

(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 10 月 19 日 (19.10.2023)



(10) 国际公布号
WO 2023/197717 A1

(51) 国际专利分类号:
H04N 19/577 (2014.01)

(21) 国际申请号: PCT/CN2023/071928

(22) 国际申请日: 2023 年 1 月 12 日 (12.01.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210397258.4 2022 年 4 月 15 日 (15.04.2022) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

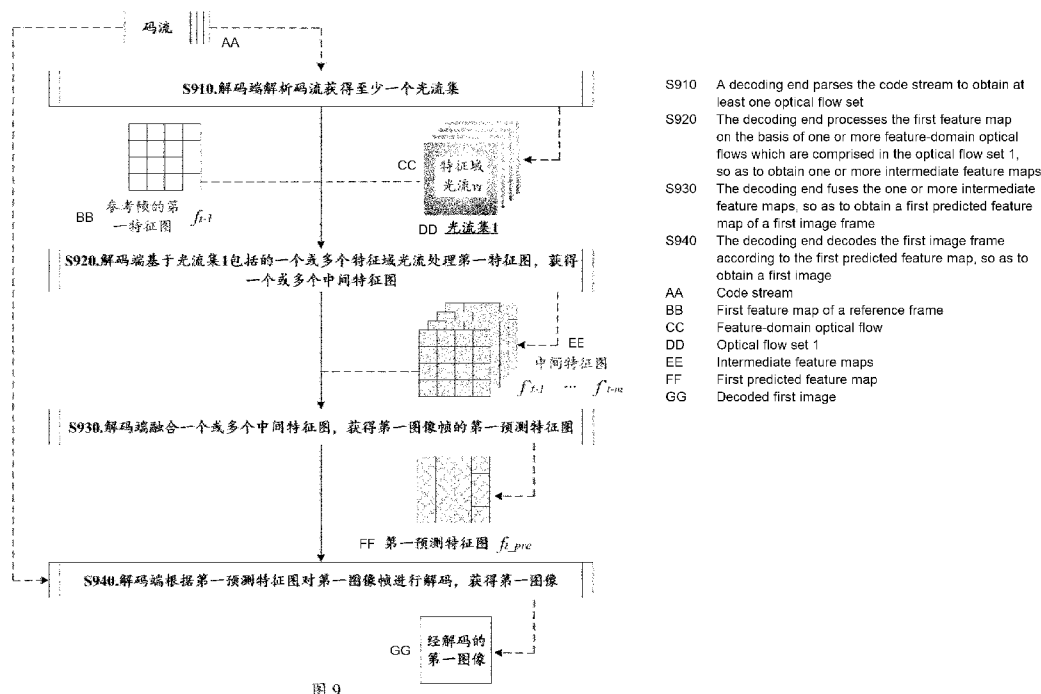
(72) 发明人: 葛运英 (GE, Yunying); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 王晶 (WANG, Jing); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 师一博 (SHI, Yibo); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京中博世达专利商标代理有限公司等 (BEIJING ZBSD PATENT & TRADEMARK AGENT LTD. et al.); 中国北京市海淀区交大东路 31 号 11 号楼 8 层, Beijing 100044 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,

(54) Title: IMAGE DECODING METHOD AND APPARATUS, AND IMAGE CODING METHOD AND APPARATUS

(54) 发明名称: 一种图像解码方法、编码方法及装置



(57) Abstract: An image decoding method and apparatus, and an image coding method and apparatus, which methods and apparatuses relate to the technical field of video processing. The image decoding method comprises: on the basis of a group of feature-domain optical flows (such as an optical flow set) corresponding to an image frame, a decoding end processing a first feature map of a reference frame, so as to obtain a group of intermediate feature maps; the decoding end fusing the group of intermediate feature maps, so as to obtain a predicted feature map; and the decoding end decoding the image frame according to the predicted feature map, so as to obtain a

[见续页]

GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

target image. The pixel error in the feature-domain optical flow is less than a pixel error in an image-domain optical flow, therefore the decoding error caused by the intermediate feature map which is determined according to the feature-domain optical flow is lower than a decoding error caused by the image-domain optical flow in a common technique. The predicted feature map of the image frame is determined by means of the decoding end fusing a plurality of intermediate feature maps, such that the predicted feature map includes more image information; and when the decoding end decodes the image frame on the basis of the predicted feature map, the problem of it being difficult for a single intermediate feature map to accurately express a target image is avoided, thereby improving the image decoding accuracy and the image quality.

(57) 摘要: 一种图像解码方法、编码方法及装置, 涉及视频处理技术领域。解码端基于图像帧对应的一组特征域光流(如光流集), 对参考帧的第一特征图进行处理获得一组中间特征图; 解码端融合该组中间特征图进行获得预测特征图; 解码端根据该预测特征图解码图像帧获得目标图像。该特征域光流中像素误差小于图像域光流中像素误差, 因此, 由特征域光流确定的中间特征图所带来的解码误差低于通常技术中图像域光流所导致的解码误差。图像帧的预测特征图是解码端融合多个中间特征图确定的, 该预测特征图包含了更多的图像信息, 解码端基于预测特征图解码图像帧时, 避免了单个中间特征图难以准确表达目标图像的问题, 提高了图像解码的准确性和图像质量。

一种图像解码方法、编码方法及装置

本申请要求于 2022 年 04 月 15 日提交国家知识产权局、申请号为 202210397258.4、申请名称为“图像解码方法、编码方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及视频处理技术领域，尤其涉及一种图像解码方法、编码方法及装置。

背景技术

在视频编解码技术中，视频压缩和视频增强技术尤为重要。视频压缩系统执行空间（图像内）预测和/或时间（图像间）预测，以减少或移除视频序列中固有的冗余信息；视频增强技术用于提高图像的显示质量。针对于视频压缩或者视频增强对应的视频解码过程，解码端采用光流映射（warping）的方法对视频包含的图像帧进行解码。光流映射是指解码端获取图像帧和参考帧之间的图像域光流，并根据该光流对图像帧进行解码。图像域光流用于指示相邻两帧图像中的相应像素之间的运动速度和运动方向。然而，光流映射对光流精度较为敏感，光流精度的细微变化会影响 warping 的准确性。由于相邻两帧图像之间的光流预测的误差较大，如 5 个像素或更多像素的误差，因此，解码端基于图像域光流对图像帧进行解码的准确性较低，导致解码获得的图像的清晰度受到影响。因此，如何提供一种更有效的图像解码方法成为目前亟需解决的问题。

发明内容

本申请提供一种图像解码方法、编码方法及装置，解决了基于图像域光流对图像帧进行解码的准确性较低，导致解码获得的图像的清晰度受到影响的问题。

本申请采用如下技术方案。

第一方面，本申请提供了一种图像解码方法，该方法应用于视频编解码系统，或支持实现该视频编解码系统实现图像解码方法的物理设备，例如该物理设备为解码端或者视频解码器（decoder），在一些情况中，该物理设备可包括芯片系统。这里以解码端执行图像解码方法为例进行说明，该图像解码方法包括：首先，解码端解析码流获得至少一个光流集；该至少一个光流集包括第一光流集，该第一光流集对应前述第一图像帧的参考帧的第一特征图，且该第一光流集包括一个或多个特征域光流，该一个或多个特征域光流中的任一个特征域光流用于指示第一图像帧的特征图与前述第一特征图之间的运动信息。其次，解码端基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图；以及，解码端融合前述的一个或多个中间特征图，获得第一图像帧的第一预测特征图。最后，解码端根据该第一预测特征图解码第一图像帧获得第一图像。

在本实施例中，对于单个特征域光流而言，该特征域光流中像素的误差小于图像域光流中像素的误差，因此，解码端根据特征域光流确定的中间特征图所带来的解码误差低于通常技术中图像域光流所导致的解码误差，换言之，解码端根据特征域光流对图像帧进行解码，降低了相邻两帧图像之间的图像域光流导致的解码误差，提高了图像解码的准确性。而且，解码端将光流集和参考帧的特征图确定的多个中间特征图进行融合，获得图像帧的预测特征图，该预测特征图包含了比单个图像域光流更多的图像信息，使得解码端基于融合获得的预测特征图对图像帧进行解码时，避免了单个图像域光流指示的信息难以准确表

达第一图像的问题，提高了图像解码的准确性和图像质量（如图像清晰度等）。

在一种可选的实现方式中，解码端基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图，包括：解码端解析码流获得前述参考帧的第一特征图；以及，针对于第一光流集包括的一个或多个特征域光流，解码端基于其中的第一特征域光流对第一特征图进行光流映射（warping），获得第一特征域光流对应的中间特征图；该第一特征域光流为前述一个或多个特征域光流中的任一个。

在本实施例中，解码端基于一组特征域光流对第一特征图进行 warping，继而，在解码端获得参考帧对应的图像后，解码端可在参考帧对应的图像中，根据一组特征域光流对该组特征域光流在参考帧对应的图像中相应位置进行插值，获得第一图像中像素值的预测值，从而获得第一图像，避免了解码端需根据图像域光流对第一图像所有像素值进行预测，减少了解码端进行图像解码所需的计算量，提高了图像解码的效率。

在另一种可选的实现方式中，前述的第一光流集还对应参考帧的第二特征图。本实施例提供的图像解码方法中，在解码端根据第一预测特征图解码第一图像帧获得第一图像之前，该图像解码方法还包括：第一步，解码端基于第一光流集包括的一个或多个特征域光流处理第二特征图，获得第二特征图对应的一个或多个中间特征图。第二步，解码端融合第二特征图对应的一个或多个中间特征图，获得第一图像帧的第二预测特征图。如此，解码端可根据第一预测特征图和第二预测特征图，解码第一图像帧获得第一图像。

在本实施例中，参考帧的多个特征图（或称一组特征图）可对应一个光流集（或称一组特征域光流），换言之，针对于属于同一组的多个特征图共用一组特征域流，解码端根据特征图对应的一组特征域光流处理该特征图，从而获得该特征图对应的中间特征图，进一步的，解码端将该特征图对应的中间特征图进行融合获得该特征图对应的预测特征图。最后，解码端根据参考帧的所有特征图对应的预测特征图对图像帧进行解码获得目标图像。如此，在图像解码过程中，若参考帧对应多个特征图，解码端可将该多个特征图划分为一个组或多个组，属于同一个组的特征图共用一个光流集，并将特征图对应的中间特征图进行融合获得预测特征图，避免了当参考帧或图像帧具有较多信息时，解码端根据特征图重构图像的精度较低、速度较慢的问题，提高了图像解码的准确性。值得注意的是，属于不同组的特征图的通道数可以不相同。

在另一种可选的实现方式中，解码端融合一个或多个中间特征图，获得第一图像帧的第一预测特征图，包括：解码端获取前述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值；以及，解码端基于一个或多个中间特征图权值处理与该一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得第一预测特征图。其中的权值用于指示中间特征图在第一预测特征图所占的权重。

在本实施例中，针对于第一特征图对应的多个中间特征图，每个中间特征图的权值可以不同，换言之，解码端可根据图像解码的需求为中间特征图设置不同的权值，示例性的，若在一些中间特征图对应的图像较为模糊的情况下，降低这些中间特征图的权值，从而提高第一图像的清晰度。

在另一种可选的实现方式中，解码端融合一个或多个中间特征图，获得第一图像帧的第一预测特征图，包括：将第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图。特征融合模型包括卷积网络层，卷积网络层用于融合中间特征图。

在通常技术中，解码端获取图像帧和参考帧之间的多个图像域光流，并根据多个图像域光流与参考帧获取到多张图像，从而将多张图像进行融合获得图像帧对应的目标图像，因此，解码端解码一帧图像需预测多张图像的像素值，并将多张图像进行融合获得目标图像，导致图像解码所需的计算资源较大，解码端根据图像域光流解码视频的效率较低。

在本实施例中，如在图像解码过程中，由于特征图融合所需的计算资源小于特征图解码所需的计算资源，因此，解码端采用特征融合模型对多个中间特征图进行融合，进而，解码端基于融合多个中间特征图获得的预测特征图对图像帧进行解码，即解码端仅需根据预测特征图对该预测特征图在图像中指示的图像位置的像素值进行预测，无需预测多张图像的所有像素值，减少了图像解码所需的计算资源，提高了图像解码的效率。

在另一种可选的实现方式中，本申请提供的图像解码方法还包括：首先，解码端获取第一图像的特征图。其次，解码端根据第一图像的特征图、第一特征图和第一预测特征图，获得增强特征图。最后，解码端根据增强特征图处理第一图像，获得第二图像。第二图像的清晰度高于第一图像的清晰度。

本实施例中，解码端可将第一图像的特征图、第一特征图和第一预测特征图进行融合，并基于融合获得的增强特征图对第一图像进行视频增强处理，获得清晰度更优的第二图像，从而提高了解码后的图像清晰度及图像显示效果。

在另一种可选的实现方式中，解码端根据增强特征图处理第一图像，获得第二图像，包括：解码端依据增强特征图获得第一图像的增强层图像，并基于增强层图像重构第一图像，获得第二图像。示例性的，该增强层图像可以是指解码端依据参考帧和该增强特征图确定的图像，解码端在第一图像的基础上加入增强层图像的部分或全部信息，获得第二图像；或者，解码端将增强层图像作为第一图像的重构图像，即前述的第二图像。在本示例中，解码端获取到图像在不同阶段的多个特征图，并获得了这多个特征图确定的增强特征图，从而依据该增强特征图对第一图像进行重构和视频增强，提高了解码后的图像清晰度和图像显示效果。

第二方面，提供了一种图像编码方法，该方法应用于视频编解码系统，或支持实现该视频编解码系统实现图像编码方法的物理设备，例如该物理设备为编码端或者视频编码器（encoder），在一些情况中，该物理设备可包括芯片系统。这里以编码端执行图像编码方法为例进行说明，该图像编码方法包括：第一步，编码端获取第一图像帧的特征图，以及第一图像帧的参考帧的第一特征图。第二步，编码端根据第一图像帧的特征图和第一特征图，获得至少一个光流集；该至少一个光流集包括第一光流集，该第一光流集对应前述第一特征图，且该第一光流集包括一个或多个特征域光流，前述一个或多个特征域光流中的任一个特征域光流用于指示第一图像帧的特征图与第一特征图之间的运动信息。第三步，编码端基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图。第四步，编码端融合前述一个或多个中间特征图，获得第一图像帧的第一预测特征图。第五步，编码端根据第一预测特征图对第一图像帧进行编码获得码流。

可选的，该码流包括：至少一个光流集对应的特征域光流码流，以及第一预测特征图对应的图像区域的残差码流。

在本实施例中，对于单个特征域光流而言，该特征域光流中像素的误差小于图像域光

流中像素的误差，因此，编码端根据特征域光流确定的中间特征图所带来的编码误差低于通常技术中图像域光流所导致的编码误差，换言之，编码端根据特征域光流对图像帧进行编码，降低了相邻两帧图像之间的图像域光流导致的编码误差，提高了图像编码的准确性。而且，编码端基于多个特征域光流处理参考帧的特征图获得多个中间特征图，以及融合该多个中间特征图确定图像帧的预测特征图，即编码端将光流集和参考帧的特征图确定的多个中间特征图进行融合，获得图像帧的预测特征图，该预测特征图包含了更多的图像信息，使得编码端基于融合获得的预测特征图对图像帧进行编码时，避免了单个中间特征图难以准确表达第一图像的问题，提高了图像编码的准确性和图像质量（如图像清晰度等）。

在一种可选的实现方式中，编码端基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图，包括：针对于第一光流集包括的一个或多个特征域光流，基于其中的第一特征域光流对第一特征图进行光流映射，获得第一特征域光流对应的中间特征图。该第一特征图为第一光流集包括的一个或多个特征域光流中任一个。

在本实施例中，如在图像编码过程中，由于特征图融合所需的计算资源小于特征图编码所需的计算资源，因此，编码端采用特征融合模型对多个中间特征图进行融合，进而，编码端基于融合多个中间特征图获得的预测特征图对图像帧进行编码，即编码端仅需根据预测特征图对该预测特征图在图像中指示的图像位置的像素值进行预测，无需预测多张图像的所有像素值，减少了图像编码所需的计算资源，提高了图像编码的效率。

在另一种可选的实现方式中，第一光流集还对应参考帧的第二特征图。在编码端根据第一预测特征图对第一图像帧进行编码获得码流之前，本实施例提供的图像编码方法还包括：第一步，编码端基于第一光流集包括的一个或多个特征域光流处理第二特征图，获得第二特征图对应的一个或多个中间特征图。第二步，编码端融合第二特征图对应的一个或多个中间特征图，获得第一图像帧的第二预测特征图。如此，前述的编码端根据第一预测特征图对第一图像帧进行编码获得码流，可包括：编码端根据第一预测特征图和第二预测特征图，编码第一图像帧获得码流。在本实施例中，参考帧的多个特征图（或称一组特征图）可对应一个光流集（或称一组特征域光流），换言之，针对于属于同一组的多个特征图共用一组特征域流，编码端根据特征图对应的一组特征域光流处理该特征图，从而获得该特征图对应的中间特征图，进一步的，编码端将该特征图对应的中间特征图进行融合获得该特征图对应的预测特征图。最后，编码端根据参考帧的所有特征图对应的预测特征图对图像帧进行编码获得目标码流。如此，在图像编码过程中，若参考帧对应多个特征图，编码端可将该多个特征图划分为一个组或多个组，属于同一个组的特征图共用一个光流集，并将特征图对应的中间特征图进行融合获得预测特征图，避免了当参考帧或图像帧具有较多信息时，编码端根据特征图编码获得码流的冗余较多、精度较低的问题，提高了图像编码的准确性。值得注意的是，属于不同组的特征图的通道数可以不相同。

