Seeing with sound; surface detection and avoidance by sensing self-generated noise

Simon Wilshin, Stephen Amos, Richard J Bomphrey
Structure and Motion Lab, Royal Veterinary College
swilshin@rvc.ac.uk

http://www.imavs.org/

ABSTRACT

Here, we demonstrate obstacle and secondary drone avoidance capability by quadcopter drones that can perceive and react to modulation of their self-generated acoustic environment when in proximity to surfaces. A ground truth for the interpretation of self-noise was established by measuring the intrinsic, three-dimensional, acoustic signature of a drone in an anechoic chamber. This was used to design sensor arrangements and machine learning algorithms to estimate the position of external features, obstacles or another drone, within the environment. Our machine learning approach took short segments of recorded sound and their Fourier transforms, fed these into a convolutional neural network, and output the location of an obstacle or secondary drone in the environment. The convolutional layers were constructed with a suitable topology that matched the physical arrangement of the sensors. Our surface detection and avoidance algorithms were refined during tethered flight within an anechoic chamber, followed by an exercise in free flight without obstacle avoidance, and finally free flight obstacle detection and avoidance. Our acoustic sense-and-avoid capability extends to vertical and horizontal planar surfaces and tethered secondary drones.

1 INTRODUCTION

Many methods exist by which a drone might sense its environment, such as LIDAR and laser range-finders[1], cameras and optical flow sensors[2], pressure sensors [3], or even bioinspired "hair sensors"[4, 5]. Each such sensor modality has strengths and weaknesses.

Another method of environmental detection is acoustic sensing with microphones. These have been used to estimate the distance[6] and bearing[7] between drones and the relative location of drones[8] in a swarm. Drone mounted acoustic sensors, when integrated MEMS and particle velocity sensors, have been shown to be capable of detecting the location of large civil aircraft[9]. Ground based acoustic sensors have also been used to localise small gasoline-powered UAV over hundreds of meters[10].

Acoustic surface detection methods have enjoyed success on robotic platforms drawing inspiration from biological systems[11]. Echolocation has been used to relocate a robot in an environment by comparison with existing maps of said environment[12]. Active echolocation on has also been employed on the crazyflie drone to identify surfaces[13], though this requires a loud on-board speaker to generate the required sound, which we wish to avoid.

Flying drones generate sound, principally from the motors and the aeroacoustic noise from the propellors (which we term the acoustic signature), and nearby surfaces in the environment reflect that sound. The spatio-temporal structure of the reflected acoustic signature contains information about the position of the drone relative to obstacles in the environment. Microphones that observe this information can also detect the acoustic signature of other drones in the environment. While physics based acoustic modelling approaches have been shown to be theoretically viable for surface detection in real time[14], here we employ a machine learning approach.

This method of obstacle and drone detection offers several advantages over others. Unlike cameras, microphones work in the dark and draw minimal power; our arrays draw only a few mA. This makes them suitable even for micro-drones weighing only a few tens of grams.

They can be used passively to listen for the sound of other drones, or actively by listening to the reflected sound of the drone's own acoustic signature. In either case, they make such a drone no more conspicuous than it would be operating in free flight, since a flying drone must operate its engines to resist gravity regardless of whether or not it is sensing the environment. This in contrast to laser range-finding or LIDAR; if a laser beam is directed at a suitable sensor it can be readily detected and the direction of the drone quickly determined.

The approach is also more general than sensing environmental changes by monitoring motor power consumption directly. Such sensing can be used to detect a surface below a drone and shut off the motors for landing (for example in[15]), but has limited utility for sensing vertical planar surfaces in the environment and no capability to detect other drones.

Here, we show how to use such information to detect surfaces, both in ideal conditions while tethered in an anechoic chamber, and in the variable, suboptimal conditions that occur in free flight. We then demonstrate how such information

