Deep Neural Network used as a Supervised Image Classifier with Noise Recognition

F First Author **S Vinutha**, Second Author **P Brinda**, Third Author **Savier Soundaraj S** V

Abstract—— **A deep neural network (DNN) is an artificial neural network (ANN) with multipl e hidden layers of units between the input and output layers. Robot navigation means the robot's ability to determine its own position in its frame of reference and then to plan a path towards some goal location. In order to navigate in its environment, th e robot or any other mobility device requires representation, i.e. a map of the environment and the ability to interpret that representation. Noise recognition using ecosystem which helps in identifying environmental disturbance. A mobile robot is an automatic machine that is capable of locomotion. Machine learning which is evolved using pattern recognition and computational learning theory in artificial intelligence.**

***Keywords*— Computational learning theory, Locomotion, noise recognition, Pattern recogni tion, Robot navigation.**

I. INT RODUCT ION

Mobile robots automatically following the trail in the forest is a challenging and mostly difficult tas for the robots. Working on such difficulties is important becaus e its application in wildernes s mapping, s earch a nd rescue of human life's and wildlife's.

Micro aerial Vehicles (MAV) flying under the tree canopy. In order to follow the forest trail a robot has to perceive where the trail is also have to recognize the nois e in the environment if neces s ary and follow the trail. In this paper we include nois e recognition us ing ecosystem which helps in identifying environmental dis as ters . Bas e paper which implements machine learning approach to the vis ual perception to the fores t trail and s how the preview results on an autonomous quad rotor. The input to the machine is a monocular image from a forward looking camera which identifies the trail forward using the pattern recognition method. It is challenging even for humans to identify the trail in the forest as it is very narrow and uneven. Tracking the trail is challenging becaus e the appearance is not s imilar in all over the fores t and the boundaries are not s caled. Often the s tructure of the s cale may keep changing as it is not well maintained with cons trained length and width. The path might fluctuate as the time elaps e. The robot has to be developed in s uch a way that it recognizes the trail and follow the nois e according to the environment changes . Several previous works deals with the forest trail but which did not implement the nois e recognition which is implemented in this paper. Deep Neural Networks (DNNs) have recently emerged as a powerful tool for computer vis ion tas k like object clas s ification, biomedical image s egmentation. One of advantage of Deep Neural Networks (DNNs) is generality.
Nois e recognition helps in tracking the region which is not reachable by the humans. If incas e of any group of people and wild animals in and around the place who are in any kind of ris k can be res cued by us ing the robot nois e s ens ing. Once the nois e is s ens ed it then the robot proceeds in that direction and captures the image of that particular region from where the noise is been sensed. This robot can als o be us ed by the cops for the detection of anti-s ocial activity (terrorist) or wildlife security and protection.

A. CONT RIBUTIONS

The main contributions are:

A perception technique bas ed on Deep Neural Network which bypas s es the challenging characteris tic feature of a trail and nois e.

A datas et required for efficiently acquired on real world hiking trails trained and tes ted by the Deep Neural Network.

The noise recognition by inverse wavelet transform.

II. . PROBLEM FORMULATION

Hearing range des cribes the range of frequencies that can be heard by humans or other animals , though it can als o refer to the range of levels . The human range is commonly given as 20 to 20,000 Hz. Several animal s pecies are able to hear frequencies well beyond the human hearing range. Some dolphins and bats for e.g. can hear frequencies up to 100 KHz, elephants can hear s ounds at 14-16Hz, while s ome whales can hear s ubsonic s ounds as low as 7Hz (in water). Whereas a humanoid robot under real-world environments usually hears mixtures of sounds and thes e three capabilities are es s ential for robot audition, s ound s ource localization, s eparation and recognition of separation sounds . A microphone array is used along with a real-time dedicated implementation.

