(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
2017 年 9 月 28 日 (28.09.2017)

WIPO | PCT

(10) 国际公布号
WO 2017/162068 A1

(51) 国际专利分类号:
H04N 19/40 (2014.01)

(21) 国际申请号: PCT/CN2017/076547

(22) 国际申请日: 2017 年 3 月 14 日 (14.03.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201610179243.5 2016 年 3 月 25 日 (25.03.2016) CN

(71) 申请人: **阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED)** [—/CN]; 开曼群岛英属开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: **徐浩晖 (XU, Haohui)**; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **梅大为 (MEI, Dawei)**; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **周昌 (ZHOU, Chang)**; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: **北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.)**; 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, 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)。

[见续页]

(54) **Title**: VIDEO TRANSCODING METHOD, DEVICE, AND SYSTEM

(54) **发明名称** : 一种视频转码方法、装置和系统

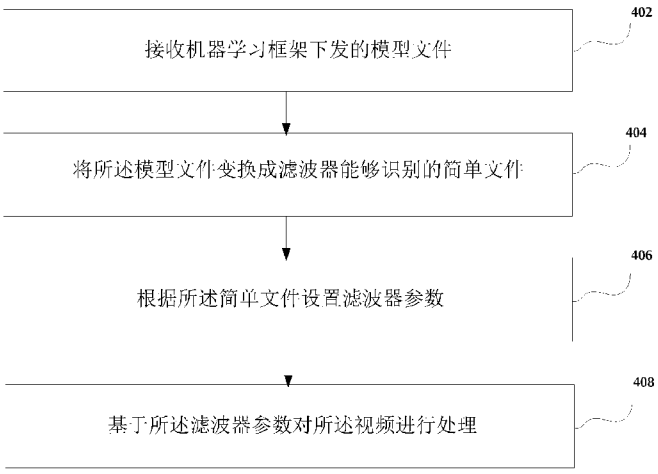


图 4

401 MODEL FILE ISSUED BY MACHINE-LEARNING FRAMEWORK IS RECEIVED
404 MODEL FILE IS CONVERTED TO SIMPLE FILE THAT CAN BE RECOGNIZED BY FILTER
406 FILTER PARAMETER IS SET ACCORDING TO SIMPLE FILE
408 VIDEO IS PROCESSED ON BASIS OF FILTER PARAMETER

(57) **Abstract**: Provided are a video transcoding method, device, and system. The method comprises: receiving a model file issued by a machine-learning framework; converting said model file to a simple file that can be recognized by a filter; according to said simple file, setting a filter parameter; on the basis of said filter parameter, processing the video. No extra storage consumption is added during the process of video transcoding, and system resource consumption is reduced to the greatest possible extent while not requiring the addition of any extra system deployment costs.

(57) **摘要**: 提供一种视频转码方法、装置和系统。所述方法包括：接收机器学习框架下发的模型文件；将所述模型文件变换成滤波器能够识别的简单文件；根据所述简单文件设置滤波器参数；基于所述滤波器参数对所述视频进行处理。在进行视频转码过程中不会增加额外存储消耗，并可以在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

WO 2017/162068 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种视频转码方法、装置和系统

本申请要求2016年03月25日递交的申请号为201610179243.5、发明名称为“一种视频转码方法、装置和系统”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，特别是涉及一种视频转码方法、装置和系统。

背景技术

随着互联网技术的迅猛发展，人们对互联网视频服务的需求也日益增加。而视频转码几乎是一切互联网视频服务的基础。然而，由于视频拍摄、版权和年代等原因，部分视频需要进一步加工，例如去噪、超分辨率等，才能获得更好的观看质量。因此，经过多年的发展，常规的视频图像处理技术已被应用。

由于近几年机器学习算法的兴起，利用机器学习尤其是对图像进行复杂处理已经在理论上证明可以取得更佳的处理效果。然而，现有机器学习在图像处理方面还仅仅局限在单张图片处理的阶段，而无法在视频中使用。分析其原因在于，现有机器学习框架本身并不支持视频的输入输出，而从图片到视频图像在处理流程上的复杂度并不只是简单的帧数增加。

因此，如何更高效地利用机器学习的优势直接对视频进行复杂处理，成为亟需本领域技术人员解决的技术问题。

发明内容

鉴于上述问题，提出了本申请实施例以便提供一种克服上述问题或者至少部分地解决上述问题的一种视频转码方法、装置及系统。

本申请公开了一种视频转码方法，包括：

接收机器学习框架下发的模型文件；

将所述模型文件变换成滤波器能够识别的简单文件；

根据所述简单文件设置滤波器参数；

基于所述滤波器参数对所述视频进行处理。

相应的，本申请还公开了一种视频转码装置，包括：

模型接收模块，用于接收机器学习框架下发的模型文件；

文件变换模块，用于将所述模型文件变换成滤波器能够识别的简单文件；

参数设置模块，用于根据所述文件变换模块变换后的简单文件，设置滤波器参数；

视频处理模块，用于基于所述参数设置模块设置的滤波器参数对所述视频进行处理。

此外，本申请还公开了一种视频转码系统，包括：

机器学习框架，用于对训练样本进行训练得到模型文件；

视频转码装置包括：

模型接收模块，用于接收所述机器学习框架下发的模型文件；

文件变换模块，用于将所述模型文件变换成滤波器能够识别的简单文件；

参数设置模块，用于根据所述文件变换模块变换后的简单文件，设置滤波器参数；

视频处理模块，用于基于所述参数设置模块设置的滤波器参数对所述视频进行处理。

根据本申请提供的具体实施例，本申请公开了以下技术效果：

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

此外，本申请实施例基于用户在云端视频服务中心通过应用程序接口提出的个性化视频处理需求，提供差异化的视频处理服务，并且可以将第三方机器学习框架训练的模板文件高效的集成到现有的视频转码系统中，将视频直接嵌入到转码系统的解码和编码流程之中实现视频处理，从而更加高效、灵活的实现视频转码的一体化处理；也更加适应日后更多业务场景的扩展，利于系统升级。

当然，实施本申请的任一产品并不一定需要同时达到以上所述的所有优点。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例中所需

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的前提下，还可以根据这些附图获得其他的附图。

图 1 是一种集成机器学习框架的视频处理示意图；

图 2 是本申请集成机器学习框架的视频处理示意图；

图 3 是本申请视频转码系统中视频滤波处理示意图；

图 4 是本申请的一种视频转码方法实施例的步骤流程图；

图 5 是本申请的另一种视频转码方法实施例的步骤流程图；

图 6 是本申请的一种视频转码装置实施例的结构框图；

图 7 是本申请的一种视频转码系统实施例的结构框图。

具体实施方式

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，所描述的实施例仅仅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所获得的所有其他实施例，都属于本申请保护的范围。

