Measurement 48 (2014) 195–207

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Optical flow background estimation for real-time pan/tilt camera object tracking ☆,☆☆

Daniel D. Doyle *, Alan L. Jennings, Jonathan T. Black

Department of Aeronautical and Astronautical Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, United States

ARTICLE INFO

Article history:
Received 24 May 2013
Received in revised form 13 August 2013
Accepted 10 October 2013
Available online 23 October 2013

Keywords:
Pan/tilt camera
Optical flow
Background subtraction
Object tracking

ABSTRACT

As Computer Vision (CV) techniques develop, pan/tilt camera systems are able to enhance data capture capabilities over static camera systems. In order for these systems to be effective for metrology purposes, they will need to respond to the test article in real-time with a minimum of additional uncertainty. A methodology is presented here for obtaining high-resolution, high frame-rate images, of objects traveling at speeds $\geqslant$1.2 m/s at 1 m from the camera by tracking the moving texture of an object. Strong corners are determined and used as flow points using implementations on a graphic processing unit (GPU), resulting in significant speed-up over central processing units (CPU). Based on directed pan/tilt motion, a pixel-to-pixel relationship is used to estimate whether optical flow points fit background motion, dynamic motion or noise. To smooth variation, a two-dimensional position and velocity vector is used with a Kalman filter to predict the next required position of the camera so the object stays centered in the image. High resolution images can be stored by a parallel process resulting in a high frame rate procession of images for post-processing. The results provide real-time tracking on a portable system using a pan/tilt unit for generic moving targets where no training is required and camera motion is observed from high accuracy encoders opposed to image correlation.

1. Introduction

1.1. Motivation

The Unmanned Aircraft System (UAS) flight plan describes Nano/Micro UAS as "aircraft capable of conducting a variety of indoor and outdoor reconnaissance sensing missions" [1]. Developing aircraft capable of these activities requires an in-depth understanding of biological counterparts and measurement systems capable of capturing in-flight dynamics. In order to gain this type of understanding for biomimetics advancement, measurement systems require the ability to measure the motion of flying organisms in an unconstrained, real-time manner.

The Air Force Institute of Technology and the Air Force Research Laboratory have worked on advancing flapping wing mechanisms through flight control research, wing design and analysis of dynamics [2–4]. There are many other facilities devoted to developing and understanding the mechanics of flapping wing flight. A sample includes: the Harvard Microrobotics Lab [5], the Korean Advanced Institute of Science and Technology (KAIST) Smart Systems and Structures Lab [6], and Georgia Institute of Technology's Intelligent Control Systems Laboratory [7]. Due to the small scale and speed involved with these vehicles, novel systems have been needed to measure forces and motion without disrupting flight.

☆☆ The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.
* Corresponding author. Tel.: +1 937 656 8743.
E-mail addresses: daniel.doyle@us.af.mil (D.D. Doyle), alan.jennings@afit.edu (A.L. Jennings), jonathan.black@afit.edu (J.T. Black).

0263-2241/$ - see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.10.025

Contact measurement methods pose many challenges when working with UASs. These methods consist of, but are not limited to, using strain gages, paint, retro-reflective markers, etc. which typically alter vehicle performance either by weight or other restrictions to an already light-weight device (typically ~15 gm). The alternative approach is using non-contact measurement methods for observing the UAS's behavior without disrupting it. Current non-contact displacement measurement methods include laser vibrometry, laser range finders, capacitance measurement, interferometry and photo/videogrammetry. Laser vibrometry is the preferred method for measuring non-contact vibrations [8]. However, only a single point is measured at a time, and alignment must be maintained through the test. Videogrammetry uses multiple, synchronized images or stereo pairs and typically retro-reflective markers or projected targets for determining surface shapes and can capture shape and motion from a single test [4]. Projected targets are not practical for use with moving targets, and adding retro-reflective markers negates non-contact measurement. Using a camera allows for more freedom of movement. However, the subject must stay in the field-of-views (FOVs) with little to no occlusion. Incorporating a pan/tilt camera with real-time tracking will allow for the camera to move so that the object remains in the FOV. This will expand the effective capture volume without a sacrifice in resolution of the object, for better flight data over a larger set of flight profiles. The ideal system would have a fast response (high frame rate and low latency allows for faster vehicles), high resolution (for better discrimination of the vehicle) and work with general dynamic objects (not require selection or training).

