**技术交底书**

|  |  |  |
| --- | --- | --- |
| **案件名称** | 基于扩散模型与光流法的首尾帧发型视频生成技术 | |
| **技术联系人** | **姓名** | 魏林锋 |
| **电话** | 18770721620 |
| **邮箱** | icecocoeci@163.com |
| **专利类型** |  | |

**注意事项：**

（1）交底书应使代理人能看懂，尤其是背景技术和详细技术方案，一定要写得全面、清楚、完整；

（2）技术的公开程度，应以本领域普通技术人员不需付出创造性劳动即可进行实施为准。

（3）在与代理人沟通时，对于代理人咨询的技术问题，应给予回答并认真讲解，并且按要求及时正确地补充相应技术材料。

**一、介绍相关技术背景，描述与本发明技术最相近的现有技术，并说明该现有技术存在的缺点**

1、技术背景(简单介绍一下相关技术背景)+现有技术(并描述已有的与本发明最相近似的实现方案)：

**1.1 技术背景**

随着人工智能技术在图像/视频生成领域的不断发展，发型生成与编辑技术逐渐成为研究热点。传统的发型处理技术如基于生成对抗网络（GAN）、基于深度学习的图像编辑技术以及3D建模与渲染技术等虽取得了一定成果，但仍面临诸多挑战。

本项目旨在利用扩散模型实现AI自动换发型，并通过首尾帧生成视频，从而实现从一张图片生成发型替换的视频。该项目应用并优化了多种先进的人工智能技术和图像处理方法，包括姿态检测、蒙版生成与处理、模型训练优化以及视频生成技术。

**1.2 现有技术**

GAN 由生成器和判别器组成，生成器负责根据输入的人脸图像生成新的发型。它通过学习大量发型数据的分布特征，尝试合成逼真的发型。判别器则对输入的图像（包括生成器生成的发型图像和真实发型图像）进行判断，区分其是真实的还是生成的。通过生成器和判别器之间的对抗训练，生成器不断优化生成结果，使其更加接近真实发型。

**1.2.1 基于深度学习的图像编辑技术**

**（1）语义分割技术**

利用卷积神经网络（CNN）对人脸图像进行语义分割，准确地将头发区域从人脸图像中分离出来。经过大量标注数据的训练，网络能够识别头发的边界和形状。例如基于U-Net等网络架构，可以实现高精度的头发区域分割。

**（2）风格迁移技术**

基于深度神经网络，通过提取内容图像（原始发型）和风格图像（目标发型风格）的特征，将目标发型的风格特征迁移到原始发型上。例如，将卷发的卷曲度和光泽等风格特征迁移到直发上，实现发型风格的转换。典型的方法如基于 VGG网络的风格迁移算法。

**1.2.2 3D建模与渲染技术**

首先通过3D扫描设备获取人脸的3D数据，或者基于2D图像进行3D重建，构建人脸的3D模型。然后在3D模型上进行发型的设计和替换，考虑发型的立体效果和光影变化。最后通过渲染技术将3D发型渲染成2D图像，使其看起来更加逼真。例如，利用Blender等3D建模和渲染软件，结合相关的插件或算法，可以实现复杂发型在3D人脸模型上的创建和渲染。

**2、现有技术存在的缺点：**

**2.1发型生成的逼真度问题**

在 GAN 生成发型的过程中，尽管通过对抗训练能够不断提高生成质量，但仍然存在发型细节不够真实的情况。例如，生成的头发丝的纹理可能不够细腻，在复杂发型（如编发、盘发）的生成上容易出现结构不准确的问题。这是因为训练数据中的发型多样性可能有限，导致生成器无法学习到所有可能的发型细节。

基于图像编辑的方法，在利用语义分割和风格迁移进行发型编辑时，由于是基于 2D 图像操作，在处理发型的空间立体感方面存在不足。比如，在将一种发型的风格迁移到另一种发型时，可能会丢失原发型的一些空间形态特征，导致生成的发型在视觉上不够自然。

**2.2计算资源消耗问题**

基于 GAN 和 3D 建模的方法。GAN 的训练过程需要大量的计算资源和时间，尤其是在处理高分辨率图像和复杂发型时。生成器和判别器的反复迭代需要强大的 GPU 支持，这对于普通用户和一些小型应用来说成本过高。

