

视频生成方法、装置、电子设备以及计算机可读介质

- 5 本专利申请要求于 2020 年 10 月 30 提交的、申请号为 202011192359.5、发明名称为“视频生成方法、装置、电子设备以及计算机可读介质”的中国专利申请的优先权，该申请的全文以引用的方式并入本申请中。

10 技术领域

本公开涉及计算机技术领域，具体涉及计算机视觉技术领域，尤其涉及视频生成方法、装置、电子设备以及计算机可读介质。

背景技术

- 15 视频生成是指将符合视频语义的视频片段剪辑为一个视频，而电商场景要求生成的视频能够多方面、多维度、多角度展示商品特性。

发明内容

- 20 本公开的实施例提出了视频生成方法、装置、电子设备以及计算机可读介质。

- 第一方面，本公开的实施例提供了一种视频生成方法，该方法包括：基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材；基于制作模板，
25 对商品图片或商品视频素材进行处理，生成商品视频。

- 在一些实施例中，上述基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频，包括：将制作模板中的音乐或商品视频素材中的音乐转到频域，计算音频能量的局部极值和错位卷积，确定重音点和节拍；将商品图片生成初始视频并从初始视频中提取多个预设时长的视频段落，或者从视频素材中提取多个预设时长的视频段落；
30

基于重音点、节拍，将多个视频段落以过渡动画的方式合并，生成商品视频。

5 在一些实施例中，上述响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材，包括：在用户登录之后，判断用户是否具有商家注册信息；响应于判断结果为用户具有商家注册信息，基于商家注册信息，获取用户在售商品的基本信息；基于用户对基本信息的操作，确定用户输入的商品信息，并获取与商品信息相关的商品图片或商品视频素材；响应于判断结果为用户不具有商家注册信息，提示用户输入商品信息，
10 响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材。

在一些实施例中，上述方法还包括：基于商品信息，获取与商品信息相关的商品详情页；提取商品详情页中的关键信息；对关键信息进行特效处理，并将特效处理后的关键信息写入商品视频；对写入关
15 键信息的商品视频进行滤镜和光影效果处理。

在一些实施例中，上述提取商品详情页中的关键信息，包括：采用语言模型提取商品详情页中的关键信息，语言模型基于制作模板的品类训练得到。

20 在一些实施例中，上述对商品图片进行处理包括：对商品图片进行预处理；识别预处理后的图片的文字区域，去除文字区域的文字内容。

在一些实施例中，上述方法还包括：将商品视频与商品的编码绑定；并将与商品的编码绑定后的商品视频上传至商品主图的宣传显示位。

25 在一些实施例中，上述方法还包括：响应于确定用户输入的商品信息所属品类与制作模板的品类不相匹配，发送提示更换制作模板的提示信息。

第二方面，本公开的实施例提供了一种视频生成装置，该装置包括：确定单元，被配置成基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；获取单元，被配置成响应于确定用户输入
30

的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材；生成单元，被配置成基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频。

5 在一些实施例中，上述生成单元包括：计算模块，被配置成将制作模板中的音乐或商品视频素材中的音乐转到频域，计算音频能量的局部极值和错位卷积，确定重音点和节拍；提取模块，被配置成将商品图片生成初始视频并从初始视频中提取多个预设时长的视频段落，或者从视频素材中提取多个预设时长的视频段落；生成模块，被配置成基于重音点、节拍，将多个视频段落以过渡动画的方式合并，生成商品视频。

10 在一些实施例中，上述获取单元包括：判断模块，被配置成在用户登录之后，判断用户是否具有商家注册信息；获取模块，被配置成响应于判断结果为用户具有商家注册信息，基于商家注册信息，获取用户在售商品的基本信息；确定模块，被配置成基于用户对基本信息的操作，确定用户输入的商品信息；响应模块，被配置成响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材；提示模块，被配置成响应于判断结果为用户不具有商家注册信息，提示用户输入商品信息，触发响应模块工作。

20 在一些实施例中，上述装置还包括：详分单元，被配置成基于商品信息，获取与商品信息相关的商品详情页；提取单元，被配置成提取商品详情页中的关键信息；特效单元，被配置成对关键信息进行特效处理，并将特效处理后的关键信息写入商品视频；处理单元，被配置成对写入关键信息的商品视频进行滤镜和光影效果处理。

25 在一些实施例中，上述提取单元进一步被配置成采用语言模型提取商品详情页中的关键信息，语言模型基于制作模板的品类训练得到。

在一些实施例中，上述生成单元包括：预处理模块，被配置成对商品图片进行预处理；识别模块，被配置成识别预处理后的图片的文字区域；去除模块，被配置成去除文字区域的文字内容。