在另一种可选的实现方式中，编码端融合一个或多个中间特征图，获得第一图像帧的第一预测特征图，包括：编码端获取前述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值；以及，编码端基于前述一个或多个权值处理与该一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得第一预测特征图。该权值用于指示中间特征图在第一预测特征图中所占的权重。在本实施例中，针对于第一

特征图对应的多个中间特征图，每个中间特征图的权值可以不同，换言之，编码端可根据图像编码的需求为中间特征图设置不同的权值，示例性的，若在一些中间特征图对应的图像较为模糊的情况下，降低这些中间特征图的权值，从而提高第一图像的清晰度。

在另一种可选的实现方式中，编码端融合一个或多个中间特征图，获得第一图像帧的第一预测特征图，包括：编码端将第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图。特征融合模型包括卷积网络层，卷积网络层用于融合中间特征图。在图像编码过程中，由于特征图融合所需的计算资源小于特征图编码所需的计算资源，因此，编码端采用特征融合模型对多个中间特征图进行融合，进而，编码端基于融合多个中间特征图获得的预测特征图对图像帧进行编码，减少了图像编码所需的计算资源，提高了图像编码的效率。

第三方面，提供了一种图像解码装置，该图像解码装置可应用于解码端，或者支持实现前述图像解码方法的视频编解码系统。该图像解码装置包括用于执行第一方面或第一方面任一种可能实现方式中的图像解码方法的各个模块。示例的，图像解码装置包括：码流单元、处理单元、融合单元和解码单元。

码流单元，用于解析码流获得至少一个光流集。该至少一个光流集包括第一光流集，该第一光流集对应第一图像帧的参考帧的第一特征图，该第一光流集包括一个或多个特征域光流，前述一个或多个特征域光流中任一个特征域光流用于指示第一图像帧的特征图与第一特征图之间的运动信息。

处理单元，用于基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图。

融合单元，用于融合一个或多个中间特征图，获得第一图像帧的第一预测特征图。

解码单元，用于根据第一预测特征图解码第一图像帧获得第一图像。

有益效果可以参见第一方面中任一方面的描述，此处不再赘述。所述图像解码装置具有实现上述第一方面中任一方面的方法实例中行为的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

在一种可选的实现方式中，处理单元，具体用于：解析码流获得参考帧的第一特征图，以及，基于第一特征域光流对第一特征图进行光流映射（warping），获得第一特征域光流对应的中间特征图；该第一特征域光流为前述第一光流集包括的一个或多个特征域光流中的任一个。

在另一种可选的实现方式中，第一光流集还对应参考帧的第二特征图。处理单元，还用于基于第一光流集包括的一个或多个特征域光流处理第二特征图，获得第二特征图对应的一个或多个中间特征图。融合单元，还用于融合第二特征图对应的一个或多个中间特征图，获得第一图像帧的第二预测特征图。解码单元，具体用于根据第一预测特征图和第二预测特征图，解码第一图像帧获得第一图像。

在另一种可选的实现方式中，融合单元，具体用于：获取前述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，该权值用于指示中间特征图在第一预测特征图所占的权重。以及，基于一个或多个权值处理与前述一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得第一预测特征图。

在另一种可选的实现方式中，融合单元，具体用于：将第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图。特征融合模型包括卷积网络层，卷积网络层用于融合中间特征图。

在另一种可选的实现方式中，图像解码装置还包括：获取单元和增强单元。获取单元，用于获取第一图像的特征图。融合单元，还用于根据第一图像的特征图、第一特征图和第一预测特征图，获得增强特征图。增强单元，用于根据增强特征图处理第一图像，获得第二图像。第二图像的清晰度高于第一图像的清晰度。

在另一种可选的实现方式中，增强单元，具体用于：依据增强特征图获得第一图像的增强层图像。以及，基于增强层图像重构第一图像，获得第二图像。

第四方面，提供了一种图像编码装置，该图像编码装置可应用于编码端，或者支持实现前述图像编码方法的视频编解码系统。该图像编码装置包括用于执行第二方面或第二方面任一种可能实现方式中的图像编码方法的各个模块。示例的，图像编码装置包括：获取单元、处理单元、融合单元和编码单元。

获取单元，用于获取第一图像帧的特征图，以及第一图像帧的参考帧的第一特征图。

处理单元，用于根据第一图像帧的特征图和第一特征图，获得至少一个光流集。该至少一个光流集包括第一光流集，该第一光流集对应第一图像帧的参考帧的第一特征图，该第一光流集包括一个或多个特征域光流，前述一个或多个特征域光流中任一特征域光流用于指示第一图像帧的特征图与第一特征图之间的运动信息。

处理单元，还用于基于第一光流集包括的一个或多个特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图。

融合单元，用于融合一个或多个中间特征图，获得第一图像帧的第一预测特征图。

编码单元，用于根据第一预测特征图对第一图像帧进行编码获得码流。

可选的，该码流包括：至少一个光流集对应的特征域光流码流，以及第一预测特征图对应的图像区域的残差码流。

有益效果可以参见第二方面中任一方面的描述，此处不再赘述。所述图像编码装置具有实现上述第一方面中任一方面的方法实例中行为的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

在一种可选的实现方式中，处理单元，具体用于：以及，针对于第一光流集包括的一个或多个特征域光流，基于其中的第一特征域光流对第一特征图进行光流映射(warping)，获得第一特征域光流对应的中间特征图；该第一特征域光流为第一光流集包括的一个或多个特征域光流中任一个。

在另一种可选的实现方式中，第一光流集还对应参考帧的第二特征图。处理单元，还用于基于第一光流集包括的一个或多个特征域光流处理第二特征图，获得第二特征图对应的一个或多个中间特征图。融合单元，还用于融合第二特征图对应的一个或多个中间特征图，获得第一图像帧的第二预测特征图。编码单元，具体用于根据第一预测特征图和第二预测特征图，编码第一图像帧获得码流。

在另一种可选的实现方式中，融合单元，具体用于：获得前述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，该权值用于指示中间特征图在

第一预测特征图所占的权重。以及，基于前述一个或多个权值处理与该一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得第一预测特征图。

在另一种可选的实现方式中，融合单元，具体用于：将第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图。特征融合模型包括卷积网络层，卷积网络层用于融合中间特征图。

第五方面，提供了一种图像译码的装置，该图像译码的装置包括至少一个处理器和存储器，存储器用于存储程序代码，当处理器调用所述程序代码时，执行第一方面或第一方面任一种可能实现方式中的图像解码方法的操作步骤，示例性的，该图像译码的装置可以是视频编解码系统包括的解码端或视频解码器。

或者，当处理器调用存储器中的程序代码时，执行第二方面或第二方面任一种可能实现方式中的图像编码方法的操作步骤。示例性的，该图像译码的装置可以是视频编解码系统包括的编码端或视频编码器。

第六方面，提供了一种计算机可读存储介质，存储介质中存储有计算机程序或指令，当所述计算机程序或指令被电子设备执行时，执行第一方面或第一方面任一种可能实现方式中的图像解码方法的操作步骤，和/或，执行第二方面或第二方面任一种可能实现方式中的图像编码方法的操作步骤。示例性的，该电子设备是指前述图像译码的装置。

第七方面，提供了另一种计算机可读存储介质，该计算机读存储介质中存储有根据第二方面或第二方面中任一种可能实现方式中的图像编码方法获取的码流。示例性的，该码流可以包括：前述第二方面中至少一个光流集对应的特征域光流码流，以及第一预测特征图对应的图像区域的残差码流。

第八方面，提供了一种视频编解码系统，包括编码端和解码端，该解码端可执行第一方面或第一方面任一种可能实现方式中的图像解码方法的操作步骤，编码端可执行第二方面或第二方面任一种可能实现方式中的图像编码方法的操作步骤。有益效果可以参见第一方面中任一方面的描述，或者，第二方面中任一方面的描述，此处不再赘述

第九方面，提供了一种计算机程序产品，当计算机程序产品在电子设备上运行时，使得电子设备执行如第一方面或第一方面任意一种可能的实现方式中所述方法的操作步骤，和/或，当计算机程序产品在计算机上运行时，使得电子设备执行如第一方面或第一方面任意一种可能的实现方式中所述方法的操作步骤。示例性的，该电子设备是指前述图像译码的装置。

第十方面，提供了一种芯片，包括控制电路和接口电路，接口电路用于接收来自芯片之外的其它设备的信号并传输至处理器，或将来自控制电路的信号发送给芯片之外的其它设备，控制电路通过逻辑电路或执行代码指令用于实现：上述第一方面或第一方面中任一种可能实现方式中方法的操作步骤，和/或，上述第二方面或第二方面中任一种可能实现方式中方法的操作步骤。示例性的，该芯片是指前述图像译码的装置。

本申请在上述各方面提供的实现方式的基础上，还可以进行进一步组合以提供更多实现方式。

附图说明

图 1 为本申请提供的视频编解码系统的示例性框图；

图 2 为本申请提供的光流和颜色的映射关系图；

图 3 为本申请提供的光流映射的示意图；
图 4 为本申请提供的视频压缩系统的架构示意图；
图 5 为本申请提供的图像编码方法的流程示意图；
图 6 为本申请提供的特征提取网络的结构示意图；
图 7 为本申请提供的光流估计网络的结构示意图；
图 8 为本申请提供的光流映射和特征融合的流程示意图；
图 9 为本申请提供的图像解码方法的流程示意图一；
图 10 为本申请提供的特征编码网络和特征解码网络的结构示意图；
图 11 为本申请提供的图像解码方法的流程示意图二；
图 12 为本申请提供的图像重构网络的结构示意图；
图 13 为本申请提供的可变性卷积网络的示意图；
图 14 为本申请提供的效率比对的示意图；
图 15 为本申请提供的图像编码装置的结构示意图；
图 16 为本申请提供的图像解码装置的结构示意图；
图 17 为本申请提供的图像译码的装置的结构示意图。

具体实施方式

本申请的实施例提供了一种图像解码（编码）方法，包括：解码端（或者编码端）获取到图像帧的一组特征域光流（或称一个光流集）与参考帧的特征图之间的多个中间特征图，并将该多个中间特征图进行融合获得图像帧的预测特征图，从而，解码端（或者编码端）根据该预测特征图对图像帧进行解码（或者编码），获得图像帧对应的目标图像（或者码流）。以图像解码过程为例，对于单个特征域光流而言，该特征域光流中像素的误差小于图像域光流中像素的误差，因此，解码端根据特征域光流确定的中间特征图所带来的解码误差低于通常技术中图像域光流所导致的解码误差，换言之，解码端根据特征域光流对图像帧进行解码，降低了相邻两帧图像之间的图像域光流导致的解码误差，提高了图像解码的准确性。而且，解码端基于多个特征域光流处理参考帧的特征图获得多个中间特征图，以及融合该多个中间特征图确定图像帧的预测特征图，即解码端将光流集和参考帧的特征图确定的多个中间特征图进行融合，获得图像帧的预测特征图，该预测特征图包含了更多的图像信息，使得解码端基于融合获得的预测特征图对图像帧进行解码时，避免了单个中间特征图难以准确表达第一图像的问题，提高了图像解码的准确性和图像质量（如图像清晰度等）。下面结合实施例对本申请提供的方案进行说明，为了下述各实施例的描述清楚简洁，首先给出相关技术的简要介绍。

图 1 为本申请提供的视频编解码系统的示例性框图，如本文所使用，术语“视频译码器”一般是指视频编码器和视频解码器两者。在本申请中，术语“视频译码”或“译码”可一般地指代视频编码或视频解码，示例性的，编码端或解码端可统称为图像译码的装置。

如图 1 中所示，视频编解码系统包含编码端 10 和解码端 20。编码端 10 产生经编码视频数据。因此，编码端 10 可被称为视频编码装置。解码端 20 可对由编码端 10 所产生的经编码的视频数据（如包含一个或多个图像帧的视频）进行解码。因此，解码端 20 可被称为视频解码装置。编码端 10、解码端 20 或两个的各种实施方案可包含一或多个处理器以及耦合到所述一个或多个处理器的存储器。所述存储器可包含但不限于随机存取存储器

(random access memory, RAM)、闪存、只读存储器 (read-only memory, ROM)、可编程只读存储器 (programmable ROM, PROM)、可擦除可编程只读存储器 (erasable PROM, EPROM)、电可擦除可编程只读存储器 (electrically EPROM, EEPROM)、快闪存储器或可用于以可由计算机存取的指令或数据结构的形式存储所要的程序代码的任何其它媒体。

编码端 10 和解码端 20 可以包括各种装置, 包含桌上型计算机、移动计算装置、笔记型 (例如, 膝上型) 计算机、平板计算机、机顶盒、例如所谓的“智能”电话等电话手持机、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、车载计算机或其类似者。

解码端 20 可经由链路 30 从编码端 10 接收经编码视频数据。链路 30 可包括能够将经编码视频数据从编码端 10 移动到解码端 20 的一或多个媒体或装置。在一个实例中, 链路 30 可包括使得编码端 10 能够实时将经编码视频数据直接发射到解码端 20 的一或多个通信媒体。在此实例中, 编码端 10 可根据通信标准 (例如无线通信协议) 来调制经编码视频数据, 且可将经调制的视频数据发射到解码端 20。所述一或多个通信媒体可包含无线和/或有线通信媒体, 例如射频 (radio frequency, RF) 频谱或一或多个物理传输线。所述一或多个通信媒体可形成基于分组的网络的一部分, 基于分组的网络例如为局域网、广域网或全球网络 (例如, 因特网)。所述一或多个通信媒体可包含路由器、交换器、基站或促进从编码端 10 到解码端 20 的通信的其它设备。

在另一实例中, 可将经编码数据从输出接口 140 输出到存储装置 40。类似地, 可通过输入接口 240 从存储装置 40 存取经编码数据。存储装置 40 可包含多种分布式或本地存取的数据存储媒体中的任一者, 例如硬盘驱动器、蓝光光盘、数字通用光盘 (digital video disc, DVD)、只读光盘 (compact disc read-only memory, CD-ROM)、快闪存储器、易失性或非易失性存储器, 或用于存储经编码视频数据的任何其它合适的数字存储媒体。

在另一实例中, 存储装置 40 可对应于文件服务器或可保持由编码端 10 产生的经编码视频的另一中间存储装置。解码端 20 可经由流式传输或下载从存储装置 40 存取所存储的视频数据。文件服务器可为任何类型的能够存储经编码的视频数据并且将经编码的视频数据发射到解码端 20 的服务器。实例文件服务器包含网络服务器 (例如, 用于网站)、文件传输协议 (file transfer protocol, FTP) 服务器、网络附接式存储 (network attached storage, NAS) 装置或本地磁盘驱动器。解码端 20 可通过任何标准数据连接 (包含因特网连接) 来存取经编码视频数据。这可包含无线信道 (例如, 无线保真 (wireless-fidelity, Wi-Fi) 连接)、有线连接 (例如, 数字用户线路 (digital subscriber line, DSL)、电缆调制解调器等), 或适合于存取存储在文件服务器上的经编码视频数据的两者的组合。经编码视频数据从存储装置 40 的传输可为流式传输、下载传输或两者的组合。

本申请提供的图像的解码方法可应用于视频编解码以支持多种多媒体应用, 例如空中电视广播、有线电视发射、卫星电视发射、串流视频发射 (例如, 经由因特网)、用于存储于数据存储媒体上的视频数据的编码、存储在数据存储媒体上的视频数据的解码, 或其它应用。在一些实例中, 视频编解码系统可用于支持单向或双向视频传输以支持例如视频流式传输、视频增强、视频回放、视频广播和/或视频电话等应用。

图 1 中所说明的视频编解码系统仅为实例, 并且本申请的技术可适用于未必包含编码装置与解码装置之间的任何数据通信的视频译码设置 (例如, 视频编码或视频解码)。在其它实例中, 数据从本地存储器检索、在网络上流式传输等等。视频编码装置可对数据进行

编码并且将数据存储到存储器，和/或视频解码装置可从存储器检索数据并且对数据进行解码。在许多实例中，由并不彼此通信而是仅编码数据到存储器和/或从存储器检索数据且解码数据的装置执行编码和解码。

在图 1 的实例中，编码端 10 包含视频源 120、视频编码器 100 和输出接口 140。在一些实例中，输出接口 140 可包含调节器/解调器（调制解调器）和/或发射器。视频源 120 可包括视频捕获装置（例如，摄像机）、含有先前捕获的视频数据的视频存档、用以从视频内容提供者接收视频数据的视频馈入接口，和/或用于产生视频数据的计算机图形系统，或视频数据的此些来源的组合。

视频编码器 100 可对来自视频源 120 的视频数据进行编码。在一些实例中，编码端 10 经由输出接口 140 将经编码视频数据直接发射到解码端 20。在其它实例中，经编码视频数据还可存储到存储装置 40 上，供解码端 20 以后存取来用于解码和/或播放。

在图 1 的实例中，解码端 20 包含输入接口 240、视频解码器 200 和显示装置 220。在一些实例中，输入接口 240 包含接收器和/或调制解调器。输入接口 240 可经由链路 30 和/或从存储装置 40 接收经编码视频数据。显示装置 220 可与解码端 20 集成或可在解码端 20 外部。一般来说，显示装置 220 显示经解码视频数据。显示装置 220 可包括多种显示装置，例如，液晶显示器（liquid crystal display, LCD）、等离子显示器、有机发光二极管（organic light-emitting diode, OLED）显示器或其它类型的显示装置。