3D建模与渲染更是对硬件有很高的要求。构建高精度的3D人脸和发型模型以及进行逼真的渲染，需要大量的内存和计算能力，导致软件运行速度慢，限制了这些技术在实时应用中的推广。

**2.3个性化不足**

现有的技术在生成或编辑发型时，往往是基于通用的发型库或风格进行操作，缺乏对用户个人面部特征和风格的充分考虑。例如，生成的发型可能在时尚杂志模特上看起来很好，但并不一定适合特定用户的脸型、气质等，导致生成结果缺乏个性化。

**2.4与原始图像自然融合问题**

在将生成或编辑后的发型与原始人脸图像进行合成时，容易出现边界不自然、颜色不协调等问题。无论是基于2D图像编辑还是 3D 建模渲染回 2D 的方法，都需要更好的融合算法来确保发型与原始图像无缝对接。

**二、针对上述现有技术的缺点，说明本发明所要解决的技术问题(一篇专利仅仅解决一个技术问题，这个技术问题就是一篇专利的纲，下面所有的材料都是为解决这个技术问题而服务的)**

本项目旨在利用扩散模型实现AI自动换发型，并通过首尾帧生成视频，从而实现从一张图片生成发型替换的视频。该项目应用并优化了多种先进的人工智能技术和图像处理方法，包括姿态检测、蒙版生成与处理、模型训练优化以及视频生成技术。

1. **本发明技术方案的详细阐述**（如本发明涉及结构，请提供结构图，并结合附图对本技术方案进行详细描述；如本发明涉及方法步骤，请提供流程图，并结合流程图详细描述本技术方案的实现步骤；附图应是线条图）(注意，流程图可以是功能流程图，但具体要落地成专利的应该是信息处理流程图，功能流程图只帮助理解方案大概实现了什么功能)

**整体方案概述**

本方案主要涵盖两个核心部分，分别为关键帧制作与基于关键帧生成视频。通过一系列技术步骤，实现从原始数据到最终流畅视频的生成。

#### 第一部分：关键帧制作

**步骤一：发型模型训练**

输入：包含各类发型的图像数据集，以及预训练的基础模型。

操作：运用正则化方法，对模型参数进行约束，防止过拟合；采用 Dropout 技术，在训练过程中随机忽略部分神经元，增强模型的泛化能力；利用 Locon 技术对模型结构进行优化，提升局部特征的提取能力；通过分层网络控制，对不同层次的网络参数进行针对性调整。

输出：训练后的发型生成模型，能够生成多样化的发型。

**步骤二：身体姿态识别**

输入：用户上传的图像。

操作：运用骨骼关键点检测技术，在用户图像中标记出人体骨骼的关键节点，如头部、肩部、肘部、手腕、髋部、膝盖、脚踝等，计算各关键点之间的相对位置和角度关系。

输出：识别出的用户图像骨骼姿态数据，包含各关键点的坐标信息。

**步骤三：头发蒙版处理**

输入：用户上传的图像以及姿态识别后的骨骼姿态数据。

操作：利用图像分割算法，基于头发与背景的颜色、纹理等特征差异，将用户图像中的头发区域分割出来，得到头发蒙版；通过高斯模糊算法，对头发蒙版边缘进行平滑处理，减少锯齿现象；运用羽化算法，使头发蒙版边缘过渡自然。

输出：经过平滑与羽化处理的头发蒙版，为新发型生成提供精准的区域界定。

**步骤四：基于控制链的关键帧生成**

输入：步骤一的发型模型，步骤二的用户图像骨骼姿态数据，步骤三的头发蒙版以及新发型提示词

操作：依托 StableDiffusion 模型，以骨骼姿态数据为约束，确保生成的关键帧姿态与用户图像一致；以头发蒙版为区域限定，在发型区域利用发型生成模型，结合新发型参考图像的特征，生成新发型；通过融合损失函数，调整新发型与身体其他部分的融合效果，使两者自然衔接。

