We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

6,600
Open access books available

177,000
International authors and editors

195M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us? Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected. For more information visit www.intechopen.com
Abstract
In this chapter we present a hidden Markov model (HMM) based framework for situational awareness that utilizes multi-sensor multiple modality data. Situational awareness is a process that comes to a conclusion based on the events that take place over a period of time across a wide area. We show that each state in the HMM is an event that leads to a situation and the transition from one state to another is determined based on the probability of detection of certain events using multiple sensors of multiple modalities - thereby using sensor fusion for situational awareness. We show the construction of HMM and apply it to the data collected using a suite of sensors on a Packbot.

1. Introduction
Situation awareness (SA) is a process of conscious effort to process the sensory data to extract actionable information to accomplish a mission over a period of time with or without interaction with the sensory systems. Most of the information is time dependent and they usually follow a sequence of states. This is where the Markov or hidden Markov models are useful in analyzing the data and to extract the actionable information from the sensors. To gain better understanding, the following section would elaborate on situation awareness.

1.1 Situation Awareness
Situational awareness means different things to different people. Experience plays a great role in the situational awareness. Based on one’s experience, the interpretation of the situation will be different. For example, in the case of animal world, the situation assessment by the predator and prey will be different. The predator assesses the situation based on the past experience, circumstances, etc., and determines when to strike. Similarly, the prey assesses its situation based on its experience and determines the best route to take to escape from the imminent danger. The origins of SA are in the military (Smith, 2003) back in 1970’s. Initial work is done in the area of analyzing and understanding what a pilot is observing and how he is making decisions based on the data provided to him in the cockpit and what he/she is able to observe outside through the windows. Some of it resulted in the design of modern cockpit and flight training facilities. The US Army defines the SA as$^1$:

$^1$http://www.army.mil/armyBTKC/focus/sa/index.htm
Situational Awareness is the ability to generate actionable knowledge through the use of timely and accurate information about the Army enterprise, its processes, and external factors.

Endsley and Garland (Endsley & Mataric, 2000) defines SA as “SA is knowing what is going around you”. There is usually a torrent of data coming through the sensors, situational awareness is sifting through all that data and extracting the information that is actionable and predicting the situation ahead. The awareness of the situation ahead lets one plan the data collection from the right set of sensors. SA allows selective attention to the information. Some other pertinent definitions are provided here (Beringer & Hancock, 1989):

SA requires an operator to “quickly detect, integrate and interpret data gathered from the environment. In many real-world conditions, situational awareness is hampered by two factors. First, the data may be spread throughout the visual field. Second, the data are frequently noisy” (Green et al., 1995).

Situation awareness is based on the integration of knowledge resulting from recurrent situation awareness (Sarter & Woods, 1991).

“Situation awareness is adaptive, externally-directed consciousness that has as its products knowledge about a dynamic task environment and directed action within that environment” (Smith & Hancock, 1995).

In a sensor world, situation awareness is obtained by gathering data using multi-modal multiple sensors distributed over an area of interest. Each modality of sensor obtains the data within its operating range. For example video observes the data within its field of view. Acoustic sensors record the sound within its audible (sensitive) range. In this chapter, several sensor modalities will be considered and the data they present. Proper information from each sensor or from a combination of sensors will be extracted to understand the scene around. Extraction of the right information depends mostly on previous knowledge or previous situation awareness. Understanding of the contribution of each sensor modality to the SA is key to the development of algorithms pertinent to the SA. Clearly, the information one would like to obtain for SA depends on the mission. In order to help us better understand the functionality of each modality, three different missions are considered as exemplars here, namely, (a) urban terrain operations, (b) difficult terrain such as tunnels, caves, etc., and (c) battlefield.

1.1.1 Urban Terrain Operations
Since World War II, nation building after war has become a common practice, partly, to ensure the vanquished country does not become a pariah nation or some dictator does not take hold of the country. After World War II, Marshal plan was developed to help the countries. Recently, after Iraq war, coalition partners (US and UK) stayed back in Iraq to facilitate smooth functioning of the Iraqi government. However, the presence of foreign troops always incite mixed feelings among some people and may become the cause for friction resulting in urban war or operations. Moreover by year 2020, 85% of world’s population live in the coastal cities (Maj. Houlgate, 2004) which cause friction among various ethnic groups that call for forces to quite the upraising necessitating the urban military operations. In general, the urban operations include (Press, 1998):

• Policing operations – to deter violence
• Raids
- Evacuation of embassies
- Seize ports and airfields
- Counter weapons of mass destruction (WMD)
- Seize enemy leaders

- Sustained urban combat

From the above list of operations that may take place in an urban area, clearing of buildings and protecting them is one of the major missions. Often, once a building is cleared, one may leave some sensors in the building to monitor the building for intruders. Another important operation is perimeter protection. In the case of perimeter protection, several sensors will be deployed around the perimeter of a building or a place. These sensors detect any person approaching the perimeter and report to the command center for further investigation and action. Next we consider operations in difficult terrain.

1.1.2 Operations in Difficult Terrain

In general, terrorists take advantage of the rugged terrain and often hide in the caves in the mountain range or bunkers in the ground. There are too many hiding places and one can not just walk in to these areas without risking their own lives. The operations required in these areas are quite different from those conducted in the urban areas. Often, one would send a robot equipped with sensors to monitor if there is any human activity in the caves/tunnels or to find any infrastructure, man made objects, etc.

Borders between warring nations and between rich and poor nations have become porous for illegal transportation of people, drugs, weapons, etc. Operations in these areas include:
(a) detection of tunnels using various sensing modalities and (b) making sure that the tunnels remain cleared once they are cleared. Detection of tunnels require different kind of sensors.

