LFIR2Pose : Pose Estimation from an Extremely Low-Resolution FIR Image Sequence

Saki Iwata†, Yasutomo Kawanishi†, Daisuke Deguchi†, Ichiro Ide†, Hiroshi Murase†, Tomoyoshi Aizawa††
†Nagoya University, ††OMRON Corporation Email: iwatas@murase.is.i.nagoya-u.ac.jp

ICPR 2020 25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION Milan, Italy 10 | 15 January 2021

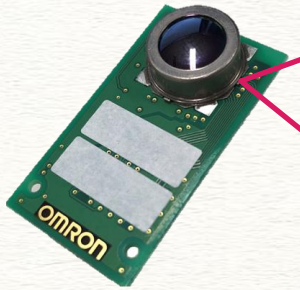
Background and Purpose

Monitoring system for elderly without privacy concerns

LFIR image

Pose

Far-InfraRed (FIR) sensor array

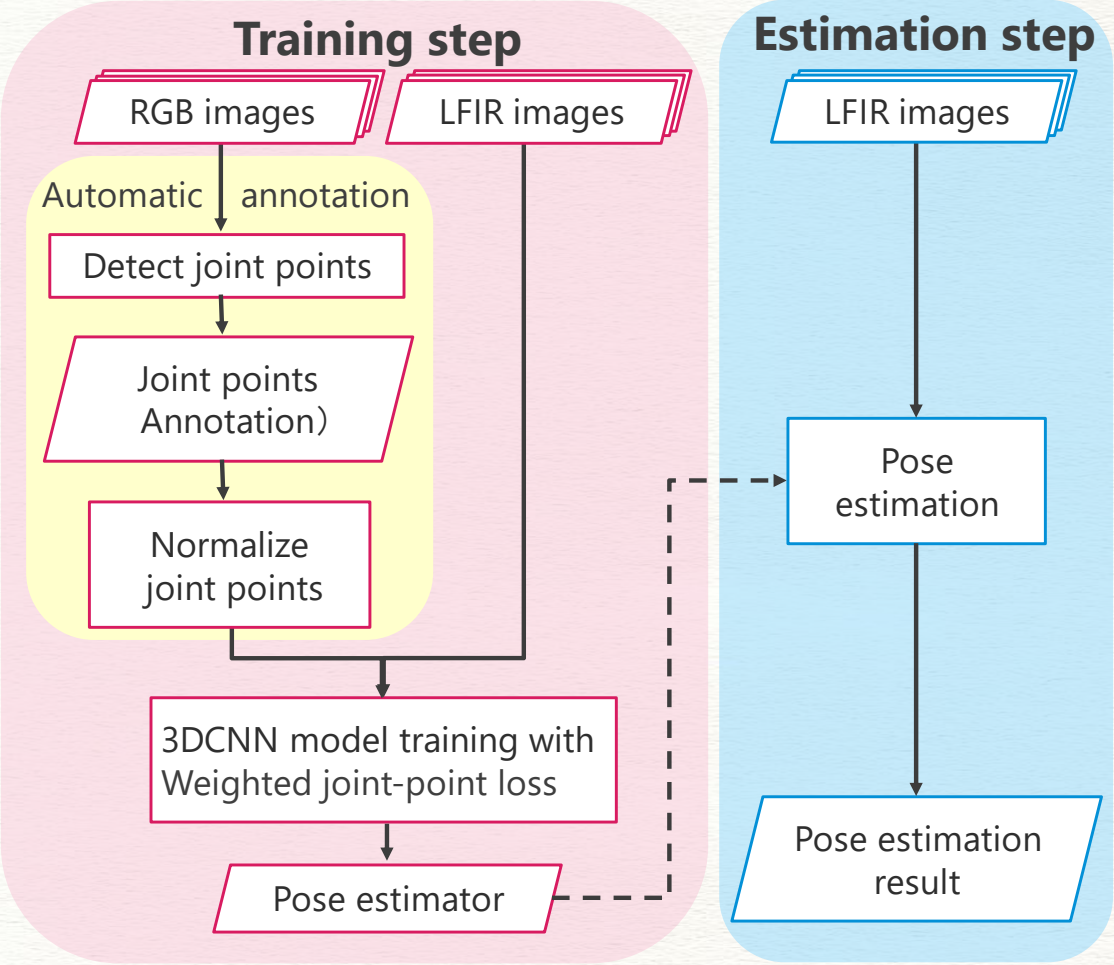


Outputs **temperature distribution** in a specific area **(256 =16 × 16 pixels)**

