Concept & background

Robust Aged Feature set

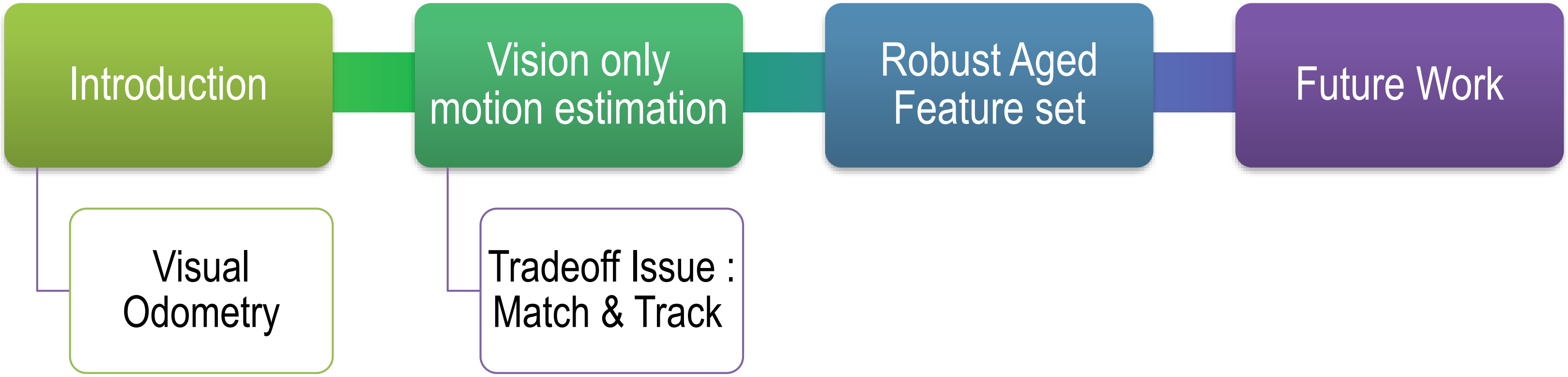
for Visual Odometry

ISL Lab Seminar

Jin Hyung Kim

2015.05.15.