1.1.3 Operations in open battlefield

This is the traditional cold war scenario where the war is fought in an open area. Here the situation awareness requires knowing where the enemy is, how big is the enemy, where the firing is coming from, and the type of weapons used, etc. Furthermore, one would like to know, not only the firing location but also the impact point of the mortars and rockets. The launch location helps in taking action to mitigate the enemy and its firing weaponry, etc., and the knowledge of impact location helps in assessing the damage to provide the necessary medical and other support to control and confine the damage.

Clearly, the requirements for different operations are different. To be successful in the operations, one need to have a clear understanding of the situation. Situation awareness comes from the sensors deployed on the ground and in the air, and human intelligence. The sensor data is processed for right information to get the correct situation awareness. The next section presents various sensors that could be used to monitor the situation.

1.2 Sensor Suite for Situational Awareness

Traditionally, when the subject of sensors comes up, immediately, Radar, and video sensors come to one’s mind. With the advent of very large scale integrated (VLSI) circuits, other sensor modalities have been developed and used extensively in modern times. Main reasons for development of new sensor modalities are: (a) limited capability of existing sensors, (b) high
power consumption by traditional sensors, (c) wide area of operation requiring many sensors, (d) limited field of view by Radar and video and (e) new modalities offer better insight into the situation. Most of the sensors for situation awareness are deployed in an area of interest and left there for days, weeks, and months before attending to them. This necessitated the need for low power, low cost and large quantities of sensors that could be deployed in the field.

Now, we will present some of the sensors that may be deployed in the field and discuss their utility.

**Acoustic Sensors:** While the imaging sensors (for example: camera, video) act as the eyes, the acoustic sensors fulfill the role of ears in the sensing world. These microphones capture the sounds generated by various events taking place in their vicinity, such as, a vehicle traveling on a nearby road, mortar/rocket launch and detonations, sound of bullets whizzing by and of course sounds made by people, animals, etc., to name few. These are passive sensors, that is, they do not transmit any signals unlike the Radar, hence they can be used for stealth operations. There are several types of microphones, namely, condenser, piezoelectric, dynamic, carbon, magnetic and micro-electro mechanical systems (MEMS) microphones. Each microphone has its own characteristic response in terms of sensitivity to the sound pressure and the frequency of operation. Each application demands a different type of microphone to be used depending on the signals that are being captured by the microphone. For example, detection of motor vehicles require the microphones that have the frequency response equal or greater than the highest engine harmonic frequency. On the other hand to capture a transient event such as a shock wave generated by a super sonic bullet require a microphone with frequency response of 100 kHz or more. When the microphones are used in an array configuration, such as, linear, circular or tetrahedral array, the signals from all the microphones can be processed for estimating the angle of arrival (AoA) of the target. Figure 1 shows a single microphone
and a tetrahedral array. The microphones in the tetrahedral array Figure 1b are covered by foam balls to reduce the wind noise.

**Seismic Sensors**: These are also called geophones. These sensors are used to detect the vibrations in the ground caused by the events taking place in the sensing range of the sensors. Just as in the case of acoustic sensors, the seismic sensors are passive sensors. Typical applications for these sensors include (a) detection of vehicles (both civilian and military vehicles) by capturing the signals generated by a moving vehicles, (b) perimeter protection – by capturing the vibrations caused by footsteps of a person walking, (c) explosion, etc. The Indonesian tsunami in December 2004 was devastating to the people. However, several animals sensed the vibrations in the ground caused by the giant waves coming to the shore and ran to the hills or elevated areas and survived the tsunami. Figure 2 shows different seismic sensors. The spikes are used to couple the the sensor to the ground by burying the spikes in the ground.

![Fig. 2. Different seismic sensors](image)

**Magnetic Sensors**: Magnetic (B-field) sensors can be used to detect ferromagnetic materials carried by people, e.g., keys, firearms, and knives. These sensors may also detect the usage of computer monitors. There are several types of magnetic sensors, namely, (a) flux gate magnetometer and (b) coil type magnetic sensor. The coil type magnetic sensor has high frequency response compared to the flux gate magnetometer. One can use multiple sensors in order to detect the flux change in all three X, Y and Z directions. The sensitivity of the magnetic sensor depends on the type and as well as the construction of the sensor. Figure 3 shows two types of magnetic sensors.

![Fig. 3. (a) Flux gate magnetometer, (b) Coil type magnetic sensor](image)
**Electrostatic or E-field Sensors:** These are passive sensors that detect static electric charge built-up on the targets or any electric field in the vicinity of the sensor. Some of the sources of the static electric charge are (a) clothes rubbing against the body, (b) combing hair, and (c) bullet or projectile traveling in the air builds up charge on the bullet, etc. All the electric transmission lines have electric field surrounding the lines – this field gets perturbed by a target in the vicinity – and can be detected by E-field sensors. Figure 4 shows some of the E-field sensors that are commercially available.

![E-field sensors](image)

**Passive Infrared (PIR) Sensor:** These are passive sensors that detect infrared radiation by the targets. These are motion detectors. If a person walks in front of them, the sensor generates an output proportional to the temperature of the body and inversely proportional to the distance between the person and the sensor. Figure 5 shows a picture of PIR sensor.

![Passive Infra Red sensor](image)

**Chemical Sensor:** These sensors are similar to the carbon monoxide detectors used in buildings. Some of the sensors can detect multiple chemicals. Usually, these sensors employ several wafers. Each wafer reacts to a particular chemical in the air changing the resistivity of the wafer. The change in the resistivity in turn changes the output voltage indicating the presence of that chemical.

