



(21) 申请号 202410092090.5

(22) 申请日 2024.01.23

(71) 申请人 北京沃东天骏信息技术有限公司
地址 100176 北京市大兴区经济技术开发区
科创十一街18号院2号楼4层A402室

(72) 发明人 王颢涵 沈俊杰 冯伟 卢杨
李耀宇 张政 吕晶晶 朱鑫

(74) 专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204
专利代理师 王达佐 王艳春

(51) Int. Cl.
G06T 11/60 (2006.01)
G06T 5/70 (2024.01)
G06Q 30/0241 (2023.01)

权利要求书3页 说明书19页 附图8页

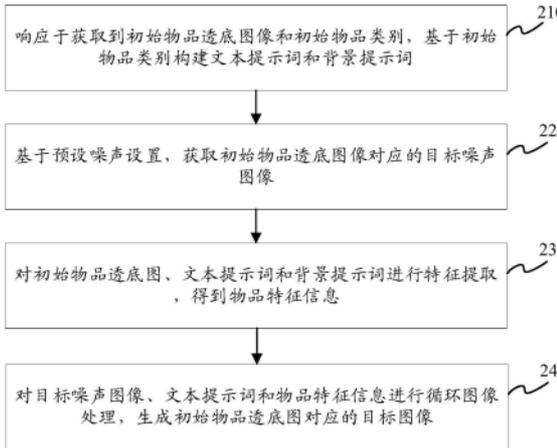
(54) 发明名称

图像生成方法和装置

(57) 摘要

本公开的实施例提供了一种图像生成方法和装置。所述图像生成方法包括：首先响应于获取到初始物品透底图像和初始物品类别，基于初始物品类别构建文本提示词和背景提示词，然后基于预设噪声设置，获取初始物品透底图像对应的目标噪声图像，之后对初始物品透底图、文本提示词和背景提示词进行特征提取，得到物品特征信息，最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理，生成初始物品透底图对应的目标图像，提出了一种端到端的物品广告背景生成方式，提高了图像生成的自动化和生成效率。

200



1. 一种图像生成方法,所述方法包括:
 - 响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;
 - 基于预设噪声设置,获取所述初始物品透底图像对应的目标噪声图像;
 - 对所述初始物品透底图、所述文本提示词和所述背景提示词进行特征提取,得到物品特征信息;
 - 对所述目标噪声图像、所述文本提示词和所述物品特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像。
2. 根据权利要求1所述的方法,其中,所述对所述目标噪声图像、所述文本提示词和所述物品特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像,包括:
 - 将所述目标噪声图像和所述文本提示词输入至文生图扩散模型中的编码器,得到背景特征信息;
 - 将所述物品特征信息和所述背景特征信息输入至所述文生图扩散模型中的解码器,得到去噪图像;
 - 判断所述去噪图像是否满足图像生成条件;
 - 响应于确定所述去噪图像满足所述图像生成条件,将所述去噪图像确定为所述初始物品透底图对应的目标图像。
3. 根据权利要求2所述的方法,其中,所述对所述目标噪声图像、所述文本提示词和所述物品特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像,还包括:
 - 响应于确定所述去噪图像不满足所述图像生成条件,将所述去噪图像作为目标噪声图像。
4. 根据权利要求1所述的方法,其中,所述方法还包括:获取参考图像和所述参考图像对应的物品掩膜;以及,
 - 所述基于预设噪声设置,获取所述初始物品透底图像对应的目标噪声图像,包括:
 - 基于预设噪声设置和所述参考图像,获取所述初始物品透底图像对应的噪声参考图像;以及,
 - 所述方法还包括:对所述参考图像和所述参考图像对应的物品掩膜进行特征提取,得到参考特征信息;以及,
 - 所述对所述目标噪声图像、所述文本提示词和所述物品特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像,包括:
 - 对所述噪声参考图像、所述文本提示词、所述物品特征信息和所述参考特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像。
5. 根据权利要求4所述的方法,其中,所述对所述噪声参考图像、所述文本提示词、所述物品特征信息和所述参考特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像,包括:
 - 将所述噪声参考图像和所述文本提示词输入至文生图扩散模型的编码器,得到背景特征信息;
 - 将所述物品特征信息、所述参考特征信息和所述背景特征信息输入至所述文生图扩散模型的解码器,得到去噪图像;

判断所述去噪图像是否满足图像生成条件；

响应于确定所述去噪图像满足所述图像生成条件，将所述去噪图像确定为所述初始物品透底图对应的目标图像。

6. 根据权利要求5所述的方法，其中，所述对所述噪声参考图像、所述文本提示词、所述物品特征信息和所述参考特征信息进行循环图像处理，生成所述初始物品透底图对应的目标图像，还包括：

响应于确定所述去噪图像不满足所述图像生成条件，将所述去噪图像作为噪声参考图像。

7. 根据权利要求4所述的方法，其中，所述方法还包括：基于所述初始物品透底图对应的物品掩膜，对所述参考特征信息进行信息过滤，得到过滤后的参考特征信息；以及，

所述对所述噪声参考图像、所述文本提示词、所述物品特征信息和所述参考特征信息进行循环图像处理，生成所述初始物品透底图对应的目标图像，包括：

对所述噪声参考图像、所述文本提示词、所述物品特征信息和所述过滤后的参考特征信息进行循环图像处理，生成所述初始物品透底图对应的目标图像。

8. 根据权利要求2-7任意一项所述的方法，其中，所述对所述初始物品透底图、所述文本提示词和所述背景提示词进行特征提取，得到物品特征信息，包括：

通过物品特征提取模型对所述初始物品透底图、所述文本提示词和所述背景提示词进行特征提取，得到物品特征信息，其中，所述物品特征提取模型包括物品掩膜的注意力模块。

9. 根据权利要求8所述的方法，其中，所述物品特征提取模型和所述文生图扩散模型基于以下步骤训练得到：

响应于获取到样本物品图像、样本物品掩膜和样本物品类别，基于所述样本物品图像和所述样本物品掩膜，获取样本物品透底图像；

基于所述样本物品类别，构建样本文本提示词和样本背景提示词；

基于所述样本物品图像，获取样本噪声图像；

将所述样本物品透底图像、所述样本文本提示词和所述样本背景提示词作为初始物品特征提取模型的输入，将所述样本噪声图像和所述样本文本提示词作为初始文生图扩散模型的输入，对所述初始物品特征提取模型和所述初始文生图扩散模型进行训练，得到所述物品特征提取模型和所述文生图扩散模型。

10. 根据权利要求9所述的方法，其中，所述方法还包括：获取参考图像对；以及，

所述基于所述样本物品图像，获取样本噪声图像，包括：

基于预设噪声设置和所述参考图像对，获取样本噪声图像；以及，

所述将所述样本物品透底图像、所述样本文本提示词和所述样本背景提示词作为初始物品特征提取模型的输入，将所述样本噪声图像和所述样本文本提示词作为初始文生图扩散模型的输入，对所述初始物品特征提取模型和所述初始文生图扩散模型进行训练，得到所述物品特征提取模型和所述文生图扩散模型，包括：

将所述样本物品透底图像、所述样本文本提示词和所述样本背景提示词作为初始物品特征提取模型的输入，将所述参考图像对作为初始参考特征提取模型的输入，将所述样本噪声图像和所述样本文本提示词作为初始文生图扩散模型的输入，对所述初始物品特征提

取模型、所述初始参考特征提取模型和所述初始文生图扩散模型进行训练,得到所述物品特征提取模型、参考特征提取模型和所述文生图扩散模型,其中,所述参考特征提取模型用于对所述参考图像和所述参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

11. 根据权利要求10所述的方法,其中,所述获取参考图像对,包括:

对所述样本物品掩膜进行膨胀,得到膨胀物品掩膜;

对所述样本物品图像进行数据增强,得到新的样本图像;

对所述新的样本图像和所述膨胀物品掩膜进行平移操作,得到平移样本图像和平移物品掩膜;

对所述平移物品掩膜进行随机掩膜处理,生成包括多个随机矩形掩膜的随机物品掩膜;

对所述平移样本图像和所述随机物品掩膜进行旋转,得到新的样本物品图像和新的样本物品掩膜,并将新的样本物品图像和新的样本物品掩膜组成所述参考图像对。

12. 一种图像生成装置,所述装置包括:

构建模块,被配置成响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;

获取模块,被配置成基于预设噪声设置,获取所述初始物品透底图像对应的目标噪声图像;

第一提取模块,被配置成对所述初始物品透底图、所述文本提示词和所述背景提示词进行特征提取,得到物品特征信息;

生成模块,被配置成对所述目标噪声图像、所述文本提示词和所述物品特征信息进行循环图像处理,生成所述初始物品透底图对应的目标图像。

13. 一种电子设备,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1-11中任一项所述的方法。

14. 一种计算机可读介质,其上存储计算机程序,其特征在于,该计算机程序被处理器执行时实现如权利要求1-11中任一项所述的方法。

图像生成方法和装置

技术领域

[0001] 本公开的实施例涉及计算机技术领域和信息处理技术领域,尤其涉及图像生成方法和装置。

背景技术

[0002] 物品广告背景生成旨在为物品透底图生成自然、逼真的背景,以构造高质量的广告图片,从而提升图片点击率。现有的背景生成方法主要分为两种方式,即“文生图”模式和“图生图”模式,其中,“文生图”模式指的是向扩散大模型(如Stable Diffusion, ControlNet)输入一段描述图片的提示词和一张商品透底图,由大模型根据提示词的内容填充商品周围的背景区域;“图生图”模式指的是在“文生图”模式的基础上,额外引入一张参考图像,并将该参考图像添加一定强度的噪声,作为扩散大模型的初始噪声,使得生成的背景区域与参考图像具备一定的相似性。