尽管图 1 中未图示，但在一些方面，视频编码器 100 和视频解码器 200 可各自与音频编码器和解码器集成，且可包含适当的多路复用器-多路分用器单元或其它硬件和软件，以处置共同数据流或单独数据流中的音频和视频两者的编码。在一些实例中，如果适用的话，那么解复用器（MUX-DEMUX）单元可符合 ITU H.223 多路复用器协议，或例如用户数据报协议（user datagram protocol, UDP）等其它协议。

视频编码器 100 和视频解码器 200 各自可实施为例如以下各项的多种电路中的任一者：一个或多个微处理器、数字信号处理器（digital signal processor, DSP）、专用集成电路（application-specific integrated circuit, ASIC）、现场可编程门阵列（field programmable gate array, FPGA）、离散逻辑、硬件或其任何组合。如果部分地以软件来实施本申请，那么装置可将用于软件的指令存储在合适的非易失性计算机可读存储媒体中，且可使用一或多个处理器在硬件中执行所述指令从而实施本申请技术。前述内容（包含硬件、软件、硬件与软件的组合等）中的任一者可被视为一或多个处理器。视频编码器 100 和视频解码器 200 中的每一者可包含在一或多个编码器或解码器中，所述编码器或解码器中的任一者可集成为相应装置中的组合编码器/解码器（编码解码器）的一部分。

本申请可大体上将视频编码器 100 称为将某些信息“发信号通知”或“发射”到例如视频解码器 200 的另一装置。术语“发信号通知”或“发射”可大体上指代用以对经压缩视频数据进行解码的语法元素和/或其它数据的传送。此传送可实时或几乎实时地发生。替代地，此通信可经过一段时间后发生，例如可在编码时在经编码码流中将语法元素存储到计算机可读存储媒体时发生，解码装置接着可在所述语法元素存储到此媒体之后的任何时间检索所述语法元素。

视频序列通常包含一系列视频帧或图像。图像群组（group of picture, GOP）示例性地包括一系列、一个或多个视频图像。GOP 可在 GOP 的头信息中、图像中的一者或多者的

头信息中或在别处包含语法数据，语法数据描述包含于 GOP 中的图像的数目。图像的每一条带可包含描述相应图像的编码模式的条带语法数据。视频编码器 100 通常对个别视频条带内的视频块进行操作以便编码视频数据。视频块可对应于 CU 内的解码节点。视频块可具有固定或变化的大小，且可根据指定解码标准而在大小上不同。

在一些可行的实施方式中，视频编码器 100 可利用预定义扫描次序来扫描经量化变换系数以产生可经熵编码的串行化向量。在其它可行的实施方式中，视频编码器 100 可执行自适应性扫描。在扫描经量化变换系数以形成一维向量之后，视频编码器 100 可根据上下文自适应性可变长度解码（context-based adaptive variable-length code, CAVLC）、上下文自适应性二进制算术解码（context-based adaptive binary arithmetic coding, CABAC）、基于语法的上下文自适应性二进制算术解码（syntax-based adaptive binary arithmetic coding, SBAC）、概率区间分割熵（probability interval partitioning entropy, PIPE）解码或其他熵解码方法来熵解码一维向量。视频编码器 100 也可熵编码与经编码视频数据相关联的语法元素以供视频解码器 200 用于解码视频数据。

为了执行 CABAC，视频编码器 100 可将上下文模型内的上下文指派给待传输的符号。上下文可与符号的相邻值是否为非零有关。为了执行 CAVLC，视频编码器 100 可选择待传输的符号的可变长度码。可变长度解码（variable-length code, VLC）中的码字可经构建以使得相对较短码对应于可能性较大的符号，而较长码对应于可能性较小的符号。以这个方式，VLC 的使用可相对于针对待传输的每一符号使用相等长度码字达成节省码率的目的。基于指派给符号的上下文可以确定 CABAC 中的概率。

本申请可将视频解码器当前在解码的图像称作当前图像。

在一种示例中，本申请提供的视频编解码系统可应用于视频压缩场景，如人工智能（artificial intelligence, AI）的视频编码/解码模块。

在另一种示例中，本申请提供的视频编解码系统可用于将压缩好的视频文件针对不同业务进行存储，如终端相册的数据（图像或视频）存储、视频监控、华为™云等。

在又一种示例中，本申请提供的视频编解码系统可用于将压缩好的视频文件进行传输，如华为™云、视频直播（live）等。示例性的，当视频编解码系统应用于直播（live）技术中时，输出接口 140 向外（如支持直播技术的服务器集群）发送数据流的过程也可称为推流，服务器集群向输入接口 240 发送数据流的过程也可以称为分发。

以上几种示例仅为本实施例提供的视频编解码系统可能的应用场景，不应理解为对本申请的限定。

在本申请提供的实施例中，为减少码流的数据量，提高码流在多个设备之间的传输速度，编码端 10 可采用光流映射的方式对图像帧进行编码和图像压缩。相对应的，解码端 20 也可采用光流映射的方式对图像帧进行解码和图像重构。

其中，光流表示的是相邻两帧图像中像素的运动速度和运动方向。示例性的，图像域光流表示的是相邻两帧图像中像素之间的运动信息，该运动信息可以包括的相邻两帧图像的像素之间的运动速度和运动方向，特征域光流表示的是相邻两帧图像的特征图之间的运动信息，该运动信息指示了：图像帧的特征图以及该图像帧的相邻图像帧的特征图之间的运动速度和运动方向。示例性的，前述图像帧（如图像帧 1）的相邻图像帧（如图像帧 2）也可称为该图像帧 1 的参考帧。

光流在时间维度上有两个方向：方向①，前一帧图像到后一帧图像的光流；方向②，后一帧图像到前一帧图像的光流。在一个方向上的光流通常进行数字化表示，如使用一个三维的数组 $[2, h, w]$ 表示，“2”表示光流包含2个通道，其中的第一通道表示图像在 x 方向的偏移方向和大小，第二通道表示图像在 y 方向的偏移方向和大小。 h 为图像的高度， w 为图像的宽度。在 x 方向上，正值表示物体向左移动，而负值表示物体向右移动；在 y 方向上，正值表示物体向上移动，而负值表示物体向下移动。

图2为本申请提供的光流和颜色的映射关系图，图2的(A)中箭头示出了图像中光流的方向，箭头的长度示出了图像中光流的大小；图2中(B)中示出了依据该光流确定的图像的颜色和亮度，示例性的，颜色表示光流的方向，亮度表示光流的大小，如亮度越大，光流对应的值越大。值得注意的是，图2中(B)以灰度为例对图像的颜色进行说明，但不应理解为对本申请的限定，关于光流和图像的映射关系的更多内容可参照通常技术的相关内容，此处不予赘述。

通过参考帧和两帧之间的光流预测当前帧的方法称为光流映射(Warping)，通常表示为 $\tilde{x}_t = W(x_{t-1}, v_t)$ 。其中， $\tilde{x}_t$ 为预测帧， x_{t-1} 为参考帧， v_t 为参考帧向预测帧的光流，解码端在已知参考帧和光流，可以推断出当前帧的像素点对应参考帧的位置，根据对应参考帧的位置进行插值就可以得到当前帧的像素值的估计值。

光流映射包括前向映射(forward warping)和后向映射(backward warping)。如图3所示，图3为本申请提供的光流映射的示意图：图3中的A为前向映射，是指解码端根据前一帧图像(图像帧1)和前后帧之间的光流预测后一帧图像(图像帧2)；图3中的B为后向映射，是指解码端根据后一帧图像(图像帧2)和前后帧之间的光流预测前一帧图像(图像帧1)。

下面在图1至图3的基础上，本申请的实施例提供了一种面向特征空间的深度视频压缩框架(framework towards deep video compression in feature space, 简称：FVC)，如图4所示，图4为本申请提供的视频压缩系统的架构示意图，该视频压缩系统400包括：运动估计(motion estimation)模块410、光流压缩(motion compression)模块420、运动补偿(motion compensation)模块430、残差压缩(residual compression)模块440、熵编码(entropy coding)模块450，以及多帧特征融合(multi-frame feature fusion)模块460。其中的运动估计模块410、光流压缩模块420和运动补偿模块430可统称为可变运动补偿组件。

图像帧 X_t 在经过特征提取(feature extraction)后获得图像帧的特征图 F_t ，参考帧 X_{t-1} 在经过特征提取后获得参考帧的特征图 F_{t-1} ，参考帧 X_{t-1} 可存储在重建帧缓存(decoded frame buffer)，重建帧缓存用于提供多个帧的数据存储空间。

运动估计模块410根据 F_t 和 F_{t-1} 确定图像帧与参考帧之间的特征域光流 O_t 。

光流压缩模块420对 O_t 进行压缩获得特征域光流码流 O'_t 。

运动补偿模块430根据参考帧的特征图 F_{t-1} 和解码特征域光流码流 O'_t ，做特征预测确定图像帧 X_t 对应的预测特征图 F_{t_pre} 。

特征域残差为 $R_t = F_t - F_{t_pre}$ ，残差压缩模块440根据获得特征域残差 R_t 输出压缩后的解码残差 R'_t 。该预测特征图 F_{t_pre} 和解码残差 R'_t 可用于确定图像帧的初始重建特征 $F'_{t_initial}$ 。

进一步的，多帧特征融合模块460根据该初始重建特征 $F'_{t_initial}$ 和多个参考帧对应的重建特征图(如图4所示出的 F_{t-1_ref} 、 F_{t-2_ref} 和 F_{t-3_ref})，确定图像帧 X_t 的最终重建特征图 F'_{t_final} 。

该最终重建特征图 F'_{t_final} 经过帧重建 (frame reconstruction) 获得图像帧 X_t 对应的重建帧 X'_t 。

熵编码模块 450 用于将特征域光流 O_t 和特征域残差 R_t 中至少一个进行编码, 获得二进制码流。

图 1 至图 4 仅为本申请提供的示例, 在一些示例中, 编码端 10、解码端 20、视频编码器 100、视频解码器 200 以及视频编解码系统可以包括更多或更少的部件或单元, 本申请对此不予限定。

下面将结合附图对本实施例提供的图像编码方法的具体实现方式进行详细描述。

如图 5 所示, 图 5 为本申请提供的图像编码方法的流程示意图, 该图像编码方法可应用于图 1 所示出的视频编解码系统或者图 4 所示出的视频压缩系统, 示例性的, 该图像编码方法可由编码端 10 或者视频编码器 100 执行, 这里以编码端 10 执行本实施例提供的图像编码方法为例进行说明。如图 5 所示, 本实施例提供的图像编码方法包括以下步骤 S510 至 S560。

S510, 编码端获取图像帧 1 的特征图, 以及图像帧 1 的参考帧的第一特征图。

该图像帧 1 和参考帧可同属于一个视频包括的 GOP。举例来说, 视频包括一个或多个图像帧, 图像帧 1 可以为视频中任一个图像帧, 该参考帧可以是视频中与图像帧 1 相邻的图像帧, 如参考帧为在图像帧 1 之前的相邻图像帧, 或者, 参考帧为在图像帧 1 之后的相邻图像帧。值得注意的是, 在一些情形中, 图像帧 1 也可称为第一图像帧。

在一种可选的示例中, 编码端获取特征图的方式可以包括, 但不限于: 编码端基于神经网络模型来实现。假设视频中每个图像帧的大小为 $[3, H, W]$, 即图像帧具有 3 个通道、高度为 H 、宽度为 W , 图像帧对应的特征图大小为 $[N, H/s, W/s]$, 其中的 s 为正整数, 这里假设 $s=2$, $N=64$, 神经网络模型为特征提取网络, 对本实施例提供的图像帧 1 的特征图和参考帧的第一特征图的特征提取过程进行说明: 如图 6 所示, 图 6 为本申请提供的特征提取网络的结构示意图, 该特征提取网络包括 1 个卷积层 (convolutional layer, conv) 和 3 个残差块 (residual block, resblock) 处理层, 卷积层的卷积核为 $64 \times 5 \times 5/2$, resblock 处理层的卷积核大小为 $64 \times 3 \times 3$, 进而, 编码端确定图像帧 1 对应的特征图为 f_i , 参考帧的第一特征图为 f_{i-1} 。

需要说明的是, 图 6 仅为本实施例提供的一种特征提取方式的示例, 不应理解为对本申请的限定, 当图像帧和参考帧的大小为其他时, 图 6 所示出的特征提取网络中各网络层的参数也可不同。在一些可选的实现方式中, 编码端还可以采用其他方式来提取图像帧 1 和参考帧的特征, 本申请对此不予限定。

S520, 编码端根据图像帧 1 的特征图和第一特征图, 获得至少一个光流集。

该至少一个光流集中的第一光流集对应前述的第一特征图, 如该第一光流集可以为图 5 所示出的光流集 1。

该第一光流集可包括一个或多个特征域光流 v_t , 该特征域光流 v_t 用于指示图像帧 1 (或称第一图像帧) 的特征图与第一特征图之间的运动信息。该运动信息可用于指示图像帧 1 的特征图和第一特征图之间的运动信息和运动方向。

在本申请的实施例中, 编码端根据获取光流集的过程实际是一种光流估计过程。作为一种可行的实现方式, 该光流估计过程可由图 4 所示出的运动估计模块 410 来实现, 该运

动估计模块 410 可由编码端支持实现。

在一些可选的示例中，编码端可利用光流估计网络来确定前述的光流集。举例来说，如图 7 所示，图 7 为本申请提供的光流估计网络的结构示意图，该光流估计网络包括上采样网络和下采样网络，其中的上采样网络包括 3 个网络层 1，网络层 1 依次包括：3 个残差块处理层（卷积核为 $64 \times 3 \times 3$ ）和 1 个卷积层（卷积核为 $64 \times 5 \times 5/2$ ）。下采样网络包括 3 个网络层 2，网络层 2 依次包括：1 个卷积层（卷积核为 $64 \times 5 \times 5/2$ ）和 3 个残差块处理层（卷积核为 $64 \times 3 \times 3$ ）。作为一种可选的情形，1 个残差块处理层依次包括：1 个卷积层（卷积核为 $64 \times 3 \times 3$ ）、一个激活层和 1 个卷积层（卷积核为 $64 \times 3 \times 3$ ）。其中的激活层可以是指线性整流层（rectified linear units layer, ReLU），或参数化修正线性单元（parametric rectified linear unit, PReLU）等。

值得注意的是，图 7 仅为本申请实施例提供的光流估计网络的示例，不应理解为对本申请的限定，其中各网络层的卷积核大小、输入网络层的特征图通道数、下采样位置、卷积层数、网络激活层均可调整。在一些较为复杂的场景中，光流估计网络还可以使用更加复杂的网络结构。

S530，编码端对 S520 获得的光流集进行特征域光流编码，获得特征域光流码流。

该特征域光流码流包括光流集 1 对应的码流。

在一些情形中，该特征域光流码流可以是二进制文件，或者，该特征域光流码流也可以是符合多媒体传输协议的其他类型文件，不予限定。

S540，编码端基于光流集 1 包括的特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图。

这里提供一种可行的处理方式，针对于光流集 1 包括的一个或多个特征域光流，编码端基于其中的第一特征域光流（ v_l ）对第一特征图进行光流映射（warping），获得第一特征域光流对应的中间特征图。值得注意的是，一个特征域光流与第一特征图 warping，会获得一个中间特征图，第一特征图对应的中间特征图数量与光流集 1 包含的特征域光流的数量一致。如图 8 所示，图 8 为本申请提供的光流映射和特征融合的流程示意图，光流集 1 包含多个特征域光流，如 v_l 至 v_m ， m 为正整数。针对于参考帧的第一特征图 f_l ，编码端将光流集 1 包含的每个特征域光流分别与第一特征图 f_l 进行 warping，获得 m 个中间特征图。

S550，编码端融合一个或多个中间特征图，获得第一图像帧的第一预测特征图。

通常技术中，解码端获取图像帧和参考帧之间的多个图像域光流，并根据多个图像域光流与参考帧获取到多张图像，从而将多张图像进行融合获得图像帧对应的目标图像，因此，解码端解码一帧图像需预测多张图像的像素值，并将多张图像进行融合获得目标图像，导致图像解码所需的计算资源较大，解码端根据图像域光流解码视频的效率较低。

相比之下，在本实施例提供的一种可能的情形中，编码端将前述的一个或多个中间特征图输入至特征融合模型，获得该第一预测特征图。如特征融合模型包括卷积网络层，该卷积网络层用于融合中间特征图。如在图像解码过程中，由于特征图融合所需的计算资源小于特征图解码所需的计算资源，因此，解码端采用特征融合模型对多个中间特征图进行融合，进而，解码端基于融合多个中间特征图获得的预测特征图对图像帧进行解码，即解码端仅需根据预测特征图对该预测特征图在图像中指示的图像位置的像素值进行预测，无需预测多张图像的所有像素值，减少了图像解码所需的计算资源，提高了图像解码的效率。

在本实施例提供的另一种可能的情形中，编码端获取前述一个或多个中间特征图的一个或多个权值，并基于该一个或多个权值处理与该一个或多个权值各自对应的中间特征图，以及将所有处理后的中间特征图相加获得第一预测特征图。该权值用于指示中间特征图在第一预测特征图所占的权重。

在本实施例中，针对于第一特征图对应的多个中间特征图，每个中间特征图的权值可以不同，换言之，编码端可根据图像编码的需求为中间特征图设置不同的权值，示例性的，若在一些中间特征图对应的图像较为模糊的情况下，降低这些模糊图像所对应的中间特征图的权值，从而提高第一图像的清晰度。

