Sensor-Aware Augmented Reality

Addressing Real-World HMI Challenges

Dr. Liu Ren

Global Head and Chief Scientist, HMI
Bosch Research North America
Palo Alto, CA

BOSCH

Sensor-Aware Augmented Reality

Bosch Overview

Bosch is one of the world's leading international providers of technology and services

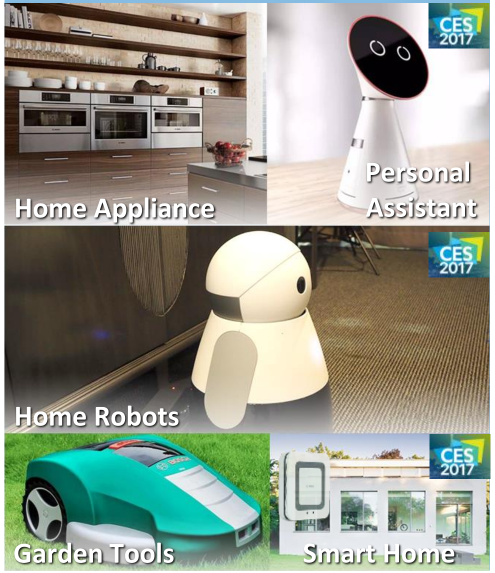
- 375,000[1] Bosch associates
- More than 440[1] subsidiary companies and regional subsidiaries in some 60[1] countries
- Including its sales and service partners, Bosch is represented in some 150[1] countries.

