Augmented Reality : Modern Trends and Technical Solutions

Ievgen Gorovyi, PhD,
founder & CEO @ It-Jim,
signal and image processing

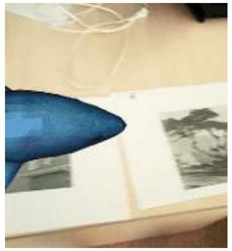
AI UKRAINE 18
V INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND DATA SCIENCE APPLICATIONS

It-Jim: Key Interests

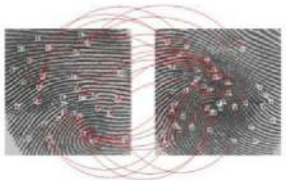
CEO & founder, PhD, signal and image processing
author of >40 publications

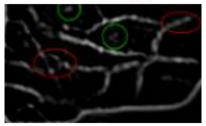
Indoor Navigation

Augmented reality



Pattern recognition

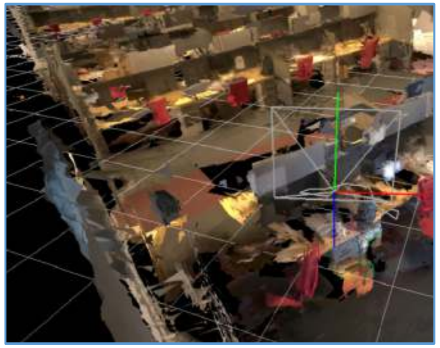




Object tracking

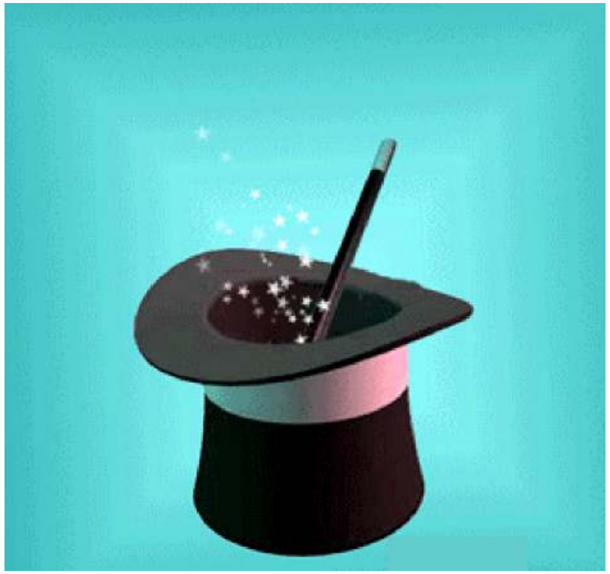
OUTLINE

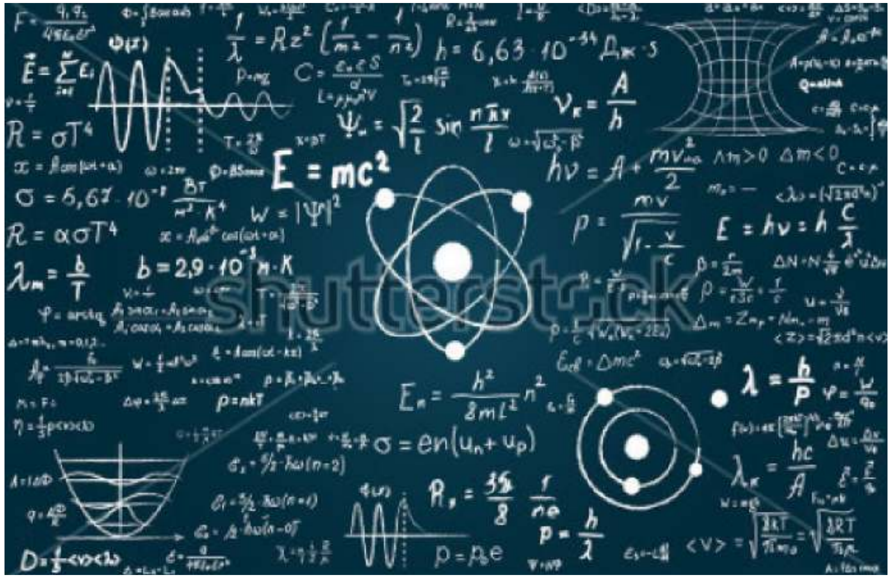
- **Introduction**
  - AR history
  - Examples, applications, market
  - Software and hardware
- **AR: What's inside?**
  - AR and camera pose
  - AR markers
  - Detection
  - Tracking
  - Image retrieval
  - SLAM for AR
- **Mobile AR in Production**
  - AR user experience
  - SDK infrastructure
  - Real examples
- **Conclusions**



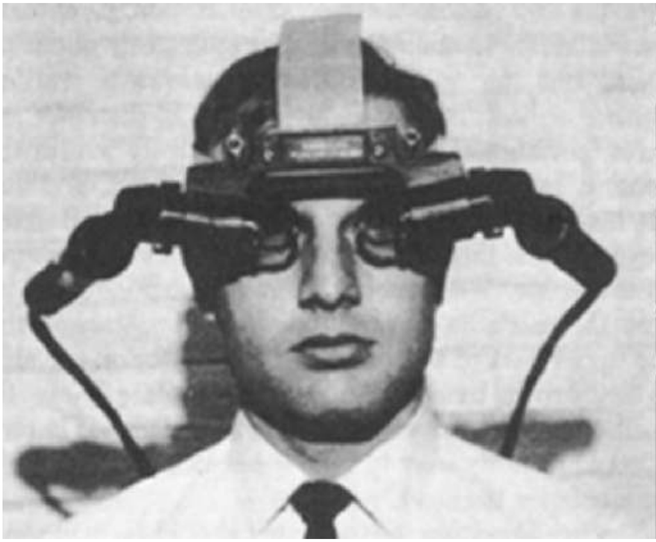
Augmented Reality: Intro

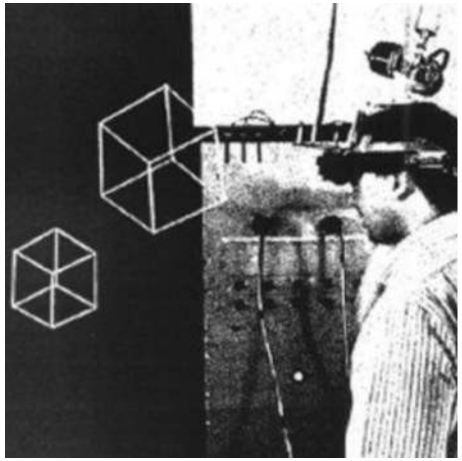
Augmented Reality: What’s It?



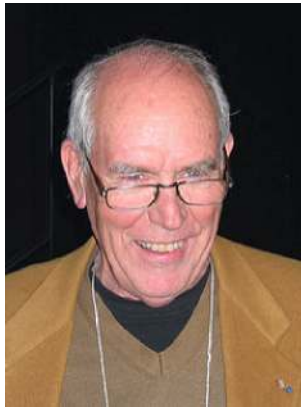


AR History: First Appearance

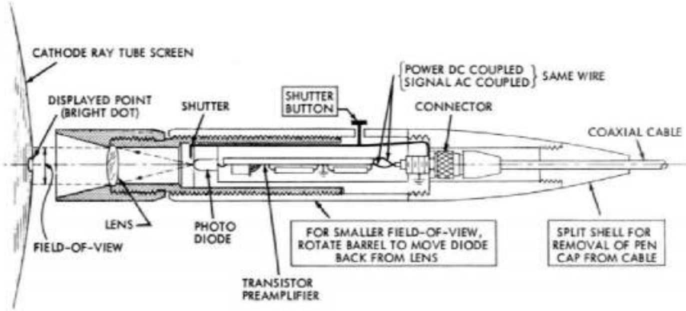





1968 1st head-mounted display – Sword of Damocles (Ivan Sutherland)



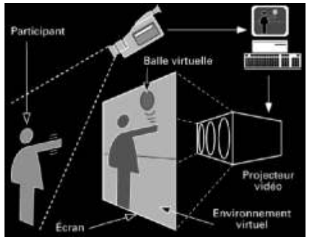
Sutherland - a father of computer graphics

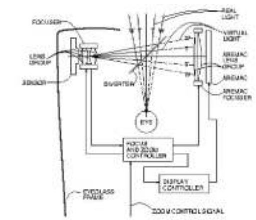



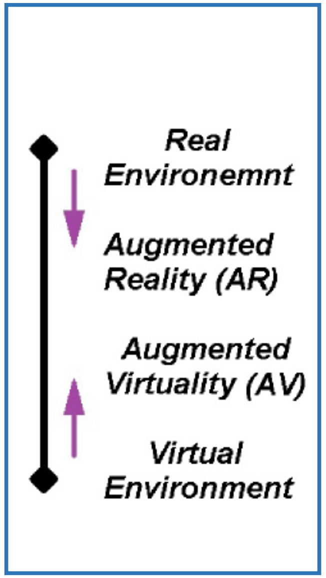
Sketchpad (first light-pen)*

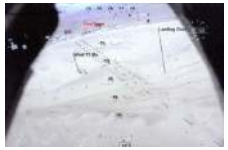
Sketchpad:A MAN-MACHINE GRAPHICAL COMMUNICATION SYSTEM, I. Sutherland, PhD thesis, MIT, 1963.

AR History: First Steps









1968
Sutherland

1978
VideoPlace
Myron Krueger

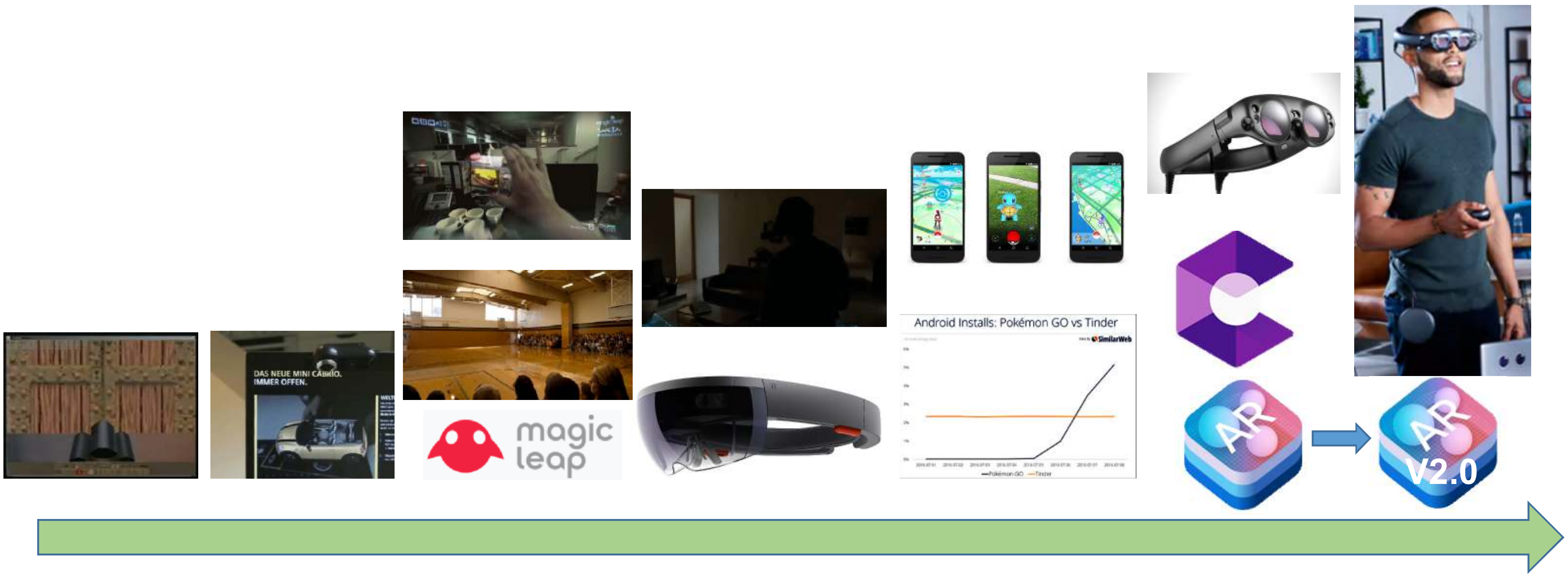
1980
EyeTap
(Steve Mann)

1990
Tom Caudell
(Boeing)

1994
Reality continuum
(Milgram)

1999
AR in NASA
ARToolKit

AR History: Modern Times



2000
AR Quake

2008
1st commercial AR app
BMW mini

2010
Magic Leap

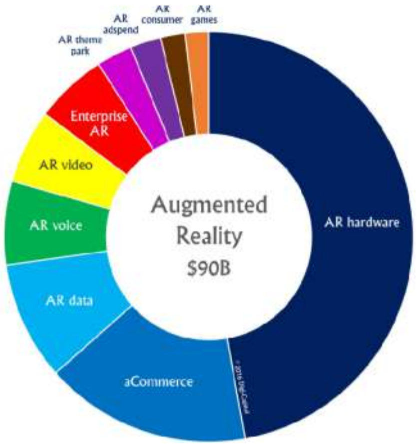
2015
MS Hololens

2016
Pokemon Go

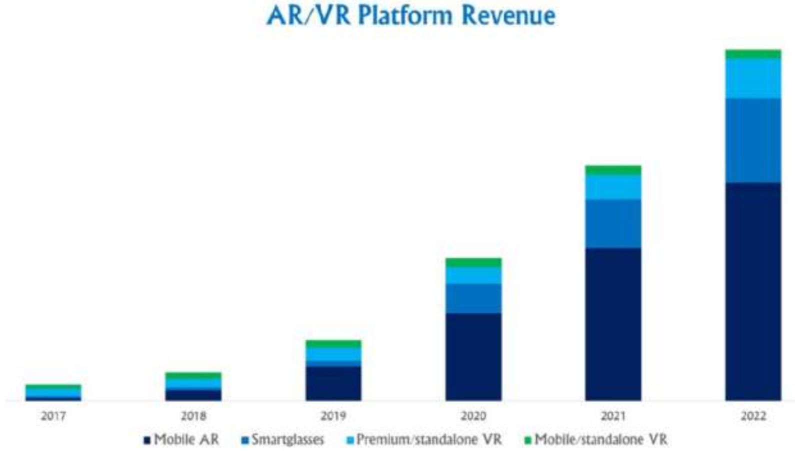
2017
ARKit,
ARCore

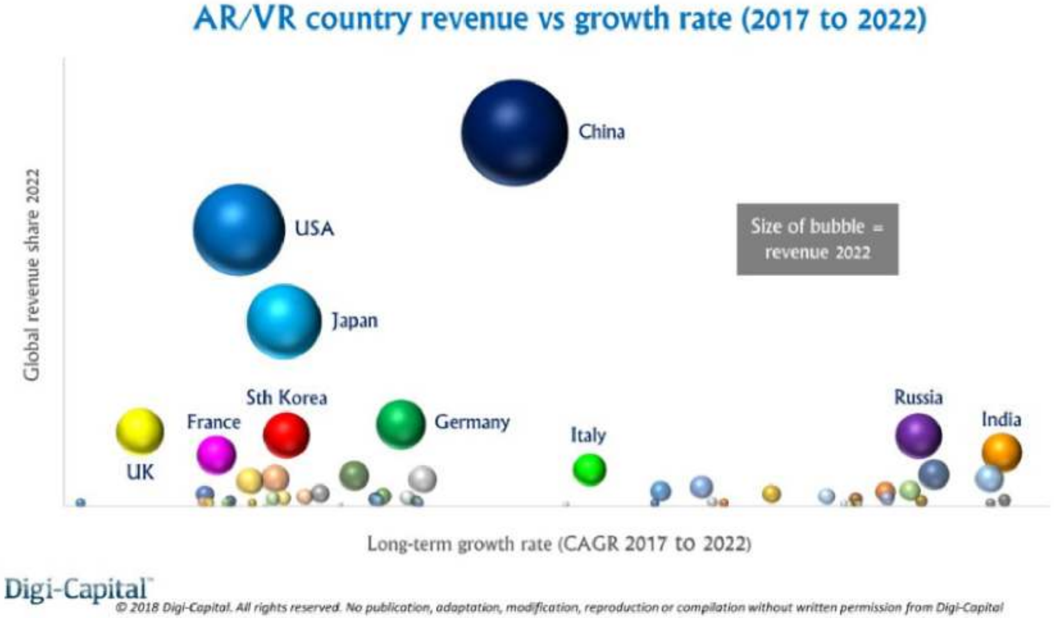
2018
ARKit V2.0
Magic Leap
One

Mobile AR Market



Market share by 2020 (Digi-Capital)





Mobile AR: Software Landscape

	country	platforms	features
vuforia	UK	iOS	Complex object recognition, smart glasses support, 0.5M community!
kudan AUGMENTED REALITY	UK, Japan	iOS	Versatile infrastructure for mobile, drones and automotive
wikitude	Austria	iOS	Unity, smart glasses support
MAXST	South Korea		Unity, Various trackers
	USA	iOS	Unity, VIO, Multiplayer
	USA	iOS	Motion tracking, environmental understanding, Light estimation

AR Use Cases

AR in sports (simplest)

AR in retail

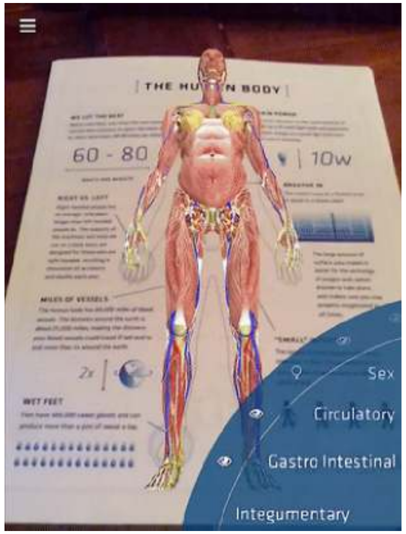
Location-based AR

AR in real estate

Furniture AR

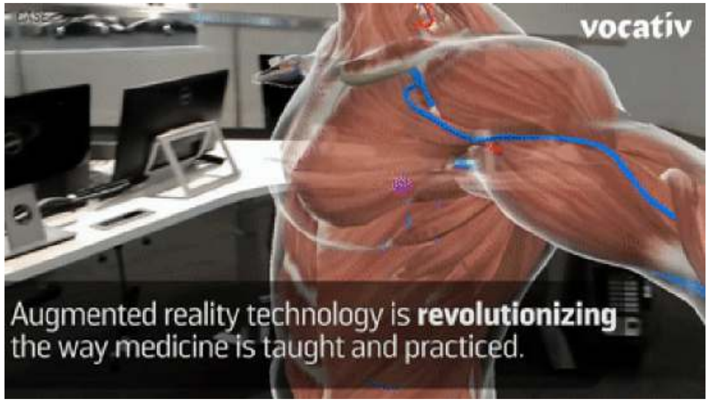
AR in Education

Basic education



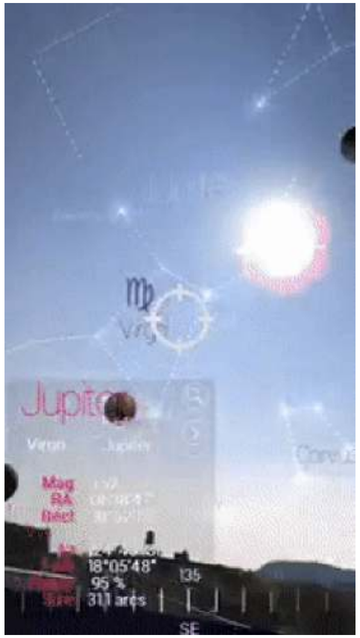
DAQRI anatomy

DAQRI elements 4D



Medical education with Hololens

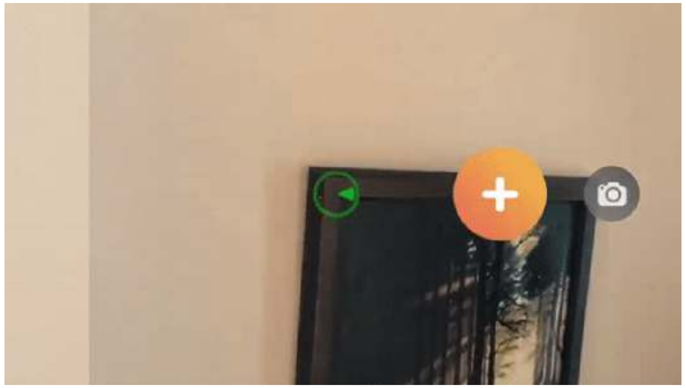
AR in Entertainment



AR astronomy (ARKit)

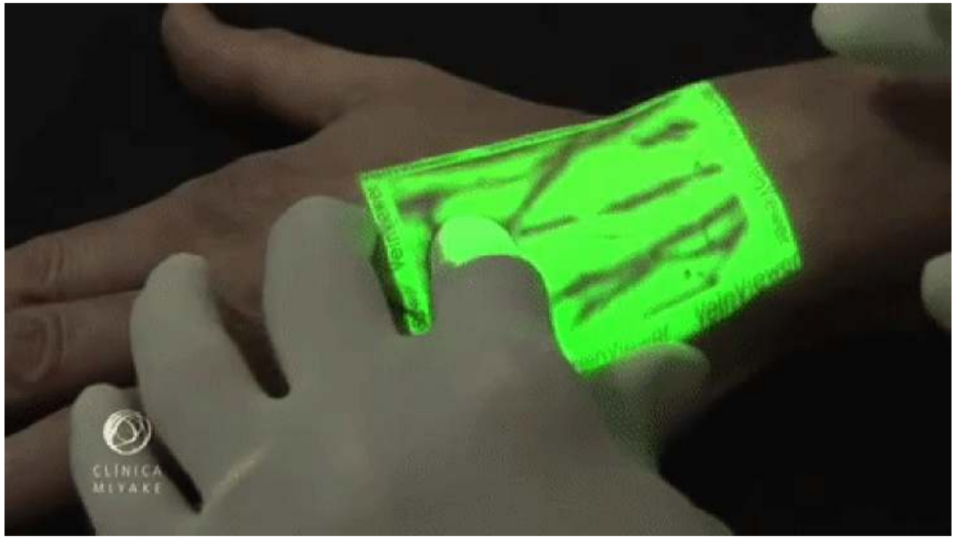
AR highway (ARCore)

AR portals (ARKit)



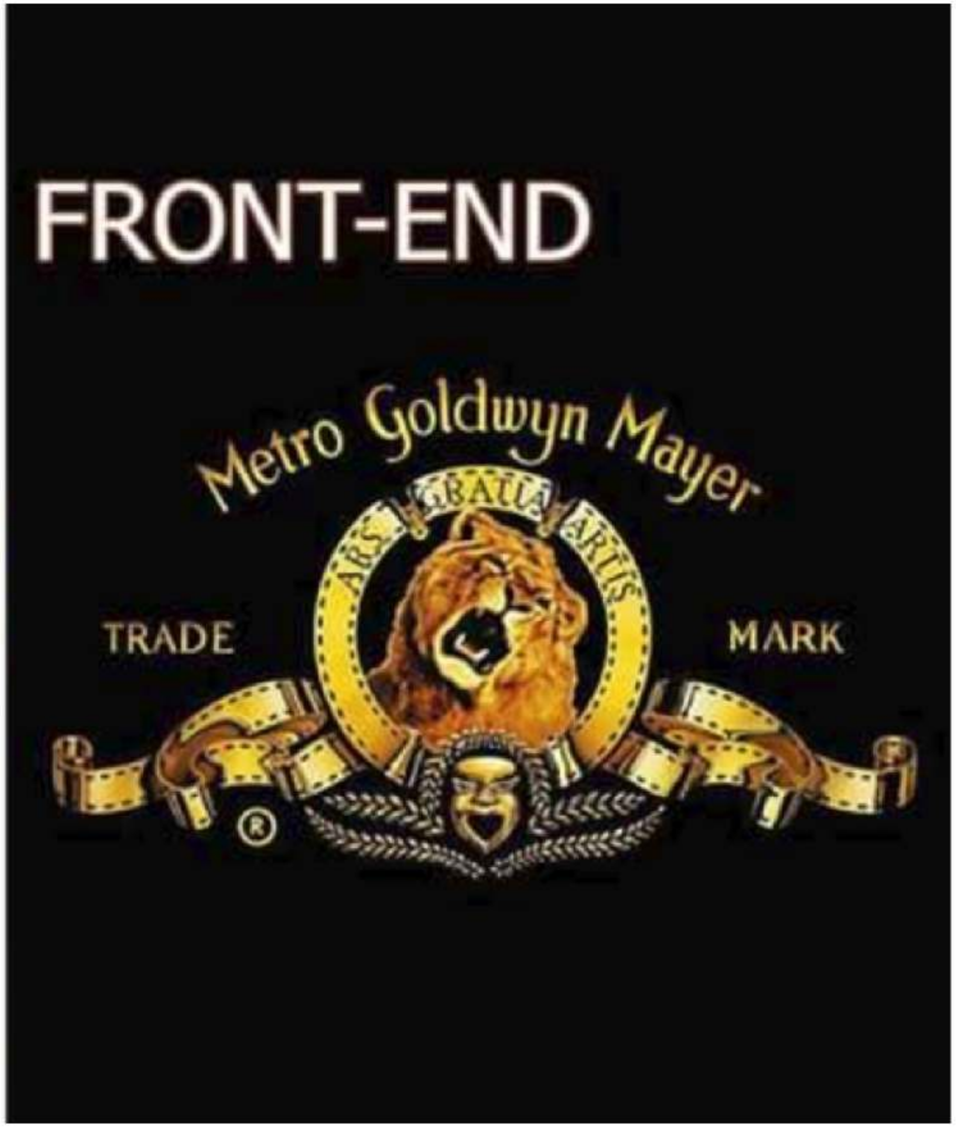
AR measure (ARKit)

AR in Medicine



Vein viewer (NIR+Optical)
Clinica Miyake

AR surgery (X-ray + optical)
Augmedics


FRONT-END
Metro Goldwyn Mayer
ARS GRATIA ARTIS
TRADE
MARK

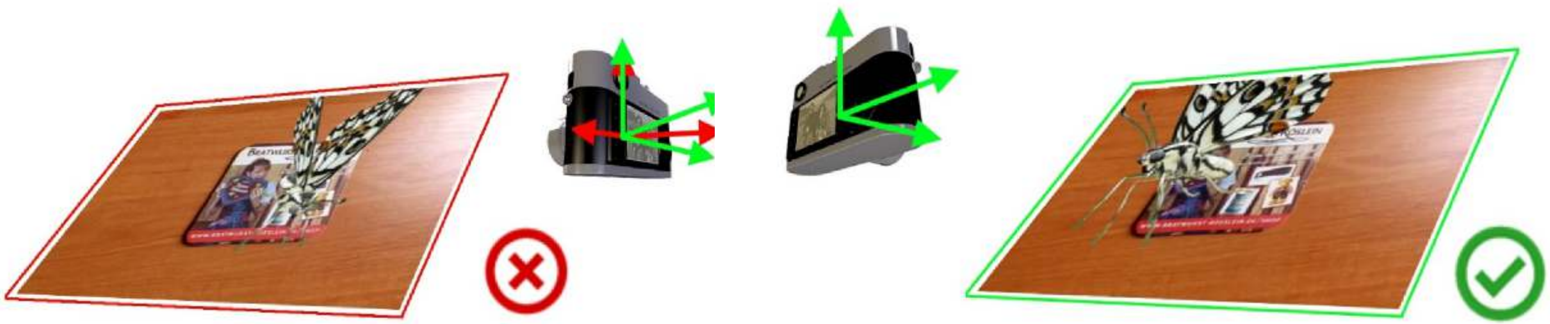

BACK-END

AR: Algorithms and Solutions

- camera pose and marker detection
- image retrieval
- tracking
- markerless and SLAM

AR Magic: What Do We Need?

