Sound source localization through shape reconfiguration in a Snake Robot

Sriranjan Rasakatla, Bipin Indurkhya, Ikuo Mizuuchi, Kazuki Sekine, Hijiri Akahane

Abstract—**This paper describes a snake robot system that uses sound source localization. We show in this paper as to how we can localize a sound source in 3D and solve the classic forward backward problem in sound source localization using minimum number of audio sensors by using the multiple degrees of freedom of the snake robot. We describe the hardware and software architecture of the robot and show the results of several sound tracking experiments we did with our snake robot. We also present biologically inspired sound tracking behavior in different postures of a biological snake robot as "Digital Snake Charming".**

Index Terms— **snake robot, sound source localization**

I. INTRODUCTION

SOUND source localization problem is a classic problem and many robots in the past have tried to implement SSL (sound source localization) in their architecture. Mostly if it is a humanoid robot then they used binaural hearing and in case of some other robots like UAVs, mobile robots etc they have used an array of 4 or more microphones. The idea was to localize a sound source using either binaural method with HRTF(head related transfer function) using ILD(interaural level difference) or IPD(interaural phase difference) and TDOA (time delay of arrival method). We here present our attempt to do SSL in a hyper-redundant snake robot. Till today there have been no snake robots which use microphones to bio-mimic the biological snake and explore its sound reactive behaviours. Snake charmers in India have been demoing this ability in street shows with biological snakes where they blow a musical instrument called "Pungi"(fig.1) to raise the attention of the snake so that it raises its hood and tracks this sound. Here we have studied different sound source localization works done in the past, the use of array of microphones on robots and we tried to implement for the first-time sound source localization with tracking based biomimicking behaviour and localize a sound source through reorganization of the hyper redundant structure of the snake robot using minimum number of microphones. We also coin the term "Digital Snake charming" to show this behaviour.

We also show how forward-backward problem in SSL is solved using our snake robot. Nobody did sound source localization on a snake robot before and gave snake robot a perception that is auditory in nature. Giving sound perception to snake robot not only solves the sound source localization problem in a resource efficient manner but also opens new interesting sound tracking behaviours as seen in real snakes. We solve the forward backward problem by getting the snake robot into a piecewise gait posture and make it scan 360 degrees with its change of posture and solve localization problem, the forward backward problem and also show biomimicking behaviours. Right now, tracking multiple sound sources is out of the scope of this paper, that in itself is a different problem. In some humanoids this was possible before with pinnae like ear structures and was not very accurate, we are +/-5 degrees accurate in our solution. The paper demonstrates this in the form of a system design, software and mechanically built solution. Forward backward problem which is commonly solved using pinnae shaped sound receivers in some papers and naturally in humans where the ears are shape so is solved here in using pose reconfiguration using minimum microphones (mics) in L shape. Earlier works required 4 or 8 or 12 localization mics but our strategy to localize the sound source works with a minimum of 3 microphones through shape reconfiguration to L shape along the 4 quadrants.

Snake robots have been widely researched by CMU Biorobotics Dr.Howie Choset [19,20], Dr. Shigeo Hirose[22] from Ti-Tech, EPFL[21] etc showcasing their capability to do all terrain locomotion both on land with rocky environments, through tight space like fitting through pipes and crevices and also swim in water. Our idea in this paper is to fit the snake robot with microphones so that we can listen for sound of the trapped survivors in a disaster scenario and get a location information as to where the sound is coming from. This we

Sriranjan Rasakatla is with Mizuuchi lab, Applied Mechanical systems, Department of Mechanical systems Engineering, Tokyo University of Agriculture and Technology. TUAT 2-24-16 Nakacho, Koganei, Tokyo, 184-0012 Japan. (email: infibit@gmail.com)

Bipin Indurkhya is with Department of computer Science, Jagiellonian University, Krakow, Poland. He was also with IIIT-H,India Cognitive science Department. (email: bipin@agh.edu.pl, bpin8@gmail.com)

Ikuo Mizuuchi is with Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology. TUAT 2-24-16 Nakacho, Koganei, Tokyo, 184-0012 Japan. (email: mizuuchi@cc.tuat.ac.jp)

Kazuki Sekine is with Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology. TUAT 2-24-16 Nakacho, Koganei, Tokyo, 184-0012 Japan.

Hijiri Akahane is with Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology. TUAT 2-24-16 Nakacho, Koganei, Tokyo, 184-0012 Japan.

believe will help in search process of a search and rescue scenario and thus help locate survivors faster.

II. RELATED WORK

We discuss about binaural array and multiple array microphone systems. We discuss about their computational aspects of sound source localization. Then we outline references of papers which used sound source localization in robots. We also discuss how sound source localization problem evolved over the years and some classical problems in localization like the forward-backward problem for which we provided a solution using a snake robot (fig.2) in this paper.