![Chemical sensor](image)
Infra Red Imagers: There are several IR imagers depending on the frequency band they operate at, namely, long wave IR, medium wave IR, and forward looking infrared (FLIR). These sensors take the thermal image of the target in their field of view. A typical IR imager’s picture is shown in Figure 6.

Visible Imagers: These are regular video cameras. They take the pictures in visible spectra and have different resolution and different field of view depending on the lens used. Figure 6 shows a picture of a typical video camera.

In the next section, we present the description of the unattended ground sensors.

1.2.1 Unattended Ground Sensors

A typical unattended ground sensor (UGS) is a suite of multi-modal sensor package with a processor facilitating the collection of data from all the sensors and having a capability to process the data and extracting the information relevant to the mission. A typical UGS sensor consists of acoustic, seismic, magnetic and both IR and visible cameras. The non-imaging sensors are often called activity detection sensors. As the name implies, these sensors are utilized to detect any activity within the receptive field of the sensors, such as a person walking/running, vehicle moving, etc. Once the activity sensors detect a target, they cue the imaging sensors to capture a picture of the target which will be sent to the command control center. Target/activity detection algorithms run on the processor in the UGS system. There are algorithms running when to cue the imagers and which one of the pictures to transmit to the command and control center in order to reduce the bandwidth of the communication channel. In general activity detection sensors consume low power, hence reduce the power consumption by the UGS prolonging the battery life.

UGS are in general placed in the area of interest conspicuously and left to operate for several days or months. In general these are low power sensors that meant to last for several days or months before replacing the batteries. There are several manufacturers that make the UGS systems.
1.3 Techniques for Situational Awareness

In order to assess the situation, sensor information is needed. Based on the history of sensor information/output when a particular event took place, one can infer same event has taken place if similar information/output is observed. Such inference can be made using Bayesian nets or hidden Markov model. If several events are observed in sequence, then such a sequence of events can be modeled using Markov or Hidden Markov chain. In the following subsection, both Bayesian nets and Hidden Markov models will be described.

1.3.1 Bayesian Belief Networks

Bayesian belief networks (BBN) are directed acyclic graphical networks with nodes representing variables and arcs (links between nodes) representing the dependency relationship between the corresponding variables. Quite often, the relationship between the variables is known but cannot quantify it in absolute terms. Hence, the relationship is described in probabilistic terms. For example, if there are clouds then there is a chance of rain. Of course, there need not be rain every time a cloud is formed. Similarly, if a person walks in front of a seismic sensor, the sensor detects periodic vibrations caused by footfalls, however, if periodic vibrations are observed it does not mean there is a person walking. One of the uses of BBN is in situations that require statistical inference.

Bayesian methods provide a way for reasoning about partial beliefs under conditions of uncertainty using a probabilistic model, encoding probabilistic information that permits us to compute the probability of an event. The main principle of Bayesian techniques lies in the inversion formula:

\[ p(H|e) = \frac{p(e|H)p(H)}{p(e)} \]

where \( H \) is the hypothesis, \( p(e|H) \) is the likelihood, \( p(H) \) is called the prior probability, \( p(H|e) \) is the posterior probability, and \( p(e) \) is the probability of evidence. Belief associated with the hypothesis \( H \) is updated based on this formula when new evidence arrives. This approach forms the basis for reasoning with Bayesian belief networks. Figure 7 shows how the evidence is collected using hard and soft methods.

Nodes in Bayesian networks (Pearl, 1986; 1988) represent hypotheses, and information is transmitted from each node (at which evidence is available or belief has been updated) to adjacent nodes in a directed graph. Use of Bayesian rule for large number of variables requires estimation of joint probability distributions and computing the conditional probabilities. For example, if no assumption on the dependencies is made, that is, all variables are dependent on each other, then

\[ p(A, B, C, D, E) = p(A|B, C, D, E) p(B|C, D, E) p(C|D, E) p(D|E) p(E) \] (1)

If the dependencies are modeled as shown in Figure 8, then the joint probability distribution is much simpler and is given by

\[ p(A, B, C, D, E) = p(A|B) p(B|C, E) p(C|D) p(D) p(E) \] (2)

Let \( G(V, E) \) is a directed acyclic graph with a set of vertices \( V = \{v_1, v_2, \ldots, v_n\} \) and a set of edges \( E = \{e_{1,2}, e_{1,3}, \ldots, e_{i,j}\} \), with \( i \neq j \in \{1, 2, \ldots, n\} \). Note that the directed edge \( e_{i,j} \)
connects the vertex \( v_i \) to vertex \( v_j \) and it exists if and only if there is a relationship between nodes \( v_i \) and \( v_j \). Node \( v_i \) is the parent of node \( v_j \) and \( v_j \) is the descendant of node \( v_i \). Let us denote the random variable associated with the node \( v_i \) by \( X_{v_i} \). For simplicity, let us denote \( X_j = X_{v_j} \). Let \( pa(v_i) \) denote the parent nodes of the node \( v_i \). For a Bayesian belief network the following properties must be satisfied:

- Each variable is conditionally independent of its non-descendants
- Each variable is dependent on its parents

This property is called the local Markov property. Then the joint probability distribution is given by

\[
p (X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} p (X_i | pa(X_i))
\] (3)
Now it is possible to associate meaning to the links in the Bayesian belief network and hence what we need to specify to turn the graphical dependence structure of a BBN into a probability distribution. In Figure 8 the nodes labeled ‘sound’ and ‘human voice’ are related. The node ‘sound’ is the parent node of ‘human voice’ node since without sound there is no human voice. The link shows that relation. Similarly nodes in Figure 8 are related to others with certain probability. Each node in the BBN represents a state and provides the situation awareness.

