Mapping II - learning for mapping

Karel Zimmermann

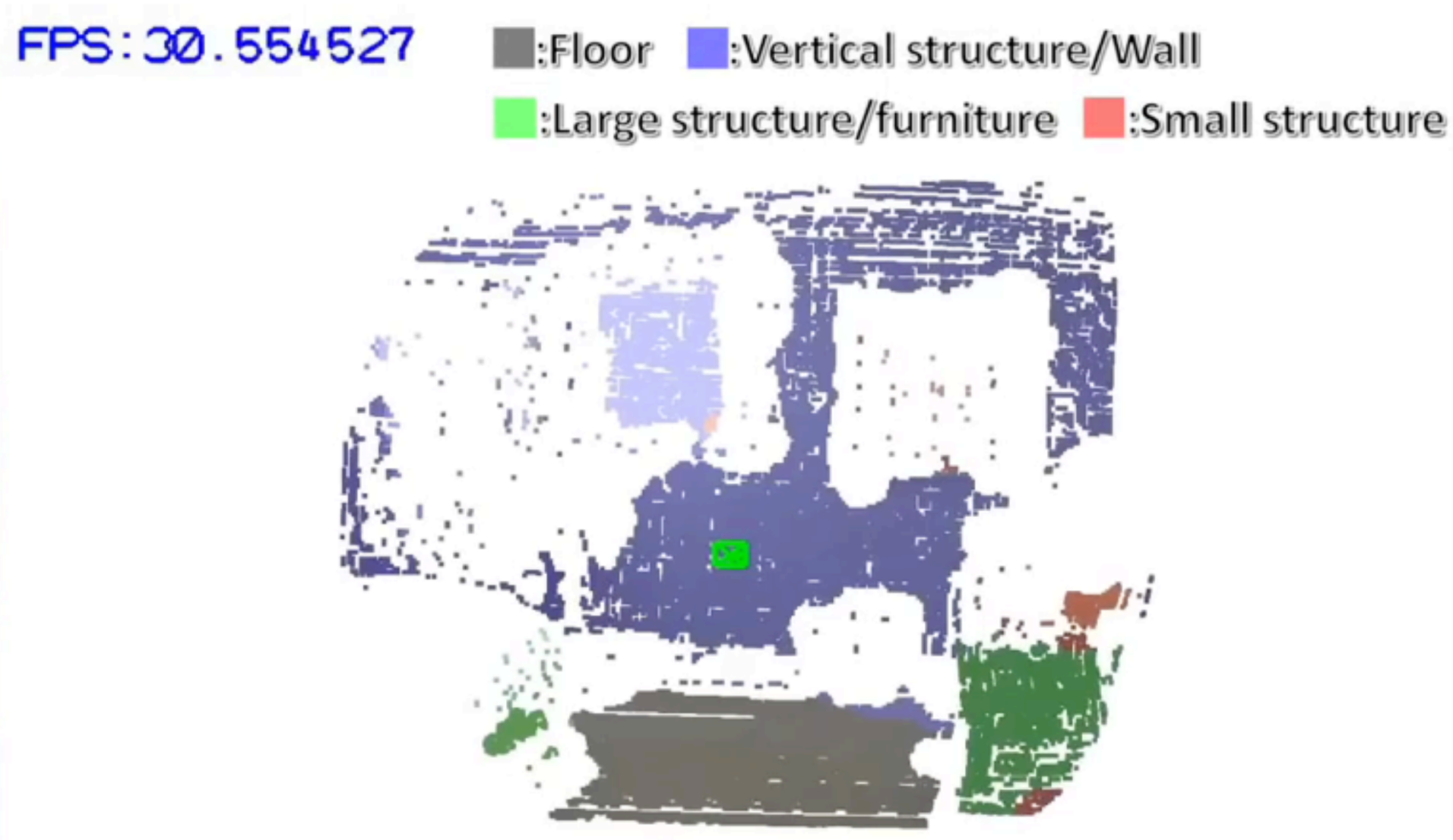
http://cmp.felk.cvut.cz/~zimmerk/

Vision for Robotics and Autonomous Systems
https://cyber.felk.cvut.cz/vras/

Center for Machine Perception
https://cmp.felk.cvut.cz

Department for Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

Semantic segmentation [Tateno CVPR 2017]



