



## MONITORING OF RAILROAD PARTS FOR THE PRESENCE OF AN OBJECTS ON THE RAILS

Petr Krug<sup>1</sup>, Andrey Ostroukh<sup>2</sup>, Tatiana Morozova<sup>3</sup>, Evgeniy Kashkin<sup>3</sup> and Irina Ivanova<sup>3</sup>

<sup>1</sup>National Research University, Moscow Power Engineering Institute, Moscow, Russia

<sup>2</sup>Department of Automated Control Systems, State Technical University-MADI, Leningradsky Prospect, Moscow, Russia

<sup>3</sup>Moscow State Institute of Radio Engineering, Electronics and Automation, Moscow, Russia

E-Mail: [ostroukh@mail.ru](mailto:ostroukh@mail.ru)

### ABSTRACT

The problems of wavelet-based data and image fusion design for on-board control equipment of the locomotive are highlighted. Development of a new algorithmic supply for the locomotive on-board hardware to process images obtained in different spectrum sub bands to create a fused image of high quality is required. It allows the algorithms for object detection and recognition to be efficiently applied to the fused image. Paper describes some our methods to the fusion of multispectral images, which are degraded not only by bad weather conditions, but also the internal noise of the sensors. The results of modeling and comparison are presented.

**Keywords:** identification of barriers, locators, infrared scanning devices, wavelet-transform, image fusion.

### INTRODUCTION

The last decades are characterized by increasing railroad traffic worldwide. Many countries with developed railroad infrastructures are compelled to implement different systems of functional requirements for monitoring and controlling train movements to provide increased safety. For example, the American Railway Engineering and Maintenance-of-Way Association (AREMA) describes Positive Train Control (PTC).

The same situation is in Russian Federation where the government payloads a lot of efforts and does constantly modify the state standards of railroad safety. One of modifications lays in the conception also known as movement authorities. According to that, the train receives information about its location and where it is allowed to safely travel. The equipment on board the locomotive must continually calculate the train's current speed relative to a speed target some distance away governed by a braking curve. If the train risks not being able to slow to the speed target given the braking curve, the brakes are automatically applied and the train is immediately slowed. The speed targets are updated by information regarding fixed and dynamic speed limits determined by the track profile and signaling system.

Most current implementations also use the speed control unit to store a database of track profiles attached to some sort of navigation system. The unit keeps track of the train's position along the rail line and automatically enforces any speed restrictions as well as the maximum authorized speed. Temporary speed restrictions can be updated before the train departs its terminal or via wireless data links. The track data can also be used to calculate braking curves based on the grade profile. The navigation system can use fixed track beacons or differential GPS stations combined with wheel rotation to accurately determine the train's location on the line accurately within a few feet.

Unfortunately, these technical and organization efforts are not sufficient today. A train traveling at high speed has a tremendous amount of inertia which must be overcome in order to quickly stop it. A quick stop for a train may take up to a half kilometer. Moreover, many barriers and crossings are situated such that a train's driver has little, if any, reaction time to stop a train if he/she is able to recognize barriers as a vehicle, person or animal on rails. Or, due to darkness, rain, or fog, the train may actually be at the crossing before an object is recognized in the crossing. In this case, it is too late to avoid an accident. And, even if a train can be significantly slowed before reaching the barriers, it still may have enough momentum to damage or destroy anything in its path. In addition to the damage or harm possibly inflicted on those outside the train, occupants of the train may also be injured. Crossing accidents often cause some of the cars of the train to derail with people in those cars being injured. People in other railroad cars are often shaken or jolted by the sudden deceleration which takes place during emergency braking resulting in injuries.

Thus, the safety of people due to damage resulting from a collision or sudden stoppage of the train "pushes" to invent new methods of the locomotive speed control. The one way out is to expand railroad safety systems with video detection hardware for viewing railroad barriers and crossings to determine if the barrier does exist or the crossing is clear or whether vehicles, persons, animals or other objects are on the rails. Further, the apparatus creates an alarm signal to a train's driver notifying him/her if there is any barrier or, one or more of these objects are identified as being in the crossing. This is done so the train can be halted prior to its reaching the crossing, thereby preventing injury to a person or animal, or damage to the vehicle, train, or other object in the crossing.



