1

Introduction to object tracking

Object/target tracking refers to the problem of using sensor measurements to determine the location, path and characteristics of objects of interest. A sensor can be any measuring device, such as radar, sonar, ladar, camera, infrared sensor, microphone, ultrasound or any other sensor that can be used to collect information about objects in the environment. The typical objectives of object tracking are the determination of the number of objects, their identities and their states, such as positions, velocities and in some cases their features. A typical example of object/target tracking is the radar tracking of aircraft. The object tracking problem in this context attempts to determine the number of aircraft in a region under surveillance, their types, such as military, commercial or recreational, their identities, and their speeds and positions, all based on measurements obtained from a radar.

There are a number of sources of uncertainty in the object tracking problem that render it a highly non-trivial task. For example, object motion is often subject to random disturbances, objects can go undetected by sensors and the number of objects in the field of view of a sensor can change randomly. The sensor measurements are subject to random noises and the number of measurements received by a sensor from one look to the next can vary and be unpredictable. Objects may be close to each other and the measurements received might not distinguish between these objects. At times, sensors provide data when no object exists in the field of view. As we will see in later chapters of this book, the problems in object tracking can be classified according to the various types of uncertainties involved.

In this chapter, we introduce Bayes' rule, a deceptively simple yet extremely powerful tool from statistical inference, which facilitates recursive reasoning and estimation in the presence of uncertainty. This, in conjuction with the Chapman–Kolmogorov theorem, provides the foundation used to derive the object tracking algorithm presented in this book.

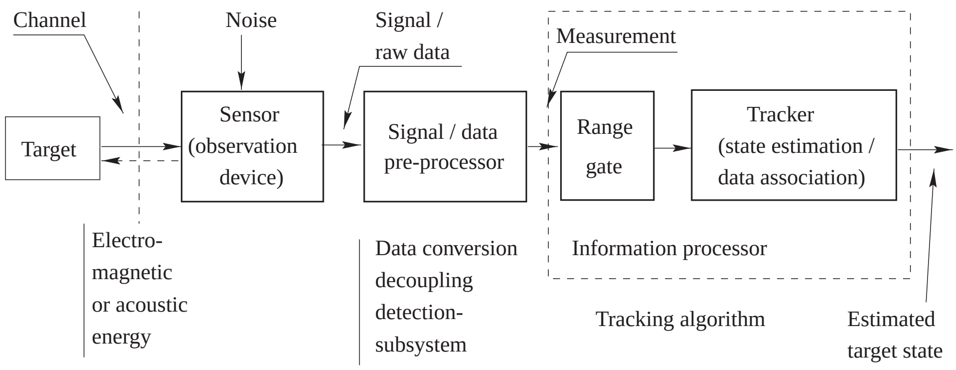

Figure 1.1 Typical object tracking system.

1.1 Overview of object tracking problems

The typical object tracking problem is essentially a state estimation problem where the object states to be estimated from noisy corrupted and false measurements are kinematic states such as position, velocity and acceleration. The tracking system consists of an object or objects to be tracked, a sensor which measures some aspect of the object, a signal processor and an information processor, as shown in Figure 1.1.

There are a number of applications for object tracking. Some of these are reviewed here.

1.1.1 Air space monitoring

An important tracking problem is the tracking of aircraft using radar (Krause, 1995), such as for air traffic control. Radar tracking is also used in military surveillance systems, where the problem involves identifying aircraft, their type, identity, speed, location, and the likely intentions of the object in order to determine, for example, if the object is a threat. Radar comes with a wide variety of measurement capabilities ranging from simple range measurements to high-resolution imaging. Radar uses reflected radio waves to measure the direction, distance and radial speed of the detected object. A radar transmitter emits electromagnetic waves, which are reflected by the object and detected by a receiver (Figure 1.2). The measured data are used to extract tracks, which are often displayed along with the object reflections on a display screen.

The tracking of aircraft using radar data is made particularly difficult because of uncertainties in the origin of the measurements. Birds, animals, vegetation, terrain, clouds, sea, rain, snow and signals generated by other radars create radar signal noise, known as clutter. Clutter consists of those detections which are not

Figure 1.2 Radar sensors.

the objects of interest. The problem of determining which detections belong to the object of interest and should be used to extract the track of the object is called the data association problem and is discussed in Chapter 4 and subsequent chapters.

Aircraft display maneuvering behaviors. Tracking such objects requires elaborate solutions that can adaptively change the model of the dynamics being used to represent the object behavior. Chapter 3 presents several approaches for dealing with maneuvering object tracking and discusses efficient algorithms to solve this problem.

Very often, air surveillance is conducted in areas where a large number of often closely spaced aircraft are present. This leads to the multi-object tracking problem (see, for example, Hwang *et al.*, 2004). A major difficulty in such situations is deciding which measurement belongs to which object in an environment where the radar produces false measurements and furthermore does not always produce a measurement for every object. This generalizes the single object tracking problem and involves consideration of more complex probabilistic models. Chapters 5 and 6 introduce solutions to multi-object tracking.

1.1.2 Video surveillance

The use of digital-video-based surveillance is growing significantly. Video surveillance is now instrumental in implementing security at airports, buildings, banks, department stores, casinos, railway stations, highways, streets, stadiums, crowd gathering places and all government institutions. In practically all sectors of society, video surveillance is used as a means to increase public safety and security and deter criminal acts. With the proliferation of high-speed broad band wired/wireless networks, many institutions now deploy large networks of cameras for surveillance. A typical large building in a major city owns a large network of cameras deployed at major entrances, floors, large gathering places, elevators, hallways, labs and offices. While such surveillance systems allow trained security guards to visually monitor the protected areas, intelligent software is needed to make use of the huge amount of information collected by the cameras. An increasingly large part of research and development is devoted to intelligent visual surveillance software (Olsen and Brill, 1997; Wren *et al.*, 1997; Boult, 1998; Lipton *et al.*, 1998; Tan *et al.*, 1998; Collins *et al.*, 2000; Haritaoglu *et al.*, 2004).

To detect and track a person or vehicle in video images, and furthermore to infer their behavior, such as unusual, loitering or even criminal behavior, one must solve non-trivial problems (Hu *et al.*, 2004). For example, the US military supported the Visual Surveillance and Monitoring (VSAM) project (Collins *et al.*, 2000) in 1997. The purpose of VSAM was to enable an operator to monitor behavior over complex areas such as battlefields. Another project, called the Human Identification at a Distance (HID) project, followed in 2000. HID aimed to develop a full range of multi-modal surveillance technologies for detecting and identifying humans from a distance. Some intelligent systems have been implemented for crowd estimation. Real-time systems for crowd estimation have been implemented based on existing closed circuit television (CCTV) at railway stations in London (Davies *et al.*, 1995) and Genova (Regazzoni *et al.*, 1993; Regazzoni and Tesei, 1996).

The tracking algorithm often breaks down when people or objects move in unusual ways. Other problems hinder traditional tracking in the visual context. Distraction occurs when some motion or light comes across the object and creates distorted measurements. Obstruction is another case when an object appears between the camera and another object. The measurements are not produced for the hidden object, which results in the termination of the track. In case both tracks do not terminate, then, as the two people separate, it is not easy to associate the new measurements with the objects.

A myriad of other problems not usually found in radar tracking situations lead to difficulties in establishing a robust approach for visual tracking. Interesting research challenges result. For example the probabilistic data association filter

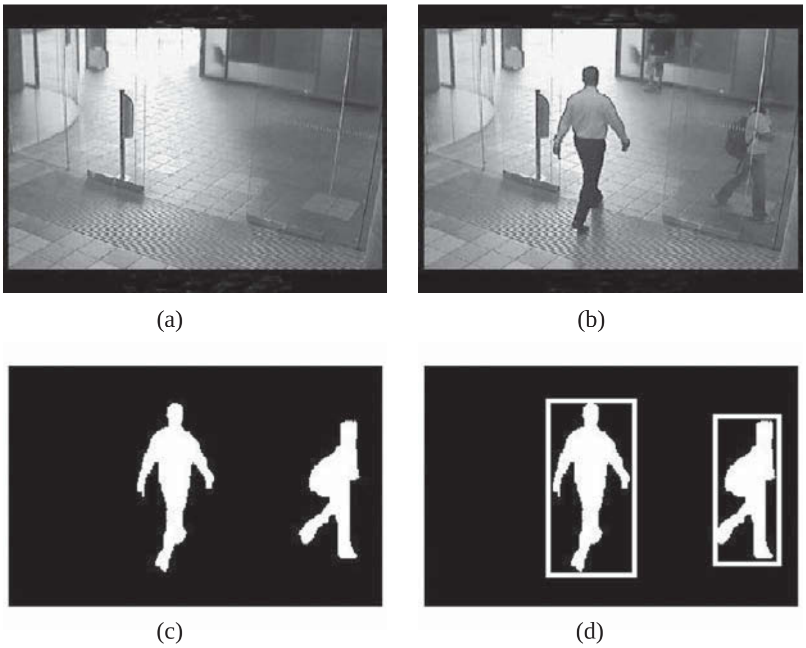

Figure 1.3 Sensor measurement extraction in video images: (a) background image, (b) background and foreground, (c) blobs, (d) bounding boxes as sensor measurements.

(PDAF) and the joint probabilistic data association filter (JPDAF) described in later chapters are excellent methods for dealing with data association problems in tracking. These are problems in which the measurements come with significant noise, or clutter. Rasmussen and Hager (2001) make good use of these filters by adapting them to visual tracking problems.

1.1.3 Weather monitoring

In order to provide weather forecasts, weather bureaus use several techniques. One technique is to track weather balloons, which provide information on high-altitude wind velocities, pressure, humidity and temperature. Weather bureaus release 50 to 70 balloons each day with the release time spaced throughout the day. In extreme weather conditions, the number of releases increases because more information is needed for accurate weather prediction. In order to get weather-related parameters at different levels of atmosphere, each weather balloon needs to be tracked.

A simple approach is to track the balloon using ground-based radars. But this approach has limitations as the initial trajectory is not in the line-of-sight of the radar. So an operator needs to track the balloon manually until it reaches a specific

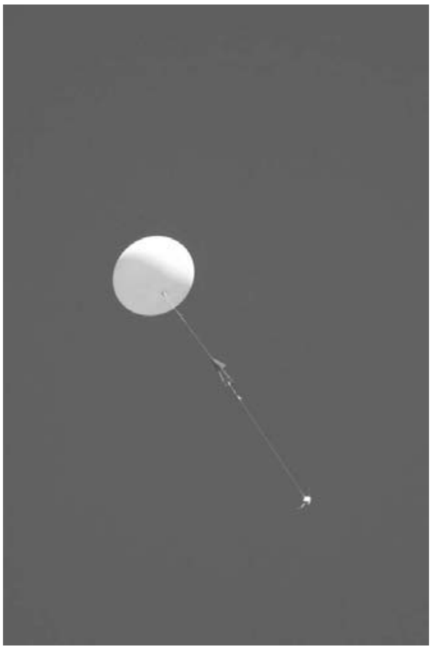

Figure 1.4 Balloon for weather monitoring (NASA/courtesy of nasaimages.org).

height from which the ground-based radar can track it automatically. A second approach is to use GPS devices in the balloon. But as the balloon is unrecoverable, the cost of the device adds to the cost of the whole process. Moreover, as this process is repeated throughout the day, using a GPS device becomes extremely costly.

A clever solution is to build a system that tracks the initial stages of the balloon trajectory and then hands it over to the ground radar once the balloon reaches a specified height. In order to do so, a radiosonde device is attached to the balloon. This device sends a synchronization pulse to three base stations. The location of the balloon is estimated by a filter that works on the principle of triangulation using the time-of-arrival of the synchronization pulse in three base stations. Once the balloon crosses a pre-specified height, a ground-based radar is locked onto the balloon and tracks it further into the atmosphere.

1.1.4 Cell biology

In studies of humans, animals, plants and insects, medical researchers and pathologists routinely study birth and death rates and the movement of biological cells. In immunology, the organism's immune response is correlated with the life cycle

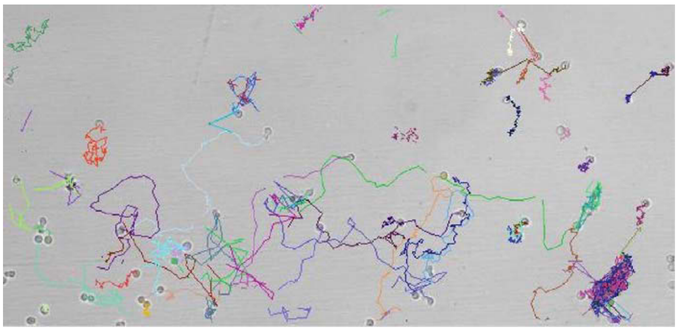

Figure 1.5 Tracking of lymphocyte cells.

of lymphocytes. The parameters of interest are the division/birth and death time of each generation of cells. In fertility studies, the interesting parameter is the velocity or shape of sperm cells. In the case of anti-inflammatory diseases, researchers look for the speed and acceleration of lymphocyte cells (Figure 1.5). Medical researchers carefully prepare the cells and take sequential images at regular intervals. In some cases, the images can be collected over several days. These images are then examined manually. The process of going through the set of images and finding out the required parameters, i.e., speed, division and death, of each cell is both time-consuming and prone to error and in some cases is almost impractical. The problem can be formulated as the problem of tracking cells over a sequence of images. The events of cell division and death are observed by looking at the track initiation and termination probabilities. The process is not only faster and reliable but also enables the discovery of previously unobserved phenomena. The key concept is to define the "data association" parameters for cell identification and tracking and then using the domain knowledge of cell immunology to model the "division/birth" and "death" of cells and integrate those into the recursive Bayesian solutions.

1.2 Bayesian reasoning with application to object tracking

The Bayesian approach is a well-developed probabilistic and statistical theory that can be applied to the modeling and solution of problems encountered in object tracking. Bayesian methods are widely used in the statistical and engineering communities, and engineers have long made use of them in solving a large variety of problems.

In most object tracking problems, like the ones introduced in Section 1.1, measurements from sensors are received in a sequential manner over time. At each measurement stage, new estimates of the object state are made by combining new information with current estimates of the object state. These latest estimates are updated in light of new information received at the next stage and so on. The recursive form of Bayes' theorem is an appropriate framework for handling the sequential nature of the measurement data and the associated uncertainties. The Bayesian recursive approach is based on Bayes' theorem, the mathematical and conceptual basis upon which the Bayesian paradigm is built. The Bayesian paradigm is a theoretical framework for reasoning under uncertainty. The history of Bayes' theorem goes back to the eighteenth and early nineteenth centuries with the work of Thomas Bayes (1764) and Pierre-Simon de Laplace (1812). The theorem was ignored for a long time. In the latter half of the twentieth century interest in the Bayesian approach grew rapidly finding application in many areas of science, engineering and statistical inference. Stigler (1986) provides an historical account of Bayesian thinking.

1.2.1 Bayes' theorem

Bayes' theorem is the encapsulation of a philosophy to consistently reconcile past and current information by exploiting the conditional probability concepts of probability theory. It may be stated as follows: given two related events $\mathbf{x}$ and $\mathbf{y}$, the conditional probability of event $\mathbf{x}$ given observation of event $\mathbf{y}$ is

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{y})}. \tag{1.1}$$

It is equal to the joint probability of events $\mathbf{x}$ and $\mathbf{y}$, $p(\mathbf{x}, \mathbf{y})$, normalized by the unconditional probability of event $\mathbf{y}$, $p(\mathbf{y})$. Using (1.1) twice, Bayes' theorem can be rewritten as

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})}. \tag{1.2}$$

Various aspects of the object tracking problem, e.g., the number of objects and their states, can be modeled as events $\mathbf{x}$ and many types of sensor outputs, e.g., radar returns or infrared images, as events $\mathbf{y}$. Bayes' theorem in (1.2) can then be applied to obtain the conditional probabilities $p(\mathbf{x}|\mathbf{y})$ as a probabilistic answer to the object tracking problem.

Interpretation of Bayes' theorem

Let $\mathbf{x}$ be a random variable of interest. In object tracking, $\mathbf{x}$ is often the state of the object under consideration. Let $\mathbf{y}$ be a measurement related to $\mathbf{x}$. Given knowledge of $\mathbf{y}$, we want to update our existing knowledge about $\mathbf{x}$. Since $\mathbf{x}$ is a random variable, and using a probabilistic approach in dealing with uncertainty, we represent our knowledge about $\mathbf{x}$ with a probability distribution or a probability density function (pdf) $p(\mathbf{x})$, depending on whether x takes discrete values or is continuous. $p(\mathbf{x})$ is a function of $\mathbf{x}$ that attributes probability values to $\mathbf{x}$. The problem at the core of most tracking problems is how to update $p(\mathbf{x})$ to take into account the new information $\mathbf{y}$. The answer is provided in (1.2) by $p(\mathbf{x}|\mathbf{y})$, the conditional distribution of $\mathbf{x}$ *given* $\mathbf{y}$. Viewing $\mathbf{x}$ as a random variable that can take a range of values, we can see that for a fixed $\mathbf{y}$, $p(\mathbf{x})$ is altered by $p(\mathbf{y}|\mathbf{x})$ to become $p(\mathbf{x}|\mathbf{y})$. We can ignore the effect of $p(\mathbf{y})$, since it is the same for all $\mathbf{x}$ values, once $\mathbf{y}$ is known. It is this, theoretically sound, alteration of $p(\mathbf{x})$ for each possible value $\mathbf{x}$ that makes us change our probability representation of $\mathbf{x}$, from $p(\mathbf{x})$ to $p(\mathbf{x}|\mathbf{y})$. It is done with a re-weighting of the initial probability distribution with the term $p(\mathbf{y}|\mathbf{x})$ for each $\mathbf{x}$ value. This deceptively simple operation is at the heart of many complex algorithms for estimating parameters of dynamic processes. The term $\mathcal{L}(\mathbf{x}) = p(\mathbf{y}|\mathbf{x})$ is known as the *likelihood function* when viewed as a function of $\mathbf{x}$. We write

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x})$$

to mean that $p(\mathbf{x}|\mathbf{y})$ is proportional to the product of $p(\mathbf{y}|\mathbf{x})$ and $p(\mathbf{x})$. The term $p(\mathbf{y})$ is a constant when viewed as a function of $\mathbf{x}$ and is called the normalizing constant or normalization factor which ensures that $p(\mathbf{x}|\mathbf{y})$ as a function of $\mathbf{x}$ sums up or integrates to 1. The initial distribution $p(\mathbf{x})$ is called the *prior distribution*. The new distribution of $\mathbf{x}$, $p(\mathbf{x}|\mathbf{y})$, is called the *posterior distribution*.

Consider the single object whose state can be represented by the one-dimensional vector $\mathbf{x}$. Suppose that given all past history, our current knowledge about $\mathbf{x}$ is contained in the probability density function

$$p(\mathbf{x}) = \frac{1}{\mathbf{\Sigma}_0\sqrt{2\pi}} e^{-\frac{(\mathbf{x}-\mathbf{x}_0)^2}{2\mathbf{\Sigma}_0^2}}.$$

This means that we think that $\mathbf{x}$ is near $\mathbf{x}_0$, with an estimation error quantified by $\mathbf{\Sigma}_0$. We may have arrived at this normal distribution (Gaussian distribution) through some previous calculations. We write $p(\mathbf{x}) = N(\mathbf{x}; \mathbf{x}_0, \mathbf{\Sigma}_o^2)$. The bigger $\mathbf{\Sigma}_0$ the less certain we are about $\mathbf{x}$. Assume that a sensor produces an observation $\mathbf{y}$ about $\mathbf{x}$. Knowing that the sensor output has measurement error, we write the following equation:

$$\mathbf{y} = \mathbf{x} + \mathbf{w}, \quad p(\mathbf{w}) = N(\mathbf{w}; 0, \mathbf{\Sigma}^2).$$

The measurement error $\mathbf{w}$ is modeled probabilistically using a Gaussian distribution with zero mean and variance $\mathbf{\Sigma}^2$. Using the basic properties of the Gaussian distribution (see Hogg and Craig, 1995), the likelihood function $\mathcal{L}(x) = p(\mathbf{y}|\mathbf{x})$

Cambridge University Press
978-0-521-87628-5 - Fundamentals of Object Tracking
Subhash Challa, Mark R. Morelande, Darko Mušicki and Robin J. Evans
Excerpt
More information

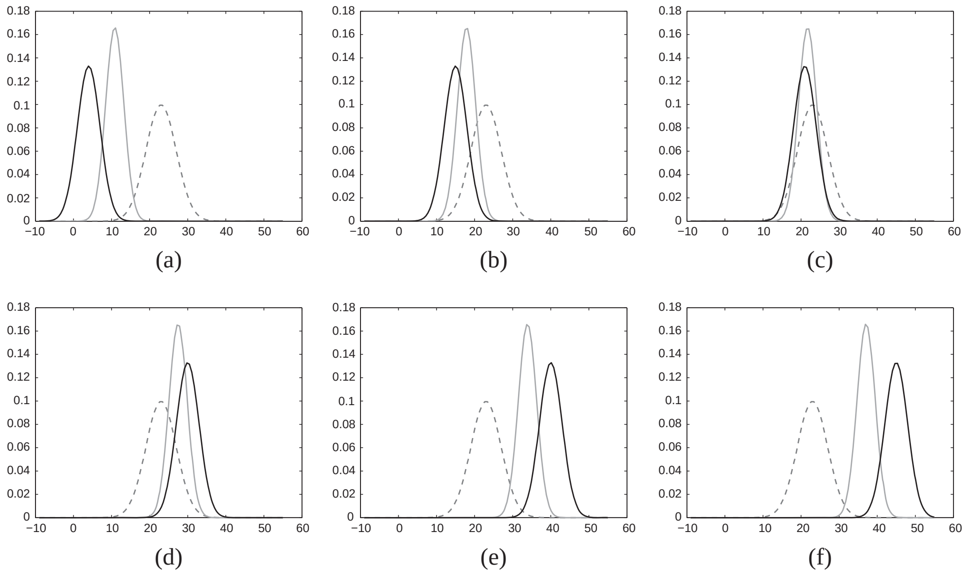

Figure 1.6 Prior $p(\mathbf{x})$ (grey line), likelihood $\mathcal{L}(\mathbf{x}) = p(\mathbf{y}|\mathbf{x})$ (black line) and posterior $p(\mathbf{x}|\mathbf{y})$ (grey line) for data values (a) $\mathbf{y} = 4$, (b) $\mathbf{y} = 15$, (c) $\mathbf{y} = 21$, (d) $\mathbf{y} = 30$, (e) $\mathbf{y} = 40$, (f) $\mathbf{y} = 45$.

becomes

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\mathbf{\Sigma}\sqrt{2\pi}} e^{-\frac{(\mathbf{y}-\mathbf{x})^2}{2\mathbf{\Sigma}^2}}.$$

Following Bayes' theorem, we can easily show that $p(x|y) = N(\mathbf{x}; \hat{\mathbf{x}}, \hat{\mathbf{\Sigma}}^2)$, where

$$\hat{\mathbf{x}} = \left(\frac{\mathbf{x}_0}{\mathbf{\Sigma}_0^2} + \frac{\mathbf{y}}{\mathbf{\Sigma}^2}\right)\left[\frac{1}{\mathbf{\Sigma}_0^2} + \frac{1}{\mathbf{\Sigma}^2}\right]^{-1},$$

$$\hat{\mathbf{\Sigma}}^2 = \left[\frac{1}{\mathbf{\Sigma}_0^2} + \frac{1}{\mathbf{\Sigma}^2}\right]^{-1},$$

where $p(\mathbf{x})$ is the prior distribution and $p(\mathbf{x}|\mathbf{y})$ is the posterior distribution. Figure 1.6 shows how the likelihood function, which differs for different values of the sensor measurement, affects the prior to produce a posterior distribution. $\mathbf{x}_0 = 23$, $\mathbf{\Sigma}_0 = 4$ and $\mathbf{\Sigma} = 3$ are assumed to be the parameter values in this example. In Figure 1.6(a), $\mathbf{y} = 4$ leads to an estimate $\hat{\mathbf{x}} = 10.84$ with a posterior variance $\hat{\mathbf{\Sigma}}^2 = 5.76$. As the value of $\mathbf{y}$ increases, the likelihood function drifts to the right, pulling with it the posterior distribution. Table 1.1 lists the estimates of $\mathbf{x}$ for different values of $\mathbf{y}$. This result is intuitive, and Bayes' theorem provides the equation that computes the posterior distribution.