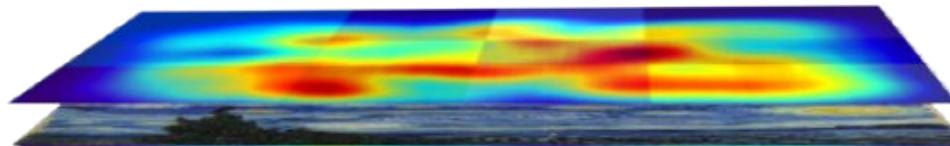


Machine Learning-based Indoor Localization for Micro Aerial Vehicles



Volker Strobel
volker.strobel87@gmail.com

14th July 2016

Radboud Universiteit

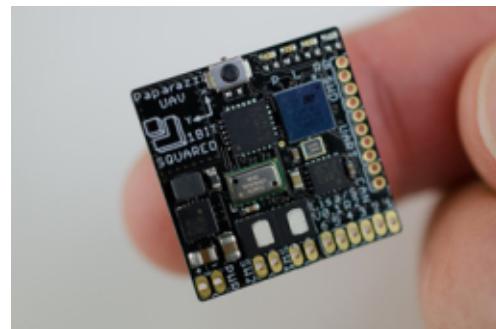


Louis Vuurpijl

TU Delft
Delft
University of
Technology

Guido de Croon
Roland Meertens

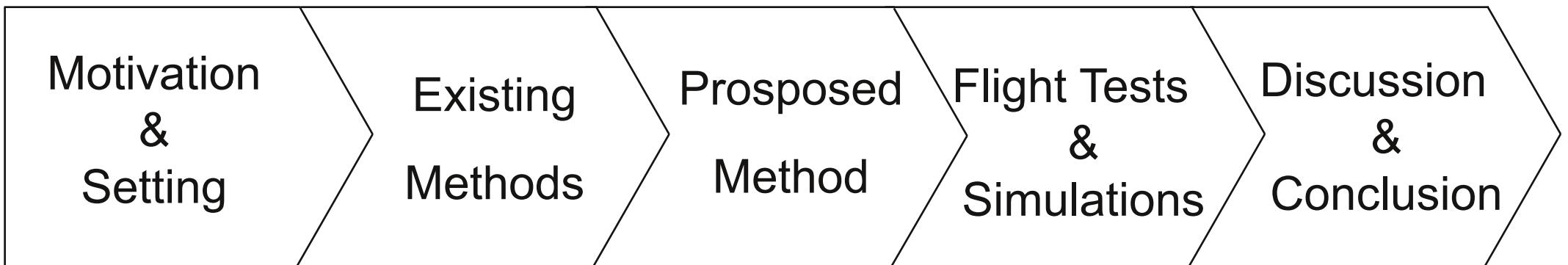
Micro Air Vehicle Lab



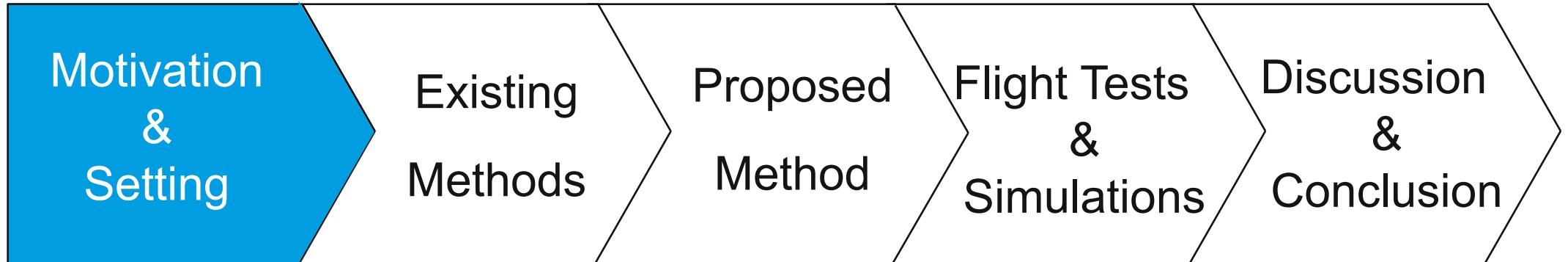
Miniaturization



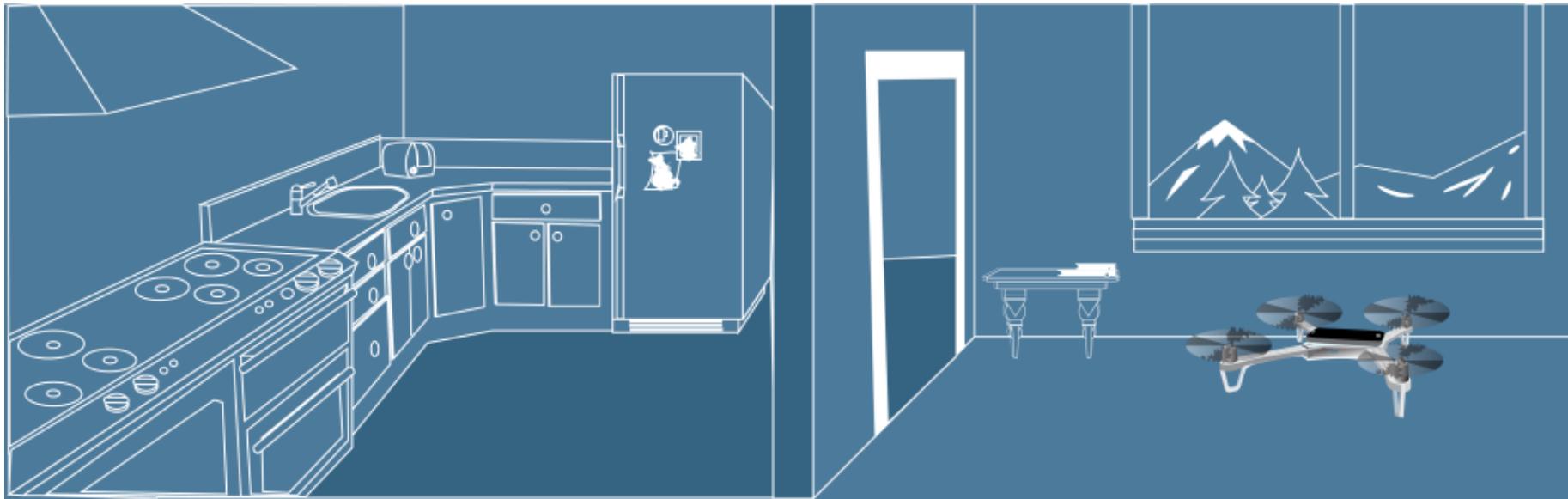
OUTLINE



OUTLINE



MOTIVATION



MOTIVATION

Tools



Environment



modifiable

known

fixed

planar

MOTIVATION

Tools



Environment



modifiable

known

fixed

planar

MOTIVATION

Tools



Environment



modifiable

known

fixed

planar

x,y-coordinates

accurate

on-board

real-time

fixed height

MOTIVATION

Research Question

Can vision-based indoor localization be done on a limited platform?

accurate

on-board

real-time

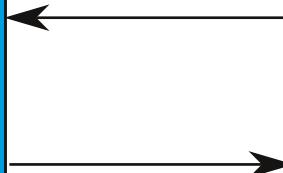
RESEARCH QUESTIONS

Research Question 1

Can vision-based indoor localization be done on a limited platform?

Research Question 2

Can we predict the suitability of an environment for the proposed localization algorithm?

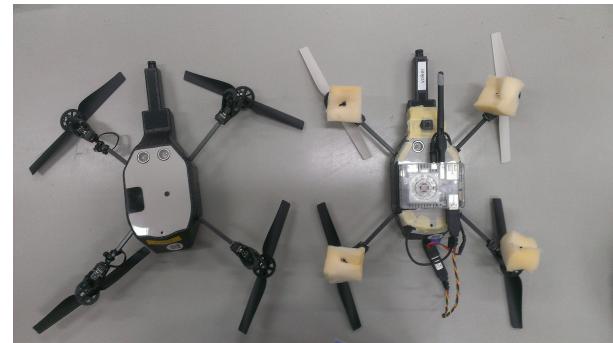


CHALLENGES

Low-performance
platform

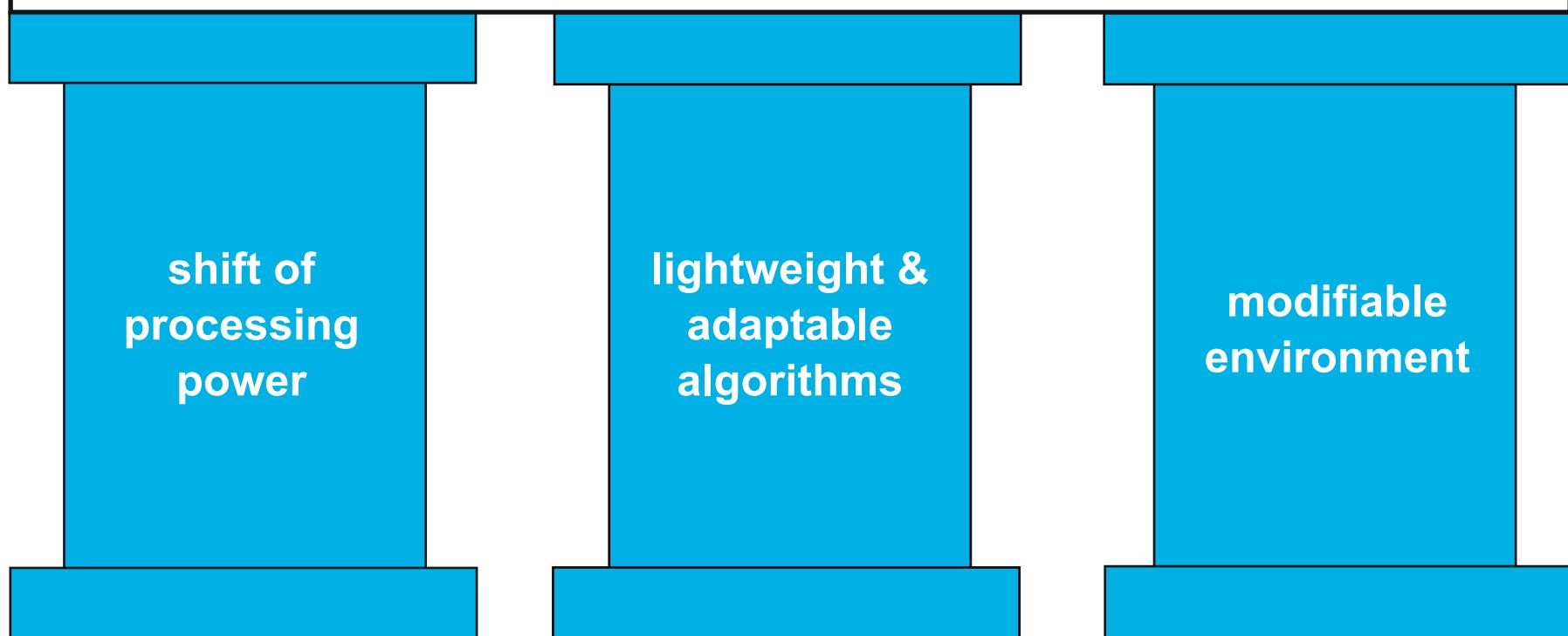


Low-level embedded
programming (C)