A closely related process to BBN is a Markov process. Both Markov and Hidden Markov process are presented in the next section.

1.3.2 Markov & Hidden Markov Models (HMM)

In probability theory, people studied how the past experiments effect the future experiments. In general, the outcome of the next experiment is dependent on the outcome of the past experiments. For example, a student’s grades in the previous tests may affect the grades in the final test. In the case of student grades, a teacher might have specified a particular formula or weightage given to each test for assessing the final grade. However, if the experiments are chance experiments, prediction of the next experiment’s outcome may be difficult. Markov introduced a new chance process where the outcome of the given experiment only influences the outcome of the next experiment. This is called the Markov process and is characterized by:

$$p(X_n \mid X_{n-1}, X_{n-2}, \ldots, X_1) = p(X_n \mid X_{n-1})$$  (4)

In real world situations, the Markov process occurs quite frequently, for example, rain falls after clouds are formed.

One of the important application of Markov model is in speech recognition where the states are hidden but the measured parameters depend on the state the model is in. This important model is called the hidden Markov model (HMM). A more detailed description of the model is presented in the next section.

2. Hidden Markov Model

Consider a scenario, where there are several sensors deployed along a road as shown in Figure 9. These sensors could be acoustic, seismic, or video sensors. For the sake of discussion, let us assume they are acoustic sensors. In the case of a tracked vehicle, for example, a tank, the track makes slap noise as each segment (shoe) of the track slaps the road as it moves. The engine of a vehicle has a fundamental frequency associated with the engine cylinder’s firing rate and its harmonics will be propagated through the atmosphere. The tires make noise due to friction between the road and the tire. These sounds will be captured by the sensors. The sound level decreases inversely proportional to the distance $R$ between the vehicle and the sensor. Moreover, there is wind noise that gets added to the the vehicle sound. As a result each sensor records the vehicle sound plus the noise as voltage; generated by the microphone associated with the sensor. Let us assume that each sensor is capable of recording ‘$M$’ discrete levels of voltage $V = \{v_1, v_2, \ldots, v_M\}$ where $V$ is called the alphabet. In this experiment, let us assume only one vehicle is allowed to pass at a time. After the first vehicle completes its run, the second vehicle is allowed to pass, and so on till all the vehicles complete their runs. Let the experiment consist of using some random process for selecting initial sensor. An observation is made by measuring the voltage level at the sensor. A new sensor is selected according to some random process associated with the current sensor. Again another
observation is made. The process is repeated with other sensors. The entire process generates a sequence of observations \( O = O_1, O_2, \ldots, O_M \), where \( O_i \in V \). This is similar to the urn and ball problem presented in (Rabiner, 1989). One of the problems could be; given the observation sequence, what is the probability that it is for car, truck or tank?

An HMM in Figure 10 is characterized by (Rabiner, 1989):

1. The number of states \( N \). Let \( S \) denote the set of states, given by, \( S = \{ S_1, S_2, \ldots, S_N \} \) and we denote the state at time \( t \) as \( q_t \in S \).

2. Size of the alphabet \( M \), that is, the number of distinct observable symbols \( V = \{ v_1, v_2, \ldots, v_M \} \).

3. The state transition probability distribution \( A = \{ a_{ij} \} \) where

\[
a_{ij} = P \left[ q_{t+1} = S_j \mid q_t = S_i \right], \quad 1 \leq i, j \leq N.
\]
4. The probability distribution of each alphabet $v_k$ in state $j$, $B = \{b_j(v_k)\}$, where

$$b_j(v_k) = P(v_k \text{ at } t \mid q_t = S_j), \ 1 \leq j \leq N; \ 1 \leq k \leq M. \quad (6)$$

5. The initial state distribution $\pi = \{\pi_i\}$ where

$$\pi_i = P[q_1 = S_i], \ 1 \leq i \leq N. \quad (7)$$

Clearly, the HMM is completely specified if $N, M, A, B, \pi$ are specified and it can be used to generate an observation sequence $O = O_1, O_2, \cdots, O_T$ (Rabiner, 1989). Three questions arise with HMMs, namely,

- **Question 1**: Given the observation sequence $O = O_1, O_2, \cdots, O_T$ and the model $\lambda = \{A, B, \pi\}$, how does one compute the $P(O \mid \lambda)$, that is, the probability of the observation sequence,
- **Question 2**: Given the observation sequence $O = O_1, O_2, \cdots, O_T$ and the model $\lambda$, how does one compute the optimal state sequence $Q = q_1q_2 \cdots q_T$ that best explains the observed sequence, and
- **Question 3**: How does one optimizes the model parameters $\lambda = \{A, B, \pi\}$ that maximizes $P(O \mid \lambda)$.

Getting back to the problem posed in Figure 9, we will design a separate $N$-state HMM for each vehicle passage. It is assumed that the vehicles travel at near constant velocity and the experiment starts when the vehicle approaches a known position on the road. For training purposes the experiment is repeated with each vehicle traveling at different positions on the road, for example, left, right, middle or some other position. Now, for each HMM a model has to be built. In section 3.4 we show how to build an HMM. This is same as finding the solution to the question 3. Answer to question 2 provides the meaning to the states. Recognition of the observations is given by the solution to the question 1.

### 2.1 Solutions to the questions

In this section we will provide the answer to question 1 as it is the most important one that most of the practical situations demand. The answers to the other questions can be found in reference (Rabiner, 1989) or books on HMM.

