UV-Based 3D Hand-Object Reconstruction with Grasp Optimization

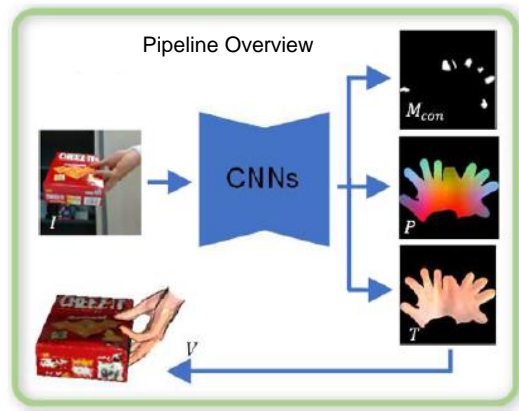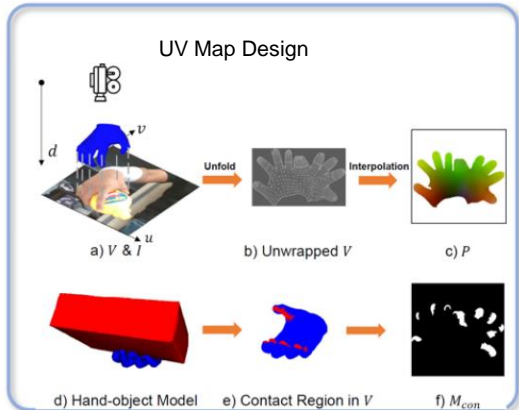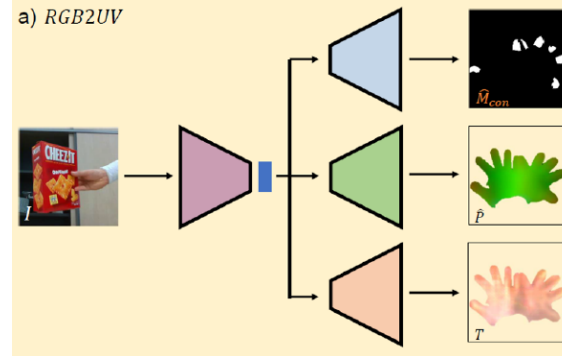Ziwei Yu and Angela Yao

National University of Singapore

Observation:

1. MANO model is convenient to use but its accuracy is limited because it cannot represent direct correspondences between the RGB input and the hand surface.
2. UV maps are idea representations as they establish dense correspondences between 3D surfaces and 2D images.
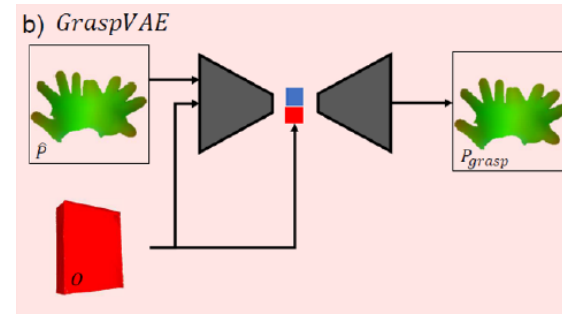
UV Map Design

a) V & I b) Unwrapped V c) P d) Hand-object Model e) Contact Region in V f) M_{con}

Pipeline Overview

Framework

a) RGB2UV

UV Estimation from an RGB image:
$$L_{RGB2UV} = L_P + L_{grad} + L_{con} + L_V + L_{texture}$$

b) GraspVAE

Encoder   Decoder   ■ Conditioning Variable   ■ Latent Variable

Grasp Optimization:
$$L_{Grasp} = L_P + L_{grad} + L_{con} + L_V + L_{pene}$$

Results

Dataset	FreiHand		Ho3D V2					Ho3D V3					DexYCB				
Method	MPJPE	MPVPE	MPJPE	MPVPE	PD	SIV	SD	MPJPE	MPVPE	PD	SIV	SD	MPJPE	MPVPE	PD	SIV	SD
Hampali[17]	-	-	1.07	1.06	12.34	17.28	4.02	-	-	-	-	-	-	-	-	-	-
Hasson[19]	1.33	1.33	1.10	1.12	-	-	-	-	-	-	-	-	-	-	-	-	-
Hasson[20]	1.33	1.33	1.14	1.09	18.44	14.35	4.10	-	-	-	-	-	-	-	-	-	-
GraspField [25]	-	-	1.38	1.36	14.61	14.92	**3.29**	-	-	-	-	-	-	-	-	-	-
Li [31]	-	-	1.13	1.10	11.32	14.67	3.94	1.08	1.09	16.78	10.87	3.90	1.28	1.33	9.47	11.02	3.64
Chen [6]	0.72	0.74	**0.99**	**1.01**	10.25	15.38	4.05	1.25	1.24	17.87	12.56	4.02	1.23	1.13	8.64	12.38	3.36
Dataset GT	-	-	-	-	-	-	-	-	-	-	-	-	0	0	4.58	7.16	1.48
MANO CNN	1.10	1.09	1.30	1.30	17.42	14.2	4.33	1.38	1.37	18.65	14.78	4.36	1.39	1.40	16.24	15.38	4.08
MANO Fit	1.37	1.37	1.58	1.61	21.37	18.0	4.82	1.66	1.65	22.04	15.60	4.65	1.49	1.43	18.56	17.60	4.86
RGB2UV	0.75	0.78	1.08	1.07	11.33	17.24	4.24	1.22	1.23	16.72	13.50	4.27	1.20	1.19	7.03	11.02	3.28
Hand Only	**0.71**	**0.73**	1.04	1.04	9.68	14.1	3.92	**1.08**	**1.04**	13.07	11.77	3.88	**1.09**	**1.02**	6.44	9.32	2.98
Hand+Object	-	-	1.25	1.33	**7.66**	**10.4**	**3.22**	1.28	1.26	**9.67**	**8.66**	**3.01**	1.29	1.27	**5.05**	**7.11**	**2.64**
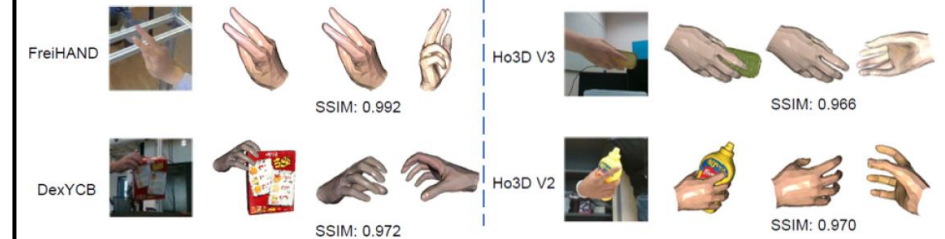
Surface Results

Interaction Refinement Results

Conclusion & Limitation:

- Explored UV coordinate maps for hand-object surface modelling and designed the first dense representation to model contact regions
- Introduce grasp optimization to improve the feasibility of the hand UV coordinate maps.

The 33rd British Machine Vision Conference
21st - 24th November 2022, London, UK