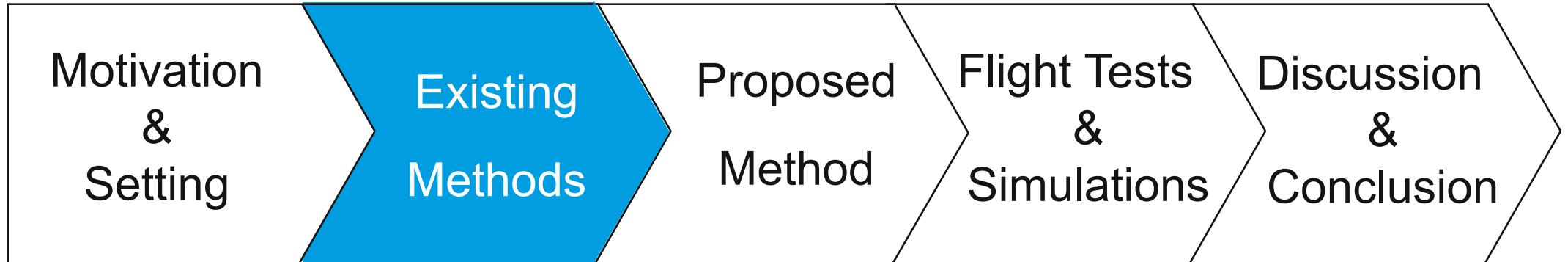


Real-world

EFFICIENT INDOOR LOCALIZATION



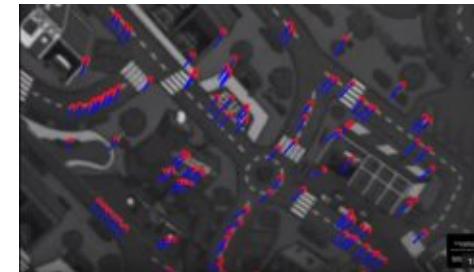
OUTLINE



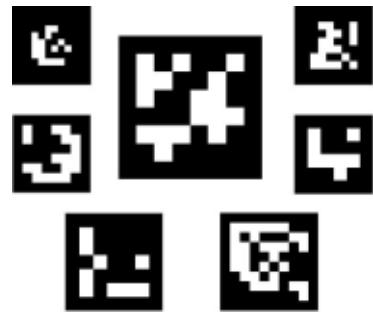
METHODS FOR ONBOARD LOCALIZATION



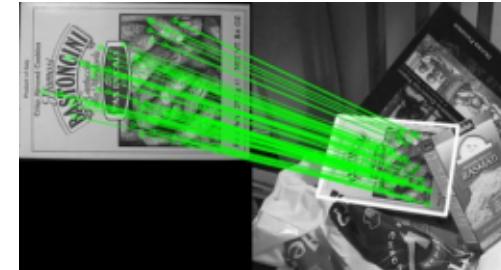
Laser range finder



Optical flow



Markers



Homography finding

DUMMY: INSERT SIFT VIDEO

APPROACH

Flight phase



APPROACH

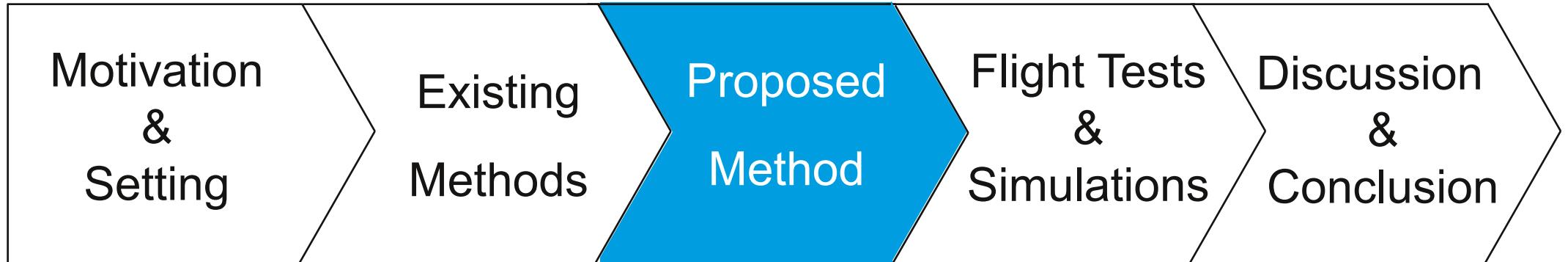
Flight phase



1 Image / sec



OUTLINE



APPROACH

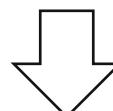
Flight phase



1 Image / sec



APPROACH

 ~~Flight phase~~
Pre-flight phase



→ **Dataset**



x y

200 300

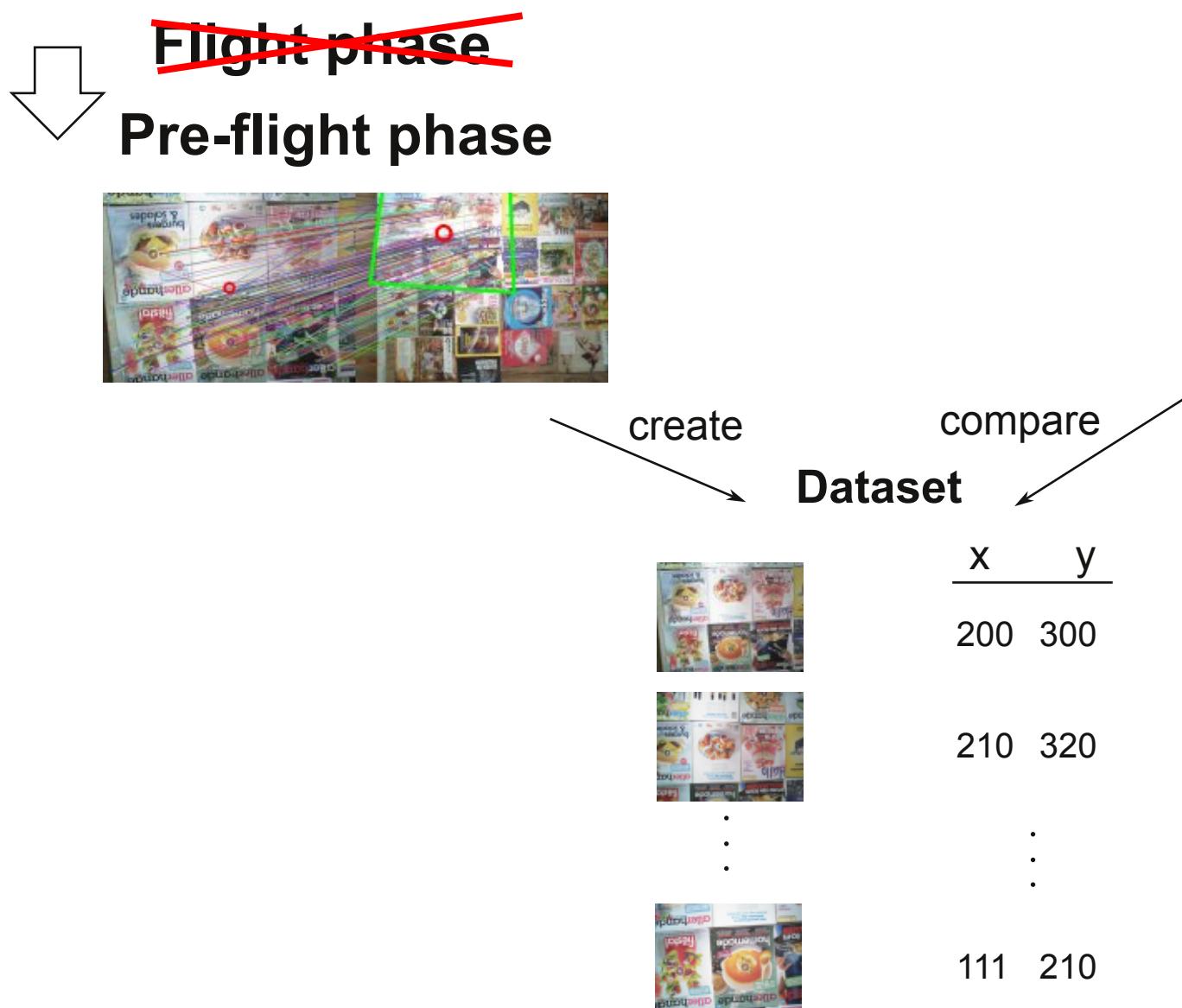


210 320



111 210

APPROACH

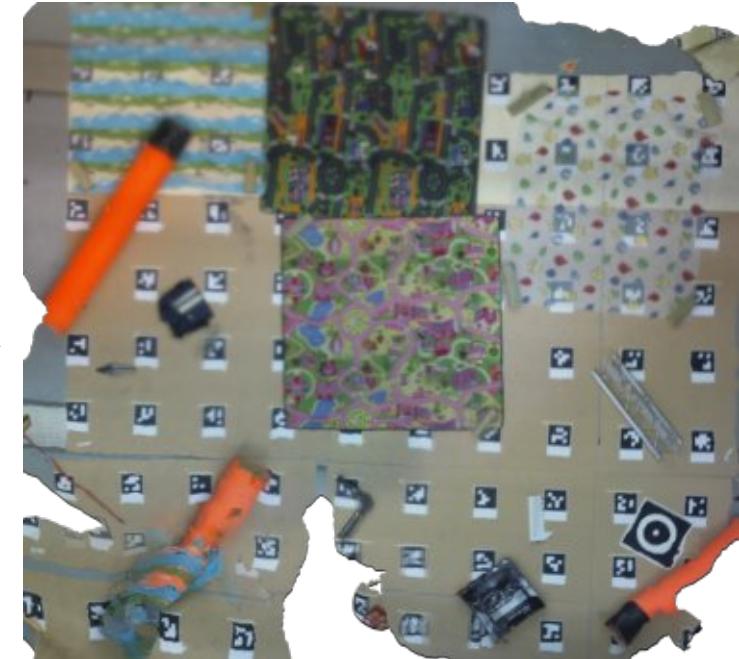


SCALABILITY



The diagram shows two sets of points arranged in a grid-like pattern. The top set consists of 10 points arranged in three columns of 3, 4, and 3 points respectively. The bottom set consists of 10 points arranged in three columns of 3, 4, and 3 points respectively, mirroring the top set.