**Solution to Question 1**: Given the observation sequence $O$ and the model $\lambda$, estimate $P(O \mid \lambda)$. Let the observed sequence is

$$O = O_1, O_2, \cdots, O_T$$

and one specific state sequence that produced the observation $O$ is

$$Q = q_1q_2 \cdots q_T$$

where $q_1$ is the initial state. Then

$$P(O \mid Q, \lambda) = \prod_{t=1}^{T} P(O_t \mid q_t, \lambda) \quad (8)$$

Invoking (6) we get

$$P(O \mid Q, \lambda) = b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdot \cdots b_{q_T}(O_T). \quad (9)$$

www.intechopen.com
The probability of the state sequence $Q$ can be computed using (5) and (7) and is given by

$$P(Q | \lambda) = \pi q_1 a_{q_1 q_2} a_{q_2 q_3} \cdots a_{q_{T-1} q_T}. \quad (10)$$

Finally, the probability of the observation sequence $O$ is obtained by summing over all possible $Q$ and is given by

$$P(O | \lambda) = \sum_{all Q} P(O | Q, \lambda) P(Q | \lambda) \quad (11)$$

There are efficient ways to compute the probability of the observation sequence given by (11) which will not be discussed here. Interested people should consult (Rabiner, 1989).

3. HMM framework for Situational Awareness

One of the advantages of using multiple sensors with multiple modalities is to detect various events with high confidence. Situational awareness is achieved based on the sequence of events observed over a period of time. These events may take place in a closed area or on a wide area. In the case of wide area, one would require multiple sensors distributed over the entire region of interest. Situational awareness leads to better response in a timely manner either to mitigate the situation or to take appropriate action proactively rather than reactively.

Since the situational awareness is achieved based on the sequence of events observed - hidden Markov model (HMM) (Rabiner, 1989) is ideally suited. Researchers used HMM for situational awareness for traffic monitoring (Bruckner et al., 2007) and learning hand grasping movements for robots (Bernardin et al., 2003).

Sensor fusion is supposed to lead to a better situational awareness. However fusion of multi-modal data is a difficult thing to do as there are few joint probability density functions exist for mixed modalities. Fusion mostly depends on the application at hand. The problem is further complicated if one has to fuse the events that take place over a period of time and over a wide area. If they are time dependent, relevance of the data observed at different times become an issue. We opted to do fusion of information, that is, probability of detection of an event. In a majority of the cases Bayesian networks (Singhal & Brown, 1997; 2000) are used for fusion. In this chapter we use Dempster-Shafer fusion (Hall & Llinas, 2001; Klein, 2004) for fusion of multi-modal multi-sensor data.

3.1 Example scenario for Situational Awareness in an urban terrain

Some of the situational awareness problems that may be of interest are discussed here. In a situation where we are monitoring a building (Damarla, 2008), we would like to know if there is any activity taking place. In particular, we placed a robot inside an office room (in stealth mode, various sensors will be placed and camouflaged to avoid detection) as shown in Figure 11.

Figure 12 shows the robot with 4 microphones, 3-axis seismic sensor, PIR, chemical sensor, 3 coil type magnetometer (one coil for each axis X, Y and Z ), three flux gate magneto meter, 3-axis E-field sensor, visible video and IR imaging sensors. The goal is to assess the situation based on the observations of various sensor modalities over a period of time in the area covered by the sensor range. We enacted the data collection scenario with several features built-in to observe the happenings inside the office room and assess the situation.
Data Collection Scenario:

- A person walks into the office room - this triggers PIR, B & E-field and seismic sensors.
- She occasionally talks - the acoustic sensor picks up the voice.
- She sits in front of a computer.
- She turns on the computer.
In the next section we present the data from various sensors and show the events detected by each sensor and also present some of the signal processing done to identify the events.

3.2 Processing of sensor data for information

We process the data from sensors in order to extract the features corresponding to various events - depending on the situation and application these extracted features will be different even for the same sensor, e.g., voice versus chime.

**Acoustic sensor data analysis:** In the case of acoustic sensors, we try to look for any human or machine activity - this is done by observing the energy levels in 4 bands, that is, 20 - 250Hz, 251 - 500Hz, 501 - 750Hz and 751 - 1000Hz corresponding to voice indicative of human presence. These four energy levels become the feature set and a classifier (Damarla et al., 2007; 2004; Damarla & Ufford, 2007) is trained with this feature set collected with a person talking and not talking. The algorithm used to detect a person is presented in the references (Damarla et al., 2007; 2004; Damarla & Ufford, 2007) and the algorithm is provided here.

**Classifier:** Let $X = [X_1, X_2, \cdots, X_N]^T$ is a vector of $N$ features, where $T$ denotes the transpose. Assuming they obey the normal distribution, then the multi-variate normal probability distribution of the pattern $X$ is given by

$$p(X) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp \left\{ -1/2 (X - M)^T \Sigma^{-1} (X - M) \right\},$$

where the mean, $M$, and the covariance matrix $\Sigma$ are defined as

$$M = E\{X\} = [m_1, m_2, \cdots, m_N]^T,$$

$$\Sigma = E\{(X - M) (X - M)^T\} = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\
\sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN}
\end{bmatrix},$$

and $\sigma_{pq} = E\{(x_p - m_p) (x_q - m_q)\}^T$, $p, q = 1, 2, \cdots, N$. We assume that for each category $i$, where $i \in \{1, \cdots, R\}$, $R$ denotes the number of classes (in our case $R = 2$, person present.

