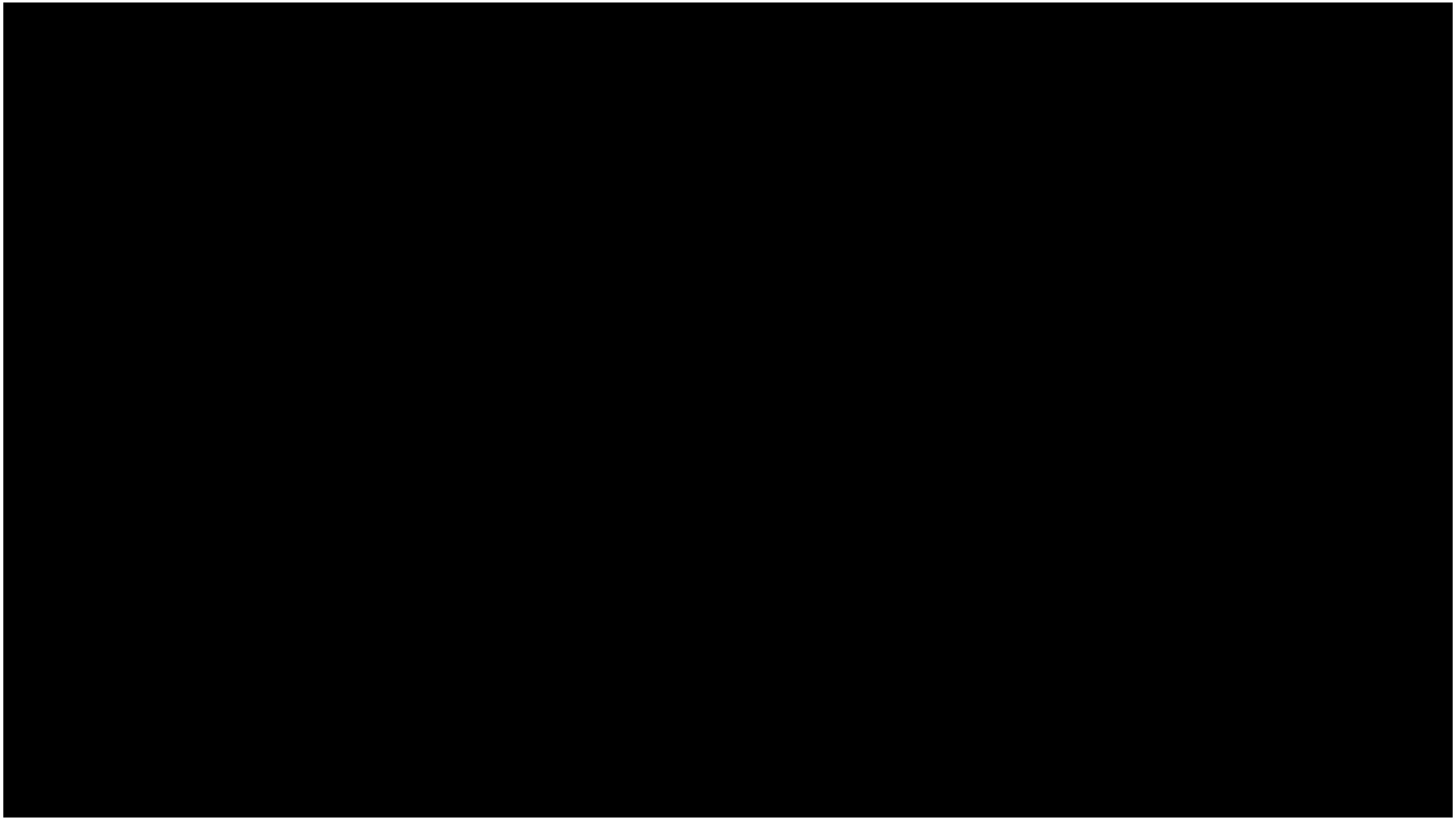
ISMAR 2019 Tutorial: Bridging the gaps between research and engineering in AR
https://ar.163.com/ismar2019_tutorial

The NEAR VI Dataset and AR oriented metrics

Cheng Wang
wangcheng@ezxr.com
Hangzhou EasyXR Co., Ltd. China
October 18, 2019

Who we are?

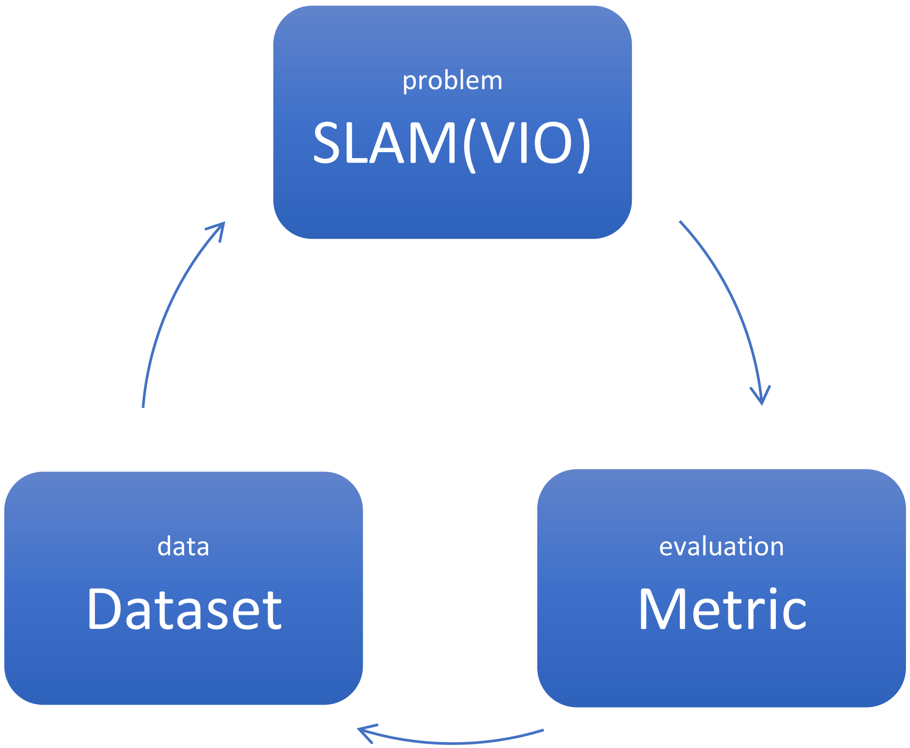
AR Tech Startup: Tracking, rendering, and AR content distribution.
https://ar.163.com/dongjian



Contents

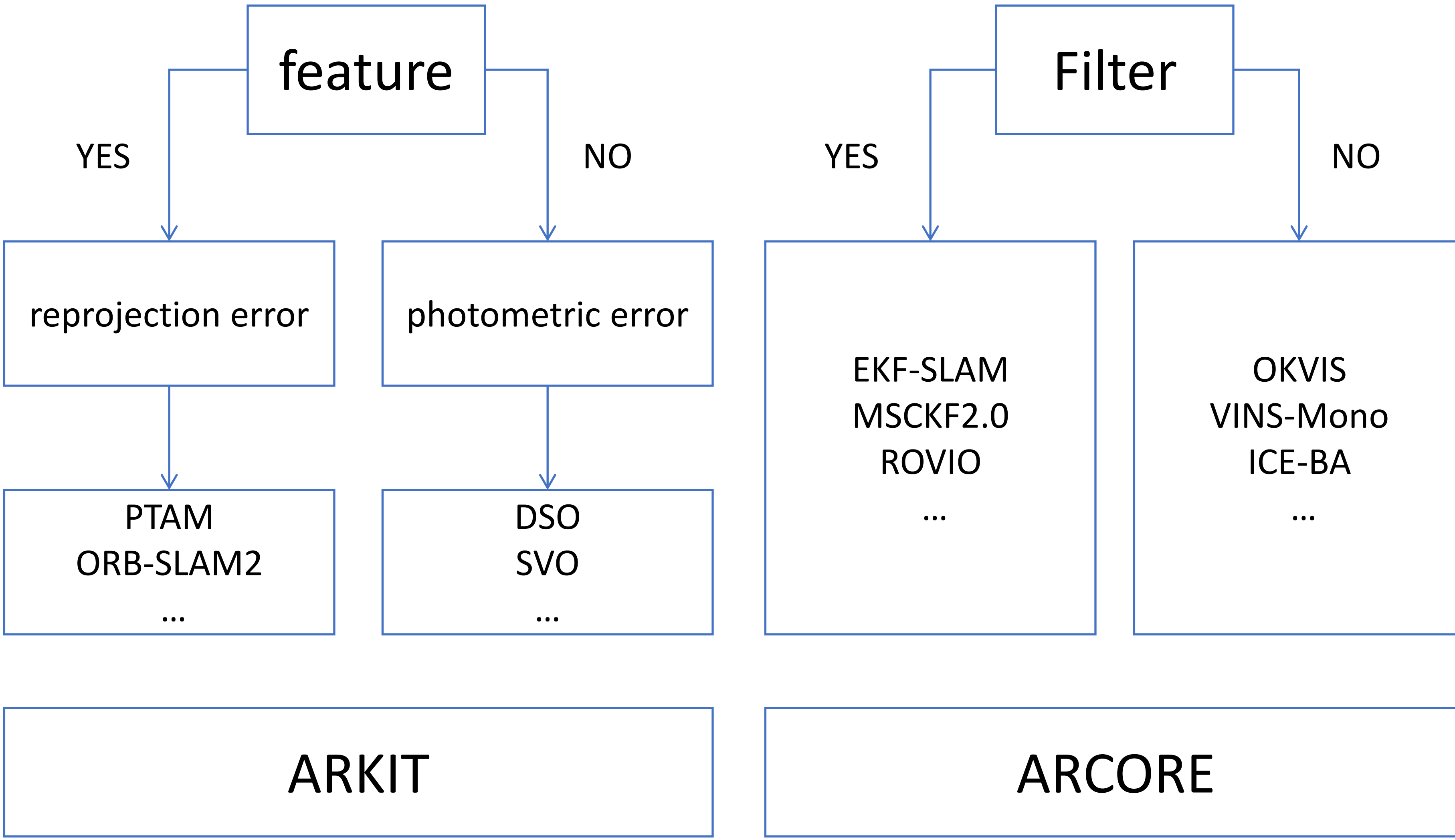
- Review the problem of VIO/VISLAM
- Gaps and bridge between research and engineering
- NEAR VI dataset
- AR oriented metric
- EZVIO
- Summary