输出：生成的关键帧图像，包含与用户姿态一致且发型自然融合的新形象。

**第二部分：基于关键帧生成视频**

**步骤一：运动向量计算**

输入：原图和第一部分步骤四生成的关键帧图像。

操作：采用光流法，计算两帧图像中每个像素点的运动位移，得到每个像素点在两帧之间的运动方向和速度，形成运动向量场。

输出：两帧关键帧之间的运动向量数据，描述了图像中物体的运动轨迹和速度。

**步骤二：帧插值**

输入：步骤一的运动向量数据以及相邻的两帧关键帧图像。

操作：将运动向量插入到中间帧，利用线性插值算法，根据相邻关键帧像素值和运动向量，计算中间帧每个像素点的像素值；或采用样条插值算法，通过构建平滑曲线，对中间帧像素进行插值计算，生成中间过渡帧。

输出：插入运动向量并经过插值计算得到的中间帧图像，增加了视频的帧率。

**步骤三：多轮帧插值**

输入：经过一次帧插值后的视频序列（包含关键帧和中间帧）。

操作：对视频序列中相邻的帧再次执行帧插值操作，不断重复上述步骤，逐步增加视频的帧率。

输出：生成的流畅视频，满足视觉上的连贯性和流畅性需求。

**3.1头发蒙版生成和处理**

**3.1.1姿态检测与保持一致**

首先利用Openpose进行原图姿态检测。Openpose是一种先进的人体姿态估计技术，它能够准确地识别图像中人体的关键点和姿态信息，包括身体躯干、四肢以及脸部。通过获取这些信息，在后续的发型生成过程中，可以确保生成的身体姿态与原图保持一致，从而提高生成效果的真实性和自然度。

**3.1.2头发蒙版识别与叠加**

在姿态保持一致的基础上，通过扩散模型stable diffusion和训练的发型Lora生成指定发型，并提取头发蒙版。这一步骤涉及到对图像中头发区域的特征提取和分析，以框定头发的边界和范围。

**3.1.3基于颜色相似度的边缘羽化**

使用Opencv Guidedfilter 根据颜色相似度对头发蒙版的边缘进行羽化处理。这种方法在边缘具有很强的颜色分离时效果最佳。Guidedfilter是一种基于引导图像的滤波方法，它通过参考原图来对目标头发蒙版进行滤波处理。通过这种方式，可以使头发蒙版的边缘更加平滑自然，减少生硬的边界感，从而在后续的发型替换过程中能够更好地与新发型融合。

**3.2发型 Lora 训练**

**3.2.1 LoRA 技术原理**

在发型模型微调阶段，大型语言模型的低秩适配（LoRA）技术发挥着关键作用。LoRA 的核心创新在于通过引入秩分解权重矩阵，对现有权重进行调整。具体而言，在不改变模型整体架构的前提下，仅针对新添加的低秩矩阵进行训练，这种方式极大地减少了训练过程中的内存消耗，显著提升了大型模型的训练效率。 从数学原理上看，LoRA利用了矩阵的低秩特性，有效地降低了需要优化的参数规模。在实际应用场景中，LoRA主要针对模型中的UNet结构，尤其是其中的 Cross-Attention层进行修改。Cross-Attention层在整个模型的风格生成机制中占据着核心地位，通过对其进行精细调整，可以精准地控制模型输出的风格特征，从而实现发型风格的多样化变换。

**3.2.2 提高模型泛化能力的策略**

**（1）正则化方法**

为增强微调模型在不同数据场景下的泛化表现，采用了正则化技术。正则化通过在损失函数中引入额外的正则项，对模型复杂度施加约束。其原理在于，当模型复杂度过高时，容易在训练数据上过度拟合，导致在未见过的数据上表现不佳。通过正则化惩罚模型的复杂度过高，能够促使模型学习到更具普遍性的特征，从而有效提升其在未知数据上的预测准确性和泛化能力。

**（2）Dropout 策略**

Dropout 是提高泛化能力的另一重要策略。在训练过程中，Dropout以一定概率随机将神经元的输出置为0，从而在每次迭代中引入随机性。这种随机性能够有效防止模型过度依赖特定的神经元，避免模型陷入局部最优解，进而提高模型对不同数据分布的适应性和鲁棒性。通过在发型模型训练中加入Dropout策略，能够使模型更好地应对各种发型特征的变化，增强其在不同发型图像数据上的泛化效果。

**（3）LoCon 网络结构**

本项目采用了 LoCon（LoRA for Convolution Network）这一优化的网络结构。相较于传统的LoRA，LoCon的显著改进在于其对卷积层的有效控制。在卷积神经网络中，卷积层是提取图像特征的关键组件。LoCon通过对卷积核进行降维操作，在保持模型效果的同时进一步减少了参数数量。这种降维操作不仅提升了模型的训练效率，还优化了模型对图像特征的提取与表示能力，使得模型能够更精准地捕捉发型图像中的细微特征和复杂结构，从而为发型训练任务提供更强大的支持。

