

Listening closely to see far away: Radar-based terrain classification from auditory signals

David Williams
University of Oxford
Department of Engineering Science
Parks Road, Oxford, UK
david.williams@keble.ox.ac.uk

Daniele De Martini
University of Oxford
Department of Engineering Science
Parks Road, Oxford, UK
daniele@robots.ox.ac.uk

Letizia Marchegiani
Aalborg University
Department of Electronic Systems
9220 Aalborg Ø - Denmark
lm@es.aau.dk

Paul Newman
University of Oxford
Department of Engineering Science
Parks Road, Oxford, UK
pnewman@robots.ox.ac.uk

Safe navigation of mobile robots in outdoor environments requires a detailed understanding of surrounding terrains, as each poses a unique challenge. Robust terrain classification is therefore a key problem to be solved for reliable outdoor deployment of robots. Previously, lidar and vision-based terrain classification systems have been investigated and despite their success in certain scenarios, they still have some significant disadvantages. Environmental factors such as heavy rain, fog and dust clouds all greatly degrade the performance of these systems, thereby limiting their range of applications. A terrain classification system using FMCW scanning radar would however be able to operate as normal in these conditions and at a longer range than all alternative systems.

To design a radar-based terrain classification system using a supervised learning approach, a labeled dataset of radar scans is required. This means each radar scan has to be labeled on a bin-by-bin basis, which is both difficult and time-consuming to do by hand. We therefore propose an alternative solution, whereby we weakly supervise the training of the radar-based classifier. This weak supervision is provided by an audio-based classifier (itself trained in a supervised manner), that labels the different terrain types on a radar scan. Audio input has been chosen in the light of previous work in the literature and on the fact that it is completely resilient to the scene's appearance and much less influenced by weather conditions than visual-based methods.

Audio-based classification is made possible by the fact that each interaction between the robot and the ground has a terrain-specific audio signature. The advantage of an audio-based classifier is that only the terrain the robot is currently on has to be labeled, as opposed to the entire surrounding area. This simplification allows us to more simply build the audio-based classifier by collecting data from each surface separately. The class labels from the trained audio-based classifier can then be used to label the radar scans. In order to do this, we need to relate the terrain at the robot's location to the radar scan. Through the use of odometry and GPS, we have an understanding of where the robot is at any given time step and where it has been. This information can then be combined with the radar scans to provide labels for any ground that the robot has traveled over and also any ground that has similar features on the radar scan.

To classify the audio data, we first preprocess the audio signal to generate time-frequency representations based on Gammatone filterbanks (modeled on the human auditory system). Once labeled, we can then use the dataset of time-frequency representations to train our classifier. For classification, we use a deep convolutional neural network and implement it using TensorFlow. For data collection, a Clearpath Husky robot is fitted with microphones by the front wheels and a Navtech CTS350-X FMCW scanning radar. At present, both radar and audio data have been collected for a small number of terrains and results from preliminary classification experiments are very promising.

Keywords

Terrain classification, weakly supervised learning, radar, audio, machine learning, computer vision, autonomous machine imaging algorithms, image & signal processing applications