Research closed loop



Note: We do not differ **VIO** and **VISLAM** in this tutorial.

Review: VIO Framework



Note: We do not differ **VIO** and **VISLAM** in this tutorial.

Review: Odometry metrics

- ATE

$$r_{\text{ate}} = \min_{\mathrm{T} \in \mathrm{SE}(3)} \sqrt{\frac{1}{|I_{gt}|} \sum_{i \in I_{gt}} \|\mathrm{T}\mathbf{p}_i - \hat{\mathbf{p}}_i\|^2},$$

- RPE

$$r_{\text{rpe}} = \sqrt{\frac{1}{|I_{\text{gt}}|} \sum_{i \in I_{\text{gt}}} \|\text{trans}(\mathbf{E}_i)\|^{\mathbf{2}}}$$

$$\mathbf{E}_i = \left(\hat{\mathrm{T}}_i^{-1} \hat{\mathrm{T}}_{i+\Delta}\right)^{-1} \left(\mathrm{T}_i^{-1} \mathrm{T}_{i+\Delta}\right)$$

reference:

[1] https://github.com/MichaelGrupp/evo

[2] A Tutorial on Quantitative Trajectory Evaluation for V(-Inertial) Odometry
github.com/uzh-rpg/rpg_trajectory_evaluation

Review: VI Datasets

Dataset	Year	Environment & Carrier	Cameras	IMUs	Time sync.	Ground truth	Stats/Props
EUROC	2016	indoors & MAV	-1 stereo gray: 2x768x480 @20Hz global shutter	-ADIS16488 3-axis acc/gyro @200Hz	hw	-6DOF (MoCap) @100Hz acc 1mm -3D (laser tracker) @20Hz acc 1mm	11 seqs 0.9 km
PennCOSYVIO	2017	indoors outdoors & handheld	-4 RGB: 1920x1080 @30Hz rolling shutter -1 stereo gray: 2x752x480 @20Hz global shutter -1 fisheye gray: 640x480 @30Hz global shutter	-ADIS16488 3-axis acc/gyro @200Hz -Tango 3-axis acc @128Hz 3-axis gyro @100Hz	hw (stereo gray/ ADIS) sw	-6DOF (visual tags) acc 15cm	4 seqs 0.6 km
TUM VI	2018	indoors outdoors & handheld	-1 stereo gray fisheye: 2x1024x1024 @20Hz global shutter	-BMI160 3-axis acc/gyro @200Hz	hw	-6DOF (MoCap) @120Hz acc 1mm (3DOF static mark)	28 seqs 20 km
ADVIO	2018	indoors outdoors & handheld	-1 RGB: 1280x720 @60Hz rolling shutter -1 gray fisheye: 640x480 @60Hz rolling shutter	3-axis acc/gyro: iphone @100Hz tango @200Hz	sw	-6DOF (INS) @100Hz acc 0.1m-1m	23 seqs 4.5 km
CVG ZJU	2018	indoors & handheld	-2 RGB: 640x480 @30Hz rolling shutter	3-axis acc/gyro: iPhone @100Hz Xiao Mi @400Hz	sw	-6DOF (MoCap) @400Hz acc 1mm	16 seqs 376 m

1. collection device
2. motion pattern
3. total sequences

SLAM challenge training dataset @ISMAR2019

Reference:

[1] NEAR: The NetEase AR Oriented Visual Inertial Dataset.

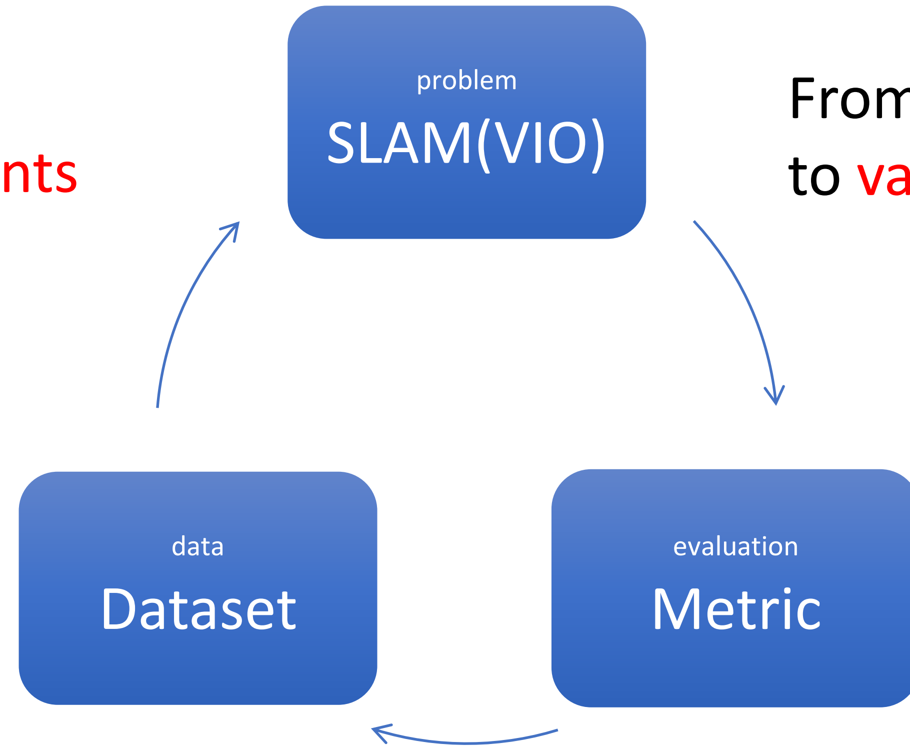
Gaps between research and engineering

- **Device price**
  - Quality of raw data
  - Computing resource
- **Device coverage**
  - Computation platform
  - algorithm optimization