例如，该权值可以是指每个光流对应的 mask 值，假设光流集 1 包括 9 个特征域光流，各特征域光流的权值依次为： $m_t^1, m_t^2, \dots, m_t^9$ ，如特征域光流的大小为 $[2, H/s, W/s]$ ，mask 大小为 $[1, H/s, W/s]$ 。进而，编码端将这 9 个特征域光流所对应的中间特征图进行融合，获得第一特征图对应的第一预测特征图。

以上两种可能的情形仅为本实施例提供的示例，不应理解为本申请中的特征图融合仅能采用上述两种方式。如图 8 所示，假设 $m=4$ ，针对于参考帧的第一特征图 f_i ，编码端可将光流集 1 包含的每个特征域光流分别与第一特征图 f_i 进行 warping，获得 4 个中间特征图；进而，编码端将 4 个中间特征图进行特征融合获得第一预测特征图。

请继续参见图 5，本实施例提供的图像编码方法还包括以下步骤 S560。

S560，编码端根据第一预测特征图对图像帧 1 对应的残差进行编码获得残差码流。

该第一预测特征图对应的图像区域的残差码流，以及前述 S530 确定的特征域光流码流可统称为图像帧 1（或称第一图像帧）对应的码流。

可选的，该码流包括：前述光流集对应的特征域光流码流和第一预测特征图对应的图像区域的残差码流。

在本实施例中，对于第一图像帧（如前述的图像帧 1）的参考帧，编码端基于该第一图像帧的特征图和参考帧的第一特征图确定一组特征域光流（如前述的光流集 1），并对该参考帧的第一特征图进行处理，获得一个或多个中间特征图；其次，在编码端获得第一特征图对应的一个或多个中间特征图后，将该一个或多个中间特征图进行融合，从而获得第一图像帧对应的第一预测特征图；最后，编码端根据第一预测特征图对第一图像帧进行编码获得码流。

如此，对于单个特征域光流而言，该特征域光流中像素的误差小于图像域光流中像素的误差，因此编码端根据特征域光流确定的中间特征图所带来的编码误差低于通常技术中图像域光流所导致的编码误差，换言之，编码端根据特征域光流对图像帧进行编码，降低了相邻两帧图像之间的图像域光流导致的编码误差，提高了图像编码的准确性。而且，编码端基于多个特征域光流处理参考帧的特征图获得多个中间特征图，以及融合该多个中间特征图确定图像帧的预测特征图，即编码端将光流集和参考帧的特征图确定的多个中间特征图进行融合，获得图像帧的预测特征图，该预测特征图包含了更多的图像信息，使得编码端基于融合获得的预测特征图对图像帧进行编码时，避免了单个中间特征图难以准确表达第一图像的问题，提高了图像编码的准确性和图像质量（如图像清晰度等）。

作为一种可选的实现方式，参考帧可对应多个特征图，如前述的第一特征图，以及第二特征图。示例性的，第一特征图和第二特征图可以是指参考帧包含的多个通道中属于不

同通道的两个特征图。

在本申请的实施例中，参考帧对应的一组特征图可对应一个光流集，如前述的光流集1还可以对应该第二特征图。结合图8所示出的内容，若参考帧的特征图 f_{t-1} 为第二特征图，则编码端还可基于光流集1包括的特征域光流处理第二特征图，获得第二特征图对应的一个或多个中间特征图（如图8所示出的 f_{t-1} 至 f_{t-m} ， m 为正整数）。此外，编码端还融合第二特征图对应的一个或多个中间特征图，获得第一图像帧的第二预测特征图。

由于参考帧对应了多个特征图，如前述的第一特征图和第二特征图，因此，编码端对第一图像帧进行编码的过程可包括以下内容：编码端根据第一预测特征图和第二预测特征图，编码第一图像帧获得码流。该码流可包括第一图像帧中与第一特征图、第二特征图对应的图像区域的残差码流，以及前述光流集1对应的特征域光流码流。

例如，假设参考帧的每8个特征图共享9个特征域光流，这8个特征图看作一个特征图的分组，在图像编码的过程中，属于同一个分组的8个特征图可以共用该9个特征域光流，从而获得 $8 \times 9 = 72$ 个中间特征图。进一步的，在编码端对中间特征图进行特征融合的过程中，将一个特征图对应的9个中间特征图进行融合获得该一个特征图对应的预测特征图。值得注意的是，在本申请中，几个特征图共享几个特征域多光流在此不加限制，每个特征图分组的特征通道数是否一致也不加限制。

在本实施例中，参考帧的多个特征图（或称：一组特征图，或一个特征图分组等）可对应一个光流集（或称一组特征域光流），例如，针对于属于同一组的多个特征图共用一组特征域流，编码端根据特征图对应的一组特征域光流处理该特征图，从而获得该特征图对应的中间特征图，进一步的，编码端将该特征图对应的中间特征图进行融合获得该特征图对应的预测特征图。最后，编码端根据参考帧的所有特征图对应的预测特征图对图像帧进行编码获得目标码流。

如此，在图像编码过程中，若参考帧对应多个特征图，编码端可将该多个特征图划分为一个组或多个组，属于同一个组的特征图共用一个光流集，并将特征图对应的中间特征图进行融合获得预测特征图，避免了当参考帧或图像帧具有较多信息时，编码端根据特征图编码获得的码流冗余较多、精度较低的问题，提高了图像编码的准确性。值得注意的是，属于不同组的特征图的通道数可以不相同。

作为一种可选的实现方式，当参考帧对应多个特征图，且这多个特征图对应一组特征域光流时，对于参考帧的不同特征图，编码端可采用不同的特征融合方式对参考帧的特征图所对应的中间特征图进行融合，从而获得参考帧的特征图所对应的预测特征图。示例性的，参考帧对应了特征图1和特征图2，编码端将特征图1对应多个中间特征图输入特征融合模型，从而确定特征图1对应的预测特征图；编码端获取特征图2对应的每个中间特征图的权值，并根据这些权值处理各自对应的中间特征图，获得特征图2对应的预测特征图。在本实施例中，针对于不同的特征融合方式，在图像的编码需求不同时，编码端可为每种特征融合方式输出的预测特征图设置置信度，从而满足用户的不同编码需求。

对应于前述实施例提供的图像编码方法，为对包含残差码流和特征域光流码流的码流进行解码，获得该码流对应的目标图像或视频，本申请的实施例还提供一种图像解码方法，如图9所示，图9为本申请提供的图像解码方法的流程示意图一，该图像解码方法可应用于图1所示出的视频编解码系统或者图4所示出的视频压缩系统，示例性的，该图像编码

方法可由解码端 20 或者视频解码器 200 执行，这里以解码端 20 执行本实施例提供的图像解码方法为例进行说明。

如图 9 所示，本实施例提供的图像解码方法包括以下步骤 S910 至 S940。

S910，解码端解析码流获得至少一个光流集。

该至少一个光流集包括第一光流集（如图 9 所示出的光流集 1），该光流集 1 对应参考帧的第一特征图，且该光流集 1 包括一个或多个特征域光流，其中该一个或多个特征域光流中任一个特征域光流用于指示第一图像帧的特征图与前述第一特征图之间的运动信息。关于光流集和特征域光流更多的内容可参照前述 S510 的相关阐述，在此不予赘述。

可选的，参考帧可以是指解码端已解码的与该第一图像帧相邻的图像帧（在第一图像帧之前的图像帧，或者在第一图像帧之后的图像帧），或者，该参考帧为码流中携带的包含一张图像的完整信息的图像帧。

S920，解码端基于光流集 1 包括的特征域光流处理第一特征图，获得第一特征图对应的一个或多个中间特征图。

值得注意的是，解码端根据特征域光流处理参考帧的特征图的过程，也称为特征对齐、特征预测、或者特征 align，本申请对此不予限定。

对应于 S520 中编码端基于光流集 1 处理参考帧的特征图的过程，解码端也可采用和前述 S520 相同的方式对第一特征图进行处理，获取到第一特征图对应的一个或多个中间特征图。例如，解码端基于一组特征域光流对第一特征图进行 warping，如解码端可根据一组特征域光流对应参考帧的位置进行插值，获得第一图像中像素值的预测值，避免了解码端需根据图像域光流对第一图像所有像素值进行预测，减少了解码端进行图像解码所需的计算量，提高了图像解码的效率。

S930，解码端融合 S920 确定的一个或多个中间特征图，获得第一图像帧的第一预测特征图。

对应于前述编码端执行的 S550，解码端可将 S920 确定的一个或多个中间特征图输入至特征融合模型，获得该第一预测特征图。如特征融合模型包括卷积网络层，该卷积网络层用于融合中间特征图。在图像解码过程中，由于特征图融合所需的计算资源小于特征图解码所需的计算资源，因此，解码端采用特征融合模型对多个中间特征图进行融合，进而，解码端基于融合多个中间特征图获得的预测特征图对图像帧进行解码，减少了图像解码所需的计算资源，提高了图像解码的效率。

或者，解码端还可获取前述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值；以及解码端基于中间特征图的权值处理中间特征图，并将所有处理后的中间特征图相加获得第一预测特征图。其中的权值用于指示中间特征图在第一预测特征图所占的权重。在本实施例中，针对于第一特征图对应的多个中间特征图，每个中间特征图的权值可以不同，换言之，解码端可根据图像解码的需求为中间特征图设置不同的权值，示例性的，若在一些中间特征图对应的图像较为模糊的情况下，降低这些中间特征图的权值，从而提高第一图像的清晰度。关于中间特征图的权值可参照 S550 的相关内容，在此不予赘述。

S940，解码端根据 S930 确定的第一预测特征图对第一图像帧进行解码，获得第一图像。

示例性的, 码流中与第一预测特征图对应的残差为 r'_i , 第一预测特征图为 f'_i , 则解码端对第一图像帧进行解码所需的重建特征图 $f_{res}=f'_i+r'_i$, 解码端可根据该重建特征图 f_{res} 对第一图像帧进行解码获得第一图像。

在本实施例中, 对于第一图像帧的参考帧, 解码端基于该第一图像帧对应的一组特征域光流(如前述的第一光流集), 对该参考帧的第一特征图进行处理, 获得一个或多个中间特征图; 其次, 在解码端获得第一特征图对应的一个或多个中间特征图后, 将该一个或多个中间特征图进行融合, 从而获得第一图像帧对应的第一预测特征图; 最后, 解码端根据该第一预测特征图对第一图像帧进行解码获得第一图像。

如此, 对于单个特征域光流而言, 该特征域光流中像素的误差小于图像域光流中像素的误差, 因此, 解码端根据特征域光流确定的中间特征图所带来的解码误差低于通常技术中图像域光流所导致的解码误差, 换言之, 解码端根据特征域光流对图像帧进行解码, 降低了相邻两帧图像之间的图像域光流导致的解码误差, 提高了图像解码的准确性。

而且, 解码端基于多个特征域光流处理参考帧的特征图获得多个中间特征图, 以及融合该多个中间特征图确定图像帧的预测特征图, 即解码端将光流集和参考帧的特征图确定的多个中间特征图进行融合, 获得图像帧的预测特征图, 该预测特征图包含了更多的图像信息, 使得解码端基于融合获得的预测特征图对图像帧进行解码时, 避免了单个中间特征图难以准确表达第一图像的问题, 提高了图像解码的准确性和图像质量(如图像清晰度等)。

可选的, 编码端对特征图进行编码可通过特征编码网络实现, 解码端对特征图对应的码流进行解码可通过特征解码网络实现。示例性的, 针对于编码端所需的特征编码网络和解码端所需的特征解码网络, 这里提供了一种可能的实施例, 如图 10 所示, 图 10 为本申请提供的特征编码网络和特征解码网络的结构示意图, 特征编码网络包括 3 个网络层 1, 网络层 1 依次包括: 3 个残差块处理层(卷积核为 $64 \times 3 \times 3$)和 1 个卷积层(卷积核为 $64 \times 5 \times 5/2$)。

特征解码网络包括 3 个网络层 2, 网络层 2 依次包括: 1 个卷积层(卷积核为 $64 \times 5 \times 5/2$)和 3 个残差块处理层(卷积核为 $64 \times 3 \times 3$)。

在一种可选的情形中, 1 个残差块处理层依次包括: 1 个卷积层(卷积核为 $64 \times 3 \times 3$)、一个激活层和 1 个卷积层(卷积核为 $64 \times 3 \times 3$)。其中的激活层可以是指 ReLU 层或 PReLU 等。

值得注意的是, 图 10 仅为本申请实施例提供的特征编码网络和特征解码网络的示例, 不应理解为对本申请的限定, 其中各网络层的卷积核大小、输入网络层的特征图通道数、下采样位置、卷积层数、网络激活层均可调整。在一些较为复杂的场景中, 光流估计网络还可以使用更加复杂的网络结构。

在一种可选的实现方式中, 参考帧可对应多个特征图, 如前述的第一特征图, 以及第二特征图。

在本申请的实施例中, 参考帧对应的一组特征图可对应一个光流集, 如前述的光流集 1 还可以对应该参考帧的第二特征图。结合图 8 所示出的内容, 若参考帧的特征图 f_{t-1} 为第二特征图, 则解码端还可基于光流集 1 包括的特征域光流处理第二特征图, 获得第二特征图对应的一个或多个中间特征图(如图 8 所示出的 f_{t-1} 至 f_{t-m} , m 为正整数)。此外, 解码端还融合第二特征图对应的一个或多个中间特征图, 获得第一图像帧的第二预测特征图。

由于参考帧对应了多个特征图，如前述的第一特征图和第二特征图，因此，解码端对第一图像帧进行解码的过程可包括以下内容：解码端根据第一预测特征图和第二预测特征图，解码第一图像帧获得第一图像。例如，假设参考帧的每8个特征图共享9个特征域光流，这8个特征图看作一个特征图的分组，在图像解码的过程中，属于同一个分组的8个特征图可以共用该9个特征域光流，从而获得 $8 \times 9 = 72$ 个中间特征图。进一步的，在解码端对中间特征图进行特征融合的过程中，将一个特征图对应的9个中间特征图进行融合获得该一个特征图对应的预测特征图。值得注意的是，属于不同组的特征图的通道数可以不相同。

在本实施例中，参考帧的多个特征图（或称一组特征图）可对应一个光流集（或称一组特征域光流），换言之，针对于属于同一组的多个特征图共用一组特征域流，解码端根据特征图对应的一组特征域光流处理该特征图，从而获得该特征图对应的中间特征图，进一步的，解码端将该特征图对应的中间特征图进行融合获得该特征图对应的预测特征图。最后，解码端根据参考帧的所有特征图对应的预测特征图对图像帧进行解码获得目标图像。如此，在图像解码过程中，若参考帧对应多个特征图，解码端可将该多个特征图划分为一个组或多个组，属于同一个组的特征图共用一个光流集，并将特征图对应的中间特征图进行融合获得预测特征图，避免了当参考帧或图像帧具有较多信息时，解码端根据特征图重构图像的精度较低问题，提高了图像解码的准确性。

在一些情形中，若第一图像的清晰度未能达到预期的清晰度，导致解码后的第一图像较为模糊，为提高第一图像的清晰度，改善视频显示效果，本申请的实施例还提供了一种视频增强的技术方案，如图11所示，图11为本申请提供的图像解码方法的流程示意图二，图11所示出的图像解码方法可以和前述实施例提供的图像编码方法和图像解码方法相结合，也可以单独实现，这里以解码端执行图11所示出的图像解码方法为例进行说明，如图11所示，本实施例提供的图像解码方法包括以下步骤S1110至S1130。

S1110，解码端获取第一图像的特征图。

示例性的，解码端可根据图6所示出的特征提取网络来获取第一图像的特征图，在此不予赘述。

S1120，解码端根据第一图像的特征图、参考帧的第一特征图 and 第一预测特征图，获得增强特征图。

示例性的，解码端可通过特征融合模型来对S1120包含的多个特征图进行融合，获得增强特征图。特征融合模型可以是指前述S550或者S930所提供的特征融合模型，也可以是指包含卷积层的其他模型，如该卷积层的卷积核为 3×3 ，本申请对此不予限定。

在一些可能的情形中，解码端还可为第一图像的特征图、参考帧的第一特征图和第一预测特征图设置不同的权值，从而将这多个特征图进行融合获得增强特征图。

在一些可能的示例中，该增强特征图可用于确定第一图像的增强层图像。该增强层图像的视频质量高于第一图像的视频质量，视频质量可以是指或包括图像信噪比（signal-to-noise ratio, SNR）、分辨率（image resolution）和峰值信噪比（Peak signal-to-noise ratio, PSNR）中至少一种。在本文中，针对于一帧图像的视频质量可称为该图像的图像质量。

其中，图像信噪比是指图像的信号均值与背景标准偏差的比值，对于图像，这里的“信号均值”一般是指图像的灰度平均值，背景标准偏差可以用图像的背景信号值的方差来表

示,背景信号值的方差是指噪声功率;对于图像而言,图像信噪比越大,图像的质量越好。分辨率是单帧图像在单位面积内的像素点数,图像的分辨率越高,图像的质量越好。PSNR用于指示图像的主观质量,PSNR越大,图像的质量越好。关于SNR、分辨率和PSNR的更多内容可以参考现有技术的相关描述,此处不予赘述。

S1130,解码端根据增强特征图处理第一图像,获得第二图像。

第二图像和第一图像指示的内容相同,但第二图像的清晰度高于第一图像的清晰度。在一些实施例中,图像的清晰度可以用峰值信噪比(PSNR)来指示,

本实施例中,解码端可将第一图像的特征图、第一特征图和第一预测特征图进行融合,并基于融合获得的增强特征图对第一图像进行视频增强处理,获得清晰度更优的第二图像,从而提高了解码后的图像清晰度及图像显示效果。

