

(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
2018 年 9 月 20 日 (20.09.2018)



(10) 国际公布号
WO 2018/166535 A1

(51) 国际专利分类号:
H04N 19/156 (2014.01) H04N 19/593 (2014.01)

(21) 国际申请号: PCT/CN2018/079378

(22) 国际申请日: 2018 年 3 月 17 日 (17.03.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710159343.6 2017年3月17日 (17.03.2017) CN

(71) 申请人: 山东科技大学(SHANDONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) [CN/CN]; 中国山东省青岛市黄岛区前湾港路 579 号, Shandong 266590 (CN)。

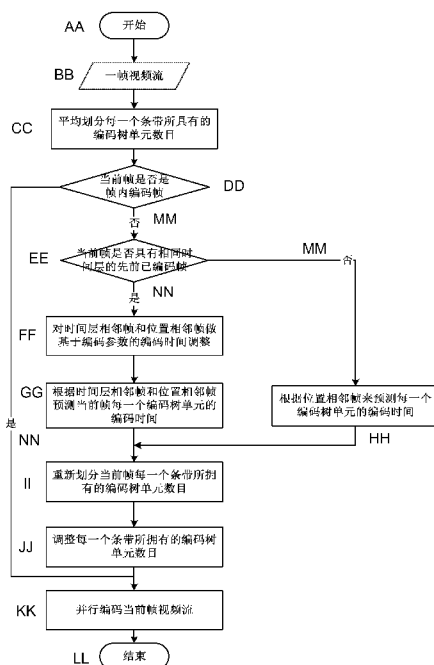
(72) 发明人: 房胜(FANG, Sheng); 中国山东省青岛市黄岛区前湾港路 579 号, Shandong 266590 (CN)。李哲(LI, Zhe); 中国山东省青岛市黄岛区前湾港路 579 号, Shandong 266590 (CN)。赵晴(ZHAO, Qing); 中国山东省青岛市黄岛区前湾港路 579 号, Shandong 266590 (CN)。赵建立(ZHAO, Jianli); 中国山东省青岛市黄岛区前湾港路 579 号, Shandong 266590 (CN)。

(74) 代理人: 青岛智地领创专利代理有限公司等(QINGDAO ZHIDILINGCHUANG PATENT AGENCY CO., LTD et al.); 中国山东省青岛市黄岛区香江路高科技创业中心 215 室联系人: 窦懋松, Shandong 266555 (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,

(54) Title: METHOD FOR LOAD BALANCING BASED ON ENCODING TIME PREDICTION MODEL

(54) 发明名称: 一种基于编码时间预测模型的负载均衡方法



AA Start
BB One frame of video stream
CC Divide the number of encoding tree units involved in each band evenly
DD Is a current frame an intra-frame encoding frame?
EE Does the current frame have a frame encoded previously of an identical time layer?
FF Adjust the encoding time of a time layer adjacent frame and a position adjacent frame based on an encoding parameter
GG According to the time layer adjacent frame and the position adjacent frame, predict the encoding time of each encoding tree unit of the current frame
HH According to the position adjacent frame, predict the encoding time of each encoding tree unit
II Re-divide the number of encoding tree units involved in each band of the current frame
JJ Adjust the number of encoding tree units involved in each band
KK Perform parallel encoding on the current frame of video stream
LL End
MM No
NN Yes

图 1

(57) Abstract: Disclosed is a method for load balancing based on an encoding time prediction model, belonging to the field of multimedia encoding. The method involves, prior to parallel encoding of a video stream, firstly predicting an encoding time of each encoding unit of a current video frame according to an encoding time prediction model, and then taking the predicted encoding time of each encoding unit of the current frame as the basis for performing a load balancing operation; load balancing mainly involves load allocation, firstly automatically determining an average load of each core according to a configuration of a current device, and then



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, 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))。

allocating a certain number of encoding units to each core according to the predicted encoding time; and finally, load balancing of an HEVC parallel encoder is realised by means of a load adjustment strategy. The method for load balancing of an HEVC parallel encoder based on an encoding time prediction model provided in the present invention can significantly increase the encoding speed while ensuring the encoding quality.

(57) 摘要: 本发明公开了一种基于编码时间预测模型的负载均衡方法, 属于多媒体编码领域, 该方法在对视频流进行并行编码之前, 首先根据编码时间预测模型预测当前视频帧每一个编码单元的编码时间, 然后以当前帧每一个编码单元的预测编码时间为依据进行负载的均衡操作; 负载均衡主要是对于负载的分配, 首先根据当前设备的配置自动确定每一个核心的平均负载, 然后根据预测的编码时间来为每一个核心分配一定的编码单元数目; 最后通过负载的调整策略实现 HEVC 并行编码器的负载均衡。本发明提出的基于编码时间预测模型的 HEVC 并行编码器的负载均衡方法能够在保证编码质量的前提下, 显著提高编码速度。

一种基于编码时间预测模型的负载均衡方法

技术领域

本发明属于多媒体编码领域，具体涉及一种基于编码时间预测模型的负载均衡方法。

背景技术

随着新一代视频编码标准 HEVC (High Efficiency Video Coding) 的发布，视频的编码效率进一步提升。相比于以前的视频编码标准 H.264/AVC，在相同编码质量的情况下，HEVC 可以节省近一半的码率，然而这是以牺牲编码复杂度为代价的。如何在不改变码流结构和保持原有码率的前提下，最大程度的降低算法复杂度，显著提高编码速度，是 HEVC 大规模推广与应用的关键所在。

