Surface Roughness Classification Using a Translation Invariant Wavelet Packet Decomposition

Dragoș NASTASIU, Denis STĂNESCU, Cristina DESPINA-STOIAN, Angela DIGULESCU, Cornel IOANA, Alexandru ȘERBĂNESCU

***Abstract*—Image pattern recognition is an automatic process capable of identifying regularities in data and to discriminate between different categories. One important challenge in pattern recognition is related to the feature extraction techniques that should be invariant to affine transformations of the images, such as translations. In this study, we propose the use a translation-invariant Wavelet Packet Decomposition to decompose an image, entropy features to characterize it and a two-layer neural network to classify it. The patterns used in this study are images of different levels of surface roughness from samples. The surface roughness is an essential aspect in determining the health of the pipe system and the possible damages that can occur, which lead to increased maintenance costs and disruption of the system. We prove the capabilities of our method with a high accuracy obtained on a separate testing set.**

***Index Terms*—pattern recognition, invariance constraints, entropy, machine learning, neural networks.**

I. Introduction

One of the most difficult tasks in industrial environments and manufacturing processes is inspecting for surface visual appearing [1]. Surface roughness is an important factor in determining the normal state and functioning of machined components and pipe systems [2]. Usually, the precision measurement of the surface roughness is done with a stylus instrument. The issue with this approach is that the process is intrusive and not suitable for convenient and facile inspection as it requires considerable time for setting up the instrument, but also it has a limitation in measuring fine surfaces.

In critical applications, it is desired to have an accurate estimation of the roughness, but this is not always possible, due to difficulties in accessing the surveillance zone, interruptions of the systems and in-situ maneuvering of the measuring instruments.

However, it is not always necessary to make an accurate prediction of the roughness coefficient. A sufficient information for operators of pipe systems is to determine if the level of roughness is critical and if a damage to the system is imminent. Therefore, in this paper, we estimate the risk of the possible damages using images with different level of degradation which are priori labeled by experts.

Nowadays, the image processing techniques along with the rapid advancements in machine learning made possible to create many approaches for surface measurements and damage prediction using: optical techniques [3], laser speckle images [4] and hybrid systems [5], which combine the capabilities of powerful image characterization techniques with feature extraction and usually, neural networks.

The disadvantage of the optical and laser approaches is that they require additional equipment and specialized personnel to use them. In this case, hybrid systems are the most suitable for our task as they usual use optical images of the surfaces that are processed using the available techniques such as Fourier or Wavelet Transform.

In this study, our final goal is to classify three image patterns that correspond to low, medium and high possibility of damage appearing in the system. An important aspect to consider is the view-point variation of the analyzed images. For humans, it is easy to identify a pattern regardless of its orientation, but for a computer, the variations of the images will result in a misinterpretation and wrong classification. Our hybrid approach uses a Translation-Invariant WPD (TI-WPD) [6] to decompose the pattern into multiple frequency bands and a basis selection algorithm that generates a subspace of minimum entropy distribution. The next step, is to extract features from the selected best basis in terms of entropy of the diagonal coefficients at different levels of the decomposition. Lastly, the created feature vector is fed to a two-layer neural network to classify the images and to estimate the possibility of damage. We demonstrate that in the aforementioned tasked, our approach shows interesting preliminary results and proves its capacity to be used in general pattern recognition tasks with radar, acoustic or other types of images.

The paper is organized as follows: Section II discusses the theoretical aspects of our method, including the WPD, feature extraction and the neural network; Section III presents the image database; Section IV addresses the experimental results obtained; Section V closes the paper with the conclusions.

D. NASTASIU is with Military Technical Academy "Ferdinand I", Communications Department, Bucharest, Romania. (e-mail: dragos.nastasiu@mta.ro).

D. STANESCU is with Military Technical Academy "Ferdinand I", Communications Department, Bucharest, Romania. (e-mail: denis.stanescu@mta.ro).