可选的,解码端根据增强特征图处理第一图像,获得第二图像,包括:解码端依据增强特征图获得第一图像的增强层图像,并基于增强层图像重构第一图像,获得第二图像。示例性的,该增强层图像可以是指解码端依据参考帧和该增强特征图确定的图像,解码端在第一图像的基础上加入增强层图像的部分或全部信息,获得第二图像;或者,解码端将增强层图像作为第一图像的重构图像,即前述的第二图像。在本示例中,解码端获取到图像在不同阶段的多个特征图,并获得了这多个特征图确定的增强特征图,从而依据该增强特征图对第一图像进行重构和视频增强,提高了解码后的图像清晰度和图像显示效果。

在一种可能的情形中,解码端对第一图像进行重构可以是基于图像重构网络实现的,如图12所示,图12为本申请提供的图像重构网络的结构示意图,该图像重构网络依次包括:3个残差块处理层(卷积核为 $64 \times 3 \times 3$)和1个反卷积层(卷积核为 $3 \times 5 \times 2$),其中的残差块处理层可以包括1个卷积层(卷积核为 $64 \times 3 \times 3$)、一个激活层和1个卷积层(卷积核为 $64 \times 3 \times 3$)。

可以理解的是,图6、图7、图10和图12所示出的各网络是本实施例以图像帧的大小是 $[3, H, W]$ 为例进行说明的,不应理解为对本申请的限定。例如,若图像帧的大小发生了变化,前述的各网络中卷积层、残差块处理层、反卷积层的参数也会根据视频编解码、视频增强等的需求而改变,本申请对此不予限定。

在本申请的实施例中,编码端/解码端是通过对参考帧的特征图进行光流映射确定的中间特征图集预测特征图,在一些情形中,编码端/解码端也可基于可变形网络(deformable convolutional networks, DCN)对特征图进行处理,DCN是基于卷积实现的,下面是一种卷积的数学表达式:

$$F(p) = \sum_{k=1}^{n^2} w(p_k) * F(p + p_k)$$

其中 n 为卷积核大小, w 为卷积核权重, F 是输入特征图(input feature map), p 是卷积位置, p_k 是相对 p 的位置在卷积核中的枚举值。DCN是基于一个网络学习偏移(offset),使得卷积核在input feature map的采样点发生偏移,集中于感兴趣的区域(region of interest, ROI)或者目标区域,其数学表达形式为:

$$F(p) = \sum_{k=1}^{n^2} w(p_k) * F(p + p_k + \Delta p_k)$$

另外带有 mask 的数学表达形式为：

$$F(p) = \sum_{k=1}^{n^2} m(p_k) * w(p_k) * F(p + p_k + \Delta p_k)$$

其中 Δp_k 为相对于 p_k 的偏移量，这样卷积采样位置就变成了不规则位置， $m(p_k)$ 表示位置 p_k 的 mask 值是对位置 p_k 的惩罚项。如图 13 所示通过邻域内的点卷积出最中心点的卷积操作：(a) 是常见的 3x3 卷积核的采样方式，(b) 是采样可变形卷积加上偏移量之后的采样点的变化，其中 (c)、(d) 是可变形卷积的特殊形式，如 (c) 中的多个像素点经过缩放后获得目标图像中相应位置的像素预测值，又如 (d) 中的多个像素点经过旋转和缩放后获得目标图像中相应位置的像素预测值。

将特征域光流设置为前述表达式中的偏移量 Δp_k ，则本申请实施例提供的图像编码方法、图像解码方法也可由 DCN 实现。可以理解的，当参考帧对应多个通道不同的特征图时，每个特征图可采用相同或不同的 DCN 处理方式，如图 13 所示出的 4 种可能的 DCN 处理方式。

针对于通常技术和本方案提供的视频增强的过程，图 14 为本申请提供的效率比对的示意图，图 14 提供了两个指标来对本申请提供的技术方案和通常技术进行比较，这两个指标为：PSNR 和每像素比特数 (bits per pixel, BPP)，BPP 是指存储每个像素所用的位数，BPP 也用于指示图像的分辨率。

参照图 14 可知，在多个码点（编码或解码的节点位置）中，通常技术所提供的方案的 BPP 高于本申请实施例提供的技术方案，即通常技术产生的码流所占用的存储空间更大，传输码流所需的网络带宽更大；但通常技术所提供的方案的 PSNR 低于本申请实施例提供的技术方案，即本申请实施例提供的技术方案中，码流解码后获得的视频（或图像）的主观质量更优。

举例来说，对于一帧分辨率为 1080P (1920×1080) 的图像，通常技术方案测试该帧图像的端到端的时间为 0.3206 秒 (second, s)，而本申请实施例提供的技术方案测试该帧图像的端到端的时间为 0.2188s。也就是说，在视频编解码过程中，本申请实施例提供的技术方案不仅可以节省码率（由 BPP 体现），又保证了质量（由 PSNR 体现），还降低了单帧图像的编解码的时延，从而提高了视频编解码的整体效率。

值得注意的是，本申请提供的图像编码方法、图像解码方法不仅可以应用于视频编解码、视频增强、视频压缩等场景，还可应用于视频预测、视频插帧、视频分析等所有需要视频帧间特征融合或特征对齐的视频处理技术领域。

可以理解的是，为了实现上述实施例中功能，编码端和解码端包括了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本申请中所公开的实施例描述的各示例的单元及方法步骤，本申请能够以硬件或硬件和计算机软件相结合的形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用场景和设计约束条件。

图 15 为本申请提供的图像编码装置的结构示意图，图像编码装置可以用于实现上述方法实施例中编码端或视频编码器的功能，因此也能实现上述图像编码方法实施例所具备的有益效果。在本申请的实施例中，该图像编码装置可以是如图 1 所示的编码端 10 或者视频编码器 100，也可以是应用于编码端 10 或者视频编码器 100 的模块（如芯片）。

如图 15 所示, 图像编码装置 1500 包括: 获取单元 1510、处理单元 1520、融合单元 1530 和编码单元 1540。图像编码装置 1500 用于实现上述图 5 和图 8 所示的方法实施例中编码端或视频编码器的功能。

当图像编码装置 1500 用于实现图 5 所示的方法实施例时, 获取单元 1510 用于执行 S510, 处理单元 1520 用于执行 S520 和 S540, 融合单元 1530 用于执行 S550, 编码单元 1540 用于执行 S530 和 S560。

当图像编码装置 1500 用于实现图 8 所示的方法实施例时, 处理单元 1520 和融合单元 1530 用于实现光流映射和特征融合的功能。

有关上述获取单元 1510、处理单元 1520、融合单元 1530 和编码单元 1540 更详细的描述可以直接参考图 5 和图 8 所示的方法实施例中相关描述直接得到, 这里不加赘述。

相应的, 本申请的实施例还提供了一种图像解码装置, 如图 16 所示, 图 16 为本申请提供的图像解码装置的结构示意图, 图像解码装置可以用于实现上述方法实施例中解码端或视频解码器的功能, 因此也能实现上述图像解码方法实施例所具备的有益效果。在本申请的实施例中, 该图像解码装置可以是如图 1 所示的解码端 20 或者视频解码器 200, 也可以是应用于解码端 20 或者视频解码器 200 的模块 (如芯片)。

如图 16 所示, 图像解码装置 1600 包括: 码流单元 1610、处理单元 1620、融合单元 1630 和解码单元 1640。图像解码装置 1600 用于实现上述图 8 和图 9 所示的方法实施例中解码端或视频解码器的功能。

当图像解码装置 1600 用于实现图 8 所示的方法实施例时, 处理单元 1620 和融合单元 1630 用于实现光流映射和特征融合的功能。

当图像解码装置 1600 用于实现图 9 所示的方法实施例时, 码流单元 1610 用于执行 S910, 处理单元 1620 用于执行 S920, 融合单元 1630 用于执行 S930, 解码单元 1640 用于执行 S940。

在一些可选的情形中, 图像解码装置 1600 还可以包括增强单元, 该增强单元用于根据增强特征图处理第一图像, 获得第二图像。在一种可能的具体示例中, 增强单元, 具体用于: 依据增强特征图获得第一图像的增强层图像。以及, 基于增强层图像重构第一图像, 获得第二图像。

有关上述码流单元 1610、处理单元 1620、融合单元 1630、解码单元 1640 和增强单元更详细的描述可以直接参考图 8 和图 9 所示的方法实施例中相关描述直接得到, 这里不加赘述。

图像编码 (或图像解码) 装置通过软件实现前述附图中任一所示的图像编码 (或图像解码) 方法时, 图像编码 (或图像解码) 装置及其各个单元也可以为软件模块。通过处理器调用该软件模块实现上述的图像编码 (或图像解码) 方法。该处理器可以是中央处理单元 (central processing unit, CPU), 特定应用集成电路 (application-specific integrated circuit, ASIC) 实现, 或可编程逻辑器件 (programmable logic device, PLD), 上述 PLD 可以是复杂程序逻辑器件 (complex programmable logical device, CPLD)、现场可编程门阵列 (field programmable gate array, FPGA)、通用阵列逻辑 (generic array logic, GAL) 或其任意组合。

有关上述图像编码 (或图像解码) 装置更详细的描述可以参考前述附图所示的实施例

中相关描述，这里不加赘述。可以理解的，前述附图所示出的图像编码（或图像解码）装置仅为本实施例提供的示例，根据图像编码（或图像解码）过程或者业务的不同，图像编码（或图像解码）装置可包括更多或更少的单元，本申请对此不予限定。

当图像编码（或图像解码）装置通过硬件实现时，该硬件可以通过处理器或芯片实现。芯片包括接口电路和控制电路。接口电路用于接收来自处理器之外的其它设备的数据并传输至控制电路，或将来自控制电路的数据发送给处理器之外的其它设备。

控制电路通过逻辑电路或执行代码指令用于实现上述实施例中任一种可能实现方式的方法。有益效果可以参见上述实施例中任一方面的描述，此处不再赘述。

可以理解的是，本申请的实施例中的处理器可以是 CPU、神经处理器（neural processing unit, NPU）或图形处理器（graphic processing unit, GPU），还可以是其它通用处理器、数字信号处理器（digital signal processor, DSP）、ASIC、FPGA 或者其它可编程逻辑器件、晶体管逻辑器件，硬件部件或者其任意组合。通用处理器可以是微处理器，也可以是任何常规的处理器。

本申请的实施例中的方法步骤可以通过硬件的方式来实现，如该硬件为图像译码的装置，如图 17 所示，图 17 为本申请提供的图像译码的装置的结构示意图，该图像译码的装置 1700 包括存储器 1710 和至少一个处理器 1720，该处理器 1720 可以实现上述实施例提供的图像编码方法和图像解码方法，该存储器 1710 用于存储上述图像编码方法和图像解码方法对应的软件指令。作为一种可选的实现方式，在硬件实现上，图像译码的装置 1700 可以是指封装有一个或多个处理器 1720 的芯片或芯片系统。示例的，当图像译码的装置 1700 用于实现上述实施例中方法步骤时，图像译码的装置 1700 包括的处理器 1720 执行上述方法的步骤及其可能的子步骤。在一种可选的情形中，图像译码的装置 1700 还可以包括通信接口 1730，该通信接口 1730 可以用于收发数据。例如，通信接口 1730 用于接收用户的编码请求、解码请求，或者发送码流、接收码流等。

本申请的实施例中，通信接口 1730、处理器 1720 以及存储器 1710 之间可通过总线 1740 连接，所述总线 1740 可以分为地址总线、数据总线、控制总线等。

值得注意的是，图像译码的装置 1700 还可以执行图 15 所示出的图像编码装置 1500 的功能，以及图 16 所示出的图像解码装置 1600 的功能，此处不予赘述。

本实施例提供的图像译码的装置 1700 可以是服务器，个人计算机，或者具有数据处理功能的其他图像译码的装置 1700，本申请对此不予限定。例如，该图像译码的装置 1700 可以是前述的编码端 10（或视频编码器 100），或者解码端 20（或视频解码器 200）。又如，该图像译码的装置 1700 也可以同时具有前述编码端 10 和解码端 20 的功能，如该图像译码的装置 1700 是指具有视频编解码功能的视频编解码系统（或者视频压缩系统）。

本申请的实施例中的方法步骤也可以由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成，软件模块可以被存放于随机存取存储器（random access memory, RAM）、闪存、只读存储器（read-only memory, ROM）、可编程只读存储器（programmable ROM, PROM）、可擦除可编程只读存储器（erasable PROM, EPROM）、电可擦除可编程只读存储器（electrically EPROM, EEPROM）、寄存器、硬盘、移动硬盘、CD-ROM 或者本领域熟知的任何其它形式的存储介质中。一种示例性的存储介质耦合至处理器，从而使处理器能够从该存储介质读取信息，且可向该存储介质写入信息。当然，存

储介质也可以是处理器的组成部分。处理器和存储介质可以位于 ASIC 中。另外，该 ASIC 可以位于网络设备或终端设备中。当然，处理器和存储介质也可以作为分立组件存在于网络设备或终端设备中。

另外，本申请还提供一种计算机可读存储介质，该计算机可读存储介质中存储有根据前述任一实施例提供的图像编码方法获得的码流。如该计算机可读存储介质可以为，但不限于：RAM、闪存、ROM、PROM、EPROM、EEPROM、寄存器、硬盘、移动硬盘、CD-ROM 或者本领域熟知的任何其它形式的存储介质。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机程序或指令。在计算机上加载和执行所述计算机程序或指令时，全部或部分地执行本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、网络设备、用户设备或者其它可编程装置。所述计算机程序或指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机程序或指令可以从一个网站站点、计算机、服务器或数据中心通过有线或无线方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是集成一个或多个可用介质的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，例如，软盘、硬盘、磁带；也可以是光介质，例如，数字视频光盘（digital video disc，DVD）；还可以是半导体介质，例如，固态硬盘（solid state drive，SSD）。

在本申请的各个实施例中，如果没有特殊说明以及逻辑冲突，不同的实施例之间的术语和/或描述具有一致性、且可以相互引用，不同的实施例中的技术特征根据其内在的逻辑关系可以组合形成新的实施例。

本申请中，“至少一个”是指一个或者多个，“多个”是指两个或两个以上。“和/或”，描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 的情况，其中 A，B 可以是单数或者复数。在本申请的文字描述中，字符“/”，一般表示前后关联对象是一种“或”的关系；在本申请的公式中，字符“/”，表示前后关联对象是一种“相除”的关系。

可以理解的是，在本申请的实施例中涉及的各种数字编号仅为描述方便进行的区分，并不用来限制本申请的实施例的范围。上述各过程的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定。

权 利 要 求 书

1.一种图像解码方法，其特征在于，所述方法包括：

解析码流获得至少一个光流集；其中，所述至少一个光流集包括第一光流集，所述第一光流集包括一个或多个特征域光流，所述第一光流集对应第一图像帧的参考帧的第一特征图，所述一个或多个特征域光流中的任一个特征域光流用于指示第一图像帧的特征图与所述第一特征图之间的运动信息；

基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一個或多个中间特征图；

融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图；

根据所述第一预测特征图解码所述第一图像帧获得第一图像。

2.根据权利要求1所述的方法，其特征在于，所述基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一個或多个中间特征图，包括：

解析所述码流获得所述第一特征图；

基于第一特征域光流对所述第一特征图进行光流映射 warping，获得所述第一特征域光流对应的中间特征图，其中，所述第一特征域光流为所述一个或多个特征域光流中的任一个。

3.根据权利要求1或2所述的方法，其特征在于，所述第一光流集还对应所述参考帧的第二特征图；

在所述根据所述第一预测特征图解码所述第一图像帧获得第一图像之前，所述方法还包括：

基于所述一个或多个特征域光流处理所述第二特征图，获得所述第二特征图对应的一個或多个中间特征图；

融合所述第二特征图对应的一個或多个中间特征图，获得所述第一图像帧的第二预测特征图；

所述根据所述第一预测特征图解码所述第一图像帧获得第一图像，包括：

根据所述第一预测特征图和所述第二预测特征图，解码所述第一图像帧获得第一图像。

4.根据权利要求1至3中任一项所述的方法，其特征在于，所述融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图，包括：

获取所述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，所述权值用于指示所述中间特征图在所述第一预测特征图中所占的权重；

基于所述一个或多个权值处理与所述一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得所述第一预测特征图。

5.根据权利要求1至3中任一项所述的方法，其特征在于，所述融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图，包括：

将所述第一特征图对应的一個或多个中间特征图输入特征融合模型，获得第一预测特征图；所述特征融合模型包括卷积网络层，所述卷积网络层用于融合中间特征图。

6.根据权利要求1至5中任一项所述的方法，其特征在于，所述方法还包括：

获取所述第一图像的特征图；

根据所述第一图像的特征图、所述第一特征图和所述第一预测特征图，获得增强特征

图；

根据所述增强特征图处理所述第一图像，获得第二图像；所述第二图像的清晰度高于所述第一图像的清晰度。

7.根据权利要求 6 所述的方法，其特征在于，所述根据所述增强特征图处理所述第一图像，获得第二图像，包括：

依据所述增强特征图获得所述第一图像的增强层图像；

基于所述增强层图像重构所述第一图像，获得所述第二图像。

8.一种图像编码方法，其特征在于，所述方法包括：

获取第一图像帧的特征图，以及所述第一图像帧的参考帧的第一特征图；

根据所述第一图像帧的特征图和所述第一特征图，获得至少一个光流集；所述至少一个光流集包括第一光流集，所述第一光流集包括一个或多个特征域光流，所述第一光流集对应所述第一特征图，所述一个或多个特征域光流中的任一个特征域光流用于指示所述第一图像帧的特征图与所述第一特征图之间的运动信息；

