COMPUTER VISION 2021 - 2022

> OPTICAL FLOW

UTRECHT UNIVERSITY

RONALD POPPE

OPTICAL FLOW

What is shown here?

OPTICAL FLOW[2]

Essential for analyzing motion in video

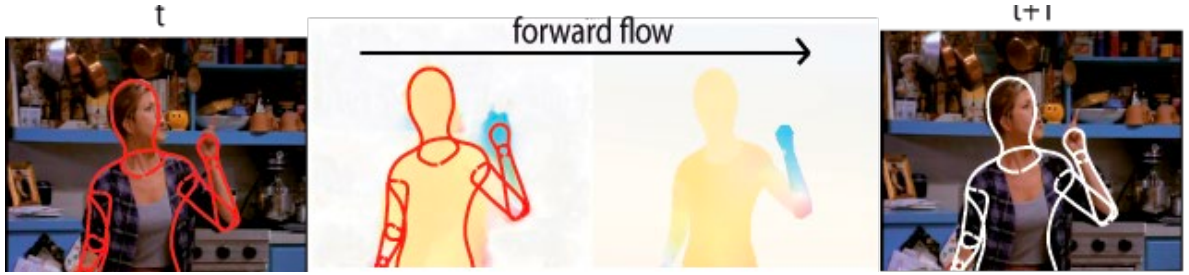
- Tracking
- Recognition (of actions, activities or scenes)
- Alignment (in time)
- Depth estimation (from a single camera)

Often used in combination with spatial image features

- Also employed in convolutional neural networks

OPTICAL FLOW[3]

Zuffi et al., ICCV 2013



Alldieck et al., GCPR 2017

OPTICAL FLOW OVER THE YEARS

1940 Term coined by JJ Gibson

1981 Lucas & Kanade: sparse optical flow

1981 Horn & Schunck: global estimations with smoothness contraint

2004 Brox et al.: warping-based

2010 Sun et al.: performance evaluation, median filtering

2011 Brox & Malik: variational optical flow

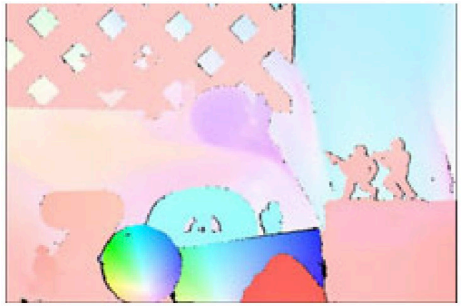
2015 Fischer et al.: FlowNet

2017 Ilg et al.: FlowNet 2.0

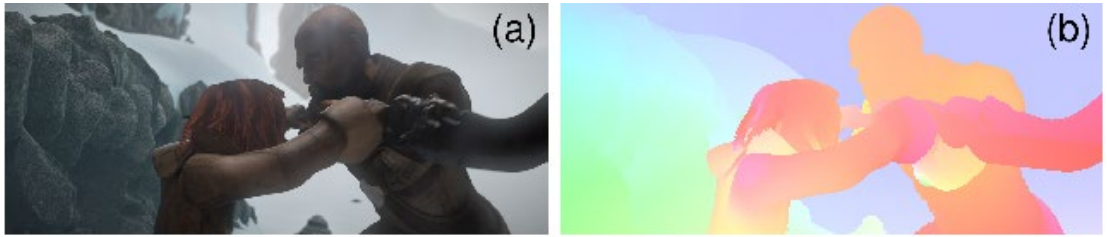
2020 Teed & Deng: multi-scale correlation

DATASETS OVER THE YEARS

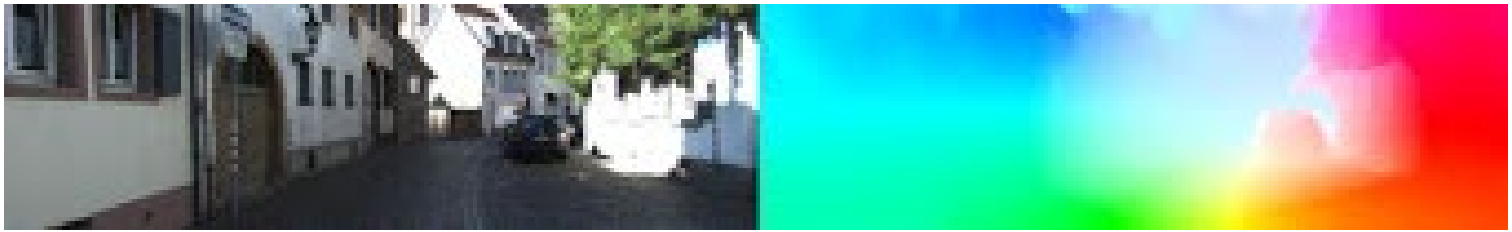
Middlebury (2007)