SCALABILITY



Orthomap

-
-
-
-
-
-
-
-
-

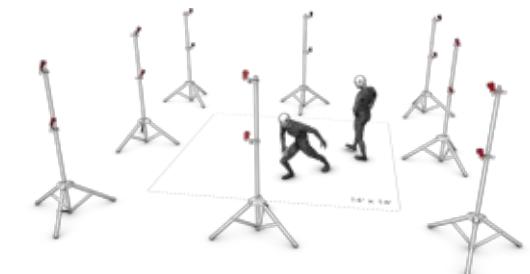
GROUND TRUTH ESTIMATION



OR



OR



Orthomap

Poster

Motion tracking

Dataset



x y

200 300



210 320

⋮

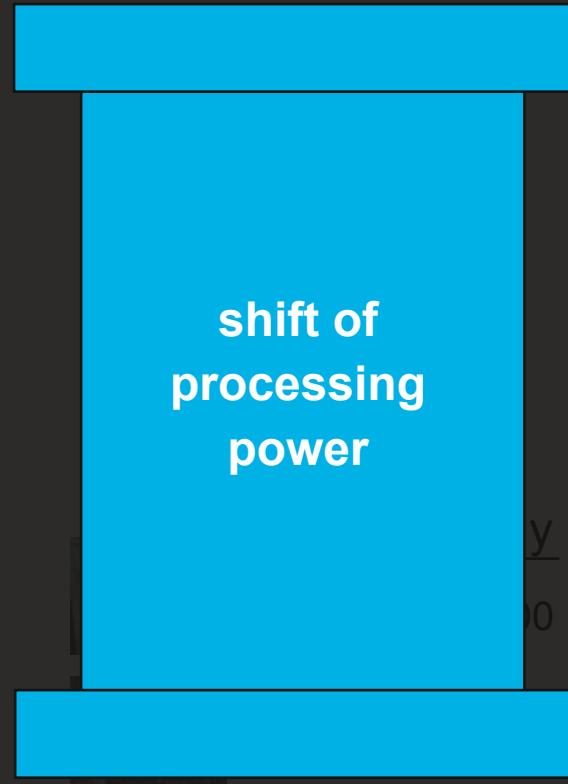
⋮

⋮

GROUND TRUTH ESTIMATION



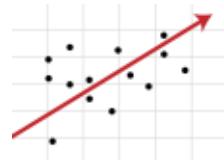
Orthomap



Motion tracking

CHALLENGES

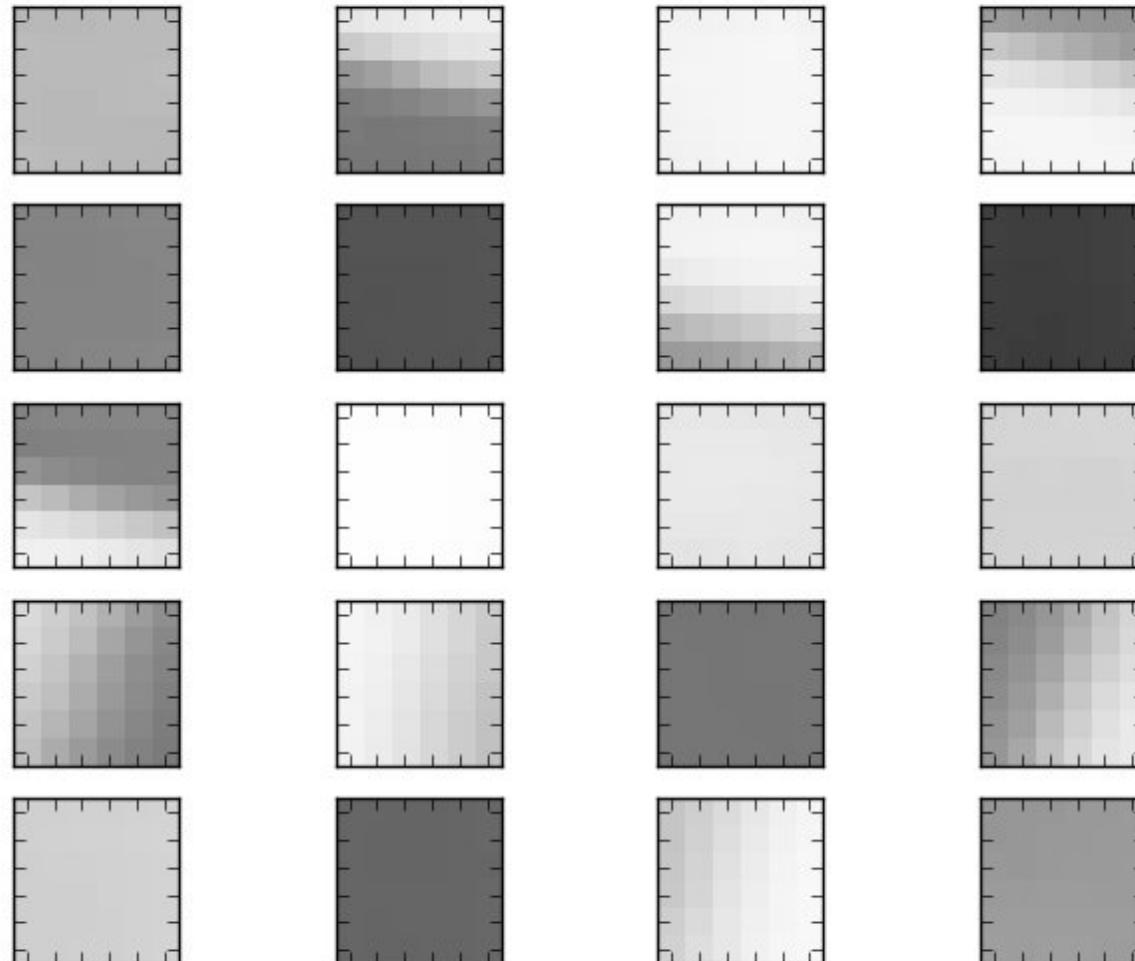
2-dimensional
regression



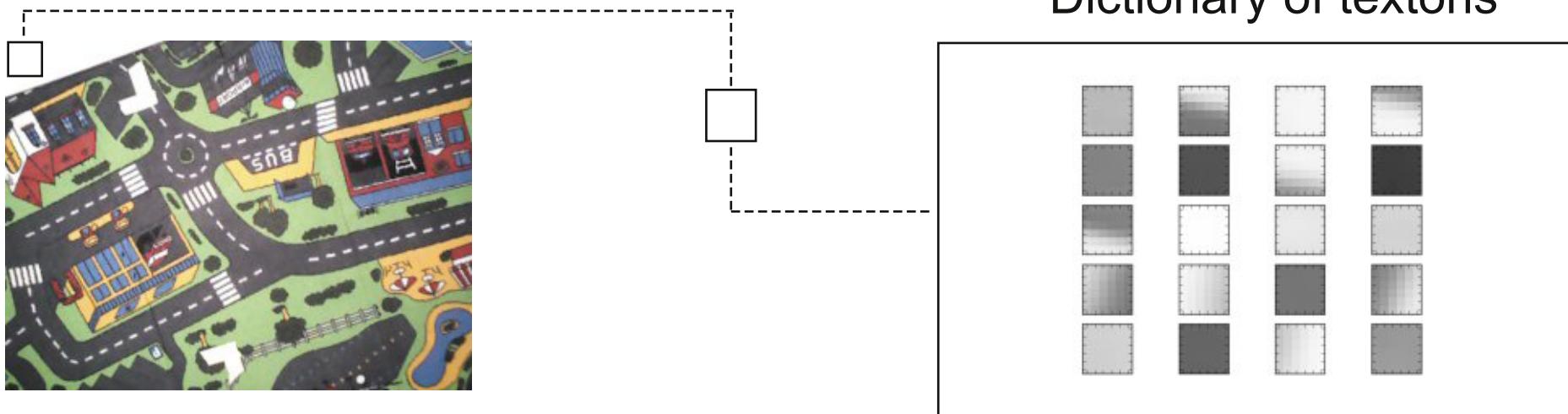
Which map is good?



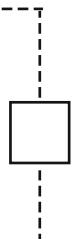
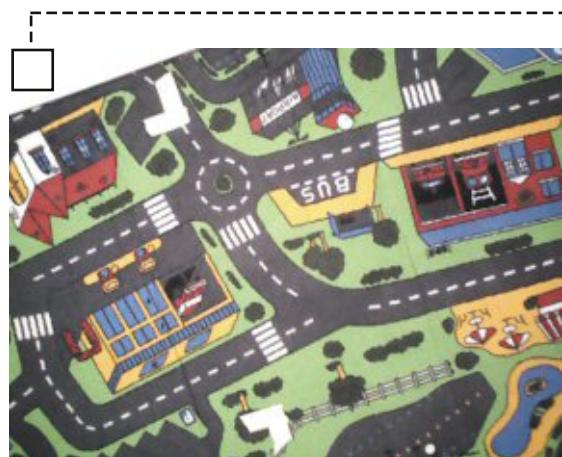
TEXTONS AS IMAGE FEATURES



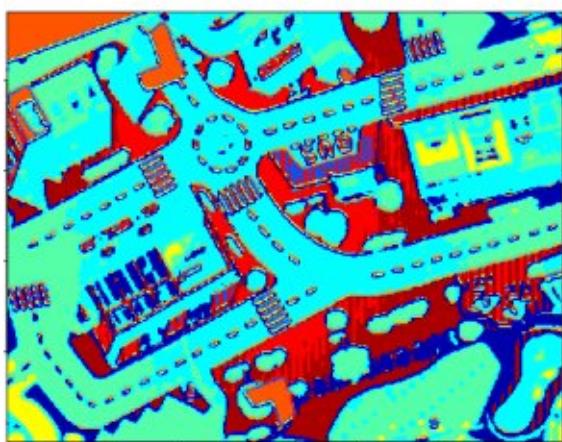
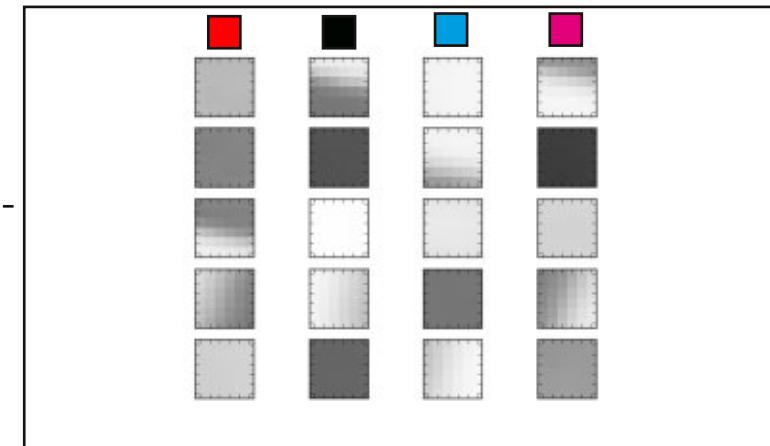
MACHINE-LEARNING APPROACH



