Introduction to Visual Simultaneous Localization and Mapping (VSLAM)

Amol Borkar

Product Marketing Director

Shrinivas Gadkari

Group Director of Engineering

Cadence Design Systems

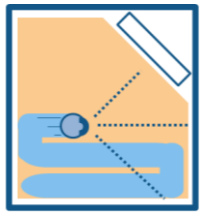
cadence®

Agenda

- What is SLAM?
- Applications of SLAM
- How it works
- Feature Detection Explained
- Combining different sensors
- Improving SLAM with AI?

What is SLAM

- What is SLAM short for?
  - Simultaneous Localization And Mapping
  - VSLAM => Visual SLAM
- One definition of SLAM?
  - “A method used for autonomous vehicles that lets you build a map and localize your vehicle in that map at the same time.”
- In simple terms:
  - Understand your (camera) 2D or 3D position in an environment and build a map of that environment
- SLAM is computationally complex
  - Previously implemented on “racks of server CPUs”
    - Darpa Grand Challenge, Darpa Urban Challenge
  - Consumer/desktop CPU have caught up in performance
    - Single core CPU implementations capable of real-time implementation around 2010



Typical Applications and Market Segments

ADAS

Underwater

Mixed Reality

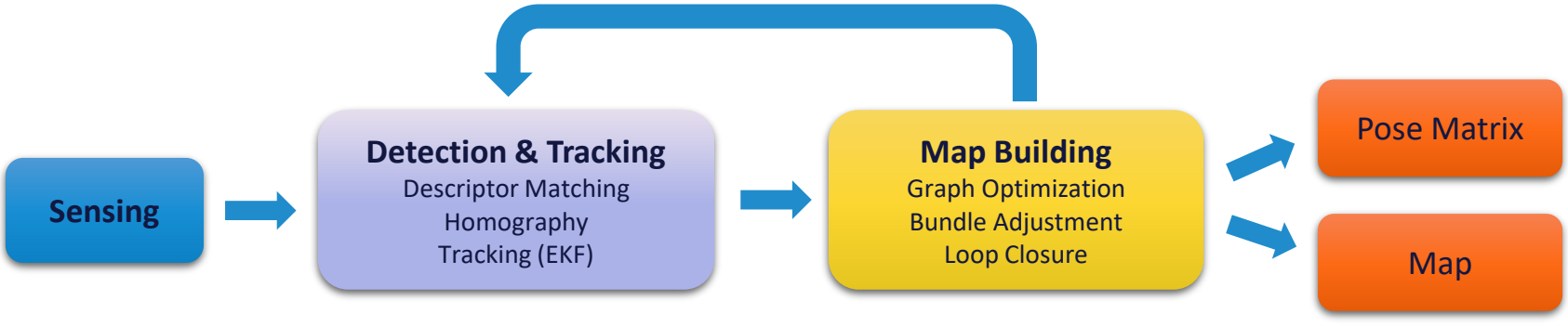
Robotics

Drones

Space

Any application that needs a 2D or 3D map of the environment

How Does SLAM Work?