基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一个或多个中间特征图；

融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图；

根据所述第一预测特征图对所述第一图像帧进行编码获得码流。

9.根据权利要求 8 所述的方法，其特征在于，所述基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一个或多个中间特征图，包括：

基于第一特征域光流对所述第一特征图进行光流映射 warping，获得所述第一特征域光流对应的中间特征图，其中，所述第一特征域光流为所述一个或多个特征域光流中的任一个。

10.根据权利要求 8 或 9 所述的方法，其特征在于，所述第一光流集还对应所述参考帧的第二特征图；

在所述根据所述第一预测特征图对所述第一图像帧进行编码获得码流之前，所述方法还包括：

基于所述一个或多个特征域光流处理所述第二特征图，获得所述第二特征图对应的一个或多个中间特征图；

融合所述第二特征图对应的一个或多个中间特征图，获得所述第一图像帧的第二预测特征图；

所述根据所述第一预测特征图对所述第一图像帧进行编码获得码流，包括：

根据所述第一预测特征图和所述第二预测特征图，编码所述第一图像帧获得码流。

11.根据权利要求 8 至 10 中任一项所述的方法，其特征在于，所述融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图，包括：

获取所述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，所述权值用于指示所述中间特征图在所述第一预测特征图中所占的权重；

基于所述一个或多个权值处理与所述一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得所述第一预测特征图。

12.根据权利要求 8 至 11 中任一项所述的方法，其特征在于，所述融合所述一个或多

个中间特征图，获得所述第一图像帧的第一预测特征图，包括：

将所述第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图；所述特征融合模型包括卷积网络层，所述卷积网络层用于融合中间特征图。

13.一种图像解码装置，其特征在于，所述装置包括：

码流单元，用于解析码流获得至少一个光流集；所述至少一个光流集包括第一光流集，所述第一光流集包括一个或多个特征域光流，其中的第一光流集对应第一图像帧的参考帧的第一特征图，所述一个或多个特征域光流中的任一个特征域光流用于指示所述第一图像帧的特征图与所述第一特征图之间的运动信息；

处理单元，用于基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一个或多个中间特征图；

融合单元，用于融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图；

解码单元，用于根据所述第一预测特征图解码所述第一图像帧获得第一图像。

14.根据权利要求13所述的装置，其特征在于，所述处理单元，具体用于：解析所述码流获得所述第一特征图；以及，基于第一特征域光流对所述第一特征图进行光流映射warping，获得所述第一特征域光流对应的中间特征图；所述第一特征域光流为所述一个或多个特征域光流中的任一个。

15.根据权利要求13或14所述的装置，其特征在于，所述第一光流集还对应所述参考帧的第二特征图；

所述处理单元，还用于基于所述一个或多个特征域光流处理所述第二特征图，获得所述第二特征图对应的一个或多个中间特征图；

所述融合单元，还用于融合所述第二特征图对应的一个或多个中间特征图，获得所述第一图像帧的第二预测特征图；

所述解码单元，具体用于根据所述第一预测特征图和所述第二预测特征图，解码所述第一图像帧获得第一图像。

16.根据权利要求13至15中任一项所述的装置，其特征在于，所述融合单元，具体用于：获取所述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，所述权值用于指示所述中间特征图在所述第一预测特征图中所占的权重；以及，基于所述一个或多个权值处理与所述一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得所述第一预测特征图。

17.根据权利要求13至15中任一项所述的装置，其特征在于，所述融合单元，具体用于：将所述第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图；所述特征融合模型包括卷积网络层，所述卷积网络层用于融合中间特征图。

18.根据权利要求13至17中任一项所述的装置，其特征在于，所述装置还包括：获取单元和增强单元；

所述获取单元，用于获取所述第一图像的特征图；

所述融合单元，还用于根据所述第一图像的特征图、所述第一特征图和所述第一预测特征图，获得增强特征图；

所述增强单元，用于根据所述增强特征图处理所述第一图像，获得第二图像；所述第

二图像的清晰度高于所述第一图像的清晰度。

19.根据权利要求 18 所述的装置，其特征在于，所述增强单元，具体用于：依据所述增强特征图获得所述第一图像的增强层图像；以及，基于所述增强层图像重构所述第一图像，获得所述第二图像。

20.一种图像编码装置，其特征在于，所述装置包括：

获取单元，用于获取第一图像帧的特征图，以及所述第一图像帧的参考帧的第一特征图；

处理单元，用于根据所述第一图像帧的特征图和所述第一特征图，获得至少一个光流集；所述至少一个光流集包括第一光流集，所述第一光流集包括一个或多个特征域光流，所述第一光流集对应所述第一特征图，所述一个或多个特征域光流中的任一个特征域光流用于指示所述第一图像帧的特征图与所述第一特征图之间的运动信息；

所述处理单元，还用于基于所述一个或多个特征域光流处理所述第一特征图，获得所述第一特征图对应的一个或多个中间特征图；

融合单元，用于融合所述一个或多个中间特征图，获得所述第一图像帧的第一预测特征图；

编码单元，用于根据所述第一预测特征图对所述第一图像帧进行编码获得码流。

21.根据权利要求 20 所述的装置，其特征在于，所述处理单元，具体用于：基于第一特征域光流对所述第一特征图进行光流映射 warping，获得所述第一特征域光流对应的中间特征图，所述第一特征域光流为所述一个或多个特征域光流中的任一个。

22.根据权利要求 20 或 21 所述的装置，其特征在于，所述第一光流集还对应所述参考帧的第二特征图；

所述处理单元，还用于基于所述一个或多个特征域光流处理所述第二特征图，获得所述第二特征图对应的一个或多个中间特征图；

所述融合单元，还用于融合所述第二特征图对应的一个或多个中间特征图，获得所述第一图像帧的第二预测特征图；

所述编码单元，具体用于根据所述第一预测特征图和所述第二预测特征图，编码所述第一图像帧获得码流。

23.根据权利要求 20 至 22 中任一项所述的装置，其特征在于，所述融合单元，具体用于：获得所述一个或多个中间特征图的一个或多个权值，其中，一个中间特征图对应一个权值，所述权值用于指示所述中间特征图在所述第一预测特征图所占的权重；以及，基于所述一个或多个权值处理与所述一个或多个权值各自对应的中间特征图，并将所有处理后的中间特征图相加获得所述第一预测特征图。

24.根据权利要求 20 至 22 中任一项所述的装置，其特征在于，所述融合单元，具体用于：将所述第一特征图对应的一个或多个中间特征图输入特征融合模型，获得第一预测特征图；所述特征融合模型包括卷积网络层，所述卷积网络层用于融合中间特征图。

25.一种图像译码的装置，其特征在于，包括：存储器和处理器；所述存储器用于存储程序代码，所述处理器用于调用所述程序代码实现权利要求 1 至 7 中任意一项所述的方法，或者实现权利要求 8 至 12 中任意一项所述的方法。

26.一种计算机可读存储介质，其特征在于，所述存储介质中存储有计算机程序或指令，

当所述计算机程序或指令被电子设备执行时,实现权利要求 1 至 7 中任意一项所述的方法,或者实现权利要求 8 至 12 中任意一项所述的方法。

27.一种计算机可读存储介质,其特征在于,所述存储介质中存储有根据权利要求 8-12 任意一项所述的方法获取的码流。

28.一种视频编解码系统,其特征在于,包括:编码端和解码端;

所述编码端用于执行权利要求 1 至 7 中任意一项所述的方法;

所述解码端用于执行权利要求 8 至 12 中任意一项所述的方法。

29.一种计算机程序产品,其特征在于,当所述计算机程序产品在电子设备上运行时,所述电子设备实现权利要求 1 至 7 中任意一项所述的方法,和/或,所述电子设备实现权利要求 8 至 12 中任意一项所述的方法。

30.一种芯片,其特征在于,包括:控制电路和接口电路;

所述接口电路用于:接收来自所述芯片之外的其它设备的信号并传输至处理器,或将来自所述控制电路的信号发送给所述芯片之外的其它设备;

所述控制电路用于:通过逻辑电路或执行代码指令来实现权利要求 1 至 7 中任意一项所述的方法,或者实现权利要求 8 至 12 中任意一项所述的方法。

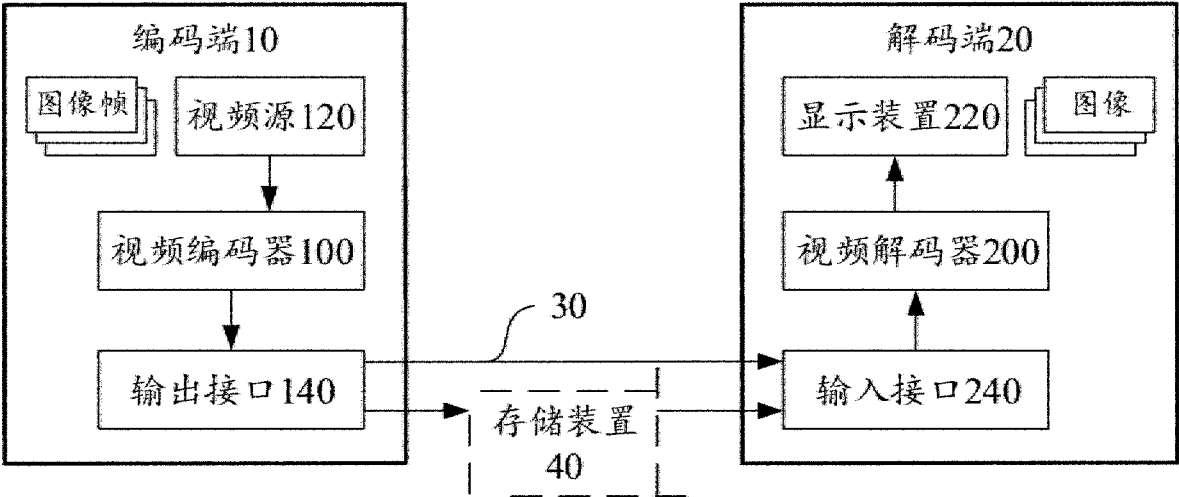


图 1

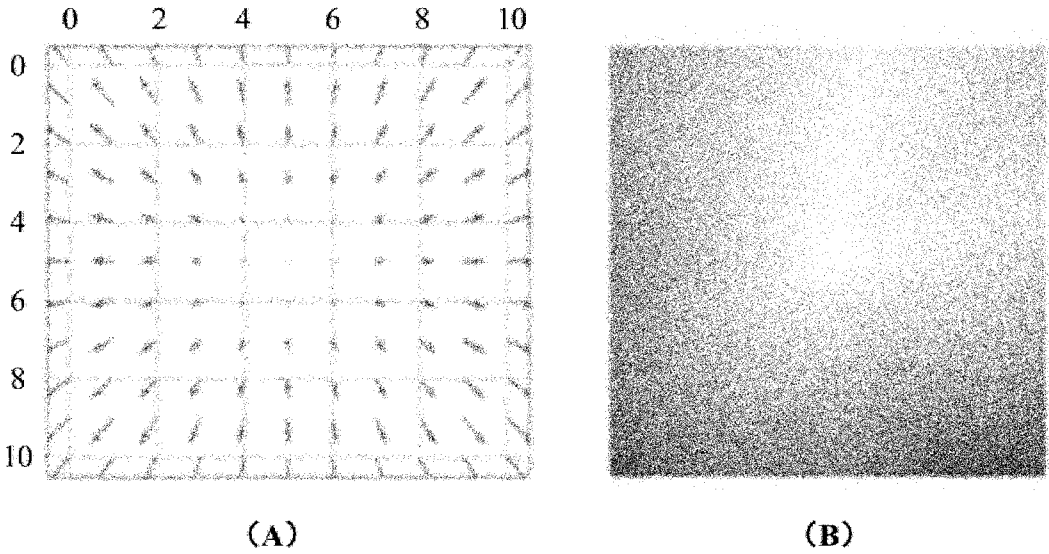
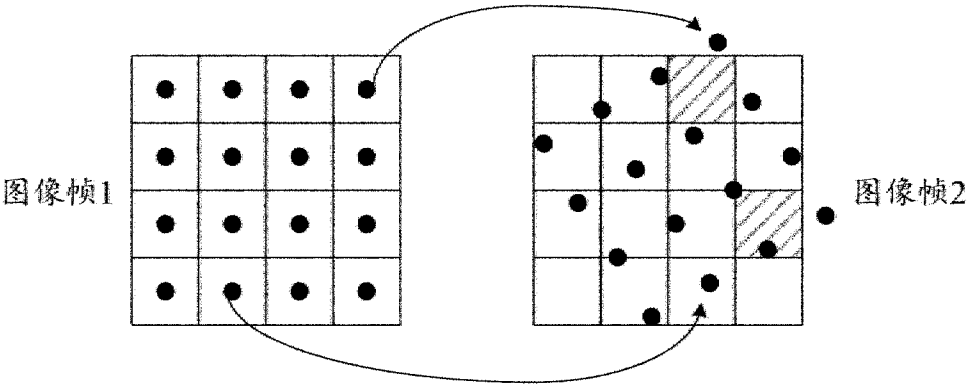
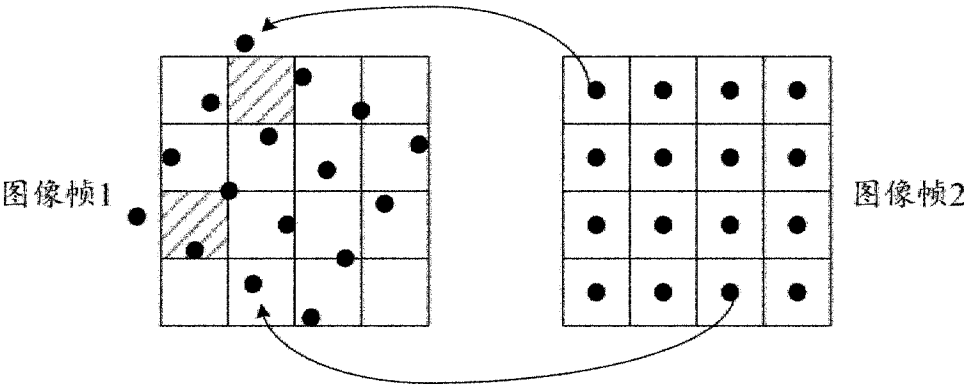