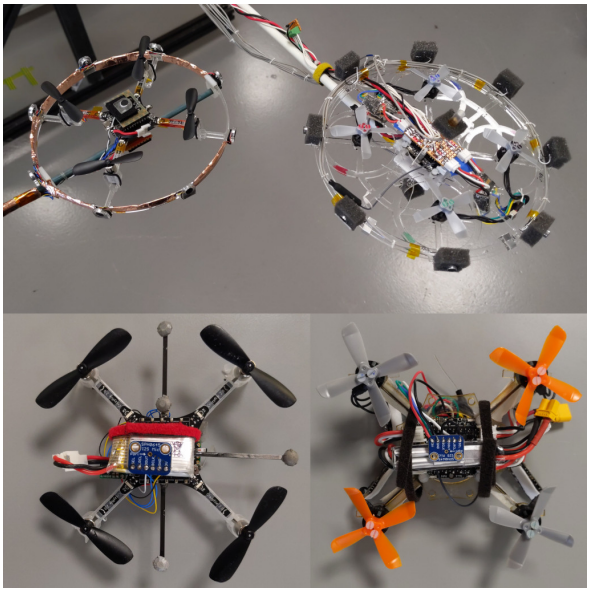

Figure 1: Various microphone arrays deployed on the Crazyflie and Mormoops prototype. At the top are the analogue microphone arrays deployed in a ring around the Crazyflie on the left and the Mormoops on the right. Below are the digital microphone arrays (in ground detection configuration with microphones on top of and below the drones body), again with the Crazyflie on the left and the Mormoops on the right). The bottom left drone was the final prototype used in the sense and avoid experiments

can be used to avoid collision with obstacles in free flight.

Finally, we demonstrate how such a sensor system can be used to detect another drone's location in the environment, while tethered, with a high degree of accuracy.

2 METHODS

2.1 Acoustic sensors

We have developed several acoustic sensor systems as required, including tethered analogue microphone arrays, on-drone analogue microphone arrays digitised by an onboard microcontroller, and onboard digital microphones managed by an onboard microcontroller ((Fig. 1)). While specialised small, directional acoustic sensors do exist we opted for an array of microphones similar to those that have already seen successful employment on drones[16]. Directional microphones were rejected because it was felt having multiple recordings on an array would allow for better detection of interference effects.

The outputs from tethered analogue microphone arrays were digitised during the characterisation of the acoustic signature of the Crazyflie, and for tethered surface detection in the anechoic chamber, and for the detection of anther drone in the environment when both drones were tethered.

Following measurements of the acoustic signature of the drones and the tethered trials, for our recordings in free flight we used digital I_2S microphones connected to a teensy 4.0. This permitted the fast onboard notch filtering and power estimation of recordings needed for in flight surface detection and avoidance.

2.2 Self-generated drone acoustic signature

In order to design and calibrate our surface detection algorithms for our drone, we needed to to obtain the self-generated acoustic signature: that is, the sound the drone made in various directions while hovering.

To do this, we mounted an array of analogue acoustic microphones on a Crazyflie and recorded the sound generated at a fixed distance at a range of angles (elevations at $45\,\mathrm{deg}$ intervals and at 18 azimuthal angles approximately $15\,\mathrm{deg}$ seperation, $n = 90$) arranged around the drone. Note these arrays are not the same as those fixed to the drone, as can be seen in figure 2, these microphones are attached to the fixed white semicircle suspended in the chamber. Recordings were taken with 6mm diameter uni-directional electret condenser microphones (RS PRO, 50 to 16000 Hz at -47dB) wired to separate signal conditioning amplifiers (Vishay 2210B), and a Powerlab DAQ (PowerLab 16/35, ADInstruments).

We took recordings in an anechoic chamber (Fig. 2) to maximise the quality of the acoustic signature signals, shown in Fig. 5.

2.3 Tethered surface detection using self-generated sound

Based on the results from the self-generated acoustic signature of our drones(Fig. 5), we designed a microphone array for detecting surfaces in the environment with a high density of sensors at eight different azimuthal angles and two sensors above and below the drone (Fig. 1, top right). We used this array to detect detect planar surfaces while tethered in an anechoic chamber using self-generated noise.

Based in part on the results from the acoustic signature of the drone we decided to implement a machine learning classifier built using PyTorch[17]. The classification network is built from a configurable input layer, two convolutional layers, two fully connected layers and an output layer.

The input layer can be adjusted to match the sensor suite being used via a human readable YAML configuration file. In this configuration the data is split into blocks.

These blocks are the microphone recordings from the eight sensors in a ring, the two microphone on the top and bottom of the drone, and two additional blocks formed from the Fast Fourier Transform (FFT) of these two sets of signals. The FFT was taken along the time axis of each block using the rfft function of tensor flow.