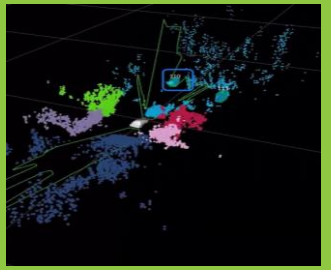
Cameras (Mono/Stereo)
IMU
Radar
Lidars

Front-end
Fixed Point Implementation
Requires real-time

Back-end
Typically in floating point
Can be offline

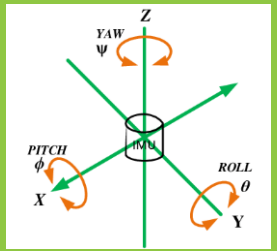
Various Options for Sensing

Camera



Radar

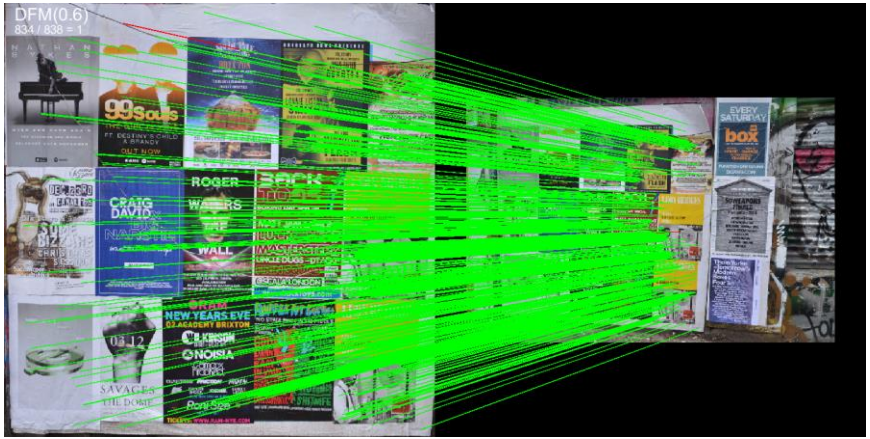
LiDAR

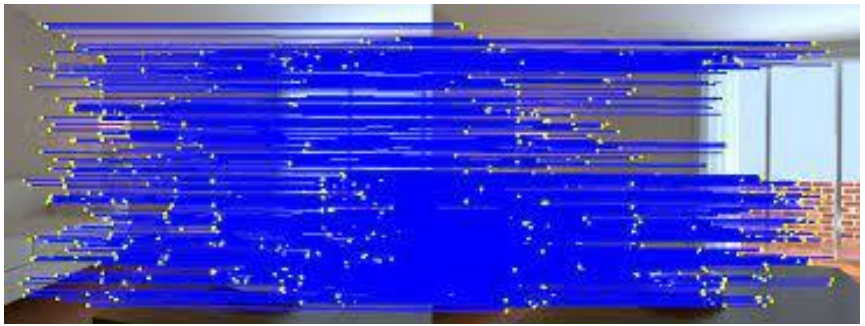


IMU

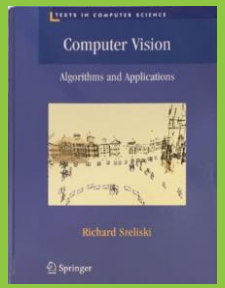
Feature Detection and Keypoint Matching

- What is this?
- What does it mean?
- What are all those lines?
- Let’s take a look at an example

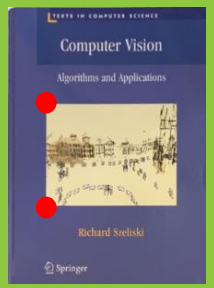


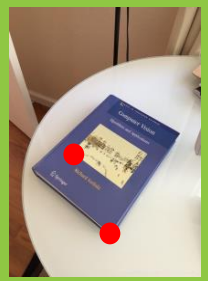


Step by Step: Feature Detection and Keypoint Matching



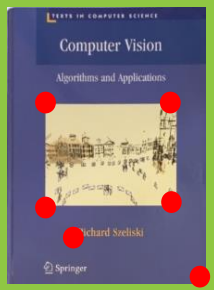
Start with 2 images



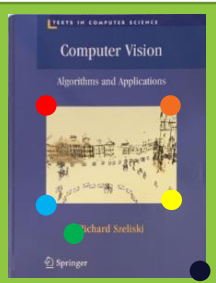


Find “feature points”

SIFT, SURF, Fast9, ORB

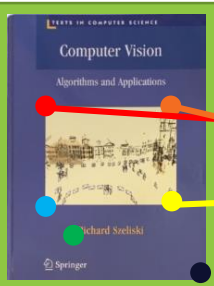


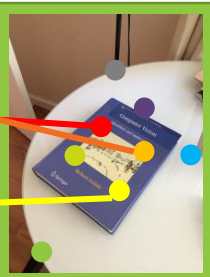
Find best 100 or 1000 points



Match features

Compare descriptors by hamming distance etc.