为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本申请作进一步详细的说明。

为了方便理解本发明实施例，首先在此介绍本发明实施例描述中会涉及的几个要素：

A、视频转码服务（Multimedia Transcoding Service，MTS）是指将已经压缩编码的视频码流转换成另一个视频码流，以适应不同的网络带宽、不同的终端处理能力和不同的用户需求。转码本质上是一个先解码，再编码的过程，因此转换前后的码流可以遵循相同的视频编码标准，也可以不遵循相同的视频编码标准。

B、机器学习：机器学习是一门人工智能的科学，主要研究如何在经验学习中改善具体算法的性能。在实际应用中，通常指通过不断尝试指定任务，让机器自动优化出更好地完成该任务的过程。

C、机器学习框架：机器学习框架是指用于机器学习的算法、库和/或软件的集合，通常体现为开源的模型、项目、产品等，可以在服务器和/客户端上运行。

D、机器学习的训练过程：机器学习的训练通常指根据特定算法，通过让机器不断尝试特定任务并得到反馈，从而使其掌握特定规律和模式，并提高机器在将来执行该任

务的效果的过程。一般是通过大量样本，例如文本、图像或视频等计算出一个模型文件。

E、预测过程：指经过学习训练后的机器根据得到的模型执行特定任务的过程。预测过程中遇到的一般是训练过程中不包含的新样本，即使用机器学习的人真正希望机器去执行处理的样本。

F、模型文件：是机器学习通过反复尝试特定任务所总结出的规律和模式的记录。机器在预测时通过该记录进行特定任务的处理。

G、视频增强：指除简单裁剪和坐标变换外其他图像级别的操作，包括去噪、去抖动、超分辨率等处理。

H、超分辨率：指从低分辨率图像重构出高分辨率图像的技术。由于源图像（即低分辨率）比目标图像（即高分辨率）所含的信息少，通常需要额外的信息才能恢复缺失的细节，简单滤波和机器学习方法都可以在一定程度上补充这些缺失的细节。

如图 1 所示，提供一种利用卷积网络进行图像的超分辨率重构方法。该技术的机器学习和预测都建立在 Torch 库，并且只支持图片处理。而对于视频的处理，则需要先将视频每帧转换成 png 图片，基于机器学习训练生成的模型将所有图片经过预测过程后，再分别进行超分辨率处理得到新的图片，最后将所有处理后的图片序列编码形成新的视频。

然而，上述方法对视频进行超分辨率处理时还存在很多缺陷。首先，对于视频处理，虽然基于转码系统而言其输入和输出都是视频，但由于中间的处理过程集成了机器学习框架，因此受到机器学习框架的局限必须将输入视频转换成图片才能进行处理。而对于视频增强处理而言，显然中间的转换图片的过程完全是冗余的，并且，将源视频转化成图片会带来大量额外的存储消耗，视频处理效率非常低。再有，机器学习和转换两个过程的耦合性很强，并且在实际视频处理时，从视频输入到最终视频输出必须经过机器学习框架和原有转码系统这两套系统，这样，会因用到第三方的机器学习框架而给原有转码系统的集成和部署带来诸多不便，不仅会增加系统部署成本，还会增加系统资源消耗。

本申请实施例可以在现有转码系统基础上集成第三方机器学习框架实现视频转码功能。如图 2 所示，是本申请实施例集成机器学习框架的视频处理示意图。本申请实施例可以基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对收集到的训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无

需额外增加系统部署成本的同时尽量减少系统资源消耗。

进一步的，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，如图 3 所示，是本申请视频转码系统中视频增强滤波处理示意图。视频输入转码系统后，依次经过解码、滤波、编码等一系列处理，最后再将增强处理后的视频输出，也就是将视频直接嵌入到转码系统的解码和编码流程之中实现视频处理。因此，与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗，并可以高效的实现集成机器学习的视频转码处理功能。

此外，本申请实施例可以基于用户在云端视频服务中心通过应用程序接口提出的个性化视频处理需求，提供差异化的视频处理服务，并且可以将第三方机器学习框架训练的模板文件高效的集成到现有的视频转码系统中，从而更加高效、灵活的实现视频转码的一体化处理；也更加适应日后更多业务场景的扩展，利于系统升级。

实施例一

参照图 4，示出了本申请的一种视频转码方法实施例的步骤流程图，具体可以包括如下步骤：

步骤 402、接收机器学习框架下发的模型文件；

优选的，所述机器学习框架可以采用 Theano 库、Caffe 库或 Torch 库中的任一个对样本进行训练从而生成模板文件。

这里需要说明的是，模型文件可以通过现有的各种机器学习框架训练得到，本申请实施例对具体采用何种机器学习框架并不做限制。

此外，模型文件是机器学习框架离线训练得到的，该过程与实际视频处理过程是完全独立的，这样，可以根据不同的需要分别通过最合适的机器学习框架生成模板文件，而无需担心将学习框架整合到转码系统的问题。

步骤 404、将所述模型文件变换成滤波器能够识别的简单文件；

优选的，在本申请的另一个实施例中，可以将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

这里需要说明的是，在实际应用中，只需要将从机器学习框架输出的模型文件变换成转码系统滤波器接受的简单文件即可，因此所述的变换只需要支持从机器学习框架的

输出模型文件到转码系统滤波器支持的简单文件的单向变换，而无需考虑其反过程，从而减少参数格式设计的复杂度和实现模型变换过程的工作量。

步骤 406、根据所述简单文件设置滤波器参数；

需要说明的是，由于机器学习框架每次下发的模型文件不同，绝大多数情况下，只需设计一次滤波器格式，包括滤波器实现，而对采用的每个机器学习框架实现一个模型文件变换即可。而滤波器参数是需要针对每次变换得到的简单文件重新设置。

步骤 408、基于所述滤波器参数对所述视频进行处理。

优选的，在本申请的另一个实施例中，所述基于所述滤波器参数对所述视频进行处理的步骤可以进一步包括：

对输入视频进行解码处理；

基于所述滤波器参数对解码后的视频进行滤波处理；

对滤波处理后的视频进行编码处理后输出。

需要说明的是，在实际应用中，原有的转码处理流程不会改变，待处理的视频先经过解码器解码后，每帧图像都会在预设的滤波器中进行视频增强处理，处理后的图像再送到编码器中进行编码输出。