C. DESPINA-STOIAN is with Military Technical Academy "Ferdinand I", Communications Department, Bucharest, Romania. (e-mail: cristina.despina@mta.ro)

A. DIGULESCU is with Military Technical Academy "Ferdinand I", Communications Department, Bucharest, Romania. (e-mail: angela.digulescu@mta.ro).

C. IOANA is with University of Grenoble-Alpes, Grenoble, France. He is within the GIPSA-LAB at the same institute. (e-mail: cornel.ioana@gipsa-lab.fr).

A. SERBANESCU is with Military Technical Academy "Ferdinand I", Communications Department, Bucharest, Romania. (e-mail: alexandru.serbanescu@mta.ro).

Digital Object Identifier 10.32754/JMT.2022.2.01

II. THEORETICAL ASPECTS

WPD is the expansion of the WT that decomposes a signal, in our case a 2-dimensional one, into an over-complete time-frequency analysis. In the case of images, a one level WPD results into four sub-images: one approximation coefficient and three detail coefficients. If the depth of the decomposition is increased, at each step, every sub-image is further decomposed into another four, and the process continues until a criterion is met. The classic version of the 2-D WPD is implemented using the principles of the multiresolution analysis, filter banks and the pyramidal image decomposition technique [7]. Therefore, we can define the 2-D WPD of an $M \times M$ image up to a depth $D+1$, where $D \le \log_2 M$ as follows:

$$C_{4k,(i,j)}^{d+1} = \sum_m \sum_m h(m)h(n)C_{k(m+2i,m+2j)}^{d} \quad (1)$$

$$C_{4k+1,(i,j)}^{d+1} = \sum_m \sum_m h(m)g(n)C_{k(m+2i,m+2j)}^{d} \quad (2)$$

$$C_{4k+2,(i,j)}^{d+1} = \sum_m \sum_m g(m)h(n)C_{k(m+2i,m+2j)}^{d} \quad (3)$$

$$C_{4k+3,(i,j)}^{d+1} = \sum_m \sum_m g(m)g(n)C_{k(m+2i,m+2j)}^{d} \quad (4)$$

where C_0^0 is the initial image, k is the node's index in the wavelet packet tree, representing each sub-band, h and g are a pair of quadrature mirror filters' impulse responses. Recursively, the image C_k^d is filtered into four images at 2 times lower resolution, C_{4k}^{d+1}, C_{4k+1}^{d+1}, C_{4k+2}^{d+1}, C_{4k+3}^{d+1}.

In Fig. 1 we depict the graph representation of an over-complete wavelet analysis.

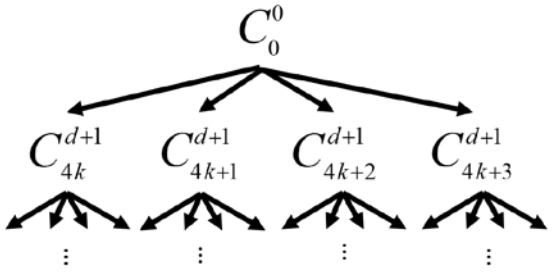


Figure 1. Over-complete WPD graph

A. *"Best basis" selection*

In general, the over-complete WPD determine many possible decompositions that are inefficient and computationally expensive to work with in practice. In this sense, there are energy or entropy optimal algorithms [8] that can generate the "best" basis from the initial WPD. In our study, the selection algorithm is based on the minimization of the nonnormalized Shannon entropy function for an image $I(x,y)$:

$$Ent\left[I(x,y)\right] = \sum_x \sum_y I(x,y)\ln\left[I(x,y)\right]. \quad (5)$$

For a better understanding of the algorithm, we summarize the steps taken in our processing in this stage. Starting from the initial image and proceeding level by level in the decomposition, we first compute the Shannon entropy for each node, ent_P (entropy of the parent node) and the entropy of its four children, ent_{Approx} (entropy of the approximation coefficients), ent_{Vert} (entropy of the vertical detail coefficients), ent_{Hor} (entropy of the horizontal detail coefficients), ent_{Diag} (entropy of the diagonal detail coefficients). Afterwards, the selection takes places by comparing the summed entropy of the children nodes with the entropy of the parent node. If the entropy of the parent is lower or equal, we keep the parent node in the best basis, otherwise we keep the children.

