UNIVERSITY OF TRENTO - Italy

i NEST

Binding Touch to Vision: A Step Toward Immersive Healthcare

Antonio Luigi Stefani, Niccolò Bisagno, Francesco De Natale, Nicola Conci

University of Trento - Department of Information Engineering and Computer Science

Context: VR for rehabilitation

Who needs this?

Patients Doctors Caregivers

To truly digitalize rehabilitation scenarios, we need to integrate **vision** with **touch**

General problems

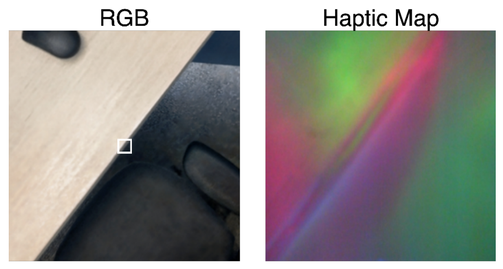
Poor quality 3D environments

No standard data representation

Few datasets that align vision & touch

Haptic data representation

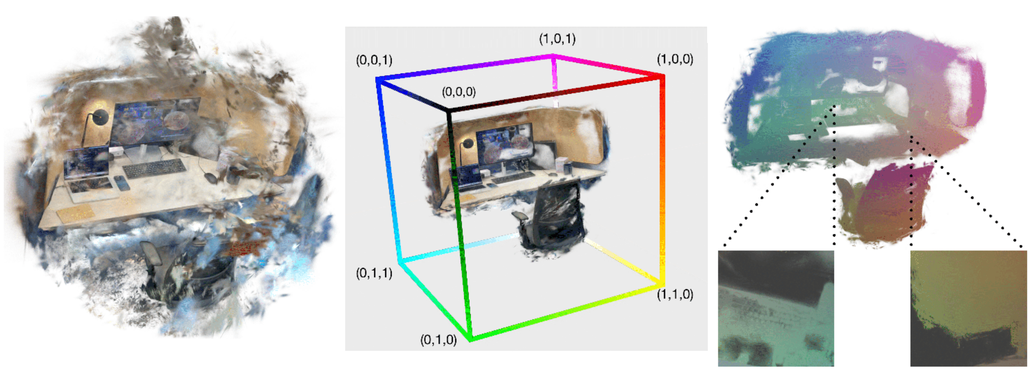
A widely used haptic data representation is the haptic map, provided by sensors like GelSight



We can use Diffusion Models to generate missing data. But...

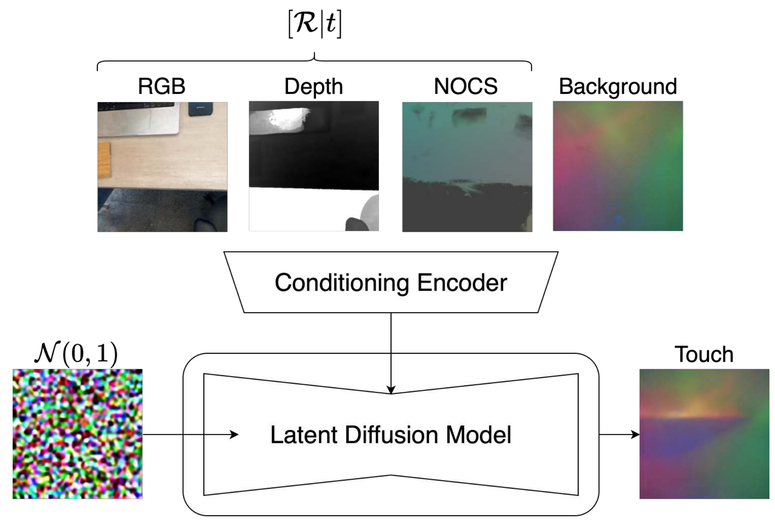
- **Technical issue**: **DMs are not 3D-aware**, resulting in poor understanding of 3D scenes and thus limiting generation capabilities
- **Research question**: 3D data are too heavy to handle. **How can we make DMs 3D-aware without using 3D data?**
- **Solution**: Employ **NOCS maps to embed 3D information onto images**

NOCS (NOrmalized Coordinate Space)

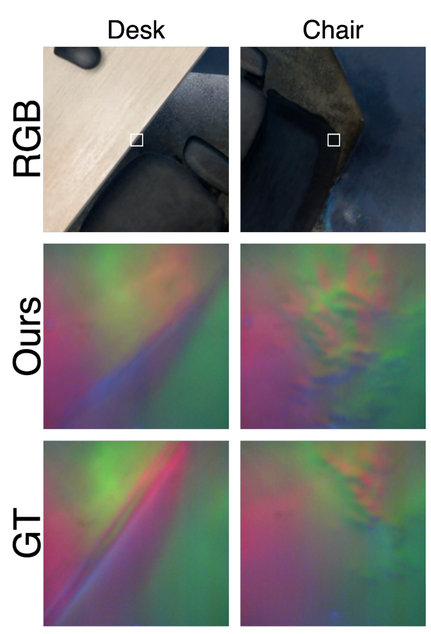


NOCS enables the estimation of size and position of objects in RGB images, providing a uniform, viewpoint-independent representation

Architecture



Results



Model	PSNR ↑	SSIM ↑	FID ↓
VisGel	24.34	0.82	97.05
TaRF	22.84	0.72	28.97
TaRF (our scene)	23.88	0.76	15.20
Ours	**30.19**	**0.84**	**10.06**

Real world scenarios