A. Binaural and Multiarray microphone based localization

Jean M.V et al[1] discuss two robots Robita and SIG humanoid where two stereo microphones have been used. One pair is for capturing external sound and the other is for capturing internal noise because of motors of the humanoid. They say it is difficult to match human hearing because we have a shadowing effect because of our ear structure. So, they used an 8 array microphone system on a pioneer robot to localize in 360 degrees. Another paper by Inkyu[2] et al mentions about microphones arranged in the form of a cube. Generally to be able to compute time delay of arrival to calculate the sound source direction one needs more than a single frame of data but here they were able to take a single frame of sound, used a reflection aware setup so that the reflected wave of sound is generated and is traced back from the cube of microphone array to get the sound source location in 3D. Shengkui[3] et al mentioned their 8 microphone array sound source localization implementation on Olivia robot which enables the robot to face/orient towards the face of the speaker even if the user is out of the field of vision of the camera and even in poor lighting conditions. This is one advantage of using audio perception for search and rescue robots like in our snake robot because under poor lighting conditions also we can detect a sound source which could be the help cry of a possible survivor. Their robot HARK uses 8 microphone robot audition system using steered beam forming. Some humanoid robots like HRP-2 combine both audio and vision to detect speech events from vision. Another paper by DeForge[4] discusses about sound source localization using software-based mapping of high dimensional interaural data. Interaural level difference and interaural audio differences using head related transfer function has been one of the most famous methods to localize sound sources in humanoid robots. Hong[5] et al mentions in his paper that in 2006 Honda performed multi source real time tracking by embedding 64 microphones in the room and 8 microphones in the robot's head. 8 number of microphones appears to be a popular choice for many SSL systems but in resource limited application or damage to the robot we show how the same can be achieved using 2 or 3 microphones but with the help of multiple degrees of freedom configuration of our snake robot in this paper. Another implementation was from Canada in 2007 where they used beam forming on a robot with 8 microphones to avoid obstacles. Hara in 2008 says 8 channel microphones had better localization accuracy in REEM-A humanoid.

Daniele[6] et al mentions his work on using 4 array microphones in his multi rotor UAV with beam forming based acoustic source localization. The idea was to suppress the noise in audio due to rotors and get sound localization information of point of interest. This paper[7] presents a micro phone array design and how they optimize it for beam forming purpose. The microphone array here has 64 microphones arranged in the shape of a sphere with 350mm diameter designed to be mounted on a mobile robot with omni directional sensitivity in azimuth and elevation. They have tested the microphone array in challenging environments and with different pressure sound sources and have provided their results. The performance of beam forming improves with the increase in the number of microphones and even their robustness to different pressure sound sources increases. In the paper [8] for generating the training data set the robot HEARBO used an audio setup which was equipped with 8 and 16 channel microphone array. A 7 channel Microcone microphone array was also used to get the impulse responses.

B. Search and Rescue snake robot systems

Search and Rescue and inspection go hand in hand where one has to search through tight spaces using a camera. In the paper by Yoshiaki[10] et al they mention about inspecting a pipe and generating maps using ultrasound. At the opening of the pipe these is a speaker and in the head of the snake there is a microphone. The snake also has an IMU, by knowing the time of flight and length of the path taken by the sound the snake robot is able to generate a map in the GPS denied pipe, it is to be noted that this is not SSL. Odometry in pipe like environment is also not accurate due to slippage. The accuracy of map generated was more than 68 percent here by snake robot using sound based online localization method. Snake robots are especially suitable for moving in pipe like environment because they can twist their multiple DOFs and generate a pipe climbing gait. P. Thanu [11] et al says that Search and rescue robots these days as the demand need to be fitted with advanced sensors, reliable communication equipment so that they can help the rescue workers in rescuing survivors buried under the debris. USAR (urban search and rescue) typically has a dynamic environment and the robots need to have a sense of sovereignty to cope with the dynamic conditions. The snake robot by Dr.Gavin miller uses infrared range finding sensors, camera and flex sensors to sense its environment. Humming bird is an advanced intelligent UAV developed at University of Oxford that has a camera and works in an autonomous way. The USAR robots need to have features like reconnaissance, search functions, rubble removal and victim extraction and telepresence. This paper also presents a survey of human detecting sensors like Binary sensors, SpO2 sensors, CO2 sensors, thermal sensors, sound sensors like microphones, thermo piles, vision sensors and radio etc.. The robots needs to have a secure and low energy consuming/ energy efficient wireless network communication equipment. Generally when the robot goes out of the sight of the operator their location is lost but by using Zigbee sensors networks and other sensing modalities for wireless networks like signal strength and 4G we

can get location information of the robot. Target tracking is one of the important features discussed in this paper which is of value in USAR scenarios. In the chapter from a book by Fumitoshi [12] et al they discuss about both wheeled and non-wheeled snake robots they developed for inspection of narrow crevices of buildings and insides and outsides of the pipes.. So snake robot especially finds application to fit through tight spaces as it can change its degrees of freedom and when it comes to narrow spaces in an urban search and rescue situation, it is an important application for snake robots. So, our idea is to combine the ability of sound-based localization and use it in snake robots for the future purpose of search and rescue, the first step towards this task we identified is to equip the snake robot with microphones and study various possible biological, bio mimicking and anti-biological behaviours.