☺ No privacy concerns
Works even in the dark

☹ Difficult to compute rich features

Proposed method : LFIR2Pose



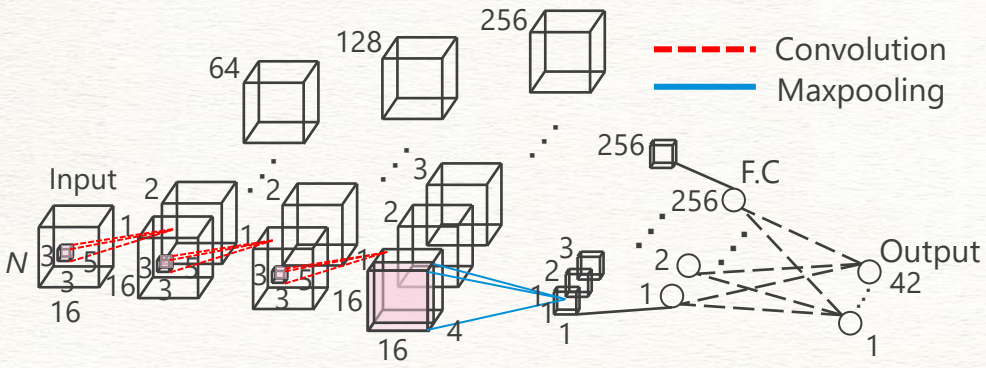
LFIR2Pose : **Pose** Estimation from an Extremely **L**ow-Resolution **FIR** Image Sequence

Saki Iwata†, Yasutomo Kawanishi†, Daisuke Deguchi†, Ichiro Ide†, Hiroshi Murase†, Tomoyoshi Aizawa††
†Nagoya University, ††OMRON Corporation Email: iwatas@murase.is.i.nagoya-u.ac.jp

Proposed method : LFIR2Pose

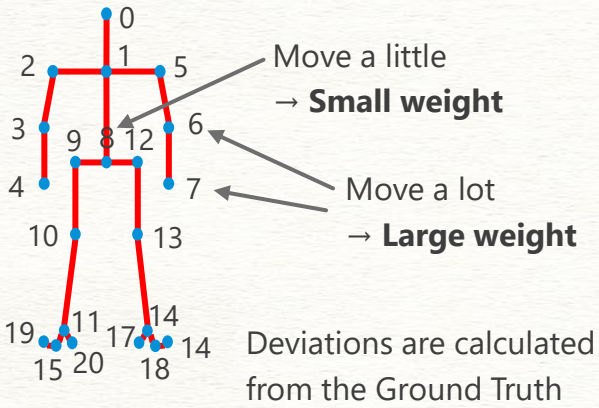
Contribution 1 : LFIR2Pose model

① 3DCNN - Utilize temporal information



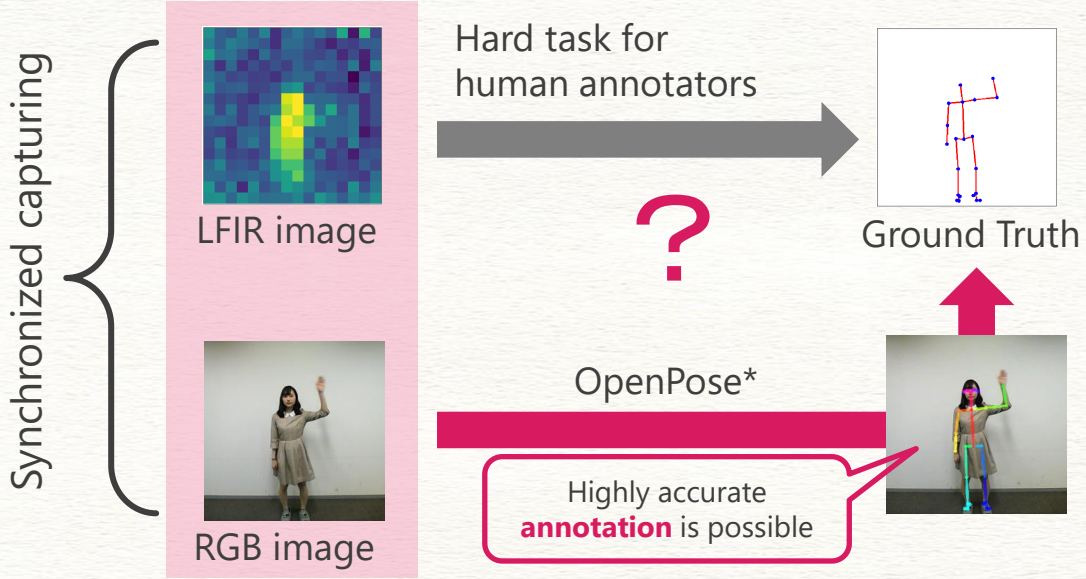
② Weighted joint point loss

- Weighting each joint point depending on the standard deviations of their movements



Contribution 2 : Automatic annotation

Capturing both **RGB** and **LFIR images**, and annotating the ground truth for each LFIR image by applying OpenPose* to the corresponding RGB image



* Z. Cao et al., "Realtime multi-person 2D pose estimation using part affinity fields", CVPR 2017.