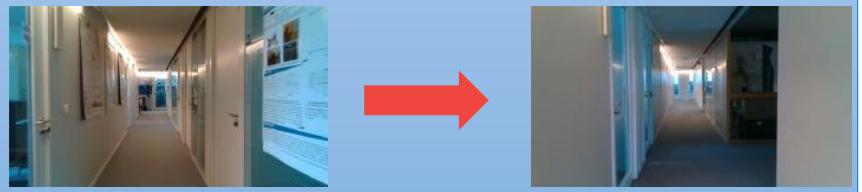




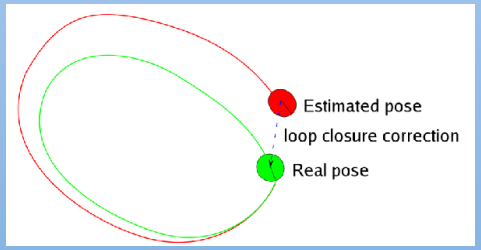
Draw Lines

How camera has moved

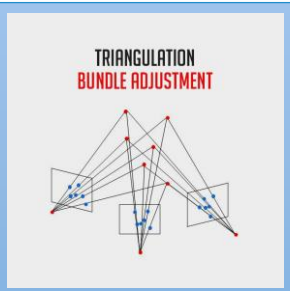
Other Commonly Used Terms



Key Frames
Sufficient movement/change between subsequent frames