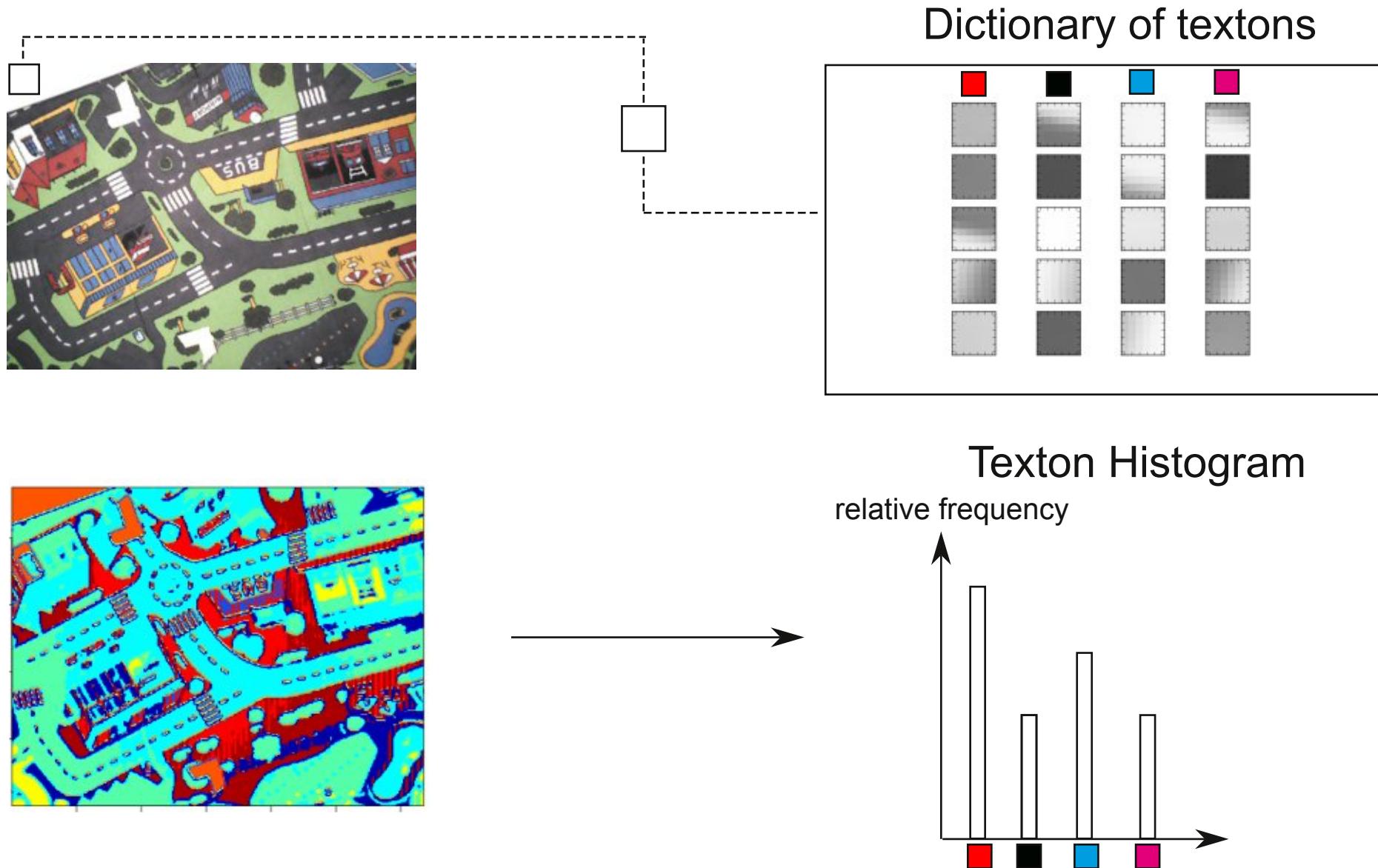
MACHINE-LEARNING APPROACH



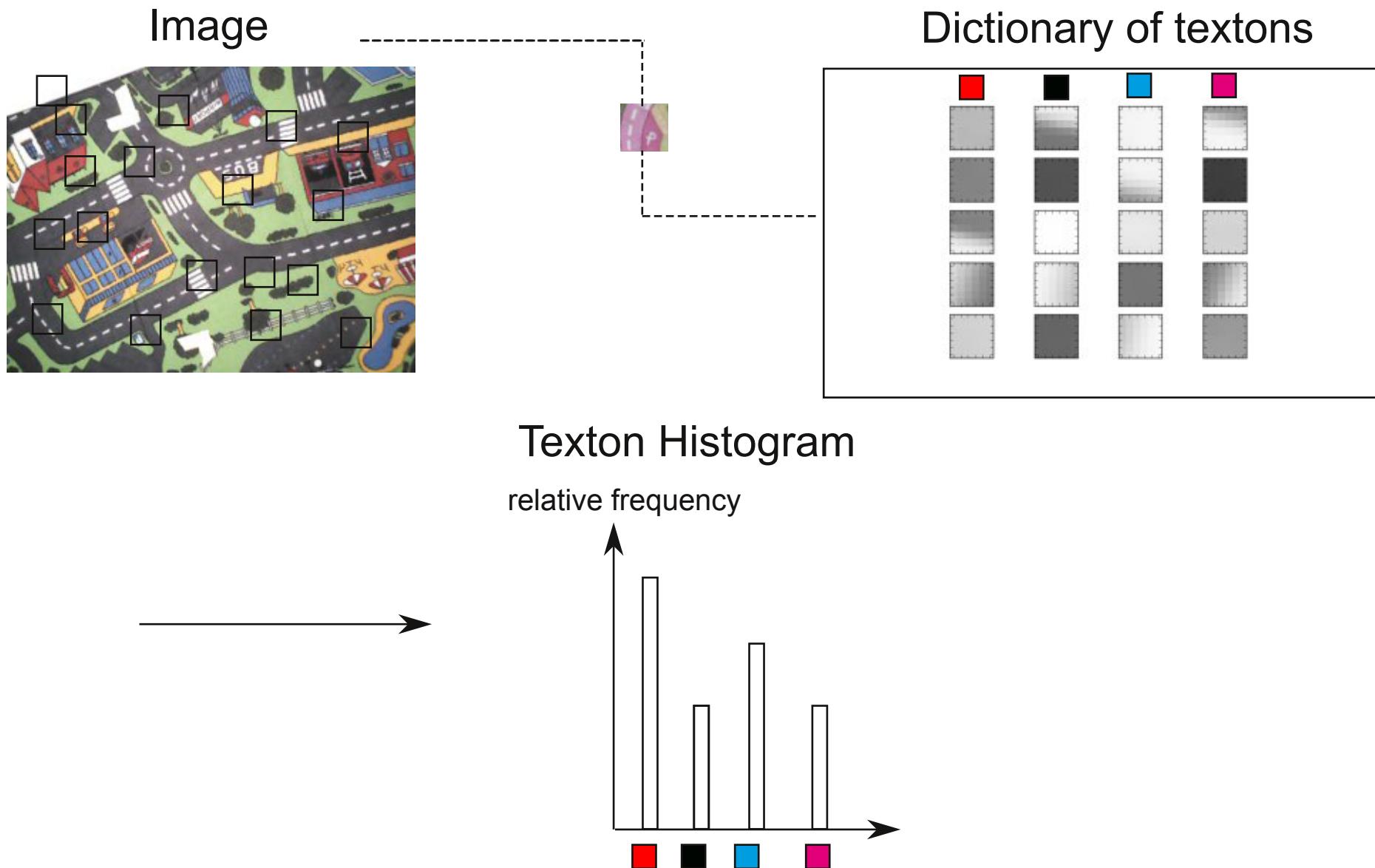
Dictionary of textons



MACHINE-LEARNING APPROACH

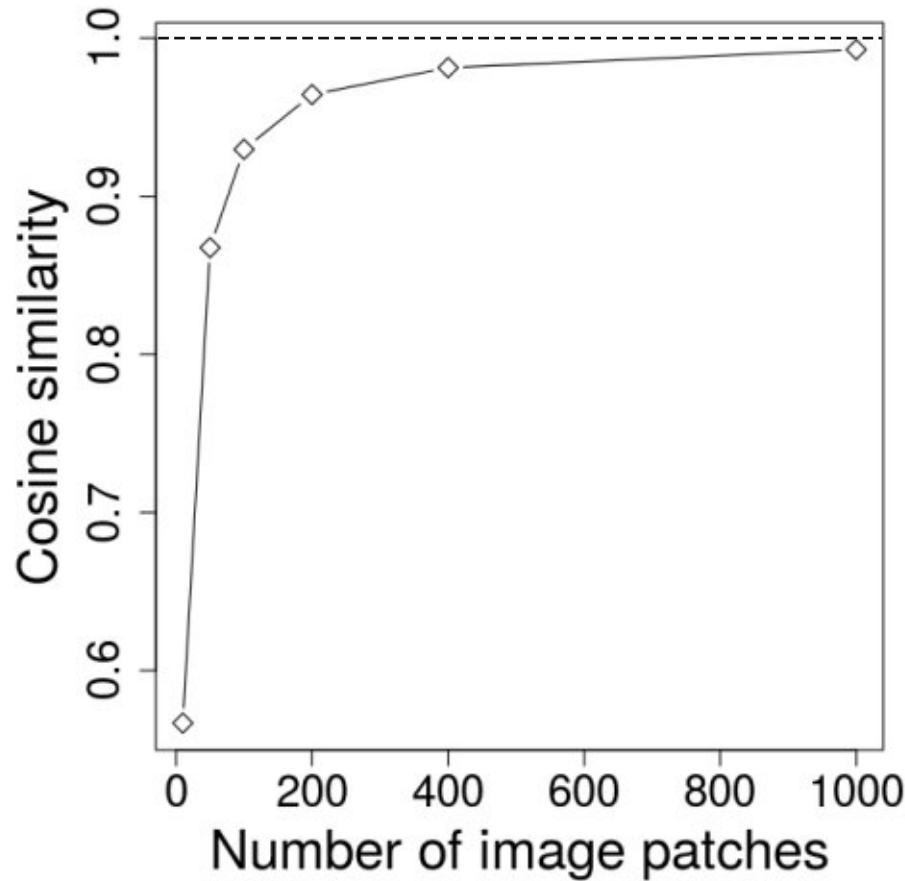


MACHINE-LEARNING APPROACH



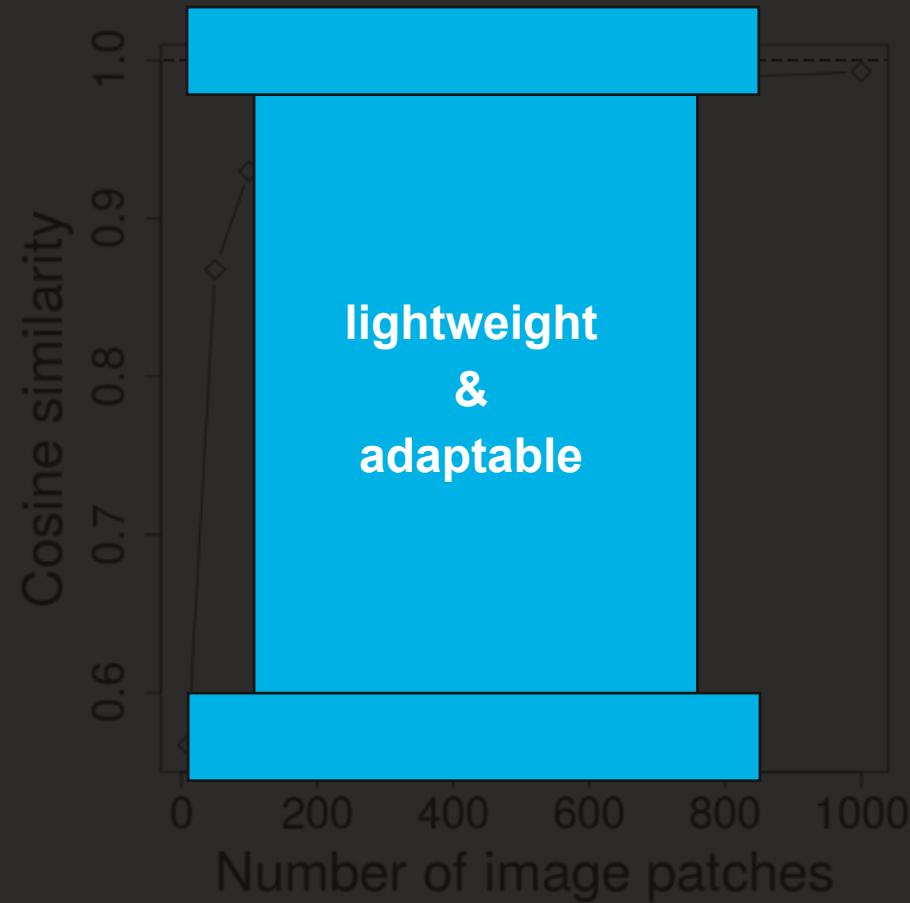
MACHINE-LEARNING APPROACH

307200!



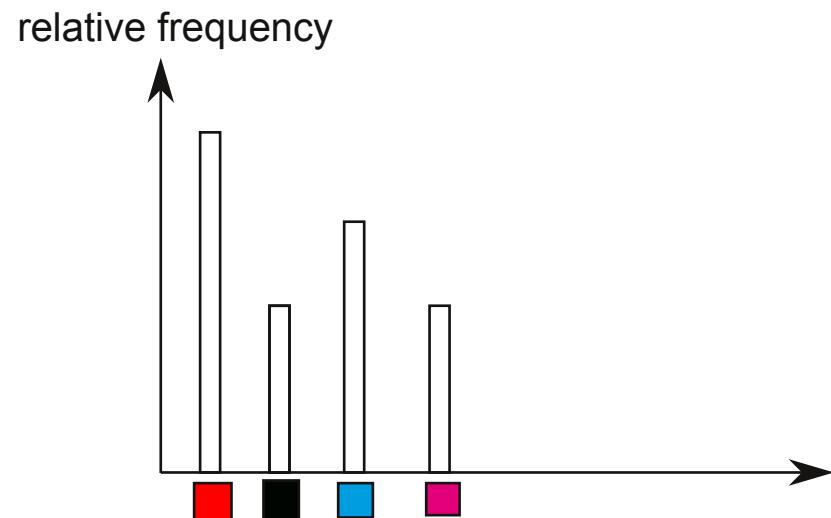
MACHINE-LEARNING APPROACH

307200!

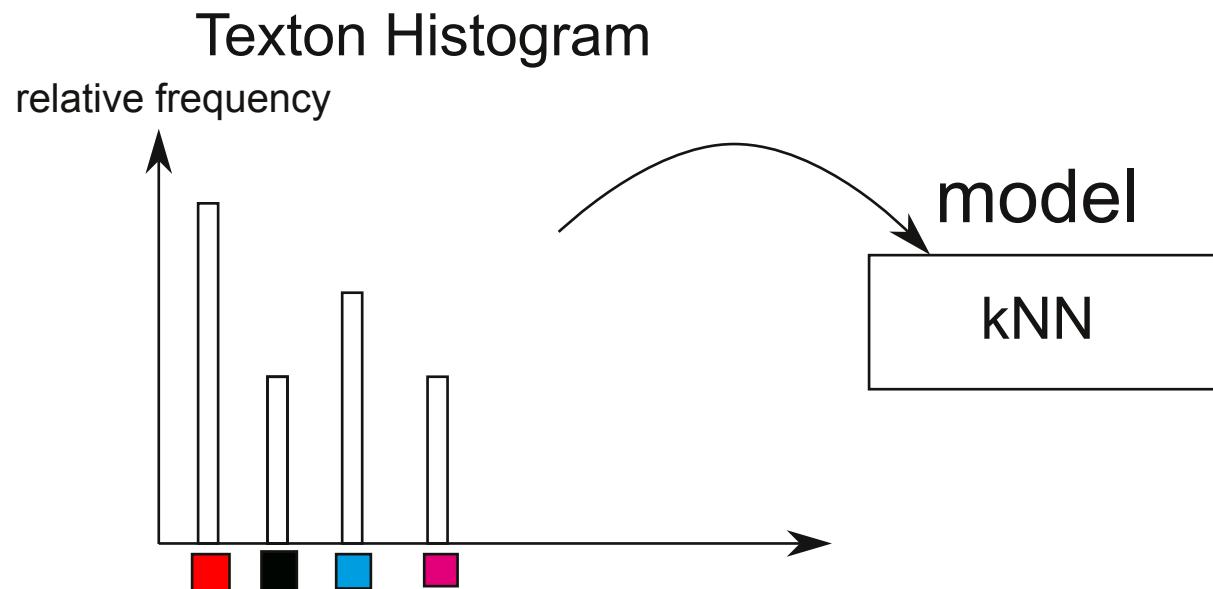


MACHINE-LEARNING APPROACH

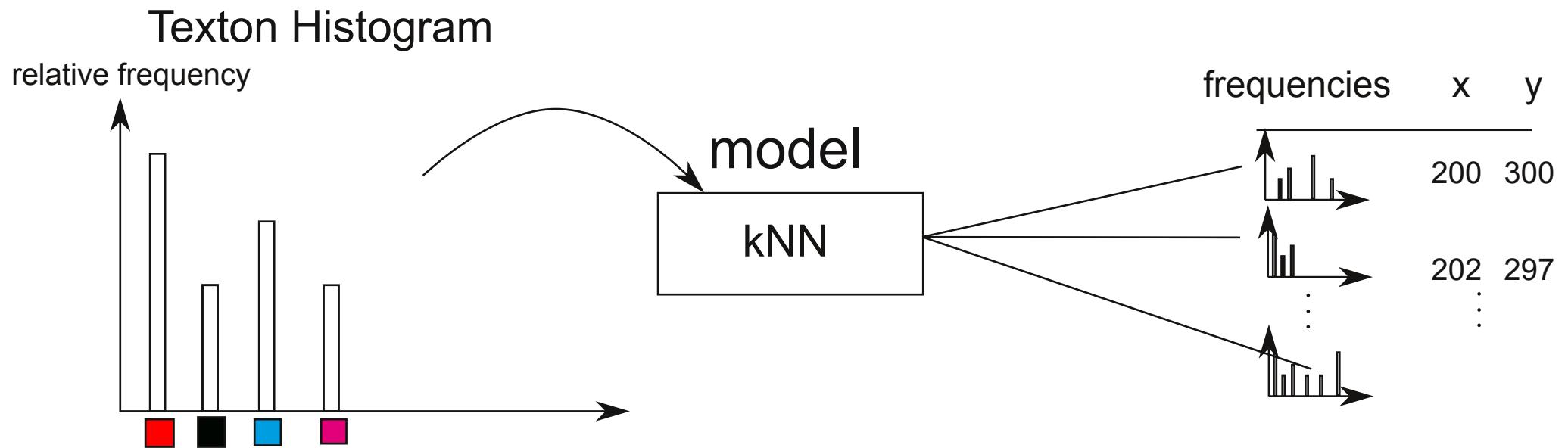
Texton Histogram



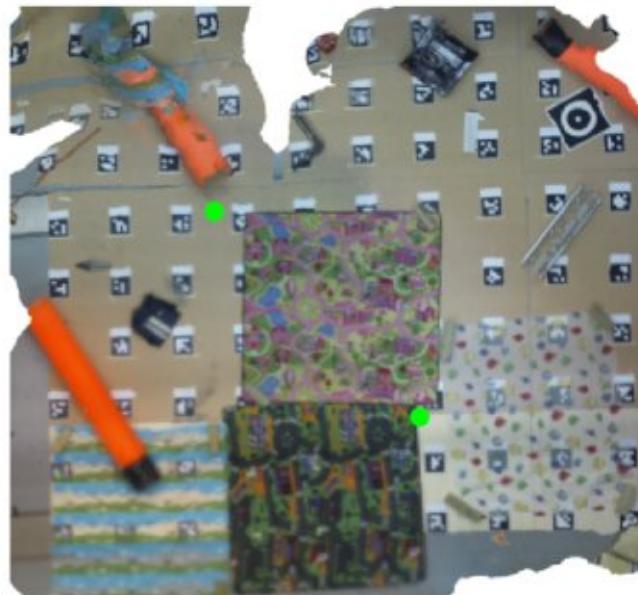
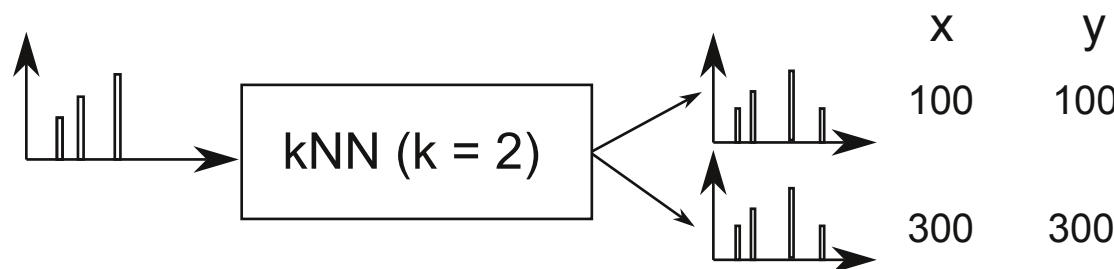
MACHINE-LEARNING APPROACH