Finally, after the selection algorithm, we determine the entropy-optimal decomposition as it is presented in Fig. 2.

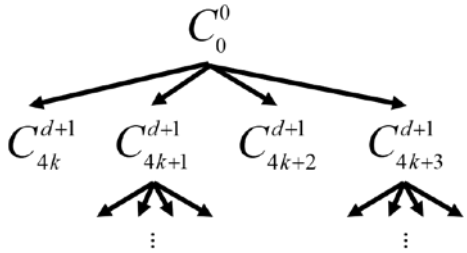
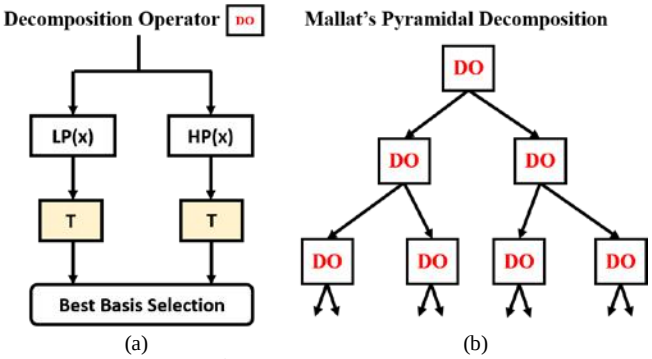


Figure 2. Entropy-optimal WPD

B. *Translation invariancy*

As we stated in the introduction, the problem tackled from a signal processing point of view, regards the translation invariancy of the decomposition, as we aim for a unique decomposition of the same patter, regardless of its position. To achieve, this the WPD can be improved by adding an additional degree of freedom, generated at the decomposition stage and incorporated into the "best" basis selection algorithm. At each decomposition step, besides the usual coefficients, we also generate their translated versions, in this manner, increasing the subspace from which we select the "best" basis, as it is depicted in Fig. 3 with the Decomposition Operator (DO) and the Mallat's pyramidal decomposition principle.



(a) (b)

Figure 3. Presentation of the: (a) translation invariant decomposition operator and (b) Mallat's pyramidal decomposition principle

The additional coefficients are generated by circularly shifting the original coefficients one pixel to the left, then one pixel to the right, and finally, combining the shifting on vertical and horizontal directions also with one pixel. It is demonstrated that these three new coefficients are sufficient to achieve a translation-invariant decomposition regardless of the shifting degree of the image to be decomposed [6].

With respect to the translations, $T \in \{0,-1\}$ the TI-WPD can be defined as follows:

 Digital Object Identifier 10.32754/JMT.2022.2.01

$$C_{4k,(i,j)}^{d+1,(T,T)} = \sum_m \sum_n h(m)h(n)C_{k(m+2i+T,n+2j)}^{d} \quad (6)$$

$$C_{4k+1,(i,j)}^{d+1,(T,T)} = \sum_m \sum_n h(m)h(n)C_{k(m+2i+T,n+2j)}^{d} \quad (7)$$

$$C_{4k+2,(i,j)}^{d+1,(T,T)} = \sum_m \sum_n g(m)h(n)C_{k(m+2i+T,n+2j)}^{d} \quad (8)$$

$$C_{4k+3,(i,j)}^{d+1,(T,T)} = \sum_m \sum_n g(m)g(n)C_{k(m+2i+T,n+2j)}^{d} \quad (9)$$

Equations (6)-(9) represent all the wavelet coefficients that appear if the original analyzed image is translated before each filtering procedure in the manner described earlier. Afterwards, we select the "best" time-frequency bands from the generated subspace based on the entropy cost function. By repeating the procedure, at each node, we can successfully select a space of minimum entropy distribution, invariant to any translations.