**（4）分层控制策略**

针对发型替换任务的独特需求，本项目实施了 LoRA 分层控制策略。该策略聚焦于头发部分的训练，旨在避免原训练集对图像其他部分的干扰，从而提升最终生成图像的质量。在训练过程中，依据发型替换任务的特点，通过精细调整分层训练参数与权重，模型能够更有效地学习头发区域的独特特征与变化规律。通过减少对图像无关区域的关注，模型的训练针对性得以显著增强，从而提高了模型在发型生成任务中的有效性和准确性，能够生成更符合预期且质量更高的发型图像。

**3.3 AI发型重绘**

**3.3.1 重绘模型与智能交互性**

通过上述的头发蒙版处理和发型Lora训练，进行基于Stable Diffusion的重绘，重绘模型采用Brushnet，这一模型在基于Stable Diffusion 的图像重绘任务中展现出卓越的性能。Brushnet能够依据输入的图像信息以及预设的发型样式，以一种高效且精准的方式对图像中的头发区域进行重绘操作。它通过对图像的语义理解和纹理生成能力，将目标发型的特征自然地融合到原始图像的头发区域中，确保重绘后的发型在视觉效果上具有高度的真实感和连贯性。 此外，本项目还创新性地借鉴了MagicQuill在图像编辑过程中的交互性和智能性，使得用户在发型替换过程中能够获得更加便捷、灵活的操作体验。

**3.3.1 多发型智能替换与个性化效果**

凭借上述的头发蒙版处理和发型 LoRA 训练，本项目成功实现了对多种发型的智能替换。在模型训练过程中，通过对丰富多样的发型数据进行广泛学习，模型不仅掌握了各种发型的典型特征，还能够理解不同发型在不同场景、不同人物形象下的适配性变化规律。基于此，用户在使用本系统时，能够根据自身独特的审美需求、特定的使用场景以及个人形象特点，从众多预设发型库中自由地选择最为合适的发型进行替换。无论是时尚的短发造型、优雅的长发款式，还是精致的卷发设计，系统均能够以高质量的重绘效果予以呈现，从而为用户带来高度个性化的 AI 换发型体验，极大地拓展了人工智能在图像编辑与个性化时尚设计领域的应用边界。

**3.4首尾帧生成视频**

采用基于光流法的插帧方法生成连贯视频。以下是基于光流法的详细技术描述，包括运动向量的计算、帧插值的具体方法以及优化细节。

**3.4.1光流法计算运动向量**

（1）光流基本原理

光流法基于图像中像素的灰度在短时间内保持不变的假设，通过光流方程来描述像素的运动。具体而言，其核心公式为：I(x, y, t) = I(x + dx, y + dy, t + dt)

通过对该公式进行一阶泰勒展开，得到光流约束方程：I\_x \* u + I\_y \* v + I\_t = 0

其中：

I\_x和I\_y为图像x和图像y方向上的空间梯度

I\_t：图像在时间方向上的梯度

u和v：像素在x和y方向上的运动向量分量

该方程为欠约束问题，需引入额外假设或优化算法来解算运动向量。

（2）光流计算方法

光流计算主要分为稀疏光流和稠密光流两类。考虑到发型需要快速变换的场景，本方案采用稠密光流的 Pyramidal Lucas - Kanade 方法。该方法结合了 Lucas - Kanade 和图像金字塔技术，其计算过程是从低分辨率到高分辨率逐层计算运动向量。在低分辨率图像上，由于像素的位移相对较小，能够快速地获得一个大致的运动估计，然后随着分辨率的逐步提高，对运动向量进行细化和精确化，从而有效地处理大位移运动的情况，提高光流计算的准确性和鲁棒性。

**3.4.2帧插值方法**

在计算得到前后帧的光流场后，根据运动向量插入中间帧，实现帧插值。具体步骤如下：

**（1）双向光流计算**

正向光流：从起始帧到目标帧的稠密光流，它描述了每个像素在前一帧中移动到后一帧的位置。通过计算正向光流，可以了解图像中像素在时间序列上向前推进时的运动轨迹和位移信息。