If the animals or humans are able to produce a nois e as sharp as clapping, then the mobile robot will follow to the

ISSN:2319-6890)(online),2347-5013(print)
20 May 2016

noise dense region. The robot will cons is ts of the hardware such as sensory circuit, microphone which can even sense e the low frequency sounds based on the sensors we can detect the sound up to 100-20m.

We us e the ultras onic s ens ors , the reflected s ound or "echo" is then received by the s ens or. Detection of the s ound generates an output s ignal for us e by an actuator, controller, or computer. The output s ignal can be analog or digital.

Ultras onic s ens ing technology is bas ed on the principle that sound has a relatively cons tant velocity. The advantage of ultras onic s ens ors is that they are affected by cons idering mois ture.

A humanoid robot under real-world environments us ually hears mixtures of s ounds, and thus three capabilities are es sential for robot audition; sound s ource localization, s eparation, and recognition of s eparated s ounds . While the firs t two are frequently addres s ed, the las t one has not been s tudied s o much. We pres ent a s ys tem that gives a humanoid robot the ability to localize, s eparate and recognize s imultaneous s ound s ources . A microphone array is us ed along with a real-time dedicated implementation of Geometric Source Separation (GSS) and a multi-channel pos t-filter that gives us a further reduction of interferences from other s ources . An automatic s peed recognizer bas ed on the Mis s ing Feature Theory (MFT) recognizes s eparated s ounds in real time by generating mis s ing feature mas ks automatically from the pos t-filtering s tep. The main advantage of this approach for humanoid robots res ides in the fact that ASR with a clean acous tic model can adapt the dis tortion of s eparated s ound by cons ulting the pos t-filter feature mas ks . Recognition rates are pres ented for three s imultaneous s peakers located at 2m from the robot. Us e of both the pos t-filter and the mis s ing feature mas k res ults in an average reduction in error rate of 42% (relative).

Ultras onic Sens or cons tructions

There are four bas ic components of an ultras onic proximity s ens or:

- ✓ Trans ducer / receiver
- ✓ Comparator
- ✓ Detector circuit
- ✓ Solid-State output.

Trans ducer receives echoes of thos e waves as reflected Off an object. When the s ens or receives the reflected e cho. The comparator calculates the dis tance by the emit-to-receive time frames to the s peed of s ound. The s ignal from digital s ens or indicates the pres ence or abs ence of an object in the s ens ing field. The s ignal from analog s ens ors indicates the dis tance to an object in the s ens ing field.

III. PERCEP T ION OF FOREST T RAILS

To s olve the problem we are us ing on extremely Challenging tas k called s upervis ed machine learning becaus e of wide trail of vis ibility in its s urroundings . We had the nois e recognition s ys tem to trail the machine to recognize the s ounds . The range varies from ultras onic to infras onic repres enting s uch large label datas et for covering trails of long dis tance.

A. DAT A SET

The hiker is equipped with three Head MountainAlong with s peakered monitor. This enables to have three views i.e. 180 degree view along with the nois e recognition s ys tem, each view overlaps each other to get a complete panorama image. The datas et is compos ed by the images and the nois e acquired by the three camera's and the nois e recognizer. The clas s ifier enables each of the image to its ground truth glas s . They were acquired while the hiker walked along the trail in a s traight ahead direction and in proportionally the other two cameras looking at left and right captures th e image.

The datas et is currently compos ed eight hours of 1920*1080 30fps video acquired us ing three cameras in the s pecified configuration of the hardware and covers approximately a dis tance of les s than 10km of the hiking trail and ranges of an altitude between 300m to 1,200m. It als o acquires the nois e us ing a microphone that ranges from the lowes t ultras onic frequency to the highes t infras onic frequency that includes human audible range.

Since, it would yield to a motion blur during day time, it would be auto corrected by the s ound dens ity rather than improving image qualities . During night times the image can be obtained by the nois e dens ity emitted from the environment rather than night image capturing cameras . To ens ure the datas et is repres entative not only of ideal, "pure" trails but als o frequency challenges obs erved in the real world. Synchronized GPS and compas s information has been recorded for s uch s equences but it is unus ed at the moment.

