www.jpis.az

An algorithm of the sequence of artificial symmetric signals for comparing and creating a new convolution method

Azer Kerimov

Institute of Control Systems, B. Vahabzadeh str., 68, AZ1141, Baku, Azerbaijan

a.k.etn00001@gmail.com

orcid.org/0000-0002-1732-2474

ARTICLE INFO

http://doi.org/10.25045/jpis.v15.i2.03

Article history:

Received 16 January 2024
Received in revised form
18 March 2024
Accepted 20 May 2024

Keywords:

Artificial signals
Discrete wavelet transform
Feature extraction
Recognition
Sequence
Mean value
Standard deviation

ABSTRACT

The objective of this article is to create an algorithm of the sequence of artificial signals that can be used to compare and create methods for processing one- and two-dimensional signals. It will then be implemented to compare feature extraction methods that rely on discrete wavelet transforms. The discrete wavelet transform is superior to other signal processing techniques in several ways. Developing a feature set is a crucial step in using the discrete wavelet transform. Mean value and standard deviation are suggested as feature extraction techniques in this study. The mean value is the only option selected for the first feature extraction method; the mean value and standard deviation are selected for the second feature extraction method. To build any number of artificial signal sequences from a single, several conditions are taken into account, for example, their symmetry, they are supposed to be located at the same distance from each other, that is, with an equal step. Symmetrical signal sequences constructed in this way differ from common well-known signal sequences, such as Fourier series, in that they converge to a given signal in equal steps.

1. Introduction

The new algorithm – sequence of artificial symmetric signals (SAS) is introduced in (Kerimov, 2022), to a comparative evaluation of the suitability of signal identification techniques. Additionally, this approach is used to create methods for additive convolution (Rzayev et al., 2023), (Rzayev et al., 2023). Usually, a certain number of signals are chosen, and recognition is done for each signal to determine how good the recognition technique is.

In order to conduct a comparison evaluation of the accuracy of discrete wavelet transform-based feature extraction techniques, the author of this paper creates a SAS algorithm for a comparative quantitative assessment of signal recognition methods and implements it.

2. Related work

The accuracy of the employed methods is determined by the quantity of recognized signals, and these methods are then utilized to evaluate the recognition methods according to the following: the greater the accuracy, the more superior the recognition method (Keogh et al., 2019), (Geler et al., 2019), (Itakura et al., 1975). However, this strategy has drawbacks, including significant processing costs and the assumption of some degree of unpredictability and uncertainty (Sakoe et al., 1978), (Akin, 2022), (Scholl, 2021).

3. Materials and methods

3.1 The algorithm of sequence of artificial signals for comparison and creating methods

This study has been made possible by a new algorithm in the field of signal processing, called as the algorithm of sequence of artificial symmetric signals for comparison and creation of new methods, which is proposed by the author in (Kerimov, 2022). The SAS algorithm's original purpose was to compare several approaches, but as it developed, it became clear that new additive convolution methods might be made using this algorithm. Assume the following analytical function illustrated in Fig. 1 describes the analog signal:

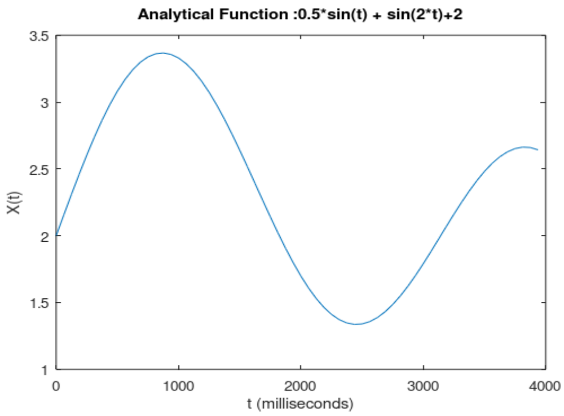

Fig. 1. The analog signal

The following formula can be used to get the Nyquist frequency given the sampling rate:

$$\text{Nyquist frequency} = \frac{\text{Sampling rate}}{2} \quad (1)$$

For instance, the Nyquist frequency equals to 16/2=8 if the sampling rate is 16 samples per second. The maximum frequency for the function x(t)=0.7*sin(t) + sin(2*t)+2 is equal to 1/3.14= 0.3185, which is lower than the Nyquist frequency.