Sintel (2012)



KITTI (2015)



FARNEBACK VS. LUCAS-KANADE

Lucas-Kanade was originally a sparse algorithm

- Based on neighborhoods
- Also exists in a dense variation

Farnebäck uses second-order polynomial expansion of neighborhoods

- Allows for the integration of prior knowledge (certainty)

Sparse/dense Lucas-Kanade and Farnebäck are supported in OpenCV

MATERIALS

FlowNet 2.0: https://www.youtube.com/watch?v=JSzUdVBmQP4

Paperswithcode: https://paperswithcode.com/task/optical-flow-estimation

Optical flow survey:
https://www.sciencedirect.com/science/article/abs/pii/S0923596518302479

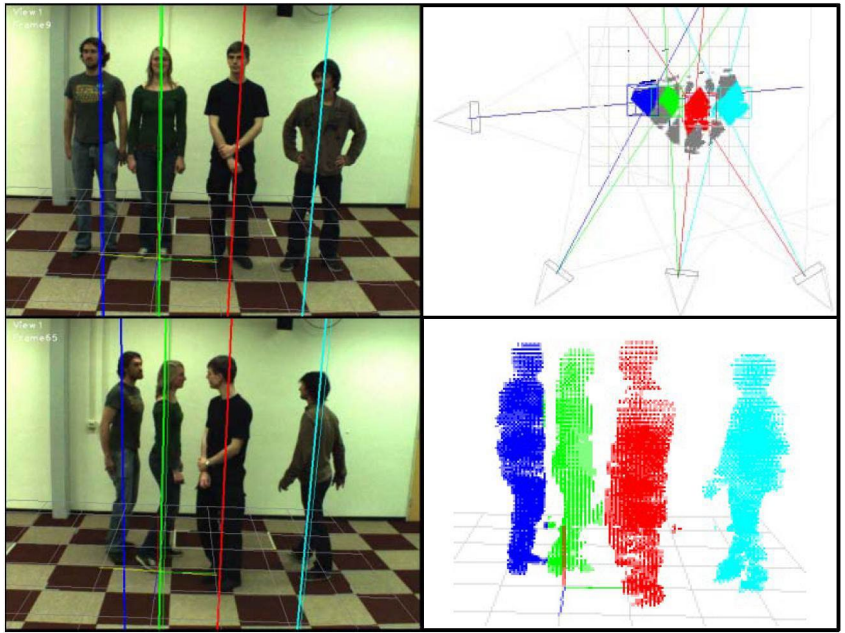
QUESTIONS?

VOXEL-BASED LABELING

VOXEL-BASED LABELING

Goal: determine which voxels belong to which object/subject

- Color indicates the subject identity



TRACKING AND LABELING

Tracking: “following” people over time

- Keeping a consistent labeling of the voxel clusters

Usually:

- Cluster and label
- Use temporal information (previous position/movement)

In Assignment 3, we use a different approach

FIRST ANALYSIS

Voxels that are close together, are likely to be from the same person

- Height does not play a role → ignore it?
- Cluster voxels based on location in 3D space?

In an image, voxels can be projected to each view

Should then be “on” a subject

- Each subject might be wearing differently colored clothing
- Label voxels based on colors?

