，其采用的是 16 比特编码，因此，经全卷积模型处理后输出的 J 帧 HDR 视频帧是 16 比特编码或者 10 比特编码的图像数据，相较于 8 比特编码的图像数据，16 比特编码或者 10 比特编码的图像数据显示的图像数据更亮。

应该理解的，可以采用抽帧工具对经 HDR 视频转换模型处理后得到的 J 帧 HDR 视频帧进行合帧，例如，采用 FFmpeg 工具对 J 帧 HDR 视频帧进行合帧。

在执行 HDR 视频转任务时，现有的模型往往在利用抽帧工具完成抽帧之后，需要将抽帧得到的 SDR 视频帧转换成 8 比特 YUV 编码帧，才能进行处理。处理后得到的 YUV 序列文件也需要进行格式转换，得到 HDR 视频。

而本申请实施例提供全卷积神经网络，利用抽帧工具抽帧得到 SDR 视频帧可以直接输入全卷积神经网络中处理，得到的 HDR 视频帧也可以直接合帧得到 HDR 视频。与现有的模型，本申请提供的方法无需多次将待处理的 SDR 视频转换成其他格式的数据，相应的，经神经网络模型处理后，也不需要多次将其他格式的数据转换成 HDR 数据，降低了方法的复杂度。

基于同一发明构思，作为对上述方法的实现，本申请实施例提供了一种 HDR 视频转换装置，该装置实施例与前述方法实施例对应，为便于阅读，本装置实施例不再对前述方法实施例中的细节内容进行逐一赘述，但应当明确，本实施例中的装置能够对应实现前述方法实施例中的全部内容。

如图 4 所示，本申请提供一种 HDR 视频转换装置，上述装置 200 包括：抽帧单元 201，用于对待处理的 SDR 视频进行抽帧处理，得到所述 SDR 视频中包含的 J 帧 SDR 视频帧，J 为大于 1 的整数；

处理单元 202, 用于分别将 J 帧 SDR 视频帧输入已训练的全卷积模型中处理, 输出 J 帧 HDR 视频帧, 全卷积模型包括 N 个卷积核大小为 1×1 的卷积层, N 个卷积层中穿插设置有 N-1 个激活函数, N 为大于或者等于 3 的正整数。

5 合帧单元 203, 用于将 J 帧 HDR 视频帧进行合帧处理, 得到与 SDR 视频对应的 HDR 视频。

可选地, $3 \leq N \leq 10$ 。

可选地, 激活函数为非线性激活函数。

可选地, 对待处理的 SDR 视频进行抽帧处理, 得到 SDR 视频中包含的
10 J 帧 SDR 视频帧, 包括: 采用 FFmpeg 工具对 SDR 视频帧进行抽帧处理, 得到 J 帧 SDR 视频帧;

将 J 帧 HDR 视频帧进行合帧处理, 得到与 SDR 视频对应的 HDR 视频, 包括: 采用 FFmpeg 工具对 J 帧 HDR 视频帧进行合帧处理, 得到与 SDR 视频对应的 HDR 视频。

15 可选地, 全卷积模型的训练方式包括:

利用预设的训练集和预设的损失函数对初始的全卷积模型进行迭代训练, 得到全卷积模型;

训练集包括多个 SDR 视频帧样本以及与 SDR 视频帧样本对应的 HDR 视频帧样本。

20 预设的损失函数用于描述预测的 HDR 视频帧和 HDR 视频帧样本之间的 L2 损失, 预测的 HDR 视频帧为全卷积模型对 SDR 视频帧样本进行处理得到的。

所属领域的技术人员可以清楚地了解到, 为了描述的方便和简洁, 仅以上述各功能单元、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单元、模块完成, 即将所述装置的内部
25 结构划分成不同的功能单元或模块, 以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块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，也可以采用软件功能单元的形式实现。另外，各功能单元、模块的具体名称也只是为了便于相互区分，并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程，
5 可以参考前述方法实施例中的对应过程，在此不再赘述。