Contents



Introduction

Vision only motion estimation

Robust Aged Feature set

Future Work

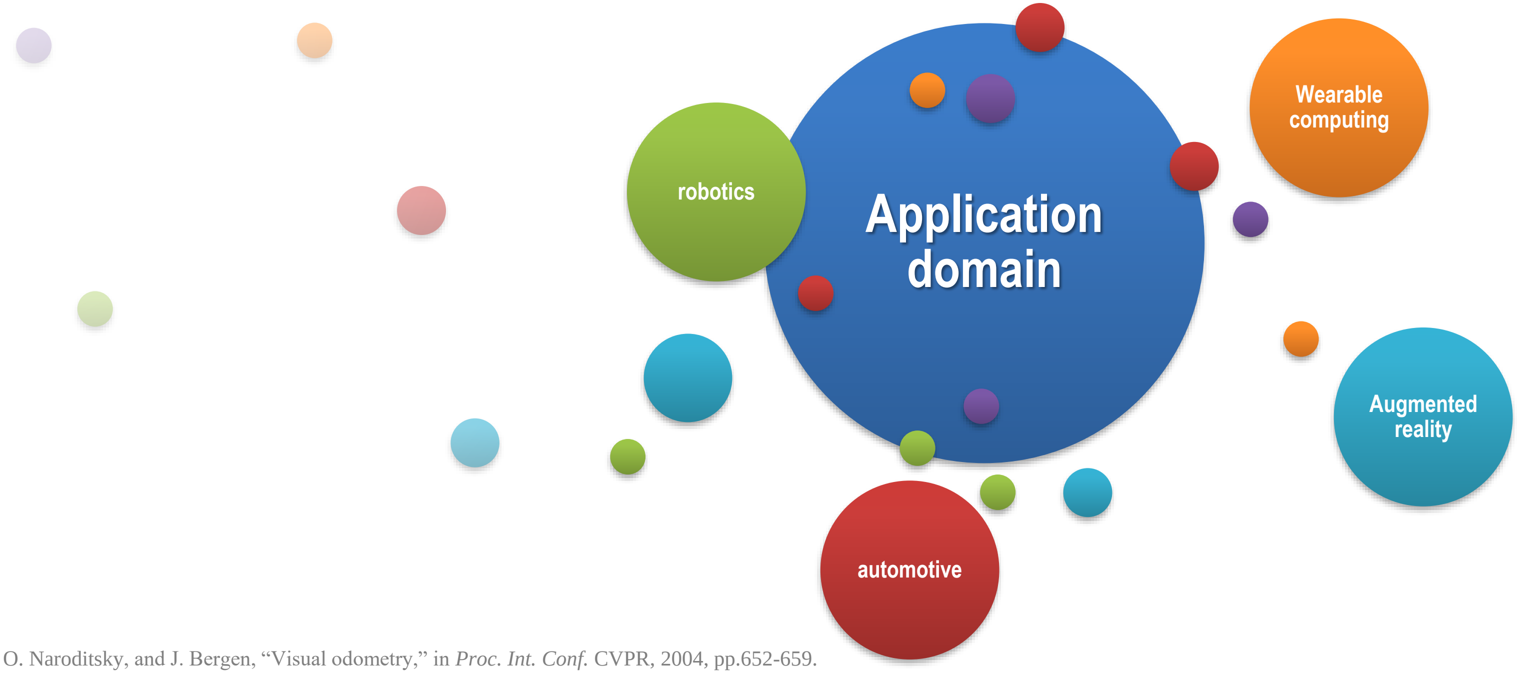
Visual Odometry

Tradeoff Issue : Match & Track

Introduction : Visual Odometry

Visual Odometry

- The process of estimating the egomotion of an agent using only the input of a single or multiple cameras attached to it
- The term VO was coined in 2004 by Nister in his paper*
  - Was chosen for its similarity to wheel odometry

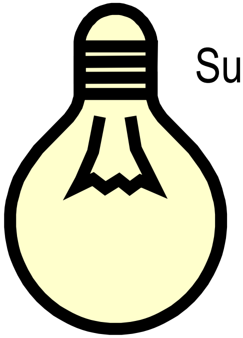


*D. Nister, O. Naroditsky, and J. Bergen, “Visual odometry,” in *Proc. Int. Conf.* CVPR, 2004, pp.652-659.

Visual Odometry

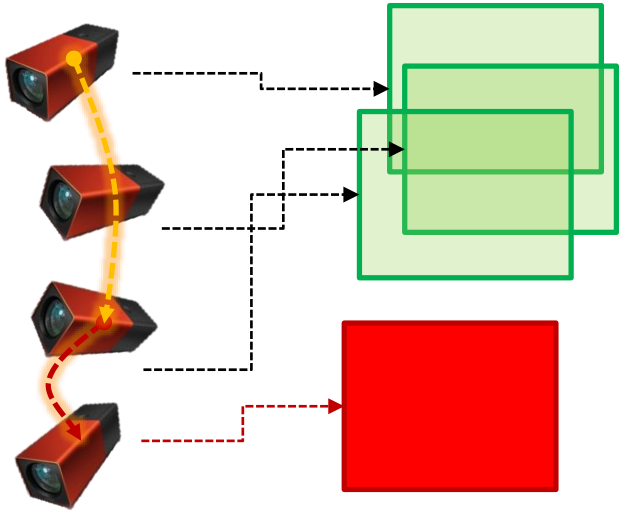
- **Advantages of VO**
  - Is not affective by wheel slip in uneven terrain or other adverse conditions
  - Provides accurate trajectory estimates
  - Additional near space information acquisition
    - ➢ IMU, GPS, Wheel Odometry : egomotion only
  - Low cost comparing to IMU, Laser Odometry
  - Capable in GPS-denied environments
    - ➢ Underwater, Aerial, indoor, another planet