www.intechopen.com
and person not present), we know the a priori probability and the particular \( N \)-variate normal probability function \( P \{ X \mid i \} \). That is, we know \( R \) normal density functions. Let us denote the mean vectors \( M_i \) and the covariance matrices \( \Sigma_i \) for \( i = 1, 2, \cdots, R \), then we can write

\[
p(X \mid i) = \frac{1}{(2\pi)^{N/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (X - M_i)^T \Sigma_i^{-1} (X - M_i) \right\}
\]  

(12)

where \( M_i = (m_{i1}, m_{i2}, \cdots, m_{iN}) \). Let us define \( H_0 \) and \( H_1 \) as the null and human present hypotheses. The likelihood of each hypothesis is defined as the probability of the observation, i.e., feature, conditioned on the hypothesis,

\[
l_{H_j}(X_s) = p(X_s \mid H_j)
\]  

(13)

for \( j = 1, 2 \) and \( s \in S \), where \( S = \{ \text{acoustic, PIR, seismic} \} \). The conditional probability is modeled as a Gaussian distribution given by (12),

\[
p(X_s \mid H_j) = \mathcal{N}(X_s; \mu_{s,j}, \sigma_{s,j}^2).
\]  

(14)

Now, (13)-(14) can be used to determine the posterior probability of human presence given a single sensor observation. Namely,

\[
p(H_i \mid X_s) = \frac{l_{H_i}(X_s) p(H_i)}{l_{H_0}(X_s) p(H_0) + l_{H_1}(X_s) p(H_1)}
\]  

(15)

where \( p(H_0) \) and \( p(H_1) \) represent the prior probabilities for the absence and presence of a human, respectively. We assume an uninformative prior, i.e., \( p(H_0) = p(H_1) = 0.5 \).

In the office room scenario, we are looking for any activity on the computer - the Windows operating system produces a distinct sound whenever a computer is turned on or off. This distinct sound has a 75-78Hz tone and the data analysis looks for this tone. The acoustic data process is depicted in the flow chart shown in Figure 13. Figure 14 shows the spectrum of the acoustic data when a person is talking and when Windows operating system comes on.

The output of the acoustic sensor is \( P_i, i = 1, 2, 3 \), corresponding to three situations, namely, (i) a person talking, (ii) computer chime and (iii) no acoustic activity.

![Flow chart for acoustic sensor data analysis](https://www.intechopen.com)

---

**Attention:** The text is truncated at 4,068 characters. To read the entire document, please follow the link provided in the text snippet.
Seismic Sensor Data Analysis: We analyze the seismic data for footfalls of a person walking. The gait frequency of normal walk is around 1-2Hz. We use the envelope of the signal instead of the signal itself to extract the gait frequency (Damarla et al., 2007; Houston & McGaffigan, 2003). We also look for the harmonics associated with the gait frequency. Figure 15 shows the flow chart for seismic data analysis. We use the 2-15Hz band to determine the probability of person walking in the vicinity. The seismic sensor provides two probabilities, (i) probability of a person walking and (ii) probability of nobody present.

PIR sensor data analysis: These are motion detectors, if a person walks in front of them, they will give an output proportional to the temperature of the body and inversely proportional to the distance of the person from the sensor. Figure 16 shows the PIR sensor data collected in the office room. Clearly, one can see a large amplitude when a person walked by the sensor. The smaller amplitudes correspond to the person seated in the chair in front of the computer and moving slightly (note that the chair is obstructing the full view of the person) and only part of the body is seen by the PIR sensor. In order to assess the situation, both seismic and PIR sensor data can be used to determine whether a person entered the office room. The seismic sensor does not require line of sight unlike the PIR sensor - they complement each other.

Fig. 14. Spectrum of voice and computer chime

Seismic Sensor Data Analysis:

Fig. 15. Flow chart for seismic sensor data analysis

PIR sensor data analysis:
Magnetic sensor (B-field sensor) Data Analysis: We used both Flux gate and coil magnetometers. The former has low frequency response while the coil magnetometer provides high frequency response. A total of six sensors: three flux gate magnetometers, one for each direction X, Y, and Z and three coil magnetometers were used. The coil magnetometers are placed in X, Y, and Z axes to measure the magnetic flux in respective direction. Figure 17 shows clearly the change in magnetic flux when a computer is turned on and off. Similar signals are observed in Y and Z axes.

E-Field Sensor data analysis: We used three E-field sensors - one in each axis. The output of X-axis E-field sensor data is shown in Figure 18. A spike appears when the computer is turned on in the E-field sensor output, however, we did not observe any spike or change in amplitude when the computer is turned off.

Visible and IR imaging sensors: Several frames of visible and IR images of the office room and its contents are taken over a period of time. In this experiment, the images are used to determine if the computers are on or off and if anybody is sitting in front of the computer to assess the situation. Due to limited field of view of these sensors, only a partial view of the room is visible – often it is difficult to observe a person in the room. Figure 19 shows a frame of visible image showing only the shoulder of a person sitting in front of a computer. Figure 20 shows an IR frame showing a thermal image of the person in front of the computer due to reflection. Most of the thermal energy radiated by the person in front of the computer monitor is reflected by the monitor and this reflected thermal energy is detected by the IR imager. The IR imager algorithm processes the silhouette reflected from the monitor – first Hough transform (Hough, 1962) is used to determine the line patterns of an object and then using elliptical and rectangular models to detect a person (Belongie et al., 2002; Dalal & Triggs, 2005; Wang et al., 2007) in front of the monitor and provide the probability of a person being present in the room. The visible imager algorithm determines the brightness of the monitor.
Hidden Markov Model as a Framework for Situational Awareness

In the next section, we present the framework for HMM.

In the next section 3.3, we present an HMM with hypothetical states and how they can be reached based on the information observed. Although we present that these states are determined based on the output of some process, hence making them deterministic rather than the

and varying patterns and provides the probability that the computer is on. In the next section we present the framework for HMM.

Fig. 17. Flux gate magnetometer output in X-axis

Fig. 18. E-field sensor output in X-axis
hidden states, it is shown like this for conceptual purposes only. In section 3.4 we present the HMM where the states are hidden and can be reached only by particular observations.

3.3 Relation between HMM states and various states of Situational Awareness

Based on the situation we are interested in assessing, the HMM is designed with four states as shown in Figure 21. The states are as follows:

- $S_0$ denotes the state when there is no person in the office room,
- $S_1$ denotes the state when a person is present in the office room,
- $S_2$ denotes the state when a person is sitting in front of a computer and
- $S_3$ denotes the state when a computer is in use.
The above mentioned states are just a sample and can be extended to any number based on the situation one is trying to assess on the basis of observations using multi-modal sensors. We now discuss how each state is reached, what sensor data is used and how they are used. This also illustrates that the HMM also achieves the sensor fusion as each state transition is made on the observations of all or a subset of sensors.

**State $S_0$:** This is the initial state of the HMM. We use acoustic, seismic, PIR and visible video data to determine the presence of a person. Each sensor gives probability of detection, probability of no detection and confidence level denoted by $(P_d, P_{nd}, P_c)$ as shown in Figure 22. These probabilities are fused using the Dempster-Shafer (Hall & Llinas, 2001; Klein, 2004) fusion paradigm to determine the overall probability. There will be transition from state $S_0$ to $S_1$ if this probability exceeds a predetermined threshold otherwise it will remain in state $S_0$. The Dempster-Shafer fusion paradigm used is presented here.

**Dempster-Shafer fusion rule:** To combine the results from two sensors ($s_1$ and $s_2$), the fusion algorithm uses the Dempster-Shafer Rule of combination (Hall & Llinas, 2001; Klein, 2004): The total probability mass committed to an event $Z$ defined by the combination of evidence
represented by \( s_1(X) \) and \( s_2(Y) \) is given by

\[
s_{1,2}(Z) = s_1(Z) \oplus s_2(Z) = K \sum_{X \cap Y = Z} s_1(x)s_2(y)
\]  

where \( \oplus \) denotes the orthogonal sum and \( K \) the normalization factor is:

\[
K^{-1} = 1 - \sum_{X \cap Y = \emptyset} s_1(X)s_2(Y)
\]

This is basically the sum of elements from the set of Sensor 1 who intersect with Sensor 2 to make \( Z \), divided by 1 minus the sum of elements from \( s_1 \) that have no intersection with \( s_2 \).

The rule is used to combine all three probabilities (Pd, Pnd, Pc) of sensors \( s_1 \) and \( s_2 \). The resultant probabilities are combined with the probabilities of the next sensor.

**State S1:** This is the state when there is a person in the room. There are three transitions that can take place while in this state, namely, (1) transition to state \( S_2 \), (2) transitions back to state \( S_0 \) and (3) stays in the same state.

**Fig. 23.** Data processing in state \( S_1 \)

Transition to \( S_2 \) happens if any one of the following takes place: (a) if the computer turn on chime is heard, (b) if magnetic and E-field sensors detect flux change and E-field by respective sensors, (c) if the IR imager detects an image on the monitor and (d) if the visible imager detects changing images that appear during the windows turning on process.

Transition to \( S_0 \) takes place if there is no activity on any of the sensors.

The HMM remain in state \( S_1 \) if there is activity in the PIR, acoustic or seismic but not any of the events described for the case of transition to \( S_2 \). Figure 23 shows the data processing in each sensor modality.
State $S_2$: This is the state where a person is in front of the computer. The transition from this state either to $S_3$ or to $S_1$ depends on the following: (a) there is keyboard activity or the IR imager detects a hand on the keyboard and the PIR detects slight motion. $S_2$ to $S_1$ takes place when the computer is turned off - as detected by acoustic and magnetic sensors.

State $S_3$: This is the state where the computer is in use. As long as keyboard activity is detected using acoustic and IR imagers the state remains in state $S_3$, if no keyboard activity is detected, it will transition to $S_2$.

Data processing in state $S_2$ is shown in Figure 24. Data processing in $S_3$ is straightforward.

We discussed what processing is done at each state and how the probabilities are estimated. The transition probabilities of HMM are generated based on several observations with people entering into the computer room, sitting in front of the computer, turning it on, using it for a period of time, turning it off and leaving the office room.

Data processing of various sensors depends on the state of the machine and the confidence levels of various sensor modalities are also changed based on the state of the HMM. For example, in state $S_2$ the PIR sensor output monitoring a person in a chair produces small amplitude changes as shown in Figure 16 - in normal processing those outputs will not result in high probability – however in this case they will be given high probability. In state $S_3$ the acoustic sensor determines the tapping on the keyboard, this sound is often very light and the sensor is given high confidence levels than normal. In order to accommodate such varying confidence levels based on the state – it is necessary the state information should be part of the processing in a deterministic system. In a HMM where the states are automatically transition based on the outputs of sensor observations. In the next section 3.4 an HMM is built for the above problem.
3.4 Generation of HMM for the Example Scenario
In the previous section, we showed how the states could be set up based on the outputs of various sensor processes. The processes used are:

<table>
<thead>
<tr>
<th>Process</th>
<th>Output random variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic data analysis for human voice</td>
<td>( X_1 )</td>
</tr>
<tr>
<td>Acoustic data analysis for computer chime</td>
<td>( X_2 )</td>
</tr>
<tr>
<td>Seismic data analysis for footstep detection</td>
<td>( X_3 )</td>
</tr>
<tr>
<td>PIR data analysis</td>
<td>( X_4 )</td>
</tr>
<tr>
<td>Magnetic sensor data analysis</td>
<td>( X_5 )</td>
</tr>
<tr>
<td>E-field sensor data analysis</td>
<td>( X_6 )</td>
</tr>
<tr>
<td>Motion detection in imagers</td>
<td>( X_7 )</td>
</tr>
<tr>
<td>Detection of image in IR data</td>
<td>( X_8 )</td>
</tr>
</tbody>
</table>

Clearly some processes can be combined to reduce the number of variables. For example, acoustic and seismic data can be processed together for detection of human presence. Less number of variables simplify the code table needed to train the HMM. Or one can use the output of process in Figure 22 as one variable, output of process in Figure 23 as another variable and so on. Let us assume that each variable gives a binary output, that is, in the case of acoustic data analysis \( X_1 = 0 \) implies no human voice, \( X_1 = 1 \) implying the presence of human voice. At each instant of time we observe \( X = \{ X_1, X_2, \ldots, X_8 \} \) which can take \( 2^8 = 256 \) different values. Each distinct vector \( X \) is an alphabet and there are 256 alphabets.

The data collection scenario in section 3.1 is enacted several times and each enactment is made with some variation. While enacting the scenario, for each time step \( t \), we make an observation \( O_t = \{ O_{t1}, O_{t2}, \ldots, O_{t8} \} \), where \( O_i = X_i \). Each observation \( O_t \) is associated with a state \( S_i \), for \( i \in \{ 1, 2, \ldots, N \} \). This can be computed by counting how many times \( S_i \) has appeared as an initial state. Let this number is denoted by \( n_{i1} \) and dividing it by the total number of experiments \( n_e \). Then

\[
\pi_i = \frac{n_{i1}}{n_e}
\]

3.5 Computation of transition probabilities for HMM
In this section we estimate the model parameters \( \pi, A, \) and \( B \). The number of states \( N = 4 \) by design. The number of alphabet, the different possible observations, \( M = 256 \).

Estimation of \( \pi \): \( \pi = \{ \pi_i \}, \forall i \in \{ 1, 2, \ldots, N \} \), where \( \pi_i \) is the initial state probability distribution (7) for the state \( S_i \), that is, \( \pi_i = p [ q_1 = S_i ] \). This can be computed by counting how many times \( S_i \) has appeared as an initial state. Let this number is denoted by \( n_{i1} \) and dividing it by the total number of experiments \( n_e \). Then

\[
\pi_i = \frac{n_{i1}}{n_e}
\]
Table 1. Exemplar observations and the state assignment

<table>
<thead>
<tr>
<th>O₁</th>
<th>O₂</th>
<th>O₃</th>
<th>O₄</th>
<th>O₅</th>
<th>O₆</th>
<th>O₇</th>
<th>O₈</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Estimation of $A$: $A$ is the state transition probability distribution $A = \{a_{ij}\}$ where

$$a_{ij} = p\{q_{t+1} = S_j \mid q_t = S_i\}, \quad 1 \leq i, j \leq N$$

In order to compute $a_{ij}$, we need to estimate how many times the state $S_i$ to $S_j$ in the Table 1, let this number is denoted by $n_{ij}$. Note that $n_{ij}$ need not be equal to $n_{ji}$. Then

$$a_{ij} = \frac{n_{ij}}{n_T} \quad (19)$$

where $n_T$ denotes the total number of rows in the Table 1.

Estimation of $B$: $B$ is the probability distribution of each alphabet $v_k$ in state $j$, $B = \{b_j(v_k)\}$, where

$$b_j(v_k) = p\{v_k \text{ at } t \mid q_t = S_j\}, \quad 1 \leq j \leq N; \quad 1 \leq k \leq M.$$ 

In order to compute $b_j(v_k)$, first we count the number of times $n_j$ the state $S_j$ has occurred in Table 1. Out of these count the number of times the pattern $v_k = \{O_1, O_2, \cdots, O_8\}$ has occurred and denote this number by $n_{v_k}$. Then

$$b_j(v_k) = \frac{n_{v_k}}{n_j} \quad (20)$$

Now we have showed how to compute the model $\lambda = \{A, B, \pi\}$ and it can be used to determine the state and hence the situation when a new pattern is observed. It is worth noting several educational institutes have developed HMM packages for the MATLAB programming language and are available on the Internet HMM Toolbox.

In this chapter we showed how the HMM can be used to provide the situational awareness based on its states. We also showed how to build a HMM. We showed that fusion happens in HMM.
4. References


www.intechopen.com


This book aims to explore the latest practices and research works in the area of sensor fusion. The book intends to provide a collection of novel ideas, theories, and solutions related to the research areas in the field of sensor fusion. This book is a unique, comprehensive, and up-to-date resource for sensor fusion systems designers. This book is appropriate for use as an upper division undergraduate or graduate level textbook. It should also be of interest to researchers, who need to process and interpret the sensor data in most scientific and engineering fields. The initial chapters in this book provide a general overview of sensor fusion. The later chapters focus mostly on the applications of sensor fusion. Much of this work has been published in refereed journals and conference proceedings and these papers have been modified and edited for content and style. With contributions from the world’s leading fusion researchers and academicians, this book has 22 chapters covering the fundamental theory and cutting-edge developments that are driving this field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