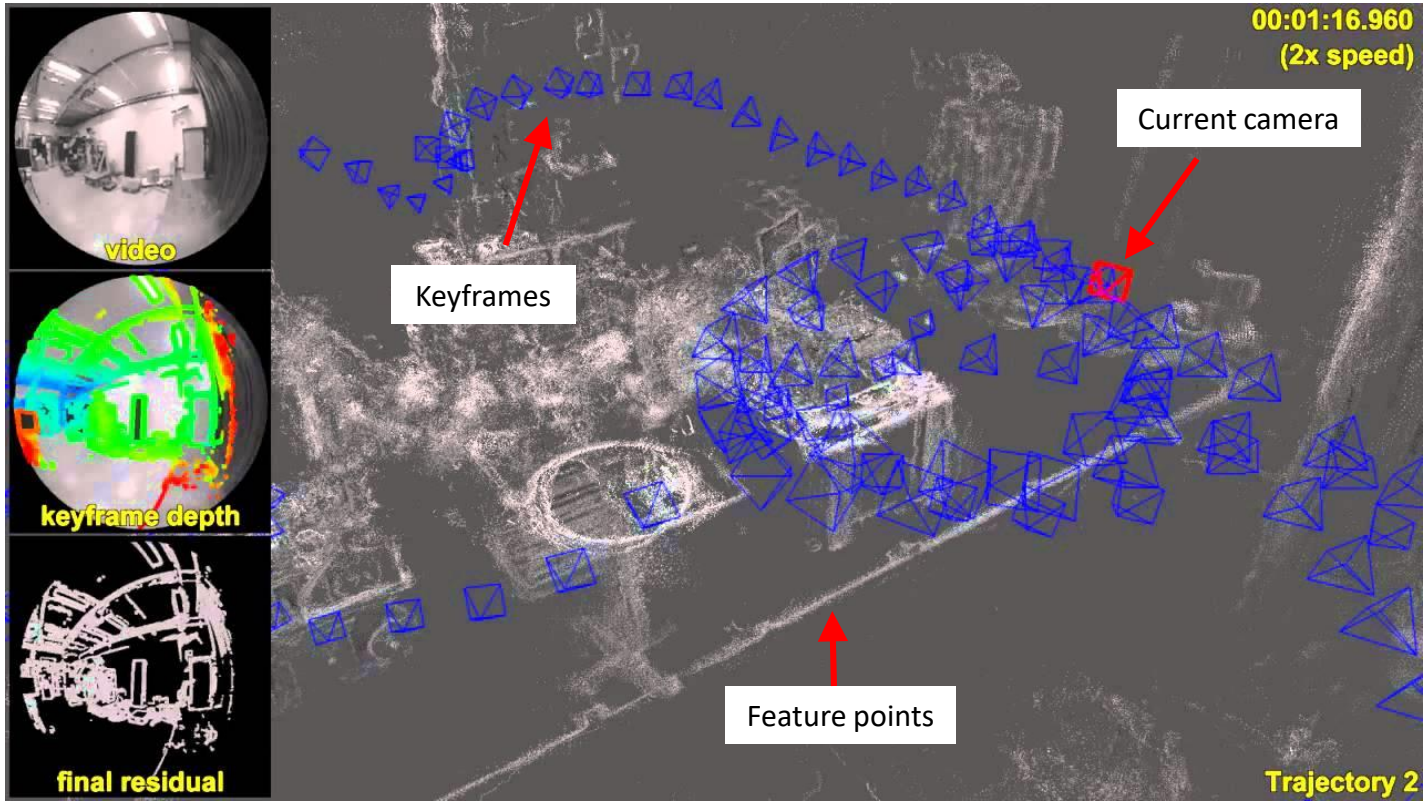
Loop Closure
Return to origin detected



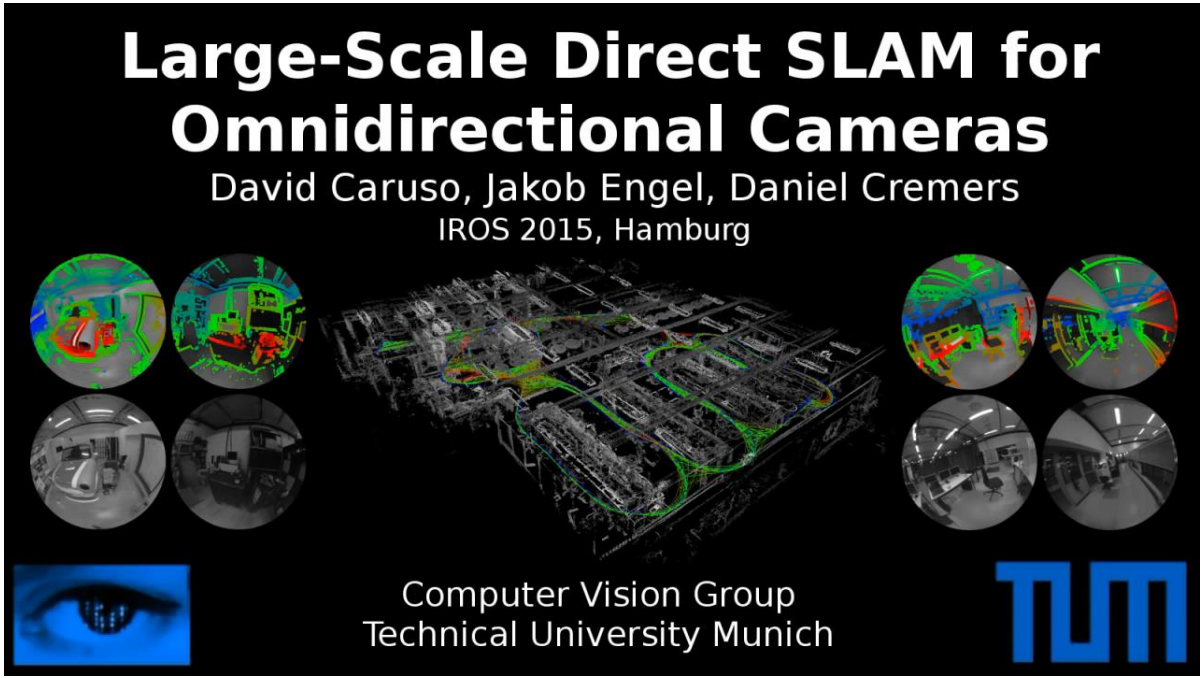
Bundle Adjustment
Redistribution of accumulated errors

Re-localization
Key Frame matching, or
“I know where I am because I have been here”