30 在一些实施例中，上述装置还包括：绑定单元，被配置成将商品

视频与商品的编码绑定；上传单元，被配置成将与商品的编码绑定后的商品视频上传至商品主图的宣传显示位。

5 在一些实施例中，上述装置还包括：发送单元，被配置成响应于确定用户输入的商品信息所属品类与制作模板的品类不相匹配，发送提示更换制作模板的提示信息。

第三方面，本公开的实施例提供了一种电子设备，该电子设备包括：一个或多个处理器；存储装置，其上存储有一个或多个程序；当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如第一方面中任一实现方式描述的方法。

10 第四方面，本公开的实施例提供了一种计算机可读介质，其上存储有计算机程序，该程序被处理器执行时实现如第一方面中任一实现方式描述的方法。

本公开的实施例提供的视频生成方法和装置，首先基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；其次响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品
15 信息相关的商品图片或商品视频素材；最后基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频；由此通过与用户交互的确定制作模板，并基于制作模板以及获取的商品图片或商品视频素材生成商品视频，减化了视频制作的流程，提高了视频制作效率。

20

附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本公开的其它特征、目的和优点将会变得更明显：

图 1 是本公开的一个实施例可以应用于其中的示例性系统架构图；
25 图；

图 2 是根据本公开的视频生成方法的一个实施例的流程图；

图 3 是根据本公开的获取商品信息相关的商品图片或商品视频素材的方法的流程图；

图 4 是根据本公开的视频生成方法的另一个实施例的流程图；

30 图 5 是根据本公开的视频生成装置的实施例的结构示意图；

图 6 是适于用来实现本公开的实施例的电子设备的结构示意图。

具体实施方式

下面结合附图和实施例对本公开作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

需要说明的是，在不冲突的情况下，本公开中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本公开。

图 1 示出了可以应用本公开的视频生成方法的示例性系统架构 100。

如图 1 所示，系统架构 100 可以包括终端设备 101、102、103，网络 104 和服务器 105。网络 104 用以在终端设备 101、102、103 和服务器 105 之间提供通信链路的介质。网络 104 可以包括各种连接类型，通常可以包括无线通信链路等等。

终端设备 101、102、103 通过网络 104 与服务器 105 交互，以接收或发送消息等。终端设备 101、102、103 上可以安装有各种通讯客户端应用，例如即时通信工具、邮箱客户端等。

终端设备 101、102、103 可以是硬件，也可以是软件；当终端设备 101、102、103 为硬件时，可以是具有通信和控制功能的用户设备，上述用户设备可与服务器 105 进行通信。当终端设备 101、102、103 为软件时，可以安装在上述用户设备中；终端设备 101、102、103 可以实现成多个软件或软件模块（例如用来提供分布式服务的软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

服务器 105 可以是提供各种服务的服务器，例如为终端设备 101、102、103 上图像处理系统提供支持的视频生成的后台服务器。后台服务器可以对网络中各在线销售商品的相关信息进行分析处理，并将处理结果（如视频生成结果）反馈给终端设备。

需要说明的是，服务器可以是硬件，也可以是软件。当服务器为

硬件时，可以实现成多个服务器组成的分布式服务器集群，也可以实现成单个服务器。当服务器为软件时，可以实现成多个软件或软件模块（例如用来提供分布式服务的软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

- 5 需要说明的是，本公开的实施例所提供的视频生成方法一般由服务器 105 执行。

应该理解，图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务

- 10 如图 2，示出了根据本公开的视频生成方法的一个实施例的流程 200，该视频生成方法包括以下步骤：

步骤 201，基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板。

- 本实施例中，视频生成方法运行于其上的执行主体可以为有视频
15 制作需要的用户提供视频生成界面，并在视频生成界面上显示多种品类模板，用户在视频生成界面上输入指令，执行主体根据用户的输入指令，确定制作模板，其中，多种品类模板用于区别不同品类的商品，多种品类模板用于展示各品类商品的不同特性。不同品类包括：运动类、休闲类等，如运动服饰使用运动模板制作主图视频，运动模板上
20 快节奏的音频加上活泼的特效，更能凸显商品的特点。

- 本实施例中，各品类模板是一种数据结构，在品类模板中定义了视频要用的音乐、多种动画的种类以及进入方式和时间、转场、文字以及特效等。各品类模板是视频制作的基础，确定制作模板则可以在制作商品视频时，复用制作模板上的预设的专场、特效、音乐、卖点
25 等信息。

- 本实施例中，用户在制作商品视频之前，可根据品类选择品类模板，可选地，执行主体还可以记录各品类模板的使用量，每个品类模板使用一次，其对应的使用量累加；进一步，执行主体还可以根据使用量（例如，使用量前三位）推荐不同品类模板，具体推荐方式可以
30 为使用量前三位品类模板添加推荐标签，以使用户按喜好或者推荐选

