UMASSCS
SCHOOL OF COMPUTER SCIENCE

MoA-Net: Self-supervised Motion Segmentation

Pia Bideau, Rakesh R Menon, Erik Learned-Miller

University of Massachusetts Amherst
College of Information and Computer Science

Motion Segmentation

- P Bideau, E Learned-Miller, ECCV 2016:
  It's moving! A probabilistic model for causal motion segmentation

- P Bideau, A RoyChoudhury, R Menon, E Learned-Miller, CVPR 2018:
  The best of both worlds: Combining CNNs and geometric constraints for hierarchical motion segmentation

- P Bideau, R Menon, E Learned-Miller, Workshop ECCV 2018:
  MoA-Net: Unsupervised Motion Segmentation

Overview

Motivation

How do humans know
what is moving in the world and what is not?

Approach: Motion Segmentation

- Rotation compensation
- Learning Motion Patterns: MoA-Net

Results

Future Research Questions

Motivation

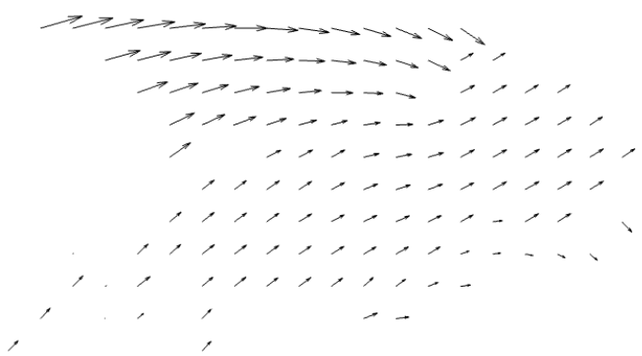
Motivation

- stationary scene
- moving object
- **no** observer motion

- stationary scene
- moving object
- observer motion

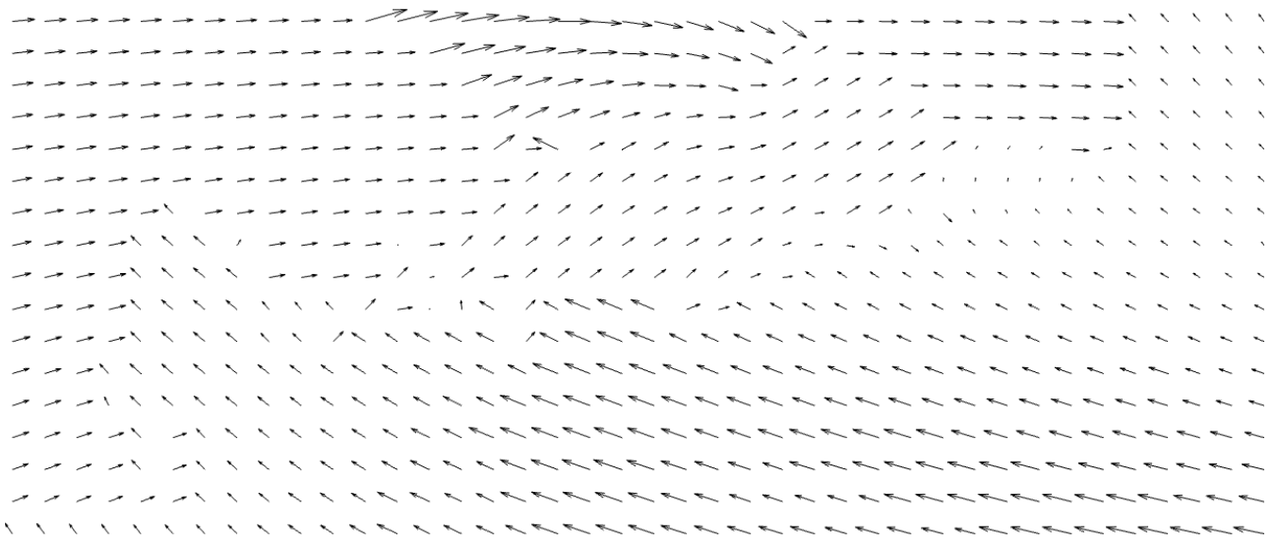
Motivation

- stationary scene
- moving object
- **no** observer motion



- stationary scene
- moving object
- observer motion

Motivation



- stationary scene
- moving object
- **no** observer motion

- stationary scene
- moving object
- observer motion

Motivation

All motions result in changes of the retinal image.

