Computationally-efficient Navigation of Intelligent Unmanned Ground vehicles

Peyman Moghadam
School of Electrical & Electronic Engineering
Nanyang Technological University, Singapore
peym0002@e.ntu.edu.sg, +65-91160043

Saba Salehi
Electrical & Computer Engineering Department
National University of Singapore, Singapore
saba.salehi@nus.edu.sg

EXTENDED ABSTRACT

Autonomous navigation system for Unmanned Ground Vehicles (UGVs) in outdoor unstructured environments have been studied for several decades for a variety of applications such as planetary exploration, military, and hazardous areas for human due to natural or unnatural events. Fundamentally, the objective is to have an autonomous vehicle which moves reliably at a desired speed from a starting location to a goal point while it avoids obstacles. This scenario in outdoor unstructured environments confronts many challenging difficulties such as highly complex scene geometry, ground cover variation, uncontrolled lighting, weather conditions and shadows.

The autonomous navigation systems are often facilitated with a set of algorithms and sensors such as a stereo vision system or Light Detection And Ranging (LIDAR) to sense their environments in order to predict the traversability of the terrain (i.e. to discriminate traversable path and non-traversable path regions). Despite the fact that LIDAR has proved to be quite reliable and accurate in range measurements, it is relatively expensive, bulky, heavy and unable to provide visual information of the scene (i.e. color and/or texture). Accordingly, color stereo vision cameras are typically used as the alternative primary sensors which can provide not only 3D perception of terrain geometry but also color and texture information. However, the problem with present stereo vision systems is the limited cameras' field of view and maximum range. As a result, they can only provide accurate geometry information from their near-field. This limitation of near-sightedness of typical off-the-shelf stereo vision systems can be overcome by associating the geometry of terrain segments close to the vehicle (near-field) with their visual appearance attributes (e.g. color and/or texture) and use this association to segment terrain and obstacles in the far-field. However, finding a global correlation between terrain geometry characteristics and appearances that can be broadly applied is formidable due to the complex variability of appearance with terrain geometry, lighting, shadows, and weather conditions, especially in outdoor environments. Therefore, most hand-designed deterministic and rule-based system has proved to be ineffective as they are not robust to changing environments due to their inability to adapt to unforeseen ground cover variations. To address this problem, one promising approach is to use Machine-Learning techniques to replace hand-designed deterministic vision-based terrain classification systems for UGVs. This problem has been addressed by several research groups in the DARPA-LAGR program by exploiting self-supervised near-to-far learning algorithms.

The self-supervised (near-to-far) learning refers to an approach of generating near-field traversability labels (road and non-road) from each incoming pair of stereo images using only the near-field stereo information. Next, these labels are linked with the visual appearance features of one reference image (e.g. right image of stereo camera system) to train a classifier, and the model obtained is used to classify the far-field terrain well beyond the stereo range. Subsequently, this model is discarded and therefore, one full training and classification cycle are completed on every incoming pair of stereo images. Hence, every processing cycle of the basic scenario of self-supervised near-to-far learning involves four main steps, (1) stereo auto-labeling, (2) feature extraction, (3) classifier training and model generation, and (4) applying that model to the entire input image and subsequently discarding it.

This paper proposes to enhance the existing methods of online, self-supervised learning (with application to autonomous navigation systems) through the introduction of a confidence-based, self-supervised auto-labeling system. A Salient-based feature selection method is also introduced to bias more attention to prominent features. The proposed method is general and can be applied to a variety of applications. Finally, a memory efficient semi-incremental SVM training is presented to decrease the computational load and time. To illustrate the effectiveness of the proposed approaches, experimental results on an autonomous navigation system are given using real datasets from the DARPA-LAGR project, which is the current gold standard for vision-based terrain classification using machine-learning techniques.

The main contribution of this study is to improve both the performance and the speed of learning algorithms through simple and efficient computational approaches which are the principal requirements in a practical system. First, confidence-based auto-labeling for self supervised online learning is introduced which detects and eliminates the input samples with low confidence level that are susceptible to be mislabeled. This technique results in less unnecessary data processing as well as reduction of support vectors in addition to an increase in prediction accuracy rate. Then, a new salient-based approach for feature weighting is presented which is able to detect the salient features through bottom-up saliency definition and devotes higher attention to them through top-down weighting. Worth to mention that top-down weights are considered to mimic the role of the brain's feedback discussed in biological contexts and are designated for discriminating the regions of interest. These weightings lead to a lower FP-rate. Finally, a semi-incremental SVM learning for autonomous navigations is proposed which suggests using the model of the current image for subsequent images on the condition that the model provides a satisfactory performance on near-field of those images.

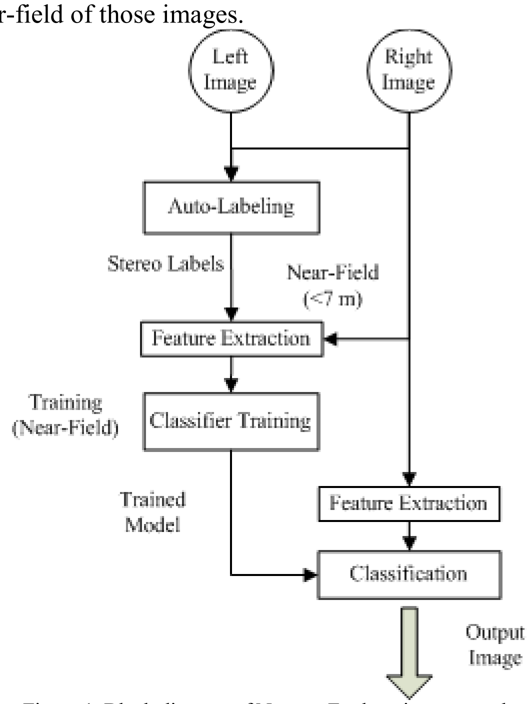

Figure 1. Block diagram of Near-to-Far learning approach.

Figure 2. Input Image (Top left), hand labeled ground truth (green: ground plane, red: obstacle, and blue: unknown) (top right), basic auto-labeling output image (bottom left), proposed auto-labeling algorithm output image (bottom right).

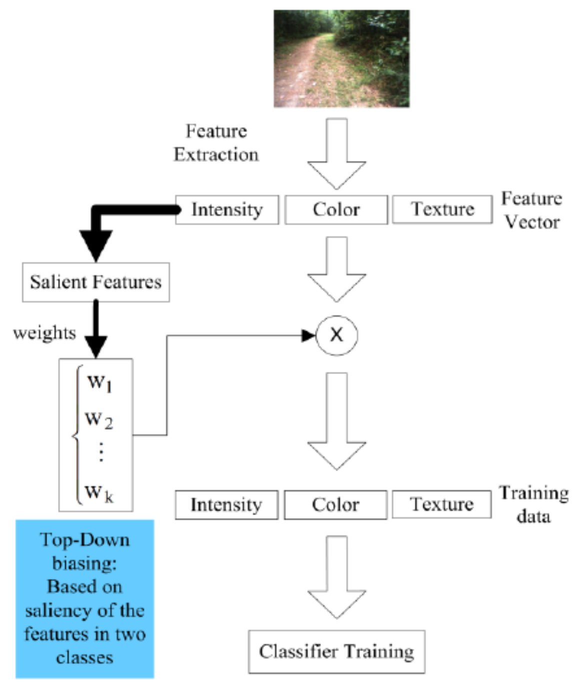

Figure 3. Block diagram of our proposed biological-based feature selective attention model.