用品类模板作为制作模板。

步骤 202，响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材。

本实施例中，在用户向视频生成界面输入指令，确定制作模板之后。视频生成方法运行与其上的执行主体还可以通过视频生成界面获取用户输入的商品信息，由商品信息确定视频生成的制作材料，该制作材料可以是商品图片或商品视频素材。

本实施例中，用户输入的商品信息包括：商品的编码、商品的名称、商品封面图、商品的图片、商品宣传视频、商品讲解视频、商品展示视频等。进一步地，执行主体可以通过用户输入的商品信息（例如，商品的编码、商品的名称、商品封面图），获取与商品信息相关的商品图片或商品视频素材。

在实际场景中，执行主体还可以基于用户输入的商品的编码，获取网页上商品的详情页，基于制作模板，经过对商品的详情页进行智能选图、图片杂质擦除、智能裁剪、卖点提取等步骤的处理，生成视频。

本实施例中，商品图片可以是商品的各种维度、各个角度的图片；商品视频素材可以是商品宣传视频、商品讲解视频、商品展示视频等素材。

本实施例的一些可选实现方式中，用户输入的商品信息可以是商品的编码，或者商品图片或者商品视频素材，用户输入的商品信息还可以包括：商品的编码、商品的编码对应的商品图片或者商品视频素材。即用户输入的商品信息可以是商品的编码、商品图片、商品视频素材中的任意一种，在本实施例中，如果不获取商品信息，则无法生成商品视频。

本实施例中，用户输入的商品信息所属品类可以是一种也可以是两种，当用户输入的商品信息所属品类是一种时，制作模板的品类需与用户输入的商品信息所属品类相同。当用户输入的商品信息所属品类是两种以上时，制作模板可以是通用模板，通用模板是所有品类商品可用的模板，无个性化品类的特征，比如非通用的运动类的模板，

会加一些运动类的元素或文案描述。

步骤 203, 基于制作模板, 对商品图片或商品视频素材进行处理, 生成商品视频。

本实施例中, 制作模板是即将生成的商品视频参照的模板, 制作模板为商品视频提供视频版式参考, 制作模板中定义了商品视频所涉及的音乐、动画的种类以及动画人物进入方式和时间、转场、文字以及特效等内容, 按照制作模板定义的内容, 对商品图片或商品视频素材进行处理, 生成商品视频。

本实施例中, 对商品图片或商品视频素材进行处理, 可以是一些简单的图片处理, 比如图片平移、缩放, 还可以是一些复杂的变换, 比如科技感的闪烁, 3D 旋转等, 还可以是设计师设计出来的动画, 这种通常是多张图片按照切割、拼接、复杂的空间运动等方式形成的动画格式。

可选地, 基于制作模板生成商品视频过程中, 可以直接在视频生成过程中增加商品卖点, 也可以在生成商品视频之后增加商品卖点。商品卖点是企业为展示自己产品的特点、优点, 而提炼的语言和演示。进一步地, 为了提升商品视频的吸引力, 可以对商品视频进行滤镜和特效处理, 能使视频呈现不同的风格, 丰富视频可观赏性。

在本实施例的一些可选实现方式中, 基于制作模板, 对商品图片或商品视频素材进行处理, 生成商品视频, 包括:

将制作模板中的音乐或商品视频素材中的音乐转到频域, 计算音频能量的局部极值和错位卷积, 确定重音点和节拍; 将商品图片生成初始视频并从初始视频中提取多个预设时长的视频段落, 或者从视频素材中提取多个预设时长的视频段落; 基于重音点、节拍, 将多个视频段落以过渡动画的方式合并, 生成商品视频。

本可选实现方式中, 可以采用多种动画生成函数将一张或商品图片生成初始视频。可以采用视频摘要提取模型从初始视频或视频素材中提取多个预设时长的视频段落。

本可选实现方式中, 过渡动画又称为转场过渡, 具体地, 可以通过 OPENGL (Open Graphics Library, 开放图形库) 实现转场过渡, 本

可选实现方式中，采用 OPENGL 实现的转场过渡可以得到近百种视频转场效果，比如，转场包括渐黑到渐亮的转场效果。

本可选实现方式中，通过计算音频能量的局部极值确定音乐的重音点，通过错位卷积确定音乐的节拍，基于重音点以及节拍将提取的多个视频段落以过渡动画的方式合并，可以保证视频转场点和视频配乐的节奏一致。