Two convolutional layers are then used. A useful summary of recent applications is provided in [18] . The surrounding sensors form a ring around the drone, and any classification network should respect this topology, i.e. there should be no 'start' or 'end' sensor, and any mathematical

http://www.imavs.org/

http://www.imavs.org/

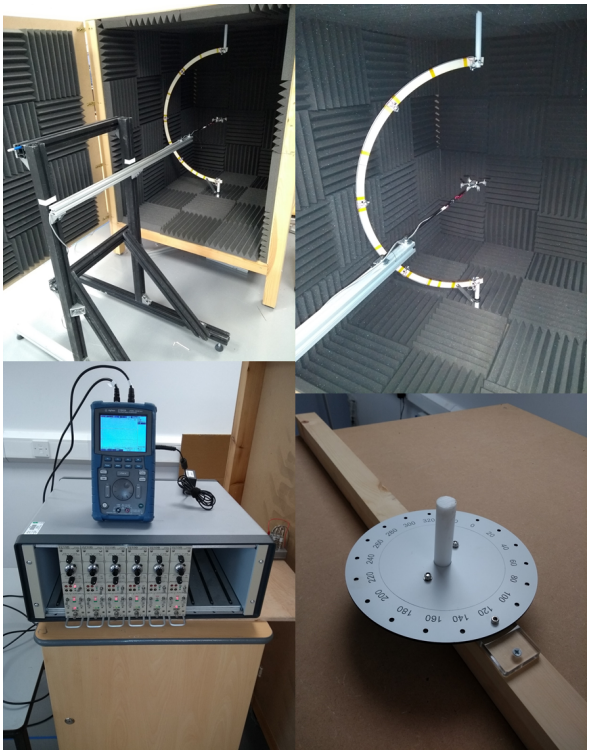

Figure 2: Apparatus for measuring quadcopter acoustic signature. The arc is fitted with five microphones and can be rotated in azimuth to repeatable angles around the tethered quadcopter. Microphones are mounted at the North and South poles, the equator and +/- 45 degrees. The microphones are fed by common excitation and the output signal is received by individual Vishay 2210B signal conditioning amplifiers. The quadcopter motors are provided with a voltage known to support the weight of the aircraft.

operation should be invariant to a transformation which shifts all the sensors around this ring.

As such we use a circular convolution for the sensor input elements on the ring. Since the top and bottom elements of the microphone are completely included in the convolution along the sensor number axis, these can be a regular convolution.

If $x_i j$ is an element of the blocks, x, discussed above, with i running from one to the sample length N, that is if we take 20 millisecond blocks at 22 kilohertz this would be 440 samples, j running from one to the number of elements in this block (for example eight if we are dealing with the eight microphones around the drone), then this layer performs the following operation:

$$y_\mu = b + w_\mu \star x \tag{1}$$

Where $\star$ is the cross-correlation operator, w_μ is the kernel associated with the μth convolution, of which there are M total, and b is a constant matrix. In the language of convolutional neural networks, this first layer assumes inputs with only one channel, and outputs M channels.

All units in the network are Rectified Linear Units [19] (ReLU) to improve training ease training and speed convergence, and so this response y is modified by:

$$y'_\mu = \max(0, y_\mu) \tag{2}$$

Our network pools after every convolutional layer to generate invariant features, which makes the network less sensitive to the precise locations of features and reduces the number of parameters in the network [20], reducing the risk of overfitting. We found that two-dimensional max pooling worked well, splitting the layer activations into $K \times L$ blocks and transforms the layer by downsampling to only the maximum value in these blocks.

Our network then proceeds with another convolutional layer with ReLU and pooling. Next, we have two fully connected layers. We treated the output of the last convolutional layer as a single vector a_i. The activation b_j of the fully connected layer is

$$b_j = c_j + \sum_i m_{ji} a_i \tag{3}$$

The units in the fully connected layer are again ReLU units, for the same reason as previously discussed. No pooling is performed here as the fully connected layer respects the topology of the input only in the most trivial way (every unit is connected to every input), meaning neighbouring units are likely to be unrelated.

Another fully connected layer is then used followed by the output layer. The output layer is, for this case, the azimuthal angle of the obstacle, and a vector pointing to the obstacle in the plane. The azimuthal angle can only be 0 or 90 degrees, and is re-interpreted after the network is trained as a class variable indicating wall position.

http://www.imavs.org/

The network is trained using Stochastic Gradient Descent (SGD) [21] to minimise the mean square error.

The network performance was evaluated on a test data set of ten thousand sections two milliseconds in length that were not used in the training data. As the testing and training data are drawn randomly from sixty second length trials there is a risk that temporal correlation within the longer trials could allow for identification of the sample.

2.4 *Surface detection in free flight*

As a precusor to surface avoidance we created a version of our prototype drone with the ability to detect surfaces in free flight. The prototype (called Mormoops) consisted of a Bolt flight controller *(Bitcraze, Malmö, Sweden)*, a custom chasis and four Betafpv 1105 5000KV brushless motors *Betafpv, Hong Kong, PRC)*. The drone was operated near a planar surface, and our neural network tasked with detecting the surface.