[0003] 然而“文生图”模式和“图生图”模式需要花费大量时间设计和修正提示词,以及若提示词在描述图片的空间位置布局或抽象风格时效果较差,给精细化定制背景带来了较大挑战等等。

发明内容

[0004] 本公开的实施例提出了一种图像生成方法、图像生成装置、电子设备和计算机可读介质。

[0005] 第一方面,本公开的实施例提供了一种图像生成方法,该方法包括:响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像;对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息;对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0006] 在一些实施例中,对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,包括:将目标噪声图像和文本提示词输入至文生图扩散模型中的编码器,得到背景特征信息;将物品特征信息和背景特征信息输入至文生图扩散模型中的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0007] 在一些实施例中,对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,还包括:响应于确定去噪图像不满足图像生成条件,将去噪图像作为目标噪声图像。

[0008] 在一些实施例中,该方法还包括:获取参考图像和参考图像对应的物品掩膜;以及,基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,包括:基于预设噪声设置和参考图像,获取初始物品透底图像对应的噪声参考图像;以及,该方法还包括:对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息;以及,对目标噪声图

像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,包括:对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0009] 在一些实施例中,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,包括:将噪声参考图像和文本提示词输入至文生图扩散模型的编码器,得到背景特征信息;将物品特征信息、参考特征信息和背景特征信息输入至文生图扩散模型的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0010] 在一些实施例中,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,还包括:响应于确定去噪图像不满足图像生成条件,将去噪图像作为噪声参考图像。

[0011] 在一些实施例中,该方法还包括:基于初始物品透底图像对应的物品掩膜,对参考特征信息进行信息过滤,得到过滤后的参考特征信息;以及,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,包括:对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0012] 在一些实施例中,对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,包括:通过物品特征提取模型对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,其中,物品特征提取模型包括物品掩膜的注意力模块。

[0013] 在一些实施例中,物品特征提取模型和文生图扩散模型基于以下步骤训练得到:响应于获取到样本物品图像、样本物品掩膜和样本物品类别,基于样本物品图像和样本物品掩膜,获取样本物品透底图像;基于样本物品类别,构建样本文本提示词和样本背景提示词;基于样本物品图像,获取样本噪声图像;将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型。

[0014] 在一些实施例中,该方法还包括:获取参考图像对;以及,基于样本物品图像,获取样本噪声图像,包括:基于预设噪声设置和参考图像对,获取样本噪声图像;以及,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型,包括:利用机器学习方法,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将参考图像对作为初始参考特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型、参考特征提取模型和文生图扩散模型,其中,参考特征提取模型用于对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

[0015] 在一些实施例中,获取参考图像对,包括:对样本物品掩膜进行膨胀,得到膨胀物品掩膜;对样本物品图像进行数据增强,得到新的样本图像;对新的样本图像和膨胀物品掩膜进行平移操作,得到平移样本图像和平移物品掩膜;对平移物品掩膜进行随机掩膜处理,生成包括多个随机矩形掩膜的随机物品掩膜;对平移样本图像和随机物品掩膜进行旋转,得到新的样本物品图像和新的样本物品掩膜,并将新的样本物品图像和新的样本物品掩膜组成参考图像对。

[0016] 第二方面,本公开的实施例提供了一种图像生成装置,该装置包括:构建模块,被配置成响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;获取模块,被配置成基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像;第一提取模块,被配置成对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息;生成模块,被配置成对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0017] 在一些实施例中,生成模块,进一步被配置成:将目标噪声图像和文本提示词输入至文生图扩散模型中的编码器,得到背景特征信息;将物品特征信息和背景特征信息输入至文生图扩散模型中的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0018] 在一些实施例中,生成模块,进一步被配置成:响应于确定去噪图像不满足图像生成条件,将去噪图像作为目标噪声图像。

[0019] 在一些实施例中,该装置还包括第二提取模块;获取模块,进一步被配置成:获取参考图像和参考图像对应的物品掩膜;基于预设噪声设置和参考图像,获取初始物品透底图像对应的噪声参考图像;以及,第二提取模块,被配置成:对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息;以及,生成模块,进一步被配置成:对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0020] 在一些实施例中,生成模块,进一步被配置成:将噪声参考图像和文本提示词输入至文生图扩散模型的编码器,得到背景特征信息;将物品特征信息、参考特征信息和背景特征信息输入至文生图扩散模型的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0021] 在一些实施例中,生成模块,进一步被配置成:响应于确定去噪图像不满足图像生成条件,将去噪图像作为噪声参考图像。

[0022] 在一些实施例中,该装置还包括过滤模块;过滤模块,被配置成:基于初始物品透底图像对应的物品掩膜,对参考特征信息进行信息过滤,得到过滤后的参考特征信息;以及,生成模块,进一步被配置成:对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0023] 在一些实施例中,第一提取模块,进一步被配置成:通过物品特征提取模型对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,其中,物品特征提取模型包括物品掩膜的注意力模块。

[0024] 在一些实施例中,物品特征提取模型和文生图扩散模型基于训练单元训练得到;

训练单元,被配置成:响应于获取到样本物品图像、样本物品掩膜和样本物品类别,基于样本物品图像和样本物品掩膜,获取样本物品透底图像;基于样本物品类别,构建样本文本提示词和样本背景提示词;基于样本物品图像,获取样本噪声图像;将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型。

[0025] 在一些实施例中,训练单元,进一步被配置成:获取参考图像对;以及,基于预设噪声设置和参考图像对,获取样本噪声图像;以及,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将参考图像对作为初始参考特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型、参考特征提取模型和文生图扩散模型,其中,参考特征提取模型用于对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

[0026] 在一些实施例中,训练单元,进一步被配置成:对样本物品掩膜进行膨胀,得到膨胀物品掩膜;对样本物品图像进行数据增强,得到新的样本图像;对新的样本图像和膨胀物品掩膜进行平移操作,得到平移样本图像和平移物品掩膜;对平移物品掩膜进行随机掩膜处理,生成包括多个随机矩形掩膜的随机物品掩膜;对平移样本图像和随机物品掩膜进行旋转,得到新的样本物品图像和新的样本物品掩膜,并将新的样本物品图像和新的样本物品掩膜组成参考图像对。

[0027] 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施例描述的图像生成方法。

[0028] 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,该计算机程序被处理器执行时实现如第一方面中任一实施例描述的图像生成方法。

[0029] 本公开的实施例提供的图像生成方法,上述执行主体首先响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词,然后基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,之后对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,提出了一种端到端的物品广告背景生成方式,可以根据物品透底图像和其所属的物品类别自动生成适合该类别物品的背景,从而自动生成包括物品和对应背景的目标图像,不需要精细设计提示词,自动根据物品类别构建相应的文本提示词和背景提示词,使得提示词更准确,当不提供参考图像时,自动为物品生成相应的背景,不需要依赖参考图像即可获取到相应的目标图像,提高了图像生成的自动化和生成效率。

附图说明

[0030] 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:

[0031] 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;

- [0032] 图2是根据本公开的图像生成方法的一个实施例的流程图；
- [0033] 图3是根据本公开的生成初始物品透底图对应的目标图像的一个实施例的流程图；
- [0034] 图4是根据本公开的图像生成方法的另一个实施例的流程图；
- [0035] 图5是根据本公开的生成初始物品透底图对应的目标图像的另一个实施例的流程图；
- [0036] 图6是根据本公开的训练物品特征提取模型和文生图扩散模型的一个实施例的流程图；
- [0037] 图7是根据本公开的获取参考图像对的一个实施例的流程图；
- [0038] 图8是根据本公开的图像生成装置的一个实施例的结构示意图；
- [0039] 图9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

[0040] 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公开,而非对该公开的限定。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公开相关的部分。

[0041] 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。

[0042] 图1示出了可以应用本申请的图像生成方法或图像生成装置的实施例的示例性系统架构100。

[0043] 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,例如有线、无线通信链路或者光纤电缆等等。

[0044] 用户可以使用终端设备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可以是用户终端设备,其上可以安装有各种客户端应用,例如,电商平台应用、图像类应用、视频类应用、搜索类应用、金融类应用等。

[0045] 终端设备101、102、103可以是具有显示屏并且支持接收服务器消息的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、电子播放器、膝上型便携计算机和台式计算机等等。

[0046] 其中,终端设备101、102、103可以获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词,然后基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,之后对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0047] 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,可以是各种电子设备,当终端设备101、102、103为软件时,可以安装在上述所列举的电子设备中。其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多个软件模块),也可以实现成单个软件或软件模块。在此不做具体限定。

[0048] 服务器105可以是提供各种服务的服务器,例如对与其建立通信连接的终端设备

发送的请求进行接收的后台服务器。后台服务器可以对终端设备发送的请求进行接收和分析等处理,并生成处理结果。

[0049] 服务器105也可以获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词,然后基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,之后对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0050] 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,可以是为终端设备提供各种服务的各种电子设备。当服务器为软件时,可以实现成为为终端设备提供各种服务的多个软件或软件模块,也可以实现成为为终端设备提供各种服务的单个软件或软件模块。在此不做具体限定。

[0051] 需要说明的是,本公开的实施例所提供的图像生成方法可以由终端设备101、102、103或服务器105执行,相应地,图像生成装置可以设置于终端设备101、102、103或服务器105中。

[0052] 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、网络和服务器。

[0053] 参考图2,示出了根据本公开的图像生成方法的一个实施例的流程图200。该图像生成方法包括以下步骤:

[0054] 步骤210,响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词。

[0055] 在本步骤中,上述图像生成方法运行于其上的执行主体(例如图1中的终端设备101、102、103或服务器105)可以通过网络读取获取初始物品图像,对初始物品图像进行物品类别分析,获取到初始物品类别,该初始物品类别可以表征初始物品图像中包括的物品类别,例如牛奶、口红等等。上述执行主体可以利用初始物品图像对应的物品掩膜对初始物品图像进行图像处理,获取到初始物品图像对应的初始物品透底图像,该初始物品透底图像可以表征只包括初始物品、不包括背景的图像。

[0056] 上述执行主体获取到初始物品透底图像和初始物品类别后,可以读取文本提示词的文本构建模版,利用文本构建模板对初始物品类别进行文本构建,生成初始物品透底图像对应的文本提示词,该文本提示词可以表征初始物品对应的物品类别,该文本构建模版可以为“A photo of 初始物品类别”。作为示例,上述初始物品类别为C,则文本提示词可以为“A photo of C”。

[0057] 上述执行主体可以根据初始物品类别与物品类别编码之间的对应关系,确定出初始物品类别对应的初始物品类别编码。上述执行主体还可以将初始物品类别编码与预设字符进行拼接,得到初始物品类别对应的背景信息,该预设字符可以是预先设定的字符,例如“sks”。上述执行主体可以读取背景提示词的背景构建模版,利用背景构建模板对背景信息进行构建,生成初始物品透底图像对应的背景提示词,该背景提示词可以表征初始物品对应的背景信息,该背景构建模板可以为“in the background of 背景信息”。作为示例,上述初始物品类别为C,背景信息可以为sks与C拼接成的字符串D,则背景提示词可以为“in the background of D”。

[0058] 步骤220,基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像。

[0059] 在本步骤中,上述执行主体可以根据预设噪声设置对图像进行噪声叠加处理,生成初始物品透底图像对应的目标噪声图像,该预设噪声设置可以包括对图像进行预设噪声叠加处理,该图像可以是与初始物品透底图像对应的参考图像、与初始物品透底图像对应的随机图像等,该目标噪声图像可以表征与初始物品透底图像对应的、已经进行噪声处理的图像。作为示例,预设噪声设置可以包括对图像进行50次噪声叠加处理,则上述执行主体可以对图像进行50次噪声叠加处理,生成初始物品透底图像对应的目标噪声图像。

[0060] 步骤230,对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息。

[0061] 在本步骤中,上述执行主体可以利用特征提取网络对初始物品透底图、文本提示词和背景提示词进行特征提取,提取初始物品透底图中的初始物品信息,用于生成适配初始物品类别的通用背景信息,得到物品特征信息。或者,上述执行主体可以对初始物品透底图、文本提示词和背景提示词进行特征分析,分析其中与初始物品相关的特征信息,并将该特征信息提取出来,得到物品特征信息。

[0062] 作为一个可选实现方式,上述步骤230,对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,可以包括以下步骤:通过物品特征提取模型对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,其中,物品特征提取模型包括物品掩膜的注意力模块。

[0063] 具体地,上述执行主体可以获取物品特征提取模型,物品特征提取模型可以用于表征初始物品透底图和文本提示词两者与物品特征信息之间的对应关系,该物品特征提取模型的具体结构与ControlNet(控制网络)基本相同。上述执行主体可以将初始物品透底图、文本提示词和背景提示词输入至物品特征提取模型,通过物品特征提取模型对初始物品透底图、文本提示词和背景提示词进行特征提取,输出物品特征信息,该物品特征信息可以用于表征提取初始物品透底图中的初始物品信息,用于生成适配初始物品类别的通用背景信息。

[0064] 上述物品特征提取模型可以包括物品掩膜的注意力模块,即物品特征提取模型的具体结构与ControlNet(控制网络)基本相同,在本公开中将物品特征提取模型中的注意力模块替换为基于物品掩膜的注意力模块,若上述初始物品透底图像中对应的物品掩膜M,基于物品掩膜的注意力模块可以表示为:

$$[0065] \quad X_{out} = M \cdot CA(X_{in}, P_{fg}) + (1 - M) \cdot CA(X_{in}, P_{bg}) + X_{in}$$

[0066] 其中, X_{in} 与 X_{out} 分别表示物品掩膜的注意力模块的输入、输出,CA()表示常规的注意力模块, \cdot 表示点乘符号, P_{fg} 表示文本提示词, P_{bg} 表示背景提示词。

[0067] 在本实现方式中,物品特征提取模型包括物品掩膜的注意力模块,在计算注意力时,令物品区域仅和文本提示词计算注意力,令背景区域仅和背景提示词计算注意力,确保背景风格相关知识被准确地编码进入背景提示词中,避免受到前景信息的干扰,使得物品特征提取模型能够更准确、更具体地对物品特征信息进行分析和提取,获取到更准确的物品特征信息。

[0068] 步骤240,对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0069] 在本步骤中,上述执行主体可以对目标噪声图像、文本提示词和物品特征信息进行图像背景生成和去噪处理,对目标噪声图像和文本提示词进行背景信息提取,获取到对应的背景信息,然后对背景信息和物品特征信息进行特征融合和去噪处理,生成一个具备初始物品和背景信息的图像,该图像可能是一个具备噪声的图像,则需要将该图像作为新的目标噪声图像,与文本提示词和物品特征信息再次进行图像背景生成和去噪处理,直至得到的图像为清晰的具备初始物品和背景信息的图像,对目标噪声图像、文本提示词和物品特征信息进行了循环图像处理,对图像进行循环去噪处理,可以执行预设数量次去噪处理,该预设数量可以与预设噪声设置相对应,直至生成初始物品透底图对应的目标图像,该目标图像可以用于表征具备初始物品和背景的图像。

[0070] 本公开的实施例提供的图像生成方法,上述执行主体首先响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词,然后基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,之后对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,提出了一种端到端的物品广告背景生成方式,可以根据物品透底图像和其所属的物品类别自动生成适合该类别物品的背景,从而自动生成包括物品和对应背景的目标图像,不需要精细设计提示词,自动根据物品类别构建相应的文本提示词和背景提示词,使得提示词更准确,当不提供参考图像时,自动为物品生成相应的背景,不需要依赖参考图像即可获取到相应的目标图像,提高了图像生成的自动化和生成效率。

[0071] 参考图3,示出了生成初始物品透底图对应的目标图像的一个实施例的流程图300,即上述步骤240,对目标噪声图像、文本提示词、背景提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,可以包括以下步骤:

[0072] 步骤310,将目标噪声图像和文本提示词输入至文生图扩散模型中的编码器,得到背景特征信息。

[0073] 在本步骤中,上述执行主体获取到文生图扩散模型,该文生图扩散模型可以用于表征目标噪声图像、文本提示词和物品特征信息三者与目标图像之间的对应关系,该文生图扩散模型可以为预训练的扩散大模型Stable Diffusion(SD),该文生图扩散模型可以包括编码器和解码器。上述执行主体可以将目标噪声图像、文本提示词和物品特征信息输入至文生图扩散模型,通过文生图扩散模型对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,对图像进行循环去噪处理,可以执行预设数量次去噪处理,该预设数量可以与预设噪声设置相对应,直至生成初始物品透底图对应的目标图像。

[0074] 上述执行主体可以将目标噪声图像和文本提示词输入至文生图扩散模型中的编码器,该编码器对目标噪声图像和文本提示词进行编码处理,提取出背景特征信息,该背景特征信息可以用于生成初始物品对应的背景。

[0075] 步骤320,将物品特征信息和背景特征信息输入至文生图扩散模型中的解码器,得到去噪图像。

[0076] 在本步骤中,上述执行主体可以将物品特征提取模型输出的物品特征信息与编码器输出的背景特征信息共同输入至文生图扩散模型中的解码器,该解码器对物品特征信息和背景特征信息进行解码处理,输出对应的去噪图像,该去噪图像可以是包括初始物品和

背景的噪声图像,也可以是包括初始物品和背景的目标图像。

[0077] 步骤330,判断去噪图像是否满足图像生成条件。

[0078] 在本步骤中,上述执行主体可以对输出的去噪图像进行判断,确定去噪图像是否满足图像生成条件,该图像生成条件可以包括图像清晰,或者图像经过预设数量次去噪处理等。

[0079] 响应于确定去噪图像满足图像生成条件,执行步骤340,将去噪图像确定为初始物品透底图对应的目标图像。

[0080] 在本步骤中,上述执行主体经判断确定去噪图像满足图像生成条件,则可以确定文生图扩散模型已经执行了预设数量次去噪处理,生成了清晰的包括初始物品和背景的图像,可以将该去噪图像确定为初始物品透底图对应的目标图像。

[0081] 在本实现方式中,可以同时考虑物品特征信息和背景特征信息,使得文生图扩散模型自动生成包括物品和对应背景的目标图像,自动生成适合该类别物品的背景。

[0082] 继续参见图3,响应于确定去噪图像不满足图像生成条件,执行步骤350,将去噪图像作为目标噪声图像。

[0083] 在本步骤中,上述执行主体经判断确定去噪图像不满足图像生成条件,则可以确定该去噪图像还存在噪声,需要将去噪图像作为目标噪声图像,再次与文本提示词输入至文生图扩散模型中的编码器进行编码,再输入至解码器进行解码,对噪声图像进行循环去噪处理,直至生成满足图像生成条件的目标图像。

[0084] 在本实现方式中,可以利用文生图扩散模型进行循环图像处理,直至生成初始物品透底图对应的目标图像,提高了目标图像的生成效率和准确性。

[0085] 参见图4,图4示出了图像生成方法的另一个实施例的流程图400,即上述图像生成方法可以包括以下步骤:

[0086] 步骤410,响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词。

[0087] 在本步骤中,步骤410与图2所示实施例中的步骤210相同,此处不赘述。

[0088] 步骤420,获取参考图像和参考图像对应的物品掩膜。

[0089] 在本步骤中,上述执行主体可以通过网络读取获取到参考图像,该参考图像可以用于作为初始物品透底图像的背景参考基础。上述执行主体可以对参考图像进行物品分析,确定出参考图像对应的物品掩膜。

[0090] 步骤430,基于预设噪声设置和参考图像,获取初始物品透底图像对应的噪声参考图像。

[0091] 在本步骤中,上述执行主体可以根据预设噪声设置对参考图像进行噪声叠加处理,生成初始物品透底图像对应的噪声参考图像,该预设噪声设置可以包括对参考图像进行预设次噪声叠加处理,该噪声参考图像可以表征与初始物品透底图像对应的、已经进行噪声处理的参考图像。作为示例,预设噪声设置可以包括对图像进行50次噪声叠加处理,则上述执行主体可以对参考图像进行50次噪声叠加处理,生成初始物品透底图像对应的噪声参考图像。

[0092] 步骤440,对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息。

[0093] 在本步骤中,步骤440与图2所示实施例中的步骤230相同,此处不赘述。

[0094] 步骤450,对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

[0095] 在本步骤中,上述执行主体可以获取预先训练的参考特征提取模型,其中,该参考特征提取模型可以用于表征参考图像和参考图像对应的物品掩膜两者与参考特征信息之间的对应关系,该参考特征提取模型的具体结构与ControlNet(控制网络)基本相同。上述执行主体可以将参考图像和参考图像对应的物品掩膜输入至参考特征提取模型,通过参考特征提取模型对参考图像和参考图像对应的物品掩膜进行特征提取,输出参考特征信息,该参考特征信息可以用于表征参考图片中的背景区域信息。

[0096] 步骤460,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0097] 在本步骤中,上述执行主体可以对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行图像背景生成和去噪处理,对噪声参考图像和文本提示词进行背景信息提取,获取到对应的参考背景信息,然后对参考背景信息、参考特征信息和物品特征信息进行特征融合和去噪处理,生成一个具备初始物品和与参考图像的背景信息相似的图像,该图像可能是一个具备噪声的图像,则需要将该图像作为新的噪声参考图像,与文本提示词、物品特征信息和参考特征信息再次进行图像背景生成和去噪处理,直至得到的图像为清晰的具备初始物品和与参考图像的背景信息相似的图像,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行了循环图像处理,对图像进行循环去噪处理,可以执行预设数量次去噪处理,该预设数量可以与预设噪声设置相对应,直至生成初始物品透底图对应的目标图像,该目标图像可以用于表征具备初始物品和背景的图像。

[0098] 在本实施例中,通过根据初始物品透底图和参考图像,为初始物品生成与参考图像布局、组成元素、色彩、风格等相似的背景,可以提取参考图像中的布局、纹理等信息来为初始物品生成相似的背景。

[0099] 参考图5,示出了生成初始物品透底图对应的目标图像的另一个实施例的流程图500,即上述步骤460,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,可以包括以下步骤:

[0100] 步骤510,将噪声参考图像和文本提示词输入至文生图扩散模型的编码器,得到背景特征信息。

[0101] 在本步骤中,上述执行主体获取到文生图扩散模型,该文生图扩散模型可以用于表征噪声参考图像、文本提示词、参考特征信息和物品特征信息几者与目标图像之间的对应关系,该文生图扩散模型可以为预训练的扩散大模型Stable Diffusion(SD),该文生图扩散模型可以包括编码器和解码器。上述执行主体可以将噪声参考图像、文本提示词、参考特征信息和物品特征信息输入至文生图扩散模型,通过文生图扩散模型对噪声参考图像、文本提示词、参考特征信息和物品特征信息进行循环图像处理,对图像进行循环去噪处理,可以执行预设数量次去噪处理,该预设数量可以与预设噪声设置相对应,直至生成初始物品透底图对应的目标图像。

[0102] 上述执行主体可以将噪声参考图像和文本提示词输入至文生图扩散模型中的编码器,该编码器对噪声参考图像和文本提示词进行编码处理,提取出背景特征信息,该背景

特征信息可以表征参考图像的背景区域信息,用于生成初始物品对应的背景。

[0103] 步骤520,将物品特征信息、参考特征信息和背景特征信息输入至文生图扩散模型的解码器,得到去噪图像。

[0104] 在本步骤中,上述执行主体可以将物品特征提取模型输出的物品特征信息、参考特征提取模型输出的参考特征信息与编码器输出的背景特征信息共同输入至文生图扩散模型中的解码器,该解码器对物品特征信息、参考特征信息和背景特征信息进行解码处理,输出对应的去噪图像,该去噪图像可以是包括初始物品和背景的噪声图像,也可以是包括初始物品和背景的目标图像。

[0105] 步骤530,判断去噪图像是否满足图像生成条件。

[0106] 在本步骤中,上述执行主体可以对输出的去噪图像进行判断,确定去噪图像是否满足图像生成条件,该图像生成条件可以包括图像清晰,或者图像经过预设数量次去噪处理等。

[0107] 响应于确定去噪图像满足图像生成条件,执行步骤540,将去噪图像确定为初始物品透底图对应的目标图像。

[0108] 在本步骤中,上述执行主体经判断确定去噪图像满足图像生成条件,则可以确定文生图扩散模型已经执行了预设数量次去噪处理,生成了清晰的包括初始物品和背景的图像,可以将该去噪图像确定为初始物品透底图对应的目标图像。

[0109] 在本实现方式中,可以同时考虑物品特征信息、参考特征信息和背景特征信息,使得文生图扩散模型自动生成包括物品和参考背景的目标图像,自动生成与参考图像相似的背景。

[0110] 继续参见图5,响应于确定去噪图像不满足图像生成条件,执行步骤550,将去噪图像作为噪声参考图像。

[0111] 在本步骤中,上述执行主体经判断确定去噪图像不满足图像生成条件,则可以确定该去噪图像还存在噪声,需要将去噪图像作为噪声参考图像,再次与文本提示词输入至文生图扩散模型中的编码器进行编码,再输入至解码器进行解码,对噪声图像进行循环去噪处理,直至生成满足图像生成条件的目标图像。

[0112] 在本实现方式中,可以利用文生图扩散模型进行循环图像处理,直至生成初始物品透底图对应的目标图像,提高了目标图像的生成效率和准确性。

[0113] 作为一个可选实现方式,上述图像生成方法还包括:基于初始物品透底图像对应的物品掩膜,对参考特征信息进行信息过滤,得到过滤后的参考特征信息;以及,上述步骤460,对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,可以包括以下步骤:对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0114] 具体地,上述执行主体获取到参考特征信息后,可以利用初始物品透底图像对应的物品掩膜对参考特征信息进行信息过滤,得到过滤后的参考特征信息,即与ControlNet(控制网络)相同,参考特征提取模型可以输出多个不同分辨率的特征图 Y_i ($i=1\sim N$),则对于 Y_i ,可以使 $Y_i = (1-M) \cdot Y_i$,其中, M 表示初始物品透底图像对应的物品掩膜。

[0115] 上述执行主体可以对噪声参考图像、文本提示词、物品特征信息和过滤后的参考

特征信息进行图像背景生成和去噪处理,对噪声参考图像和文本提示词进行背景信息提取,获取到对应的参考背景信息,然后对参考背景信息、过滤后的参考特征信息和物品特征信息进行特征融合和去噪处理,生成一个具备初始物品和与参考图像的背景信息相似的图像,该图像可能是一个具备噪声的图像,则需要将该图像作为新的噪声参考图像,与文本提示词、物品特征信息和过滤后的参考特征信息再次进行图像背景生成和去噪处理,直至得到的图像为清晰的具备初始物品和与参考图像的背景信息相似的图像,对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行了循环图像处理,对图像进行循环去噪处理,可以执行预设数量次去噪处理,该预设数量可以与预设噪声设置相对应,直至生成初始物品透底图对应的目标图像。

[0116] 或者,上述执行主体可以通过文生图扩散模型对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0117] 在本实现方式中,通过对参考特征信息进行信息过滤,使得过滤后的参考特征信息可以更关注于参考图像的背景信息,提高了参考特征信息的准确性和针对性,从而使得目标图像更准确。

[0118] 参考图6,示出了训练物品特征提取模型和文生图扩散模型的一个实施例的流程图600,即上述物品特征提取模型和文生图扩散模型可以基于以下步骤训练得到:

[0119] 步骤610,响应于获取到样本物品图像、样本物品掩膜和样本物品类别,基于样本物品图像和样本物品掩膜,获取样本物品透底图像。

[0120] 在本步骤中,上述执行主体可以通过网络读取获取到样本物品图像,对样本物品图像进行物品类别分析,获取到样本物品类别和样本物品掩膜。上述执行主体可以利用样本物品掩膜对样本物品图像进行图像处理,获取到样本物品图像对应的样本物品透底图像,该样本物品透底图像可以表征只包括样本物品、不包括背景的图像。

[0121] 步骤620,基于样本物品类别,构建样本文本提示词和样本背景提示词。

[0122] 在本步骤中,上述执行主体可以读取文本提示词的文本构建模版,利用文本构建模板对样本物品类别进行文本构建,生成样本物品透底图像对应的文本提示词,该文本提示词可以表征初始物品对应的物品类别,该文本构建模版可以为“A photo of初始物品类别”。