反向光流：从目标帧到起始帧的稠密光流，它表示每个像素在后一帧中移动到前一帧的位置。反向光流与正向光流相互补充，能够为插帧过程提供更全面的像素运动信息，尤其是在处理复杂运动和遮挡情况时具有重要意义。  
 **（2）插帧位置估计**

对于插帧时间，其时间比例α=（t-t0）/（t1-t0）。通过该比例对光流进行插值，从而计算插帧时像素的位置。具体来说，利用正向光流和反向光流以及时间比例信息，能够确定在插帧时刻像素应该处于的空间位置，这是后续准确计算插帧图像像素值的关键步骤。

**（3）像素值插值**

线性插值：对前后两帧像素值按时间比例进行线性混合。即根据插帧时间比例，将前一帧像素值与后一帧像素值进行加权求和，以得到插帧图像中对应像素的灰度值。这种方法简单直接，在像素运动较为平滑且无明显遮挡的情况下能够取得较好的效果。  
 双边插值：当正向和反向位置不一致时（可能由于遮挡或运动估计误差），采用双边插值对两个位置进行加权融合，以提高插值精度。双边插值不仅考虑了像素的空间位置关系，还考虑了像素值的相似性，能够在处理复杂场景下的像素插值时更好地保留图像的边缘和细节信息，减少因位置不一致而导致的插值误差。

**3.4.3遮挡处理**

在插帧过程中，可能出现像素遮挡现象，即某些区域在插帧时无对应的运动向量。针对此问题，采用以下方法进行处理：

**遮挡区域检测：**通过正向光流和反向光流的不一致性来检测遮挡区域。当正向光流和反向光流所指示的像素位置差异较大时，可判定该区域可能存在遮挡情况。这种基于光流不一致性的检测方法能够有效地识别出图像中由于物体运动而产生遮挡的区域，为后续的遮挡处理提供依据。

**邻域插值：**对遮挡区域的像素值使用周围非遮挡区域的像素值进行插值。在确定了遮挡区域后，利用周围非遮挡区域像素的信息，通过一定的插值算法（如双线性插值或其他合适的邻域插值方法）来估算遮挡区域的像素值，从而尽可能地恢复被遮挡区域的图像信息，减少遮挡对插帧图像质量的影响。

**深度信息辅助：**结合深度信息进行遮挡处理。深度信息能够提供图像中不同物体的空间位置关系，在处理遮挡问题时，可以根据深度信息判断遮挡物与被遮挡物的前后关系，从而更准确地进行遮挡区域的像素值估计。例如，对于被遮挡的背景区域，可以根据其深度信息和周围背景区域的像素值进行更合理的插值计算。

**3.4.4 优化措施**

**（1）多尺度光流估计**

使用金字塔分层方法，从低分辨率到高分辨率逐层估计光流。在低分辨率层，由于图像尺寸较小，像素的位移相对较大，能够快速捕捉到图像的整体运动趋势，减少因快速运动和大位移带来的误差。随着分辨率的逐步提高，对光流估计进行细化，能够更精确地处理图像中的细节运动信息，从而提高光流计算的准确性和稳定性。

**（2）正则化约束**

在光流计算中加入边缘保持或全局光滑性约束。在图像的纹理区域，光流容易出现过度抖动的情况，通过引入边缘保持约束，可以使光流在物体边缘处保持连续性，避免因纹理变化而导致的错误运动估计；同时，全局光滑性约束能够使整个光流场更加平滑，减少局部异常光流的出现，提高光流计算的可靠性和视觉效果。

**（3）边缘插值平滑**

对插帧中出现的边缘伪影进行后处理，采用高斯平滑或双边滤波等方法。在插帧过程中，由于像素插值和运动估计等操作，可能会在图像边缘产生一些伪影，这些伪影会影响图像的视觉质量。通过高斯平滑或双边滤波等边缘插值平滑技术，可以有效地去除这些边缘伪影，使插帧后的图像边缘更加自然流畅，提高视频的整体视觉效果。

1. **与现有技术相比，本发明技术方案具有哪些优点？**

本项目通过先进的人工智能技术和图像处理方法，实现了从一张图片生成发型替换的视频。本项目提出一系列创新技术方案，涵盖头发蒙版生成与处理、发型 Lora 训练、发型重绘以及视频生成等多个关键环节。