为了降低 HEVC 的计算复杂度，并将其应用于实时应用，很多学者都致力于这个方向。降低计算复杂度的优化方法主要分为两种：串行优化和并行优化。

串行优化主要是针对于编码流程的优化，主要分为两类，一类是通过编码单元的快速划分和编码模式的快速选择来降低计算复杂度，一类是通过减少运动估计的时间来降低计算复杂度。这两类优化的效果是明显的，通过这些编码流程的优化，大约可以降低 HEVC 接近一半的计算复杂度。但是，随着社会的进步，人们对于视频的清晰度有了更高的要求。仅仅降低一半的计算复杂度依然满足不了人们对于超高清视频的需求。于是就出现了关于并行优化的研究。

并行优化也分为很多类，从小到大主要包括图像组级的并行、帧级的并行、条带或 tile 级的并行，以及编码树单元级的并行。图像组级和帧级的并行由于需要缓存大量的视频数据，时延大，不适应于实时应用；编码树单元级的并行由于相邻编码树单元间的相关性较强，并行编码存在很大的限制，并行效率不高。条带和 tile 都是一帧中的编码树单元的集合，没有时延和制约，且相关性限制不大。但是条带的划分较 tile 更加的灵活，且具有更好的并行性能。因此，本专利主要是针对条带级别的并行。

目前有关条带并行优化的方法主要有两种，一种是根据相邻已编码帧的编码时间来预测当前帧的编码时间，一种是建立一种归一化的计算复杂度来确定当前帧的编码复杂度。第一种方案仅仅使用了相邻已编码帧的信息，而忽略了分层时间编码结构，第二种方案建立的归一化的计算复杂度则由于图像的纹理特性和运动特性不相似而预测精度较差，不能适应于所有的视频序列。因此我们亟需找到一种能够同时使用分层时间编码结构和相邻已编码帧的方法，来更好的预测当前帧中每一个编码树单元的计算复杂度。

发明内容

为了找到一种更加有效的条带级并行的负载均衡方法，本发明提出了一种基于编码时间

预测模型的负载均衡方法，设计合理，克服了现有技术的不足，具有良好的效果。

为了实现上述目的，本发明采用如下技术方案：

一种基于编码时间预测模型的负载均衡方法，包括如下步骤：

步骤 1：进入一帧视频流；

步骤 2：平均划分当前帧中每一个条带所具有的的编码树单元数目；

步骤 3：判断当前帧是否是帧内编码帧；

若：判断结果是当前帧是帧内编码帧，则执行步骤 10；

或判断结果是当前帧不是帧内编码帧，则执行步骤 4；

步骤 4：判断当前帧是否有同一时间层的已编码非帧内编码帧；

若：判断结果是当前帧有同一时间层的已编码非帧内编码帧，则执行步骤 5；

或判断结果是当前帧没有同一时间层的已编码非帧内编码帧，则执行步骤 7；

步骤 5：通过编码时间预测模型将位置相邻已编码帧的编码时间同步到与同一时间层的相邻已编码帧相同的水平上；

步骤 6：通过同步后的两个帧的编码时间共同预测当前帧的编码时间，然后执行步骤 8；

步骤 7：根据位置相邻已编码帧的编码时间来预测当前帧的编码时间；

步骤 8：以当前帧的预测编码时间为基准来分配每一个条带所拥有的编码树单元数目；

步骤 9：根据条带的分配结果进行微调，使每一个条带获得相同的计算负载；

步骤 10：并行编码当前帧的每一个条带，结束一帧。

优选地，在步骤 5 中，具体包括如下步骤：

步骤 5.1：将位置上的相邻已编码帧做基于其在图像组中位置的调整，使其同步到与同一时间层的相邻已编码帧相同的图像组位置水平上；

步骤 5.2：将位置上的相邻已编码帧的编码时间做基于量化参数大小的调整，使其同步到与同一时间层的相邻已编码帧相同的量化参数水平上；

步骤 5.3：将位置上的相邻已编码帧的编码时间做基于参考帧数目的调整，使其同步到与同一时间层的相邻已编码帧相同的参考帧水平上。

优选地，在步骤 8 中，具体包括如下步骤：

步骤 8.1：计算当前帧的总预测编码时间；

步骤 8.2：根据每一个设备的 CPU 核心数分配条带的数目；

步骤 8.3：通过步骤 8.1 和步骤 8.2 的结果确定每一个条带的平均负载；

步骤 8.4：根据步骤 8.3 中的平均负载为每一个条带分配一定数目的编码树单元。

优选地，在步骤 9 中，具体包括如下步骤：步骤 9.1：进入一帧视频流；

步骤 9.2: 预测当前帧的总编码时间和各个编码树单元的编码时间;

步骤 9.3: 计算每一个条带的平均负载并划分每一个条带的编码树单元数目;

步骤 9.4: 计算划分后的各个条带的负载并排序,

步骤 9.5: 记录最大计算负载, 并将其赋值给一个变量 a;

步骤 9.6: 判断 a 的负载值是否大于其前或后的计算负载;

若: 判断结果是 a 的负载值大于其前或后的计算负载, 则执行步骤 9.7;

或判断结果是 a 的负载值小于其前或后的计算负载, 则执行步骤 9.11;

