Piotr BOJARCZAK, Piotr LESIAK
TECHNICAL UNIVERSITY OF RADOM, FACULTY OF TRANSPORT AND ELECTRICAL ENGINEERING
Malczewskiego 29, 26-600 Radom, Poland

Visual system diagnosing the state of elements fastening the rail to the sleepers

Ph.D. Piotr BOJARCZAK

He received M.Sc and Ph.D from the Radom University of Technology in Transport and from the Warsaw University of Technology in Electrical Engineering respectively. He has been working at the Faculty of Transport and Electrical Engineering Technical University of Radom since 1990. His research interests are in neural networks and signal and image processing.

e-mail: bojarczpp@wp.pl

Ph.D. Piotr LESIAK

He received M.Sc and Ph.D from the Gdańsk University of Technology and the Warsaw University of Technology respectively all in Electrical Engineering. He received Dr.Sc from the Warsaw University of Technology in Transport. He has been working at the Faculty of Transport and Electrical Engineering of Radom University of Technology for many years. His research interests are in development of diagnostic and measuring systems for railway transport.

e-mail: p.lesiak@pr.radom.pl

Abstract

The paper presents an attempt of application of image processing algorithms along with neural networks to diagnosis of the state of elements fastening the rail to the sleepers in a railway line. The system allows for detection of the presence or the absence of fastening elements on a basis of the sleeper view and its neighborhood. The system effectiveness is equal to 90%.

Keywords: railway line, image processing, wavelet transform, neural networks.

Wizyjny system diagnostyki stanu elementów mocujących szynę do podkładu

Streszczenie

Artykuł przedstawia próbę wykorzystania algorytmów przetwarzania obrazów oraz sieci neuronowych do diagnostyki stanu elementów mocujących szynę do podkładu w torze kolejowym. System na postawie obrazu podkładu wraz z podsypką wykrywa obecność lub nieobecność elementu mocującego podkład do szyny. System składa się z trzech bloków: bloku zawężającego obszar poszukiwania elementów mocujących, do obszaru podkładu, bloku określającego cechy charakterystyczne elementu mocującego oraz bloku klasyfikatora (sieci neuronowej). Blok pierwszy przeprowadza segmentację obrazu w oparciu o jego teksturę z wykorzystaniem filtru entropijnego w celu wyodrębnienia podkładu z analizowanego obrazu. Blok drugi wykorzystując dwuwymiarową transformatę falkową z falką Coif1 wydobywa cechy charakterystyczne elementów mocujących szynę. Blok trzeci w oparciu o wydobyte cechy przeprowadza klasyfikację - element mocujący dobry, uszkodzony bądź jego brak. Jako klasyfikator została użyta sieć neuronowa hybrydowa. Skuteczność zaproponowanego rozwiązania wyniosła 90%.

Słowa kluczowe: tor kolejowy, przetwarzanie obrazów, transformata falkowa, sieci neuronowe.

1. Introduction

The state of elements fastening the rail to the sleepers is a crucial issue having the significant influence on the railway traffic safety of . Missing a few subsequent fastening elements can pose a threat to the train moving on this rail. Often it can lead to the train derailment what in turn causes huge financial and human losses. At present in the Polish Railway Lines the state of fastening elements is checked by a trained staff. This approach is highly inefficient, one person can check a very limited distance of the railway line. In addition to that, this method is highly dependent on fatigue and physical condition of a controlling person. It made the authors to deal with this problem. The authors tried to choose the best image processing algorithms allowing for automatic visual inspection of fastening elements. On the basis of view of a sleeper along with its surrounding neighborhood, this method is able to presence or the absence of fastening element. Algorithm consists of three major blocks:

- determine the area in which fastening element is to be searched,
- determine the salient features of fastening element,
- detect a fastening element.

The paper presents in details this algorithm.

2. Determination of the area for searching fastening element

As it was mentioned above, the task of the algorithm first block is restriction of the area in which a fastening element is to be searched to the minimal small range. Because fastening elements are placed on the sleeper, therefore the system should firstly extract the sleeper from the image. This extraction can be treated as a typical image segmentation. There are many image segmentation techniques. The authors decided to use an entropy filter [1] as a tool for image segmentation. This approach is based on texture recognition and image segmentation. It turns out that the textures of ballast and sleepers differ significantly. It allows us to use the entropy filter having some advantages towards other techniques. Its main advantages are relative simple implementation and fast operation. The entropy filter uses the pixel intensity distribution (histogram) to determine the texture. The area covered with ballast is characterized by a uniform ("flat") distribution of the pixel intensity, whereas the area containing the sleeper is characterized by a much more sharp (spike-like) distribution of the pixel intensity. These two different distributions can be described through the entropy. The higher entropy corresponds to a flat distribution, whereas the lower entropy corresponds to a sharp distribution. In case of the entropy filter, the pixel intensity distribution and its corresponding entropy is not calculated for the whole image but only for a subsequent small fixed size area of the analyzed image. The size and shape (usually rectangular) of this area is determined by the filter frame. After determining the distribution, its corresponding entropy is calculated according to the formula:

$$E = \sum_{i=0}^{N} P(i) \frac{1}{\log\left(P(i)\right)} \tag{1}$$

where: N is the maximal pixel intensity for the whole analyzed image (N=255 for grayscale image), i is i-th level of the pixel intensity, $P(i)$ is the probability of occurring a pixel with i-th intensity level in the image area determined by the filter frame.

During filtering process the filter frame is successively moved by 1 pixel through the whole analyzed image and the intensity distribution along with its corresponding entropy is calculated for every position of the filter frame. After completion of the filtering process, the new image of the same size as the analyzed one is obtained. Every pixel of this image contains the entropy of the intensity distribution corresponding to a certain position of the filter frame. Therefore the areas of different textures will have different entropy in the generated image. In order to extract the sleeper from the image, it is sufficient to determine the global threshold value T for the entropy. The entropy lower than threshold T will correspond to the area of sleeper, whereas the entropy higher than threshold T will correspond to the area covered with ballast. There are two crucial parameters having the influence on the quality of the extraction process [2, 3]. First is the filter frame size. There exists an optimal size at which the separation between the sleeper and ballast is best. In order to estimate this optimal size, the authors chose randomly 60 images of sleepers with ballast. The resolution of images taking part in the experiment was equal to 400x300 pixels. Next, in each image the sleeper and ballast were manually extracted. In the next step the filtering process with the filter frame of size 9x9, 15x15, 19x19, 25x25 and 33x33 pixels was performed for all manually extracted sleepers and ballasts. Additionally, for each size of the frame, the entropy distributions for sleeper and the ballast were calculated. The optimal filter frame is calculated on a basis of the maximal distance between these distributions- each of them is unimodal. The distance between these distributions is defined as [2, 3]:

$$d = \left|(\mu_1 - \sigma_1) - (\mu_2 + \sigma_2)\right| \tag{2}$$

where: μ_1 and σ_1 are the mean and standard deviation of the entropy distribution for ballast and μ_2 and σ_2 are the mean and standard deviation of the entropy distribution for the sleeper. Fig. 1 shows the distance *d* between these distributions versus the filter frame length.

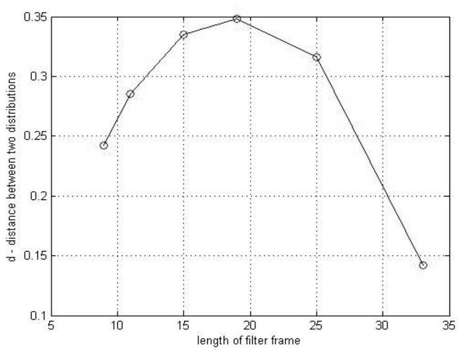

Fig. 1. Distance between two distributions vs. the filter frame length
Rys. 1. Zależność odległości między dwoma rozkładami a długością ramy filtru

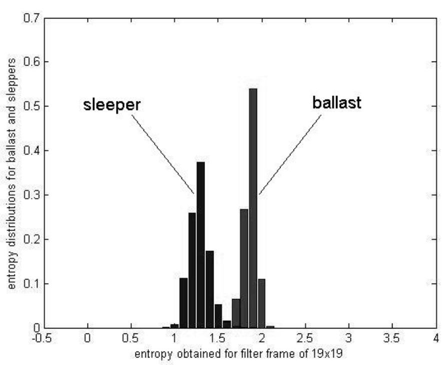

Fig. 2. Entropy distributions .for sleeper and ballast for filter frame of 19x19
Rys. 2. Rozkład entropii dla podkładu i podsypki dla filtru o ramie 19x19

On a basis of Fig. 1 there was chosen the filter of size 19x19 pixels. In the next step we determined the optimal threshold T – the second crucial parameter. Fig. 2 presents the entropy distributions for the sleeper and ballast for the filter frame size of 19x19. In this approach determination of the sleeper or ballast can be regarded as a classification process (distinction between two groups). Examining the area characterized by the entropy belongs to this group for which the corresponding value of entropy distribution is larger. According to [4], the optimal classification (segmentation) is performed for the entropy threshold value T corresponding to the intersection of these two distributions. Hence, the threshold T is equal to 1.6.

