Modelling performing calculations over the data presented in a probabilistic form

Nikolay Sapozhnikov[1], *Aleksey* Bryukhovetskiy[2], *Aleksandr* Polyakov[1], and *Dmitriy* Moiseev[2,*]

[1] Black Sea Higher Navy Order of the Red Star Academy by P.S.Nakhimov, Federal State Owned Military Educational Institution of the Higher Professional Education, Russian Federation, 299028, Sevastopol
[2] Federal State Autonomous Educational Institution of Higher Education «Sevastopol State University», Russian Federation, 299053, Sevastopol

Abstract. The article presents the results of comparison of different forms of representation and transformation of information in the form of non-positional probabilistic display. Assessment of the hardware of the compute nodes and the analysis of the relationship between accuracy of representing the information and their performance have been carried out.

1 Introduction

Currently, the period of computer technology development is characterized by intensive search for fundamentally new methods of information processing and storage, construction of universal and specialized computer architectures and systems based on them with the involvement of modern technologies, among which digital signal processing (DSP) is one of the most popular [1].

Thanks to modern achievements in the field of nano- and microelectronics in the latest information systems, the DSP is known to be successfully applied for filtering and coding speech and audio signals, processing images and measuring information, spectral analysis of digital audio and video recording in radio engineering systems and telecommunication systems, control and robotics, information protection in such areas as communications, multimedia, telephony and television, radar and radio navigation, hydro acoustics, medicine, etc.[2].

The DSP used in recent decades is characterized by numerous calculations on data clusters of high bit-rates carried out in real time.

Striving to improve energy efficiency and universality of DSP devices leads to complication of the computational algorithms, an escalation of the problem of hardware cost, operating speed and accuracy, and increase of requirements to fault tolerance of the devices and interference immunity of communication channels during data transmission [3].

Thus, research in the field of one of the directions of development of parallel computing technologies connected with development of methods, algorithms and devices of computer facilities for digital processing of signals with application of non-positional numeration

---

* Corresponding author: dmitriymoiseev@mail.ru

systems, the most perspective of which are probabilistic and pseudo-probabilistic forms of information presentation, acquires special relevance [3–7].

This work is devoted to the comparison of different forms of representation and transformation of information into probabilistic display.

Advantages of applying non-positional probabilistic representation of information. With the development of digital computer technology and simulation modelling methods, the method of statistical tests got into wide use, the main idea of which is the relationship between the probabilistic characteristics of random processes and the values which are the solutions to problems of mathematical analysis [4].

As the main advantage of the use of non-positional probabilistic representation of information in the implementation of digital signal processing, the following significant reduction in the hardware capacity of computing devices can be highlighted [5]:

– when performing addition of two operands, the hardware capacity of the probabilistic adder will be 13 basic logical elements, while the hardware capacity of the parallel 16-bit combinational adder will be 181 elements, which at the level of the logical element exceeds the proposed solution by 14 times;

– when multiplying two double-byte binary numbers, a multiplier device is required, the number of Boolean basis elements in which will be about 700, while the probabilistic multiplier device will be implemented on three AND gates;

– the squaring scheme in the probabilistic form is realized on two elements, and with the increase of the degree the number of elements will be $2 \cdot n$, where n is the degree, which gives an advantage in comparison to the digital device by about 300 times;

– the hardware capacity of the proposed probability divider is approximately 6 times smaller than the digital one.

The advantages discussed above remain valid when using parallel processing of the information presented in the form of non-positional probabilistic display [6].

Thus, we can conclude that development of universal and specialized devices of processing information represented in non-positional probabilistic displays is especially well suited for application in measuring systems and data processing systems in which the important parameter is the hardware capacity of the device.

Representation of information in the form of non-positional probabilistic display. In general, the essence of stochastic or probabilistic conversion is that any value of the converted parameter can be adjusted to a certain probability [3].

Depending on the rule according to which it occurs, the conversion methods are divided into single-line unipolar, single-line bipolar, two-line bipolar, and pseudo-probabilistic representations [4].

In the case of a single-line unipolar representation of information, the value of the parameter of the converted value is either always positive or always negative, and the conversion process itself is performed in accordance with the rule [3]

$$y_{ij} = \begin{cases} 1 \text{ when } x_i > R(t_{ij}) \\ 0 \text{ when } x_i \leq R(t_{ij}) \end{cases}, \tag{1}$$

where $x_i - i$-th value of the parameter of the converted signal $X(t)$, $R(t_{ij})$ – j-th value of the parameter of the auxiliary random signal $R(t)$, changed in the change interval, $i = \overline{1, N}$ – number of signal conversion cycles $X(t)$, $j = \overline{1, K}$ – number of statistical tests of every value x_i within the time interval $\Delta t_i = t_{i+1} - t_i$, y_{ij} – value of the probabilistic display of the signal x_i parameter out of the consequence

$$Y_i(t) = \{ y_{i1}; y_{i2}; \ldots y_{ij}; \ldots y_{iK} \} . \tag{2}$$

As shown in [3], probabilistic display has the properties of synchronicity and independence of each display element from any other.

When considering the change range $X(t)$ equal one, the maximum value $x_i = 1$, will be represented in the probabilistic display by values $y_{ij} = 1$ with the probability equal one in every step, and the minimum value $x_i = 0$ by the same values $y_{ij} = 0$ with zero probability.

Considering (1), probabilities of appearance of «1» and «0» in the probabilistic display are

$$\begin{aligned} P(y_{ij} = 1) &= P[R(t_{ij}) < x_i] \\ P(y_{ij} = 0) &= 1 - P[R(t_{ij}) < x_i] \end{aligned} , \tag{3}$$

Mathematical expectancy (ME) of probabilistic display is determined through a series of distributions for a discrete random variable y_{ij} the following way:

$$M[Y_i(t)] = \sum_{l=1}^{2} y_{ijl} P_l = P(y_{i,j} = 1) = P[R(t) < x_i = r)] = F_{x_i}(R) \tag{4}$$

It follows from (4), that the probability of occurrence of "1" "in the probabilistic display" is ME display, which, in turn, is numerically equal to the value of the integral law of distribution $F_{x_i}(R)$ of the auxiliary signal $R(t)$ at the level of comparison x_i.

In the case of subordination of the auxiliary random signal $R(t)$ to the uniform distribution law, in accordance with [3]

$$F(R) = \begin{cases} 0 \text{ when } r < 0 \\ r \text{ when } 0 \leq r \leq 1 \\ 1 \text{ when } r > 1 \end{cases} . \tag{5}$$

In this case (5) ME of the probabilistic will have the form of:

$$M[Y_i(t)] = P(y_{i,j} = 1) = x_i . \tag{6}$$

It can be seen from the expression (6) that the most important factor is that the value of the original parameter x_i is recoverable from its probabilistic display y_i, which in its turn makes possible a reversal transformation of "probability-value". Knowledge of the law of the auxiliary random signal $R(t)$ distribution and calculation of ME probabilistic display allows to determine the value x_i^*, which is the assessment x_i, by means of functional transformation [5].

The assessment satisfying the requirements of unbiasedness, consistency and efficiency, in accordance with Chebyshev's theorem, has the form [5]

$$x_i^* = \left(M[Y_i(t)] \right)^* = \frac{1}{K} \sum_{j=1}^{K} y_{ij} . \tag{7}$$

From (7) it follows that in order to obtain the original value, it is necessary to calculate the number of units in its probabilistic representation and refer to the number of statistical tests.

For the transition to a two-line bipolar representation, it is enough to multiply logically the values y_{ij} of the probability display obtained in accordance with rule (1) by a variable Z taking only two values: +1 or -1 for the distribution on "positive" or "negative" line.

In this case the expression (2) for a two-line bipolar representation will take the form of:

$$Y_i(t)z_i = \{y_{i1}z_i; y_{i2}z_i; ...y_{ij}z_i; ...y_{iK}z_i\} \tag{8}$$

In the case of the hardware implementation of a two-line bipolar representation, we have two negative lines, on which the values y_{ij} will be submitted with $z_i = -1$ and positive, on which the values y_{ij} will be submitted with $z_i = +1$. The maximum positive value will be introduced in the last expression with the single probability by the values $y_{ij}z_i = 1$ on the positive line and the values $y_{ij}z_i = 0$ on the negative one; the maximum negative value – vice versa.

As a consequence, the expression for the ME of the probabilistic display for a two-line bipolar representation remains valid, and both probabilistic display properties considered earlier are valid.

The third form of probabilistic representation under study is one-line bipolar representation. The value of the signal parameter $X(t)$ for it varying in the interval [-1; +1], i.e. from $x_{min} = -1$ to $x_{max} = +1$ is implemented by hardware as well as for single-line unipolar probabilistic representation on one line. Herewith the expression (4) for ME is:

$$M[Y_i(t)] = 0.5(1 + x_i) . \tag{9}$$

In this case, the number "+1" will be represented by a series of impulses with a magnitude equal to one in each cycle, i.e. P=1, the number "0" by a random series consisting of ones and zeros, with an equal probability of occurrence P = 0,5, and the number "-1" by a series of impulses with a probability P = 0.

As can be seen from (1) and (4) the probabilistic display of the probabilistically transformed value of the signal will correspond to the transformed parameter (excluding the methodical error) only in the case of an infinite number of independent statistical tests K [5].

The analysis of errors of probabilistic transformation is carried out on the basis of the formula (9), so let's use the considered method.

Assuming some value x_i within the interval [0:1], $0 \le x_i \le 1$, it is necessary to perform the operation of probabilistic transformation. Then, for a single-line unipolar representation according to (1), we obtain a probabilistic representation in the form of:

$$Y_{x_i}(t) = \left\{ y_{x_i 1}; y_{x_i 2}; ...; y_{x_i j}; ...; y_{x_i K} \right\} . \tag{10}$$

ME of this probabilistic display (10), according to (2) will be defined the following way [3]:

$$M\left[Y_{x_i}(t) \right] = F_{x_i}(R) ,$$

And as assessment to determine ME we assume

$$\left\{ M\left[Y_{x_i}(t) \right] \right\}^* = \frac{1}{K} \sum_{j=1}^{K} y_{x_i} = x_i^*. \tag{11}$$

Let's determine the ME of the assessment x_i^*:

$$M(x_i^*) = M(\frac{1}{K} \sum_{j=1}^{K} y_{ij}) = \frac{1}{K} \sum_{j=1}^{K} M(y_{ij}) = F_{x_i}(R) = x_i. \tag{12}$$

To calculate the error of the probabilistic transformation, let's define dispersion of the assessment according to the formula (13):

$$D(x_i^*) = \frac{1}{K^2} \sum_{j=1}^{K} \sum_{l=1}^{2} [y_{ijl} - M(y_{ijl})]^2 P_l = \frac{[F_{x_i}(R) - F_{x_i}^2(R)]}{K}. \tag{13}$$

Thus, the expression for the root-mean-square deviation (RMSD) according to formula (13) will take the form of:

$$\sigma(x_i^*) = \sqrt{\frac{F_{x_i}(R) - F_{x_i}^{\,2}(R)}{K}}. \tag{14}$$

As can be seen from (14), the value of the error of the probabilistic transformation depends on the type of the auxiliary random signal distribution law $R(t)$, and is inversely proportional to the square root of the number of independent statistical tests K [5].

However, the expression for RMSD (14) cannot be used directly to determine the error, since the assessments of the probabilistic characteristics are random variables. It is necessary to know the law of the assessment distribution to determine the quantitative dependence, which is required to determine the absolute and reduced measurement errors - the value of its difference from the true value of the probabilistic characteristic with a predetermined probability.

It is known that in accordance with the central limit theorem, the distribution law of the considered assessment at K=(20.30) will be close to normal, which makes it possible to determine the absolute error of the probability transformation by the formula

$$\Delta = \sqrt{2}\Phi^{-1}(P)\sigma(x_i^*) = \frac{\sqrt{2}\Phi^{-1}(P)}{\sqrt{K}} \sqrt{F_{x_i}(R) - F_{x_i}^{\,2}(R)}, \tag{15}$$

where Φ^{-1} is the inverse function of the Laplace, P is the probability of that the true value $F_{x_i}(R)$ is within the interval I_p,

$$I_p = \left[F_{x_i}^*(R) - \Delta; F_{x_i}^*(R) + \Delta \right].$$

Rate setting in the single range of the value x_i undergoing probabilistic transformation leads to the calculation of the reduced error by the formula:

$$\gamma = \frac{\Delta}{1} \cdot 100 \quad \%. \tag{16}$$

To reveal quantitative regularities we consider the case of subordination of the auxiliary random signal $R(t)$ to the uniform distribution law.

In this case, expression (15) will take the form of the absolute error:

$$\Delta = \sqrt{2}\Phi^{-1}(P)\sigma(x_i^*) = \frac{\sqrt{2}\Phi^{-1}(P)}{\sqrt{K}}\sqrt{x_i(1-x_i)} . \qquad (17)$$

As we know from [4], the maximum error in the probabilistic transformation occurs when the value of the converted value x_i is in the middle of the dynamic range [0;1], i.e. when $x_i = 0,5$.

In this case, the top rating will be:

$$\Delta \le 0,7\Phi^{-1}(P)\sqrt{K} .$$

In case of using the confidence interval P, equal to $P = 0,9973$, we have $0,7\Phi^{-1}(P) = 1,645$, hence, the absolute error of the probability transformation is:

$$\Delta_{max} \le \frac{1,645}{\sqrt{K}} .$$

Then the reduced error of the probabilistic transformation will be equal to:

$$\gamma \le \frac{1,645}{\sqrt{K}} \cdot 100\% . \qquad (18)$$

Table 1 gives the calculated values γ showing the dependence of the reduced error on the number of statistical tests K for different values of the parameter X.

Table 2. Values of the reduced error at probabilistic transformation

Number of tests K	Value of the parameter X			
	0,1	0,3	0,4	0,5
1	98.700	150.767	161.176	164.500
2	69.791	106.608	113.969	116.319
4	49.350	75.383	80.588	82.250
8	34.896	53.304	56.984	58.160
16	24.675	37.692	40.294	41.125
32	17.448	26.652	28.492	29.080
64	12.338	18.846	20.147	20.563
128	8.724	13.326	14.246	14.540
256	6.169	9.423	10.074	10.281
512	4.362	6.663	7.123	7.270

In Fig. 1 the graphs of dependence γ calculated in accordance with the formula (18) on the number of statistical tests K at different values of the parameter X are given (Table 1).
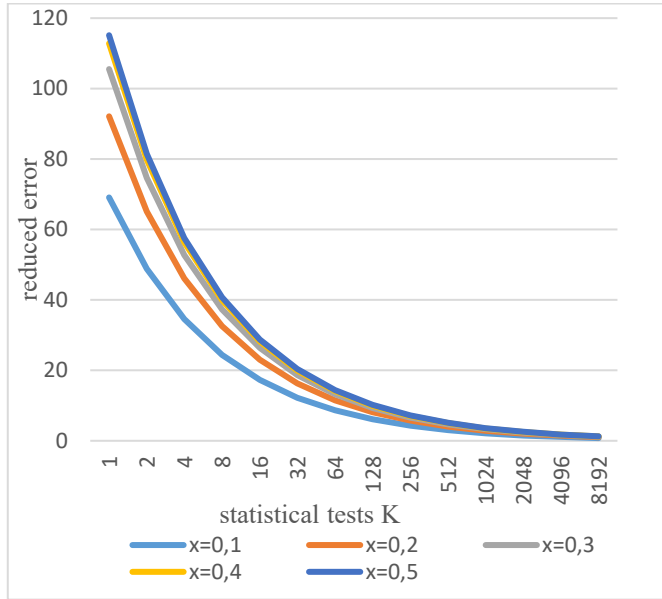
Fig. 1. Dependence of the reduced error on the number of statistical tests K for different values of the parameter X

Thus, the probability that the value $F_{x_i}(R)$ falls within the confidence interval, as before, is assumed to be equal $P = 0,9973$. From the graphs in Fig. 1 we can seen that the reduced error of the probabilistic transformation in the uniform distribution $R(t)$ is less dependent on the value of the converted parameter *X*.

With a sequential method of converting analog or digital information into a probabilistic form of representation, discussed above, methodical errors arise, which are the greater, the smaller the interval of probabilistic transformation. When operating information measuring systems with DSP in real time, this leads to a noticeable reduction in their frequency range and, as a consequence, to the limitation of the scopes of their application [6].

The solution to the problem of increasing the operating speed can be the use of the parallel probabilistic transformation of the measured signal [6].

With the development of modern technology and methods of simulation modelling, the method of statistical tests got into wide use, the main idea of which is the relationship between the probabilistic characteristics of random processes and values that are solutions to problems of mathematical analysis. The considered algorithms for converting any signal into a probabilistic display have a common drawback - the inverse relationship between the transformation accuracy and operating speed.

Representation of information in the form of non-positional pseudo-probabilistic display. To solve the problem of reducing the errors of information representation in the form of non-positional probabilistic display, we shall consider its main properties. The first of them is in synchrony or clock cycle character of members of the probabilistic representation and reduced to the fact that the formation y_{ij} is realized in a steady time interval $\Delta t_j = t_{j+1} - t_j$. The second is the property of independence of each member of the probabilistic representation from any other. This follows from the fact that the generation of the probabilistic display corresponds to the Bernoulli test scheme, and for the random sequence obtained in accordance with this scheme, the autocorrelation function is δ-function

with $\tau = 0$. Thus, the probabilistic display is characterized by the probabilistic nature of formation of the sequence of "1" and "0" and the probabilistic nature of the number of "1" in the sequence, which leads to the error of the probabilistic transformation. Hence, the conclusion: for decrease, and in the boundary case for elimination of the error of probabilistic transformation, it is necessary to abandon the second characteristic feature of the probabilistic display, which in this case will be called "pseudo-probabilistic" [7]. This can be achieved by changing the algorithm of formation of a valid display (1), pre-setting the number of "1" in the probabilistic display in accordance with the expression:

$$P(y_{ij} = 1) = x_i. \tag{19}$$

At the same time, to preserve the properties of probabilistic display, the distribution of units in pseudo-probabilistic display should be equal-dimensional. To achieve this, use the "signal – pseudo-probabilistic display" converter (Fig. 2). The number x_i in a binary positional code, which is to be converted, is entered in the register of the original data RG of the converter.
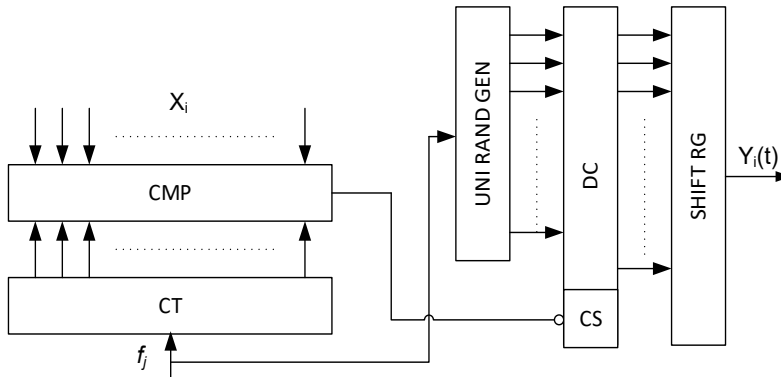

Fig. 2. "Signal – pseudo-probabilistic display" converter

It follows from the expression that the faster the device is, the more its clock frequency is, and the less it is, the greater the number of "1" must be recorded in the register. For the above example, at the same clock frequency of the probabilistic processor and the maximum value x_i, it appears possible to achieve the comparable error of the probabilistic transformation at the operating speed of 700 times more [6, 8].

2 Conclusion

Thus, it turns out obvious that applying non-positional information display in DSP in the form of a probabilistic representation allows minimize the hardware of computing devices twice in average, as a result causing increase in survivability, reliability and energy efficiency, and consequently reduce equipment, design, manufacture, and maintenance costs.

The considered algorithms for converting any signal into a probabilistic representation have a common disadvantage - the inverse interdependence between the transformation accuracy and operating speed.

The application of the proposed algorithm for converting a signal into a pseudo-probabilistic representation enables to increase the speed of the probabilistic processor three times at a given accuracy, while maintaining the reliability and hardware capacity.

References

1. E. Lebedev, N. Galanina, Dynamics of Nonlinear Discrete Electrical and Electronic Systems, (2003)

2. I. Dzegelenok, S. Ocokov, Journal of MEI, **3,** 7 (2010)

3. D. Moiseev, N. Sapozhnikov, P. Bejner, Journal of NTU "HPI", **38(1011)**, 3 (2013)

4. N. Sapozhnikov, D. Moiseev, O. Chuzhikova-Proskurnina, Collection of Scientific Works of SNTUEI, **3 (47)**, 15 (2013)

5. N. Sapozhnikov, D. Moiseev, A. Shokin, Collection of Scientific Works of SNTUEI, **1 (45)**, 8 (2013)

6. D. Moiseev, N. Sapozhnikov, P. Bejner, Technological Audit and Production Reserves]. – Har'kov: ehnologichnij centr, **3/1(11)**, 1 (2013)

7. D. Moiseev, N. Sapozhnikov, P. Bejner, East European Journal of Advanced Technologies, **№ 3/4 (63)**, 6 (2013)

8. N. Sapozhnikov, D. Moiseev, Collection of Scientific Works of SNTUEI, **2 (50)**, 17 (2014)