

(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
2025 年 6 月 5 日 (05.06.2025)



(10) 国际公布号
WO 2025/113271 A1

(51) 国际专利分类号:
G06F 16/432 (2019.01)

(21) 国际申请号: PCT/CN2024/133084

(22) 国际申请日: 2024 年 11 月 20 日 (20.11.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311597109.3 2023 年 11 月 27 日 (27.11.2023) CN

(71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园4号楼2层0207 100190 (CN)。

(72) 发明人: 张子健(**ZHANG, Zijian**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028

(CN)。林佩吟(**LIN, Peiyin**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。

(74) 代理人: 中国贸促会专利商标事务所有限公司(**CCPIT PATENT AND TRADEMARK LAW OFFICE**); 中国北京市西城区复兴门内大街158号远洋大厦F10层 100031 (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, 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, MU, 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。

(54) **Title:** METHOD AND APPARATUS FOR GENERATING INTERACTIVE MULTIMEDIA CONTENT, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 互动多媒体内容的生成方法、装置、电子设备和存储介质

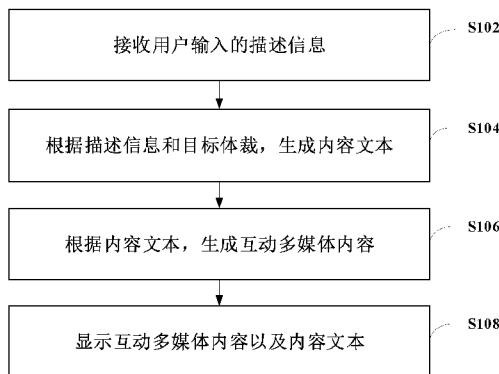


图 1

S102 Receive description information inputted by a user
S104 On the basis of the description information and a target genre, generate a content text
S106 On the basis of the content text, generate interactive multimedia content
S108 Display interactive multimedia content and the content text

(57) **Abstract:** The present disclosure relates to a method and apparatus for generating interactive multimedia content, an electronic device, and a storage medium, and relates to the technical field of computers. The method of the present disclosure comprises: receiving description information inputted by a user; on the basis of the description information and a target genre, generating a content text, the content text comprising a description text of plot content and a description text of a visual transformation corresponding to the target genre; on the basis of the content text, generating interactive multimedia content; and displaying the interactive multimedia content and the content text, the interactive multimedia content presenting a visual transformation effect corresponding to the description text of the visual transformation.

(57) **摘要:** 本公开涉及一种互动多媒体内容的生成方法、装置、电子设备和存储介质, 涉及计算机技术领域。本公开的方法包括: 接收用户输入的描述信息; 根据所述描述信息和目标体裁, 生成内容文本, 其中, 所述内容文本包括剧情内容的描述文本以及所述目标体裁对应的视觉变换的描述文本; 根据所述内容文本, 生成互动多媒体内容; 显示所述互动多媒体内容以及所述内容文本, 其中, 所述互动多媒体内容呈现与所述视觉变换的描述文本对应的视觉变换效果。

[见续页]

WO 2025/113271 A1

(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))。

互动多媒体内容的生成方法、装置、电子设备和存储介质

相关申请的交叉引用

本申请是以 CN 申请号为 202311597109.3，申请日为 2023 年 11 月 27 日的申请
5 为基础，并主张其优先权，该 CN 申请的公开内容在此作为整体引入本申请中。

技术领域

本公开涉及计算机技术领域，特别涉及一种互动多媒体内容的生成方法、装置、
电子设备和存储介质。

10 背景技术

互动多媒体内容是一种集成了图像、声音、文字等多种元素、并且能够为其观看
者（或称为使用者）提供交互接口的内容，例如，游戏、互动电影、互动电视剧等。

互动多媒体内容一般由专业的设计和开发人员进行制作，需要经过剧情的创作、
美术人员绘制、开发人员的编写代码或组织演员拍摄、后期处理等一系列的过程才能
15 完成，并且不同体裁的互动多媒体内容的制作过程所需的专业人员和制作流程是不同。

发明内容

提供该发明内容部分以便以简要的形式介绍构思，这些构思将在后面的具体实施
方式部分被详细描述。该发明内容部分并不旨在标识要求保护的技术方案的关键特征
20 或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

根据本公开的一些实施例，提供了一种互动多媒体内容的生成方法，包括：接收
用户输入的描述信息；根据所述描述信息和目标体裁，生成内容文本，其中，所述内
容文本包括剧情内容的描述文本以及所述目标体裁对应的视觉变换的描述文本；根据
所述内容文本，生成互动多媒体内容；显示所述互动多媒体内容以及所述内容文本，
25 其中，所述互动多媒体内容呈现与所述视觉变换的描述文本对应的视觉变换效果。

根据本公开的另一些实施例，提供了一种互动多媒体内容的生成方装置，包括：
接收模块，被配置为接收用户输入的描述信息；内容生成模块，被配置为根据所述描
述信息和目标体裁，生成内容文本，其中，所述内容文本包括剧情内容的描述文本以
及所述目标体裁对应的视觉变换的描述文本；多媒体生成模块，被配置为根据所述内
30 容文本，生成互动多媒体内容；显示模块，被配置为显示所述互动多媒体内容以及所

述内容文本，其中，所述互动多媒体内容呈现与所述视觉变换的描述文本对应的视觉变换效果。

根据本公开的一些实施例，提供一种电子设备，包括：存储器；和耦接至存储器的处理器，所述处理器被配置为基于存储在所述存储器中的指令，执行本公开中任一
5 实施例的互动多媒体内容的生成方法。

根据本公开的一些实施例，提供一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时执行本公开中任一实施例的互动多媒体内容的生成方法。

通过以下参照附图对本公开的示例性实施例的详细描述，本公开的其它特征、方面及其优点将会变得清楚。

附图说明

下面参照附图说明本公开的优选实施例。此处所说明的附图用来提供对本公开的进一步理解，各附图连同下面的具体描述一起包含在本说明书中并形成说明书的一部分，用于解释本公开。应当理解的是，下面描述中的附图仅仅涉及本公开的一些实施例，而非对本公开构成限制。在附图中：

图 1 示出本公开的一些实施例的互动多媒体内容的生成方法的流程示意图；

图 2 示出本公开的一些实施例的创作界面的示意图；

图 3 示出本公开的另一些实施例的互动多媒体内容的生成方法的流程示意图；

图 4 示出本公开的另一些实施例的创作界面的示意图；

图 5 示出本公开的又一些实施例的创作界面的示意图；

图 6 示出本公开的一些实施例的互动多媒体内容的画面的示意图；

图 7 示出本公开的再一些实施例的创作界面的示意图；

图 8 示出本公开的一些实施例的互动多媒体内容的生成装置的结构示意图；

图 9 示出本公开的一些实施例的电子设备的结构示意图；

图 10 示出本公开的一些实施例的计算机系统的结构示意图。

应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不一定是按照实际的比例关系绘制的。在各附图中使用了相同或相似的附图标记来表示相同或者相似的部件。因此，一旦某一项在一个附图中被定义，则在随后的附图中可能不再对其进行进一步讨论。

具体实施方式

下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，但是显然，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。以下对实施例的描述实际上也仅仅是说明性的，决不作为对本公开及其应用或使用的任何限制。应当理解的是，本公开可以通过各种形式来实现，而且不应该被解

5 释为限于这里阐述的实施例。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值应被解释为仅仅是示例性的，不限制本公开

10 的范围。

本公开中使用的术语“包括”及其变型意指至少包括后面的元件/特征、但不排除其他元件/特征的开放性术语，即“包括但不限于”。此外，本公开使用的术语“包含”及其变型意指至少包含后面的元件/特征、但不排除其他元件/特征的开放性术语，即“包含但不限于”。因此，包括与包含是同义的。术语“基于”意指“至少部分地基于”。

15

整个说明书中所称“一个实施例”、“一些实施例”或“实施例”意味着与实施例结合描述的特定的特征、结构或特性被包括在本发明的至少一个实施例中。例如，术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。而且，短语“在一个实施例中”、“在一些实施例中”或“在实施例中”在整个说明书中各个地方的出现不一定全都指的是同一个实施例，但是也可以指同一个实施例。

20

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。除非另有指定，否则“第一”、“第二”等概念并非意图暗示如此描述

25 的对象必须按时间上、空间上、排名上的给定顺序或任何其他方式的给定顺序。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

30

下面结合附图对本公开的实施例进行详细说明，但是本公开并不限于这些具体的实施例。下面这些具体实施例可以相互结合，对于相同或者相似的概念或过程可能在某些实施例不再赘述。此外，在一个或多个实施例中，特定的特征、结构或特性可以由本领域的普通技术人员从本公开将清楚的任何合适的方式组合。

5 目前，互动多媒体内容属于专业生产内容（**Professional Generated Content**，简称：**PGC**），其生成过程具有较高的门槛，并且通常涉及多个领域的专业人员的合作。互动多媒体内容难以以用户生成内容（**User Generated Content**，简称：**UGC**）的方式进行制作，其生成效率比较低。由于不同体裁的互动多媒体内容的表现形式不同，更加难以高效的以用户生成内容的方式，制作不同体裁的互动多媒体内容。

10 为了提高不同体裁的互动多媒体内容的生成效率，本公开提出一种互动多媒体内容的生成方法、装置、电子设备和存储介质。下面结合图 1~7 描述本公开的互动多媒体内容的生成方法。

图 1 为本公开互动多媒体内容的生成方法一些实施例的流程图。如图 1 所示，该实施例的方法包括：步骤 S102~S108。

15 在步骤 S102 中，接收用户输入的描述信息。

例如，显示互动多媒体内容的创作界面，在创作界面中显示输入区域，接收用户输入的描述信息。描述信息可以包括文本、语音、图片中至少一项。在描述信息为语音的情况下，可以进行语音识别将语音转换为文本。在描述信息为图片的情况下，可以进行图像识别，将图片转换为文本。

20 用户输入的描述信息包括可以用于描述待生成的互动多媒体内容的关键信息。例如，描述信息可以包括剧情的描述信息、人物（或角色）的描述信息、场景的描述信息、配乐的描述信息中至少一项。不同类型的描述信息可以通过不同的界面输入。例如，描述信息可以描述待生成的互动多媒体内容的故事背景，主角和相关的角色，故事大概的脉络和情节以及最后的结局等。

25 在步骤 S104 中，根据描述信息和目标体裁，生成内容文本。

目标体裁（或目标媒体类型）即最后生成的互动多媒体内容的体裁（或媒体类型）。在一些实施例中，目标体裁包括：游戏、互动电视剧、互动电影中至少一项，不限于所举示例。

例如，采用生成式模型根据描述信息和目标体裁，生成内容文本。生成式模型用于基于输入的信息，输出目标内容。输入的信息包括生成式模型在生成过程中的处理

依据，例如参考哪些信息执行生成过程、输出的目标内容的要求等等。生成式模型例如包括基于文本进行生成的模型、或基于图像进行生成的模型，生成式模型的输出可以包括文本、图像或二者的结合。当然，生成式模型的输入或输出也可以是其他模态的数据，如音频、视频或者多种类型的数据的结合。生成式模型可以为单一模态的模型，例如基于文本生成文本的模型（简称“文生文模型”）、基于图像生成图像的模型（简称“图生图模型”）；或者，生成式模型也可以是跨模态的模型，即输入和输出属于不同模态的模型，例如基于文本生成图像的模型（简称“文生图模型”）；再或者，生成式模型的输入可以包括多个模态，输出也可以包括多个模态。

内容文本用于描述待生成的互动多媒体内容的剧情、呈现形式等。在一些实施例中，内容文本包括剧情内容的描述文本、目标体裁对应的视觉变换的描述文本中至少一项。

例如，剧情内容的描述文本包括基于用户输入的描述信息中剧情的描述信息采用生成式模型扩写后的剧情信息，相对于用户输入的描述信息，故事内容更加的丰富，细节更多，更加接近真实的剧本中的剧情信息。