6 DOF



Wrong augmentation

Correct augmentation

Marker-based vs Markerless

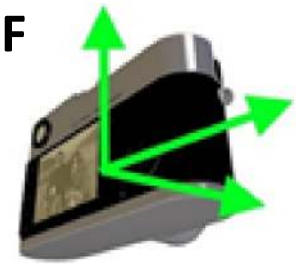
2

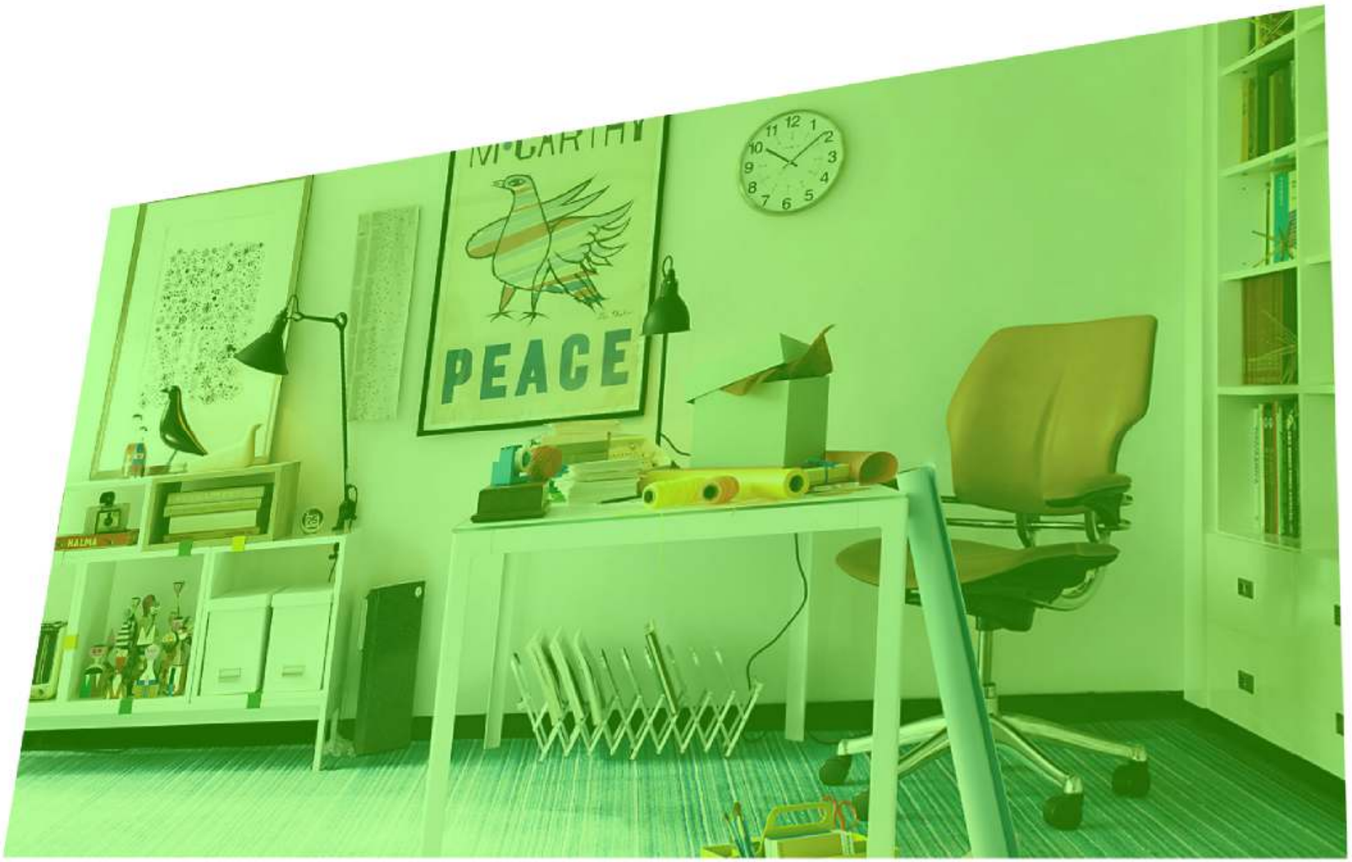
Marker-based

Markerless

$\Delta R|t$

6 DOF

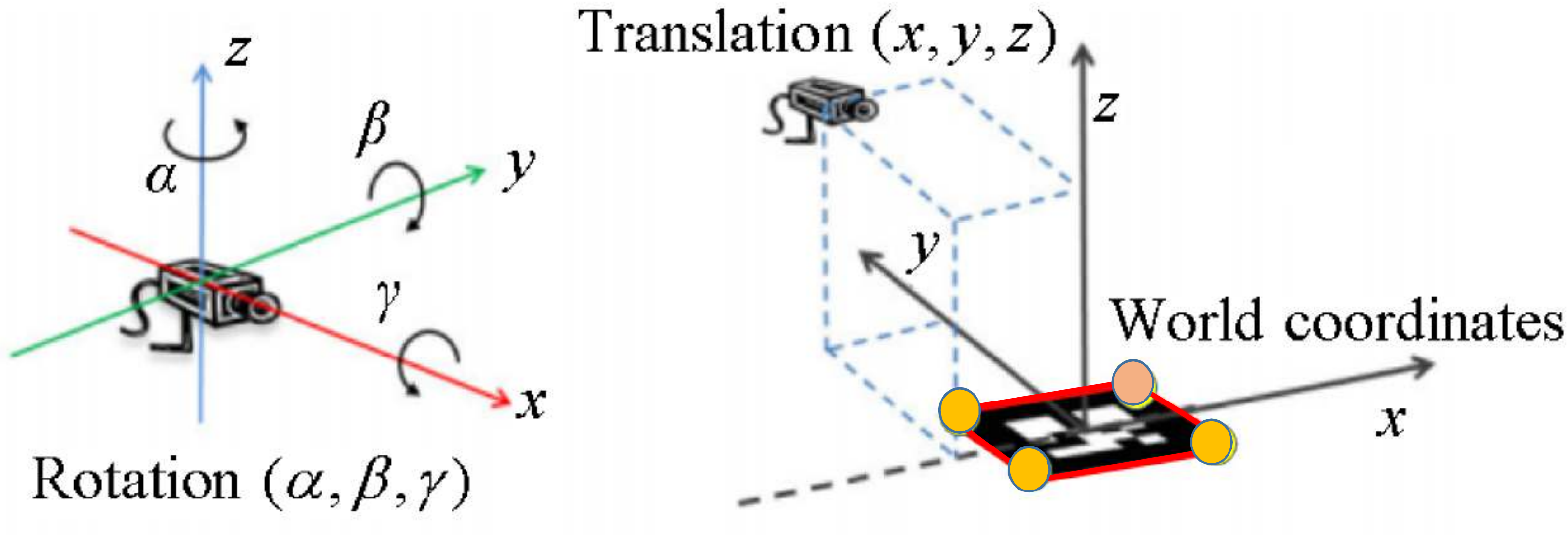


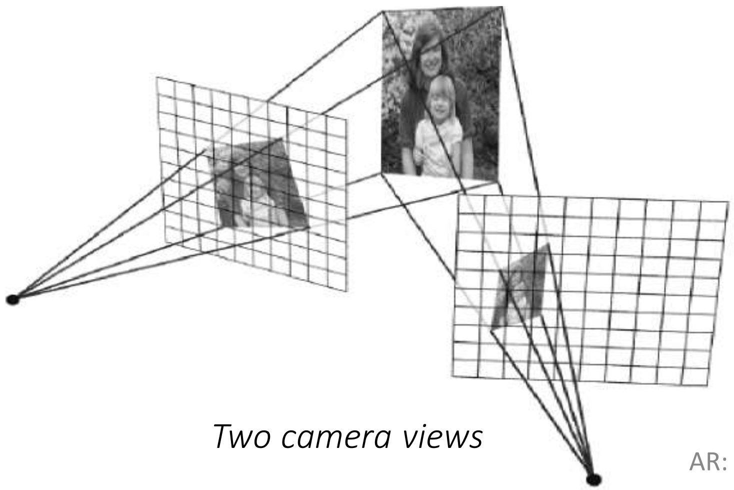


- Easy to detect
- No need for markers
- Textureless scene is not a problem

- Occlusions
- High complexity
- Works at limited distances
- Fails for textureless

Marker-based AR and Camera Pose



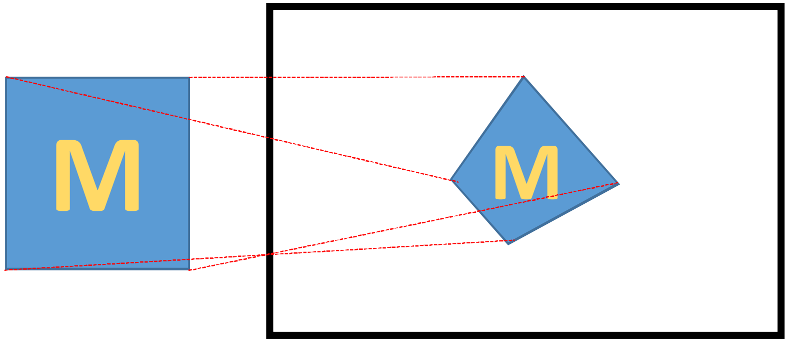


Two camera views

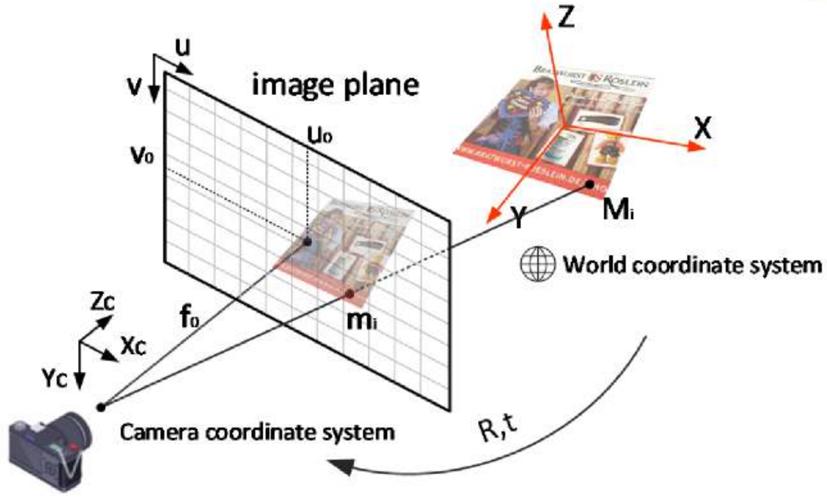
$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} \cong \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

Homography

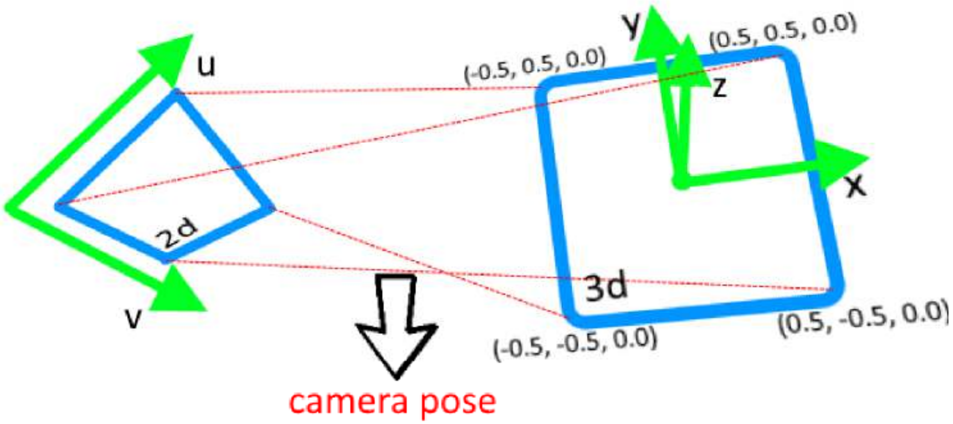
Camera Pose Estimation Steps



Homography estimation



Projective geometry



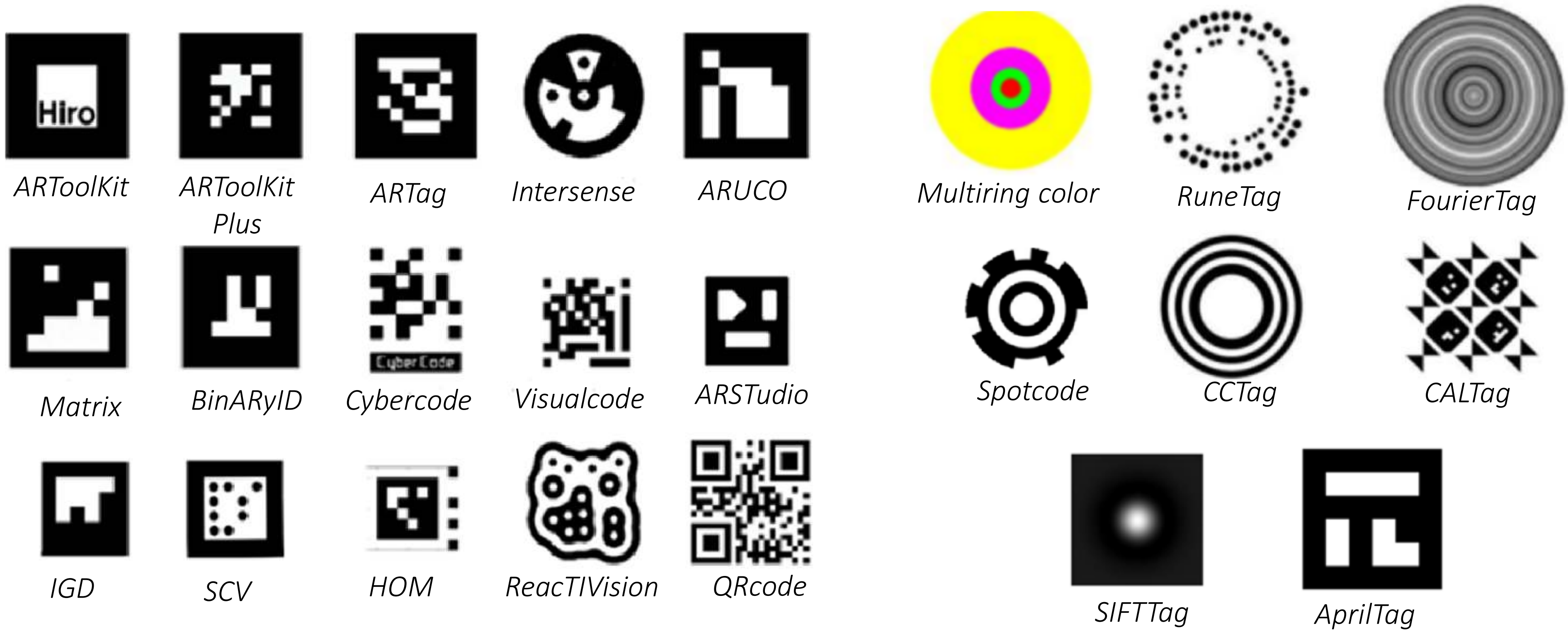
Camera pose extraction (2D-3D)

K $\qquad$ $R|t$

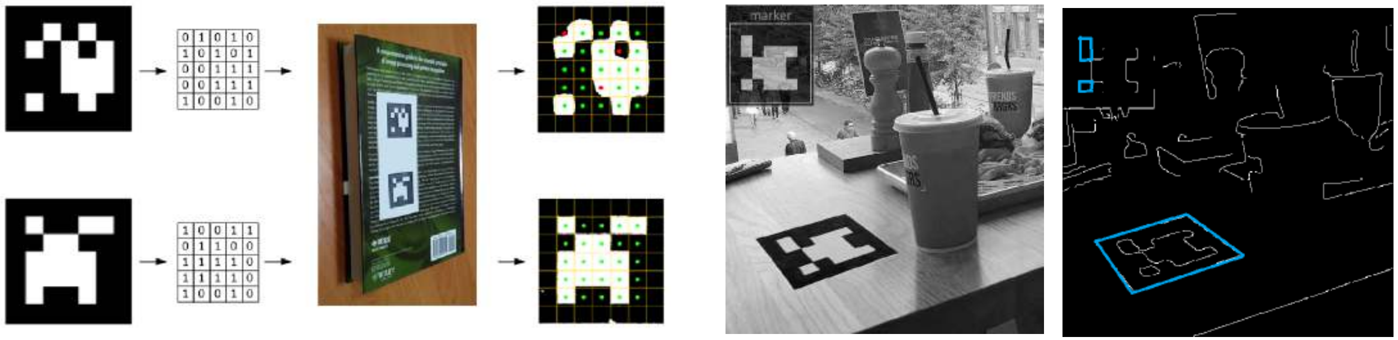
$$\left(\ \right) = \begin{bmatrix} f_x & 0 & u_x \\ 0 & f_y & v_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{32} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ \end{bmatrix}$$

Fiducial Markers Diversity

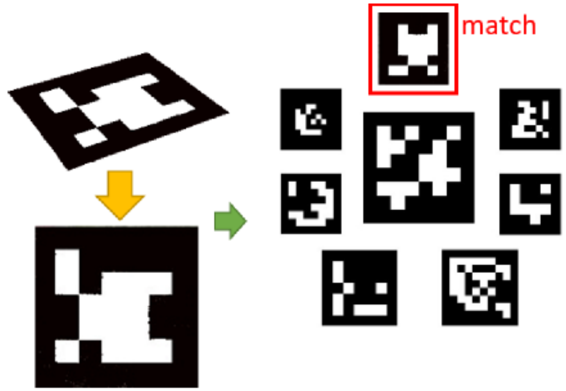
Template-based markers



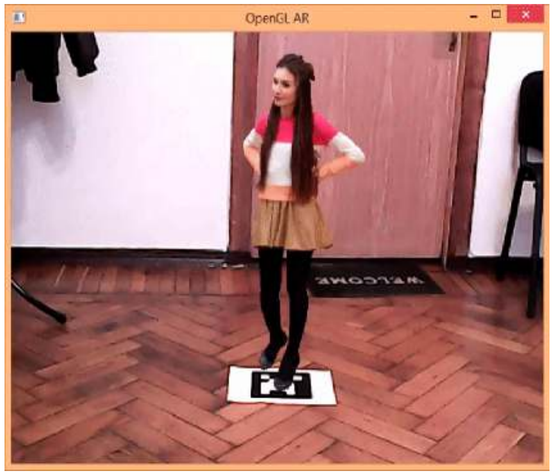
ArUCO Markers



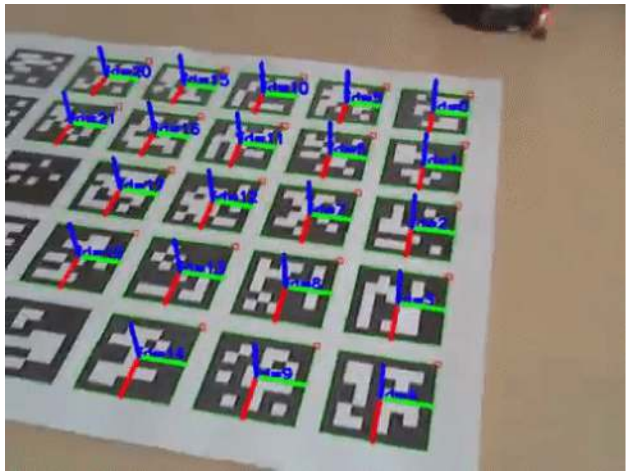
Binary marker detection



Binary marker recognition



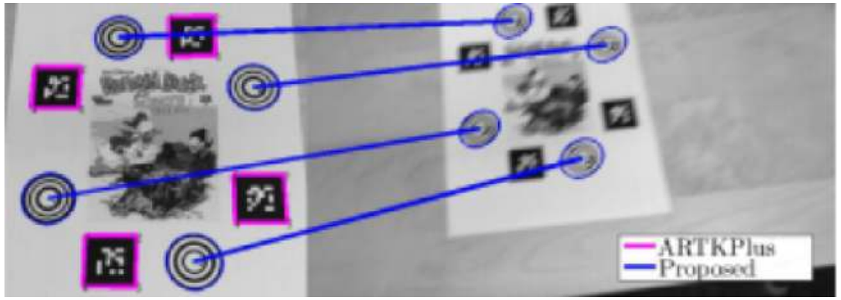
Fiducial marker tracking



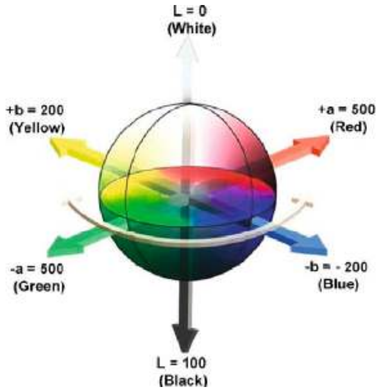
Multiple marker tracking

S.Garrido-Jurado et. al. Generation of fiducial marker dictionaries using Mixed Integer Linear Programming.Pattern Recognition. Volume 51, March 2016, Pages 481-491

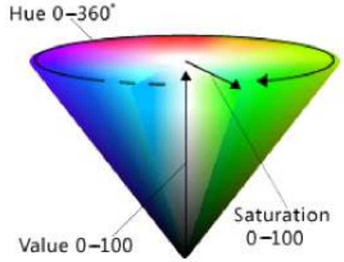
Fiducial Markers: Interesting Examples



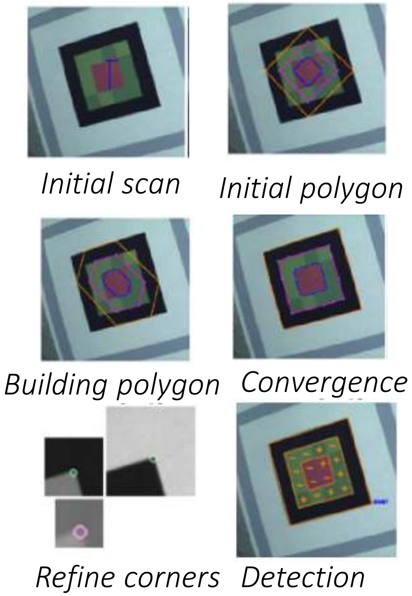
Resistance to motion blur

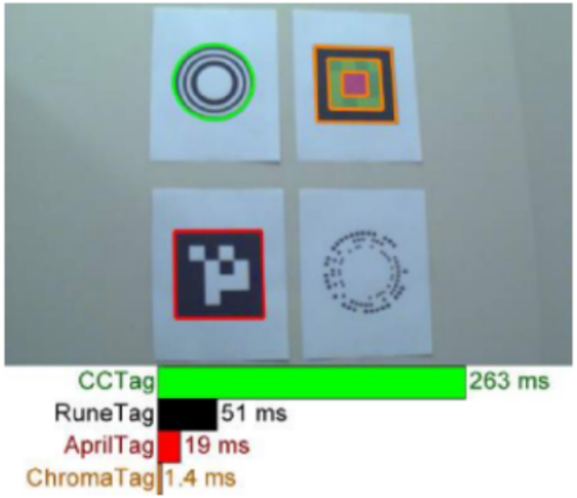


LAB color space



HSV color space





ChromaTag: A Colored Marker and Fast Detection Algorithm. Joseph DeGol,Timothy Bretl,Derek Hoiem, ICCV-2017, pp.1472-1481.