In accordance with the generally stated situation, the proposed hardware is provided for monitoring a railroad path and crossings for the presence of an object on the rails. To determine if the objects (people, animals, vehicles, other objects) are of a sufficient size to cause damage to the train a television (video) camera and other electro-optical devices (sensors) working in infrared and ultrasound bandwidths have to be installed onto the locomotive cabin. An image processor processes the images obtained to form a fused image of high quality, and then analyzes the barrier appearance. An alarm unit is responsive to an output from the processor to provide an alarm signal to the train's driver as to the presence of the object on the rails. An alarm on the train is sounded in response to the alarm signal obtained from the on-board equipment. In response to the alarm signal, the train may be halted so that it stops short of the barrier or the crossing. This would prevent injury to people in the crossing, people on the train, the train itself, or vehicles or other objects in the crossing.

## MODELS AND METHODS

The method of extracting coherent structures [1] using the library of local trigonometric bases and wavelet packets has been successfully applied to recovering the corrupted musical records. For the sake of simplicity, this scheme was modified for fusion of 2D signals where noisy images form the library and an estimator of the correlation coefficient is computed with only the one wavelet basis chosen by the user before. To avoid artifacts in the fused image a procedure known as the Shift Invariant Discrete Wavelet Transform (SI-DWT) [2] is applied.

Let  $Y = X + Z$  be a noised original image  $X$  containing  $I$  pixels which is undergone to be decomposed in several levels by WT with the given wavelet basis.

Noise  $Z$  in  $Y$ , with variance  $\sigma_z^2$  can be determined as components of the decomposed image  $w_{\xi_k}$ ,  $k=1, \dots, I$ , not having a strong correlation with the given basis. If  $I$  is relatively large then for any orthonormal wavelet basis there is existing the probability up to 1 that [3]

$$\frac{\max_{1 \leq k \leq I} |w_{\xi_k}|}{\|Z\|} \leq \frac{\sqrt{2 \ln I} \sigma_z}{\sqrt{I} \sigma_z} = \frac{\sqrt{2 \ln I}}{\sqrt{I}} = \rho_l \quad (1)$$

Here is assumed that  $Z$  is normal (Gaussian) process (white noise) with zero mean. The parameter  $\rho_l$  can be considered as the high bounded estimator of the correlation coefficient for correlation between Gaussian noise and any given wavelet basis, and as it is seen from (1) this estimator does not depend on noise variance. It has been experimentally proven by the first author that the theoretical value of  $\rho_l$  is also the highest asymptote of the correlation coefficients for correlation between Daubechies wavelets, symmlets, few biorthogonal wavelets and noises

having different probability density functions (pdf) and their combinations, such as log-normal, exponential, Gamma-distributed, etc.

On the other hand, the correlation coefficient for correlation between an input image and the given wavelet basis can be computed using wavelet coefficients  $w_{Y_i}$ ,  $1 \leq i \leq I$ , of the observed image  $Y$ :

$$\rho(Y) = \frac{\sup_{1 \leq i \leq I} |w_{Y_i}|}{\|Y\|} \quad (2)$$

The correlation coefficient (2) is computed on the data of the noised image and it can help to choose the best basis from the library of bases only. In order to extract useful components from the noised image it is necessary to have more flexible scheme. It is known from the theory of wavelets [4] that the nonlinear approximation of any signal being decomposed by wavelet transform is referred as inverse wavelet transform  $W^{-1}$  for first (significant)  $M$  wavelet coefficients after their sorting:

$$\hat{X} = \sum_{k=1}^M W^{-1} \{w_{Y_k}\}, \quad (3)$$

where

$$|w_{Y_k}| \geq |w_{Y_{k+1}}|, \quad \forall k \in [1, \dots, I] \quad (4)$$

Hence, a coherent structure of the input image is determined according to the rule [3]:

$$\rho(Y_k) = \frac{|w_{Y_k}|}{\|Y\|} > \rho_l, \quad 1 \leq k \leq M \quad (5)$$

Having any value of  $M$  the rest image  $Y_M$  can be calculated as the difference between the input image and the pseudo image obtained by means of the inverse wavelet transform for the coherent structures:

$$Y_M = Y - \sum_{k=1}^M W^{-1} \{w_{Y_k}\} = \sum_{k=M+1}^I W^{-1} \{w_{Y_k}\} \quad (6)$$