C. More sound localization works

Kazuhiro[13] et al in paper says that in the real word when capturing audio data there will be many sources of noise in the environment, reverberations and noise from the motors of the robot. This paper uses active direction pass filter, interaural phase difference, interaural intensity difference to localize the sound. Human tracking and sound source localization is implemented here in this paper in the upper torso of a humanoid robot. Yosuke[14] et al in their paper mentions that localization was implemented in the humanoid robot Robita. Here microphones have been arranged on a 2DOF head neck. Here stop-look-act form of tracking is implemented. This is one of the seminal works in which sound localization system using binaural audio has been setup on a 2DOF movable head neck system. Sound direction here is estimated using spectral properties and HRTFs. Francois et at [15] says that sound localization becomes difficult where there are noise sources in the environment. It is difficult to estimate the direction from where the sound is coming from the intended source when there is noise. So, this paper describes a binary weighted frequency mask on generalized cross correlation with phase transform. We in our Snake robot Z3(fig.1) use a Fast Fourier transform in our detection system that differentiates between predator sound and sound from prey and uses this frequency domain info in multiple sound source tracking and invoking the appropriate gait behavior in moving the snake robot. SSL can be used to give the location and number of sound sources which later can be used for sound separation. Nakamura[16] et al says singular eigen vector decomposition methods have been used earlier but its performance can be improved against noise and high power using generalized eigen vector decomposition. John C Murray[17] et al in their paper describes a sound localization system which uses only 2 microphone to closely replicate the mammalian hearing system. They also used audio cues to help in localization. Mammals can sound localize with an angular resolution of 1 deg and azimuthal resolution of 5 degrees. This is because the audio cues like interaural time difference (ITD) and interaural phase difference (IPD) are encoded in the lower brain stem regions. A multi-processing model has been used to do the sound tracking here. We in our snake robot use different processing pipelines for gait, sound localization and IMU heading calculation to achieve better sound localization performance. Looking at all this we chose minimum mic

configuration and used near real time cross-correlation method to sample and estimate the global minima to find the sound source. Xavier[18] et al discusses about non arbitrary shapes to do sound source localization using time delay of arrival estimate. But none of these papers show if by changing multiple degree of freedom where the microphone's geometric positioning is reconfigured to get a better estimate of the sound sources position or direction and thus the localization. This is especially of good value as snake robots are multi DOF reconfigurable. Their geometric analysis helps in determining a position of mics in geometric space with global optimization that corresponds to the sound source being localized. Such TDOAs determined are called feasible sets which correspond to a unique positioning in the source space.

Figure 1: This is a musical instrument call the Pungi used by the Snake charmers in India.

III. System Description

A. Hardware description

The snake robot Z3(fig2, fig.5 shows hardware schema) consists of 12 servo motor modules using high torque HiTech7995 TG 34kg-cm motors. Each motor under loaded condition can consume a total current of 3.5A. The 12 modules are linked successively with each other in an orthogonal fashion. The linkage is done through custom made servo brackets using aluminium sheets through sheet medal bending. Before the sheet metal bending process, the footprint of the bracket is made using a CNC machine. As they are orthogonally linked, 6 servo modules have their axes parallel to the ground and the rest of the 6 servo modules have their axis parallel to each other by perpendicular to ground. This way a 3D skeletal structure of the snake robot was built. After this the modules are covered with foam rubber adhesive skin. The foam rubber skin helps in giving traction from ground to the snake robot and also helps in absorbing the impact when the snake locomotes on the ground. A 40A 7.4V adjustable power supply is use. Under different postures of the snake robot, we observed a maximum current between 5A to 10A for gaits like caterpillar crawling and holding a static position in raised hood mode. The servos are analog that means they cannot be networked onto a single BUS like RS-485 or TTL. The reason we chose analog servos is because if we use networked servos like Dynamixels from Bioloid company we observed that there is bus delay in servo command packet transmission and only one servo can be commanded at an instant, whereas in analog servos all the servos can be simultaneously controlled at a time using pulse width modulated signal on each port pin of the microcontroller. The microcontroller used is Teensy 3.2 Arduino. We are using an inertial measurement unit in the head to get the absolute pose estimate of the robot. We know the commanded angle of each

servo arm (eventually each servo without a mechanical obstruction almost reaches the commanded position with a back lash error in servo output shaft and skeletal structure) and this relative to the head we will know the angular position of each of the module and thus the entire pose of the snake robot can be estimated. The IMU is connected over the I2C bus to the microcontroller. There are 2 pairs of MEMS microphones in the head and there are 3 microphones along the middle of the body of the snake robot. We are also using a Zigbee module in the head of the snake robot and one receiver module at the laptop computer end. Zigbee's use will be described in the follow section. The micro-controller and the microphones are powered using an off the shelf DC-DC buck which reduces the input voltage from 7.4V to 3.3V which is the compatible voltage for the MEMs mics, microcontroller board and Zigbee.

Figure 2: Snake robot Z3 with foam adhesive skin

The head of the snake robot was 3D printed(fig.4). An IMU was assembled on the top of the head along with 4 mics on the top, bottom, left and right faces each of the 3D printed cube. The mics were assembled such that the glue which holds the mics to the face does not dampen the sensitivity of the MEMs mics. Earlier we noticed that gluing directly at the back of the MEMs mic PCB reduced the sensitivity to $1/10^{th}$. So, we just glued at the point of the signal wires to the mic so that the microphone is left free to vibrate similar to how human ear drum is left free to vibrate by its biological design. These are omnidirectional mics and thus no sound shadowing effects like human ears were there to be taken advantage off. All the wires going into the head of the robot can be seen as similar to nerves of human body converging into the head from the neck and below.
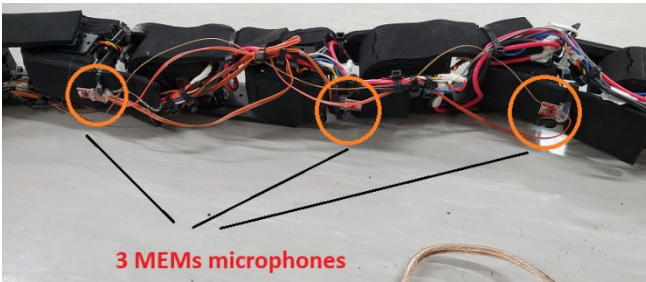