Image Markers

Film billboard

Book covers

Newspapers

Menu, food, etc.

Image Description: First Approaches

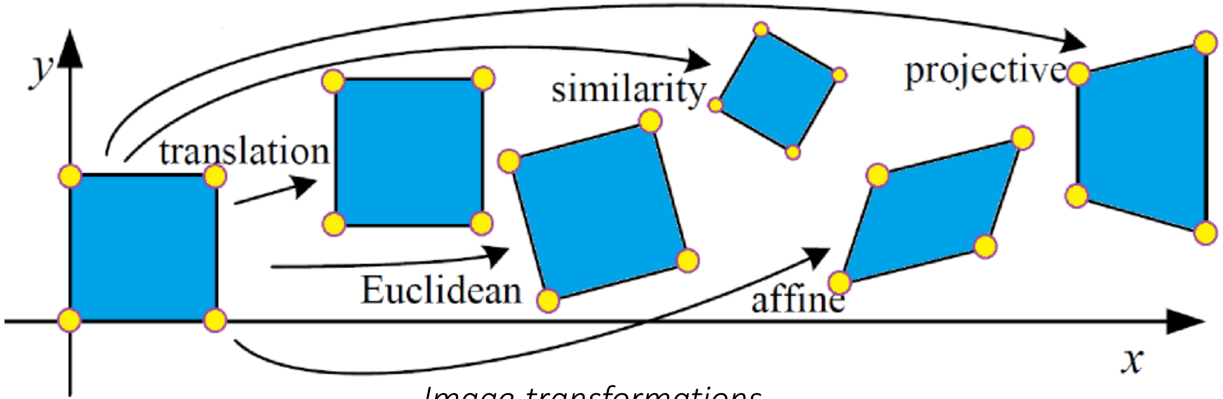
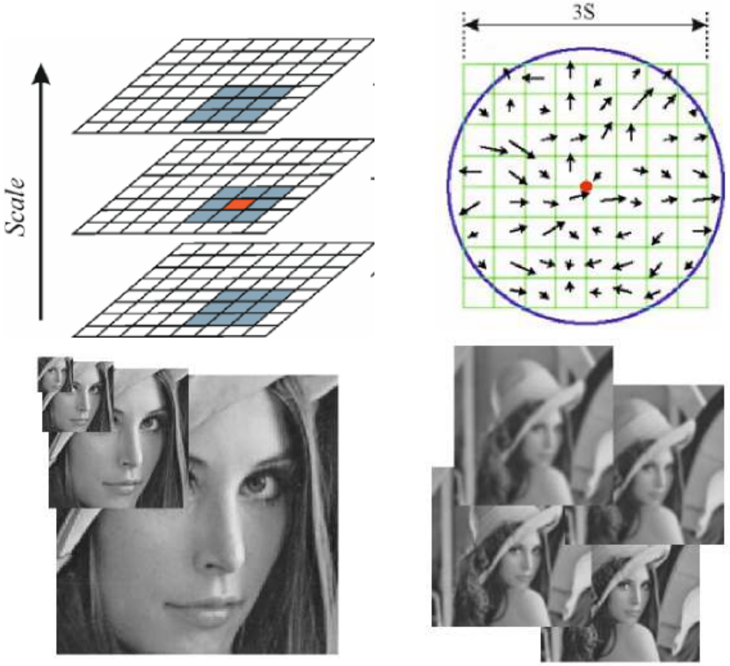
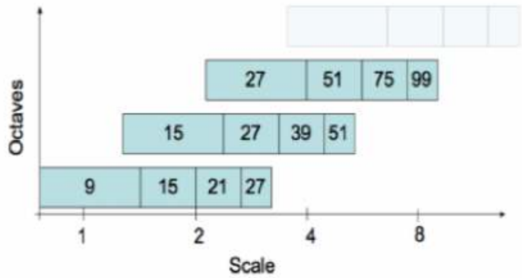


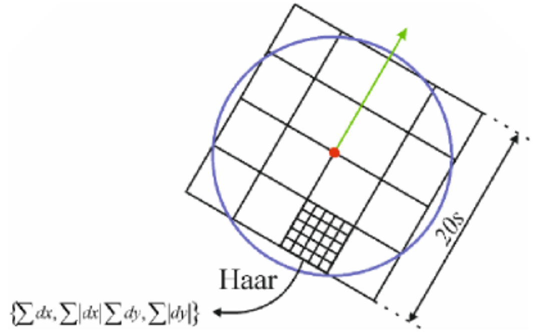
Image transformations



SIFT (1999)



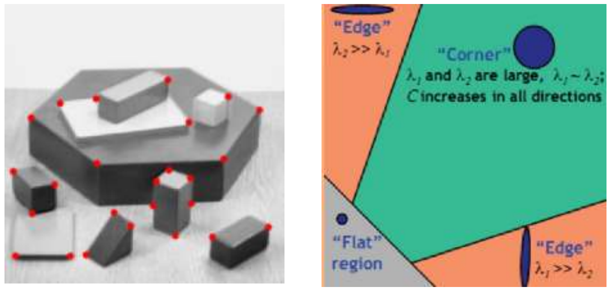
scale space with box filters



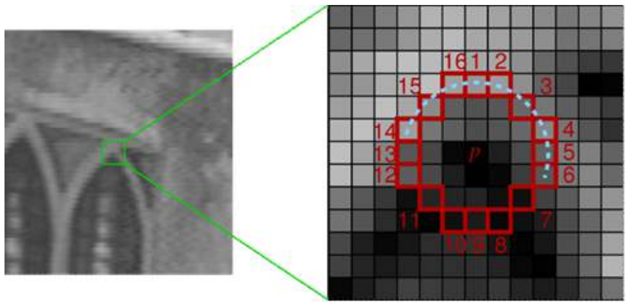
SURF descriptor (2006)

Local Feature Descriptors

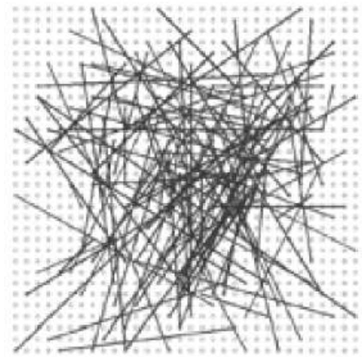
Keypoint detection → **Keypoint description** → **Matching**



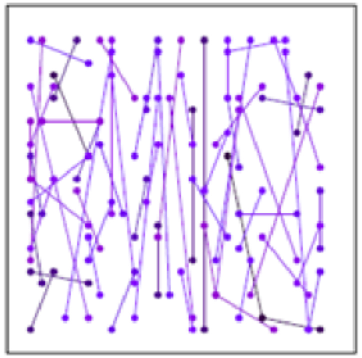
Corner detection



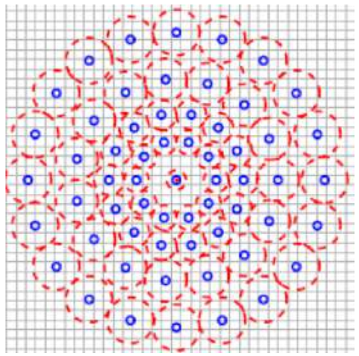
FAST



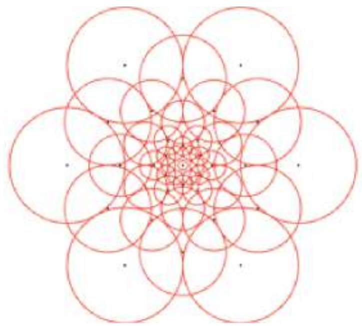
BRIEF



ORB



BRISK

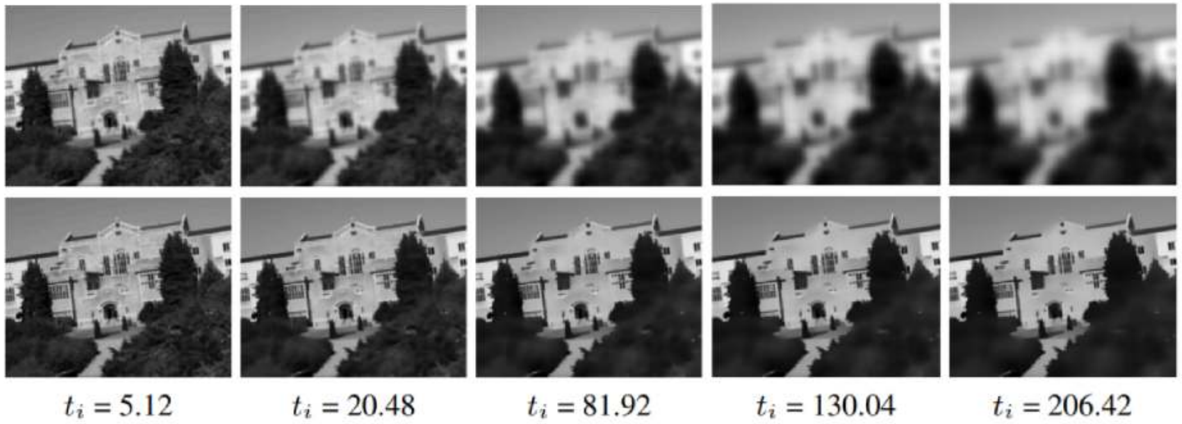


FREAK

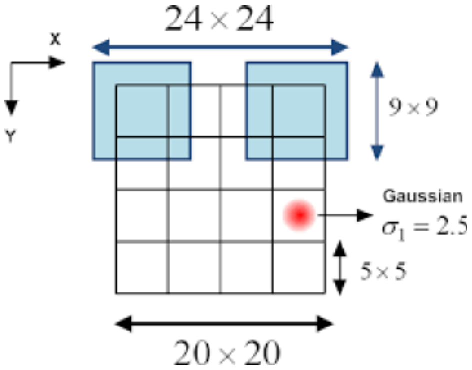
KAZE and A-KAZE

linear diffusion (SIFT) - slow

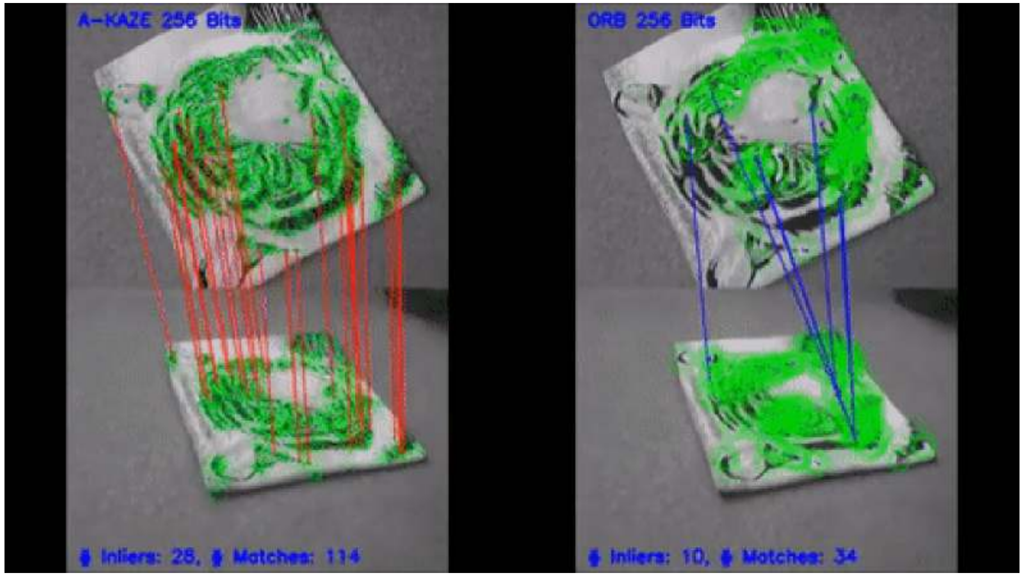
non-linear diffusion (KAZE)



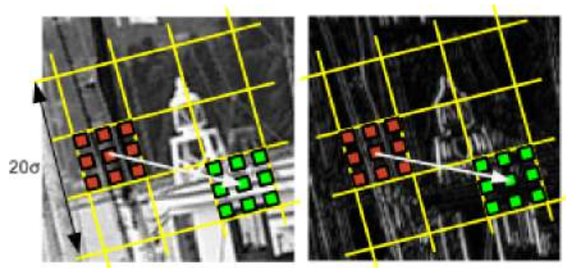
Scale-space comparison



M-SURF (KAZE)

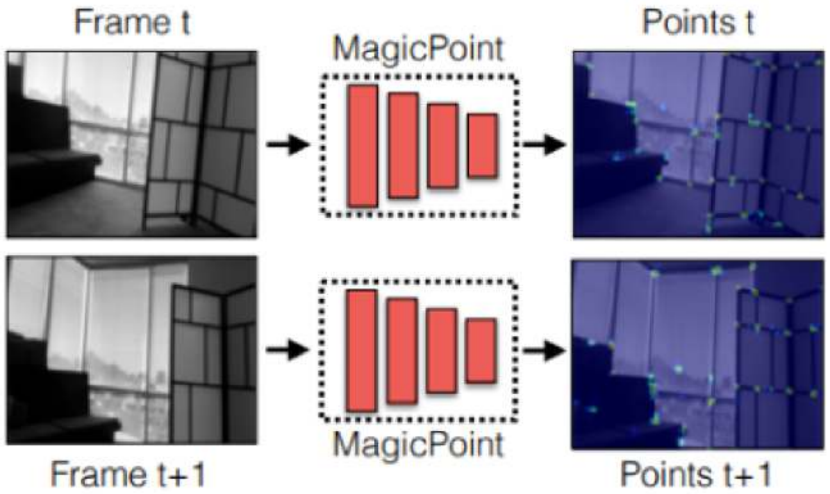


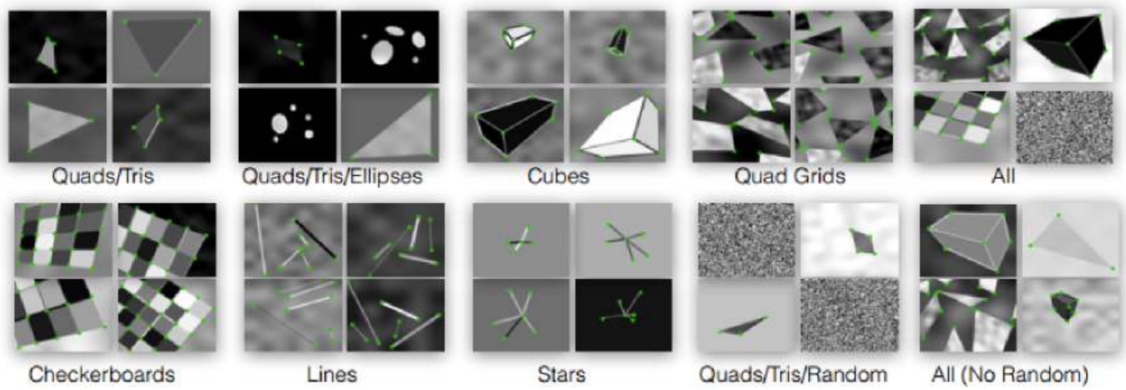
A-KAZE (2013) vs ORB (2011)



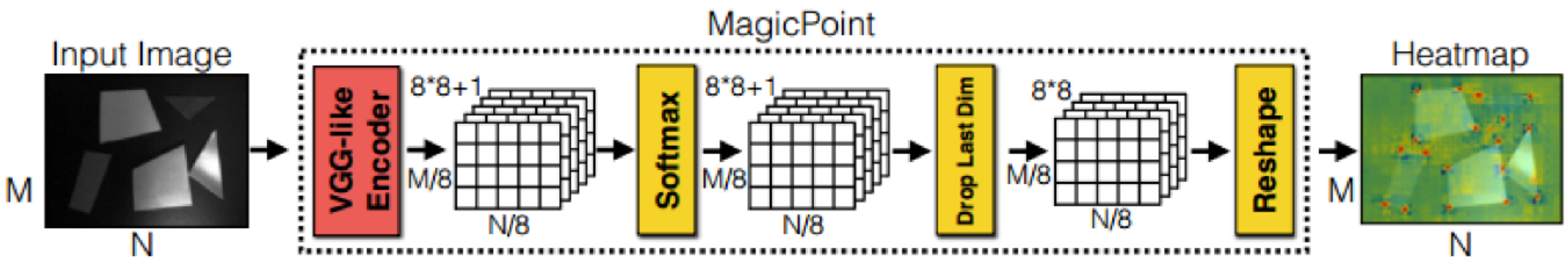
M-LDB (A-KAZE)

MagicPoints (Magic Leap)





Synthetic data for training



Magic point detection

D. DeTonem T. Malisiewicz, A.Rabinovich - ”Toward Geometric Deep SLAM” / https://arxiv.org/pdf/1707.07410.pdf

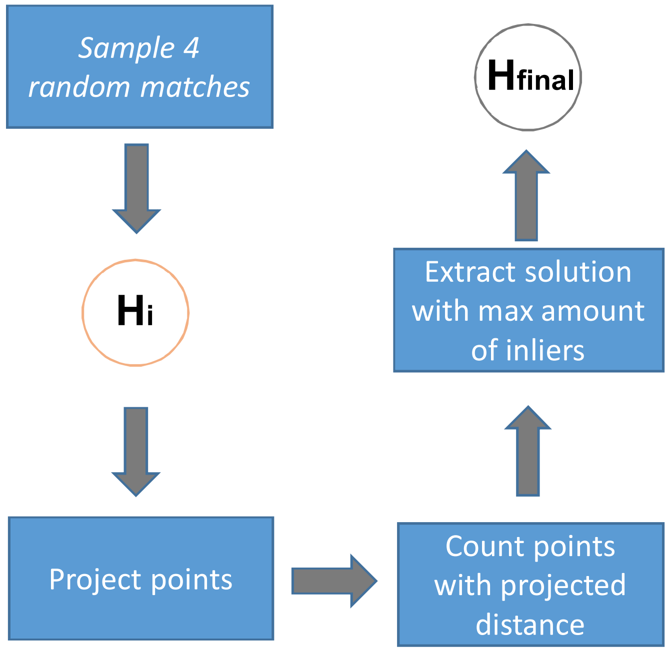
Matching

Matches before filtering

Incorrect pose

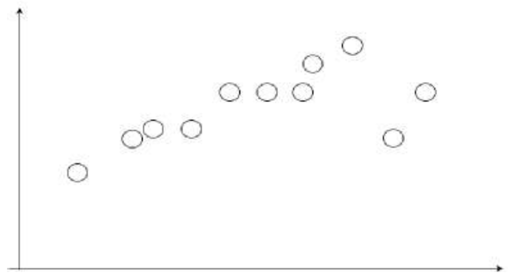
Matches after filtering

Correct pose estimation



Random sample consensus scheme

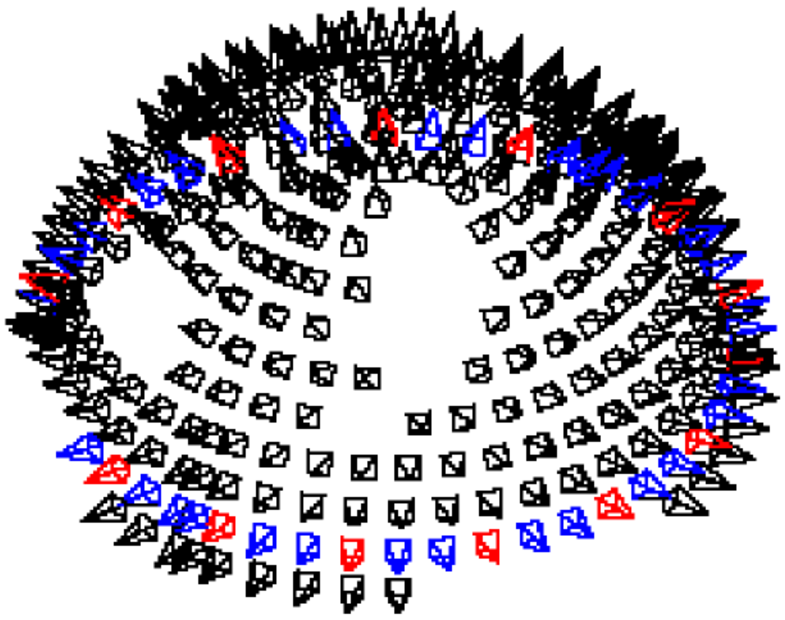
- **Ratio test**
- **Keypoints spread**
- **Geometry validation**



RANSAC for line fitting

Scale and Rotation Invariance

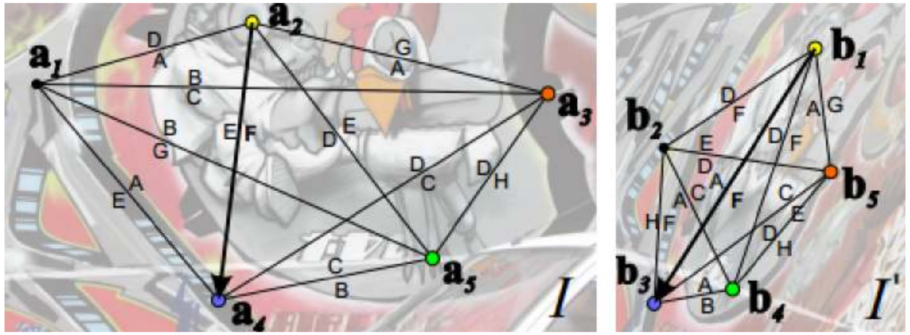
Problem: Local feature descriptors (SIFT, SURF, ORB, FREAK, etc.) do not handle extremal angles