基于同一发明构思，本申请实施例还提供了一种终端设备。图 5 为本申请实施例提供的终端设备的示意图，如图 5 所示，本实施例提供的终端设备 300 包括：存储器 302 和处理器 301，存储器 302 用于存储计算机程序；处理器 301 用于在调用计算机程序时执行上述方法实施例所述的方法，例
10 如图 2 所示的步骤 S101 至步骤 S103。或者，所述处理器 301 执行所述计算机程序时实现上述各装置实施例中各模块/单元的功能，例如图 4 所示单元 201 至单元 203 的功能。

示例性的，所述计算机程序可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器 302 中，并由所述处理器 301
15 执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描述所述计算机程序在所述终端设备中的执行过程。

本领域技术人员可以理解，图 5 仅仅是终端设备的示例，并不构成对终端设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件，
20 或者不同的部件，例如所述终端设备还可以包括输入输出设备、网络接入设备、总线等。

所述处理器 301 可以是中央处理单元（Central Processing Unit, CPU），还可以是其它通用处理器、数字信号处理器（Digital Signal Processor, DSP）、专用集成电路（Application Specific Integrated Circuit, ASIC）、现场可编程
25 门阵列（Field-Programmable Gate Array, FPGA）或者其它可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

所述存储器 302 可以是所述终端设备的内部存储单元，例如终端设备的硬盘或内存。所述存储器 302 也可以是所述终端设备的外部存储设备，例如所述终端设备上配备的插接式硬盘，智能存储卡（Smart Media Card，SMC），安全数字（Secure Digital，SD）卡，闪存卡（Flash Card）等。进一步地，所述存储器 302 还可以既包括所述终端设备的内部存储单元也包括外部存储设备。所述存储器 302 用于存储所述计算机程序以及所述终端设备所需的其它程序和数据。所述存储器 302 还可以用于暂时地存储已经输出或者将要输出的数据。

本实施例提供的终端设备可以执行上述方法实施例，其实现原理与技术效果类似，此处不再赘述。

本申请实施例还提供一种计算机可读存储介质，其上存储有计算机程序，计算机程序被处理器执行时实现上述方法实施例所述的方法。

本申请实施例还提供一种计算机程序产品，当计算机程序产品在终端设备上运行时，使得终端设备执行时实现上述方法实施例所述的方法。

本申请实施例还提供一种芯片系统，包括处理器，所述处理器与存储器耦合，所述处理器执行存储器中存储的计算机程序，以实现上述方法实施例所述的方法。其中，所述芯片系统可以为单个芯片，或者多个芯片组成的芯片模组。

上述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读存储介质中。基于这样的理解，本申请实现上述实施例方法中的全部或部分流程，可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读存储介质至少可以包括：能够将计算机程序代码携带到拍照装置/终端设备的任何实体或装置、记录介质、计算机存储器、只读存储器

(Read-Only Memory , ROM)、随机存取存储器(Random Access Memory, RAM)、电载波信号、电信信号以及软件分发介质。例如 U 盘、移动硬盘、磁碟或者光盘等。在某些司法管辖区, 根据立法和专利实践, 计算机可读介质不可以是电载波信号和电信信号。

5 在本申请中描述的参考“一个实施例”或“一些实施例”等意味着在本申请的一个或多个实施例中包括结合该实施例描述的特定特征、结构或特点。由此, 在本说明书中的不同之处出现的语句“在一个实施例中”、“在一些实施例中”、“在其他一些实施例中”、“在另外一些实施例中”等不是必然都参考相同的实施例, 而是意味着“一个或多个但不是所有的实施例”, 除非是以其他方式另外特别强调。术语“包括”、“包含”、
10 “具有”及它们的变形都意味着“包括但不限于”, 除非是以其他方式另外特别强调。