Putting it all together

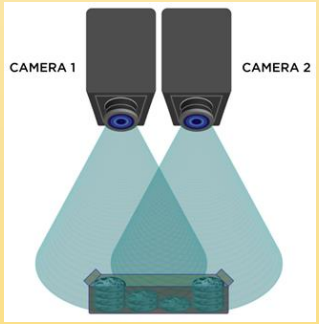


Putting it all together (2)

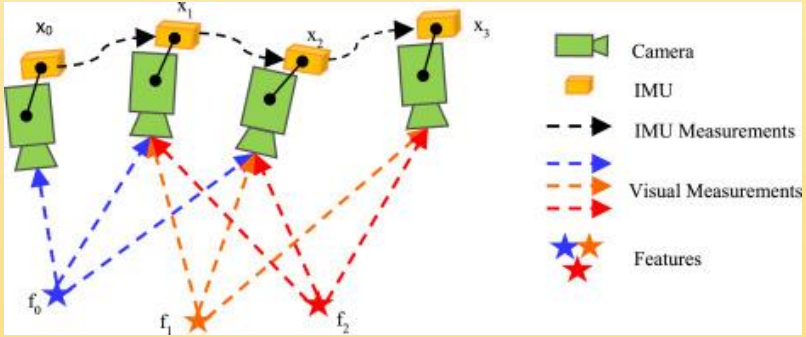


https://www.youtube.com/watch?app=desktop&v=v0NqMm7Q6S8

Combining Sensors to Make VSLAM Better?



2 or more Cameras



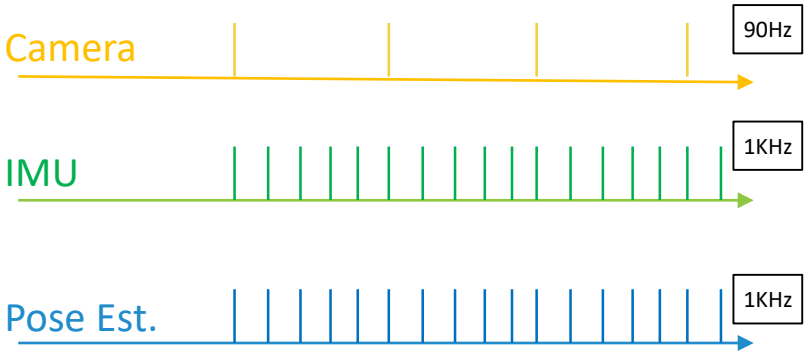
Camera + IMU

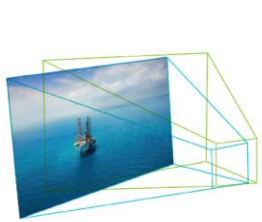
- Each sensor compensates for the other
- Camera has holistic view of environment
  - Calculate movement is relatively costly
  - Lower refresh rate

- High precision rotational movement with IMU
- Inertial Measurement Unit contains Gyro, Magnetometer, Accelerometer
  - IMU can refresh very fast

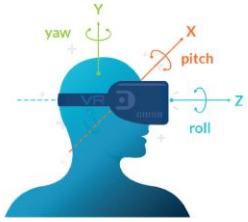
Combining 2 sensors for better results?

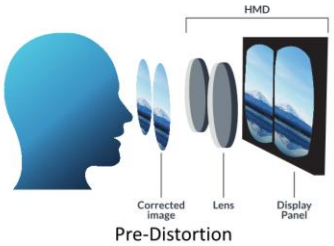
Use case of Camera + IMU





Re-Projection





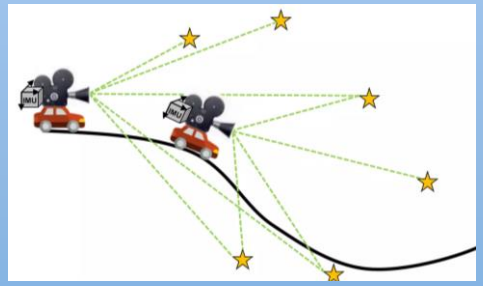
Pre-Distortion

- Camera FPS between 90-120Hz
- IMU FPS over 1KHz
- IMU can help track “micro-movements”
- Pose estimation at 1KHz (or higher)