步骤 9.7: 将 a 的编码树单元数目减一, 比其小的条带的编码树单元数目加一;

步骤 9.8: 计算变化后的最大负载;

步骤 9.9: 判断变化后的最大计算负载是否小于之前的最大负载;

若: 判断结果为变化后的最大计算负载小于之前的最大负载, 则执行步骤 9.10;

或判断结果为变化后的最大计算负载大于之前的最大负载, 则执行步骤 9.11;

步骤 9.10: 重新排序各条带的计算负载, 然后执行步骤 9.5;

步骤 9.11: 判断 a 是否是计算负载;

若: 判断结果是 a 是计算负载, 则结束;

或判断结果是 a 不是计算负载, 则执行步骤 9.12;

步骤 9.12: 将 a 赋值为次小的计算负载, 然后执行步骤 9.6;

优选地, 步骤 5 中编码时间预测模型为:

$$T_{pred} = a * T_{temp} + b * \beta * T_{pos} * (1 + \alpha)^{ref_{temp} - ref_{pos}} * (1 - \frac{qp_{temp} - qp_{pos}}{qp_{pos}});$$

式中, T_{pred} 是最终的预测时间, T_{pos} 是位置相邻帧的编码时间, qp_{pos} 是位置相邻帧的量化参数, ref_{pos} 是位置相邻帧的参考帧数目, a 是位置相邻帧的权重; T_{temp} 是时间层相邻帧的编码时间, qp_{temp} 是时间层相邻帧的量化参数, ref_{temp} 是时间层相邻帧的参考帧数目, b 是时间层相邻帧的权重; α 是我们从统计数据计算得到的变化因子, β 是基于图像组之中编码帧的位置计算得到的比率。

本发明所带来的有益技术效果: 本发明从编码配置出发, 通过研究三种编码配置参数对于编码时间的影响程度, 提出了三种基于编码参数的调整方案, 并将其综合成为一种编码时

间预测模型。

本发明根据上述编码时间预测模型，将位置上的相邻已编码帧与同一时间层上的相邻已编码帧结合起来考虑，利用这两种已编码帧共同预测当前帧的编码时间，建立了一种基于编码时间预测模型的负载均衡方法，本方法能够在保证编码质量的前提下，很好的适应高清、超高清视频的实时压缩应用，大大提高了编码速度。

附图说明

图 1 为基于编码时间预测模型的条带并行 HEVC 编码器的负载均衡方法的流程图。

图 2 为基于编码参数的编码时间预测方法的流程图。

图 3 为负载均衡方法的具体流程图。

具体实施方式

下面结合附图和具体实施方式对本发明作进一步的说明。

为了满足高清、超高清视频图像的实时压缩应用需求，提高 HEVC 视频编码器的编码速度，本发明提出了一种基于编码预测模型的条带并行 HEVC 编码器的负载均衡方法。该方法主要包括两部分：基于编码参数的编码时间调整方法和基于编码时间预测模型的负载均衡方法，基于编码参数的编码时间调整方法涉及发明内容中的步骤（5）和步骤（6），基于编码时间预测模型的负载均衡方法涉及发明内容中的步骤（8）和步骤（9）。

该发明的主要流程如图 1 所示，首先判断当前视频流是不是帧内编码帧，是否具有同一时间层的相邻已编码帧，当满足不是帧内编码帧，且具有同一时间层的相邻已编码帧的时候，根据同一时间层的相邻已编码帧（简称为时间层相邻帧）的配置参数，将位置上的相邻已编码帧（简称为位置相邻帧）分别做基于其在图像组位置、量化参数大小和参考帧数目的调整；然后根据编码时间预测模型，用时间层相邻帧和经过调整后的位置相邻帧共同预测当前帧每一个编码树单元的编码时间和总编码时间；再根据电脑配置确定条带的个数，并据此确定每一个条带的平均负载；最后根据编码树单元的预测时间和每一个条带的平均负载来确定每一个条带的编码树单元数目，并在此基础上进行微调，从而使得每一个条带获得相同的计算负载，以达到条带并行的负载均衡。

下面展开具体说明：

1、基于编码参数的编码时间调整方法

新一代视频编码标准 HEVC 中出现了固定的编码配置模板，主要包括三种：全帧内编码帧配置（AI）、低时延配置（LP）和随机存取配置（RA）。每一种配置文件下都有相似的编码配置，但是编码配置中有很多值是可以改变的，这些变化的配置参数往往会对编码的时间和编码的效率产生影响。同时这些不同的配置参数也严重影响了将两种不同类型的帧联合起来

用于实际预测的准确度。因此，为了同时使用两种不同的已编码帧，我们需要研究这些编码参数对于编码时间的影响。

1) 图像组中的位置

通过实验统计发现，在每一个图像组中都会存在一个编码时间明显大于其余帧的帧，在改变了图像组的大小之后，结果依然是这样，这个帧就是每一个图像组中的最后一帧。通过研究发现，这一帧的编码时间严重大于此图像组中其余帧的编码时间的原因是，这一帧的时间层 ID 为 0， λ 较小。此外，研究表明在同一个视频序列中，图像组中的最后一帧的编码时间与其余帧的编码时间的比率在每一个图像组中都是相似的。为此，我们定义了一个比率 β ，计算公式如下式 (1) 所示。

$$\beta = T_{other} \div T_{GOPsize} \quad (1)$$

式中， $T_{GOPsize}$ 是图像组中最后一帧的编码时间， T_{other} 是图像组中其他帧的编码时间。通过使用这个比率，可以将图像组中不同位置水平的编码帧同步到一个位置水平上。