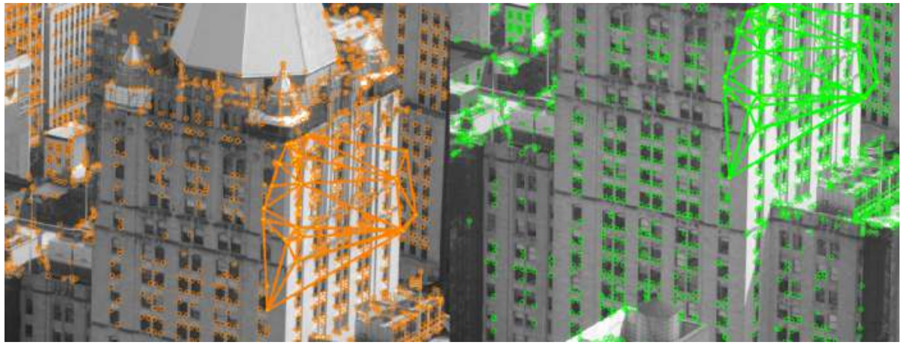
Synthetic views warping

Generated views

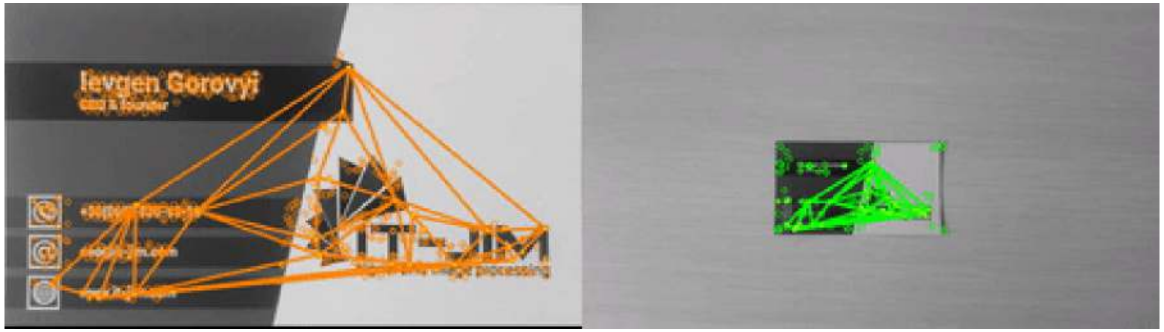
D-Nets



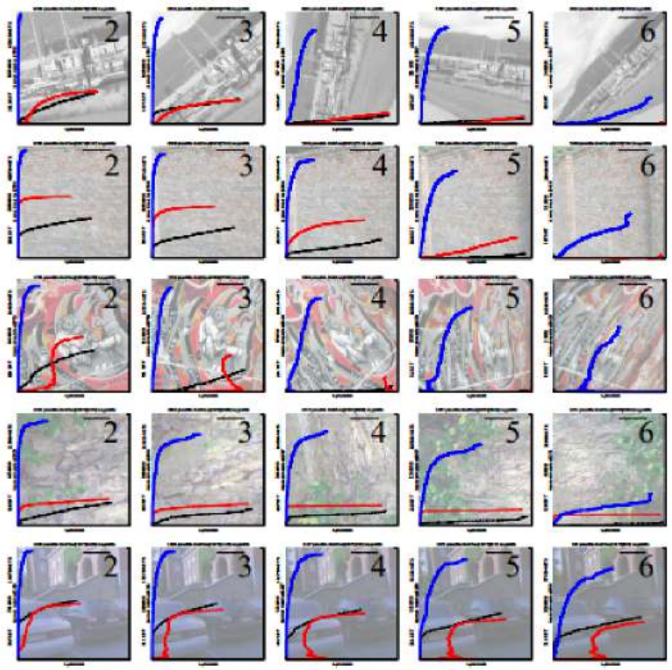
Geometric principles



Two-view matching

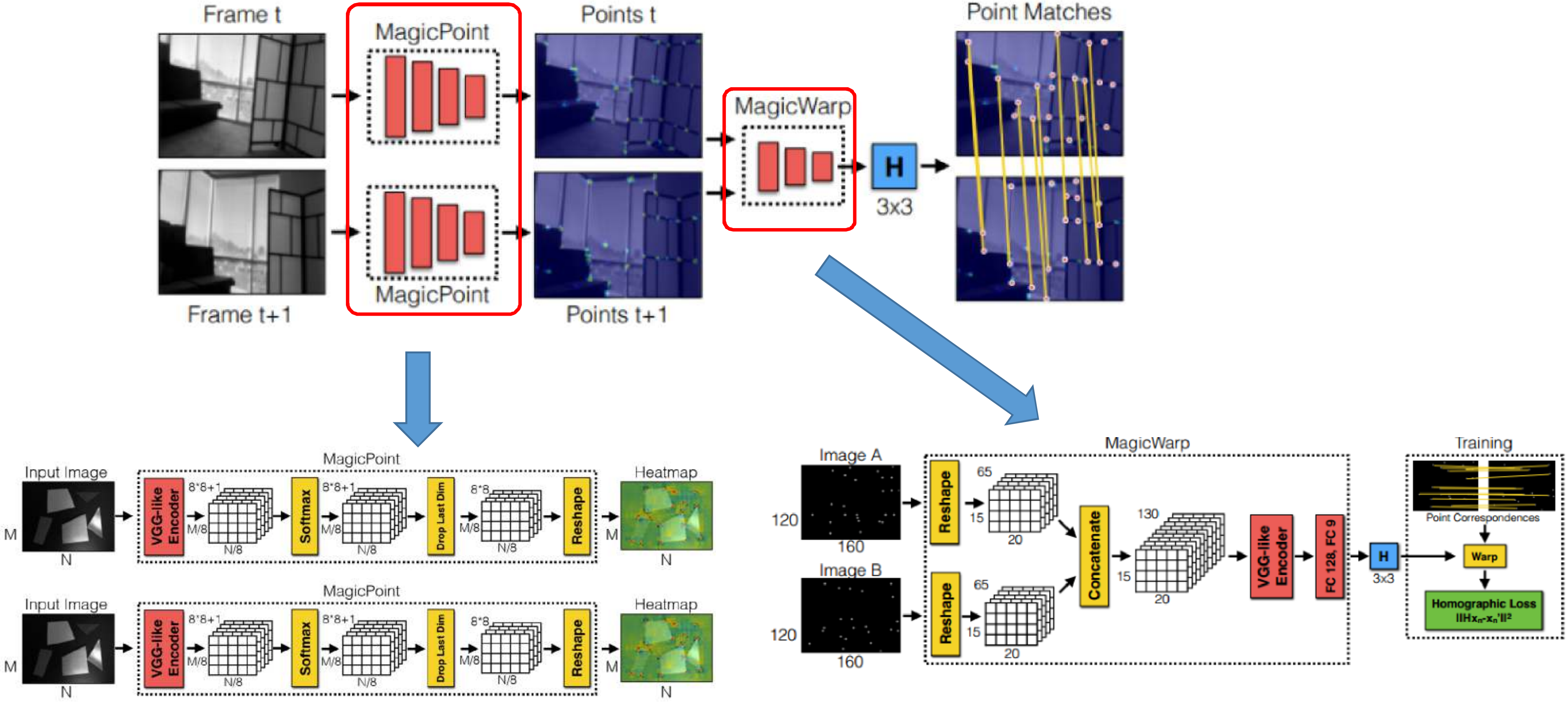


Live example



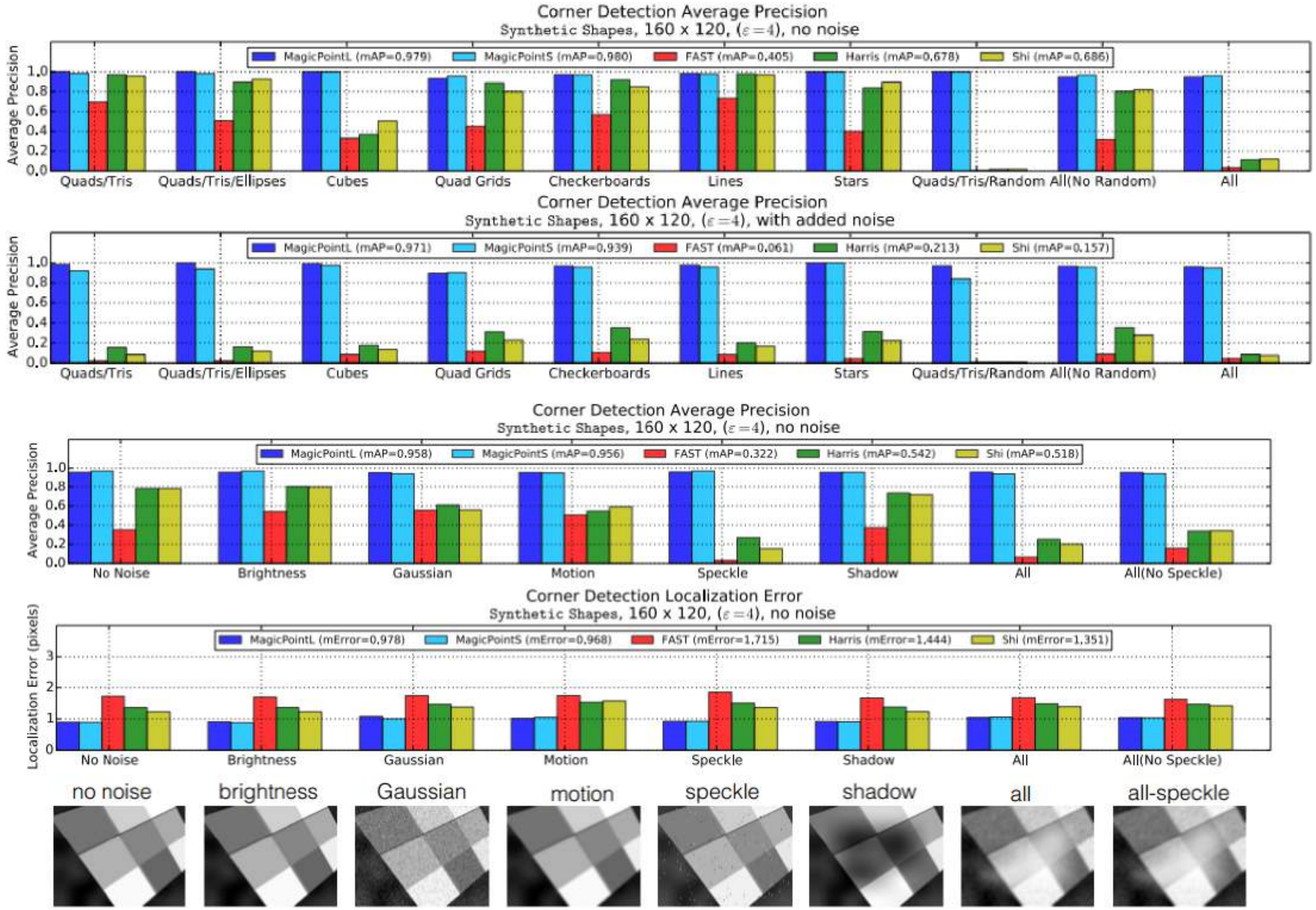
D-Nets vs ORB, SIFT

MagicWarp

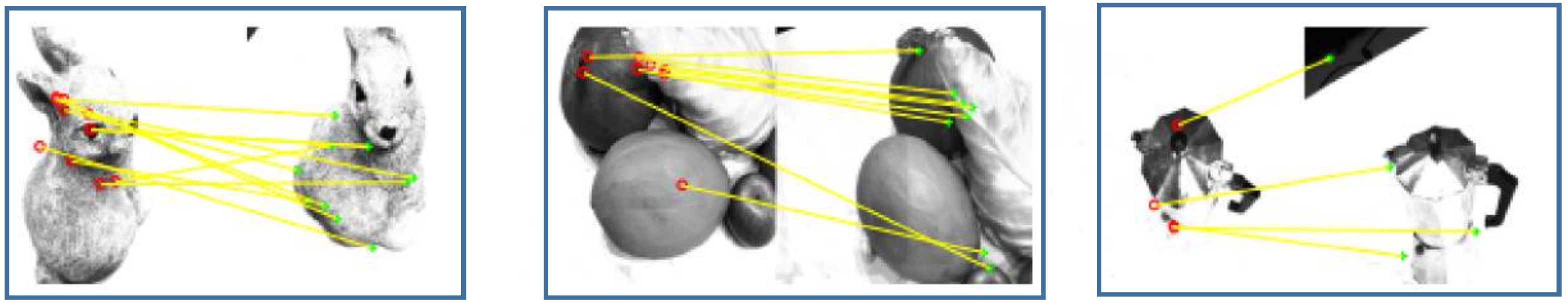


D. DeTonem T. Malisiewicz, A.Rabinovich - "Toward Geometric Deep SLAM" / https://arxiv.org/pdf/1707.07410.pdf

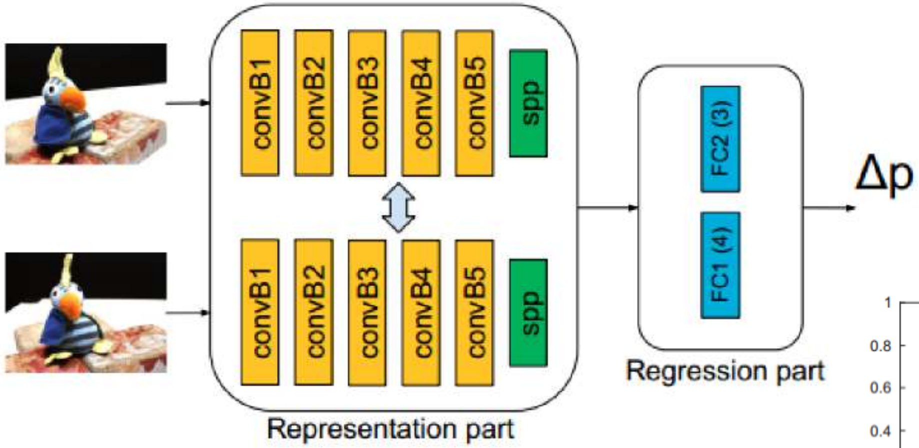
MagicWarp Evaluation



Relative Camera Pose using CNN

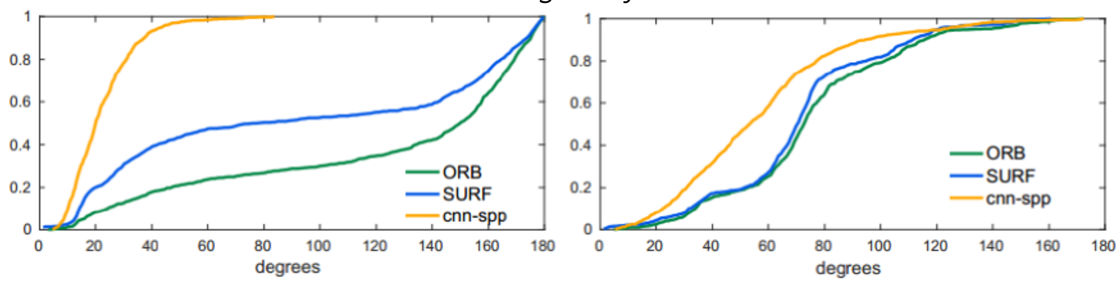


Challenging matching conditions



Network architecture

Textureless and light reflections



Cumulative hist

Iaroslav Melekhov, Juha Ylioinas, Juho Kannala, Esa Rahtu
Relative Camera Pose Estimation Using Convolutional Neural Networks, International Conference on Advanced Concepts for Intelligent Vision Systems, 2017

AR: Algorithms and Solutions

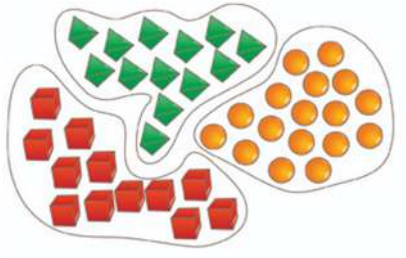
- camera pose and marker detection
- image retrieval
- tracking engine
- markerless and SLAM

Image Retrieval: BoW

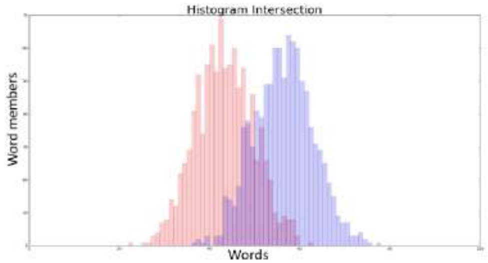
Question: How to track 1000 markers in real-time?

Image Retrieval: d-BoW

BOW **K-means**

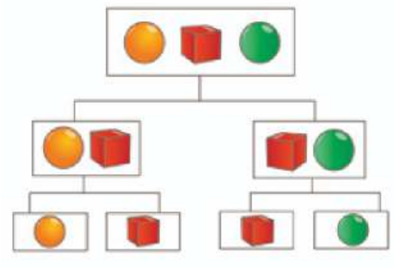


Algorithm:
1. Build feature -histogram
2. Match Histograms

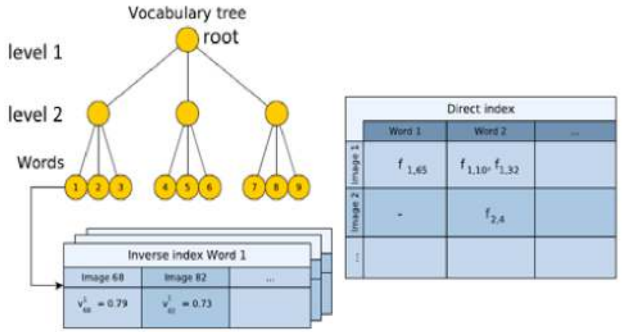


- ✓ Simple
- ✗ For small datasets
- ✗ Fixed feature vector size
- ✗ Computation time

d-BOW **Hierarchical K-means**



Algorithm:
1. Define tree structure
2. Build Voc + DB
3. Match BoW feature vectors



- ✗ More complex solution
- ✓ For large datasets
- ✓ feature vector size < #clusters
- ✓ Computation time is low

Real-Time Marker Detection (1000 markers)

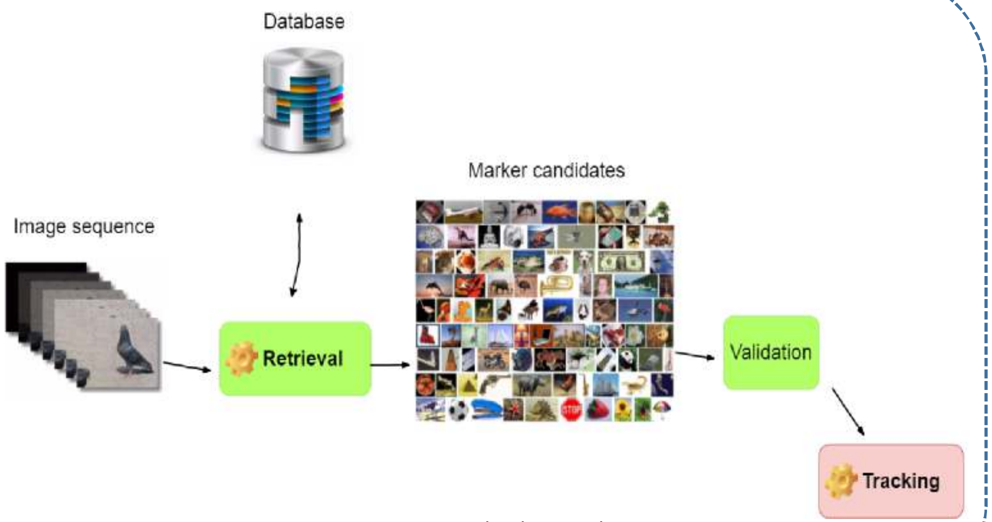
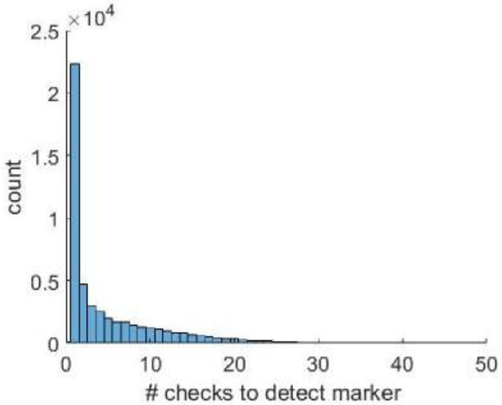
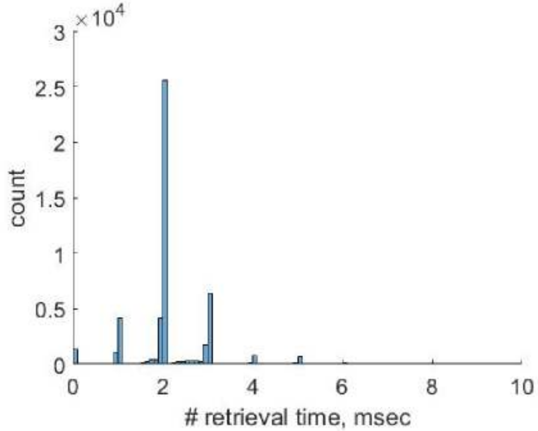


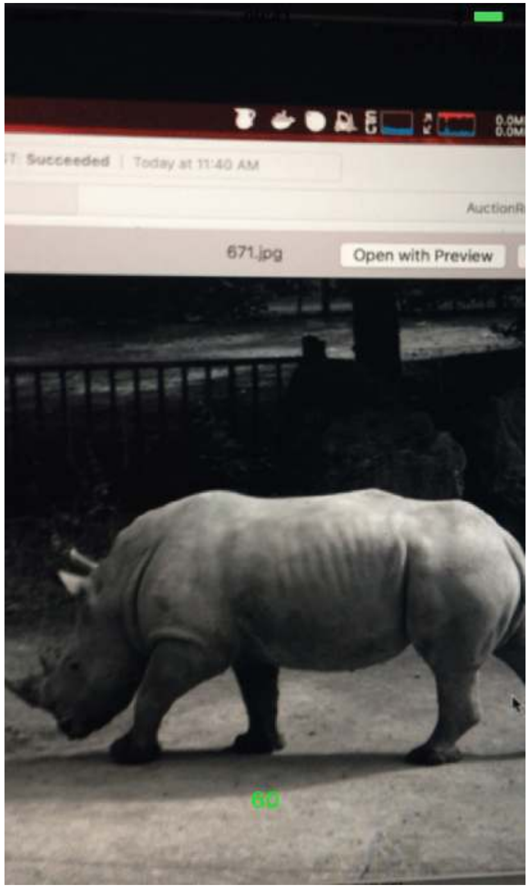
Image retrieval algorithm



#checks histogram



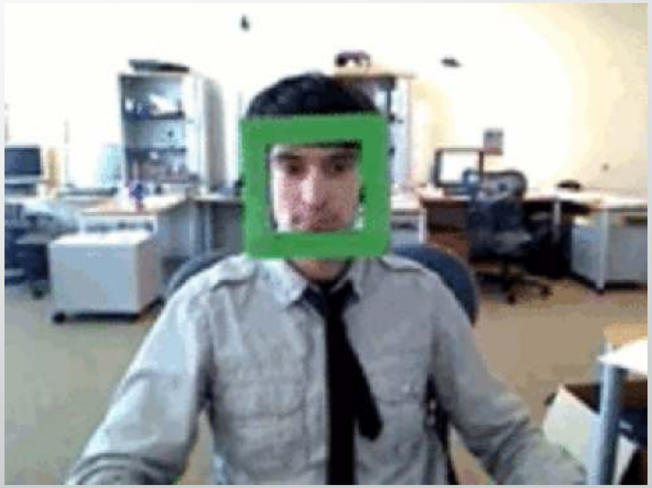
Retrieval time distribution



Retrieval example (iPhone 6)

AR: Algorithms and Solutions

- camera pose and marker detection
- image retrieval
- tracking
- markerless and SLAM



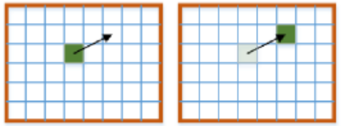
Detection vs Tracking

Problem: Tracking-by-detection is not good idea

- Slow
- Fails for extremal angles

- Fast
- More robust to geometry

Marker Tracking: Optical Flow (LK)



$$I_x u + I_y v + I_t = 0$$

$$\begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_y I_x & \sum I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