[0123] 上述执行主体可以根据样本物品类别与物品类别编码之间的对应关系,确定出样本物品类别对应的初始物品类别编码。上述执行主体还可以将样本物品类别编码与预设字符进行拼接,得到样本物品类别对应的背景信息,该预设字符可以是预先设定的字符,例如“sks”。上述执行主体可以读取背景提示词的背景构建模版,利用背景构建模板对背景信息进行构建,生成样本物品透底图像对应的背景提示词,该背景提示词可以表征初始物品对应的背景信息,该背景构建模版可以为“in the background of背景信息”。

[0124] 步骤630,基于样本物品图像,获取样本噪声图像。

[0125] 在本步骤中,上述执行主体可以根据预设噪声设置对样本物品图像进行噪声叠加处理,生成样本物品图像对应的样本噪声图像,该预设噪声设置可以包括对图像进行预设次噪声叠加处理。作为示例,预设噪声设置可以包括对图像进行50次噪声叠加处理,则上述执行主体可以对样本物品图像进行50次噪声叠加处理,生成样本物品图像对应的样本噪声

图像。

[0126] 步骤640,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像、样本文本提示词和初始物品特征提取模型的输出结果作为初始文生图扩散模型的输入,对初始物品特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型。

[0127] 在本步骤中,上述执行主体可以获取初始物品特征提取模型和初始文生图扩散模型,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像、样本文本提示词和初始物品特征提取模型的输出结果作为初始文生图扩散模型的输入,采用自监督的方式对初始物品特征提取模型和初始文生图扩散模型同时进行训练,得到物品特征提取模型和文生图扩散模型。

[0128] 具体地,上述执行主体可以将样本物品透底图像、样本文本提示词和样本背景提示词输入至初始物品特征提取模型,初始物品特征提取模型对样本物品透底图像、样本文本提示词和样本背景提示词进行特征提取,输出样本物品特征信息。

[0129] 上述执行主体可以将样本噪声图像和样本文本提示词输入至初始文生图扩散模型,初始文生图扩散模型对样本噪声图像和样本文本提示词进行特征提取,并对提取的特征和初始物品特征提取模型的输出结果进行解码处理,输出预测图像,不断调整初始物品特征提取模型和初始文生图扩散模型的网络参数,直至预测图像满足训练条件,则模型训练完成,得到物品特征提取模型和文生图扩散模型。

[0130] 在本实现方式中,通过模型采用自监督的方式进行训练,使得生成的物品特征提取模型和文生图扩散模型能够同时考虑物品特征信息和背景特征信息,从而可以更准确地生成目标图像。

[0131] 作为一个可选实现方式,上述方法还可以包括以下步骤:获取参考图像对;以及,上述步骤630,基于样本物品图像,获取样本噪声图像,可以包括以下步骤:基于预设噪声设置和参考图像对,获取样本噪声图像;以及,上述步骤640,将样本物品透底图像和样本文本提示词作为物品特征提取模型的输入,将样本噪声图像、样本文本提示词和样本背景提示词作为文生图扩散模型的输入,对物品特征提取模型和文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型,可以包括以下步骤:将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将参考图像对作为初始参考特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型、参考特征提取模型和文生图扩散模型,其中,参考特征提取模型用于对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

[0132] 具体地,上述执行主体可以预先获取参考图像对,该参考图像对可以包括物品图像和物品掩膜,上述执行主体可以根据预设噪声设置对参考图像对进行噪声叠加处理,生成样本噪声图像,该预设噪声设置可以包括对参考图像对进行预设次噪声叠加处理。

[0133] 上述执行主体可以获取初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将参考图像对作为初始参考特征提取模型的输入,将样本噪声图像、样本文本提示词和初始物品特征提取模型的输出结果作为初始文生图扩散模型的输入,采用

自监督的方式对初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型同时进行训练,得到物品特征提取模型、参考特征提取模型和文生图扩散模型。其中,该物品特征提取模型用于对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息;该参考特征提取模型用于对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息;该文生图扩散模型用于对目标噪声图像、文本提示词、参考特征信息和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0134] 具体地,上述执行主体可以将样本物品透底图像、样本文本提示词和样本背景提示词输入至初始物品特征提取模型,初始物品特征提取模型对样本物品透底图像、样本文本提示词和样本背景提示词进行特征提取,输出样本物品特征信息。

[0135] 上述执行主体可以将参考图像对输入至初始参考特征提取模型,初始参考特征提取模型对参考图像对进行特征提取,输出样本参考特征信息。

[0136] 上述执行主体可以将样本噪声图像和样本文本提示词输入至初始文生图扩散模型,初始文生图扩散模型对样本噪声图像和样本文本提示词进行特征提取,并对提取的特征、样本物品特征信息和样本参考特征信息进行解码处理,输出预测图像,不断调整初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型的网络参数,直至预测图像满足训练条件,则模型训练完成,得到物品特征提取模型、参考特征提取模型和文生图扩散模型。

[0137] 在本实现方式中,通过模型采用自监督的方式进行训练,使得生成的物品特征提取模型、参考特征提取模型和文生图扩散模型能够同时考虑物品特征信息、参考特征信息和背景特征信息,从而可以更准确地生成目标图像。

[0138] 参考图7,示出了获取参考图像对的一个实施例的流程图700,即上述获取参考图像对可以包括以下步骤:

[0139] 步骤710,对样本物品掩膜进行膨胀,得到膨胀物品掩膜。

[0140] 在本步骤中,上述执行主体可以利用第一预设概率对样本物品掩膜进行膨胀,得到膨胀物品掩膜,该膨胀可以对二值化物体边界点进行扩充,将与物体接触的所有背景点合并到该物体中,使边界向外部扩张,可以包括对样本物品掩膜进行水平膨胀、垂直膨胀、全方向膨胀等操作。

[0141] 步骤720,对样本物品图像进行数据增强,得到新的样本图像。

[0142] 在本步骤中,上述执行主体可以对样本物品图像进行数据增强,以预设概率对样本物品图像执行mixup操作,即可以对样本物品图像和一张随机广告图像在每个位置上的像素值进行加权求和,从而得到新的样本图像。

[0143] 步骤730,对新的样本图像和膨胀物品掩膜进行平移操作,得到平移样本图像和平移物品掩膜。

[0144] 在本步骤中,上述执行主体可以利用第二预设概率对新的样本图像和膨胀物品掩膜进行平移操作,在平移过程中新的样本图像和膨胀物品掩膜保持同步,形成的空白用0进行填充,得到平移样本图像和平移物品掩膜。

[0145] 步骤740,对平移物品掩膜进行随机掩膜处理,生成包括多个随机矩形掩膜的随机物品掩膜。

[0146] 在本步骤中,上述执行主体可以利用第三预设概率对平移物品掩膜进行随机掩膜

处理,即在平移物品掩膜上,以第三预设概率额外生成三个随机位置的矩形掩膜,得到包括多个随机矩形掩膜的随机物品掩膜,该随机物品掩膜既包括样本物品掩膜,又包括多个随机矩形掩膜。

[0147] 步骤750,对平移样本图像和随机物品掩膜进行旋转,得到新的样本物品图像和新的样本物品掩膜,并将新的样本物品图像和新的样本物品掩膜组成参考图像对。

[0148] 在本步骤中,上述执行主体可以利用第四预设概率对平移样本图像和随机物品掩膜进行旋转,在旋转过程中平移样本图像和随机物品掩膜保持同步,形成的空白用0进行填充,得到新的样本物品图像和新的样本物品掩膜。上述执行主体可以将新的样本物品图像和新的样本物品掩膜组成参考图像对。

[0149] 在本实现方式中,通过对样本物品图像进行一系列数据增强处理,构造出“新的样本物品图像和新的样本物品掩膜”的参考图像对,避免人工收集参考图像对,降低训练成本,提高了模型训练效率。

[0150] 参考图8,作为对上述各图所示方法的实现,本公开提供了一种图像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。

[0151] 如图8所示,本实施例的图像生成装置800可以包括:构建模块810、获取模块820、第一提取模块830和生成模块840。

[0152] 其中,构建模块810,被配置成响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;

[0153] 获取模块820,被配置成基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像;

[0154] 第一提取模块830,被配置成对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息;

[0155] 生成模块840,被配置成对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0156] 在本实施的一些可选的实现方式中,生成模块840,进一步被配置成:将目标噪声图像和文本提示词输入至文生图扩散模型中的编码器,得到背景特征信息;将物品特征信息和背景特征信息输入至文生图扩散模型中的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0157] 在本实施的一些可选的实现方式中,生成模块840,进一步被配置成:响应于确定去噪图像不满足图像生成条件,将去噪图像作为目标噪声图像。

[0158] 在本实施的一些可选的实现方式中,该装置还包括第二提取模块;获取模块820,进一步被配置成:获取参考图像和参考图像对应的物品掩膜;基于预设噪声设置和参考图像,获取初始物品透底图像对应的噪声参考图像;以及,第二提取模块,被配置成:对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息;以及,生成模块840,进一步被配置成:对噪声参考图像、文本提示词、物品特征信息和参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0159] 在本实施的一些可选的实现方式中,生成模块840,进一步被配置成:将噪声参考图像和文本提示词输入至文生图扩散模型的编码器,得到背景特征信息;将物品特征信息、

参考特征信息和背景特征信息输入至文生图扩散模型的解码器,得到去噪图像;判断去噪图像是否满足图像生成条件;响应于确定去噪图像满足图像生成条件,将去噪图像确定为初始物品透底图对应的目标图像。

[0160] 在本实施的一些可选的实现方式中,生成模块840,进一步被配置成:响应于确定去噪图像不满足图像生成条件,将去噪图像作为噪声参考图像。

[0161] 在本实施的一些可选的实现方式中,该装置还包括过滤模块;过滤模块,被配置成:基于初始物品透底图像对应的物品掩膜,对参考特征信息进行信息过滤,得到过滤后的参考特征信息;以及,生成模块840,进一步被配置成:对噪声参考图像、文本提示词、物品特征信息和过滤后的参考特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0162] 在本实施的一些可选的实现方式中,第一提取模块,进一步被配置成:通过物品特征提取模型对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,其中,物品特征提取模型包括物品掩膜的注意力模块。