Fig. 3 presents the image after application of segmentation.

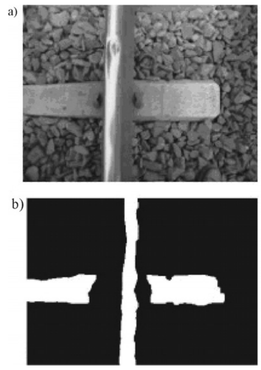

Fig. 3. Results of segmentation: a) original image, b) image after segmentation
Rys. 3. Efekt segmentacji: a) obraz oryginalny, b) obraz po segmentacji

3. Determination of salient features of fastening element

After restricting the area where fastening elements are searched, the method allowing for choice of the best salient features describing a fastening element has to be determined. Detection of the fastening element is performed through moving a rectangular window of size approximately equal to that of the fastening element in the area restricted to the sleeper. Recognition of the fastening element is based on its view, thus it seems to be reasonable to use the wavelet transform as a tool for extraction of salient features of the fastening element. It allows for reduction of the size of this rectangular window at preservation of the most salient features of the fastening element. The wavelet transform decomposes the analyzed function into finite lasting components Ψ(t) called wavelets [5]. The continuous-time wavelet transform (CWT) of function $f(t)$ generates CWT coefficients $W(a,b)$ according to the formula:

$$W(a,b) = \int_{-\infty}^{\infty} f(t) * \Psi_{a,b}(t)dt \tag{3}$$

where:

$$W_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \tag{4}$$

is a wavelet function of time scale (dilation) equal to a and time shift (translation) described by b. For a =1 and b = 0, $\Psi(t)$ is called the mother wavelet. If the wavelet function is only chosen for a=2k and b=2kn where k and n are integer number, then formula (3) describes the discrete wavelet transform (DWT) of original function $f(t)$. The wavelet transformation is reversible and the original function can be reconstructed on a basis of the values of $W(a,b)$ coefficients. DWT can split the original function $f(t)$ into two parts: $f_0(t)$ corresponding to coarser approximation of $f(t)$ and $g_0(t)$ corresponding to the high frequency detail function, defined as the difference between $f(t)$ and its approximated version. The approximated version $f_0(t)$ can be further split into two parts - coarse approximation $f_1(t)$ and the detail part $g_1(t)$. This process can be continuously performed up to the assumed level. The splitting process is often called decomposition level. For instance the first splitting process generating approximation function $f_0(t)$ corresponds to the first decomposition level, the second one generating approximation function $f_1(t)$ corresponds to the second decomposition level and so on. In practice the DWT is calculated with use of two orthogonal (or biorthogonal) low-pass l(n) and high-pass h(n) filters [6]. Different wavelets correspond to appropriate filter forms. The low-pass and high-pass filtering is performed step by step producing, respectively, the coarse approximation of the function and the coefficients containing high frequency details. In case of 2-dimensional image in each step the wavelet decomposition is performed twice: first on the rows of the image and then on its columns, thus 2-dimensional (2D) DWT generates 4 sub-images for each decomposition level. Fig. 4 shows the process of generating 4 sub-images for the first decomposition level.

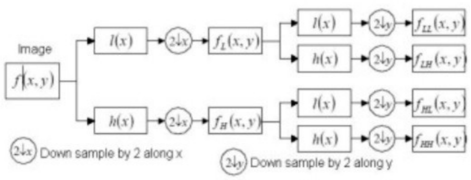

Fig. 4. Illustration of first level decomposition performed by 2D DWT
Rys. 4. Zobrazowanie pierwszego poziomu dekompozycji dla 2 wymiarowej DWT

Thanks to down sampling by 2 in x and y axes, each next decomposition level generates sub-images of size of approximately 4 times smaller than that of the image generated in the previous decomposition level. The first level decomposition of Fig. 4 generates $f_{LL}(x,y)$, $f_{LH}(x,y)$, $f_{HL}(x,y)$ and $f_{HH}(x,y)$ corresponding to sooth sub-image, horizontal details sub-image, vertical details sub-image and diagonal details sub-image, respectively.