Bridge from research to engineering

Standing on the shoulders of giants

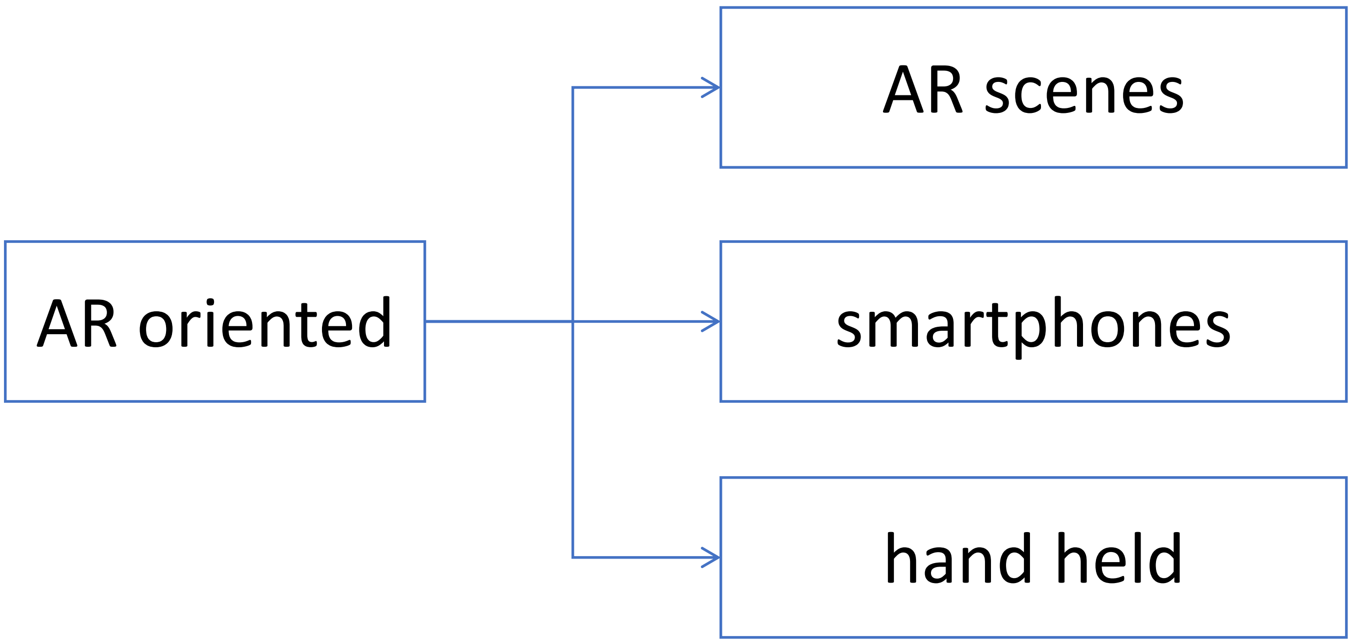
95%



From the specific device to various devices

5%

NEAR: An AR oriented dataset

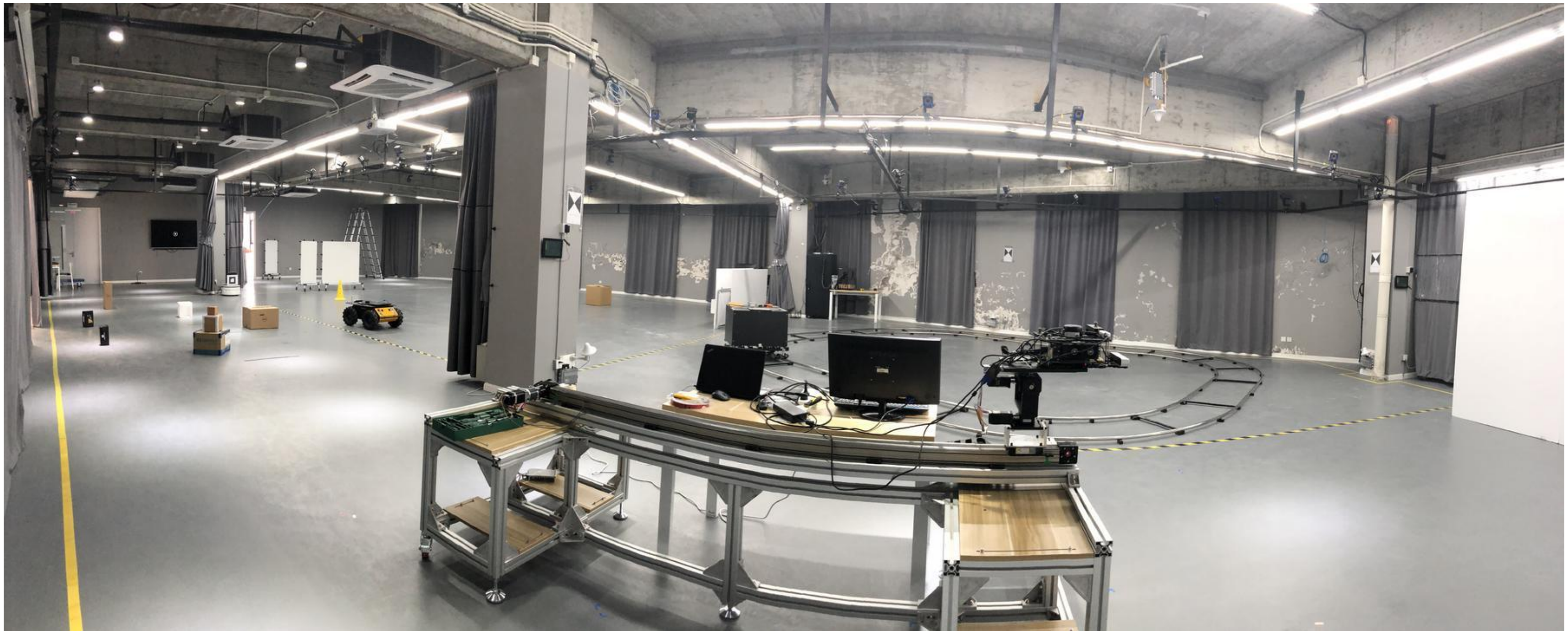


113 sequences, 49 cases , 4.8 km   Largest dataset for cellphone AR

https://github.com/EZXR-Research/NEAR-VI-Dataset

NEAR Design: Ground Truth

- Shanghai Beidou Research Institute @ 10x10m
- RealVision Tracking System(RTS) 4 @ 1mm, 120Hz, 6DOF



NEAR Design: Scene Setup

Table area

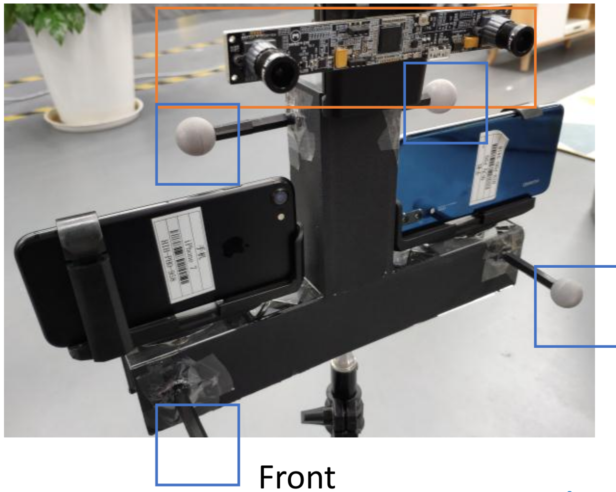
Living area

Rich texture

Medium texture

High texture

NEAR Design: Device Selection and Setting



MYNT

Front

Default setting

capturing @60Hz, saving @30Hz

Device /for short	Camera	Shutter	IMU(Hz)
iPhone 7 /ip7	1280x720@60fps	Rolling	100
iPhone XR /ipxr	1280x720@60fps	Rolling	100
OnePlus 5T /1p5t	1280x720@30fps	Rolling	400
Huawei P20 /p20	1280x720@30fps	Rolling	250
MYNT /mynt	752x480@30fps	Global	250

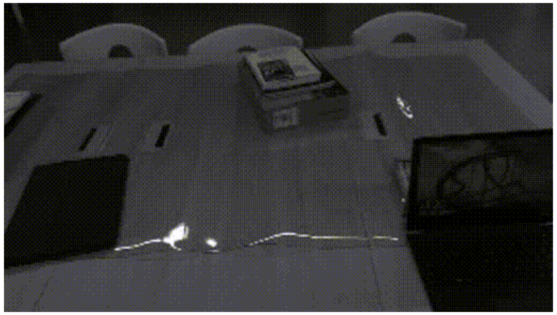
Rigid: 6 reflection balls

MYNT

Back

	Default	Test
Focus Mode	Fixed	Auto
Sampling Rate	60Hz@iPhone	30Hz@iPhone

NEAR Design: Motion Pattern



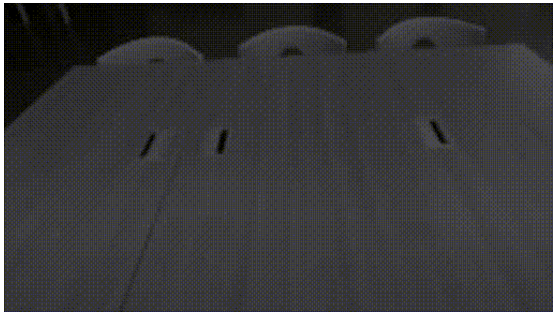
Case 006 Circular motion

Case 008 Translational motion

Case 34 Hybrid motion

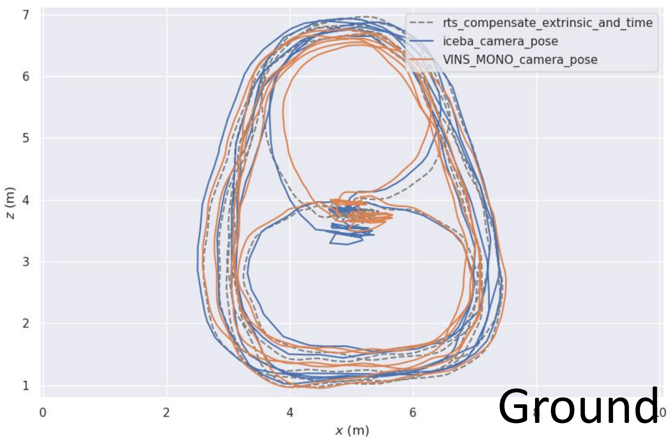
Case 037 Free motion

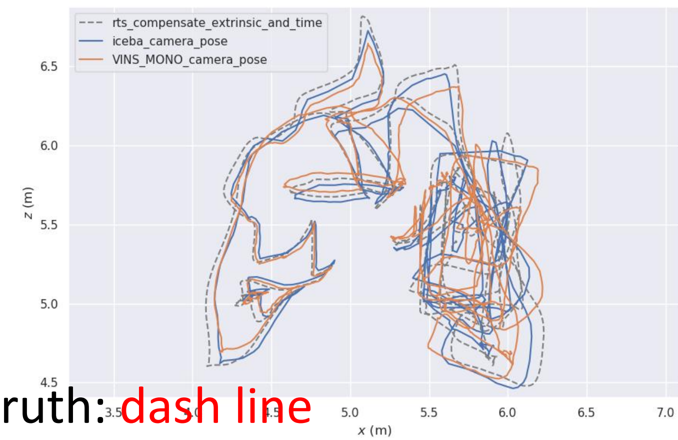
Case 039 Gaming motion(Shoting)