Mobility Solutions

Industrial Technology

Energy and Building Technology

Consumer Goods

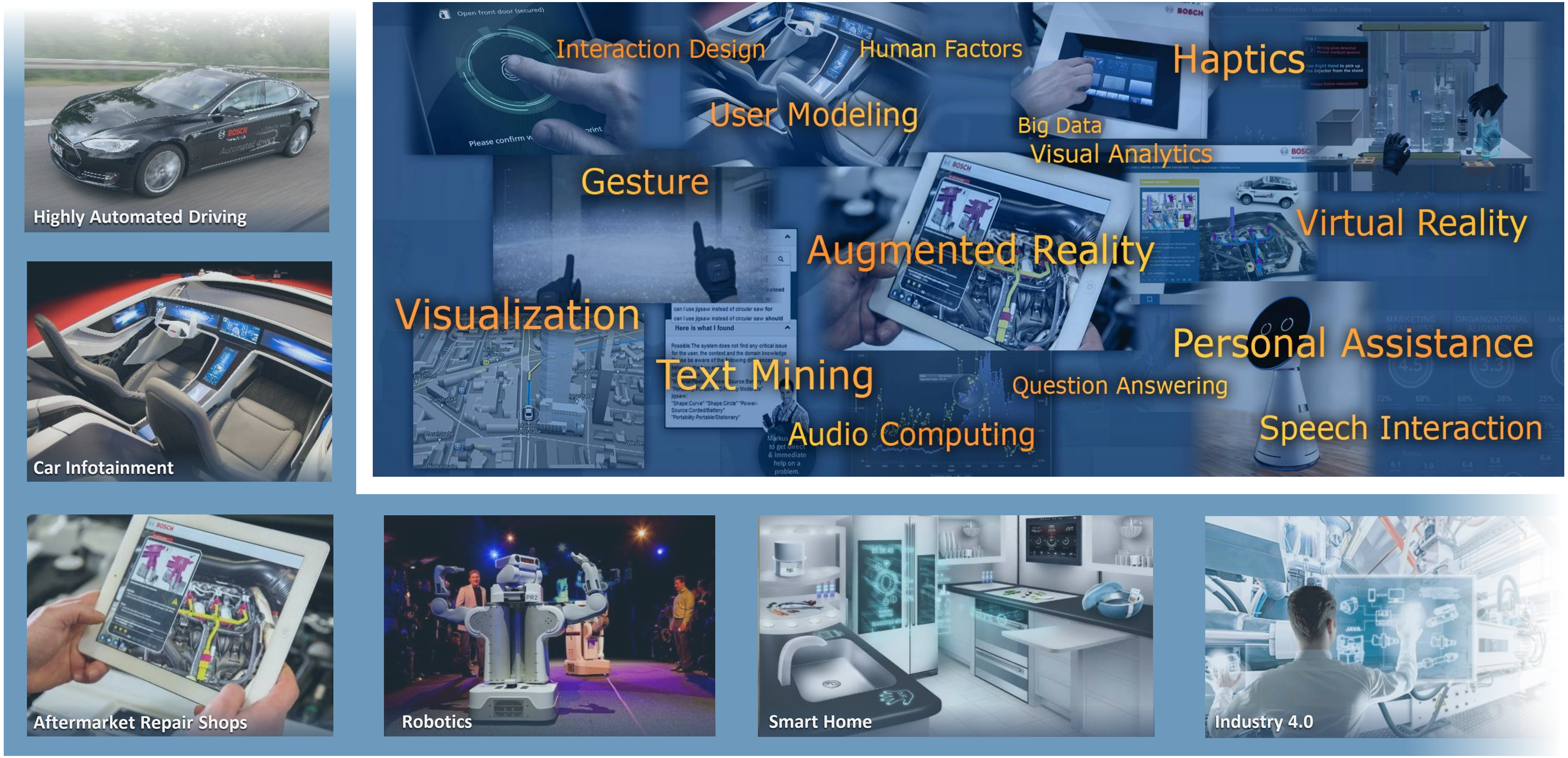


[1] *As of Dec. 2015*

BOSCH

Sensor-Aware Augmented Reality

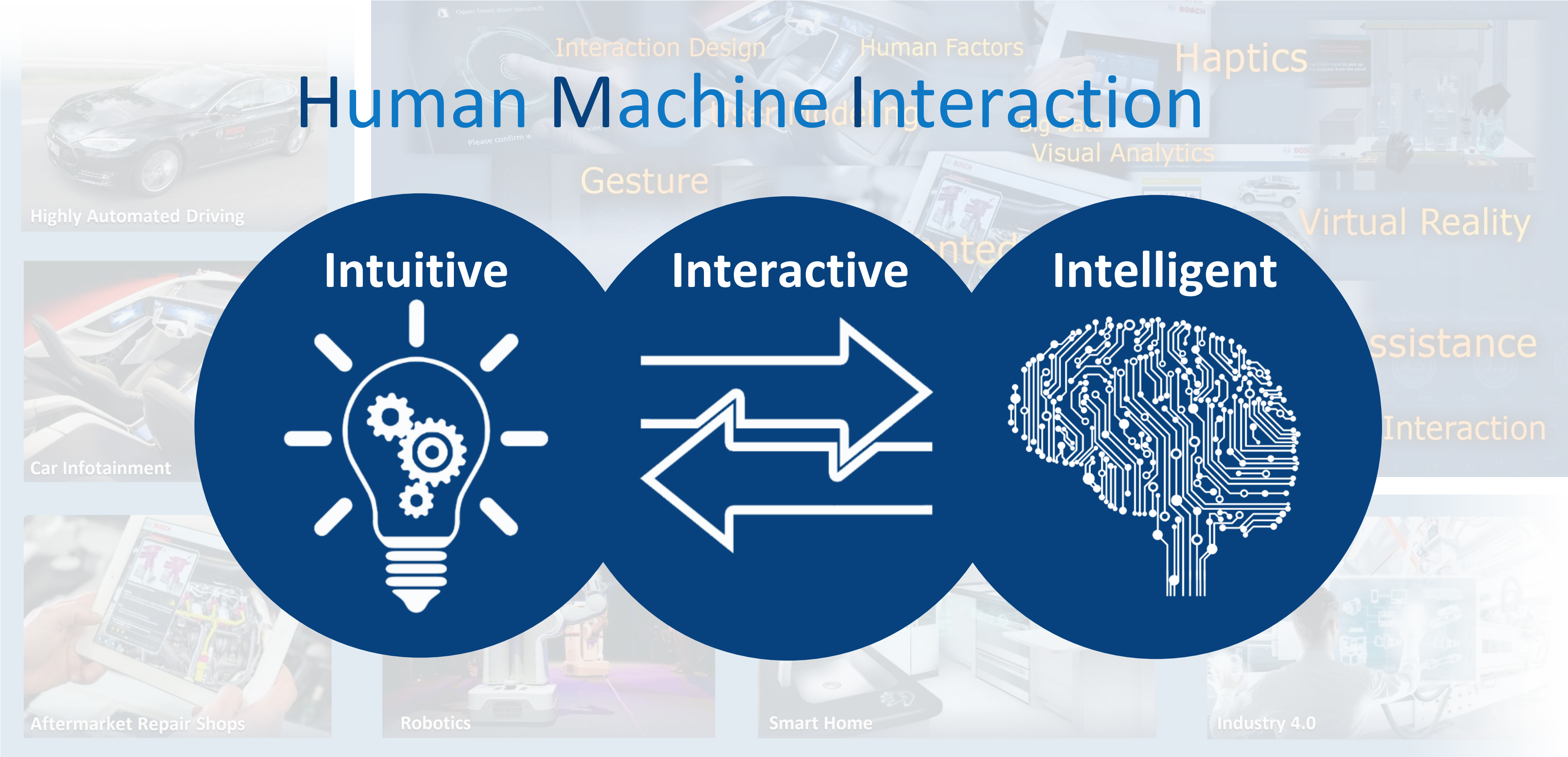
Human Machine Interaction (HMI) Research in Bosch



BOSCH

Sensor-Aware Augmented Reality

Key Success Factors of HMI Products



BOSCH

Sensor-Aware Augmented Reality

Global HMI Research Team

Liu Ren Short Bio

BOSCH

- Global Head and Chief Scientist, HMI, **Bosch Research**
- Ph.D. and M.Sc. in Computer Science, **Carnegie Mellon University**
- B.Sc. in Computer Science, **Zhejiang University**, P.R. China