2) 量化参数大小

为了研究量化参数大小对于编码时间的影响水平，在 HEVC 的测试序列中选择了六个测试序列作为训练集，分别是 B 类测试序列中的 ParkScene 和 Cactus、C 类测试序列中的 BQMall 和 PartyScene 以及 D 类测试序列中的 BasketballPass 和 RaceHorses。在参考帧数目为 1，图像组大小为 4，其余配置参数为默认配置的条件下，通过改变量化参数（范围为 20-48，也就是视频编码中常用的量化参数范围）的大小来记录上述六个测试序列的编码时间。每一种情况都测试了 3 次，最后通过取平均值来作为每一帧的编码时间。

实验数据表明：编码时间与量化参数大小之间存在如下式 (2) 所示的规律。

$$T_{qp_{pred}} = T_{qp_{reference}} * (1 - \frac{QP_{pred} - QP_{reference}}{QP_{reference}}) \quad (2)$$

式中， $T_{qp_{pred}}$ 是需要预测的编码时间， $T_{qp_{reference}}$ 是已经知道并且将用于预测的时间， QP_{pred} 是需要预测的帧的量化参数， $QP_{reference}$ 是用于预测的帧的量化参数。通过使用这个规律，可以将具有不同量化参数大小的帧同步到一种量化参数水平上。

3) 参考帧数目

运动估计是 HEVC 帧间预测中最耗时的操作，当增加一个参考帧的时候，相应的也就要多做一个帧的运动估计。为了研究参考帧数目对编码时间的影响，我们依然选择 2) 中所用到的 6 个测试序列作为训练集，在保持所有配置参数都不改变的条件下，通过只改变每一帧

的参考帧的数目来测试上述每一个测试序列，并记录每一帧的编码时间，其中，参考帧的数目分别被设定为 1、2、3 和 4。在测试 4 次之后，取平均值作为最后的测试结果。

统计数据表明编码时间与参考帧数目之间的关系可以用下式 (3) 来描述。

$$T_{ref_j} = T_{ref_i} * (1 + \alpha)^{j-i} \quad (3)$$

式中， i ， j 是参考帧的数目， T_{ref_j} 是需要预测的编码时间，其拥有参考帧的数目为 j ， T_{ref_i} 是用于预测的编码时间，其拥有参考帧数目为 i ， α 是从数据中分析得出的变化因子。

通过上述成果，我们提出了一种基于编码参数的编码时间预测模型，其公式如下式 (4) 所示。

$$T_{pred} = a * T_{temp} + b * \beta * T_{pos} * (1 + \alpha)^{ref_{temp} - ref_{pos}} * (1 - \frac{qp_{temp} - qp_{pos}}{qp_{pos}}) \quad (4);$$

式中， T_{pred} 是最终的预测时间， T_{pos} 是位置相邻帧的编码时间，其量化参数是 qp_{pos} ，参考帧数目是 ref_{pos} ，权重是 a ； T_{temp} 是时间层相邻帧的编码时间，量化参数是 qp_{temp} ，参考帧数目是 ref_{temp} ，权重是 b ； α 是 3) 中的变化因子， β 是 1) 中的比率。

基于编码参数的编码时间调整方法的流程如图 2 所示。对我们所获取到的两种编码帧的编码时间，首先将两者同步到同一个图像组的位置水平上。然后对调整后的位置相邻帧的编码时间分别进行基于量化参数的调整和基于参考帧的调整。从而我们可以得到两个在统一编码水平上的帧。调整的过程描述如下：

(1) 对当前帧，我们可以获得两种已编码帧，即时间层相邻帧和位置相邻帧。首先判断这两种帧是否在同一个图像组位置水平上，如果在，则不需要调整；否则，将位置相邻帧进行如式 (1) 所示的调整。

(2) 若这两种帧不在同一量化参数水平上，我们将位置相邻帧进行如式 (2) 所示的调整；否则，不需要调整。

(3) 若这两种帧不在同一参考帧数目水平上，我们将位置相邻帧进行如式 (3) 所示的调整；否则，不需要调整。

通过上述调整，我们可以得到在同一编码参数水平的两个不同类别的已编码帧。通过使用这两种已经同步到同一编码参数水平上的已编码帧的编码时间信息，我们可以预测得到当前帧的编码时间。

2. 基于编码时间预测模型的负载均衡方法

负载均衡是一种我们能获得最高加速比的理想状态。但是在实际的编码中，由于自然图像的随机性和最优编码模式的选择问题，一帧图像的计算复杂度分布并不均匀。因此，当我们给每一个条带均匀分配编码单元数目的时候，会出现负载失衡。负载失衡是降低并行处理器系统性能的一个很重要的因素。为了实现条带水平并行的负载均衡，我们需要找到一种能够有效预测每一个编码树单元计算复杂度的方法。这样，我们就可以平均划分每一个条带的计算负载，使得每一个编码核心同时完成。通过这样做，并行处理系统就不会存在空闲等待现象，我们也将获得最高的加速比。

本发明利用前面的基于编码参数的编码时间预测方法来得到当前帧中每一个编码单元的预测编码时间，进而使用这些信息来实现条带水平并行的负载均衡。负载分配和调整的流程如图3所示，具体实施方式如下：