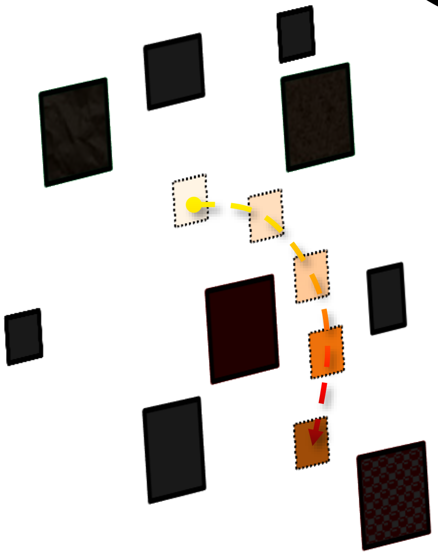
Positive condition for VO



Sufficient illumination

Consecutive frames
that have sufficient scene overlap





Static scene with enough textured features

VO Pipeline (2D to 2D)

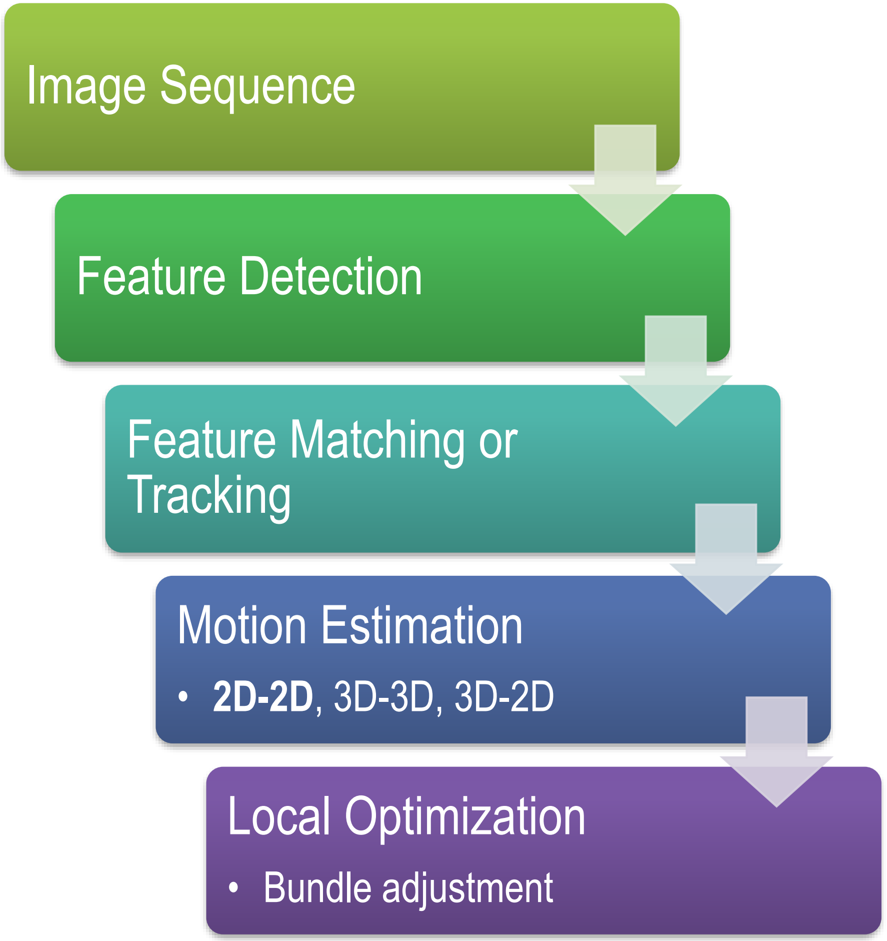


Image Sequence

Degradation(Noise, Gain), Calibration(Lens Distortion)

Feature Detection

SIFT, SURF, Harris Corner, GFTT, FAST, etc.

