

DOI: 10.19650/j.cnki.cjsi.J2311948

# 基于纹理图像和用户动作信息的纹理 触觉再现方法研究\*

陈大鹏<sup>1,2,3</sup>, 陈庚<sup>1,2</sup>, 刘佳<sup>1,2,3</sup>, 方映平<sup>1,2</sup>, 张运杰<sup>1,2</sup>

(1. 南京信息工程大学自动化学院 南京 210044; 2. 江苏省智能气象探测机器人工程研究中心 南京 210044;  
3. 江苏省大气环境与装备技术协同创新中心 南京 210044)

**摘要:**近年来,对虚拟表面纹理进行工具介导的振动触觉再现已经成为触觉领域的一个研究热点。针对已有的纹理触觉渲染方法还存在适用范围窄、泛化能力弱、交互真实感低等问题,以改进后的 MelGAN 为基本网络,构建了一种新的纹理触觉渲染模型。该模型以纹理图像和用户的实时动作信息为输入,在能够高保真地生成振动触觉信号的同时,对常见纹理图像拥有更好的泛化能力。此外,设计了一款具有动作信息实时采集和振动触觉表达功能的笔式装置。在采集了数据库之外的真实纹理表面的振动触觉信号后,比较了提出的模型与已有方法在信号生成方面的性能差异。结果表明,模型实现了最低的均方根误差(0.173),验证了其对未建模纹理进行触觉渲染的能力。最后,使用笔式装置进行了两项用户实验。平均 6.01 的主观相似度评分说明,即使是数据库之外的新纹理,模型也能为用户提供较高的纹理交互真实感。

**关键词:**纹理触觉渲染;深度学习;动作信息;笔式交互;振动触觉反馈

中图分类号: TH7 TP391 文献标识码: A 国家标准学科分类代码: 520.60

## Research on texture haptic display method based on texture image and user action information

Chen Dapeng<sup>1,2,3</sup>, Chen Geng<sup>1,2</sup>, Liu Jia<sup>1,2,3</sup>, Fang Yingping<sup>1,2</sup>, Zhang Yunjie<sup>1,2</sup>

(1. School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Jiangsu Province Engineering Research Center of Intelligent Meteorological Exploration Robot (C-IMER), Nanjing 210044, China; 3. Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAET), Nanjing 210044, China)

**Abstract:** In recent years, tool-mediated haptic feedback for virtual surface textures has become a hot topic in the field of haptics. In view of the problems of narrow application range, weak generalization ability and low interactive realism of the existing haptic texture rendering methods, a new texture haptic rendering model is constructed in this paper based on the improved MelGAN. This model takes texture image and real-time user action information as inputs, which can generate vibrotactile signals with high fidelity and has better generalization ability for common texture images. Furthermore, this paper designs a pen-type device with real-time action data acquisition and vibrotactile expression. After collecting vibrotactile signals from real texture surfaces outside the database, this paper compared the performance differences between the proposed model and existing methods in signal generation. The results indicate that the model in this paper achieved the lowest root mean square error (0.173), verifying its ability to perform haptic rendering on unmodeled textures. Finally, this paper conducted two user experiments using a pen-type device. A subjective similarity score of 6.01 on average indicates that even for new textures outside of the database, our model can provide users with a high level of texture interaction realism.

**Keywords:** texture haptic rendering; deep learning; action information; pen-based interaction; vibrotactile feedback

收稿日期: 2023-09-20 Received Date: 2023-09-20

\* 基金项目: 国家自然科学基金(62003169)、江苏省自然科学基金青年基金(BK20200823)、江苏产业前瞻与关键核心技术重点项目(BE2020006-2)、江苏省研究生科研与实践创新计划项目(SJCX22\_0348, KYCX22\_1202, SJCX22\_0351)资助

## 0 引言

作为物体表面普遍存在的一种具有视-触觉双重属性的特征<sup>[1]</sup>,纹理对人区分、识别和理解物体起到非常重要的作用<sup>[2]</sup>。纹理材料的触觉感知维度主要包括粗糙度、硬度、摩擦等<sup>[3]</sup>。其中粗糙度是纹理最突出的感知属性之一,且已经被大多数关于纹理的心理物理学研究所关注<sup>[4]</sup>。当用户手持刚性工具(一般为刚性探针)在纹理表面拖动时,尖端与表面之间物理交互产生的高频振动会经过工具传导至皮肤并激活皮肤中的机械感受器<sup>[5]</sup>。大脑通过对振动信号的时空编码最终形成对纹理特征的触觉感知<sup>[6]</sup>。这种工具介导的高频振动反馈为用户提供了关于纹理粗糙度的直观感受<sup>[4]</sup>。基于以上理论,一些研究已经探索了通过力反馈或振动触觉反馈帮助用户感知虚拟纹理粗糙度的方法<sup>[7-10]</sup>。这些对虚拟纹理进行触觉建模与渲染的方法能够模拟用户与真实纹理交互的过程和感受,显著增强了人与虚拟物体交互的真实感<sup>[4]</sup>。

传统对图像纹理进行力触觉再现的研究已经使用了统计法、结构法、模型法等方法来提取纹理特征<sup>[11]</sup>,并在对特征进行建模后通过力触觉反馈来表达。Tian等<sup>[12]</sup>采用局部全变分模型分解出图像的结构和纹理特征,并对两种图像特征分别进行建模和表达。然而,这些纯粹基于图像的纹理特征提取方法受到图像的质量和成像条件的影响,且普遍具有较高的计算复杂度。