[0163] 在本实施的一些可选的实现方式中,物品特征提取模型和文生图扩散模型基于训练单元训练得到;训练单元,被配置成:响应于获取到样本物品图像、样本物品掩膜和样本物品类别,基于样本物品图像和样本物品掩膜,获取样本物品透底图像;基于样本物品类别,构建样本文本提示词和样本背景提示词;基于样本物品图像,获取样本噪声图像;将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型和文生图扩散模型。

[0164] 在本实施的一些可选的实现方式中,训练单元,进一步被配置成:获取参考图像对;以及,将样本物品透底图像、样本文本提示词和样本背景提示词作为初始物品特征提取模型的输入,将参考图像对作为初始参考特征提取模型的输入,将样本噪声图像和样本文本提示词作为初始文生图扩散模型的输入,对初始物品特征提取模型、初始参考特征提取模型和初始文生图扩散模型进行训练,得到物品特征提取模型、参考特征提取模型和文生图扩散模型,其中,参考特征提取模型用于对参考图像和参考图像对应的物品掩膜进行特征提取,得到参考特征信息。

[0165] 在本实施的一些可选的实现方式中,训练单元,进一步被配置成:对样本物品掩膜进行膨胀,得到膨胀物品掩膜;对样本物品图像进行数据增强,得到新的样本图像;对新的样本图像和膨胀物品掩膜进行平移操作,得到平移样本图像和平移物品掩膜;对平移物品掩膜进行随机掩膜处理,生成包括多个随机矩形掩膜的随机物品掩膜;对平移样本图像和随机物品掩膜进行旋转,得到新的样本物品图像和新的样本物品掩膜,并将新的样本物品图像和新的样本物品掩膜组成参考图像对。

[0166] 本公开的上述实施例提供的图像生成装置,上述执行主体首先响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词,然后基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像,之后对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息,最后对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像,提出了一种端到端的物品广告背景生成方式,可以根据物品透底图像和其所属的物品类别自动生成适

合该类别物品的背景,从而自动生成包括物品和对应背景的目标图像,不需要精细设计提示词,自动根据物品类别构建相应的文本提示词和背景提示词,使得提示词更准确,当不提供参考图像时,自动为物品生成相应的背景,不需要依赖参考图像即可获取到相应的目标图像,提高了图像生成的自动化和生成效率。

[0167] 本领域技术人员可以理解,上述装置还包括一些其他公知结构,例如处理器、存储器等,为了不必要时模糊本公开的实施例,这些公知的结构在图8中未示出。

[0168] 需要说明的是,本公开的技术方案中,所涉及的用户个人信息的采集、收集、更新、分析、处理、使用、传输、存储等方面,均符合相关法律法规的规定,被用于合法的用途,且不违背公序良俗。对用户个人信息采取必要措施,防止对用户个人信息数据的非法访问,维护用户个人信息安全、网络安全和国家安全。

[0169] 下面参考图9,其示出了适于用来实现本公开的实施例的电子设备900的结构示意图。本公开的实施例中的终端设备可以包括但不限于诸如智能屏、笔记本电脑、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。图9示出的终端设备仅仅是一个示例,不应对本公开的实施的功能和使用范围带来任何限制。

[0170] 如图9所示,电子设备900可以包括处理装置(例如中央处理器、图形处理器等)901,其可以根据存储在只读存储器(ROM)902中的程序或者从存储装置908加载到随机访问存储器(RAM)903中的程序而执行各种适当的动作和处理。在RAM 903中,还存储有电子设备900操作所需的各种程序和数据。处理装置901、ROM 902以及RAM 903通过总线904彼此相连。输入/输出(I/O)接口905也连接至总线904。

[0171] 通常,以下装置可以连接至I/O接口905:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置906;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置907;包括例如磁带、硬盘等的存储装置908;以及通信装置909。通信装置909可以允许电子设备900与其他设备进行无线或有线通信以交换数据。虽然图9示出了具有各种装置的电子设备900,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9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。

[0172] 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909从网络上被下载和安装,或者从存储装置908被安装,或者从ROM 902被安装。在该计算机程序被处理装置901执行时,执行本公开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(射频)等等,或者上述的任意合适的组合。

[0173] 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)。

[0174] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、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,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。

[0175] 描述于本申请实施例中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:一种处理器包括构建模块、获取模块、第一提取模块和生成模块,其中,这些模块的名称在某种情况下并不构成对该模块本身的限定。

[0176] 作为另一方面,本申请还提供了一种计算机可读介质,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。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电子设备执行时,使得该电子设备:响应于获取到初始物品透底图像和初始物品类别,基于初始物品类别构建文本提示词和背景提示词;基于预设噪声设置,获取初始物品透底图像对应的目标噪声图像;对初始物品透底图、文本提示词和背景提示词进行特征提取,得到物品特征信息;对目标噪声图像、文本提示词和物品特征信息进行循环图像处理,生成初始物品透底图对应的目标图像。

[0177] 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,本公开的实施例中涉及到的发明范围,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。

