WILEY

HANDBOOK OF Measuring System DESIGN

EDITORS **PETER H. SYDENHAM** **RICHARD THORN**

ARTICLE OFFPRINT

108: Transducer Fundamentals

Paul P.L. Regtien

University of Twente, Enschede, The Netherlands

1 DEFINITION OF A TRANSDUCER

A transducer is an essential part of any information processing system that operates in more than one physical domain. These domains are characterized by the type of quantity that provides the carrier of the relevant information. Examples are the optical, electrical, mechanical, and magnetic domains. A transducer is the part of the instrument that converts information about a measurand from one domain into another, ideally without information loss. So, a transducer can formally be defined as *a device that converts one form of energy into another*, with the intention of preserving information.

A transducer has at least one input and one output. In measuring instruments, where information processing is performed by electrical signals, either the output or the input is of electrical nature (voltage, current, resistance, capacitance, etc.), whereas the other is a nonelectrical signal (displacement, temperature, elasticity, etc.). A transducer with a *nonelectrical input* is an *input transducer*, intended to convert a nonelectrical quantity into an electrical signal so as to measure that quantity. A transducer with a *nonelectrical output* is called an *output transducer*, intended to convert an electrical signal into a nonelectrical quantity so as to control that quantity. So, a more explicit definition of a transducer is *an electrical device that converts one form of energy into another*, with the intention of preserving information.

According to common terminology, these transducers are also called *sensor* and *actuator* respectively (see Figure 1).

So, a sensor is an input transducer and an actuator is an output transducer. It should be noted, however, that this terminology is not standardized. In literature, other definitions are found. In particular, some authors make a clear differentiation between a *sensor* and a (input) *transducer*, stressing a distinction between the element that performs the physical conversion and the complete device, for instance, a strain gauge (transducer) and a load cell with one or more strain gauges (sensor), or even vice versa. Modern sensors not only contain the converting element but also contain part of the signal processing (analog processing such as amplification and filtering, Analogue-to-Digital-conversion and even some digital electronics). Many such sensors have the electronics integrated with the transducing part onto a single chip. Present day sensors may have a bus compatible output, implying full signal conditioning on board. There is a trend to also include transmission electronics within the device, for instance, in biomedical applications.

Signal conditioning may be included

- to protect the sensor from being loaded or to reduce loading errors;
- to fit the sensor output to the input range of the Analogue-to-Digital Converter (ADC);
- to enhance the S/N (signal-to-noise ratio) prior to further signal processing;
- to generate a digital, bus compatible electrical output;
- to transmit measurement data for wireless applications.

Handbook of Measuring System Design, edited by Peter H. Sydenham and Richard Thorn.
 ISBN: 0-470-02143-8.

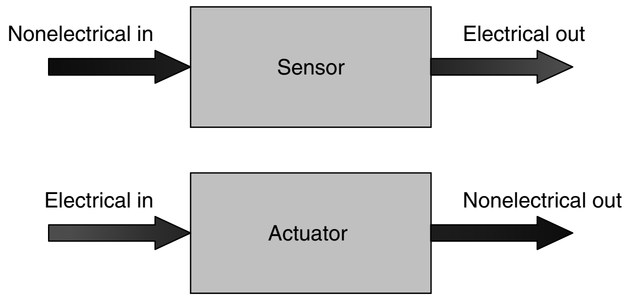

Figure 1. Input transducer (sensor) and output transducer (actuator).

In conclusion, the boundaries between sensor and transducer, as proclaimed in many sensor textbooks, are disappearing or losing their usefulness: the user buys and applies the sensor system as a single device, with a nonelectrical input and an electrical (analog, digital, bus compatible) output.

2 CATEGORIZING SENSORS

A sensor (or input transducer) performs the conversion of information from the physical domain of the measurand to the electrical domain. Many authors have tried, more or less successfully, to build up a consistent categorization of sensors (*see also* **Article 112, Systematic Description of Sensors, Volume 2**). It is not easy to create a consistent systematic description encompassing all sensor principles. There is at least consensus on a division into two groups of sensors: direct and modulating sensor types (see Figure 2). The distinguishing property is the need for auxiliary energy.

Direct sensors do not require additional energy for conversion. As information transport cannot exist without energy transport, a direct sensor withdraws the output energy directly from the measurement object. As a consequence, loss of information about the original state of the object may occur. There might also be energy loss, for instance, heat. An important advantage of a direct sensor is its freedom from offset: at zero input, the output is essentially zero. Examples of direct sensors are the thermocouple and the piezoelectric force and acceleration sensor.

Indirect sensors or *modulating sensors* use an additional energy source that is modulated by the measurand; the sensor output energy mainly comes from this auxiliary source, and just a fraction of energy is withdrawn from the measurement object. In this respect, modulating sensors do not significantly load the measurement object and hence are more accurate than direct sensors. Most sensors are of the modulating type, for instance, all resistive and capacitive sensors and many inductive sensors.

There are many sensors on the market, over a million types worldwide. A categorization of sensors would help make the proper choices and also make it easier to understand, but a useful basis for categorization is difficult to define. There are various possibilities such as

- according to the measurand,
- according to the conversion principle,
- according to the domain of the measurand,
- according to application fields.

All of them have their limitations. For instance, the number of measurands is rather large, making the first option not very practical. Figure 3 gives an overview of the most common physical quantities for which sensors are available, after Middelhoek and Noorlag (1981). However, the list is not complete. Moreover, many quantities can be measured by a variety of sensor types. For example, position can be measured using resistive, capacitive, inductive, acoustic, and optical methods.

The second option, according to the conversion principle, is often used for the reason that the sensor performance is mainly determined by the physics of the underlying principle of operation. On the other hand, a particular type of sensor might be suitable for a variety of physical quantities and in many different applications. For instance, a magnetic sensor of a particular type could be applied as a displacement sensor, a velocity sensor, a tactile sensor, and so on. For all these applications, the performance is limited by the physics of this magnetic sensor, but the limitations manifest in completely different ways.

A closer look at the various conversion effects may lead to the observation that the electrical output of a sensor depends either on a material property, or the geometry, or a movement. Figure 4 tabulates these three phenomena, for various types of sensors. The figure gives the material parameter, the geometric parameter, and the velocity induced parameter, together with associated sensors.

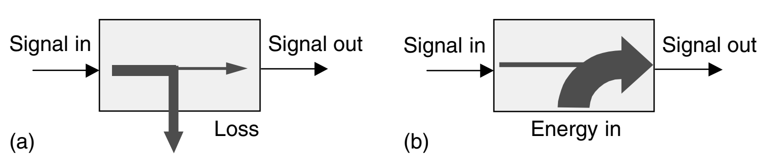

Figure 2. (a) Direct sensor and (b) modulating sensor.

1. Mechanical, solids
acceleration
angle
angular velocity
area
diameter
distance
elasticity
expansion
filling level
force
gradient
hardness
height
length
mass
moment
movement
orientation
pitch
position
pressure
proximity
rotation
roughness
shape
tension
torque
torsion
velocity
vibration
weight

2. Mechanical, fluids
density
flow direction
flow velocity
level
pressure
rate of flow
viscosity
volume

3. Thermal
enthalpy
entropy
temperature
thermal capacity
thermal conduction
thermal expansion
thermal radiation

4. Optical
color
light polariaztion
light wavelength
luminance
luminous intensity
reflection
refractive index

5. Acoustic
sound frequency
sound intensity
sound polarization
sound pressure
sound velocity
time of flight

6. Nuclear radiation
ionization degree
mass absorption
radiation dose
radiation energy
radiation flux
radiation type

7. Chemical
cloudiness
composition
concentration
electrical conductivity
humidity
impurity
ionization degree
moisture
molar weight
particle form
particle size
pH
polymerization degree
reaction rate
redox potential
thermal conductivity
water content

8. Magnetic and electrical
capacity
charge
current
dielectric constant
electric field strength
electric power
electric resistance
frequency
inductivity
magnetic field strength
phase
pulse duration
signal distortion

9. Time
time

Figure 3. Parameters for which sensors are on the market (after Middelhoek and Noorlag (1981)).

Type	Material property	Geometry	Relative movement
Resistive	Resistivity (piezoresistor,LDR)	Relative length (potentiometer; metal strain gauge)	
Capacitive	Permittivity (fluid level sensor)	Relative electrode distance capacitive displacement (LVDC)	
Magnetic	Permeability (magnetoresistor)	Distance source-detector (magnetic displacement sensor)	Induction (magnetic velocity sensors)
Inductive		Inductance, self-inductance, mutual inductance reluctance (inductive displacement sensors, LVDT, resolver)	Induction (inductive velocity sensors)
Optical	Index of refraction; absorptivity (fibre optic sensors)	Distance transmitter–receiver (intensity modulation sensors; interferometer; TOF sensor) transmissivity, reflectivity (optical encoder, tachometer)	Doppler frequency (Doppler velocimeter)
Acoustic	Acoustic impedance	Distance transmitter–receiver (TOF displacement sensors)	Doppler frequency (Doppler velocimeter)
Piezoelectric	Polarization (piezoelectric sensors)	Deformation (piezoelectric sensors)	

Figure 4. Three groups of parameters, with examples of sensors.

A categorization based on the domain of the measurand is too coarse, and domain definitions are not unambiguous.

Finally, an application field provides no restricted set of sensors, since in each field (biomedical, automotive, agriculture), almost all types of sensors could be applied.

3 TERMINOLOGY OF SENSORS

We have defined a sensor as a device that performs the conversion of information from the physical domain of the measurand to the electrical domain. For instance, a position sensor converts position information into an electrical signal (a voltage, current, etc.). Obviously, a position sensor measures position. However, there are different names for different applications and situations, for instance,

Distance sensor: measures the length of the straight line between two defined points;
Position sensor: measures the coordinates of a specified point of an object in a specified reference system;
Displacement sensor: measures the change of position relative to a reference point;
Proximity sensor: (i) determines the sign (positive or negative) of the linear distance between an object point and a fixed reference point; (ii) a contact free displacement or distance sensor for short distances (down to zero);
Level sensor: measures the distance of the top level of a liquid or a granular substance in a container with respect to a specified horizontal reference plane;
Angular sensor: measures the angle of rotation relative to a reference position;
Encoder: displacement sensor (linear or angular) with digital output:
Tilt sensor: measures the angle relative to the earth's normal;
Tachometer: measures the rotational speed;
Vibration sensor: measures the motion of a vibrating object in terms of displacement, velocity or acceleration;
Accelerometer: measures acceleration.

Sensors for the measurement of force and related quantities are as follows:

Pressure sensor: measures pressure difference, relative to either vacuum (absolute pressure), a reference pressure or ambient pressure;
Force sensor: measures the (normal and/or shear) force exerted on the active point of the transducer;
Torque sensor: measures torque (moment);
Force–torque sensor: measures both forces and torques (up to six components);
Load cell: force or pressure sensor, for measuring weight;
Strain gauge: measures linear relative elongation (positive or negative) of an object, caused by compressive or tensile stress.

Many sensors have been given names according to their operating principle or construction. Examples are as follows:

LVDT: or linear variable displacement transformer, a device that is basically a voltage transformer, with a linearly movable core;
Gyroscope: a device for measuring the angular velocity, based on the gyroscopic effect occurring in rotating or vibrating structures;
Hall sensor: measures magnetic field on the basis of the Hall effect, after the American physicist Edwin Hall (1855–1938).

Some sensors use a *concatenation* of conversion steps. A displacement sensor combined with a spring can act as a force sensor. The measurand (force) is converted to a displacement, which in turn is converted into an electrical signal. In combination with a calibrated mass, a displacement sensor can serve as an accelerometer: the measurand (acceleration) is converted into a force (by the inertial effect), the force into a displacement (by the spring), and the displacement into an electrical signal. The performance of such transducers not only depends on the original sensor but also on the added components: in the case of the accelerometer, it depends on the spring compliance and the seismic mass respectively.

Information about a particular quantity can also be obtained by *calculation* or by additional electronic *signal processing* using relations between quantities. The accuracy of the result not only depends on the errors in the quantities that are measured directly but also on the accuracy of the parameters in the model that describes the relation between the quantities involved. For instance, in an acoustic distance measurement, the distance is calculated from the measured time-of-flight (with associated errors) and the sound velocity. An accurate measurement result requires knowledge of the acoustic velocity of the medium at the prevailing temperature.

Speed and acceleration can be measured using a displacement sensor, by differentiating its output signal once or twice respectively, and vice versa: by integrating the output signal of an accelerometer, a velocity signal is obtained, and by a second integration, a position signal is obtained. Obviously, the performance of the final result depends on the quality of the signal processing. The main problem with differentiation is the increased noise level (in particular in the higher frequency range), and integration may result in large drift due to the integration of offset.

4 SENSOR PERFORMANCE

Imperfections of a sensor are usually listed in the data sheets provided by the manufacturer. These sensor specifications inform the user about deviations from the ideal behavior. The user has to accept technical imperfections, as long as they do not exceed the specified values.

Any measuring instrument, and hence any sensor, has to be fully specified with respect to its performance. Unfortunately, many data sheets show lack of clarity and completeness. Gradually, international agreements about formal error descriptions are being established. An extensive description of measurement errors and error terminology can be found in IOS (1995). Further, there is an international standard on transducer nomenclature and terminology ISA (1975). Finally, various international committees are working toward a uniform framework for specifying sensors.

The characteristics that describe sensor performance can be classified into four groups:

- static characteristics, describing the performance with respect to very slow changes;
- dynamic characteristics, specifying the sensor response to variations in time and in the measurand (the quantity that has to be measured);
- environmental characteristics, relating the sensor performance after or during exposure to specified external conditions (pressure, temperature, vibration, radiation);
- reliability characteristics, describing the sensor life expectancy.

First, we define some common specifications that apply to sensors:

- sensitivity;
- nonlinearity and hysteresis;
- resolution;
- accuracy;
- offset and zero drift;
- noise;
- response time;
- frequency response.

4.1 Sensitivity

The sensitivity of a sensor is defined as the ratio between a change in the output value and the change in the input value that causes that output change. Mathematically, the sensitivity is expressed as $S = \partial y/\partial x$, where x is the input signal (measurand) and y is the output (an electrical signal). Usually, a sensor is also sensitive to changes in quantities other than the intended input quantity, such as the ambient temperature or the supply voltage. These unwelcome sensitivities should be specified as well, for a proper interpretation of the measurement result. To have a better insight in the effect of such unwanted sensitivities, it is often related to the sensitivity of the measurement quantity itself.

Example 1 The sensitivity of a particular displacement sensor with voltage output is specified as 10 mV mm^{-1}. Its specified temperature sensitivity is 0.1 mV K^{-1}. Since 0.1 mV corresponds to a displacement of 10 mm, the temperature sensitivity can also be expressed as 10 mm K^{-1}. A temperature rise of 5 °C results in an apparent displacement of 50 mm.

Example 2 The sensitivity of a particular type of temperature sensor is 100 mV K^{-1}, including the signal conditioning unit. The signal conditioning part itself is also sensitive to (ambient) temperature and appears to create an extra output voltage of 0.5 mV for each °C rise in ambient temperature (not necessarily the sensor temperature). So, the unwanted temperature sensitivity is 0.5 mV K^{-1} or 0.5/100 = 5 mK K^{-1}. A change in ambient temperature of ±10 °C gives an apparent change in sensor temperature equal to ±50 mK.

4.2 Linearity and hysteresis

If the output y is a linear function of the input x, the sensitivity S does not depend on x. In the case of a nonlinear transfer function $y = f(x)$, S does depend on the input or output value. Often, a linear response is preferred to reduce computational burden in, for instance, multisensor control systems. In that case, the sensitivity can be expressed with a single parameter. The transfer of a sensor with a slight nonlinearity may be approximated by a straight line, to specify its sensitivity by just one number. The user should be informed about the deviation from the actual transfer; this is specified by the nonlinearity.

The linearity error of a system is the maximum deviation of the actual transfer characteristic from a prescribed straight line. Manufacturers specify linearity in various ways, for instance, as the deviation in input or output units: Δx_{max} or Δy_{max}, or as a fraction of FS (full scale): $\Delta x_{max}/x_{max}$. Nonlinearity should always be given together with a specification of the straight line. The following definitions are in use:

- *Terminal nonlinearity*: based on the terminal line–a straight line between 0 and 100% theoretical full-scale points.
- *End-point nonlinearity*: based on the end-point line–the straight line between the calibrated end points of the range; coincides with the terminal (theoretical) line after calibration of zero and scale.
- *Independent nonlinearity*: referring to the best-fit straight line, according to a specified error criterion, for instance, the line midway between two parallel lines enclosing all calibration points; if the least-square error criterion for the best-fit straight line is used, this linearity error is as follows.
- *Least-square nonlinearity*: based on the least square line, the line for which the summed squares of the residuals is minimized.

Hysteresis is the maximum difference in output signal when the measurand first increases over a specified range, and next returns to the starting value. The traveled range should be specified because hysteresis strongly depends on it.

4.3 Resolution

The resolution indicates the smallest detectable increment of the input quantity. When the measurand varies continuously, the sensor output might show discontinuous steps. The value of the corresponding smallest detectable change in the input variable is the resolution: Δx_{min}. Sometimes this parameter is related to the maximum value x_{max} that can be processed (full-scale value), resulting in the resolution expressed as $\Delta x_{min}/x_{max}$.

Example 3 The resolution of a particular type of wire-wound linear potentiometer with range 10 cm is specified as 10^{-4}; assuming this is relative to the full-scale value, it means that the output changes discontinuously in steps equivalent to input displacements of 10 mm.

Example 4 A particular type of optical encoder that has a resolution of 14 bit. The smallest change in angle that can be detected by this encoder is $2\pi/2^{14} \approx 1.9 \times 10^{-4}$ rad or 0.022°.

4.4 Accuracy

Formally, the accuracy reflects the closeness of the agreement between the actual measurement result and a true value of the measurand. The accuracy specification should include relevant conditions and other quantities. Many sensor manufacturers specify the sensor performance in terms of accuracy. This specification should be viewed with suspicion, because it may or may not include particular imperfections of the sensor (nonlinearity, hysteresis, drift), and may be only valid under strict conditions.

4.5 Offset and zero drift

Most sensors are designed such that the output is zero at zero input. If the transfer characteristic does not intersect the origin (x,y = 0,0) the system is said to have offset. The offset is expressed in terms of the input or the output quantity. Specifying the input offset is preferred to facilitate a comparison with the measurand.

Example 5 The sensitivity of a particular type of force sensing system is $0.1\,\mathrm{V\,N^{-1}}$. At zero force, the output appears to be 3 mV. The (input) offset of this system is the output offset divided by the sensitivity, so 0.03 N.

A nonzero offset arises mainly from component tolerances. Offset compensation can be performed in the interface electronics or the signal processing unit. Once adjusted to zero, the offset may nevertheless change, owing to temperature variations, changes in the supply voltage or aging effects. This relatively slow change in the offset is called *zero drift*. In particular, the temperature induced drift (the *temperature coefficient* or t.c. of the offset) is an important item in the specification list.

Sometimes a system is deliberately designed with offset. Many industrial transducers have a current output ranging from 4 to 20 mA. This facilitates the detection of cable fractures or a short circuit, producing a zero output clearly distinguishable from a zero input.

4.6 Noise

Electrical noise is also specified in terms of the input quantity, to show its effect relative to that of the measurand. *White noise* (noise with constant power over a wide frequency range) is usually expressed in terms of *spectral noise power* (W/Hz), *spectral noise voltage* (V/$\sqrt{}$Hz) or *spectral noise current* (A/$\sqrt{}$Hz). Thermal noise is an example of 'white noise'.

Another important type of noise is *1/f noise* (one-over-f noise), a collection of noise phenomena with a spectral noise power that is proportional to f^{-n}, with $n = 1$ to 2.

Quantization noise is the result of quantizing an analog signal. The rounding off results in a (continuous) deviation from the original signal. This error can be considered as a 'signal' with zero mean, and a standard deviation determined by the resolution of the AD converter.

4.7 Response time

The response time is associated with the speed of change in the output upon a stepwise change of the measurand. The specification of the response time always needs to be accompanied by an indication of the input step (for instance, FS – full scale) and the output range for which the response time is defined, for instance, 10 to 90%. Creep and oscillations may make the specification of the response time meaningless or at least misleading.

4.8 Frequency response

The sensitivity of a system depends on the frequency or rate of change of the measurand. A measure for the useful frequency range is the *frequency band*. The upper and lower limit of the frequency band are defined as those frequencies for which the output signal has dropped to half the nominal value, at constant input power. For voltage or current quantities, the criterion is $1/2\sqrt{2}$ of the nominal value. The lower limit of the frequency band may be zero; the upper limit has always a finite value. The extent of the frequency band is called the *bandwidth* of the system, expressed in Hz.

All specification items only apply within the operating range of the system, which should also be specified correctly. It is given by the measurement range, the required supply voltage, the environmental conditions and possibly other parameters.

Despite the specified limitations of sensors, a sensing system can be configured in such a way that the effect of some of these limitations are eliminated or at least reduced. Various possibilities of error reduced designs are described in **Article 16, Reduction of Influence Factors, Volume 1**.

5 ROLE AND USE OF ACTUATORS

Similar to the definition of a sensor (or input transducer) an actuator (or output transducer) can be defined as *a device that converts information from the electrical domain to another physical domain*, with preservation of information. Actuators and sensors can be considered as each other's opposites. Actually, many transduction principles are reversible (notably the direct conversion types), which means that the same device can act as both – a sensor as well as an actuator (for instance, a piezoelectric sensor and actuator). However, an actuator should, in general, be able to deliver output energy, whereas a sensor operates best when the amount of energy taken from the measurement object is close to zero. Reversibility does not imply that both options can be combined in a single device.

Since an actuator has to produce output energy, its construction and dimensions differ largely from that of sensors. There are far less physical conversion principles suitable for actuation than for sensing. However, the variability in types is larger, according to the application and the required energy.

Some actuator types, grouped according to the physical domain of the output are as follows:

- *Mechanical output*: induction motor; permanent magnet electromotor; stepper motor, piezoelectric actuators (linear and rotating movements), relay;
- *Optical output*: light emitting diode (LED), solid state laser diode (SSLD); incandescent lamp;
- *Thermal output*: electric heater; (thermo)electric cooler;
- *Acoustic output*: piezoelectric ultrasound transmitter; electrostatic ultrasound transmitter.

As with sensors, actuators too can operate either in a direct way or in a modulating mode. Most actuators are of the direct type: no auxiliary energy is required. Examples of modulating actuators are an oscilloscope (the electrical signal modulates the deflection of an electron beam and hence the position of the light spot on the screen) and an LCD (the electrical signal modulates the transmission of an auxiliary light source by influencing the transmittance of the crystal material).

The performance of a sensor system can sometimes be improved by including the sensor in a feedback loop. The electric output of the sensor is amplified and supplied to an actuator, having the inverse conversion characteristic. So its output is of the same nature as that of the measurand, or at least is closely related.

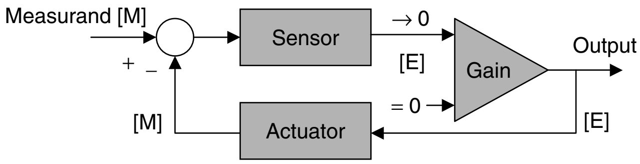

Figure 5. Feedback sensor configuration; [M] denotes the domain of the measurand, [E] the electrical domain.

In the steady state, and assuming stability and a high loop gain, the input of the sensor is compensated (made zero) by the feedback action, see Figure 5. The compensating signal is delivered by the actuator, and hence its input equals the output of the measurement system. Since the input signal of the sensor is zero, its static characteristics (sensitivity, nonlinearity, limited range) are irrelevant to the measurement result. The only requirement is zero offset. The performance of the measurement completely relies on the characteristics of the actuator. Further details of feedback sensing systems are discussed in more detail in the article **Article 113, Force-feedback Sensors, Volume 2**.

Like with sensors, also some actuators operate on the basis of a concatenation of conversions. Examples are as follows:

- *Thermal microactuator*: a particular element of the microstructure that is heated by an electric current upon which it expands, producing a displacement of the tip.
- *Reed switch*: a coil around the switch that is activated by a current, producing a magnetic field, which closes the contacts of the switch.
- *Incandescent lamp*: electric current heats up a filament that emits visible light above a particular temperature.

In the past decades, sensors tended to get smaller and actuators larger. Development in microtechnology, however, allows the creation of microactuators as well see, for instance, Tabib-Azar (1997). Consequently, such microactuators can produce little energy, and are therefore only useful as part of a microsystem, including (micro)sensors, signal processing, and actuation.

The role of an actuator in a measurement system can be

- indication of the measurement result (optical display),
- registration of the measurement result (magnetic or optical head; driving a plotter pen),
- control of the measurement system (scanning in 1, 2, or 3 dimensions),
- control of a process to perform measurements (for instance, in wear detection, fatigue tests)
- control of the sensing part of the measurement system (feedback sensors).

REFERENCES

Middelhoek, S. and Noorlag, D.J. (1981) Three-Dimensional Representation of Input and Output Transducers. *Sensors and Actuators*, **2**, 29–41.

Tabib-Azar, M. (1997) *Microactuators – Electrical, Magnetic, Thermal, Optical, Mechanical, Chemical and Smart Structures*, Kluwer Academic Publishers, Boston, MA.

IOS (1995) *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva.

ISA (1975) S37.1: Electrical Transducer Nomenclature and Terminology.