What is the problem about retinal image motion?

- photoreceptors are slow
- motion detection in our brain is challenging

Need to stabilize the image, to reduce retinal image motion

Motivation

Overview

Motivation

How do humans know
what is moving in the world and what is not?

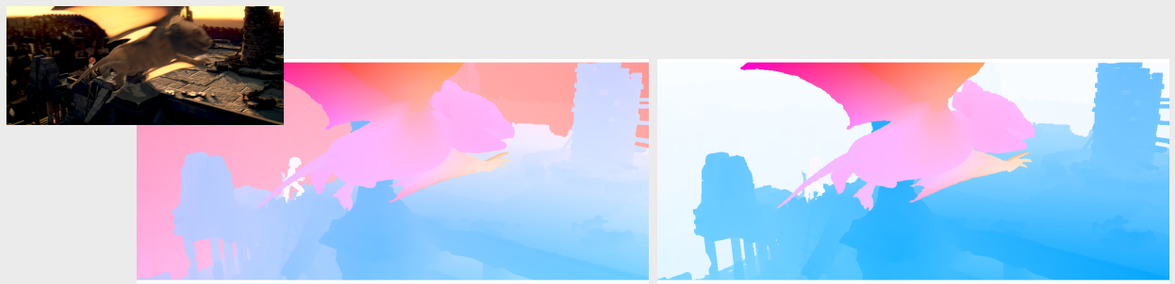
Approach: Motion Segmentation

- Rotation compensation
- Learning Motion Patterns: MoA-Net

Results

Future Research Questions

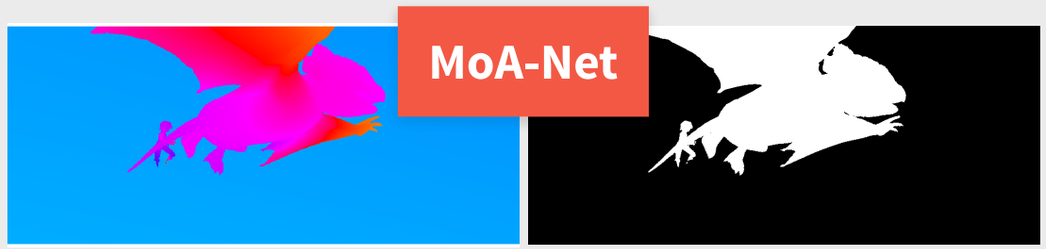
Approach:
Motion Segmentation



optical flow

rotation compensated flow

step 1: rotation compensation

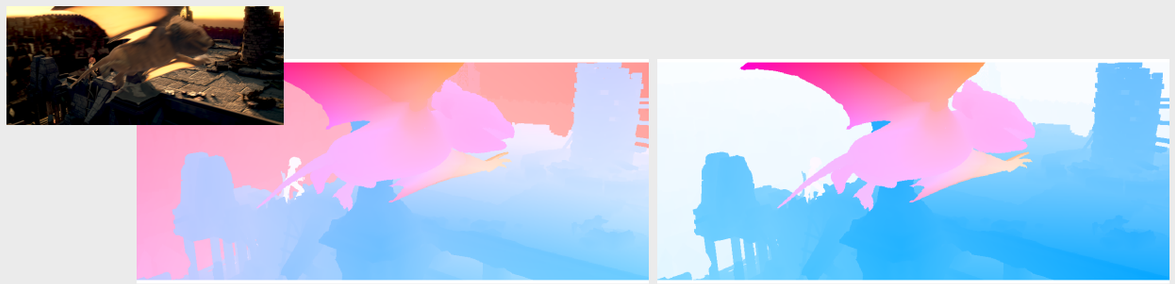


angle field

motion segmentation

step 2: motion segmentation

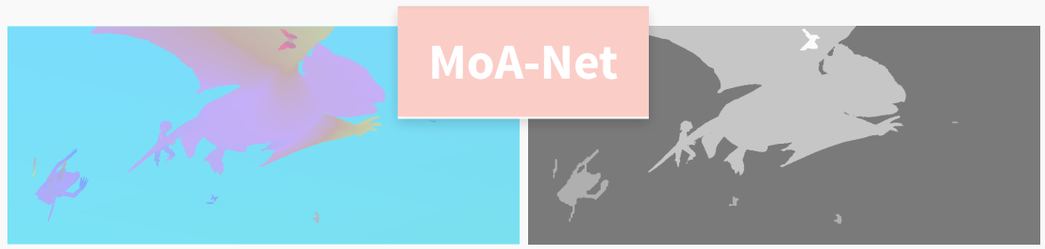
Approach:
Motion Segmentation



optical flow

rotation compensated flow

step 1: rotation compensation



angle field

motion segmentation

step 2: motion segmentation

Rotation Compensation

- rotation + translation
- optical flow **magnitude** is **dependent** on scene depth
- optical flow **angle** is **dependent** on scene depth