LFIR2Pose : Pose Estimation from an Extremely Low-Resolution FIR Image Sequence

Saki Iwata†, Yasutomo Kawanishi†, Daisuke Deguchi†, Ichiro Ide†, Hiroshi Murase†, Tomoyoshi Aizawa††
†Nagoya University, ††OMRON Corporation Email: iwatas@murase.is.i.nagoya-u.ac.jp

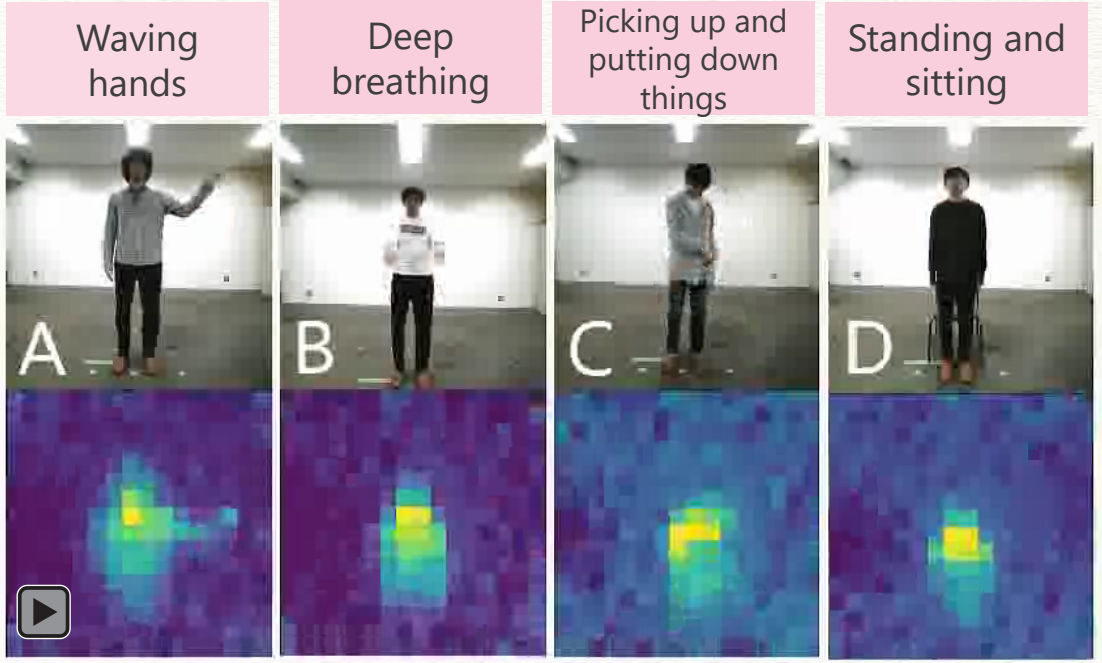
Experiment

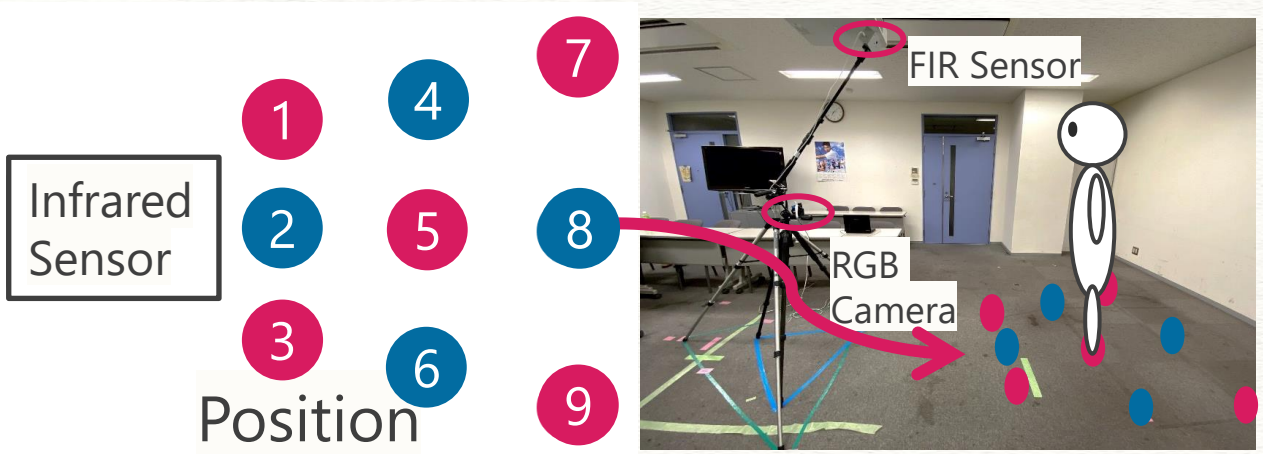
Dataset

The number of actions	4
The positions of a subject	9
The number of frames	About 170 / Video
The number of subject	11
The number of videos	99 / Action
Training data	Positions 1,3,5,7,9
Test data	Positions 2,4,6,8

Method

Comparative mothod 1 : 2DCNN
Comparative mothod 2 : 2DCNN (global max pooling)
Proposed mothod 1 : 3DCNN
Proposed mothod 2 : 3DCNN + Weighted joint point loss



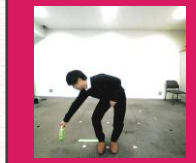


LFIR2Pose : Pose Estimation from an Extremely Low-Resolution FIR Image Sequence

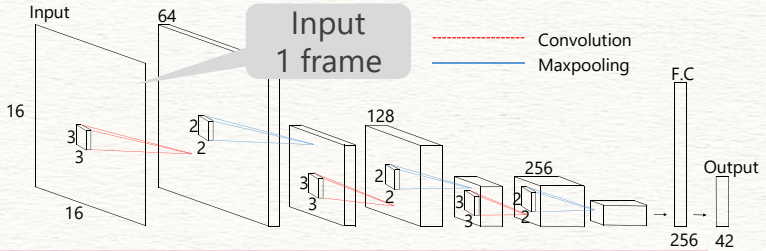
Saki Iwata†, Yasutomo Kawanishi†, Daisuke Deguchi†, Ichiro Ide†, Hiroshi Murase†, Tomoyoshi Aizawa††
†Nagoya University, ††OMRON Corporation Email: iwatas@murase.is.i.nagoya-u.ac.jp