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

实施例二

参照图 5，示出了本申请的另一种视频转码方法实施例的步骤流程图，具体可以包括如下步骤：

步骤 500、通过机器学习框架对训练样本进行训练得到模型文件。

优选的，所述通过机器学习框架对训练样本进行训练得到模型文件的步骤还可以进一步包括：

S 510、基于视频处理需求收集训练样本；

优选的，基于用户对视频增强的具体需求，例如超分辨率、去噪等。所述训练样本可以是收集到的大量图片，例如数万张图片等。

S 520、对所述训练样本进行离线训练得到模型文件。

优选的，所述机器学习框架可以采用 Theano 库、Caffe 库或 Torch 库中的任一个对样本进行训练从而生成模板文件。这里需要说明的是，模型文件可以通过现有的各种机器学习框架训练得到，本申请实施例对具体采用何种机器学习框架并不做限制。

此外，模型文件是机器学习框架离线训练得到的，该过程与实际视频处理过程是完全独立的，这样，可以根据不同的需要分别通过最合适的机器学习框架生成模板文件，而无需担心将学习框架整合到转码系统的问题。

步骤 502、接收机器学习框架下发的模型文件；

需要说明的是，由于模型文件是机器学习框架训练得到的，因此在实际进行视频处理时，只需要从机器学习框架接收离线训练得出的模型文件即可。

步骤 504、将所述模型文件变换成滤波器能够识别的简单文件；

优选的，在本申请的另一个实施例中，可以将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

这里需要说明的是，在实际应用中，只需要将从机器学习框架输出的模型文件变换成转码系统滤波器接受的简单文件即可，因此所述的变换只需要支持从机器学习框架的输出模型文件到转码系统滤波器支持的简单文件的单向变换，而无需考虑其反过程，从而减少参数格式设计的复杂度和实现模型变换过程的工作量。

步骤 506、根据所述简单文件设置滤波器参数；

需要说明的是，由于机器学习框架每次下发的模型文件不同，绝大多数情况下，只需设计一次滤波器格式，包括滤波器实现，而对采用的每个机器学习框架实现一个模型文件变换即可。而滤波器参数是需要针对每次变换得到的简单文件重新设置。

步骤 508、基于所述滤波器参数对所述视频进行处理。

优选的，在本申请的另一个实施例中，所述基于所述滤波器参数对所述视频进行处理的步骤可以进一步包括：

对输入视频进行解码处理；

基于所述滤波器参数对解码后的视频进行滤波处理；

对滤波处理后的视频进行编码处理后输出。

需要说明的是，在实际应用中，原有的转码处理流程不会改变，待处理的视频先经过解码器解码后，每帧图像都会在预设的滤波器中进行视频增强处理，处理后的图像再送到编码器中进行编码输出。

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

此外，本申请实施例基于用户在云端视频服务中心通过应用程序接口提出的个性化视频处理需求，提供差异化的视频处理服务，并且可以将第三方机器学习框架训练的模板文件高效的集成到现有的视频转码系统中，将视频直接嵌入到转码系统的解码和编码流程之中实现视频处理，从而更加高效、灵活的实现视频转码的一体化处理；也更加适应日后更多业务场景的扩展，利于系统升级。

实施例三

下面，结合一实际业务场景来说明本申请实施例的一种具体应用。

在 MTS 上实现视频转高清业务的功能，在具体视频处理时，例如，用户欲对某部动画片进行高清重制，则可以将对该部动画片视频的增强需求通过 MTS 在线提出，例如对该视频进行超分辨率处理，也就是将该动画片转为高清视频。

基于用户的需求在收集到数万张图片后，机器学习框架运行特定的算法对数万张图片进行离线训练得到对应的模型文件。由于训练过程与实际视频增强处理过程是完全独立的，因此训练机器可以直接部署在实验环境中。

在本申请实施例的具体应用中，该模型文件通过采用基于 Lua 语言的 Torch 机器学

习框架训练得到。此处需要说明的是，现有常用的机器学习框架有很多，例如采用 Theano、Caffe 或 Torch 库的机器学习框架都可以在本申请实施例中应用，这样就可以根据不同的需要分别选择最合适的框架，而无需担心将机器学习框架整合到转码系统的问题。

MTS 接收到 Torch 机器学习框架训练得到的二进制模型文件，将其变换成转码滤波器支持或识别的简单文件。这里，MTS 采用的是 Ffmpeg 工具进行滤波器处理，因此需要将二进制模型文件变换成简单的文本文件。这里需要说明的是，在实际应用中，只需要将从机器学习框架输出的模型文件变换成转码系统中滤波器能够识别的简单文件即可，因此所述的变换只需要支持从机器学习框架的输出模型文件到转码系统滤波器支持的简单文件的单向转换，而无需考虑其反过程，从而减少参数格式设计的复杂度和实现模型变换过程的工作量。需要说明的是，由于机器学习框架每次下发的模型文件不同，绝大多数情况下，只需设计一次滤波器格式，包括滤波器实现，而对采用的每个机器学习框架实现一个模型文件变换即可。而滤波器参数是需要针对每次变换得到的简单文件重新设置。

MTS 对源视频进行解码处理，然后通过滤波器对每帧视频进行滤波处理，即完成超分辨率处理过程，然后再对处理后的视频进行编码处理，最后输出增强后的视频。

这里需要说明的是，在 MTS 上实现视频转高清业务时，Ffmpeg 工具中的滤波器的格式只需要部署一次，后面即便接收到的模型文件有所改动，也只需重新将更新后的模型文件下发到转码系统即可完成更新升级。

因此，本申请实施例具备以下优点：

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

此外，本申请实施例基于用户在云端视频服务中心通过应用程序接口提出的个性化

视频处理需求，提供差异化的视频处理服务，并且可以将第三方机器学习框架训练的模板文件高效的集成到现有的视频转码系统中，将视频直接嵌入到转码系统的解码和编码流程之中实现视频处理，从而更加高效、灵活的实现视频转码的一体化处理；也更加适应日后更多业务场景的扩展，利于系统升级。

需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请实施例所必须的。

实施例四

参照图 6，示出了本申请的一种视频转码装置实施例的结构框图，具体可以包括如下模块：

模型接收模块 610，用于接收机器学习框架下发的模型文件；

优选的，所述机器学习框架可以采用 Theano 库、Caffe 库或 Torch 库中的任一个对样本进行训练从而生成模板文件。

这里需要说明的是，模型文件可以通过现有的各种机器学习框架训练得到，本申请实施例对具体采用何种机器学习框架并不做限制。

此外，模型文件是机器学习框架离线训练得到的，该过程与实际视频处理过程是完全独立的，这样，可以根据不同的需要分别通过最合适的机器学习框架生成模板文件，而无需担心将学习框架整合到视频转码装置的问题。