A Centeye optical flow sensor, *(Centeye, Washington, DC, USA)* was used to measure the change in height of the drone over one minute trials. The flight controller maintained constant height based on optical flow odometry and no obvious drift was observed over these short flights. The drone was flown at three different heights, one deep in ground effect at approximately 5cm height and two outside at 15 and 20cm.

Surface detection in free flight benefits from higher bandwidth microphone recording and processing. As such we decided to use a digital microphone platform combining SPH0645LM4H I^2S microphones *(Adafruit, New York City, USA)* with a Teensy 4.0) microcontroller *(PJRC, Sherwood, OR, USA)* with custom SD card reader for storage.

This would allow us to use the Teensy microcontroller's specialised I^2S audio processing hardware that enables high quality recording, fast filtering, Fourier transforms, and other signal processing.

Our neural network was then required to distinguish these states using the sound recordings from $20ms$ samples.

2.5 *Free flight surface avoidance*

Tests of wall detection capabilities were performed during the previous free flight experiments. However, this version of our drone could not safely resist wall effect, so insufficient training data could be gathered.

We modified our algorithms to detect and avoid both horizontal (ground) and vertical (wall) planar surfaces in free flight. We used the difference in power of the recording from two notch filtered microphones, a method that was based on examination of the structure of the machine learning approach - specifically the frequency dependency - and required minimal tethered training data.

The controller requires platform-specific calibration. The frequency range of the notch filter is specific to the drone that carries the sensor package, as are the gains on the microphones, and the threshold for detection. Ground detection is insensitive to these parameters, with simple benchtop tests sufficient for a good calibration. Wall detection requires in-flight refinement and iteration and is less reliable. In both instances the threshold for engaging the avoidance behaviour had to be set prior to the test.

Figure 3: Image of the autonomous flight arena during its early development showing Qualisys motion tracking cameras and a Crazyflie UAV on the floor in the centre of a netted enclosure. This was used to track and control the Crazyflie during sense-and-avoid flight trials.

At present, for avoidance purposes the system behaves as a classifier (the conditions being "wall present" and "no wall"). For the ground plane this is a simplification of the multiple heights that can be discerned by the machine learning controller, but for wall detection little evidence of continuous signal was apparent beyond the simple detection of wall effect.

Once an estimate of the surface position is available on the Teensy microcontroller, the information is communicated to the onboard Crazyflie flight controller by a parallel digital on the GPIO pins (Fig. 1). This distance estimate is then transmitted, via the Crazyradio, to a computer that tracks the Crazyflie using a Qualisys motion capture system *(Qualisys, Göteborg, Sweden)*. (Fig. 3)

If an obstace is detected, evasive action is taken to reposition the drone by updating the target position set on the control computer.

2.6 *Sensing other drones*

We determined to what extent it was possible to identify a second drone using the same acoustic sensor array. Initially, we confined our considerations to detecting drones that are physically different, the Crazyflie and our new Mormoops drone.

We ran both drones simultaneously in tethered flight and recorded the acoustic signature of each at a range of distances and angles, as in Fig. 4. The objective was to use our microphone array and machine learning approach to detect and locate the secondary drone, from the observer drone.

The analogue array was used to record from ten microphones distributed as for tethered surface detection. Recordings were performed in the same manner, though outside of

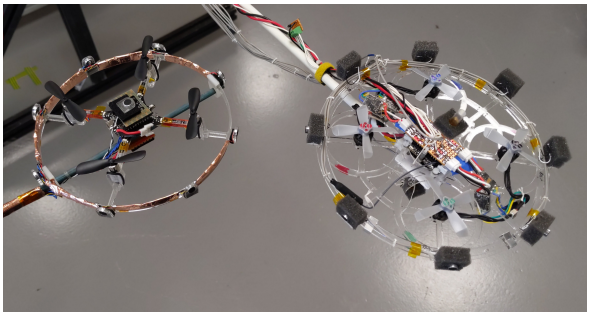

Figure 4: Experimental configuration for detecting and localising a prototype Mormoops drone from a Crazyflie, and *vice versa*. Recordings were taken from the microphone array on each drone, and used in conjunction with our machine learning approach to identify and localise each drone.

the anechoic chamber due to space constraints.

3 RESULTS

3.1 *Self-generated drone acoustic signature*

Acoustic signatures are shown in Fig 5. Position dependent differences are apparent in these signals, which suggested that discerning positional information about quadcopters from external acoustic recordings was highly feasible. The spatial distribution of the signal suggests the need for multiple sensors around the body of the drone. These results suggested, and later observations confirmed, that a single pair of microphones mounted above and below the body would be sufficient in these directions for planar surface detection.

3.2 *Tethered surface detection through self-generated acoustics*

Performance of the network was excellent (see Fig. 6 and Fig. 7).