目前更受关注的纹理触觉建模方法有数据驱动模型<sup>[13]</sup>、随机模型<sup>[14]</sup>和基于几何的确定性模型<sup>[15]</sup>等。其中,数据驱动的方法由于能自然地保持物理交互的真实性以及增强纹理触觉再现的真实感而被广泛关注<sup>[13]</sup>。早期的数据驱动方法通过对预先测量数据的插值,向用户传达与真实纹理交互相类似的触觉感受。Kuchenbecker团队<sup>[16-17]</sup>先后使用线性预测编码、分段自回归模型等方法来建立纹理触觉模型,并在用户与虚拟纹理交互时驱动音圈致动器呈现逼真的纹理感受。进一步探究了工具介导的纹理交互方式所产生的多模态反馈信号的最佳带宽和抗噪声能力<sup>[18]</sup>,这一结果对设计用于虚拟纹理感知的力触觉交互装置具有重要指导意义。此外,Kuchenbecker团队<sup>[19]</sup>还开发了包含有100个样本的Penn Haptic纹理数据库(HaTT),用于帮助触觉领域的研究人员比较和验证他们的纹理建模和渲染方法。随后,还有一些新的数据驱动模型被开发出来用于纹理的触觉建模与渲染<sup>[20-21]</sup>。然而,由于需要为每种纹理建立单独的模型,上述方法难以对数据库之外未建模的新纹理进行触觉渲染,故泛化能力受到限制。

随着深度学习技术的发展,一些研究在建立的纹理

数据库基础上,已经将跨模态生成对抗网络(generative adversarial networks, GAN)<sup>[22]</sup>应用于数据驱动的纹理触觉建模之中。此类方法能学习不同纹理图像的特征以及它们与数据库中对应的触觉表示之间的关系,并依据此关系将数据库之外的新纹理转化成对应的触觉信号。Cai等<sup>[23]</sup>在HapTex织物数据库上训练了用于摩擦系数生成的条件生成对抗网络模型,该模型首先将纹理图像转化成对应的触觉频谱,再利用由Griffin-Lim算法转换得到的一维信号生成触觉反馈。还有Song等<sup>[24]</sup>也做了类似的研究。Li等<sup>[25]</sup>利用LMT数据库训练了一种端到端的GAN,其能消除Griffin-Lim算法在转化信号时带来的偏差。以上方法虽然解决了早期数据驱动方法泛化能力不强的问题,但仅在预定义和受约束的动作条件下对虚拟纹理进行了触觉渲染,并未考虑交互时用户的实时动作信息(如滑动速度和按压力)。已有的研究表明,工具介导的纹理交互所产生的振动信号,其幅度和频率不但受到工具尖端和纹理表面几何特性的影响,还与探测纹理表面的动作因素密切相关<sup>[5]</sup>。因此,不考虑用户动作信息的触觉渲染无法确保纹理再现的真实感。

最近,一些工作尝试同时结合纹理图像和用户的动作信息对纹理进行触觉建模。此类方法需要在包含有多种纹理图像以及对应的动作信息和振动加速度信息的纹理数据库(如HaTT)上训练一个联合模型,并学习数据库中不同纹理之间的关联关系。该联合模型所生成的触觉反馈不仅与纹理的特征相关,还受到用户探索纹理时所施加的力和速度的影响。同时,该模型可以将新纹理与数据库中的已有纹理相关联,从而对未建模的新纹理进行触觉渲染。因此,该类方法能为用户提供具有真实感的触觉反馈,并确保对新纹理具有的较好的泛化能力。Heravi等<sup>[26]</sup>设计了一种基于学习的动作条件模型,并在扩充后的HaTT数据库上进行训练。该模型以用户的实时动作信息和纹理的GelSight图像为输入,通过编码器预测下一时刻振动加速度的离散傅里叶变换,再利用快速Griffin-Lim算法获得一维的振动信号。然而,该模型在训练时需要使用传感器重新采集HaTT数据库中93种真实材料的GelSight图像来对数据库进行扩充,具有较大的工作量。此外,该模型仅能对GelSight图像进行触觉建模,而在实际交互中,虚拟纹理通常以RGB图像的形式呈现给用户,这使得模型的适用范围受到一定的限制。

为了帮助人们实现对常见RGB纹理图像具有真实感的触觉感知,本文提出一种融合纹理的RGB图像和用户实时动作信息的纹理触觉渲染模型。该模型在将不同时刻的动作信息(速度信息和法向力信息)降采样后,依次与经VGG16<sup>[27]</sup>提取后的纹理粗糙度特征串联。然后,利用改进的MelGAN<sup>[28]</sup>网络生成对应的振动加速度信

号。接着,本文在 HaTT 纹理数据集上训练了该模型,并通过波形相似度计算评估了模型的生成效果。此外,为了进一步验证模型的性能,本文利用一款具有动作信息检测与振动触觉表达功能的笔式装置采集了数据库之外的真实纹理表面的触觉信号,并在频域中计算了采集的信号与模型生成的信号在频域方面的均方根误差 (root mean square error, RMSE)。最后,本文使用笔式装置执行了两项用户实验。结果表明,本文提出的模型不但能对未建模的新纹理进行触觉渲染,还能够为用户带来更加逼真的纹理交互体验。

### 1 结合动作信息的纹理触觉渲染模型

本文的目标是建立多模态信息 (RGB 图像  $x$ 、扫描速度  $v$  和法向力  $f$ ) 与振动加速度信号  $a$  之间的映射关系,可表示为  $g(x, v, f) \rightarrow a$ 。受现有工作的启发,本文设计了一个基于动作信息的纹理触觉渲染模型。该模型以改进后的 MelGAN 网络为主体,包含生成器和鉴别器两部分。模型的整体结构如图 1 所示。

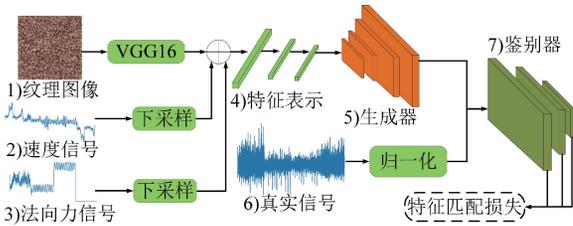


图 1 基于动作信息的纹理触觉渲染模型  
Fig. 1 Texture haptic rendering model based on action information