Feature Matching or Tracking

Matching(descriptor), Tracking(Optical Flow)

Motion Estimation

- **2D-2D**, 3D-3D, 3D-2D

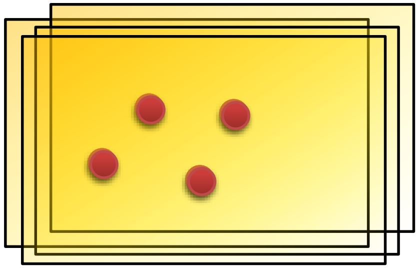
Essential matrix from image feature correspondences

Local Optimization

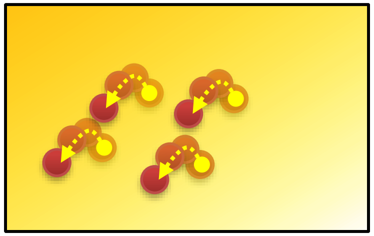
- Bundle adjustment

Iterative refinement to increase accuracy of trajectory

Tradeoff : baseline & correspondence

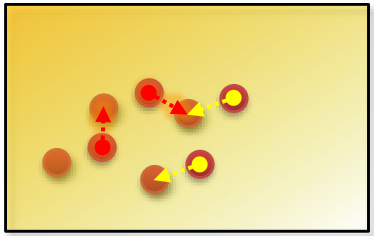


Camera movement



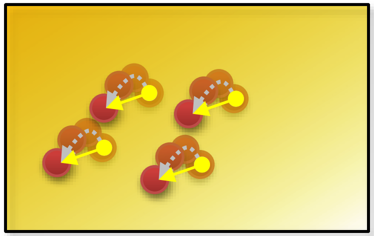
Feature Tracking

Short baseline : low ME accuracy
Guaranteed feature correspondence



Feature Matching

long baseline : High ME accuracy
Poor feature correspondence
(even heavy computation for descriptor)

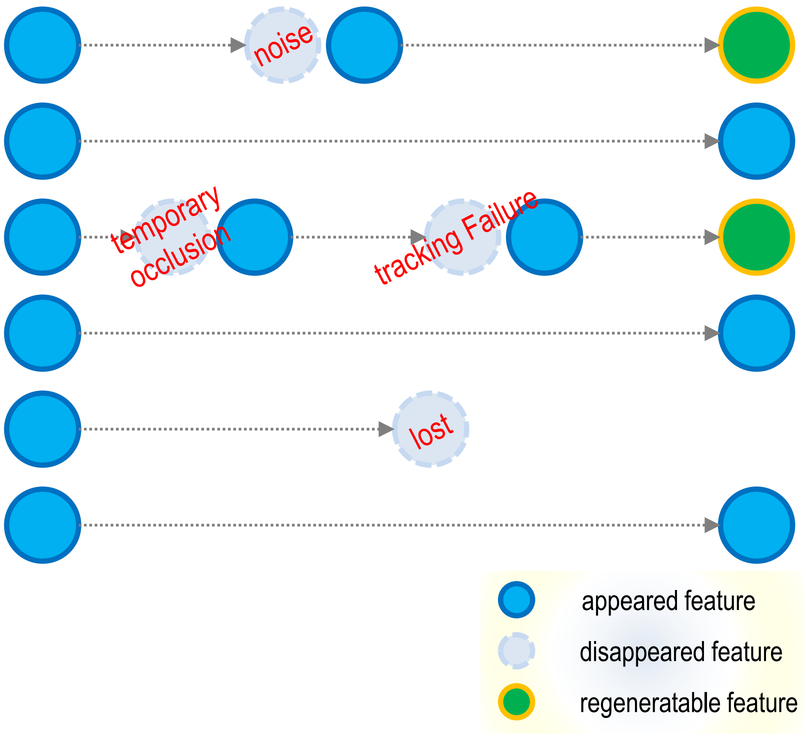


Match after track

long baseline : High ME accuracy
Guaranteed feature correspondence

Limit of "Match after Track"

- Feature blinking problem : decreasing number of feature



Robust Aged Feature set

- **How to obtain accurate Essential matrix.**
  - Sufficient reliable correspondences
    - ➢ Repeatable & Traceable feature extraction(Aging & tracking)
    - ➢ Adding feature of new part of scene(Feature detection & matching)
    - ➢ Outlier rejection
  - Sufficient long baseline length
    - ➢ Sufficient pixel movement
- **VO : Real–time**
  - Low-complex feature detector
  - Low-computational descriptor
  - Simple tracking algorithm