The drone was able to distinguish perfectly between the two conditions (wall to the side, floor below) for all examples in the test data set. When estimating the height of the drone above the surface the standard error was 36.1mm. This error was relatively consistent over the range of distances measured, meaning that when deep in ground effect the detection of the ground was relatively inaccurate (although amply sufficient to trigger a surface avoidance response), and surprisingly good at large distances (although we note that these constitute close to ideal conditions). The limits of this floor detection capability were not found as we reached the maximum height of the anechoic chamber before such a limit was reached. For practical purposes this is sufficient, although detecting such a limit for a free-flying drone is highly desirable and is a priority future objective.

Similarly, the horizontal distance at which performance begins to degrade was not detected, as the algorithm maintained a good performance out to 250mm, with a standard error on the distance to the wall at 39.5mm and an angular

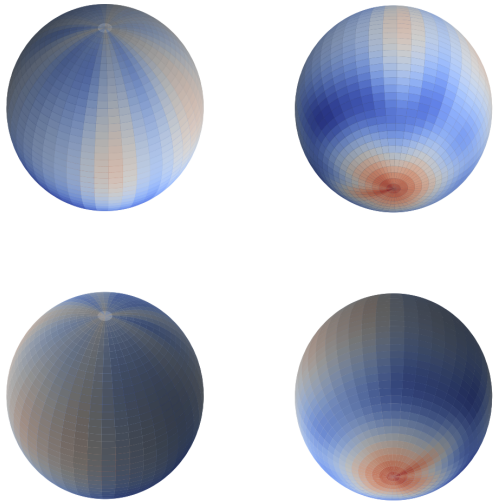

Figure 5: The external acoustic signature of a crazyflie. Here is shown a sphere around the drone and the acoustic power recorded at that location around the drone is displayed, cool colours indicate low acoustic power while hot colours indicate high acoustic power. The engines generate high pressure below the drone, and this can be seen as the red region on the bottom of the sphere. The four sphere are four different views, left and right are different elevations and the plots are spheres are rotated half a circle azimuthally going from the top to bottom row. The covariance matrix used in our anomaly detection algorithm can likely be inferred (or at least constrained) through these observations. This has the virtue that a physically derived estimate is guaranteed positive definite.

standard error (assuming a Gaussian distribution, rather than von Mises, which is a reasonable approximation here as the error is small), of $16.4\,\mathrm{deg}$.

3.3 *Surface detection in free flight*

For the trials in ground effect, the sound recordings were clearly different, close approach to surfaces can be reliably detected from acoustic signatures. The drone was set to hover at three discrete heights, 5cm, 15cm and 20cm. The diameter of the rotos was 5.08mm. For the case where the drone was at 5cm height, deep in ground effect, 96% of trials were successfully identified. For the 15cm and 20cm cases 46% and 48% were successfully identified. Close detection of planar horizontal surfaces is reliable, especially if integrated over several seconds.

3.4 *Free flight surface avoidance*

The drone sucessfully takes action to avoid collisions with planar surfaces. It manoeuvres away from them as they are brought close. Examples of such behaviour can be seen in

http://www.imavs.org/

http://www.imavs.org/

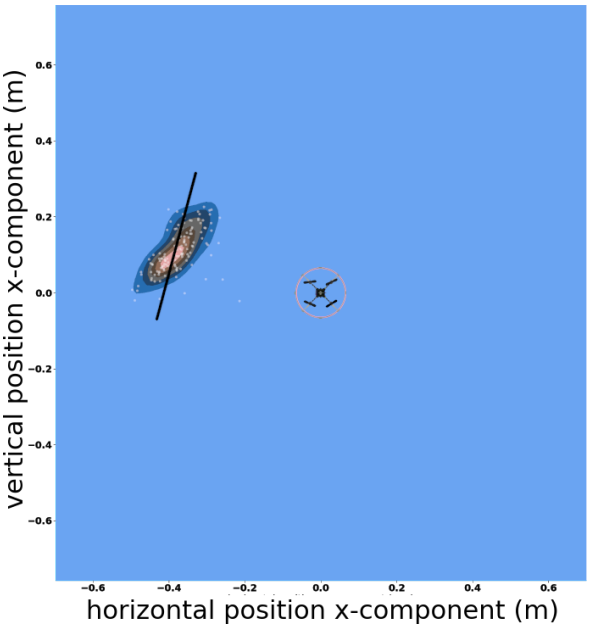

Figure 6: Performance of the machine learning sensory integrator in detecting the position of a surface to the side of the drone. The drone is shown inside the red circle, the wall is the solid black line, the x- and y-axis are position in meters. The white dots show the estimated position of the wall according to the algorithm, with the kernel density plot [22] in blue, black and red constructed from this data. There is strong agreement between these estimates and the actual wall position. The standard error in the estimated position by shortest distance to the wall was 39.5mm, and the standard error on the angle was 16.4 deg