 在本申请的描述中, 需要理解的是, 术语“第一”、“第二”仅用于描述目的, 而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此, 限定有“第一”、“第二”的特征可以明示或者隐
15 含地包括至少一个该特征。还应当理解, 在本申请说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合, 并且包括这些组合。

 此外, 在本申请中, 除非另有明确的规定和限定, 术语“连接”、“相
20 连”等应做广义理解, 例如可以是机械连接, 也可以是电连接; 可以是直接连接, 也可以通过中间媒介间接相连, 可以是两个元件内部的连通或两个元件的相互作用关系, 除非另有明确的限定、对于本领域的普通技术人员而言, 可以根据具体情况理解上述术语在本申请中的具体含义。

 以上各实施例仅用以说明本申请的技术方案, 而非对其限制; 尽管参
25 照前述各实施例对本申请进行了详细的说明, 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申请各实施例技术方案的范围。

权 利 要 求 书

1.一种 HDR 视频转换方法，其特征在于，包括：

对待处理的 SDR 视频进行抽帧处理，得到所述 SDR 视频中包含的 J 帧 SDR 视频帧，J 为大于 1 的整数；

5 分别将 J 帧所述 SDR 视频帧输入已训练的全卷积模型中处理，输出 J 帧 HDR 视频帧，所述全卷积模型包括 N 个卷积核大小为 1×1 的卷积层，N 个所述卷积层中穿插设置有 N-1 个激活函数，N 为大于等于 3 的整数；

将所述 J 帧 HDR 视频帧进行合帧处理，得到与所述 SDR 视频对应的 HDR 视频。

10 2.根据权利要求 1 所述的方法，其特征在于， $3 \leq N \leq 10$ 。

3.根据权利要求 1 所述的方法，其特征在于，所述激活函数为非线性激活函数。

4.根据权利要求 1 所述的方法，其特征在于，对待处理的 SDR 视频进行抽帧处理，得到所述 SDR 视频中包含的 J 帧 SDR 视频帧，包括：

15 采用抽帧工具对所述 SDR 视频帧进行抽帧处理，得到所述 J 帧 SDR 视频帧；

所述将所述 J 帧 HDR 视频帧进行合帧处理，得到与所述 SDR 视频对应的 HDR 视频，包括：

采用所述抽帧工具对所述 J 帧 HDR 视频帧进行合帧处理，得到与所述 SDR 视频对应的 HDR 视频。

20 5.根据权利要求 1-3 任一项所述的方法，其特征在于，所述全卷积模型的训练方式包括：

利用预设的训练集和预设的损失函数对初始的全卷积模型进行迭代训练，得到所述全卷积模型；

25 所述训练集包括多个 SDR 视频帧样本以及与所述 SDR 视频帧样本对应的 HDR 视频帧样本；

所述预设的损失函数用于描述预测的 HDR 视频帧和所述 HDR 视频帧样本之间的 L2 损失,所述预测的 HDR 视频帧为所述全卷积模型对所述 SDR 视频帧样本进行处理得到的。

6.一种 HDR 视频转换装置,其特征在于,包括:

5 抽帧单元,用于对待处理的 SDR 视频进行抽帧处理,得到所述 SDR 视频中包含的 J 帧 SDR 视频帧,J 为大于 1 的整数;

处理单元,用于分别将 J 帧所述 SDR 视频帧输入已训练的全卷积模型中处理,输出 J 帧 HDR 视频帧,所述全卷积模型包括 N 个卷积核大小为 1×1 的卷积层,N 个所述卷积层中穿插设置有 N-1 个激活函数,N 为大于等于 3
10 的整数;

合帧单元,用于将所述 J 帧 HDR 视频帧进行合帧处理,得到与所述 SDR 视频对应的 HDR 视频。

7.根据权利要求 6 所述的装置,其特征在于, $3 \leq N \leq 10$ 。

8.根据权利要求 6 所述的装置,其特征在于,所述激活函数为非线性激
15 活函数。

9.一种终端设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求 1 至 5 任一项所述的方法。

10.一种计算机可读存储介质,所述计算机可读存储介质存储有计算机
20 程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求 1 至 5 任一项所述的方法。