(1) 首先根据上述编码时间预测方法来得到当前帧的总编码时间和每一个编码单元的编码时间。

(2) 根据所用设备的配置自动确定条带的数目。

(3) 根据条带的数目和总预测编码时间计算出每一个条带的平均负载。

(4) 根据每一个条带的平均负载和所预测出来的每一个编码单元的预测编码时间，来确定每一个条带所应该具有的编码单元的数目。

(5) 根据每一个条带划分的编码单元的数目，计算划分后的各个条带的计算负载并排序，记录最大负载，将其赋值给一个变量 a 。

(6) 将变量 a 与其前后相邻的条带的负载相比，且向比其小的条带转移一个编码树单元；

(7) 若变化后的最大负载小于之前的最大计算负载，或变化后的最小负载大于之前的最小负载，则确定此新的划分，并更新条带计算负载的排序，将最大负载赋值给变量 a ；执行步骤(6)；否则保持原划分；

(8) 若负载 a 不为最小，则将其赋值为次大计算负载，继续执行步骤(6)；否则开始编码。

当然，上述说明并非是对本发明的限制，本发明也并不仅限于上述举例，本技术领域的技术人员在本发明的实质范围内所做出的变化、改型、添加或替换，也应属于本发明的保护范围。

权 利 要 求 书

1、一种基于编码时间预测模型的负载均衡方法，其特征在于，包括如下步骤：

步骤 1：进入一帧视频流；

步骤 2：平均划分当前帧中每一个条带所具有的编码树单元数目；

步骤 3：判断当前帧是否是帧内编码帧；

若：判断结果是当前帧是帧内编码帧，则执行步骤 10；

或判断结果是当前帧不是帧内编码帧，则执行步骤 4；

步骤 4：判断当前帧是否有同一时间层的已编码非帧内编码帧；

若：判断结果是当前帧有同一时间层的已编码非帧内编码帧，则执行步骤 5；

或判断结果是当前帧没有同一时间层的已编码非帧内编码帧，则执行步骤 7；

步骤 5：通过编码时间预测模型将位置相邻已编码帧的编码时间同步到与同一时间层的相邻已编码帧相同的水平上；

步骤 6：通过同步后的两个帧的编码时间共同预测当前帧的编码时间，然后执行步骤 8；

步骤 7：根据位置相邻已编码帧的编码时间来预测当前帧的编码时间；

步骤 8：以当前帧的预测编码时间为基准来分配每一个条带所拥有的编码树单元数目；

步骤 9：根据条带的分配结果进行微调，使每一个条带获得相同的计算负载；

步骤 10：并行编码当前帧的每一个条带，结束一帧。

2、根据权利要求 1 所述的基于编码时间预测模型的负载均衡方法，其特征在于，在步骤 5 中，具体包括如下步骤：

步骤 5.1：将位置上的相邻已编码帧做基于其在图像组中位置的调整，使其同步到与同一时间层的相邻已编码帧相同的图像组位置水平上；

步骤 5.2：将位置上的相邻已编码帧的编码时间做基于量化参数大小的调整，使其同步到与同一时间层的相邻已编码帧相同的量化参数水平上；

步骤 5.3：将位置上的相邻已编码帧的编码时间做基于参考帧数目的调整，使其同步到与同一时间层的相邻已编码帧相同的参考帧水平上。

3、根据权利要求 1 所述的基于编码时间预测模型的负载均衡方法，其特征在于，在步骤 8 中，具体包括如下步骤：

步骤 8.1：计算当前帧的总预测编码时间；

步骤 8.2：根据每一个设备的 CPU 核心数分配条带的数目；

步骤 8.3：通过步骤 8.1 和步骤 8.2 的结果确定每一个条带的平均负载；

步骤 8.4：根据步骤 8.3 中的平均负载为每一个条带分配一定数目的编码树单元。

4、根据权利要求 1 所述的基于编码时间预测模型的负载均衡方法，其特征在于，在步骤

9 中, 具体包括如下步骤:

步骤 9.1: 进入一帧视频流;

步骤 9.2: 预测当前帧的总编码时间和各个编码树单元的编码时间;

步骤 9.3: 计算每一个条带的平均负载并划分每一个条带的编码树单元数目;

步骤 9.4: 计算划分后的各个条带的负载并排序;

步骤 9.5: 记录最大计算负载, 并将其赋值给一个变量 a ;

步骤 9.6: 判断 a 的负载值是否大于其前或后的计算负载;

若: 判断结果是 a 的负载值大于其前或后的计算负载, 则执行步骤 9.7;

或判断结果是 a 的负载值小于其前或后的计算负载, 则执行步骤 9.11;

步骤 9.7: 将 a 的编码树单元数目减一, 比其小的条带的编码树单元数目加一;

步骤 9.8: 计算变化后的最大计算负载;

步骤 9.9: 判断变化后的最大计算负载是否小于之前的最大计算负载;

若: 判断结果为变化后的最大计算负载小于之前的最大计算负载, 则执行步骤 9.10;

或判断结果为变化后的最大计算负载大于之前的最大计算负载, 则执行步骤 9.11;

步骤 9.10: 重新排序各条带的计算负载, 然后执行步骤 9.5;

步骤 9.11: 判断 a 是否是最低负载;