ICPR 2020 25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION Milan, Italy 10 | 15 January 2021

RMSE of the Ground Truth and the estimation results ($\times 10^{-2}$)

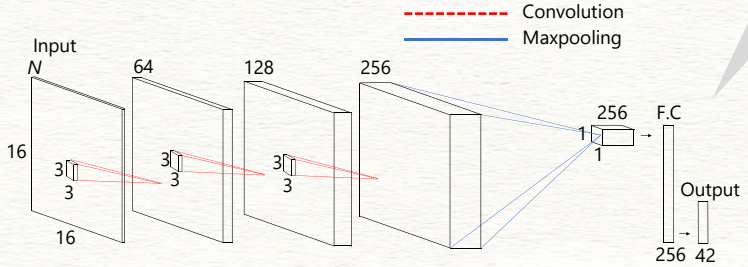
Position		2	4	6	8	Average
Action A 	Comp1	4.33	4.85	5.35	5.11	4.91
	Comp2	4.63	4.70	5.04	5.22	4.90
	Prop1	2.88	3.46	3.45	3.63	3.36
	Prop2	2.65	3.12	3.32	3.68	3.19
Action B 	Comp1	9.65	10.4	8.28	12.6	10.2
	Comp2	7.10	5.62	5.52	7.19	6.36
	Prop1	4.97	4.35	4.32	4.85	4.59
	Prop2	4.53	4.29	4.10	4.59	4.38
Action C 	Comp1	11.2	10.0	16.8	10.9	12.2
	Comp2	7.35	7.32	7.24	7.31	7.30
	Prop1	5.13	5.12	5.07	4.99	5.07
	Prop2	4.91	4..96	5.04	4.96	4.97
Action D 	Comp1	8.32	7.25	6.57	7.19	7.33
	Comp2	5.53	5.67	5.01	6.34	5.64
	Prop1	3.82	4.34	4.08	5.29	4.38
	Prop2	3.73	4.15	4.06	5.40	4.33