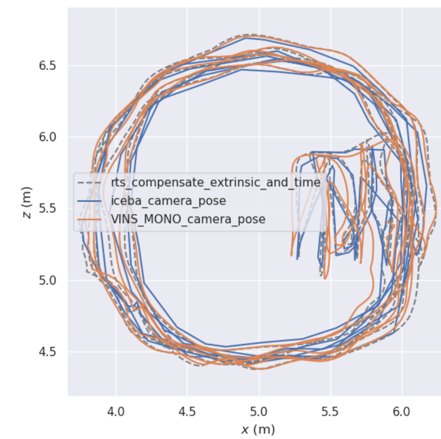


Gaming motion(YuME)

NEAR Design: Typical trajectories





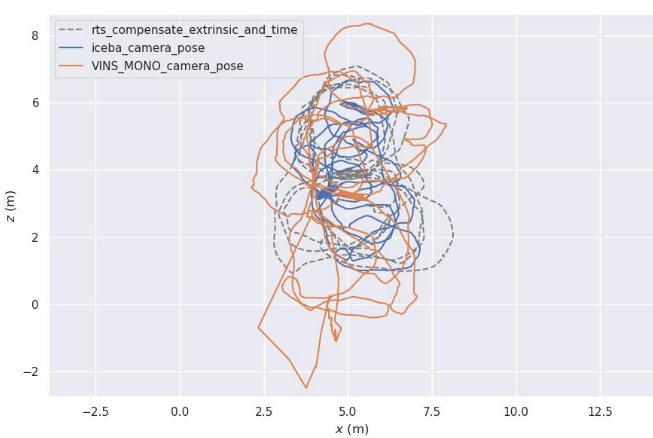


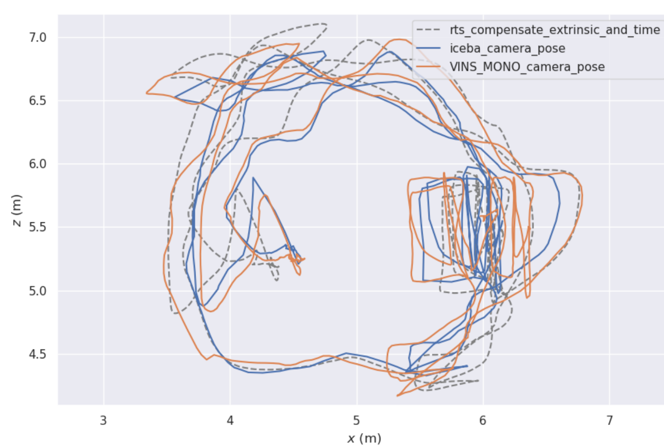
Ground truth: dash line

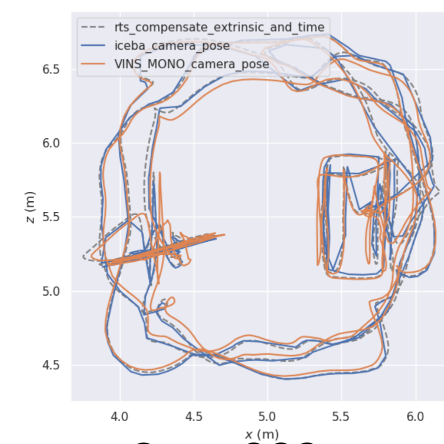
Case 009_control
(circular motion)

Case 019
(YuME imitation)

Case 003
(circle motion)





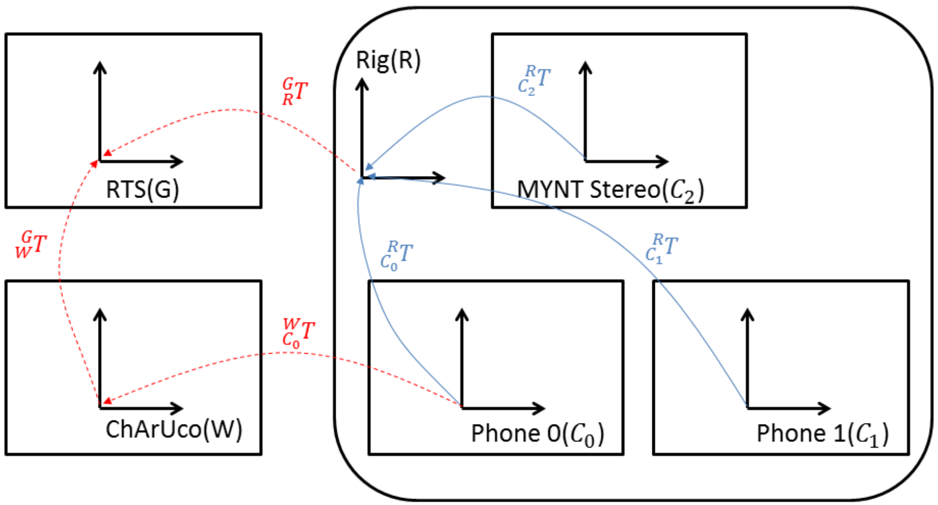


Case 037
(free motion)

Case 048
(hybrid motion)

Case 002
(hybrid motion)

NEAR Design: Calibration



Transformation between MoCap-Cam:

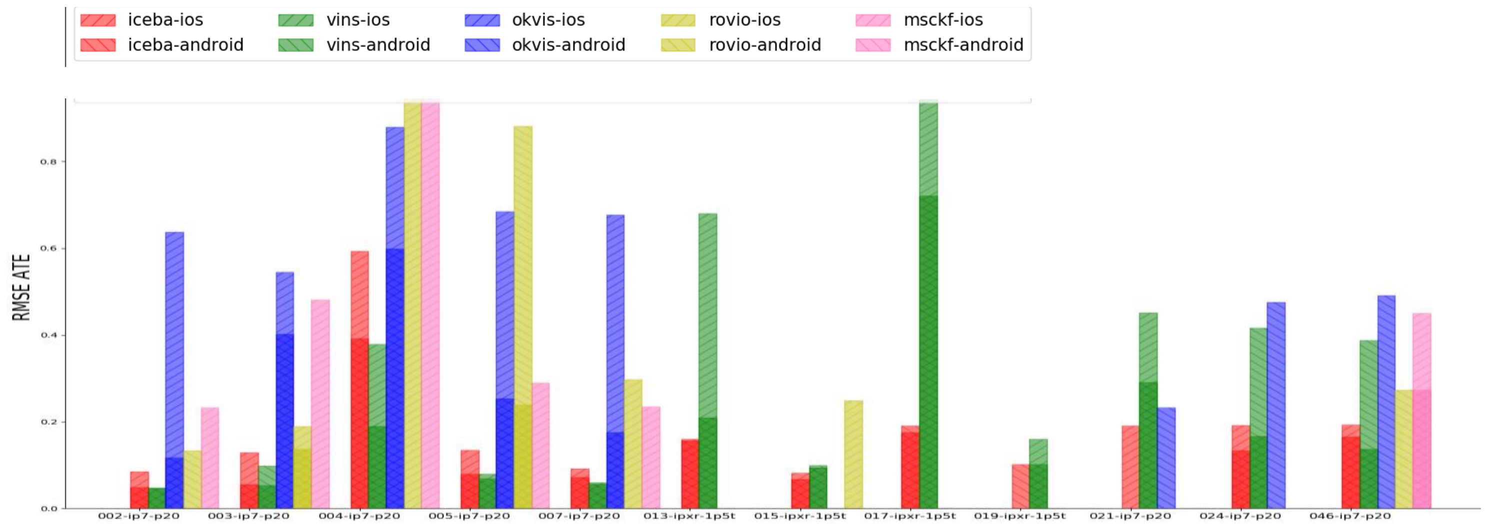
Marker Odometry $\quad {}^{G}_{W}\mathrm{T}\,\boxed{{}^{W}_{C}\mathrm{T}} \equiv \boxed{{}^{G}_{R}\mathrm{T}}\,{}^{R}_{C}\mathrm{T}$ $\quad$ Motion Capture System

Formulate the optimization problem:

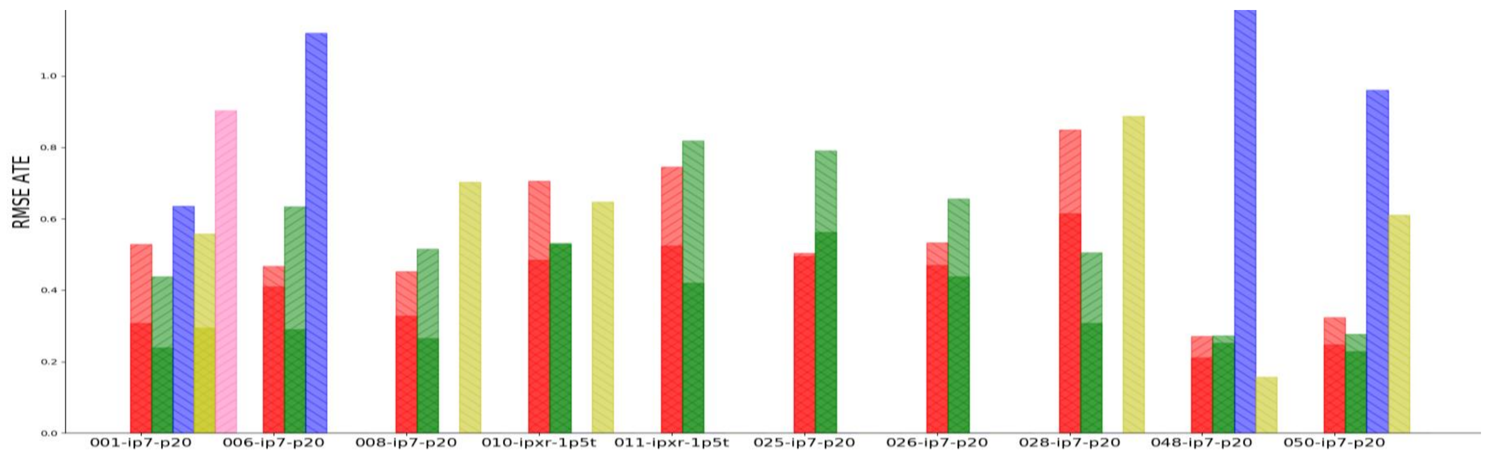
$$\underset{{}^{G}_{W}\mathrm{T},\ {}^{R}_{C}\mathrm{T}}{\arg\min} \sum_{i} \| \boxed{{}^{G}_{W}\mathrm{T}}\,{}^{W}_{C}\mathrm{T}_{i} - {}^{G}_{R}\mathrm{T}_{j(i)}\,\boxed{{}^{R}_{C}\mathrm{T}} \|_{F}^{2}$$

1.Linear solver $\quad$ 2.least square solver

NEAR Validation: Overall ATE



Easy group



Medium group

NEAR Validation: Brief

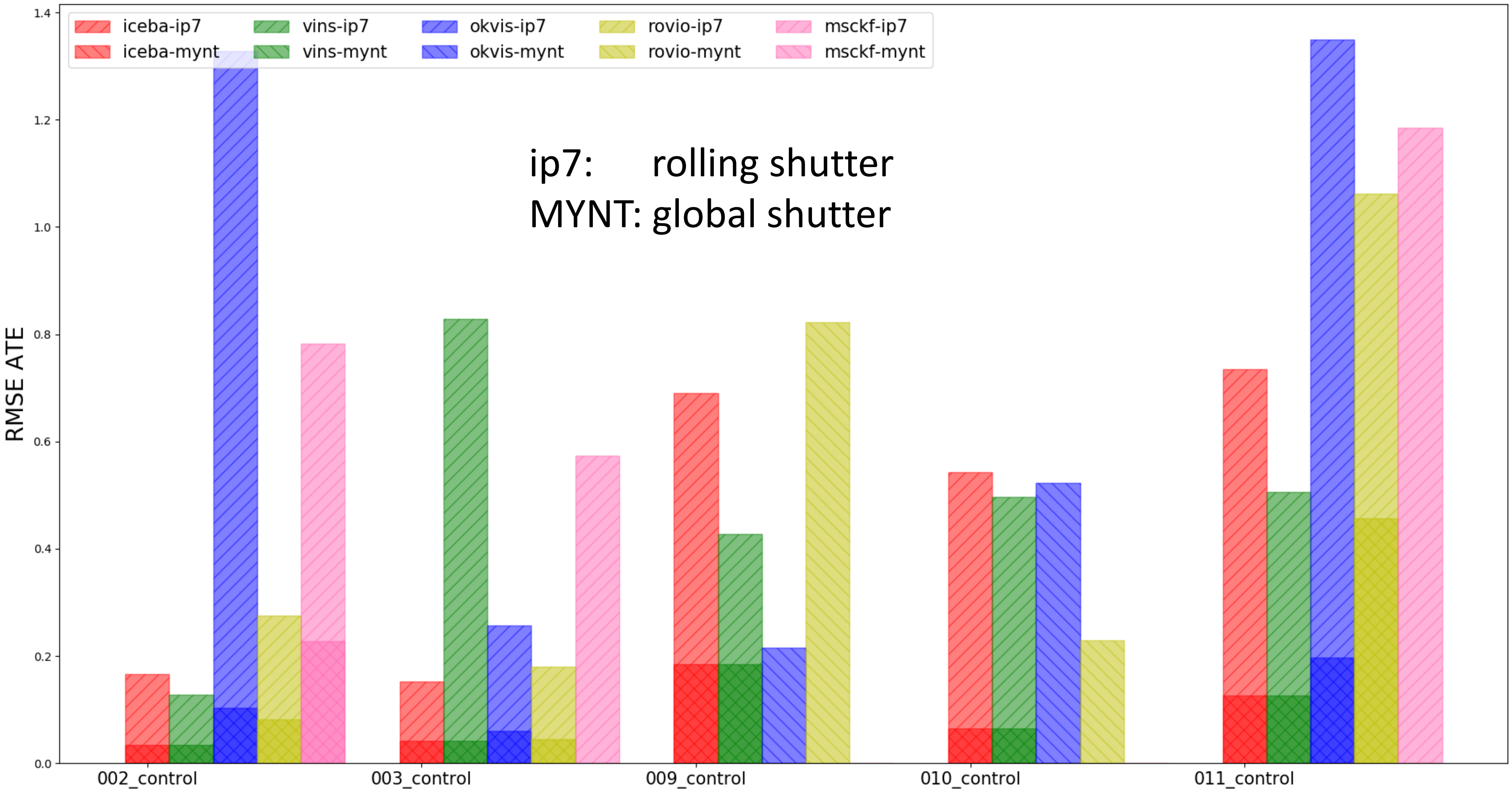
Group	ATE range	Cases
Control 5 Cases	-	002_control, 003_control, 009_control, 010_control, 011_control
Easy 12 Cases	[0.0, 0.2)	002, 003, 004, 005, 007, 013, 015, 017, 019, 021, 024, 046
Medium 10 Cases	[0.2, 0.5)	001, 006, 008, 010, 011, 025, 026, 028, 048, 050
Hard 27 Cases	$[0.5, +\infty)$	the rest

Note:

Grouped by the best performance of the five state-of-the-art VIOs for each case.

002_control is the same as 002 except an additional MYNT camera.

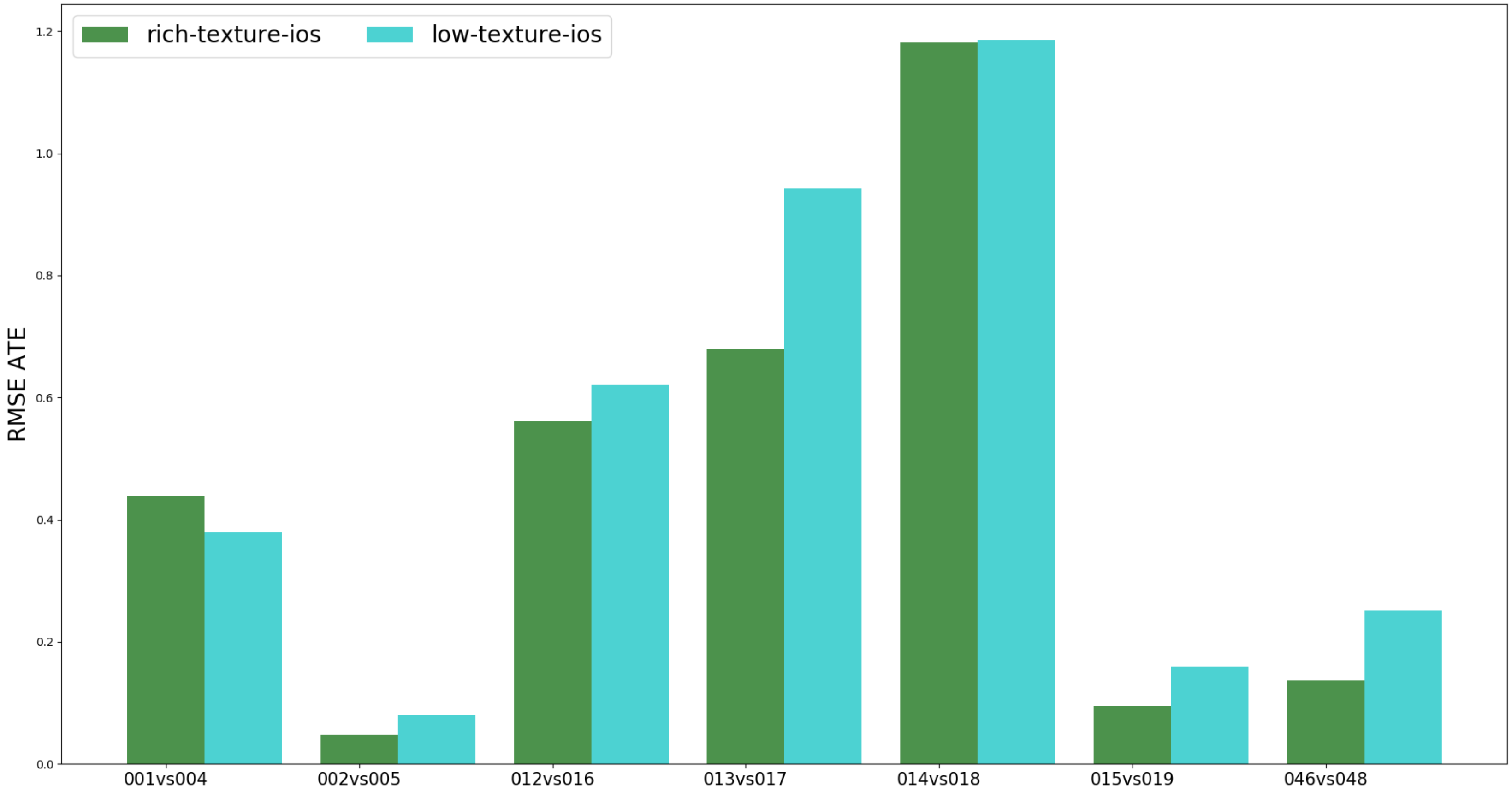
NEAR Validation: Rolling Shutter Effect



Note: The IMU of MYNT is different with the IMU of Iphone7, but the prices are both civil level.