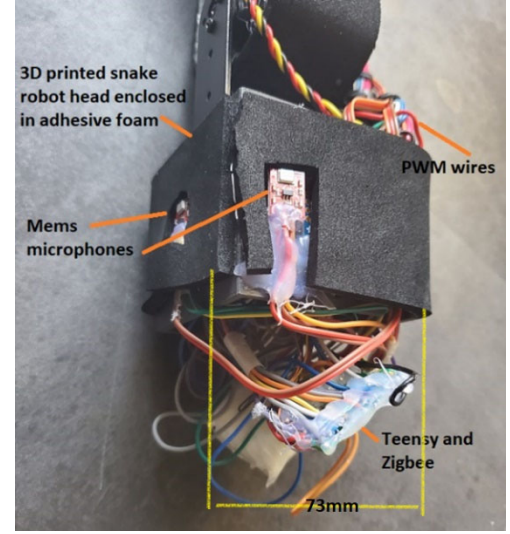Figure 3: 3 MEMs microphones the only ones used for 360 localization and forward backward problem.

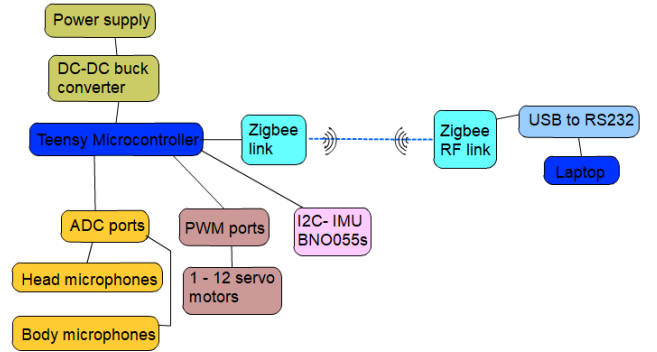Figure 4: Head module of the snake robot.

Figure 5: Hardware Schematic

We use 3 MEMs microphones(fig.3) on the middle portion of the body of the snake robot used for 360-degree localization of the sound source.

B. Software description

The snake robot has some functional modules running in the firmware that sits inside the Teensy microcontroller. All the computation runs onboard the snake robot head unlike our earlier snake robots (Snake P3[9]) whose gaits were calculated remotely on a computer and sent to the snake robot controller wirelessly. One sub routine is the sensory perception layer which gather MEMS microphone audio data from all the 7 microphones on individual analog to digital converter ports. This sensory layer also gets head orientation from the IMU so that the posture of the snake can be mapped. There is a sound source localization module function which calculates the direction of the sound coming from a pair of microphones each. The sound localization module uses cross correlation function which runs at 50Hz. The localization data coming from the pair of microphones are interpolated to get an actual and a valid location of the sound source relative to the position and orientation of the snake robot. We have also coded two Kalman filters one for each pair of microphones along the X axis and Y axis.

Here is the pseudo code for the 1 Dimensional Kalman filter.

mea3 = micpos3;

kalman_gain3 = err_estimate3/(err_estimate3 + err_measure3);

*current_estimate3 = last_estimate3 + kalman_gain3 * (mea3 - last_estimate3);*

*err_estimate3 = (1.0 - kalman_gain3)*err_estimate3 + fabs(last_estimate3-current_estimate3)*q3;*

last_estimate3=current_estimate3;

micpos3k = current_estimate3;

"micpos3" is the initial microphone pair localization value. "micpos3k" is the estimate value after the 1D Kalman filter is applied. This value is robust to noise and jitter in the microphone signal. We also use a threshold filter at the ADC to prevent false noise capture leading to ambiguous localization values. err_estimate is the average error calculated after applying the Kalman gain.

Another routine captures microphone samples on multiple ADC ports in parallel for multiple microphones at minimum 100Khz. Then we compute cross correlation between the two microphone signals. Now cross correlation is an expensive function so we used minimum of sum of least squares after shift through time axis of the captured array of microphone data

*sum[j] = sum[j] + ((a1[i + j - mid] – a2[i]) * (a1[i + j - mid] – a2[i])), where a1 and a2[] represent captured mic data.*

A pair of Zigbee modules is used to communicate between the snake robot and the laptop computer. Over this wireless link we send information like computed sound localization values, head IMU yaw, pitch, and roll values so that we can use this to update our remote visualization user interface software. We also send some condition flags to update the pose of the robot and sound source detection confirmation flags as well.

