Draft Chapter for Second Edition of Lee & Seshia, Introduction to Embedded Systems, a Cyber-Physical Systems Approach. Version 0.01, January 5, 2012. Contributor: Jeff C. Jensen.

7 Sensors and Actuators

Contents

A **sensor** is a device that measures a physical quantity. An **actuator** is a device that alters a physical quantity. In electronic systems, sensors usually produce a voltage that is proportional to the physical quantity being measured. The voltage may then be converted to a number by an **analog-to-digital converter** (**ADC**). A sensor that is packaged with an ADC is called a **digital sensor**, whereas a sensor without an ADC is called an **analog sensor**. A digital sensor will have a limited **precision**, determined by the number of bits used to represent the number (this can be as few as one!). Conversely, an actuator is commonly driven by a voltage that may be converted from a number by a **digital-to-analog converter** (**DAC**). An actuator that is packaged with a DAC is called a **digital actuator**.

Key properties of sensors and actuators include the proportionality constant that relates the physical quantity to the measurement or control signal, the offset or bias, and the dynamic range. These properties are discussed in section 7.1.1 and the sidebar on page 179. For many sensors and actuators, it is useful to model the degree to which a sensor or actuator deviates from a proportional measurement (its **nonlinearity**), and the amount of random variation introduced by the measurement process (its **noise**).

7.1 Models of Sensors and Actuators

Sensors and actuators connect the world of computation with the physical world. Numbers in the first world bear a relationship with quantities in the second. In this section, we provide models of that relationship. Having a good model of a sensor or actuator is essential to effectively using it.

7.1.1 Linear and Affine Models of Sensors

A function $f\colon \mathbb{R} \to \mathbb{R}$ is a **linear function** if there exists a **proportionality constant** $a \in \mathbb{R}$ such that for all $x \in \mathbb{R}$

$$f(x) = ax.$$

It is an **affine function** if there exists a **proportionality constant** $a \in \mathbb{R}$ and a **bias** $b \in \mathbb{R}$ such that

$$f(x) = ax + b.$$

Clearly, every linear function is an affine function (with $b = 0$), but not vice versa.

Many sensors may be approximately modeled by an affine function. Interpreting the readings of a sensor requires knowledge of the proportionality constant and bias. The proportionality constant represents the **sensitivity** of the sensor, since it specifies the degree to which the measurement changes when the physical quantity changes.

7.1.2 Dynamic Range of Sensors

The **range of a sensor** is the set of values of a physical quantity that it can measure. For example, a thermometer designed for weather monitoring may have a range of -10° centigrade to 45° centigrade. Physical quantities outside this range will typically **saturate**, meaning that they yield a maximum or a minimum reading outside their range. An affine function model of a sensor may be augmented to take this into account as follows,

$$f(x) = \begin{cases} ax+b & \text{if } L \leq x \leq H \\ aH+b & \text{if } x > H \\ aL+b & \text{if } x < L, \end{cases} \tag{7.1}$$

where $L, H \in \mathbb{R}$, $L < H$, are the low and high end of the sensor range, respectively.

A relation between a physical quantity and a measurement given by (7.1) is not an affine relation. In fact, this is a simple form nonlinearity that is shared by all sensors. The sensor is reasonably modeled by an affine function within an **operating range** (L, H), but outside that operating range, its behavior is distinctly nonlinear.

7.1.3 Quantization and Dynamic Range

Digital sensors are unable to distinguish between two closely-spaced values of the physical quantity. The **precision** $p \in \mathbb{R}_+$ of a sensor is the smallest absolute difference between two values of a physical quantity whose sensor readings are distinguishable. The **dynamic range** $D \in \mathbb{R}_+$ of a digital sensor is the ratio

$$D = \frac{H-L}{p}.$$

Dynamic range is usually measured in decibels (see sidebar on page 179), as follows:

$$D_{dB} = 20 \log_{10}\left(\frac{H-L}{p}\right). \tag{7.2}$$

7.1.4 Noise

For an analog sensor, the precision is harder to characterize precisely. Instead, we introduce noise: if the ideal (noiseless) measurement is $f(x)$, but additive noise $n \in \mathbb{R}$ is present, then the actual measurement is

$$f'(x) = f(x) + n.$$

The root-mean-square (RMS) noise $N \in \mathbb{R}_+$ is equal to the square root of the average value of n^2. This is a measure of (the square root of) **noise power**. The dynamic range (in decibels) is defined to in terms of RMS noise,

$$D_{dB} = 20 \log_{10} \left(\frac{H - L}{N} \right).$$

7.1.5 Sampling

7.1.6 Harmonic Distortion

A form of nonlinearity that occurs even within the operating range of sensors and actuators is **harmonic distortion**. Harmonic distortion is a nonlinear effect that can be modeled by powers of the physical quantity. Specifically, **second harmonic distortion** is a dependence on the square of the physical quantity. That is, given a physical quantity x, the measurement is modeled as

$$f(x) = ax + b + d_2 x^2, \tag{7.3}$$

where d_2 is the amount of second harmonic distortion. If d_2 is small, then the model is nearly affine. If d_2 is large, then it is far from affine. The $d_2 x^2$ term is called second harmonic distortion because of the effect it has the frequency content of a signal x that is varying in time.

Example 7.1: Suppose that a microphone is stimulated by a purely sinusoidal input sound

$$x(t) = \cos(\omega_0 t),$$

Decibels

The term "decibel" is literally one tenth of a **bel**, which is named after Alexander Graham Bell. This unit of measure was originally developed by telephone engineers at Bell Telephone Labs to designate the ratio of the **power** of two signals.

Power is a measure of energy dissipation (work done) per unit time. It is measured in **watts** for electronic systems. One bel is defined to be a factor of 10 in power. Thus, a 1000 watt hair dryer dissipates 1 bel, or 10 dB, more power than a 100 watt light bulb. Let $p_1 = 1000$ watts be the power of the hair dryer and $p_2 = 100$ be the power of the light bulb. Then the ratio is

$$\log_{10}(p_1/p_2) = 1 \text{ bel}, \quad \text{or}$$

$$10\log_{10}(p_1/p_2) = 10 \text{ dB}.$$

Comparing against (7.2) we notice a discrepancy. There, the multiplying factor is 20, not 10. That is because the ratio in (7.2) is a ratio of amplitude (magnitude), not powers. In electronic circuits, if an amplitude represents the voltage across a resistor, then the power dissipated by the resistor is proportional to the *square* of the amplitude. Let a_1 and a_2 be two such amplitudes. Then the ratio of their powers is

$$10\log_{10}(a_1^2/a_2^2) = 20\log_{10}(a_1/a_2).$$

Hence the multiplying factor of 20 instead of 10 in (7.2). A 3 dB power ratio amounts to a factor of 2 in power. In amplitudes, this is a ratio of $\sqrt{2}$.

In audio, decibels are used to measure sound pressure. A statement like "a jet engine at 10 meters produces 120 dB of sound," by convention, compares sound pressure to a defined reference of 20 micropascals, where a pascal is a pressure of 1 newton per square meter. For most people, this is approximately the threshold of hearing at 1 kHz. Thus, a sound at 0 dB is barely audible. A sound at 10 dB has 10 times the power. A sound at 100 dB has 10^{10} times the power. The jet engine, therefore, would probably make you deaf without ear protection.

where t is time in seconds and ω_0 is the frequency of the sinusoid in radians per second. If the frequency is within the human auditory range, then this will sound like a pure tone.

A sensor modeled by (7.3) will produce at time t the measurement

$$\begin{aligned}(f(x))(t) &= ax(t)+b+d_2(x(t))^2\\ &= a\cos(\omega_0 t)+b+d_2\cos^2(\omega_0 t)\\ &= a\cos(\omega_0 t)+b+\frac{d_2}{2}+\frac{d_2}{2}\cos(2\omega_0 t),\end{aligned}$$

where we have used the trigonometric identity

$$\cos^2(\theta)=\frac{1}{2}(1+\cos(2\theta)).$$

To humans, the constant term $b+d_2/2$ is not audible. Hence, this signal consists of a pure tone, scaled by a, and a distortion term at twice the frequency, scaled by $d_2/2$. This distortion term is audible as harmonic distortion as long as $2\omega_0$ is in the human auditory range.

A cubic term will introduce **third harmonic distortion**, and higher powers will introduce higher harmonics.

The importance of harmonic distortion depends on the application. The human auditory system is very sensitive to harmonic distortion, but the human visual system much less so, for example.

7.2 Common Sensors

In this section, we describe a number of sensors and show how to obtain and use reasonable models of these sensors.

7.2.1 Measuring Sound

microphone

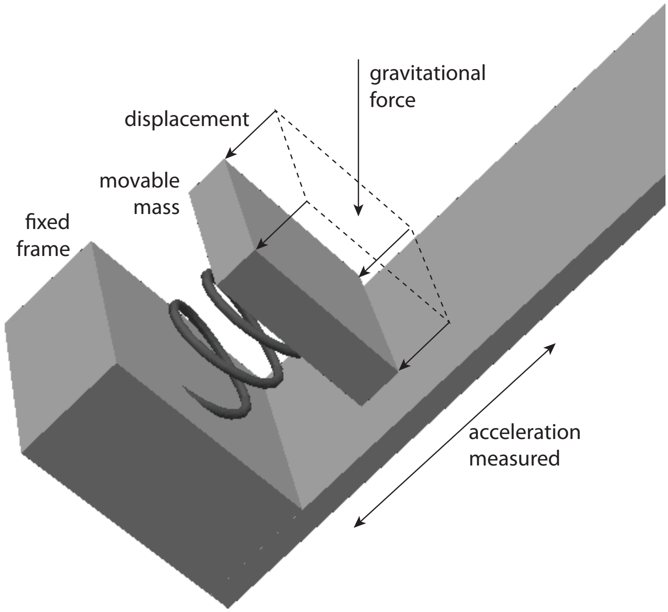

Figure 7.1: A schematic of an accelerometer as a spring-mass system.

7.2.2 Measuring Tilt and Acceleration

An **accelerometer** is a sensor that measures **proper acceleration**, which is the acceleration of an object as observed by an observer in free fall. As we explain here, gravitational force is indistinguishable from acceleration, and therefore an accelerometer measures not just acceleration, but also gravitational force. This result is a precursor to Albert Einstein's Theory of General Relativity and is known as Einstein's **equivalence principle** (Einstein, 1907).

A schematic view of an accelerometer is shown in Figure 7.1. A movable mass is attached via a spring to a fixed frame. Assume that the sensor circuitry can measure the position of the movable mass relative to the fixed frame (this can be done, for example, by measuring capacitance). When the frame accelerates in the direction of the double arrow in the figure, the acceleration results in displacement of the movable mass, and hence this acceleration can be measured.

The movable mass has a neutral position, which is its position when the spring is not deformed at all. It will occupy this neutral position if the entire assembly is in free fall, or if the assembly is lying horizontally. If the assembly is instead aligned vertically, then gravitational force will compress the spring and displace the mass. To an observer in free fall, this looks exactly as if the assembly were accelerating upwards at the **acceleration of gravity**, which is approximately $g = 9.8\text{meters}/\text{second}^2$.

An accelerometer, therefore, can measure the tilt (relative to gravity) of the fixed frame. Any acceleration experienced by the fixed frame will add or subtract from this measurement. It can be challenging to separate these two effects, gravity and acceleration. The combination of the two is what we call proper acceleration.

Assume x is the proper acceleration of the fixed frame of an accelerometer at a particular time. A digital accelerometer will produce a measurement $f(x)$ where

$$f\colon (L,H) \to \{0,\ldots,2^B-1\}$$

where $L \in \mathbb{R}$ is the minimum measurable proper acceleration and $H \in \mathbb{R}$ is the maximum, and $B \in \mathbb{N}$ is the number of bits of the ADC.

Today, accelerometers are typically implemented in silicon (see Figure 7.2), where silicon fingers deform under gravitational pull or acceleration (see for example Lemkin and Boser (1999)). Circuitry measures the deformation and provides a digital reading. Often, three accelerometers are packaged together, giving a three-axis accelerometer. This can be used to measure orientation of an object relative to gravity, plus acceleration in any direction in three-dimensional space.

7.2.3 Measuring Temperature

Measurement of temperature is central to HVAC systems, automotive engine controllers, overcurrent protectection, and many industrial chemical processes. This section, when written, will give an overview of such technologies.

(FIXME: See: http://en.wikipedia.org/wiki/Thermostat)

7.2.4 Measuring Motion

This section, when written, will give an overview of accelerometers, **gyroscope**s, and **encoder**s for measuring motion.

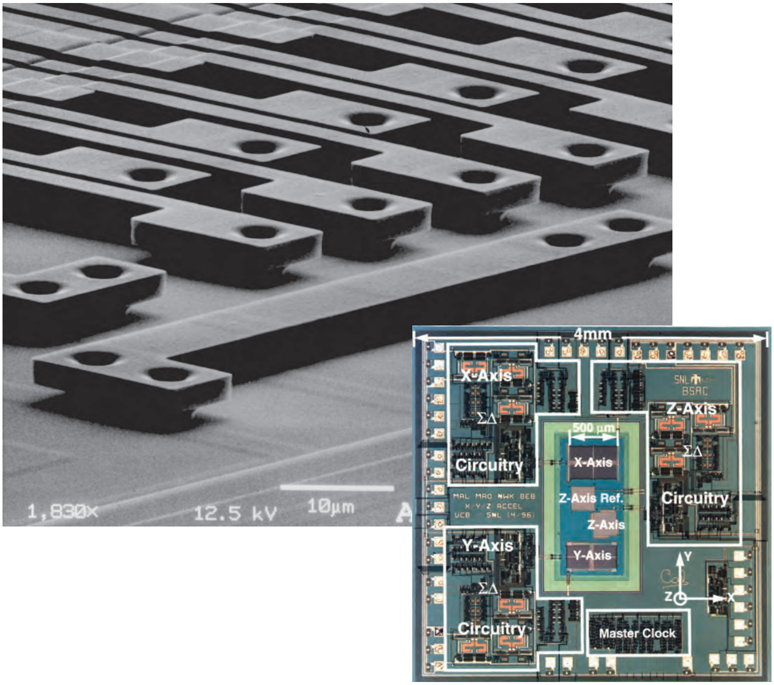

Figure 7.2: A silicon accelerometer consists of flexible silicon fingers that deform under gravitational pull or acceleration (Lemkin and Boser, 1999).

7.2.5 Measuring Time

This section, when written, will give an overview of technologies for measuring the passage of time, including **crystal oscillators**.

7.3 Actuators

7.3.1 Motion Control

pulse width modulation (**PWM**).

7.3.2 Producing Sound

7.3.3 Heating and Cooling

7.3.4 Light and Displays

Exercises

1. Show that the composition $f \circ g$ of two affine functions f and g is affine.

2. The following questions are about how to determine the function

$$f\colon (L,H) \to \{0,\ldots,2^B-1\},$$

for an accelerometer, which given a proper acceleration x yields a digital number $f(x)$. We will assume that x has units of "g's," where 1g is the acceleration of gravity, approximately $g = 9.8\text{meters}/\text{second}^2$.

   (a) Let the bias $b \in \{0,\ldots,2^B-1\}$ be the output of the ADC when the accelerometer measures no proper acceleration. How can you measure b?

   (b) Let $a \in \{0,\ldots,2^B-1\}$ be the *difference* in output of the ADC when the accelerometer measures 0g and 1g of acceleration. This is the ADC conversion of the sensitivity of the accelerometer. How can you measure a?

   (c) Suppose you have measurements of a and b from parts (2b) and (2a). Give an affine function model for the accelerometer, assuming the proper acceleration is x in units of g's. Discuss how accurate this model is.

   (d) Given a measurement $f(x)$ (under the affine model), find x, the proper acceleration in g's.

   (e) The process of determining a and b by measurement is called **calibration** of the sensor. Discuss why it might be useful to individually calibrate each particular accelerometer, rather than assume fixed calibration parameters a and b for a collection of accelerometers.

   (f) Suppose you have an ideal 8-bit digital accelerometer that produces the value $f(x) = 128$ when the proper acceleration is 0g, value $f(x) = 1$ when the proper acceleration is 3g to the right, and value $f(x) = 255$ when the proper acceleration is 3g to the left. Find a and b. What is the dynamic range (in decibels) of this accelerometer? Assume the accelerometer never yields $f(x) = 0$.