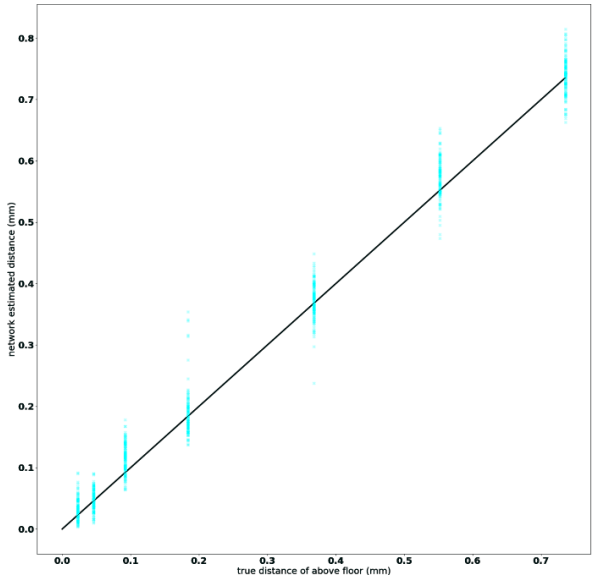

Figure 7: Performance of the machine learning sensory integrator in detecting the position of a surface to the below of the drone. On the x-axis is the true position of the surface below the drone in meters. On the y-axis is the estimated position of the surface according to the algorithm. The standard error on the estimated distance was 36.1mm.

the in Fig. 8 and supplemental movies.

Detection of horizontal planar surfaces was reliable. However, surface avoidance of vertical planar surfaces was extremely unreliable. Only two successful trials (defined by no premature triggering of the avoidance behaviour and correct avoidance when presented with a surface) were recorded, out of dozens of such trials.

3.5 Sensing other drones

The Mormoops prototype was always reliably detected by the Crazyflie observer, and both distance (Fig. 9) and orientation in the Crazyflie frame of reference (Fig. 10) could readily be identified. The reverse scenario was not the case, and no reliable estimates of the position of the Crazyflie by the Mormoops observer could be obtained due to signal being overwhelmed by self-noise. Adjustment of gains and more sophisticated filtering may improve detection capability.

The bearing to a secondary drone was always reliably extracted, with a standard error around 30 degrees. The distances were somewhat reliable out to around 20 cm, with a resolution of around 7cm at this distance. Resolution improved as the target drone got closer. At distances greater than one meter, estimates of distance could no longer be obtained with accuracy.

Our Mormoops prototype can thus be detected reliably and consistently from a Crazyflie observer when tethered. The reverse was not the case, largely because the Crazyflie is approximately two orders of magnitude quieter than the pro-

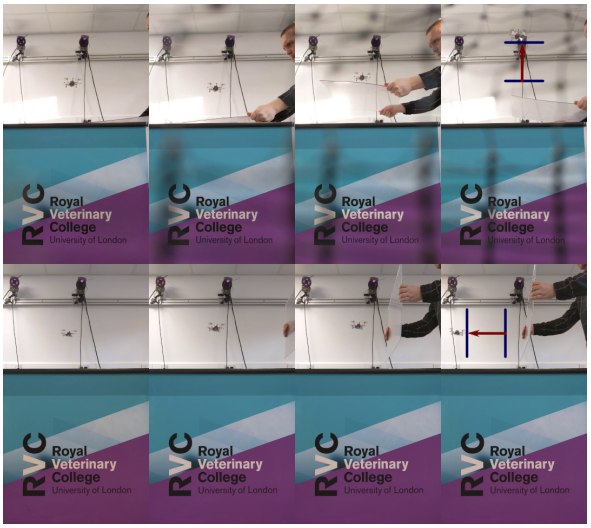

Figure 8: Still images from movie of surface avoidance behaviour. Above: horizontal planar surface avoidance as a surface approaches. Below: wall avoidance as a vertical planar surface approaches.

totype. No signal was therefore detected for our microphones when gains were configured to avoid clipping on the prototype. In short, the configuration for obstacle avoidance on the Mormoops was not compatible with detection of a secondray, much quieter, drone.