**Increase the number of features as many as possible &
Select the reliable features and keep traceable…**

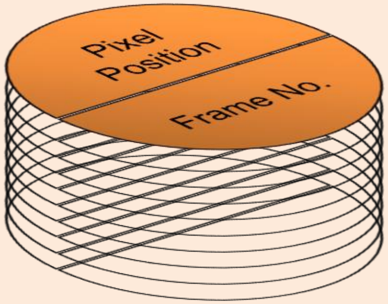
Robust Aged Feature set

- **Structure of RAF**

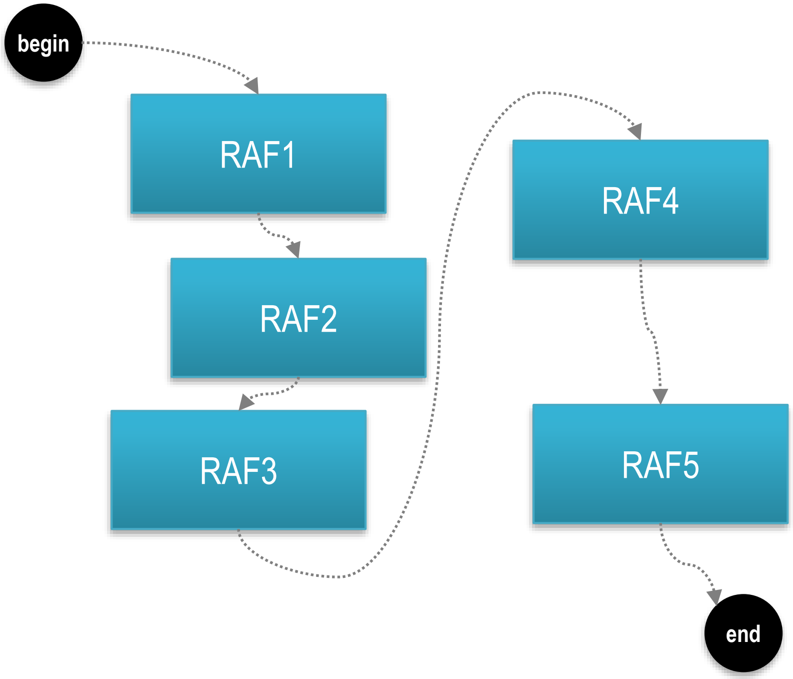
Descriptor : Local histogram of gradient