HMI teams in Bosch Research

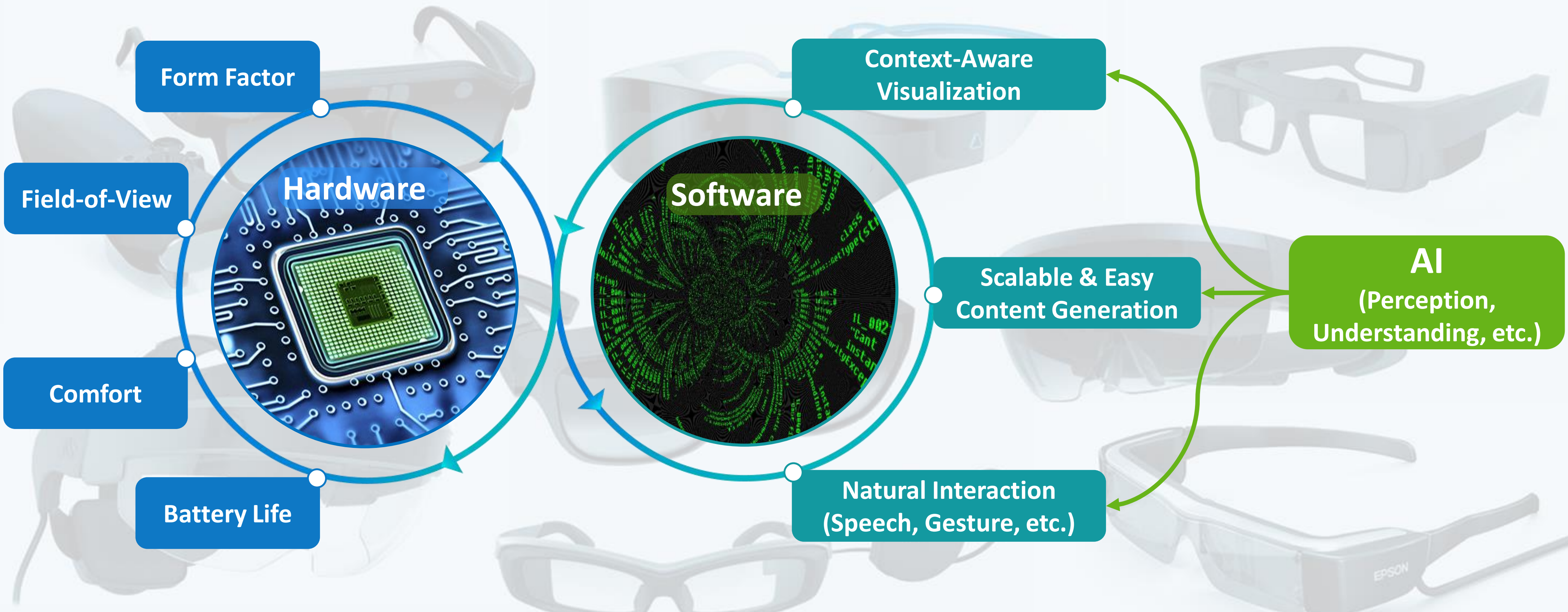
Headquarters
Palo Alto, USA

Renningen, Germany

Shanghai, China

Sensor-Aware Augmented Reality

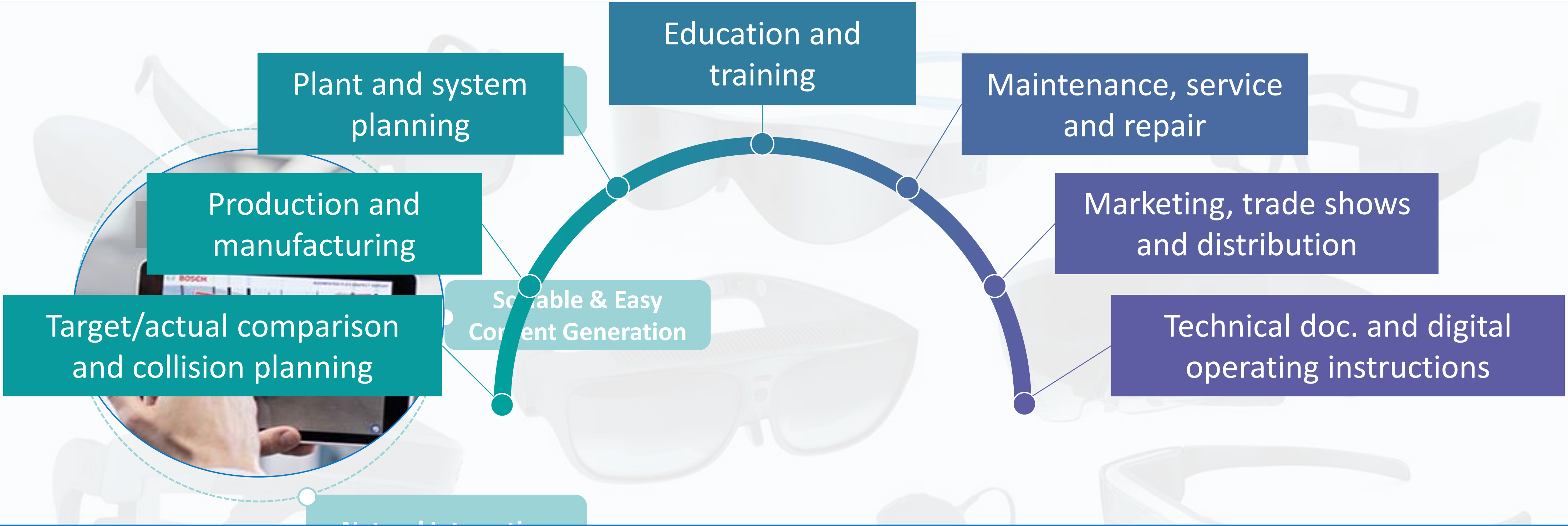
Real-World HMI Challenges for (Wearable) AR



BOSCH

Sensor-Aware Augmented Reality

Bosch Product: CAP (Common Augmented Reality Platform)



Bosch CAP enables implementation of complete enterprises AR solutions

- Integrates the production of visual and digital content directly into the authoring process.
- Existing CAD, image and video data were used and save the expense of creating new content.

Sensor-Aware Augmented Reality

Our Sensor-Aware Solutions

Software

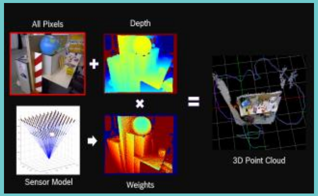
Context-Aware Visualization

Scalable & Easy Content Generation

Natural Interaction (Speech, Gesture, etc.)

Dynamic occlusion handling
ISMAR 2016[1]

- Enhance realistic depth perception



Robust visual tracking
ISMAR 2016[2]

- Improve tracking robustness and accuracy

[1] Chao Du, Yen-Lin Chen, Mao Ye, and Liu Ren, "Edge Snapping-Based Depth Enhancement for Dynamic Occlusion Handling in Augmented Reality", IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2016.

[2] Benzun Wisely Babu, Soohwan Kim, Zhixin Yan, and Liu Ren, "σ-DVO: Sensor Noise Model Meets Dense Visual Odometry", IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2016.

Sensor-Aware Augmented Reality
Dynamic Occlusion Handling

Software

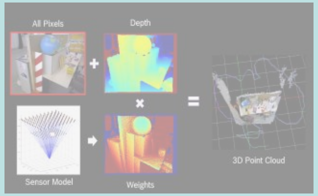
Context-Aware Visualization

Scalable & Easy Content Generation

Natural Interaction (Speech, Gesture, etc.)

