1. Introduction

A sensor is defined as a device that provides a usable electrical output signal in response to a specified measurand. An actuator, the reverse of a sensor, is a device that converts an electrical signal to an action, while a transducer can be considered as the device that transforms one form of signal or energy into another form. Therefore, the term transducer can be used to include both sensors and actuators. However, these definitions vary with different researchers, times and regions. The definitions stated above are used throughout this review.

A number of criteria can be used to classify sensors, including (a) transduction principle - the physical, chemical or biological effects; (b) measurands - pressure, acceleration, gas concentration, ion concentration, etc.; (c) fabrication technology and materials - thin film, semiconductors, ceramics, etc.; (d) applications - automotive, medical, aerospace, etc.; and (e) cost and performance. In most of the literature, a combination of principle, measurand, and application is used to define sensors and actuators. This section will follow the popular approach of classifying sensors into three major groups: physical, chemical and biological (including medical). Transduction principles or effects will be used to subdivide the material on physical sensors.

Microsensor development started with physical sensors, partly because of existing market demand and partly because mechanical sensors can be easily scaled or transferred from conventional sensors. Furthermore, the fabrication technology developed for ICs was easily adapted to fabricate the mechanical components for physical sensors. However, for chemical and biological sensors, there were many materials issues that needed to be resolved for commercial silicon base sensors to impact the market. Thousands of papers have been published on physical microsensors since 1961 when the Kulite Company introduced a pressure sensor made from diffused resistors on a silicon diaphragm. Since then, hundreds of physical sensors have been introduced into the commercial market. This section will review the present status of solid state physical microsensors. The discussion will include: (a) classification of sensors, (b) integrated or intelligent sensors, (c) sensor principles and examples, and (d) future trends and challenges.

2. Classification of Physical Sensors

Sensor principles are based on physical or chemical effects. These effects have been compiled in the literature. More than 350 effects are known, most of which can be exploited for sensors. These physical principles or effects can be grouped according to the five forms of physical energy in which the signals are received or generated, as shown...
in Figure 1. The chemical domain has been purposefully omitted, since Section 4 will review chemical and biological sensors. Examples of signal parameters or measurands in the five physical domains are listed in Table 1. A typical sensor employs one or more transduction principles in different domains. Therefore, the classification presented here is for convenience of discussion and is neither absolute nor complete. Some sensors that can generate energy with a signal, such as photoelectric or piezoelectric devices, are called active or self-generating sensors. Other sensors require an auxiliary energy source to generate the desired signal. These are called passive or modulating sensors.

![Figure 1. Signal domains of physical sensors](image)

### 3. Integrated, Intelligent or Smart Sensors

When a microsensor is integrated with signal processing circuits in a single package, it is referred to as an integrated sensor. A monolithic integrated sensor has the signal processing circuit fabricated on the same chip as the sensor, while a hybrid integrated sensor has the signal processing circuit on the same hybrid substrate as the sensor chip. Thus, the packaged sensor not only transduces the measurand into electrical signals, but also may have other signal processing and decision making capabilities. Thus, the term intelligent sensors was coined. In trade journals and popular news articles they are also called smart sensors.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanical</td>
<td>Length, area, volume, all time derivatives (linear/angular velocity/acceleration), mass flow, force, torque, pressure, acoustic wavelength and intensity</td>
</tr>
<tr>
<td>Thermal</td>
<td>Temperature, (specific) heat, entropy, heat flow</td>
</tr>
<tr>
<td>Electrical</td>
<td>Voltage, current, charge, resistance, inductance, capacitance, dielectric constant, polarization, electric field, frequency, dipole moment</td>
</tr>
<tr>
<td>Magnetic</td>
<td>Field intensity, flux density, magnetic moment, permeability</td>
</tr>
<tr>
<td>Radiant</td>
<td>Intensity, phase, wavelength, polarization, reflectance, transmittance, refractive index</td>
</tr>
</tbody>
</table>
Integrated sensors with specific types of on-chip circuits may have the following advantages:

(1) **Better signal-to-noise ratio:**

On-chip amplification can boost the signal level before transmission to other system blocks, thus reducing the effects of interference and transmission noise, especially when analog-to-digital (A-to-D) conversion is made on-chip.

(2) **Improved characteristics:**

Signal processing can provide: (a) on-chip feedback system or look-up tables to improve the output linearity, (b) compensation circuits to reduce cross-sensitivity to temperature, strain, or other known interfering effects, (c) on-chip accurate current and voltage sources and associated circuits to provide automatic and periodic self-checking and calibration, and (d) feedback and other circuits to improve or compensate for the frequency response of the sensor.

(3) **Signal conditioning and formatting:**

Signal conditioning can be performed by integrated circuits prior to the output, such as: (a) A-to-D conversion, (b) impedance matching, (c) output formatting to a standard, and (d) signal averaging.

(4) **Improved signals representing the measurands:**

Decision-making and computing circuits as well as memory devices may be used with redundant sensors to exclude noisy or failed sensors from the array of devices, thus improving the signal beyond a single sensor's capability. Multiple sensors, each with poor selectivity to a family of measurands, can be used with pattern recognition circuits to obtain accurate signals for each of the measurands. Neural networks may be used to train the sensors to recognize the desired measurands among several interfering effects. Many of the signal processing techniques used so successfully in communications and defense applications may be incorporated with the sensors to improve the quality of the sensor output beyond the individual device's capability.

With the advances of micro-actuators and MEMS, the goal is to integrate a complete system for a given application with built-in self-calibration and self-diagnostics through special actuators.

4. **Sensor Principles and Examples**

As discussed above, physical sensors are classified into thermal, electrical, mechanical, optical, and magnetic groups according to the sensor principles. The present status of these sensors is summarized as follows.
4.1 Thermal sensors

The best known thermal sensors are thermistors and thermocouples for measuring the temperature of the environment. Thermistors and RTDs (resistive temperature detectors) are based on the change of mobility and carrier density with temperature. These changes are represented by temperature coefficients which may be constants or nonlinear functions of temperature. The resistance of a thermistor is an exponential function of the temperature. Linearization networks may be used to make the output a linear function of the temperature over a specific range, with some sacrifice in sensitivity.

The thermocouple is based on the Seebeck effect, one of the three thermo-electric effects (Seebeck, Peltier and Thompson effects). Two different materials (usually metals) are joined at one point to form a thermocouple. The ideal reference junction thermocouple measurement system which consists of two metal 1/metal 2 thermocouple junctions in series, as shown in Figure 2. An ideal voltmeter, with copper leads, is connected to metal 2 of each junction, which forms two more thermocouples. The reference junction is held at a known temperature, such as the ice-water equilibrium point (i.e., known thermocouple voltage), and the difference between this fixed voltage and the thermocouple voltage at the measurement junction is measured by the voltmeter. The two Cu/metal 2 junction voltages will cancel as long as they are isothermal. In practice this method is not always used, since a stable reference temperature is not usually available. Also, since the standard thermocouple wire is expensive, the added cost of a second junction may be prohibitive to large volume products.

A practical way to measure temperature with only one thermocouple is to hold the terminal connections between the thermocouple wires and the voltmeter leads at a constant temperature. If this temperature is measured independently, (e.g., with an RTD), the system can be calibrated to remove the temperature dependent offset voltage due to these two thermocouple junctions.

![Figure 2. Ideal reference junction thermocouple measurement](image)

Various metal wires can be used as thermocouples for different temperature ranges and sensitivity. Common wire materials are: iron/constantan, chromel/constantan, and platinum/platinum-10% rhodium. Semiconductors, including silicon, may also be used with a metal to form a microthermocouple. Many of these microthermocouples may be
connected in series to form a thermopile which is much more sensitive than a single
thermocouple. Thermopiles are sensitive enough to measure the temperature rise as a
result of incident infrared radiation (IR), making thermopiles useful as passive IR
sensors.

If the current through a p-n junction is kept constant, the junction voltage is a linear
function of temperature. This principle has been used for many commercial temperature
sensors. When two bipolar transistors are operated at a constant ratio of emitter current
densities, the difference in base emitter voltages is proportional to the absolute
temperature. This principle is the basis of many precise commercial temperature sensor
ICs.

The mass flow of fluids can be measured by the temperature difference between two
sensors, one on the down-stream and one on the up-stream of a heating (or cooling)
source. This is the principle of micro-anemometers and is also a good example of how the
mechanical measurand-flow is transduced to thermal parameters which are then
transduced to electrical signals.

4.2 Electrical sensors

Electrical signals can be picked up by ideal ohmic connections without special sensors.
However, most connections and electrical probes are not ideal and, sometimes, the
magnitude of the charge, current and voltage may not be convenient to measure directly.
For example, Hall effect sensors are used to measure the current and polarity in motor
windings where the current value is very large or a contact is difficult to make. Similarly,
sensors may be used to measure very high voltages or very large charges.

The flow of electrical currents occurs through the motion of charge carriers which may be
electrons (and holes), ions and charged defects in materials. These different charge
carriers can have very different charge-to-mass ratios and mobilities. When a current has
to pass through the surface of a material with one type of charge carrier (e.g., electrons)
to the surface of another material with a different type of charge carrier (e.g., ions), there
are complex reactions occurring at the interface associated with the transfer of charges.
The charge carriers in metals are electrons, and in semiconductors may be electrons or
holes (the collective behavior of electrons). In electrolytes, the movement of ions carries
the current. Dielectrics are much more complicated because electrons, ions, charged atom
groups, and combinations of these can carry the current.

The electrodes that are used to measure electrical parameters in electrolytes or dielectrics
are electrical sensors. These electrodes are used in electrochemistry, material science, and
biology. The equivalent circuit of such electrodes may include several resistance-
capacitance (RC) networks to model the several layers of carriers near the interface. The
characterization of the electrode may involve many techniques. A commonly used
method is to plot the imaginary part of the impedance versus the real part. In all
applications, it is important to know the RC equivalent circuit of the electrodes, the
contact potential (work function difference) between the materials in contact, and the
thermoelectric effects between the materials in the measurement loop. Using specialized computer programs, the equivalent circuit of the electrode may be found. Noble metal (platinum) wires coated with various materials may be used to measure the potential of the electrolyte, or the ionic concentration of the ions in the electrolyte. These chemical sensors are called ion selective electrodes and are used in many chemical and biomedical applications.

Microelectrodes were designed for the measurement of electrical potentials on the surfaces of tissues, cells, and other materials. Vibrating microelectrodes can be used to measure the potential or charge at the surface without making direct contact to it.

4.3 Mechanical sensors

There are many mechanical sensors described in the literature and commercially available. Mechanical parameters may be converted to other energy domains and then sensed or measured directly. For direct sensing, the parameters are related to strain or displacement. Silicon, a brittle material, will fail (break) at a maximum strain of about 2%. However, below the elastic limit the strain is related to the stress by a nearly constant Young's modulus, which is much larger than steel. The principles used to sense strain are piezoelectricity, piezo-resistivity, and capacitive or inductive impedance.

The piezoelectric effect relates the elastic strain, $S$, (or stress, $T$) in one orientation to displacement charge density, $D$, (or electric field intensity, $E$) in another orientation which may or may not be the same as the orientation of the strain, through the piezoelectric coefficients. The piezoelectric coefficients are, in general, a $(n \times m)$ matrix, where $n$ is the number of orientations for strain-the mechanical parameter, and $m$ is the number of orientations for $D$ or $E$-the electrical parameters. The four parameters $S$, $T$, $D$, and $E$ are interrelated by various coefficients or constants. For most applications, simplifying assumptions can be made so that only parallel and perpendicular orientations between the mechanical and electrical parameters are considered. For example, $d_{ij}$ is the strain constant or charge constant of piezoelectricity relating $D$ and $T$ when $E$ is constant, or $S$ and $E$, when $T$ is constant:

$$D_i = d_{ij}T_j + eE_i \quad (1)$$

and

$$S_j = YT_j + D_{ji}E_i \quad (2)$$

where $e$ is the dielectric constant or permittivity; and $1/Y$ is the Young's modulus of the material. The coefficient

$$d_{33} = ([\Delta]D/[[\Delta]T])E=k \quad \text{or} \quad d_{33} = ([\Delta]S/[[\Delta]E]T = k$$

relates $D$ and $T$ or $S$ and $T$ when they are both oriented in the "z"-"3" direction; thus, they are in parallel. $d_{31}$ relates the electrical parameters $D_3$ and $E_3$ to the perpendicular mechanical parameters $T_1$ and $S_1$. These coefficients can be used to estimate the characteristics of sensors and actuators.
When used for sensors, the piezoelectric effect is used to measure various forms of strain or stress. Examples are microphones for strains generated by acoustic pressure on a diaphragms; ultrasonic sensors for high frequency strain waves arriving at or propagating through the sensors; and pressure sensors for AC pressures on a silicon diaphragm coated with piezoelectric materials. The piezoelectric effect can also be used to sense small displacements, bending, rotations, etc. These measurements require a high input impedance amplifier to measure the surface charges or voltages generated by the strain or stress.

The piezoresistive effect in conductors and semiconductors is used for commercial pressure sensors and strain gauges. The strain on the crystal structure deforms the energy band structure and, thus, changes the mobility and carrier density which changes the resistivity or conductivity of the material. The strain in one orientation may affect the conductivity in another orientation. Therefore, similar to the piezoelectric effect, piezoresistivity can be described as a \( n \times n \) matrix. For most applications, the piezoresistive material is a crystal and has symmetry in its structure. For cubic crystals, \( n = 6 \), (three for axial strain, three for shear strain, three for voltage and three for current). Using some simplifying assumptions, the matrix of piezoresistivity coefficients can be reduced to parallel and perpendicular coefficients just as for the piezoelectric coefficients. However, the coefficients are not only orientation dependent, but are also affected by doping and temperature. The piezoresistivity coefficients of silicon and germanium were measured by C.S. Smith.

A practical piezoresistive pressure sensor can be built by fabricating four sensing resistors at the edges or at the center of a thin silicon diaphragm which acts as a mechanical amplifier to increase the stress and strain at the sensor site. Usually, the four sensing resistors are connected in a bridge configuration with push-pull signals to increase the sensitivity. The measurable pressure range for such a sensor can be from 1 to \( 10^6 \) Torr. Bulk and surface micro-machined pressure sensors are currently at various stages of commercialization.

Capacitive or inductive impedances can also be used to measure displacements and strains. Capacitive pressure sensors present some advantages over piezoresistive devices. Capacitive devices integrate the change of elementary capacitive areas while piezoresistive devices take the difference of the resistance changes of the bridge arms. Therefore, capacitive sensors are less sensitive to the sideways forces and are more stable. Furthermore, capacitive changes can be much larger as a percentage of the reference value than the 5% maximum resistance change found in piezoresistive devices. However, capacitive sensors require a capacitance-to-voltage (C-to-V) converter on or near the chip to avoid the effects of stray capacitances. Therefore, the device operation becomes complicated unless the required C-to-V converter can be fabricated on the chip or packaged in the same encapsulation. The measurement circuit also must be stable and have low noise.

Figure 3 shows a cross-sectional schematic drawing of pressure sensors based on these principles which can be used to measure acceleration, force, flow and small
displacements. Other sensors which can measure torque, rotation, touch, etc. have also been reported in the literature.

### 4.4 Optical sensors

Optical sensors include photoconductors, photovoltaic devices and fiber optic sensors. The conductivity of photoconductors changes under optical radiation due to changes in the charge carrier population. Photovoltaic devices involve a p-n junction where radiation generated carriers may cross the junction to form currents and a self generated voltage. Photovoltaic devices, such as solar cells, can supply energy to external circuits. However, the current is proportional to the radiation intensity, but the voltage is not (it is very nonlinearly related to the intensity of the radiation).

![Figure 3](image)

**Figure 3.** Schematic drawings of the different types of silicon pressure sensors: (a) piezoelectric, (b) piezoresistive, (c) capacitive, and (d) surface micromachined capacitive
When strained, a fiber optic cable changes the intensity or the phase delay of the output optical wave relative to a reference. Using an optical detector and an interference measuring technique, small strains can be measured with high sensitivity. Intensity sensors can detect changes in optical intensity which is related to an applied measurand. Examples include: underwater acoustic sensors, fiber micro-bend sensors, evanescent or coupled waveguide sensors, moving fiber optic hydrophones, grating sensors, polarization sensors and total internal reflection sensors.

Optical interference sensors have been developed for interferometer acoustic sensors, fiber optic magnetic sensors with a magnetostrictive jacket, and fiber optic gyroscopes. Specially doped or coated optical fibers have been shown to have great versatility for physical sensors of various types and configurations. They have been used for radiation dosimeters, current sensors, accelerometers, temperature sensors, as well as detectors for liquid level, displacement, strain, torque, fluid flow, etc. Many chemical and biomedical materials can be sensed with fiber optic devices.

4.5 Magnetic sensors

Magnetic sensors may utilize any of the following effects: (a) the magneto-optic effect based on the Faraday rotation of the polarization plane of linearly polarized light due to the Lorentz force on bound electrons; (b) the magnetostrictive effect where the magnetic field causes strain on the material; (c) the galvanomagnetic effect that shows up as a Hall field and carrier deflection and magnetoresistance with different sample configurations. Sensing devices include: Hall effect devices-bulk Hall plate and magnetic field sensitive field effect transistors (MAGFETS), magneto-transistors, magneto-diodes, and current domain magnetometers. A comprehensive review paper by H.P. Baltes and volume 5 of are suggested for further reading.

5. Future trends

Physical microsensors have been developed over the past 25 years and nearly all possible sensors have been explored. Future research and development will be focused on the following:

1. New materials, principles, and technology that push sensors beyond current limitations, such as high temperature devices, submicron devices, self-powered sensors, etc.
2. Sensor arrays and multiple sensors to improve the sensitivity, selectivity and stability, as well as to sense the distribution of the measurands.
3. Integrated sensors with built-in intelligence to recognize desired signals in noise.
4. Integrated sensors with actuators to provide self-checking and self-calibration functions, as a step towards integrated systems.