Age

Trajectory vector



- **RAF set**
  - List formation
    - ➢ to erase & add frequently



VO Procedure using RAFset

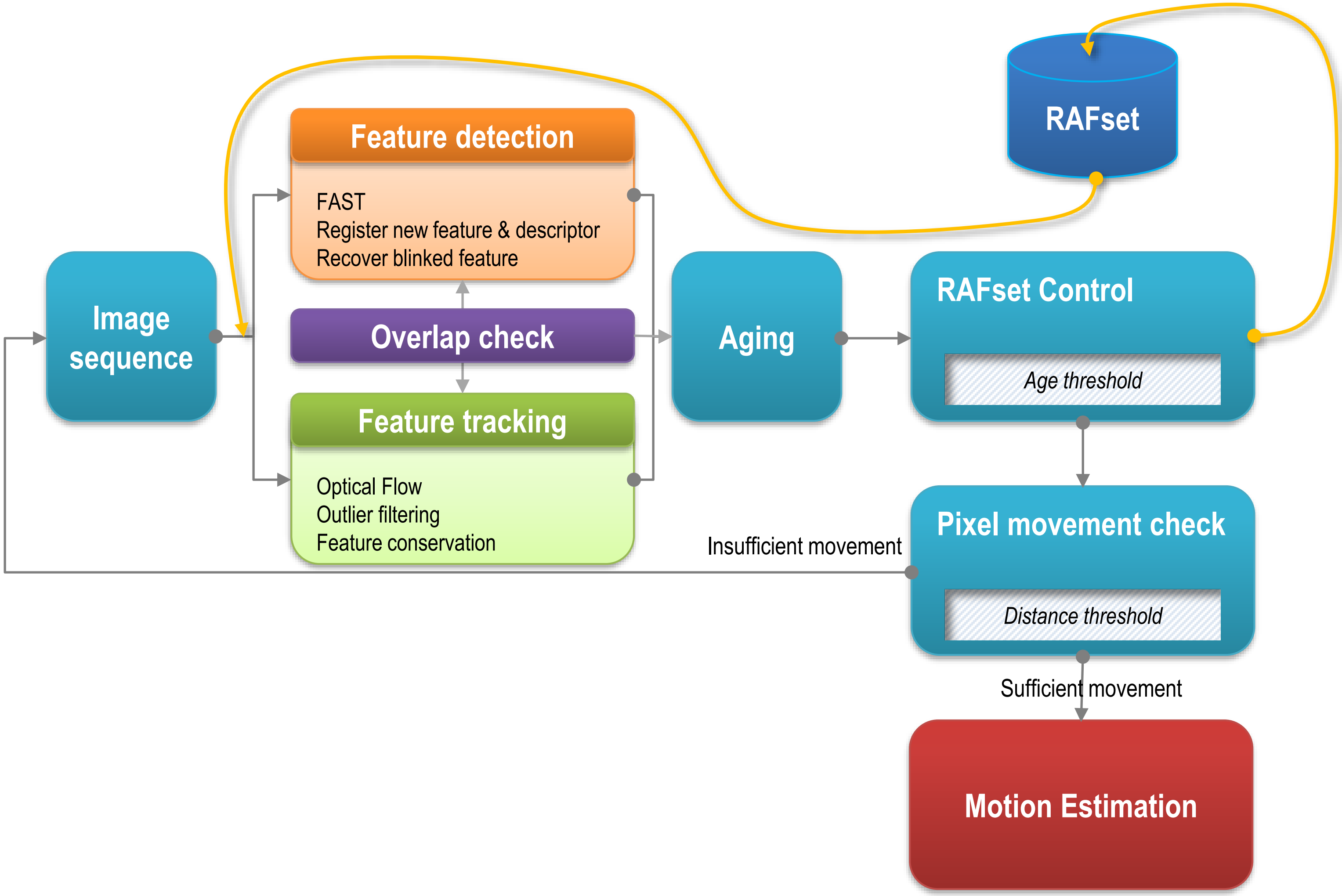


Image sequence

Feature detection

FAST
Register new feature & descriptor
Recover blinked feature

Overlap check

Feature tracking

Optical Flow
Outlier filtering
Feature conservation

Aging

RAFset

RAFset Control

Age threshold

Pixel movement check

Distance threshold

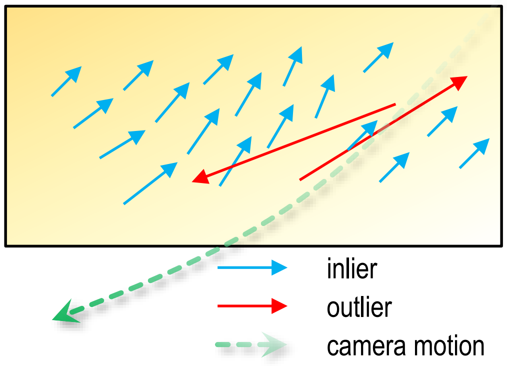
Insufficient movement

Sufficient movement

Motion Estimation

Outlier Filter for Optical Flow output

- **Correspondence outlier**



Caused by
Noise, blinking feature, moving object, etc.