Result of dense 3D reconstruction and semantic label fusion

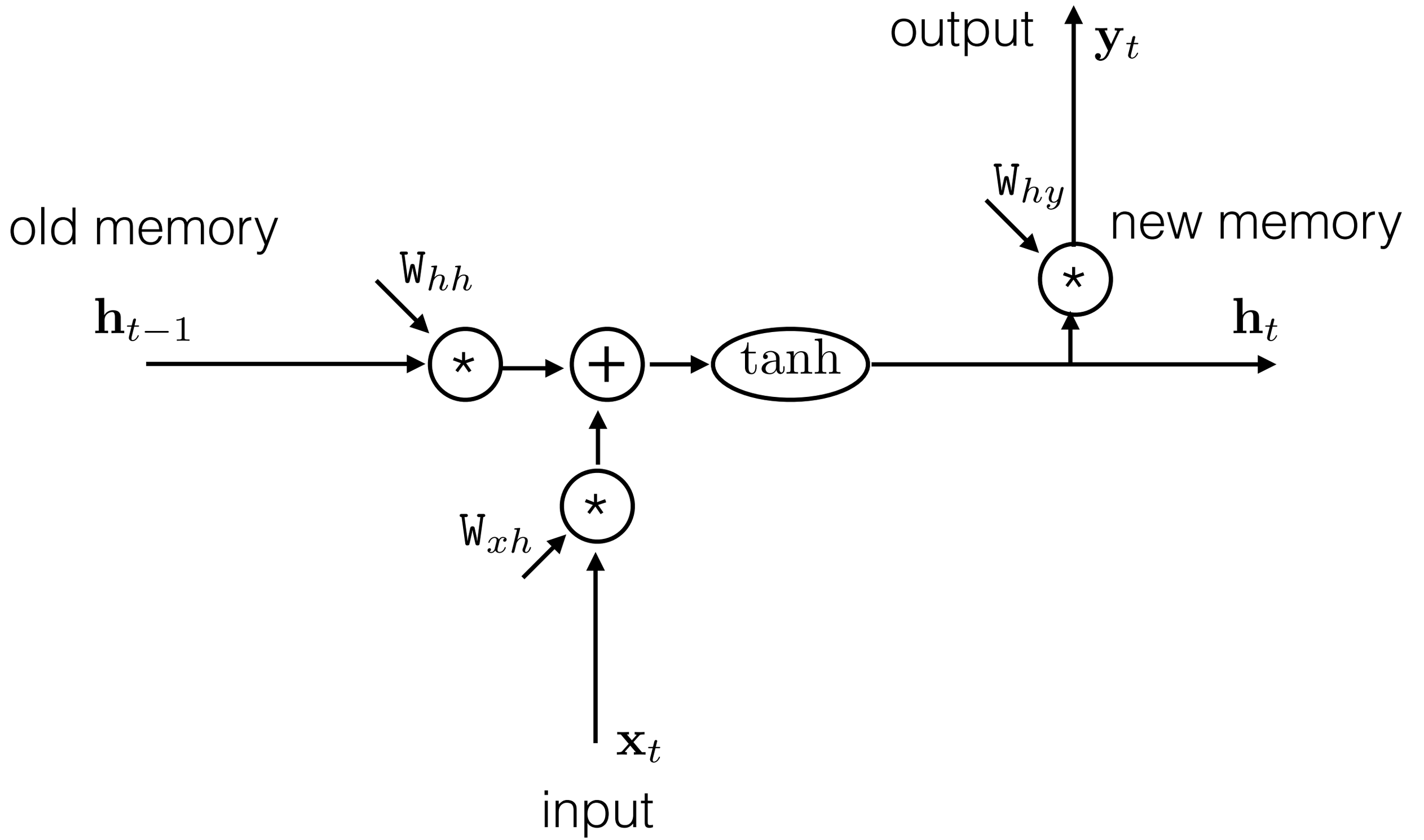
https://arxiv.org/pdf/1704.03489.pdf

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

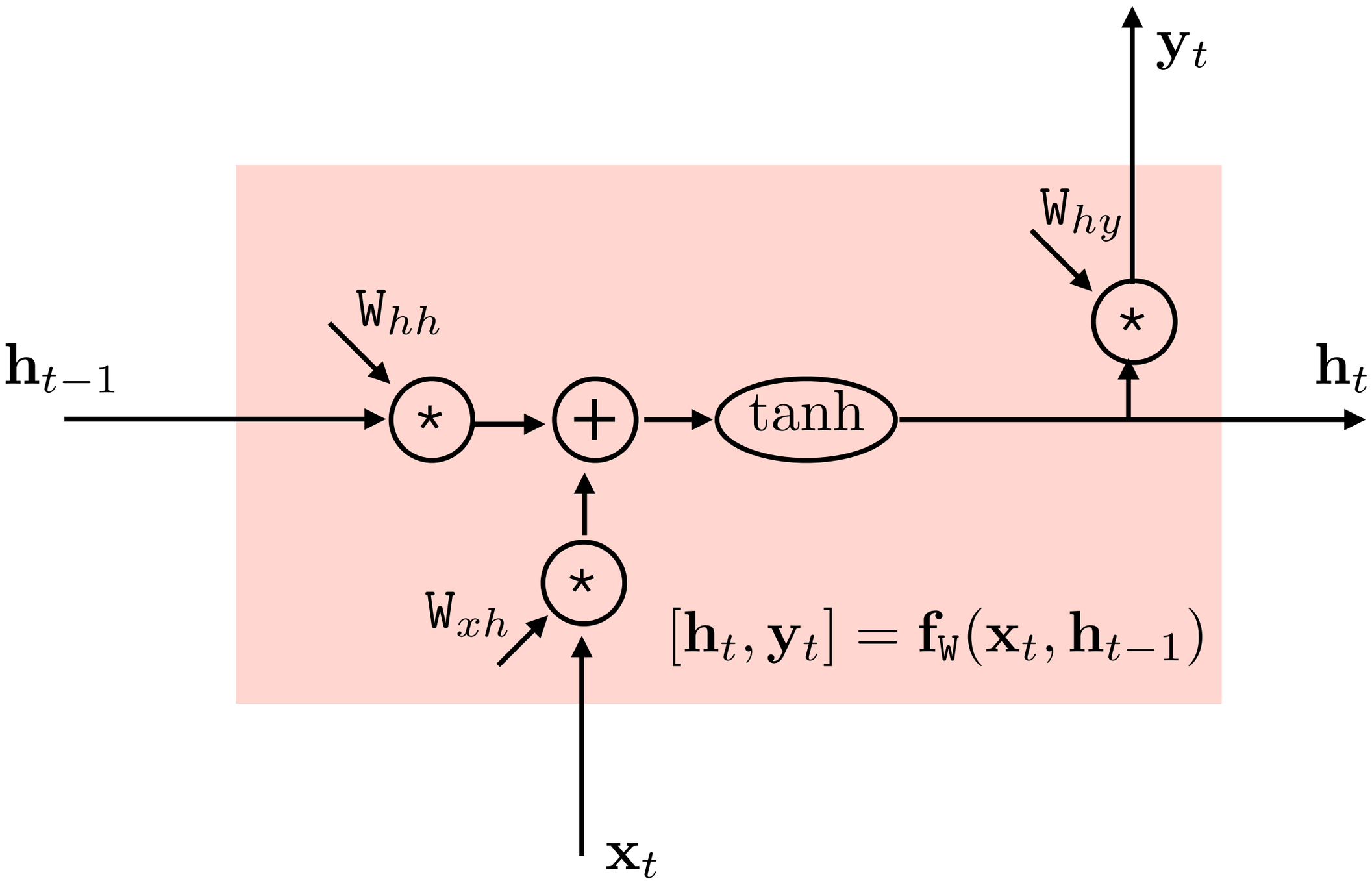
Mapping summarises measurements into a map

- Suitable learnable structure for this summarisation process is a recurrent network
- Other applications of RNN:
  - Natural language translation
  - Image captioning
  - Video action recognition / labelling
  - Motion control (e.g. motion model or policy)
    - Everywhere, where
      - a simple memory (i.e. last 4 frames) is not enough
      - or reasonable explicit representation (i.e. map) is not available