Comparative method 1 network

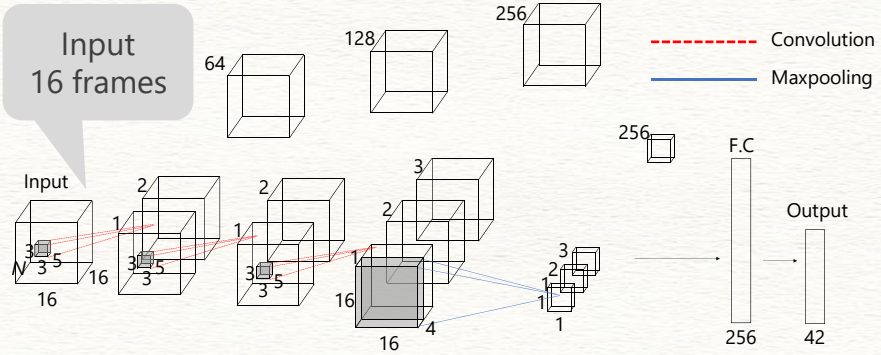


Comparative method 2 network

Use global max pooling



Proposed method network

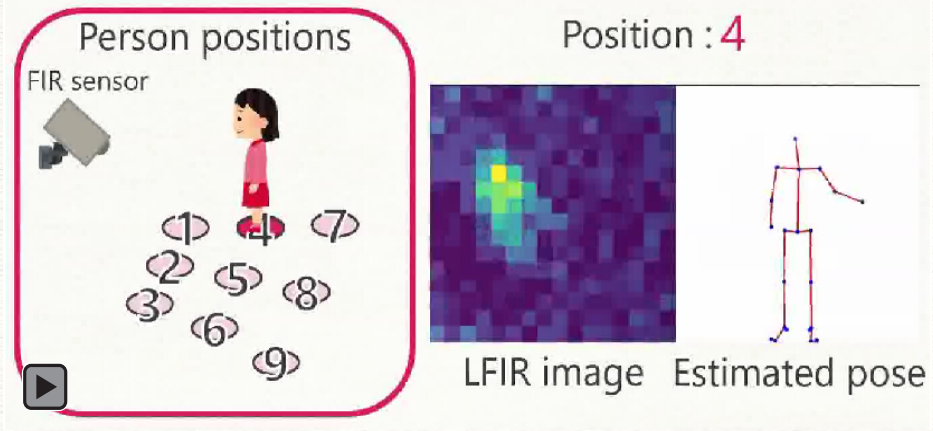


LFIR2Pose : Pose Estimation from an Extremely Low-Resolution FIR Image Sequence

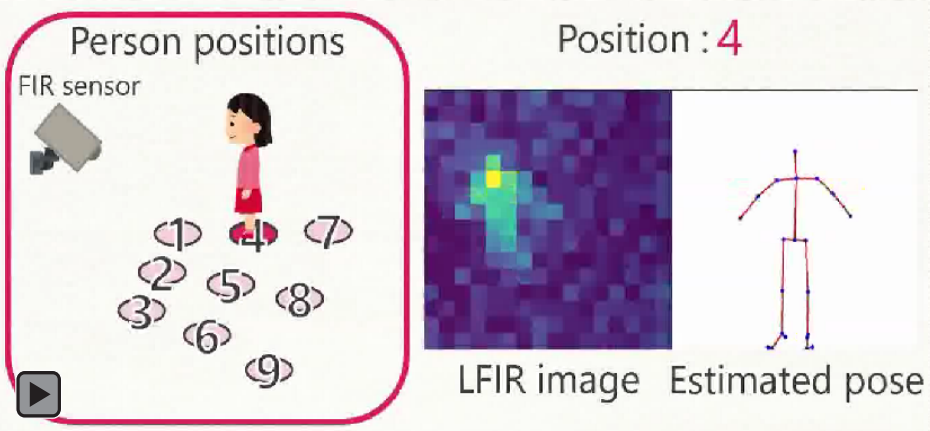
Saki Iwata†, Yasutomo Kawanishi†, Daisuke Deguchi†, Ichiro Ide†, Hiroshi Murase†, Tomoyoshi Aizawa††
†Nagoya University, ††OMRON Corporation Email: iwatas@murase.is.i.nagoya-u.ac.jp