Async Time Warp (ATW)

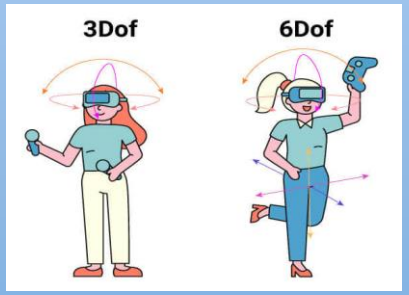
- Need extremely fast refresh rates, reduce dizziness
- GPU rendering pipeline may not achieve 90+ fps
- ATW rotates/re-projects/warps rendered frame (not creating a new render)
- Human brain gets impression of faster rendering

Other Commonly Used Terms



Visual Inertial Odometry (VIO)

Path estimation, no map creation



6 Degrees of Freedom (6DoF) = Pose Matrix

Translation (X,Y,Z) + Rotation (yaw, pitch, roll)

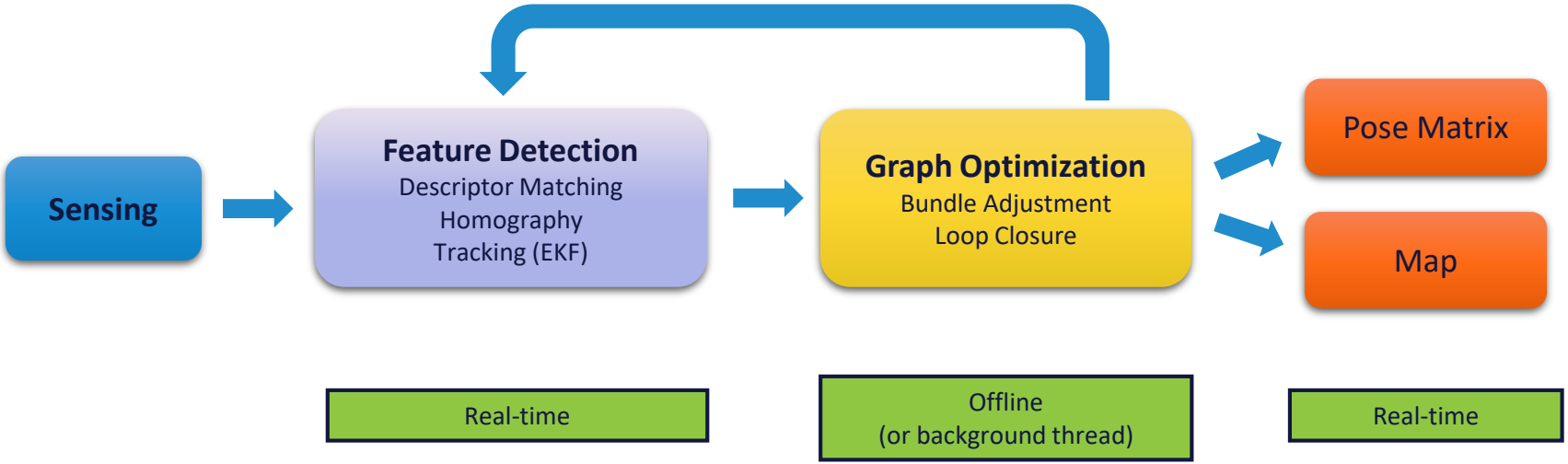
Inside-out-tracking

Position/location information is computed on the device

Outside-In-tracking

External sensors tell you your position

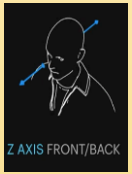
To be Real-time or Not to be?

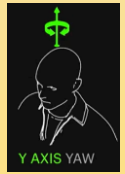


Application-based trade-offs?

Can we make trade-offs?