图 2



A.前向映射



B.后向映射

图 3

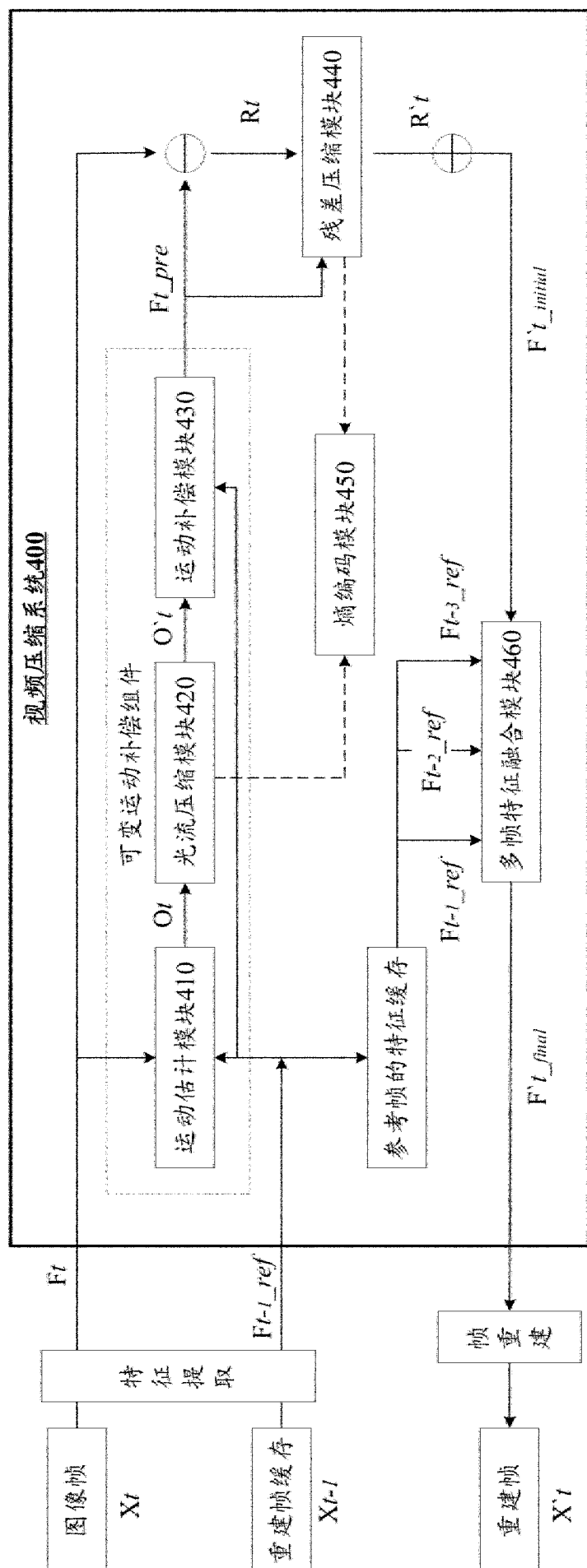


圖 4

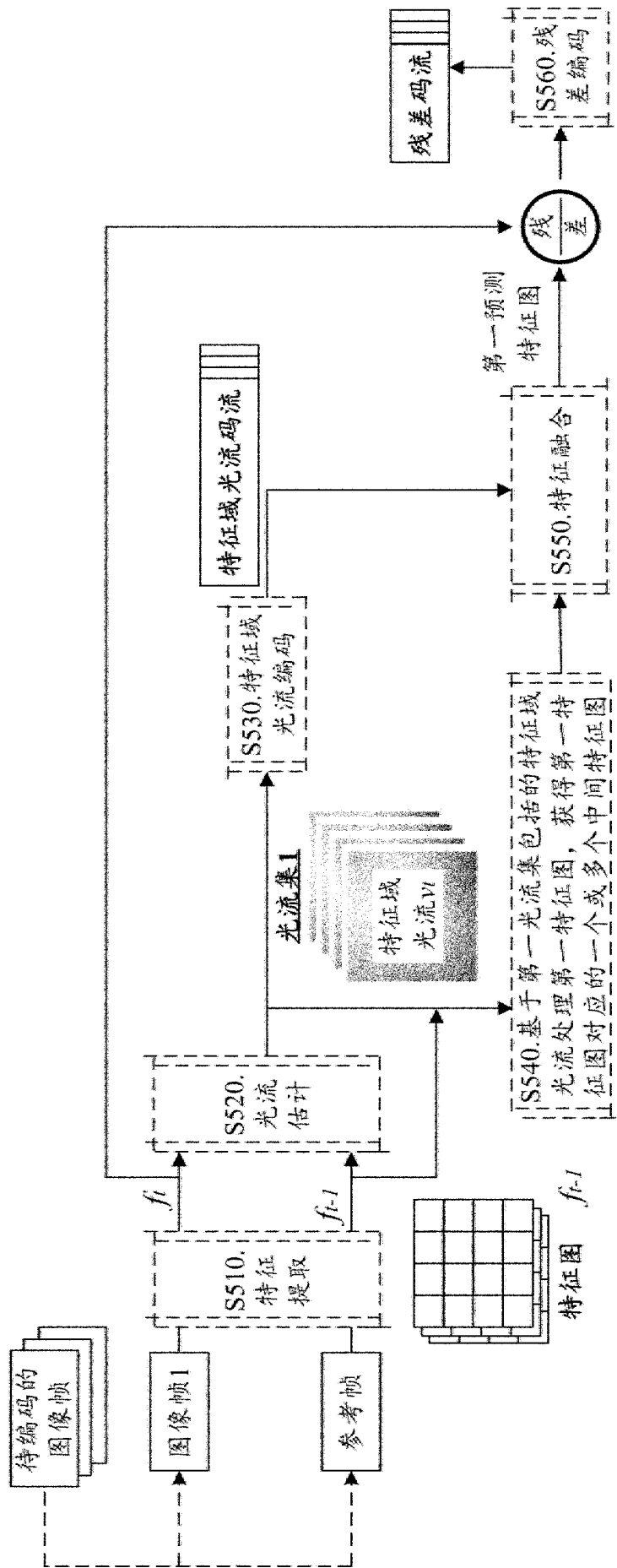


图5

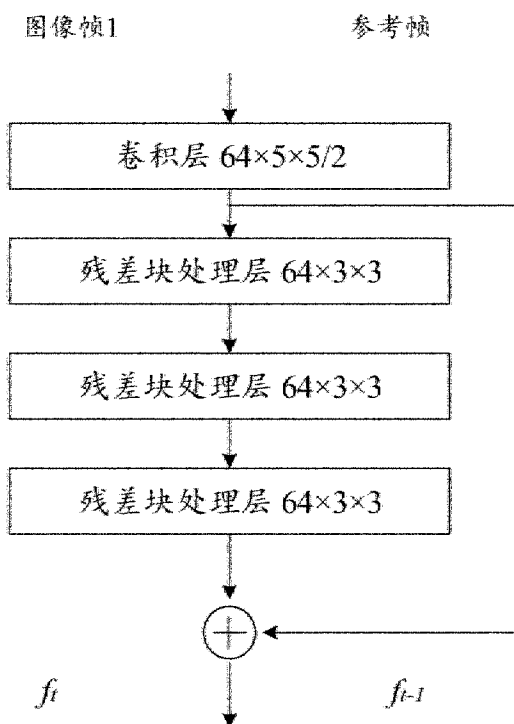


图 6

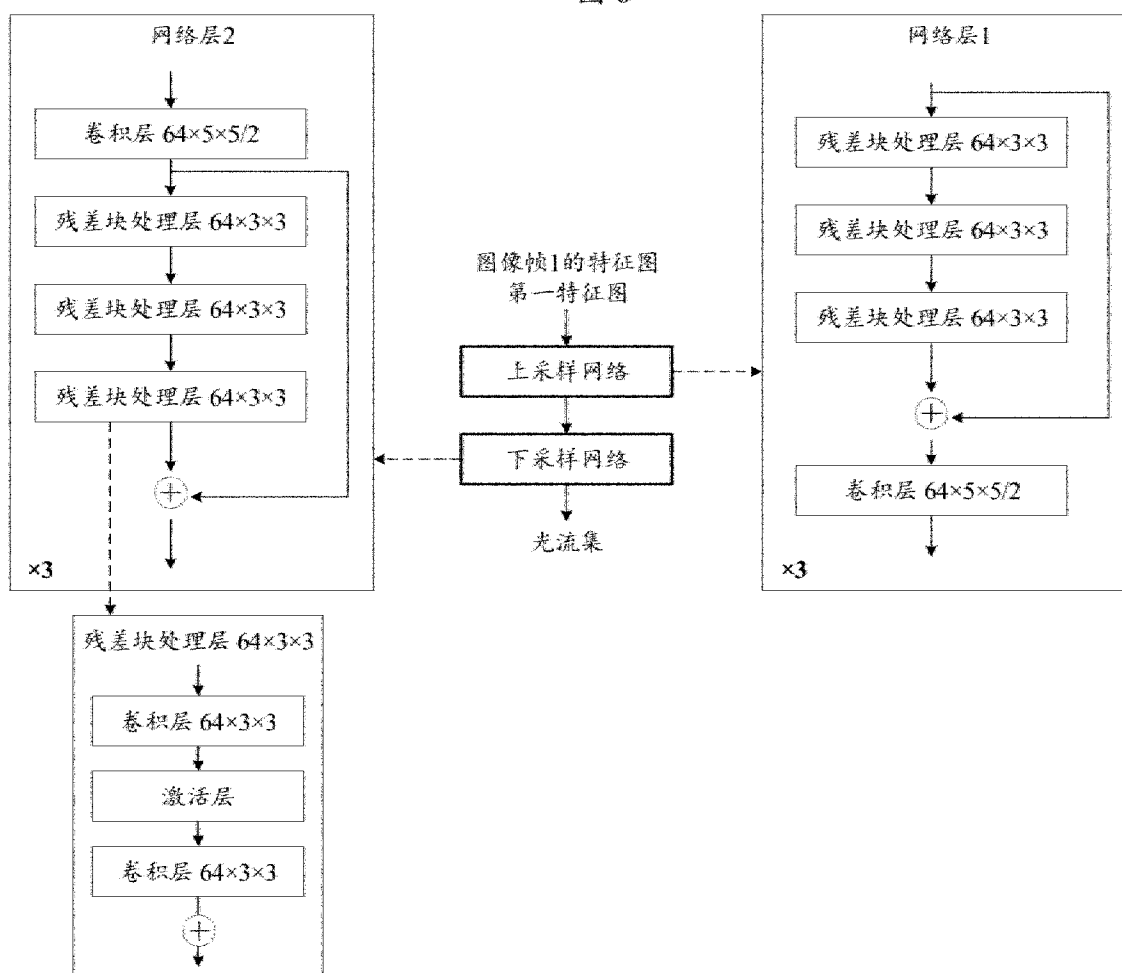


图 7

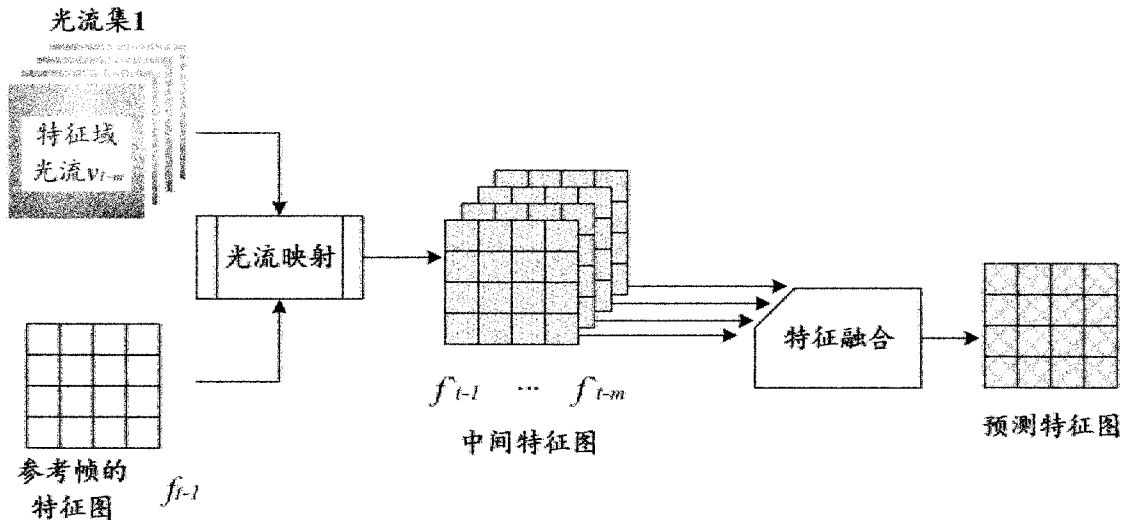


图 8

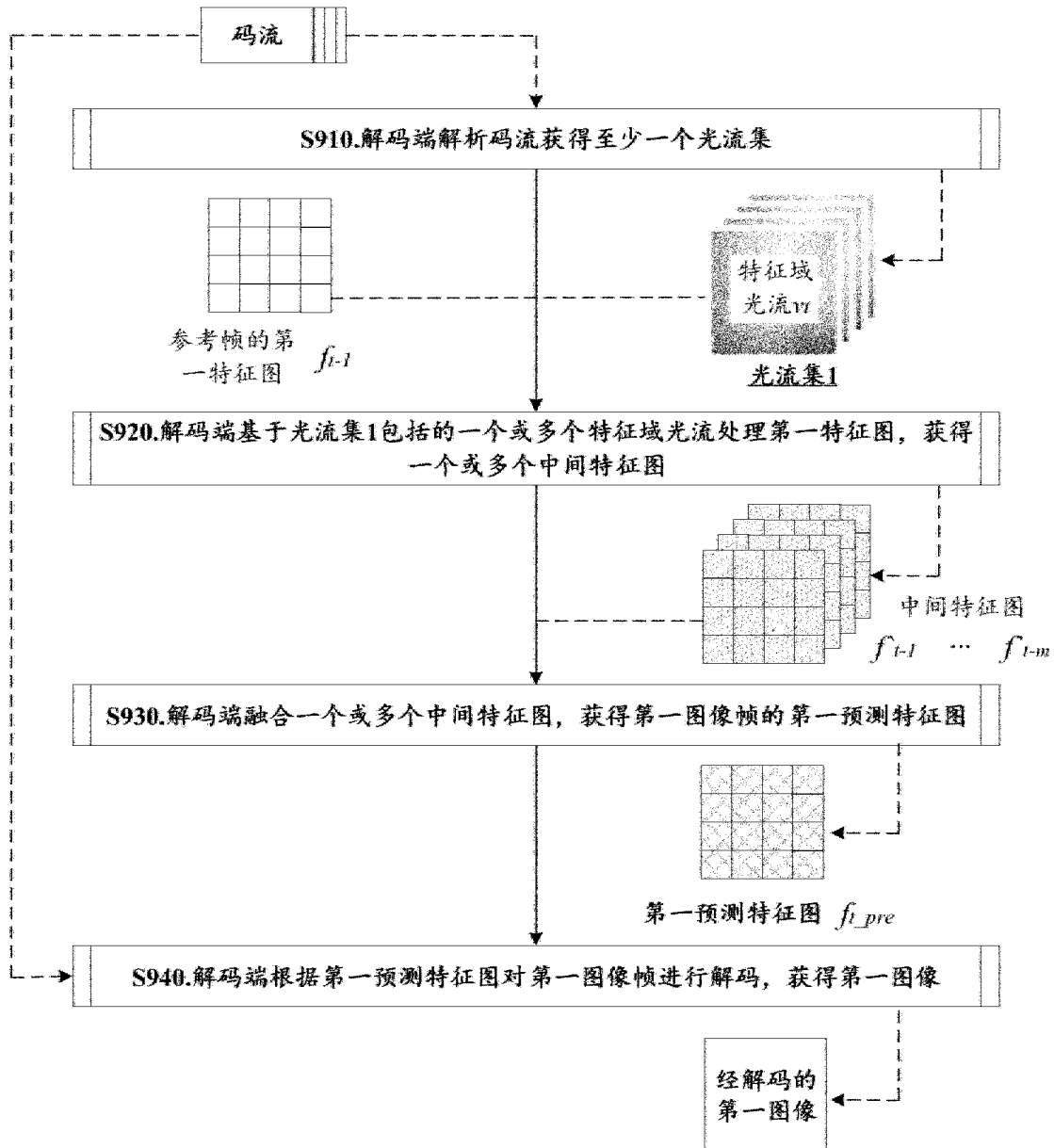


图 9

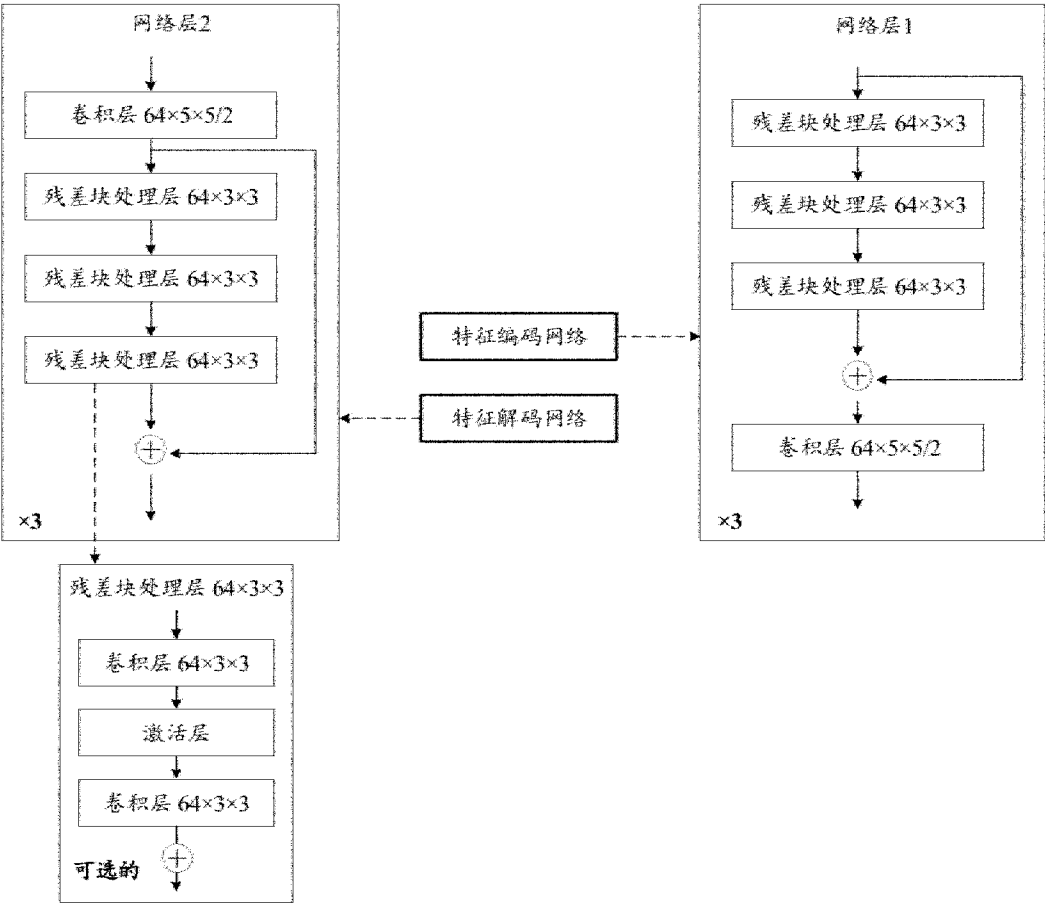


图 10

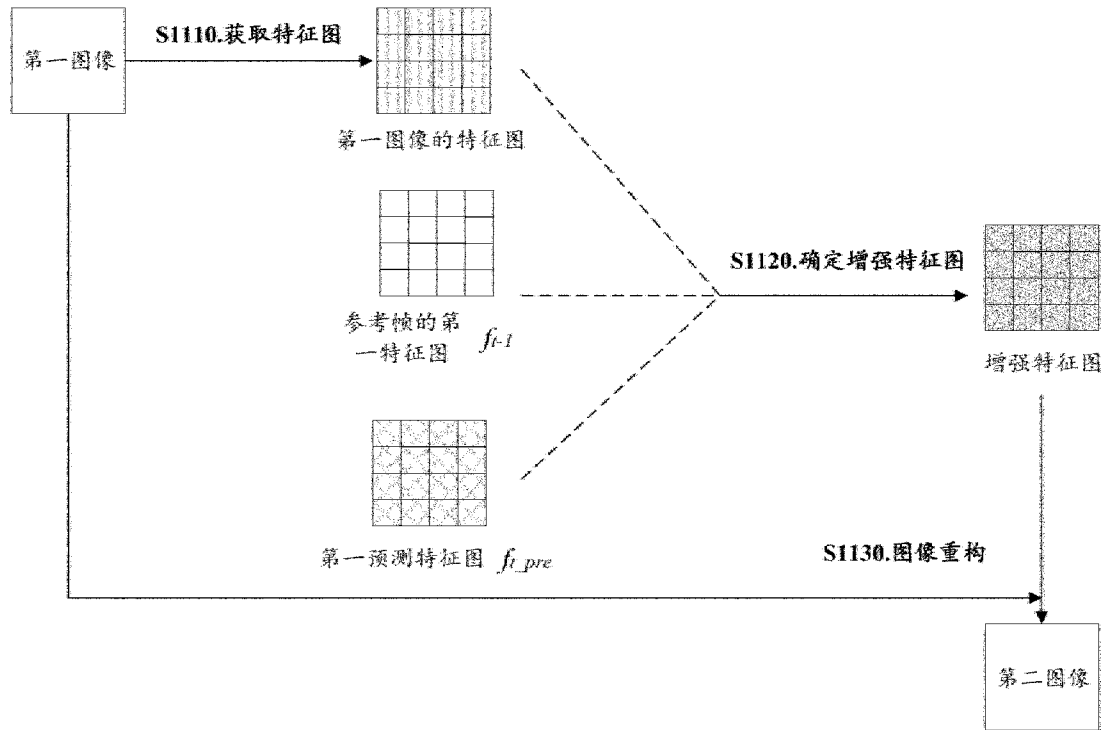


图 11

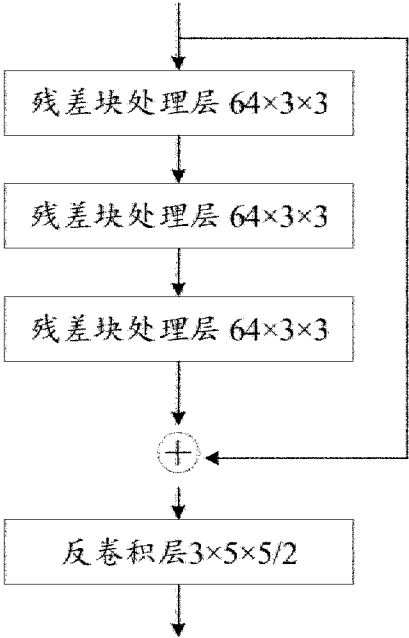


图 12

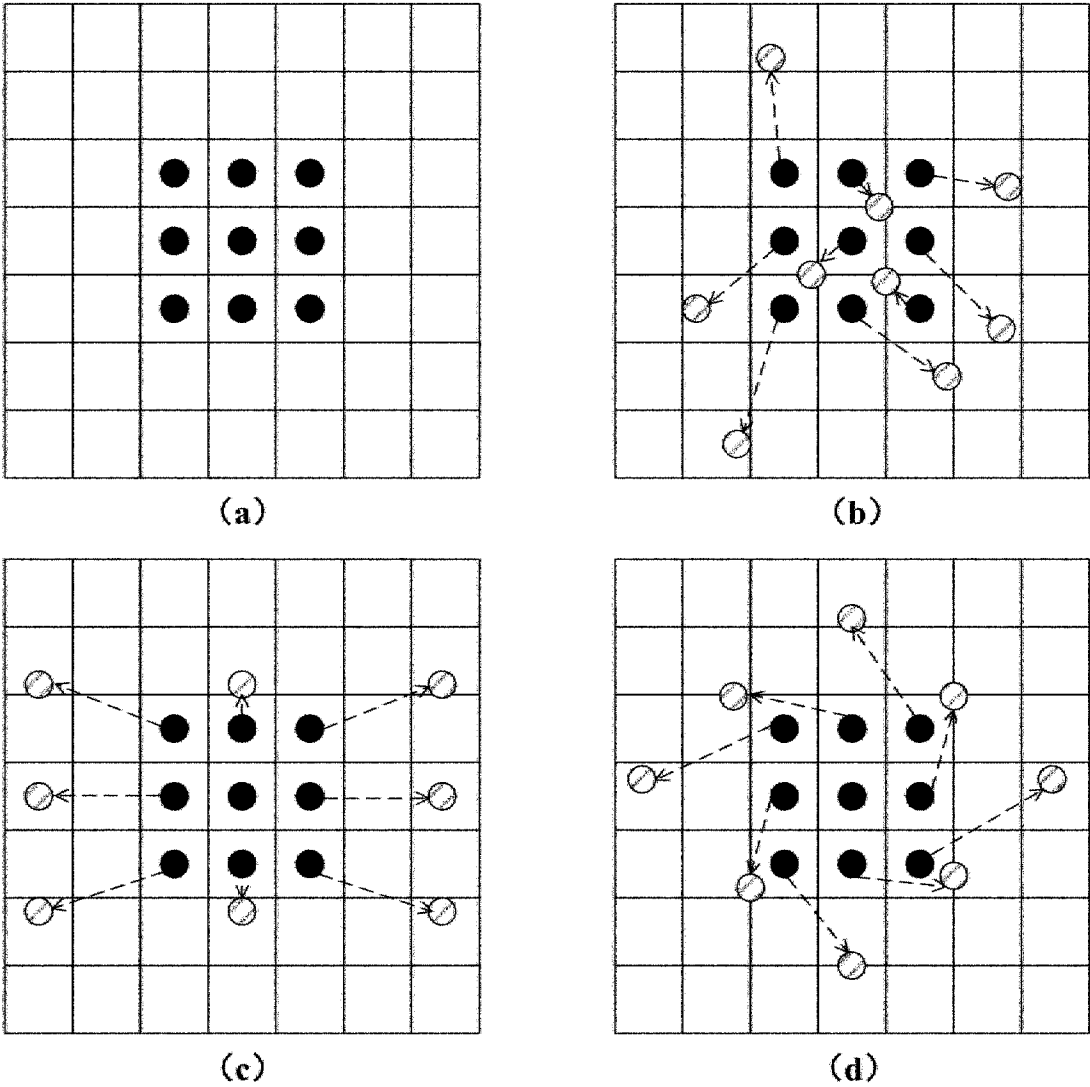


图 13

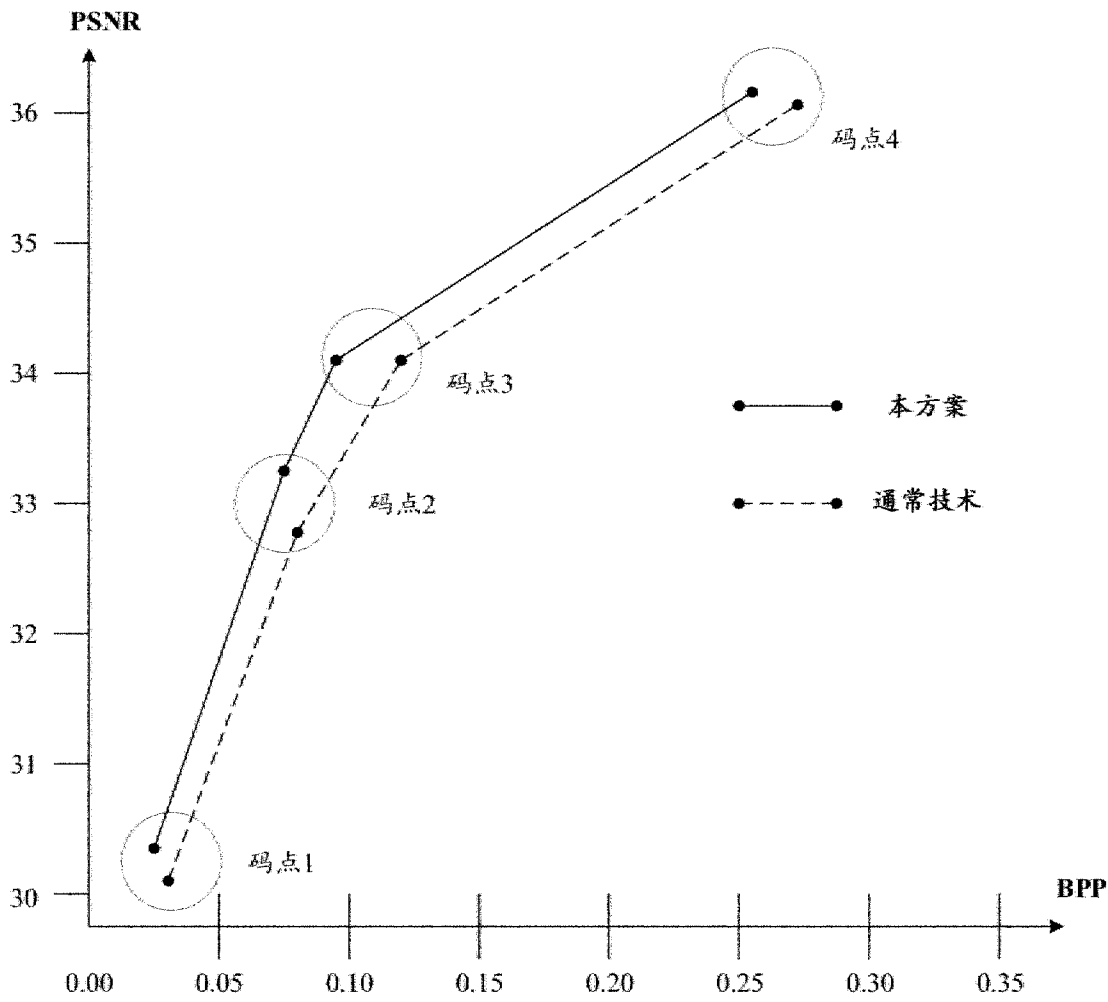


图 14

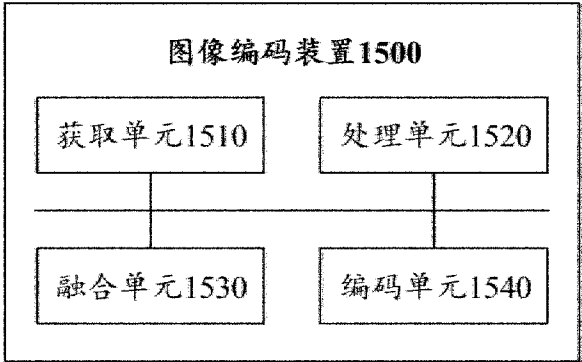


图 15

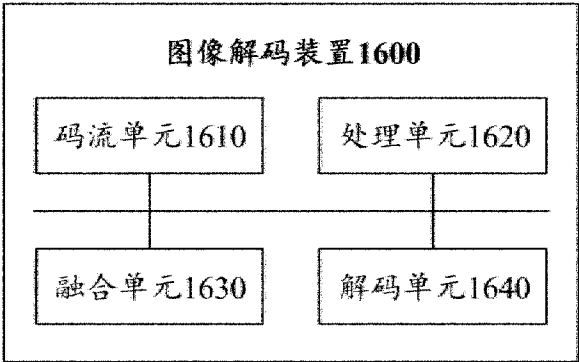


图 16

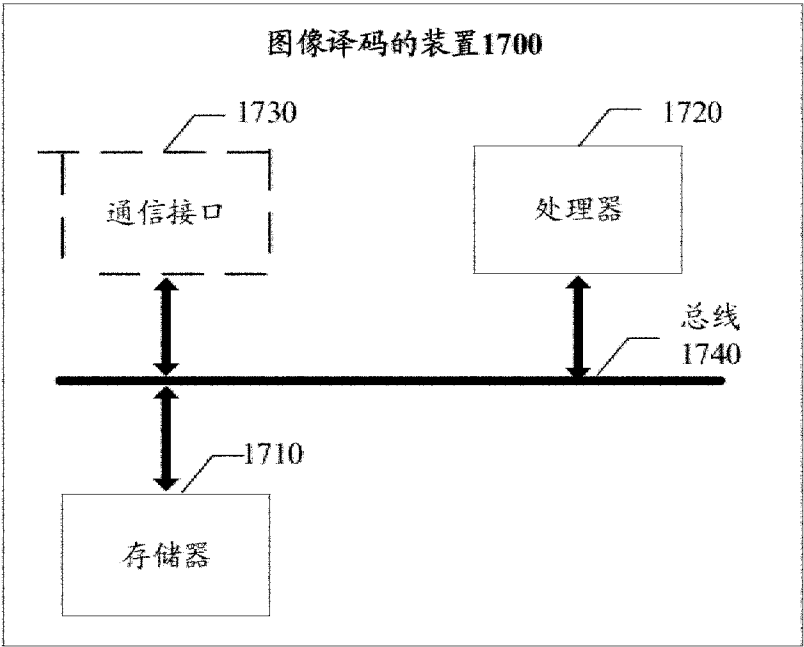


图 17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/071928

A. CLASSIFICATION OF SUBJECT MATTER

H04N19/577(2014.01)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)

IPC:H04N19

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)

CNTXT, ENTXTC, ENTXT, DWPI, CJFD, CNKI: 光流, 参考帧, 特征, 运动, 融合, 加权, 面向特征空间的深度视频压缩, warp, reference frame, feature, movement, combine, weight, FVC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 113949883 A (WUHAN TCL GROUP INDUSTRIAL RESEARCH INSTITUTE CO., LTD.) 18 January 2022 (2022-01-18) description, paragraphs 0063-0144, and figures 1-10	1, 2, 4, 5, 8, 9, 11-14, 16, 17, 20, 21, 23-30
A	CN 110913218 A (HEFEI TUCODEC INFORMATION TECHNOLOGY CO., LTD.) 24 March 2020 (2020-03-24) entire document	1-30
A	CN 113542651 A (BEIJING MEGVII TECHNOLOGY CO., LTD.) 22 October 2021 (2021-10-22) entire document	1-30
A	CN 114339219 A (ZHEJIANG DAHUA TECHNOLOGY CO., LTD.) 12 April 2022 (2022-04-12) entire document	1-30
A	US 2018098066 A1 (QUALCOMM INC.) 05 April 2018 (2018-04-05) entire document	1-30
A	US 2021314474 A1 (SAMSUNG ELECTRONICS CO., LTD.) 07 October 2021 (2021-10-07) entire document	1-30



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

“D” document cited by the applicant in the international application

“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

18 April 2023

Date of mailing of the international search report

24 April 2023

Name and mailing address of the ISA/CN

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

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/071928

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113949883	A	18 January 2022	None			