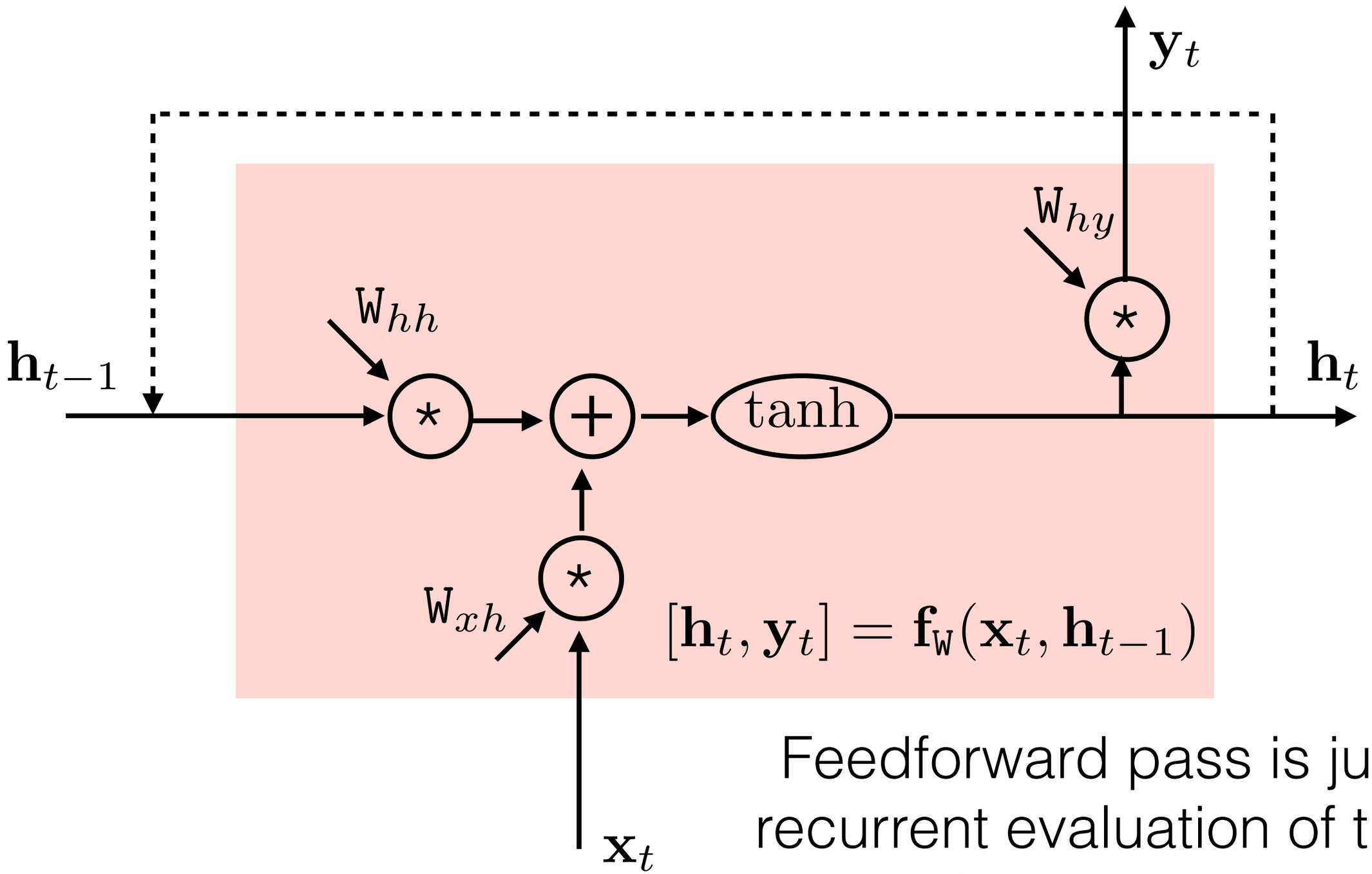
Simple block



Simple block

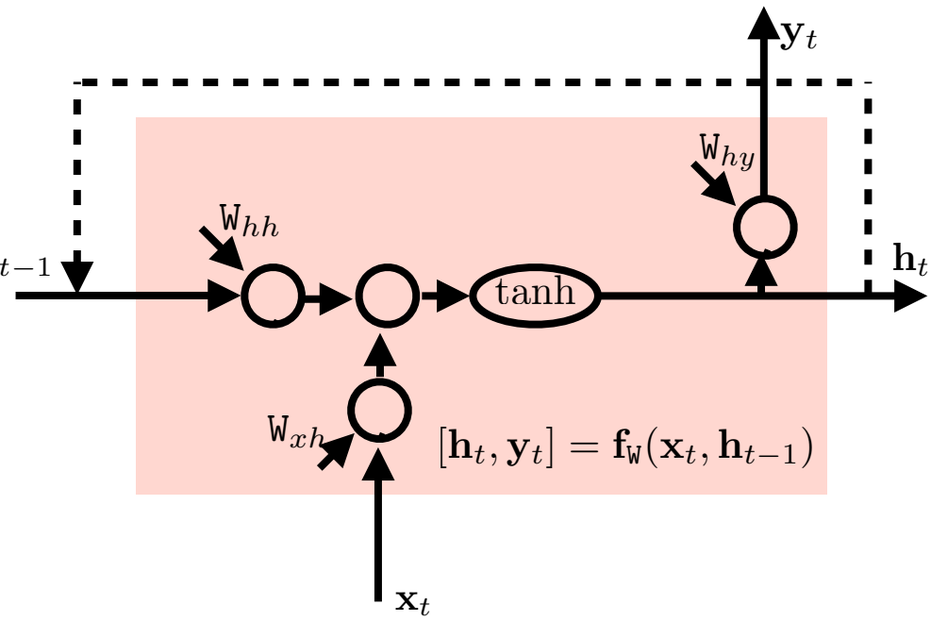


Simple recurrent block



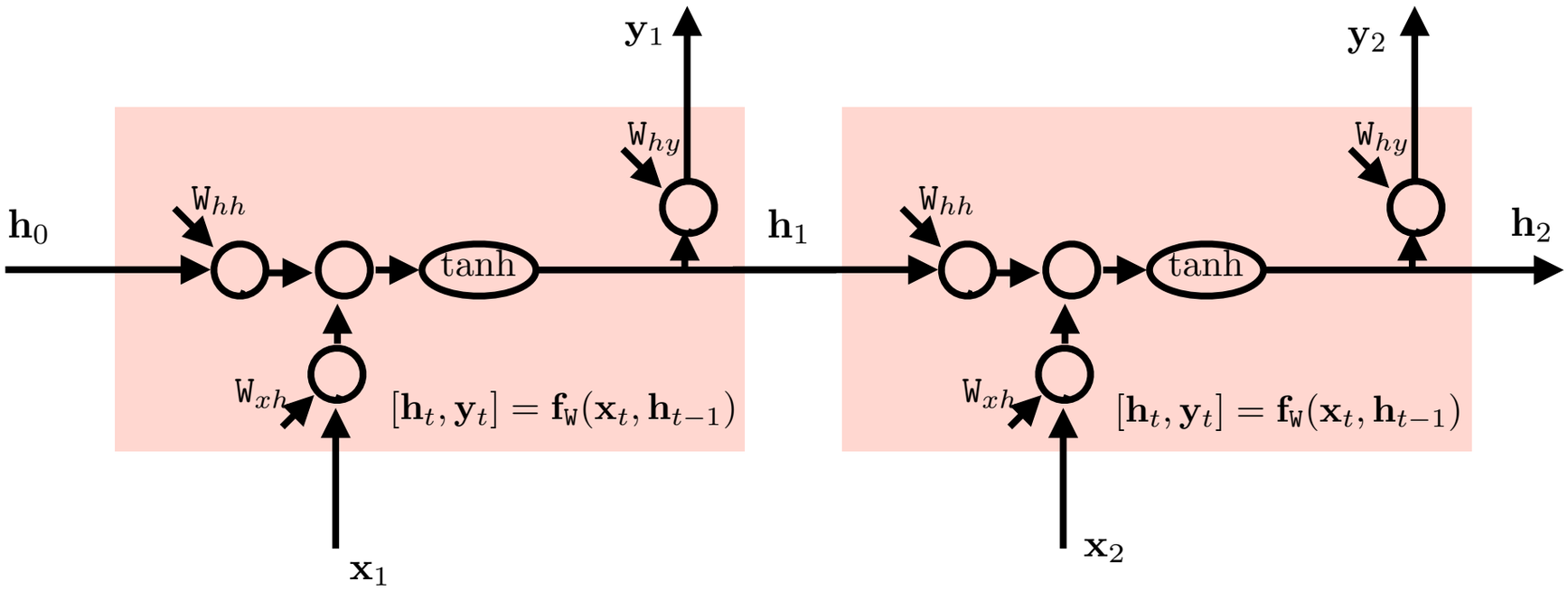
Feedforward pass is just recurrent evaluation of this expression from t=1 to t=T.

Simple recurrent block



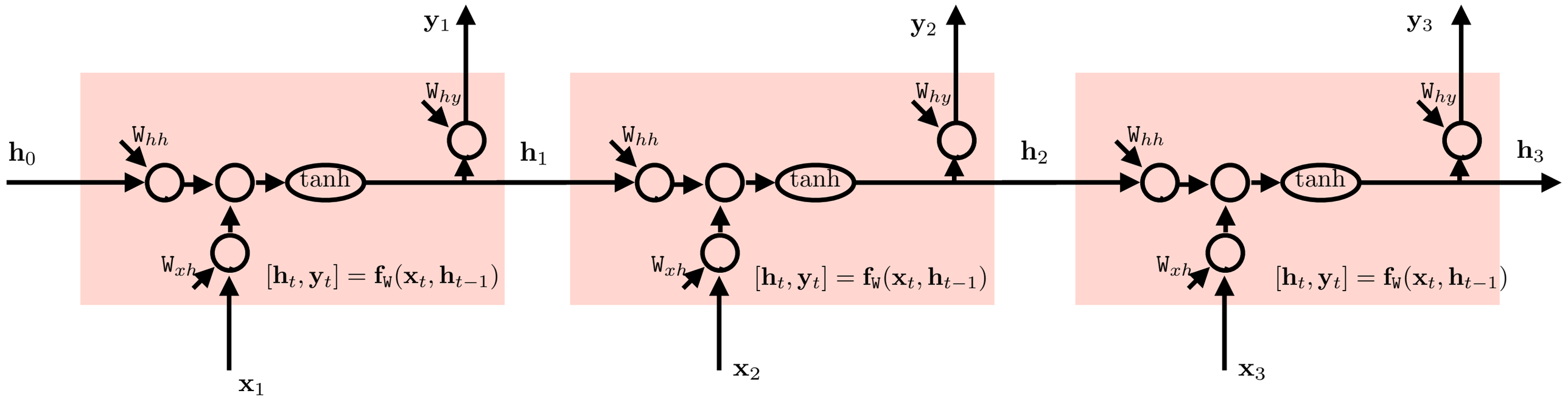
We remove recurrent connection by unrolling the net

Simple recurrent block unrolled



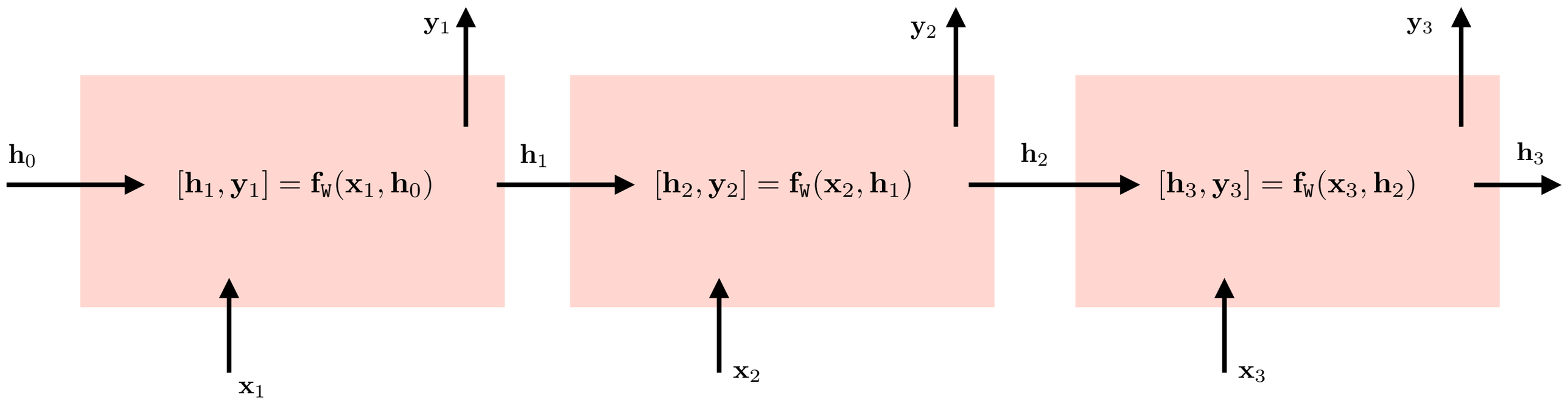
- by successively repeating evaluation of
  - the same function $[\mathbf{h}_t, \mathbf{y}_t] = \mathbf{f}_\mathtt{W}(\mathbf{x}_t, \mathbf{h}_{t-1})$
  - on different inputs/outputs $\mathbf{h}_0$ $\mathbf{h}_1$ $\mathbf{h}_2$ $\mathbf{x}_1$ $\mathbf{x}_2$