在一些实施例中，视觉变换包括镜头的运转、背景的切换、人物姿态的变化中至少一项。例如，游戏的画面相对固定，人物姿态比较单一，镜头的运转也很少甚至没有，互动电视剧的叙事速度较慢，一个镜头持续的时间可能比较长，镜头的变换频率较低，样式也较少，同样人物姿态和背景持续的时间都比较长，互动电影的叙事速度较快，镜头的变换频率比较高，样式也比较多，人物和背景更加的动态。因此，可以针对不同的目标体裁生成不同的视觉变换的描述文本。

在步骤 S106 中，根据内容文本，生成互动多媒体内容。

例如，采用生成式模型根据内容文本，生成互动多媒体内容。这里的生成式模型可以是文生图模型或文生视频模型等。由于内容生成文本已经包括了剧情内容的描述文本和视觉变换的描述文本，基于这些文本可以生成对应的画面、人物、声音等并呈现出对应的视觉效果。

在步骤 S108 中，显示互动多媒体内容以及内容文本。

可以将互动多媒体内容和内容文本同时显示，例如在界面中不同区域分别显示，并且对应（或关联）显示。例如，基于互动多媒体内容显示的进度内容，文本可以滚动显示。也可以在不同界面中分别显示互动多媒体内容和内容文本。例如，首先显示互动多媒体内容，响应于对内容文本的显示的触发操作，在浮层或新的窗口或新的页

面中显示内容文本。

在一些实施例中，互动多媒体内容可以被划分为一个或多个剧情节点的多媒体内容，内容文本包括每个剧情节点的内容文本。将每个剧情节点的多媒体内容和该剧情节点的内容文本关联显示。

5 步骤 S104 和 S106 可以在客户端实现，也可以在服务器端实现，在服务器端实现的情况下，客户端将描述信息发送至服务器端，并接收服务器端返回的内容文本和互动多媒体内容，进行显示。本公开中显示相关步骤都可以通过图形用户界面(**Graphical User Interface, GUI**)实现。

10 上述实施例的方法，基于用户输入的描述信息和目标体裁可以自动生成内容文本，并基于内容文本自动生成互动多媒体内容。针对目标体裁还可以生成与之匹配的视觉变换的描述文本，从而生成的互动多媒体内容呈现出与目标体裁对应的视觉变换效果。通过上述实施例的方法，用户只需要简单的输入描述信息，就可以生成扩写后的内容文本，并进一步生成不同体裁的互动多媒体内容，提高了互动多媒体内容的创作效率，节约了资源。普通用户也可以参与故事的创作和发展，以更好地与角色互动，创造属于自己的独特故事，提高了互动多媒体内容的创作的可行性，可以将更多的创意转换为互动多媒体产品。

15 在一些实施例中，对描述信息进行解析，确定描述信息对应的体裁(或媒体类型)，作为目标体裁；或者，接收用户对多个体裁中一个或多个体裁的选择操作，将一个或多个体裁确定为目标体裁。

20 例如，描述信息中可以包括用户希望创作的互动多媒体内容的体裁对应的关键词，例如，游戏、电影、电视剧等，可以对这些关键词进行识别确定目标体裁。又例如，不同体裁的描述信息的内容可能不同，游戏包含的互动情节更多，剧情更简单，互动电视剧的剧情更长，情节更多等。因此，可以利用机器学习模型对描述信息进行解析和理解，确定描述信息对应的体裁，作为目标体裁。机器学习模型可以是 **AI(Artificial Intelligence, 人工智能)** 模型、大模型等。具体的可以是神经网络、深度学习模型、生成式模型等。

25 在对描述信息进行解析确定描述信息对应的体裁后，可以显示确定的体裁，响应于用户的确认操作，将确认的体裁作为目标体裁。

30 用户可以在客户端的应用(**APP**)或 **Web** 页面中进行操作，创作互动多媒体内容。例如，响应于应用的开启，显示首页，首页中可以设置创作控件和引导信息，引导信

息例如为“创作你的互动故事”，响应于创作控件的触发，进入创作界面。首页或创作界面中可以包括多个体裁的选择控件，用户可以对一个或多个体裁的选择操作，将一个或多个体裁确定为目标体裁。如果用户选择了多个体裁，可以分别生成多个互动多媒体内容。

5 在一些实施例中，描述信息可以包括剧情的描述信息、人物（或角色）的描述信息、场景的描述信息、配乐的描述信息中至少一项。

例如，剧情的描述信息包括：剧情概述信息以及一个或多个剧情节点中每个剧情节点的剧情信息中至少一项。剧情概述信息可以用于描述待生成的互动多媒体内容的故事内容，例如，剧情概述信息可以包括故事的全局背景、开端、经过、结局中至少
10 一项的简要描述信息。故事的经过和结局可以有多种情况，可以分别描述。

创作界面可以区分不同的面板，例如，如图 2 所示，创作界面中包括故事面板、人物面板等。可以响应于用户对故事面板的触发操作，显示故事面板，在故事面板中可以输入剧情的描述信息。例如，可以在故事面板中显示故事内容的输入区域，接收用户输入的剧情概述信息，例如，故事内容的输入区域可以输入故事的全局背景和/或
15 开端，还可以设置结局（或成功条件）的输入区，用于输入故事的结局（或成功条件）。

可以将待生成的互动多媒体内容的整体剧情划分为一个或多个剧情节点（也可以称为情节或章节）。在一些实施例中，可以根据剧情概述信息生成一个或多个剧情节点，以及每个剧情节点的剧情信息。可以采用文本生成模型对剧情概述信息进行拆分，生成每个剧情节点的剧情信息，进而根据每个剧情节点的剧情信息生成每个剧情节点
20 的剧情内容的描述文本。即生成的剧情内容的描述文本包括一个或多个剧情节点中每个剧情节点的剧情内容的描述文本。

用户输入的剧情的描述信息也可以包括一个或多个剧情节点中每个剧情节点的剧情信息。故事面板中可以区分剧情节点（也可以称为情节或章节）分别接收用户的输入，每个剧情节点的剧情信息用于描述每个剧情节点的故事内容。每个剧情节点
25 的故事内容可以包括该节点的背景、开端、经过、结局中至少一项的简要描述信息。

如图 2 所示，显示故事面板，并在故事面板中显示情节添加控件 201，用于添加一个或多个剧情节点，通过每个剧情节点的输入区域接收用户输入的每个剧情节点的剧情信息。每个剧情节点的输入区域可以包括该剧情节点的故事内容描述区域 202、该剧情节点的结局描述区域 203，还可以包括发言角色选择区域 204、音乐选择（上传）区域 205、图片选择（上传）区域 206 中至少一项。不同区域中可以设置相对应
30

的引导信息，用于引导用户输入。用户选择或上传的音乐或图片，可以作为背景音乐或背景（场景）图片用于生成剧情节点对应的多媒体内容。如果描述信息中不包含这些具体的信息，则可以通过生成式模型自动生成。

下面结合图 3 描述本公开互动多媒体内容的生成方法的另一些实施例。

图 3 为本公开互动多媒体内容的生成方法另一些实施例的流程图。如图 3 所示，该实施例的方法包括：步骤 S302~S310。

在步骤 S302 中，接收用户输入的描述信息。

在步骤 S304 中，采用文本生成模型对描述信息进行解析，生成一个或多个剧情节点中每个剧情节点的剧情内容的描述文本。

如果描述信息中只包括剧情概述信息，可以采用文本生成模型先根据剧情概述信息生成一个或多个剧情节点中每个剧情节点的剧情信息，再根据每个剧情节点的剧情信息，生成每个剧情节点的剧情内容的描述文本。

如果描述信息只包括每个剧情节点的剧情信息，可以采用文本生成模型根据每个剧情节点的剧情信息生成每个剧情节点的剧情内容的描述文本。

如果描述信息同时包括剧情概述信息和每个剧情节点的剧情信息，可以采用文本生成模型结合剧情概述信息和每个剧情节点的剧情信息，生成每个剧情节点的剧情内容的描述文本。

在一些实施例中，采用文本生成模型根据描述信息和目标体裁，生成每个剧情节点的剧情内容的描述文本。每个剧情节点的剧情内容的描述文本可以参考目标体裁进行不同程度的扩写，例如，在目标体裁为游戏的情况下，每个剧情节点的剧情内容的描述文本可以比较简单，在目标体裁为互动电视剧的情况下，每个剧情节点的剧情内容的描述文本会更加复杂。

在一些实施例中，采用文本生成模型，根据描述信息和剧情生成提示信息，生成每个剧情节点的剧情内容的描述文本。剧情生成提示信息用于辅助生成每个剧情节点的剧情内容的描述文本。例如，剧情生成提示信息包括与剧情类型和目标体裁相关的提示信息等，剧情类型可以是根据描述信息确定的。例如，探险类游戏的提示信息可以包括探险类剧情的示例，动作类游戏的提示信息可以包括角色打斗的示例。

在一些实施例中，在剧情的描述信息包括多个剧情节点的剧情信息的情况下，用户输入的描述信息包括剧情的描述信息还可以包括：多个剧情节点之间的逻辑关系。多个剧情节点之间的逻辑关系用于各个剧情节点的先后顺序、由一个剧情节点进入下

一个剧情节点的条件信息等。

交互操作节点可以作为一类特殊的剧情节点，包括在剧情的发展过程中，需要玩家（或观看用户或操作用户）进行互动的情节。不同的交互操作节点可以引导剧情走向不同的分支。在一些实施例中，描述信息还包括：交互操作节点的描述信息和/或剧情节点和交互操作节点之间的逻辑关系。

如果用户输入的描述信息不包括多个剧情节点之间的逻辑关系、交互操作节点的描述信息、剧情节点和交互操作节点之间的逻辑关系，这些信息可以由生成式模型根据用户输入的描述信息自动生成。最终生成互动多媒体内容时可以根据多个剧情节点之间的逻辑关系，剧情节点和交互操作节点之间的逻辑关系，将每个剧情节点的多媒体内容和每个交互操作节点的多媒体内容进行整合。

在一些实施例中，显示设置界面，接收用户对剧情节点和交互操作节点之间的逻辑关系的设置，显示设置后的剧情节点和交互操作节点之间的逻辑关系。

设置界面中可以显示根据用户输入的描述信息自动生成的初始设置信息，例如，初始设置信息包括各个剧情节点、各个交互操作节点以及剧情节点和交互操作节点之间的逻辑关系。用户可以对初始设置信息进行调整。用户也可以直接在设置界面中设置各个剧情节点，各个交互操作节点以及剧情节点和交互操作节点之间的逻辑关系。

如图 4 所示，故事设置界面中以流程图的形式显示，更加便于用户清楚准确的设置剧情节点和交互操作节点之间的逻辑关系，但不限于所举示例。流程图中包括多个剧情节点 401 以及多个交互操作节点 402，节点之间的连线表示它们之间的逻辑关系。每个节点可以以节点（情节或章节）名称或简介显示。该流程图可以是通过生成式模型根据用户输入的描述信息自动生成的，用户可以进行调整或确认。该流程图也可以是根据用户输入的一个或多个剧情节点的剧情信息（例如，图 2 所示）和/或一个或多个交互操作节点的描述信息生成的。或者该流程图可以由用户直接配置的。

如图 4 所示，交互操作节点用于配置可以供玩家（或观看用户或操作用户）交互的剧情，通过不同的交互操作节点可以走向不同的剧情分支。针对剧情节点或交互操作节点，不仅可以以流程图的形式显示，还可以响应于用户对每个剧情节点或交互操作节点的配置或选择，显示生成的任意剧情节点的剧情内容的描述文本和/或任意交互操作节点的描述文本，以及生成的任意剧情节点的多媒体内容和和/或任意交互操作节点的多媒体内容。