- **Outlier filtering**

$$e_i = \sqrt{(x_i - x_i')^2 + (y_i - y_i')^2}$$

$$where,\ (\mathrm{x},\mathrm{y}) \xleftrightarrow[correspondence]{} (x', y')$$

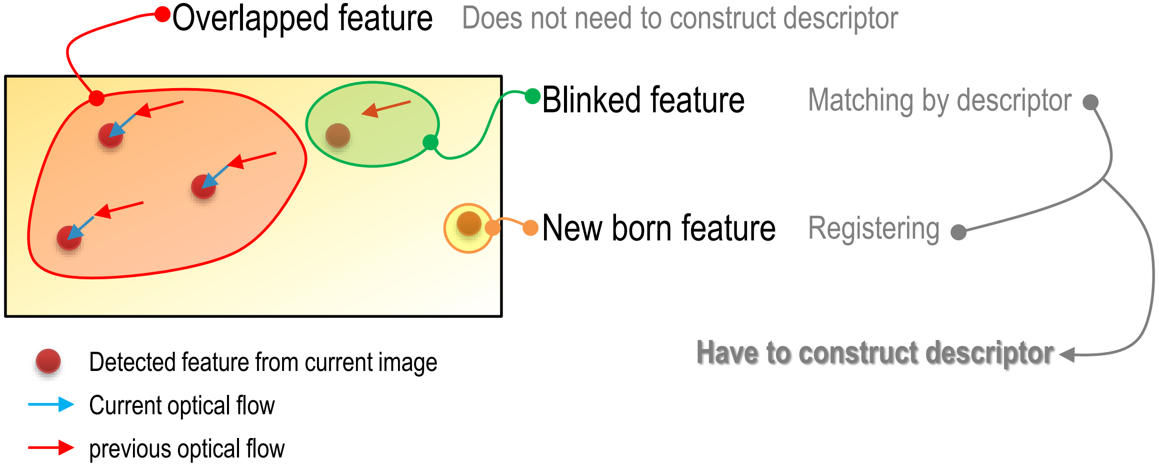
$$\bar{e} = \sum_{i=0}^{n} e_i, \qquad \sigma_e = \sqrt{\sum_{i=0}^{n} e_i^2 - \bar{e}^2}$$

$$T_{outlier} = \bar{e} + \sigma_e$$

Inliers should be smaller than $T_{outlier}$

Overlap check

In order to reduce computation, check overlapped features.
(Near distance)



