MODULE 2: IoT SENSING AND ACTUATION

2.1. INTRODUCTION

A major chunk of IoT applications involves sensing in one form or the other. Almost all the applications in IoT—be it a consumer IoT, an industrial IoT, or just plain hobby-based deployments of IoT solutions—sensing forms the first step. Incidentally, actuation forms the final step in the whole operation of IoT application deployment in a majority of scenarios. The basic science of sensing and actuation is based on the process of transduction. Transduction is the process of energy conversion from one form to another. A transducer is a physical means of enabling transduction. Transducers take energy in any form (for which it is designed) electrical, mechanical, chemical, light, sound, and others—and convert it into another, which may be electrical, mechanical, chemical, light, sound, and others. Sensors and actuators are deemed as transducers. For example, in a public announcement (PA) system, a microphone (input device) converts sound waves into electrical signals, which is amplified by an amplifier system (a process). Finally, a loudspeaker (output device) outputs this into audible sounds by converting the amplified electrical signals back into sound waves. Table outlines the basic terminological differences between transducers, sensors, and actuators.

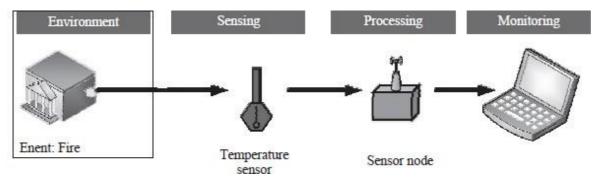
Basic outline of the differences between transducers, sensors, and actuators

Parameters	Transducers	Sensors	Actuators
Definition	Converts energy from one form to another.	Converts various forms of energy into electrical signals.	Converts electrical signals into various forms of energy, typically mechanical energy.
Domain	Can be used to represent a sensor as well as an actuator.	It is an input transducer.	It is an output transducer.
Function	Can work as a sensor or an actuator but not simultaneously.	Used for quantifying environmental stimuli into signals.	Used for converting signals into proportional mechanical or electrical outputs.
Examples	Any sensor or actuator	Humidity sensors, Temperature sensors, Anemometers (measures flow velocity), Manometers (measures fluid pressure), Accelerometers (measures the acceleration of a body), Gas sensors (measures	Motors (convert electrical energy to rotary motion), Force heads (which impose a force), Pumps (which convert rotary motion of shafts into either a pressure or a fluid velocity).

concentration of	
specific gas or	
gases), and others	

2.2. SENSORS

Sensors are devices that can measure, or quantify, or respond to the ambient changes in their environment or within the intended zone of their deployment. They generate responses to external stimuli or physical phenomenon through characterization of the input functions (which are these external stimuli) and their conversion into typically electrical signals. For example, heat is converted to electrical signals in a temperature sensor, or atmospheric pressure is converted to electrical signals in a barometer. A sensor is only sensitive to the measured property (e.g., a temperature sensor only senses the ambient temperature of a room). It is insensitive to any other property besides what it is designed to detect (e.g., a temperature sensor does not bother about light or pressure while sensing the temperature). Finally, a sensor does not influence the measured property (e.g., measuring the temperature does not reduce or increase the temperature). Figure shows the simple outline of a sensing task. Here, a temperature sensor keeps on checking an environment for changes. In the event of a fire, the temperature of the environment goes up. The temperature sensor notices this change in the temperature of the room and promptly communicates this information to a remote monitor via the processor.



The outline of a simple sensing operation

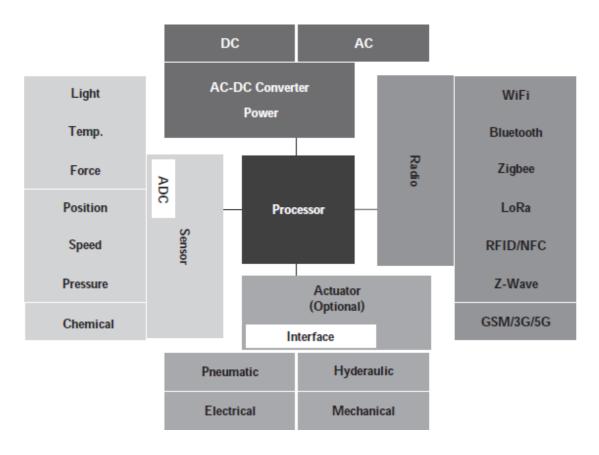
The various sensors can be classified based on: 1) power requirements, 2) sensor output, and 3) property to be measured.

- **Power Requirements**: The way sensors operate decides the power requirements that must be provided for an IoT implementation. Some sensors need to be provided with separate power sources for them to function, whereas some sensors do not require any power sources. Depending on the requirements of power, sensors can be of two types.
 - (i) Active: Active sensors do not require an external circuitry or mechanism to provide it with power. It directly responds to the external stimuli from its ambient environment and converts it into an output signal. For example, a photodiode converts light into electrical impulses.
 - (ii) Passive: Passive sensors require an external mechanism to power them up. The sensed properties are modulated with the sensor's inherent characteristics to generate patterns in the output of the sensor. For example, a thermistor's resistance can be detected by applying voltage difference across it or passing a current through it.

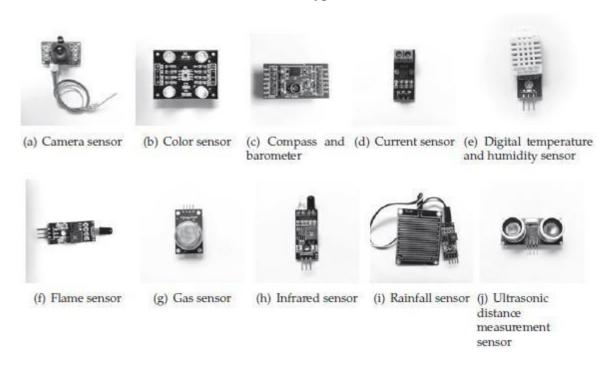
Output: The output of a sensor helps in deciding the additional components to be integrated with an IoT node or system. Typically, almost all modern-day processors are digital; digital sensors can be directly integrated to the processors. However, the integration of analog sensors to these digital processors or IoT nodes requires additional interfacing mechanisms such as analog to digital converters (ADC), voltage level converters, and others.

Sensors are broadly divided into two types, depending on the type of output generated from these sensors, as follows.

- (i) Analog: Analog sensors generate an output signal or voltage, which is proportional (linearly or non-linearly) to the quantity being measured and is continuous in time and amplitude. Physical quantities such as temperature, speed, pressure, displacement, strain, and others are all continuous and categorized as analog quantities. For example, a thermometer or a thermocouple can be used for measuring the temperature of a liquid (e.g., in household water heaters). These sensors continuously respond to changes in the temperature of the liquid.
- (ii) Digital: These sensors generate the output of discrete time digital representation (time, or amplitude, or both) of a quantity being measured, in the form of output signals or voltages. Typically, binary output signals in the form of a logic 1 or a logic 0 for ON or OFF, respectively are associated with digital sensors. The generated discrete (non-continuous) values may be output as a single "bit" (serial transmission), eight of which combine to produce a single "byte" output (parallel transmission) in digital sensors.
 - **Measured Property**: The property of the environment being measured by the sensors can be crucial in deciding the number of sensors in an IoT implementation. Some properties to be measured do not show high spatial variations and can be quantified only based on temporal variations in the measured property, such as ambient temperature, atmospheric pressure, and others. Whereas some properties to be measured show high spatial as well as temporal variations such as sound, image, and others. Depending on the properties to be measured, sensors can be of two types.
 - (i) Scalar: Scalar sensors produce an output proportional to the magnitude of the quantity being measured. The output is in the form of a signal or voltage. Scalar physical quantities are those where only the magnitude of the signal is sufficient for describing or characterizing the phenomenon and information generation. Examples of such measurable physical quantities include color, pressure, temperature, strain, and others. A thermometer or thermocouple is an example of a scalar sensor that has the ability to detect changes in ambient or object temperatures (depending on the sensor's configuration). Factors such as changes in sensor orientation or direction do not affect these sensors (typically).
 - (ii) Vector: Vector sensors are affected by the magnitude as well as the direction and/or orientation of the property they are measuring. Physical quantities such as velocity and images that require additional information besides their magnitude for completely categorizing a physical phenomenon are categorized as vector quantities. Measuring such quantities are undertaken using vector sensors. For example, an electronic gyroscope, which is commonly found in all modern aircraft, is used for detecting the changes in orientation of the gyroscope with respect to the Earth's orientation along all three axes.



The functional blocks of a typical sensor node in IoT



Some common commercially available sensors used for IoT-based sensing applications

2.3. SENSOR CHARACTERISTICS

All sensors can be defined by their ability to measure or capture a certain phenomenon and report them as output signals to various other systems. However, even within the same sensor type and class, sensors can be characterized by their ability to sense the phenomenon based on the following three fundamental properties.

- Sensor Resolution: The smallest change in the measurable quantity that a sensor can detect is referred to as the resolution of a sensor. For digital sensors, the smallest change in the digital output that the sensor is capable of quantifying is its sensor resolution. The more the resolution of a sensor, the more accurate is the precision. A sensor's accuracy does not depend upon its resolution. For example, a temperature sensor $\bf A$ can detect up to 0.5° C changes in temperature; whereas another sensor $\bf B$ can detect up to 0.25° C changes in temperature. Therefore, the resolution of sensor $\bf B$ is higher than the resolution of sensor $\bf A$.
- Sensor Accuracy: The accuracy of a sensor is the ability of that sensor to measure the environment of a system as close to its true measure as possible. For example, a weight sensor detects the weight of a $100 \, \text{kg}$ mass as $99.98 \, \text{kg}$. We can say that this sensor is 99.98% accurate, with an error rate of $^{\pm 0.02\%}$.
- Sensor Precision: The principle of repeatability governs the precision of a sensor. Only if, upon multiple repetitions, the sensor is found to have the same error rate, can it be deemed as highly precise. For example, consider if the same weight sensor described earlier reports measurements of 98.28 kg, 100.34 kg, and 101.11 kg upon three repeat measurements for a mass of actual weight of 100 kg. Here, the sensor precision is not deemed high because of significant variations in the temporal measurements for the same object under the same conditions.

2.4. SENSORIAL DEVIATIONS

Most of the sensing in IoT is non-critical, where minor deviations in sensorial outputs seldom change the nature of the undertaken tasks. However, some critical applications of IoT, such as healthcare, industrial process monitoring, and others, do require sensors with high-quality measurement capabilities. As the quality of the measurement obtained from a sensor is dependent on a large number of factors, there are a few primary considerations that must be incorporated during the sensing of critical systems. In the event of a sensor's output signal going beyond its designed maximum and minimum capacity for measurement, the sensor output is truncated to its maximum or minimum value, which is also the sensor's limits. The measurement range between a sensor's characterized minimum and maximum values is also referred to as the full-scale range of that sensor. Under real conditions, the sensitivity of a sensor may differ from the value specified for that sensor leading to sensitivity error. This deviation is mostly attributed to sensor fabrication errors and its calibration.

If the output of a sensor differs from the actual value to be measured by a constant, the sensor is said to have an offset error or bias. For example, while measuring an actual temperature of 0^{0} C, a temperature sensor outputs 1.1^{0} C every time. In this case, the sensor is said to have an offset error or bias of 1.1^{0} C. Similarly, some sensors have a non-linear behavior. If a sensor's transfer function (TF) deviates from a straight-line transfer function, it is referred to as its non-linearity.

The amount a sensor's actual output differs from the ideal TF behavior over the full range of the sensor quantifies its behavior. It is denoted as the percentage of the sensor's full range. Most sensors have linear behavior. If the output signal of a sensor changes slowly and independently of the measured property, this behavior of the sensor's output is termed as drift.

Physical changes in the sensor or its material may result in long-term drift, which can span over months or years. Noise is a temporally varying random deviation of signals. In contrast, if a sensor's output varies/deviates due to deviations in the sensor's previous input values, it is referred to as hysteresis error. The present output of the sensor depends on the past input values provided to the sensor. Typically, the phenomenon of hysteresis can be observed in analog sensors, magnetic sensors, and during heating of metal strips. One way to check for hysteresis error is to check how the sensor's output changes when we first increase, then decrease the input values to the sensor over its full range. It is generally denoted as a positive and negative percentage variation of the full-range of that sensor.

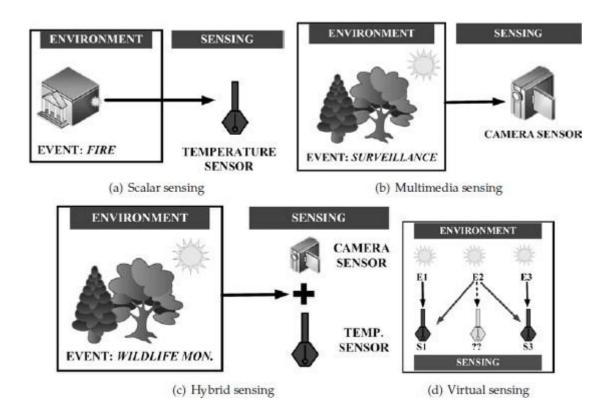
Focusing on digital sensors, if the digital output of a sensor is an approximation of the measured property, it induces quantization error. This error can be defined as the difference between the actual analog signal and its closest digital approximation during the sampling stage of the analog to digital conversion. Similarly, dynamic errors caused due to mishandling of sampling frequencies can give rise to aliasing errors. Aliasing leads to different signals of varying frequencies to be represented as a single signal in case the sampling frequency is not correctly chosen, resulting in the input signal becoming a multiple of the sampling rate. Finally, the environment itself plays a crucial role in inducing sensorial deviations. Some sensors may be prone to external influences, which may not be directly linked to the property being measured by the sensor. This sensitivity of the sensor may lead to deviations in its output values. For example, as most sensors are semiconductor based, they are influenced by the temperature of their environment.

2.5. SENSING TYPES

Sensing can be broadly divided into four different categories based on the nature of the environment being sensed and the physical sensors being used to do so (Figure 2.4): 1) scalar sensing, 2) multimedia sensing, 3) hybrid sensing, and 4) virtual Sensing

Scalar sensing

Scalar sensing encompasses the sensing of features that can be quantified simply by measuring changes in the amplitude of the measured values with respect to time. Quantities such as ambient temperature, current, atmospheric pressure, rainfall, light, humidity, flux, and others are considered as scalar values as they normally do not have a directional or spatial property assigned with them. Simply measuring the changes in their values with passing time provides enough information about these quantities. The sensors used for measuring these scalar quantities are referred to as scalar sensors, and the act is known as scalar sensing. A simple scalar temperature sensing of a fire detection event is shown in Figure 2.4(a).



The different sensing types commonly encountered in IoT

Multimedia sensing

Multimedia sensing encompasses the sensing of features that have a spatial variance property associated with the property of temporal variance. Unlike scalar sensors, multimedia sensors are used for capturing the changes in amplitude of a quantifiable property concerning space (spatial) as well as time (temporal). Quantities such as images, direction, flow, speed, acceleration, sound, force, mass, energy, and momentum have both directions as well as a magnitude. Additionally, these quantities follow the vector law of addition and hence are designated as vector quantities. They might have different values in different directions for the same working condition at the same time. The sensors used for measuring these quantities are known as vector sensors. A simple camera-based multimedia sensing using surveillance as an example is shown in Figure

Hybrid sensing

The act of using scalar as well as multimedia sensing at the same time is referred to as hybrid sensing. Many a time, there is a need to measure certain vector as well as scalar properties of an environment at the same time. Under these conditions, a range of various sensors are employed (from the collection of scalar as well as multimedia sensors) to measure the various properties of that environment at any instant of time, and temporally map the collected information to generate new information. For example, in an agricultural field, it is required to measure the soil conditions at regular intervals of time to determine plant health. Sensors such as soil moisture and soil temperature are deployed underground to estimate the soil's water retention capacity and the moisture being held by the soil at any instant of time. However, this setup only determines whether the plant is getting enough water or not. There may be a host of other factors besides water availability, which may affect a plant's health. The additional inclusion of a camera sensor with the plant may be able to determine the actual condition of a plant by additionally determining the color of leaves. The aggregate information from soil moisture, soil temperature, and the camera sensor will be able to collectively determine a plant's health at any instant of time. Other common examples of hybrid sensing include smart parking systems, traffic management systems, and others. Figure shows an example of hybrid sensing, where a camera and a temperature sensor are collectively used to detect and confirm forest fires during wildlife monitoring.

Virtual sensing

Many a time, there is a need for very dense and large-scale deployment of sensor nodes spread over a large area for monitoring of parameters. One such domain is agriculture. Here, often, the parameters being measured, such as soil moisture, soil temperature, and water level, do not show significant spatial variations. Hence, if sensors are deployed in the fields of farmer A, it is highly likely that the measurements from his sensors will be able to provide almost concise measurements of his neighbour B's fields; this is especially true of fields which are immediately surrounding A's fields. Exploiting this property, if the data from A's field is digitized using an IoT infrastructure and this system advises him regarding the appropriate watering, fertilizer, and pesticide regimen for his crops, this advisory can also be used by **B** for maintaining his crops. In short, A 's sensors are being used for actual measurement of parameters; whereas virtual data (which does not have actual physical sensors but uses extrapolation-based measurements) is being used for advising **B**. This is the virtual sensing paradigm. Figure shows an example of virtual sensing. Two temperature sensors S1 and S3 monitor three nearby events E1, E2, and E3 (fires). The event E2 does not have a dedicated sensor for monitoring it; however, through the superposition of readings from sensors S1 and S3, the presence of fire in E2 is inferred.

2.6. SENSING CONSIDERATIONS

The choice of sensors in an IoT sensor node is critical and can either make or break the feasibility of an IoT deployment. The following major factors influence the choice of sensors in IoT-based sensing solutions: 1) sensing range, 2) accuracy and precision, 3) energy, and 4) device size. These factors are discussed as follows:

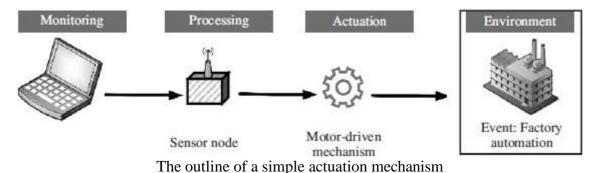
- Sensing Range: The sensing range of a sensor node defines the detection fidelity of that node. Typical approaches to optimize the sensing range in deployments include fixed k-coverage and dynamic k-coverage. A lifelong fixed k-coverage tends to usher in redundancy as it requires a large number of sensor nodes, the sensing range of some of which may also overlap. In contrast, dynamic kcoverage incorporates mobile sensor nodes post detection of an event, which, however, is a costly solution and may not be deployable in all operational areas and terrains. Additionally, the sensing range of a sensor may also be used to signify the upper and lower bounds of a sensor's measurement range. For example, a proximity sensor has a typical sensing range of a couple of meters. In contrast, a camera has a sensing range varying between tens of meters to hundreds of meters. As the complexity of the sensor and its sensing range goes up, its cost significantly increases.
- Accuracy and Precision: The accuracy and precision of measurements provided by a sensor are critical in deciding the operations of specific functional processes. Typically, off-the-shelf consumer sensors are low on requirements and often very cheap. However, their performance is limited to regular application domains. For example, a standard temperature sensor can be easily integrated with conventional components for hobby projects and day-to-day applications, but it is not suitable for industrial processes. Regular temperature sensors have a very low-temperature sensing range, as well as relatively low accuracy and precision. The use of these sensors in industrial applications, where a precision of up to 3-4 decimal places is required, cannot be

facilitated by these sensors. Industrial sensors are typically very sophisticated, and as a result, very costly. However, these industrial sensors have very high accuracy and precision score, even under harsh operating conditions.

- **Energy**: The energy consumed by a sensing solution is crucial to determine the lifetime of that solution and the estimated cost of its deployment. If the sensor or the sensor node is so energy inefficient that it requires replenishment of its energy sources quite frequently, the effort in maintaining the solution and its cost goes up; whereas its deployment feasibility goes down. Consider a scenario where sensor nodes are deployed on the top of glaciers. Once deployed, access to these nodes is not possible. If the energy requirements of the sensor nodes are too high, such a deployment will not last long, and the solution will be highly infeasible as charging or changing of the energy sources of these sensor nodes is not an option.
- Device Size: Modern-day IoT applications have a wide penetration in all domains of life. Most of the applications of IoT require sensing solutions which are so small that they do not hinder any of the regular activities that were possible before the sensor node deployment was carried out. Larger the size of a sensor node, larger is the obstruction caused by it, higher is the cost and energy requirements, and lesser is its demand for the bulk of the IoT applications. Consider a simple human activity detector. If the detection unit is too large to be carried or too bulky to cause hindrance to regular normal movements, the demand for this solution would be low. It is because of this that the onset of wearables took off so strongly. The wearable sensors are highly energyefficient, small in size, and almost part of the wearer's regular wardrobe.

2.7. ACTUATORS

An actuator can be considered as a machine or system's component that can affect the movement or control the said mechanism or the system. Control systems affect changes to the environment or property they are controlling through actuators. The system activates the actuator through a control signal, which may be digital or analog. It elicits a response from the actuator, which is in the form of some form of mechanical motion. The control system of an actuator can be a mechanical or electronic system, a software-based system (e.g., an autonomous car control system), a human, or any other input. Figure shows the outline of a simple actuation system. A remote user sends commands to a processor. The processor instructs a motor controlled robotic arm to perform the commanded tasks accordingly. The processor is primarily responsible for converting the human commands into sequential machine-language command sequences, which enables the robot to move. The robotic arm finally moves the designated boxes, which was its assigned task.



2.8. ACTUATOR TYPES

Broadly, actuators can be divided into seven classes: 1) Hydraulic, 2) pneumatic, 3) electrical, 4) thermal/magnetic, 5) mechanical, 6) soft, and 7) shape memory polymers. Figure 2.6 shows some of the commonly used actuators in IoT applications.

Hydraulic actuators

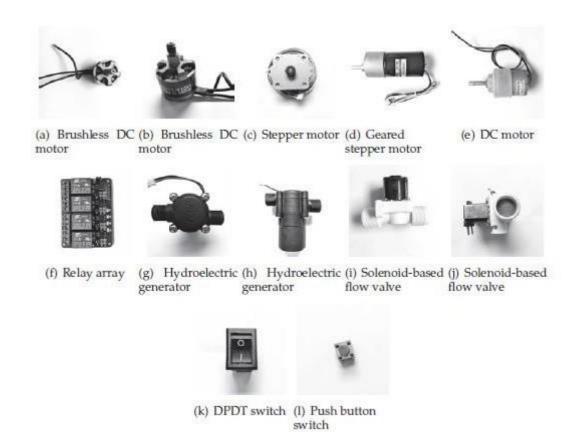
A hydraulic actuator works on the principle of compression and decompression of fluids. These actuators facilitate mechanical tasks such as lifting loads through the use of hydraulic power derived from fluids in cylinders or fluid motors. The mechanical motion applied to a hydraulic actuator is converted to either linear, rotary, or oscillatory motion. The almost incompressible property of liquids is used in hydraulic actuators for exerting significant force. These hydraulic actuators are also considered as stiff systems. The actuator's limited acceleration restricts its usage.

Pneumatic actuators

A pneumatic actuator works on the principle of compression and decompression of gases. These actuators use a vacuum or compressed air at high pressure and convert it into either linear or rotary motion. Pneumatic rack and pinion actuators are commonly used for valve controls of water pipes. Pneumatic actuators are considered as compliant systems. The actuators using pneumatic energy for their operation are typically characterized by the quick response to starting and stopping signals. Small pressure changes can be used for generating large forces through these actuators. Pneumatic brakes are an example of this type of actuator which is so responsive that they can convert small pressure changes applied by drives to generate the massive force required to stop or slow down a moving vehicle. Pneumatic actuators are responsible for converting pressure into force. The power source in the pneumatic actuator does not need to be stored in reserve for its operation.

Electric actuators

Typically, electric motors are used to power an electric actuator by generating mechanical torque. This generated torque is translated into the motion of a motor's shaft or for switching (as in relays). For example, actuating equipments such as solenoid valves control the flow of water in pipes in response to electrical signals. This class of actuators is considered one of the cheapest, cleanest and speedy actuator types available.



Some common commercially available actuators used for IoT-based control Applications

Thermal or magnetic actuators

The use of thermal or magnetic energy is used for powering this class of actuators. These actuators have a very high power density and are typically compact, lightweight, and economical. One classic example of thermal actuators is shape memory materials (SMMs) such as shape memory alloys (SMAs). These actuators do not require electricity for actuation. They are not affected by vibration and can work with liquid or gases. Magnetic shape memory alloys (MSMAs) are a type of magnetic actuators.

Mechanical actuators

In mechanical actuation, the rotary motion of the actuator is converted into linear motion to execute some movement. The use of gears, rails, pulleys, chains, and other devices are necessary for these actuators to operate. These actuators can be easily used in conjunction with pneumatic, hydraulic, or electrical actuators. They can also work in a standalone mode. The best example of a mechanical actuator is a rack and pinion mechanism.

Soft actuators

Soft actuators (e.g., polymer-based) consists of elastomeric polymers that are used as embedded fixtures in flexible materials such as cloth, paper, fiber, particles, and others. The conversion of molecular level microscopic changes into tangible macroscopic deformations is the primary working principle of this class of actuators. These actuators have a high stake in modern-day robotics. They are designed to handle fragile objects such as agricultural fruit harvesting, or performing precise operations like manipulating the internal organs during robot-assisted surgeries.

Shape memory polymers

Shape memory polymers (SMP) are considered as smart materials that respond to some external stimulus by changing their shape, and then revert to their original shape once the affecting stimulus is removed. Features such as high strain recovery, biocompatibility, low density, and biodegradability characterize these materials. SMP-based actuators function similar to our muscles. Modern-day SMPs have been designed to respond to a wide range of stimuli such as pH changes, heat differentials, light intensity, and frequency changes, magnetic changes, and others. Photopolymer/light-activated polymers (LAP) are a particular type of SMP, which require light as a stimulus to operate. LAP-based actuators are characterized by their rapid response times. Using only the variation of light frequency or its intensity, LAPs can be controlled remotely without any physical contact. The development of LAPs whose shape can be changed by the application of a specific frequency of light have been reported. The polymer retains its shape after removal of the activating light. In order to change the polymer back to its original shape, a light stimulus of a different frequency has to be applied to the polymer.

2.9. ACTUATOR CHARACTERISTICS

The choice or selection of actuators is crucial in an IoT deployment, where a control mechanism is required after sensing and processing of the information obtained from the sensed environment. Actuators perform the physically heavier tasks in an IoT deployment; tasks which require moving or changing the orientation of physical objects, changing the state of objects, and other such activities. The correct choice of actuators is necessary for the long-term sustenance and continuity of operations, as well as for increasing the lifetime of the actuators themselves. A set of four characteristics can define all actuators:

- Weight: The physical weight of actuators limits its application scope. For example, the use of heavier actuators is generally preferred for industrial applications and applications requiring no mobility of the IoT deployment. In contrast, lightweight actuators typically find common usage in portable systems in vehicles, drones, and home IoT applications. It is to be noted that this is not always true. Heavier actuators also have selective usage in mobile systems, for example, landing gears and engine motors in aircraft.
- Power Rating: This helps in deciding the nature of the application with which an actuator can be associated. The power rating defines the minimum and maximum operating power an actuator can safely withstand without damage to itself. Generally, it is indicated as the power-to-weight ratio for actuators. For example, smaller servo motors used in hobby projects typically have a maximum rating of 5 VDC, 500 mA, which is suitable for an operations-driven battery-based power source. Exceeding this limit might be detrimental to the performance of the actuator and may cause burnout of the motor. In contrast to this, servo motors in larger applications have a rating of 460 VAC, 2:5 A, which requires standalone power supply systems for operations. It is to be noted that actuators with still higher ratings are available and vary according to application requirements.
- Torque to Weight Ratio: The ratio of torque to the weight of the moving part of an instrument/device is referred to as its torque/weight ratio. This indicates the sensitivity of the actuator. Higher is the weight of the moving part; lower will be its torque to weight ratio for a given power.

Stiffness and Compliance: The resistance of a material against deformation is known as its stiffness, whereas compliance of a material is the opposite of stiffness. Stiffness can be directly related to the modulus of elasticity of that material. Stiff systems are considered more accurate than compliant systems as they have a faster response to the change in load applied to it. For example, hydraulic systems are considered as stiff and non-compliant, whereas pneumatic systems are considered as compliant.