Achieving reliable, precise sensor outputs in linear motion control systems

*By Josef Janisch*
*Senior Product Marketing Manager, austriamicrosystems AG*

In order to implement a motion control system, it is necessary to sense the movement of a device that is rotating or travelling along a straight line. This requirement is common in a growing number of applications, from X-Y tooling tables and pneumatic pistons in industrial applications, to sliding doors and zooming lenses in commercial or consumer products.

Key requirements for such motion control systems are that they should be robust and compact. This article compares three commonly-used methods of implementing a linear motion sensor: resistive, optical and magnetic. To optimise motion-control designs, engineers should consider how each of these sensing technologies affects long-term, extended operation.

## *Resistive linear motion sensing*

The technology that has traditionally been used to implement a linear motion sensor is based on a sliding potentiometer. This method has been used for many decades, as these contacting devices are comprised of low-cost components and are easy and cheap to assemble. They can be manufactured in virtually any length and shape, and offer high-resolution position feedback.

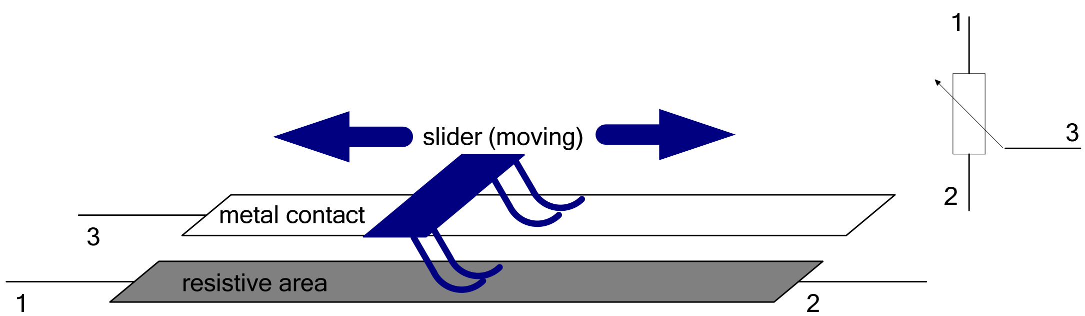


*Fig. 1: Linear motion sensing using a variable resistor*

As shown in Figure 1, the linear potentiometer consists of a moving slider with two contacts: one running along a resistive area and one along a metal contact. As the slider is moved from left to right the resistance between terminals 1 and 3 increases; at the same time the resistance between terminals 2 and 3 decreases. These resistance values can be used to calculate the position of the slider.

The resistive area is typically made of either conductive plastic or Cermet (a mixture of Ceramic and Metal). Both materials provide mechanical and thermal stability with lifetimes of several hundred thousand cycles. However, this lifetime decreases in the presence of dirt and dust. A speck of dust between the resistive area and the wiper can thus severely affect the reliability of the potentiometer. For this reason, a tight dust-proof sealing of the sensor is required. In practice, however, a long-lasting hermetic sealing is often not achievable. This means that for high-reliability operation, particularly in harsh environments, other methods of motion sensing must be considered.

## *Optical linear motion sensing*

One way to overcome the effect of dirt and dust on physical contacts is to adopt a contactless sensor. For this reason, optical sensors are now commonly used in motion detection. An optical motion sensor uses one of two methods: transmissive or reflective. The fundamental principles of each are the same.

A transmissive sensor is shown in Figure 2. A light source, usually an LED, is shone across an optical strip, consisting of equally spaced and equally sized slits of clear and opaque material. As the strip moves, the light path is interrupted.

Behind this optical patterned strip is a grid mask with small openings. The width of each opening is matched to the pattern on the optical strip, so that light is fully transmitted in one position and fully blocked in another. The two openings in the mask are shifted by half a strip width, with a photodetector placed behind each. As the optical strip moves, these photodetectors therefore create two phase-shifted signals that can be used to sense not only motion but also the direction of motion.

A pulse-shaping circuit, usually a Schmitt trigger, is used to convert the nearly sinusoidal signals from the photodetectors to rectangular pulses, commonly known as a quadrature signal. By calculating which signal has a rising edge while the other signal is low, the system can compute the direction of movement, while the frequency of pulses indicates speed of motion.

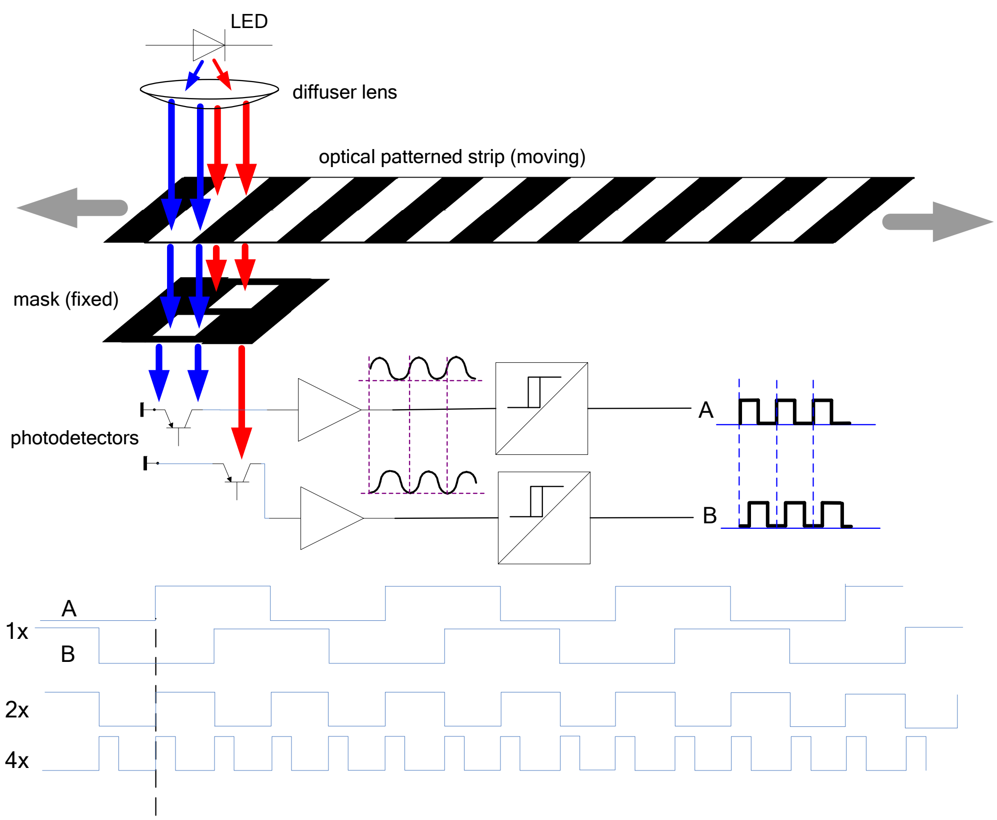


*Fig. 2: Linear motion sensing using an optical grid array*

Reflective optical sensors work in a similar way, but instead of using an optical patterned strip to vary the transmitted light they use a patterned reflective strip to vary the reflected light. The reflective strip typically consists of a pattern of black and white marks. While a white mark reflects the light rays to the photodetectors, a black mark absorbs the light. The photodetectors are placed on the same plane as the LED; picking up the reflected light, they generate a quadrature output.

This technique is more robust than the resistive contact technology described above, as it is a contactless sensing method, and so does not suffer from mechanical stresses. Nevertheless, at a certain level contamination on the optical path will give rise to errors in the output. As with the resistive contact technique, sealing of the sensor housing provides an imperfect solution to this problem, as the housing is expensive to assemble, and provides no guarantee that contamination will not enter the device during its expected lifetime.

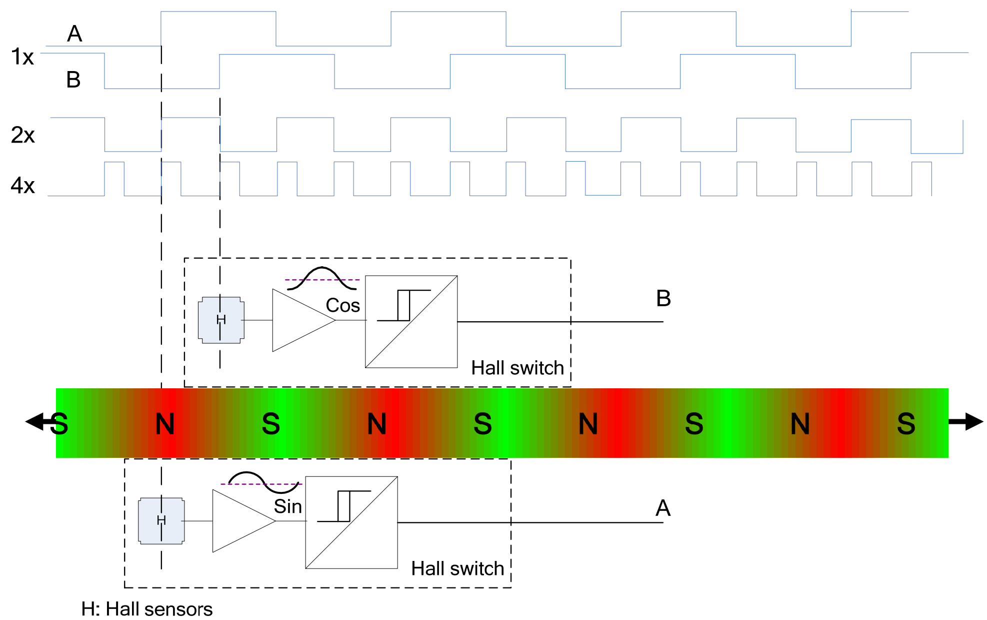


*Fig. 3: Linear motion sensing using Hall switches*

Using the same basic principle as an optical encoder, contactless linear motion sensors can be built using Hall switches and a multi-pole magnetic strip (see Figure 3).

A Hall-effect sensor is a transducer that varies its output voltage in response to changes in magnetic field. Two Hall switches replace the photodetectors shown in Figure 2. The magnetic strip is used in place of the optical patterned strip. Each Hall switch turns on and off depending on the polarity of the magnetic strip passing it, as shown by signal A in Figure 3. Since the two Hall switches are placed half a pole length apart, they produce the same signal but electrically phase-shifted by 90°. The phase of these signals and the frequency of pulses can be used to determine the direction and speed of motion respectively.

The magnetic-sensing method described above is easy to build, cheap and very robust. It is insensitive to dirt, moisture and other contaminants and so is especially useful in rough or hazardous environments. A drawback of this approach however is the limited resolution, which is restricted by the achievable pole length of the magnetic strip. The practical limit for pole length in this configuration is around 0.5mm. Furthermore, since this Hall-sensor design relies on detecting magnetic fields, its operation can be affected by unwanted external magnetic fields.

One way to shield the sensor from unwanted external fields is to encase it in a ferro-magnetic housing, such as sheet metal. This, however, adds component and assembly cost, and can be too bulky for many applications. To use the robustness offered by magnetic sensors, without the limitations of resolution and magnetic interference, designers can consider a new type of sensor device: the integrated Hall encoder.

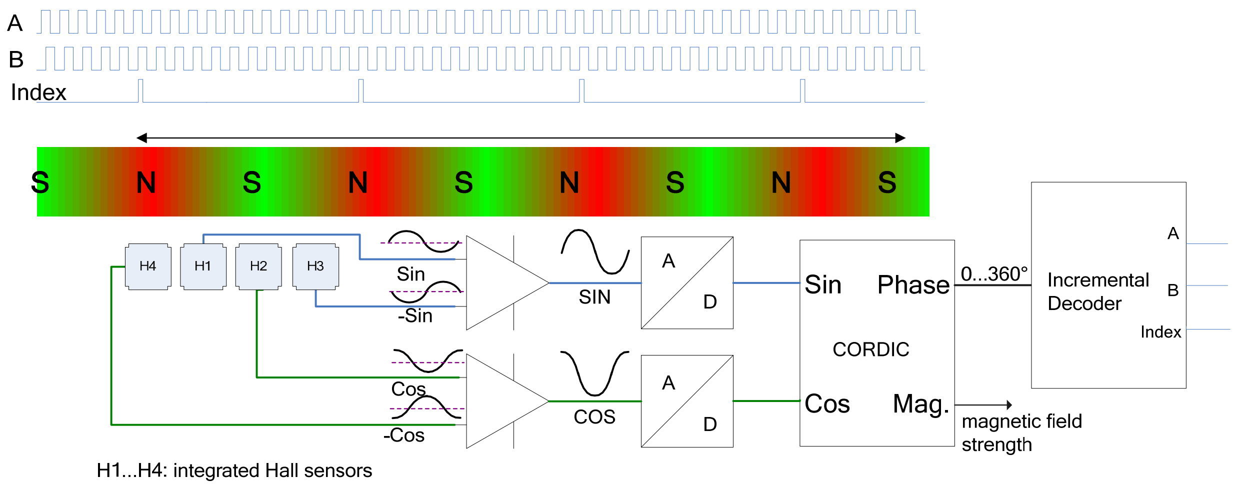


*Fig. 4: Integrated Hall-sensor array with interpolator*

Integrated magnetic sensors use linear Hall sensors instead of Hall switches. Linear Hall sensors provide an analogue output that is proportional to the strength of the magnetic field perpendicular to the Hall sensor. Sliding a multi-pole magnetic strip over a Hall sensor generates a sinusoidal signal at the output of the sensor.

By placing four Hall sensors exactly half a pole length apart, four sinusoidal signals are generated as the magnet slides over the sensors. Each signal is phase shifted by 90° from its neighbouring sensor, as shown in Figure 4. In mathematical terms, the four signals generated (H1, H2, H3 and H4) represent sine, cosine, inverted sine and inverted cosine.

Combining sine with inverted sine, and cosine with inverted cosine, provides another sine and cosine signal of double amplitude. This combination requires one of the input signals to be inverted, which therefore inverts interference from external magnetic fields. Any common-mode interference is therefore cancelled out when sine and inverted sine are combined, and when cosine and inverted cosine are combined. Using this method, no magnetic shielding is required, and even strong magnetic fields in the vicinity of the sensor cannot disturb it.

The two resulting signals are then digitised by high resolution Analogue-to-Digital converters (ADCs) and used as inputs for a high-performance signal processor, which converts sine and cosine signals to a high-resolution digital angle and magnitude output. The precision of the digital outputs from the motion sensor depends on how well the signal processor resolves the analogue sine and cosine values.

## High-resolution magnetic motion sensing in practice

The market provides a number of options for the design engineer wanting to implement motion control with a high-resolution magnetic motion sensor. All magnetic sensors offer immunity to the contamination that bedevils resistive contact and optical sensing devices.

Austriamicrosystems provides devices aimed at designers who require a high-resolution output that can be interfaced directly to a microcontroller, in a compact package. The latest such devices are the AS5304 and AS5306 (see Figure 5), for linear and off-axis rotary motion sensing, in conjunction with multi-pole magnets. These devices offer resolutions down to 15µm.

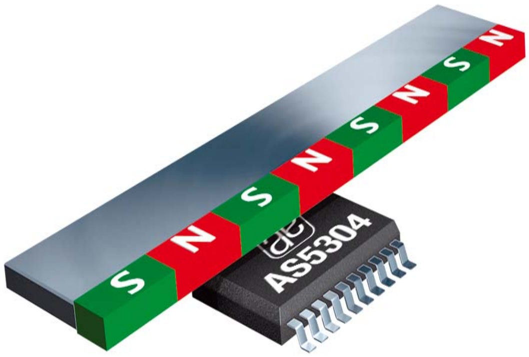


*Fig. 5: AS5304 - integrated single-chip device for high-resolution linear and off-axis rotary motion sensing*