- translation
- optical flow **magnitude** is **dependent** on scene depth
- optical flow **angle** is **independent** of scene depth

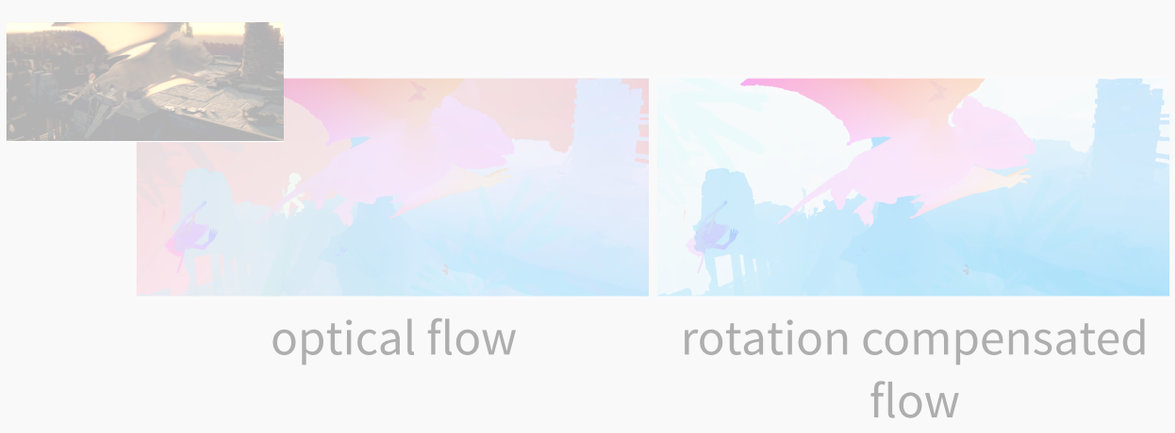
Rotation Compensation

only camera translation and object motion

optical flow

angle field

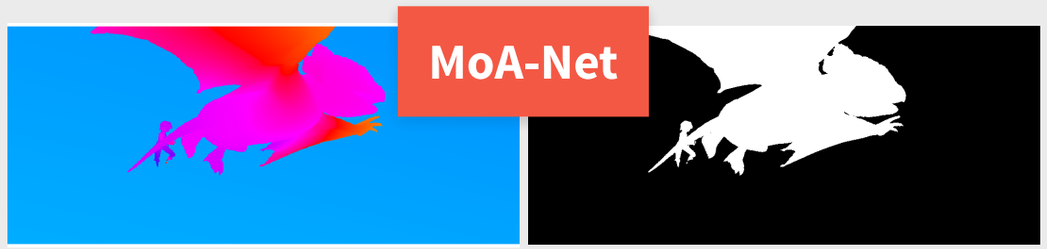
Motion Segmentation



optical flow

rotation compensated flow