The authors took into consideration several families of wavelets such as Haar, Daublets and Coif. Different decomposition levels were performed for each wavelet. The best results were obtained for the Coif1 wavelet and smooth sub-image generated at the second decomposition level. Fig. 5a) shows the rectangular window with the fastening element – the window size is equal to 45x37 and Fig. 5b) presents its corresponding smooth sub-image obtained at the second decomposition level- its size was reduced to 15x13. Fig. 5c) and Fig. 5d) present the rectangular window with a broken fastening element and its corresponding smooth sub-image at the second decomposition level, respectively.

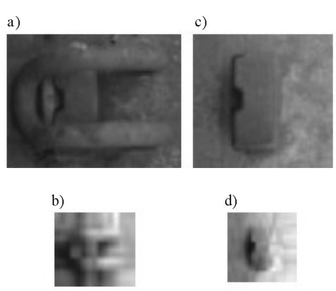

Fig. 5. Original images of fastening elements and corresponding smooth sub-images
Rys. 5. Obrazy elementów mocujących wraz z odpowiadającymi podobrazami

4. Detection of fastening element

After generating salient features describing a fastening element effectively, it is necessary to determine if the calculated smooth sub-image presents or not the fastening element. This process can be regarded as the classification- distinction between two groups: the first comprising the fastening element presence and the second comprising the fastening element absence. There are many classifiers, for example decision trees, neural networks, statistic classifiers. The authors decided to use the neural network classifier. A neural network is a common tool used to design classifiers. Its main advantage is generalization [7]. It is the ability of a neural network to infer information about the whole group on a basis of the limited number of its representatives [6]. This inference is performed during teaching phase of the network. We chose a hybrid neural network [7]. This network needs much less training data to proper operation. Additionally, the teaching process is much faster than that of other networks. A hybrid network is the fusion of the Kohonen network and Multi-layer Perceptron (MLP). They are connected serially, the Kohhonen network output is fed to the MLP input. The teaching process is performed separately, first for the Kohonen network and after its completion, the MLP network is taught. During teaching process of Kohonen network, smooth sub-images (each containing 15x13 = 195 wavelet coefficients) are given to its input neurons (the number of input neurons equals the number of wavelet coefficients - 195). After completion, the MLP network is taught. During this phase signals from the output neurons of the Kohonen network are given to MLP input neurons. Additionally, destination signals in the form of the vector of two elements are given to the MLP output neurons (the number of MLP output neurons equals 2). The destination signals correspond to the desired output values that should be generated by the network. For example if a smooth sub-image given to the input neurons of Kohonen's corresponds to fastening element then the Kohonen network generates the corresponding output signal on its output neurons. This signals is given to the input neurons of the MLP network which, in turn, should generate the output signal in the form of [0 1]. If a smooth sub-image does not contain the fastening element, the output neurons of the MLP network should generate the signal [1 0]. The hybrid network was taught on 100 images with fastening elements and 100 images without fastening elements. The classifier effectiveness of classifier was checked on 30 images with fastening elements and 30 images without fastening elements, not taking part in the teaching process. The system misclassified 6 out of 60 images. It yields the system effectiveness on the level of 90%.

5. Conclusion

The presented system is able to detect absence or presence of the fastening element on a basis of the sleeper image along with its neighborhood . The system effectiveness was equal to 90%. Thanks to application of a block restricting the area of searching the fastening element to the sleeper, the system effectiveness could be increased. In the further stage the authors will try to merge the proposed system and the visual system diagnosing the state of wooden sleepers presented in [3]. It seems this fusion can be a significant alternative for the traditional method used in Polish Railway Line, that is manual inspection by a trained staff.

6. References

[1] Bovik A.: Handbook of image and video processing. Academic PressSan Diego, 2000.
[2] Bojarczak P., Lesiak P.: Application of neural networks into automatic visual diagnostic of railway wooden sleepers. Międzynarodowa Konferencja Naukowa Transport XXI wieku, Białowieża 2010. Logistyka 4/2010 (płyta CD).
[3] Bojarczak P., Lesiak P.: Zastosowanie hybrydowej sieci neuronowej do klasyfikacji uszkodzeń drewnianych podkładów kolejowych, Prace Naukowe Politechniki Warszawskiej, Transport, zeszyt 78, 2011.
[4] Bishop C: Neural networks for pattern recognition, Oxford University Press, 1996.
[5] Daubechies I: Ten lectures on wavelets, SIAM Press, 1988.
[6] Mallat S: A theory for multiresolution signal decomposition:the wavelet representation, IEEE Trans PAMI, 1989 vol.11.
[7] Haykin S: Neural networks a comprehensive foundation, Prentice Hall Press, 1996.

otrzymano / received: 14.09.2011
przyjęto do druku / accepted: 02.11.2011 *artykuł recenzowany*