该模型以 3 个数据源 (纹理 RGB 图像、速度信号、法向力信号) 作为输入,从而得到完备的特征表示。在该模型的实现过程中,由于图像是二维数据,故先使用预训练的 VGG16 网络对图像进行特征提取,得到纹理图像的粗糙度表征向量  $M_i$ 。同时,对实时的动作信息 (一维速度信号及法向力信号) 通过降采样映射到低维的特征空间,得到实时的动作信息表征向量  $M_s$  和  $M_f$ 。将  $M_s$ 、 $M_f$  与  $M_i$  串联,得到总的特征向量  $M_{i,a}$ :

$$M_{i,a} = M_i \oplus M_s \oplus M_f \quad (1)$$

接着,通过两个全连接的神经网络进一步学习特征表示规则  $M_F$ 。 $M_F$  表示为:

$$M_F = W_1^T f(W_2^T M_{i,a} + b_2) + b_1 \quad (2)$$

式中:  $W_1$ 、 $W_2$  是具有相应偏差  $b_1$ 、 $b_2$  的投影矩阵;  $f(\cdot)$  为激活函数。

在得到  $M_F$  后,将其输入到改进的 MelGAN 网络的生成器部分,生成器的结构如图 2 所示。

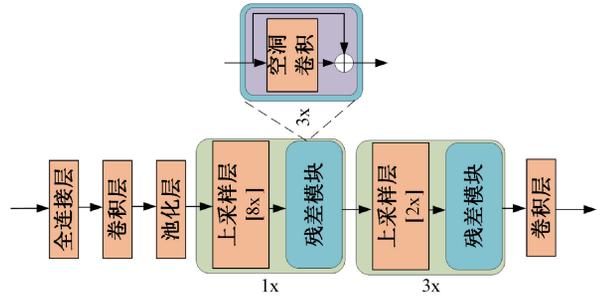


图 2 生成器结构

Fig. 2 Generator structure diagram

改进的 MelGAN 生成器由全连接层、卷积层、池化层、1 个 8 倍和 3 个两倍的上采样模块组成。其中,每个上采样模块中包括扩展率分别为 1、3、9 的空洞卷积,并使用 Leaky-ReLU 激活函数。池化层用于对特征图进行数据压缩,提取最显著的特征,以减少计算量、防止过拟合和提高模型泛化能力。 $M_F$  输入到生成器后,经全连接层、卷积层和池化层进一步提取最显著的特征,接着将特征输入上采样层,得到对应的触觉信号,并使用卷积层将其重塑为目标大小。

为了获得更多的输入信息,本文选用具有 3 个相同结构的鉴别模块 (D1、D2、D3) 作为微调后的 MelGAN 网络的鉴别器,每个鉴别模块均由卷积层、激活层和下采样层组成。鉴别器的结构如图 3 所示。其中 D1 按原始序列的比例工作,D2、D3 分别以原频率的 1/2、1/4 对原始序列进行下采样。将这种不同频率的信号送给 3 个鉴别模块,从不同频率获得信息,以提高网络的准确性。

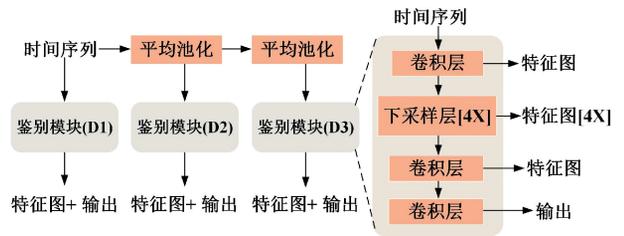


图 3 鉴别器结构

Fig. 3 Discriminator structure diagram

将生成的触觉信号和归一化后的真实信号送入鉴别器,计算二者之间的特征匹配损失,并据此不断优化模型。

## 2 模型训练与振动触觉信号生成

### 2.1 数据集

目前常用的纹理触觉数据库有 HaTT<sup>[19]</sup> 和 LMT<sup>[29]</sup> 两种。由于 LMT 数据库缺乏位置的追踪数据,无法满足本文的研究需求,故使用 HaTT 数据库作为本文中纹理触

觉渲染模型的测试和验证库。HaTT 数据库中包含有 10 种类别的 100 张不同的纹理图像,以及实验者手持工具以自然方式在每个纹理表面上移动 10 s 记录的法向力、速度和振动加速度等数据。为进一步优化模型的计算性能,本文采用基于滑动窗口的数据增强方式扩展数据集。如图 4 所示,假设信号的总数据量为  $L$ ,采样窗口的长度为  $l$ ,每次移动的距离为  $s$ ,则按照此类采样方式得到的样本总数  $n$  的表达式为:

$$n = (L - l) / s \quad (3)$$

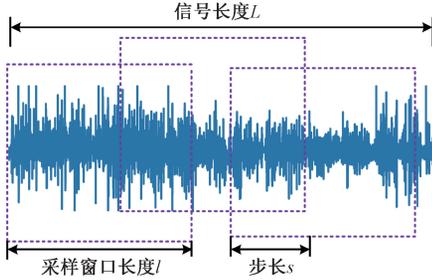


图 4 滑动窗口采样  
Fig. 4 Sliding window sampling

本文将 HaTT 数据库中的时间序列数据分割为 128 ms 的独立序列,并以 50% 的重合点数进行滑动采样,实现数据集的扩展。

## 2.2 模型训练过程

在完成数据集准备后,进行数据集的划分。从 HaTT 数据库中随机选择 10 张不同类别的纹理图像及其对应的动作信息作为测试集,其余作为训练集。在训练时,首先将 HaTT 库中的加速度信号归一化至  $[-1, 1]$ ,并将归一化后的加速度信号作为生成模块的生成对象。对得到的特征表示  $M_F$ , 对其进行一维卷积后送入到改进的 MelGAN 网络的生成器中,以便生成相应的振动加速度信号。在模型中,生成器和鉴别器的优化目标分别为:

$$\min_G \mathbb{E}_{s,z} \left[ \sum_{k=1,2,3} -D_k(G(t,z)) \right] \quad (4)$$

$$\min_{D_k} \mathbb{E}_x [\min(0, 1 - D_k(x))] +$$

$$\mathbb{E}_{t,z} [\min(0, 1 + D_k(G(t,z)))] , \forall k = 1, 2, 3 \quad (5)$$

式中:  $x$  代表原始的加速度序列;  $t$  代表条件信息 ( $M_F$ );  $z$  代表高斯噪声向量。

为了最小化真实加速度信号与生成加速度信号之间的 L1 距离,本文选用特征匹配函数 (Larsen) 来训练生成器。特征匹配函数的表达式为:

$$\mathcal{L}_{FM}(G, D_k) =$$

$$\mathbb{E} \mathcal{L}_{x,t-p_d} \left[ \sum_{i=1}^r \frac{1}{N_i} \| D_k^{(i)}(x) - D_k^{(i)}(G(t)) \|_1 \right] \quad (6)$$

