Depth Estimation using Optical Flow and CNN for the NAO Robot

Oswualdo Alquisiris-Quecha[1], Jose Martinez-Carranza[2]

{[1]oswualdoaq, [2]carranza}@inaoep.mx

Department of Computer Science, Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México

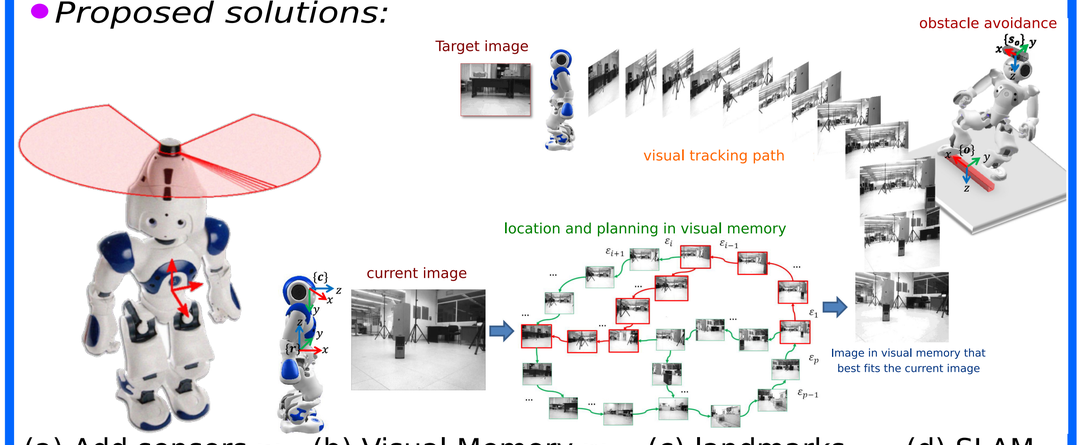
MICAI Mexican International Conference on Artificial Intelligence 2019

Abstract

In this work, we focus on the problem of depth estimation in a single image for the NAO robot. For the depth estimation, we argue that the erratic movement exhibited by the walking motion of the robot could be exploited to obtain optical flow vectors, which are strongly related to depth observed by the NAO's camera. Thus, we present a real-time system based on a CNN architecture that uses optical flow as input channels in order to estimate depth. To this aim, we present a new dataset that includes optical flow images associated to depth images for training. Our results indicate that optical flow can be exploited in humanoid robots such as NAO, but we are confident that it could be used in other platforms with erratic motion.

Introduction

- *One of the main challenges in the field of robotics is the autonomous navigation with a single camera.*
  - *Erratic movements of moving robots.*
  - *It is not possible to obtain clear and noise-free images.*
- *Proposed solutions:*

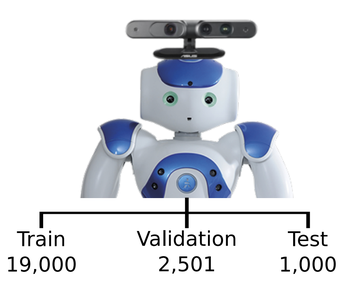


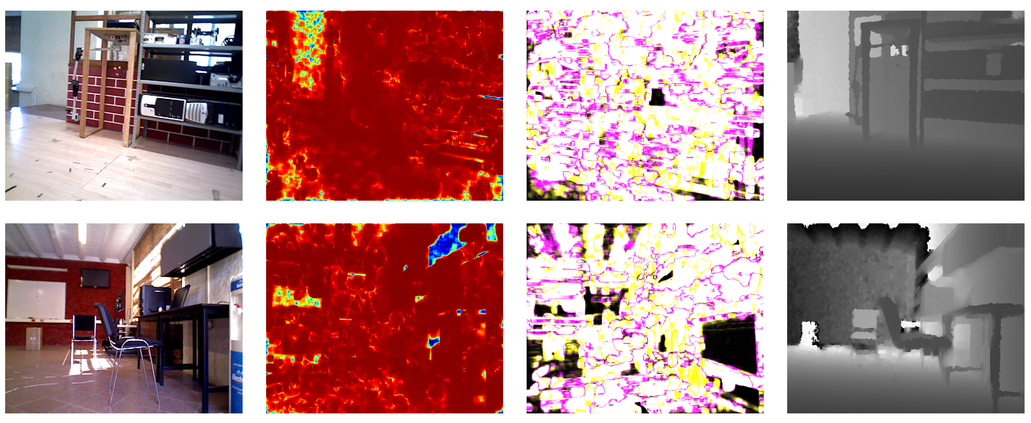
(a) Add sensors [1] (b) Visual Memory [2] (c) landmarks (d) SLAM