在一些实施例中，采用文本生成模型对描述信息进行解析，生成每个交互操作节

点的描述文本，其中，每个交互操作节点的描述文本包括对操作用户显示的交互操作选项的描述文本。

如图 4 所示，针对交互操作节点“岔路口”可以基于用户的配置或选择，显示该交互操作节点的内容界面 403。内容界面可以以窗口或浮层等形式与节点对应显示。

内容界面中可以包括该交互操作节点的描述文本。例如，在该交互操作节点的剧情中，宇航员们走到岔路口，宇航员 x 发言“这里看起来是一个岔路口，我们接下来该往哪个方向去？”等。进一步，还可以显示每个交互操作选项的描述文本，例如，A. 探索高山，B. 探索盆地…。这样可以使用户非常直观准确的看到每个节点对应的描述文本，并且可以进行调整。

每个交互操作节点的描述文本可以属于剧情内容的描述文本中的一部分。采用文本生成模型对描述信息进行解析，不仅可以生成每个剧情节点的剧情内容的描述文本，还可以生成每个交互操作节点的描述文本。

互动多媒体内容中人物（或角色）的设计是非常重要的部分，人物的发言或对白也是剧情的重要组成部分。例如，生成的剧情内容的描述文本中每个剧情节点的剧情内容的描述文本可以包括不同人物的发言，这些发言一方面需要基于剧情的发展确定，另一方面也要根据人物的特点、人物关系、形象等确定。

人物的描述文本可以由生成式模型根据描述信息生成。人物的描述文本可以包括名称、性格、身份、人物关系、说话风格、声音特点中至少一项，但不限于所举示例。在一些实施例中，用户输入的描述信息包括一个或多个个人物的人物设定信息。每个人物的人物设定信息也可以包括名称、性格、身份、人物关系、说话风格、声音特点、形象中至少一项，但不限于所举示例。人物的设定信息相对于人物的描述文本更加简略，例如，采用文本生成模型对描述信息进行解析，生成每个人物的描述文本。

创作界面中可以包括人物面板，接收用户对人物面板的触发操作，显示人物面板，在人物面板中可以输入一个或多个个人物的设定信息。如图 5 所示，显示人物面板，并在人物面板中显示角色添加控件 501，用于添加一个或多个个人物（或角色），通过每个人物的输入区域接收用户输入的人物的设定信息。每个人物的输入区域可以包括该人物的名称设置区域 502、该人物的基础设定区域 503，还可以包括台词风格设置区域 504、人物音乐设置区域 505、人物头像设置区域 506、人物立绘设置区域 507 中至少一项。不同区域中可以设置相对应的引导信息，用于引导用户输入。图 5 中每个人物的输入区域可以显示根据描述信息生成的初始描述文本，用户可以进行调整。

在一些实施例中，采用文本生成模型，根据描述信息和人物生成提示信息，生成每个人物的描述文本。人物生成提示信息可以与不同人物类型相对应，例如，人物生成提示信息包括不同类型人物的语言示例等。

用户输入的描述信息中还可以包括场景设定信息、画面风格设定信息中至少一项。这些也可以辅助生成剧情内容的描述文本。场景设定信息例如可以描述背景中景物等。画面风格设定信息例如为动漫风格、中国风等。

通过上述实施例的方法，基于用户输入的描述信息，可以自动生成剧情内容的描述文本，提高了互动多媒体内容的生成效率，降低了创作门槛，使普通用户可以自由的创造互动多媒体内容。

在步骤 S306 中，采用文本生成模型，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息，并生成每个剧情节点对应的视觉变换的描述文本。

在一些实施例中，描述信息包括视觉变换的描述信息和/或每个剧情节点对应的视觉变换的描述信息。例如，视觉变换的描述信息可以包括在角色说话或动作时采用近景镜头，在进入一个情节或章节的情况下，采用由远拉近的方式运转镜头，角色说话时配合肢体动作，背景随着时间的变化进行变化等概括性的描述信息。每个剧情节点对应的视觉变换描述信息可以更具一些，例如，设置了一个剧情节点中说话的角色，结合剧情信息该剧情节点对应的视觉变换描述信息可以包括：角色边跑边说话，对角色进行特写，背景随着角色奔跑变化。

如果描述信息中包括视觉变换的描述信息和/或每个剧情节点对应的视觉变换的描述信息，可以采用文本生成模型根据描述信息和目标体裁进行扩展，确定每个剧情节点中不同画面之间的视觉变换的信息。

普通用户可能并不善于描述视觉变换的方式，因此，描述信息中不包括视觉变换的描述信息和/或每个剧情节点对应的视觉变换的描述信息的情况会比较多。可以采用文本生成模型直接根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息。

在一些实施例中，采用文本生成模型，根据描述信息、目标体裁和视觉变换提示信息，确定每个剧情节点中不同画面之间的视觉变换的信息，并生成每个剧情节点对应的视觉变换的描述文本。视觉变换提示信息用于辅助生成视觉变换的描述文本。例如，视觉变换提示信息包括与目标体裁对应的提示信息。例如，互动电影的一些镜头

运用的示例等。

每个剧情节点的剧情内容的描述文本和视觉变换的描述文本可以采用文本生成模型同时生成。剧情内容的描述文本和视觉变换的描述文本可以在同一个文本中，并不是完全割裂的。例如，在剧情内容的描述文本中穿插视觉变换的描述文本。

5 不同画面之间的视觉变换的信息具体可以包括每个画面的视觉效果信息。

在一些实施例中，视觉变换包括镜头的运转，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中不同画面中运用不同类型的镜头效果；针对确定运用不同类型的镜头效果的剧情节点，根据描述信息，确定该剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，作为该剧情节点中不同画面之间的镜头的运转的信息。

对于一些简单的游戏，可能每个剧情节点（情节或章节）中不需要变换镜头效果，对于互动电视剧或互动电影则需要变换镜头效果。或者，对于一些剧情节点需要变换镜头效果，对于另一些剧情节点不需要变换镜头效果。不需要镜头变换的剧情节点对

15 应的镜头的运转的信息可以为保持镜头效果不变。

镜头效果根据视觉距离的不同可以分为近景、全景、中景、远景等，根据镜头的运动方式不同，可以分为推、拉、摇、升、降等。针对每个画面（或每帧图像），该画面所运用的目标镜头效果可以是独立的，例如目标镜头效果是近景；该画面所运用的目标镜头效果也可以是与其它画面关联的，例如，在连续多个画面中采用逐渐升高镜头的目标镜头效果。目标镜头效果对应的目标对象可以是特定的人、特定的物或者全景等。

在一些实施例中，根据描述信息，确定该剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象包括：根据描述信息，生成每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本，针对运用不同类型的镜头效果的剧情节点，根据该剧情节点的剧情信息或该剧情节点的剧情内容的描述文本，确定该剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象。每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本包括根据描述信息采用文本生成模型进行扩写后的剧情。由于剧情会影响镜头效果，例如，人物说话时更多的会采用近景镜头。因此，根据每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本确定的目标镜头效果和目标对象更加准确，并且效果更加丰富。

25

30

在一些实施例中，视觉变换包括人物姿态的变化，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中变化人物的姿态；针对确定人物的姿态的剧情节点，根据描述信息，确定该剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，作为剧情节点中不同画面之间每个人物的姿态的变化信息。

人物（或角色）的姿态可以包括姿势、动作、表情中至少一项。对于一些简单的游戏，可能大部分剧情节点不需要变换人物的姿态或者人物姿态的变化很简单，而对于互动电视剧或互动电影，人物姿态的变化比较复杂。用户输入的描述信息中也可能包括人物的姿势、动作、表情等信息，如果不包括，可以采用生成式模型直接生成。因此，结合描述信息和目标体裁可以确定每个剧情节点中每个画面中的人物的目标姿态。

在一些实施例中，根据描述信息，生成每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本，针对变化人物的姿态的剧情节点，根据该剧情节点的剧情信息或该剧情节点的剧情内容的描述文本，确定该剧情节点中每个人物的目标姿态。每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本包括根据描述信息采用文本生成模型进行扩写后的剧情。由于剧情会影响人物的姿态，例如，人物完成一项任务时可以表现出成功的表情和动作。因此，根据每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本确定的目标姿态更加准确，并且效果更加丰富。

用户输入的描述信息中也包括人物姿态的描述信息，人物姿态的描述信息可能比较简单，通过生成式模型可以生成更加丰富的描述文本。

在一些实施例中，视觉变换包括背景的切换，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中切换背景；针对确定切换背景的剧情节点，根据描述信息，确定剧情节点中每个画面的目标背景，作为每个剧情节点中不同画面之间背景的切换信息。

对于一些简单的游戏，可能每个剧情节点（情节或章节）中不需要切换背景，对于互动电视剧或互动电影则背景的切换则更加频繁。或者，对于一些剧情节点需要切换背景，对于另一些剧情节点不需要切换背景。

针对每个画面（或每帧图像），该画面所运用的目标背景可以是独立的，也可以是与其他画面关联的，例如，在连续多个画面中根据人物的运动不断地变换背景等。

在一些实施例中，根据描述信息，确定该剧情节点中每个画面中所运用的目标背景包括：根据描述信息，生成每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本，针对切换背景的剧情节点，根据该剧情节点的剧情信息或该剧情节点的剧情内容的描述文本，确定该剧情节点中每个画面中的目标背景。每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本包括根据描述信息采用文本生成模型进行扩写后的剧情。由于剧情会影响背景，因此，根据每个剧情节点的剧情信息或每个剧情节点的剧情内容的描述文本确定的目标背景更加准确，并且效果更加丰富。

用户输入的描述信息中也包括背景的切换的描述信息，背景的切换的描述信息可能比较简单，通过生成式模型可以生成更加丰富的描述文本。

在一些实施例中，采用文本生成模型，根据描述信息和视觉变换的提示信息，生成每个剧情节点的视觉变换的描述文本。视觉变换的提示信息用于辅助生成每个剧情节点的视觉变换的描述文本。例如，视觉变换的提示信息包括与目标体裁相关的提示信息等。例如，针对互动电影或互动电视剧，在人物有大段对白时更多的采用特写镜头。

通过上述实施例的方法，根据描述信息和目标体裁，针对不同类型的视觉变换可以生成响应的描述文本，使得生成的互动多媒体内容可以更加准确的与目标体裁相匹配，并且提高了互动多媒体内容的生成效率。

在步骤 S308 中，根据内容文本，生成互动多媒体内容。

可以采用文生图和/或文生视频模型，基于内容文本生成互动多媒体内容。生成的互动多媒体媒体可以包括：完整的互动多媒体内容、每个剧情节点的多媒体内容、每个交互操作节点的多媒体内容、每个人物的多媒体内容、每个场景的多媒体内容等。完整的互动多媒体内容可以根据每个剧情节点的多媒体内容、每个交互操作节点的多媒体内容、每个人物的多媒体内容、每个场景的多媒体内容以及剧情节点和交互操作节点中各个节点之间的逻辑关系生成，不同节点之间可以加入转场特效等。

在一些实施例中，根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的视觉变换的描述文本，生成每个剧情节点的多媒体内容，其中，每个剧情节点的多媒体内容呈现与该剧情节点对应的视觉变换的描述文本相应的视觉效果。

剧情内容的描述文本和视觉变换的描述文本可以是穿插在一起的整段文本，以自然语言描述的，可以直接用于生成多媒体内容。

在一些实施例中，在每个剧情节点对应的视觉变换的描述文本包括每个剧情节点

中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象的情况下，根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，生成每个剧情节点的每个画面，其中，每个画面中的目标对象采用对应的目标镜头效果进行显示。

5 在一些实施例中，在每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面中一个或多个个人物中每个人物的目标姿态的情况下，根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面中一个或多个个人物中每个人物的目标姿态，生成每个剧情节点的每个画面，其中，每个画面中的每个人物采用对应的目标姿态进行显示。

10 在一些实施例中，在每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面的目标背景的情况下，根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面的目标背景，生成每个剧情节点的每个画面，其中，每个画面的背景为对应的目标背景。