step 1: rotation compensation



angle field

motion segmentation

step 2: motion segmentation

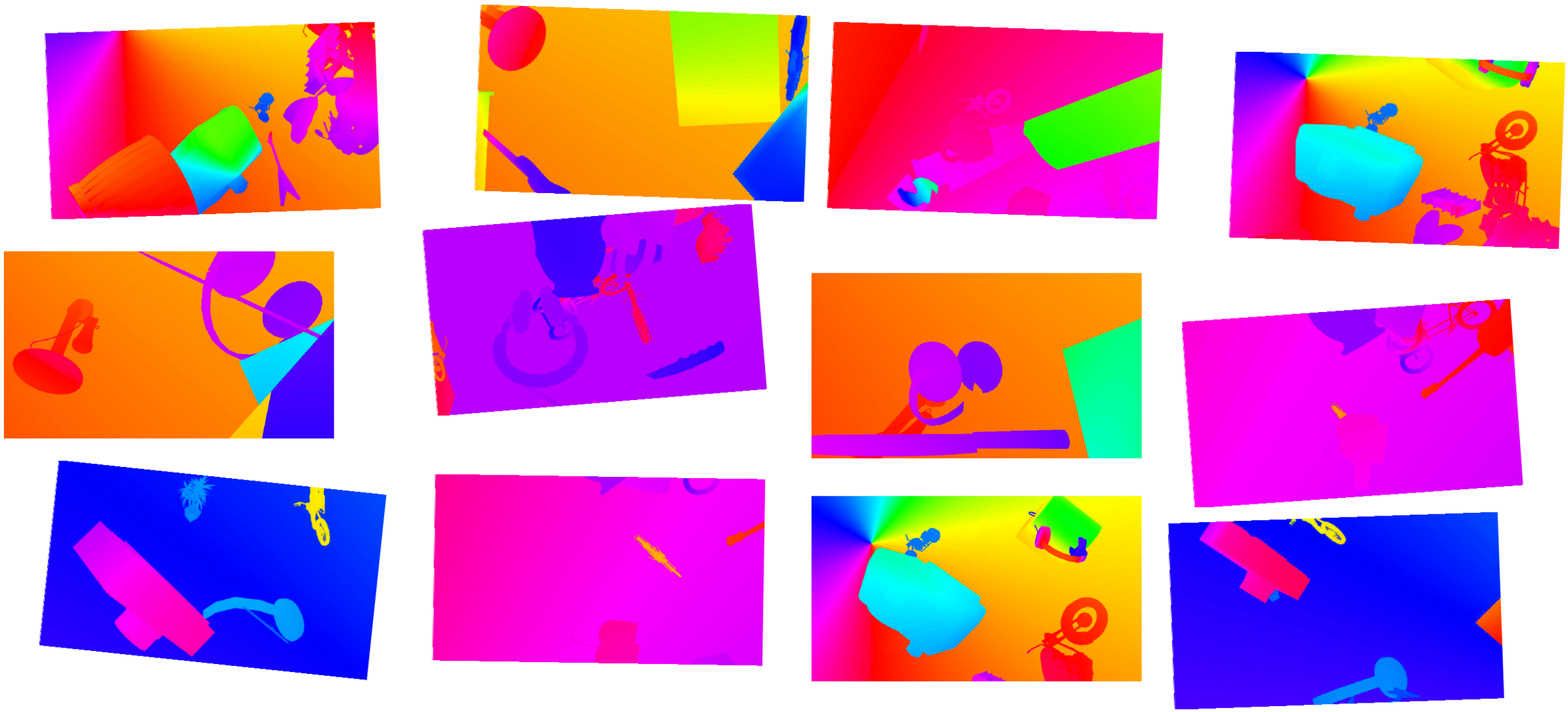
Motion Segmentation
Definition

Def.: Moving Object

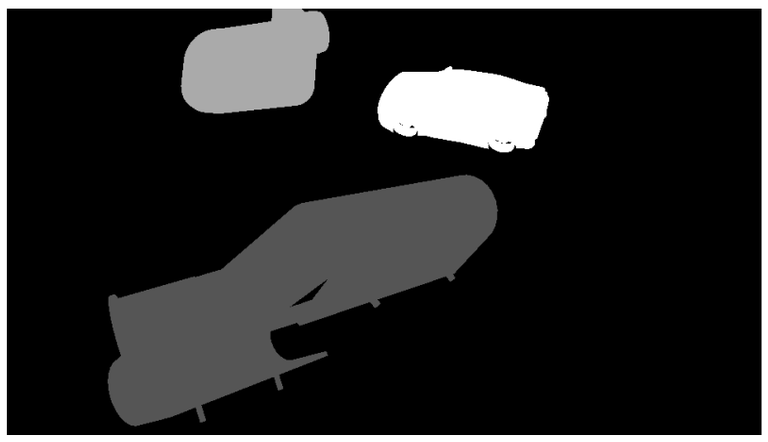
A moving object is a connected image region that undergoes some independent motion. The connected image region can be of any size and shape.

