# $\mu$-UAV Based Dynamic Target Tracking for Surveillance and Exploration

Harish Bhaskar*, Jorge Dias, Lakmal Seneviratne, and Mohammed Al-Mualla

Khalifa University of Science, Technology and Research, Abu Dhabi, U.A.E.
harish.bhaskar@kustar.ac.ae
http://www.kustar.ac.ae

**Abstract.** This paper presents an autonomous computer vision system for tracking multiple dynamic ground targets, from images acquired by a camera onboard of a $\mu$-UAV. The method proposed is a self adaptive technique that seamlessly integrates ego-motion compensation with target detection and tracking to provide robust localisation of ground targets. Ego-motion compensation is achieved through establishing homographies using target independent invariant feature descriptors. Targets are then detected using a novel background learning strategy where the optical flow field is fused together with a dynamic background model for accurate foreground extraction. In addition, the paper also reports the use of a Monte Carlo joint probabilistic data association filter for tracking multiple unknown targets. The field tests demonstrate the capabilities of the vision system based on experimental results on images captured by a camera on-board of quadrators ($\mu$-UAV).

## 1 Introduction

The use of small-unmanned air vehicles ($\mu$-UAV) as active explorers in surveillance operations, three-dimensional mapping or detection of ground targets opens new scientific and technological challenges on autonomous robots. These $\mu$-UAV's rely on visual sensing using an on-board camera for control operations during autonomous take-off, landing, stabilisation and navigation [1] mostly exploiting a-priori knowledge of the target, or the environment, differential motion characteristics of the camera to the target through optical flow computation. From a computer vision perspective, such a vision system is complicated by the presence of large displacements of targets, partial or full occlusions, image noise, clutter and blur, perspective changes of the targets and rapid changes in illumination and pose. Several state-of-the-art visual tracking methodologies capable of addressing some of the aforementioned challenges have been proposed in recent years. Many of these methods are target specific that heavily rely on a robust representation that mirrors the appearance of the target accurately enough to be able to recover correspondences over long range of motion [2]. For this purpose, primitive appearance descriptors based on colour or luminance [2] have

* Corresponding author.

G. Bebis et al. (Eds.): ISVC 2014, Part I, LNCS 8887, pp. 445–456, 2014.

been proven to be highly sensitive to pose,perspective changes and lighting variations whereas density based descriptors such as histogram features [3] together with invariant feature descriptors are an appealing alternative to model dynamic changes. In the context of using such vision systems in $\mu$-UAV applications, large simplifications are often made such as dealing with fixed targets, usually with good mathematical model of its representation and within well known or constrained environments [2]. Such systems generally assume slight or no changes to perspective and are not capable of handling occlusions etc. It is important to note here that much less research efforts has been spent on $\mu$-UAV-based tracking of multiple ground targets in an unknown environment, without sufficient a-priori information of the targets. A number of reasons contribute to this, including but not restricted to: a) the challenge in establishing accurate homographies of the visual plane from images captured on an unconstrained moving sensor; b) the problem of having very little or no knowledge of the targets being tracked adds to the previous challenge; and c) finally uncalibrated motion of both the sensor and the target renders mutual dependencies between the subsystems to exploited correctly to improving adaptation.

**Novelty and Contributions:** In this paper, one of the most challenging scientific pursuits of detecting and tracking multiple unknown ground targets using a moving visual sensor (from the $\mu$-UAV) in an unconstrained environment is addressed. The solution is based on distinguishing between camera and target motion by combining global and local motion estimation with background learning type target detection technique. The main novelty of the proposed vision system is its ability to remain self-adaptive to the unconstrained motion of the moving sensor and to multiple unknown ground targets simultaneously in any environment. Such a self-adaptive capability is by the exploitation of tracked local motion descriptors through optical flow measurements: a) in smoothing global motion parameters for improving the accuracy of ego-motion compensation and b) within the background learning framework for facilitating foreground detection through noise and clutter suppression. In addition, the combination of joint feature density and data association further enhances the results of tracking.

## 2 Vision System: An Overview

The proposed dynamic visual system is composed of three phases: a) ego-motion compensation, b) target detection and c) target tracking.

### 2.1 Ego-motion Compensation (EmC)

A good technique for EmC allows accurately estimating the projective transformations (homographies) $T_{t,t+\Delta t}$ linking each pair of consecutive frames $I_t$ and $I_{t+\Delta t}$ in a video. Such EmC techniques aim to accomplish correspondences between robust feature descriptor sets $\boldsymbol{F}_t^s$ and $\boldsymbol{F}_{t+\Delta t}^s$ using an appropriate transformation model and an optimisation mechanism. In this study an affine transformation model is assumed due to adequate decimation of frames that minimise