Dynamic occlusion handling
ISMAR 2016[1]

- Enhance realistic depth perception



Robust visual tracking
ISMAR 2016[2]

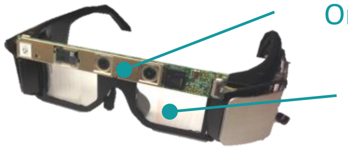
- Improve tracking robustness and accuracy

BOSCH

Sensor-Aware Augmented Reality
Dynamic Occlusion Handling: Motivation

Goals

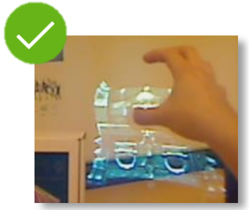
1. Compact setup: single sensor



One near-range RGBD sensor

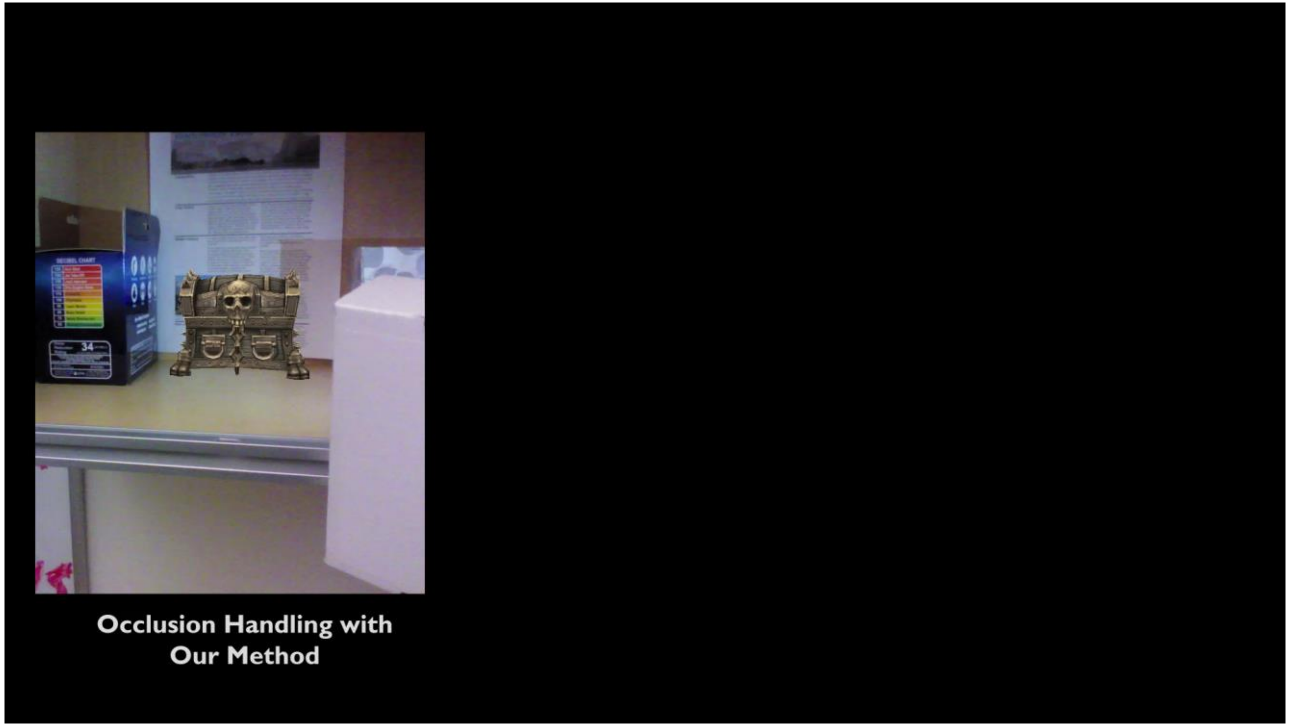
Optical see-through head-mounted display (HMD)

2. Dynamic occlusion handling



Challenges

- Performance requirements for real-time AR applications
- Limited computational resources (e.g., on tablet)



Sensor-Aware Augmented Reality

Dynamic Occlusion Handling: Our Sensor-Aware Solution

Knowledge on RGBD Sensor

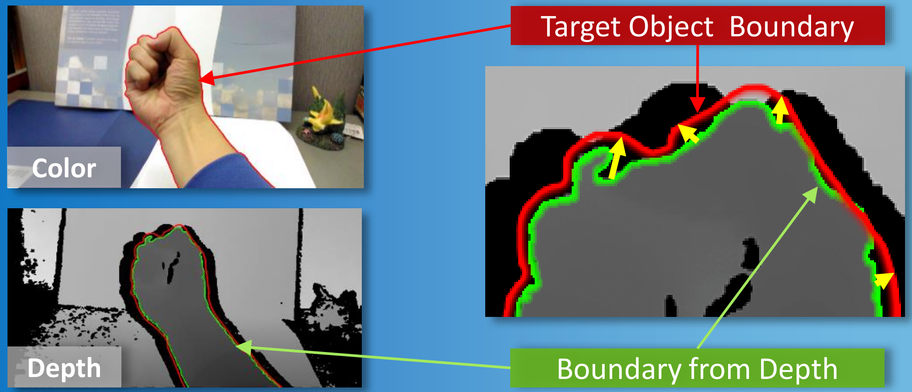
Depth data not reliable at the object boundary

- Structured Light/Stereo: matching is not accurate at the boundary
- Time of Flight: light signal reaching object boundary barely bounce back to the sensor

Edge Snapping-Based Algorithm

- Use color images as guidance
- Snap object boundaries in depth data towards the edges in color images
- Formulated as an optimization problem, efficiently solved via dynamic programming