We also use a FFT (fast fourier transoform) filter to eliminate other environmental noise sources but detect only the human sound through the way of bandpass FFT.
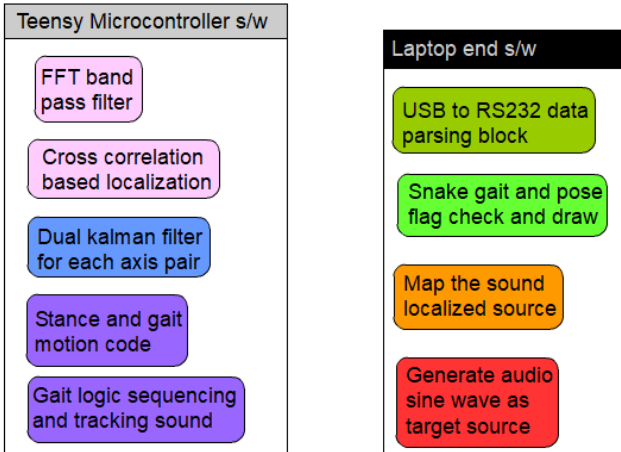

Figure 6: Software schema

IV. EXPERIMENTS AND RESULTS

A. Digital Snake Charming/ Head tracking

Biological snakes like the "Ophiophagus Hannah- King cobra" snake exhibits locking and tracking a particular object of interest or danger through vision and sound. Snake charmer's always use this to raise the attention of the snake using their musical instrument Pungi. We bio-mimic this and show digital form of snake charming here called "Digital Snake charming". Christian Christensen[23] a biologist from Aarhus University in Denmark attached electrodes to the snake's brain and found that the neural activity peaked when there was an auditory stimulus. They found out that the skull of the snake is very sensitive and can pick up vibrations caused by the sounds traveling in air. So that widely accepted myth that snakes are deaf is wrong and we thus want to use microphones in our snake robot and give it ears.

Earlier it was thought that the biological snakes do not have ears and thus cannot sense sounds but can sense vibrations with their bodies coming from the ground. But, a recent work shows that the snake behind their head have an auditory sensory system that are like ears hidden under the skin on either side of their head. So potentially they can track sounds as well. To mimic this capability, we have equipped the snake robot's head with 4 microphones arranged with a phase gap of 90 degree. This gives our snake robot the ability to track the sounds and follow the target by swaying its head in 3D (fig.7 and fig.8).
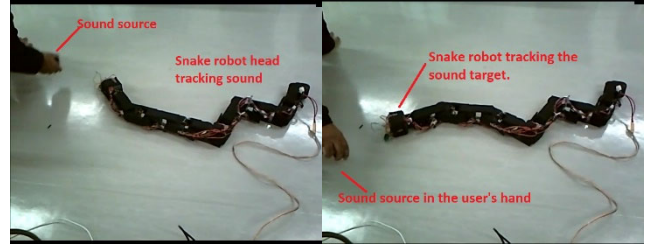

Figure 7: Snake robot first half of the body tracking the sound source in user's hand. The sound source is a Bluetooth speaker playing a sinusoidal sound waveform.
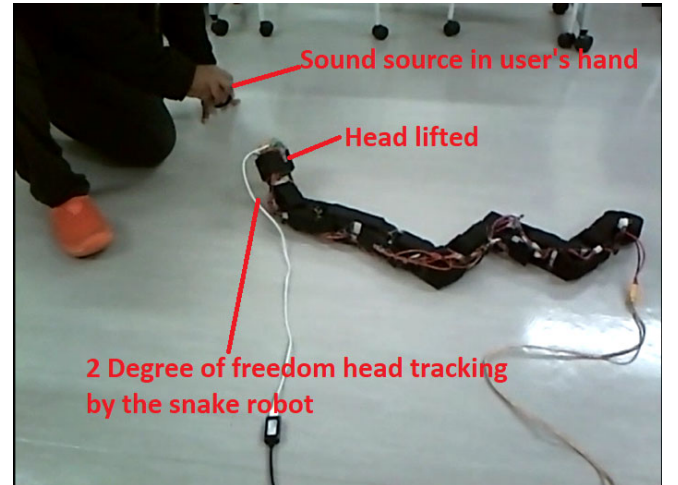

Figure 8: Head lift and track by snake robot with sound source in user's hand.

B. Hood raise and track

In this mode the snake robot takes a stance like the biological "Ophiophagus Hannah-king cobra" snake and tracks the auditory target. In this mode balancing the weight of the portion of the body is difficult and so we had to vary the link lengths that raises above the ground and take a backward stand such that weight of the raised portion is supported by the curled tail on the ground. Even the real biological cobra snake takes support from the curled part of its body and the tail to raise its hood (fig.9).
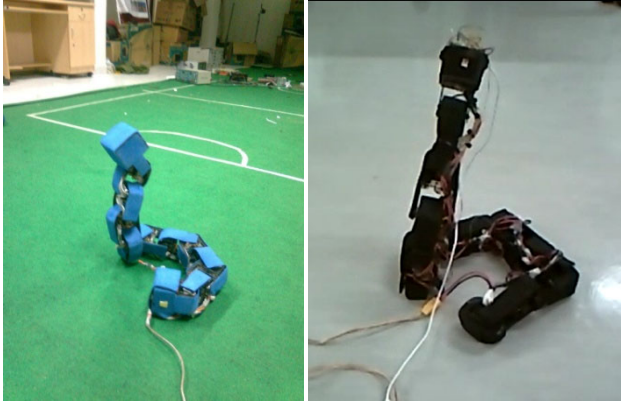

Figure 9: Hood raise posture of Snake P3[9] (left) and Hood raise for Snake Z3 (right) and listen to sound source.