式中:  $D_k^{(i)}$  代表第  $k$  个鉴别模块的第  $i$  层的特征图输出;  $N_i$  表示每层的单元数。

在此基础上,选择以下目标函数对生成器进行训练:

$$\min_G \left( \mathbb{E}_{s,z} \left[ \sum_{k=1,2,3} -D_k(G(t,z)) \right] \right) + \lambda \sum_{k=1}^3 \mathcal{L}_{FM}(G, D_k) \quad (7)$$

式中:  $\lambda$  为特征匹配损失的权重;  $G(t,z)$  为生成器生成的加速度信号;  $\mathbb{E}_{t,z}$  为输入  $t$  和高斯噪声向量  $z$  的数学期望;  $\mathbb{E}_{x,t-p_d}$  为真实的加速度信号和生成器输入的数学期望。

在训练时,选择的批次大小为 256,设置学习率为  $1 \times 10^{-4}$ ,训练次数为 1 000 次,并采用 Adam 优化器<sup>[30]</sup> 对网络模型进行优化。本文的模型是使用 PyTorch 框架在 Intel i5-13400F CPU 和 Nvidia RTX 3070 Ti GPU 上进行训练的。

## 2.3 触觉信号的生成结果

训练结束后,本文从测试集中随机选择 3 张纹理图像,并在对应的动作信息下生成振动加速度信号,结果如图 5 所示。

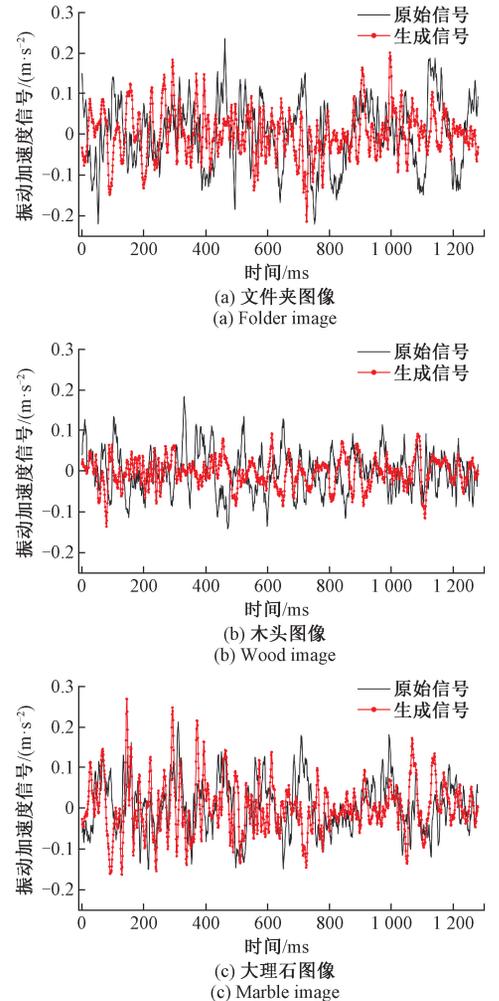


图 5 模型对测试图像的加速度信号生成结果  
Fig. 5 The acceleration signal generation result of the model on the test images

为了更好的评估模型的生成效果,本文计算了生成信号波形与原始信号波形之间的相似度。当需要比较两波形  $x_1(t)$  和  $x_2(t)$  之间的相似度时,其方法是选择一个参数  $c$ , 让  $c \cdot x_2(t)$  无限逼近  $x_1(t)$ , 接着使用误差能量分析来测量波形的相似度。误差能量可以在时域中用  $x_1(t) - c \cdot x_2(t)$  的积分来表示,即:

$$e = \int_{-\infty}^{+\infty} (x_1(t) - c \cdot x_2(t))^2 dt \quad (8)$$

对函数求导,当  $c = \frac{\int_{-\infty}^{+\infty} (x_1(t) \cdot x_2(t)) dt}{\int_{-\infty}^{+\infty} (x_2(t) \cdot x_2(t)) dt}$  时,能量误差最小。

设  $p_{x_1x_2}$  为  $x_1(t)$  与  $x_2(t)$  之间的相关系数,其平方与 1 之间的误差为误差能量。因此,  $p_{x_1x_2}$  的表达式为:

$$p_{x_1x_2} = \frac{\int_{-\infty}^{+\infty} (x_1(t) \cdot x_2(t)) dt}{\sqrt{\int_{-\infty}^{+\infty} (x_1(t))^2 dt} \cdot \sqrt{\int_{-\infty}^{+\infty} (x_2(t))^2 dt}} \quad (9)$$

由于 Griffin-Lim 算法及其变体形式被广泛应用于从频域到时域的触觉信号转化<sup>[23,26]</sup>, 本文将由 Griffin-Lim 算法得到的触觉信号作为比较对象。具体的, 本文将 CGAN<sup>[31]</sup> 作为对比网络。CGAN 网络以 HaTT 数据集中的纹理图像作为网络的输入, 以加速度信号的时频图像作为输出, 然后利用 Griffin-Lim 算法将时频图像恢复成时域的触觉信号。经计算, 对于从测试集中随机选取的 3 张纹理图像, 本文提出的渲染模型所生成的振动加速度与原始的振动加速度之间的平均相似度为 0.763, 而 Griffin-Lim 算法生成的振动加速度与原始振动加速度之间的平均相似度仅为 0.148。由此可见, 本文的模型生成的振动触觉信号具有更高的准确性。