Figure 1: navigation methods with the NAO robot

Dataset

The dataset consists of 22,501 elements:





RGB Image Color map Augmented image Depth map

Figure 2:proposed dataset structure

Methodology

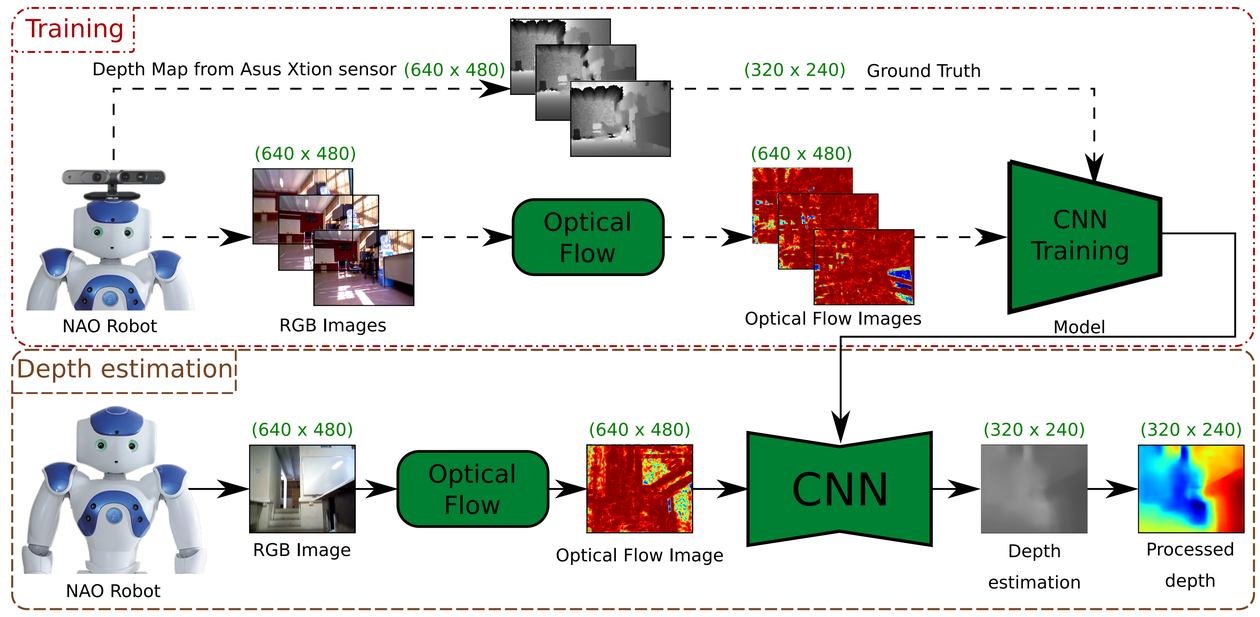


Figure 3: Process of training and estimating the depth of the system.