Motion Segmentation

Generating training data

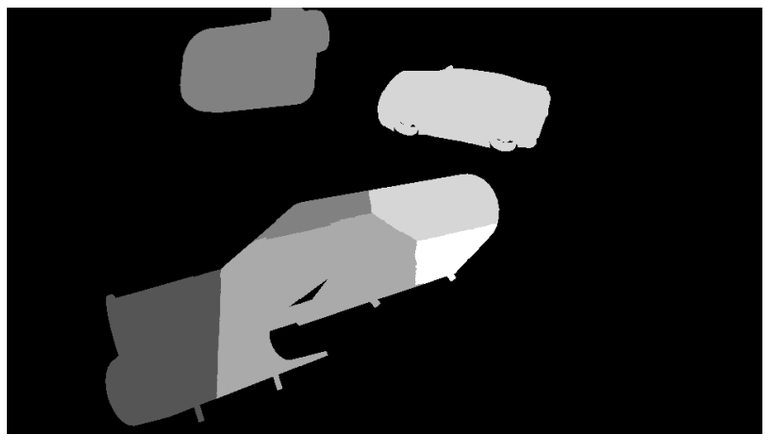


Motion Segmentation
Generating training data



- Generating connected object regions.
- Splitting each object into n subregions.
- Assigning to each motion region a translational 3D direction.
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

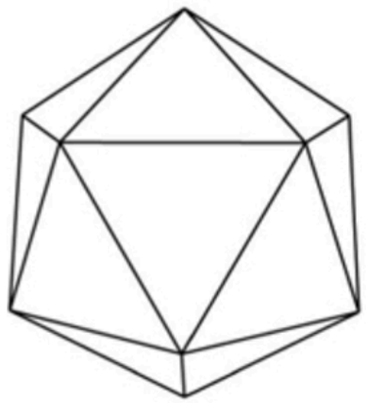
Motion Segmentation
Generating training data



- Generating connected object regions.
- Splitting each object into n subregions.
- Assigning to each motion region a translational 3D direction.
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

Motion Segmentation

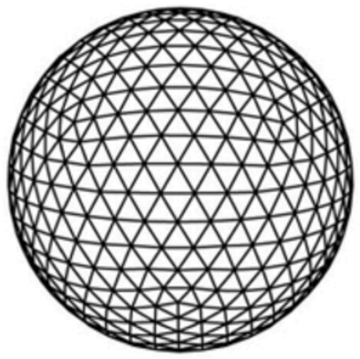
Generating training data



$$\theta = \texttt{atan}(-fV + yW, -fU + xW)$$
$$= \texttt{atan}(-V' + yW, -U' + xW)$$

- Generating connected object regions.
- Splitting each object into n subregions.
- **Assigning to each motion region a translational 3D direction.**
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

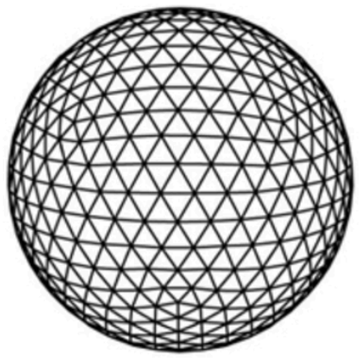
Motion Segmentation
Generating training data



$$\theta = \texttt{atan}(-fV + yW, -fU + xW)$$
$$= \texttt{atan}(-V' + yW, -U' + xW)$$

- Generating connected object regions.
- Splitting each object into n subregions.
- **Assigning to each motion region a translational 3D direction.**
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

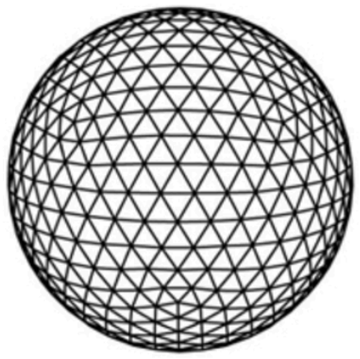
Motion Segmentation
Generating training data



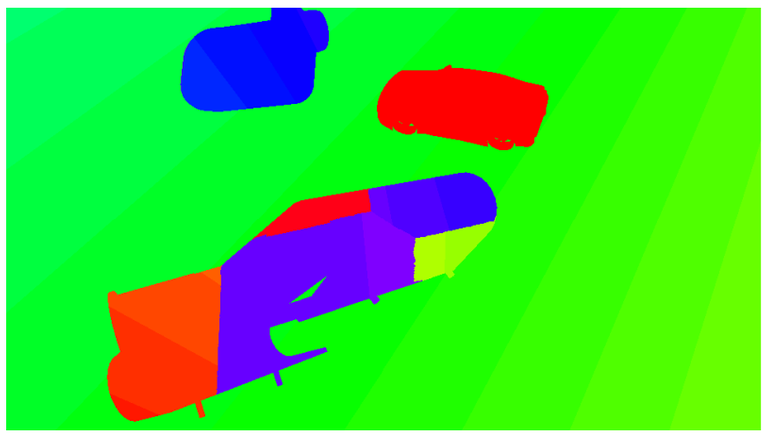
$$\theta = \texttt{atan}(-fV + yW, -fU + xW)$$
$$= \texttt{atan}(-V' + yW, -U' + xW)$$