在本实施例的一些可选实现方式中，对商品图片进行处理包括：对商品图片进行预处理；识别预处理后的图片的文字区域，去除文字区域的文字内容。

本可选实现方式中，预处理包括：被切割图片的拼接，多主体图片切割，图片筛选过滤，图片杂质擦除、图片的智能裁剪和拼接以及图片尺寸统一设计。本可选实现方式中，可以采用深度学习 OCR（Optical Character Recognition，光学字符识别）识别预处理后的图片的文字区域和文字内容。在文字区域采用常规的文字擦图模型擦去文字内容，以保证生成的视频清晰整洁。

本可选实现方式中，首先对商品图片进行预处理，可以在商品图片在具有多张时，具有统一的规格尺寸。识别预处理后的图片的文字区域，去除文字区域的文字内容，可以保证生成的商品视频的清晰整洁。

在本实施例的另一些可选实现方式中，在去除文字区域的文字内容之前，还可以采用语言模型提取文字内容中的关键信息，并将提取出的关键信息写商品视频。本可选实现方式中，关键信息为文字形式的商品的卖点，将商品的卖点写入商品视频，可以便于用户在商品视频中快速发现商品的卖点信息。

在本实施例的一些可选实现方式中，响应于确定用户输入的商品信息所属品类与制作模板的品类不相匹配，发送提供更换制作模板的提示信息。

本可选实现方式中，在用户输入的商品信息所属品类与制作模板的品类不相匹配时，及时通过提示信息提示用户重新输入指令，保证了生成视频的风格与用户需求风格最大化匹配，使后续生成的商品视

频可以达到最优的制作效果。

在本实施例的另一些可选实现方式中，响应于确定用户输入的商品信息所属品类与制作模板的品类不相匹配，还可以向用户推荐可匹配的模板，从而达到更优的制作效果。

- 5 在本实施例的一些可选实现方式中，生成商品视频之后，还可以将商品视频与商品的编码绑定；并将与商品的编码绑定后的商品视频上传至商品主图的宣传显示位。

10 本可选实现方式中，通过将商品视频与商品的编码绑定，提高商品视频的制作效率，可以使单个商品视频制作时间在 40s 左右。同时商品视频可批量制作，大幅提升效率。

15 本公开的实施例提供的视频生成方法，首先基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；其次响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材；最后基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频；由此通过与用户交互的确定制作模板，并基于制作模板以及获取的商品图片或商品视频素材生成商品视频，简化了视频制作的流程，提高了视频制作效率。

20 本实施例中，商品信息可以是用户的线上商品或者用户的在售商品，视频生成方法运行于其上的执行可以根据用户的商品注册信息，确定是否向用户自动推荐商品信息，在本实施了一些可选实现方式中，获取商品信息相关的商品图片或商品视频素材的方法，包括以下步骤：

25 步骤 301，在用户登录之后，判断用户是否具有商家注册信息；在判断结果为用户具有商家注册信息时，之后，执行步骤 302；在判断结果为用户不具有商家注册信息时，之后，执行步骤 306。

 本可选实现方式，用户需要在执行主体提供的视频生成系统注册，并注册成功之后，登录视频生成系统，通过视频生成界面选择制作模板，输入商品信息。

30 本可选实现方式中，商家注册信息是指用户在视频生成系统中注册有商家账号，通过该商家账号，可以确定用户是否有在售商品，并

且在确定有在售商品之后可以得到在售商品的基本信息。

步骤 302, 基于商家注册信息, 获取用户在售商品的基本信息, 之后, 执行步骤 303。

本可选实现方式中, 在售商品的基本信息是指在售商品的编码
5 (SKU, Stock Keeping Unit)、商品名称、商品封面图等信息。

步骤 303, 基于用户对基本信息的操作, 确定用户输入的商品信息, 之后, 执行步骤 304。

本可选实现方式中, 在售商品的基本信息可以直接显示在视频生成界面, 用户参照视频生成界面上的显示内容, 进行直接点选或者输入
10 在售商品的编码、商品封面图、商品名称等操作, 确定用户输入的商品信息。当然, 在售商品的基本信息还可以显示在用户登录的操作界面上, 在用户对在售商品的基本信息进行直接点选或者输入在售商品的编码、商品封面图、商品名称等操作, 该操作界面不再显示。

本可选实现方式中, 在用户具有商家账号时, 执行主体可以从后台服务器获取商家在售商品, 让商家用户在制作视频时直接从在售商品
15 的基本信息中选择需要输入的商品信息, 而不用输入商品信息。

进一步地, 用户输入的商品信息是由在售商品的基本信息中得到, 提高了用户的选择的便利性。在用户忘记或者无法确定商品信息时。

步骤 304, 响应于确定用户输入的商品信息所属品类与制作模板
20 的品类匹配, 获取与商品信息相关的商品图片或商品视频素材, 之后, 执行步骤 305。

本可选实现方式中, 在用户不具有商家账号时, 则需用户直接输入商城在售商品的商品信息, 该商品信息包括商品的编码、商品的链接等, 进一步地, 基于商品信息生成视频。可选的, 可以根据商品
25 SKUID、SKU 链接生成视频。可选的, 还可以根据用户添加的图片或视频素材生成视频。