文件变换模块 620，用于将所述模型文件变换成滤波器能够识别的简单文件；

优选的，在本申请的另一个实施例中，可以将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

这里需要说明的是，在实际应用中，只需要将从机器学习框架输出的模型文件变换成转码系统滤波器接受的简单文件即可，因此所述的变换只需要支持从机器学习框架的输出模型文件到转码系统滤波器支持的简单文件的单向变换，而无需考虑其反过程，从而减少参数格式设计的复杂度和实现模型变换过程的工作量。

参数设置模块 630，用于根据所述文件变换模块 620 变换后的简单文件，设置滤波

器参数；

需要说明的是，由于机器学习框架每次下发的模型文件不同，绝大多数情况下，只需设计一次滤波器格式，包括滤波器实现，而对采用的每个机器学习框架实现一个模型文件变换即可。而滤波器参数是需要针对每次变换得到的简单文件重新设置。

视频处理模块 640，用于基于所述参数设置模块 630 设置的滤波器参数对所述视频进行处理。

优选的，在本申请的另一个实施例中，所述视频处理模块 640 可以包括：

解码单元 641，用于对输入视频进行解码处理；

滤波单元 642，用于基于所述滤波器参数对所述解码单元解码后的视频进行滤波处理；

编码单元 643，用于所述滤波单元滤波处理后的视频进行编码处理后输出。

需要说明的是，在实际应用中，原有的转码处理方式不会改变，待处理的视频先经过解码器解码后，每帧图像都会在预设的滤波器中进行视频增强处理，处理后的图像再送到编码器中进行编码输出。

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

实施例五

参照图 7，示出了本申请的另一种视频转码系统实施例的结构框图，具体可以包括：

机器学习框架 700，用于对训练样本进行训练得到模型文件；

优选的，在本申请的另一个实施例中，所述机器学习框架 700 还可以进一步包括：

收集模块 701，用于基于视频处理需求收集训练样本；

训练模块 702，用于对所述收集模块收集到的训练样本进行离线训练得到模型文件。

优选的，所述机器学习框架可以采用 Theano 库、Caffe 库或 Torch 库中的任一个对样本进行训练从而生成模板文件。这里需要说明的是，模型文件可以通过现有的各种机器学习框架训练得到，本申请实施例对具体采用何种机器学习框架并不做限制。

此外，模型文件是机器学习框架离线训练得到的，该过程与实际视频处理过程是完全独立的，这样，可以根据不同的需要分别通过最合适的机器学习框架生成模板文件，而无需担心将学习框架整合到转码系统的问题。

视频转码装置包括：

模型接收模块 710，用于接收所述机器学习框架下发的模型文件；

需要说明的是，由于模型文件是机器学习框架训练得到的，因此在实际进行视频处理时，只需要从机器学习框架接收离线训练得出的模型文件即可。

文件变换模块 720，用于将所述模型文件变换成滤波器能够识别的简单文件；

优选的，在本申请的另一个实施例中，可以将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

这里需要说明的是，在实际应用中，只需要将从机器学习框架输出的模型文件变换成转码系统滤波器接受的简单文件即可，因此所述的变换只需要支持从机器学习框架的输出模型文件到转码系统滤波器支持的简单文件的单向变换，而无需考虑其反过程，从而减少参数格式设计的复杂度和实现模型变换过程的工作量。

参数设置模块 730，用于根据所述文件变换模块 720 变换后的简单文件，设置滤波器参数；

需要说明的是，由于机器学习框架每次下发的模型文件不同，绝大多数情况下，只需设计一次滤波器格式，包括滤波器实现，而对采用的每个机器学习框架实现一个模型文件变换即可。而滤波器参数是需要针对每次变换得到的简单文件重新设置。

视频处理模块 740，用于基于所述参数设置模块 730 设置的滤波器参数对所述视频进行处理。

优选的，在本申请的另一个实施例中，所述视频处理模块 740 可以进一步包括：

解码单元，用于对输入视频进行解码处理；

滤波单元，用于基于所述滤波器参数对所述解码单元解码后的视频进行滤波处理；

编码单元，用于所述滤波单元滤波处理后的视频进行编码处理后输出。

需要说明的是，在实际应用中，原有的转码处理流程不会改变，待处理的视频先经

过解码器解码后，每帧图像都会在预设的滤波器中进行视频增强处理，处理后的图像再送到编码器中进行编码输出。

本申请实施例，通过转码系统将机器学习框架下发的模型文件变换成转码滤波器能够识别的简单文件，根据该简单文件设置滤波器参数，再基于该滤波器参数对视频进行滤波处理，从而可以高效的实现集成机器学习的视频处理功能。与现有技术相比，本申请实施例在处理过程中无需受到机器学习框架的限制将视频转化成图片，而是直接对视频进行处理，因此在进行视频转码过程中不会增加额外的存储消耗。

进一步的，本申请实施例基于用户的视频处理需求收集训练样本，再通过对应的机器学习框架对该训练样本进行离线训练得到模型文件。因此，本申请实施例通过将机器学习和实际视频处理两个过程进行独立设置，而无需将第三方机器学习框架部署在转码系统上，从而在无需额外增加系统部署成本的同时尽量减少系统资源消耗。

此外，本申请实施例基于用户在云端视频服务中心通过应用程序接口提出的个性化视频处理需求，提供差异化的视频处理服务，并且可以将第三方机器学习框架训练的模板文件高效的集成到现有的视频转码系统中，将视频直接嵌入到转码系统的解码和编码流程之中实现视频处理，从而更加高效、灵活的实现视频转码的一体化处理；也更加适应日后更多业务场景的扩展，利于系统升级。

对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。

本领域内的技术人员应明白，本申请实施例的实施例可提供为方法、装置、或计算机程序产品。因此，本申请实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD@ROM、光学存储器等)上实施的计算机程序产品的形式。

在一个典型的配置中，所述计算机设备包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM) 和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存

是计算机可读介质的示例。计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD@ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非持续性的电脑可读媒体(transitory media)，如调制的数据信号和载波。

本申请实施例是参照根据本申请实施例的方法、终端设备(系统)、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上，使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本申请实施例的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例做出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本申请实施例范围的所有变更和修改。

最后，还需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何

其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者终端设备中还存在另外的相同要素。

以上对本申请所提供的一种视频转码方法、装置及系统，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均可有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权 利 要 求 书