- Generating connected object regions.
- Splitting each object into n subregions.
- **Assigning to each motion region a translational 3D direction.**
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

Motion Segmentation
Generating training data

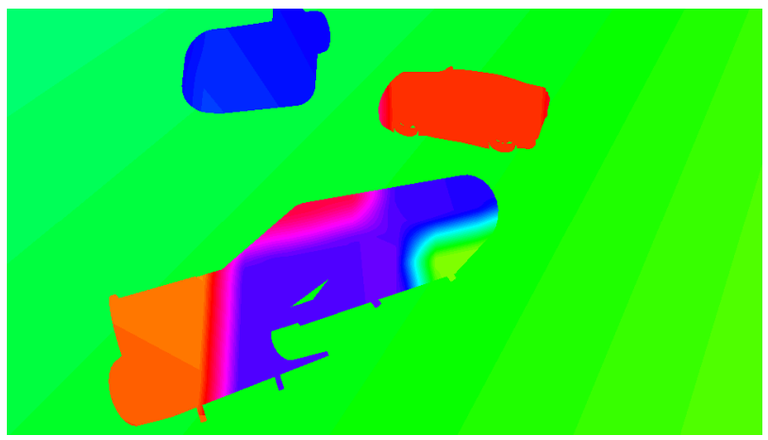


$$\theta = \texttt{atan}(-fV + yW, -fU + xW)$$
$$= \texttt{atan}(-V' + yW, -U' + xW)$$



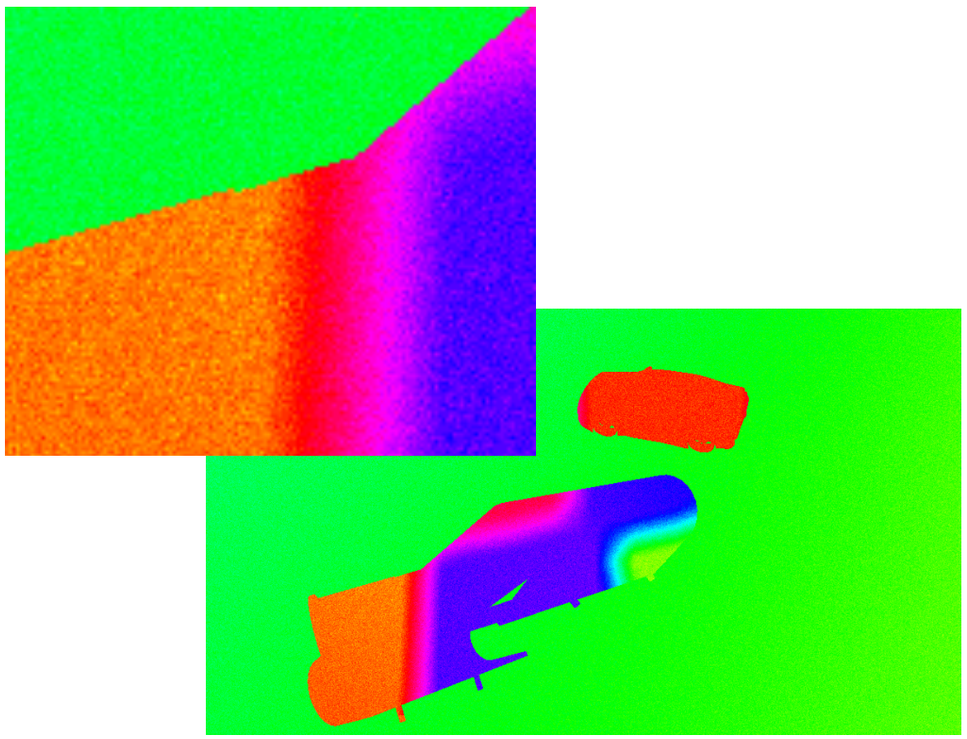
- Generating connected object regions.
- Splitting each object into n subregions.
- **Assigning to each motion region a translational 3D direction.**
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