若: 判断结果是 a 是最低负载, 则结束;

或判断结果是 a 不是最低负载, 则执行步骤 9.12;

步骤 9.12: 将 a 赋值为次小的计算负载, 然后执行步骤 9.6。

5、根据权利要求 1 所述的基于编码时间预测模型的负载均衡方法, 其特征在于, 步骤 5 中编码时间预测模型为:

$$T_{pred} = a * T_{temp} + b * \beta * T_{pos} * (1 + \alpha)^{ref_{temp} - ref_{pos}} * (1 - \frac{qp_{temp} - qp_{pos}}{qp_{pos}});$$

式中, T_{pred} 是最终的预测时间, T_{pos} 是位置相邻帧的编码时间, qp_{pos} 是位置相邻帧的量化参数, ref_{pos} 是位置相邻帧的参考帧数目, a 是位置相邻帧的权重; T_{temp} 是时间层相邻帧的编码时间, qp_{temp} 是时间层相邻帧的量化参数, ref_{temp} 是时间层相邻帧的参考帧数目, b 是时间层相邻帧的权重; α 是我们从统计数据计算得到的变化因子, β 是基于图像组之中编码帧的位置计算得到的比率。

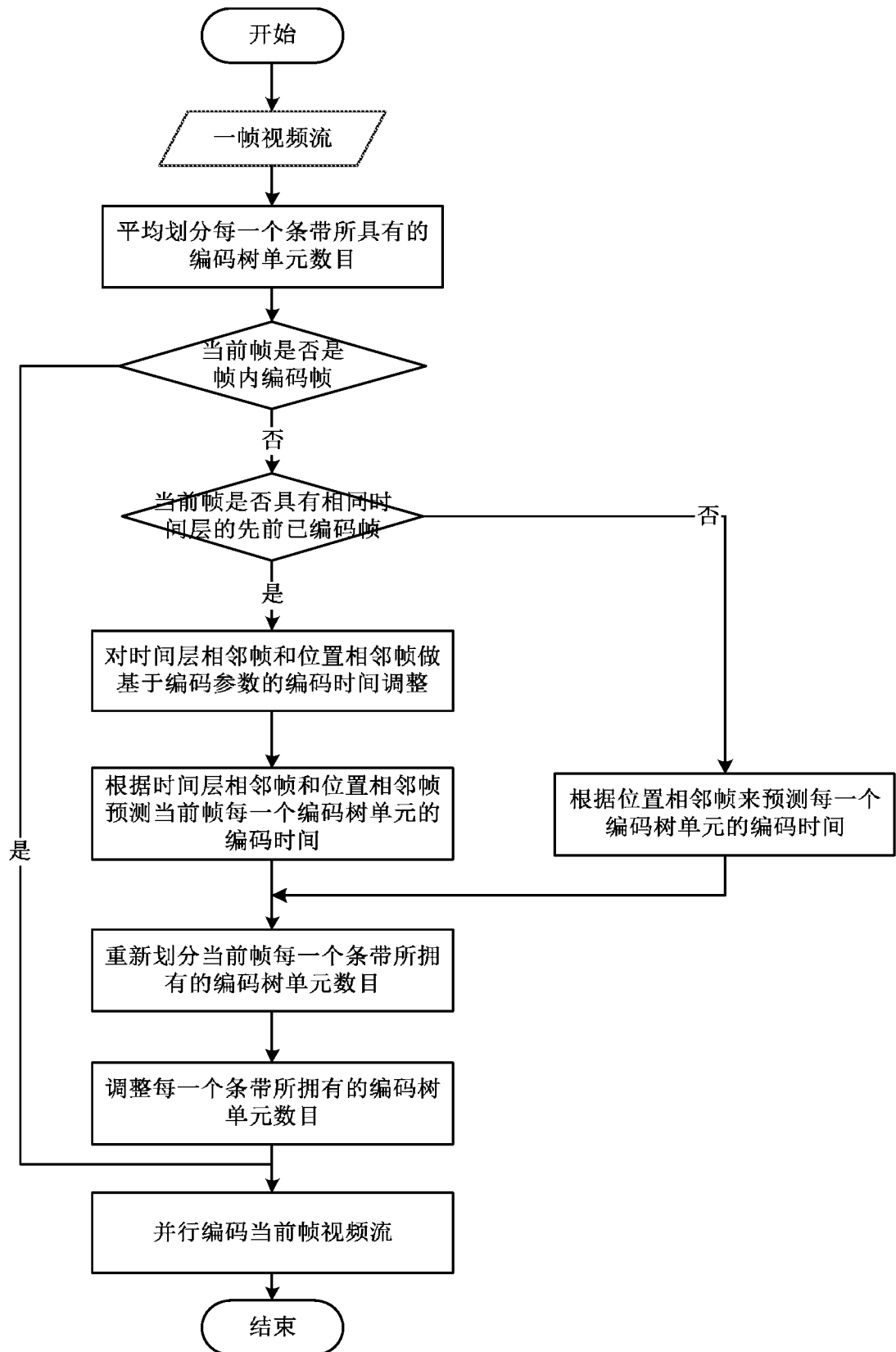


图 1

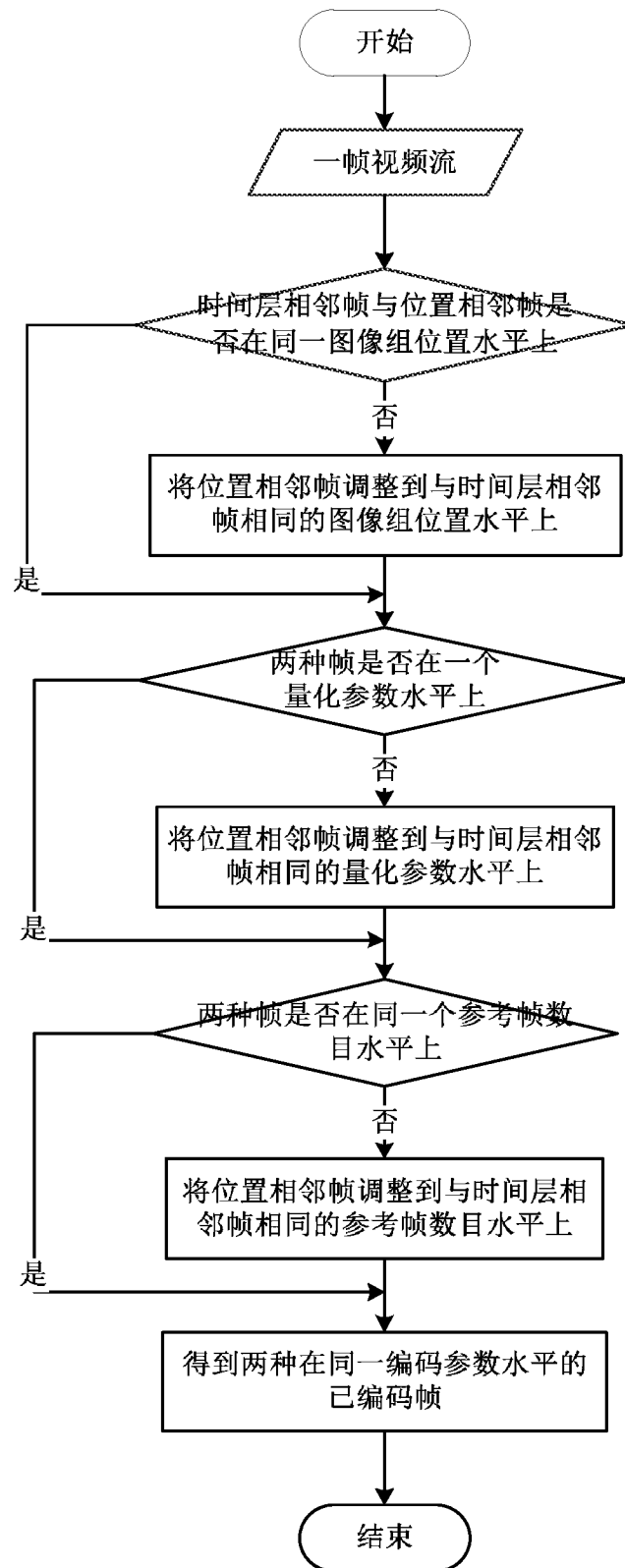


图 2

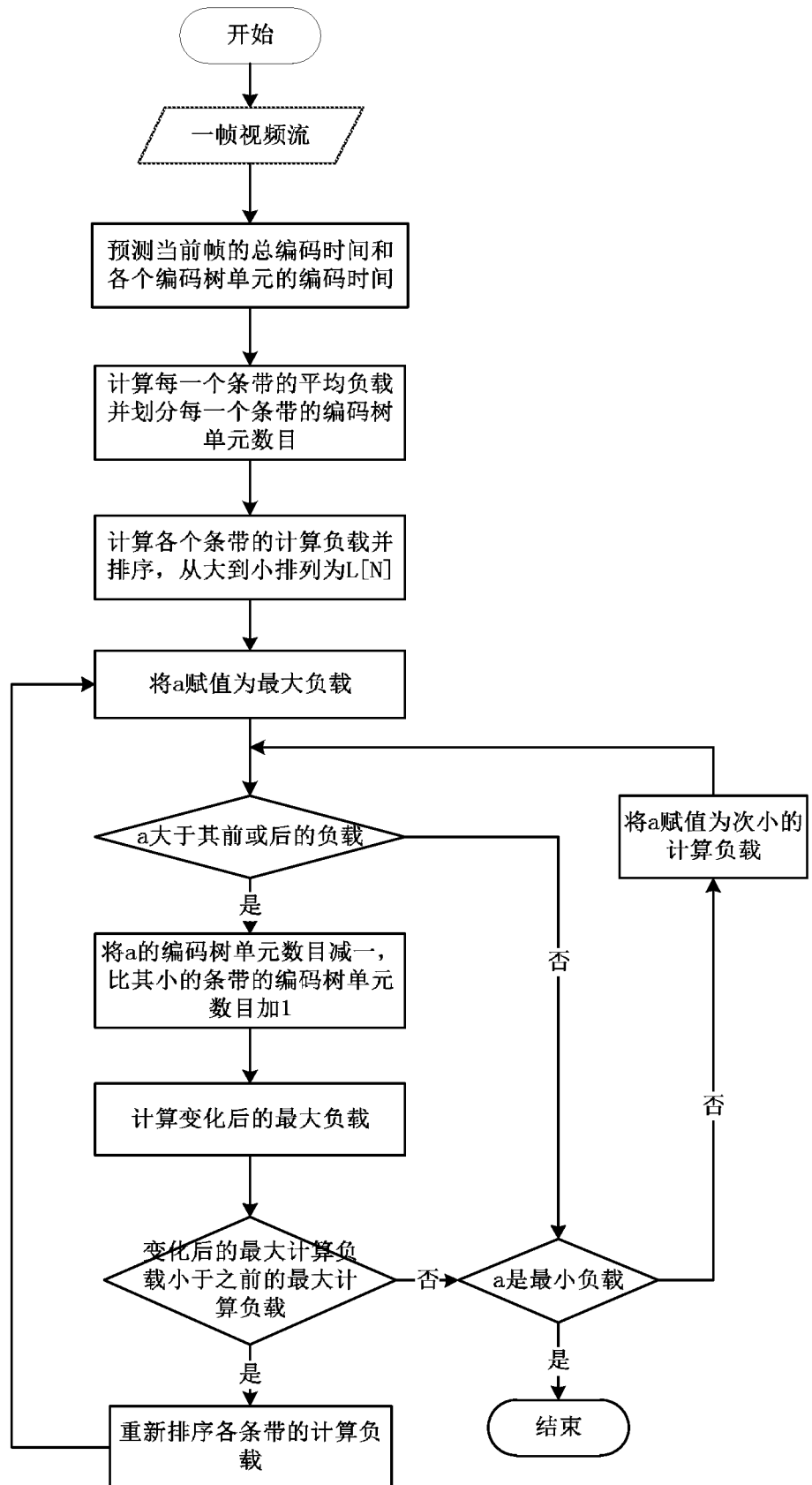


图 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/079378

A. CLASSIFICATION OF SUBJECT MATTER

H04N 19/156 (2014.01) i; H04N 19/593 (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)

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)

CNTEXT; CNABS; CNKI: 编码, 解码, 时间, 负载均衡, 优化, 复杂度, 条带, 编码树单元, 划分, 并行, 帧内;

DWPI; SIPOABS: encode, code, time, temporal, load balance, optimize, complexity, slice, code tree unit, CTU, split, divide, parallel, intra

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107071424 A (SHANDONG UNIVERSITY OF SCIENCE AND TECHNOLOGY), 18 August 2017 (18.08.2017), claims 1-5	1-5