Fig. 4 (a), (d) shows an image and its translated version. The 256×256 pixels initial image is translated vertically with 128 pixels. The filters used in this example are 'coiflet3'. For both images we provide the corresponding WPD and TI-WPD analysis for a depth $D = 3$ in Fig. 4 (b), (c), (e) and (f), respectively. From a short visual inspection, we observe that the "best" basis coming from the TI-WPD is the same regardless of the translation of the initial image, while in the classic WPD the basis changes. This feature of TI-WPD allows us to provide the same unique decomposition of a pattern regardless of the possible translations that can occur in real world images.

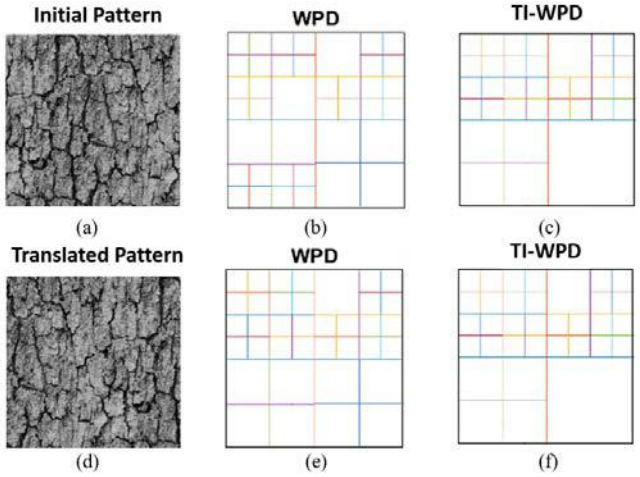


Figure 4. Over-complete WPD graph

C. *Feature Extraction*

In terms of feature extraction, we introduce a novel approach to characterize an image with a feature vector with the length equal to the decomposition depth. Thus, after a translation invariant decomposition of the image, we compute the sum of all diagonal coefficients for each level of the decomposition, as follows:

$$F(d) = \sum_x \sum_y C_{4k+3,(i,j)}^{d+1}(x,y) \ln\left(C_{4k+3,(i,j)}^{d+1}(x,y)\right). \quad (10)$$

In other words, we compute an entropy value that characterizes each level of decomposition in terms of diagonal details. The same method can be applied, for example, for the vertical, horizontal or approximation wavelet coefficients, but the discrimination between classes is not always sufficient for an accurate classification, as we observe in the following figures. Fig. 5 shows the image with high roughness, Fig. 6 presents the TI-WPD "best" basis of this image and Fig. 7 depicts a comparison for a 4-level TI-WPD for surface roughness images, where on Ox axis we have each one of the 4 levels of wavelet decomposition and on the Oy the computed entropy.

Figure 5. High roughness pattern

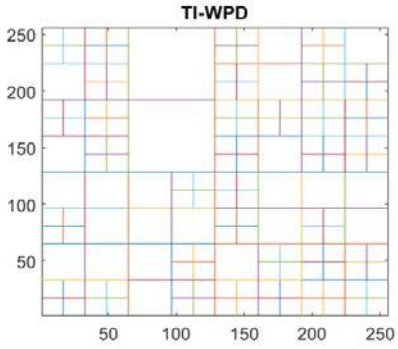


Figure 6. High roughness pattern

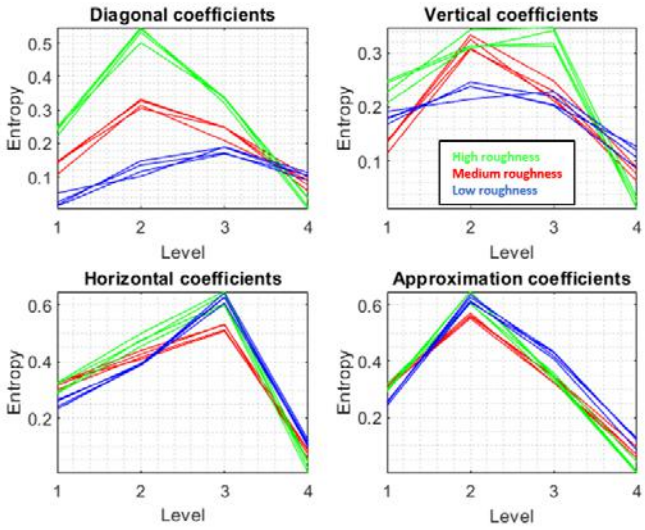


Figure 7. Comparison of entropy computation for each one of the four types of coefficients, starting from upper-left: diagonal, vertical, horizontal and approximation coefficients