MACHINE-LEARNING APPROACH

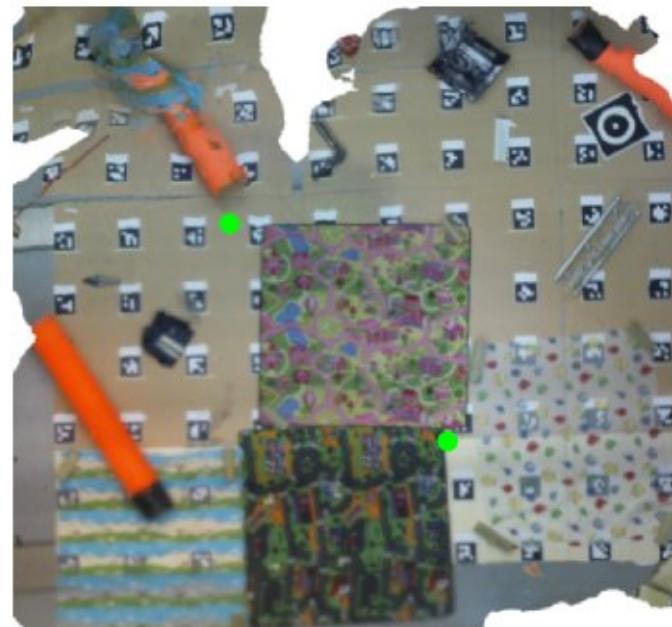


FILTERING



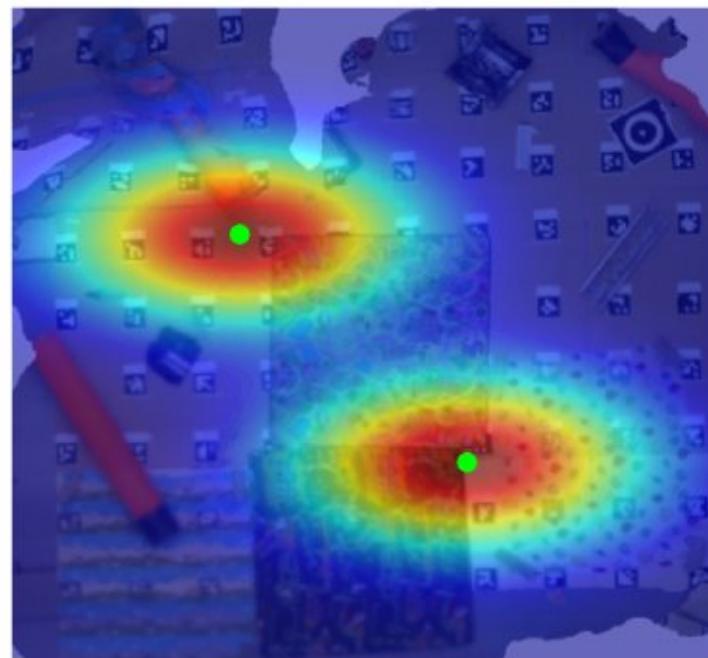
FILTERING

Sensor model (Likelihood)



FILTERING

Sensor model (Likelihood)



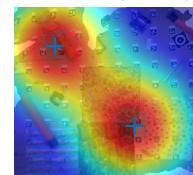
2-D Gaussian mixture model

FILTERING

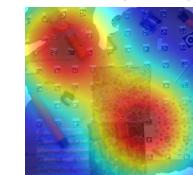
Prior ($t = 1$)



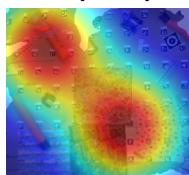
Likelihood ($t = 1$)



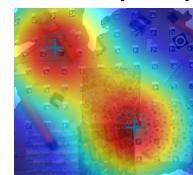
Posterior ($t = 1$)



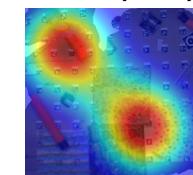
Prior ($t = 2$)



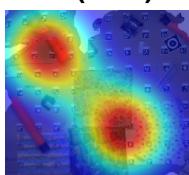
Likelihood ($t = 2$)



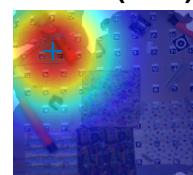
Posterior ($t = 2$)



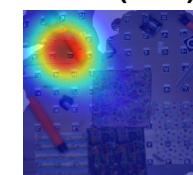
Prior ($t = 3$)



Likelihood ($t = 3$)



Posterior ($t = 3$)

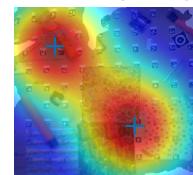


FILTERING

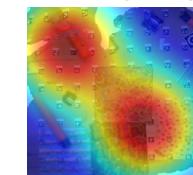
Prior (t = 1)



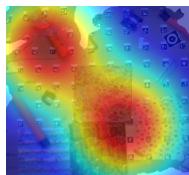
Likelihood (t = 1)



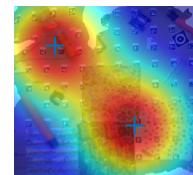
Posterior (t = 1)



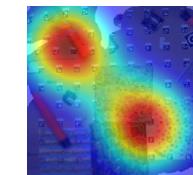
Prior (t = 2)



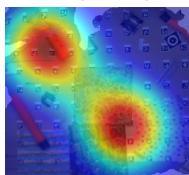
Likelihood (t = 2)



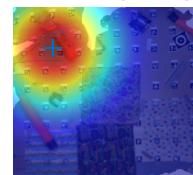
Posterior (t = 2)



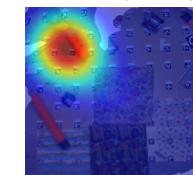
Prior (t = 3)



Likelihood (t = 3)

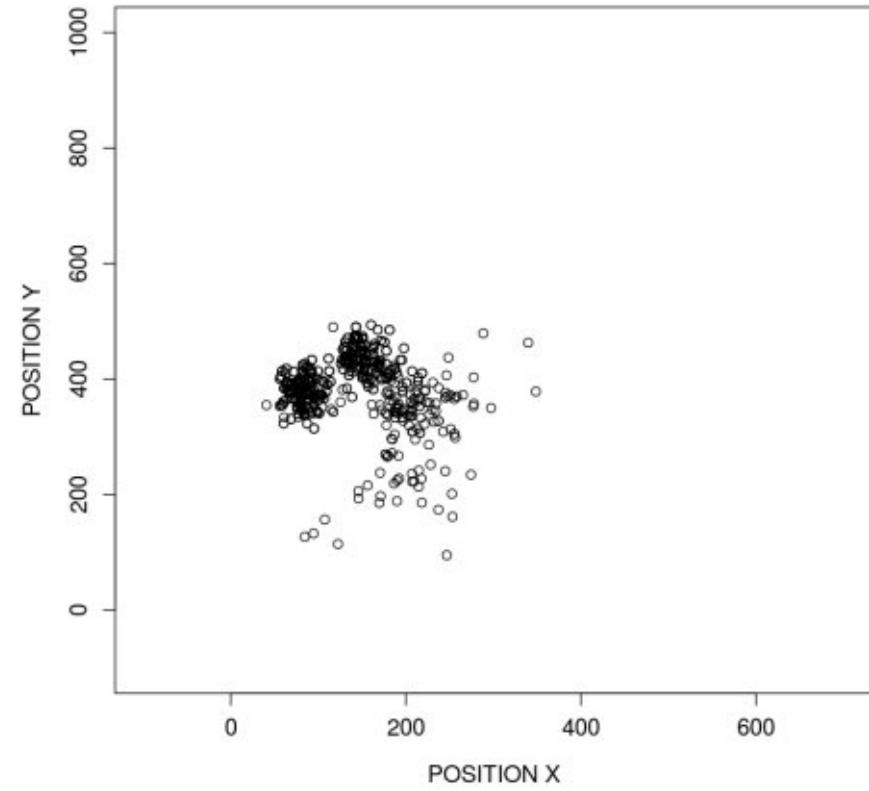
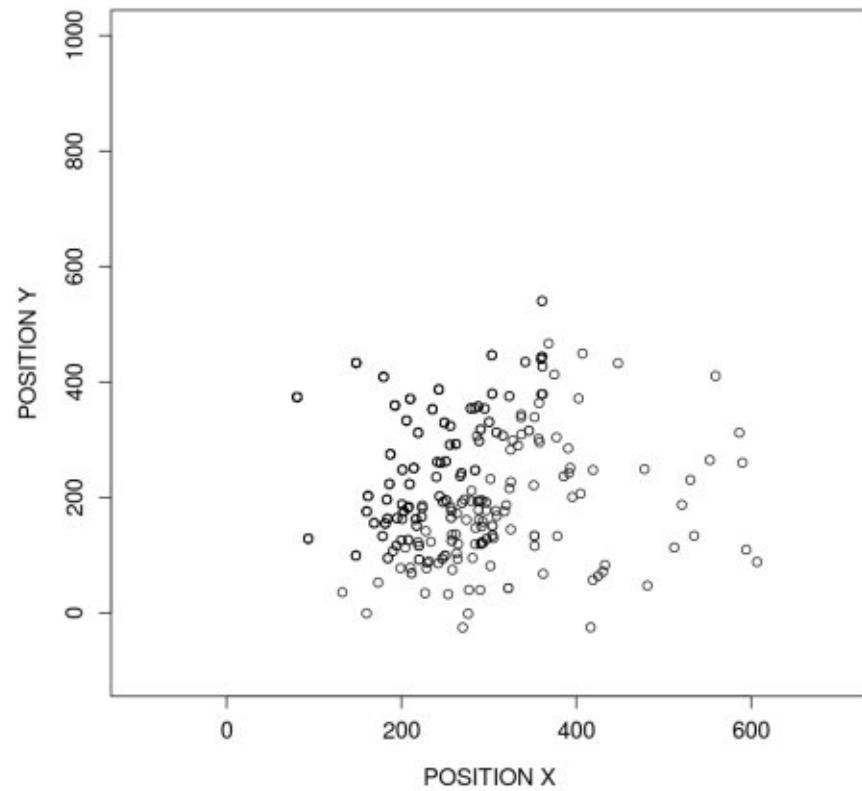


Posterior (t = 3)