### 3 纹理触觉渲染的有效性评估实验

本文通过客观的数据评估实验和主观的用户感知实验来检验本文提出的触觉渲染模型在纹理再现方面的真实感和有效性。

#### 3.1 数据评估实验

该实验旨在利用 RMSE 在频域中调查本文提出的模型以及其他 3 种触觉渲染方法 (Griffin-Lim 算法<sup>[23]</sup>、自回归模型 (AR)<sup>[17]</sup>、波形段表 (WST)<sup>[21]</sup>) 生成的触觉信号序列与真实触觉信号序列之间的差异。

由图 1 可知, 本文提出的模型需要实时采集工具与真实纹理交互时的动作信息和加速度信号。因此, 本文制作了一款 3D 打印的具有动作信息 (速度、法向力) 采集和振动触觉表达功能的笔式装置, 如图 6 所示, 其与本文提出的模型共同构成纹理触觉再现系统。

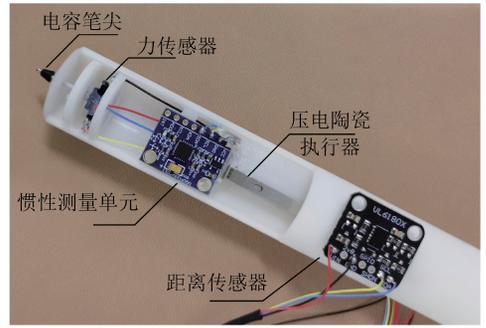


图 6 定制的笔式装置  
Fig. 6 Customized pen-type device

该笔式装置的内部集成有惯性测量单元 (MPU6050, TDK InvenSense, USA)、力传感器 (FSS1500NGT, Honeywell International Inc., USA)、激光测距传感器 (VL53L0X, STMicroelectronics, Italy) 和 Samsung 公司生产的压电陶瓷执行器。其中, 惯性测量单元位于笔的内部, 其  $y$  轴与笔身保持在同一方向。惯性测量单元通常测量 3 个轴 (即  $x$  轴、 $y$  轴、 $z$  轴) 的加速度信号。然而, 由于人手无法识别振动方向<sup>[32]</sup>, 故可以将检测的三轴信号通过 DFT321 算法<sup>[17]</sup> 转化成单轴信号, 并将此信号用于触觉建模和驱动压电陶瓷执行器产生一维振感。力传感器位于笔的下端, 垂直于笔尖放置, 用于测量实时的法向力数据。激光测距传感器位于笔的外表面, 其通过测量笔与固定物体的相对距离间接测量运动时的速度。收集到的力和速度信号以 25 Hz 的截止频率进行低通滤波, 加速度信号则以 10 Hz 的截止频率进行高通滤波。在此基础上, 选择 6 种生活中常见的真实纹理样本, 如图 7 所示, 包括海绵 T1、铜片 T2、棉布 T3、人造皮革 T4、塑料 T5, 以及纸盒 T6。这 6 个真实纹理样本均未出现在 HaTT 数据库中, 且每个样本均不重叠地固定在亚克力板上。



图 7 六种真实纹理样本

Fig. 7 Six kinds of real texture samples

对每个真实纹理样本, 本文收集了笔式装置在其上运动 10 s 产生的速度、按压力和振动加速度等原始信号数据。然后将收集到的相相关数据分割成长度为

1 280 的独立序列,分段后输入到纹理触觉渲染模型中。模型生成的加速度信号与原始的加速度信号在时域和频域方面的对比结果如图 8 所示。可以看出,生成信号的频谱对原始信号的频谱具有较好的拟合度。

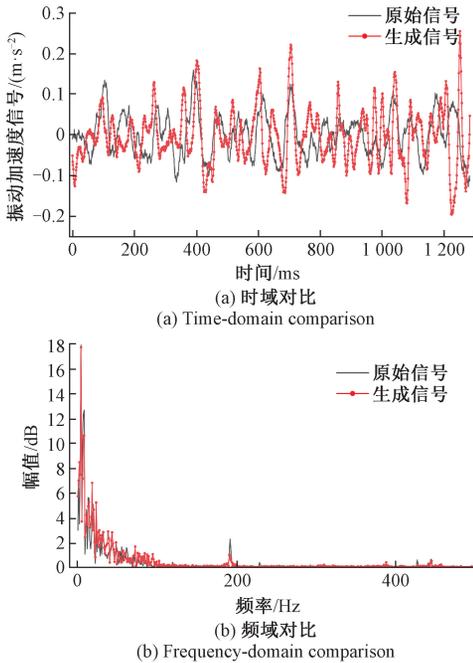


图 8 生成的和原始的加速度信号对比

Fig. 8 The comparison of original and generated acceleration signals

鉴于信号频谱随输入参数变化,本文采用文献[13]提出的方法,使用 500 个序列和 50% 重叠的滑动窗口比较原始信号与生成信号,并在频域中使用 RMSE 进行评估。表 1 为 4 种方法的 RMSE 平均值。

表 1 本文的模型与其他方法的 RMSE 比较

Table 1 The comparison of RMSE between the model of this paper and other methods

方法	RMSE 平均值
Griffin-Lim <sup>[23]</sup>	0.587
自回归模型 <sup>[17]</sup>	0.348
波形段表 <sup>[21]</sup>	0.198
本文	0.173

从表 1 可以看出,本方法的 RMSE 最低,因而具有最好的信号生成效果,可以在用户与虚拟纹理交互时提供类似于与真实纹理交互的体验。

### 3.2 用户感知实验

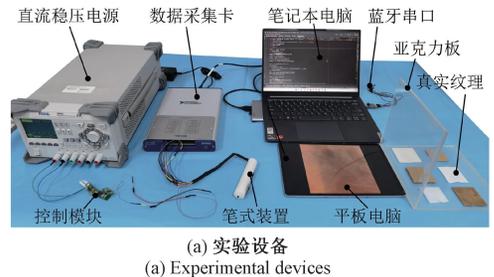
该实验旨在通过受试者的主观感受来评估 4 种触觉渲染方法所呈现的虚拟纹理的真实感及表现差异。