Simple recurrent block unrolled



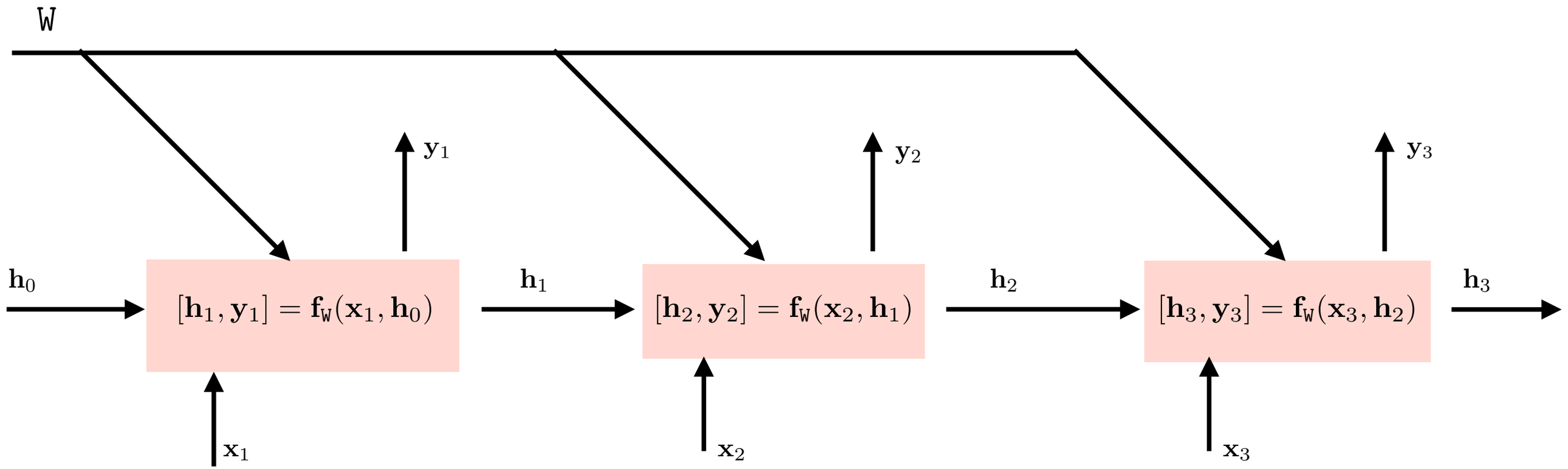
- Unrolled computational graph:
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!

Simple recurrent block unrolled



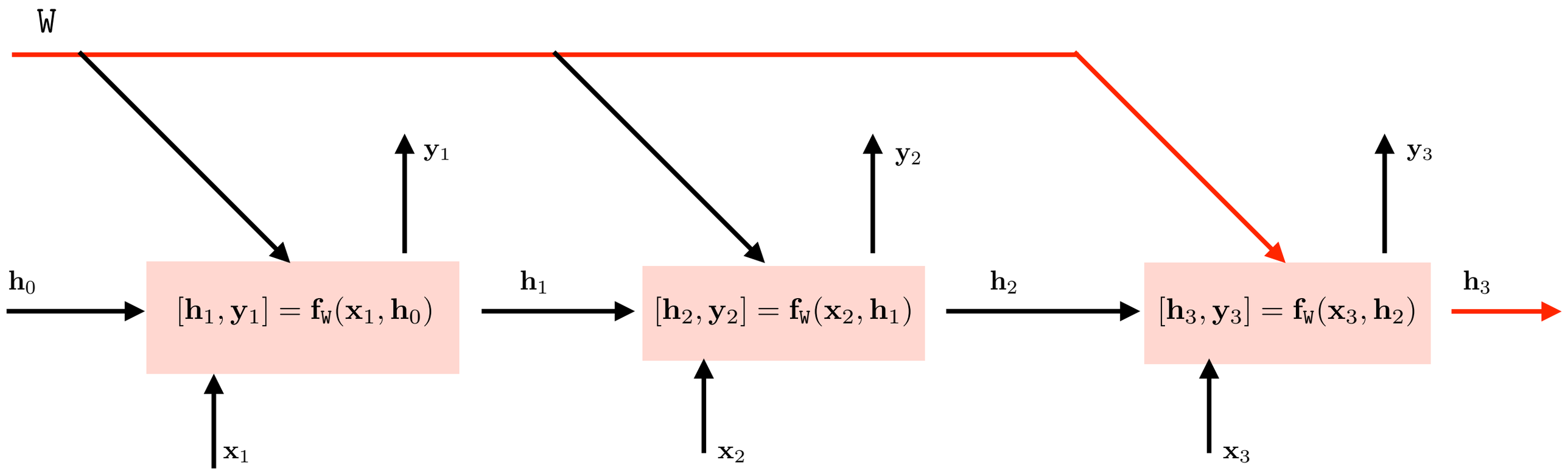
- Unrolled computational graph:
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!

Simple recurrent block unrolled



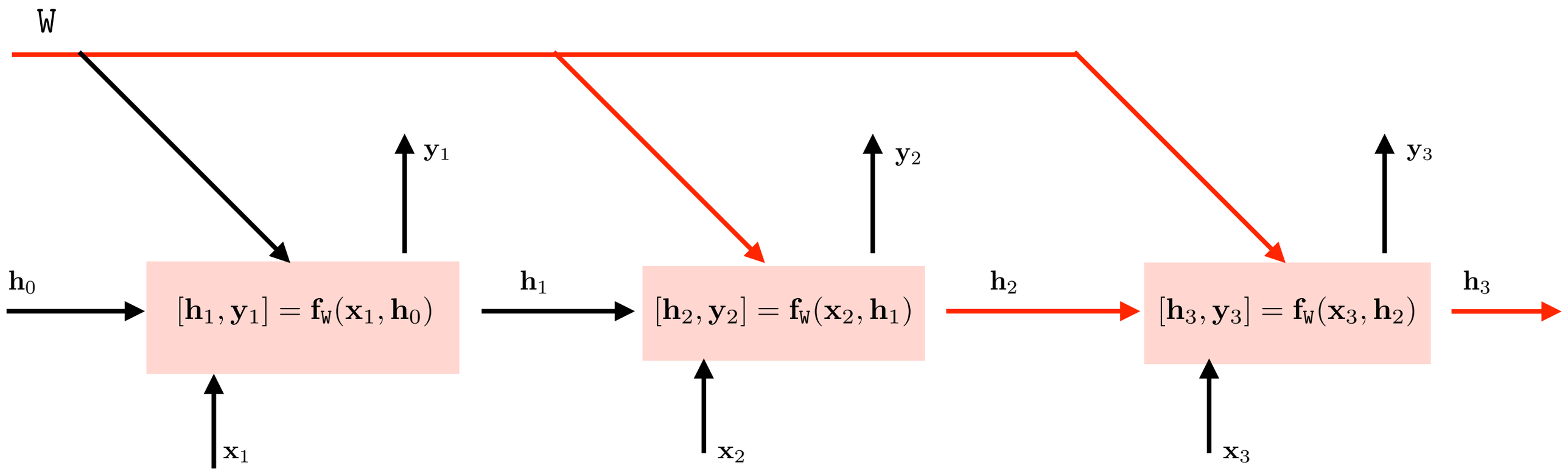
- Unrolled computational graph:
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!

Simple recurrent block unrolled - backpropagation



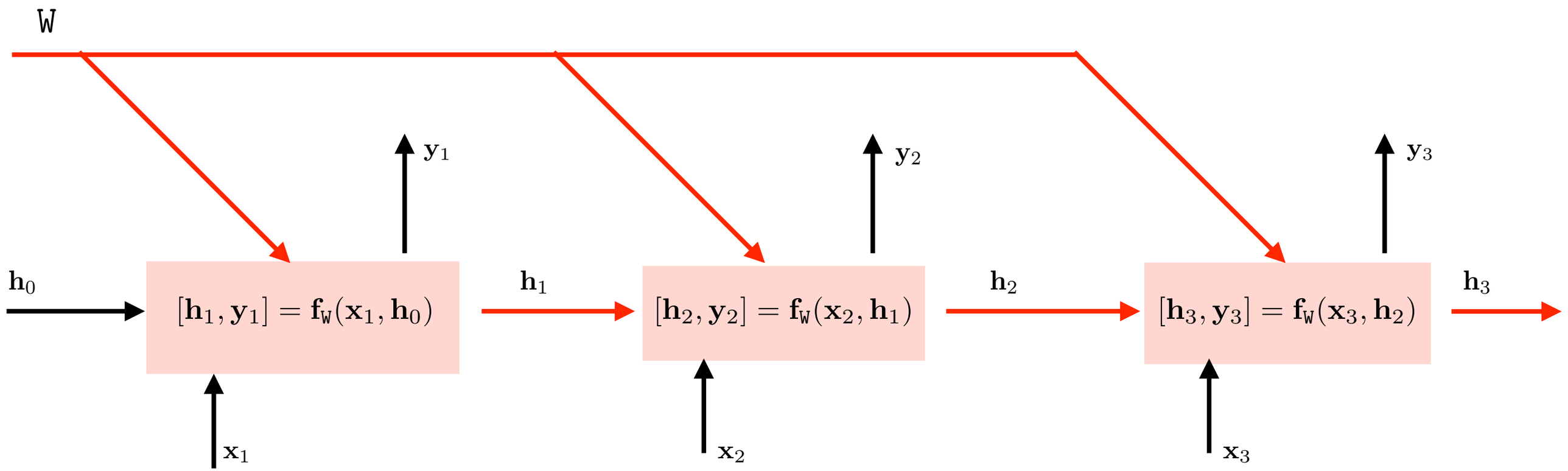
- Backpropagation
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!