上述各个实施例可以任意结合，基于剧情内容的描述文本和视觉变换的描述文本，
15 可以生成每个画面（图像）后形成视频，也可以直接生成视频包含多个画面。

在一些实施例中，根据每个交互操作节点的描述文本，生成每个交互操作节点的多媒体内容。交互操作节点可以作为一种特殊的剧情节点，也可以包含对应的视觉变换，生成交互操作节点的多媒体内容的方法，可以参考生成剧情节点的多媒体内容的方法，在此不再赘述。

20 在生成每个剧情节点的多媒体内容时，还会用到人物的描述文本、场景的描述文本、画面风格的描述文本，这些描述文本都可以根据用户输入的描述信息生成。人物的描述文本可以用于生成人物的立绘、声音音频等，场景的描述文本和画面风格的描述文本可以用于生成背景、画面中的景物等并用于确定整个互动多媒体内容的风格。

25 在一些实施例中，根据每个人物的描述文本，生成每个人物的立绘、声音音频中至少一项。人物的立绘可以是体现人物整体形象的图像。人物的立绘包括：正面静态人物立绘、多种不同人物姿态的立绘、添加动作（比如嘴巴的开合、裙摆或发丝的摆动等）的动图或视频等。针对互动电影/互动电视剧，背景场景以及其中的人物没有固定的姿态和画面布局。

互动多媒体内容还会包括背景音乐，例如，增加气氛的音乐、人物出场的音乐以及特效音乐等。
30

在一些实施例中，采用文本生成模型对描述信息进行解析，生成每个剧情节点对应的背景音乐的描述文本。例如，背景音乐的描述文本可以包括音乐风格、时长等。

在一些实施例中，根据每个剧情节点对应的背景音乐的描述文本，生成每个剧情节点对的背景音乐。

5 可以采用生成式模型直接生成背景音乐，也可以从数据库中选取与背景音乐的描述文本相匹配的背景音乐。

在生成互动多媒体内容时，内容文本可以包括前述各个实施例中的各种类型的描述文本的组合，根据内容文本可以直接生成互动多媒体内容。

10 例如，图 6 所示为生成的互动电影中的一幅画面，画面中包括人物、背景以及可以由观看用户进行互动的交互操作选项。该画面可以采用近景镜头，并且可以配有紧张氛围的背景音乐等。

在步骤 S310 中，显示互动多媒体内容以及内容文本。

可以区分多个剧情节点，将每个剧情节点的内容文本与对应的互动多媒体内容进行关联显示。

15 在一些实施例中，将每个剧情节点的剧情描述文本、每个剧情节点中不同画面之间的视觉变换的描述文本和每个剧情节点的多媒体内容进行关联显示。

在一些实施例中，将每个人物的描述文本和每个人物的立绘、声音音频中至少一项进行关联显示；将每个交互操作节点的描述文本和每个交互操作节点的多媒体内容进行关联显示。

20 人物的描述文本和人物的立绘、声音音频等可以单独进行显示。每个剧情节点或每个交互操作节点的描述文本和多媒体内容，可以分别单独显示，也可以与图 4 的流程图关联显示。

将内容文本与互动多媒体内容关联显示，便于用户进行调整。如果用户对生成的内容文本或互动多媒体内容不满意，可以通过调整内容文本重新生成互动多媒体内容。

25 如图 7 所示，为预览界面，针对探索高山这一剧情节点 701，在该剧情节点的内容界面 702 中，可以显示该剧情节点的内容文本 703，该剧情节点的内容文本包括剧情内容的描述文本、视觉效果的描述文本等，同时可以显示多媒体内容，多媒体内容包括场景图像 704、人物图像 705、背景音乐 706，还可以包括对应的视频等。如图 7 所示，还可以显示各个剧情节点之间的逻辑关系。

30 在一些实施例中，针对每个剧情节点，接收用户对剧情节点对应的视觉变换的描

述文本的调整；根据调整后的剧情节点对应的视觉变换的描述文本，生成剧情节点的多媒体内容；显示调整后的剧情节点对应的视觉变换的描述文本和调整后的剧情节点的多媒体内容。

还可以接收用户对剧情节点的剧情内容的描述文、交互操作节点的描述文本、人物
5 物的描述文本、场景的描述文本、画面风格的描述文本等进行调整，重新生成多媒体内容。

本公开的方法根据用户（作者）的描述信息可以生成不同体裁的互动多媒体内容。用户可以输入故事背景、人物设定、基本摘要等，生成式模型可以根据描述信息扩写出完整的剧本，并生成对应内容下的画面和人物，以生成剧本中每个剧情节点对应的
10 场景和人物。场景背景图、人物立绘、人物姿态、背景音乐等都可以基于用户输入的描述信息生成，并且在生成时可以生成多个选项以供用户选择。

生成式模型能够通过用户输入的描述信息，判断当前背景以及当前场景中的人物，并且能根据文本内容判断当前镜头的运转等。互动多媒体内容中的人物由 AI 模型扮演，玩家（或观看用户或操作用户）可以与其进行交互，用户（作者）还可以设置玩
15 家不同的选项对应的不同剧情分支，同时也支持玩家在游戏中进行互动时不仅依用户（作者）作者提供的选项，也能够自行输入内容与 NPC（Non-Player Character，非玩家角色）进行交互。

本公开的方法可以提高互动多媒体内容的生成效率和多样性。

本公开还提供一种互动多媒体内容的生成装置，下面结合图 8 进行描述。

图 8 为本公开互动多媒体内容的生成装置一些实施例的结构图。如图 8 所示，该
20 实施例的装置包括：接收模块 810，内容生成模块 820，多媒体生成模块 830，显示模块 840。

接收模块 810 被配置为接收用户输入的描述信息。

内容生成模块 820 被配置为根据描述信息和目标体裁，生成内容文本，其中，内
25 容文本包括剧情内容的描述文本以及目标体裁对应的视觉变换的描述文本。

多媒体生成模块 830 被配置为根据内容文本，生成互动多媒体内容。

显示模块 840 被配置为显示互动多媒体内容以及内容文本，其中，互动多媒体内容呈现与视觉变换的描述文本对应的视觉变换效果。

在一些实施例中，互动多媒体内容的生成装置 80 还包括体裁确定模块 850，被配
30 置为对描述信息进行解析，确定描述信息对应的体裁，作为目标体裁；或者接收模块

810 还被配置为接收用户对多个体裁中一个或多个体裁的选择操作，体裁确定模块 850，被配置为将一个或多个体裁确定为目标体裁。

5 在一些实施例中，剧情内容的描述文本包括一个或多个剧情节点中每个剧情节点的剧情内容的描述文本，视觉变换的描述文本包括每个剧情节点对应的视觉变换的描述文本，内容生成模块 820 被配置为采用文本生成模型对描述信息进行解析，生成每个剧情节点的剧情内容的描述文本；采用文本生成模型，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息，并生成每个剧情节点对应的视觉变换的描述文本。

10 在一些实施例中，视觉变换包括镜头的运转，内容生成模块 820 被配置为根据描述信息和目标体裁，确定是否在每个剧情节点中不同画面中运用不同类型的镜头效果；针对确定运用不同类型的镜头效果的剧情节点，根据描述信息，确定剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，作为剧情节点中不同画面之间的镜头的运转的信息。

15 在一些实施例中，视觉变换包括人物姿态的变化，内容生成模块 820 被配置为根据描述信息和目标体裁，确定是否在每个剧情节点中变化人物的姿态；针对确定变化人物的姿态的剧情节点，根据描述信息，确定剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，作为剧情节点中不同画面之间每个人物的姿态的变化信息。

20 在一些实施例中，视觉变换包括背景的切换，内容生成模块 820 被配置为根据描述信息和目标体裁，确定是否在每个剧情节点中切换背景；针对确定切换背景的剧情节点，根据描述信息，确定剧情节点中每个画面的目标背景，作为每个剧情节点中不同画面之间背景的切换信息

25 在一些实施例中，多媒体生成模块 830 被配置为根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的视觉变换的描述文本，生成每个剧情节点的多媒体内容，其中，每个剧情节点的多媒体内容呈现与该剧情节点对应的视觉变换的描述文本相应的视觉效果。

30 在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，多媒体生成模块 830 被配置为根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，生成每个剧情节点的每个画面，其中，每个画面中的目标对象采用对应的目标镜头效果进行显示。

在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面中一个或多个个人物中每个人物的目标姿态，多媒体生成模块 830 被配置为根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面中一个或多个个人物中每个人物的目标姿态，生成每个剧情节点的每个画面，其中，每个画面中的每个
5 人物采用对应的目标姿态进行显示。

在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面的目标背景，多媒体生成模块 830 被配置为根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面的目标背景，生成每个剧情节点的每个画面，其中，每个画面的背景为对应的目标背景。

10 在一些实施例中，显示模块 840 被配置为将每个剧情节点的剧情描述文本、每个剧情节点中不同画面之间的视觉变换的描述文本和每个剧情节点的多媒体内容进行关联显示。

在一些实施例中，针对每个剧情节点，接收模块 810 还被配置为接收用户对剧情节点对应的视觉变换的描述文本的调整；多媒体生成模块 830 还被配置根据调整后的
15 剧情节点对应的视觉变换的描述文本，生成剧情节点的多媒体内容；显示模块 840 还被配置为显示调整后的剧情节点对应的视觉变换的描述文本和调整后的剧情节点的多媒体内容。

在一些实施例中，内容文本还包括一个或多个个人物中每个人物的描述文本和一个或多个交互操作节点中每个交互操作节点的描述文本，内容生成模块 820 被配置为采用
20 文本生成模型对描述信息进行解析，生成每个人物的描述文本；采用文本生成模型对描述信息进行解析，生成每个交互操作节点的描述文本，其中，每个交互操作节点的描述文本包括对操作用户显示的交互操作选项的描述文本。

在一些实施例中，每个人物的描述文本包括每个人物的形象、声音特点中至少一项的描述信息，内容生成模块 820 被配置为根据每个人物的描述文本，生成每个人物
25 的立绘、声音音频中至少一项；根据每个交互操作节点的描述文本，生成每个交互操作节点的多媒体内容。

在一些实施例中，显示模块 840 被配置为将每个人物的描述文本和每个人物的立绘、声音音频中至少一项进行关联显示；将每个交互操作节点的描述文本和每个交互操作节点的多媒体内容进行关联显示。

30 在一些实施例中，内容文本还包括一个或多个剧情节点中每个剧情节点对应的背

景音乐的描述文本，内容生成模块 820 被配置为采用文本生成模型对描述信息进行解析，生成所每个剧情节点对应的背景音乐的描述文本。

在一些实施例中，多媒体生成模块 830 被配置为根据每个剧情节点对应的背景音乐的描述文本，生成每个剧情节点对的背景音乐。

5 在一些实施例中，描述信息包括剧情概述信息以及一个或多个剧情节点中每个剧情节点的剧情信息中至少一项。

在一些实施例中，描述信息还包括：各个剧情节点之间的逻辑关系、人物设定信息、场景设定信息、画面风格设定信息中至少一项。

在一些实施例中，目标体裁包括：游戏、互动电视剧、互动电影中至少一项。

10 应注意，上述各个单元仅是根据其所实现的具体功能划分的逻辑模块，而不是用于限制具体的实现方式，例如可以以软件、硬件或者软硬件结合的方式来实现。在实际实现时，上述各个单元可被实现为独立的物理实体，或者也可由单个实体（例如，处理器（CPU 或 DSP 等）、集成电路等）来实现。此外，上述各个单元在附图中用虚线示出指示这些单元可以并不实际存在，而它们所实现的操作/功能可由处理电路本身来实现。

15 此外，尽管未示出，该设备也可以包括存储器，其可以存储由设备、设备所包含的各个单元在操作中产生的各种信息、用于操作的程序和数据、将由通信单元发送的数据等。存储器可以是易失性存储器和/或非易失性存储器。例如，存储器可以包括但不限于随机存储存储器（RAM）、动态随机存储存储器（DRAM）、静态随机存取存储器（SRAM）、只读存储器（ROM）、闪存存储器。当然，存储器可也位于该设备之外。可选地，尽管未示出，但是该设备也可以包括通信单元，其可用于与其它装置进行通信。在一个示例中，通信单元可以被按照本领域已知的适当方式来实现，例如包括天线阵列和/或射频链路等通信部件，各种类型的接口、通信单元等等。这里将不再详细描述。此外，设备还可以包括未示出的其它部件，诸如射频链路、基带处理单
20 元、网络接口、处理器、控制器等。这里将不再详细描述。