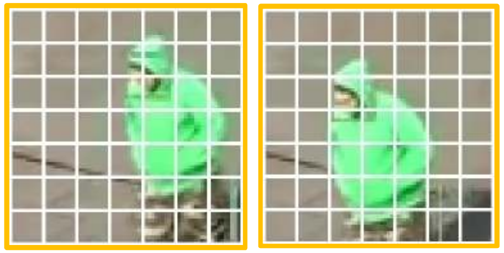


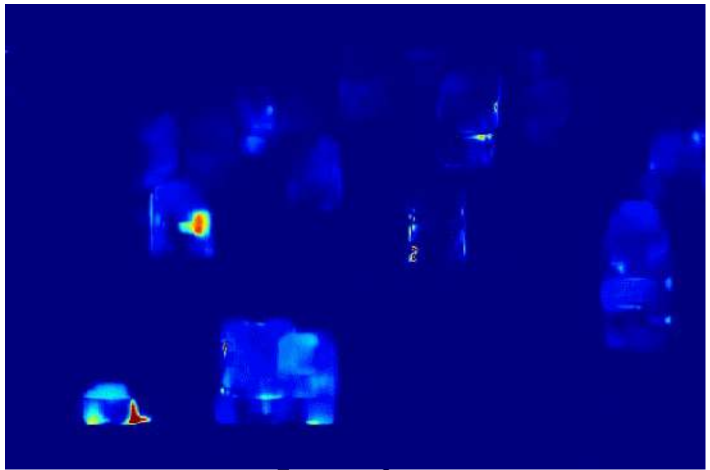
Image patch

Ix, Iy, It

Example of keypoints tracking

Video stream

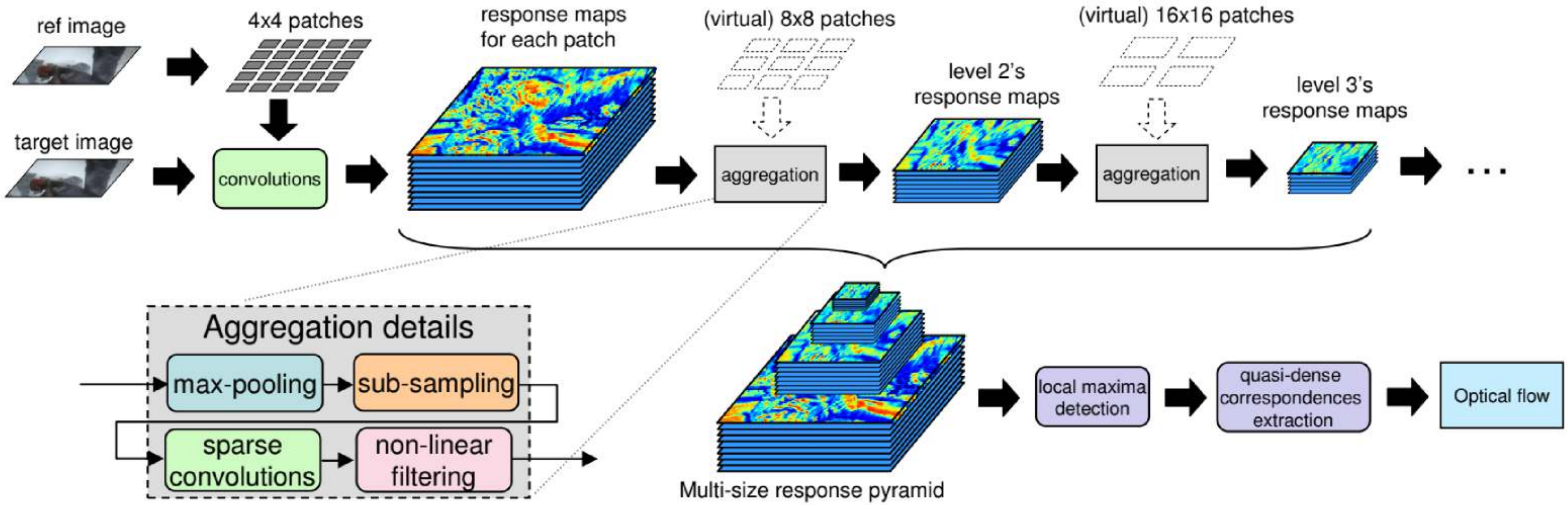
Frame 1



Optical flow magnitude

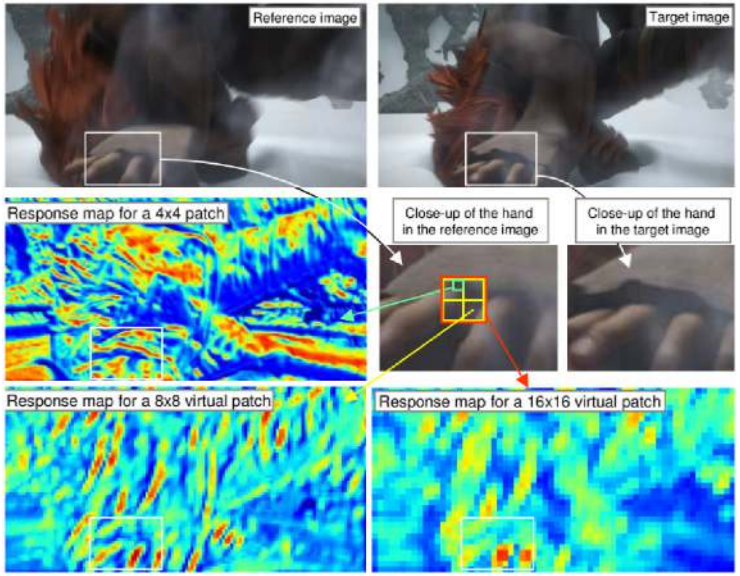
Frame 2

DeepFlow

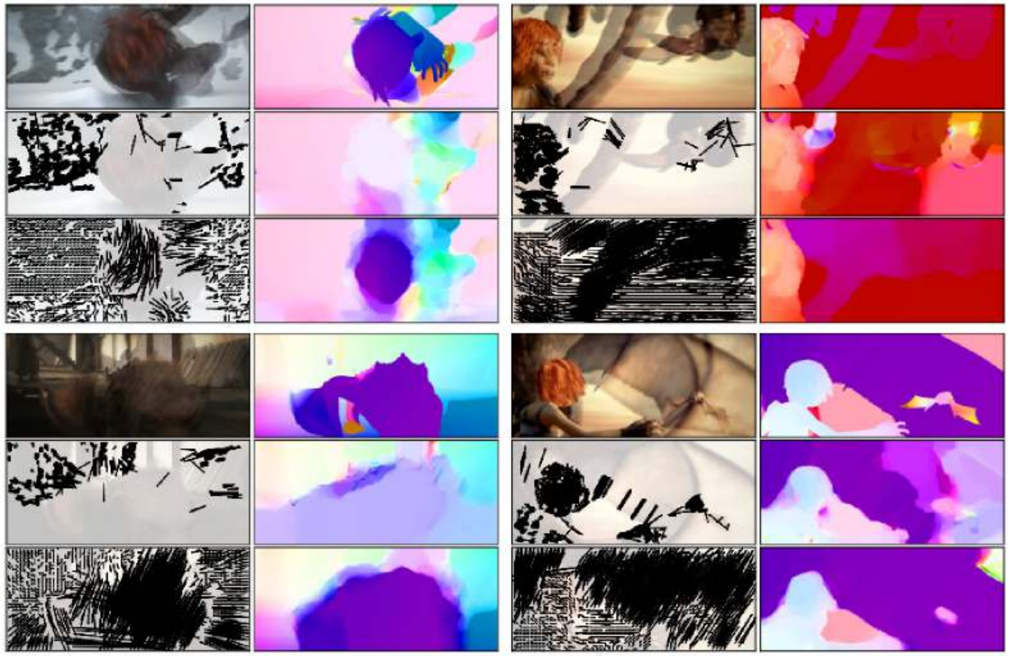


P. Weinzaepfel, J. Revaud, Z. Harchaoui, C. Schmid – " DeepFlow: Large displacement optical flow with deep matching " / ICCV-2013, Dec 2013, Sydney, Australia

DeepFlow Evaluation

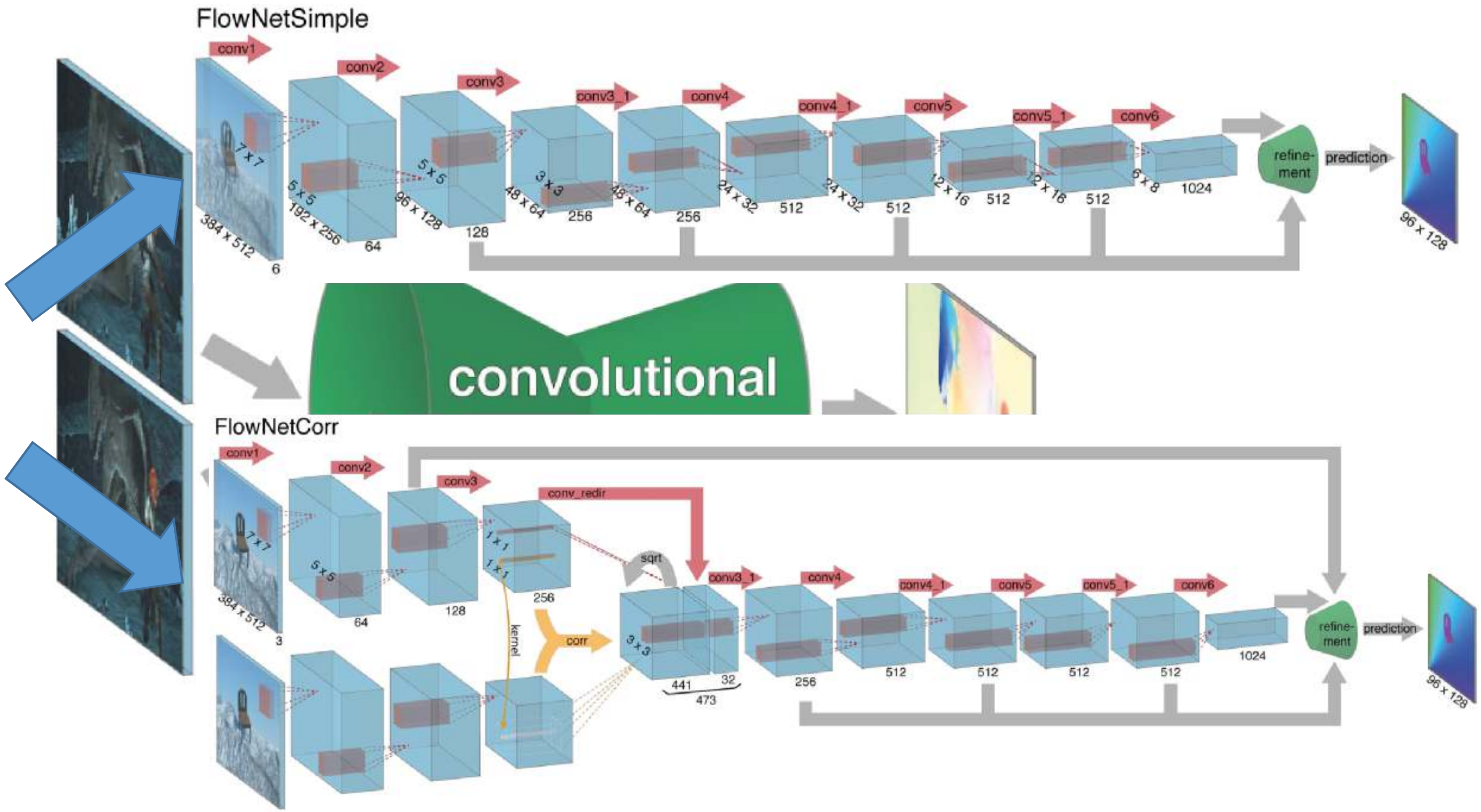


Response maps



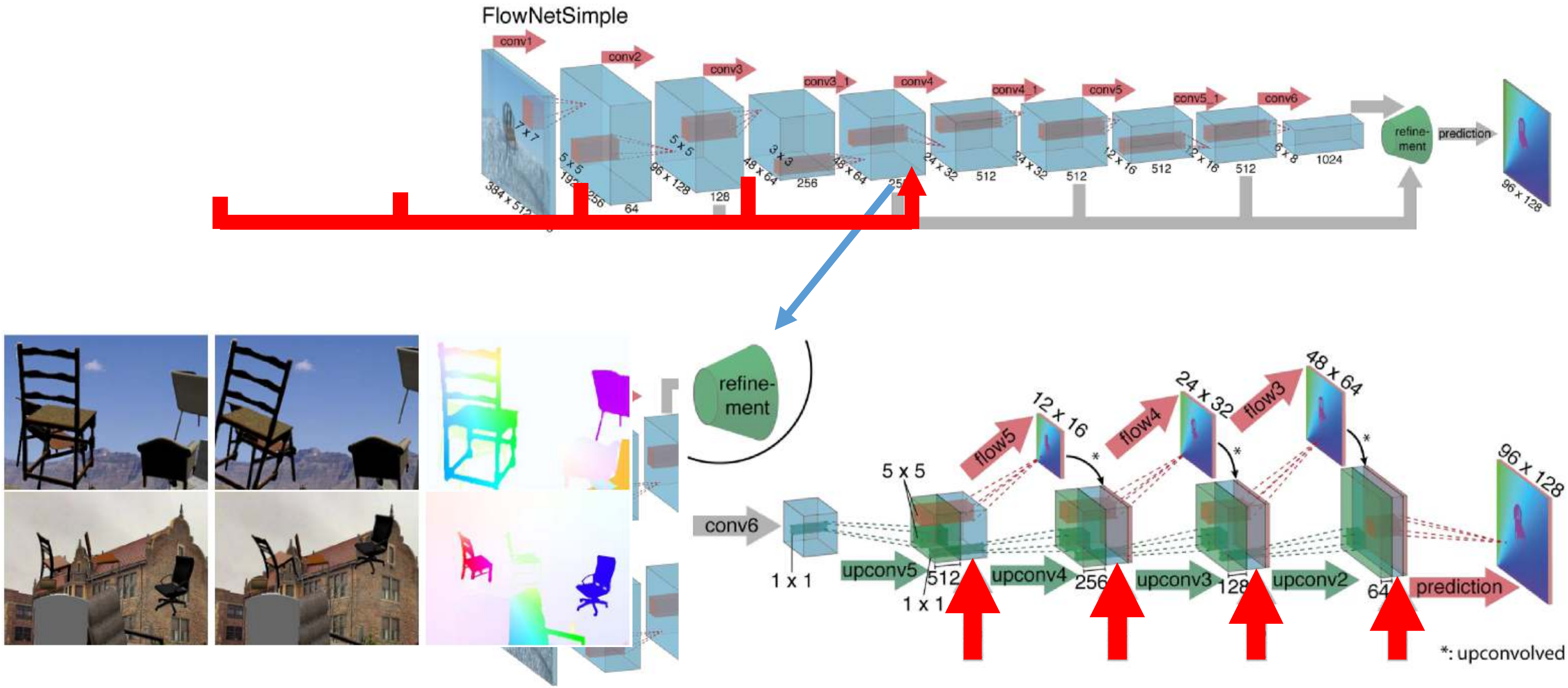
P. Weinzaepfel, J. Revaud, Z. Harchaoui, C. Schmid – " DeepFlow: Large displacement optical flow with deep matching " / ICCV-2013, Dec 2013, Sydney, Australia

DL for Optical Flow: FlowNet



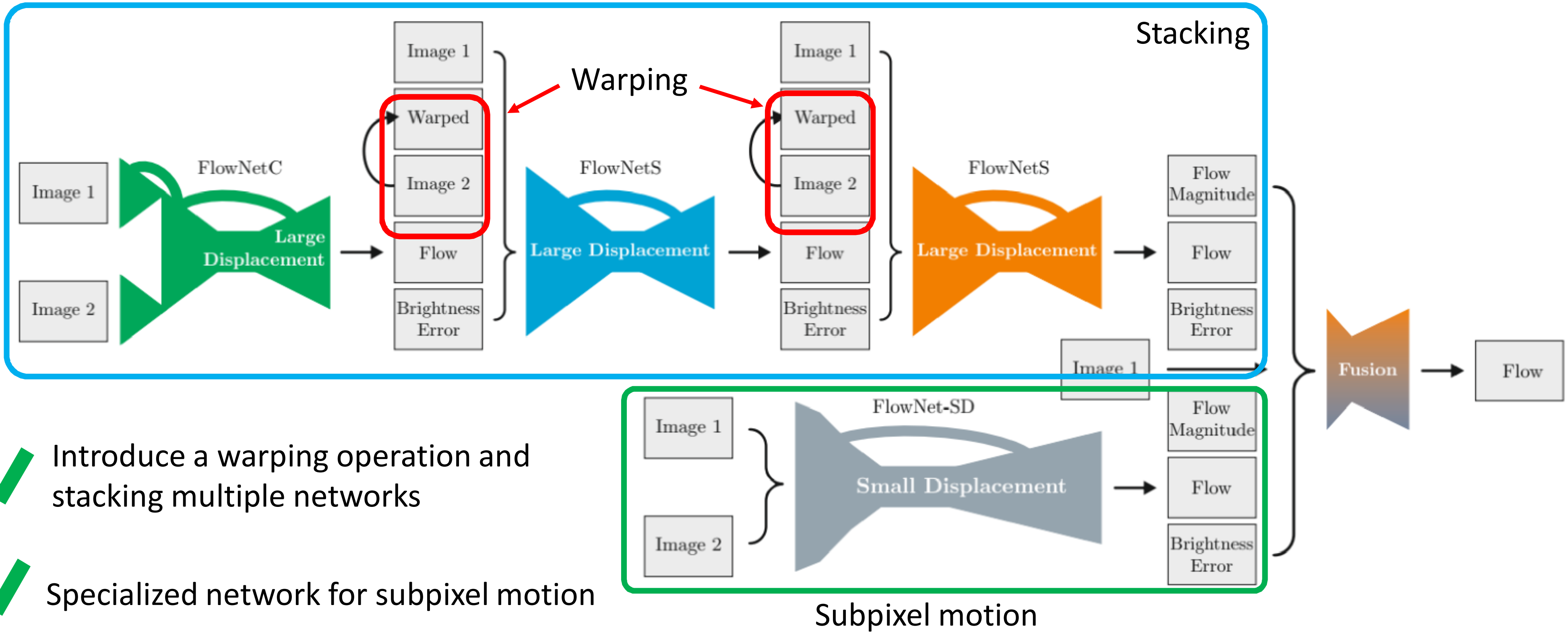
A. Dosovitskiy *et al* – "FlowNet: Learning Optical Flow with Convolutional Networks" / ICCV-2015, December 11-18, 2015, Santiago, Chile

DL for Optical Flow: FlowNet



A. Dosovitskiy *et al* – "FlowNet: Learning Optical Flow with Convolutional Networks" / ICCV-2015, December 11-18, 2015, Santiago, Chile

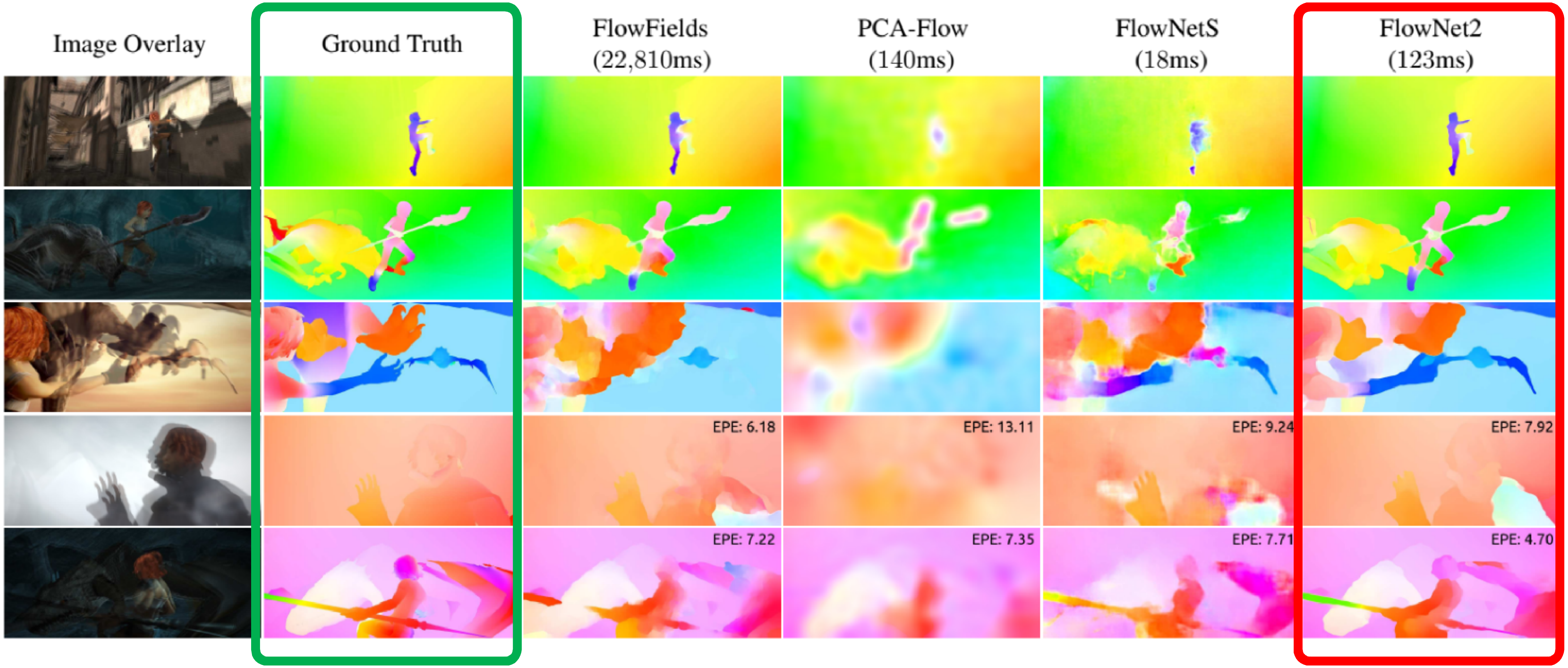
DL for Optical Flow: FlowNet 2.0



- Introduce a warping operation and stacking multiple networks
- Specialized network for subpixel motion

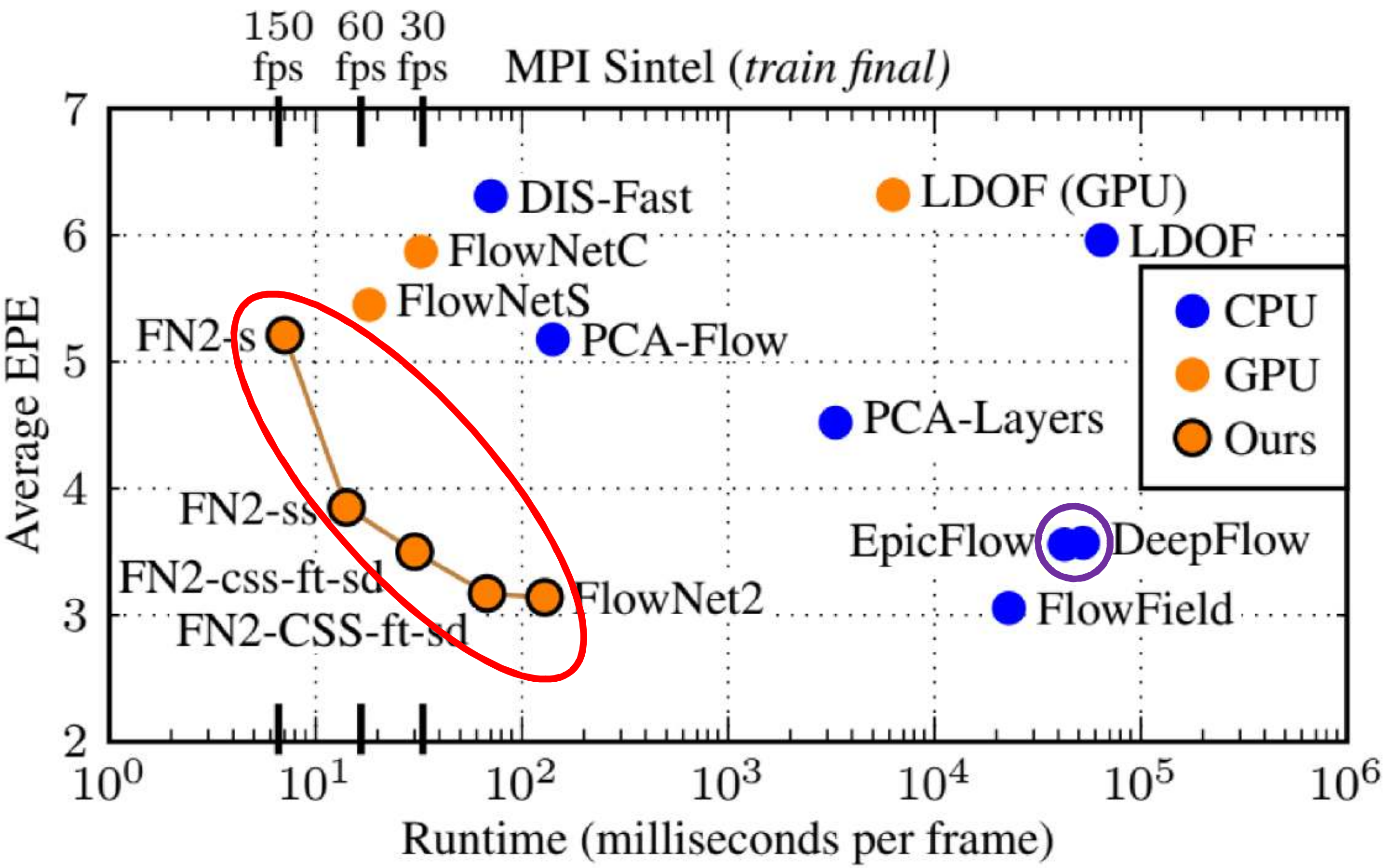
E. Ilg et al – " FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks" / CVPR-2017, July 22-25, 2017, Honolulu, Hawaii