### 1) 实验装置与人员

实验设备如图 9(a) 所示,包括平板电脑、笔记本电脑、直流稳压电源(DP831, RIGOL Technologies, CHN)、数据采集卡(USB-6341, National Instruments Inc., USA)、亚克力板、笔式装置,以及 6 种真实纹理。15 名(年龄为  $23.15 \pm 2.41$  岁,  $\text{mean} \pm \text{SD}$ ) 视力正常且右手为优势手的受试者参加了该实验。没有受试者报告他们的触觉感知能力存在任何缺陷,且他们不了解研究的目的。所有受试者均知情同意,并为他们的参与付费。在正式实验之前,通过训练示例让所有受试者熟悉笔的使用方法以及实验的流程。

### 2) 实验过程

实验开始时,用于固定真实纹理的亚克力板水平放置在水平桌面上,平板电脑位于真实纹理的左侧。受试者在与虚拟纹理交互时需要保持笔竖直,右手在握持笔时不能触碰笔外表面的距离传感器。距离传感器需要正对垂直于桌面放置的障碍物(亚克力板)。受试者在实验过程中需全程佩戴眼罩和播放白噪声的耳机,并仅靠触觉去感受两个表面之间的纹理相似度。实验场景如图 9(b) 所示。



(a) 实验设备  
(a) Experimental devices



(b) 实验场景  
(b) Experimental scenario

图 9 实验环境

Fig. 9 Experimental apparatus

实验分为两部分。第 1 部分,参考文献<sup>[33]</sup>的方法,本文向受试者提供成对的真实纹理样本,受试者被要求评价它们的相似性。这有助于避免极端评分,同时帮助受试者建立感知标准和相似度评分尺度。实验中,本文将图 7 中的 6 个真实纹理组合成 12 对,其中 6 对相同,6 对不同。评分范围为 0~9 分,0 分表示纹理对的触感完全不同,9 分则表示完全相同。第 2 部分,受试者需要评

估虚拟纹理与其对应的真实纹理之间的相似性。该实验将图7中的6个真实纹理拍摄成图像,并分别使用表1的4种方法对它们进行触觉渲染。

在3种已有方法中,自回归模型和波形段表缺乏泛化能力。为了与这两种方法进行比较,本文采集了6个真实纹理表面的触觉信息,然后为每个纹理建立了在某一渲染方法下的独立模型。在交互时,本文根据触觉渲染方法和虚拟纹理编号选择对应的模型进行振动触觉表达。对于Griffin-Lim算法,本文使用在HaTT数据库上训练好的CGAN模型生成新虚拟纹理对应的振动信号的时频图像,再利用Griffin-Lim算法将其恢复成时域信号,用于振动触觉表达。实验过程中,6个真实纹理以随机的顺序呈现给受试者。受试者首先持笔与真实纹理交互,并仅通过笔尖与真实纹理接触产生的高频振动去感知其表面的触觉属性,交互时间为10s。接着,受试者在实验者的帮助下将笔移动至触摸屏上的虚拟纹理。当受试者与虚拟纹理交互时,纹理触觉再现系统会实时采集受试者的动作信息,并随机选择4种触觉渲染方法中的一种来生成触觉反馈。受试者通过压电陶瓷执行器感知每种触觉渲染方法生成的振动触觉反馈,交互时间为10s。受试者可反复对比由触觉渲染方法生成的触觉反馈与真实纹理交互产生的触觉反馈之间的感觉差异,并在确认完成评价后进行感知相似性的主观评分。评分范围与第1部分相同,受试者完成一次评分后可以休息几分钟。当受试者完成对4种触觉渲染方法的评分后,对该纹理的用户感知实验结束。然后,受试者继续对下一个真实纹理样本进行感知和评分。当所有受试者都完成了对6个真实纹理在四种触觉渲染条件下的评分后,整个实验结束。

### 3) 实验结果

在用户实验的第1部分,本文收集了180组数据(15个受试者 $\times$ (6+6)对)。在此基础上,本文得到了相同真实纹理与不同真实纹理之间的相似度分数,统计结果如图10所示。从图10可知,铜片-铜片的相似度评分

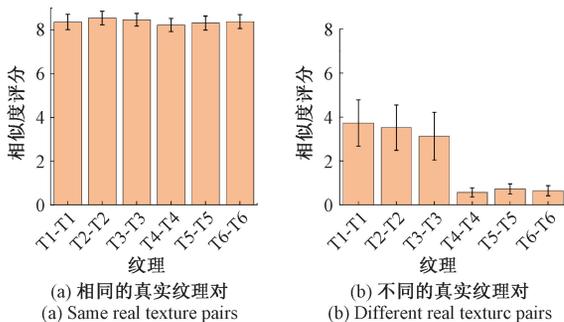


图10 真实纹理间的平均相似度评分

Fig. 10 Average pairwise similarity scores between real textures

最高(8.55),铜片-棉布之间的相似度评分最低(0.57)。故受试者打分的上界为8.55,下界为0.57。通过该实验,受试者的评分在0.57~8.55的范围内是合理的。

在用户实验的第2部分,共收集了360组数据(15个受试者 $\times$ 6种纹理 $\times$ 4种渲染方法)。去掉15个受试者评分中的最高分和最低分后,计算得到的平均值即为每个纹理获得的评分。由此,得到4种渲染条件对每种纹理主观评分的均值,统计结果如图11所示。

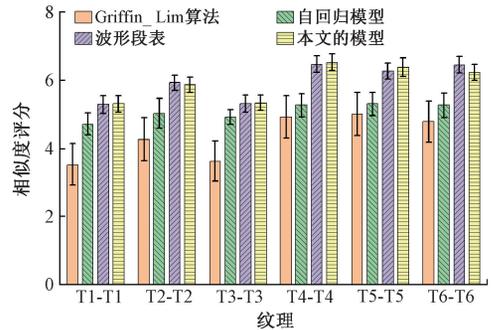


图11 平均相似度评分

Fig. 11 Average similarity scores

从图11可知,本文的模型稍优于现有的方法,只在部分情况下略低于WST方法得到的结果。但相比于WST方法,本文的模型具有一定的泛化能力,无需对数据库之外的新纹理进行单独建模,因而更具通用性。然后,对相似度评分进行关于纹理种类和渲染方法的双因素方差分析。结果表明,纹理种类( $F(5,336) = 5.27, p < 0.001$ )和渲染方法( $F(3,336) = 6.01, p < 0.001$ )均对相似度评分有高度显著影响。