25 本公开的一些实施例还提供一种电子设备。图 9 示出本公开的电子设备的一些实施例的框图。例如，在一些实施例中，电子设备 9 可以为各种类型的设备，例如可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA(个人数字助理)、PAD(平板电脑)、PMP(便携式多媒体播放器)、车载终端(例如车载导航终端)等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。例如，电子设备 9 可以包括
30

显示面板，以用于显示根据本公开的方案中所利用的数据和/或执行结果。例如，显示面板可以为各种形状，例如矩形面板、椭圆形面板或多边形面板等。另外，显示面板不仅可以为平面面板，也可以为曲面面板，甚至球面面板。

如图 9 所示，该实施例的电子设备 9 包括：存储器 91 以及耦接至该存储器 91 的处理器 92。应当注意，图 9 所示的电子设备 9 的组件只是示例性的，而非限制性的，根据实际应用需要，该电子设备 9 还可以具有其他组件。处理器 92 可以控制电子设备 9 中的其它组件以执行期望的功能。

在一些实施例中，存储器 91 用于存储一个或多个计算机可读指令。处理器 92 用于运行计算机可读指令时，计算机可读指令被处理器 92 运行时实现根据上述任一实施例的方法。关于该方法的各个步骤的具体实现以及相关解释内容可以参见上述的实施例，重复之处在此不作赘述。

例如，处理器 92 和存储器 91 之间可以直接或间接地互相通信。例如，处理器 92 和存储器 91 可以通过网络进行通信。网络可以包括无线网络、有线网络、和/或无线网络和有线网络的任意组合。处理器 92 和存储器 91 之间也可以通过系统总线实现相互通信，本公开对此不作限制。

例如，处理器 92 可以体现为各种适当的处理器、处理装置等，诸如中央处理器(CPU)、图形处理器(Graphics Processing Unit, GPU)、网络处理器(NP)等；还可以是数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。中央处理元(CPU)可以为 X86 或 ARM 架构等。例如，存储器 91 可以包括各种形式的计算机可读存储介质的任意组合，例如易失性存储器和/或非易失性存储器。存储器 91 例如可以包括系统存储器，系统存储器例如存储有操作系统、应用程序、引导装载程序 (Boot Loader)、数据库以及其他程序等。在存储介质中还可以存储各种应用程序和各种数据等。

另外，根据本公开的一些实施例，根据本公开的各种操作/处理在通过软件和/或固件实现的情况下，可从存储介质或网络向具有专用硬件结构的计算机系统，例如图 10 所示的计算机系统（或电子设备）1000 安装构成该软件的程序，该计算机系统在安装有各种程序时，能够执行各种功能，包括诸如前文的功能等等。图 10 是示出根据本公开的实施例的中可采用的计算机系统的示例结构的框图。

在图 10 中，中央处理单元 (CPU) 1001 根据只读存储器 (ROM) 1002 中存储的程序或从存储部分 1008 加载到随机存取存储器 (RAM) 1003 的程序执行各种处理。

在 RAM 1003 中，也根据需要存储当 CPU 1001 执行各种处理等时所需的数据。中央处理单元仅仅是示例性的，其也可以是其它类型的处理器，诸如前文的各种处理器。ROM 1002、RAM 1003 和存储部分 1008 可以是各种形式的计算机可读存储介质，如下文。需要注意的是，虽然图 10 中分别示出了 ROM 1002、RAM 1003 和存储装置 1008，但是它们中的一个或多个可以合并或者位于相同或不同的存储器或存储模块中。

CPU 1001、ROM 1002 和 RAM 1003 经由总线 1004 彼此连接。输入/输出接口 1005 也连接到总线 1004。

下述部件连接到输入/输出接口 1005：输入部分 1006，诸如触摸屏、触摸板、键盘、鼠标、图像传感器、麦克风、加速度计、陀螺仪等；输出部分 1007，包括显示器，比如阴极射线管（CRT）、液晶显示器（LCD），扬声器，振动器等；存储部分 1008，包括硬盘，磁带等；和通信部分 1009，包括网络接口卡比如 LAN 卡、调制解调器等。通信部分 1009 允许经由网络比如因特网执行通信处理。容易理解的是，虽然图 10 中示出计算机系统 1000 中的各个装置或模块是通过总线 1004 来通信的，但它们也可以通过网络或其它方式进行通信，其中，网络可以包括无线网络、有线网络、和/或无线网络和有线网络的任意组合。

根据需要，驱动器 1010 也连接到输入/输出接口 1005。可拆卸介质 1011 比如磁盘、光盘、磁光盘、半导体存储器等等根据需要被安装在驱动器 1010 上，使得从中读出的计算机程序根据需要被安装到存储部分 1008 中。

在通过软件实现上述系列处理的情况下，可以从网络比如因特网或存储介质比如可拆卸介质 1011 安装构成软件的程序。

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部分 1009 从网络上被下载和安装，或者从存储部分 1008 被安装，或者从 ROM 1002 被安装。在该计算机程序被 CPU 1001 执行时，执行本公开实施例的方法中限定的上述功能。

需要说明的是，在本公开的上下文中，计算机可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。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(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电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

在一些实施例中，还提供了一种计算机程序，包括：指令，指令当由处理器执行时使处理器执行上述任一个实施例的方法。例如，指令可以体现为计算机程序代码。

在本公开的实施例中，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操作的计算机程序代码，上述程序设计语言包括但不限于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。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。

5 描述于本公开实施例中所涉及到的模块、部件或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。其中，模块、部件或单元的名称在某种情况下并不构成对该模块、部件或单元本身的限定。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示例性的硬件逻辑部件包括：现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

10 根据本公开的一些实施例，提供一种互动多媒体内容的生成方法，包括：接收用户输入的描述信息；根据描述信息和目标体裁，生成内容文本，其中，内容文本包括剧情内容的描述文本以及目标体裁对应的视觉变换的描述文本；根据内容文本，生成互动多媒体内容；显示互动多媒体内容以及内容文本，其中，互动多媒体内容呈现与视觉变换的描述文本对应的视觉变换效果。

在一些实施例中，该生成方法还包括：对描述信息进行解析，确定描述信息对应的体裁，作为目标体裁；或者接收用户对多个体裁中一个或多个体裁的选择操作，将一个或多个体裁确定为目标体裁。

20 在一些实施例中，剧情内容的描述文本包括一个或多个剧情节点中每个剧情节点的剧情内容的描述文本，视觉变换的描述文本包括每个剧情节点对应的视觉变换的描述文本，根据描述信息和目标体裁，生成内容文本包括：采用文本生成模型对描述信息进行解析，生成每个剧情节点的剧情内容的描述文本；采用文本生成模型，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息，并生成每个剧情节点对应的视觉变换的描述文本。

25 在一些实施例中，视觉变换包括镜头的运转，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中不同画面中运用不同类型的镜头效果；针对确定运用不同类型的镜头效果的剧情节点，根据描述信息，确定剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，作为剧情节点中不同画面之间的镜头的运

转的信息。

在一些实施例中，视觉变换包括人物姿态的变化，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中变化人物的姿态；针对确定变化人物的姿态的剧情节点，
5 根据描述信息，确定剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，作为剧情节点中不同画面之间每个人物的姿态的变化信息。

在一些实施例中，视觉变换包括背景的切换，根据描述信息和目标体裁，确定每个剧情节点中不同画面之间的视觉变换的信息包括：根据描述信息和目标体裁，确定是否在每个剧情节点中切换背景；针对确定切换背景的剧情节点，根据描述信息，
10 确定剧情节点中每个画面的目标背景，作为每个剧情节点中不同画面之间背景的切换信息。

在一些实施例中，根据内容文本，生成互动多媒体内容包括：根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的视觉变换的描述文本，生成每个剧情节点的多媒体内容，其中，每个剧情节点的多媒体内容呈现与该剧情节点对应的视觉变
15 换的描述文本相应的视觉效果。

在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的视觉变换的描述文本，生成每个剧情节点的多媒体内容包括：根据每个剧情节点的剧情内容的描述文本和每个剧情节点
20 中每个画面中所运用的目标镜头效果以及目标镜头效果对应的目标对象，生成每个剧情节点的每个画面，其中，每个画面中的目标对象采用对应的目标镜头效果进行显示。

在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的视觉变换的描述文本，生成每个剧情节点的多媒体内
25 容包括：根据每个剧情节点的剧情内容的描述文本和每个剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，生成每个剧情节点的每个画面，其中，每个画面中的每个人物采用对应的目标姿态进行显示。

在一些实施例中，每个剧情节点对应的视觉变换的描述文本包括每个剧情节点中每个画面的目标背景，根据每个剧情节点的剧情内容的描述文本和每个剧情节点对应的
30 视觉变换的描述文本，生成每个剧情节点的多媒体内容包括：根据每个剧情节点的

剧情内容的描述文本和每个剧情节点中每个画面的目标背景，生成每个剧情节点的每个画面，其中，每个画面的背景为对应的目标背景。

5 在一些实施例中，显示互动多媒体内容以及内容文本包括：将每个剧情节点的剧情描述文本、每个剧情节点中不同画面之间的视觉变换的描述文本和每个剧情节点的多媒体内容进行关联显示。

在一些实施例中，该生成方法还包括，针对每个剧情节点：接收用户对剧情节点对应的视觉变换的描述文本的调整；根据调整后的剧情节点对应的视觉变换的描述文本，生成剧情节点的多媒体内容；显示调整后的剧情节点对应的视觉变换的描述文本和调整后的剧情节点的多媒体内容。

10 在一些实施例中，内容文本还包括一个或多个人物中每个人物的描述文本和一个或多个交互操作节点中每个交互操作节点的描述文本，根据描述信息和目标体裁，生成内容文本包括：采用文本生成模型对描述信息进行解析，生成每个人物的描述文本；采用文本生成模型对描述信息进行解析，生成每个交互操作节点的描述文本，其中，每个交互操作节点的描述文本包括对操作用户显示的交互操作选项的描述文本。

15 在一些实施例中，每个人物的描述文本包括每个人物的形象、声音特点中至少一项的描述信息，根据内容文本，生成互动多媒体内容包括：根据每个人物的描述文本，生成每个人物的立绘、声音音频中至少一项；根据每个交互操作节点的描述文本，生成每个交互操作节点的多媒体内容。

20 在一些实施例中，显示互动多媒体内容以及内容文本包括：将每个人物的描述文本和每个人物的立绘、声音音频中至少一项进行关联显示；将每个交互操作节点的描述文本和每个交互操作节点的多媒体内容进行关联显示。

在一些实施例中，内容文本还包括一个或多个剧情节点中每个剧情节点对应的背景音乐的描述文本，根据描述信息和目标体裁，生成内容文本包括：采用文本生成模型对描述信息进行解析，生成所每个剧情节点对应的背景音乐的描述文本。

25 在一些实施例中，根据内容文本，生成互动多媒体内容包括：根据每个剧情节点对应的背景音乐的描述文本，生成每个剧情节点对的背景音乐。

在一些实施例中，描述信息包括剧情概述信息以及一个或多个剧情节点中每个剧情节点的剧情信息中至少一项。

30 在一些实施例中，描述信息还包括：各个剧情节点之间的逻辑关系、人物设定信息、场景设定信息、画面风格设定信息中至少一项。

在一些实施例中，目标体裁包括：游戏、互动电视剧、互动电影中至少一项。

根据本公开的另一些实施例，提供一种互动多媒体内容的生成装置，包括：接收模块，被配置为接收用户输入的描述信息；内容生成模块，被配置为根据描述信息和目标体裁，生成内容文本，其中，内容文本包括剧情内容的描述文本以及目标体裁对应的视觉变换的描述文本；多媒体生成模块，被配置为根据内容文本，生成互动多媒体内容；显示模块，被配置为显示互动多媒体内容以及内容文本，其中，互动多媒体内容呈现与视觉变换的描述文本对应的视觉变换效果。