C. Snooping

The biological snake always does not take a raised hood stance. It takes a raised hood stance when it wants to warn the human or animal that is trying to cause it harm. Sometimes the snake only lifts a part of its body and looks around generally when the snake is on the move and exploring. This is the classing snooping action which was also integrate a hybrid snake-wheeled robot for the purpose of search and rescue where the wheeled portion of the robot goes to a target site and the snake arm tried to look for objects of interest with a camera. But we have integrated a audio snooping capability in our snake robot where a part of the body is above the ground in the air and the rest of the part is curled in a zig zag /half rounded fashion to get some support(fig.10). The footprint is maintained during the posture transition in one half of the snake robot's body to avoid flipping or rolling on one axis.
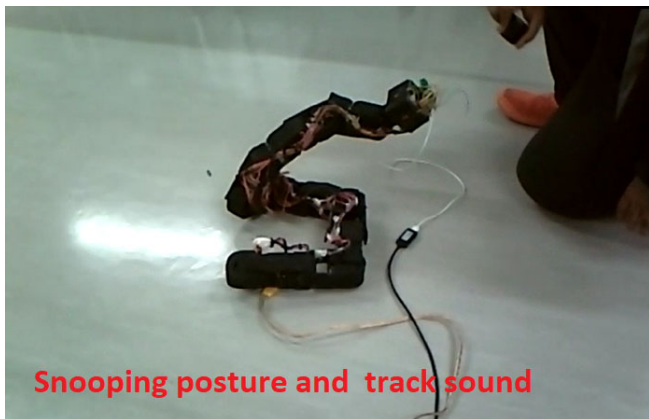

Figure 10: Snake tail curl and half body lifted above the ground like a snooping posture and track sound. The snake robot can

also lift its tail and get into a rattle position but still track the sound in 3D.

Due to one particular coding error in sound tracking to invoke the gait we observed the snake robot would move in unpredicted and random ways and it appeared as if a real biological snake is being handled by a person.

D. Track and Move

In this mode of gait, the snake robot loops in cycles to scan for auditory source with its head and then curve its body appropriately to trace a curved path. The direction of motion is governed by the curvature of the snake's skeletal body which is formed by the servos with their axis perpendicular to the ground termed vertical servos. The horizontal servos (which are the servos whose axis of rotation is parallel to the ground) execute a differential curve namely the caterpillar [9] gait here. This motion scenario is as if the snake robot is searching for prey/survivor in a search and rescue scenario and senses the sound and then moves towards it. Such behaviour has been observed in biological snake. Here in our snake robot we use stop-hear-act approach as seen in some animals like squirrels etc. Stop-hear-act allows us not only to eliminate motor noise but to also get the stance of the snake robot like the raised hood position of a biological snake to listen to the sound, we also turn the motors at low speed.

E. Localization through shape reconfiguration

Here is the idea behind designing the snake robot with microphones. Generally, if one uses two omnidirectional microphones then one can use Time delay of arrival and localize along on axis only. But if the object is above or below the line joining the two mics then one cannot differentiate if the source of sound is above or below that line. This is called the classical forward-backward problem in sound source localization. Some have increased the number of mics to 4 or more like have provided an array of mics along the 3 axis to solve this problem, some researchers have made the mic directional using pinnae (ear like structures), some have used directional microphones etc. For this differentiation we need to add a 3rd mic that is along the perpendicular to the line joining the two mics, but the problem is we cannot identify all the four quadrants or in 3D with just 3 mics, we will need 4 or more to achieve better localization accuracy. We mitigate this problem by using only 3 mics but since the snake robot is a hyper redundant structure and has the advantage that its degree of freedom can be controlled, we used this strategy. This solution uses the snake robot's mechanism as an advantage to minimize the mics. Compared to wheeled robots it is difficult to estimate the position kinematically and predict its position after some gait execution because like all other non-wheeled snake robots this robot also slips. If is a wheeled robots it is possible to roughly estimate its position with the use of encoders indoors or use SLAM based navigation, but it is not possible to mount some sort of encoding sensors on snake robot to estimate its slippage. However, if used outdoors the snake robots' position can be estimated using GPS and relative posture of its body can be estimated using servo encoders (not wheeled encoders) and indoors if we use optical motion tracking system or external camera-based tracking system it is possible. SLAM based navigation might also be possible on the snake robot but SLAM

for snake robot is out of the scope of this paper. We can get the snake posture information from the IMU in the head and in the L shape the snake other half of the body is almost 90 degrees to the first half. So even though there is slippage and snake position information in world frame is erroneous, the posture information can be estimated well with relative angles of the snake servo modules thus its body in the world frame and then the sound source localization with respect to the revised posture of the snake body can be found. Here we estimate the global pose of its head using IMU and snake relative pose of its successive modules using motion commands. Use of onboard encoders in the servo modules or the use of smart servo modules will give more accurate posture information, but it is not needed for the current sound source localization application. The idea is we roughly need to get the snake form an L shape and make this L shape rotate in 360 degrees to calculation localization information as it undulates out of L shape to line shape and gets back to L shape (fig.11, fig.12 and fig.13,fig 14).