III. Neural Network Architecture

The network we propose has 2 fully connected layers with 4 neurons, with ReLU activations and one final dropout layer as depicted in Fig. 8. The choice of the activation functions is argued by its simplicity, its accelerated rate of convergence when computing gradients and by its characteristic of not saturating the positive gradients.

The regularization layer comes in to apply a dropout with a rate of 0.2 on the final fully connected layer. This means that a fraction of the neurons is dropped and not considered in the training. Overall, the regularization layer leads to reduced overfitting by making the presence of any hidden

neuron unreliable. The last layer, has 3 neurons, corresponding to each one of the 3 classes of surface roughness and it uses Softmax in order to offer a probabilistic interpretation of the values resulting through feed-forwarding.

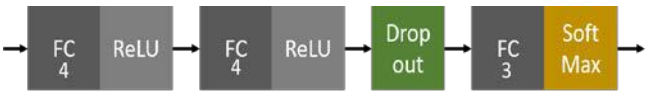


Figure 8. Neural classifier with fully connected layers, dropout and Softmax

IV. Results

The database of images used in this study was created by capturing different zones of three surfaces with roughness classified as low, medium or high. The photos were taken at 10 cm distance from the samples with the camera of a mobile phone. The images are transformed to grayscale and then resized to 256×256 pixels with values varying between 0 and 255. The database consists of 20 images per class, thus 60 images. Some of the samples are presented in Fig. 9.

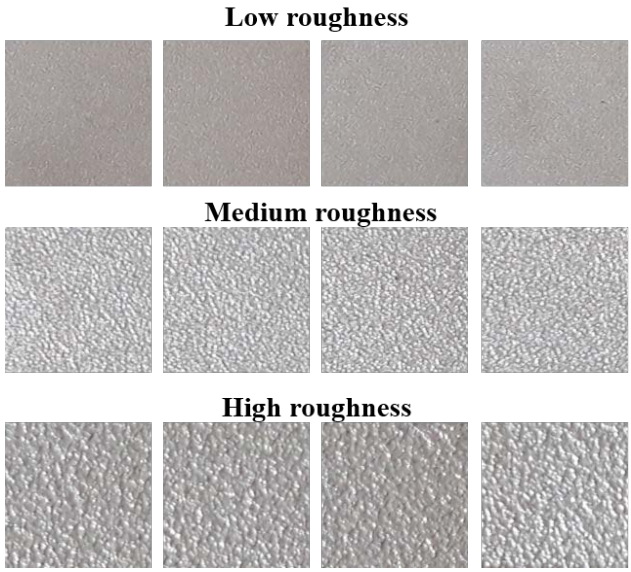


Figure 9. The 3 levels of roughness from our database

The set is divided into training, validating and testing sets with a ratio of 2:1:1, therefore having 10 images per class to train, 5 images to validate during training and 5 to test the performances of the classifier.

The neural network presented in Section III is trained for 10 epochs and then evaluated on the testing set. Fig. 10 present the evolution of the learning curves, the train/validation loss and accuracy. Having a steady convergence means that the network has easily learnt the dataset. Comparing the differences between the validation and train learning curves, we conclude that there is little to no overfitting.

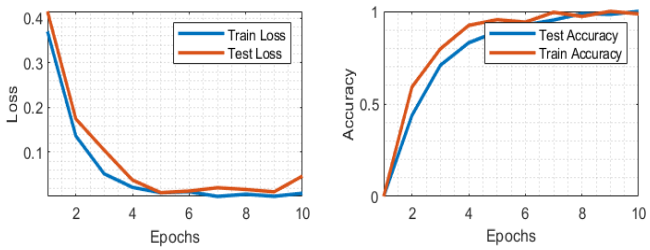


Figure 10. Learning curves resulted from training the proposed network

The confusion plot in Fig. 11 shows how well the classifier performed on the testing set on the testing set. The overall accuracy is 93.3%. Considering the small testing set, the statistical error is increased. However, the preliminary results are promising and prove the robustness of our approach in this roughness classification challenge.