CONCEPT

For our voxel-based tracking, we use two types of input:

- The voxel model
- The 2D images from each view

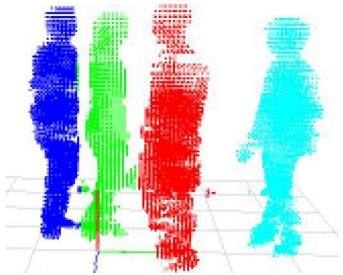
We'll use concepts we've discussed previously:

- Clustering of voxels (3D)
- Color models (image)

CONCEPT²

Initialization (offline):

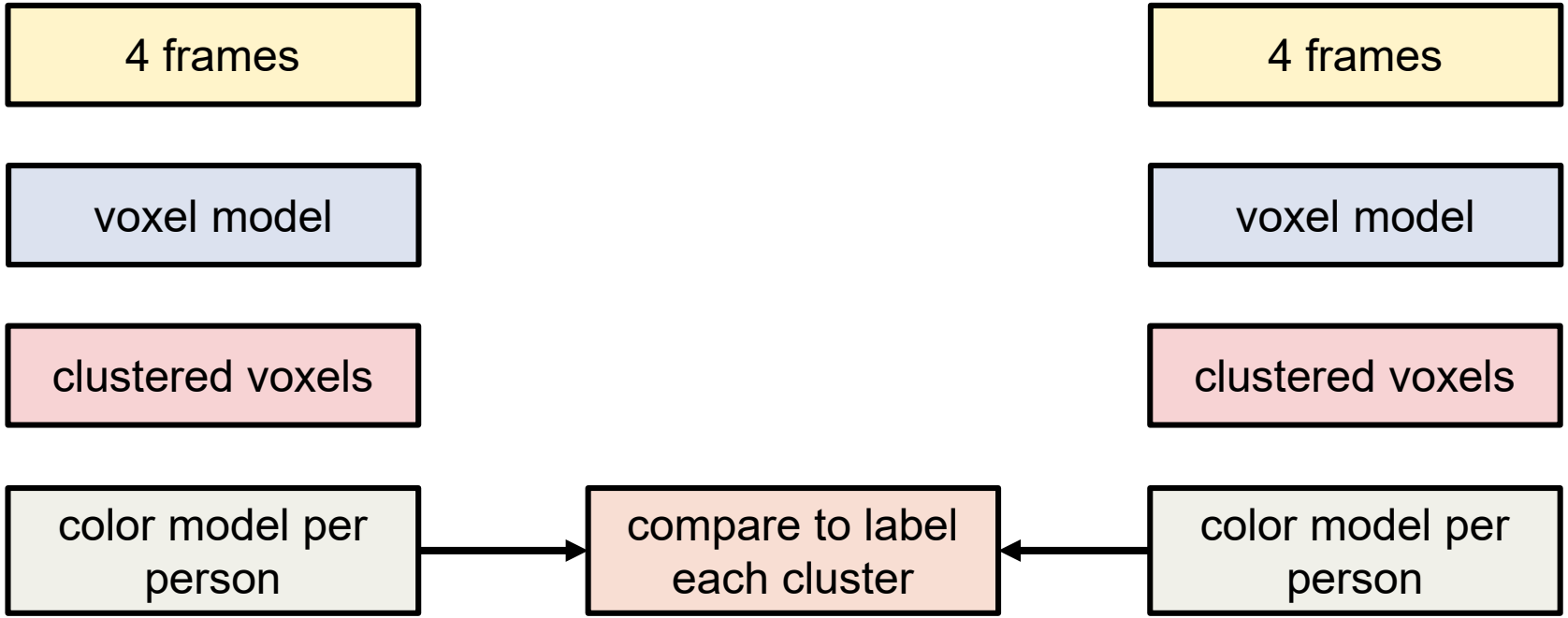
1. Cluster voxels into persons using only voxel locations
2. Make a color model for each person (project to image)



For each next frame (online):

3. Cluster these voxels to form the new person locations
4. Label voxels to persons using color models (compare to offline model)

CONCEPT[3]


4 frames
voxel model
clustered voxels
color model per person
compare to label each cluster
4 frames
voxel model
clustered voxels
color model per person

INITIALIZATION

When clustering the voxels into persons:

- Choose a frame where everyone is visible and well-seperated
- The number of clusters is equal to the number of subjects
- Run a clustering algorithm based on the location of the voxel
- Ignore the height!
- Check if clusters are close to each other (you're stuck in a local minimum)

Output:

- Cluster centers corresponding to the location of each person (on the ground plane, so (x,z) location)
- A label of each voxel to which person it belongs

INITIALIZATION^2

Next step is to make a color model for each person

We need to know which pixels belong to which person:

- Project the labeled voxels to view (multiple views for additional points)
- Make sure that the pixels are visible (occlusion!)