The key idea is to place one mic in the middle of the snake robot's body which becomes the origin and orient the other two mics one each along X axis and Y axis. To align these mics along these axes we pivot the body about the origin mic module and then make an L shape with each mic on one each of the two legs of L shape. Then we do cross correlation based TDOA estimation. Then the snake robot is commanded to form a straight line which is the normal shape. By this time the snake robot would have rotated by 60-90 degrees on the ground. The exact angle by which the snake robot's body has turned cannot be told accurately in world reference frame because of slippage and backlash however a rough estimate can be done using the IMU in the head. Then the robot is commanded again to form an L shape and localize. The transition from one L shaped posture to another L shape postured which is expected to cover another quadrant is also not perfect because of slippage and lack of snake robot kinematic estimation under slippage. Pose transition was also learned after repeated trials as to how the snake robot's body gets into a posture destabilizes and then gets into the final required posture. Then relaxed to normal shape and this is repeated till the IMU values report a full 360 degrees turn completing one revolution. This mixed piece wise gaits are sufficient to scan the surrounding in full and thus we can accurately track where the target sound source is and thus solve the forward back problem and SSL using minimum mics as possible. A demo of this work in video is giving in the link [24].
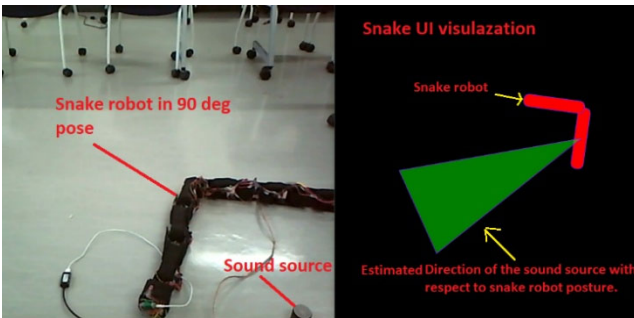

Figure 11: Figure of actual Snake robot in L shape pose with sound source., the user interface showing the direction of the source calculated/estimated with respect to the snake robot.
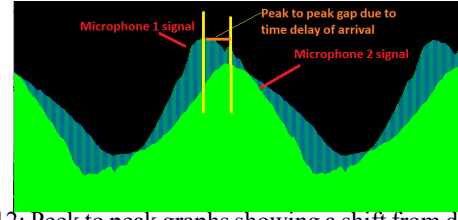

Figure 12: Peek to peak graphs showing a shift from data received in Mic1 and Mic2 as a result of time delay of arrival.

Figure 13: Snake robot not in L shape, speaker to left and right of the snake robot. Localization graph for both left and right shown
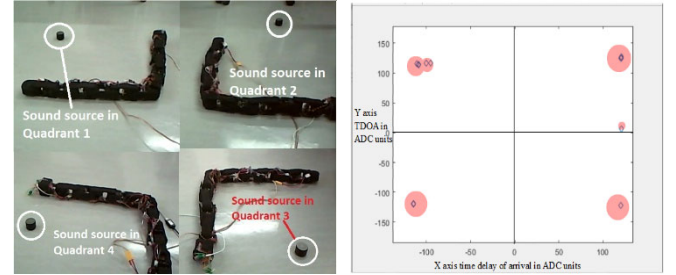

Figure 14: Shows the non-ambiguous of sound source detection in 4 quadrants, with only 1 false positive among 25 readings. The radius of circle indicates the density of similar reading.

The sound source when placed in left or right of snake when the snake robot uses only 2-3 mics (fig.13) and is straight, gives an ambiguous solution but when the snake robot takes an L shape and localizes with one pair of mics along X axis L leg and Y axis L leg (fig.14) gives a unique solution and thus solving the ambiguity in solution. This is the proof that using minimum mics but by posture reconfiguration we can get better results in SSL.

V. CONCLUSIONS AND FUTURE WORK

We showed that by using just a minimum of 3 mics and using pose reconfiguration in multiple degrees of freedom of the snake robot we could localize with reasonable accuracy of +/- 5 degrees. We also showed interesting sound tracking behaviour mimicking the biological snake. In the next iteration we plan to use a differential gait to localize more naturally with differential gaits. But this later attempt would need more microphones and a better robust snake robot hardware with networked motors and higher torque. We will also expand our work on track and move feature in the next paper The main objective of this paper was to show that using posture reconfiguration which is natural to multi DOF robots we can do 3D sound source localization using a minimum number of mics. Other works used pinnae and an array of mics which were more than 3 in number to localize. Thus, our system is more resource minimal, resource effective and is a novel way of sound source localization using multiple DOFs. We also solved the classic

forward backward problem found in sound source localization system through our snake robot solution approach.

REFERENCES

[1] Valin, J-M., et al. "Robust sound source localization using a microphone array on a mobile robot." *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*. Vol. 2. IEEE, 2003

[2] An, Inkyu, et al. "Reflection-aware sound source localization." *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018.

[3] Zhao, Shengkui, et al. "A robust real-time sound source localization system for olivia robot." *2010 APSIPA annual summit and conference*. 2010.

[4] Deleforge, Antoine, and Radu Horaud. "2D sound-source localization on the binaural manifold." *2012 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, 2012.

[5] Liu, Hong, and Miao Shen. "Continuous sound source localization based on microphone array for mobile robots." *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010.

[6] Salvati, Daniele, et al. "Beamforming-based acoustic source localization and enhancement for multirotor UAVs." *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE, 2018.