1 *Align Object Boundary (Edge-Snapping)*

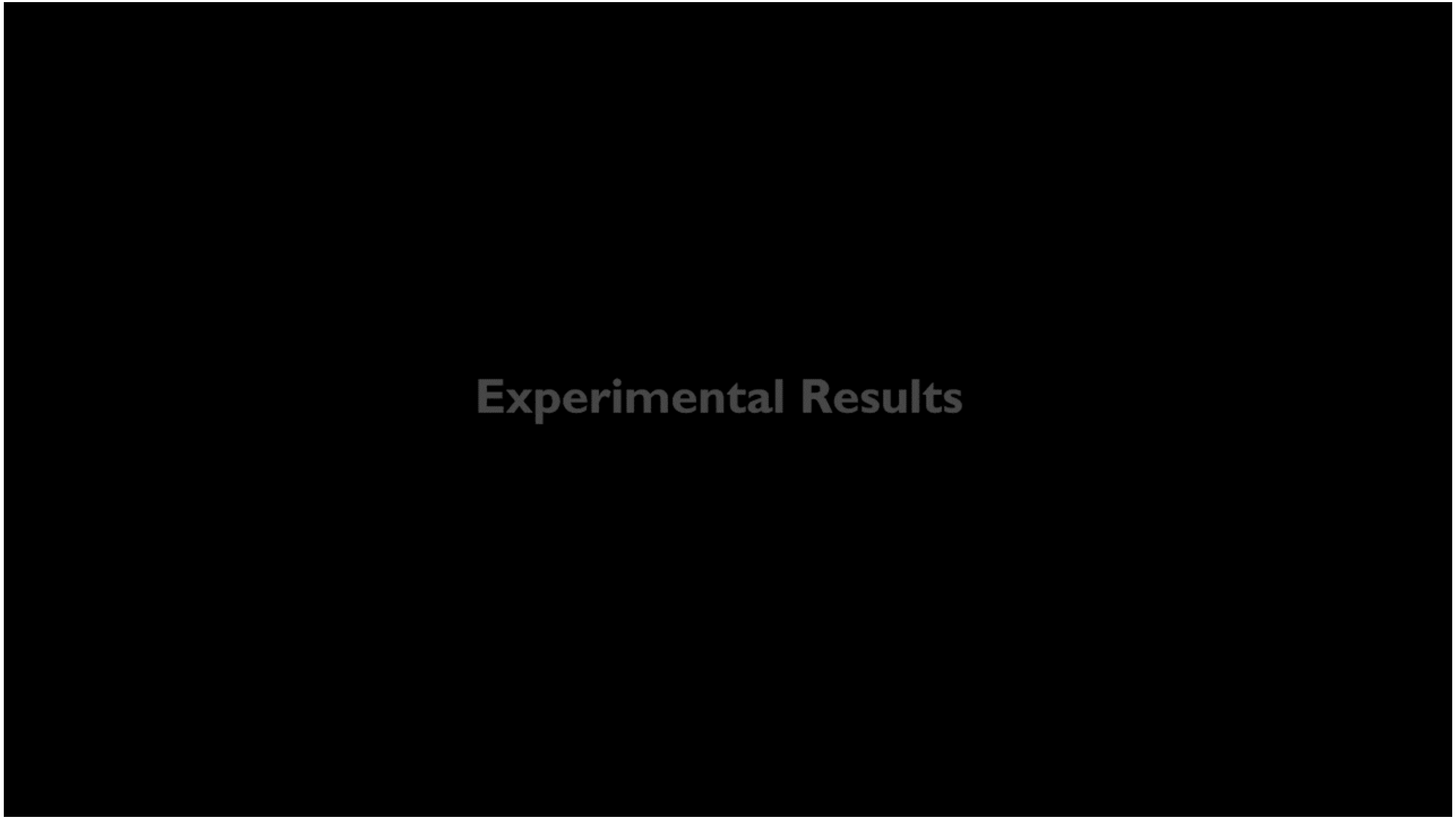


2 *Enhance Depth Map*

Raw Depth Map **Enhanced Depth Map**

Sensor-Aware Augmented Reality

Dynamic Occlusion Handling: Experimental Results



BOSCH

Sensor-Aware Augmented Reality

Robust Visual Tracking

Software

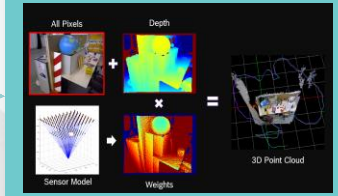
Context-Aware Visualization

Scalable & Easy Content Generation

Natural Interaction (Speech, Gesture, etc.)

Dynamic occlusion handling
ISMAR 2016[1]

- Enhance realistic depth perception



Robust visual tracking
ISMAR 2016[2]