## 4 讨 论

本文通过数据评估实验和用户感知实验检验了所提模型对数据库之外的新虚拟纹理进行触觉渲染的有效性和真实感。在数据评估实验中,本文利用笔式装置采集了用户分别与6种真实纹理交互时的动作信息及对应的振动加速度信号,并在时域和频域中比较了提出的模型生成的加速度信号与原始的加速度信号在数值方面的差异。

用户感知实验是检验本文提出的模型进行虚拟纹理触觉渲染有效性的核心内容,即直接通过用户的主观感知去评价触觉渲染的效果。在用户实验的第2部分中,本文模型所得到的平均相似度评分为6.01,远远高于用户评分0.57的区间下限。

1) 当用户手持笔与虚拟纹理交互时,本文的模型能给受试者带来更加逼真的触觉感受,尽管在部分场

景下略低于现有方法,但本文的模型具有一定的泛化能力,即使对于未建模的新纹理,也能提供较好的触觉渲染效果。

2)在将本文提出的模型用于虚拟纹理渲染时,所得到的相似度评分与用户评分的区间上限仍然有较大差异,这说明本文的模型仍然具有较大的改进空间。同时,在实验中,一些受试者反映笔产生的振感弱于在真实纹理表面上滑动时的振感,这可能是由于压电陶瓷执行器安装在笔的内部导致的。

由图 11 可知,受试者在本文的模型条件下对海绵(5.21)和棉布(5.32)的评分较低,这可能受到了两种材料的柔软性的影响,因为其他 4 种材料的表面均较硬。当笔尖(刚性材料)在柔软的真实纹理表面滑动时,用户施加的按压力太大可能使笔尖突然被卡住。而当用户与虚拟纹理交互时,则不存在这个问题。这种突然被卡住的情况的缺失会给用户带来不真实的触觉感受。

## 5 结 论

为了进一步提升纹理触觉渲染的适用范围、泛化能力和交互真实感,本文以纹理的 RGB 图像作为微调后的 MelGAN 网络的输入之一,建立了融合纹理图像和用户动作信息(速度信号和法向力信号)的纹理触觉渲染模型,以便在交互过程中高保真地生成随动作条件变化的振动触觉信号。相比于 GelSight 图像,本文的模型拥有对常见的纹理图像更好的泛化能力。此外,本文还设计了一款具有动作信息(速度、法向力)实时采集和振动触觉表达功能的笔式装置。在采集了数据库之外的真实纹理表面的触觉信号后,本文通过比较提出的模型与其他 3 种已有方法在信号生成结果方面的差异,检验了模型对未建模的纹理表面进行触觉渲染的能力。最后,本文使用提出的纹理触觉再现系统进行了两项用户实验。在用户实验中,本文首先确定了受试者对纹理的相似度进行评分的合理区间,然后比较了 4 种触觉渲染算法为受试者提供的纹理交互的真实感。平均 6.01 的主观相似度打分结果说明本文提出的模型在保证对新纹理具有一定泛化能力的同时,还能够帮助受试者准确地感知虚拟纹理。

本文的模型虽然在泛化能力和交互真实感方面有了一定的提升,但仍可从如下两方面进行改进。首先,本文将时域信号之间的误差作为损失函数的优化指标,这样

会忽略掉对触觉信号生成产生很大影响的频域特征。其次,本文的模型是在 HaTT 数据库上训练得到的,仅针对正常交互条件下的触觉信号生成,难以在一些边界条件(如施加的压力较小、移动缓慢或没有移动等)下进行触觉信号预测。后续研究可以融合触觉信号的时域和频域特征,继续对模型进行优化,并研究边界条件下的纹理触觉再现问题。

## 参考文献

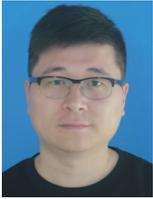
- [ 1 ] ECK J, KAAS A L, GOEBEL R. Crossmodal interactions of haptic and visual texture information in early sensory cortex [J]. *Neuroimage*, 2013, 75: 123-135.
- [ 2 ] LEDERMAN S J, KLATZKY R L. Haptic perception: A tutorial [J]. *Attention, Perception, & Psychophysics*, 2009, 71(7): 1439-1459.
- [ 3 ] OKAMOTO S, NAGANO H, YAMADA Y. Psycho-physical dimensions of tactile perception of textures [J]. *IEEE Transactions on Haptics*, 2012, 6(1): 81-93.
- [ 4 ] CULBERTSON H, KUCHENBECKER K J. Importance of matching physical friction, hardness, and texture in creating realistic haptic virtual surfaces [J]. *IEEE Transactions on Haptics*, 2016, 10(1): 63-74.
- [ 5 ] JOHANSSON R S, VALLBO Å B. Tactile sensory coding in the glabrous skin of the human hand [J]. *Trends in Neurosciences*, 1983, 6: 27-32.
- [ 6 ] JOHANSSON R S, FLANAGAN J R. Coding and use of tactile signals from the fingertips in object manipulation tasks [J]. *Nature Reviews Neuroscience*, 2009, 10(5): 345-359.
- [ 7 ] 李佳璐, 宋爱国, 吴涓, 等. 基于 SFS 技术的纹理力触觉再现方法研究 [J]. *仪器仪表学报*, 2010, 31(4): 812-817.  
LI J L, SONG A G, WU J, et al. Research on the method of haptic display based on SFS [J]. *Chinese Journal of Scientific Instrument*, 2010, 31(4): 812-817.
- [ 8 ] LU S, ZHENG M, FONTAINE M C, et al. Preference-driven texture modeling through interactive generation and search [J]. *IEEE Transactions on Haptics*, 2022, 15(3): 508-520.
- [ 9 ] ALMA U A, ROSENKRANZ R, ALTINSOY M E. Perceptual Substitution based haptic texture rendering for narrow-band reproduction [J]. *IEEE Transactions on*