根据本公开的又一些实施例，提供一种电子设备，包括：存储器；以及耦接至存储器的处理器，处理器被配置为基于存储在存储器中的指令，执行前述任意实施例的互动多媒体内容的生成方法。

根据本公开的再一些实施例，提供一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现前述任意实施例的互动多媒体内容的生成方法。

根据本公开的又一些实施例，提供计算机程序，包括：指令，指令当由处理器执行时使处理器执行本公开中任意实施例的互动多媒体内容的生成方法。

根据本公开的再一些实施例，提供一种计算机程序产品，包括指令，指令当由处理器执行时实现本公开中任意实施例的互动多媒体内容的生成方法。

以上描述仅为本公开的一些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，本公开中所涉及的公开范围，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公开构思的情况下，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。

在本文提供的描述中，阐述了许多特定细节。然而，理解的是，可以在没有这些特定细节的情况下实施本发明的实施例。在其他情况下，为了不模糊该描述的理解，没有对众所周知的方法、结构和技术进行详细展示。

此外，虽然采用特定次序描绘了各操作，但是这不应当理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

虽然已经通过示例对本公开的一些特定实施例进行了详细说明，但是本领域的技术人员应该理解，以上示例仅是为了进行说明，而不是为了限制本公开的范围。本领域的技术人员应该理解，可在不脱离本公开的范围和精神的情况下，对以上实施例进行修改。本公开的范围由所附权利要求来限定。

权 利 要 求

1. 一种互动多媒体内容的生成方法，包括：

接收用户输入的描述信息；

根据所述描述信息和目标体裁，生成内容文本，其中，所述内容文本包括剧情内容的描述文本以及所述目标体裁对应的视觉变换的描述文本；

根据所述内容文本，生成互动多媒体内容；

显示所述互动多媒体内容以及所述内容文本，其中，所述互动多媒体内容呈现与所述视觉变换的描述文本对应的视觉变换效果。

2. 根据权利要求 1 所述的生成方法，还包括：

对所述描述信息进行解析，确定所述描述信息对应的体裁，作为所述目标体裁；
或者

接收用户对多个体裁中一个或多个体裁的选择操作，将所述一个或多个体裁确定为所述目标体裁。

3. 根据权利要求 1 或 2 所述的生成方法，其中，所述剧情内容的描述文本包括一个或多个剧情节点中每个剧情节点的剧情内容的描述文本，所述视觉变换的描述文本包括所述每个剧情节点对应的视觉变换的描述文本，所述根据所述描述信息和目标体裁，生成内容文本包括：

采用文本生成模型对所述描述信息进行解析，生成所述每个剧情节点的剧情内容的描述文本；

采用所述文本生成模型，根据所述描述信息和所述目标体裁，确定所述每个剧情节点中不同画面之间的视觉变换的信息，并生成所述每个剧情节点对应的视觉变换的描述文本。

4. 根据权利要求 3 所述的生成方法，其中，所述视觉变换包括镜头的运转，所述根据所述描述信息和所述目标体裁，确定所述每个剧情节点中不同画面之间的视觉变换的信息包括：

根据所述描述信息和所述目标体裁，确定是否在所述每个剧情节点中不同画面中运用不同类型的镜头效果；

针对确定运用不同类型的镜头效果的剧情节点，根据所述描述信息，确定所述剧情节点中每个画面中所运用的目标镜头效果以及所述目标镜头效果对应的目标对象，

作为所述剧情节点中不同画面之间的镜头的运转的信息。

5. 根据权利要求 3 或 4 所述的生成方法, 其中, 所述视觉变换包括人物姿态的变化, 所述根据所述描述信息和所述目标体裁, 确定所述每个剧情节点中不同画面之间的视觉变换的信息包括:

根据所述描述信息和所述目标体裁, 确定是否在所述每个剧情节点中变化人物的姿态;

针对确定变化人物的姿态的剧情节点, 根据所述描述信息, 确定所述剧情节点中每个画面中一个或多个个人物中每个人物的目标姿态, 作为所述剧情节点中不同画面之间所述每个人物的姿态的变化信息。

6. 根据权利要求 3-5 任一项所述的生成方法, 其中, 所述视觉变换包括背景的切换, 所述根据所述描述信息和所述目标体裁, 确定所述每个剧情节点中不同画面之间的视觉变换的信息包括:

根据所述描述信息和所述目标体裁, 确定是否在所述每个剧情节点中切换背景;

针对确定切换背景的剧情节点, 根据所述描述信息, 确定所述剧情节点中每个画面的目标背景, 作为所述每个剧情节点中不同画面之间背景的切换信息。

7. 根据权利要求 3-6 任一项所述的生成方法, 其中, 所述根据所述内容文本, 生成互动多媒体内容包括:

根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点对应的视觉变换的描述文本, 生成所述每个剧情节点的多媒体内容, 其中, 所述每个剧情节点的多媒体内容呈现与该剧情节点对应的视觉变换的描述文本相应的视觉效果。

8. 根据权利要求 7 所述的生成方法, 其中, 所述每个剧情节点对应的视觉变换的描述文本包括所述每个剧情节点中每个画面中所运用的目标镜头效果以及所述目标镜头效果对应的目标对象, 所述根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点对应的视觉变换的描述文本, 生成所述每个剧情节点的多媒体内容包括:

根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点中每个画面中所运用的目标镜头效果以及所述目标镜头效果对应的目标对象, 生成所述每个剧情节点的每个画面, 其中, 所述每个画面中的目标对象采用对应的目标镜头效果进行显示。

9. 根据权利要求 7 或 8 所述的生成方法, 其中, 所述每个剧情节点对应的视觉变换的描述文本包括所述每个剧情节点中每个画面中一个或多个个人物中每个人物的目

标姿态，所述根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点对应的视觉变换的描述文本，生成所述每个剧情节点的多媒体内容包括：

根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点中每个画面中一个或多个人物中每个人物的目标姿态，生成所述每个剧情节点的每个画面，其中，所述每个画面中的每个人物采用对应的目标姿态进行显示。

10. 根据权利要求 7-9 任一项所述的生成方法，其中，所述每个剧情节点对应的视觉变换的描述文本包括所述每个剧情节点中每个画面的目标背景，所述根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点对应的视觉变换的描述文本，生成所述每个剧情节点的多媒体内容包括：

根据所述每个剧情节点的剧情内容的描述文本和所述每个剧情节点中每个画面的目标背景，生成所述每个剧情节点的每个画面，其中，所述每个画面的背景为对应的目标背景。

11. 根据权利要求 7-10 任一项所述的生成方法，其中，显示所述互动多媒体内容以及所述内容文本包括：

将所述每个剧情节点的剧情描述文本、每个剧情节点中不同画面之间的视觉变换的描述文本和所述每个剧情节点的多媒体内容进行关联显示。

12. 根据权利要求 3-11 任一项所述的生成方法，还包括，针对每个剧情节点：

接收所述用户对所述剧情节点对应的视觉变换的描述文本的调整；

根据调整后的所述剧情节点对应的视觉变换的描述文本，生成所述剧情节点的多媒体内容；

显示调整后的所述剧情节点对应的视觉变换的描述文本和调整后的所述剧情节点的多媒体内容。

13. 根据权利要求 1-12 任一项所述的生成方法，其中，所述内容文本还包括一个或多个人物中每个人物的描述文本和一个或多个交互操作节点中每个交互操作节点的描述文本，所述根据所述描述信息和目标体裁，生成内容文本包括：

采用文本生成模型对所述描述信息进行解析，生成所述每个人物的描述文本；

采用文本生成模型对所述描述信息进行解析，生成所述每个交互操作节点的描述文本，其中，所述每个交互操作节点的描述文本包括对操作用户显示的交互操作选项的描述文本。

14. 根据权利要求 13 所述的生成方法，其中，所述每个人物的描述文本包括所述

每个人物的形象、声音特点中至少一项的描述信息，所述根据所述内容文本，生成互动多媒体内容包括：

根据所述每个人物的描述文本，生成所述每个人物的立绘、声音音频中至少一项；
根据所述每个交互操作节点的描述文本，生成所述每个交互操作节点的多媒体内容。

15. 根据权利要求 14 所述的生成方法，其中，所述显示所述互动多媒体内容以及所述内容文本包括：

将所述每个人物的描述文本和所述每个人物的立绘、声音音频中至少一项进行关联显示；

将所述每个交互操作节点的描述文本和所述每个交互操作节点的多媒体内容进行关联显示。

16. 根据权利要求 1-15 任一项所述的生成方法，其中，所述内容文本还包括一个或多个剧情节点中每个剧情节点对应的背景音乐的描述文本，所述根据所述描述信息和目标体裁，生成内容文本包括：

采用文本生成模型对所述描述信息进行解析，生成所述每个剧情节点对应的背景音乐的描述文本。

17. 根据权利要求 16 所述的生成方法，其中，所述根据所述内容文本，生成互动多媒体内容包括：

根据所述每个剧情节点对应的背景音乐的描述文本，生成每个剧情节点对的背景音乐。

18. 根据权利要求 1-17 任一项所述的生成方法，其中，所述描述信息包括剧情概述信息以及一个或多个剧情节点中每个剧情节点的剧情信息中至少一项。

19. 根据权利要求 18 所述的生成方法，其中，所述描述信息还包括：各个剧情节点之间的逻辑关系、人物设定信息、场景设定信息、画面风格设定信息中至少一项。

20. 根据权利要求 1-19 任一项所述的生成方法，其中，所述目标体裁包括：游戏、互动电视剧、互动电影中至少一项。

21. 一种互动多媒体内容的生成装置，包括：

接收模块，被配置为接收用户输入的描述信息；

内容生成模块，被配置为根据所述描述信息和目标体裁，生成内容文本，其中，所述内容文本包括剧情内容的描述文本以及所述目标体裁对应的视觉变换的描述文

本；

多媒体生成模块，被配置为根据所述内容文本，生成互动多媒体内容；

显示模块，被配置为显示所述互动多媒体内容以及所述内容文本，其中，所述互动多媒体内容呈现与所述视觉变换的描述文本对应的视觉变换效果。

22. 一种电子设备，包括：

存储器；以及

耦接至所述存储器的处理器，所述处理器被配置为基于存储在所述存储器中的指令，执行如权利要求 1 至 20 中任一项所述的互动多媒体内容的生成方法。

23. 一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现权利要求 1 至 20 中任一项所述的互动多媒体内容的生成方法。