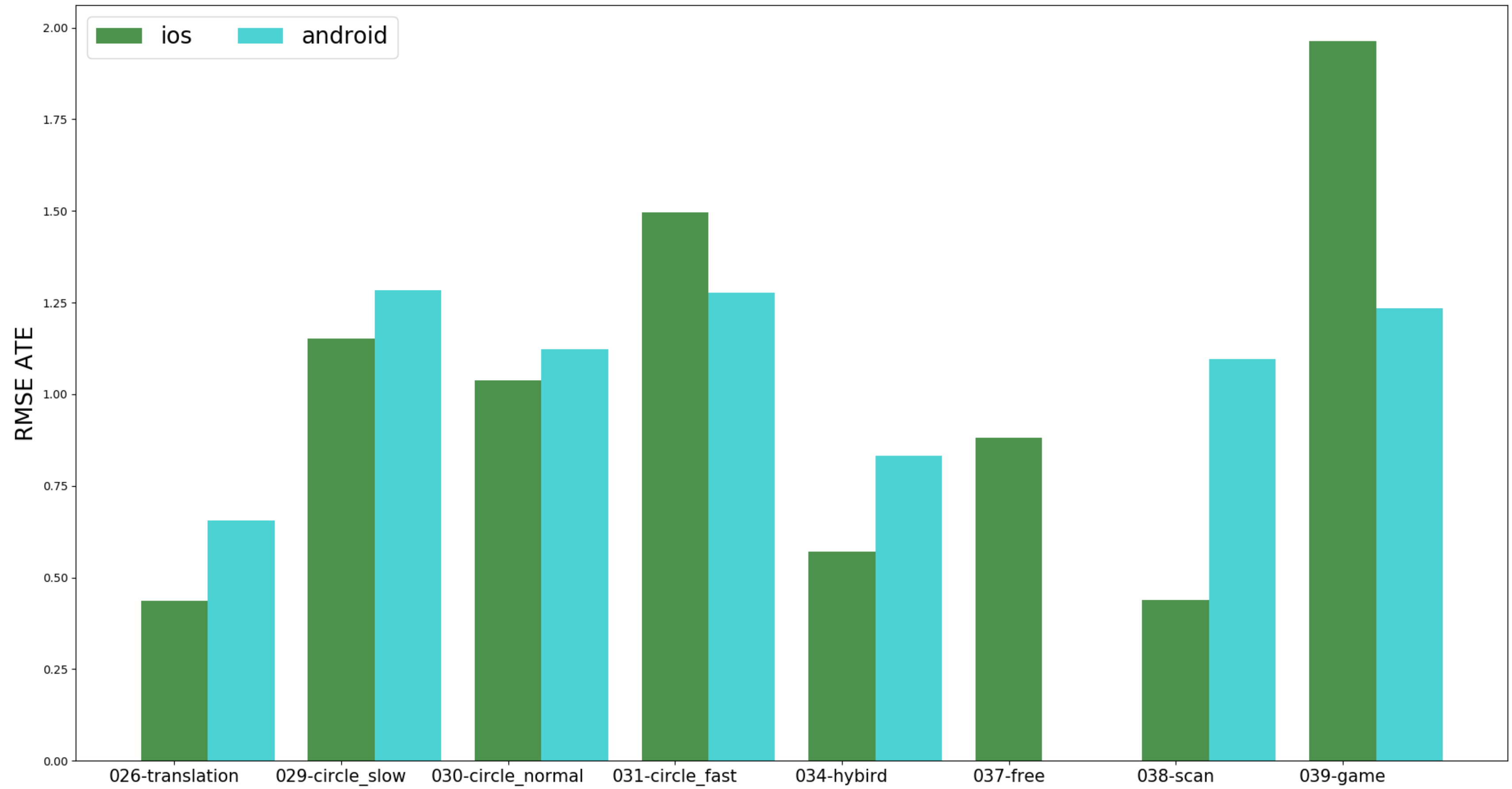
NEAR Validation: Scene Texture

Example: VINS-Mono

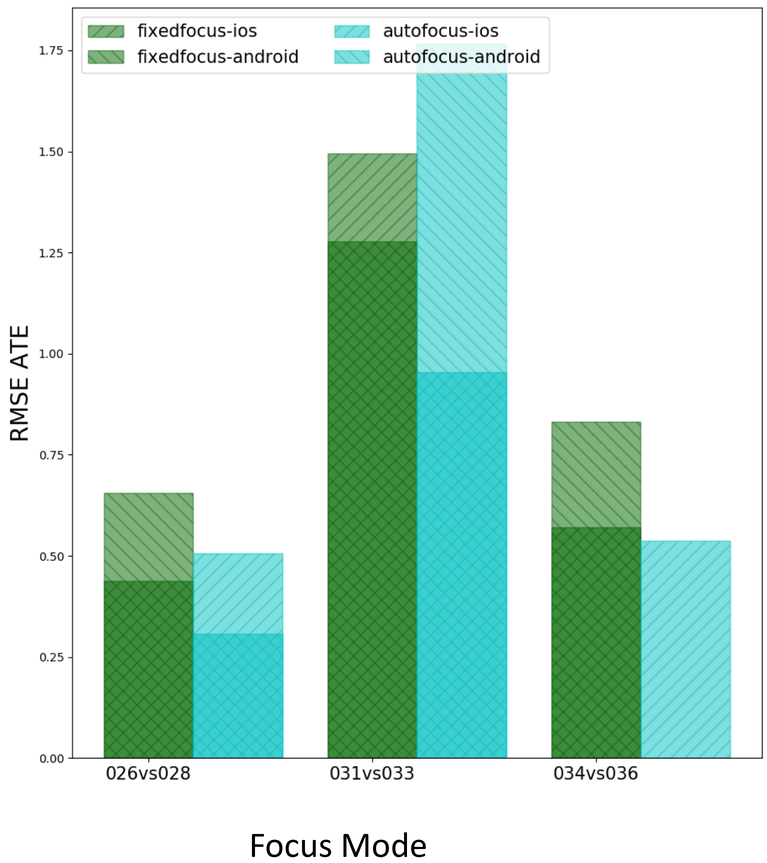


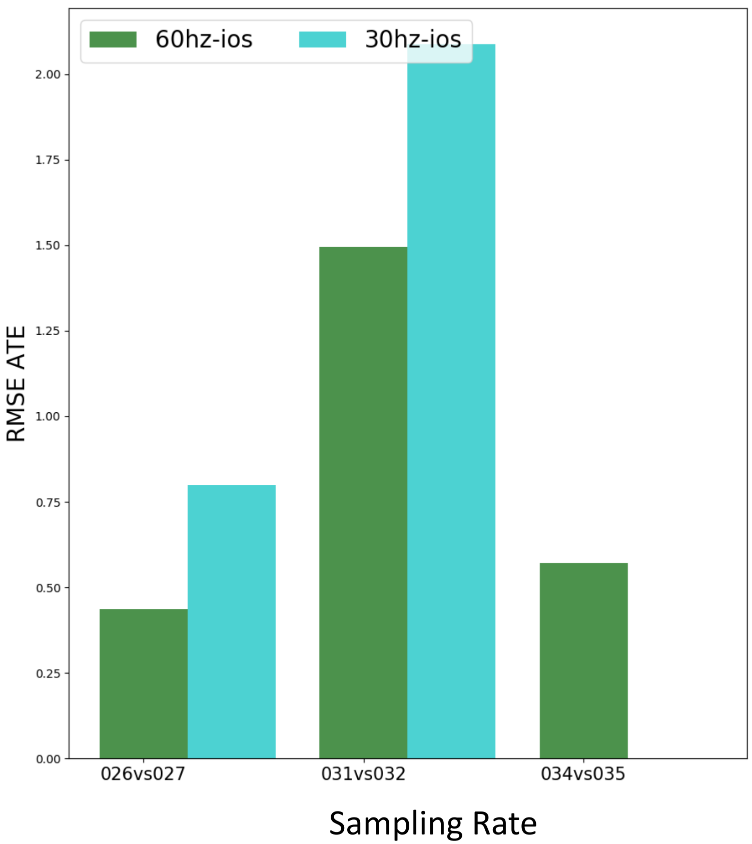
NEAR Validation: Motion Pattern

Example: VINS-Mono

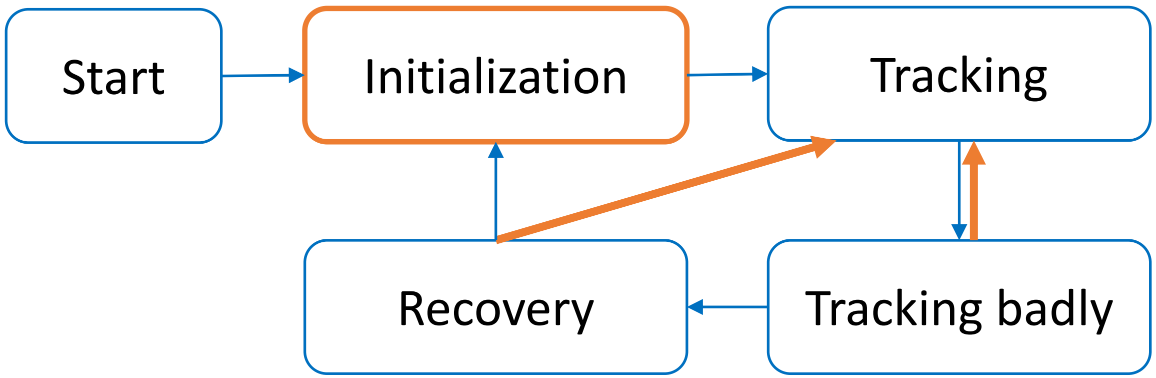


NEAR Validation: Device Setting





AR Metric: AR oriented metrics



Safe AR

User Experience

FAST initialization
ROBUST, ACCURATE and SMOOTH tracking
FAST recovery
EFFICIENT computation

[1] J. LI, B. YANG, D. CHEN, N. WANG, G. ZHANG, and H. BAO. Survey and evaluation of monocular visual-inertial slam algorithms for augmented reality. *Virtual Reality Intelligent Hardware*