100

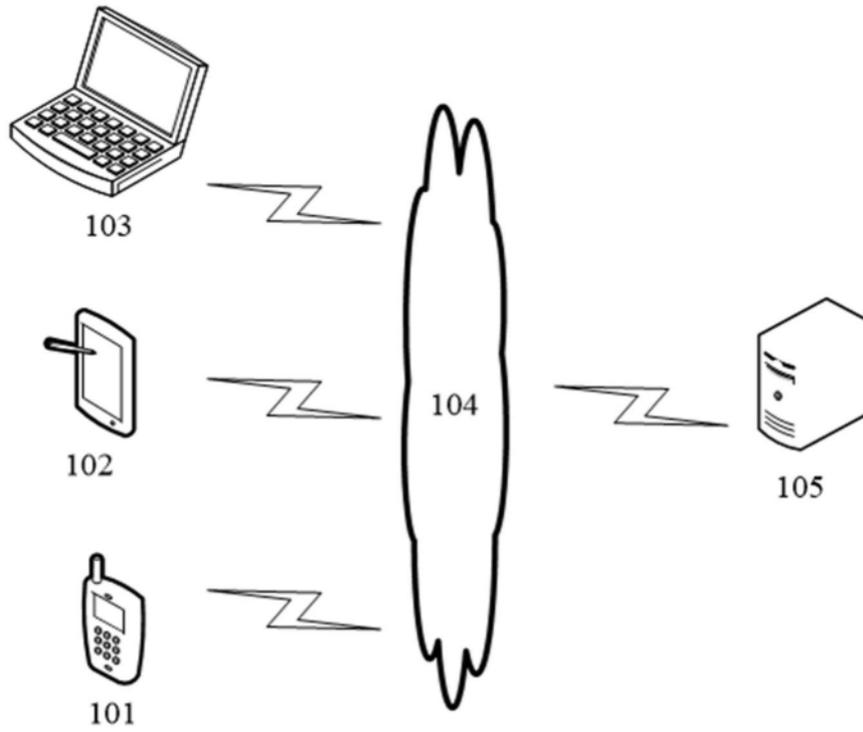


图1

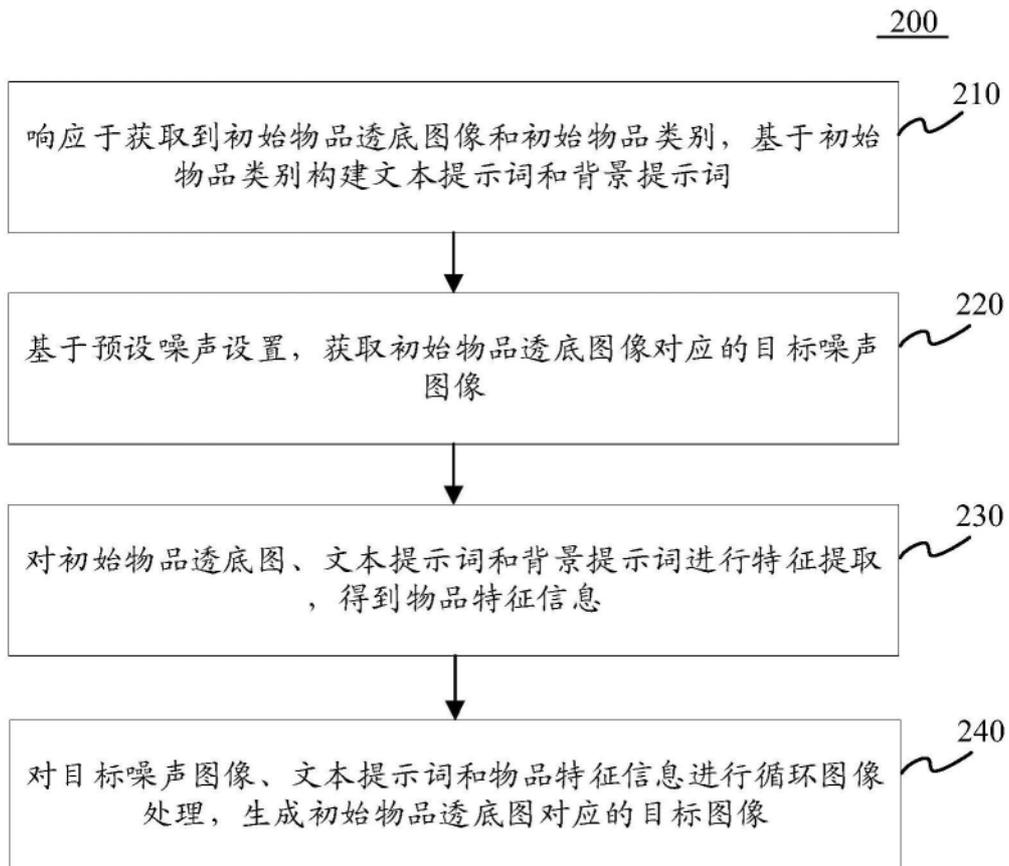


图2

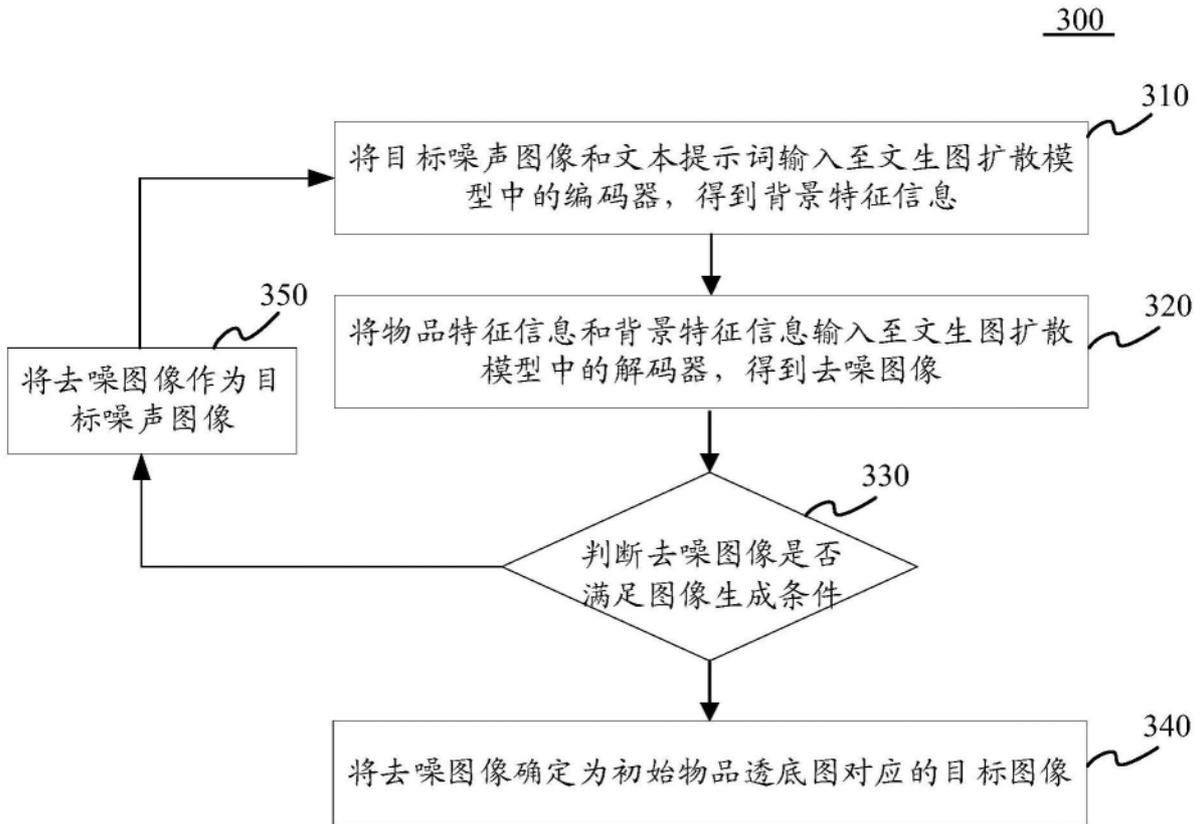


图3

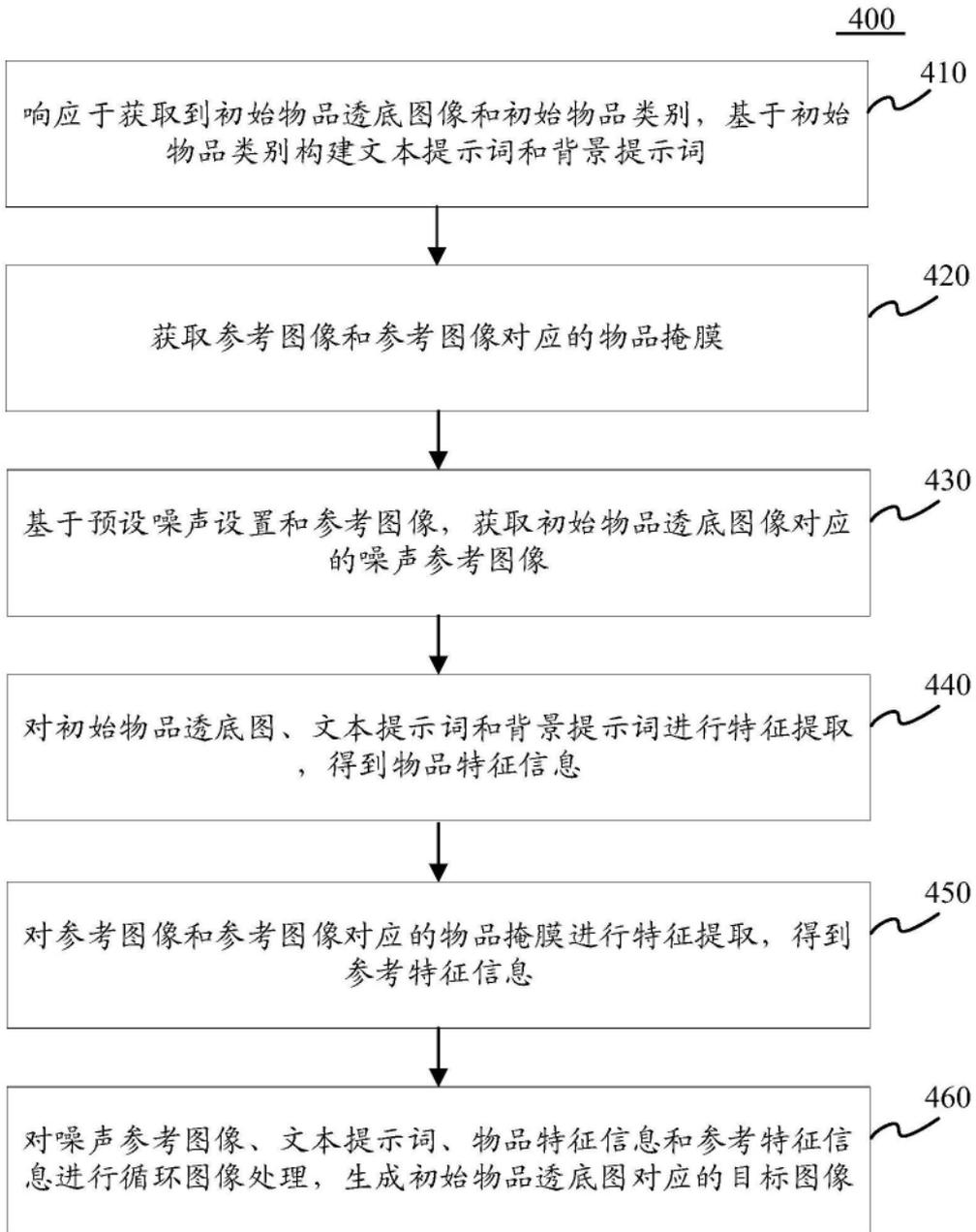


图4