4 CONCLUSION

Detection of horizontal planar surfaces is very reliable and robust across modes, tethered, post free flight recording and in drones in flight engaging in surface avoidance. Detection of vertical planar surfaces is reliable while tethered, but far less reliable in the other testing modes. The detection of a larger, louder, more powerful drone from a smaller drone while tethered is possible, and the angle and distance of this drone can be estimated somewhat reliably via acoustic recordings. Detection of the louder drone might be improved by the implementation of motor denoising techniques[23].

For free flight surface avoidance, latency was an issue due to the long chain between sensor systems and flight control (recording to on board microcontroller, to flight control board, to radio, to flight control computer). A reduction in the latency from detection to avoidance could see this evolve into a practical system. At present the latency is around 250ms. Larger spikes in latency are also seen, and arguably the most pressing challenge for further use of the wall detection method is variability in latency as wall effect quickly renders a drone in an unrecoverable state. Reducing latency is achievable and would substantially improve drone sense-and-avoid performance.

These trials were conducted under good conditions, indoors and free from wind, both of which could negatively

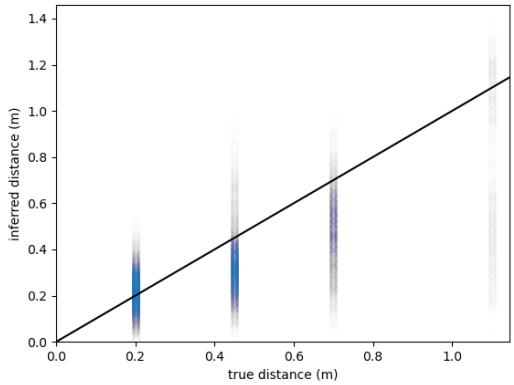

Figure 9: Performance of our machine learning approach when estimating the distance of a target drone from the sensing drone. The x-axis shows the actual distance (ground truth); the y-axis shows the estimate inferred from the microphone recordings and machine learning algorithm. Performance is reasonable for shorter distances, with the error at 20cm approximately 7cm. Shorter range recordings (not shown), suggest accuracy is improved as distance decreases. Above $20cm$ accuracy deteriorates with a bias towards underestimates although estimates of distance are possible and may be useful.

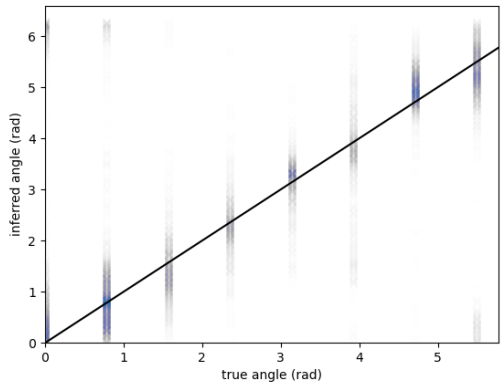

Figure 10: Performance of our machine learning approach when estimating the orientation of a target drone (angle relative to forward on the observer drone). The x-axis shows the actual orientation (ground truth); the y-axis shows the estimate inferred from the microphone recordings and machine learning algorithm. Performance is excellent, with a typical error around 30 degrees, although some variation is observed by angle.

http://www.imavs.org/

http://www.imavs.org/

contribute to performance of both surface detection and avoidance. It is unclear how robust a neural network based approach will be to the additional sounds that occur outdoors, including wind. Unlike physics based approaches which have predictable responses to new conditions (and which allow for those new condition to be incorporated into the physical model), neural network based approaches may not perform stably in conditions that are not in the training data. One approach to this problem would be to intentionally create noise in the training data to simulate wind or external noise in an effort to make the algorithm more robust. Wind would also likely make avoidance manoeuvres more difficult.

The multiplatform capability of our sensor, and sensor integration, packages mean that multi-drone operation with drone identification are achievable with the current system. We have high temporal resolution digital microphones using the specialised audio processing available on the Teensy should frequency range and audio processing be a limiting factor and we have analogue microphone arrays should spatial resolution and multi-direction sampling be critical.

A quiet, light drones capable of supporting these sensors like the Crazyflie could be used for future development should background motor noise be an issue, and a larger, more capable, but louder, like our Mormoops platform could be employed should extra sensors and equipment be needed for drone identification. This capability could be complimented by integrating acoustic surface detection with surface detection using other sensor modalities such as optical flow or pressure sensors.

This work demonstrates novel capabilities of surface detection via self-noise can be extended in exciting new directions including wider sensor integration and in-flight drone detection.

ACKNOWLEDGEMENTS

This work was supported principally by the Defence Science and Technology Laboratory (grant DSTLX-1000143653 to RJB) and, in part, by AFOSR European Office for Aerospace Research and Development (FA9550-19-1-7040 to RJB). We would like to thank Dr. Geoffrey Barrows of Centeye Inc. for useful conversations about drone design and configuration. We would like to thank William Wilson for his efforts in data collection. We would also like to thank the team at the Naval Research Laboratory, particularly Dr. Alisha Sharma, for useful conversations and feedback on the design of neural networks.