在头发蒙版生成和处理阶段，采用了基于扩散模型 stable diffusion 的姿态检测技术，借助 Openpose 精准识别原图人体姿态信息，并确保在后续发型生成过程中姿态的一致性，极大地提升了生成效果的真实性与自然度。同时，运用 opencv guidedfilter 依据颜色相似度对头发蒙版边缘进行羽化处理，有效解决了头发蒙版边缘生硬、与原始图像融合不自然的问题，为发型的精准替换和自然融合奠定了坚实基础。 发型 Lora 训练环节引入了多项创新技术以提升训练效果和模型性能。为增强模型的泛化能力，综合运用了多种策略，包括正则化方法、Dropout 策略、LoCon网络结构、LoRA 分层控制策略等。在维持模型效果的同时进一步减少参数数量，不仅提升了模型训练效率，还优化了对发型图像特征的提取与表示能力，使其能够更敏锐地捕捉发型图像中的细微特征和复杂结构。发型重绘方面，采用在基于Stable Diffusion 的图像重绘任务中表现卓越的 Brushnet 模型进行重绘操作。Brushnet 凭借其强大的图像语义理解和纹理生成能力，能够依据输入的图像信息以及预设的发型样式，将目标发型特征自然流畅地融合到原始图像的头发区域，确保重绘后的发型在视觉效果上呈现出高度的真实感和连贯性。同时，创新性地借鉴 MagicQuill 在图像编辑过程中的交互性和智能性，赋予用户在发型替换过程中更加便捷、灵活的操作体验，实现了技术与用户体验的深度融合。

在视频生成环节，基于光流法进行了全面创新和优化。采用稠密光流的 Pyramidal Lucas - Kanade 方法计算运动向量，该方法结合了 Lucas - Kanade 和图像金字塔技术，从低分辨率到高分辨率逐层计算运动向量。在低分辨率图像上，能够快速获取大致的运动估计，随着分辨率逐步提高，对运动向量进行细化和精确化，有效处理大位移运动情况，显著提高光流计算的准确性和鲁棒性。在帧插值方法上，首先进行双向光流计算，即计算从起始帧到目标帧的正向光流以及从目标帧到起始帧的反向光流，两者相互补充，为插帧过程提供全面的像素运动信息，尤其在处理复杂运动和遮挡情况时发挥关键作用。然后，依据插帧时间比例对光流进行插值以确定插帧时像素的位置，并采用线性插值和双边插值相结合的方式计算插帧图像像素值。线性插值在像素运动较为平滑且无明显遮挡的情况下能够简单有效地获取插帧像素值，而双边插值则在正向和反向位置不一致（如因遮挡或运动估计误差）时，通过考虑像素的空间位置关系和像素值的相似性，对两个位置进行加权融合，有效提高插值精度，更好地保留图像边缘和细节信息。针对插帧过程中可能出现的像素遮挡现象，通过正向光流和反向光流的不一致性检测遮挡区域，利用周围非遮挡区域像素值进行邻域插值估算遮挡区域像素值，并结合深度信息辅助判断遮挡物与被遮挡物的前后关系，从而更精准地进行遮挡区域的像素值估计，最大限度减少遮挡对插帧图像质量的影响。此外，还采用了一系列优化措施，如多尺度光流估计，使用金字塔分层方法从低分辨率到高分辨率逐层估计光流，先捕捉整体运动趋势后细化细节运动信息，提高光流计算准确性和稳定性；在光流计算中加入边缘保持或全局光滑性约束，避免光流在物体边缘处因纹理变化导致的错误运动估计，使光流场更平滑，减少局部异常光流，提升光流计算的可靠性和视觉效果；对插帧产生的边缘伪影采用高斯平滑或双边滤波等方法处理，使插帧后的图像边缘更自然流畅，提高视频整体视觉效果。这些创新技术在视频生成过程中的综合应用，有效解决了传统视频生成技术在处理发型相关视频时面临的诸多问题，显著提升了视频质量和视觉效果。

**五、本发明的技术关键点和欲保护点是什么？**

**（1）头发蒙版生成和处理技术**

**（2）发型Lora训练技术**

**（3）AI发型重绘技术**

**（4）首尾帧生成视频技术**

**六、其它**

流程图

图示

描述已自动生成