步骤 305, 退出。

步骤 306, 提示用户输入商品信息, 之后, 执行步骤 304。

本可选实现方式中, 可以向用户登录的操作界面中展示预设的提示
30 信息, 以提示用户输入商品信息。其中, 该操作界面还可以是用户

输入商品信息的界面。

本可选实现方式中，在具有商家账号的用户添加商品信息后，系统会再次判断商品信息所属品类与制作模板的品类是否匹配，不匹配则会推荐匹配模板，达到商品视频更优的制作效果。

- 5 本可选实现方式提供的获取商品信息相关的商品图片或商品视频素材的方法，在用户具有商家注册信息时，基于商家注册信息，确定用户在售商品的基本信息，基于用户对基本信息的操作，确定用户输入的商品信息。从而在与用户交互时，基于用户在售商品的基本信息，向用户自动推荐商品信息，提高了视频制作效率。

10

为了生成的视频具有更好的展示效果，进一步参考图 4，其示出了本公开的视频生成方法的另一个实施例的流程 400。该视频生成方法，包括以下步骤：

- 15 步骤 401，基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板。

步骤 402，响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材。

步骤 403，基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频。

- 20 应当理解，上述步骤 401-步骤 403 中的操作和特征，分别与步骤 201-步骤 203 中的操作和特征相对应，因此，上述在步骤 401-步骤 403 中对于操作和特征的描述，同样适用于步骤 201-步骤 203，在此不再赘述。

步骤 404，基于商品信息，获取与商品信息相关的商品详情页。

- 25 本实施例中，商品详情页即商品卖家制作的商品详情页，该商品详情页上描述有商品的产地、厂家、规格、适用范围等等详细描述信息，该详细描述信息可以包含商品的视频、图片、文字描述。例如，在某网页点击某个商品宣传显示位的图片可以查看到的商品详情页。

步骤 405，提取商品详情页中的关键信息。

- 30 本实施例中，关键信息可以是表征商品的特征的信息，比如，文

字、图片、视频，通过提取的关键信息可以体现商品的卖点。

本实施例的一些可选实现方式中，关键信息可以为文字，提取商品详情页中的关键信息，包括：采用语言模型提取商品详情页中的关键信息，语言模型基于制作模板的品类训练得到。

- 5 针对不同商品品类可以训练不同的 BERT (Bidirectional Encoder Representations from Transformers)语言模型，采用语言模型可以对 OCR 识别的文字进行分词、权重设置以及语义理解，从大段文字或者多句文字中提炼出数条简练的短语作为商品卖点宣传文案。其中，权重设置是指：在语言模型样本选出一部分商品文案作为标定样例，分
- 10 词后给标定团队按照业务的需求标定，再用语言模型在这个数据上训练出每个词的权重。

本可选实现方式中，通过语言模型可以有效提取商品详情页中的文字的关键信息，提高了商品卖点提取的效率。

步骤 406，对关键信息进行特效处理。

- 15 本实施例中，特效处理可以根据用户指定的业务需求而进行设置，以达到多种展示效果的目的，例如，特效处理为关键信息增加一些光影或者粒子。

步骤 407，将特效处理后的关键信息写入商品视频。

- 20 本实施例中，特效处理后的关键信息实现了卖点的多种样式展示效果。例如，关键信息为文字时，实现了多种文字展示效果。

步骤 408，对写入关键信息的商品视频进行滤镜和光影效果处理。

- 25 本实施例中，滤镜，主要是用来实现图像的各种特殊效果。光影效果处理，使图像中物体具有在阳光照射下形成的阴影效果。对商品视频进行数个常用滤镜和光影效果，能使商品视频呈现不同的风格，
- 丰富商品视频可观赏性。

步骤 409，将写入关键信息的商品视频与商品的编码绑定。

- 30 本实施例中，将商品视频与商品的编码进行绑定，便于查找到与商品编码相对于的商品。例如，批量一同生成 5 个商品的编码的商品视频，则制作的视频自动关联至这 5 个商品。通过将商品视频与商品的编码绑定，提高商品视频的制作效率，可以使单个商品视频制

作时间在 40s 左右。同时商品视频可批量制作，大幅提升效率。

步骤 410，将与商品的编码绑定后的商品视频上传至商品主图的宣传显示位。

本实施例中，已生成的商品视频将会被上传至商品主图的宣传显示位。用户在浏览商品时可通过浏览商品视频全方位了解商品外观、特性和卖点。进一步地，可以通过视频审核系统将商品视频上传至商品主图的宣传显示位，视频审核系统用于审核商品是否符合预设的视频规范，该预设的视频规范有专门的规范文件进行约束。