1.2. Related work

Surveillance and robotics have driven development of computer vision (CV) tools. The CV tools related to this work involve camera motion compensation, object detection, and object tracking. Features discriminate a point from others, such as differences in color or intensity gradients. Different methods can be employed to detect specific features or objects in the scene; such as point detection, background modeling, image segmentation or classifiers based on supervised learning. Tracking can be done by matching points, or objects can be tracked as a whole by employing silhouettes or kernels. See Yilmaz et al. [9] for a detailed survey on object tracking.

The system presented here pulls existing methods in a novel combination for general purpose. It uses known camera motion with exact pixel displacement, corner detection and optical flow tracking. Image processing is performed on the graphics processing unit (GPU) for faster speed. The resulting system is able to track a single moving object using a single camera.

Camera motion can either be undesired, requiring image stabilization [10], or desired, such as for tracking or building a background mosaic [11]. The principle involved is the same: finding the transform from an image to a reference image [12]. Transforms are typically found between frames by fitting or matching point correspondences [13], phase correlation [14] or block matching [15]. An alternate technique is to use a projective texture as a Ref. [12]. The use of the known camera motion is surprisingly not used more often. Though the exact model is developed [16], often simplified models are used, such as the small angle approximation [17]. Algorithms that do not require known camera motion have the advantage that they can be used with hand-controlled cameras or other moving mounts. However, estimating motion comes at the cost of increased error, increased processing time, the object appearing smaller in the image (because the background must be viewed) and requiring texture, or the lack of texture, on the background and foreground. This work uses the exact model based on encoder positions for high accuracy predictions, even with large displacements (~1/7th of the image).

An alternative to compensating for tracking the camera's motion is to use one or more static cameras to direct the moving camera. This has the advantage that moving objects are simpler to detect with a static camera. Applications include controlling a remote vehicle [18,19] or following sporting events [20]. For this alternative to be effective, the FOV must cover the range of the moving camera, increasing setup time and calibration. However, stationary, smart cameras can be used to self-calibrate with fast setup and reasonable accuracy [21]. The work presented here uses a single pan/tilt camera to track a moving object.

As processing power has developed, new and existing tools have been applied in real-time, and have been used for both tracking an object within a static camera's view or tracking an object with a moving camera.

Heuristics can be used to design detectors for increased accuracy or speed. General purpose feature descriptors can be used with training data, tuning or selection for a wide variety of objects. Some specific applications using a pan/tilt camera include tracking the color on a ball [22,23], human positioning/tracking [24–26], faces [23,26,27], and moving vehicles/ships [13,27–29]. Various methods are used in these applications for detecting objects such as mean shift and its variant continuously adaptive mean shift (CAMShift) [22,25,26,28,29], background modeling using a Gaussian mixture model (GMM) [25,29], skin color segmentation [23], training/learning an object's appearance using principle component analysis (PCA) and sum of squared difference (SSD) [27], and template-based detection using speeded up robust features (SURF) [13]. The general purpose methods often require selection, a pre-defined capture area (i.e., alert zone), training, and/or tuning. The cost of using application specific algorithms include decreased system versatility, increased setup/calibration time, and limitations on feature changes such as illumination, object pose, shape, and size.

Features can also be non-specific, such as motion-based features, which are useful in general purpose applications. General purpose (motion-based) tracking allows for more flexibility in handling multi-purpose problems and allows tracking of an object of any size or shape [17]. This is well-suited for a metrology system capable of tracking a multitude of objects without having the concern of training, or changes in illumination or appearance associated with application specific methods. Short duration/sudden events are easily accommodated. Operator delay is