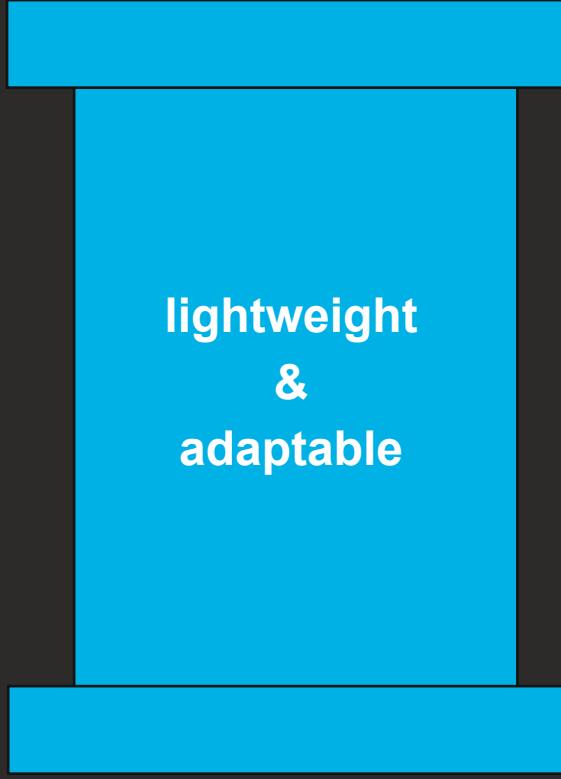
+ proximity-based motion model

PARTICLE FILTER



DUMMY: VIDEO FAST FLIGHT

DUMMY: PARTICLE FILTER



**lightweight
&
adaptable**

MAP EVALUATION



MAP EVALUATION



**IDEAL SIMILARITY OF
HISTOGRAMS FOR FIXED
POSITION**

MAP EVALUATION - SYNTHETIC FLIGHT



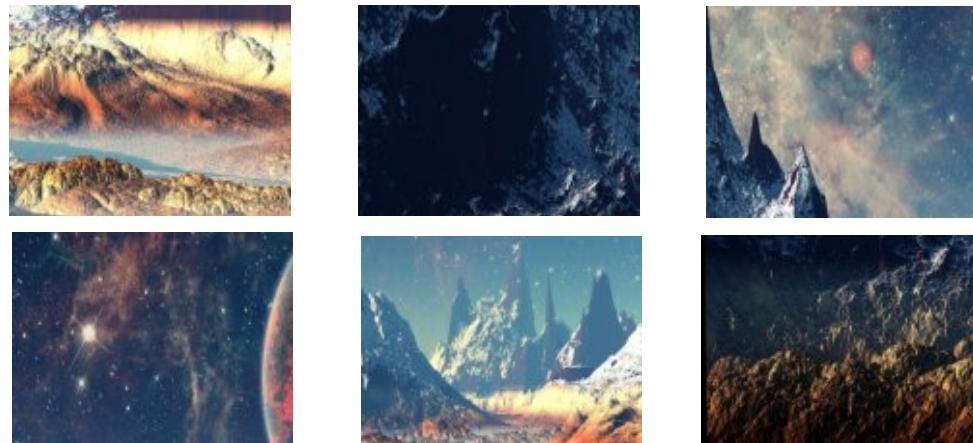
MAP EVALUATION - SYNTHETIC FLIGHT



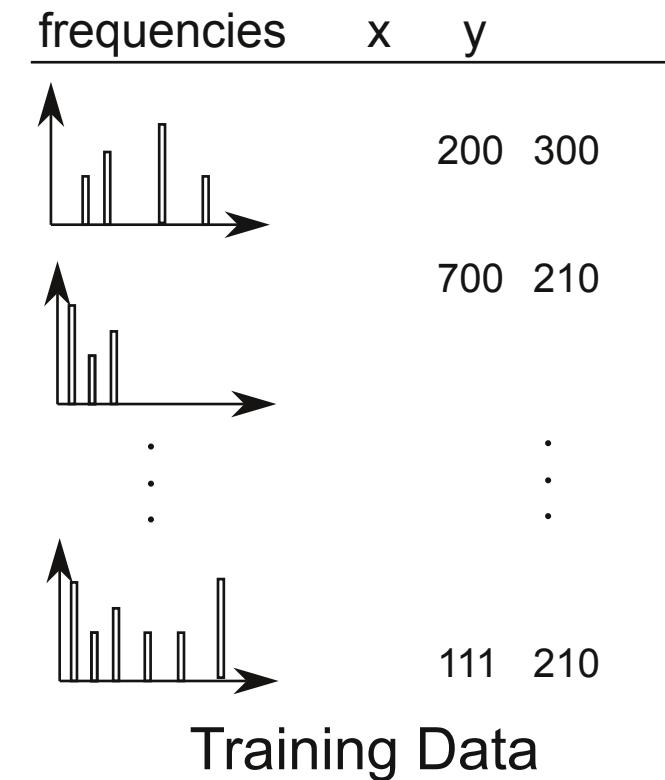
synthetic flight



MAP EVALUATION - SYNTHETIC FLIGHT



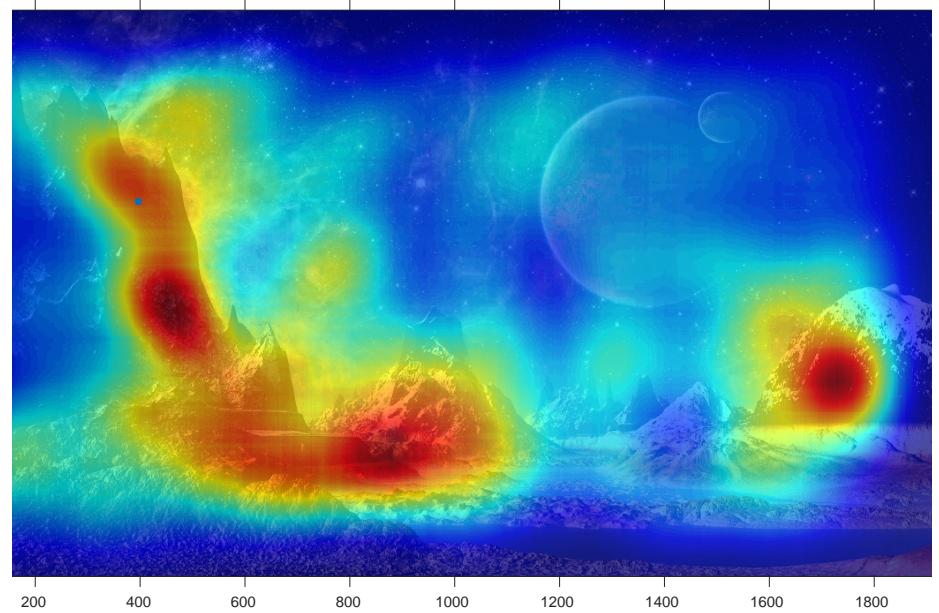
1000 patches



MAP EVALUATION

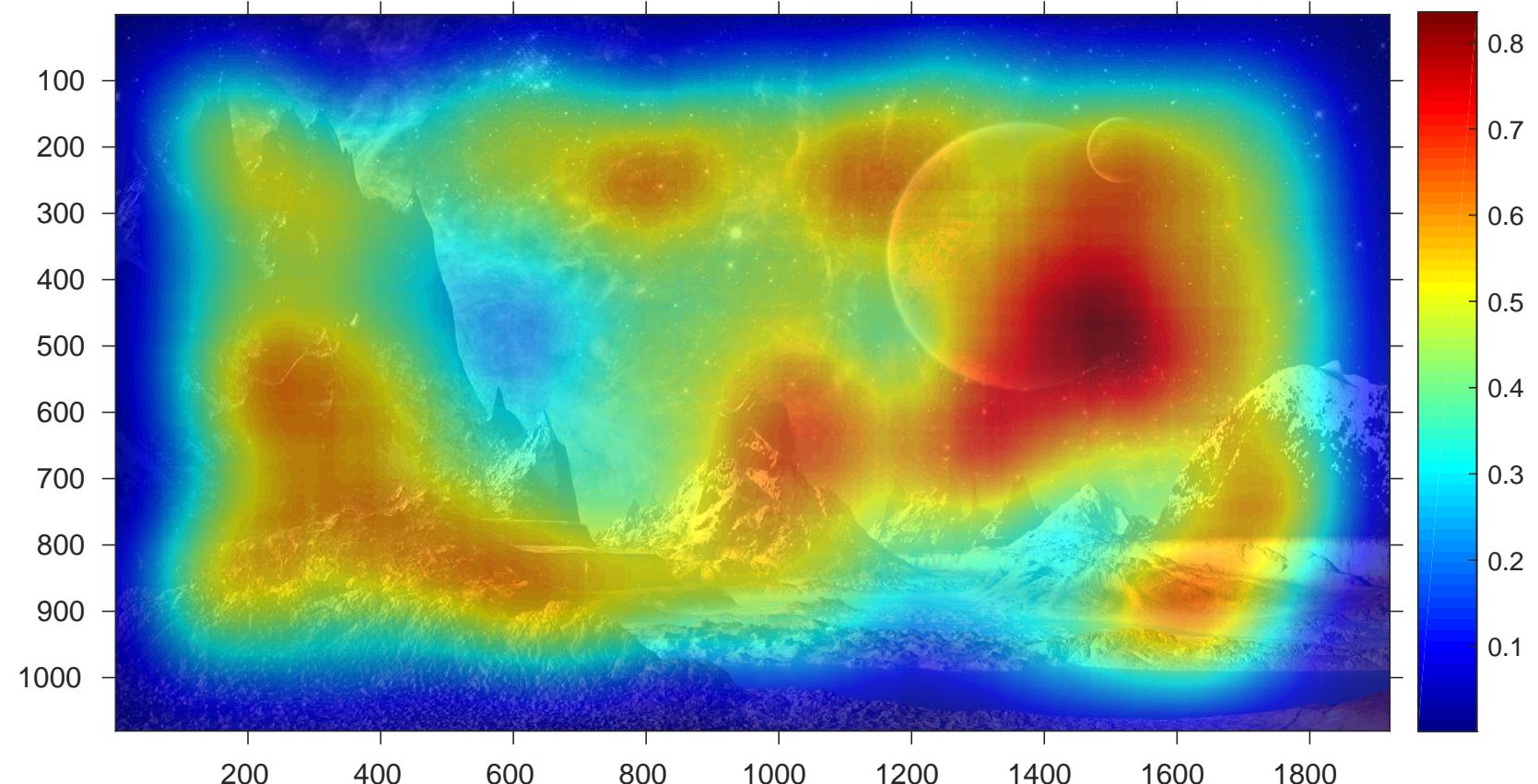


**IDEAL SIMILARITY OF
HISTOGRAMS FOR FIXED
POSITION**



**ACTUAL SIMILARITY
(Gaussian smoothing)**

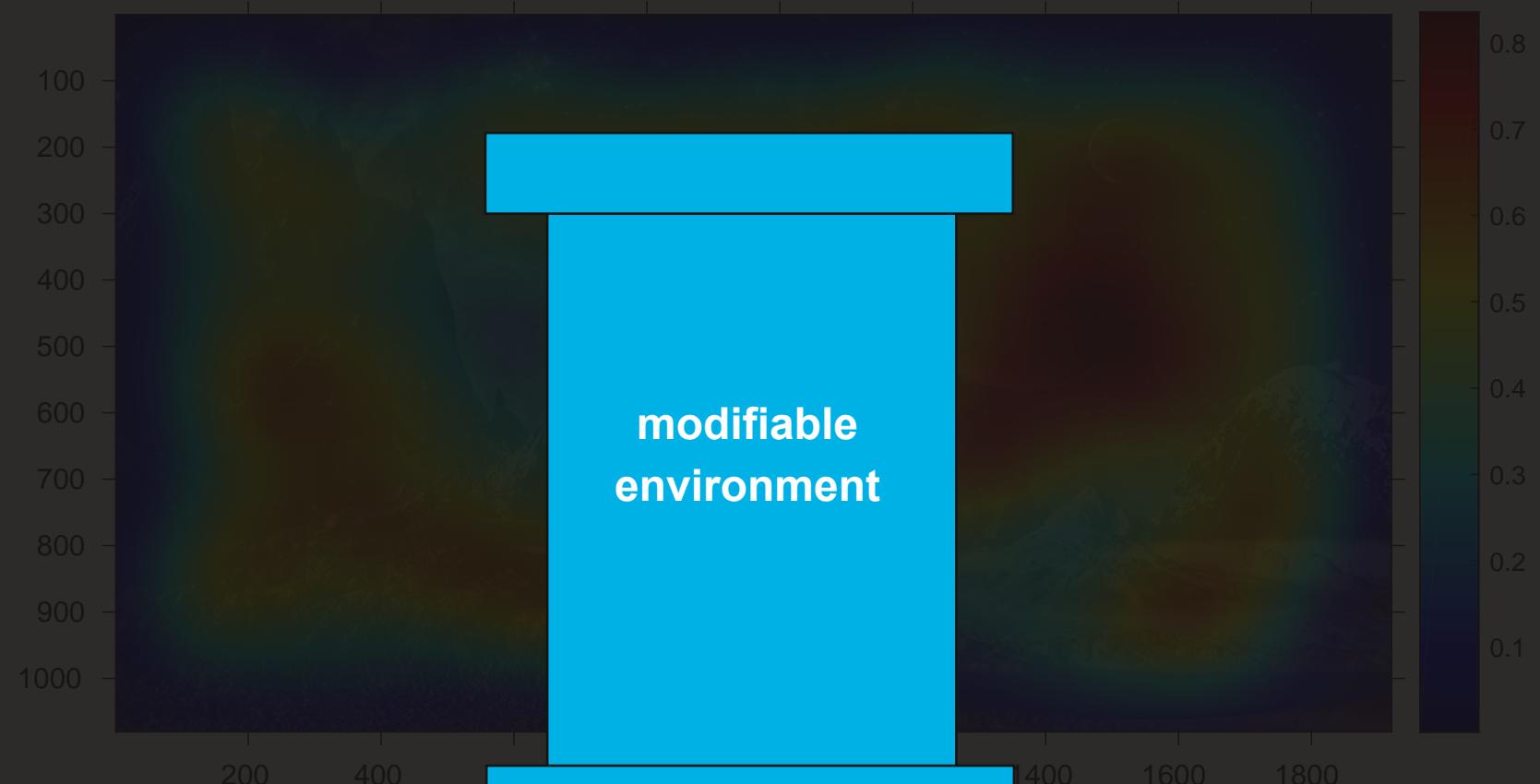
MAP EVALUATION



GLOBAL LOSS (1000 patches)