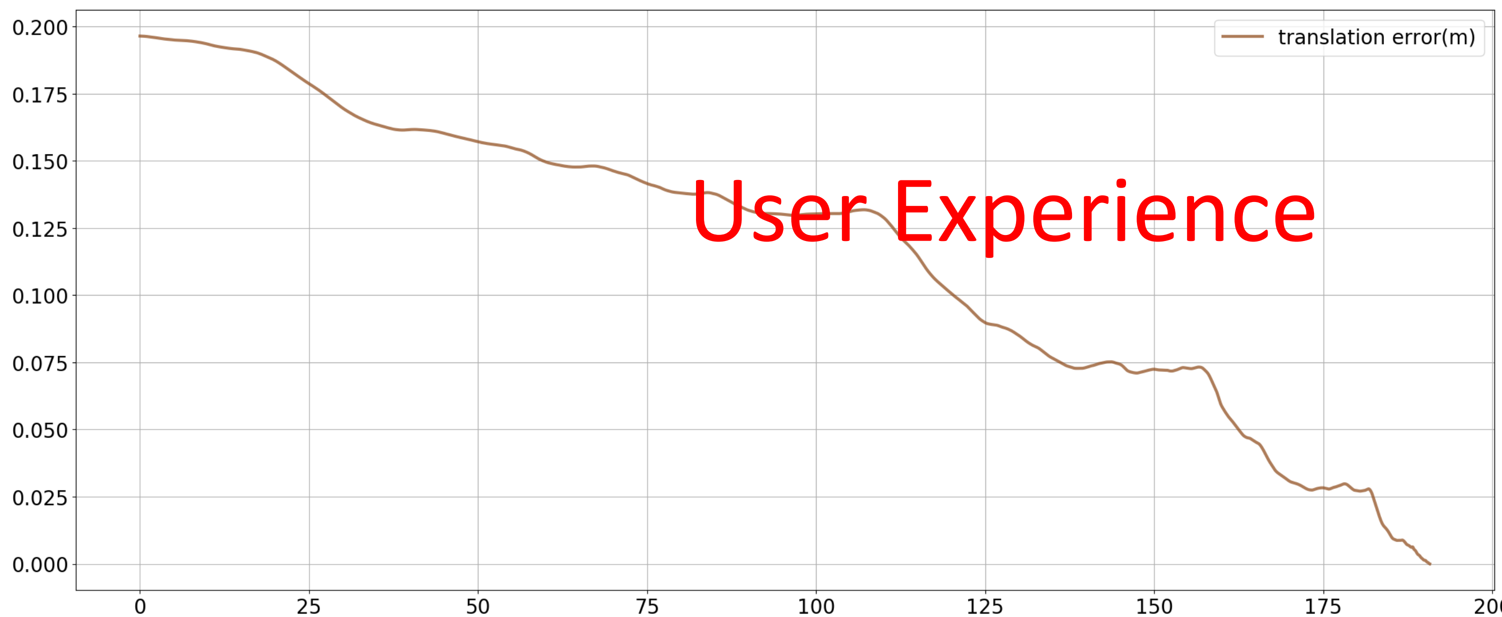
易现 EZXR

AR Metric: AR oriented metrics

ATE

Initial time

$$r_{\text{init}}\boxed{(t)} = \min_{T \in SE(3)} \sqrt{\frac{1}{|I_{gt}(t)|} \sum_{i \in I_{gt}(t)} \|\mathbf{T}\mathbf{p}_i - \hat{\mathbf{p}}_i\|^2}$$



The curve of time-ATE

EZVIO-008_ip7 (no loop closure)

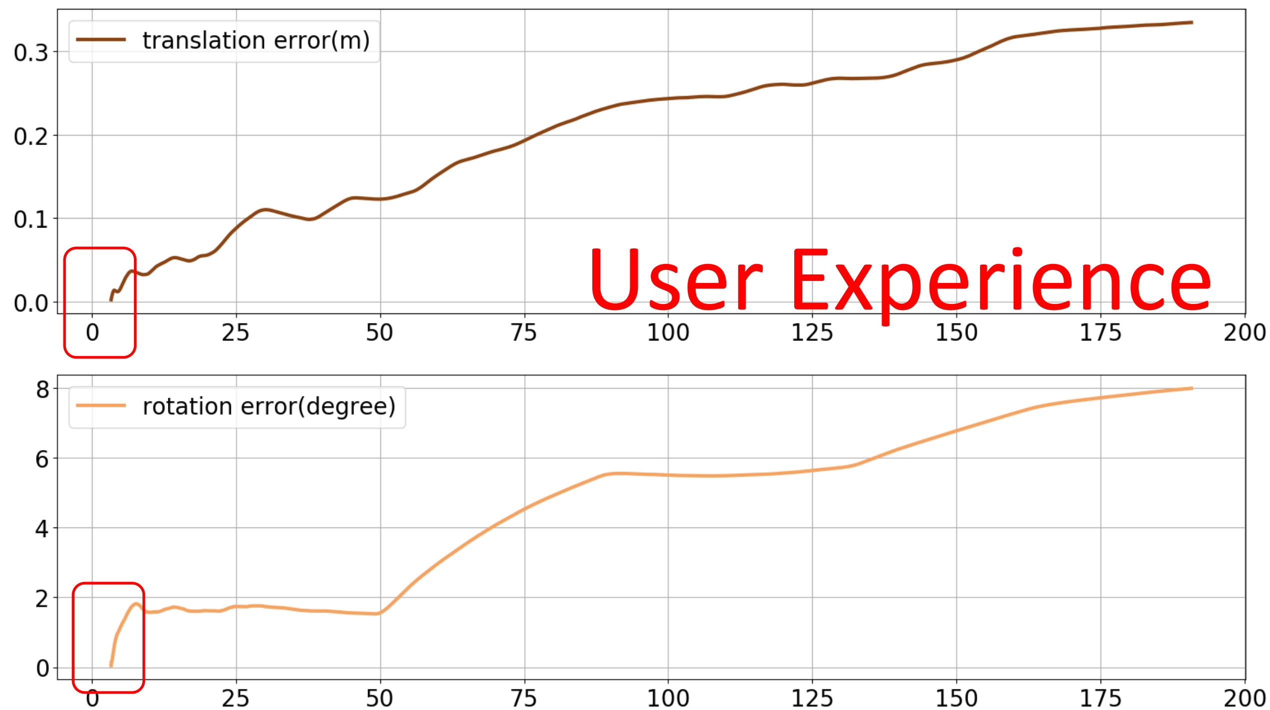
易现
E Z X R

AR Metric: Drift Error

RPE

anchor

$$\mathbf{E}_j = \left(\boxed{\hat{\mathrm{T}}_i^{-1}}\hat{\mathrm{T}}_{i+j}\right)^{-1}(\mathrm{T}_i^{-1}\mathrm{T}_{i+j}) \qquad r_{\text{drift\_trans}} = \sqrt{\frac{1}{|I_{\text{gt}}|}\sum_{j\in I_{\text{gt}}}\|\text{trans}(\mathbf{E}_j)\|^2}$$