The datas et has been s plit into dis joint train ing frames for both image as well as nois e. The s plit was defined by carefully avoiding the s ame trail s ection appearing in both the training tes ts . The four clas s es are evenly repres ented in the datas et.

B. NOISE RECOGNIT ION

The nois e recognition cannot s elect a s tandalone s ound. The form of s ound is very dependent on the s ound environment of a s ignal that goes before the hiker, it is known that the wave form is a s mooth trans ition frame one s ound to another. The clas s ifier converts the nois e to wavelets by us ing invers e wavelet trans form algorithm.

ISSN:2319-6890)(online),2347-5013(print)
20 May 2016

Again the nois e is made into a datas et that is added along with the image datas et that is controlled and handled by the clas s ifier. Improving recognition quality in s tandard approach is as s ociated us ually with manipulation on the input s ignal, or s electing the convers ion and the improvement of pre -proces s ing. The clas s ifier has a s ignificant impact view output in the training s et of the clas sifier its elf. This is done to reduce the redundancy s ounds affecting the clas sifier. The clas sifier of nois e recognition implies a number of clas s es of recognitionentry given to each clas s es are a probability belonging to the given clas s . This fragment of nois e to s ound.

Us ually the s ound s ignal is broken into s mall pie ces- frames each of the frame is converted us ing a Fourier trans fer s uch as the wavelet algorithm.

C. DEEP NEURAL NETWORK FOR TRAI L PERCEPTION

The Deep Neural Network is an image and sound Clas s ifier a matrix of 3*101*101 neurons as the input and output neurons for image and the s ound matrix is defined during its creation due to varied value s ize of the s ound. The image matrix has a value becaus e it is clas sified us ing RGB.

The Deep Neural Network us es the wavelet to convert both image and as well as s ound to produce the datas et into two parts one for image and the other for s ound res pectively and individually been divided by the Deep Neural Network in its res pective modules.

Deep Neural Network us es GS training method to train
the images and GMM for training the s ound.

Drawbacks :

Ultras onic s ens ors mus t view a s urface (es pecially a hard, flat s urface) s quarely (perpendicularly) to receive ample
s ound echo. Als o, reliable s ens ing requires a minimum target s urface area, which is s pecified for each s ens or type.

While ultras onic's exhibit good immunity to backgroundnois e, thes e s ens ors are s till likely to fals ely respond to
s ome loud nois es , like the "his s ing" s ound produced by air hos es and relief valves.

Proximity s tyle ultras onic s ens ors require time for the trans ducer to s top ringing after each trans mis s ion burs t before they are ready to receive returned echoes. As a res ult, s ens or respons e times are typically s lower than other technologies at about 0.1 s econd. This is generally not a dis advantage in mos t level s ens ing and dis tance meas urement applications . Extended res pons e times are even advantageous in s ome applications . Trans mitted beam s tyle ultras onic s ens ors are much fas ter with res pons e times on the order of 0.002 or 0.003 s econds .

Ultras onic s ens ors have a minimum s ens ing dis tance.Changes in the environment, s uch as temperature,pres s ure, humidity, air turbulence, and airborne particles affect ultras onic res pons e.

Targets of low dens ity, like foam and cloth, tend to abs orb s ound energy; thes e materials may be difficult to s ens e at long range.

Smooth s urfaces reflect s ound energy more efficiently than rough s urfaces; however, the s ens ing angle to a s mooth s urface is
generally more critical than to a rough s urface.

IV. CONCLUSION

We trained a Deep Neural Network for vis ually Perceiving the direction of the trail bas ed on the nois e obs ervation. The s ys tem can perform better obs erver and can detect very accurately bas ed on the nois e frequency and dens ity. The s ens or can s ens e more than the human frequency that it can hear. Here it is eas y to capture the image and unders tand the s ituation bas ed on the image along with the nois e which is been produced which will give better utilization of the mobile robots.