EXPERIMENTS AND RESULTS

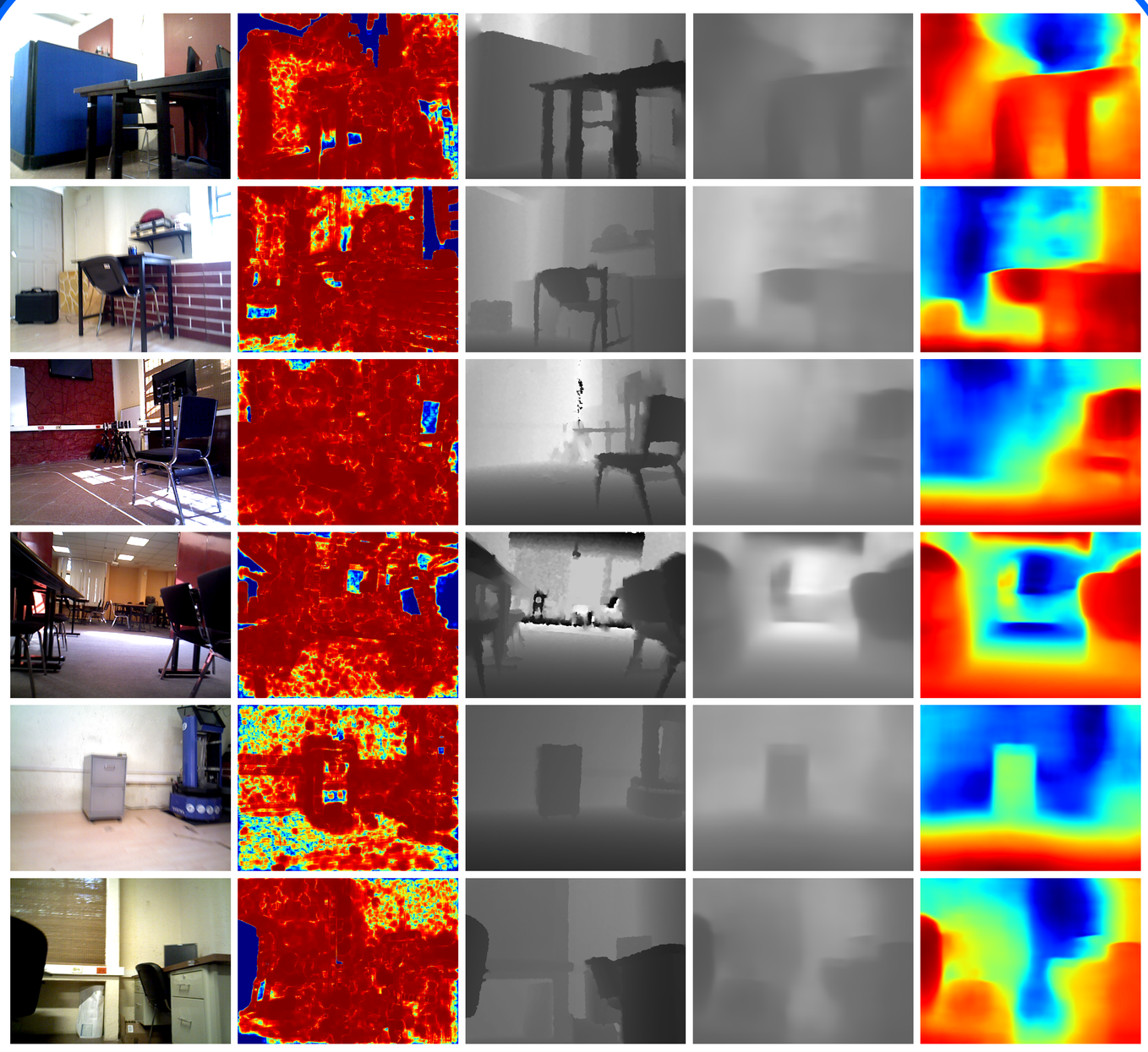


Figure 4: Qualitative results from proposed method

Table 1: Comparison with other methods of depth estimation

Method	$\delta_1 \uparrow$	$\delta_2 \uparrow$	$\delta_3 \uparrow$	REL $\downarrow$	RMS $\downarrow$	$Log_{10} \downarrow$
Eigen et al. [3]	0.769	0.950	0.988	0.158	0.641	–
Laina et al. [4]	0.811	0.953	0.988	0.127	0.573	0.055
Alhashim et al. [5]	0.846	0.974	0.994	0.123	0.465	0.053
Ours	0.401	0.600	0.760	0.490	1.080	0.170

CONCLUSIONS

- It is possible to use optical flow techniques with deep learning, for the estimation of depth maps, exploding the erratic movement in a humanoid robot through the optical flow vectors.
- Our results indicate that our approach is feasible and it compares to state of the art methods on depth estimation in a single image.

References

[1] Hornung, A., Wurm, K. M., and Bennewitz, M. *Humanoid robot localization in complex indoor environments*. In IROS (2010), pp. 1690–1695.

[2] Delfin, J., Becerra, H. M., and Arechavaleta, G. *Humanoid navigation using a visual memory with obstacle avoidance*. Robotics and Autonomous Systems (2018).

[3] David Eigen, Christian Puhrsch and Rob Fergus. *Depth map prediction froma a single image using a multi-scale deep network*. In Advances in neural information processing systems, pages 2366-2374, 2014.

[4] Iro Laina, Christian Rupprecht, Vasileios Belagiannis, Federico Tombari and Nassir Navad. *Deeper depth prediction with fully convolutional resual networks*. In 2016 Fourth international conference on 3D vision (3DV), pages 239-248. IEEE, 2016.

[5] Ibraheem ALhashim and Peter Wonnka. High quality monocular depth estimation via transfer learning. arXiv preprint arXiv:1812.11941, 2018.