$0 < \text{global loss} < 1$

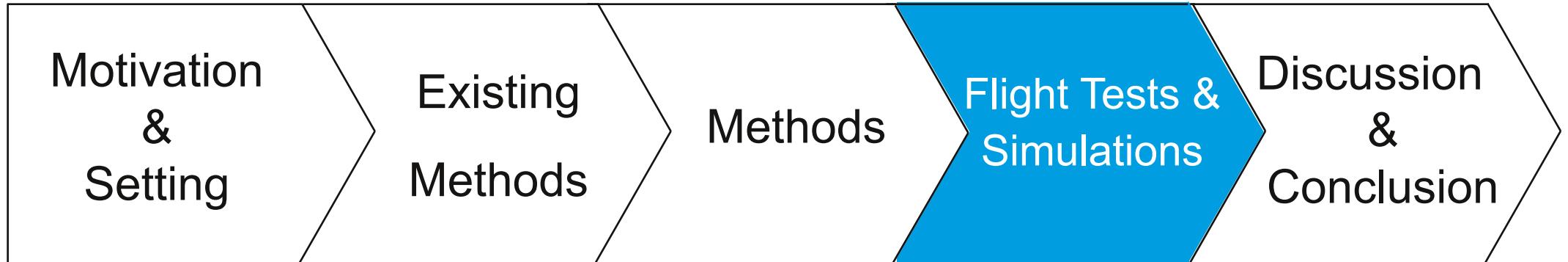
MAP EVALUATION



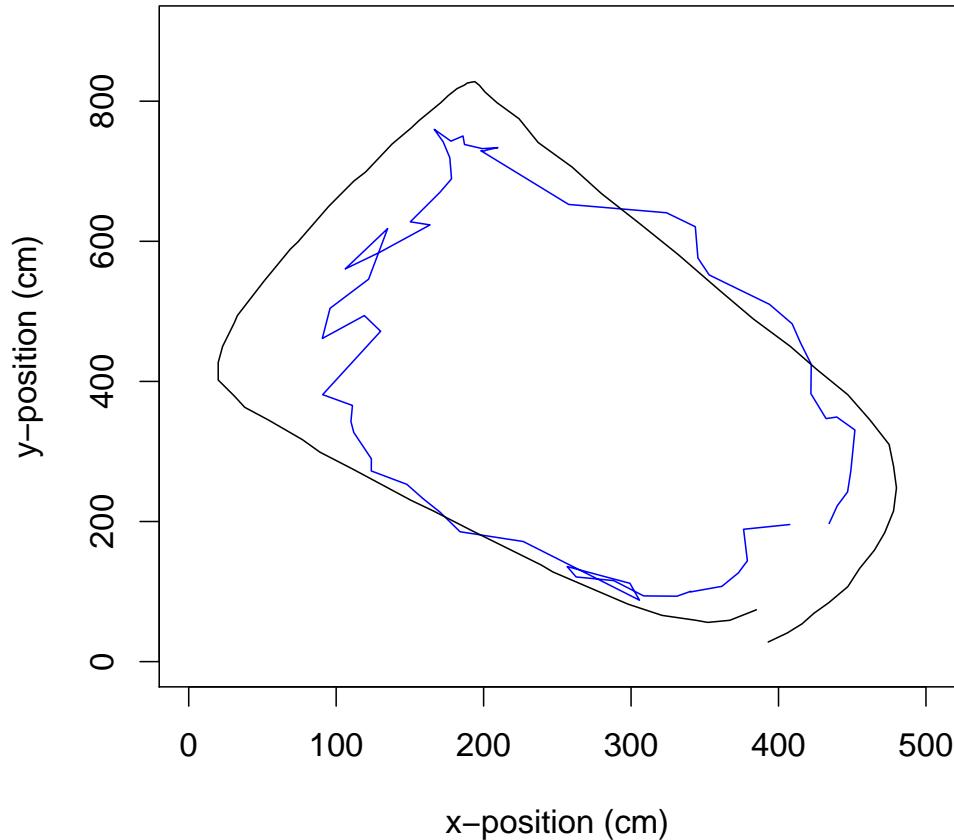
GLOBAL LOSS (1000 patches)

$0 < \text{global loss} < 1$

OUTLINE



FLIGHT ACCURACY

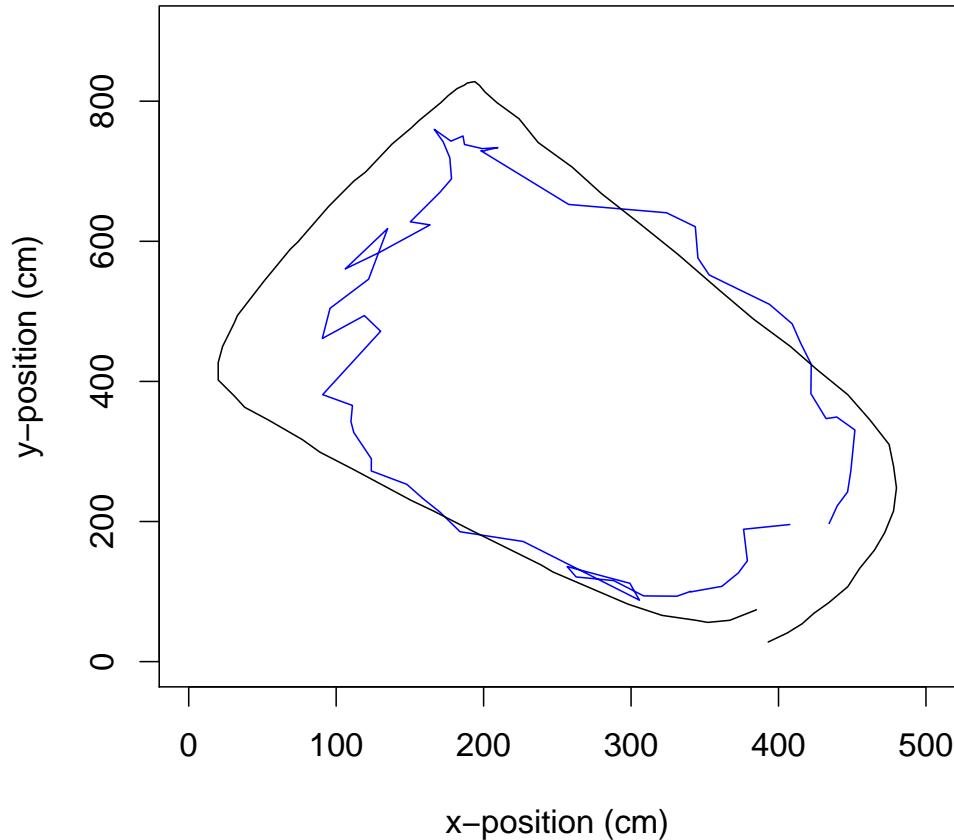


Flight with 400 images

Mean distance x: 46 cm

Mean distance y: 54 cm

FLIGHT ACCURACY



Flight with 400 images

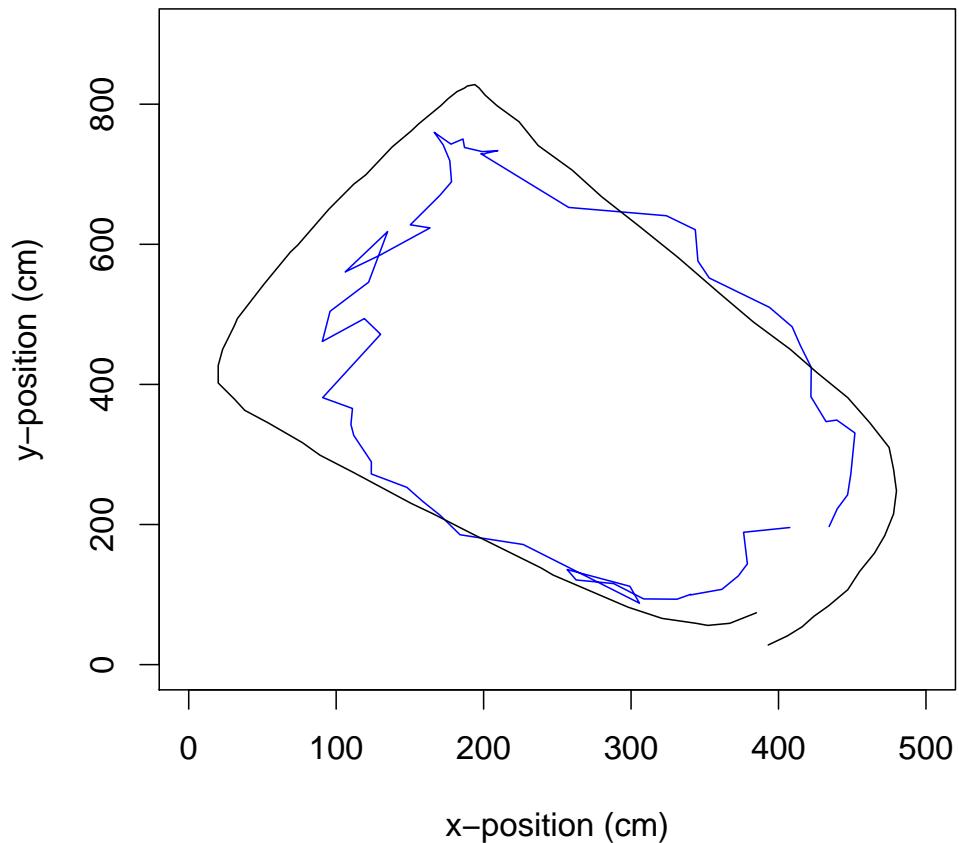
Mean distance x: 46 cm

Mean distance y: 54 cm

SD distance x: 56 cm

SD distance y: 71 cm

FLIGHT ACCURACY



Flight with 400 images

Mean distance x: 46 cm

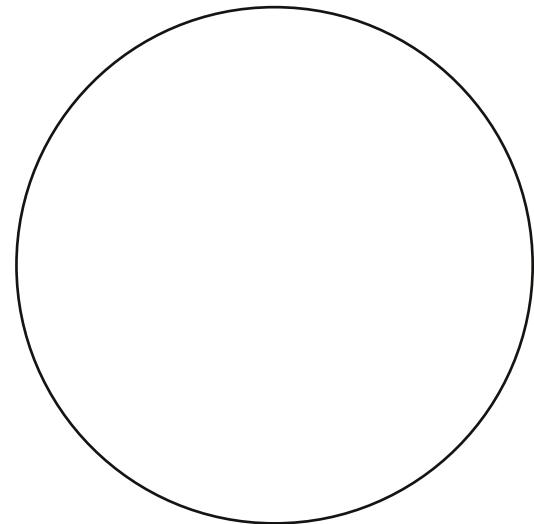
Mean distance y: 54 cm

SD distance x: 56 cm

SD distance y: 71 cm

Frequency: 12 Hz

TRIGGERED LANDING



6 Landings

criterion:
distance < 60 cm

safety criterion:
low variance of particles

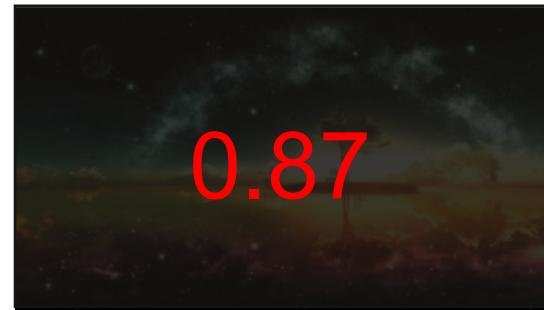
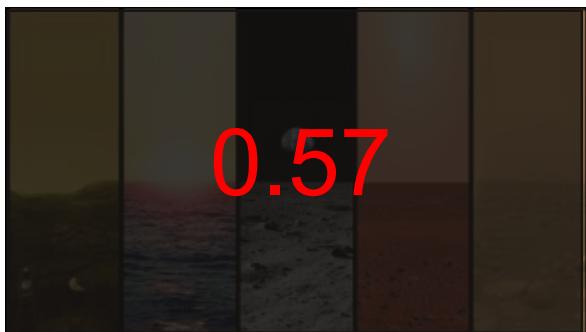
DUMMY: VIDEO TRIGGERED LANDING

MAP EVALUATION

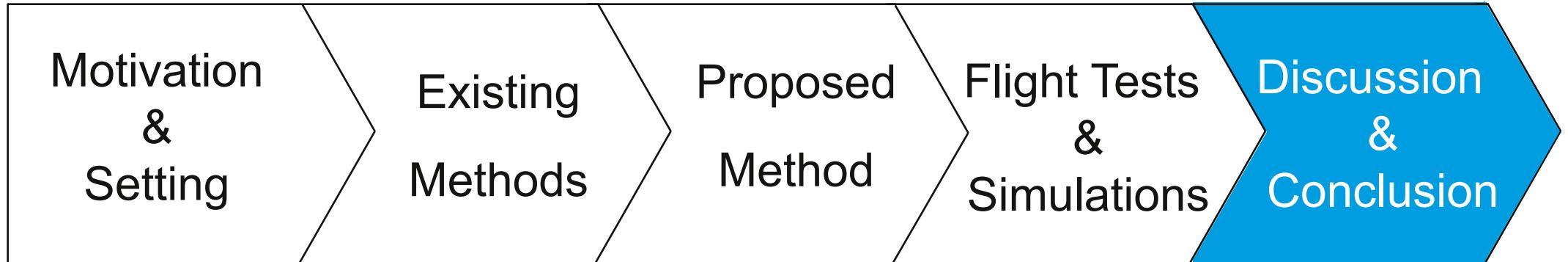


46 images

MAP EVALUATION - LOSSES



OUTLINE



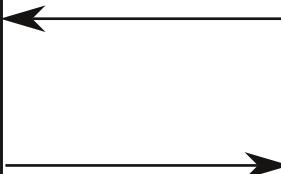
RESEARCH QUESTIONS

Research Question 1

Can vision-based indoor localization be done on a limited platform?

