Redundancy-enabled stabilisation of linear encoder performance: The biSLIDER

Original

Availability:
This version is available at: 11696/58305 since: 2018-07-10T19:09:09Z

Publisher:
Elsevier

Published
DOI:10.1016/j.cirp.2018.04.004

Terms of use:
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)
1. Introduction

Linear encoders are widely used to measure relative displacements, e.g. a carriage relative to a base of a machine tool. They are robust and reliable even in non-cooperative environments, and inexpensive compared with other devices of similar performance [1]. Apart from local errors due to e.g. interpolation and noise (usually negligible in medium-long encoders), their accuracy is dominated by thermal expansion and strain. The former can be either compensated by measuring the temperature or minimised by resorting to expensive low expansion materials. The strain can be minimised by suitably decoupling the scale tensioning from the mount. In all cases, only a global linear compensation is achieved, while local distortions – due to e.g. a concentrated heat source or to gluing [2] – remain.

The technique described in this paper (biSLIDER - bi-Sensor for Locally Interpolated Differential Error Recovery [3]) is based on the simultaneous use of two reading heads per encoder, and aims at stabilising the performance over time, by recovering global and local perturbations. It is not intended for traceability or for compensating the static errors of the encoder: a prior traceable compensation according to conventional techniques, e.g. with a laser interferometer, is assumed.

The use of two simultaneous heads in linear encoders has been proposed for several applications [we will not consider rotary and areal encoders]. In Sakagami et al. it joins two consecutive legs of a same long scale [4]. Günter exploits the known difference in the coefficients of thermal expansion (CTE) of the head spacer and of the scale to derive the temperature and then compensate [5]. Schuchardt et al. rely on accurate knowledge of the head separation, on a special movable pattern and on a dedicated phase detection to compensate the encoder [6]. Li et al. [7] use three heads to make the encoder absolute.

With the biSLIDER, the thermal expansion and the strain are compensated simultaneously, with no need for calibration or accurate knowledge of the head spacer length and CTE, and standard components only are required.

The following sections complement the biSLIDER concept [3] with its mathematical model, the analysis of the error sources and the experiments validation.

2. The biSLIDER concept and theory of operation

Two conventional reading heads are mounted on a moving spacer (the bislider) and read a same scale simultaneously (Figure 1). In normal use, a head is redundant and not considered. A reference state is chosen, at which the encoder exhibits reference performance, typically immediately after conventional error compensation, e.g. by a laser interferometer. Let us call this conventional compensation – not a part of the biSLIDER – the static compensation, and assume it is active throughout.

The bislider is assumed invariant in time. This is much more easily achieved for the short and free-standing bislider than it is for the long scale attached to a machine base. The thermal invariance can be achieved either by low CTE material or by temperature compensation, or both.

At reference state, the heads H2 and H1 are subsequently homed at the same home. At the second homing, \( x_1 = 0 \) and \( x_2 = b \), where \( b \) is the head spacing. Then the bislider quickly scans a full stroke in steps of length \( b \), so that at each one the H1 is at the same position as H2 was at the previous step. The difference \( x_2 - x_1 \) at each step is recorded in a reference state table. The deviations of the entries from the nominal value \( b \) portrait the encoder performance at the reference state. This reference state procedure is done preliminarily once for all.

At will, e.g. periodically or when encoder perturbations are suspected, a recovery procedure is performed. This is identical to the reference state procedure, but the resulting table is separately recorded as a recovery table. Because of the invariance of the bislider, any difference between the reference state and the recovery tables is due to the scale. A third table, the dynamic correction table, is then calculated as the accumulated difference of the reference state to the recovery tables:

\[
dyn\alpha_k = \sum_{i=1}^{k} (\text{ref}_i - \text{rec}_i) = \text{ref}_k - \text{rec}_k + \text{dyn}_{k-1} \quad (1)
\]

where \( \text{ref}_i, \text{rec}_i, \text{dyn}_i \) are the \( i \)-th entries of the reference state, recovery and dynamic correction tables, respectively. The dynamic correction table is a lookup table to compensate the encoder errors and to recover its performance to the reference state. It applies to either head, shifted one-step back for H2. The reference state and recovery procedures are very quick and require standard components only. The bislider is a simple spacer. Most machine tool CNC’s offer spare encoder channels, and the extra computation required is trivial (sum and differences of table entries). This makes the biSLIDER suitable for retrofitting existing machines, too. The presence of two spaced heads on a same scale reduces the stroke of a value \( b \) from the scale length. In newly designed applications, a longer scale can be easily fit; in retrofitting, the stroke is reduced.

Figure 1. Schematic representation of the biSLIDER concept.
3. Error sources and uncertainties

The entries $\delta a_k$ in eq. (1) are a cumulative sum of elementary compensations at the preceding individual legs, $a_i = a_i - a_k$, $i \leq k$. The overall error will also be a sum of elementary errors incurred in individual legs. The uncertainty sources can be divided into two groups: those whose leg-specific components are uncorrelated and those that are (fully) correlated. The leg-specific components sum up quadratically in the former and linearly in the latter case ([8] § 5.2.2 NOTE 1). If we assume that the leg-specific components of a same uncertainty source are all equal (i.e. there is no better or worse leg than any other), $u_i^2 = u^2 \forall i$, then the combined uncertainty $u_k$ for each source is $u_k = \sqrt{u_i^2}$ and $u_k = k u_i$ in the former and latter cases, respectively. At the leg boundaries, $x_k$, the number of preceding legs is $k = x_k/b$, resulting in $u_k = (u/\sqrt{b})x_k$ and $u_k = (u/b)x_k$ (Table 1). For each source, the combined uncertainty is proportional either to the absissa $x_k$ or to its square root, depending on the correlation of the leg-specific components. In all cases, the spacing $b$ is at the denominator: the longer the better. We will discuss the design choice of $b$ later in the conclusions.

<table>
<thead>
<tr>
<th>Uncorrelated</th>
<th>Fully correlated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td>$u_k = (u/\sqrt{b})x_k$</td>
</tr>
<tr>
<td>Resolution</td>
<td>$u = \varepsilon/(\sqrt{4})$</td>
</tr>
<tr>
<td>Homing</td>
<td>$u = \varepsilon/(\sqrt{4})$</td>
</tr>
<tr>
<td>Thermal expansion</td>
<td>$u = b_\gamma$</td>
</tr>
</tbody>
</table>

$\varepsilon$: Resolution of the encoder; $b$: head separation (bislider length); $u_\gamma$: relative uncertainty due to the bislider thermal expansion.

3.1. Calibration of the bislider (uncorrelated)

The value $b$ does not enter the eq. (1) but only sets the carriage steps during the reference state and recovery procedures. If the actual steps differ from $b$, then H2 in a step is not exactly at the same position as H1 in the subsequent step. The quantities $\delta a_i$ and $\delta a_k$ are insensitive to the actual position of the bislider being reading differences. The smoothness of the scale error function well tolerates imperfect coincidences between opposite ends of subsequent legs. The calibration of the bislider is not needed, as the resulting uncertainty is negligible; the bislider does not intend to provide traceability. This uncertainty component is of the first group (uncorrelated: the error at each step is local) and negligible in practical cases.

3.2. Resolution (uncorrelated)

The entries $\delta a_i$ and $\delta a_k$ suffer finite resolution of each head; $\delta a_i$ is the difference of the two and then involves four readings. As resolution errors are uncorrelated, the uncertainty is $u = \varepsilon/(\sqrt{4})$ (Table 1), where $\varepsilon$ is the resolution of the encoder.

3.3. Homing (correlated)

The homing sets the scale origins as read by either head. Any error occurring at homing, due to e.g. parallax and finite resolution, results in a zero error. What affects the bislider is the difference between recovery and reference states, of the differences at the two heads. This highly differential scheme cancel most errors out, leaving the resolution as the predominant one. The coefficient $\delta a_i$ depends on the bislider CTE and on the change in temperature, or on the errors in the CTE and temperature values used for compensation. Due to the short time taken by the reference state and recovery procedures, $\gamma$ is usually almost equal for all legs, and the error accumulates, thus projecting the proportional error $\gamma$ of the bislider onto the scale. The uncertainty components of each leg are $u = b_\gamma$, where $u_\gamma$ is the relative uncertainty due to the bislider expansion, and are correlated.

3.4. Bislider thermal expansion (correlated)

The bislider is supposed invariant. If a differential (uncompensated) expansion occurs at the recovery state relative to the reference state, then $\delta a_i$ is affected by an error $\varepsilon = \gamma b$ proportional to the bislider spacing $b$. The coefficient $\gamma$ depends on the bislider CTE and on the change in temperature, or on the errors in the CTE and temperature values used for compensation. Due to the short time taken by the reference state and recovery procedures, $\gamma$ is usually almost equal for all legs, and the error accumulates, thus projecting the proportional error $\gamma$ of the bislider onto the scale. The uncertainty components of each leg are $u = b_\gamma$, where $u_\gamma$ is the relative uncertainty due to the bislider expansion, and are correlated.

4. Experimental validation

Two independent set ups were used to validate the bislider, in the field and in lab. In each one, a linear encoder was perturbed and the bislider applied to recover the perturbation. A laser interferometer served as independent reference. The bislider does not distinguish between thermal expansion and strain. As thermal perturbations are difficult to control and prone to risk of cross effects on the reference measurements, the perturbation was induced by stressing the scale. Static compensations were performed before the tests, using the same interferometer as a reference. This resulted in nominally null errors at the reference states. In the following, $E_{ax}$ denotes the encoder’s error of indication, by analogy with the linear positioning error motion of a machine carriage [9].
4.1. Validation in the field

A bislider was mounted on the 1 µm resolution Heidenhain linear encoder (steel graduated tape) of the x axis of a boring/milling machine Alesamonti MAF45 at the Alesamonti premises (Figure 2). A Renishaw XL80 laser interferometer with weather station was used as reference. The retroreflector position complied with the Abbe principle [10] along the y axis only: for practical reasons, a 78 mm Abbe arm resulted along z. Consequently, the effect of the yaw was preliminary measured and compensated. Four machine states T0÷T3 were induced for the validation, by adjusting the scale preloading and by loading a mass onto the machine (Table 2). For each state, the full stroke was scanned twice in either direction in 25 mm steps, resulting in four measurements of which the mean was taken as a result.

T0 was taken as reference state. The scale was then tensioned abnormally (T1), and relaxed back (T2). Finally, the machine was loaded (T3). The $E_{xx}$ measured in T2 was indistinguishable from that in T0 within the experimental limits, as expected. The machine proved very tolerant to the load and the $E_{xx}$ measured in T3 was unexpectedly indistinguishable from that in T2 within the experimental limits. T1 was then the focus. Figure 3 shows the effect of the compensation at T1. The red and blue curves are the measured $E_{xx}$ before and after the biSLIDER compensation. The residual error was within ± 1 µm.

Table 2

<table>
<thead>
<tr>
<th>Machine load</th>
<th>Scale tension</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0 (reference state)</td>
<td>unloaded Normal</td>
</tr>
<tr>
<td>T1</td>
<td>unloaded ~50 ppm</td>
</tr>
<tr>
<td>T2</td>
<td>unloaded Normal, nominally as in T0</td>
</tr>
<tr>
<td>T3</td>
<td>4000 kg As in T2 (normal)</td>
</tr>
</tbody>
</table>

4.2. Validation in lab

A 1000 mm Heidenhain LB382C linear encoder (steel graduated tape) was used (Figure 4). The bislider was realised with two Heidenhain AE LB382C heads separated by a stainless steel plate (Figure 5). The excellent laboratory environment provided thermal stability; in addition, the bislider was equipped with a calibrated Pt100 for thermal compensation. A Renishaw RLU-RLD laser interferometer with refractivity compensation was used as reference. A Heidenhain ND1203 counter was used to treat the head signals. Its original 1 µm resolution was improved to 0.1 µm by an in-house phasemeter reading the 1 Vpp head signals. In this set up, the bislider was kept still and the scale moved, to achieve a full Abbe configuration. Neither the laser nor the scale were reset during each measurement.

Figure 3. Effect of the biSLIDER compensation (T1): linear strain.

Figure 4. Measurement setup. Triangles: supports; rectangular frames: straining devices.

Figure 5. Bislider close-up. (a)(b) Heads; (c) Pt100. Spacing $b = 137$ mm.

Figure 6. Torque-free straining device: (a) tape handle; (b) graduated tape; (c) head; (d) leverage driving the tape handle; (e) housing; (f) preloaded slidingguides. The red arrow indicates the stressing force.
The actual stroke is \( b \) shorter than the 1 000 mm encoder.

The residual error is small even at the end of the stroke: there is no (or small) error accumulation, even in the presence of abnormal perturbations and even if the integral of the uncompensated \( E_{xx} \) curve is not null.

The residual error in the presence of highly non-linear strain is comparable with that obtained with merely linear strain.

The dynamic correction table is in fact a look up table with spatial resolution equal to \( b \). The biSLIDER is unable to detect any perturbation occurring inside a same leg. This is the case about the two perturbation points, where the change in slope is sudden: the compensation does not help.

5. Discussion and conclusions

The biSLIDER was validated in two independent set ups, with uniform and highly nonuniform strains. In either case, it proved suitable for compensating linear encoders, with substantial error compressions. The biSLIDER is all based on commercially available cheap components, and can be used to retrofit existing machines. The one specific component is the bislider, which is a simple spacer.

When designing a biSLIDER application, the choice of the spacing value, \( b \), is important and based on a trade-off. On one hand, \( b \) enters the uncertainty equations in the denominator (Table 1). On the other hand, \( b \) sets the sampling rate of \( E_{xx} \) and then the cut-off length for the sensed error wavelengths: high frequency errors with \( \lambda \leq 2b \) are not detected. In addition, \( b \) reduces the machine stroke and sizes the bislider, which must fit the set up.

The biSLIDER assumes an invariance point at the scale home. Where a set up results in, or an application requires, a different invariant point, additional precautions should be taken. An easy option for encoders with multiple homes is to choose the closest to the intended invariant point.

A possible biSLIDER improvement (not tested yet) is based on the homing. As described in § 2, the actual value \( b \) results from the H2 reading when H1 is homed. This applies at the reference and at the recovery states, resulting in possibly slightly different values: their difference detects a possible bislider expansion. The requirement on the bislider invariance can be much relaxed or even removed by exploiting this information. This can be done by subtracting the actual values \( b_{cor} \) and \( b_{rec} \) from the reference state and the recovery tables, respectively, or equivalently by using the H1 home signal to reset both counts.
Acknowledgments

The research presented in this paper was part of the project SOMMACT [11]. The authors thank Mr. Emanuele Audrito for his valuable contribution to the project.

References