Estimation results at various person positions

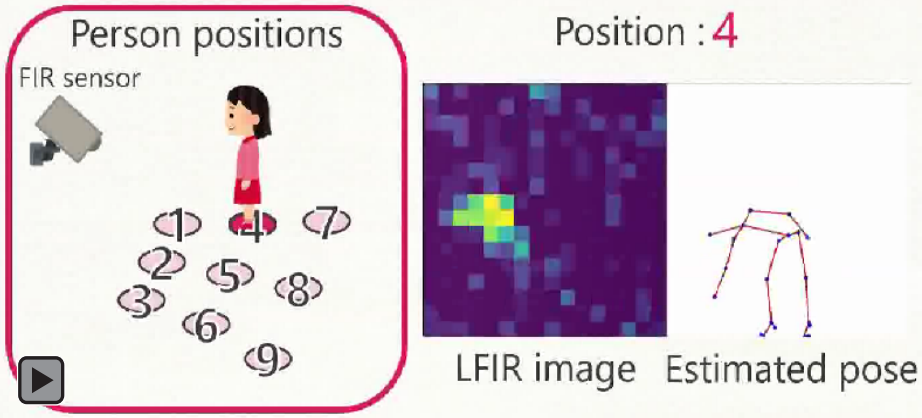
Action A Waving hands



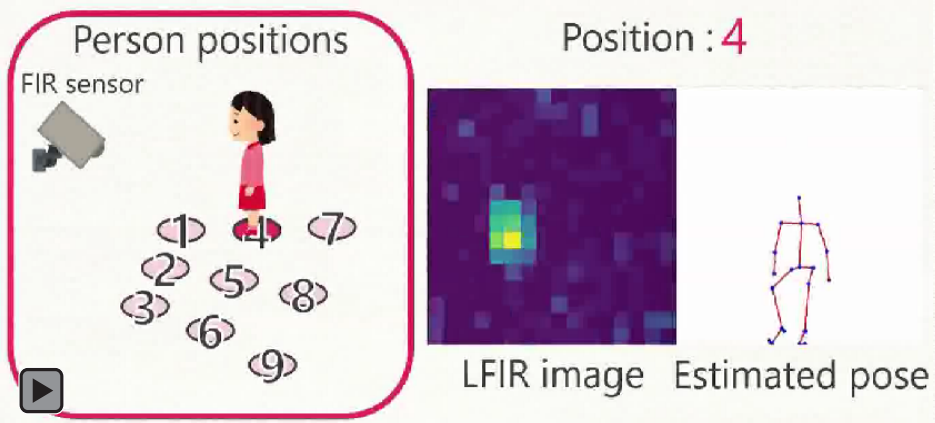
Action B Deep breathing



Action C Picking up and putting down things



Action D Standing and sitting



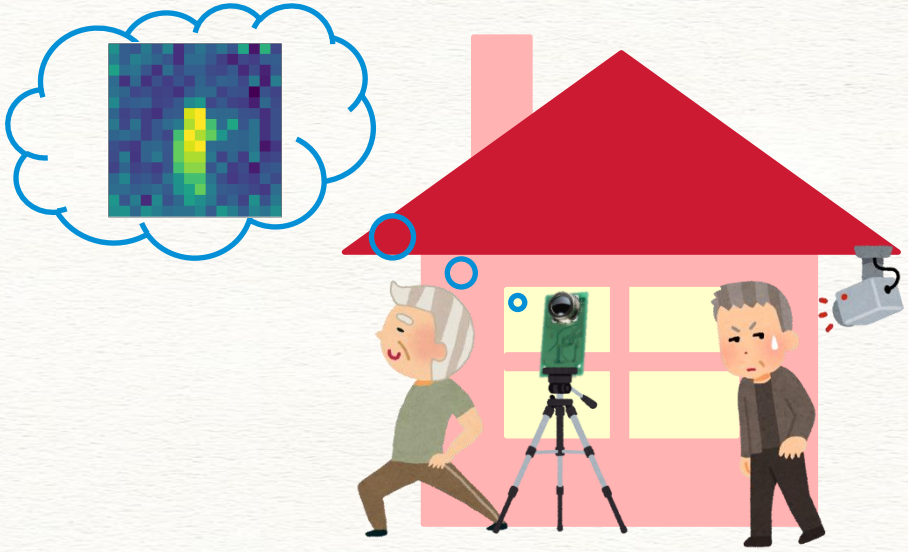
Background and Purpose

◆Aging society

Increase in the number of elderly living alone

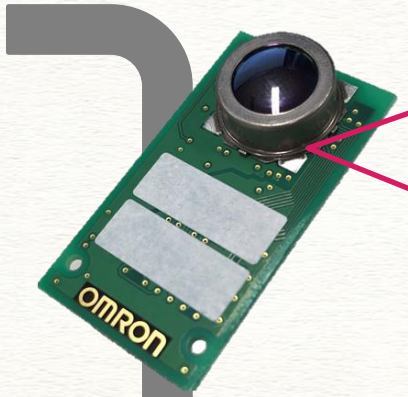
– Systems are needed to
  - Keep them healthy and safe
  - Be free from privacy concerns

  → **Use FIR sensor**



LFIR2Pose

Actual scene



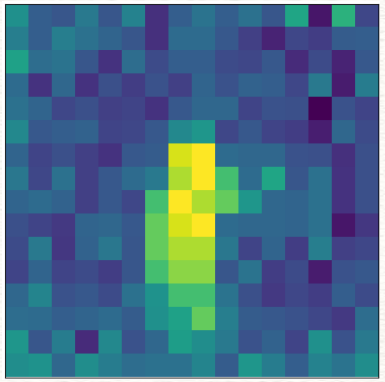
FIR sensor

Outputs **temperature distribution** in a specific area **(256 =16 × 16 pixels)**

☺ No privacy concerns
Works even in the dark

☹ Difficult to compute rich features

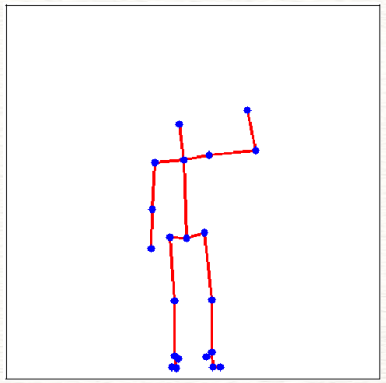
16



16

LFIR image

LFIR2Pose



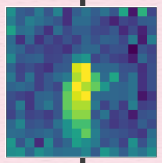
Pose

Process flow of the proposed method

Training step

RGB images

LFIR images



Automatic annotation

Detect joint points

Joint points Annotation)

