Redundancy-enabled stabilisation of linear encoder performance: The biSLIDER

Original
Redundancy-enabled stabilisation of linear encoder performance: The biSLIDER / Balsamo, Alessandro; Francese, Claudio; Ottone, Renato; Piccato, Aline. - In: CIRP ANNALS. - ISSN 0007-8506. - 67:1(2018), pp. 531-534. [10.1016/j.cirp.2018.04.004]

Publisher:
Elsevier

Published
DOI:10.1016/j.cirp.2018.04.004

Redundancy-enabled stabilisation of linear encoder performance: the biSLIDER

Alessandro Balsamo, Claudio Francese, Renato Ottone, Aline Piccato

Linear encoders are widely used in industry particularly for machine tools. Their performance may suffer thermal and mechanical instability. This paper presents a technique to stabilise the performance over time by recovering to a reference state. It is based on the simultaneous readings of two heads separated by an invariant spacer. It requires off-the-shelf components only and is widely applicable in industry. Experimental results in the field and in lab show excellent error recovery even in the presence of highly nonlinear errors.

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 encodes 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:

$$ {}_{\mathrm{dyn}}a_k = \sum_{i=1}^{k} \left({}_{\mathrm{ref}}a_i - {}_{\mathrm{rec}}a_i \right) = {}_{\mathrm{ref}}a_k - {}_{\mathrm{rec}}a_k + {}_{\mathrm{dyn}}a_{k-1} \qquad (1) $$

where ${}_{\mathrm{ref}}a_i, {}_{\mathrm{rec}}a_i, {}_{\mathrm{dyn}}a_i$ 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.

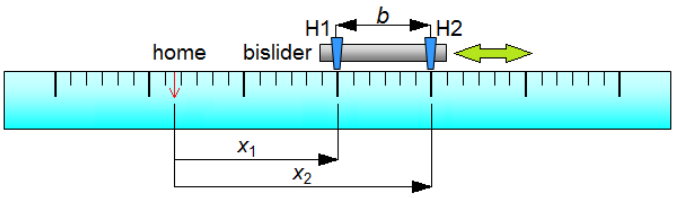

Figure 1. Schematic representation of the biSLIDER concept.

3. Error sources and uncertainties

The entries ${}_{\mathrm{dyn}}a_k$ in eq. (1) are a cumulative sum of elementary compensations at the preceding individual legs, ${}_{\mathrm{ref}}a_i - {}_{\mathrm{rec}}a_i, i \leq k$. The overall error will also be a sum of elementary errors incurred in individual legs. The uncertainty sources can be divided in 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), $u_i = u, \forall i$, then the combined uncertainty for each source is $u_k = \sqrt{k}u$ and $u_k = ku$ 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 $u_k = \left(u/\sqrt{b}\right)\sqrt{x_k}$ and $u_k = (u/b)x_k$ (Table 1). For each source, the combined uncertainty is proportional either to the abscissa 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 1
Uncertainty components of the biSLIDER compensation

Uncorrelated	$\boldsymbol{u_k = \left(u/\sqrt{b}\right)\sqrt{x_k}}$
Calibration	Negligible
Resolution	$u = \varepsilon/\left(4\sqrt{3}\right)$
Fully correlated	$\boldsymbol{u_k = (u/b)x_k}$
Homing	$u = \varepsilon/\left(4\sqrt{3}\right)$
Thermal expansion	$u = bu_\gamma$

ε: Resolution of the encoder; b: head separation (bislider length); u_γ: 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 ${}_{\mathrm{ref}}a_i$ and ${}_{\mathrm{rec}}a_i$ 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 ${}_{\mathrm{ref}}a_i$ and ${}_{\mathrm{rec}}a_i$ suffer finite resolution of each head; ${}_{\mathrm{dyn}}a_i$ is the difference of the two and then involves four readings. As resolution errors are uncorrelated, the uncertainty is $u = \varepsilon/\sqrt{4 \times 12} = \varepsilon/\left(4\sqrt{3}\right)$ ([8] § F2.2.1), where ε 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. This error is set once in all at the reference state and recovery procedures, and is *exactly* the same at any leg. As above, the uncertainty is $u = \varepsilon/\left(4\sqrt{3}\right)$, but here this component is fully correlated, resulting in an accumulated relative uncertainty $u_k/x_k = \varepsilon/\left(4\sqrt{3}b\right)$.

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 ${}_{\mathrm{dyn}}a_i$ is affected by an error $e = \gamma b$ proportional to the bislider spacing b. The coefficient γ 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, γ is usually almost equal for all legs, and the error accumulates, thus projecting the proportional error γ of the bislider onto the scale. The uncertainty components of each leg are $u = bu_\gamma$, where u_γ 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_{xx} denotes the encoder's error of indication, by analogy with the linear positioning error motion of a machine carriage [9].

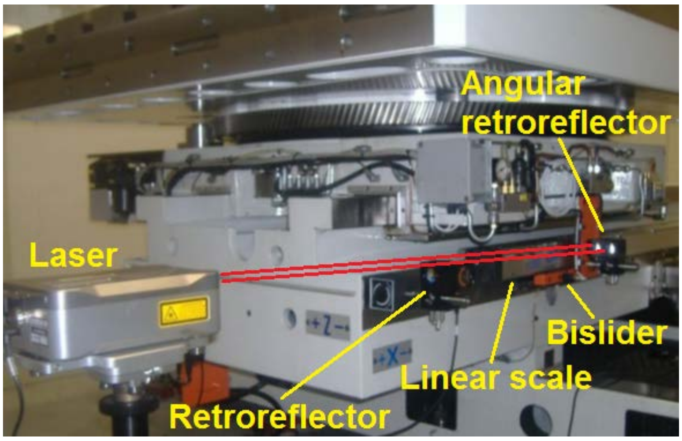

Figure 2. Set up in the field. The yaw is being compensated. The (linear) retroreflector is parked, awaiting the E_{xx} measurement.

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
Validation in the field - perturbation states

	Machine load	**Scale tension**
T0 (reference state)	unloaded	Normal
T1	unloaded	~ 50 ppm
T2	unloaded	Normal, nominally as in T0
T3	4 000 kg	As in T2 (normal)

4.2. Validation in lab

A 1 000 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 V_{pp} 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

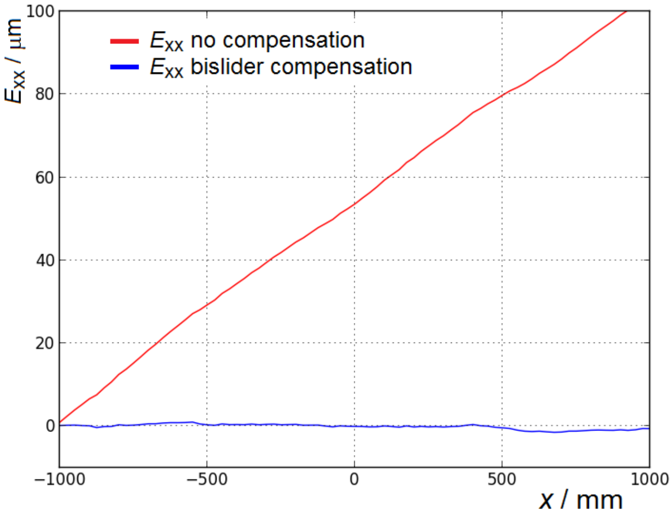

Figure 3. Effect of the biSLIDER compensation (T1): linear strain.

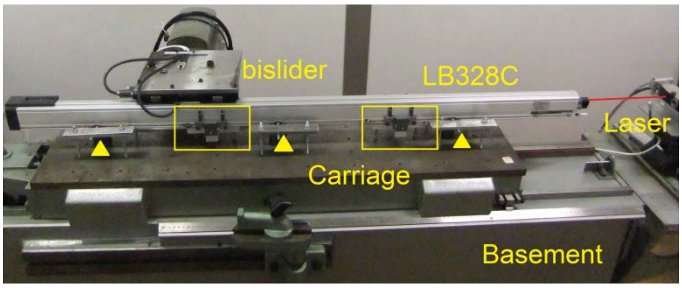

Figure 4. Measurement setup. Triangles: supports; rectangular frames: straining devices.

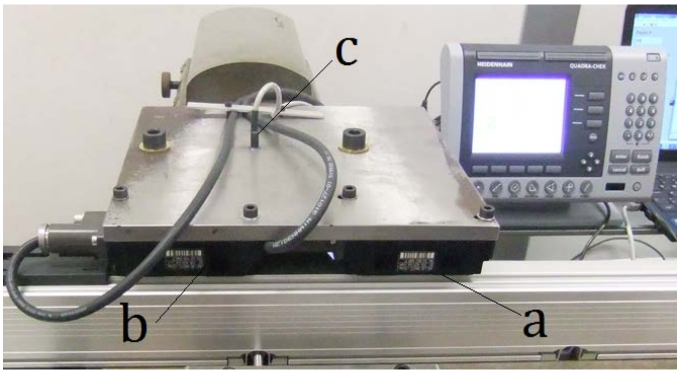

Figure 5. Bislider close-up. (a)(b) Heads; (c) Pt100. Spacing b = 137 mm.

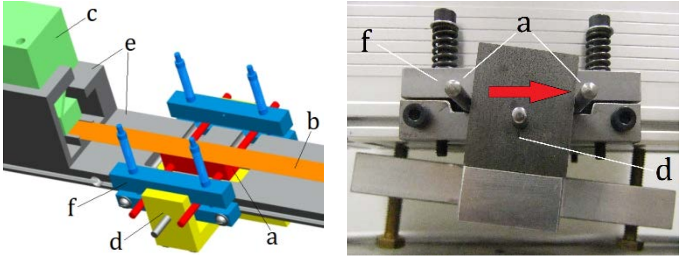

Figure 6. Torque-free straining device: (a) tape handle; (b) graduated tape; (c) head; (d) leverage driving the tape handle; (e) housing; (f) preloaded sliding guides. The red arrow indicates the stressing force.

measurement session. The tension of the scale was preliminary adjusted to minimise E_{xx}.

The same test as on the machine tool was repeated. The scale was tensioned in steps up to 50 ppm. The results were similar to those of the validation in the field. The residual errors in the range (0 - 3) μm were due to the unwanted permanent perturbations in preparation for the subsequent test (see below).

As a further step, we wanted to challenge the biSLIDER with a nonuniform perturbation. Unfortunately, a nonlinear strain is not trivial to induce: the graduated tape is hidden in the housing, and any off-axis stressing force applied at intermediate scale points, results in torque bending the graduated tape. This may disturb the signals to the point of a possible count loss.

Two dedicated straining mechanism were then designed, manufactured and used (Figure 6). A handle was glued to the back face of the internal graduated tape, with pins constrained to slide on guides and kept in contact by preloaded counterparts. This pin-guide coupling prevented undesired rotations and effectively balanced the torque generated by the off-axis stressing force. The drive was provided by a tilting lever in between the pins, slightly slimmer than the pin opening. The lever was simply actuated by two screws, one for positive and one for negative strains. When the active screw was turned, the lever pushed a pin apart and translated the handle with no rotation, thus dragging the graduated tape to generate the sought strain. The two straining devices were glued at intermediate points of the graduated tape through windows milled in the encoder housing. This way, the scale resulted subdivided in three segments – in between and on either side of the straining devices – whose strain could be individually controlled. The boundaries of this portions were at $x = 270$ mm and $x = 710$ mm. An abnormal nonlinear saw profile of E_{xx} was generated (Figure 7 red line), with values in the range [-25, +8] μm, and sharp slope changes at the handles.

In spite of all cares, the milling and the gluing were not equally successful at the two windows. While the right one resulted smooth and nice, the interior rail of the left one was slightly damaged with obvious perturbation of the tape, resulting in an undesired ripple. In addition, the perturbation induced unexpected friction. The heads are designed to adjust to scale imperfections and misalignments by means of a sophisticated decoupling mount. In the presence of unusual friction, a slight parasitic movement was induced in the measurement direction, resulting in a hysteresis (Figure 8). All this was incidental and unfortunate, but not related to the biSLIDER, rather to its testing set up.

Even in the presence of a highly nonuniform perturbation, the biSLIDER was able to recover the reference state with residual errors within ± 3 μm (apart from the ripple at the left perturbation point). Figure 7 shows that:

- 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

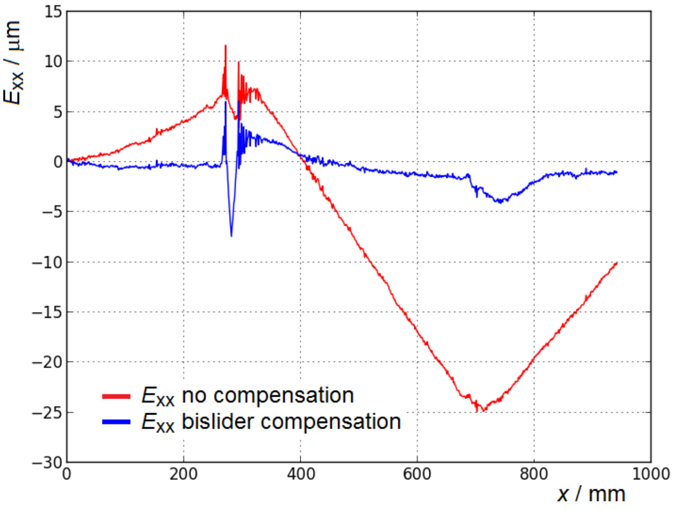

Figure 7. Effect of the biSLIDER compensation: nonlinear strain.

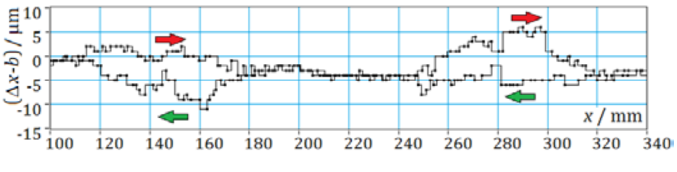

Figure 8. Ongoing (red) and backward (green) values of the difference of the two head readings, Δx. The two opposite hysteretic patterns are separated by $b = 137$ mm, in coincidence with either head at the left perturbation point.

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 and 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_{ref} 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

[1] Kunzmann, H., Pfeifer, T., Flügge, J., 1993, Scales vs. Laser Interferometers – Performance and Comparison of Two Measuring Systems, Annals of the CIRP, 42/2:743-767.
[2] Alejandre, I., Artes, M., 2005, Thermal non-linear behaviour in optical linear encoders, International Journal of Machine Tools & Manufacture, 46:1319-1325.
[3] Balsamo, A., Ottone, R., 2012, Apparatus and method for the stabilization of linear encoders, PCT/IT2012/000259.
[4] Sakagami, S., Teraguchi, M., 2007, Absolute linear encoder, Patents JP2006259621, EP1770374, US2007069117.
[5] Günter, N., Thermisch kompensiertes Meßsystem, Patent DE19919042.
[6] Schuchardt, G., Freitag H.-J., Tzschach, F., Thiemer, R., Voigt, L., 1988, Device and method for measuring lengths, Patent GB2156989.
[7] Li, X., Wang, H., Ni, K., Zhou, Q., Mao, X., Zeng, L., Wang, X., Xiao, X., 2016, Two-probe optical encoder for absolute positioning of precision stages by using an improved scale grating, Optics Express, 24/19:21378-21391.
[8] JCGM 100:2008, Evaluation of measurement data — Guide to the expression of uncertainty in measurement.
[9] ISO 230-1:2012, Test code for machine tools – Part 1: Geometric accuracy of machines operating under no-load or quasi-static conditions.
[10] Abbe, E., 1890, Measuring instruments for physicists, Journal for instrumental information, 10:446-448.
[11] SOMMACT – Self-Optimising Measuring Machine Tools, 2009-2012, FP7-NMP project 229112, http://cordis.europa.eu/project/rcn/92739_en.html.