Simple recurrent block unrolled - backpropagation



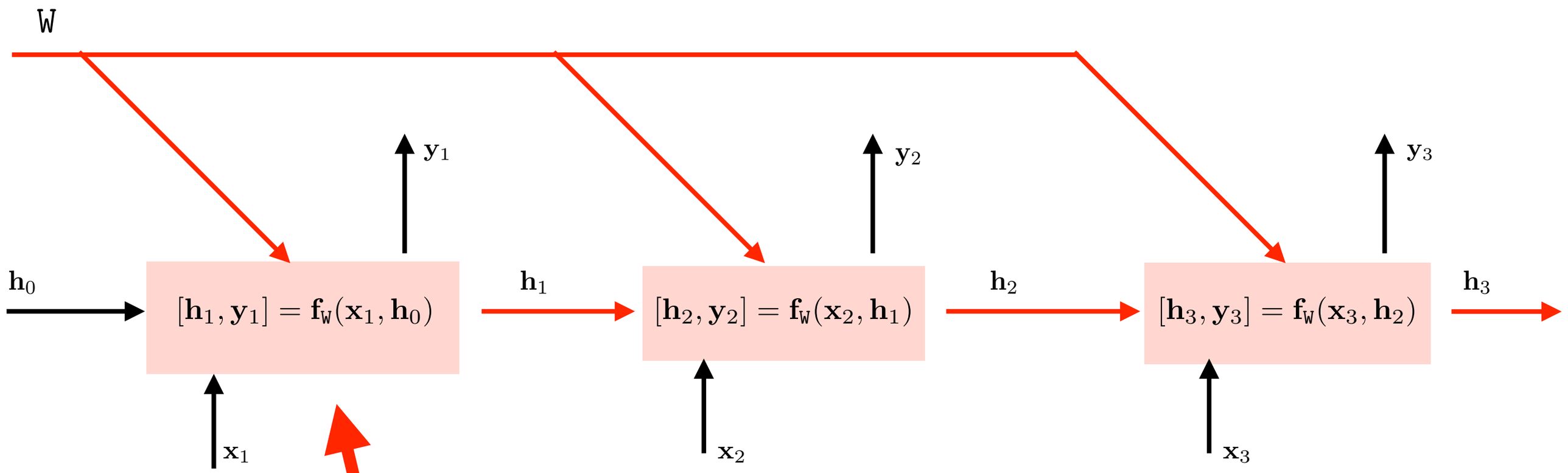
- Backpropagation
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!
  - gradient summed over all unrolled blocks (over time)

Simple recurrent block unrolled - backpropagation



- Backpropagation
  - it is normal feedforward network
  - it consists of several same blocks with the same weights!
  - gradient summed over all unrolled blocks (over time)

Simple recurrent block unrolled - backpropagation



deep blocks often suffer from vanishing gradient
=> better structure needed

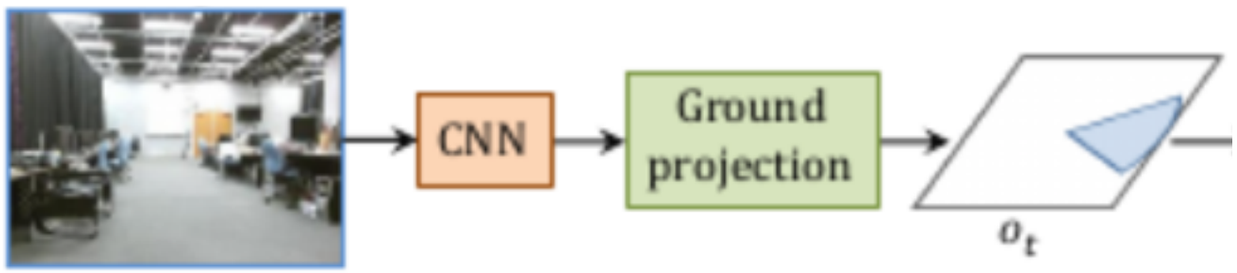
the most general structure is LSTM
(kind of ResNet for recurrent networks)

MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/

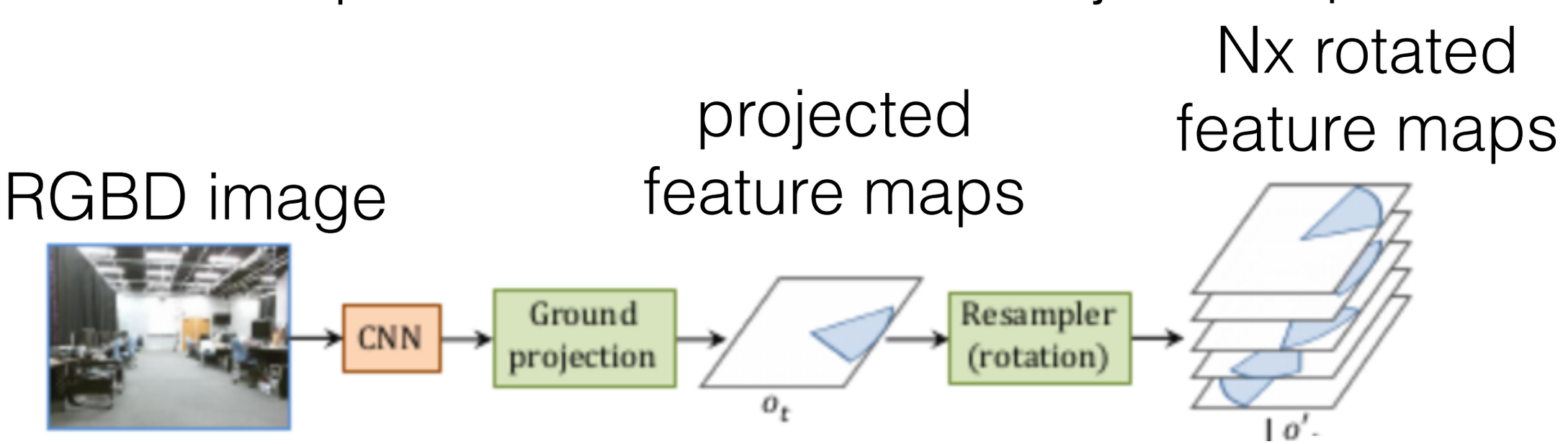
RGBD image

projected
feature maps



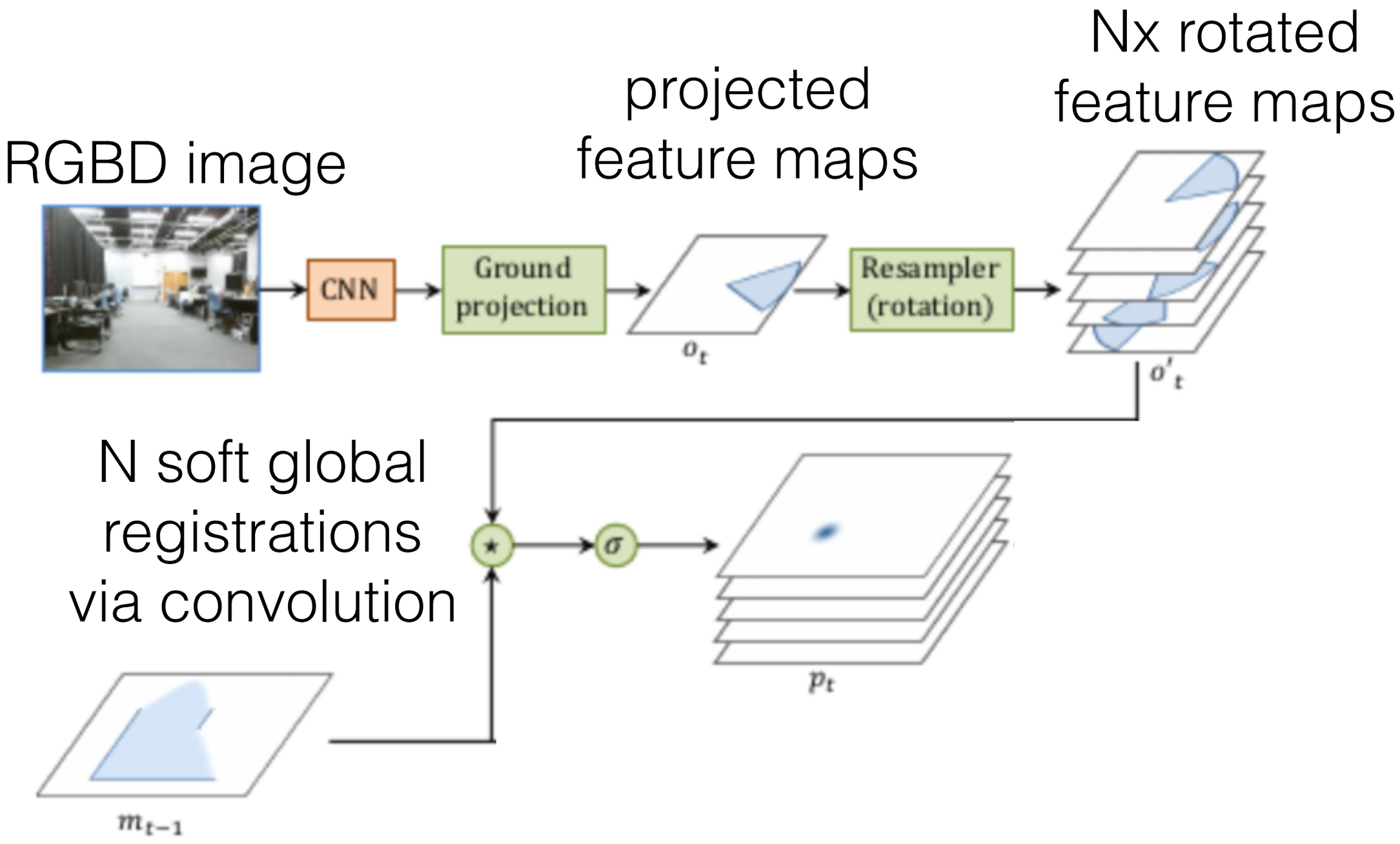
MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/



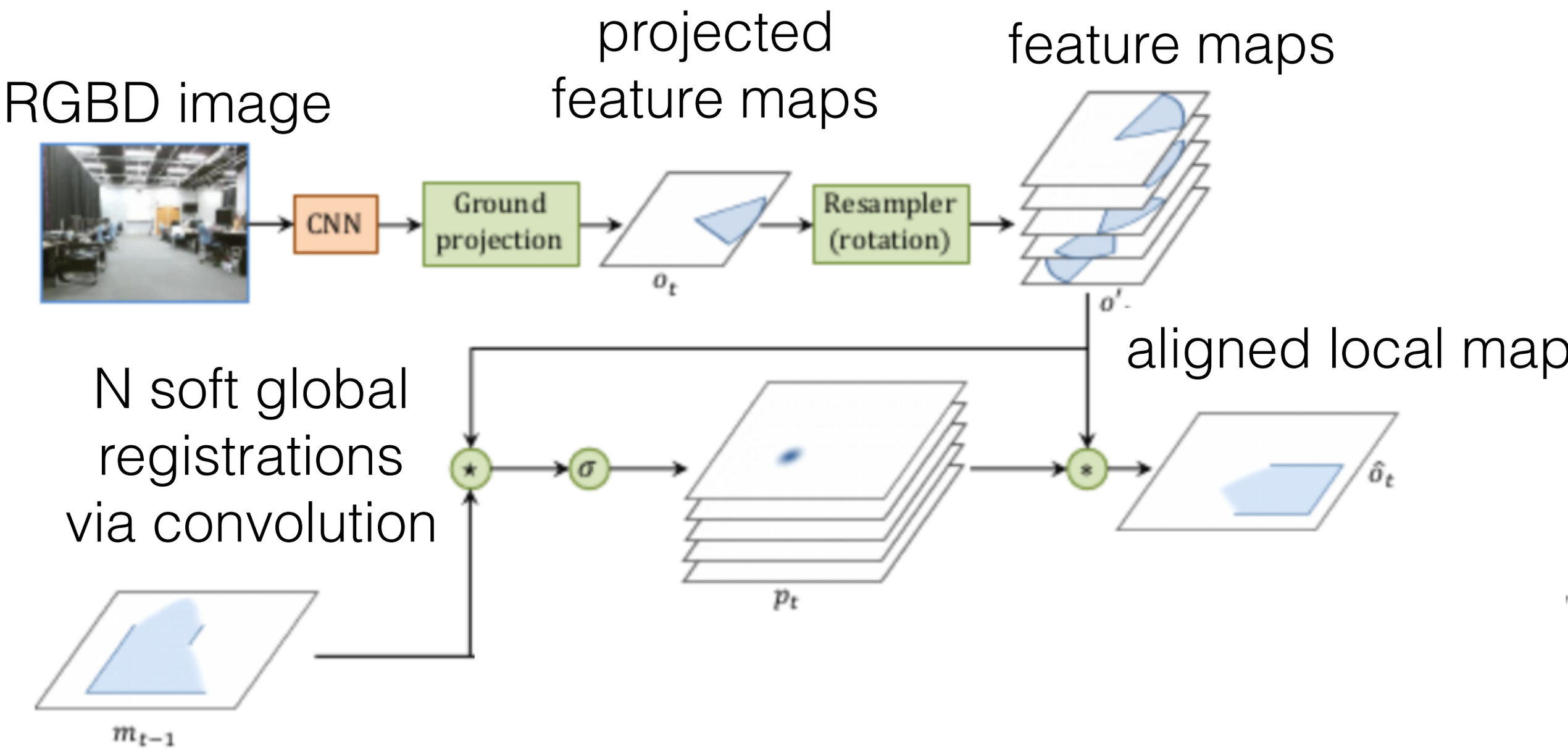
MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/



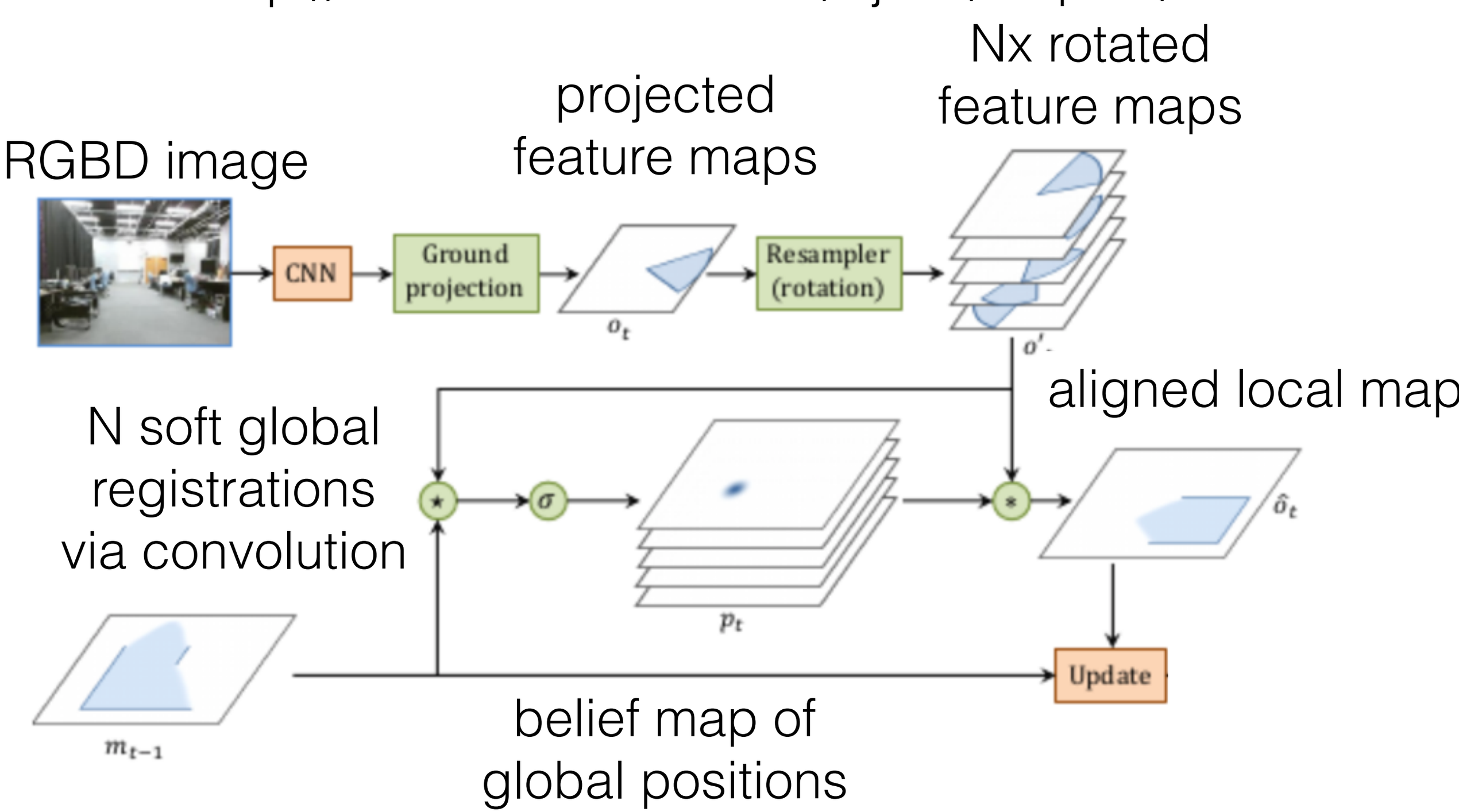
MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/



MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/



MapNet [Henriques CVPR 2018]

http://www.robots.ox.ac.uk/~joao/mapnet/

RGBD image

projected feature maps

Nx rotated feature maps

CNN

Ground projection

o_t

Resampler (rotation)

o'_t

aligned local map

N soft global registrations via convolution

$\star$

σ

p_t

$\times$

$\hat{o}_t$

m_{t-1}

belief map of global positions

Update

m_t

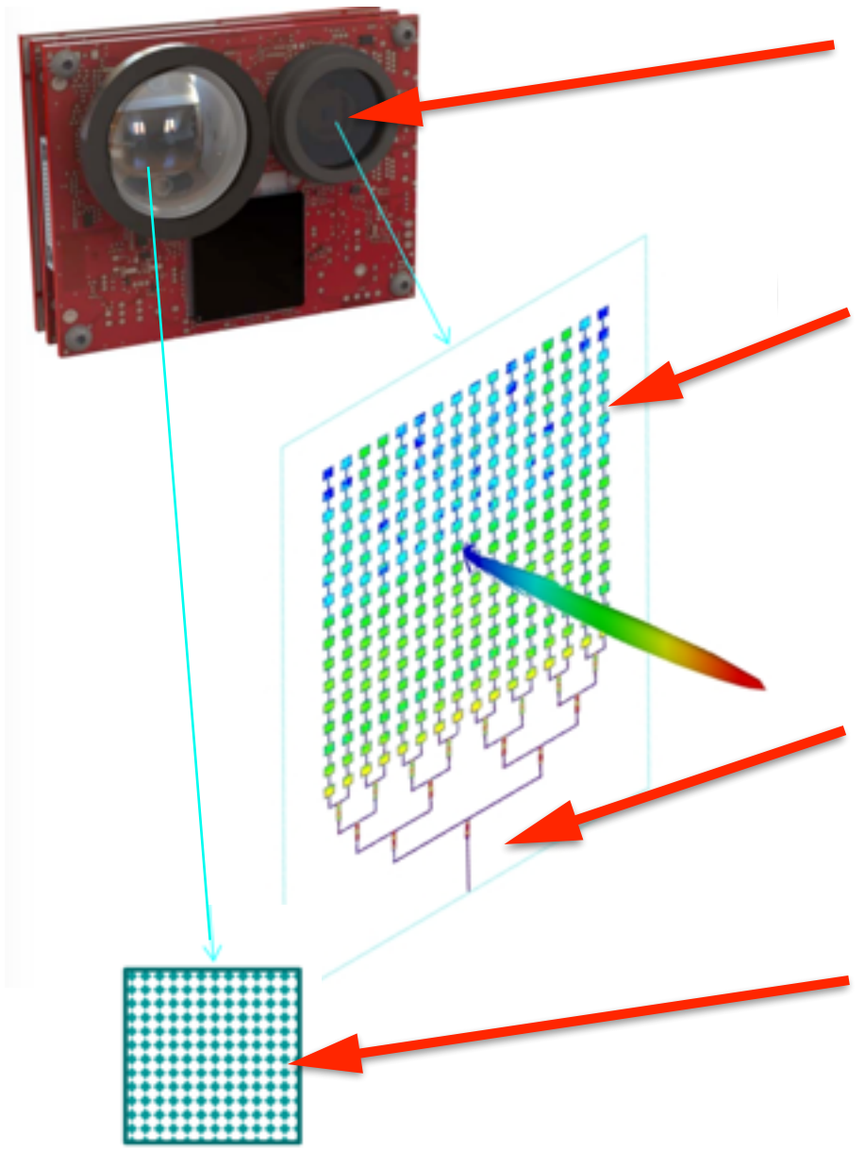
merged global map

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

Active mapping [Zimmermann, Petricek et al. ICCV 2017]
http://cmp.felk.cvut.cz/~zimmerk/

Lidar with independent steering of depth-measuring rays

QUANERGY S3 principle



Emitted laser beams

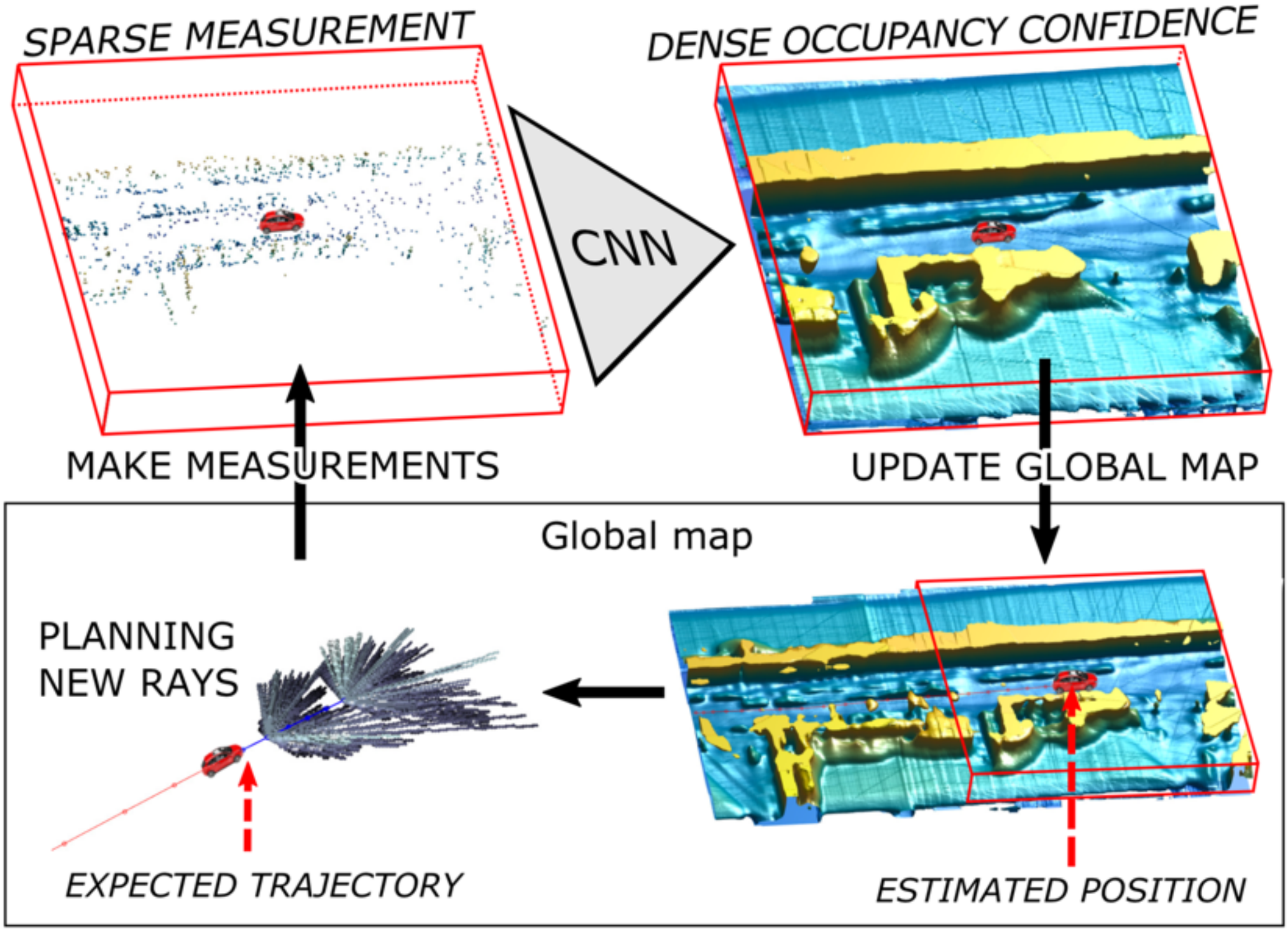
Transmitted through Optical Phased Array