- Haptics, 2023, 16(2): 204-214.
- [10] 邹垂国, 宋爱国, 吴涓. 基于 DELTA 手控器的纹理的力触觉表达方法[J]. 仪器仪表学报, 2008, 27(11): 2310-2314.
- ZOU CH G, SONG A G, WU J. Haptic rendering method for texture image based on DELTA manipulator[J]. Chinese Journal of Scientific Instrument, 2008, 27(11): 2310-2314.
- [11] 刘丽, 赵凌君, 郭承玉, 等. 图像纹理分类方法研究进展和展望[J]. 自动化学报, 2018, 44(4): 584-607.
- LIU L, ZHAO L J, GUO CH Y, et al. Texture classification: State-of-the-art methods and prospects [J]. Acta Automatica Sinica, 2018, 44(4): 584-607.
- [12] TIAN L, SONG A, CHEN D. Image-based haptic display via a novel pen-shaped haptic device on touch screens [J]. Multimedia Tools and Applications, 2017, 76: 14969-14992.
- [13] SHIN S, OSGOUEI R H, KIM K D, et al. Data-driven modeling of isotropic haptic textures using frequency-decomposed neural networks [C]. IEEE World Haptics Conference (WHC), 2015: 131-138.
- [14] FRITZ J P, BARNER K E. Stochastic models for haptic texture [C]. Tele-Manipulator and Telepresence Technologies, 1996: 34-44.
- [15] SHIN S, CHOI S. Hybrid framework for haptic texture modeling and rendering [J]. IEEE Access, 2020, 8: 149825-149840.
- [16] ROMANO J M, KUCHENBECKER K J. Creating realistic virtual textures from contact acceleration data[J]. IEEE Transactions on Haptics, 2011, 5(2): 109-119.
- [17] CULBERTSON H, UNWIN J, KUCHENBECKER K J. Modeling and rendering realistic textures from unconstrained tool-surface interactions [J]. IEEE Transactions on Haptics, 2014, 7(3): 381-393.
- [18] KHOJASTEH B, SHAO Y, KUCHENBECKER K J. Robust surface recognition with the maximum mean discrepancy: Degrading haptic-auditory signals through bandwidth and noise [J]. IEEE Transactions on Haptics, 2024, 17(1): 58-65.
- [19] CULBERTSON H, DELGADO J J L, KUCHENBECKER K J. One hundred data-driven haptic texture models and open-source methods for rendering on 3D objects [C]. IEEE Haptics Symposium (HAPTICS), 2014: 319-325.
- [20] JOOLEE J B, JEON S. Data-driven haptic texture modeling and rendering based on deep spatio-temporal networks [J]. IEEE Transactions on Haptics, 2021, 15(1): 62-67.
- [21] NAI W, LIU J, SUN C, et al. Vibrotactile feedback rendering of patterned textures using a waveform segment table method [J]. IEEE Transactions on Haptics, 2021, 14(4): 849-861.
- [22] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Communications of the ACM, 2020, 63(11): 139-144.
- [23] CAI S, ZHAO L, BAN Y, et al. GAN-based image-to-friction generation for tactile simulation of fabric material[J]. Computers & Graphics, 2022, 102: 460-473.
- [24] SONG R, SUN X, LIU G. Cross-modal generation of tactile friction coefficient from audio and visual measurements by transformer [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.
- [25] LI Y, ZHAO H, LIU H, et al. Research on visual-tactile cross-modality based on generative adversarial network[J]. Cognitive Computation and Systems, 2021, 3(2): 131-141.
- [26] HERAVI N, CULBERTSON H, OKAMURA A M, et al. Development and evaluation of a learning-based model for real-time haptic texture rendering [J]. Arxiv Preprint Arxiv: 2212.13332, 2023.
- [27] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Arxiv Preprint Arxiv: 1409.1556, 2014.
- [28] KUMAR K, KUMAR R, DE BOISSIERE T, et al. Melgan: Generative adversarial networks for conditional waveform synthesis [C]. 33rd Conference on Neural Information Processing Systems (NeurIPS), 2019: 14910-14921.
- [29] STRESE M, BRUDERMUELLER L, KIRSCH J, et al. Haptic material analysis and classification inspired by human exploratory procedures [J]. IEEE Transactions on Haptics, 2019, 13(2): 404-424.
- [30] KINGMA D P, BA J. Adam: A method for stochastic

optimization [J]. Arxiv Preprint Arxiv: 1412.6980, 2014.

- [31] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. Arxiv Preprint Arxiv: 1411.1784, 2014.
- [32] BELL J, BOLANOWSKI S, HOLMES M H. The structure and function of Pacinian corpuscles: A review [J]. *Progress in Neurobiology*, 1994, 42(1): 79-128.
- [33] OSGOUEI R H, KIM J R, CHOI S. Data-driven texture modeling and rendering on electrovibration display [J]. *IEEE Transactions on Haptics*, 2019, 13(2): 298-311.

### 作者简介



**陈大鹏**, 2011 年于安徽理工大学获得学士学位, 2019 年于东南大学获得博士学位, 现为南京信息工程大学副教授、硕士生导师, 主要研究方向为力触觉再现、人工智能、人机交互。

E-mail: dpchen@nuist.edu.cn

**Chen Dapeng** received his B.Sc. degree from Anhui University of Science and Technology in 2011, and Ph. D. degree from Southeast University in 2019. He is currently an associate professor and a M.Sc. supervisor at Nanjing University of Information Science & Technology. His main research interests include haptic display, artificial intelligence, and human-computer interaction.



**刘佳**(通信作者), 2008 年于东南大学获得博士学位, 现为南京信息工程大学自动化学院教授、博士生导师, 主要研究方向为力触觉再现、虚拟/增强现实、人机交互。

E-mail: liujia@nuist.edu.cn

**Liu Jia** (Corresponding author) received her Ph. D. degree from Southeast University in 2008. She is currently a professor and a Ph. D. supervisor at Nanjing University of Information Science & Technology. Her main research interests include haptic display, virtual/augmented reality, and human-computer interaction.