- Improve tracking robustness and accuracy

BOSCH

Sensor-Aware Augmented Reality
Robust Visual Tracking: Motivation

Background

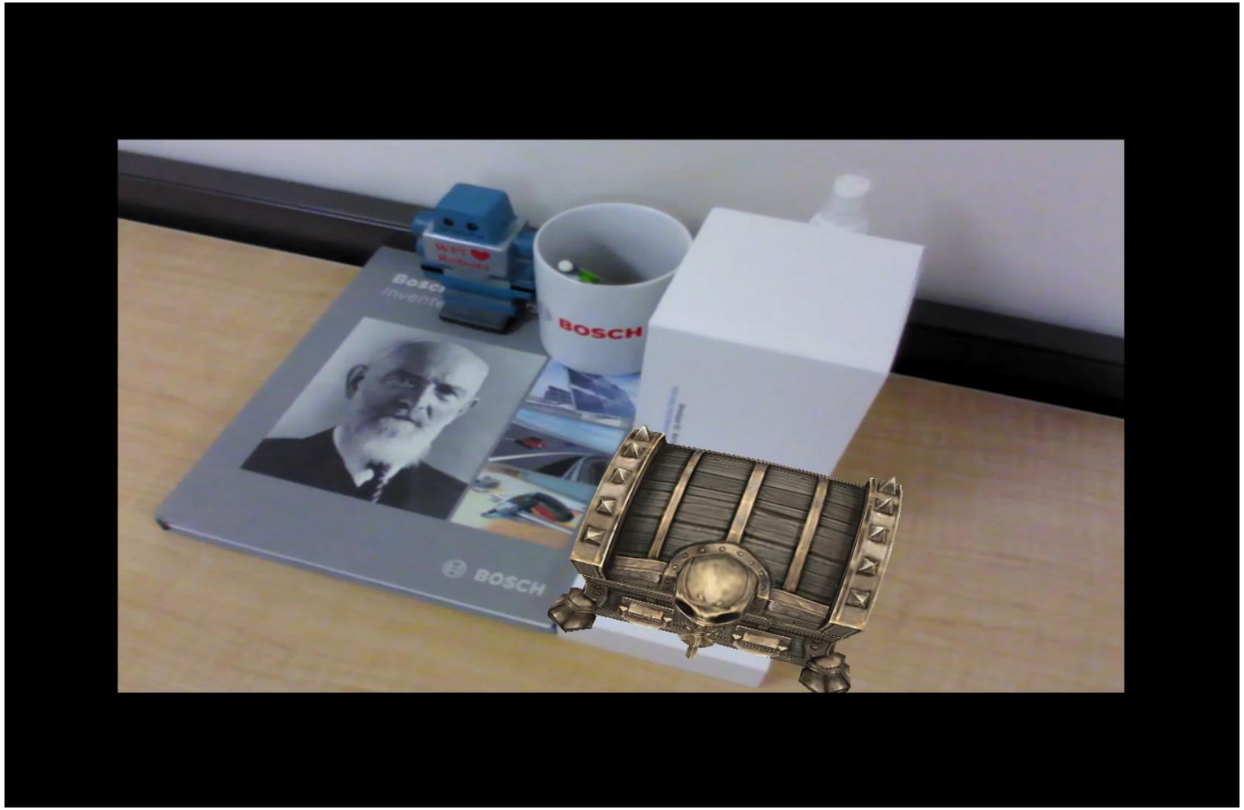
1. Visual tracking is an essential AR component
   - 6 DoF camera pose
   - Correctly place virtual objects in real world
2. Markerless Visual Tracking:
   - **Visual SLAM** (simultaneous localization and mapping)

Challenges


Textureless


Blurry image


Lighting condition change



BOSCH

Sensor-Aware Augmented Reality

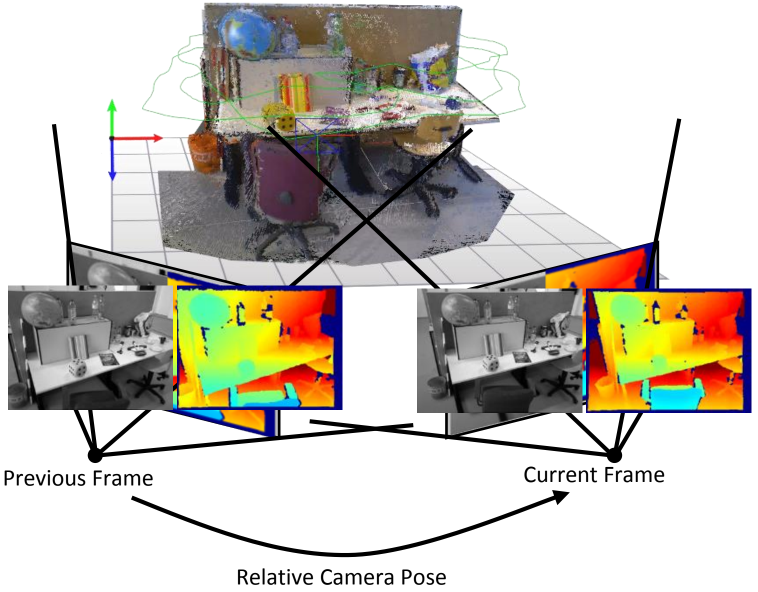
Robust Visual Tracking: Our Sensor-Aware Solution (σ-DVO)

Knowledge on RGBD Sensor

- Working well with textureless environments
- Less sensitive to lighting condition changes
- Noise of depth measurement grows quadratically as depth increases

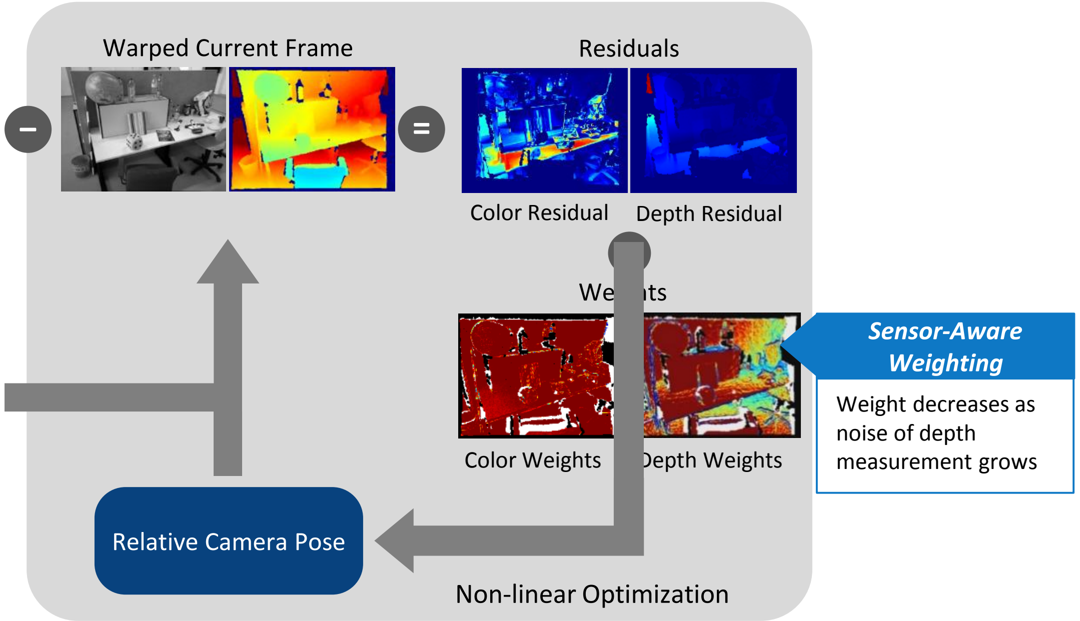
RGBD Dense Visual Odometry

- Estimate the relative pose between two given frames based on residuals (front-end of visual SLAM)
- Utilize all pixels from RGBD images
- **Incorporate sensor noise model to guide pose optimization**