Following sampling, the digital signal s0 is acquired, from which successive artificial signals are constructed in relation to s0 (Kerimov, 2022). As an illustration, the artificial signals {s0, s1, s2, ..., s6} are produced as depicted in Fig. 2.

The SAS algorithm has four criteria. Evaluation criteria are given to compare the adequacy of recognition algorithms based on artificial signals.

Criteria 1 (uniformity of the sequence) states that the distances between the identified signals and the standard should increase steadily over time rather than suddenly.

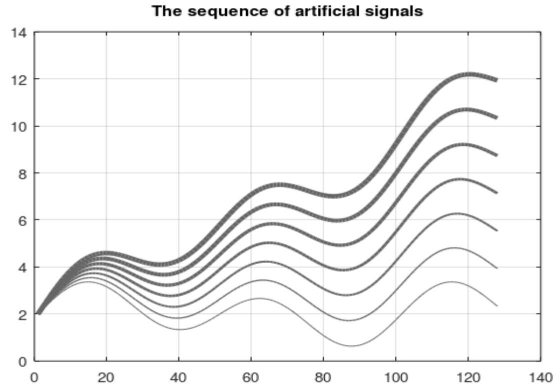

Fig. 2. Sequence of the artificial signals

Criteria 2 (method symmetricity) states that the method distances for a given recognized signal from the left standing and from the right standing signals should be roughly equal, or that their ratio should be roughly equal to one. The distances between the standing signals on the left and right will obviously be exactly same if they are parallel to this signal; that is, their pairwise ratios will equal to one.

Criteria 3 (method performance speed) states the rate of convergence of distance values for a particular method is determined as the recognized signals "approach" the standard.

Criteria 4 (sensitivity of the method) states that adjusting the signal generation phase will alter recognition method results in relation to evaluation criteria 1, 2, and 3. In other words, the recognition method's accuracy will be proportionate to the ratio of altered results.

4. Digitizing of the criterias using mathematics

Mean value (MV) and standard deviation (SD) assessment for methods are used. For each $k = 1 \div N_m$ (k is a number of methods), let us introduce the following notation: $D_k^h(s_i, s_j)$ is the distance between signals $\mathrm{i}, j = 1 \div N_s$ ($N_s$is a number of signals);

4.1 *Evaluation for first criteria*

$$CV1_k^h(i,j) = \frac{D_k^h(s_1,s_{j+1}) - D_k^h(s_1,s_j)}{\nabla t}, \; k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s, \quad (2)$$

$$MCV1_k^h = \max(\mathrm{MV}(CV1_k^h(i,j)) - CV1_k^h(i,j)), \; k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s, \quad (3)$$

$MCV1_k^h$ is the value (first variant) of first criteria which means maximum deviation of the derivatives of distances from mean value or

$$MCV1_k^h = \mathrm{SD}(CV1_k^h(i,j)), \; k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s. \quad (4)$$

$MCV1_k^h$ is the value (second variant) of first criteria which means standard deviation of the derivatives of distances.

4.2. *Evaluation for Second Criteria*

$$CV2_k^h(i,j) = \frac{D_k^h(s_i,s_{j+1})}{D_k^h(s_{i-1},s_i)},\ k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s \quad (5)$$

is the ratio between adjacent distances (that is, between the distances from the right standing (i+1)-th and from the left standing (i-1)-th signals to the i-th signal):

$$MCV2_k^h = \max(\mathrm{MV}(CV2_k^h(i,j)) - CV2_k^h(i,j)),\ k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s. \quad (6)$$

$MCV2_k^h$ is the value (first variant) of the second criteria which means maximum deviation of the ratio between adjacent distances from mean value or

$$MCV2_k^h = \mathrm{SD}(CV2_k^h(i,j)),\ k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s, \quad (7)$$

is the value (second variant) of the second criteria which means standard deviation of the ratio between adjacent distances from mean value.