CN	110913218	A	24 March 2020	None			
CN	113542651	A	22 October 2021	None			
CN	114339219	A	12 April 2022	None			
US	2018098066	A1	05 April 2018	JP	2019535198	A	05 December 2019
				KR	20190059282	A	30 May 2019
				SG	11201901631	SA	29 April 2019
				US	10341659	B2	02 July 2019
				EP	3523966	A1	14 August 2019
				WO	2018067848	A1	12 April 2018
				TW	201826791	A	16 July 2018
				TWI	775780	B	01 September 2022
US	2021314474	A1	07 October 2021	BR	112019006882	A2	25 June 2019
				US	2023022753	A1	26 January 2023
				WO	2021201438	A1	07 October 2021
				US	11503221	B2	15 November 2022

国际检索报告

国际申请号

PCT/CN2023/071928

A. 主题的分类

H04N19/577 (2014. 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)

IPC:H04N19

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))

CNTXT, ENTXTC, ENTXT, DWPI, CJFD, CNKI: 光流, 参考帧, 特征, 运动, 融合, 加权, 面向特征空间的深度视频压缩, warp, reference frame, feature, movement, combine, weight, FVC

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
X	CN 113949883 A (武汉TCL集团工业研究院有限公司) 2022年1月18日 (2022 - 01 - 18) 说明书第0063-0144段, 附图1-10	1, 2, 4, 5, , 8, 9, 11-14, 16, 17, 20, 21, 23-30
A	CN 110913218 A (合肥图鸭信息科技有限公司) 2020年3月24日 (2020 - 03 - 24) 全文	1-30
A	CN 113542651 A (北京迈格威科技有限公司) 2021年10月22日 (2021 - 10 - 22) 全文	1-30
A	CN 114339219 A (浙江大华技术股份有限公司) 2022年4月12日 (2022 - 04 - 12) 全文	1-30
A	US 2018098066 A1 (QUALCOMM INC) 2018年4月5日 (2018 - 04 - 05) 全文	1-30
A	US 2021314474 A1 (SAMSUNG ELECTRONICS CO LTD) 2021年10月7日 (2021 - 10 - 07) 全文	1-30

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2023年4月18日

国际检索报告邮寄日期

2023年4月24日

ISA/CN的名称和邮寄地址

中国国家知识产权局

中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

王芳

电话号码 (+86) 010-62412006

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/071928

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	113949883	A	2022年1月18日	无	
CN	110913218	A	2020年3月24日	无	
CN	113542651	A	2021年10月22日	无	
CN	114339219	A	2022年4月12日	无	
US	2018098066	A1	2018年4月5日	JP 2019535198 A	2019年12月5日
				KR 20190059282 A	2019年5月30日
				SG 11201901631 SA	2019年4月29日
				US 10341659 B2	2019年7月2日
				EP 3523966 A1	2019年8月14日
				WO 2018067848 A1	2018年4月12日
				TW 201826791 A	2018年7月16日
				TWI 775780 B	2022年9月1日
				BR 112019006882 A2	2019年6月25日
US	2021314474	A1	2021年10月7日	US 2023022753 A1	2023年1月26日
				WO 2021201438 A1	2021年10月7日
				US 11503221 B2	2022年11月15日