Measurement 42 (2009) 1118–1121

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

The application of the bivariate distribution and mutual information in measurement

M. Rybokas [a], K.T.V. Grattan [b], V. Giniotis [a,*]

[a] Vilnius Gediminas Technical University, Sauletekio al. 11, 10223 Vilnius-40, Lithuania

[b] School of Engineering & Mathematical Sciences Electrical, Electronic & Information Eng, City University, Northampton Square, London EC1V 0HB, UK

ARTICLE INFO

Article history:
Received 23 March 2007
Received in revised form 5 April 2009
Accepted 16 April 2009
Available online 3 May 2009

Keywords:
Measurement
Information entropy
3D measurements

ABSTRACT

The measuring systems, such as those used in coordinate measuring machines (CMMs), laser interferometers, linear or rotary encoders, etc., feature of huge amount of information indicating the position of the object under control. This information is subject for verification or metrological calibration during some periods in service. On the other hand, there are no means for verifying every digit of output information, and the great quantity of information consisting of millions of values is left with its errors undetermined. Expression of the result of measurement (including the calibration) of a measuring system supplementing it by the parameter of information entropy is proposed in the paper. The uncertainty expression in the result of measurement in the plane and in the volume is presented here with the parameter of information entropy that shows the portion of data assessed.

1. Mutual information of measurement data

The established standards [1,2] contain the main rules for expressing the uncertainty of a measurement result. The main basis for this procedure is probability theory giving a statistical means for expression of the result of a measurement, the measurand or estimate of the value of the specific quantity subject to measurement, beginning from choosing a sample for assessment, creating a set of trials, calculation of the measurand value estimator, and the evaluation of the expanded uncertainty on the basis of the desired level of confidence (type "A" method used to estimate the numerical values [1,2]). This approach is widely used in all kinds and branches of metrology. Nevertheless, there is no reference, which part of total information from the information measuring system is assessed during this procedure. It seems, there is a possibility for the further development of more complex and more informative uncertainty estimate for multidimensional measuring systems for normal or uniform probability distribution of random values. The information entropy [3,4] parameter enables a presentation to the user of an estimate of the measurand giving additional information about the sample value on which the estimate is based. That is, for the measurand estimate of rotary encoder's output information of 1,296,000 (1 s of arc) using a 12-sided polygon as a reference measure or the standard of angle, it is quite important to know that only 12 from 1,296,000 points of information from the encoder's output are assessed. Sampling strategy analysis is also important item for measurements based on the decision and possibilities to select an appropriate number of points to be measured and the strategy how these points must be placed in the line, area or volume of trials [2,6,7].

A new approach to evaluate the measurement data was presented by the authors [5] including introducing an information entropy parameter into the expression of the result of a measurement. This evaluation is valid for use in one-dimensional or one parameter measuring system analysis. A further development of this approach is presented here by applying it to multi-coordinate measuring systems and using the relative entropy and mutual entropy

* Corresponding author. Fax: +370 5 2744 705.
E-mail addresses: vg@ai.vtu.lt, vg@ai.vgtu.lt (V. Giniotis).

0263-2241/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.measurement.2009.04.007

evaluations for this purpose. Additional information on the sampling value in the result of a measurement would expand the knowledge of the extent of measurement evaluation for the object. It also would help in data comparison, the traceability of measurements and in achieving more information measurements performed on the object.

The entropy is the uncertainty of a single random variable [3]. The reduction ***I*** in the uncertainty due to the information assessed (for example, the information received after the calibration of the information-measuring system of the machine) is $I = H_0 - H_1$, where H_0 is the entropy before receiving the information, and H_1 is the entropy after receiving it.

The real technically available means of calibration permit to assess information about the linear or angular displacement of the object only at restricted intervals of information measurement system. Let the total amount of the information be m number of signals, digital information or points indicating the position of the object. The quantity of statistically determined information during the system's calibration process is b (a sample, depending of available technical means for calibration) and measurement of each variable (point) is performed c times. Then the total number of measurements performed is $n = bc$. Information assessment will be $I = H_0 - H_1 = \log_a m - \log_a b$. After some transformations the expression for the measurement result of the variable x yields to [5]:

$$X = \overline{X} \pm \frac{t \cdot s}{\sqrt{a^{-I} mc}}\bigg|_{p, I(H_0, H_1)}; \tag{1}$$

i.e., a measurand with a level of confidence p and information on the system before and after the calibration equal to $I(H_0, H_1)$.

Here t – t-distribution for v degrees of freedom; $v = n - 1$; s – an estimate of standard deviation; a is a basis of logarithm for calculations of information entropy H.

This means that the measurement result is determined with the expanded uncertainty evaluated with a level of confidence p and with the indeterminacy of the result assessed by the entropy $I(H_1, H_0)$ having evaluated a portion (a random sample, chosen according to the available points for calibration) of all the data in question. Then the general process of measurement with the increase of information processed can be portrayed as shown in Fig. 1. A material, object or specimen Q_j having the parameters (some number of attributes, i.e., geometric dimensions, mass, volume, etc.) $Q_j = f_j(n_k, v_l, \varsigma_p, \ldots)$ with its relevant values $n_k, v_l, \varsigma_p, \ldots$ is to be measured using the standard of measure $Q_i = f_i(n_k, v_l, \varsigma_p, \ldots)$ with the same features that are to be measured. As a first step, the sampling procedure $X_1, \ldots, X_{n,m}$ is undertaken. Here X is a variable, n is the number of samples taken from the specimen and the standard,m is the number of trials to give the sample estimation. In case of measurements during the calibration process, size of the sample is predetermined by the number of the standard of measure (etalon, reference standard) in use, for example, 12, 24 or 36-sided multiangle prism (polygon).

The uncertainty constituents may be determined as $u_i = s_{i,j} t_{i,j} / \sqrt{n,m}$, where u is a measure of uncertainty, named expanded uncertainty and obtained by multiplying $u_c(x)$ by a coverage factor chosen. $s_{i,j}$ is the estimate standard deviation, $t_{i,j}$ is the Student coefficient, n is the number of attributes i, and m is the number of attributes j. As a result, the information about the attributes related to the sampling can be expressed as$I_{i,j}(H_{0i,0j}, H_{1i,1j})$, where $I_{i,j}$ is the information entropy of attributes i and j; $H_{1i,1j}$ is the information before and $H_{0i,0j}$ after the measurement. The more complicated case of the assessment of the result is during multivariable or, in technical terms, multi-coordinate measurements, where information assessment is essential.

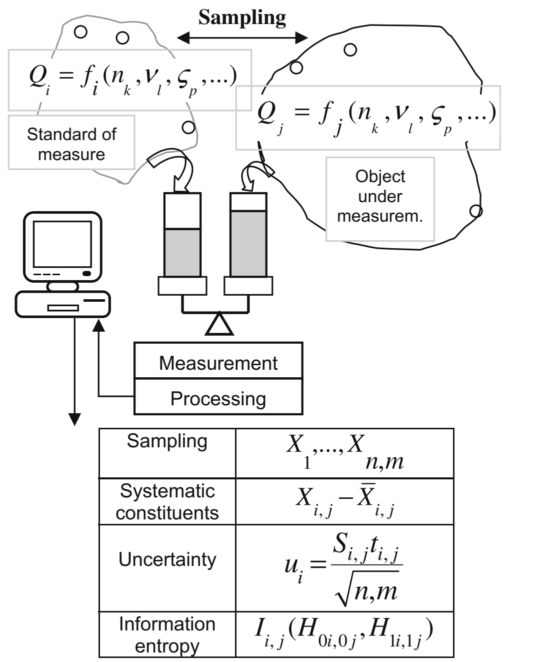

Fig. 1. General diagram of the measurement process and the presentation of the result.

Mutual information is a measure of the amount of information that one random variable contains about another random variable [3]. In other words, it is the reduction in the uncertainty of one random variable because of the existence of knowledge about the other. The relative entropy of independent variables X and Y in terms of their probabilities is [4]:

$$H(X/Y) = -\sum_i \sum_j p(X_i, Y_j) \log p(X_i/Y_j) \tag{2}$$

Both the relative entropy and the mutual entropy show a specific evaluation of the random variables. The relative entropy helps to assume a distribution of the random value (as q) when the true distribution of this variable is p. This parameter is assumed as a measure of the distance between the two distributions of one variable. So, assuming the explanations presented above, the mutual information appears as the most acceptable method to assess the information in measurement or in calibration of the multi-coordinate measuring systems [6,7].

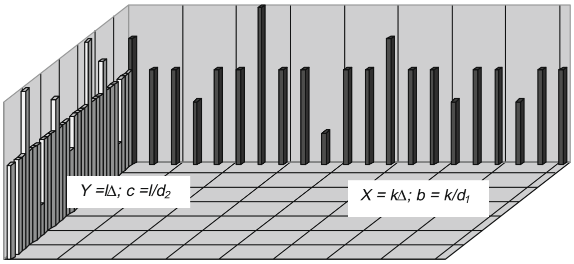

Fig. 2. Determined and undetermined data series in the 3D space.

2. Mutual information assessment for 3D measurements

The mutual information model of 3D measurement is investigated below. The x, y and z axes are subdivided into k, l and m steps (divisions), respectively (see Fig. 2 for 2D case), and Δ is the length of a single step of the scale [8,9] or the interval between the signal from the transducer of the information-measuring system. The histogram demonstrates the calibrated values among all the available values, where the latter are assumed to be just the average of the total values. The symbols b and c indicate the number of calibrated steps for axes x and y, respectively, as a result of performing a calibration at intervals of d_1 and d_2. Therefore, the intervals of the measurement values extend to

$$\begin{aligned} 0 &\leqslant x \leqslant k,\\ 0 &\leqslant y \leqslant l \text{ and}\\ 0 &\leqslant z \leqslant m. \end{aligned} \tag{3}$$

If all dimensions are independent of each other, no information is gained about any of the variables by fixing the value in one dimension. For the sake of economising on space only the example of fixing dimension z is provided.

$$\begin{aligned} I(X;Y|Z) &= \sum p(x,y,z)\log\frac{p(X,Y|Z)}{p(X|Z)p(Y|Z)}\\ &= \sum\frac{1}{klm}\log\frac{1/kl}{1/k\cdot 1/l} = \sum\frac{1}{klm}\log 1 = 0. \end{aligned} \tag{4}$$

If, on the other hand, taking one-dimension as given influences the measurement in the other two-dimensions in some generalized way, it is possible that the process of fixing the value in one dimension would result in some amount of mutual information between the measurements of the remaining two dimensions. The equations below demonstrate the expressions for the mutual information when fixing a measurement value in one dimension changes the pitch of the calibration in the other two to k_1 and k_2, l_1 and l_2, m_1 and m_2 for the x-axis, y-axis and z-axis, respectively.

$$I(X;Y|Z) = \sum\frac{1}{klm}\log\frac{k_2 l_2}{k_1 l_1} \tag{5}$$

$$I(Y;Z|X) = \sum\frac{1}{klm}\log\frac{l_2 m_2}{l_1 m_1} \tag{6}$$

$$I(X;Z|Y) = \sum\frac{1}{klm}\log\frac{k_2 m_2}{k_1 m_1} \tag{7}$$

If the pitch of the calibration is the same for the two dimensions of the plane, the following variants of the equations result for the respective equal pairs of the coordinates.

The three variants of the equations are presented below:

(1) when the pitch of the calibration (determined with appropriate uncertainty) is equal for the axes of the xy-plane or $k = l,\ l \neq m,\ m \neq k$

$$I(X;Y|Z) = \sum\frac{1}{2lm}\log\frac{l_2}{l_1} \tag{8}$$

(2) when the pitch of calibration is equal for the axes of the yz-plane or $k \neq l,\ l = m,\ m \neq k$

$$I(Y;Z|X) = \sum\frac{1}{2kl}\log\frac{l_2}{l_1} \tag{9}$$

(3) when the pitch of calibration is equal for the axes of the xz-plane or $k \neq l,\ l \neq m,\ m = k$

$$I(X;Z|Y) = \sum\frac{1}{2kl}\log\frac{k_2}{k_1} \tag{10}$$

For the continuous version of Eq. (4) it is necessary to combine the expressions for the trivariate normal density and multivariate conditional normal distributions. These expressions are provided and simplified below.

The general expression for a bivariate normal distribution is given by the formula below:

$$p(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{xy}^2}} \times e^{-\frac{1}{2(1-r_{xy}^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r_{xy}(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(x-m_y)^2}{\sigma_y^2}\right]}. \tag{11}$$

The expressions for the bivariate conditional normal distributions to be placed in the denominator of Eq. (4) are provided below.

$$p(y|z) = \frac{1}{\sigma_y\sqrt{1-r_{yz}^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r_{yz}^2)}\left(\frac{y-m_y}{\sigma_y} - r_{yz}\frac{z-m_z}{\sigma_z}\right)^2}, \tag{12}$$

$$p(x|z) = \frac{1}{\sigma_x\sqrt{1-r_{xz}^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r_{xz}^2)}\left(\frac{x-m_x}{\sigma_x} - r_{xz}\frac{z-m_z}{\sigma_z}\right)^2}, \tag{13}$$

where the values of r are the correlation coefficients between the corresponding random variables, while the values of m and σ are means and standard deviations of the random variables indicated by the indices, respectively.

The general expression for a trivariate normal distribution is given by the formula below:

$$p(x,y,z) = \frac{e^{-w/[2(r_{xy}^2 + r_{xz}^2 + r_{yz}^2 - 2r_{xy}r_{xz}r_{yz} - 1)]}}{2\sqrt{2}\pi^{3/2}\sqrt{1-(r_{xy}^2 + r_{xz}^2 + r_{yz}^2) + 2r_{xy}r_{xz}r_{yz}}}, \tag{14}$$

where from the literature [mathworld.worlfram.com]

$$\begin{aligned} w =& x^2(r_{yz}^2 - 1) + y^2(r_{xz}^2 - 1) + z^2(r_{xy}^2 - 1) + 2[xy(r_{xy} - r_{xz}r_{yz})\\ &+ xz(r_{xz} - r_{xy}r_{yz}) + yz(r_{yz} - r_{xy}r_{xz})]. \end{aligned} \tag{15}$$

3. The application of the trivariate conditional distribution

The trivariate conditional distribution is produced by combining Eq. (14) and the formula for the univariate normal distribution as shown here:

$$p(x,y|z) = \frac{p_1(x,y,z)}{p_2(z)} \tag{16}$$

After inserting the corresponding probability densities and simplifying, the above expression becomes:

$$\begin{aligned}
p(x,y|z) &= \frac{\dfrac{e^{-w/\left[2\left(r_{xy}^2+r_{xz}^2+r_{yz}^2-2r_{xy}r_{xz}r_{yz}-1\right)\right]}}{2\sqrt{2}\pi^{3/2}\sqrt{1-\left(r_{xy}^2+r_{xz}^2+r_{yz}^2\right)+2r_{xy}r_{xz}r_{yz}}}}{\dfrac{e^{-\frac{(z-m_z)^2}{2\sigma_z^2}}}{\sigma_z\sqrt{2\pi}}} \\
&= \frac{e^{-w/\left[2\left(r_{xy}^2+r_{xz}^2+r_{yz}^2-2r_{xy}r_{xz}r_{yz}-1\right)\right]}}{2\sqrt{2}\pi^{3/2}\sqrt{1-(r_{xy}^2+r_{xz}^2+r_{yz}^2)+2r_{xy}r_{xz}r_{yz}}} \\
&\quad \times \frac{\sigma_z\sqrt{2\pi}}{e^{-\frac{(z-m_z)^2}{2\sigma_z^2}}} \\
&= \frac{\sigma_z e^{-w/\left[2\left(r_{xy}^2+r_{xz}^2+r_{yz}^2-2r_{xy}r_{xz}r_{yz}-1\right)\right]}}{e^{-\frac{(z-m_z)^2}{2\sigma_z^2}}2\pi\sqrt{1-(r_{xy}^2+r_{xz}^2+r_{yz}^2)+2r_{xy}r_{xz}r_{yz}}}
\end{aligned} \tag{17}$$

And so the full expression under the logarithm in the continuous variant of Eq. (4) becomes:

$$\begin{aligned}
\frac{p(x,y|z)}{p(x|z)p(y|z)} &= \frac{\sigma_z e^{-w/\left[2\left(r_{xy}^2+r_{xz}^2+r_{yz}^2-2r_{xy}r_{xz}r_{yz}-1\right)\right]}}{e^{-\frac{(z-m_z)^2}{2\sigma_z^2}}2\pi\sqrt{1-(r_{xy}^2+r_{xz}^2+r_{yz}^2)+2r_{xy}r_{xz}r_{yz}}} \\
&\quad \times \frac{2\pi\sigma_x\sigma_y\sqrt{1-r_{xz}}\sqrt{1-r_{yz}}}{e^{-\left[\frac{1}{2\left(1-r_{yz}^2\right)}\left(\frac{y-m_y}{\sigma_y}-r_{yz}\frac{z-m_z}{\sigma_z}\right)^2+\frac{1}{2\left(1-r_{xz}^2\right)}\left(\frac{x-m_x}{\sigma_x}-r_{xz}\frac{z-m_z}{\sigma_z}\right)^2\right]}} \\
&= \frac{1}{e^{-\left[\frac{1}{2\left(1-r_{yz}^2\right)}\left(\frac{y-m_y}{\sigma_y}-r_{yz}\frac{z-m_z}{\sigma_z}\right)^2+\frac{1}{2\left(1-r_{xz}^2\right)}\left(\frac{x-m_x}{\sigma_x}-r_{xz}\frac{z-m_z}{\sigma_z}\right)^2+\frac{(z-m_z)^2}{2\sigma_z^2}\right]}} \\
&\quad \times \frac{\sigma_x\sigma_y\sigma_z\sqrt{1-r_{xz}}\sqrt{1-r_{yz}}e^{-w/\left[2\left(r_{xy}^2+r_{xz}^2+r_{yz}^2-2r_{xy}r_{xz}r_{yz}-1\right)\right]}}{\sqrt{1-(r_{xy}^2+r_{xz}^2+r_{yz}^2)+2r_{xy}r_{xz}r_{yz}}}.
\end{aligned} \tag{18}$$

The equations derived permit an assessment of the results of a calibrated multi-coordinate information measuring system (for example, a coordinate measuring machine) more completely. They show which part from a multi-coordinate system is assessed and how the results can be taken as reliable from all signals (values) in the volume of a measuring system of the machine or robot, that extend to millions in modern automated equipment. This new approach to the assessment gives more information about the measuring process and the accuracy of these measurements.

4. Conclusions

As a result of the analysis carried out above, several important points and general conclusions should be stressed. The new approach given to the evaluation of measurement data gives full information on the measurement process performed and the quantity of data assessed during this process. Further, the bivariate and trivariate cases of conditional entropy are developed and presented, permitting its application to an assessment of the conditions in multi-coordinate measuring systems.

Acknowledgement

This work has been funded by Lithuanian State Science and Studies Foundation, Project No. B-32/2008.

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, OIML. Guide to the expressions of uncertainty in measurement. ISO Publishing, 1995.

[2] Sampling procedures for inspection by attributes. Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection (ISO 2859-1:1999+ISO 2859-1:1999/Cor.1:2001).

[3] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.

[4] R.B. Ash, Information Theory, Dover Publications, New York, 1990.

[5] V. Giniotis, K.T.V. Grattan, M. Rybokas, R. Kulvietiene, Uncertainty and indeterminacy of measurement data, Measurement 36 (2004) 195–202.

[6] F.M.M. Chan, T.G. King, K.J. Stout, The influence of sampling strategy on a circular feature in coordinate measurements, Measurement 19 (2) (1996) 73–81.

[7] G. Lee, J. Mou, Y. Shen, Sampling strategy design for dimensional measurement of geometric features using coordinate measuring machines, Int. J. Mach. Tools Manufact. 37 (7) (1997) 917–934.

[8] V. Giniotis, K.T.V. Grattan, Optical method for the calibration of raster scales, Measurement 32 (2002) 23–29.

[9] V. Giniotis, K.T.V. Grattan, M. Rybokas, R. Kulvietienė, Uncertainty analysis using information entropy, In: VDI – Berichte Nr. 1860, VDI Verlag GmbH, Dusseldorf, 2004, pp. 505–510.