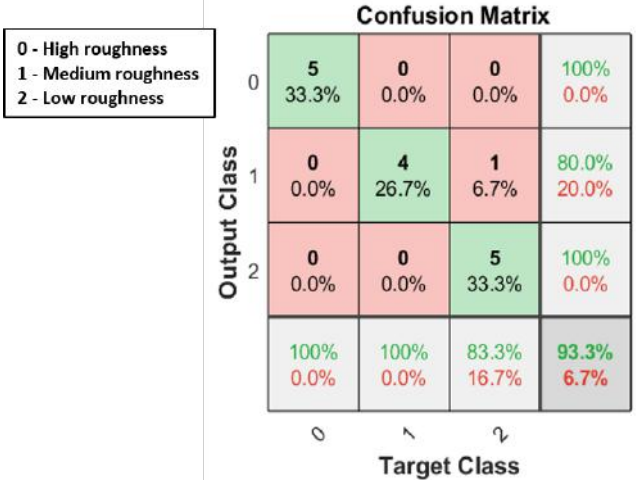


Figure 11. Confusion matrix

V. Conclusions

The proposed approach was proved to be of use in the roughness classification of different surfaces. TI-WPD increases the classifier capacity to learn features coming from images with view-point variation, such as translation, otherwise complex to learn by itself and also not suitable for such a compact neural network such as ours.

Further work will include increasing the database of images and thus increasing the statistical stability of the preliminary results.

In perspective, public benchmark databases containing patterns will be tested with our method to observe the performance of the method.

References

[1] W. Jiao, "The Classification of Surfaces," *ACM Int. Conf. Proceeding Ser.*, pp. 66–79, Jun. 2021, doi: 10.1145/3475827.3475837.

[2] T. Jeyapoovan and M. Murugan, "Surface roughness classification using image processing," *Measurement*, vol. 46, no. 7, pp. 2065–2072, Aug. 2013, doi: 10.1016/J.MEASUREMENT.2013.03.014.

[3] F. Gao, A. G. Marrugo, and S. Zhang, "State-of-the-art active optical techniques for three-dimensional surface metrology: a review [Invited]," *JOSA A*, vol. 37, no. 9, pp. B60–B77, Sep. 2020, doi: 10.1364/JOSAA.398644.

[4] D. Joseph and P. Bisnoi, "Analysis of surface roughness using laser speckle interferometry," *Saratov Fall Meet. 2012 Opt. Technol. Biophys. Med. XIV; Laser Phys. Photonics XIV*, vol. 8699, p. 869911, Feb. 2013, doi: 10.1117/12.2018867.

[5] M. Elangovan, N. R. Sakthivel, S. Saravanamurugan, B. B. Nair, and V. Sugumaran, "Machine Learning Approach to the Prediction of Surface Roughness Using Statistical Features of Vibration Signal Acquired in Turning," *Procedia Comput. Sci.*, vol. 50, pp. 282–288, Jan. 2015, doi: 10.1016/J.PROCS.2015.04.047.

[6] I. Cohen, S. Raz, and D. Malah, "Orthonormal shift-invariant wavelet packet decomposition and representation," *Signal Processing*, vol. 57, no. 3, pp. 251–270, Mar. 1997, doi: 10.1016/S0165-1684(97)00007-8.

[7] S. G. Mallat and G. Peyré, "A wavelet tour of signal processing : the sparse way," p. 805, 2009, Accessed: Jun. 05, 2022. [Online]. https://books.google.com/books/about/A_Wavelet_Tour_of_Signal_Processing.html?id=5qzeLJljuLoC.

[8] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, 1992, doi: 10.1109/18.119732.