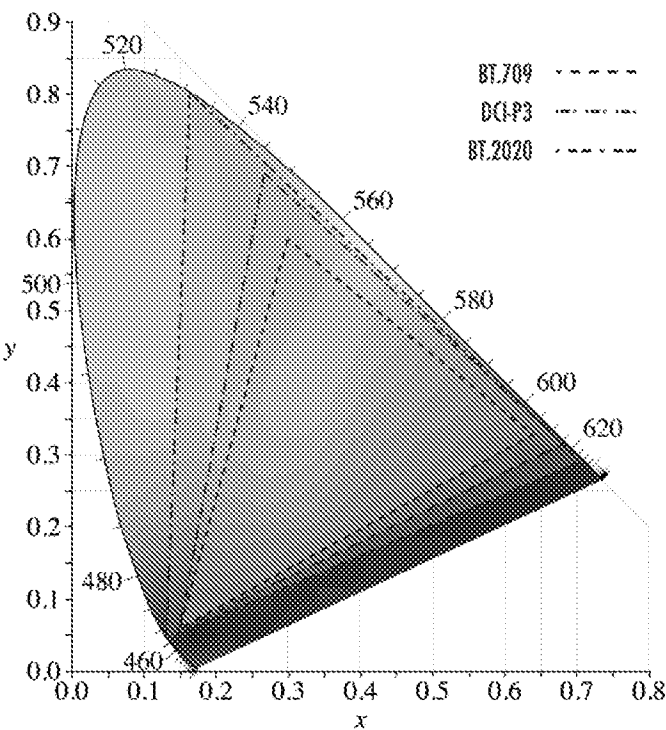


图 1

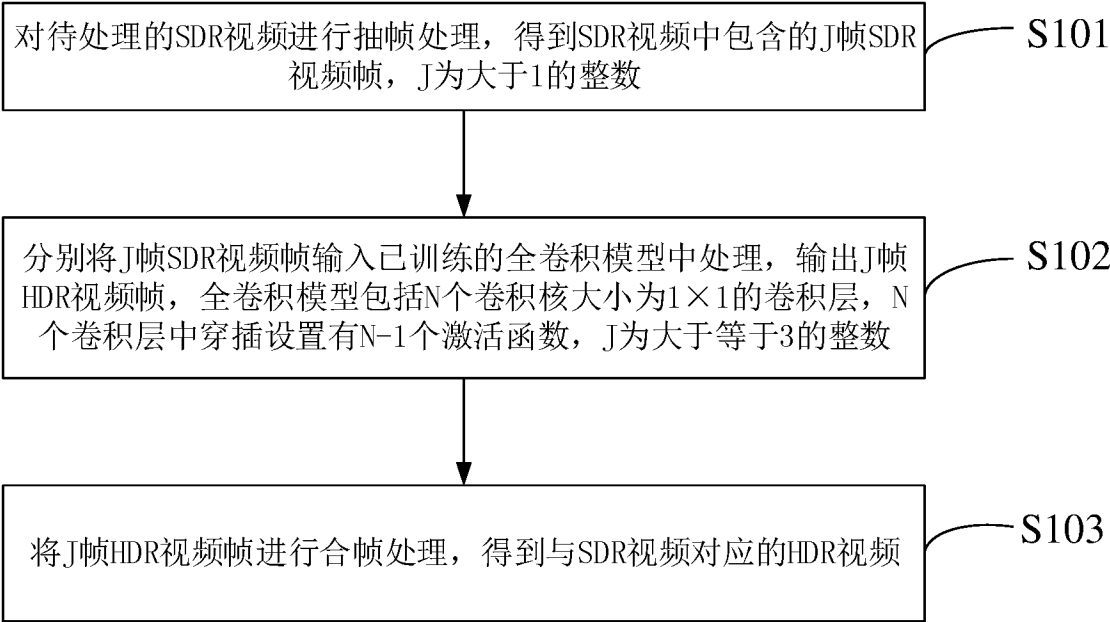


图 2

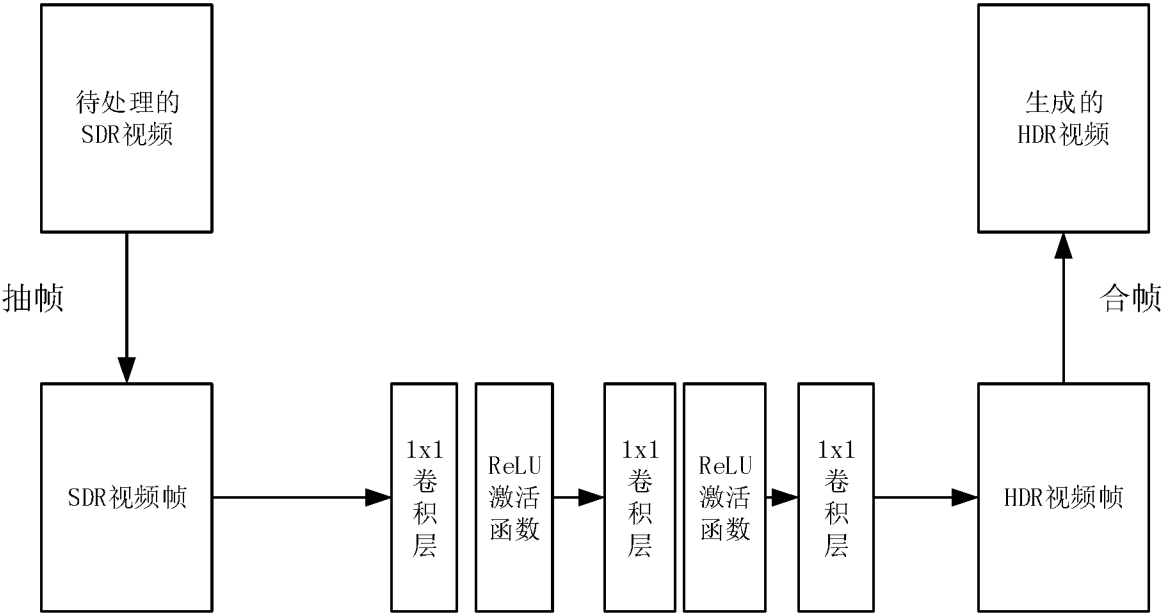


图 3

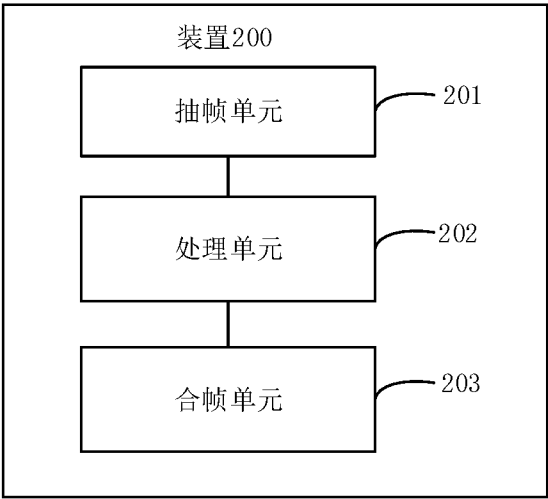


图 4

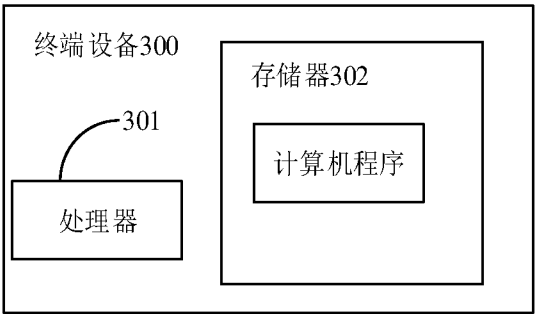


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/137979

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/2343(2011.01)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)