A	CN 105245896 A (CHUANXIAN NETWORK TECHNOLOGY (SHANGHAI) CO., LTD.), 13 January 2016 (13.01.2016), entire document	1-5
A	CN 104539972 A (CHINA SECURITY & FIRE TECHNOLOGY CO., LTD. et al.), 22 April 2015 (22.04.2015), entire document	1-5
A	CN 103702126 A (GRADUATE SCHOOL AT SHENZHEN, TSINGHUA UNIVERSITY), 02 April 2014 (02.04.2014), entire document	1-5
A	CN 106454354 A (SUN YAT-SEN UNIVERSITY), 22 February 2017 (22.02.2017), entire document	1-5
A	US 2015016503 A1 (QUALCOMM INC.), 15 January 2015 (15.01.2015), entire document	1-5

☐ 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	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“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
“O” document referring to an oral disclosure, use, exhibition or other means	“&” document member of the same patent family
“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 May 2018	Date of mailing of the international search report 06 June 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 ZHANG, Zhihua Telephone No. 86-(010)-62089470

International application No.
PCT/CN2018/079378

Form PCT/ISA/210 (patent family annex) (January 2015)

国际检索报告

国际申请号

PCT/CN2018/079378

A. 主题的分类

H04N 19/156(2014.01)i; H04N 19/593(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)