1. 一种视频转码方法，其特征在于，包括：

接收机器学习框架下发的模型文件；

将所述模型文件变换成滤波器能够识别的简单文件；

根据所述简单文件设置滤波器参数；

基于所述滤波器参数对所述视频进行处理。

2. 根据权利要求 1 所述的方法，其特征在于，所述接收机器学习框架下发的模型文件的步骤之前还包括：

通过机器学习框架对训练样本进行训练得到模型文件。

3. 根据权利要求 2 所述的方法，其特征在于，所述通过机器学习框架对训练样本进行训练得到模型文件的步骤进一步包括：

基于视频处理需求收集训练样本；

对所述训练样本进行离线训练得到模型文件。

4. 根据权利要求 1 所述的方法，其特征在于，所述将所述模型文件变换成滤波器能够识别的简单文件的步骤包括：

将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

5. 根据权利要求 1 所述的方法，其特征在于，所述基于所述滤波器参数对所述视频进行处理的步骤包括：

对输入视频进行解码处理；

基于所述滤波器参数对解码后的视频进行滤波处理；

对滤波处理后的视频进行编码处理后输出。

6. 根据权利要求 1-5 任一项所述的方法，其特征在于，所述机器学习框架采用 Theano 库、Caffe 库或 Torch 库中的任一个。

7. 一种视频转码装置，其特征在于，包括：

模型接收模块，用于接收机器学习框架下发的模型文件；

文件变换模块，用于将所述模型文件变换成滤波器能够识别的简单文件；

参数设置模块，用于根据所述文件变换模块变换后的简单文件，设置滤波器参数；

视频处理模块，用于基于所述参数设置模块设置的滤波器参数对所述视频进行处理。

8. 根据权利要求 7 所述的装置，其特征在于，所述模型文件是通过机器学习框架对

训练样本进行训练得到的。

9. 根据权利要求 8 所述的装置，其特征在于，所述模型文件是通过机器学习框架对基于视频处理需求收集到的训练样本进行离线训练得到的。

10. 根据权利要求 7 所述的装置，其特征在于，所述文件变换模块，用于将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