REFERENCES

[1] Ales s andro Gius ti, Jerome Guzzi: "A machine learning approach to Vis ual perception of fores t trails for mobile robots".

[2] Google, "Maps trekker." www.google.com/maps /about/partners / Street view/trekker.

[3] M. Hutter et al., StarlETH: A compliant quadrupedal robot for fas t,efficient and vers atile locomotion. 2012.

[4] S. Ross, N. Melik-Barkhudarov, K. S. Shankar, A. Wendel, D. Dey, J. A.Bagnell, and M. Hebert, "Learning monocular Reactive uav control in cluttered natural environments ," in Proc. ICRA, 2013.

[5] A. J. Barry, A. Majumdar, and R. Tedrake, "Safety verification of reactive controllers for uav flight in cluttered environments us ing barrier certificates ," in Proc. ICRA, 2012.

[6] A. Briod, P. Kornatows ki, J.-C. Zufferey, and D. Floreano, "A collis ionres ilient flying robot," Journal of Field Robotics , vol. 31, no. 4, 2014.

[7] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progres s in road and lane detection: a s urvey," Machine Vis ion And Applications , pp. 1–19, 2012.

[8] E. D. Dickmanns and A. Zapp, "A curvature -bas ed s cheme for improving road vehicle guidance by computer vis ion," in Mobile Robots , SPIE Proc. Vol. 727, Cambridge, Mas s ., pp. 161–168, 1986.

[9] E. D. Dickmanns , R. Behringer, D. Dickmanns , T. Hildebrandt, M.Maurer, F. Thomanek, and J. Schiehlen, "The seeing pas s engercar 'VaMoRs -P'," in Proc. Int. Symp. on Intelligent Vehicles '94, Paris, pp. 68–73, 1994.

[10] E. D. Dickmanns , Dynamic Vis ion for Perception and Control of Motion. Springer, 2007.

[11] S. Thrun et al., "Stanley: The robot that won the darpa grand challenge," Journal of field Robotics , vol. 23, no. 9, 2006.

[12] C. Ras mus s en, Y. Lu, and M. Kocamaz, "Appearance contras t for fas t, robus t trail-following," in Proc. IROS, pp. 3505– 3512,2009.

[13] P. Santana, L. Correia, R. Mendonça, N. Alves , and J.Barata, "Tracking natural trails with s warm-bas ed vis ual s aliency,"Journal of Field Robotics , vol. 30, no. 1, pp. 64–86, 2013.

[14] A. Toet, "Computational vers us ps ychophys ical bottom- up image aliency: A comparative evaluation s tudy," IEEE Trans actions on PAMI, vol. 33, no. 11, pp. 2131–2146, 2011.

[15] A. Levin and Y. Weis s , "Learning to combine bottom- up and top-down s egmentation," in Proc. ECCV, pp. 581–594, 2006.

[16] C. Rasmussen, Y. Lu, and M. Kocamaz, "A trail-following robot which us es appearance and s tructural cues ," in Field and Service Robotics , pp. 265–279, Springer, 2014.

[17] A. Krizhevs ky, I. Suts kever, and G. E. Hinton, "Imagenet clas s ification with deep convolutio nal neural networks ," in Proc.NIPS, pp. 1106– 1114, 2012.

[18] D. C. Cires an, U. Meier, and J. Schmidhuber, "Multi- column Deep Neural Networks for Image Clas s ification," in Proc. CVPR, pp. 3642– 3649, 2012.

[19] D. C. Ciresan, A. Gius ti, L. M. Gambardella, and J.Schmidhuber, "Deep Neural Networks Segment Neuronal Membranes in Electron Micros copy Images ," in Proc. NIPS, pp. 2852–2860, 2012.

[20] J. Schmidhuber and R. Huber, "Learning to generate artificial foveatrajectories for target detection," International Journal of Neural Sys tems ,vol. 2, no. 1 & 2, pp. 135–141, 1991.2014.