Effects of Markers in Training Datasets on the Accuracy of 6D Pose Estimation

Janis Rosskamp, Rene Weller, Gabriel Zachmann
University of Bremen, Faculty of Computer Science, Germany

Introduction

Background:

- 6D pose estimation requires labelled training data which is usually task specific
- Data labelling is tedious and possibly imprecise

Solution: We use optical motion capturing for easy labelling of poses

Research Questions:

1. Do markers in training data reduce accuracy of pose estimation?
2. When is high-precision marker labelling superior to imperfect manual labeling?

Method

Identical synthetic datasets (Figure 1) for a comparison without any additional influences, i.e. scene differences

- Our teabox dataset has 20k images in 800 scenes
- YCB-V dataset with 80k images for five objects.
- Pose Estimator: GDR-Net with standard parameters

Figure 1: Images of the same scene in our four training sets for the teabox dataset

UBremen Research Alliance / AI Center for Health Care

Universität Bremen

Code available!

Do markers in training data affect pose estimation?

	Dataset	AUC of ADD-S	AUC of ADD-(S)	ADD-(S)
I	w/o marker	97.1	90.0	84.8
II	w marker	96.2	87.8	76.6
III	inpainting	97.0	90.0	84.8
IV	marker aug.	97.0	89.8	84.2
V	aug. + marker	96.7	88.8	80.0

Standard Metrics

Table 1: Evaluation on a markerless test set with GDR-Net as pose estimator. The pose estimators were trained using five datasets to compare the influence of markers.

Evaluation on YCB-V Dataset

	GDR-Net				ZebraPose			
	I		III		I		III	
Evaluation on	syn	real	syn	real	syn	real	syn	real
002_master_chef_can	81.1	92.1	80.6	94.3	94.8	90.0	94.5	87.4
004_sugar_box	79.6	96.8	80.0	96.1	94.1	88.6	93.0	88.6
008_pudding_box	79.2	89.1	80.0	89.1	93.4	41.1	92.8	36.9
025_mug	78.4	65.3	84.0	52.1	93.7	73.3	93.8	61.5
036_wood_block*	81.4	17.7	81.8	4.3	88.4	11.9	87.5	13.6
Average	79.9	92.6	**81.3**	**93.2**	**92.9**	60.9	92.3	**61.8**

Table 2: Evaluation on real and synthetic test data. The pose estimators were trained on markerless images (I) and images where markers were removed with inpainting (III).

Marker Impact on Real Data

		AUC of ADD-S	AUC of ADD-(S)	ADD-(S)
I	w/o marker	93.6	79.0	45.8
III	inpainting	93.2	79.1	47.5
VIII	w/o finetuning	88.8	71.9	20.3

Table 3: The pose estimator was trained on real and synthetic images. Evaluation was done on real images.

- Training on a mixture of both real and synthetic images

High precision labelling vs. imperfect labelling

- We manually label synthetic data with known ground truth to estimate labeling errors

Δt	0-2 mm	2-4 mm	4-6 mm	6+ mm
N	79	20	9	2
$\Delta\sigma$	0-2 °	2-4 °	4-6 °	6+ °
N	76	12	14	10

Table 4: Distribution of labelling errors compared to the exact ground truth

- For small labelling errors, pose estimation with a marker training set (black line) is preferable.

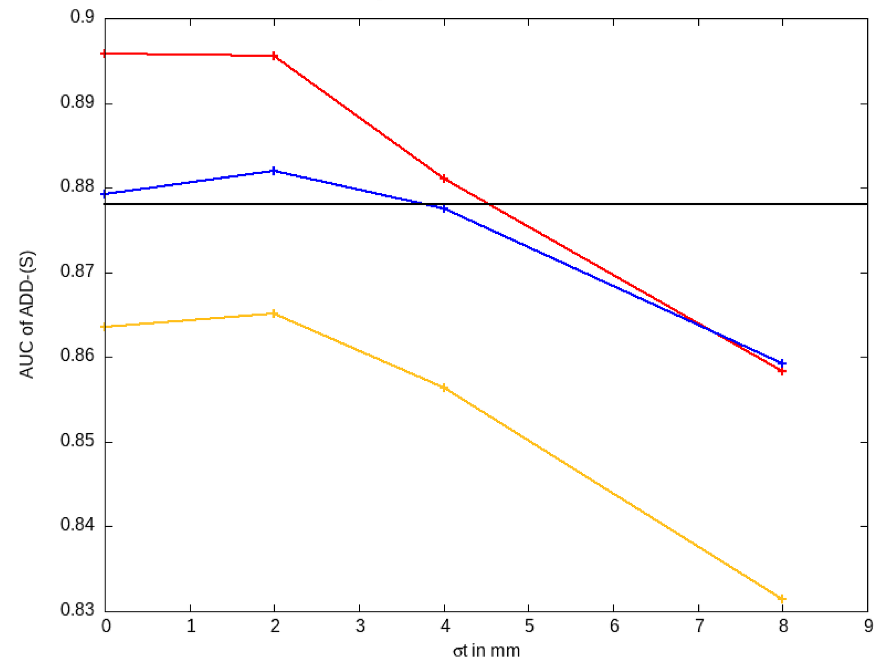

Figure 2: Accuracy of pose estimation for different labelling errors. The labelling errors were modelled using normal distributions. The black line shows the worst estimation accuracy (method II) when markers are used in the training set.

Conclusion

1. **Markers on objects reduce the accuracy of pose estimation.**
2. **With our method, we achieve the same accuracy of pose estimation compared to markerless images**
3. **We can automate labelling 6D poses with high precision**
4. **This results in better pose estimation compared to imprecise manual labelling**
5. **With our method, training data can be labelled easily and even from objects in motion**

Acknowledgements

The research in this paper was supported by the U Bremen Research Alliance/AI Center for Health Care, which is financially supported by the Federal State of Bremen.