



M^3VSNet: Unsupervised Multi-metric Multi-view Stereo Network

Baichuan Huang^[1,3], Hongwei Yi^[2], Can Huang^[3],Yijia He^[3], Jingbin Liu^[1], Xiao Liu^[3]

1. Wuhan University 2. Peking University 3. Megvii Technology Limited

huangbaichuan@whu.edu.cn https://whubaichuan.github.io/

Introduction



M^3VSnet

Depth estimation

Depth fusion





- 1. Determine the final depth map based on the constraints of photometric consistency.
- 2. Use geometric consistency to measure whether the depth between multiple views is consistent or not.

Innovation



A novel unsupervised multi-metric MVS network for dense point cloud reconstruction without any supervision. In non-ideal environments, the same color can't be guaranteed in multi-view images. We extract semantic feature from VGG network. We combine the pixellevel and feature level unsupervised loss function.

Code is available: https://github.com/whubaichuan/M3VSNet

Result

DTU dataset

Method	Mean Distance(mm)							
	Acc.	Comp.	overall					
Furu	0.612	0.939	0.775					
Tola	0.343	1.190	0.766					
Colmap	0.400	0.664	0.532					
SurfaceNet	0.450	1.043	0.746					
MVSNet(D=192)	0.444	0.741	0.592					
Unsup_MVS	0.881	1.073	0.977					
MVS2	0.760	0.515	0.637					
M3VSNet(D=192)	0.636	0.531	0.583					



lower is better

more efficient

Tanks & Temples dataset

	Method	Mean	Family	Francis	Horse	Light	M60	Panther	Play-	Train
						house			around	
	M3VSNet	37.67	47.74	24.38	18.74	44.42	43.45	44.95	47.39	30.31
	MVS2	37.21	47.74	21.55	19.50	44.54	44.86	46.32	43.48	29.72
	SurfaceNet	26.30	36.97	11.01	7.25	37.55	26.94	26.18	42.98	21.53







Thanks for your attention!

Baichuan Huang <u>huangbaichuan@whu.edu.cn</u> https://whubaichuan.github.io/ Wuhan University, China