Gaming / XR



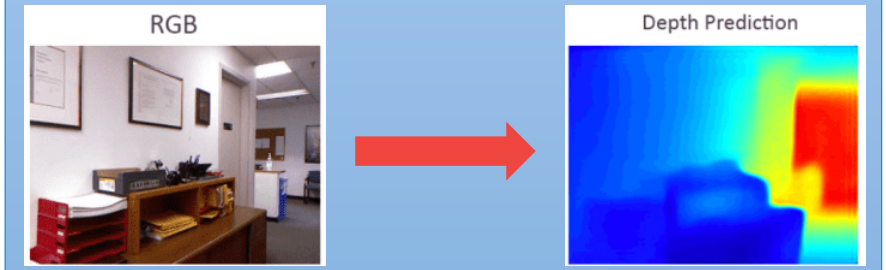


Vacuum Cleaner

- Very demanding specs:
  - 120FPS, 6DoF, ATW, 4K, 3D Graphics, Ray tracing
- Mostly real-time:
  - $1200+ (equipment only), PC $ extra

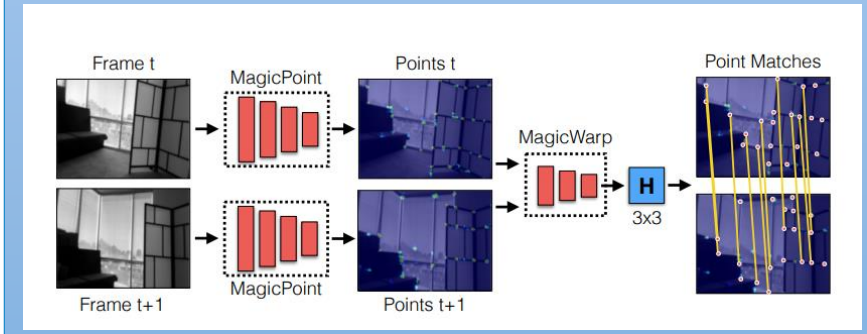
- Moderate specs:
  - 5-10FPS, no ATW and 3D graphics
  - 3DoF (X-Y Translation & Rotation)
- Lots of offline processing:
  - Approx. $400

Improving SLAM with AI?



Depth estimation and triangulation

Single camera-based depth estimation



Improved feature detection

CNN based approaches replacing SIFT, SURF

Summary

- Introduction to SLAM
- High level flow and different sensor options
- Example of Feature Detection and Keypoint matching
- Example of combining various sensors
- Trade offs for different applications
- Sufficient knowledge to understand mechanics of SLAM

Resources

- What Is SLAM?
- Location in the Visual World
- CMU Tartan Racing Wins DARPA Urban Challenge
- A vision system based real-time SLAM applications
- Motor Controllers for Autonomous Mobile Robots
- 5 Ways XR Consulting Can Help Your Business
- Monocular SLAM
- All-Environment 360 Perception
- Design and implementation the stability and direction of hexapod robot motion
- Velodyne Lidar
- 4D Imaging Radar Overview

2024 Embedded Vision Summit

- **Cadence Booth**
  #518 Exhibition Hall

Resources (2)

- DFM: A Performance Baseline for Deep Feature Matching
- FastORB-SLAM: Fast ORB-SLAM method with Coarse-to-Fine Descriptor Independent Keypoint Matching
- Computer Vision: Algorithms and Applications by Richard Szeliski
- Homography Transform — Image Processing
- Indoor visual SLAM dataset with various acquisition modalities
- How Does Triangulation in Photogrammetry Work?
- What's the relationship between "data association" and "loop closure" in SLAM?
- Stereo Vision for 3D Machine Vision Applications
- 3D World = 6 Degrees of Freedom
- Visual–inertial state estimation with camera and camera–IMU calibration
- Async Time Warp
- A Beginner's Guide To Virtual Reality
- Inside-out versus Outside-in tracking for HMDs (Acer, 2019)
- Vehicle Positioning in the Absence of GNSS Signals: Potential of Visual-Inertial Odometry
- Popular VR Party Games That Get Competitive
- Benefits of Having a Robot Vacuum For Your Home
- Magic Leap Researchers Reveal "Deep SLAM" Tracking Algorithm
- Monocular depth estimation