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)

CNABS; CNTXT; VEN; ENTXTC; CNKI: 高动态范围, 低动态范围, 标准动态范围, 1×1, 激活函数, 卷积, 抽帧, 视频, High Dynamic Range, HDR, Low Dynamic Range, LDR, Standard Dynamic Range, SDR, activation function, convolution, frame, extraction, video

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108681991 A (SHANGHAI JIAO TONG UNIVERSITY) 19 October 2018 (2018-10-19) entire document	1-10
A	CN 112200719 A (MIGU VIDEO TECHNOLOGY CO., LTD. et al.) 08 January 2021 (2021-01-08) entire document	1-10
A	CN 112102166 A (SHANGHAI JIAO TONG UNIVERSITY) 18 December 2020 (2020-12-18) entire document	1-10
A	CN 111145097 A (HUAWEI TECHNOLOGIES CO., LTD.) 12 May 2020 (2020-05-12) entire document	1-10
A	US 2019080440 A1 (INTERDIGITAL VC HOLDINGS INC.) 14 March 2019 (2019-03-14) entire document	1-10
PX	CN 113784175 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 10 December 2021 (2021-12-10) description, paragraphs 5-25	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

08 April 2022

Date of mailing of the international search report

15 April 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/137979

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108681991	A	19 October 2018	None			
CN	112200719	A	08 January 2021	None			
CN	112102166	A	18 December 2020	None			
CN	111145097	A	12 May 2020	WO	2021135657	A1	08 July 2021
US	2019080440	A1	14 March 2019	EP	3454294	A1	13 March 2019
CN	113784175	A	10 December 2021	None			

国际检索报告

国际申请号

PCT/CN2021/137979

A. 主题的分类

H04N 21/2343 (2011.01) 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)

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))

CNABS; CNTXT; VEN; ENTXT; CNKI: 高动态范围, 低动态范围, 标准动态范围, 1×1, 激活函数, 卷积, 抽帧, 视频, High Dynamic Range, HDR, Low Dynamic Range, LDR, Standard Dynamic Range, SDR, activation function, convolution, frame, extraction, video

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	CN 108681991 A (上海交通大学) 2018年10月19日 (2018 - 10 - 19) 全文	1-10
A	CN 112200719 A (咪咕视讯科技有限公司等) 2021年1月8日 (2021 - 01 - 08) 全文	1-10
A	CN 112102166 A (上海交通大学) 2020年12月18日 (2020 - 12 - 18) 全文	1-10
A	CN 111145097 A (华为技术有限公司) 2020年5月12日 (2020 - 05 - 12) 全文	1-10
A	US 2019080440 A1 (INTERDIGITAL VC HOLDINGS INC) 2019年3月14日 (2019 - 03 - 14) 全文	1-10
PX	CN 113784175 A (中国科学院深圳先进技术研究院) 2021年12月10日 (2021 - 12 - 10) 说明书第5-25段	1-10

☐ 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

2022年4月8日

国际检索报告邮寄日期

2022年4月15日

ISA/CN的名称和邮寄地址

中国国家知识产权局(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

王倩

电话号码 86-(010)-62412164

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/137979

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	108681991	A	2018年10月19日	无			
CN	112200719	A	2021年1月8日	无			
CN	112102166	A	2020年12月18日	无			
CN	111145097	A	2020年5月12日	WO	2021135657	A1	2021年7月8日
US	2019080440	A1	2019年3月14日	EP	3454294	A1	2019年3月13日
CN	113784175	A	2021年12月10日	无			