11. 根据权利要求 7 所述的装置，其特征在于，所述视频处理模块包括：

解码单元，用于对输入视频进行解码处理；

滤波单元，用于基于所述滤波器参数对所述解码单元解码后的视频进行滤波处理；

编码单元，用于所述滤波单元滤波处理后的视频进行编码处理后输出。

12. 一种视频转码系统，其特征在于，包括：

机器学习框架，用于对训练样本进行训练得到模型文件；

视频转码装置包括：

模型接收模块，用于接收所述机器学习框架下发的模型文件；

文件变换模块，用于将所述模型文件变换成滤波器能够识别的简单文件；

参数设置模块，用于根据所述文件变换模块变换后的简单文件，设置滤波器参数；

视频处理模块，用于基于所述参数设置模块设置的滤波器参数对所述视频进行处理。

13. 根据权利要求 12 所述的系统，其特征在于，所述机器学习框架还包括：

收集模块，用于基于视频处理需求收集训练样本；

训练模块，用于对所述收集模块收集到的训练样本进行离线训练得到模型文件。

14. 根据权利要求 12 所述的系统，其特征在于，所述机器学习框架采用 Theano 库、Caffe 库或 Torch 库中的任一个。

15. 根据权利要求 12 所述的系统，其特征在于，所述文件变换模块用于将所述模型文件变换成 Ffmpeg 工具中的滤波器能够识别的文本文件。

16. 根据权利要求 12 所述的系统，其特征在于，所述视频处理模块包括：

解码单元，用于对输入视频进行解码处理；

滤波单元，用于基于所述滤波器参数对所述解码单元解码后的视频进行滤波处理；

编码单元，用于所述滤波单元滤波处理后的视频进行编码处理后输出。

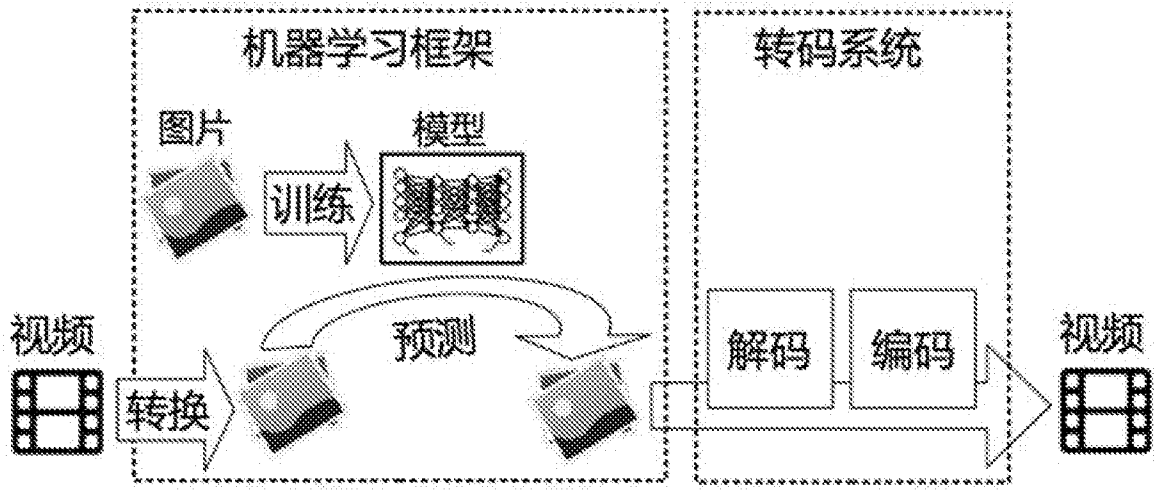

机器学习框架
转码系统
图片
模型
训练
视频
预测
转换
解码
编码
视频

图 1

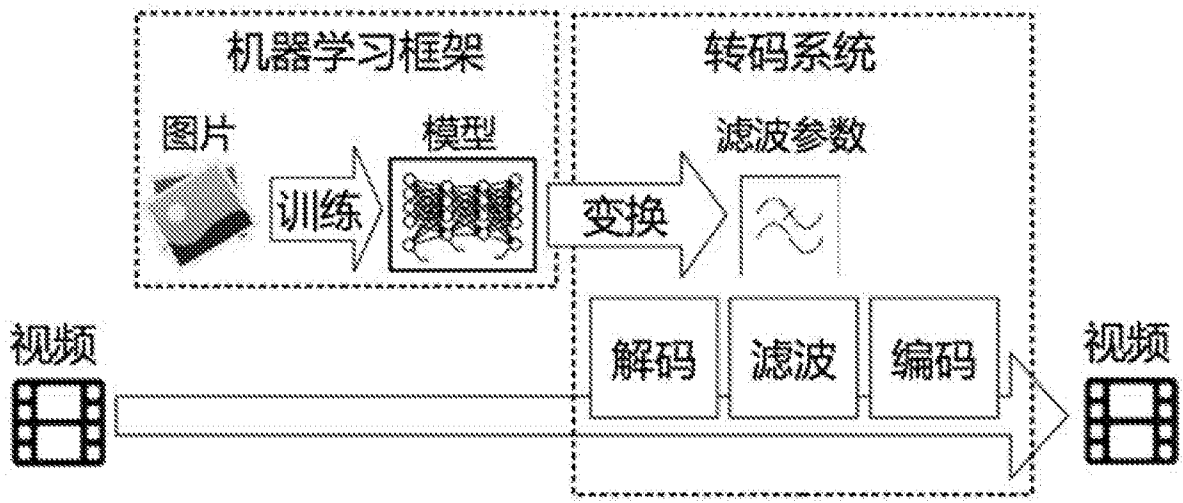

机器学习框架
转码系统
图片
模型
滤波参数
训练
变换
视频
解码
滤波
编码
视频

图 2

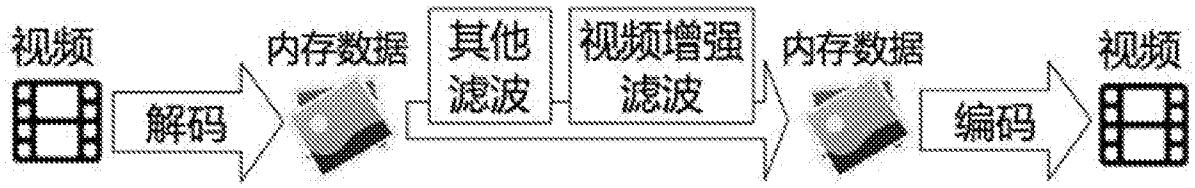

视频
解码
内存数据
其他
滤波
视频增强
滤波
内存数据
编码
视频

图 3

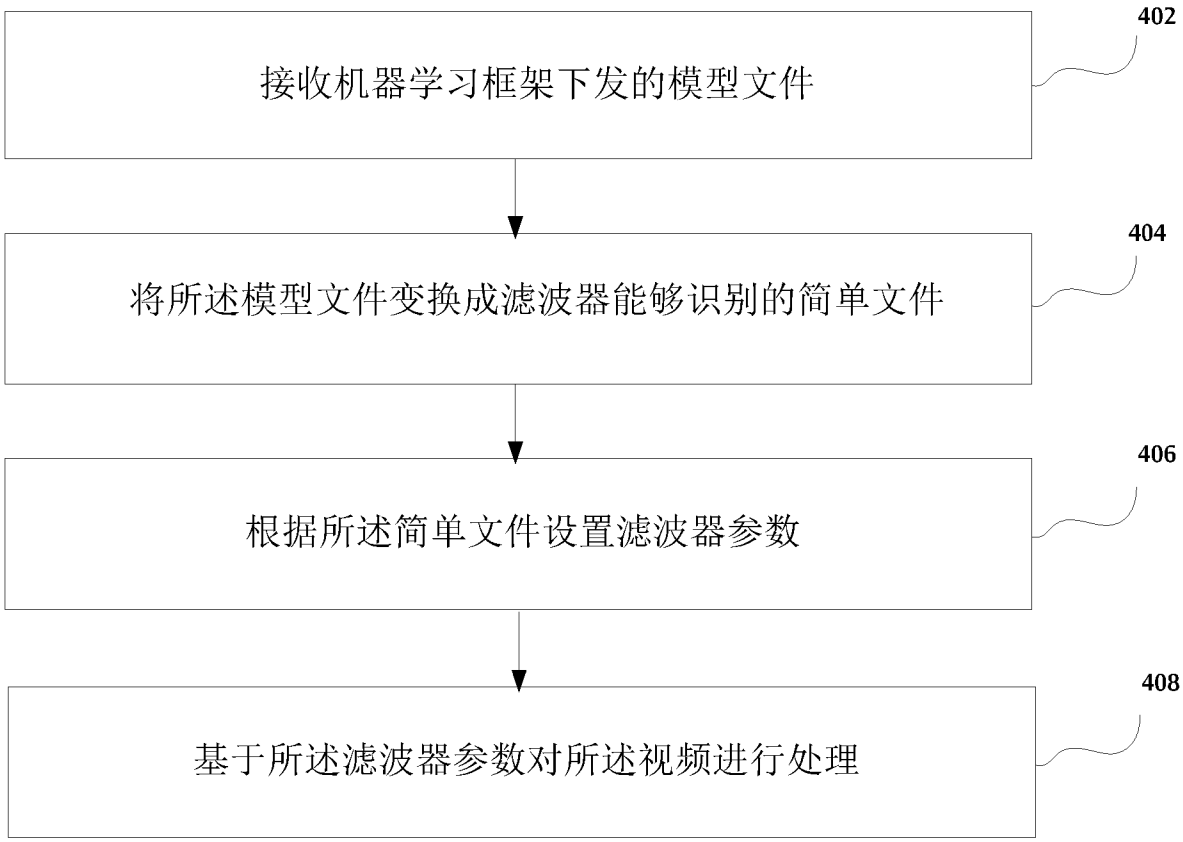

402
接收机器学习框架下发的模型文件
404
将所述模型文件变换成滤波器能够识别的简单文件
406
根据所述简单文件设置滤波器参数
408
基于所述滤波器参数对所述视频进行处理