视频生成完成后，视频审核系统判断视频生成完成，对商品视频进行审核判断，审核通过则在商品详情页的位置（即商品主图的宣传显示位）展示。

本实施例中，通过将商品视频上传至商品主图的宣传显示位，能够加强商品特性的宣传，吸引买家，加强向订单引流，直至转化。

本实施例提供的视频生成方法，基于商品信息，获取与商品信息相关的商品详情页，提取商品详情页中的关键信息，对关键信息进行特效处理，并将特效处理后的关键信息写入商品视频，并对写入关键信息的商品视频进行滤镜和光影效果处理，通过提取的关键信息得到商品的卖点，通过对关键信息进行特效处理，提高了卖点的显示效果；通过对写入关键信息的商品视频进行滤镜和光影效果处理，提高了卖点在商品视频中整体协调度，并且还提高了商品视频的显示效果。

进一步参考图 5，作为对上述各图所示方法的实现，本公开提供了视频生成装置的一个实施例，该装置实施例与图 2 所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图 5 所示，本公开的实施例提供了一种视频生成装置 500，该装置 500 包括：确定单元 501、获取单元 502、生成单元 503。其中，确定单元 501，可以被配置成基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板。获取单元 502，可以被配置成响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材。生成单元 503，可以被配置

成基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频。

在本实施例中，视频生成装置 500 中，确定单元 501、获取单元 502、生成单元 503 的具体处理及其所带来的技术效果可分别参考图 2 对应实施例中的步骤 201、步骤 202、步骤 203。

在一些实施例中，上述生成单元 503 包括：计算模块（图中未示出）、提取单元（图中未示出）、生成模块（图中未示出）。其中，计算模块，可以被配置成将制作模板中的音乐或商品视频素材中的音乐转到频域，计算音频能量的局部极值和错位卷积，确定重音点和节拍。

10 提取模块，可以被配置成将商品图片生成初始视频并从初始视频中提取多个预设时长的视频段落，或者从视频素材中提取多个预设时长的视频段落。生成模块，可以被配置成基于重音点、节拍，将多个视频段落以过渡动画的方式合并，生成商品视频。

在一些实施例中，上述获取单元 502 包括：判断模块（图中未示出）、获取模块（图中未示出）、确定模块（图中未示出）、响应模块（图中未示出）、提示模块（图中未示出）。其中，判断模块，可以被配置成在用户登录之后，判断用户是否具有商家注册信息。获取模块，可以被配置成响应于判断结果为用户具有商家注册信息，基于商家注册信息，获取用户在售商品的基本信息。确定模块，可以被配置成基于

20 用户对基本信息的操作，确定用户输入的商品信息。响应模块，可以被配置成响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材。提示模块，可以被配置成响应于判断结果为用户不具有商家注册信息，提示用户输入商品信息，触发响应模块工作。

25 在一些实施例中，上述装置 500 还包括：详分单元（图中未示出）、提取单元（图中未示出）、特效单元（图中未示出）、处理单元（图中未示出）。其中，详分单元，可以被配置成基于商品信息，获取与商品信息相关的商品详情页。提取单元，可以被配置成提取商品详情页中的关键信息。特效单元，可以被配置成对关键信息进行特效处理，并将特效处理后的关键信息写入商品视频。处理单元，可以被配置成对

30

写入关键信息的商品视频进行滤镜和光影效果处理。

在一些实施例中，上述提取单元，进一步被配置成采用语言模型提取商品详情页中的关键信息，语言模型基于制作模板的品类训练得到。

5 在一些实施例中，上述生成单元 503 包括：预处理模块（图中未示出）、识别模块（图中未示出）、去除模块（图中未示出）。其中，预处理模块，可以被配置成对商品图片进行预处理。识别模块，可以被配置成识别预处理后的图片的文字区域。去除模块，可以被配置成去除文字区域的文字内容。

10 在一些实施例中，上述装置 500 还包括：绑定单元（图中未示出）、上传单元（图中未示出）。其中，绑定单元，可以被配置成将商品视频与商品的编码绑定。上传单元，可以被配置成将与商品的编码绑定后的商品视频上传至商品主图的宣传显示位。

在一些实施例中，上述装置 500 还包括：发送单元（图中未示出）。
15 上述发送单元，可以被配置成响应于确定用户输入的商品信息所属品类与制作模板的品类不相匹配，发送提示更换制作模板的提示信息。