Motion Segmentation
Generating training data



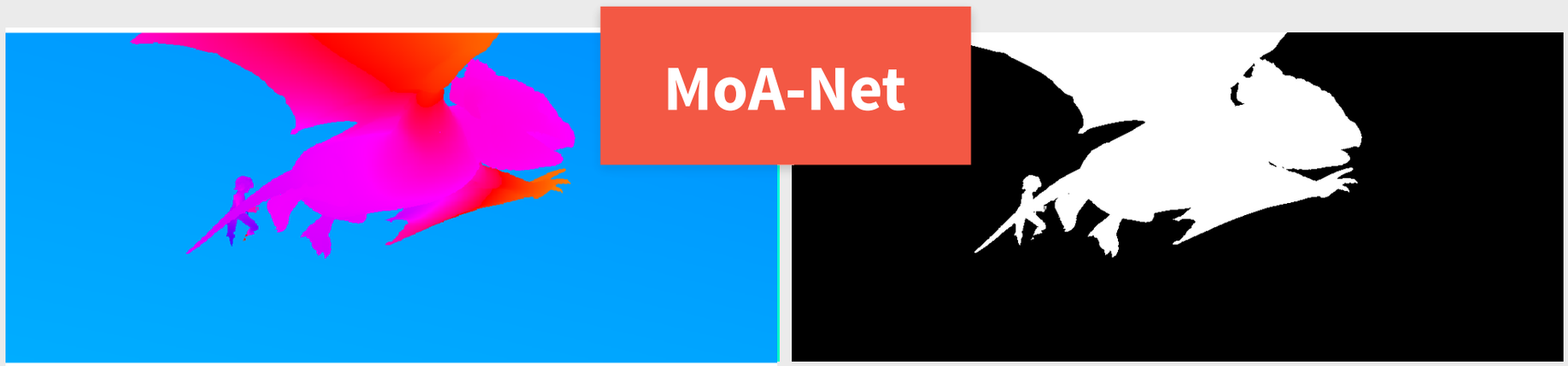
- Generating connected object regions.
- Splitting each object into n subregions.
- Assigning to each motion region a translational 3D direction.
- **Smoothing motion boundaries inside moving objects.**
- Adding random gaussian noise.

Motion Segmentation
Generating training data



- Generating connected object regions.
- Splitting each object into n subregions.
- Assigning to each motion region a translational 3D direction.
- Smoothing motion boundaries inside moving objects.
- Adding random gaussian noise.

Motion Segmentation



angle field | motion segmentation

step 2: motion segmentation

Overview

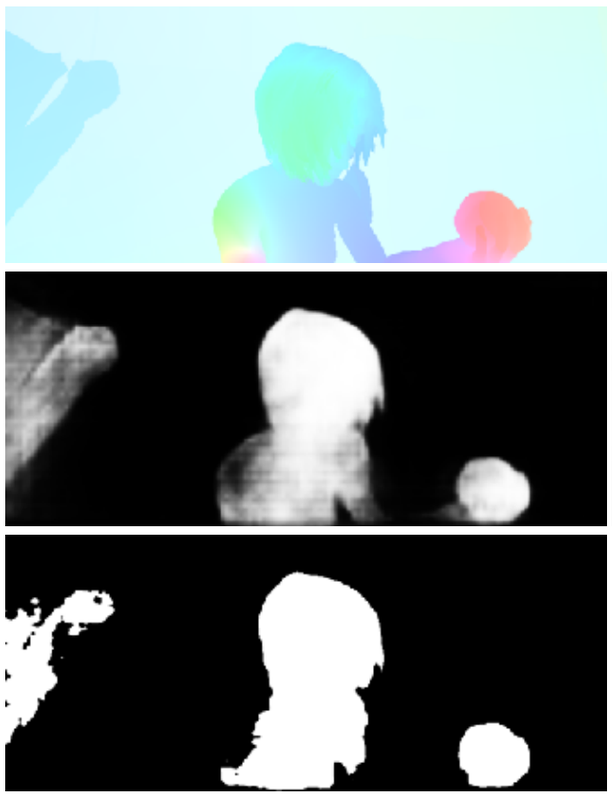
- Motivation
  How do humans know
  what is moving in the world and what is not?
- Approach: Motion Segmentation
  - Rotation compensation
  - Learning Motion Patterns: MoA-Net
- Results
- Future Research Questions