Research Question 2

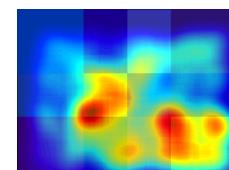
Can we predict the suitability of an environment for the proposed localization algorithm?



DISCUSSION

Implications:

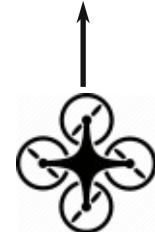
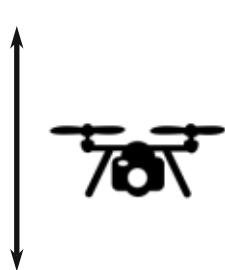
- paves the way to indoor flight
- adaptable to different platforms
- detect safe landing spots



DISCUSSION

Limitations:

- assumes constant height and no rotations



- robustness to different lighting conditions

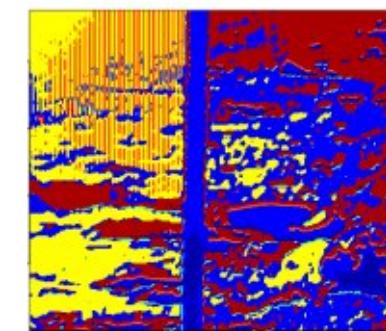
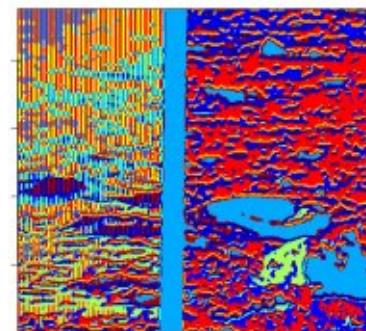
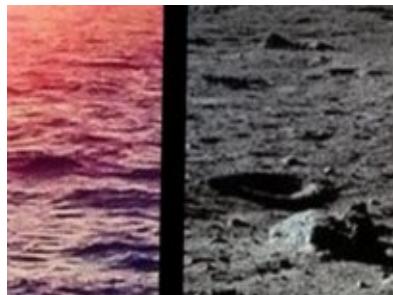


- particle filter does not include velocity or heading

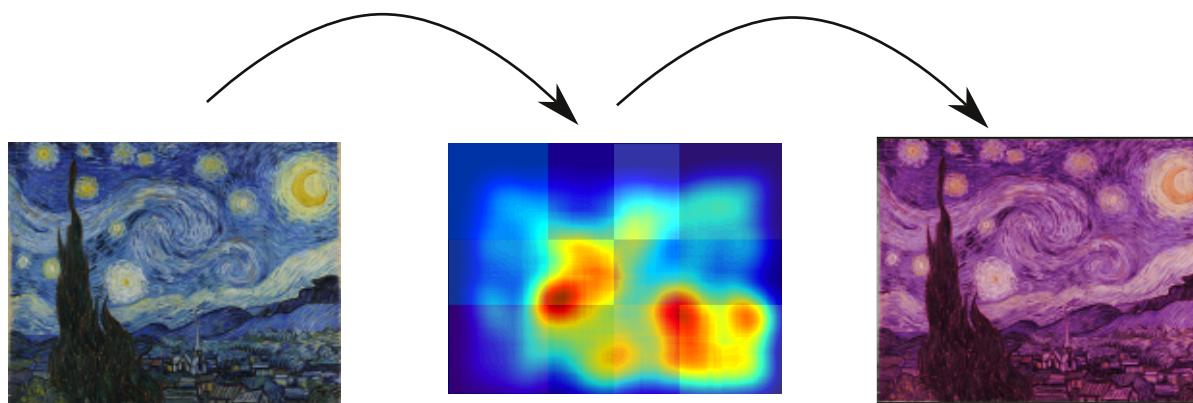
DISCUSSION

Future research:

- bridge reality gap



- automatic map generation (evolutionary algorithm)



CODE CONTRIBUTIONS

- draug: Image augmentation with synthetic views (C++)

<https://github.com/Pold87/draug>

- Map evaluation (MATLAB)

<https://github.com/Pold87/evaluation-thesis>

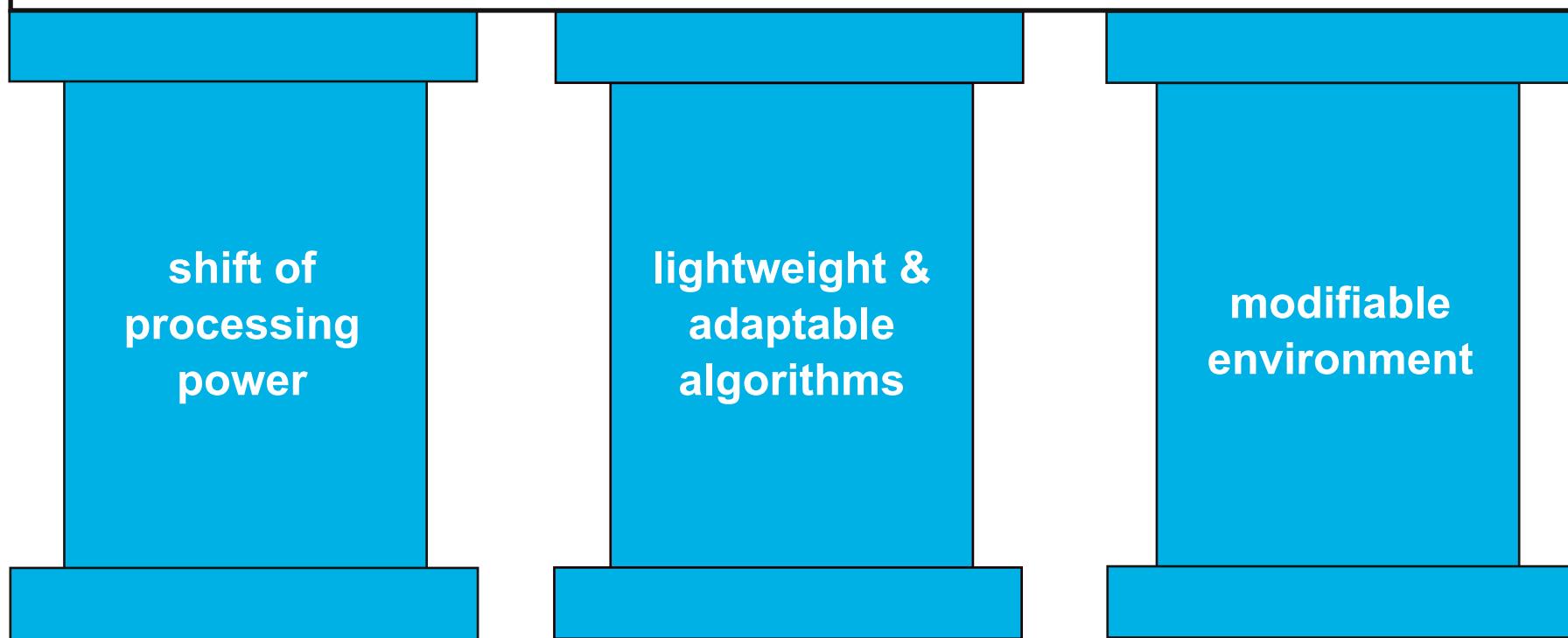
- Localization: SIFT matching (C++), particle filter (C),
texton-based approach (C)

TODO: PULL REQUEST C: <https://github.com/Pold87/paparazzi> Python: <https://github.com/Pold87/treXton>



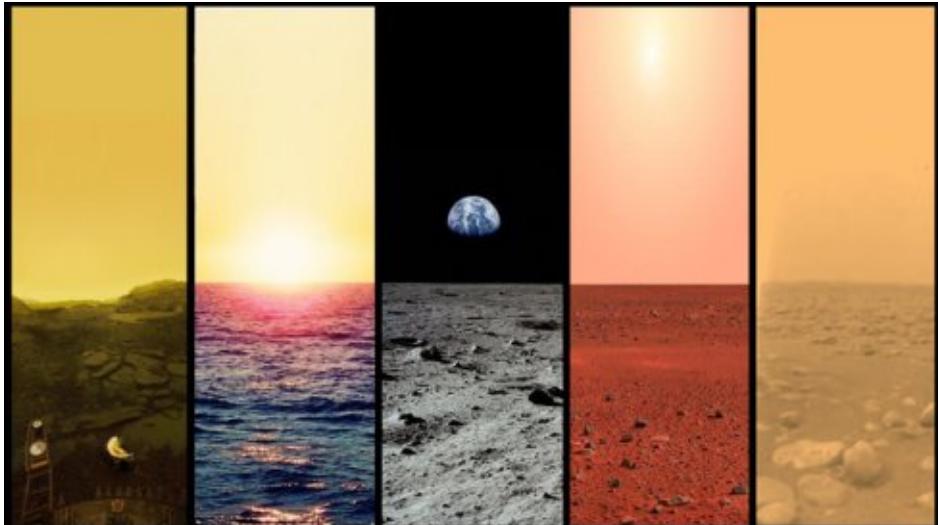
CONCLUSION

EFFICIENT INDOOR LOCALIZATION



MAP EVALUATION

GOOD



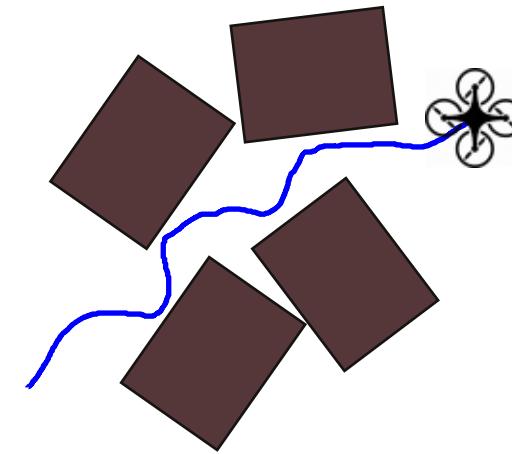
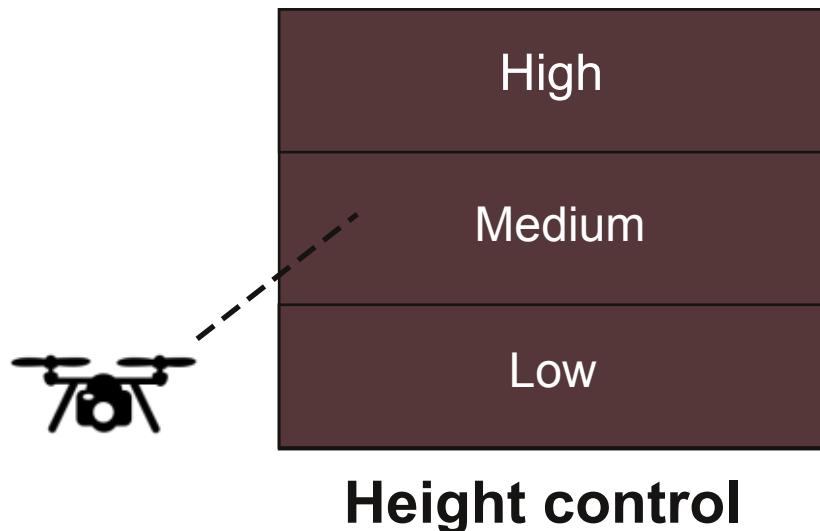
0.57

BAD



0.98

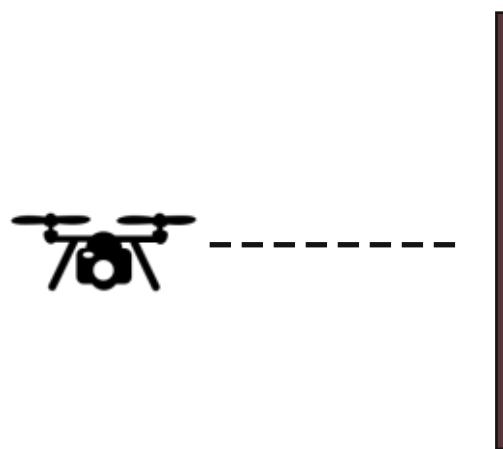
FOUNDATION



Obstacle Avoidance



Safe Landing Spot Detection



Distance Measurement

PARTICLE FILTER

Motion model

2D-Gaussian noise