4.3 *Evaluation for third criteria*

$$CV3_k^h = \frac{D_k^h(s_i,s_j)}{D_k^h(s_i,s_{j+1})},\quad k = 1 \div N_m, i, j = 1 \div N_s.$$

$$\mathrm{MCV3_k^h} = \min(\mathrm{CV3_k^h}), \quad (8)$$

$\mathrm{MCV3_k^h}$ is the value (first variant) of the third criteria which means speed of convergence of the distances $D_k^h(s_i, s_j)$ converging to zero (https) or

$$MCV1_k^h = \mathrm{SD}(CV1_k^h(i,j)),\ k = 1 \div N_m, \mathrm{i}, j = 1 \div N_s. \quad (9)$$

$MCV1_k^h$ is the value (second variant) of the third criteria which means standard deviation of the derivatives of distances.

4.4. *Evaluation for Fourth Criteria*

For fourth criteria there are two set of sequences s_i, $i = 1 \div N_s$ for different steps h_1, h_2.

Further, for each three criteria for different steps h_1, h_2 the ratio (greater value to smaller) of values u_{ik}, $i = 1,2,3, k = 1 \div N_m$ (i number of criteria, N_m is the number of methods being compared) are calculated, reflecting the "sensitivity" of the recognition method. The calculation of these sensitivity for $i = 1,2,3, k = 1 \div N_m$, $h_1 < h_2$ is carried out as follows:

$$u_{ik} = \begin{cases} \frac{CVi_k^{h_1}}{CVi_k^{h_2}}, & CVi_k^{h_2} < CVi_k^{h_1} \\ \frac{CVi_k^{h_2}}{CVi_k^{h_1}}, & CVi_k^{h_2} \geq CVi_k^{h_1} \end{cases},\ i = 1,2,3\,. \quad (10)$$

In the end by summering u_{1k}, u_{2k} and u_{3k} resulting "sensitivity" is calculated:

$$u_k = u_{1k} + u_{2k} + u_{3k}. \quad (11)$$

5. Wavelet based feature extraction methods

Generally speaking, the primary characteristics for recognition are determined in the early stages of signal recognition. The next step in the recognition process is to compute the pairwise distances between the recognized signals using the chosen metric and compare them to the standard. The type of problem being solved and the applied strategy influence the selection of recognition features.

Each signal is broken down into high-frequency and low-frequency components using this recognition approach (Saraswat et al., 2017; Song et al., 2021). The values of the so-called detailed and approximating coefficients define each component. For instance, the wavelet transform (WT) at four levels for one chosen signal from signals shown in Fig. 2, which comprise 128 samples (Yakovlev, 1998) appears as presented in Fig.3:

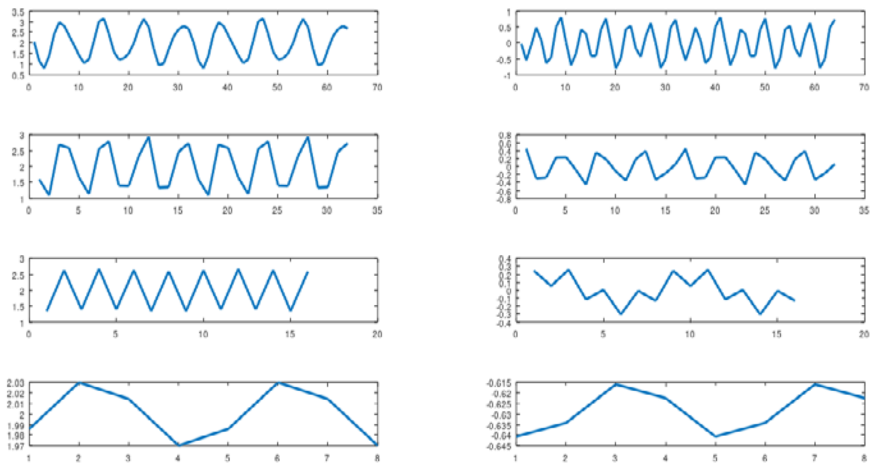

Fig. 3. Wavelet transform at 4 levels

Here, mean values of coefficients in each filtering band are chosen as recognition features. Let's assume the signal has the coefficients in the high-frequency d_{ij} and low-frequency a_{ij} bands respectively, $(i = 1, \dots, N_L)$, $(j = 1, \dots, N_{LP})$, where N_L is the number of decomposition levels, N_{LP} is the number of points in decomposition level N_L.