Normalize joint points

3DCNN model training with Weighted joint-point loss

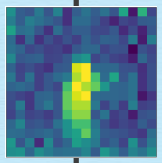
Pose estimator

Contribution 2

Contribution 1

Estimation step

LFIR images

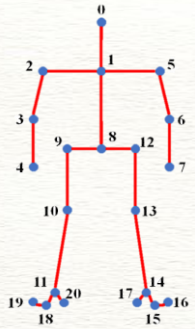


Pose estimation

Pose estimation result

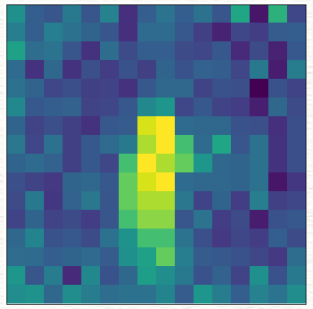
Estimated pose

x, y coordinates for 21 joint points



LFIR2Pose

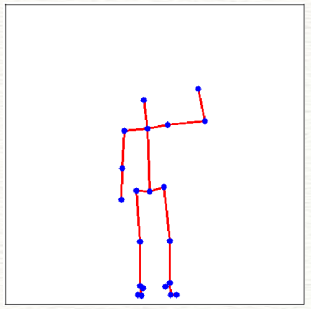
◆Contribution 1: LFIR2Pose model



LFIR image

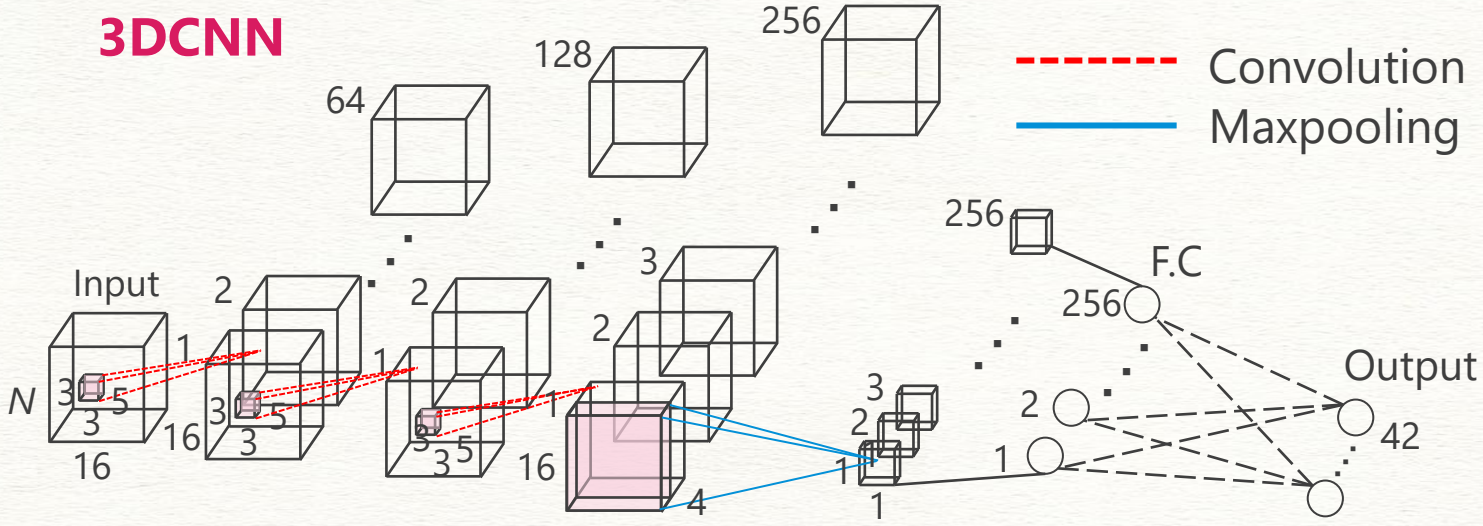
LFIR2Pose

- 3DCNN
- Weighed joint-point loss

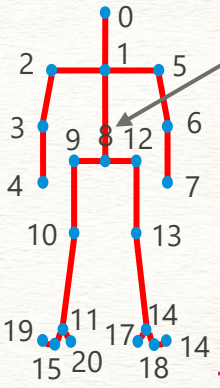


Pose

3DCNN



Weighted joint-point loss



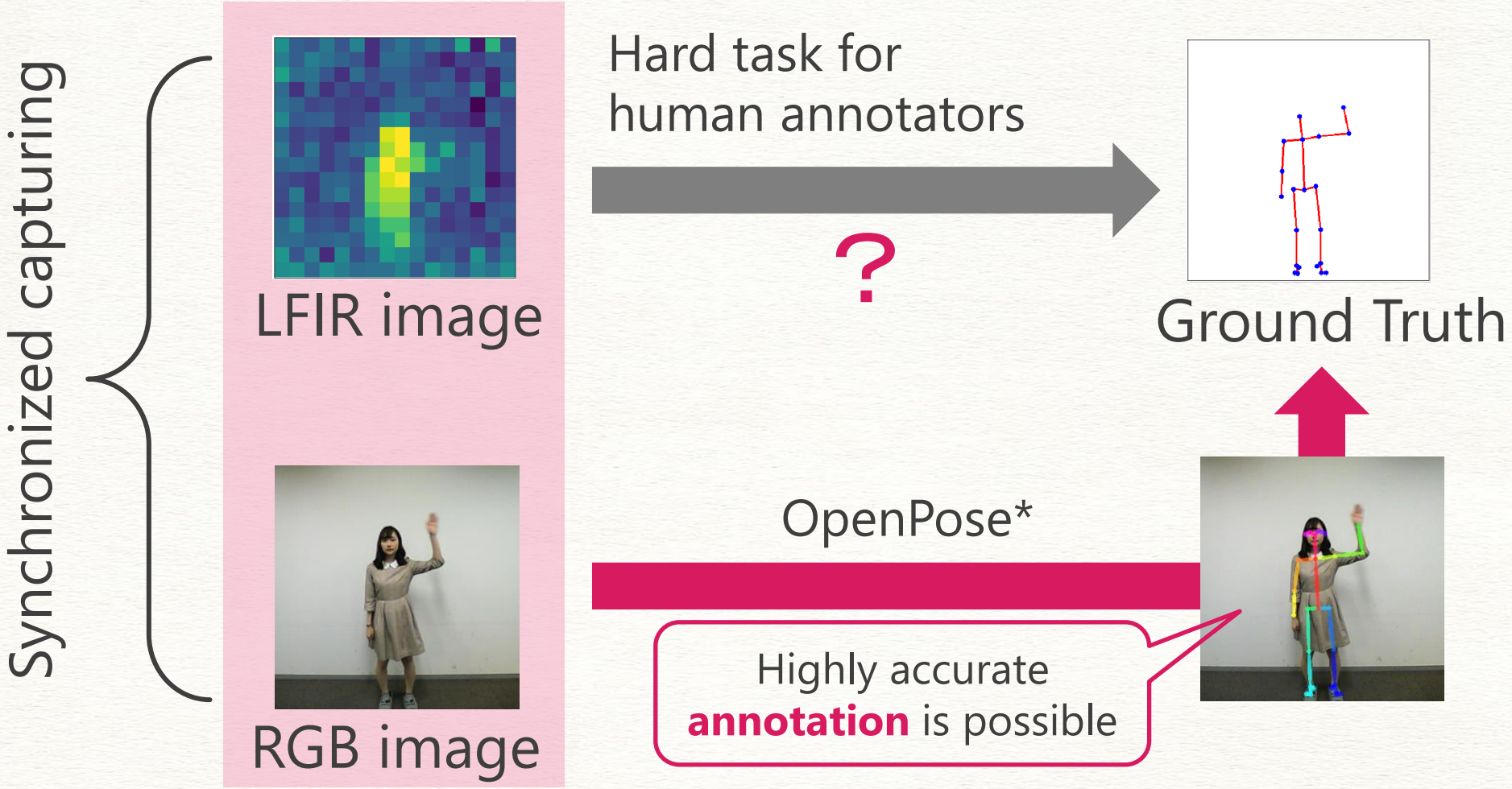