DL for Optical Flow: FlowNet 2.0



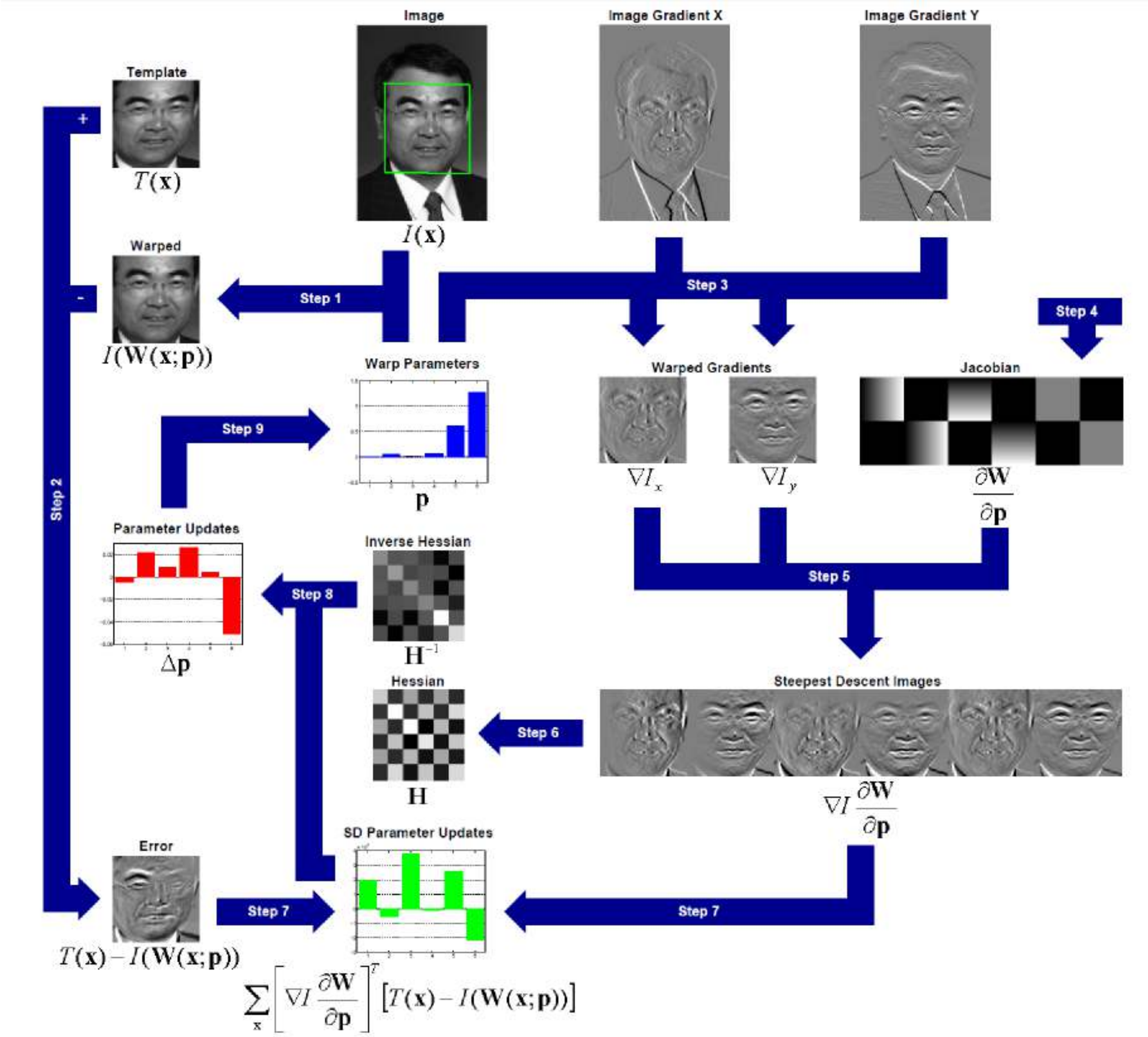
E. Ilg et al – " FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks" / CVPR-2017, July 22-25, 2017, Honolulu, Hawaii

DL for Optical Flow: FlowNet 2.0



E. Ilg et al – " FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks" / CVPR-2017, July 22-25, 2017, Honolulu, Hawaii

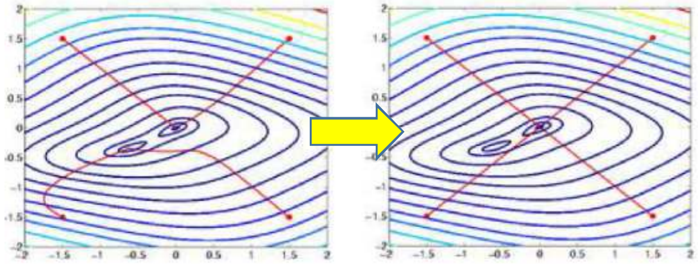
Template-Based Tracking: Lucas-Kanade



Template-based tracking (Gauss-Newton gradient descent)

Template-based tracking: Improvements

Efficient second-order minimization (ESM)



- ✓ *less iterations needed*
- ✓ *larger area of convergence*
- ✓ *avoid local minima*
- ✓ *second-order Taylor approximation*
- ✗ *Jacobian is calculated each time*

Sum of conditional variance (SCV)

ESM, but:

current image — *reference image*

$$\Delta = \sum_{\mathbf{x}} \left(I(w(\mathbf{x},\mathbf{p})) - \cancel{T(\mathbf{x})} \right)^2$$

$$\mathscr{E}(I(w(\mathbf{x},\mathbf{p}))|T(\mathbf{x}))$$

expectation operator

SSD = 2.4e+9
SCV = 0

✓ *more robust*

ESM with gradient orientations (ESM-GO, 2017)

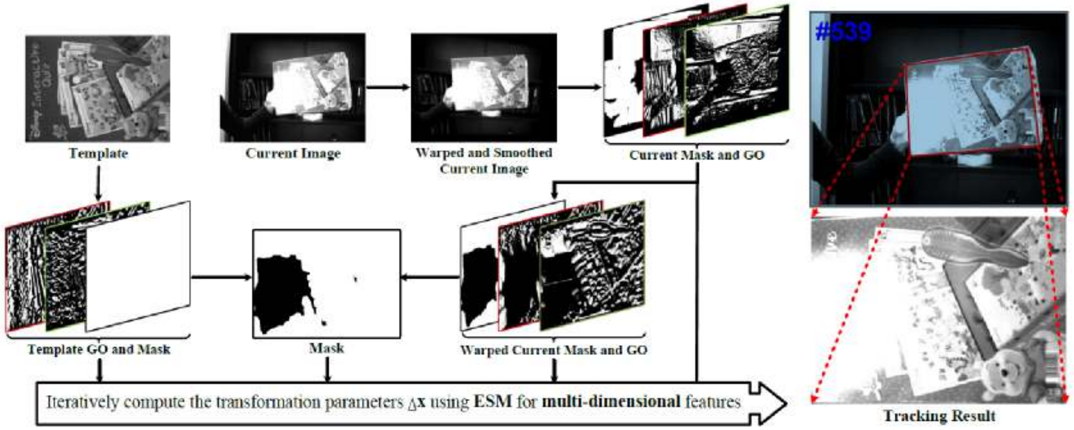
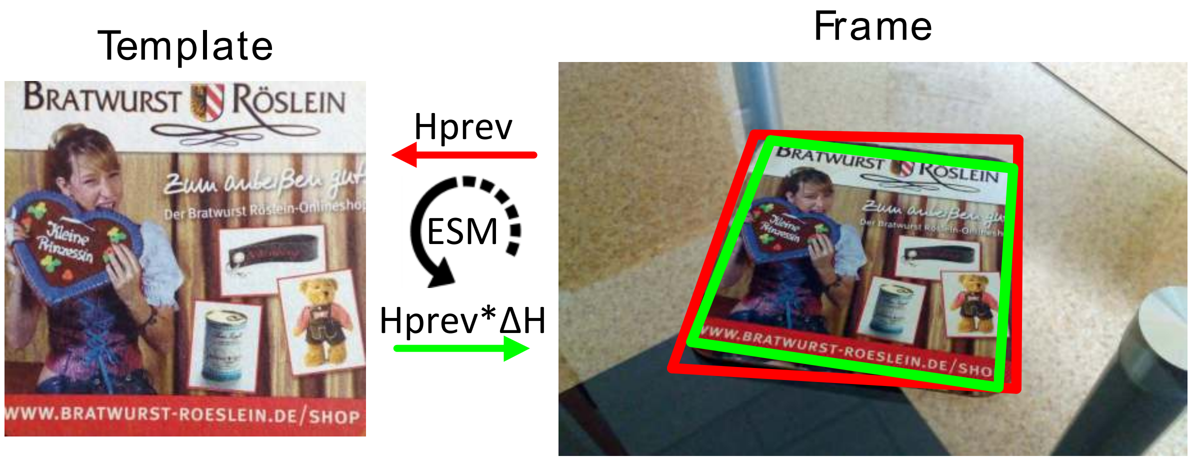


image gradients ⇩ *illumination robustness*

anisotropic diffusion ⇩ *noise robustness*

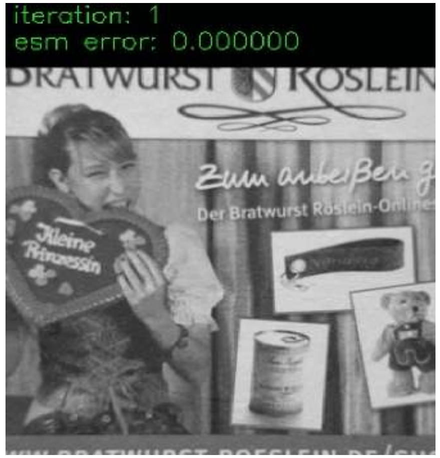
masking ⇩ *low texture robustness*

Fusion of Trackers



Homography refinement

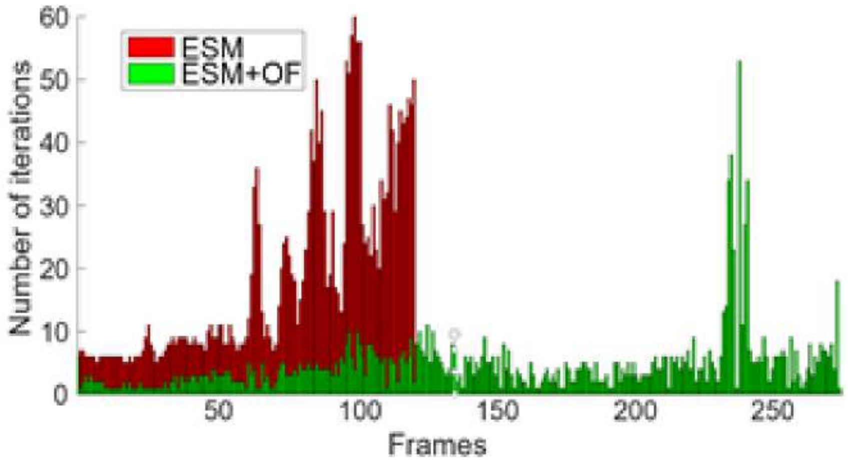
$$H_{cur} = H_{prev} \times \Delta H_{1..N}$$



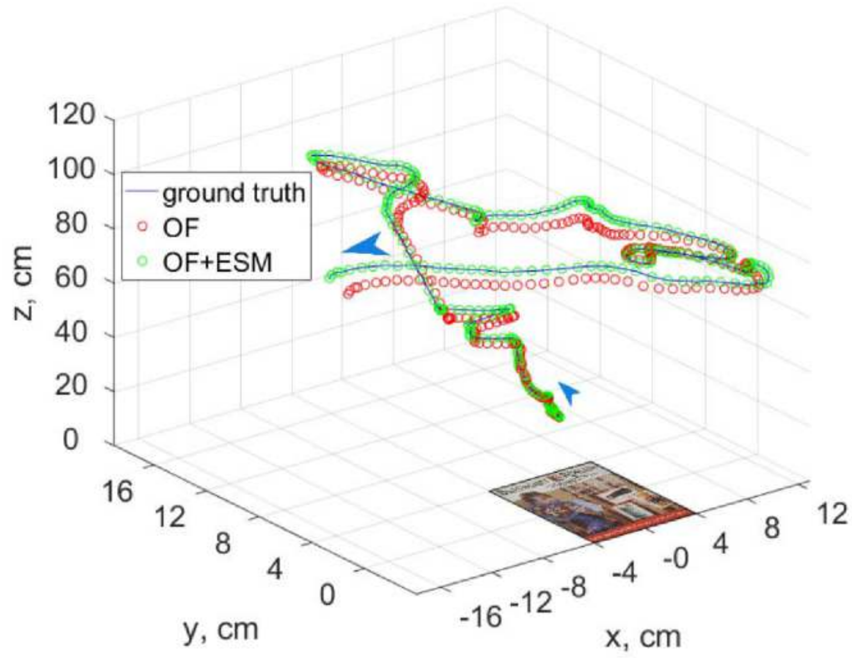
Convergence steps

ESM tracking

Hybrid tracking



Hybrid Tracking*



✓ Drift-free

✓ Extremal angles

Hybrid tracking and camera pose estimation

*Ievgen M. Gorovyi, Dmytro S. Sharapov. Advanced Image Tracking Approach for Augmented Reality Applications. Proceedings of Signal Processing Symposium (SPSympo-2017), 12-14 September, Jachranka, Poland, pp.266-270 (2017) **1st Prize for the best paper**

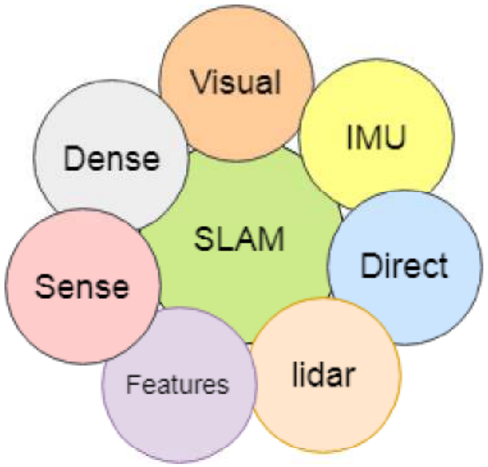
AR: Algorithms and Solutions

- marker detection
- image retrieval
- tracking engine
- markerless and SLAM

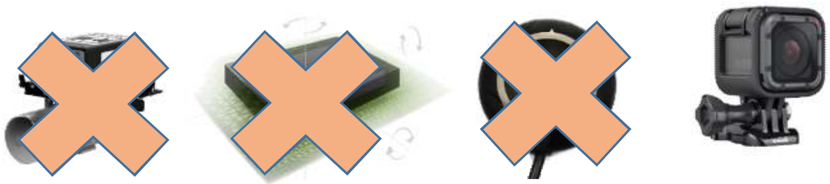
Markerless: VIO and SLAM

VO: Visual Odometry.
Goal: camera tracking using the scene. Local map optimization

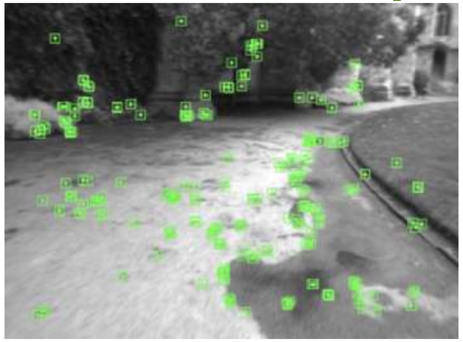
SLAM: Simultaneous Localization and Mapping.
Goal: persistent 3d and camera location simultaneously. Global map optimization



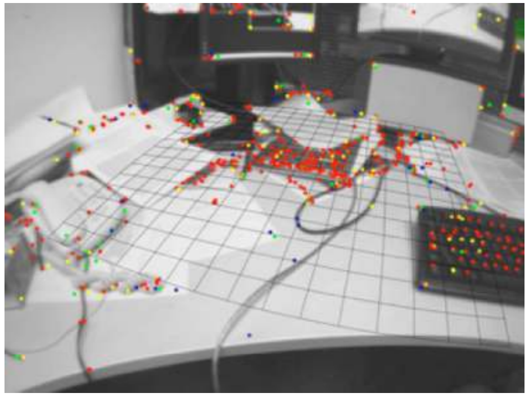
SLAM for AR



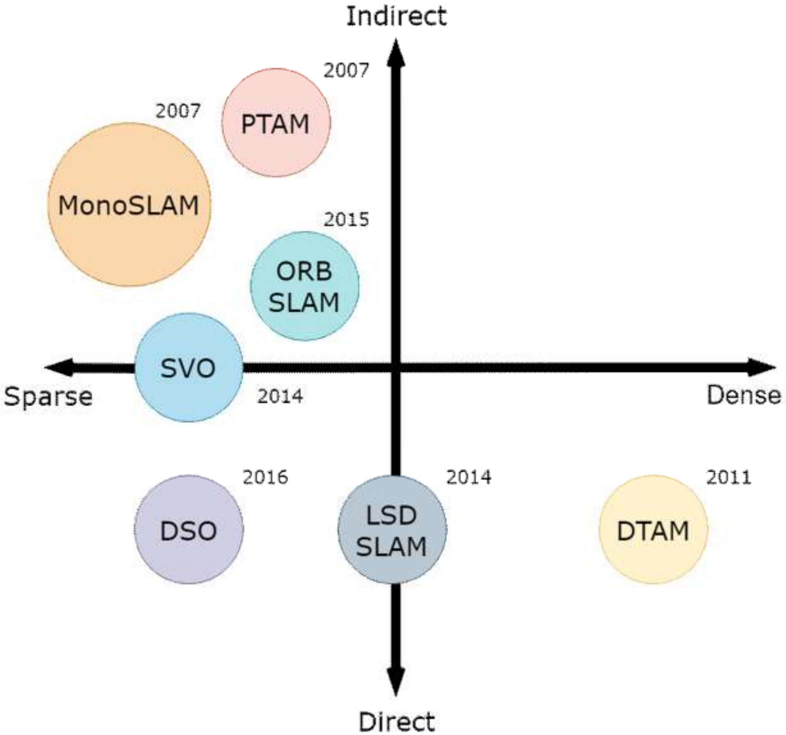
SLAM: Direct/Indirect, Sparse/Dense

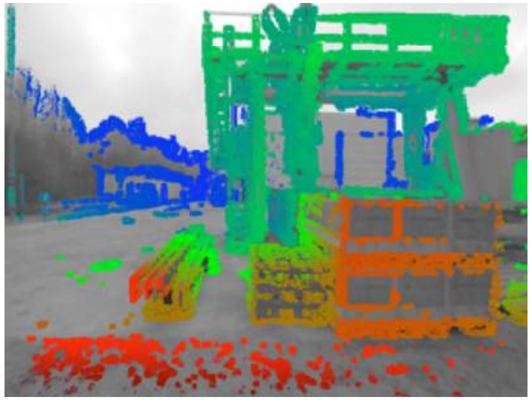


Local features



Sparse map





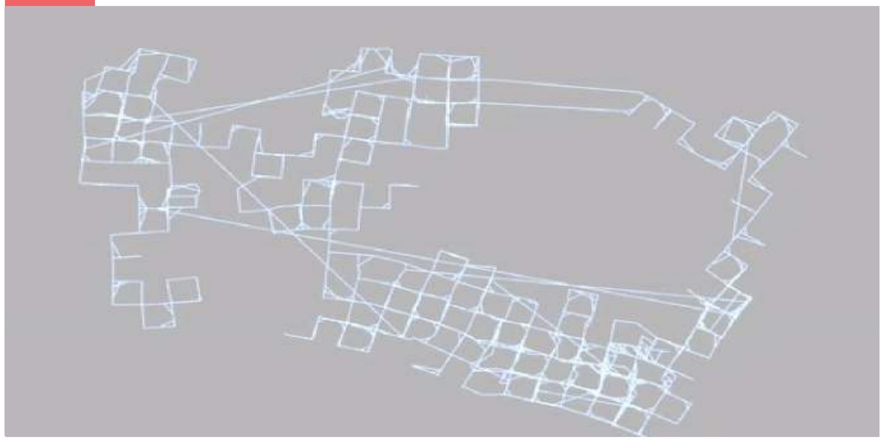
Dense map

Slam System Components

01 Initialization

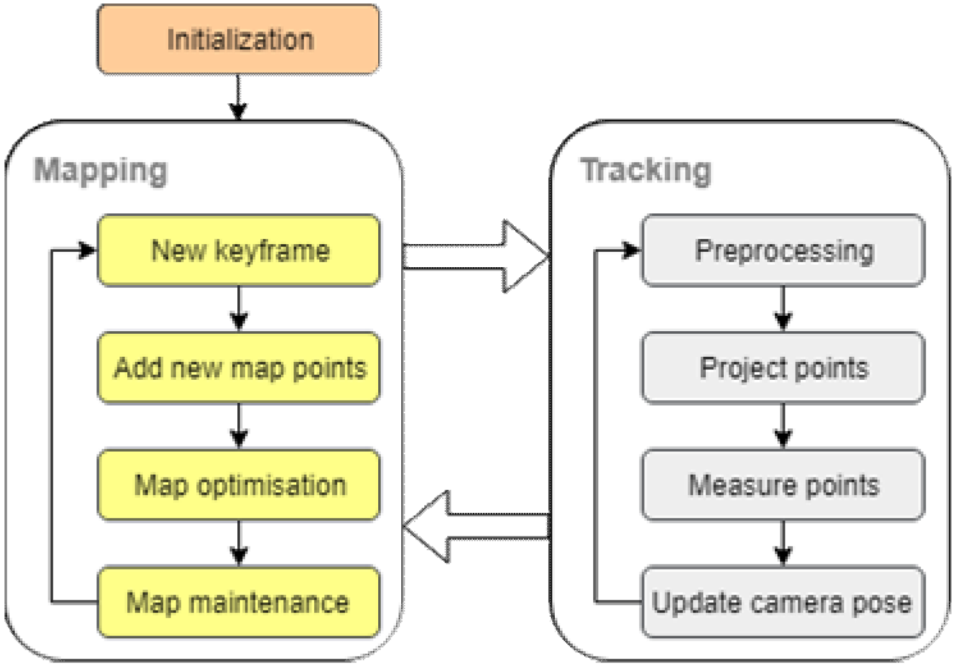
02 Tracking & mapping

03 Optimization



04 AR interaction

Parallel Tracking and Mapping (PTAM)



PTAM workflow

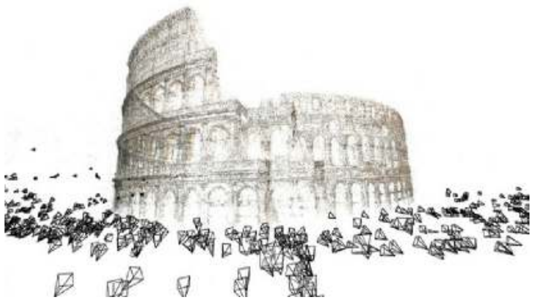
Demo (iPhone 3!)