图 4

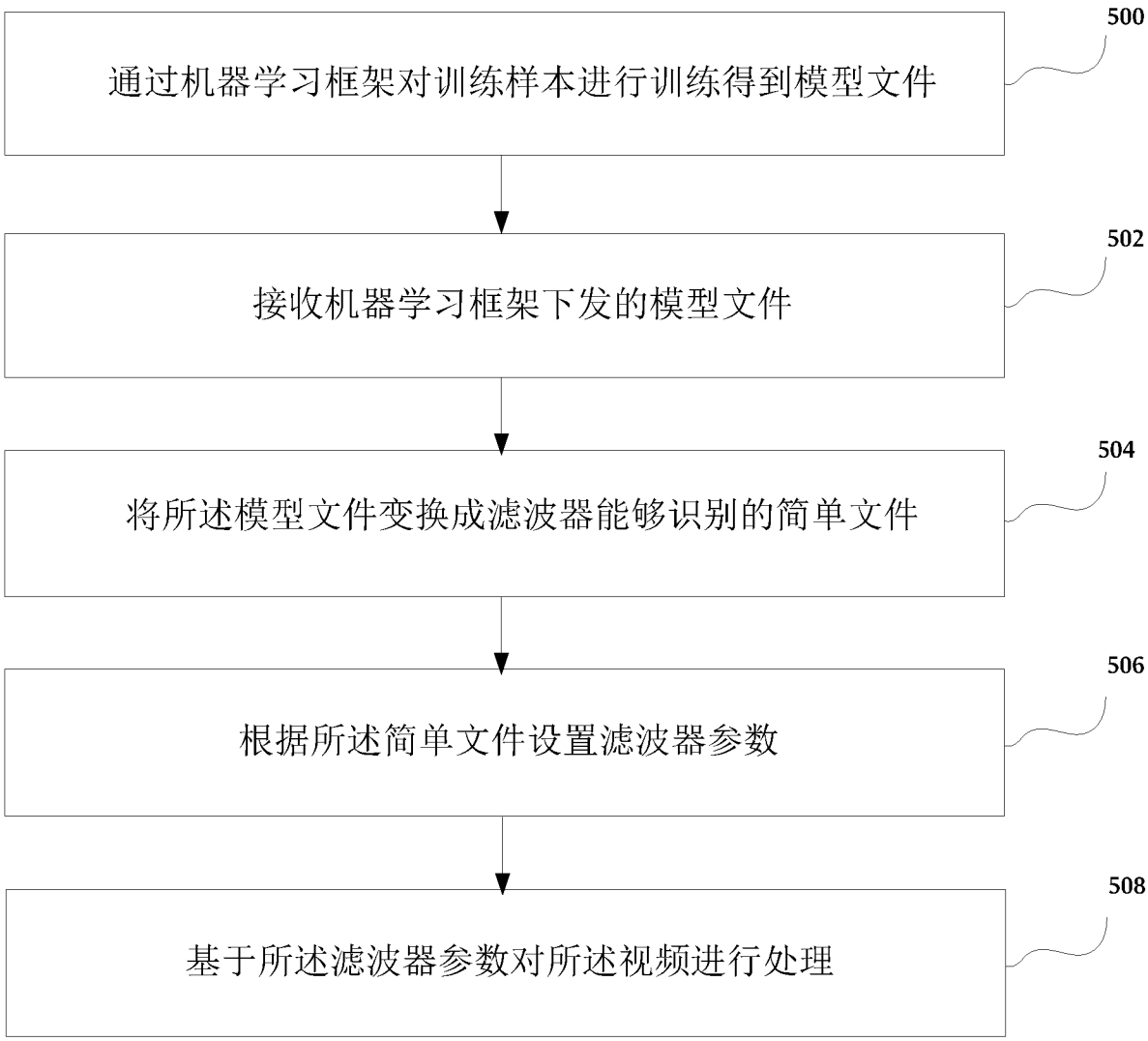

500
通过机器学习框架对训练样本进行训练得到模型文件
502
接收机器学习框架下发的模型文件
504
将所述模型文件变换成滤波器能够识别的简单文件
506
根据所述简单文件设置滤波器参数
508
基于所述滤波器参数对所述视频进行处理