24. 一种计算机程序，包括指令，当所述指令被处理器执行时，使所述处理器实现权利要求 1-20 任一项所述的互动多媒体内容的生成方法。

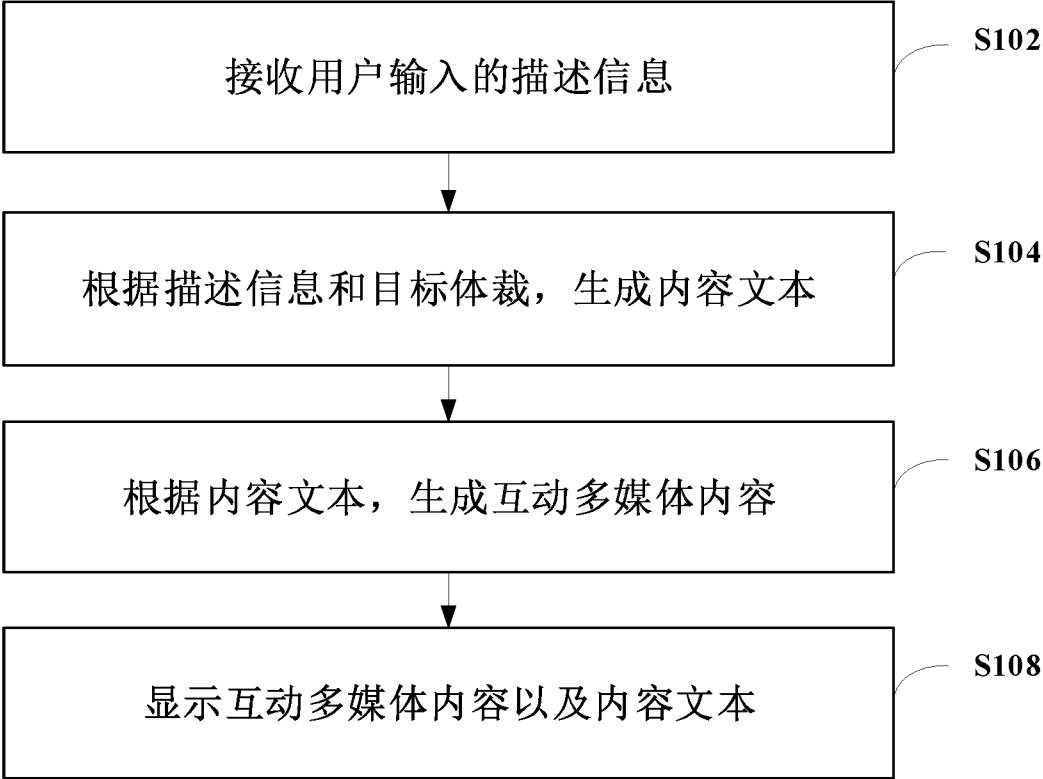


图 1

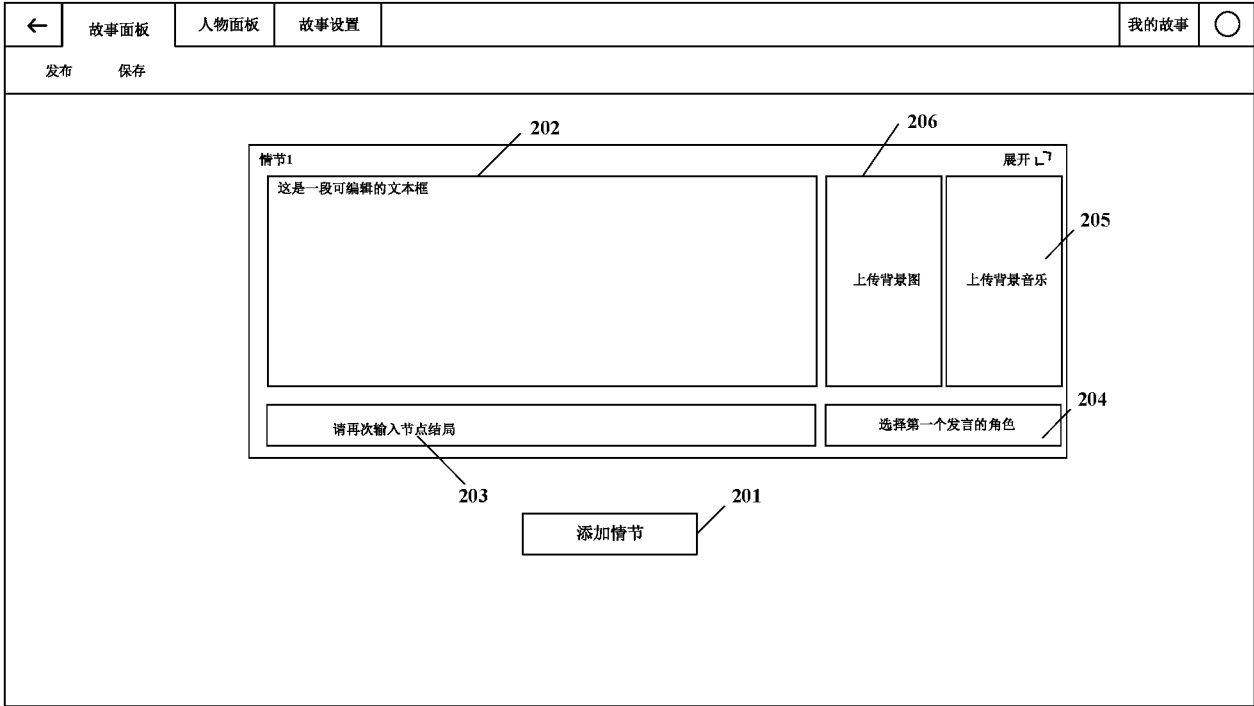


图 2

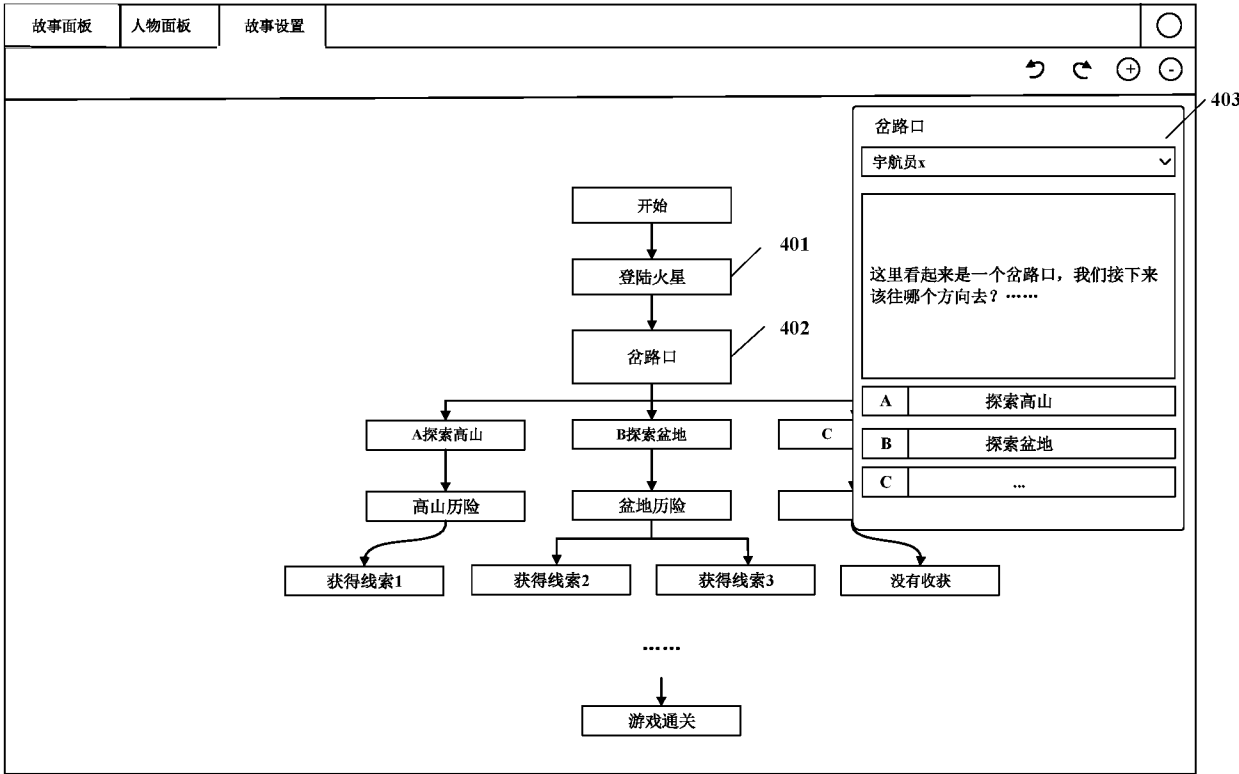
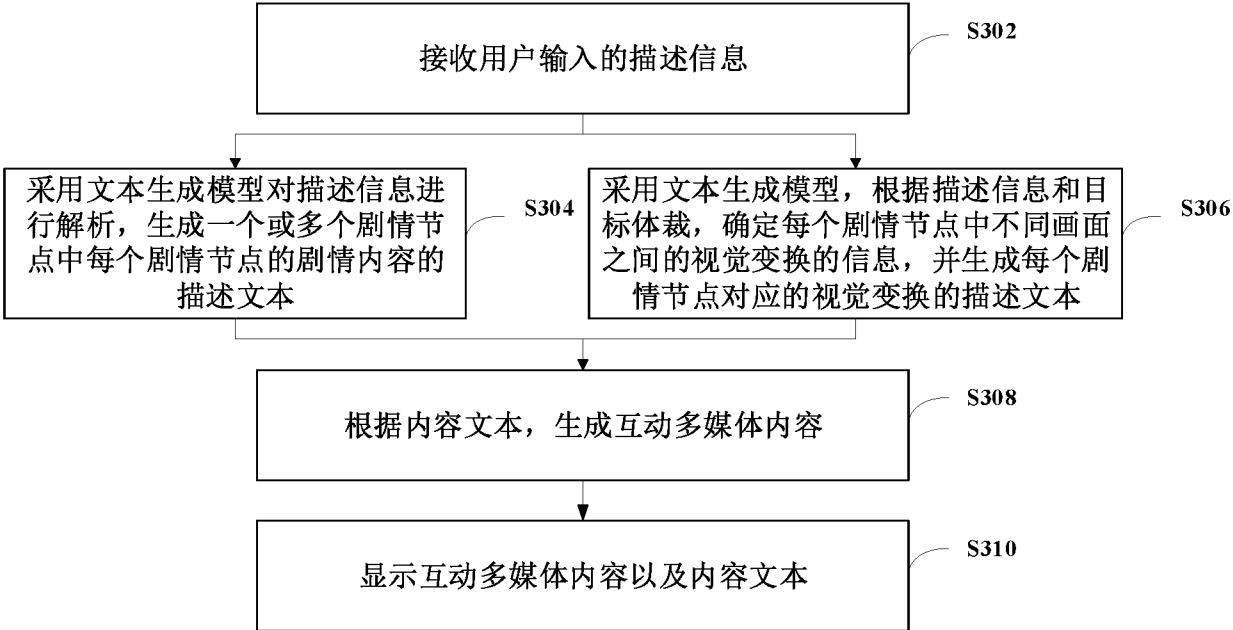