Parallel tracking and mapping for small AR workspaces, Klein G, Murray D, ISMAR 2007

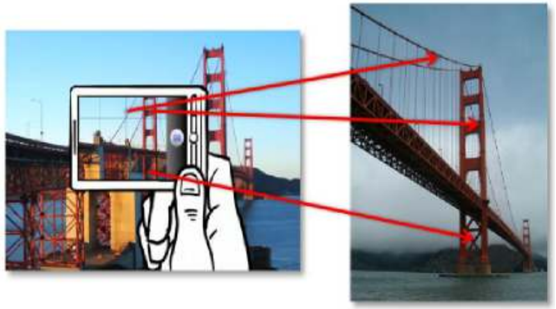
ORB-SLAM

Tracking ✓

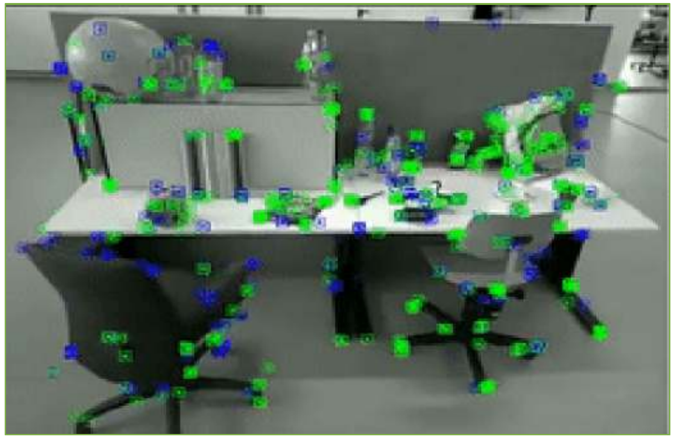
Map ✓



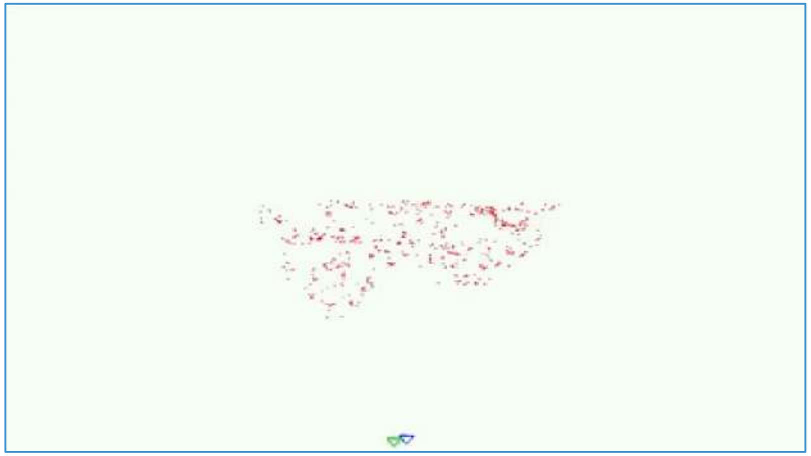
Matching ✓



Feature tracking



Camera tracking

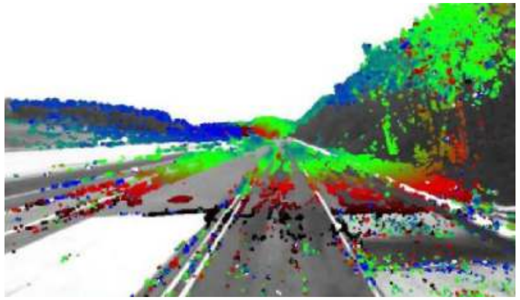


Raúl Mur-Artal, J. M. M. Montiel and Juan D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, 2015.

LSD SLAM

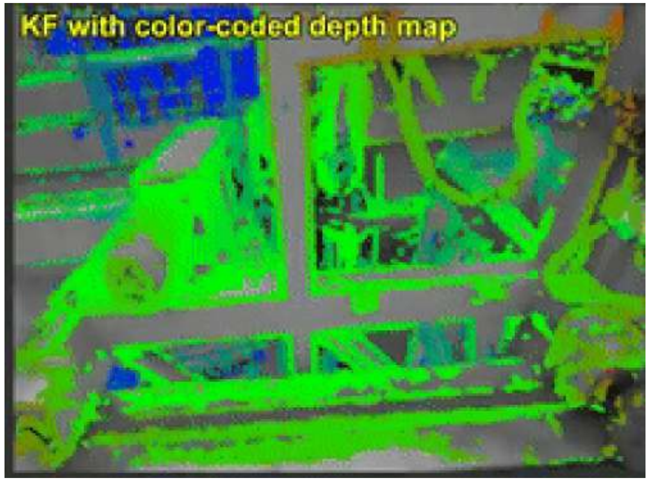
Tracking ✓

High gradients ✓

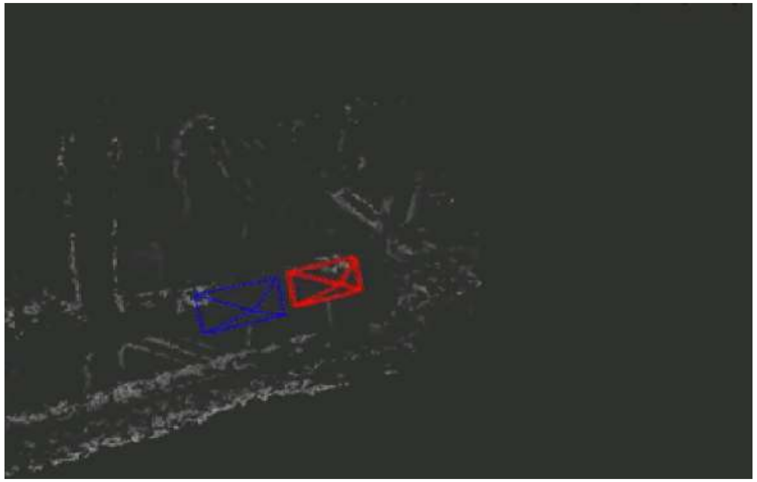


Map ✓

Image tracking



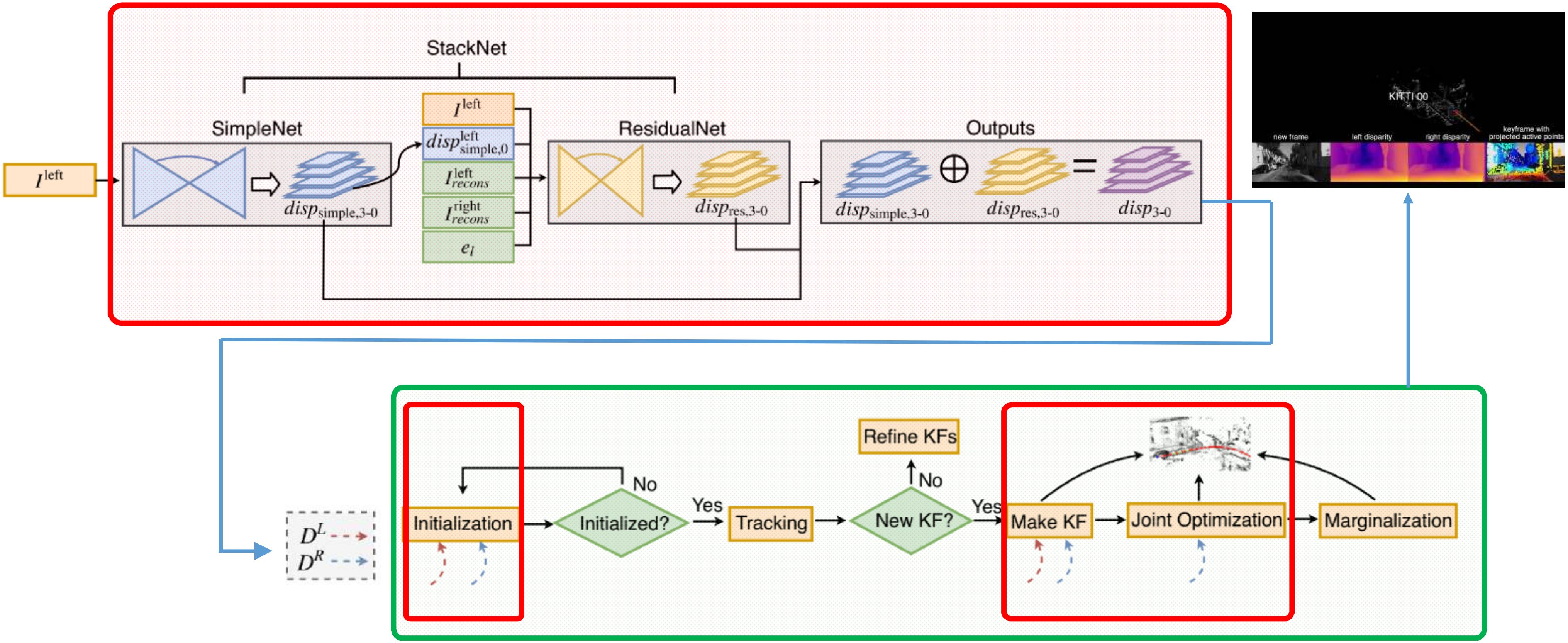
Mapping



LSD-SLAM: Large-Scale Direct Monocular SLAM, J. Engel, T. Schöps, D. Cremers, ECCV '14

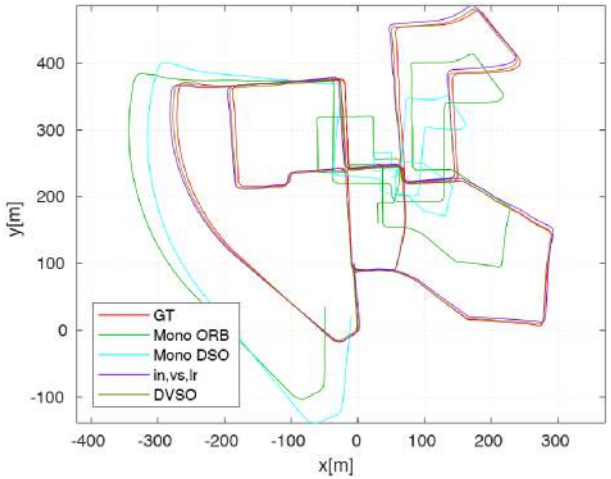
Deep Virtual Stereo Odometry

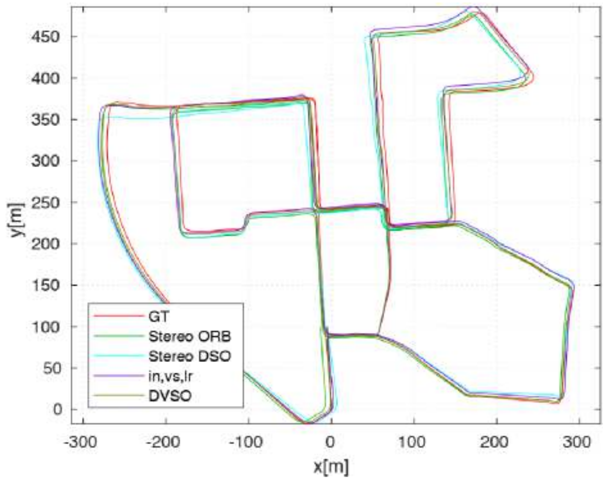
StackNet: generation of stereo from mono



N.Yang, R.Wang, J.Stuckler and D.Cremers – "Deep Virtual Stereo Odometry: Leveraging Deep Depth Prediction for Monocular Direct Sparse Odometry" / ECCV-2018, September 8-14 2018, Munich, Germany

Deep Virtual Stereo Odometry





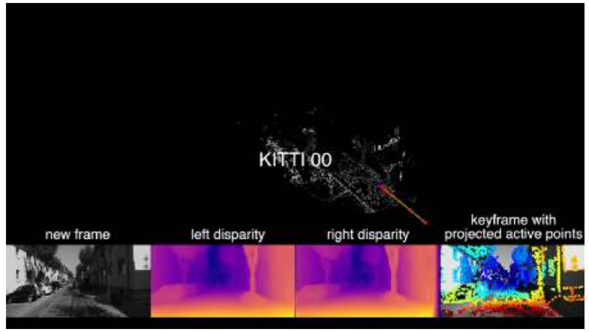
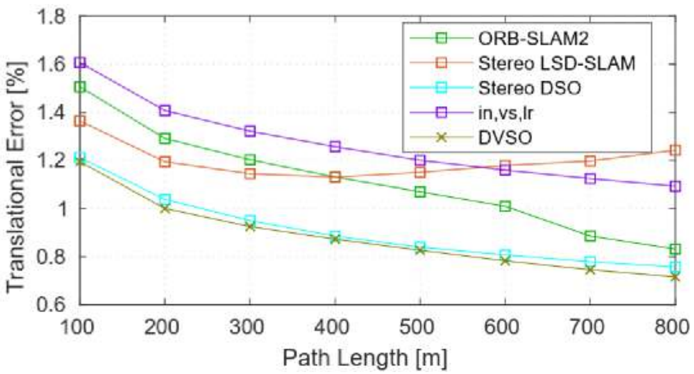
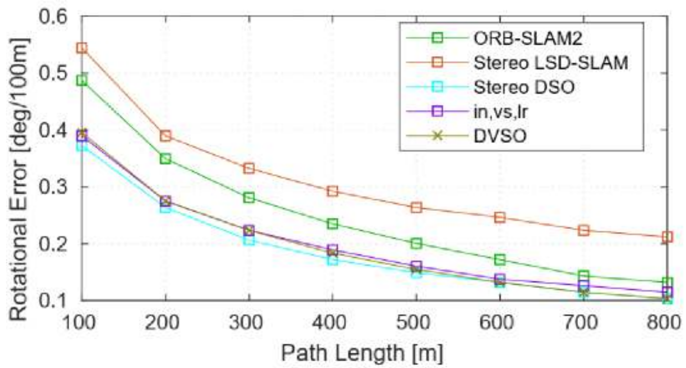


Image size: 512 × 256

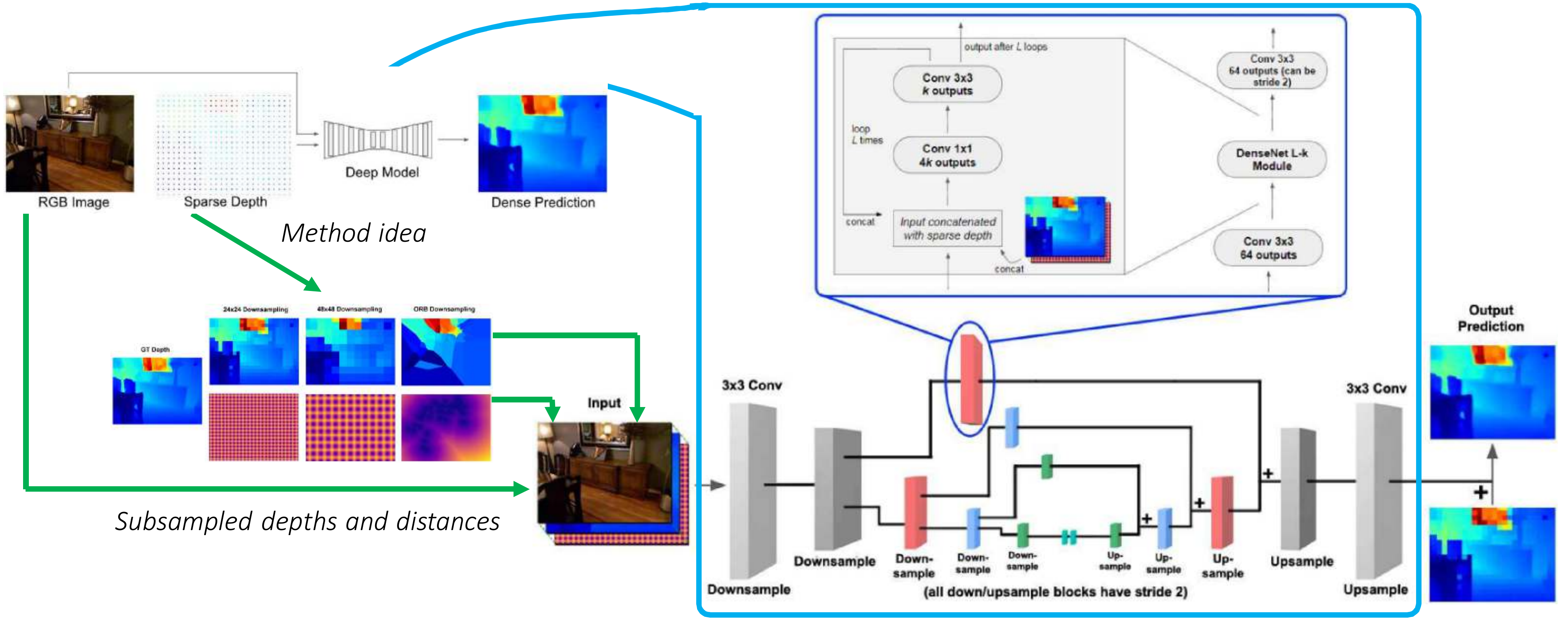
Inference time: 40 ms
(Titan X Pascal GPU)





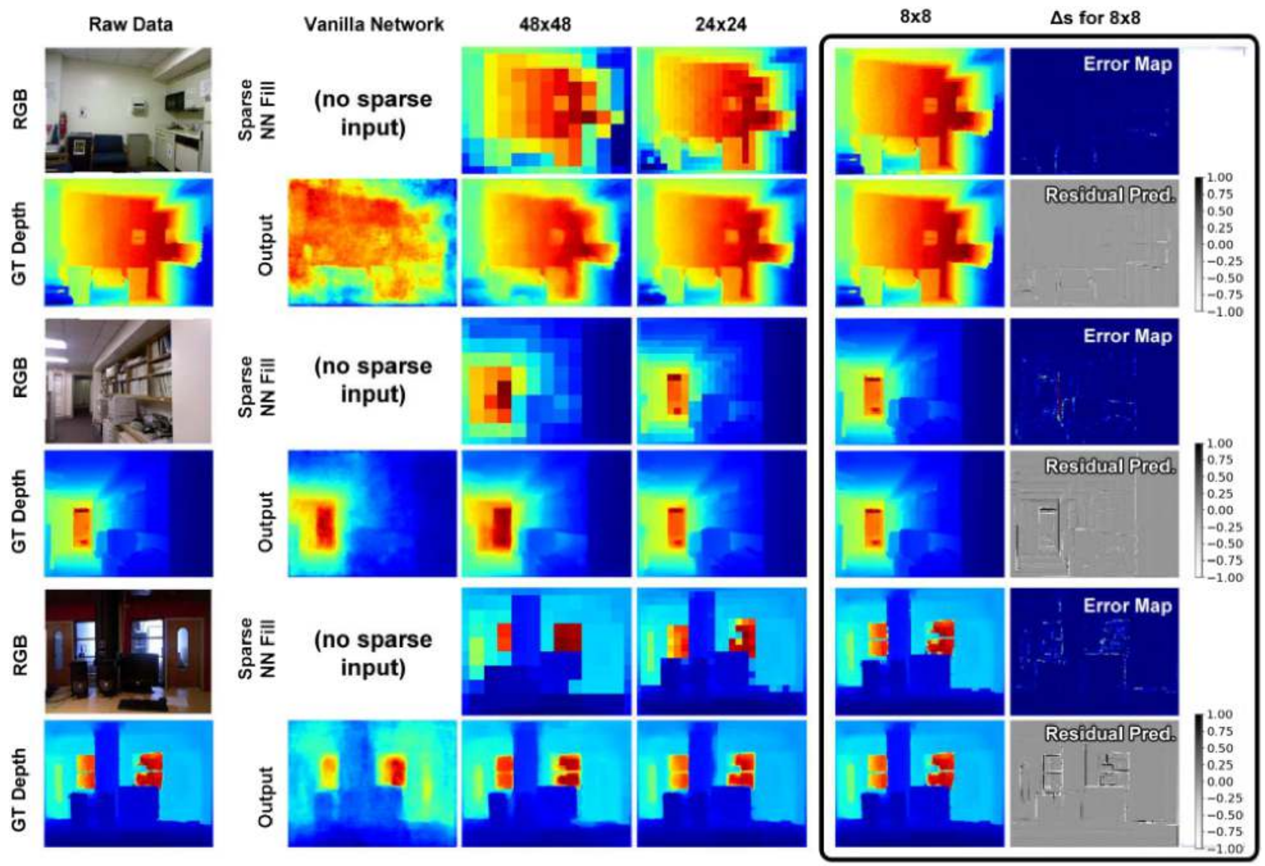
N.Yang, R.Wang, J.Stuckler and D.Cremers – "Deep Virtual Stereo Odometry: Leveraging Deep Depth Prediction for Monocular Direct Sparse Odometry" / ECCV-2018, September 8-14 2018, Munich, Germany

Deep Depth Densification (D³)

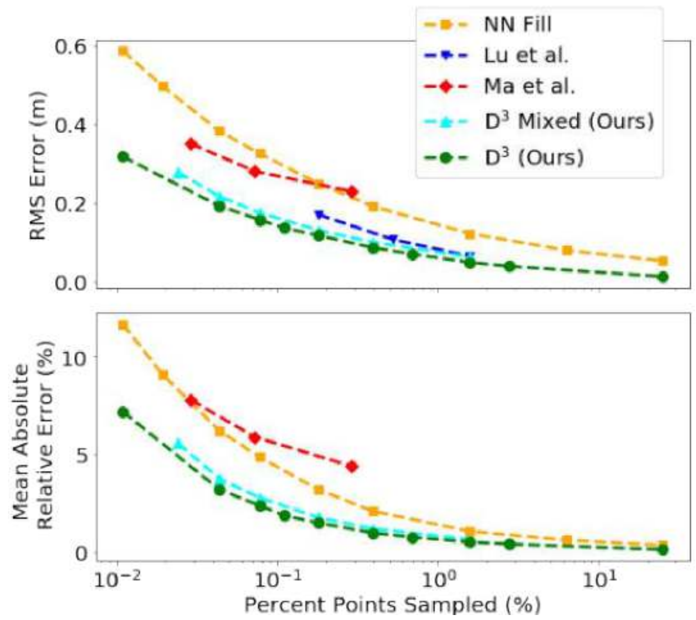


Z.Chen, V.Badrinarayanan, G.Drozdov, A.Rabinovich – "Estimating Depth from RGB and Space Sensing" / ECCV-2018, September 8-14 2018, Munich, Germany

Deep Depth Densification (D^3)



Comparative results



RMS errors

Z.Chen, V.Badrinarayanan, G.Drozdov, A.Rabinovich – "Estimating Depth from RGB and Space Sensing" / ECCV-2018, September 8-14 2018, Munich, Germany

Mobile AR in Production

AR Data Flow: User's Side

1

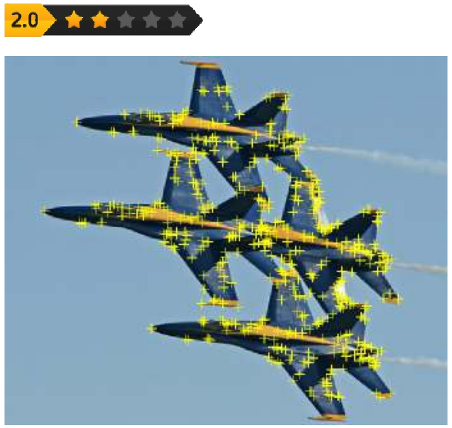
2

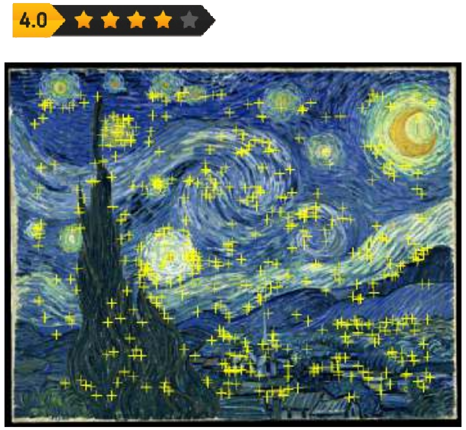
3

AR Marker Quality

Quality check list:

- ✓ Image contrast
- ✓ Image details
- ✓ Features distribution
- ✓ Features repeatability and uiqueness
- ✓ Image viewing angles and scale



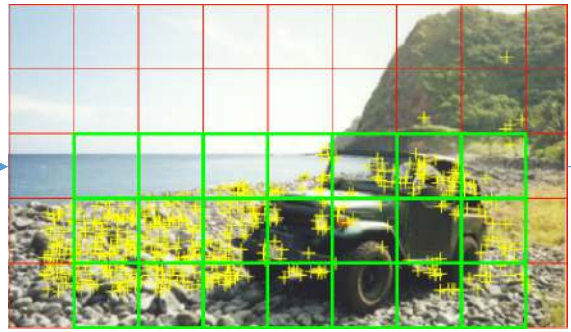


Best ROI selection:

Criteria:
1. Number of interest points in window
2. Interest points distribution in window

Original image

Sliding window processing



Marker

Preprocessing: Text Extraction

Three main concepts:

High-Pass filtering	Laplacian of Gaussian	Laplacian

Initial Grayscale Image

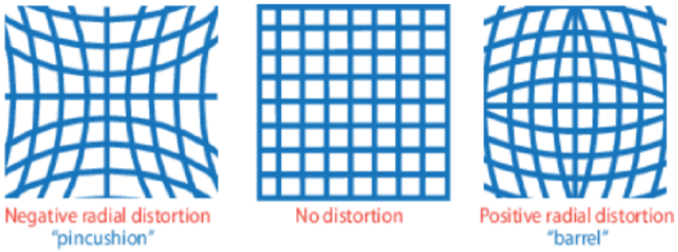
Initial Grayscale Image

Initial Grayscale Image

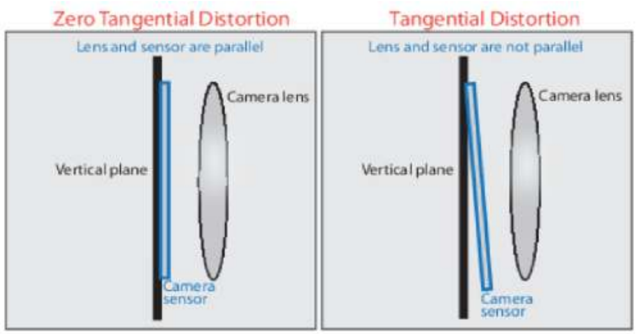
Camera Calibration

K

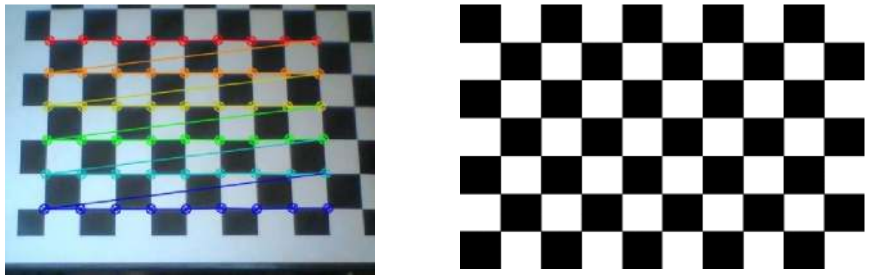
$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & u_x \\ 0 & f_y & v_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{32} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$



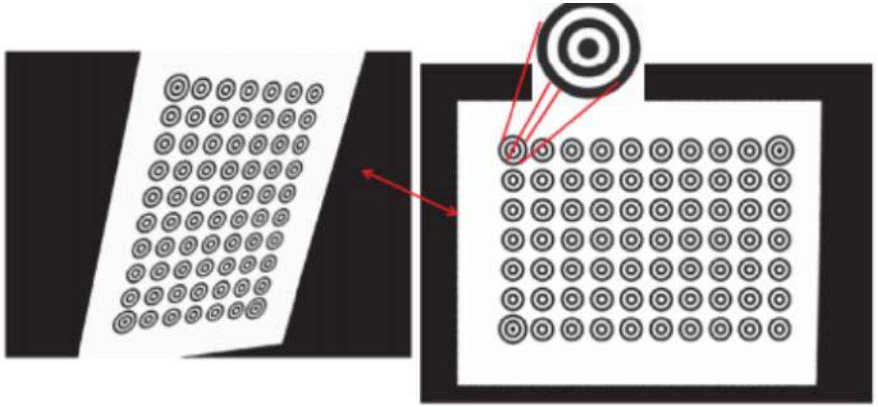
Radial distortion



Tangential distortion

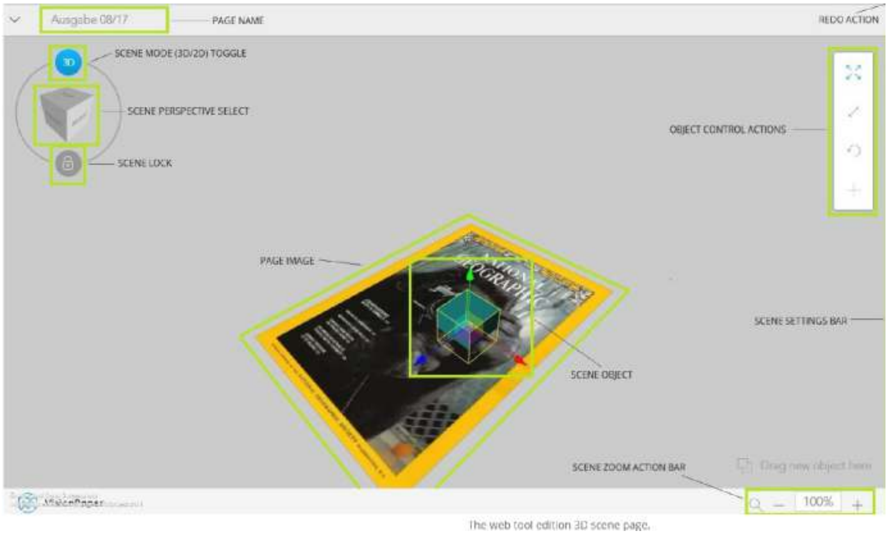


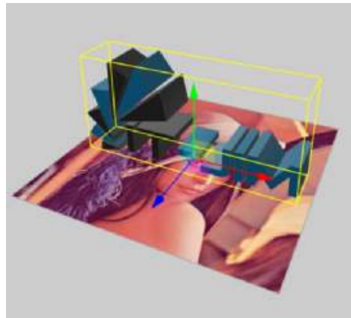
Checkboard



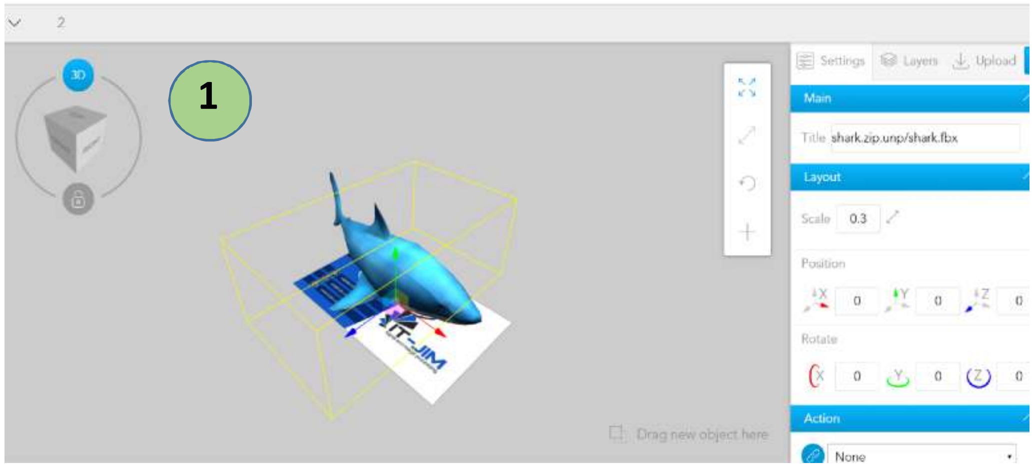
Calibration pattern example

Web AR Tool





AR WebTool



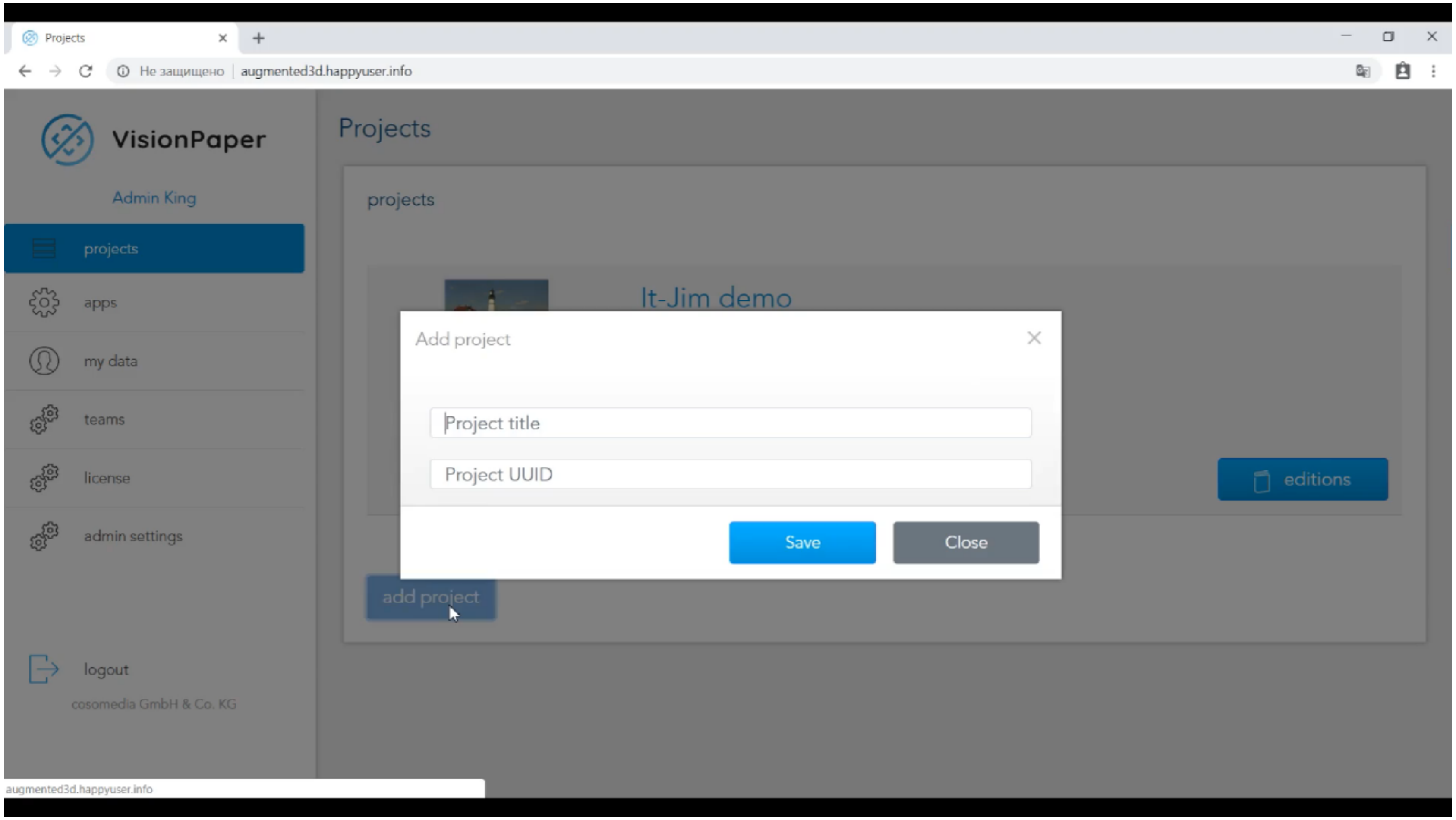
2

3D model in WebTool

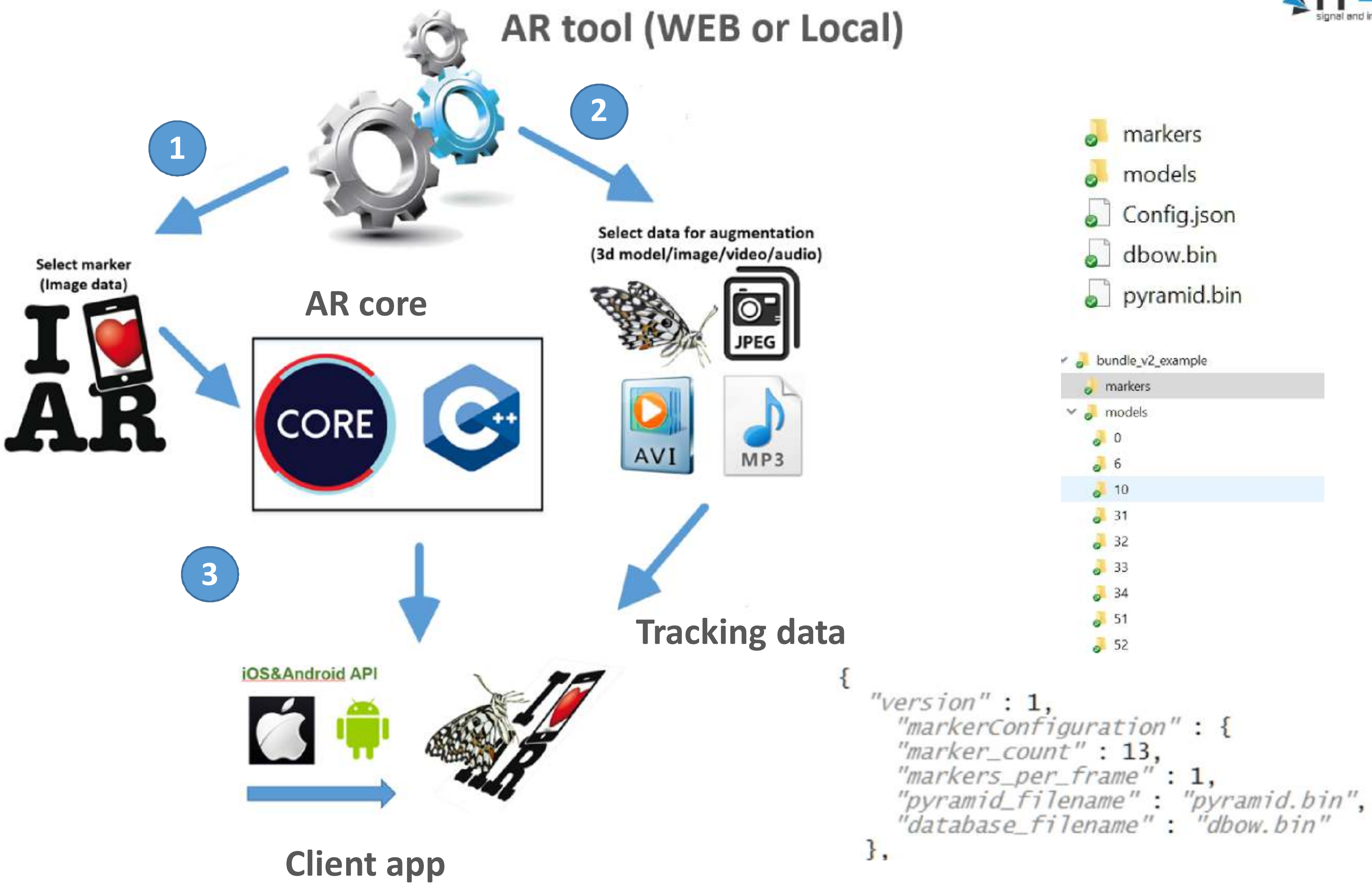
Mobile AR in production

3D model in mobile APP

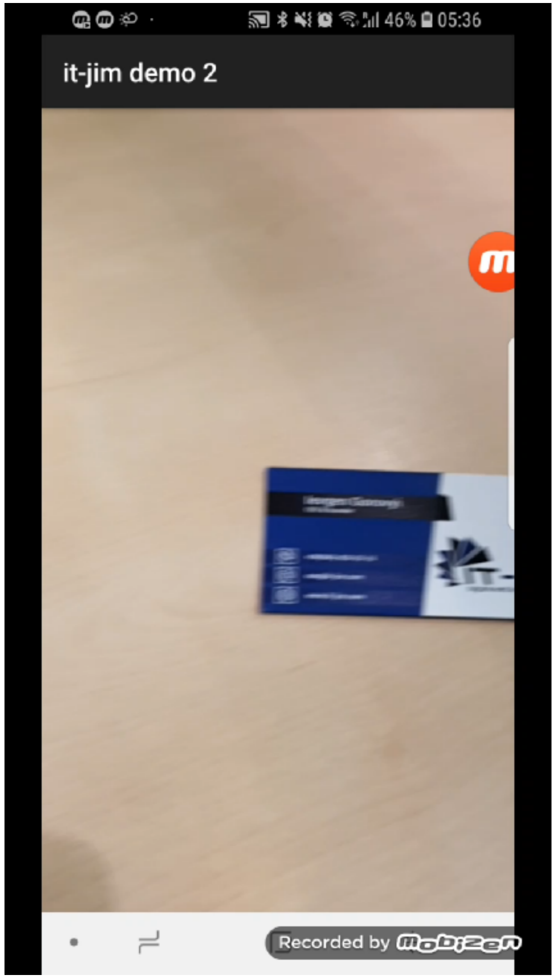
Web AR Tool: How it Works?



AR Data Flow: SDK Side



AR It-Jim: Buttons demo

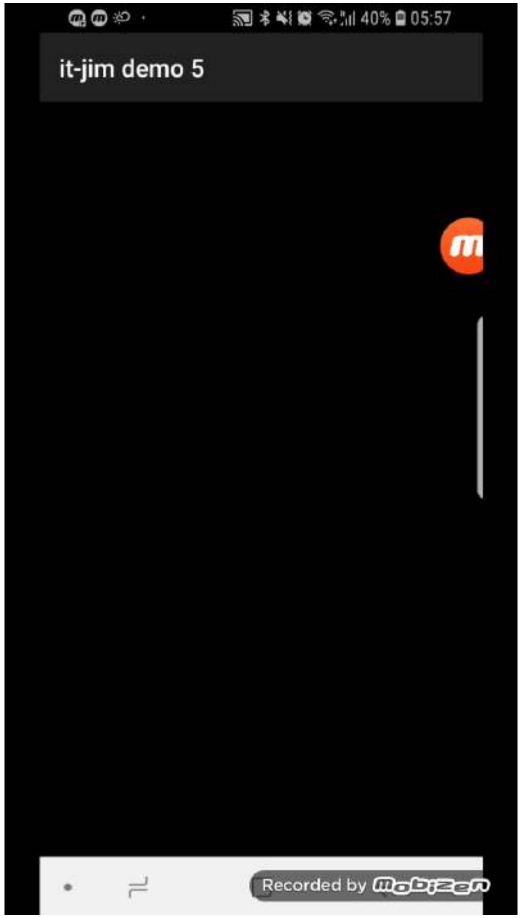


Buttons augmentation

Models augmentation

AR It-Jim: Cars Cube

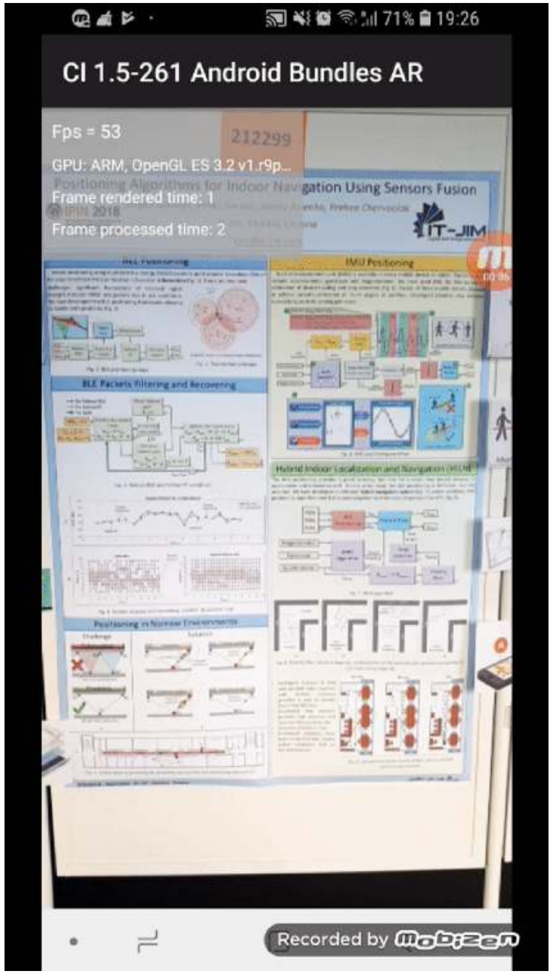
Screenshot

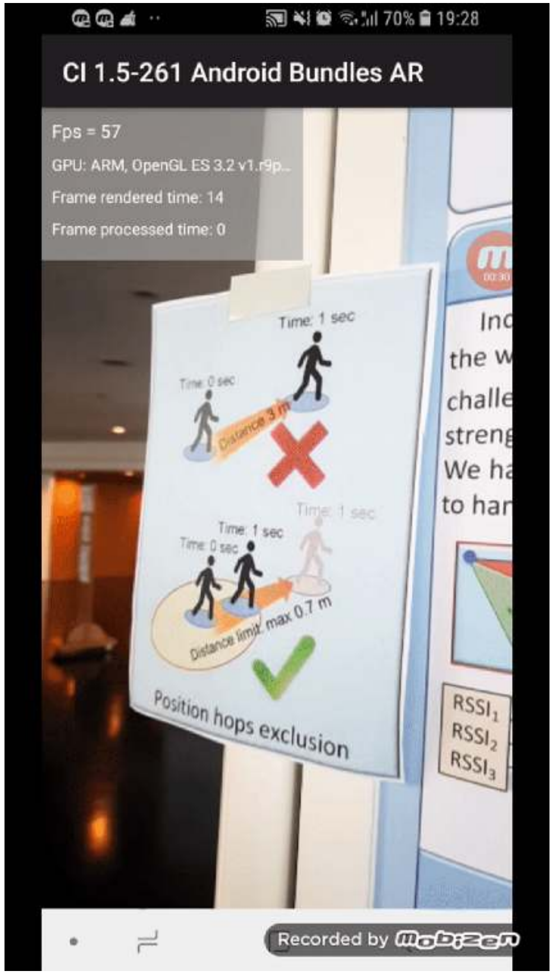


Cars cube example

AR Poster

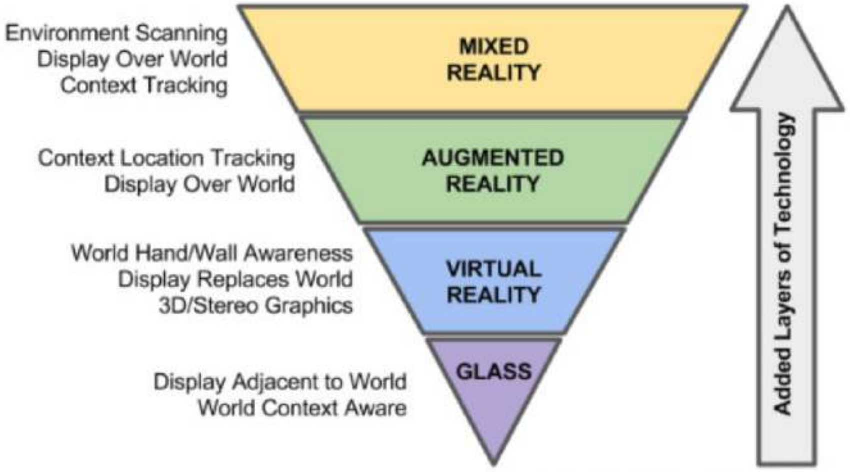
IPIN 2018

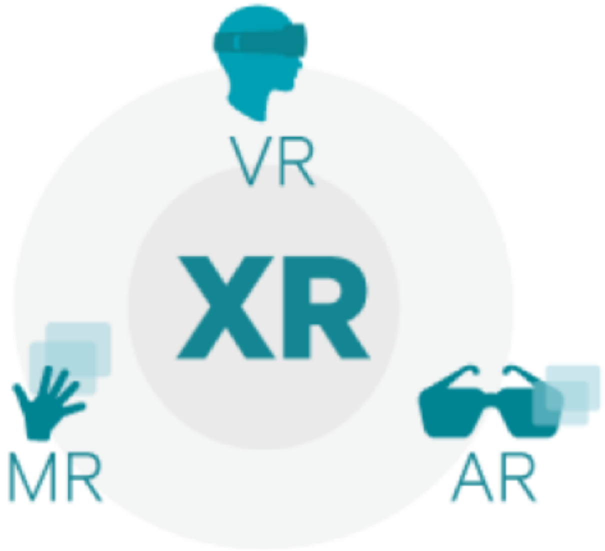




AR on poster

Toward XR


Environment Scanning
Display Over World
Context Tracking
MIXED REALITY
Context Location Tracking
Display Over World
AUGMENTED REALITY
World Hand/Wall Awareness
Display Replaces World
3D/Stereo Graphics
VIRTUAL REALITY
Display Adjacent to World
World Context Aware
GLASS
Added Layers of Technology


VR
XR
MR
AR


Fasten Halo
to Optical
Bracket

Robots vs Humans

Matt Novak
@paleofuture
Читать

That scary robot video is fake
gizmodo.com/the-viral-robo ...

7:46 - 20 авг. 2018 г.

Conclusions

- Augmented reality is truly immersive field
- AR is not only about data, AR is about robust computer vision solutions
- Real-time performance on mobile is challenging
- AR is cool, but still not ready for a mass market

AI UKRAINE 18

V INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND DATA SCIENCE APPLICATIONS

THANK YOU!
Questions?

ceo@it-jim.com