Construct the color model (per view or all views together):

- Histogram
- GMM
- Mean color (will not give you the full points)

INITIALIZATION^3

You might want to make a smart color model:

- Use only the lower/upper part?
- Discard dark pixels?
- Etc.

ONLINE PROCESSING

The online processing deals with labeling the voxels in subsequent frames

We will use color cues!

- Normally, you would track the position, based on the previous position and an estimate of the movement
- E.g., using Kalman filters

ONLINE PROCESSING²

Online processing: label the voxels in subsequent frames using color

First, cluster the voxels. Each cluster should correspond to a person, but we do not know which (because of the random initialization)

- And we're not going to cheat by using the previous clusters as initialization!
- So this step is exactly the same as in the offline phase

Then, find out which cluster belongs to which person:

- Project voxels of one cluster to one camera (occlusion!)
- Determine the color of the pixels
- Use a suitable measure for distance between pixels and model

ONLINE PROCESSING[3]

This will give you (per voxel or per cluster), a person label

- We need to combine all those labels into a coherent (final) labels
- Majority voting might work
- But avoid assigning the same label to two people!

ONLINE PROCESSING[4]

Once we know which voxels (or cluster) belongs to each person, we determine the position of each person:

- Outliers have quite a large impact on the estimation
- Optionally: iteratively filter out outliers
- Optionally: if you can improve the tracks, you can earn additional points (in case you might use tracking or smoothing)

CHOICE TASKS

Using multiple cameras to increase the robustness: 10

- Instead of projecting to a single view, use multiple cameras
- Helps when a person is occluded in your view

Finding a way to deal with people outside the voxel space (finding out when K=4 clusters is too much): 10

- Occasionally, a person leaves the view area
- Determine when this happens and adjust the number of clusters accordingly

CHOICE TASKS^2

Implementing tracking based on previous position and movement: 10 (in addition, not as replacement, to the color-based labeling)

- Think about using the position, direction or speed
- Using the previous positions as a starting point for your clustering will give you max 5

Smoothing of trajectories: 5 (can be a post-processing step)

- Get rid of noise in the estimation of the ground location

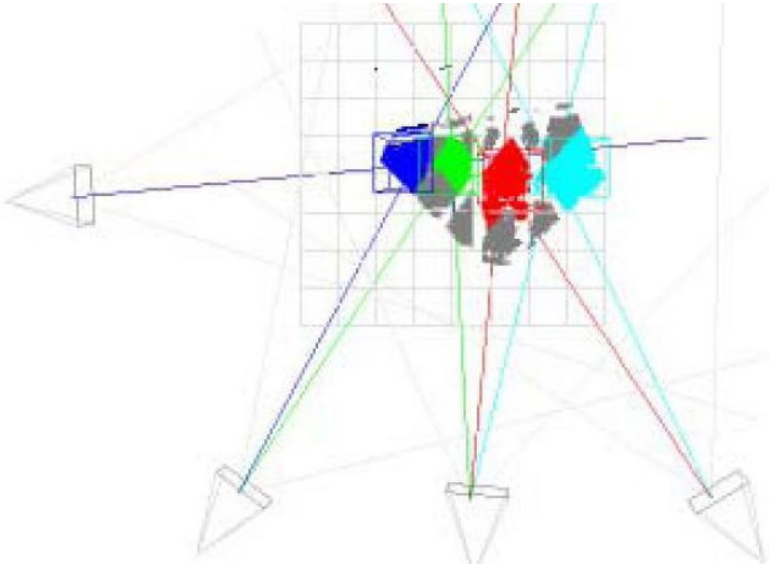
CHOICE TASKS[3]

Getting rid of outliers and ghost voxels: 10 (based on the clustering, find out which voxels should be removed altogether so they don't influence the calculation of the center)

- Only filtering of outliers: max 5

Updating your color model over time: 10

- Adapt the color model of each person



ASSIGNMENT

ASSIGNMENT

Assignment 2:

- Deadline is Wednesday March 2, 23:00 (tomorrow)

Assignment 3:

- Deadline is Sunday March 13, 23:00

NEXT LECTURE

Next lecture: Training, classification, detection

- Thursday March 3, 11:00-12:45, BOL-0.204

QUESTIONS?