Overlapped feature — Does not need to construct descriptor

Blinked feature — Matching by descriptor

New born feature — Registering

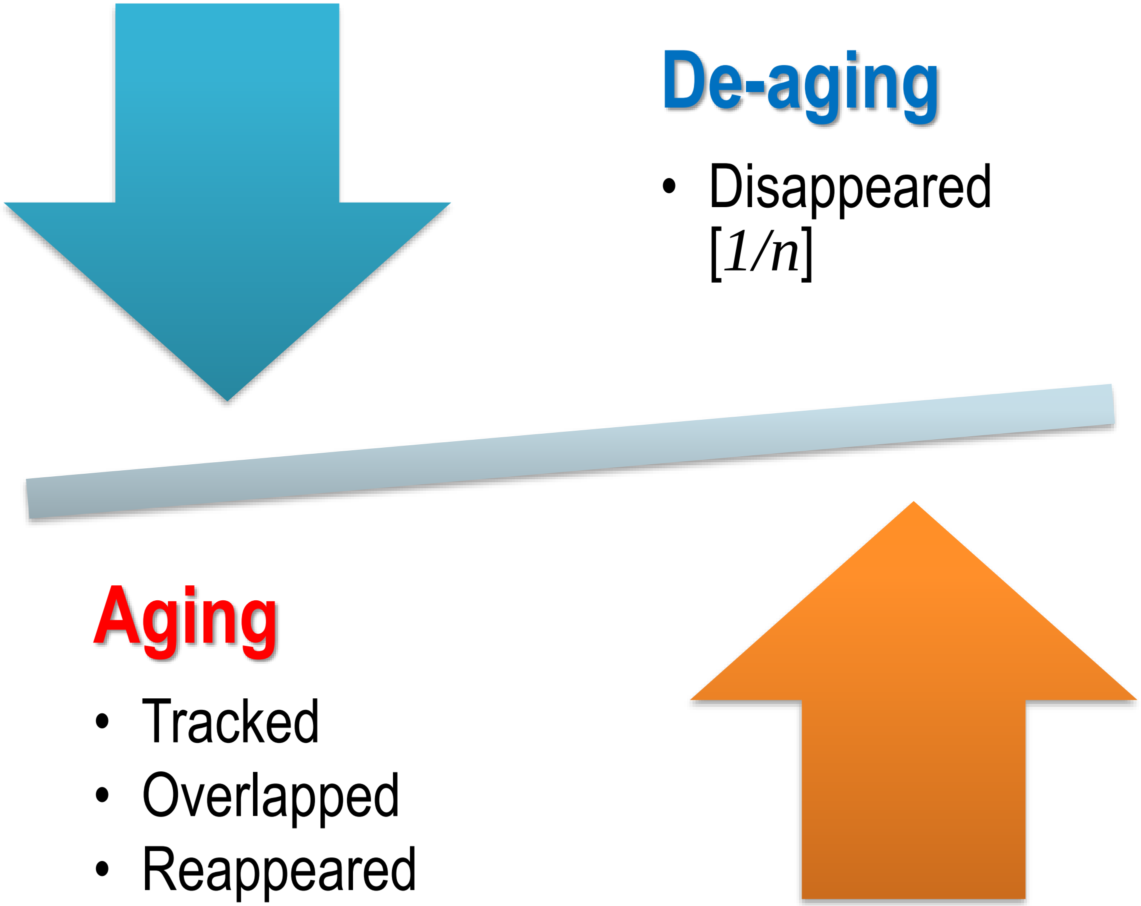
Have to construct descriptor

● Detected feature from current image

→ Current optical flow

→ previous optical flow

Aging Strategy



De-aging

- Disappeared [*1/n*]

Aging

- Tracked
- Overlapped
- Reappeared

Stop state detection

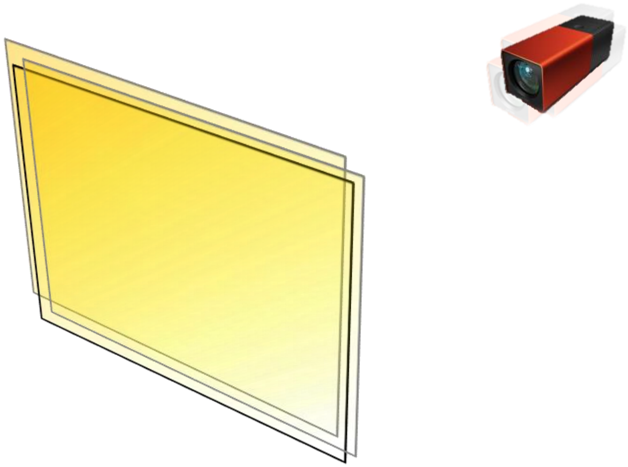
- **Even static situation, Motion vector occurs in image**
  - Vibration of camera mount
  - Scattering noise
  - Illumination blinking

- **Using average amplitude of Motion vectors**

$$e_i = \sqrt{\left(x_i - x_i^{'}\right)^2 + \left(y_i - y_i^{'}\right)^2} \qquad where,\ (\mathrm{x}, \mathrm{y}) \xleftrightarrow[correspondence]{} \left(x^{'}, y^{'}\right)$$

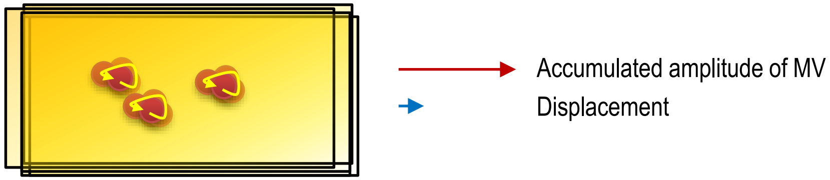
$$\bar{e} = \sum_{i=0}^{n} e_i$$

$$if \quad \bar{e} < T_m \text{ , } than\ the\ cam\ is\ stopped$$



Sufficient pixel movement detection

- **Also, using Accumulated Average amplitude of MVs**
  - However, amplitude of MV is scalar.
    - ➢ Recognizes periodic movement to large movement



  - Therefore it is required to use displacement btw. Initial frame to current frame
  - In order to reduce computation load,
    - ➢ Calculates displacement when sufficient amplitude of MV accumulated.
    - ➢ Not per frame

Future work

- Descriptor management
- Scale problem
- Visualization
- Experiment