图 5

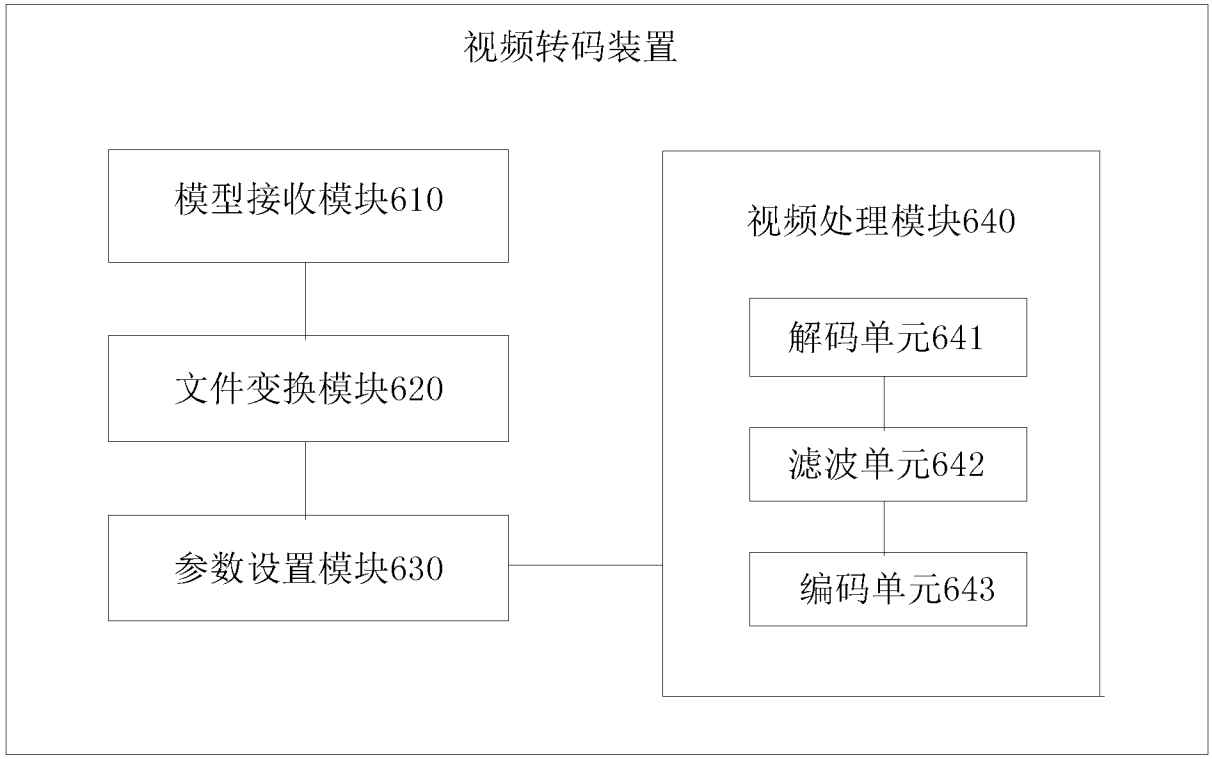

视频转码装置
模型接收模块610
文件变换模块620
参数设置模块630
视频处理模块640
解码单元641
滤波单元642
编码单元643

图 6

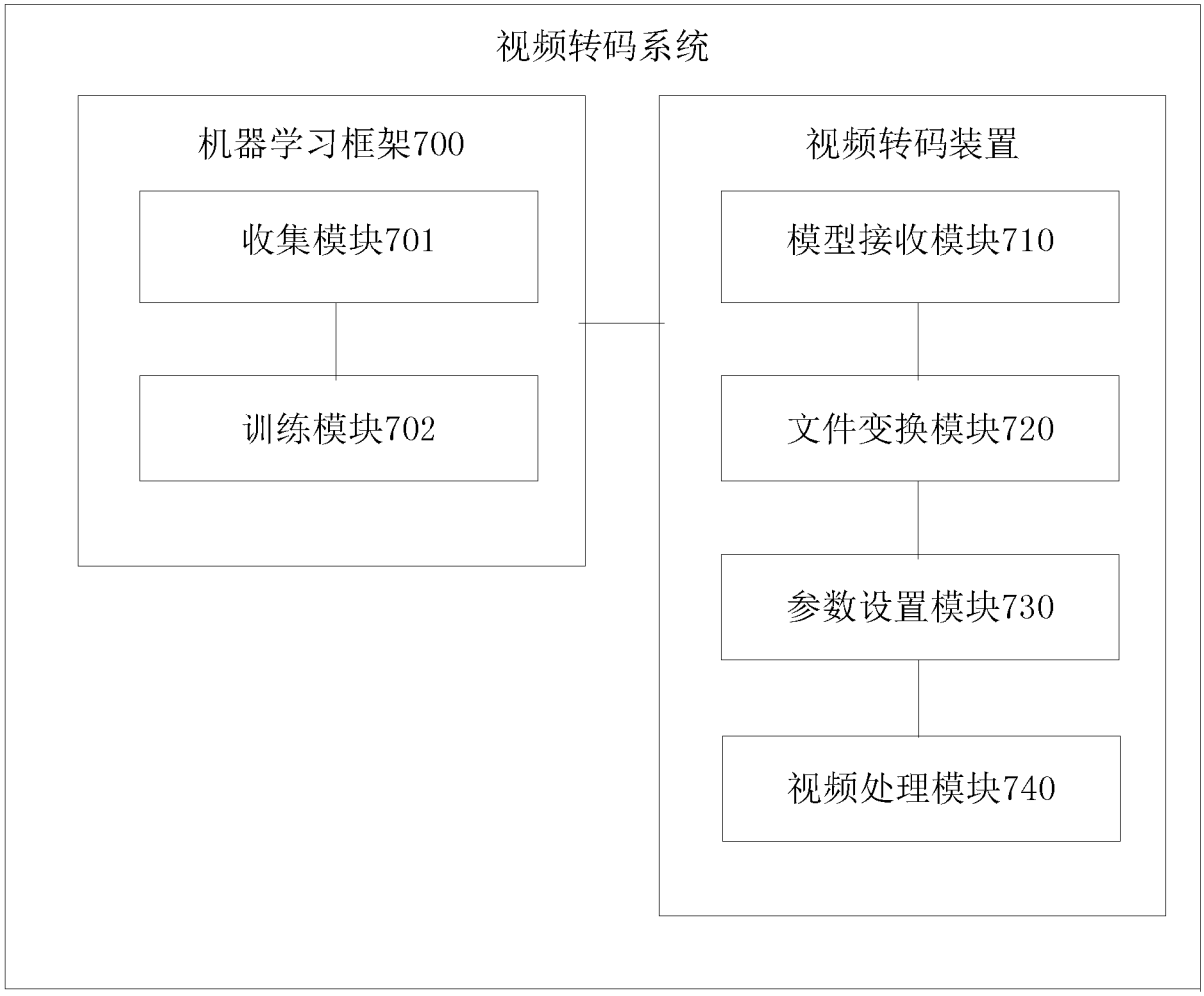

视频转码系统
机器学习框架700
收集模块701
训练模块702
视频转码装置
模型接收模块710
文件变换模块720
参数设置模块730
视频处理模块740

图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/076547

A. CLASSIFICATION OF SUBJECT MATTER

H04N 19/40 (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)

CNPAT, CNKI, EPODOC, WPI: Machine w learning, train+, transcod+, video, off w line, filter, parameter, model, file?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 105306947 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 03 February 2016 (03.02.2016) description, paragraphs [0006]-[0011], and figure 4	1-16
A	CN 101068355 A (NANJING UNIVERSITY) 07 November 2007 (07.11.2007) the whole document	1-16
A	CN 101924943 A (GUO, Min) 22 December 2010 (22.12.2010) the whole document	1-16
A	CN 103475876 A (BEIJING UNIVERSITY OF TECHNOLOGY) 25 December 2013 (25.12.2013) the whole document	1-16
A	US 2008212682 A1 (KALVA, HARI) 04 September 2008 (04.09.2008) the whole document	1-16

☐ 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	Date of mailing of the international search report
09 May 2017	05 June 2017
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, Wei Telephone No. (86-10) 62413228

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/076547

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105306947 A	03 February 2016	None	
CN 101068355 A	07 November 2007	None	
CN 101924943 A	22 December 2010	None	
CN 103475876 A	25 December 2013	None	
US 2008212682 A1	04 September 2008	US 2008205515 A1	28 August 2008
		WO 2008091687 A2	31 July 2008

国际检索报告

国际申请号
PCT/CN2017/076547

A. 主题的分类

H04N 19/40(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））

CNPAT,CNKI:机器学习,训练,视频,转码,离线,滤波,参数,模型,模板,文件。EPODOC,WPI:Machine w learning,train+,transcod+,video,off w line,filter, parameter,model,file?.

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 105306947 A (中国科学院深圳先进技术研究院) 2016年 2月 3日 (2016-02-03) 说明书第[0006]-[0011]段，附图4	1-16
A	CN 101068355 A (南京大学) 2007年 11月 7日 (2007-11-07) 全文	1-16
A	CN 101924943 A (郭敏) 2010年 12月 22日 (2010-12-22) 全文	1-16
A	CN 103475876 A (北京工业大学) 2013年 12月 25日 (2013-12-25) 全文	1-16
A	US 2008212682 A1 (KALVA, HARI) 2008年 9月 4日 (2008-09-04) 全文	1-16

☐ 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	国际检索报告邮寄日期
2017年 5月 9日	2017年 6月 5日

ISA/CN的名称和邮寄地址	受权官员
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	张伟
传真号 (86-10)62019451	电话号码 (86-10)62413228

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/076547

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	105306947	A	2016年 2月 3日	无			
CN	101068355	A	2007年 11月 7日	无			
CN	101924943	A	2010年 12月 22日	无			
CN	103475876	A	2013年 12月 25日	无			
US	2008212682	A1	2008年 9月 4日	US	2008205515	A1	2008年 8月 28日
				WO	2008091687	A2	2008年 7月 31日