1) Wavelet based feature extraction method mean value (MV) is built as follows (feature vector):

$$MV(d_{1j}),\ MV(d_{2j}),, \dots, MV(d_{N_Lj}),\\ MV(a_{N_Lj}), \quad j = 1, \dots, N_{LP}.$$

2) Wavelet based feature extraction method mean value (MV) and standard deviation (SD) (feature vector) is built as follows:

$$MV(d_{1j}), SD(d_{1j}), MV(d_{2j}), SD(d_{2j}), \dots, MV(d_{N_Lj}),\\ SD(d_{N_Lj}), MV(a_{N_Lj}),\ j = 1, \dots, N_{LP}.$$

6. Generating methods for forming the sequences of artificial signals

6.1. Generating by line function

This method means adding to given signal the line function:

$$x(t) = kt \tag{12}$$

For example, in our case (Fig.2) from signal x(t) =0.7*sin(t)+sin(2*t)+2 the sequences of artificial signals can be generated by sampling in each step the following function x(t) =0.7*sin(t)+sin(2* t)+2+kt, that is in each step to k is added some increment and then sampling is done.

6.2. Generating by shifting last part of signal

This method means shifting last part of signal horizontally to the right the given number of samples illustrated in Fig.4:

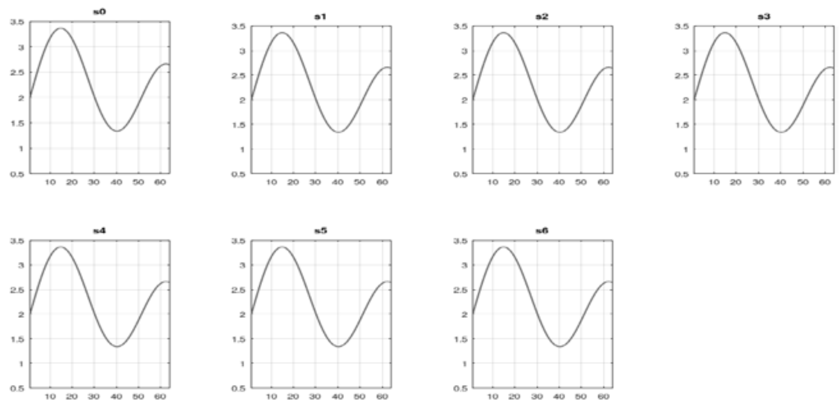

Fig. 4. Generating by shifting last part of signal to the right

7. Calculation results

7.1. Calculation results in two sequences generated by line function $x(t) = kt$

1) The first step is done with following parameters: number of samples n=64, frequency Fs =16, amount of wavelet decomposition levels 4, sampling time T = 1/Fs, k= 1.6/((n-1)*T). Results are shown in tables Table 1 -Table 5.

Table 1. Distance Matrix by the Wavelet MV

	s	s1	s2	s3	s4	s5	s6
s	0	3.2357	6.4814	9.8204	13.207	16.617	20.039
s1	3.2357	0	3.315	6.7081	10.129	13.562	17
s2	6.4814	3.315	0	3.4	6.8258	10.261	13.702
s3	9.8204	6.7081	3.4	0	3.4262	6.8623	10.303
s4	13.207	10.129	6.8258	3.4262	0	3.4361	6.8769
s5	16.617	13.562	10.261	6.8623	3.4361	0	3.4408
s6	20.039	17	13.702	10.303	6.8769	3.4408	0

Table 2. Distance Matrix by the Wavelet MV and SD

	s	s1	s2	s3	s4	s5	s6
s	0	0.95858	1.8073	2.6767	3.618	4.5978	5.5978
s1	0.95858	0	0.89595	1.8484	2.8482	3.8661	4.892
s2	1.8073	0.89595	0	0.98453	2.0028	3.0313	4.0639
s3	2.6767	1.8484	0.98453	0	1.0195	2.0487	3.0818
s4	3.618	2.8482	2.0028	1.0195	0	1.0293	2.0625
s5	4.5978	3.8661	3.0313	2.0487	1.0293	0	1.0332
s6	5.5978	4.892	4.0639	3.0818	2.0625	1.0332	0

Table 3. Values of criteria for the first step

Method	Fist Criteria	Second Criteria	Third Criteria
MV	0.11501	0.0058975	0.49922
MV and SD	0.079081	0.020232	0.53038

2) The second step is done with following parameters:
number of samples n=64, frequency Fs =16, amount of wavelet decomposition levels 4, sampling time T = 1/Fs, k= 5.2/ ((n-1)*T);

Table 4. Values of criteria for the second step

Method	Fist Criteria	Second criteria	Third criteria
MV	0.078725	0.01042	0.49022
MV and SD	0.12203	0.033334	0.47633

Table 5. Ratio of values for all three criteria

Method	Fist Criteria	Second Criteria	Third Criteria	All Three Criteria
MV	1.4609	1.7669	1.0184	4.2462
MV and SD	1.5431	1.6476	1.1135	4.3042

7.2. Calculation results in two sequences generated by line function with number of samples n=128

1) The first step is done with following parameters: number of samples n=128, frequency Fs =16, amount of wavelet decomposition levels 4, sampling time T = 1/Fs, k= 1.6/((n-1)*T). Results are shown in tables Table 6 -Table 8.

Table 6. Values for the first step criteria

Extracting Method	Fist Criteria	Second Criteria	Third Criteria
MV	0.0012603	4.8964e-05	0.50072
MV and SD	0.00019809	0.0003897	0.49996

2) The second step is done with following parameters:
Number of samples n=128, frequency Fs =16, amount of wavelet decomposition levels 4, sampling time T = 1/Fs, k= 5.2/((n-1)*T);

Table 7. Values for the second step criteria

Extracting Method	Fist Criteria	Second Criteria	Third Criteria
MV	0.0068304	0.001885	0.50305
MV and SD	0.016848	0.0043079	0.50162

Table 8. Ratio of values for all three criteria

Extracting Method	Fist Criteria	Second Criteria	Third Criteria	All Three Criteria
MV	1.0146	1.5785	1.0026	3.5957
MV and SD	1.0876	1.5141	1.0274	3.6291

7.3. Calculation results in two sequences generated by shifting last part of signal horizontally to right

The first and second steps with h=10,15 are done with following parameters:
number of samples n=128, frequency Fs =16, amount of wavelet decomposition levels 4, sampling time T = 1/Fs. Results are shown in tables Table 9 -Table 11.

Table 9. Distance matrix by the wavelet MV

	s	s1	s2	s3	s4	s5	s6
s	0	0.042906	0.028071	0.04563	0.060521	0.035399	0.024104
s1	0.042906	0	0.023594	0.064069	0.085364	0.060987	0.028266
s2	0.028071	0.023594	0	0.052115	0.070243	0.044417	0.015959
s3	0.04563	0.064069	0.052115	0	0.028944	0.017239	0.039319
s4	0.060521	0.085364	0.070243	0.028944	0	0.028468	0.061248
s5	0.035399	0.060987	0.044417	0.017239	0.028468	0	0.03446
s6	0.024104	0.028266	0.015959	0.039319	0.061248	0.03446	0

Table 10. Distance matrix by the wavelet MV and SD

	s	s1	s2	s3	s4	s5	s6
s	0	0.028531	0.035503	0.052843	0.056971	0.059274	0.059671
s1	0.028531	0	0.01859	0.038745	0.041728	0.047109	0.049523
s2	0.035503	0.01859	0	0.037643	0.042127	0.047154	0.05011
s3	0.052843	0.038745	0.037643	0	0.022632	0.031721	0.039746
s4	0.056971	0.041728	0.042127	0.022632	0	0.024688	0.036495
s5	0.059274	0.047109	0.047154	0.031721	0.024688	0	0.027825
s6	0.059671	0.049523	0.05011	0.039746	0.036495	0.027825	0

Table 11. Ratio of values for all three criteria

Extracting Method	Fist Criteria	Second Criteria	Third Criteria	All Three Criteria
MV	1.1111	1.6854	1.6342	4.4306
MV and SD	1.8086	1.5012	1.452	4.7619

8. Discussion

Three criteria were used in the calculations to demonstrate the correctness of the fake signal algorithm sequence for comparison and method creation.

The calculations also demonstrated that the accuracy feature extraction approach that was anticipated yielded better results when both the mean value and the standard deviation were selected, as opposed to selecting only the mean value.

This article's development objective was also met. The criteria were defined, and the term “deterioration” was replaced with “sensitivity”, which serves as the primary evaluation in the algorithmic sequence of false signals used to create techniques and compare them.

9. Conclusion

The analysis of tables 5, 8, and 11 revealed that the values for the sensitivity of the mean value and standard deviation feature extraction method, which were 4.3042, 3.6291, and 4.7619, respectively, were greater than the values for the sensitivity of the mean value feature extraction method, which were 4.2462, 3.5957, and 4.4306, in all calculations performed in different sequences of artificial signals with different numbers of samples. Since the mean value and standard deviation feature extraction approach was more accurate than the mean value feature extraction method, these results were anticipated.

References

Akin, M. (2002). Comparison of Wavelet Transform and FFT Methods in the Analysis of EEG Signals. Journal of Medical Systems, 26, 241-247.

Blatter, K. (2006) Wavelet analysis. basic theory. Translated from English, Tekhnosfera, Moscow.

Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition, IEEE Transactions on Acoustics, Speech and Signal Processing, 23(1), 67-72.

Geler, Z., Kurbalija, V., Ivanović, M., Radovanović, M., Dai, W. (2019). Dynamic time warping: itakura vs sakoe-chiba, IEEE International Symposium On Innovations in Intelligent SysTems and Applications (INISTA).

https://en.wikipedia.org/wiki/rate_of_convergence

Keogh, E., Pazzani, M. (2017). Derivative dynamic time warping. In Proceedings, 2001th SIAM International Conference on Data Mining (SDM), https://doi.org/10.1137/1.9781611972719.1

Kerimov, A. (2022). Accuracy comparison of signal recognition methods on the example of a family of successively horizontally displaced curves. Informatics and Control Problems, 42(2), 80–91.

Kerimov, A. (2022). Comparison of some signal recognition methods for their adequacy. In Proceedings, 8th International Conference on Control and Optimization with Industrial Applications. 1, Baku, Azerbaijan.

Rzayev, R., Kerimov, A. (2023). Comparison of signal recognition methods by combined use of appropriate evaluation criteria within the additive convolution. Problems of Information Society, 14 (2), 24–31.

Rzayev, R., Kerimov, A. (2023). Signal recognition by using addictive convolution criteria. Proceedings of IAM, 12(1), 52–64.

Rzayev, R., Kerimov, A. (2023). Signal recognition using weighted additive convolution of evaluation criteria. The Springer Series "Lecture Notes in Net-works and Systems", 758(2), 407–416.

Rzayev, R., Kerimov, A, Gurbanli, U., Salmanov, F. (2024). Criteria for assessing the adequacy of image recognition methods and their verification using examples of artificial series of signals. Problems of Information Society, 15(1), 10-17, http://doi.org/10.25045/jpis.v15.i1.02

Sakoe, H., Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing, 26(1). https://ieeexplore.ieee.org/document/1163055

Santos, M., Morais, C., Nascimento, Y., Araujo, J., Lima, K. (2017). Spectroscopy with computational analysis in biological studies: a decade (2006–2016). TrAC Trends Anal Chem, 97, 244–256.

Saraswat, S., Srivastava, G., Sachchidanand, N. (2017). Wavelet transform based feature extraction and classification of atrial fibrillation arrhythmia. Biomed Pharmakoi Journal 10(4) http://biomedpharmajournal.org/?p=17470

Scholl, S. Fourier, Gabor, Morlet or Wigner (2021): Comparison of Time-Frequency Transforms, https://arxiv.org/abs/2101.06707

Yakovlev, A.N. (2024) Introduction to wavelet transforms, textbook, Novosibirsk, NSTU Publishing House (in Russian).