本公开的实施例提供的视频生成装置，首先确定单元 501 基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；其次获取单元 502 响应于确定用户输入的商品信息所属品类与制作模板的
20 品类匹配，获取与商品信息相关的商品图片或商品视频素材；最后生成单元 503 基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频；由此通过与用户交互的确定制作模板，并基于制作模板以及获取的商品图片或商品视频素材生成商品视频，降低了视频制作的门槛，为用户提供了简单便捷的操作方式，方便了用户的使用，提
25 高了用户的体验。

下面参考图 6，其示出了适于用来实现本公开的实施例的电子设备 600 的结构示意图。

如图 6 所示，电子设备 600 可以包括处理装置（例如中央处理器、
30 图形处理器等）601，其可以根据存储在只读存储器（ROM）602 中的

程序或者从存储装置 608 加载到随机访问存储器 (RAM) 603 中的程序而执行各种适当的动作和处理。在 RAM 603 中, 还存储有电子设备 600 操作所需的各种程序和数据。处理装置 601、ROM 602 以及 RAM 603 通过总线 604 彼此相连。输入/输出 (I/O) 接口 605 也连接至总线 604。

通常, 以下装置可以连接至 I/O 接口 605: 包括例如触摸屏、触摸板、键盘、鼠标、等的输入装置 606; 包括例如液晶显示器 (LCD, Liquid Crystal Display)、扬声器、振动器等的输出装置 607; 包括例如磁带、硬盘等的存储装置 608; 以及通信装置 609。通信装置 609 可以允许电子设备 600 与其他设备进行无线或有线通信以交换数据。虽然图 6 示出了具有各种装置的电子设备 600, 但是应理解的是, 并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。图 6 中示出的每个方框可以代表一个装置, 也可以根据需要代表多个装置。

特别地, 根据本公开的实施例, 上文参考流程图描述的过程可以被实现为计算机软件程序。例如, 本公开的实施例包括一种计算机程序产品, 其包括承载在计算机可读介质上的计算机程序, 该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中, 该计算机程序可以通过通信装置 609 从网络上被下载和安装, 或者从存储装置 608 被安装, 或者从 ROM 602 被安装。在该计算机程序被处理装置 601 执行时, 执行本公开的实施例的方法中限定的上述功能。

需要说明的是, 本公开的实施例的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件, 或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于: 具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (ROM)、可擦式可编程只读存储器 (EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开的实施例中, 计算机可

读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开的实施例中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（Radio Frequency，射频）等等，或者上述的任意合适的组合。

上述计算机可读介质可以是上述服务器中所包含的；也可以是单独存在，而未装配入该服务器中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该服务器执行时，使得该服务器：基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；响应于确定用户输入的商品信息所属品类与制作模板的品类匹配，获取与商品信息相关的商品图片或商品视频素材；基于制作模板，对商品图片或商品视频素材进行处理，生成商品视频。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的实施例的操作的计算机程序代码，程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开的各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点

上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。

- 5 例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

- 10 描述于本公开的实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中，例如，可以描述为：一种处理器，包括确定单元、获取单元、和生成单元。其中，这些单元的名称在某种情况下并不构成对该单元本身的限定，例如，确定单元还可以被描述为“被配置成基于用户的
- 15 输入指令，从多种品类模板中确定一个模板作为制作模板”的单元。

- 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开的实施例中所涉及的发明范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述发明构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开的实施例中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。
- 20

权 利 要 求 书

1、一种视频生成方法，所述方法包括：

5 基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；

响应于确定用户输入的商品信息所属品类与所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材；

基于所述制作模板，对所述商品图片或商品视频素材进行处理，生成商品视频。

10

2、根据权利要求1所述的方法，其中，所述基于所述制作模板，对所述商品图片或商品视频素材进行处理，生成商品视频，包括：

将所述制作模板中的音乐或所述商品视频素材中的音乐转到频域，计算音频能量的局部极值和错位卷积，确定重音点和节拍；

15 将所述商品图片生成初始视频并从所述初始视频中提取多个预设时长的视频段落，或者从所述视频素材中提取多个预设时长的视频段落；

基于所述重音点、所述节拍，将多个所述视频段落以过渡动画的方式合并，生成商品视频。

20

3、根据权利要求1-2任一项所述的方法，其中，所述响应于确定用户输入的商品信息所属品类与所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材，包括：

在用户登录之后，判断所述用户是否具有商家注册信息；

25 响应于判断结果为所述用户具有商家注册信息，基于所述商家注册信息，获取所述用户在售商品的基本信息；

基于所述用户对所述基本信息的操作，确定用户输入的商品信息；响应于确定用户输入的商品信息所属品类与所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材；