The curve of time-RPE EZVIO-008_ip7 (no loop closure)

AR Metric: Jitter Error & Jerkiness Rate

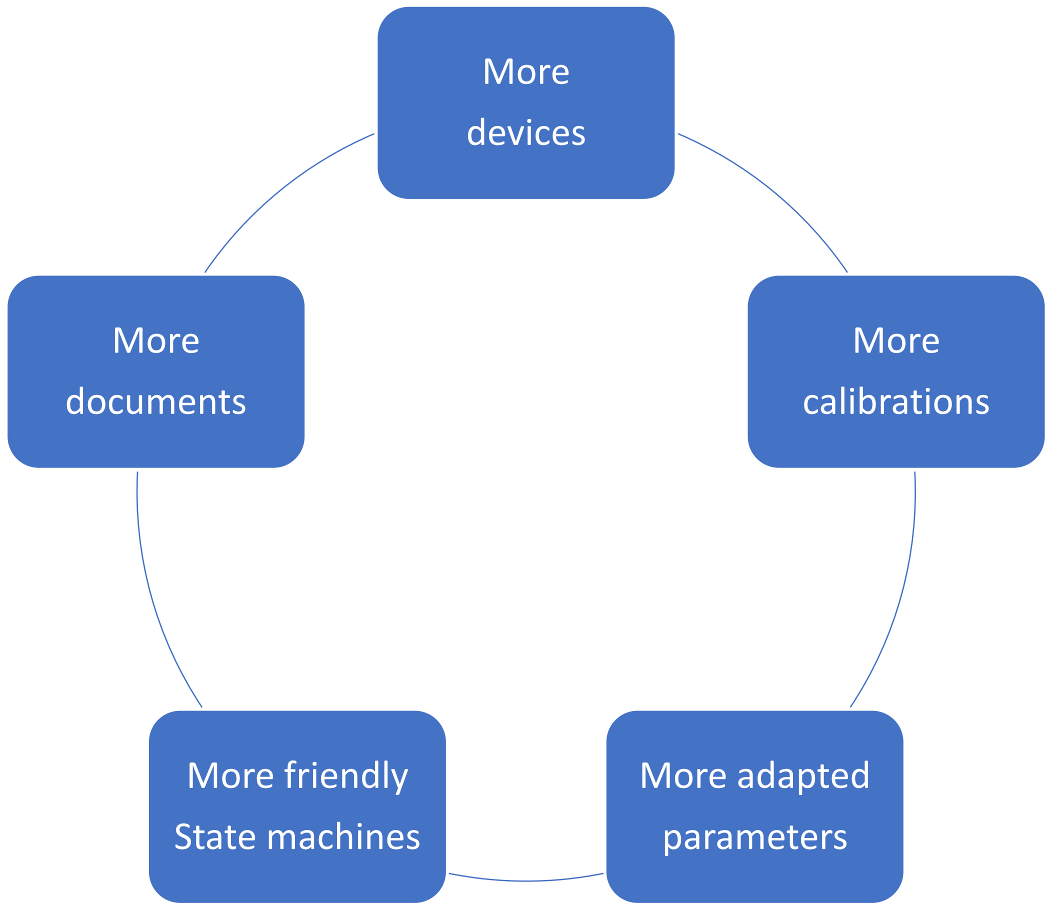
RPE

$$r_{\text{jitter}} = \mathrm{RMSE}\left(\|\mathrm{trans}\mathbf{E}_i\|, i \in \left(1 \le \|(\mathbf{E}_i)\| \boxed{\le \mathrm{u}}\right)\right)$$

$$r_{\text{jerk}} = \sum \mathrm{sign}\left(\left(\|(\mathbf{E}_i)\| > \mathrm{u}\right)\right)/\mathrm{N}$$

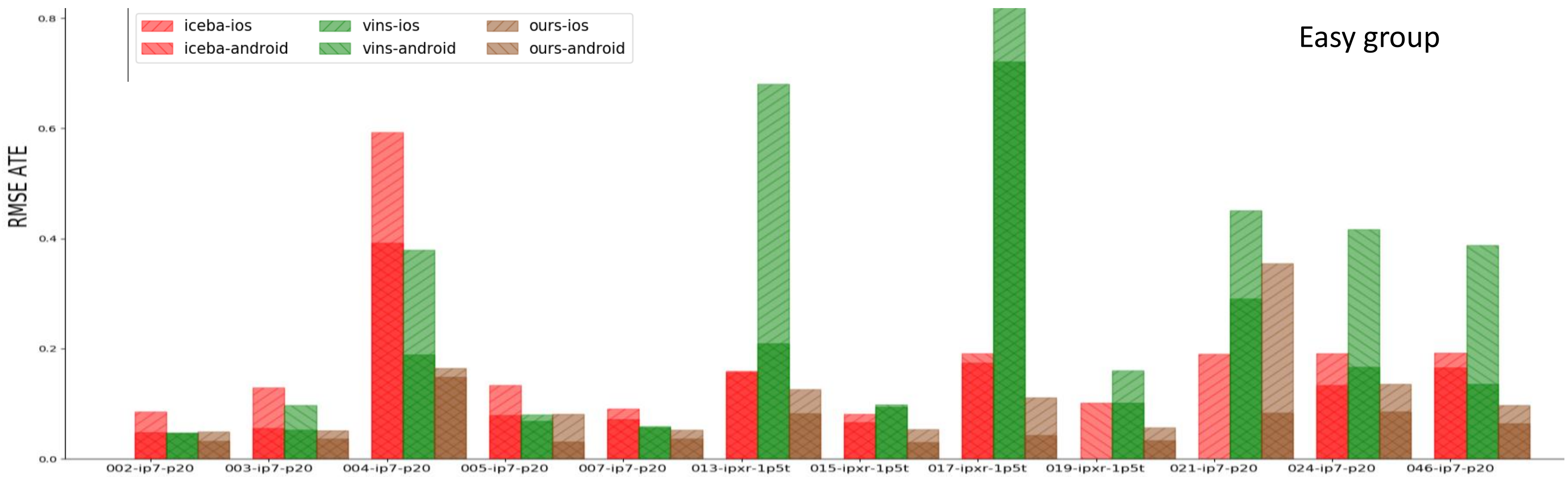
RPE curve
EZVO-008_ip7

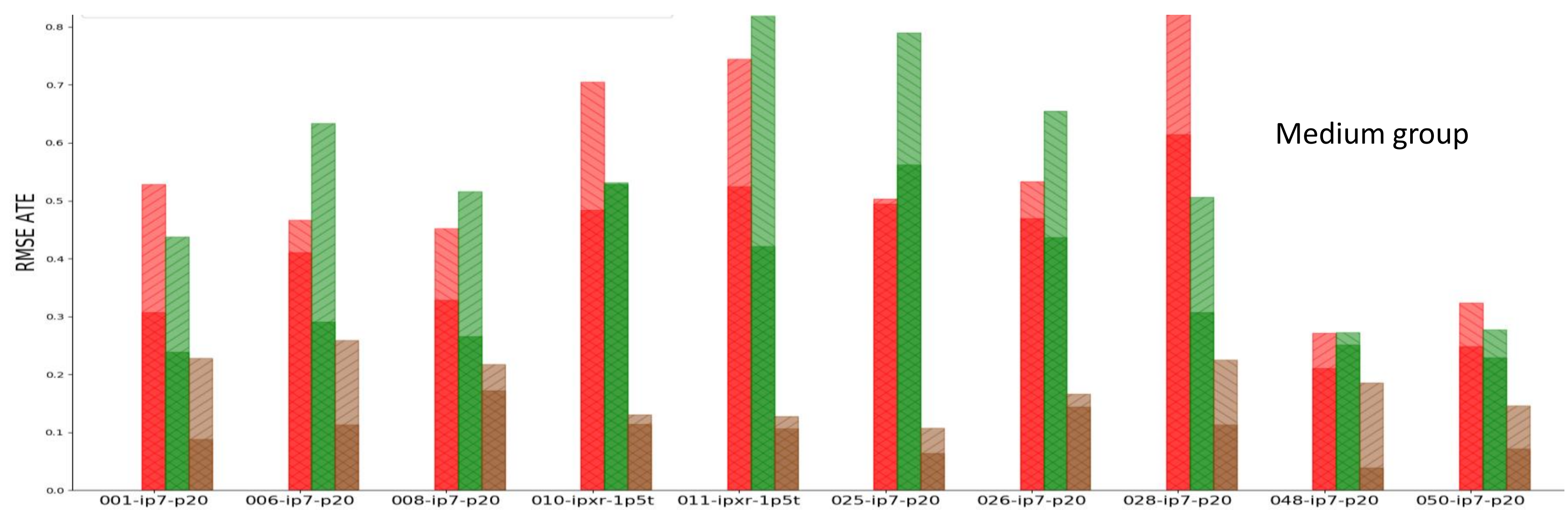
EZVIO: MSCKF2.0 for cellphones



https://github.com/EZXR-Research/ezvio

EZVIO: overall ATE





Thank You!
Q & A

Cheng Wang
wangcheng@ezxr.com
Hangzhou EasyXR Co., Ltd. China