

Mobile Robot-Assisted Mapping of Materials in Unknown Environments

Shyam Sundar Kannan, Wonse Jo, Ramviyas Parasuraman, and Byung-Cheol Min

Abstract—Active perception by robots of surrounding objects and environmental elements may involve contacting and recognizing material types such as glass, metal, plastic, or wood. This perception is especially beneficial for mobile robots exploring unknown environments, and can increase a robot’s autonomy and enhance its capability for interaction with objects and humans. This paper introduces a new multi-robot system for learning and classifying object material types through the processing of audio signals produced when a controlled solenoid switch on the robot is used to tap the target material. We use Mel-Frequency Cepstrum Coefficients (MFCC) as signal features and a Support Vector Machine (SVM) as the classifier. The proposed system can construct a material map from signal information using both manual and autonomous methodologies. We demonstrate the proposed system through experiments using the mobile robot platforms installed with Velodyne LiDAR in an exploration-like scenario with various materials. The material map provides information that is difficult to capture using other methods, making this a promising avenue for further research.

Index Terms—*Robots, Material Classification, Support Vector Machine, Tapping Sound Recognition, Environment Perception*

I. INTRODUCTION

The utility of robots in exploration and mapping applications has gained significant interest and advancements in recent years. Mobile robots are used in a variety of situations such as search and rescue scenarios, firefighting aids, service and logistics, domestic aids, and more. Nevertheless, numerous unsolved problems remain when deploying robots in unknown environments. In particular, it is important for the robot to perceive and learn environment properties to increase its autonomy and effectively execute its mission [1]. For instance, in search and rescue operations, it is essential to know the location of doors and other access points to preplan the operation.

To address this problem, researchers have utilized several sensing modalities and machine learning algorithms to classify different objects and materials. Of these modalities, computer vision has become prominent because of the availability of public image datasets and recent advances in deep learning algorithms [2]. There are several search and rescue robots in use that depend on image processing and vision-based techniques [3]–[5], but they tend to fail when the lighting conditions are bad or the environment is smoky which significantly reduces the visibility of the scene.

Tactile and acoustic sensing techniques that are robust to poor lighting conditions have been well studied and proven

S.S. Kannan, W. Joe, and B-C. Min are with SMART Lab, Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47907, USA. e-mails: jow, kannan9, minb@purdue.edu.

R. Parasuraman is with Department of Computer Science, University of Georgia, Athens, GA 30605, USA. email: ramviyas@uga.edu.



Fig. 1: A picture of our mobile robot exploring and tapping the objects in order to perceive its environment.

effective in similar applications [6], [7]. While tactile sensors alone require a variety of contact motions and potentially lengthy contact duration with the surface material, a combination of tactile and acoustic (sound) signals reduces the complexity when employing a simple interaction such as tapping. In fact, elucidating the properties of a target material through machine analysis of sounds generated from it is a well-studied topic. In particular, studies using the sound of tapping to identify material type date back to the work of Durst and Krotkov in 1995 [8], [9] where peaks in the frequency domain were used for the classification.

The objective of our work is to develop an intelligent system of robots that can identify surrounding materials and perceive their environment through the analysis of sound signals produced from tapping on the target objects. From these signals, a material map is created that gives information about the locations and types of surrounding objects. This will greatly help the robot perceive its environment irrespective of lighting conditions. This idea was inspired by how blind persons learn about objects in their surroundings through tapping objects using a cane [10], [11]. The core idea underlying the approach is that the tapping sound produced by each object depends on its density, shape, size, and, importantly, material composition; these features can be captured by variation in the spectral distribution of tapping sound waveforms.

As we take inspiration from human senses, we use the Mel-frequency cepstrum (MFC), which aligns better with human auditory perception than does a Fourier spectrogram. Material type is determined from the MFC coefficients (MFCC) with the help of machine learning methods, specifically a Support Vector Machine (SVM), which has shown to achieve high classification accuracy in similar work [12]. We create a

material map by overlaying material information as markers on the occupancy grid, and present two methods for its creation. In the first, the robot is teleoperated and interrogating tap triggered manually. The second is a multi-agent autonomous method in which one agent moves from point to point and samples objects in the environment based on a map built by another agent. In both methods, the outcome is to enhance environmental perception for mobile robots in the search, exploration, and mapping of unknown, collapsed, or damaged physical infrastructure. Fig. 1 depicts our robot exploring an unknown environment and tapping objects to identify their constituent materials.

The contributions of the paper are outlined below.

- 1) We propose and validate a simple and effective solution to the problem of classifying the material types of surrounding objects through a tapping mechanism and the machine learning analysis of associated sounds.
- 2) We construct material maps by overlaying material data on the occupancy grid map.
- 3) We present a multi-agent system for sampling and identifying materials in a region autonomously.

II. RELATED WORK

Several stand-alone and hybrid (multi-modal) sensing modalities have been used in the literature to learn and identify the abstract material and surface types (or textures) [13]–[15]. Although recent years have witnessed attention to visual sensing medium that uses images from the camera to perceive contact material properties for robot manipulators [1], [16], tactile and haptic interactions and acoustic sensing have been well-researched and proven to produce robust features for machine learning solutions.

For instance, high-pressure signals generated by one-leg hopper robots are used to classify the terrain surface type such as grass, carpet, or wood with a near-perfect accuracy of up to 99% [17]. In [6], the authors used a pad of 18 tactile sensors on the robot hand. Using tactile data from five different linear motions of the sensor pad on an object surface, they obtained a classification accuracy of up to 97% accuracy for 49 different objects using an SVM classifier. Similarly, the authors in [18] used data from a 3D tactile sensor to classify texture of the contact material with up to 89% accuracy. A study in [19] used both tactile data and the rate at which heat is transferred from the tactile sensor to the material in contact to classify the type of materials with an SVM classifier. They obtained up to 98% accuracy for longer contact duration (1.5 seconds). However, such tactile-based sensing alone requires sophisticated and longer contact and interaction with the target materials. Addition of audio signals to the tactile (and haptic) interactions reduces such limitations and enables faster and a more practical solution for material identification.

The authors in [7] used NAO humanoid robots to manipulate target objects (picking up and forcefully hitting it) and used the dominant frequency of the recorded sounds to classify the objects. Similar work was performed in [20]. In [21], using the Fourier analysis of the sounds resulting from a robot

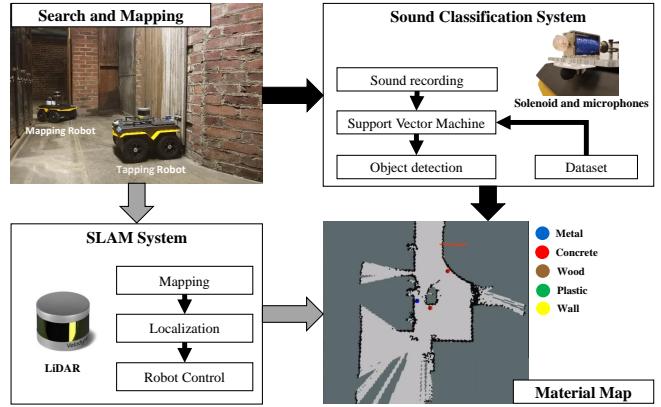


Fig. 2: Generalized architecture of the proposed system.

manipulator performing several actions on the target object (grasp, push, or drop), the authors can accurately ($\approx 97\%$) classify up to 18 objects with a Bayesian classifier.

Motion aided audio signal analysis has also been used to detect touch gestures [22], and terrain and surface types [23]. For example, in [12], the spectral and temporal features including MFCC of sound signals captured during locomotion of legged robots are used to perform terrain classification, achieving an accuracy of up to 95% for seven terrain types using SVM classifier. Therefore, inspired by the above works, we use tapping sounds and an SVM classifier to recognize material types.

Moreover, the integration of such sound-based analysis to robot exploration is still an evolving research area. For instance, in [24], the authors used tapping sounds along with a LiDAR scan from a mobile robot to create a map of the impact locations for assisting the human inspector during hammer sounding inspections of the concrete walls and buildings.

We depart from the previous works in three important ways. First, we aim for an environmental mapping system that can label objects on a map according to their material types using our recognition method. Therefore, we integrate our system with an existing SLAM (Simultaneous Localization and Mapping) algorithm on the robot. Second, we employ an adaptive noise cancellation filter through our dual microphone setup to realize practical robot applications in scenarios where environmental noise has a high influence on the recorded tapping sounds. Third, we use low dimensional cepstral features to achieve an accurate and real-time recognition circuit with low computation complexity.

III. PROPOSED SYSTEM

The proposed system in this paper identifies materials in an unknown environment and creates a material map which localizes the various materials in the environment over the occupancy grid. Fig. 2 gives a general architecture of the proposed system. The SLAM system helps in map creation, localization, and navigation of the agents. The sound classification system recognizes the various tapping sounds and identifies the corresponding material. These data are later merged into

a material map which localizes the various materials in the environment.

The system can be operated in two modes. Manual teleoperation mode where a mobile robot is teleoperated manually. Here a tapping mobile robot equipped with a solenoid for tapping and a LiDAR for mapping is used. The robot moves around building the map and also tapping various objects around and hence creating the material map. In autonomous mode, multiple agents are used. A mapping robot equipped only with a LiDAR is used for mapping, and a tapping robot (like the one used in manual mode) is used for identifying the various materials. The tapping robot functions autonomously on the map created by the mapping robot which is teleoperated.

A. Hardware Configuration

An iRobot Create 2 and Jackal from Clearpath Robotics [25] were used as the mobile base for the implementation; we used an iRobot Create 2 robot for the manual teleoperation while the Jackal robots were used for autonomous mode of materials mapping. The tapping robot for the both modes consists of a linear solenoid and two microphones in addition to the LiDAR. The hardware configuration of the tapping robot with the Jackal platform has been depicted in Fig. 3. The mapping robot consists of a LiDAR (Velodyne VLP-16) for mapping the environment. A 3D LiDAR was used, but only 2D data was used for mapping.

A linear solenoid switch that can be controlled to extend/retract was used as the tapping device in our prototype. Fig. 4 shows the solenoid in pull (retract) and push (extend) modes. We applied a plastic cap on the solenoid tip to make the solenoid tip acoustically compatible as the elasticity of the tip plays an important factor [10]. The solenoid used had a stroke length of about 15 mm and applied a force of about 45N. The force of the solenoid is good enough to produce sound but not too large that it causes damage to the environment.

The sound produced by the tap of the solenoid was recorded using a dual microphone set up. Among the two microphones, one (on the front side of the robot) is placed closer to the solenoid tapper so that it obtains the tapping sound with high sound clarity. The other microphone (on the rear side of the robot) is placed away from the solenoid tapper to capture the noise in the environment (including the noise sounds created by the robot's movement) with less influence of the tapping sounds. This system reduces the impact of the environmental noise.

IV. MATERIAL CLASSIFICATION

This section describes the various components of the proposed system to recognize material types of surrounding objects in the environment using the tapping sounds.

A. Active Noise Reduction

To deploy the proposed system in real-world, a recording system needs to remove or reduce the ambient acoustic noise for accurate analysis. We employ a method similar to the

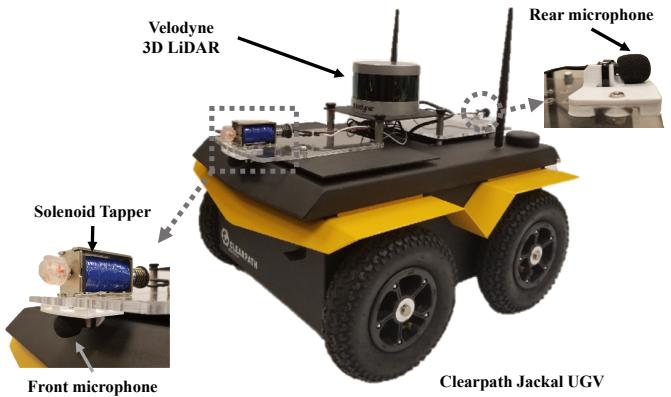
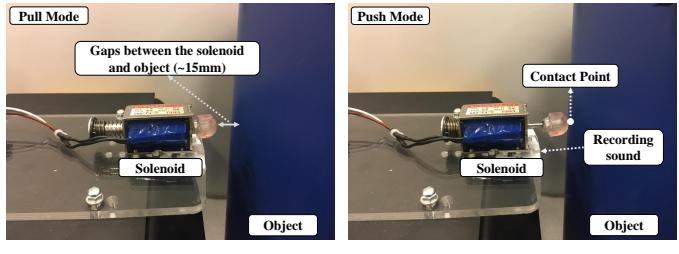


Fig. 3: Hardware configuration of the mobile tapping robot



ones used in noise canceling headset, hearing aids, and mobile phones [26].

The synchronized sound signals from the front microphone (tapping sound) and the rear microphone (background noise) are sent to the adaptive filter, which efficiently matches the unknown noise characteristics with a Finite Impulse Response (FIR) model and applies the error correction through Normalized Least Mean Squares (NLMS) algorithm that is generally used in signal enhancement algorithms. The anti-noise is generated by inverting the FIR output and then combined with the signal from the front microphone to effectively remove the background noise [27].

B. Feature Extraction

A common feature vector used in speech recognition applications and audio signal classification is the Mel-Frequency Cepstral Coefficients (MFCC) [28]. A Cepstrum is called nonlinear "spectrum-of-a-spectrum" because it captures the spectral shape of a given audio signal by calculating the discrete cosine transform (DCT) of the spectrogram in the log (power) scale. Thus, they provide more information on the signal characteristics than a spectrogram. Since our motivation comes from blind persons effectively recognizing the type of objects and materials through sounds, we focus on using the non-linear Mel-frequency bands, which resemble the response of the human auditory system. Therefore, we use Mel-frequency Cepstral Coefficients (MFCCs) as the features of the input signals compared to the spectrogram based features in [13], [15].

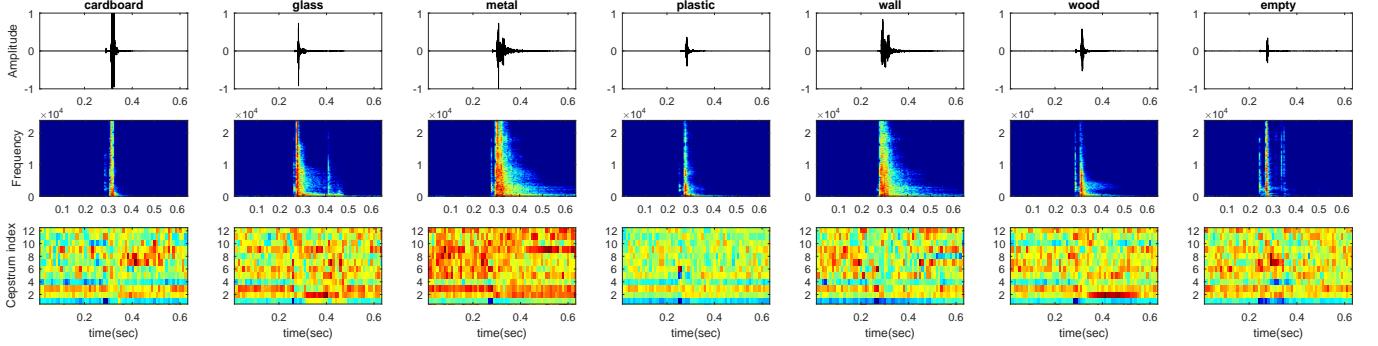


Fig. 5: Tapping sound signal characteristics in time domain (amplitude), Fourier domain (frequency), and cepstral domain (MFCC indices) are shown for the following materials (in the order from left to right): cardboard, glass, metal, plastic, wall, wood, and empty (background noise).



Fig. 6: Sample objects used in the construction of the dataset

In case of spectrum, they tend to have different shapes, duration, peak frequency, and magnitude depending on the material types and shape, but some case, like glass and metal, has similar shape and length of the spectrogram, which we believe is due to the effects of echo sounds. On the other hand, the cepstral graph is distinctly different for all the materials and hence could be considered a robust feature vector for material classification. Fig. 5 clearly shows that the cepstral features are unique for all the material later used in building the dataset.

C. Machine Learning of Tapping Sound

The support vector machines (SVMs) is a widely-used and well-studied probabilistic machine learning method to learn and classify supervised data. Due to brevity, we refer the readers to [29] for more information on the SVM classification algorithm.

The proposed system utilizes the SVM classifier with MFCC as the input feature vector and the material type as the classified output label. The classification algorithm, multiple ($\frac{n(n-1)}{2}$, where n is the number of classes) linear SVM classifiers (performing one versus one classification) uses a decision function of shapes to obtain the final classification output along with confidence values.

D. Selection of Classes and Dataset Creation

The classification of the tapping sound needs a reference dataset based on which a classification model can be developed. To the best our knowledge, there are no publicly available datasets consisting of tapping sounds of various materials. Hence, we created a dataset with an extensive collection of tapping sounds from various objects, such as, wood, metal, glass, plastic, cardboard, wood, concrete, and wall (hardboard), which are commonly found in everyday life. We recorded the tapping sound from objects like trash bins, storage cabinets, wall, cardboard boxes, doors, tables, and so on which are made with these materials (as their principal composition), and build our dataset. Fig. 6 shows some of the objects used in the construction of the dataset.

We also included a class which consists of the empty tap (with no target object) to account for tapping issues such as the solenoid tip unable to contact the target material or an accidental trigger. So the total number of classes in our experiments is 8 (7 materials + 1 Empty).

The dataset consists of an average of around 100 sample sounds per class. In our learning stage, we observed that training:test dataset split of 70:30 produces optimal results and either lower or higher split ratio resulted in either underfitting or overfitting, respectively.

E. Classification results

As discussed in the previous section, a SVM classification model was trained using the recorded tapping sounds. In the Fig. 7, we present the confusion matrix of the material classification results which has been normalized, so that the column sum to 100%. The mean accuracy of the classification results is 91.54%.

From the confusion matrix, it can be observed that there is a reasonable misclassification between plastic and wall, metal and wood, and plastic and cardboard. These are due to the similarities in the sound.

V. ROBOT EXPLORATION

In this section, we present two methods of operating the robots (1) manual teleoperation and (2) an autonomous mode

Mean Accuracy: 91.54%								
glass	metal	cardboard	concrete	wall	plastic	wood	empty	
glass	100.0%	2.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
18	1	0	0	0	0	0	0	0
metal	0.0%	80.0%	4.5%	0.0%	0.0%	0.0%	1.6%	0.0%
0	28	1	0	0	0	0	1	0
cardboard	0.0%	2.9%	95.5%	0.0%	0.0%	13.3%	0.0%	0.0%
0	1	21	0	0	10	0	0	0
concrete	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
0	0	0	15	0	0	0	0	0
wall	0.0%	0.0%	0.0%	0.0%	86.4%	13.3%	0.0%	0.0%
0	0	0	0	19	10	0	0	0
plastic	0.0%	0.0%	0.0%	0.0%	13.6%	72.0%	0.0%	0.0%
0	0	0	0	3	54	0	0	0
wood	0.0%	14.3%	0.0%	0.0%	0.0%	1.3%	98.4%	0.0%
0	5	0	0	0	1	61	0	0
empty	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
0	0	0	0	0	0	0	0	31

Fig. 7: The confusion matrix of our classifier.

where the robot functions based on a map constructed by another robot.

A. Manual Teleoperation

In this mode, the robot¹ is teleoperated manually by an operator. The online map constructed by the robot is used as the reference for teleoperation. GMapping [30] is used to simultaneously localize and map the robot's position in the environment. The robot is controlled to move towards an object such that the robot faces it and the solenoid tap is triggered manually. This tapping sound is recorded, filtered and then processed as described in the previous section, to identify the material. The identified materials are added to the map as markers, hence giving the operator better insights on the environment. This method is ideal when the number of places to be sampled is less and for smaller environments, and there are no other robots available to map the region. Fig. 8 shows a map constructed using manual teleoperation along with the ground truth for three different locations.

B. Autonomous Mapping

It can be tedious to teleoperate the robot and to sample the environment, especially when a large number of sample points are needed or if the environment is large. To remedy this, we introduce a method, where the robot navigates autonomously and detects the materials in the environment based on a map built by another agent (mapping robot).

First, the mapping robot is teleoperated to build the map of the environment. Then, a central server processes the map built by the mapping robot. From this map, the locations where the

¹Note, for the manual teleoperation experiments, we used an iRobot Create 2 robot (instead of Jackal robot) mounted with the same tapping system. The Jackal robot based system design is an evolved version of our proposed hardware design and is intended for autonomous mode of materials mapping.

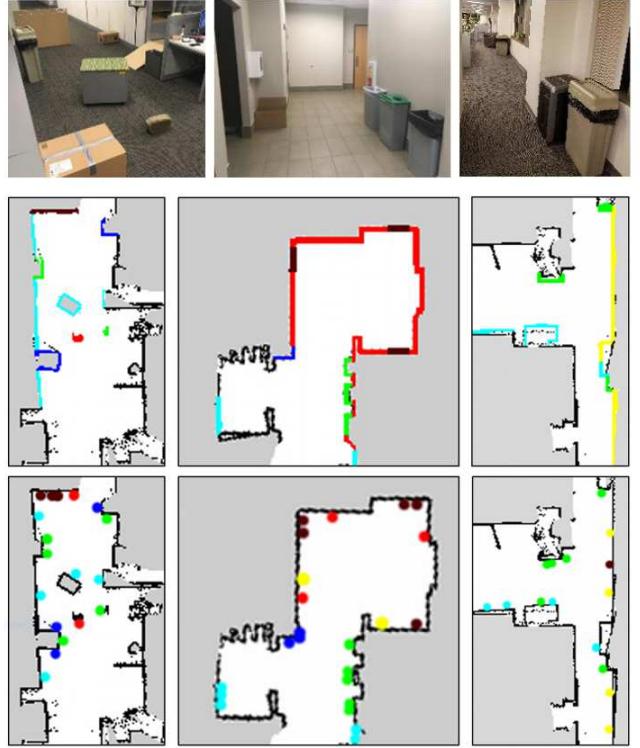


Fig. 8: Material maps constructed using manual teleoperation. (Images of the environment, the ground truth and the material map created are shown for three different locations.) The color coding are as follows: metal (cyan), plastic (green), concrete (red), cardboard (blue), wall (yellow), and wood (brown).

tapping robot needs to visit and sample the tapping sounds are extracted. A human operator needs to trigger the deployment of the tapping robot when a good map has been built by the mapping robot. The map constructed for the environment in Fig. 10a has been shown in Fig. 10b. These points are later communicated to the tapping robot, and it moves from point to point, tapping and inspecting those locations and finally creating a material map. Fig. 9 shows the overall architecture of the autonomous mapping system.

1) *Extraction of Points to sample:* The points the tapping robots needs are obtained by processing the map obtained from the mapping robot. The map is preprocessed and the points are extracted based on the geometry as described below.

Map Smoothing using Delaunay Triangulation: The map obtained from the mapping robot contains noise and some irregularities and the map needs to be smoothed for easy extraction of the points to be sampled. The noise and irregularities in the map can deter the extraction of correct points for sampling. We use a Delaunay Triangulation based method for removing noise and smoothening the map based on the cleaning method proposed in [31]. In this method, a point set is constructed by considering each occupied pixel (black pixels in the occupancy grid) as a point. Then, Delaunay Triangulation is applied to this point set. From the triangulated result, all the

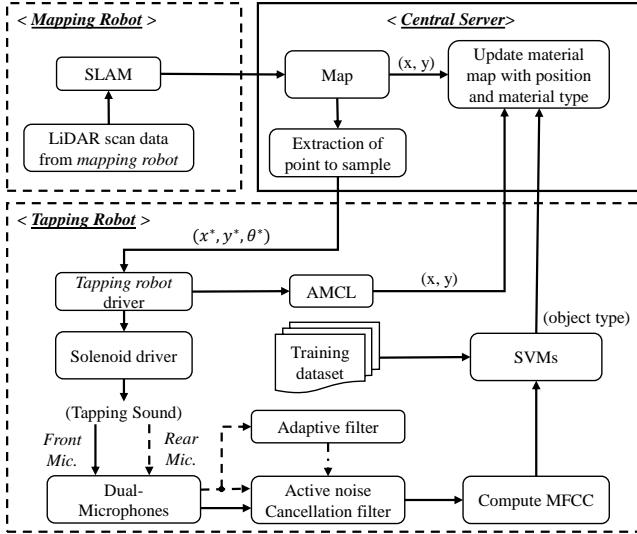


Fig. 9: Implementation of the proposed system for autonomous mapping.

triangles which have at least one edge whose length is greater than a user-defined parameter γ are removed. This process tends to remove all the big triangles in the triangulation. Based on experimentation, we found that $\gamma = 5$ looks to be ideal.

If the map has any noise, they are usually isolated and are linked to the other points by long edges in the triangulation, and hence they are removed. In case, if there is any irregularity, the region is filled with tiny triangles which are retained. These tiny triangles which are retained tend to smoothen the map.

Extracting Boundaries: The points to be sampled by the robot should lie on the border between the empty and filled regions (white and black pixels in the occupancy grid). To extract these points, we first extract all the points which form the boundary between the filled and empty region from the smoothed map obtained in the previous step. These points are obtained by looking at the eight neighboring pixels of all the occupied pixels. If a pixel is surrounded by at least one pixel corresponding to the free space (white pixel), then that pixel is considered to be a part of the boundary. A graph is constructed using the various boundary pixels and with neighboring pixels forming edges. The connected components found from this graph gives the various segments in the map are extracted. Fig. 10d shows the segmented boundary map for the map shown in Fig. 10b. Small segments are not considered. During experimentation, small segments which had less than 20 points were not considered for the later steps.

Sampling the Boundary: Based on the boundaries extracted in the previous step, points are extracted on the boundary lines at regular intervals. An interval of 1 meter was used in the experiments. Fig. 10e shows the various points to be sampled (in purple). For the robot to reach the point, the destination pose is also needed. Ideally, the robot should be perpendicular to the object surface for proper taps. The orientation is computed from the gradient direction computed

TABLE I: Accuracy of the proposed material recognition system in the exploration experiments.

Exp. #	1	2	3	4	Success rate
Metal	5/5	5/5	4/5	0/0	93%
Plastic	5/5	5/5	5/5	1/1	100%
Wood	4/5	4/5	0/0	2/2	83%
Cardboard	2/5	4/5	0/0	0/0	60%
Concrete	5/5	3/5	0/0	0/0	80%
Wall	0/0	0/0	4/5	12/12	94%
Empty	5/5	4/5	4/5	0/0	87%
Total	26/30	25/30	17/20	15/15	87%

using Sobel operator [32].

The horizontal and vertical derivative approximations G_x and G_y for a given image A is computed as

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A \quad (1)$$

$$G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (2)$$

where $*$ is the convolution operator.

Based on this with the smoothen map as the input, we can compute the gradient direction (Θ) as

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (3)$$

From the gradient direction(Θ), the required direction of the robot at point can be computed.

2) *Autonomous Control:* The positions with orientation extracted in the previous step are used as an input for this step. The Robot Operating System (ROS) navigation stack is used for localization (AMCL package) and for moving the robot from one point to another (move_base). The points are sorted based on the Euclidean distance between them so that the robot moves from one point to the next nearest point. The robot moves from point to point and tap the object and identifies the material and hence constructing the material map. Fig. 10f shows the material map for the environment in Fig. 10a and the corresponding ground truth in Fig. 10g.

C. Results and Discussions

The materials classified by the robot in both manual and autonomous modes were compared with the ground truth in order to validate the performance of the robot. The results of experiments are summarized in Table I and the LiDAR map with object markers are shown in Fig. 8 (column 1-3 in table) and Fig. 10 (column 4 in table). The Table I gives the measures of the accuracy of our algorithm measured under various experimental setups and modes of operation. Note, the first three experiments are manual teleoperation scenarios and the fourth experiment is the autonomous navigation scenario. From the table, it can be seen that the algorithm can identify most of the materials accurately at various locations and that the proposed system was able to detect most of the material

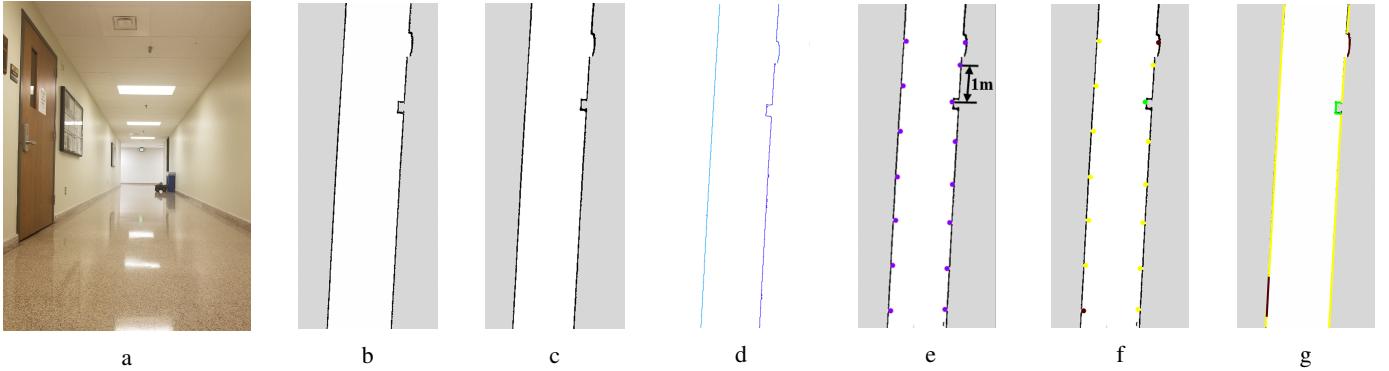


Fig. 10: Construction of material map for an unknown environment: (a) Picture of the hallway used for the experiment, (b) Original map created using LiDAR from mapping robot, (c) Smoothed map , (d) Segmented boundary map, (e) Points to sample at 1m interval (in purple), (f) Material Map created, (g) Ground Truth material map. The color coding for material map and ground truth are as follows: **plastic** (green), **wall** (yellow), and **wood** (brown).

using tapping sounds. This is because each object being tapped by the robot has different sound properties depending on the composition of the materials the object is made of. Also, the system is robust with regard to environmental noise due to NLMS algorithm for the active noise reduction and MFCCs for feature extraction. To be specific, the detection accuracy of metal, plastic, wood, cardboard, concrete, and wall is approximately 93%, 100%, 83%, 60%, 80%, and 94%, respectively. From the experiments, we found that the material can be identified with higher accuracy by multiple taps on the same material/object at various locations. This helps to reduce the number of false classifications.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have presented a tapping system for a mobile robot that can be used for mapping various materials such as wood, plastic, metal, glass, wall, etc. in an unknown environment. We have discussed the various components of the system including the hardware design, the sound classification method, and the robot control algorithm for autonomous navigation and mapping.

Through real experiments, we have demonstrated that using our proposed tapping system, we can classify the materials with an accuracy of approximately 92% in identifying materials of known objects and an average accuracy of 87% in identifying materials in unknown environments. The obtained materials map integrated with the SLAM map of the robot is not only useful for improving the robot's autonomy but also useful for planning robotic search and rescue operations in advance, for example.

As a future work, we plan to extend this idea to the 3D material map for use with drones and we intend to improve the proposed tapping system with a height adjustable solenoid leading to its ability in aiding the 3D material maps.

REFERENCES

- [1] D. Katz, A. Venkatraman, M. Kazemi, J. A. Bagnell, and A. Stentz, “Perceiving, learning, and exploiting object affordances for autonomous