30 响应于判断结果为所述用户不具有商家注册信息，提示用户输入

商品信息，响应于确定用户输入的商品信息所属品类与所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材。

- 4、根据权利要求 1-3 任一项所述的方法，所述方法还包括：
- 5 基于所述商品信息，获取与所述商品信息相关的商品详情页；
提取所述商品详情页中的关键信息；
对所述关键信息进行特效处理，并将特效处理后的关键信息写入所述商品视频；
对写入关键信息的商品视频进行滤镜和光影效果处理。

10

5、根据权利要求 4 所述的方法，其中，所述提取所述商品详情页中的关键信息，包括：

采用语言模型提取所述商品详情页中的关键信息，所述语言模型基于所述制作模板的品类训练得到。

15

6、根据权利要求 1-5 任一项所述的方法，其中，所述对所述商品图片进行处理包括：

- 对所述商品图片进行预处理；
识别预处理后的图片的文字区域；
20 去除所述文字区域的文字内容。

- 7、根据权利要求 1-6 任一项所述的方法，所述方法还包括：
将所述商品视频与商品的编码绑定；
并将与商品的编码绑定后的商品视频上传至所述商品主图的宣传
25 显示位。

8、根据权利要求 1-7 之一所述的方法，所述方法还包括：
响应于确定用户输入的商品信息所属品类与所述制作模板的品类
不相匹配，发送提示更换所述制作模板的提示信息。

30

9、一种视频生成装置，所述装置包括：

确定单元，被配置成基于用户的输入指令，从多种品类模板中确定一个模板作为制作模板；

获取单元，被配置成响应于确定用户输入的商品信息所属品类与
5 所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材；

生成单元，被配置成基于所述制作模板，对所述商品图片或商品视频素材进行处理，生成商品视频。

10 10、根据权利要求 9 所述的装置，其中，所述生成单元包括：

计算模块，被配置成将所述制作模板中的音乐或所述商品视频素材中的音乐转到频域，计算音频能量的局部极值和错位卷积，确定重音点和节拍；

提取模块，被配置成将所述商品图片生成初始视频并从所述初始
15 视频中提取多个预设时长的视频段落，或者从所述视频素材中提取多个预设时长的视频段落；

生成模块，被配置成基于所述重音点、所述节拍，将多个所述视频段落以过渡动画的方式合并，生成商品视频。

20 11、根据权利要求 9-10 任一项所述的装置，其中，所述获取单元进一步包括：

判断模块，被配置为：在用户登录之后，判断所述用户是否具有商家注册信息；

获取模块，被配置为：响应于判断结果为所述用户具有商家注册
25 信息，基于所述商家注册信息，获取所述用户在售商品的基本信息；

确定模块，被配置为：基于所述用户对所述基本信息的操作，确定用户输入的商品信息；

响应模块，被配置为：响应于确定用户输入的商品信息所属品类与
30 所述制作模板的品类匹配，获取与所述商品信息相关的商品图片或商品视频素材；以及

提示模块，被配置为：响应于判断结果为所述用户不具有商家注册信息，提示用户输入商品信息，以及触发响应模块工作。

12、根据权利要求 9-11 任一项所述的装置，其中，所述装置还包括：

5 详分单元，被配置为基于所述商品信息，获取与所述商品信息相关的商品详情页；

提取单元，被配置为提取所述商品详情页中的关键信息；

特效单元，被配置为对所述关键信息进行特效处理，并将特效处理后的关键信息写入所述商品视频；以及

10 处理单元，被配置为对写入关键信息的商品视频进行滤镜和光影效果处理。

13、根据权利要求 12 所述的装置，其中，所述提取单元进一步被配置为采用语言模型提取所述商品详情页中的关键信息，所述语言模型基于所述制作模板的品类训练得到。

14、根据权利要求 9-13 任一项所述的装置，其中所述生成单元包括：

预处理模块，被配置为对所述商品图片进行预处理；

20 识别模块，被配置为识别预处理后的图片的文字区域；以及

去除模块，被配置为去除所述文字区域的文字内容。

15、根据权利要求 9-14 任一项所述的装置，其中，所述装置还包括：

25 绑定单元，被配置为将所述商品视频与商品的编码绑定；以及

上传单元，被配置为将与商品的编码绑定后的商品视频上传至所述商品主图的宣传显示位。

16、根据权利要求 9-15 任一项所述的装置，其中，所述装置还包括：

发送单元，被配置为响应于确定用户输入的商品信息所属品类与所述制作模板的品类不相匹配，发送提示更换所述制作模板的提示信息。

5 17、一种电子设备，包括：

一个或多个处理器；

存储装置，其上存储有一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-8 中任一所述的方法。

10

18、一种计算机可读介质，其上存储有计算机程序，其中，该程序被处理器执行时实现如权利要求 1-8 中任一所述的方法。