1. 一种基于共享注意力的多模态人物图像生成算法，其特征在于，包括如下步骤：

S1、对现有的多模态人物数据集进行预处理，构建人物图像、文本信息、语义信息数据组；

S2、在现有数据组下，构建包含基于小波变换的离散变分自编码器模型和基于共享注意力的多模态条件扩散模型的多模态人物图像生成模型，并初始化网络权重。

S2-1、通过所述基于小波变换的离散变分自编码器模型重构输入图像，重构的图像为，同时学习人物图像的离散码本。

S2-2、通过所述多模态条件扩散模型，将文本描述，人物分割图和人物素描图作为多模态输入。离散扩散过程中，多模态条件扩散模型在多模态控制下最大化条件概率，估计图像离散空间上的潜在先验分布，包括在每一步添加噪声的前向扩散过程和从噪声状态恢复样本的反向去噪过程，从而实现基于多模态输入的可控人物图像生成。

S3、利用反向传播算法学习模型参数，直至网络收敛；

S4、输出模型。

2.根据权利要求1所述的一种多模态引导的人物图像生成方法，其特征在于S1所述的构建人物图像、文本信息、语义信息数据组，是指对现有的多模态人物数据集的预处理方法是，使用先进的人体解析器来计算一个包含24个标签的人体语义信息，每个标签代表人体的特定部分，比如脸、头发、胳膊、腿和衣服区域。关于文本信息，本发明基于DeepFashion-MultiModal数据集中的独热编码形状属性和纹理属性为每个人物图像创建了一个文本模板。例如：一个穿着纯色长袖上衣和牛仔长裤的男人；

3.根据权利要求2所述的一种多模态引导的人物图像生成方法，其特征在于S2-1所述的基于小波变换的离散变分自编码器模型由小波增强编码器、解码器和离散码本组成。码本包含一组码本条目，其中是码本的大小，是编码的维度。小波增强编码器接受一个全身人物图像作为输入，并输出混合特征。首先，将输入图像输入到空间编码器中提取空间特征。为了恢复丢失的高频信息，对应用离散小波变换，得到四个离散小波子带。然后，将这些离散小波子带串联起来，并将其输入到小波编码器中提取小波域特征。最后，应用交叉注意力块将空间特征和小波域特征融合，得到混合特征，其中本文将小波域特征作为“查询”（），将空间特征作为“键”（）和“值”（）。将输入图像压缩为潜在向量后，接下来的空间量化器将每个空间网格特征映射到最近的码本条目：

其中，是量化操作，是量化特征。表示中特定空间位置上的向量。解码器经过训练，可以从量化编码中重构出人物图像，即。

4.根据权利要求3所述的一种姿态和纹理引导的时尚服装设计合成的方法，其特征在于S2-2所述的多模态条件扩散模型由多模态特征融合模块（MFFB）和半解耦去噪解码器（S3D）组成。MFFB 模块将用来处理前期多种视觉模态信息的融合，尤其针对人物分割图与人物素描图两种模态创新地提出了共享多头注意力机制( SMHA)，通过多头结构和自适应层归一化（AdaLN），捕捉模态数据之间的复杂关联信息,将两种视觉模态的信息进行融合。将经由MFFB模块处理得到的视觉信息和文本信息利用半解耦多头注意力机制输入到 S3D中通过与图像特征的先验噪声和时间步融合，逐步预测图像每一时间步的离散先验。

5.根据权利要求4所述的一种姿态和纹理引导的时尚服装设计合成的方法，其特征在于所述的多模态图像生成模型的目标函数包括重构损失、码本损失、约束损失、最小化对抗损失及交叉熵损失，利用收集到的数据组训练基于小波变换的离散变分自编码器模型和基于共享注意力的多模态人物图像生成模型。