Segmentation Results

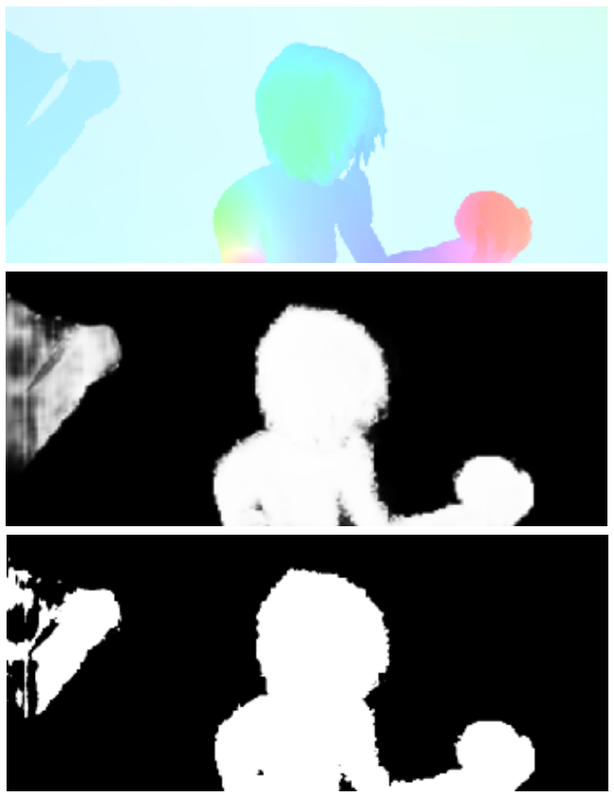
video frame

ground truth

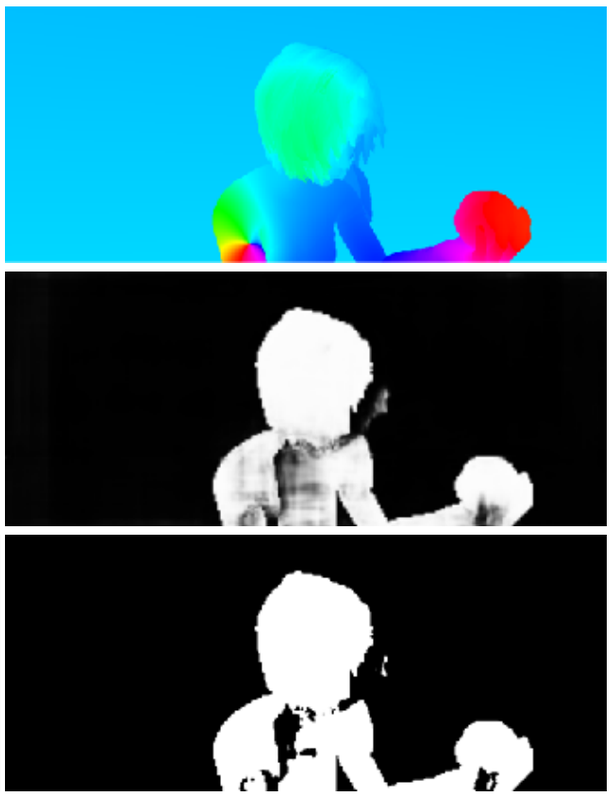
Jain et al.



Tokmakov et al.



ours



Segmentation Results

	Motion Segmentation: Sintel					
	J Mean ↑	J Recall ↑	J Decay ↓	F Mean ↑	F Recall ↑	F Decay ↓
Tokmakov [1, 2]	50.46	**55.43**	44.50	53.43	35.04	39.75
Jain et al. [3]	29.63	24.98	36.07	28.65	14.70	31.20
ours	**54.77**	54.47	**26.57**	**59.71**	**61.38**	**14.79**

Segmentation Results

Overview

Motivation

How do humans know
what is moving in the world and what is not?

Approach: Motion Segmentation

- Rotation compensation
- Learning Motion Patterns: MoA-Net

Results

Future Research Questions

Future Research Questions

angle field

magnitude

- Importance of the flow **magnitude** for
    - estimating the *scene depth*
    - dealing with *estimated (noisy) optical flow*

-QUESTIONS-