Controlling optical properties of OPA elements, allows to steer laser beams in desired directions

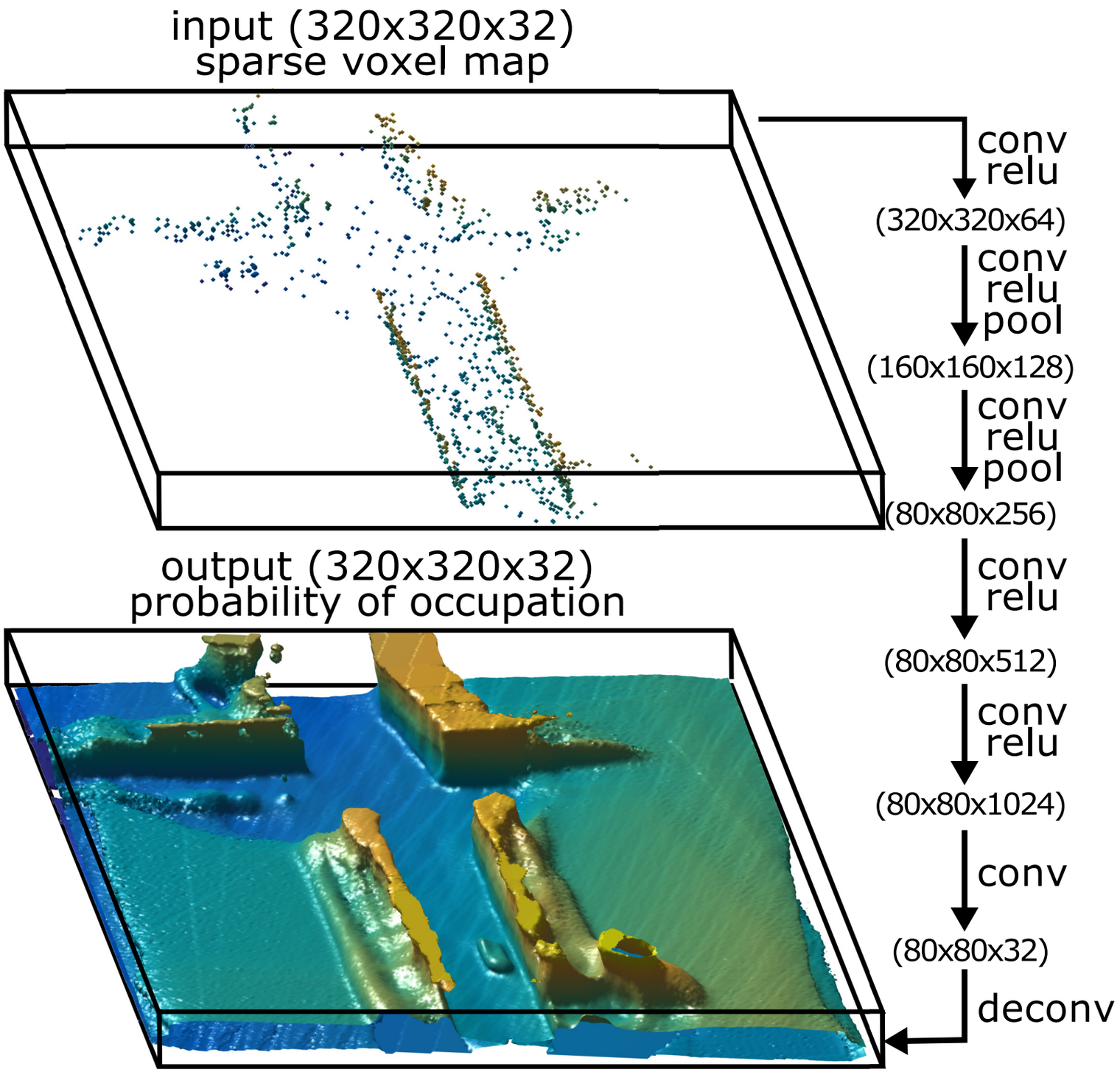
Reflected laser beams are captured by SPAD array

Images of S3 Lidar redistributed with permission of Quanergy Systems (http://quanergy.com)

Active mapping [Zimmermann, Petricek et al. ICCV 2017]
http://cmp.felk.cvut.cz/~zimmerk/



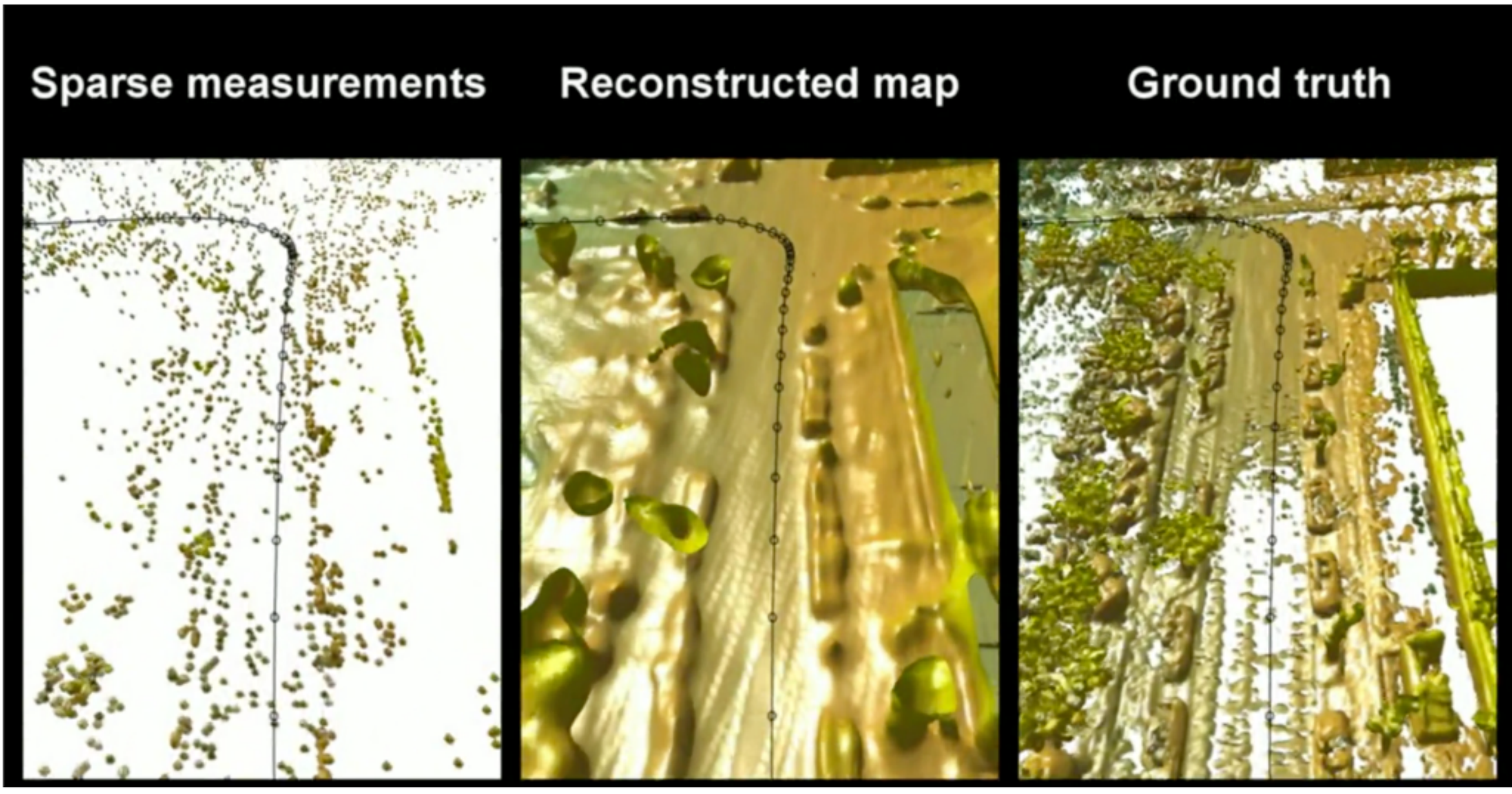
Structure of 3D mapping network



$$\theta \in \mathcal{R}^{20\mathrm{M}}$$

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

Active mapping [Zimmermann, Petricek et al. ICCV 2017]



[1] Zimmermann, Petricek, Salansky, Svoboda, Learning for Active 3D Mapping, **ICCV oral**, 2017 https://arxiv.org/abs/1708.02074

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics