CrossMark

Human-Like Sensing for Robotic Remote Inspection and Analytics

Arpan Pal[1] · Ranjan Dasgupta[1] · Arindam Saha[1] · Bhaskar Nandi[1]

Abstract In this paper we introduce Human-like Sensing or "5 senses computing" as a natural futuristic extension of Internet-of-Things. We discuss how 3D optical vision, thermal vision, acoustic profiling, olfaction and tactile sensing can help in remote inspection and analytics solutions. We propose a robot mounted opto-thermal and acoustic sensing system as a possible integrated system to gather such data. We present results of field experiments conducted with the proposed system and show how such systems can provide acceptable solutions for remote inspection and analytics.

1 Introduction

Connected Sensor Systems (also called Internet-of-Things or IoT) are poised for a disruptive growth in near future. The "Sense-Analyze-Respond" paradigm of IoT system is providing insight, deriving value and creating new ways to do business across multiple verticals be it manufacturing, healthcare, transportation, energy etc.

A typical IoT system has three components—a sensor system to collect data about physical events, a data transport system to send the collected sensor data to cloud and a

✉ Arpan Pal
  arpan.pal@tcs.com

  Ranjan Dasgupta
  ranjan.dasgupta@tcs.com

  Arindam Saha
  ari.saha@tcs.com

  Bhaskar Nandi
  bhaskar.nandi2@tcs.com

[1] TCS Innovation Labs, Tata Consultancy Services Limited, Ecospace, 4th Floor, Bldg 1B, New Town, Kolkata, West Bengal 700156, India

learning and knowledge based analytics system to derive insights about the physical event from the collected data. So far IoT systems have focused more on using specialized sensors to collect more and more information about the physical events. However, if we look at human beings as connected entities, we, since our birth continuously learn about our surrounding through our five senses—the sensed data is converted into knowledge and wisdom via complex learning systems of our brain. In the same analogy, if the human five senses can be imparted on machines along with sophisticated machine learning techniques, it would be possible to create smarter IoT systems than what is available today. This concept of sensing five human senses on machines followed by advanced machine learning based analytics to derive knowledge and insights is popularly known as "5 senses computing".[1] Here we try to present a glimpse of the "art of the possible" in the area of sensing these five senses, which can either be stand-alone sensors or can be mounted on robots—each of these have significant application in remote inspection and predictive analytics, especially for manufacturing and shop-floor scenarios.

Sight Even though camera existed as a sensor for a long time, intelligent processing of image and videos can take vision processing to a new level. This will include 3D vision, vision based measurements, automatic object recognition and semantic understanding of images and video content [1]. New research in this field is also enabling unobtrusive use of camera to sense physical events like vibration, micro motion movements etc.

Hearing There will also be significant improvements in computers' ability to hear and understand sound. Source localization and analysis of sound captured unobtrusively from microphones can provide insights into a lot of physical events spanning across verticals [2].

Touch Touch sensing and haptic feedback is poised to become core component of human–computer interaction (HCI). Going beyond, our skin not only senses touch, but also senses heat. Heat sensors in form of Infra-red camera and thermal cameras can provide a lot of insight into a machine's condition. This information, when coupled with 3D vision information can be used for predictive analytics [3].

Smell Olfactory sensing in form of gas sensors and other similar sensors are slowly becoming main-stream. These sensors, backed up by requisite analytics will be able to check for molecular biomarkers in future [4].

Taste Taste sensors that can break down ingredients to their respective chemicals are also being prototyped [5]. While they have more direct applications in food and beverage industry, their application in other areas cannot be ruled out. Both olfactory and taste sensing can be potentially used in future for composition analysis and quality check of components and materials.

Fusion of the above five senses and the associated analytics is also termed as "cognitive computing" [6]. In time, cognitive computing will be able to unobtrusively observe and model complex interactions in complex systems—such a sensory-aware machine can be used the model to predict the system condition and health leading towards predictive maintenance and quality improvement.

In this paper, we first discuss the Current Application Areas in Sect. 2. Then in Sect. 3, we present related work and in Sect. 4 we describe our contribution in 3D vision based measurement and vibration sensing, acoustic source localization and thermal sensing and discuss how it adds value in remote inspection and analytics space. In Sect. 5 we present the results and discuss their outcome. Finally in Sect. 6 we summarize the work and discuss about possible future work.

---

[1] http://readwrite.com/2012/12/18/ibms-cognitive-computing-plans-giving-smartphones-5-senses.

2 Current Application Areas

The current application areas focus mainly on optical vision systems. Industrial Machine Vision has evolved as a major contributor in automating High end factory operations like Quality Control, Packaging Inspection, Identification, Measurement, Counting and Tracking in different industries. Machine Vision has added speed and efficiency to the manual methods of Inspection. This has improved reliability and precision. The philosophies of Track–Trace-Control Operations have drastically improved the factory efficiency of assembly/manufacturing, Supply Chain and Local Warehouse Management.[2, 3] Image Capturing, Image Processing and Output generation are the standard roles being played by Industrial Cameras and Machine Vision Instruments that is capturing the imagination of the manufacturers. Common machine vision technology applications are used in sorting, quality assurance, robotic guidance, material handling and optical gauging. Some specific usages are Automated PCB Inspection, Sub-Assembly Inspection, Robotic Guidance, Packaging Inspection and Sorting, Reading of Serial numbers, Molding Flash Detection etc. The advantages offered by these technologies has led to semi-automation and full-automation of production environments and manufacturing industries. The benefits include faster processing speeds, automated decision making, high end finished goods, easy transshipment of goods and items, elimination of waste and reduction of idle time. Smart and easily configurable cameras have taken the role of Industrial Eyes which is being promoted as replacement to human based inspection.

Earlier Machine Vision was limited to 2-dimensional images and videos only. The ease of availability of 3D vision software tools and algorithms has opened the possibility of 3D Image Processing and Object Identification. Optical three-dimensional (3-D) profilometry[4] is an important application for 3D vision because of its simplicity, flexibility, high accuracy, and non-contact nature. Latest researches in imaging sensors and digital projection technology further its progress in high-speed, real-time applications, enabling the reconstruction of 3-D shapes of moving objects and dynamic scenes.

Online Real Time Vision sensors may be used to measure and inspect process quality. This may also be used to check the dimensions of parts/components, compare them with the drawing dimensions and take real-time decision on quality. Overcoming the limitations of the Camera, Parts Inspection based on Image Processing can open a new area in the Process of Inspection and Quality Control. Vision Sensors have advanced to the stage where they are able to replicate the functions of Human Behavior or may be even achieve functions which the human eye cannot detect. In next chapter we explore the technology of sensing beyond optical vision.

3 Sensing Capability: Related Work

Machines and Robots are substituting humans in performing everyday tasks and are better suited to completing routine and repetitive tasks than humans. In these situations, it requires robots to be equipped with human-like sensory inputs as well as a robust decision support system to achieve their desired utility. Most of the commercially deployable robots

---

[2] http://www.industrialvision.co.uk/vision-systems.

[3] http://www.microscan.com/en-us/Technology/MachineVisionSystems/machine-vision.aspx.

[4] http://www.laserfocusworld.com/articles/print/volume-46/issue-1/features/optical-surface-profiling.html.

are equipped with proprioceptive and exteroceptive sensors mainly for autonomous navigation, obstacle avoidance and path planning. But robots having capability of human-like sensing & cognition is a next horizon problem. 2D/3D machine vision is a quite explored area but how to work in extremely smoky dark environments where visibility is partially or completely occluded is a challenge. Acoustic imaging capability, thermal sensing to localize and achieve surface temperature profiles and augment them with 3D visual information is a state-of-the art limitation which eventually can resolve uncertainty problems of unknown environmental mapping. Sound source detection is quite a researched area but applying it on robotics to make them working under various acoustic conditions and noisy environments is a pertaining research area. E-Noses are effective inside laboratory environments but put them altogether to build machine olfaction capability to detect, differentiate and localize gases, fumes, steams, haze efficiently in real-time is miles to go effort. Multi-sensing next generation robots to create a novel symbiotic autonomy in which machines are aware of their perceptual, physical and reasoning limitations and proactively act as human is a real necessity.

3.1 Machine Vision

As outlined in Sect. 2, the field of machine vision, or computer vision, has been growing at a fast pace and has numerous applications in automation of manufacturing systems. The imaging system can be composed of a camera (2D/3D) and capturing system. The capturing and processing is entirely dependent on type of camera is used (2D or 3D). Camera selection depends on the desired use cases and accuracy. We describe the 3D systems in detail below.

Automatic remote monitoring of an enclosed space is of compelling research nowadays. Constructing a 3D map/perception model of an unknown indoor or outdoor environment using a robotic platform is needed for such monitoring system. Available IMU sensors and mobile robot kinematics allow 3D reconstruction to be finished in near real-time using a very low cost robotic platform. In a recent work [7], Pradeep et al. has described a methodology for markerless tracking and 3D reconstruction in scenes of smaller size using RGB camera. It generates high quality 3D model reconstruction using a webcam. Pizzoli et al. [8] proposed a solution by adapting a probabilistic approach in which depth map is computed by combining Bayesian estimation and convex optimization techniques. All these implementations are limited to a small scene reconstruction and not suitable for an entire 3D environment creation. The 3D reconstruction of an environment from multiple images or video captured by a single moving camera has been studied for several years and is well known as Structure-from-Motion (SfM) and quite powerful SfM pipeline readily available for 3D reconstructing as shown in [9, 10]. Recently, smart phones are used for image acquisition due to its low cost and easy availability. So researchers used smart phones sensors like accelerometer, magnetometer for data collection and 3D reconstruction, it reduces computation [11, 12] and few works such as [13, 14] have accomplished this, but the output is noisy due to a fast and course reconstruction. A system capable of dense 3D reconstruction of an unknown environment in real-time through a mobile robot requires simultaneous localization and mapping (SLAM) [15].

Another interesting application area for 3D optical vision is Quality Control. Measurement using computer vision is studied extensively in the last decade. Researchers has measured straightness defect in hot rolling steel sheets [16] in real-time. The system is designed for easy implementation in the actual plant but the result presented without any bench marking. The system is specially designed only for flatness measurement of rolling

sheets and not applicable directly for auto component measurement. 3D vision based measurement system is proposed using a single moving camera [17] but that is limited to estimate the robot calibration in offline. The trend of using camera for vision based measurement is studied in [18]. Rövid proposed [19] a vision based measurement system for vehicle body inspection where a rotating gray code pattern projection and a multi-camera based tracking system is proposed for measurement of large surfaces. The main drawback of that system is the complexity of the multiple cameras and projector based complex system.

3.2 Thermal Perception for Machines

Unobtrusive heat measurement and monitoring is well accepted in manufacturing, chemical, automobile, construction industries. Conventional industrial thermal cameras are still not in affordable range for everyday life usage. Conventional thermography for energy measurement and non-invasive assessments relies on 2D thermal images, which have significant limitations like lack of information on the shape and geometry or location of the object of interest in the scene. So there is growing interest on representing the environment in 3D which also integrates the temperature information. The combined information will help to detect the object of interest and volumetric measurement precisely. FLIR[5] lunched a low cost and affordable thermal sensor [20] as smart-phone attachment which increases the possibility of monitoring and verification of heated region using such hand held mobile low cost sensors.

Several studies are performed to explore the potential of 3D thermal mapping and volumetric inspection. The studies are mostly focused on monitoring building power consumption. ThermalMapper [21] is a well-known project which uses a terrestrial laser scanner and thermal infrared camera on a wheel robot. The result from ThermalMapper is a dense 3D point cloud which can be visualized in both RGB and thermal. Volumetric heat measurement and analysis is not part of the presented system. There is significant cost and mobility difference between the presented systems with our proposed system due to the usage of a light weight (approximately 78 g) low cost FLIR attachment with smart-phone. In a recent work [22], Vidas et al. represent a 3D thermal mapping to monitoring building interiors using Microsoft Kinect [23] and a thermal camera. In computer vision and robotics, the use of RGBD cameras like Microsoft Kinect facilitates the development of techniques for highly-detailed and spatially-extended reconstructions [24, 25]. Such costly and bulky coupled sensors are capable of reconstructing in real-time [26], but the use of structured light pattern make the product usage limited within indoor environment and short range measurements. The working environment along with cost, dimension and weight are the main drawbacks for Kinect to be used as a light weight low cost system. Though active depth sensors have many advantages, there are certain scenario where passive RGB cameras are preferred due to its low power consumption, outdoor capable and form factor. This has motivated many researchers to investigate methods for 3D reconstruction using only passive cameras.

Industrial thermal cameras are capable of measuring the temperature accurately from a specified distance and few costly cameras provide dimension of the heated regions in 2 dimensions. FLIR smart-phone thermal attachment is also providing information in 2 dimensional spaces. Volumetric measurements are limited due to 2 dimensions. The cost of thermal cameras is another metric which restrict these products to be used only in industrial segment. FLIR smart-phone thermal attachment brings the opportunity to be used as house hold product for everyday life usage due to the cost reduction, increase mobility for small

---

[5] http://www.flir.com.hk/flirone/.

dimensions and weight and finally user friendly instead of expensive or bulky thermal systems. Automatic Volumetric measurement requires the heat analysis on 3 dimensions, so there are limitations in state of the art for an autonomous affordable system which is capable of area or volumetric measurement of any heated regions.

3.3 Machine Audition

Machine audition is of great importance when machine vision and visual odometry cannot work effectively because of the poor lighting condition or the target is not within the field of view. Acoustic source detection, localization and profiling can be an important part of unobtrusive sensing. In general the sound that a machine listens to consists of not a single sound source but multiple sound sources and finally should process all of these data to extract the pertinent information. Biologically inspired sound localization systems can be built by making use of an array of microphones, which are connected to a processor. In addition, such a system of microphones can be made to extract any particular sound from multiple sources produced simultaneously by several sources. Although the spatial resolution is relatively low compared with that of vision, audition has several unique properties—(a) vision occlusion never happens in audition (b) it works well equally in darkness (c) it is omnidirectional and (d) it has high time resolution and low computation overhead. Because of these unique properties mentioned, a 3D audio sensing system can be of great utility. Acoustic 3D imaging using the audio sensing can be another application. The techniques for acoustic based imaging of objects is well established and takes precedence in situations where optic based systems fail due to strong attenuation of EM waves e.g. underwater SONAR and imaging in dispersive medium e.g. smoke haze environment.

3.4 Tactile Sensing and Machine Olfaction

Once a region-of-interest has been identified in 3D space using techniques outlined the above sections, a moving machine with all these sensors (e.g. a robot or a drone) can be navigated to go in proximity of the region. The optical, thermal, audio, and smell sensors would get better resolution data as the robot moves nearer. However, at a touching distance from the region, a new sensor in form of tactile sensing can come in play to deduce further details just like a human being does by moving closer and touching by hand. Newly reported opto-tactile sensors can be used for object shape, size and surface assessment and can enhance 3D object model in mobile robots.

Smell sensing can be applied to Gas localization and gas distribution mapping. Such systems can comprise of an on-board array of gas sensors in a mobile robot equipped with electronic noses. According to state of the art, statistical methods to build 3D-DepthMap using three e-noses mounted at different heights in a mobile robot is becoming possible. The gas distribution model (GDM) can be improved with wind measurement obtained by ultrasonic anemometer.

4 Proposed Integrated Sensing System

As part of the fusion of multi-sensor information [27, 28], we propose to augment 3D robotic vision using odometry data and ultrasonic/laser range finder readings together with gas sensing and develop a probabilistic framework for simultaneous localization and
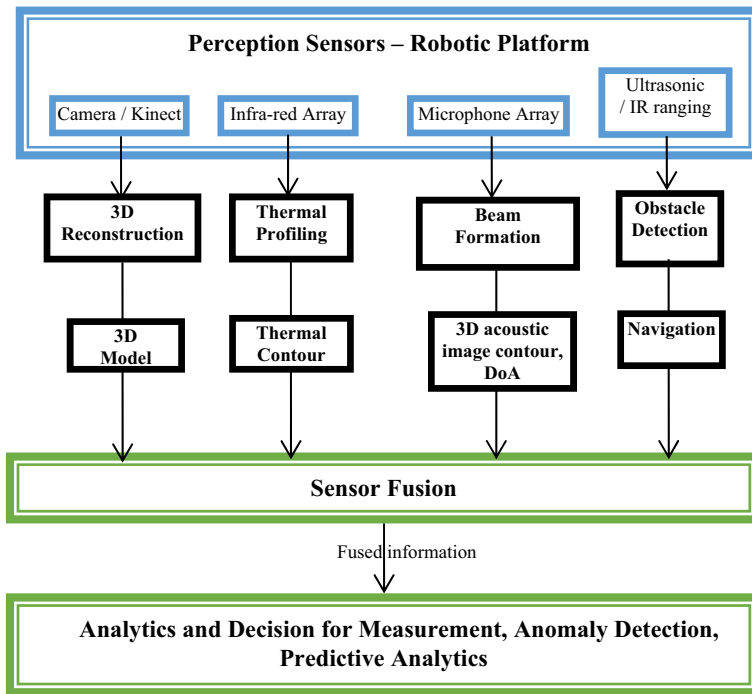
Fig. 1 Proposed robotic platform with multi-sensory perception

mapping [29]. Schematic representation of a multi sensor robotic platform is depicted in Fig. 1.

In our system, we present an end to end framework capable of generating 3D reconstruction of an environment based on the image/video captured through a remote platform mounted on a two wheel based robot. This work is a core part of our system presented in [27]. Our experimental results show that our technique is efficient and robust to a variety of indoor and outdoor environment scenarios with different scale and size. In our work, Firebird VI robot [30] (refer Fig. 2) is used whose all operations are controlled through Robot Operating System (ROS) [31]. The framework is capable to work on any robotic platform that supports ROS and Firebird VI is chosen due to its low cost and ready availability.

In the system an off-the-shelf webcam is used as 2D image capture device. The camera is calibrated and mounted on a servo on the robot as shown in Fig. 2. The servo allows the camera to pan and tilt which, in turn helps to capture the surrounding environment. Odometer and IMU sensor data is also captured simultaneously with the images in a time synchronized way. The captured information is pushed back to a backend server from where user is controlling the robot. Odometry and IMU sensor data are used for robot localization along with camera pose estimation [13]. Multi-view geometry [32] is used for creating a 3D map [11] of the environment. The details of the entire process are explained in [33].

Additionally for thermal imaging, we present a cost effective 3D thermal mapping system capable of area or volumetric measurement of heat in a continuous and non-

Fig. 2 Firebird VI robotic
platform

invasive way. Initial work with handheld smart-phone based system is presented in [34].
FLIR thermal attachment with a smartphone is mounted on the Firebird VI robotic plat-
form and it is calibrated. So the webcam and the FLIR thermal attachment pair are
calibrated to produce a three dimensional environment with thermal annotation on the
surface areas. This system is thus capable of measuring the thermal area or volume at any
abnormally heated region. Three dimensional reconstructions are carried out using pre-
viously described method as in Sects. 3.1 and 3.2. Back projection is used for thermal
overlay on surfaces of the reconstructed environment.

For an automatic quality inspection system for manufacturing industry, we use our
proposed robotic platform with a Kinect mounted as three dimensional depth sensor. The
capture process is executed with a Kinect connected with an onboard computing unit and
the captured three dimensional data is stored in the form of point cloud. Kinect is kept at a
distance about 0.8 m for best performance along with controlled lighting. The component
objects are with different shapes and size with variety of measurable attributes, so the
implemented algorithms are customized according with the shape and geometry of the
objects. The object is segmented from the background by color and shape after proper
noise removal and clean-up. We are able to measure different geometrical properties of
objects and Point Cloud Library (PCL) is used intensively at the different steps of our
implementation.

We also added an audition system [35] on top of the Fire Bird VI based mobile robot
platform by using Time Delay of Arrival (TDOA) information. We used a four-micro-
phone array with 24-bit analog-to-digital converter (ADC) including acoustic echo can-
cellation and noise suppression producing 16 kHz; 24-bit pulse code modulation (PCM)
modulated audio. The system receives time series signals and processes TDOA between
adjacent microphones using cross-correlation, phase transformation and maximum likeli-
hood techniques. It has been observed that classical cross correlation method is easy to
implement but it is sensitive to noise and also performance degrades if reverberation time is
>150 ms. PHAT normalizes the cross power spectrum of the signals. It offers the
advantage by minimizing the spreading of the peak of the correlation function. It removes
all energy content from the cross spectrum and provide very sharp correlation peak.

Finally, for acoustic imaging, we use an acoustic array in our system [27] which
localizes acoustic sources and provides visualization augmentation for the existing
machine vision capability. The basic principle lies in electronic beam-steering and beam-
shaping where the beam is steered in both azimuthal and elevation directions. Using 2D
acoustic array, all ultrasonic waveforms are transmitted using proper phase-shifts at each
antenna element so that the combined beam is steered in both azimuth and elevation

directions, covering the complete target. On the other hand ultrasonic receivers collect the reflected wave across the two distinct dimensions. Reconstruction is carried out using phase correction factors.

5 Results and Discussion

5.1 Optothermal Vision

3D environment creation implementation environment consisted of Firebird VI robot as shown in Fig. 1 and a back-end system having Intel(R) Xeon(R) E5606 processor running at 2.13 GHz along with a NVIDIA Tesla C2050 Graphic Card. One ZOTAC ZBOXHD-ID11 is mounted on top of the robot. ROS hydro is installed inside Ubuntu 12.04 LTS in all the systems. The entire capture task was run on the ZOTAC box mounted on the Firebird VI. Image is captured with 640×480 resolution using the Logitech C920 webcam. The 3D model reconstruction is carried out on the backend system due to less processing power of ZOTAC box.

In Fig. 3 we present result of the 3D reconstruction where the data is captured in a living room. The three sides of the room are captured where different objects are placed. The dimension of the room is about 13×11 feet. The user can guide the robot to go closer and capture the frames to produce a more accurate and dense points in any required zone.

For opto-thermal mapping, we present a sample heat measurement to demonstrate the usability of our solution. Figure 4 shows a sample 3D thermal point cloud of a mug having hot water inside. The presence of water is not detectable through RGB image but corresponding thermal image shows distinct temperature differences in hot regions. The testing is performed using 11 images with two iterations. The 3D thermal cloud shows the structure of the mug along with the hot region. The idea is to measure the volume of hot water present inside. The mug is segmented by the knowledge of its cylindrical shape. The segmentation finally is used to detect the dimension of the mug. Mostly heated region is extracted from the temperature and the dimension is calculated from 3D structure. The volume of hot water present is calculated as 63.7 cc by our method against actual volume of 68 cc.

For 3D vision based measurement for quality control, we present a comparative measurement result of different automobile component parts using Kinect. The experiments are carried out with every single object several times to test the stability of the implemented system. The presented results are the average of the repeated measurements and are presented in Table 1. We could reach the sub millimeter level accuracy. The tolerance of our measurement method is seen to be about ±0.5 mm.

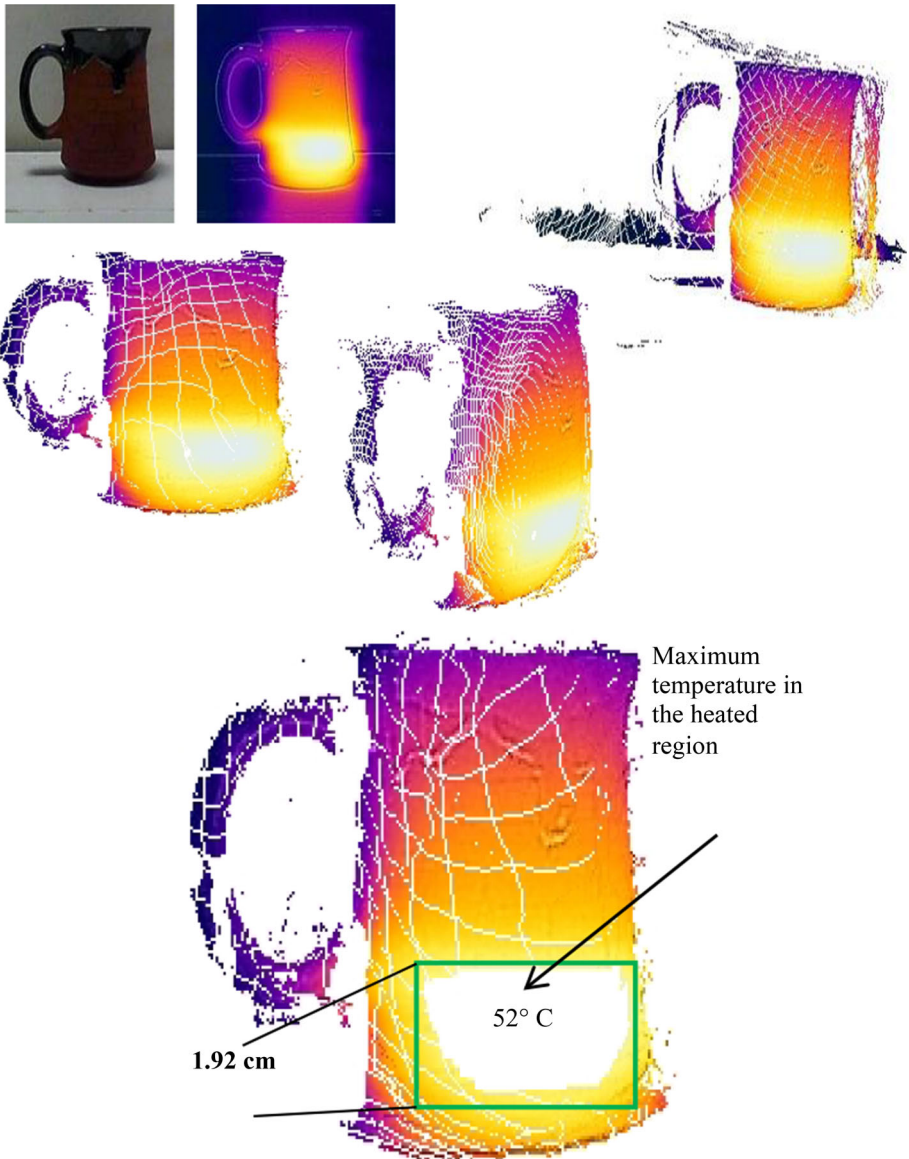**Fig. 3** 3D reconstruction using robot locomotion

Fig. 4 Volumetric measurement: *top to bottom* shows captured RGB, thermal image, segmented 3D thermal model, detected dimension of the heated region

5.2 Acoustic Sensing

We varied frequencies, ping durations, source distances and azimuths as given in Table 2 to know the effect of these parameters in finding DOA. From Table 2, the parameter 'Pings-Pause' of value 'f0.006, 0.003 g' indicates the time series signal generated with pause duration 0.003 s (48 samples) followed by sinusoidal ping signal of duration 0.006 s

Table 1 Comparative result for auto component measurement

Objects	Dimensions	Ground truth using vernier (mm)	Measured value using Kinect (mm)
	OD outer	67.992	68.54
	ID inner	16.97	16.26
	Length	109.78	110.39
	Width	71.306	71.08
	ID	50.661	50.76
	ID1	10.439	10.15
	OD1	30.056	29.47
	ID	11.937	11.73
	OD	32.009	32.17

(96 samples). The accuracy of the localization techniques is dependent on the relative bearing of the acoustic source and greatly improves for sources placed at an angle of 90° wr.t. microphone array i.e. broadside as compared to others. The platform locomotion and

Table 2 Different parameters of signals

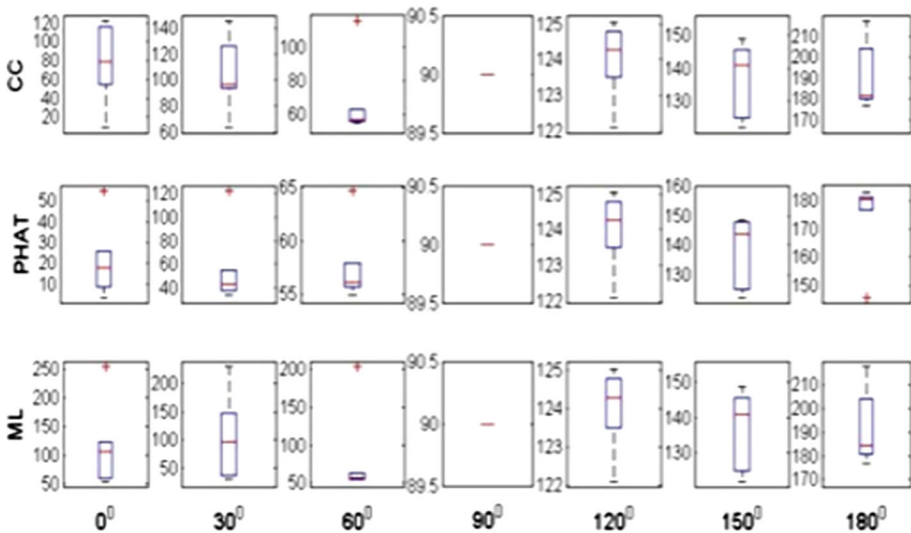| Parameters	Values
Frequency (Hz)	100, 300, 500, 1000, 3000, 5000
Pings-pause (s)	{0.003, 0.003}, {0.006, 0.003}, {0.003, 0.006}, {0.03, 0.03}, {0.036, 0.03}, {0.03, 0.06}, {0.3, 0.3}, {0.6, 0.3}, {0.3, 0.6}, {3, 3} {6, 3}, {3, 6}
Distance (m)	1, 1.5, 2
Azimuths (°)	0, 30, 60, 90, 120, 150, 180
Microphone pairs and spacing (m)	{1, 2} > 0.149, {1, 3} > 0.189, {1, 4} > 0.226, {2, 3} > 0.040, {2, 4} > 0.077, {3, 4} > 0.037
Algorithms	CC, PHAT, ML

Fig. 5 *Boxplots* of DoA's for CC, PHAT and ML. *Signal Frequency* 1 kHz, *Distance* 1 m, *Samples* 4800, *Pause* 9800

a prior 3D environmental modeling can now be leveraged to localize and track the source efficiently. The stopping criteria of the mobile robot towards reaching the source is carried out by using the range information (i.e. distance of microphones from the acoustic source) obtained from 3D model structure generated from low cost optic camera.

The performance of each localization algorithm is evaluated by comparing the original DOA and the estimated DOA for all ping durations at different frequencies and for different distances. It is evident from Fig. 5, that PHAT algorithm always outperforms CC and ML.

For acoustic imaging, a fully populated 2D planar array ($\lambda/2$ spacing) of ultrasonic transducers has been designed in a 4×4 grid with electronic beam steering in both azimuth and elevation. A target is acoustically reconstructed at a distance of 5.0 m from the acoustic sensor array (ASA) using pulsed-CW 40 kHz signal. The simulation results of the performance of such a system are represented Figs. 6 and 7 below. Performing regression tests on various combinations of parameters like total number of array elements,
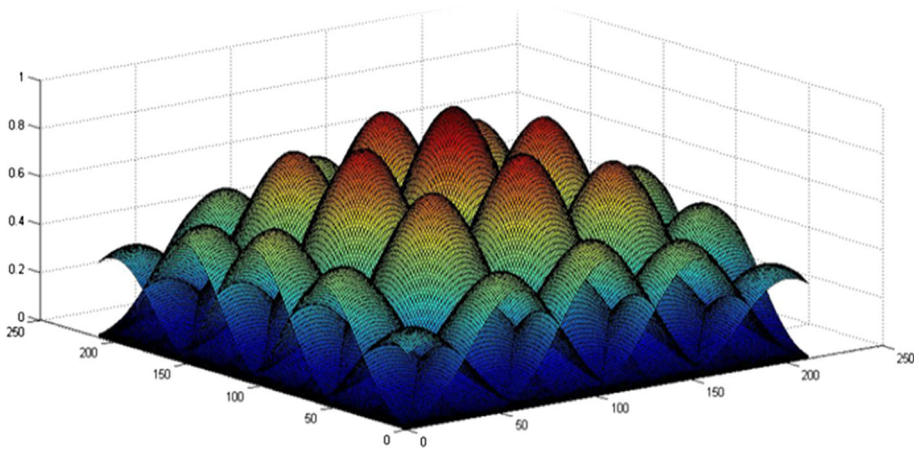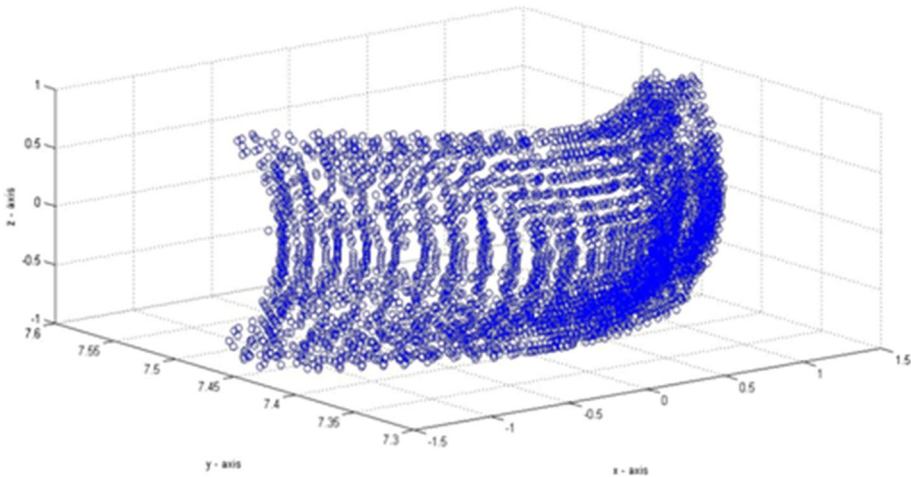
Fig. 6 Sample beam pattern of the entire array

Fig. 7 Acoustic shape of a surface. *No of elements* 16, *Frequency of operation* 40 kHz, *Pulse duration* 1 ms

frequencies of insonification etc. an optimal set of values of the parameters are arrived at. Such an opto-acoustic 3D machine vision can provide a new dimension of sensing mechanism when robots will be deployed in an uncertain environment which is visually occluded due to smoke, fumes, dust and chemical vapor contents.

6 Conclusion

In this paper, we have first introduced the concept of human-like five senses computing and have outlined how it can be applied for remote inspection and analytics. We have discussed specific scenarios where 3D vision, audio and thermal sensing can be used for

measurement, quality control and predictive analytics use cases and have presented results of some working prototypes. As seen from the results, the sensing part yields good outcome. In future the collected data can be fed into artificial intelligence engines backed by machine learning and that can indeed herald a new era of automation in remote inspection and predictive analytics space.

References

1. Shirmohammadi, S., & Ferrero, A. (2014). Camera as the instrument: The rising trend of vision based measurement. *Instrumentation & Measurement Magazine, IEEE, 17*(3), 41–47.
2. Kundu, T. (2014). Acoustic source localization. *Ultrasonics, 54*(1), 25–38.
3. Ibarra-Castanedo, C., Sfarra, S., Genest, M., & Maldague, X. (2015). Infrared vision: Visual inspection beyond the visible spectrum. In *Integrated imaging and vision techniques for industrial inspection* (pp. 41–58). London: Springer.
4. Araki, H., & Omatu, S. (2015). A proposition and an evaluation for compact multiple odor scan system. In *Control conference (ASCC), 2015 10th Asian* (pp. 1–6). IEEE.
5. Ha, D., Sun, Q., Kaiqi, S., Wan, H., Li, H., Xu, N., et al. (2015). Recent achievements in electronic tongue and bioelectronic tongue as taste sensors. *Sensors and Actuators B: Chemical, 207*, 1136–1146.
6. Noor, A. K. (2015). Potential of cognitive computing and cognitive systems. *Open Engineering, 5*(1), 75–88.
7. Pradeep, V., Rhemann, C., Izadi, S., Zach, C., Bleyer, M., & Bathiche, S. (2013). MonoFusion: Real-time 3D reconstruction of small scenes with a single web camera. In *The 13th IEEE international symposium on mixed and augmented reality* (pp. 83–88).
8. Pizzoli, M., Forster, C., & Scaramuzza, D. (2014). REMODE: Probabilistic, monocular dense reconstruction in real time. In *IEEE international conference on robotics and automation (ICRA), Hong Kong* (pp. 2609–2616).
9. Pollefeys, M., Nister, D., Frahm, D. J. M., Akbarzadeh, A., Mordohai, P., Clipp, B., et al. (2008). Detailed real-time urban 3D reconstruction from video. *International Journal of Computer Vision, 78*(2–3), 143–167.
10. Newcombe, R. A., Lovegrove, S. J., & Davison, A. J. (2011). DTAM: Dense tracking and mapping in real-time. In *ICCV* (pp. 2320–2327).
11. Saha, A., Bhowmick, B., & Sinha, A. (2014). A system for near real-time 3D reconstruction from multi-view using 4G enabled mobile. In *Proceedings of the 2014 IEEE international conference on mobile services*. IEEE Computer Society.
12. Tanskanen, P., Kolev*, K., Meier, L., Paulsen, F. C., Saurer, O., & Pollefeys, M. (2013). Live metric 3D reconstruction on mobile phones. In *ICCV* (pp. 65–72).
13. Bhowmick, B., Mallik, A., & Saha, A. (2014). Mobiscan3D: A low cost framework for real time dense 3D reconstruction on mobile devices. In *IEEE 11th international conference on Ubiquitous intelligence and computing* (pp. 783–788).
14. Mallik, A., Bhowmick, B., & Alam, S. (2015). A multi-sensor information fusion approach for efficient 3D reconstruction in smart phone. In *International conference on image processing, computer vision, and pattern recognition (IPCV)* (pp. 291–298).
15. Davison, A. (2003). Real-time simultaneous localisation and mapping with a single camera. In *IEEE international conference on computer vision* (pp. 1403–1410).
16. Gonzalez, R. C., Valdés, R., & Cancelas, J. A. (2001). Vision based measurement system to quantify straightness defect in steel sheets. In *9th International conference on computer analysis of images and patterns* (pp. 427–4341). Berlin: Springer, ISBN 3-540-42513-6.
17. Mottaa, J. M. S. T., Carvalhob, G. C., & McMasterc, R. S. (2001). Robot calibration using a 3D vision-based measurement system with a single camera. doi:10.1016/S0736-5845(01)00024-2.
18. Shirmohammadi, S., & Ferrero, A. (2014). Camera as the instrument: The rising trend of vision based measurement. *Instrumentation & Measurement Magazine, IEEE, 17*(3), 41–47.
19. Rövid, A. (2013). Machine vision-based measurement system for vehicle body inspection. *Acta Polytechnica Hungarica*. doi:10.12700/APH.10.05.2013.5.9.
20. Bhanu, B., & Holben, R. D. (1990). Model-based segmentation of FLIR images. *Aerospace and Electronic Systems, IEEE Transactions on, 26*(1), 2–11.

21. Borrmann, D., Nuchter, A., Dakulovic, M., Maurovic, I., Petrovic, I., Osmankovic, D., & Velagic, J. (2012). The project ThermalMapper thermal 3D mapping of indoor environments for saving energy. In *Proceedings of the 10th international IFAC symposium on robot control (SYROCO)* (Vol. 10).
22. Vidas, S., Moghadam, P., & Bosse, M. (2013). 3D thermal mapping of building interiors using an RGB-D and thermal camera. In *Proceedings of the IEEE international conference on robotics and automation*.
23. https://dev.windows.com/en-us/kinect. Accessed on November 26, 2015.
24. Meilland, M., & Comport, A. (2013). Super-resolution 3D tracking and mapping. In *IEEE international conference on robotics and automation*.
25. Whelan, T., Johannsson, H., Kaess, M., Leonard, J., & McDonald, J. (2013). Robust real-time visual odometry for dense RGB-D mapping. In *IEEE international conference on robotics and automation (ICRA)*.
26. Newcombe, R. A., Davison, A. J., Izadi, S., Kohli, P., Hilliges, O., Shotton, J., Molyneaux, D., Hodges, S., Kim, D., & Fitzgibbon, A. (2011). KinectFusion: Real-time dense surface mapping and tracking. In *Proceedings of the ISMAR*.
27. Deshpande, P., Reddy, V. R., Saha, A., Vaiapury, K., Dewangan, K., & Dasgupta, R. (2015). A next generation mobile robot with multi-mode sense of 3D perception. In *Advanced robotics (ICAR), 2015 international conference on* (pp. 382–387).
28. Vempada, R., Deshpande, P., Vaiapury, K., Saha, A., Dewangan, K., Das Gupta, R., & Pal, A. (2015). Sound source localization with 3D optical fusion for hazardous area surveillance using autonomous ground vehicles. In *Conference: ICRA 2015, at Seattle USA*.
29. Lilienthal, A. J. et al. (2009). A statistical approach to gas distribution modeling with mobile robots—the Kernel DM + V algorithm. In *Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS)* (pp. 570–576).
30. Firebird VI. http://www.nex-robotics.com/fire-bird-vi-robot-platform.html. Access date October 20, 2015.
31. Martinez, A., & Fernández, E. (2013). *Learning ROS for robotics programming*. Packt Publishing Ltd., ISBN 978-1-78216-144-8.
32. Hartley, R., & Zisserman, A. (2003). Multiple view geometry in computer vision. Cambridge, MA: Cambridge University Press, ISBN 0-521-54051-8.
33. Dewangan, K., Saha, A., Vaiapury, K., & Dasgupta, R. (2015). 3D environment reconstruction using mobile robot platform & monocular vision. In *Proceedings of international conference on advanced computing and communication technologies*, Panipat, India.
34. Saha, A., Dewangan, K., Dasgupta, R. (2016). 3D thermal monitoring and measurement using smartphone and IR thermal sensor. In *11th joint conference on computer vision, imaging and computer graphics theory and applications (VISIGRAPP)* (Vol. 3, pp. 696–702). ISBN 978-989-758-175-5.
35. Reddy, V. R., Deshpande, P., & Dasgupta, R. (2015). Robotics audition using Kinect. In *Automation, robotics and applications (ICARA), 2015 6th international conference on* (pp. 34–41).

Arpan Pal received both his B.Tech. and M.Tech. from Indian Institute of Technology, Kharagpur, India in Electronics and Telecommunications and Ph.D. from Aalborg University Denmark. He has more than 23 years of experience in the area of Signal Processing, Communication and Real-time Embedded Systems. Currently he is with Tata Consultancy Services (TCS), where, as Principal Scientist, he is heading research at Innovation Lab, Kolkata. He is also a member of Systems Research Council of TCS. His current research interests include Mobile phone and Camera based Sensing and Analytics, Physiological Sensing, M2M communications and Internet-of-Things based Platforms and Systems. He had worked earlier in Missile Seeker Signal Processing at DRDO and also in Interactive TV and Set-top boxes. Arpan has over 75 publications till date in reputed Journals and Conferences along with a couple of Book Chapters. He has also filed for more than 70 patents and has 10 patents granted to him. He is on the editorial board for reputed journals like ACM Transactions on Embedded Computing Systems, IEEE Transactions on Emerging Topics in Computing and IT Professional Magazine from IEEE Computer Society. He is in the program committee of various eminent conferences and is a senior member of IEEE.

Ranjan Dasgupta received M.E. degree in Software Engineering from Jadavpur University. Starting his carrier as a R&D engineer he has been working with various embedded system design, its hardware and software implementation in the area of robotics control, medical electronics and wireless applications at consumer, healthcare and telecommunication industry in Philips Medical Systems, Philips Telecommunications. Since last 10 years he is working in TCS Innovation Lab and playing a role of Senior Scientist where he is initiating, coordinating several research projects in the area of IoT, robotics and Embedded Systems. With his 18 years of R&D expertise, he has also contributed in IEEE/ACM transactions and conference papers, technical documents, and company internal white papers. He used to be a speaker in several national and international conferences. He is a reviewer of ACM Transactions on Embedded Computing Systems and program/technical or organizing committee member of several national and international conferences. He is a Member of IEEE, ACM, CSI and RSI and NZRAS.

Arindam Saha has received B.Tech. in Computer Science and Engineering from Kalyani University, West Bengal, India, in July-2004 and M.Tech. in Software Engineering from Jadavpur University, Kolkata, India, in July-2014. Since 2008, he is associated with the Innovation Labs, TCS, Kolkata, working as a scientist R&D in the field of Computer vision & Robotic vision, Embedded Signal Processing. He worked in Tata Elxsi Ltd and Ingenient Technology in the field of embedded domain of set-top-boxes for totally almost 4 years. He has published 9 papers in referred international conferences and journals in which most of them are IEEE sponsored. His current research interest includes computer and robotic 3D vision. Mr. Saha is a member of IEEE.

Bhaskar Nandi has received his B.E. Degree in Production Engineering from BIT Mesra, Ranchi in the 1996–1997. He has worked in the different areas of the Automotive Industries namely Plant Quality, Vendor Development and Supplier Quality, New Product Development, Product Innovation & Product Quality, Market Quality, Rejection Analysis and Cost Control, Plant Maintenance Processes, Field Service and Customer Care. He has served noted OEMs like Honda Motors, Hyundai Motors, Mahindra and Mahindra Ltd, and Hitachi Tohoku. He was also attached as Research Associate with IIT Delhi on Testing and Development of IC Engines. His interest area lies in Exploring and Developing New Technologies and Inventive Processes for the Improvement of the Automotive Industry. He is currently working with the Innovation Lab at Tata Consultancy Services, providing Industry Solutions for Smart Manufacturing Processes. His interest area lies in Exploring and Developing New Technologies and Inventive Processes for the Improvement in the Automotive Industry.