Move a little → **Small weight**

Move a lot → **Large weight**

Deviations are calculated from the Ground Truth

LFIR2Pose

◆Contribution 2: Automatic annotation



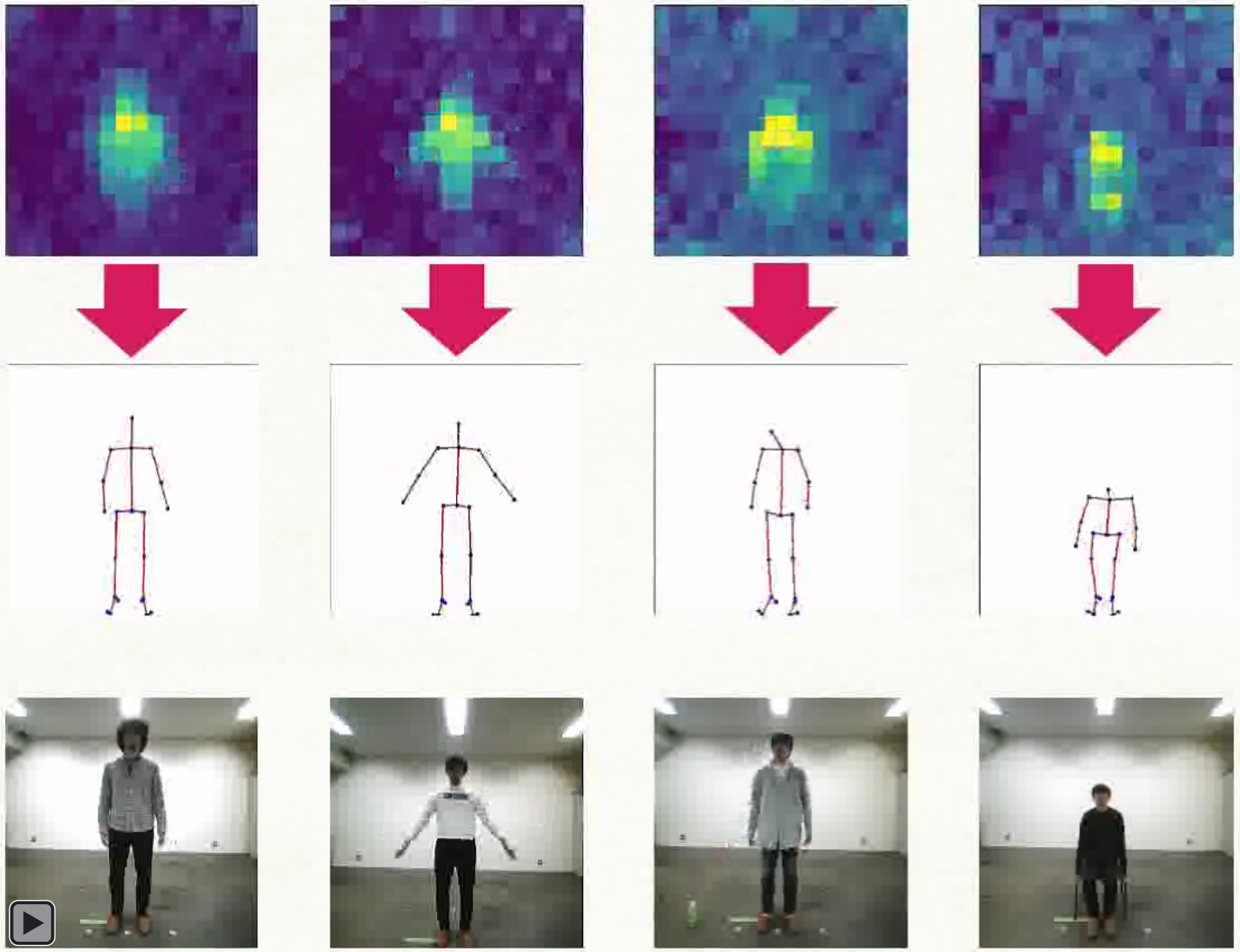
* Z. Cao et al., "Realtime multi-person 2D pose estimation using part affinity fields", CVPR 2017.

Estimation result examples

LFIR image

Estimated pose

Actual human motion



Conclusion

- **LFIR2Pose** : Estimate human pose from LFIR images
  - Model
    - 3DCNN and weighted joint point loss function
  - Automatic annotation
    - Synchronized capturing
- Future work
  - Applications to more diverse pose estimation
    - Estimation of 3D human pose
    - Quantifying the amount of activity from poses