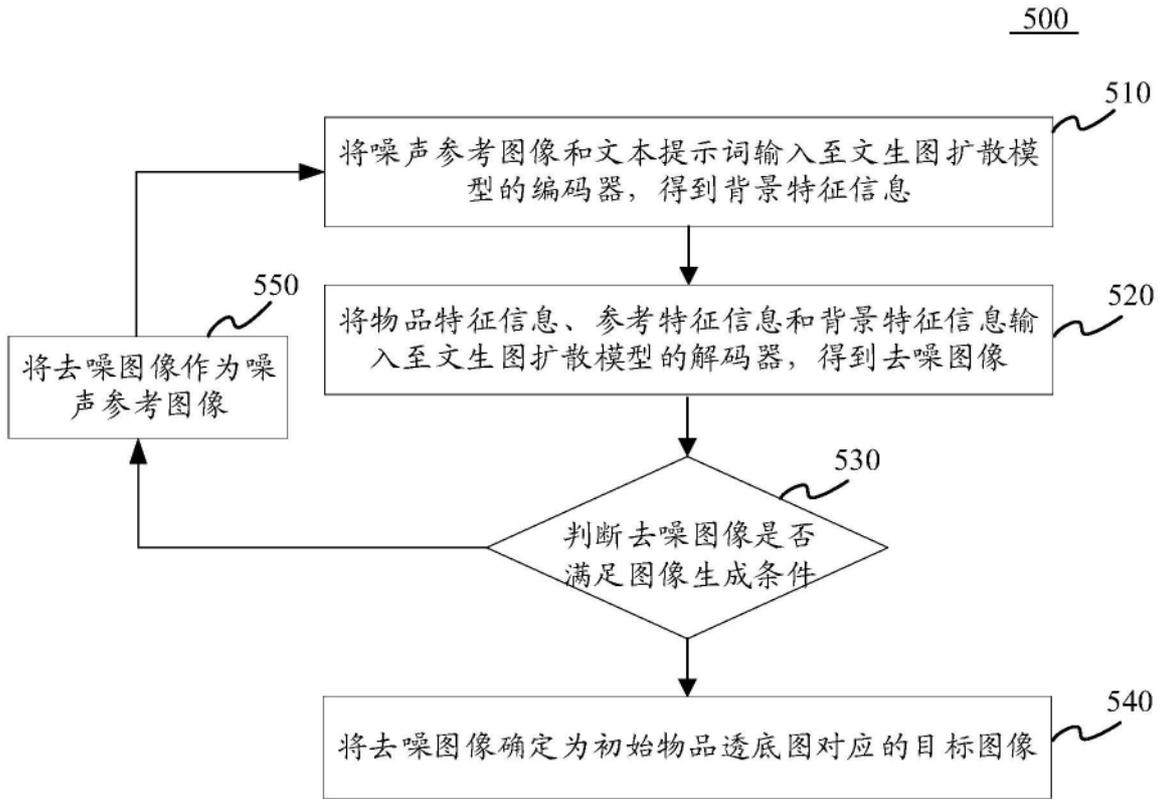


图5

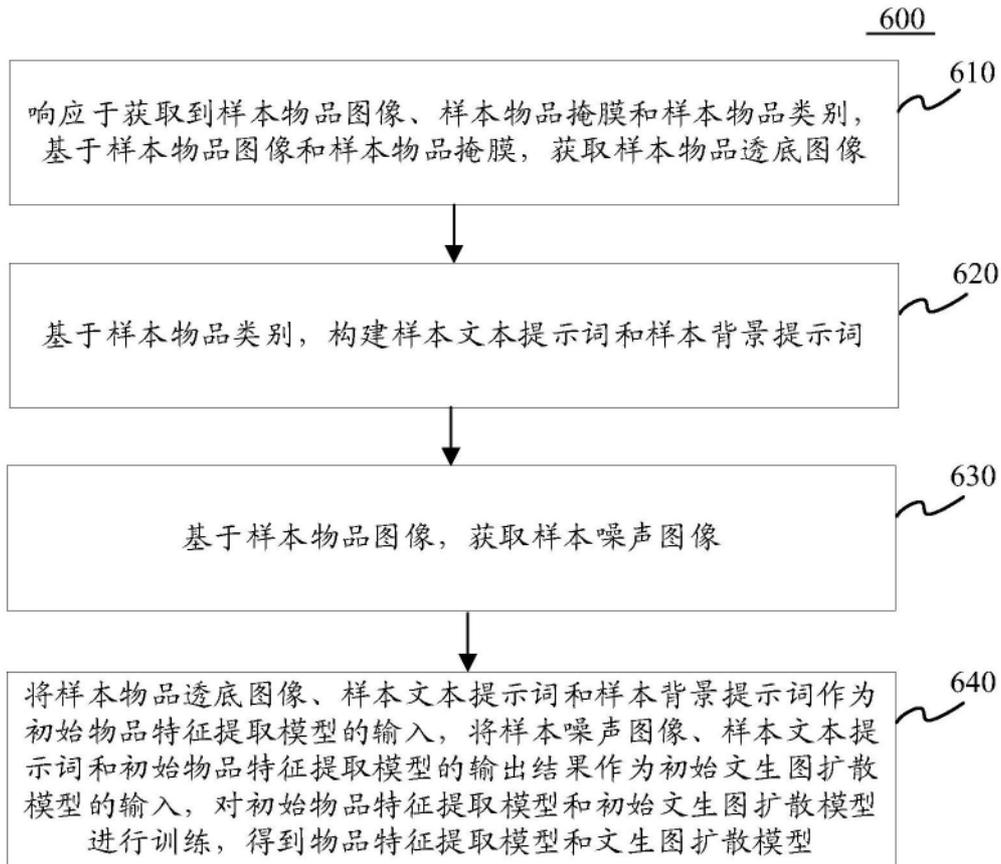


图6

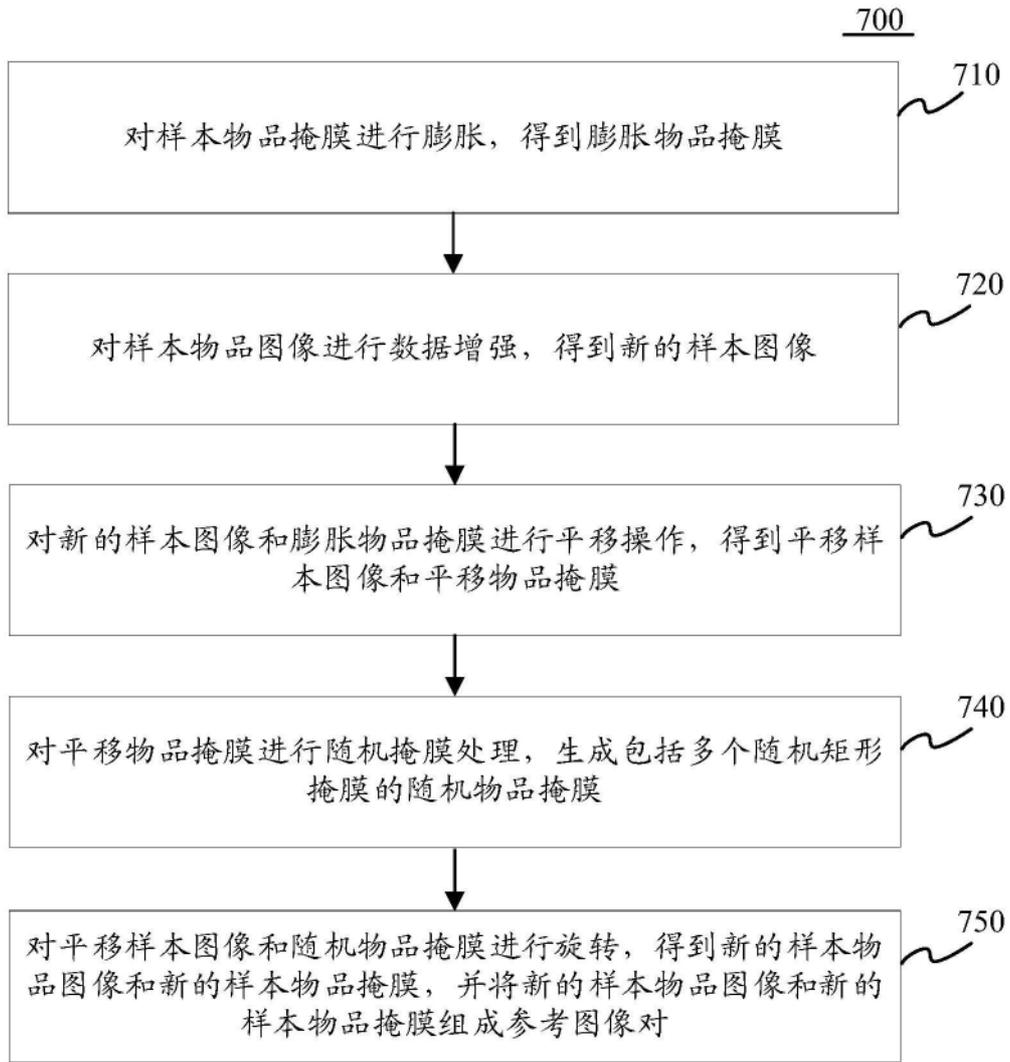


图7

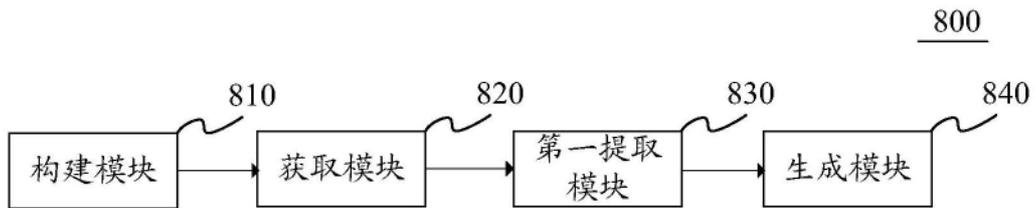


图8

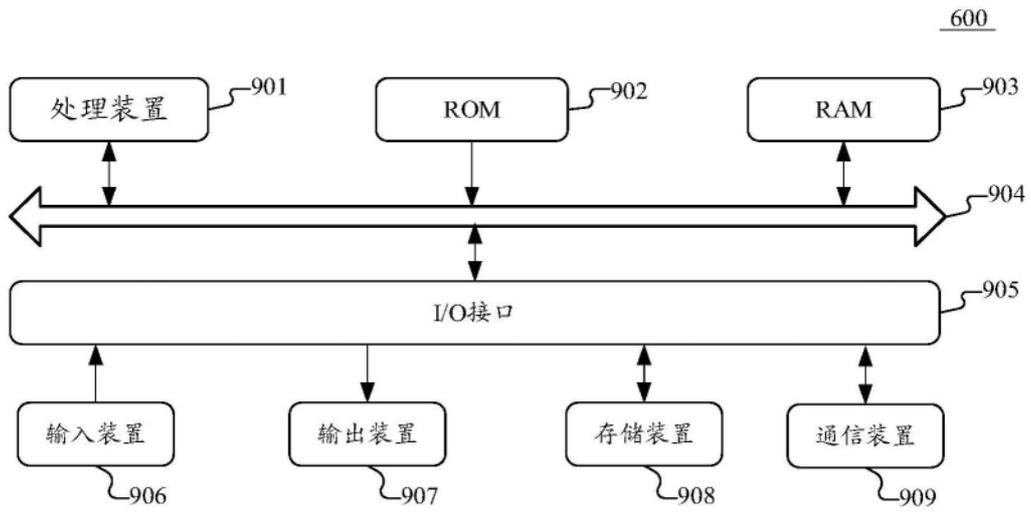


图9