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))

CNTXT;CNABS;CNKI: 编码, 解码, 时间, 负载均衡, 优化, 复杂度, 条带, 编码树单元, 划分, 并行, 帧内 DWPI;
SIPOABS: encode, code, time, temporal, load balance, optimize, complexity, slice, code tree unit, CTU,
split, divide, parallel, intra

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
PX	CN 107071424 A (山东科技大学) 2017年 8月 18日 (2017 - 08 - 18) 权利要求1-5	1-5
A	CN 105245896 A (传线网络科技上海有限公司) 2016年 1月 13日 (2016 - 01 - 13) 全文	1-5
A	CN 104539972 A (中安消技术有限公司 等) 2015年 4月 22日 (2015 - 04 - 22) 全文	1-5
A	CN 103702126 A (清华大学深圳研究生院) 2014年 4月 2日 (2014 - 04 - 02) 全文	1-5
A	CN 106454354 A (中山大学) 2017年 2月 22日 (2017 - 02 - 22) 全文	1-5
A	US 2015016503 A1 (QUALCOMM INC.) 2015年 1月 15日 (2015 - 01 - 15) 全文	1-5

☐ 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 5月 19日

国际检索报告邮寄日期

2018年 6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

张志华

电话号码 86-(010)-62089470

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/079378

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	107071424	A	2017年 8月 18日	无			
CN	105245896	A	2016年 1月 13日	无			
CN	104539972	A	2015年 4月 22日	无			
CN	103702126	A	2014年 4月 2日	CN	103702126	B	2017年 1月 18日
CN	106454354	A	2017年 2月 22日	无			
US	2015016503	A1	2015年 1月 15日	TW	201515440	A	2015年 4月 16日
				WO	2015009712	A1	2015年 1月 22日

表 PCT/ISA/210 (同族专利附件) (2015年1月)