[7] Sasaki, Yoko, et al. "Spherical microphone array for spatial sound localization for a mobile robot." *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012.

[8] Yalta, Nelson, Kazuhiro Nakadai, and Tetsuya Ogata. "Sound source localization using deep learning models." *Journal of Robotics and Mechatronics* 29.1 (2017): 37-48.

[9] Rasakatla, Sriranjan, and K. Madhava Krishna. "Snake P3: A semi-autonomous Snake robot." *2010 IEEE International Conference on Robotics and Biomimetics*. IEEE, 2010.

[10] Bando, Yoshiaki, et al. "Sound-based online localization for an in-pipe snake robot." *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2016

[11] Thavasi, P. Thanu, and C. D. Suriyakala. "Sensors and Tracking Methods Used in Wireless Sensor Network Based Unmanned Search and Rescue System-A Review." *Procedia engineering* 38 (2012): 1935-1945.

[12] Matsuno, Fumitoshi, et al. "Development of tough snake robot systems." *Disaster Robotics*. Springer, Cham, 2019. 267-326.

[13] Nakadai, Kazuhiro, Hiroshi G. Okuno, and Hiroaki Kitano. "Real-time sound source localization and separation for robot audition." *INTERSPEECH*. 2002.

[14] Matsusaka, Yosuke, et al. "Multi-person conversation via multi-modal interface-a robot who communicate with multi-user." *Sixth European Conference on Speech Communication and Technology*. 1999.

[15] Grondin, François, and François Michaud. "Time difference of arrival estimation based on binary frequency mask for sound source localization on mobile robots." *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015.

[16] Nakamura, Keisuke, Kazuhiro Nakadai, and Gökhan Ince. "Real-time super-resolution sound source localization for robots." *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012

[17] Murray, John C., Harry Erwin, and Stefan Wermter. "Robotics sound-source localization and tracking using interaural time difference and cross-correlation." *AI Workshop on NeuroBotics*. 2004.

[18] Alameda-Pineda, Xavier, and Radu Horaud. "A geometric approach to sound source localization from time-delay estimates." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.6 (2014):

[19] CMU biorobotics lab, Snake robot http://biorobotics.ri.cmu.edu/

[20] Whitman, Julian, et al. "Snake robot urban search after the 2017 Mexico City earthquake." *2018 IEEE international symposium on safety, security, and rescue robotics (SSRR)*. IEEE, 2018.

[21] Crespi, Alessandro, and Auke Jan Ijspeert. "AmphiBot II: An amphibious snake robot that crawls and swims using a central pattern generator." *Proceedings of the 9th international conference on climbing and walking robots (CLAWAR 2006)*. No. CONF. 2006.

[22] Yamada, Hiroya, and Shigeo Hirose. "Development of practical 3-dimensional active cord mechanism ACM-R4." *Journal of Robotics and Mechatronics* 18.3 (2006): 305-311.

[23] Sarah C. P. Williams,"Vibrating Skulls Help Snakes Hear", Science, 22 Dec 2011

[24] Demo video of Snake robot sound soruce localization through shapre reconifguration, February 6th 2022 Available

Sriranjan Rasakatla (PhD student and Research) is currently a 3rd PhD student in the Department of Mechanical Systems Engineering with Applied Mechanical Systems at TUAT is also a Research Intern at SONY AI. During his PhD now at Tokyo University of Agriculture and Technology he has developed force feedback compliant surgical arm, camera enable remote guided cyborg insect systems, Snake robots and Robotic leg for amputees with EMG based control. He was also selected for WISE program at TUAT for Excellent Leader Development for Super Smart Society by New Industry Creation and Diversity. Prior to this he has worked for 10 years in the industry and Academia including 3 years Uurmi Systems (currently Uurmi systems is Mathworks Hyderabad, India) where he worked on Defense projects like robots with all terrain mobility, Autonomous cars drive by wire control technology, driver assistance systems. He was enrolled in Masters by Research program with IIIT-H, India where he worked on developing Semi-autonomous snake robots, Legged robot systems and Human computer interface systems like camera mouse. His major was in Computer science with research in Robotics and HC. He also worked on developing Sixth sense like systems. His Bachelors in Engineering was from Osmania University with Major in Information Technology. He has published works in the past in Intl' conference in Robotics and Biomimetics(ROBIO 2010,2012,2021) and in SIGGRPAH(2010,2021) and SIGGRAPH ASIA.(2010,2021)

Dr.Bipin Indurkhya is a professor of Cognitive Science at the Jagiellonian University, Krakow, Poland. His main research interests are social robotics, usability engineering, affective computing and creativity. He received his Master's degree in Electronics Engineering from the Philips International Institute, Eindhoven (The Netherlands) in 1981, and PhD in Computer Science from University of Massachusetts at Amherst in 1985. He has taught at various universities in the US, Japan, India, Germany and Poland; and has led national and international research projects with collaborations from companies like Xerox and Samsung.

Dr.Ikuo Mizuuchi is a Professor, Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology2002 Doctor of Engineering from The University of Tokyo2002-2006 Project Assistant Professor, Grad. School of Info. Sci. and Tech., The Univ. of Tokyo.2006-2009 Lecturer, Dept. of Mechano-Informatics, Grad. School of Info. Sci. and Tech., The Univ. of Tokyo.2009.