



HAL
open science

PhD Forum: GPU-based Visual Odometry for Autonomous Vehicle Applications

J. A. de Jesús Osuna-Coutiño, Abiel Aguilar-González, Miguel Arias-Estrada

► **To cite this version:**

J. A. de Jesús Osuna-Coutiño, Abiel Aguilar-González, Miguel Arias-Estrada. PhD Forum: GPU-based Visual Odometry for Autonomous Vehicle Applications. Proceedings of the 11th International Conference on Distributed Smart Cameras, Sep 2017, Stanford, United States. pp.2010-211, 10.1145/3131885.3135083 . hal-01658813

HAL Id: hal-01658813

<https://hal.science/hal-01658813>

Submitted on 7 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PhD Forum: GPU-based Visual Odometry for Autonomous Vehicle Applications

J. A. de Jesús Osuna-Coutiño¹, Abiel-Aguilar-González^{1,2}, Miguel Arias-Estrada¹

1. Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Tonantzintla, Mexico

2. Université Clermont Auvergne (UCA), Institut Pascal, Clermont-Ferrand, France

Contact: abiel@inaoep.mx

ABSTRACT

One popular task under several computer vision applications is camera pose estimation under video sequences. In previous work, several camera pose estimations approaches have been developed and several algorithms have been proposed. Unfortunately, most previous formulations iterative behavior and depend on nonlinear optimizations applied over some geometric constraint and this limits the performance under embedded applications. For real-time embedded applications, another approach, more efficient in terms of computational size and processing speed could be reached via hardware-based solutions, for example GPU-based solutions. In this work, we present a GPU-based solution for the camera pose problem. As early formulation we focused our algorithm for an autonomous vehicle application. Preliminary results are encouraging and show the viability of the proposed approach.

Keywords

Camera pose, embedded systems, GPU

1. INTRODUCTION

In several computer vision applications, camera pose estimation across video sequences appears repeatedly under several contexts: visual servoing [2], Simultaneous Localization and Mapping (SLAM) [7], augmented reality [5], etc. In previous work, primary interest was for an accurate/generalized solution for the camera pose problem [4, 6]. Unfortunately, these solutions have iterative behavior and requires relatively high computational resources. Then, several tasks have limitations for embedded applications. In this work, we believe that a smart camera that delivers camera pose at video cadence could be an important contribution for current embedded computer vision applications. In general, several computer vision applications such as SLAM, augmented reality, etc., could reach onboard camera pose processing.

2. CAMERA POSE ESTIMATION BASED ON 2D PIXEL DISPLACEMENTS

As early formulation, we consider an autonomous vehicle scenario. In this case we suppose that all possible camera movements are linear dependent of the x, y pixel displacements whitening the scene. i.e., given the input image divided in four quadrants (as illustrated in **Fig. 1**) and, consider four different cases: turn left, turn right, forward and backward; for the turn left case: all pixels within all the quadrants must have right displacement under the horizontal axis and similar displacement for the vertical axis, see **Fig. 1a**. For the turn right case: all pixels for all the quadrants must have left displacement under the horizontal axis and similar displacement for the vertical axis, as shown in **Fig. 1b**. For the forward case: pixel displacements are to outside the image, see **Fig. 1c**. Finally, for the backward case: pixel displacements are to the image center, as illustrated in **Fig. 1d**.

3. PRELIMINARY RESULTS

As early validation, we implemented our algorithm in a MatLab R2017a code that uses the KITTI dataset [3] as input sequences. First, feature points (Harris corners) are tracked along the video sequence. Then, several 2D pixel displacements (considering four different quadrants, as shown in **Fig. 1**) are correlated with its corresponding camera movement (taken from the ground truth in the KITTI dataset). Finally, unknown camera movements were estimated by comparing 2D pixel displacements in the input image with the previously stored pixel displacements and its corresponding camera movement. In **Fig. 2**, preliminary results are shown. In this case, pixel displacements and camera movements for 800 random frames for the first video sequence from the KITTI dataset were stored, then, camera pose for the third and fourth sequences were computed by comparing the pixel displacements for the input images with the pixel displacements and camera movements previously stored.

4. CONCLUSIONS AND FUTURE WORK

In this work, we have demonstrated that given an autonomous vehicle scenario camera movements are linear dependent with the 2D pixel displacements within a video sequence. As work in progress, we are generalizing our algorithmic approach and implementing it inside the NVIDIA Jetson TX1 board.