- pile manipulation,” *Autonomous Robots*, vol. 37, no. 4, pp. 369–382, 2014.
- [2] P. Wieschollek and H. Lensch, “Transfer learning for material classification using convolutional networks,” *arXiv preprint arXiv:1609.06188*, 2016.
- [3] J.-H. Kim, S. Jo, and B. Y. Lattimer, “Feature selection for intelligent firefighting robot classification of fire, smoke, and thermal reflections using thermal infrared images,” *Journal of Sensors*, vol. 2016, 2016.
- [4] F. Matsuno and S. Tadokoro, “Rescue Robots and Systems in Japan,” in *2004 IEEE International Conference on Robotics and Biomimetics*, Aug. 2004, pp. 12–20.
- [5] J. Casper and R. R. Murphy, “Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 33, no. 3, pp. 367–385, Jun. 2003.
- [6] J. Hoelscher, J. Peters, and T. Hermans, “Evaluation of tactile feature extraction for interactive object recognition,” in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE, 2015, pp. 310–317.
- [7] E. Lopez-Caudana, O. Quiroz, A. Rodríguez, L. Yépez, and D. Ibarra, “Classification of materials by acoustic signal processing in real time for nao robots,” *International Journal of Advanced Robotic Systems*, vol. 14, no. 4, p. 1729881417714996, 2017.
- [8] R. S. Durst and E. P. Krotkov, “Object classification from analysis of impact acoustics,” in *Intelligent Robots and Systems 95.'Human Robot Interaction and Cooperative Robots', Proceedings. 1995 IEEE/RSJ International Conference on*, vol. 1. IEEE, 1995, pp. 90–95.
- [9] E. Krotkov, “Robotic perception of material,” in *IJCAI*, 1995, pp. 88–95.
- [10] B. N. Schenkman, “Identification of ground materials with the aid of tapping sounds and vibrations of long canes for the blind,” *Ergonomics*, vol. 29, no. 8, pp. 985–998, 1986, pMID: 3758024. [Online]. Available: <https://doi.org/10.1080/00140138608967212>
- [11] K. Nunokawa, Y. Seki, S. Ino, and K. Doi, “Judging hardness of an object from the sounds of tapping created by a white cane,” in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 5876–5879.
- [12] J. Christie and N. Kottege, “Acoustics based terrain classification for legged robots,” in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3596–3603.
- [13] M. Strese, C. Schuwerk, A. Iepure, and E. Steinbach, “Multimodal feature-based surface material classification,” *IEEE transactions on haptics*, vol. 10, no. 2, pp. 226–239, 2017.
- [14] H. Zheng, L. Fang, M. Ji, M. Strese, Y. Özer, and E. Steinbach, “Deep learning for surface material classification using haptic and visual information,” *IEEE Transactions on Multimedia*, vol. 18, no. 12, pp. 2407–2416, 2016.
- [15] W. Fujisaki, M. Tokita, and K. Kariya, “Perception of the material properties of wood based on vision, audition, and touch,” *Vision research*, vol. 109, pp. 185–200, 2015.

- [16] L. Sharan, C. Liu, R. Rosenholtz, and E. H. Adelson, "Recognizing materials using perceptually inspired features," *International journal of computer vision*, vol. 103, no. 3, pp. 348–371, 2013.
- [17] J. J. Shill, E. G. Collins, E. Coyle, and J. Clark, "Terrain identification on a one-legged hopping robot using high-resolution pressure images," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4723–4728.
- [18] D. S. Chathuranga, Z. Wang, Y. Noh, T. Nanayakkara, and S. Hirai, "Robust real time material classification algorithm using soft three axis tactile sensor: Evaluation of the algorithm," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 2093–2098.
- [19] T. Bhattacharjee, J. Wade, and C. C. Kemp, "Material recognition from heat transfer given varying initial conditions and short-duration contact." Georgia Institute of Technology, 2015.
- [20] E. Torres-Jara, L. Natale, and P. Fitzpatrick, "Tapping into touch," 2005.
- [21] J. Sinapov, M. Weimer, and A. Stoytchev, "Interactive learning of the acoustic properties of objects by a robot," in *Proceedings of the RSS Workshop on Robot Manipulation: Intelligence in Human Environments, Zurich, Switzerland*, 2008.
- [22] F. Alonso-Martín, J. J. Gamboa-Montero, J. C. Castillo, Á. Castro-González, and M. Á. Salichs, "Detecting and classifying human touches in a social robot through acoustic sensing and machine learning," *Sensors*, vol. 17, no. 5, p. 1138, 2017.
- [23] N. Roy, G. Dudek, and P. Freedman, "Surface sensing and classification for efficient mobile robot navigation," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, vol. 2. IEEE, 1996, pp. 1224–1228.
- [24] A. Watanabe, J. Even, L. Y. Morales, and C. Ishi, "Robot-assisted acoustic inspection of infrastructures-cooperative hammer sounding inspection," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 5942–5947.
- [25] Clearpath, "Jackal ugv - small weatherproof robot - clearpath." [Online]. Available: <https://www.clearpathrobotics.com/jackal-small-unmanned-ground-vehicle/>
- [26] Z. H. Fu, F. Fan, and J. D. Huang, "Dual-microphone noise reduction for mobile phone application," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 7239–7243.
- [27] S. Haykin, *Adaptive Filter Theory* (2Nd Ed.). Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1991.
- [28] N. Morgan, H. Bourlard, and H. Hermansky, "Automatic speech recognition: An auditory perspective," in *Speech processing in the auditory system*. Springer, 2004, pp. 309–338.
- [29] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [30] G. Grisetti, C. Stachniss, and W. Burgard, "Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, April 2005, pp. 2432–2437.
- [31] A. D. Parakkat, U. B. Pundrikaksha, and R. Muthuganapathy, "A delaunay triangulation based approach for cleaning rough sketches," *Computers & Graphics*, vol. 74, pp. 171 – 181, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0097849318300761>
- [32] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," *a talk at the Stanford Artificial Project in*, pp. 271–272, 1968.