REFERENCES

[1] Abhishek Kasturi, Veljko Milanovic, Bryan H Atwood, and James Yang. Uav-borne lidar with mems mirror-based scanning capability. In *Laser Radar Technology and Applications XXI*, volume 9832, pages 206–215. SPIE, 2016.

[2] Joseph Conroy, Gregory Gremillion, Badri Ranganathan, and J Sean Humbert. Implementation of wide-field integration of optic flow for autonomous quadrotor navigation. *Autonomous robots*, 27(3):189–198, 2009.

[3] Toshiyuki Nakata, Nathan Phillips, Patrício Simões, Ian J Russell, Jorn A Cheney, Simon M Walker, and Richard J Bomphrey. Aerodynamic imaging by mosquitoes inspires a surface detector for autonomous flying vehicles. *Science*, 368(6491):634–637, 2020.

[4] Kaman Thapa Magar, Gregory W Reich, Corey Kondash, Keith Slinker, Alexander M Pankonien, Jeffery W Baur, and Brian Smyers. Aerodynamic parameters from distributed heterogeneous cnt hair sensors with a feed-forward neural network. *Bioinspiration & biomimetics*, 11(6):066006, 2016.

[5] Keith Slinker, Corey Kondash, Benjamin T Dickinson, and Jeffery W Baur. High-bandwidth and sensitive air flow sensing based on resonance properties of cnt-on-fiber hairs. *C*, 3(1):6, 2017.

[6] Meysam Basiri, Felix Schill, Dario Floreano, and Pedro Lima. Audio-based relative positioning system for multiple micro air vehicle systems. Technical report, 2013.

[7] Meysam Basiri, Felix Schill, Pedro Lima, and Dario Floreano. On-board relative bearing estimation for teams of drones using sound. *IEEE Robotics and Automation letters*, 1(2):820–827, 2016.

[8] Meysam Basiri, Felix Schill, Dario Floreano, and Pedro U Lima. Audio-based localization for swarms of micro air vehicles. In *2014 IEEE international conference on robotics and automation (ICRA)*, pages 4729–4734. IEEE, 2014.

[9] E Tijs, GCHE de Croon, J Wind, B Remes, C De Wagter, HE de Bree, and R Ruijsink. Hear-and-avoid for micro air vehicles. In *Proceedings of the International Micro Air Vehicle Conference and Competitions (IMAV), Braunschweig, Germany*, volume 69, 2010.

[10] Brendan Harvey and Siu O'Young. Acoustic detection of a fixed-wing uav. *Drones*, 2(1):4, 2018.

[11] DA Waters. Echolocation in air: biological systems, technical challenges, and transducer design. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 221(10):1165–1175, 2007.

[12] Jong Hwan Lim and John J. Leonard. Mobile robot relocation from echolocation constraints. *IEEE Transactions on pattern analysis and machine intelligence*, 22(9):1035–1041, 2000.

[13] Frederike Dümbgen, Adrien Hoffet, Mihailo Kolundžija, Adam Scholefield, and Martin Vetterli. Blind as a bat: Audible echolocation on small robots. *IEEE Robotics and Automation Letters*, 2022.

[14] Luke Calkins, Joseph Lingevitch, Joe Coffin, Loy McGuire, Jason Geder, Matthew Kelly, Michael M Zavlanos, Donald Sofge, and Daniel M Lofaro. Distance estimation using self-induced noise of an aerial vehicle. *IEEE Robotics and Automation Letters*, 6(2):2807–2813, 2021.

[15] Honeywell, 2011. Patent number EP 2 386 925A1.

[16] Hans-Elias De Bree and Guido De Croon. Acoustic vector sensors on small unmanned air vehicles. *the SMi Unmanned Aircraft Systems*, 2011.

[17] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[18] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

[19] Abien Fred Agarap. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018.

[20] Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*, pages 92–101. Springer, 2010.

[21] B. T. Polyak and A. B. Juditsky. Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.*, 30(4):838–855, July 1992.

[22] David W Scott. *Multivariate density estimation: theory, practice, and visualization*, volume 383. John Wiley & Sons, 2009.

[23] Patrick Marmaroli, Xavier Falourd, and Hervé Lissek. A uav motor denoising technique to improve localization of surrounding noisy aircrafts: proof of concept for anti-collision systems. In *Acoustics 2012*, 2012.

http://www.imavs.org/