This CVPR2015 extended abstract is the Open Access version, provided by the Computer Vision Foundation.

Flying Objects Detection from a Single Moving Camera

Artem Rozantsev[1], Vincent Lepetit[2], Pascal Fua[1]
[1] Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL).
[2] Institute for Computer Graphics and Vision, Graz University of Technology.

We are headed for a world in which the skies are occupied not only by birds and planes but also by unmanned drones ranging from relatively large Unmanned Aerial Vehicles (UAVs) to much smaller consumer ones. Some of these will be instrumented and able to communicate with each other to avoid collisions but not all. Therefore, the ability to use inexpensive and light sensors such as cameras for collision-avoidance purposes will become increasingly important.

This problem has been tackled successfully in the automotive world and there are now commercial products [2, 5] designed to sense and avoid both pedestrians and other cars. In the world of flying machines most of the progress is achieved in the accurate position estimation and navigation from single or multiple cameras [1, 3, 4, 6], while not so much is done in the field of visual-guided collision avoidance [7]. On the other hand, it is not possible to simply extend the algorithms used for pedestrian and automobile detection to the world of aircrafts and drones, as flying object detection poses some unique challenges:

- The environment is fully 3D dimensional, which makes the motions more complex.
- Flying objects have very diverse shapes and can be seen against either the ground or the sky, which produces complex and changing backgrounds, as shown in Fig. 1.
- Given the speeds involved, potentially dangerous objects must be detected when they are still far away, which means they may still be very small in the images.

As a result, motion cues become crucial for detection. However, they are difficult to exploit when the images are acquired by a moving camera and feature backgrounds that are difficult to stabilize because they are non-planar and fast changing. Furthermore, since there can be other moving objects in the scene, for example, the person in Fig. 1, motion by itself is not enough and appearance must also be taken into account. In these situations, state-of-the-art techniques that rely on either image flow or background stabilization lose much of their effectiveness.

Figure 1: Detecting a small drone against a complex moving background. (Left) It is almost invisible to the human eye and hard to detect from a single image. (Right) Yet, our algorithm can find it by using motion clues.

We suggest solving the fast small flying objects detection problem by classifying 3D descriptors computed from spatio-temporal image cubes (st-cubes). These st-cubes are formed by stacking motion-stabilized image windows over several consecutive frames, which gives more information than using a single image. What makes this approach both practical and effective is a regression-based object-centric motion-stabilization algorithm. Unlike those that rely on optical flow, it remains effective even when the shape of the object to be detected is blurry or barely visible.

Fig. 2 illustrates some examples of the st-cubes with and without motion compensation. Implementation and more thorough description of the method for regression-based motion compensation is described in the paper.

Our conclusion is that temporal information from a sequence of frames plays a vital role in detection of small fast moving objects like UAVs or air-

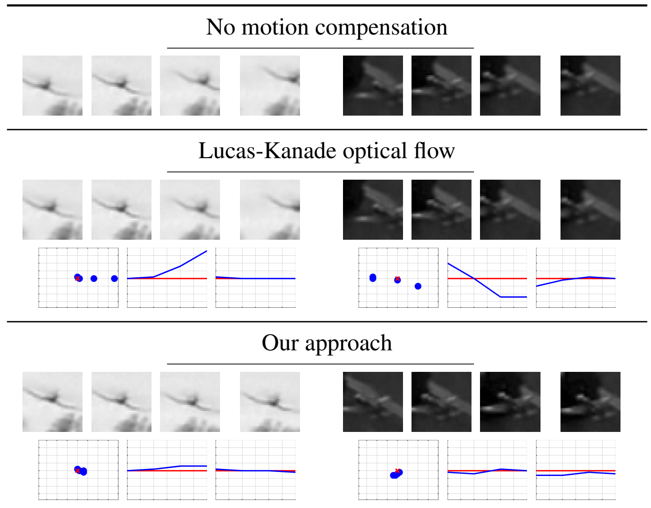

Figure 2: Compensation for the apparent motion of different aircrafts inside the st-cube allows to decrease in-class variation of the data, used by the machine learning algorithms. For each st-cube, we also provide three graphs: The blue dots in the first graph indicate the locations of the center of the drone throughout the st-cube, the red cross indicates the patch center. The next two graphs plot the variations of the x and y coordinates of the center of the drone respectively, compared to the position of the center of the patch. We can see that our method keeps the drone close to the center even for complicated backgrounds and when the drone is barely recognizable as in the right column.

crafts in complex outdoor environments. We therefore developed an object-centric motion compensation approach that is robust to changes of the appearances of both the object and the background. This approach allows us to outperform state-of-the-art techniques on two challenging datasets. Motion information provided by our method has a variety of applications, from detection of potential collision situations to improvement of vision-guided tracking algorithms.

[1] C. Forster, M. Pizzoli, and D. Scaramuzza. SVO: Fast Semi-Direct Monocular Visual Odometry. In *International Conference on Robotics and Automation*, 2014.

[2] IntelligentDrive. Mercedes-Benz Intelligent Drive. http://techcenter.mercedes-benz.com/en/intelligent_drive/detail.html/.

[3] S. Lynen, M. Achtelik, S. Weiss, M. Chli, and R. Siegwart. A Robust and Modular Multi-Sensor Fusion Approach Applied to MAV Navigation. In *Conference on Intelligent Robots and Systems*, 2013.

[4] L. Meier, P. Tanskanen, F. Fraundorfer, and M. Pollefeys. PIXHAWK: A System for Autonomous Flight Using Onboard Computer Vision. In *IEEE International Conference on Robotics and Automation*, 2011.

[5] Mobileeye. Mobileeye Inc. http://us.mobileye.com/technology/.

[6] S. Weiss, M. Achtelik, S. Lynen, M. Achtelik, L. Kneip, M. Chli, and R. Siegwart. Monocular Vision for Long-Term Micro Aerial Vehicle State Estimation: A Compendium. *Journal of Field Robotics*, 30:803–831, 2013.

[7] T. Zsedrovits, A. Zarándy, B. Vanek, T. Peni, J. Bokor, and T. Roska. Visual Detection and Implementation Aspects of a UAV See and Avoid System. In *European Conference on Circuit Theory and Design*, 2011.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.