图 4

←

故事面板

人物面板

故事设置

我的故事

○

发布

保存

主要角色

502

输入姓名

505

请选择人物音乐

503

504

请输入基础设定

请输入台词风格

507

上传人物立绘

506

上传人物头像

501

添加角色

图 5

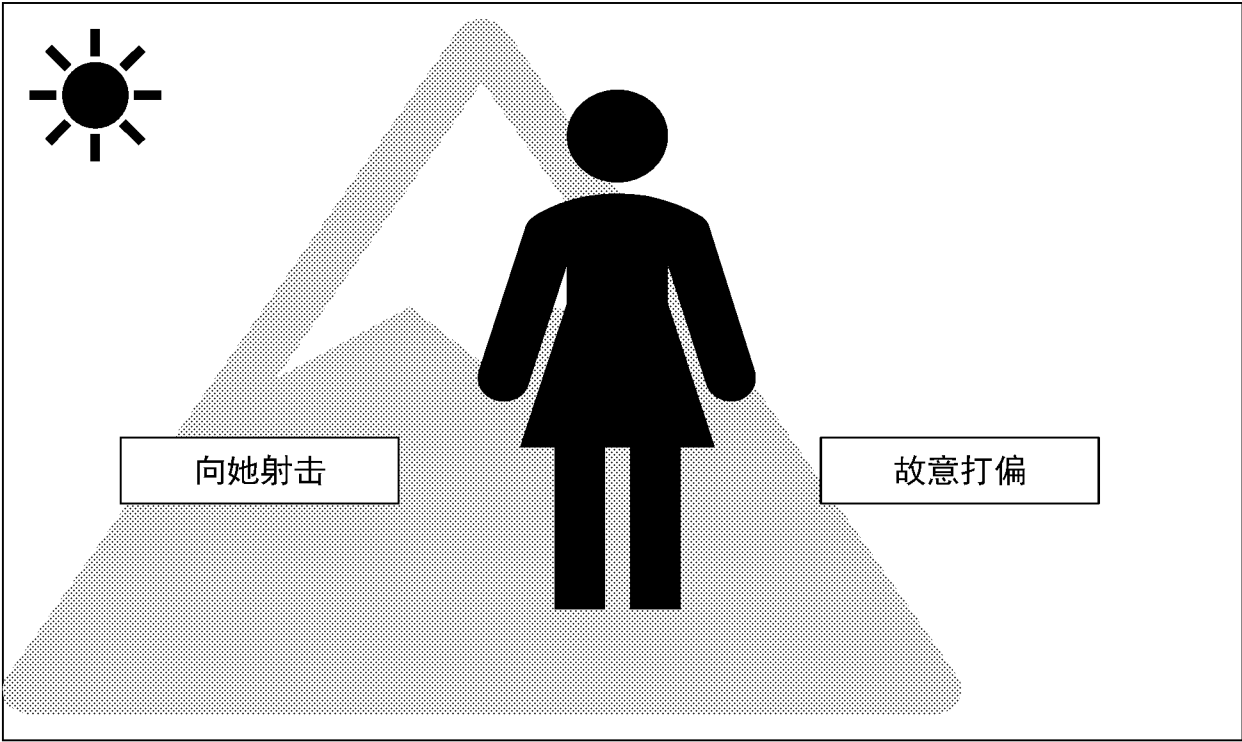


图 6

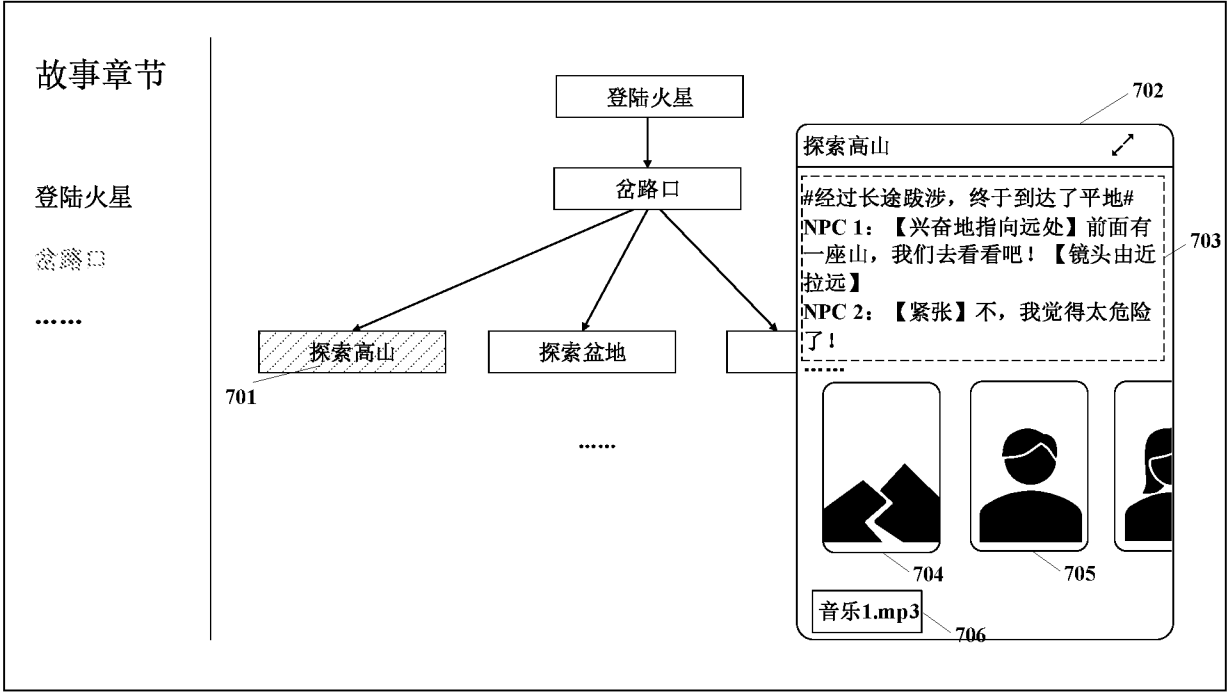


图 7

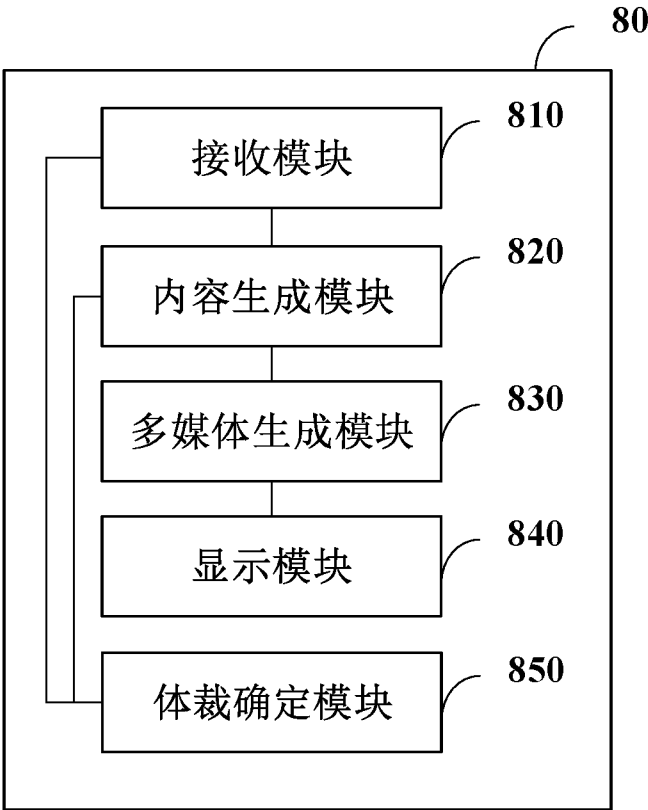


图 8

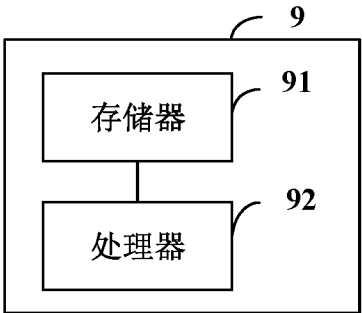


图 9

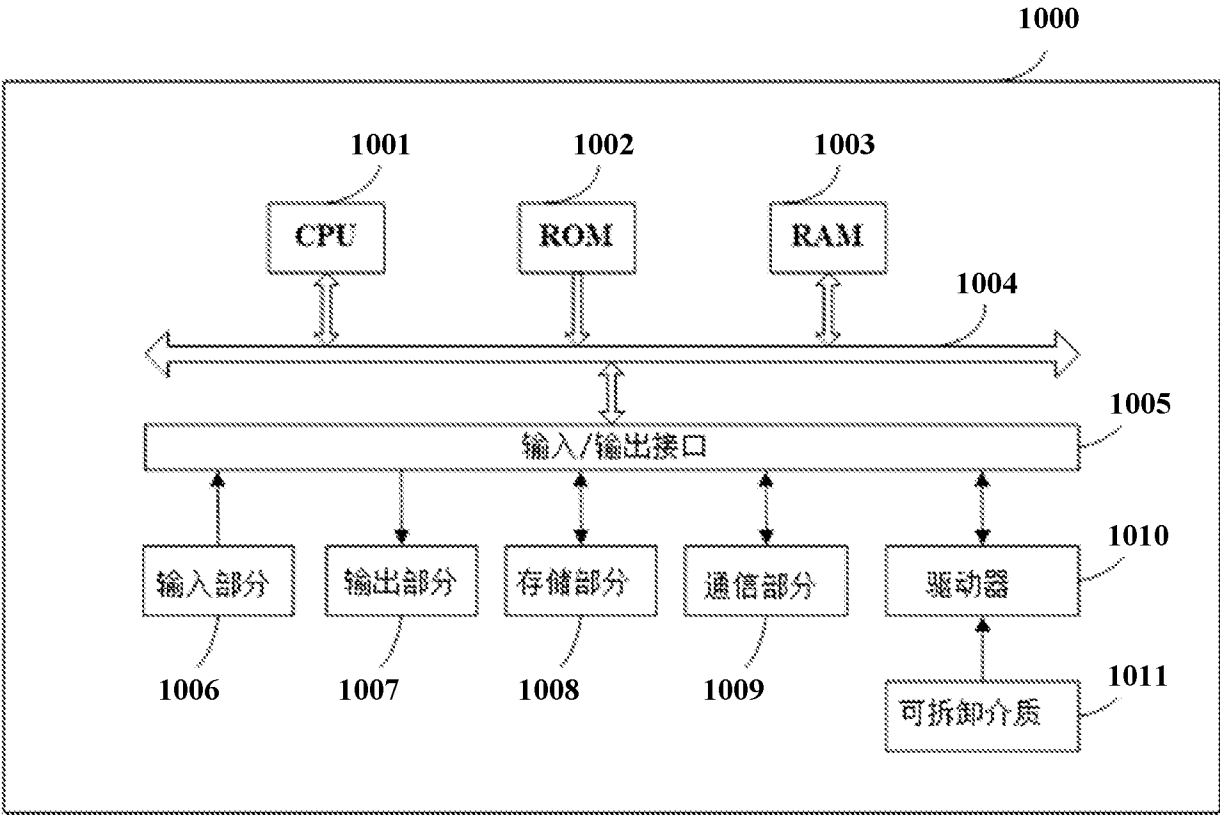


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/133084

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/432(2019.01)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: G06F

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)

VEN, CNABS, CNTXT, WOTXT, USTXT, EPTXT, CNKI, IEEE: 多媒体, 文本, 描述, 剧情, 节点, 变换, 变化, 目标, 模型, 背景, 镜头, interactive, multimedia, content, text, visual, transformation, description

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117633258 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 01 March 2024 (2024-03-01) description, paragraphs 23-210	1-24
X	CN 109493402 A (NETEASE (HANGZHOU) NETWORK CO., LTD.) 19 March 2019 (2019-03-19) description, paragraphs 6-93 and 125-128	1-24
A	CN 113778419 A (BEIJING YOUZHUJU NETWORK TECHNOLOGY CO., LTD.) 10 December 2021 (2021-12-10) entire document	1-24
A	WO 2019169068 A1 (DAVIS, GEORGE BERNARD) 06 September 2019 (2019-09-06) entire document	1-24



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

21 January 2025

Date of mailing of the international search report

05 February 2025

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/CN2024/133084

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	117633258	A	01 March 2024	None	
CN	109493402	A	19 March 2019	None	
CN	113778419	A	10 December 2021	None	
WO	2019169068	A1	06 September 2019	None	

A. 主题的分类

G06F 16/432(2019.01)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: G06F

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))

VEN, CNABS, CNTXT, WOTXT, USTXT, EPTXT, CNKI, IEEE: 多媒体, 文本, 描述, 剧情, 节点, 变换, 变化, 目标, 模型, 背景, 镜头, interactive, multimedia, content, text, visual, transformation, description

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 117633258 A (北京字跳网络技术有限公司) 2024年3月1日 (2024 - 03 - 01) 说明书第23-210段	1-24
X	CN 109493402 A (网易 (杭州) 网络有限公司) 2019年3月19日 (2019 - 03 - 19) 说明书6-93、125-128段	1-24
A	CN 113778419 A (北京有竹居网络技术有限公司) 2021年12月10日 (2021 - 12 - 10) 全文	1-24
A	WO 2019169068 A1 (DAVIS, GEORGE BERNARD) 2019年9月6日 (2019 - 09 - 06) 全文	1-24

☐ 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

2025年1月21日

国际检索报告邮寄日期

2025年2月5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

吴瑶

电话号码 (+86) 010-53961339

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/133084

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	117633258	A	2024年3月1日	无	
CN	109493402	A	2019年3月19日	无	
CN	113778419	A	2021年12月10日	无	
WO	2019169068	A1	2019年9月6日	无	