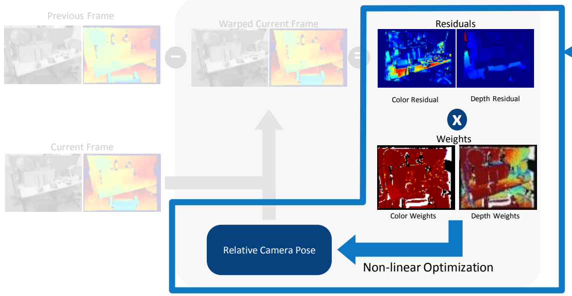
Previous Frame

Current Frame

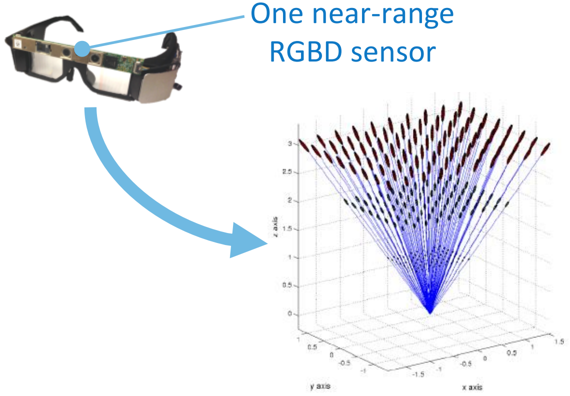


Sensor-Aware Augmented Reality

Robust Visual Tracking: Our Sensor-Aware Solution (σ-DVO)



Sensor Measurement Noise



Bayesian Framework

- For better robustness and accuracy, we formulated the **optimization** problem in a **Bayesian framework**

$$p(pose|residuals) \propto \boldsymbol{p(residuals|pose)} \cdot p(pose)$$

Early approaches

- Assume **uniform distribution** of residuals
  - ✓ All the pixels share the same weight

The state-of-the-art approach (DVO[1])

- Find an **empirical distribution** via experiments
  - ✓ Weights only depends on residuals

Our sensor-aware approach (σ-DVO)

- Explore the source of residuals, especially based on sensor characteristics
- Develop a **sensor noise model** to generate distribution of residuals
  - ✓ Decrease weights of pixels with either noisy sensor measurement or high residuals
- Easily incorporate sensor-specific noise model for different sensors to customize pose optimization for best performance

[1] Christian Kerl, Jürgen Sturm, and Daniel Cremers. "Dense visual SLAM for RGB-D cameras." 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2013.

Sensor-Aware Augmented Reality

Robust Visual Tracking: Experimental Results

Visual Odometry

ATE: Absolute Tracking Error [m], RPE: Relative Pose Error [m/s]

Dataset	DVO[1]		Our approach (σ-DVO)	
	ATE	RPE	ATE	RPE
fr1/360	0.415	0.153	**0.229**	**0.110**
fr1/desk	0.109	0.048	**0.067**	**0.039**
fr1/desk2	0.261	0.074	**0.088**	**0.065**
fr1/floor	0.242	0.070	**0.226**	**0.053**
fr1/room	0.459	0.092	**0.314**	**0.063**
fr1/rpy	0.216	0.065	**0.072**	**0.046**
fr1/xyz	0.102	0.05	**0.052**	**0.036**
fr2/desk	0.561	0.038	**0.184**	**0.016**
fr2/large	4.370	0.240	**0.724**	**0.134**
fr2/rpy	0.501	0.039	**0.188**	**0.012**
fr2/xyz	0.497	0.030	**0.188**	**0.010**
fr3/office	0.485	0.044	**0.164**	**0.014**
average	0.684	0.067	**0.208**	**0.050**

70% 25%

*Visual SLAM **

Absolute Tracking Error [m]

Dataset	RGB-D SLAM[2]	MRSMap[3]	Kintinuous[4]	ElasticFusion[5]	DVO SLAM[1]	Our SLAM approach
r1/desk	0.023	0.043	0.037	0.020	0.021	**0.019**
fr2/xyz	**0.008**	0.020	0.029	0.011	0.018	0.018
fr3/office	0.032	0.042	0.030	0.017	0.035	**0.015**
fr1/360	0.079	0.069	-	-	0.083	**0.061**