5. REFERENCES

- [1] M. Birem and F. Berry. Dreamcam: A modular fpga-based smart camera architecture. *Journal of Systems Architecture*, 60(6):519–527, 2014.
- [2] F. Chaumette, S. Hutchinson, and P. Corke. Visual servoing. In *Springer Handbook of Robotics*, pages 841–866. Springer, 2016.
- [3] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013.
- [4] R. I. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on pattern analysis and machine intelligence*, 19(6):580–593, 1997.
- [5] I. Hong, G. Kim, Y. Kim, D. Kim, B.-G. Nam, and H.-J. Yoo. A 27 mw reconfigurable marker-less logarithmic camera pose estimation engine for mobile augmented reality processor. *IEEE Journal of Solid-State Circuits*, 50(11):2513–2523, 2015.
- [6] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133–135, 1981.
- [7] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015.
- [8] Y. Nakajima and H. Saito. Robust camera pose estimation by viewpoint classification using deep learning. *Computational Visual Media*, 3(2):189–198, 2017.
- [9] A. Rubio Romano. 3d pose estimation using convolutional neural networks. Master’s thesis, Universitat Politècnica de Catalunya, 2015.
- [10] F. Shi and Y. Liu. Estimation of camera pose using 2d to 3d corner correspondence. In *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, volume 2, pages 805–809. IEEE, 2004.

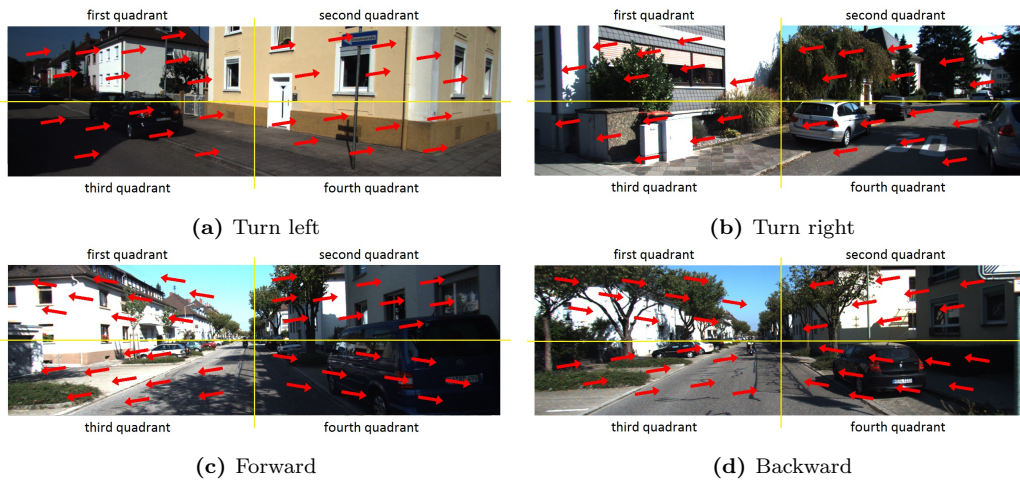


Figure 1: Given an autonomous vehicle scenario. Four different cases: turn left, turn right, forward and backward have to fulfill with all possible movements from a single moving camera installed over the vehicle. In such scenario, pixel displacements inside the images (red arrows) have different directions for each case.

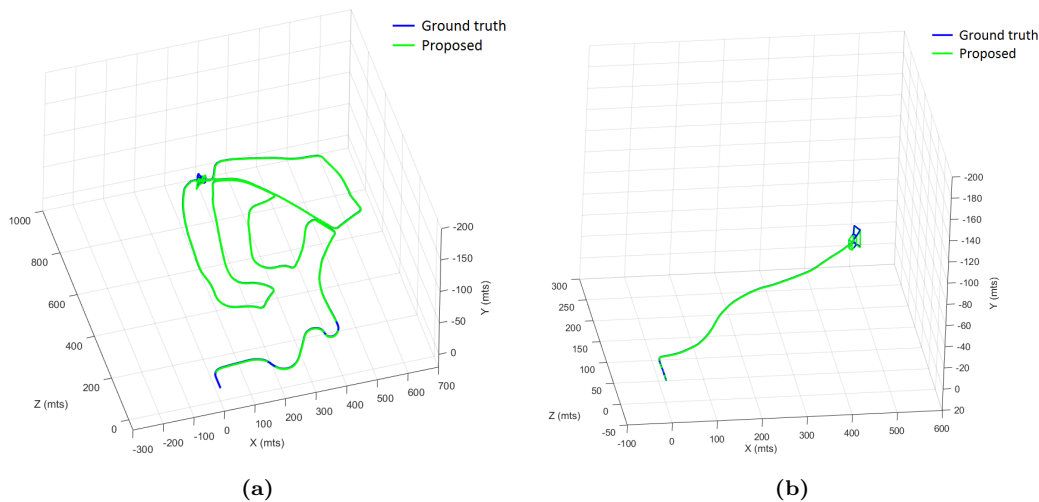


Figure 2: Camera pose estimation for the first two video sequences of the KITTI dataset [3]. In both cases, the proposed algorithm achieves accurate results (similar with the ground truth).