25%

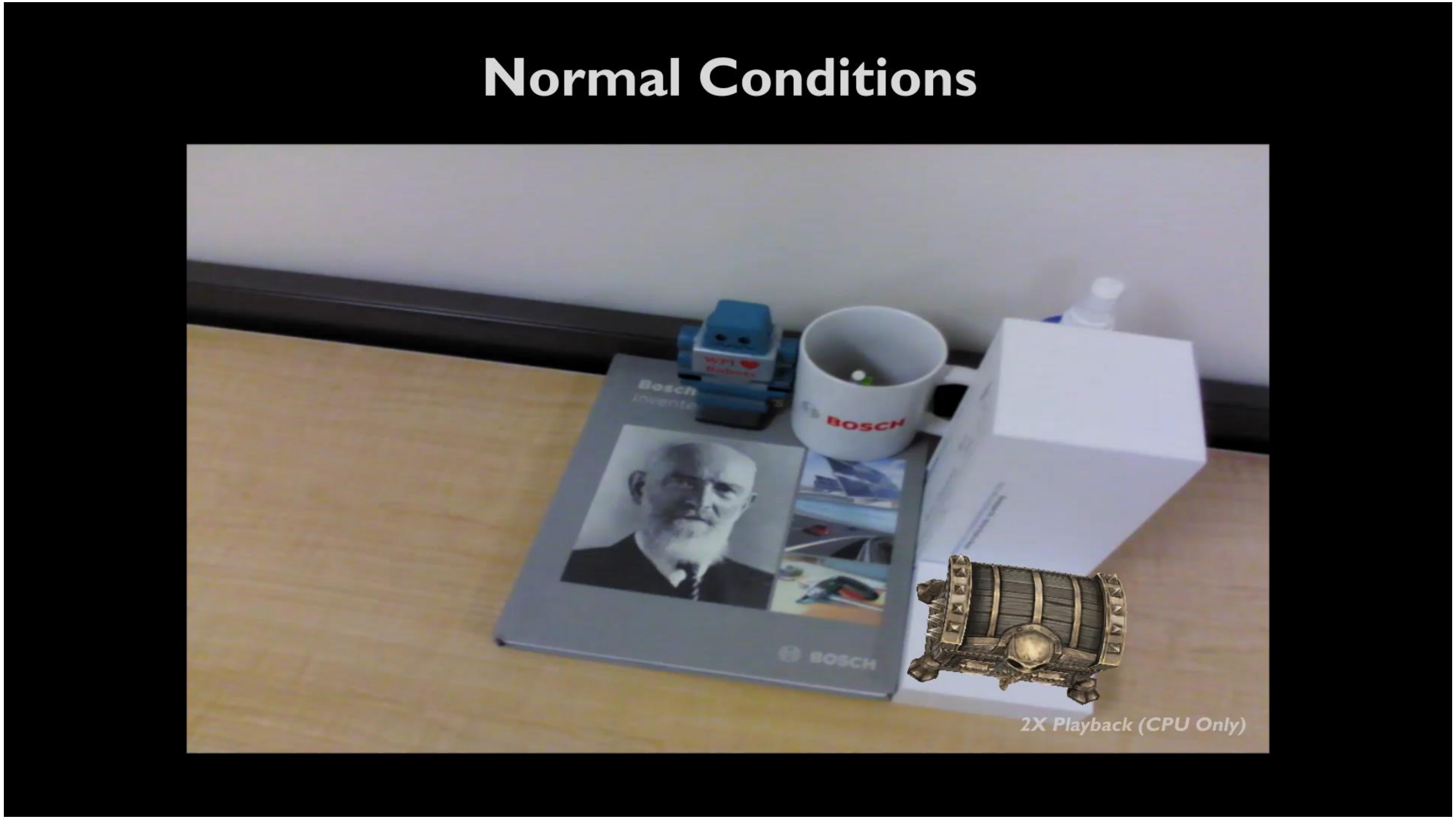
* σ-DVO is extended to σ-DVO SLAM by combining the front end (visual odometry) with the backend (pose-graph optimization)

- Our σ-DVO outperforms DVO significantly in all the datasets. On average, **70% reduction** in ATE and **25% reduction** in RPE
- σ-DVO SLAM outperforms the state-of-the-art SLAM algorithms in most of the RGB-D datasets. On average **25% reduction** in ATE

[1] Kerl, et al. "Dense visual SLAM for RGB-D cameras." 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems.
[2] Endres, et al. "An evaluation of the RGB-D SLAM system." Robotics and Automation (ICRA), 2012.
[3] Stückler, et al. "Model Learning and Real-Time Tracking using Multi-Resolution Surfel Maps". AAAI, 2012.
[4] Whelan, et al. "Kintinuous: Spatially extended kinectfusion." Proc. Workshop RGB-D, Adv. Reason. Depth Cameras, 2012.
[5] Whelan, et al. "ElasticFusion: Dense SLAM without a pose graph." Proc. Robotics: Science and Systems, 2015.

Sensor-Aware Augmented Reality

Robust Visual Tracking: Experimental Results



BOSCH

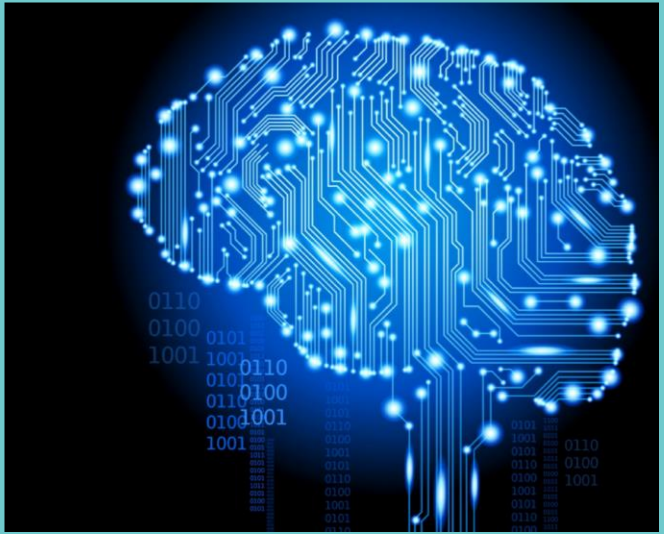
Sensor-Aware Augmented Reality

Deep Learning for Augmented Reality?

Context-Aware Visualization

Scalable & Easy Content Generation

Natural Interaction (Speech, Gesture, etc.)



- **Require semantic understanding of the environments & context**
- **Modern AI technologies, e.g., Deep Learning, could be effective approaches**

Sensor-Aware Augmented Reality
Summary and Outlook

1	The three "I"s (Intuitive, Interactive, Intelligent) are key success factors of Human Machine Interaction (HMI) solutions.
2	Sensor-aware approaches that leverage sensor knowledge and machine learning are effective to address real-world HMI challenges.
3	Using the right AI technology to address the right problem. Deep Learning, could be effective for core AR solutions.

Thank You

We’re looking for good research scientists and interns!

(Send CV to rbvisualcomp@gmail.com)