The rest image  $Y_M$  cannot be recognized as a noise if the next inequality is true:



$$\rho^2(Y_M) = \frac{|w_{Y_M}|^2}{\sum_{k=M+1}^L |w_{Y_k}|^2} > \rho_{I-M}^2 \quad (7)$$

The estimator  $\hat{X}$  of the original image is the sum of  $M$  coherent structures (3). It is worth to mark that pursuit of the coherent structures looks like hard thresholding of the wavelet coefficients with the threshold [3]:

$$\tau = \rho_{I-M} \sqrt{\sum_{k=M+1}^L |w_{Y_k}|^2} \quad (8)$$

It is assumed that there are  $L$  noisy images of the same scene registered by sensors operating in different spectral bandwidths. They form the library of the input images  $\Lambda = \{Y_1, Y_2, \dots, Y_L\}$ . In order not to have a routine procedure each image should be decomposed with the one wavelet-basis chosen by the end-user from the library of wavelet bases. Choice of the suitable wavelet basis can be done by different methods based on cost functions, e.g. using the estimator (2)  $\rho(Y^i)$ ,  $i=1, \dots, L$ , of the coefficient correlation between the input image  $Y_i$ ,  $i=1, \dots, L$ , and any wavelet basis from the library of bases. Because the input images contain one and the same scene there are cases when one and the same wavelet basis can be chosen for few images. It usually happens if the quantity of images in the library  $\Lambda$  is relatively small, e.g.  $L=2$ ,  $L=3$ . and/or images are registered from the sensors working in tight bandwidths without wide guard gaps. In spite, our experience has shown that extracting of the coherent structures is stable in such cases.

The input images have low resolutions because of sensors' noises. When the properties of the noises are unknown the pursuit of the coherent structures can be considered as thresholding of wavelet coefficients with the threshold (8) the value of that is changing adaptively. The main idea of the proposed fusion scheme is to keep up those of coherent structures of images which are the best under some criterion and the result fused image is the sum of the saved coherent structures (3).

To pursuit the best coherent structure of the input images such coherent structure is chosen corresponding to the image with number  $\alpha$  in the library which is minimizing the cost function:

$$C(Y_\alpha) = \min_{i=1, \dots, L} C(Y_i), \quad i=1, \dots, L, \quad (9)$$

where

$$C(Y_i) = \sum_{i=1}^L \Phi \left( \frac{|w_{Y_i}^{(i)}|^2}{\|Y^i\|^2} \right). \quad (10)$$

As it is seen from (9) the choice of the best coherent structure depends on the kind of the function  $\Phi$ . We choose the entropy function:

$$\Phi(u) = -u \ln u, \quad u \geq 0. \quad (11)$$

During analysis of noisy images it is impossible to exactly indicate which image along with its components is dominating in typical fusion schemes. The proposed fusion scheme can be iteratively applied to all of images from the library if we put  $Y_M = Y$  for the next step. It allows to consequently correct the fused image because the best coherent structure can be extracted from any rest image among the rest images on the current iteration.

The proposed fusion algorithm contains the next steps.

To form the library of multispectral images where all or some of them are corrupted by specific noises of sensors.

To choose the best wavelet basis for each image under some criterion by the end-user.

To put  $Y'_M = Y^i$ ,  $i=1, \dots, L$ .

To perform SI-DWT for each image from the library with the given wavelet bases and to save arrays of wavelet coefficients.

To sort the wavelet coefficients in accordance with (4).

To pursuit coherent structures of the images using (7) and to find the value of  $M$ .

To restore the nonlinear approximations  $\hat{X}$  and the rest images  $Y'^M$  by the inverse wavelet transform.

To compute the values of the cost function (10), to choose using (9) and keep up the best coherent structure  $\hat{X}^\alpha$  in memory.

To correct the saved image calculated on the previous iteration with the new nonlinear approximation  $\hat{X}^\alpha$ .

If  $M=0$  then to stop, the fused image has been already cumulated in memory and can be displayed; otherwise go to the step 3.

## RESULTS AND DISCUSSIONS

At this moment large amount of computer modeling data proves the effectiveness of the suggested fusion schemes. Images registered by video cameras, locators with microwave illumination, infrared scanning devices have been used. Some pictures were obtained in different neighboring bandwidths. Visual observations showed that there is significant improvement of the fused

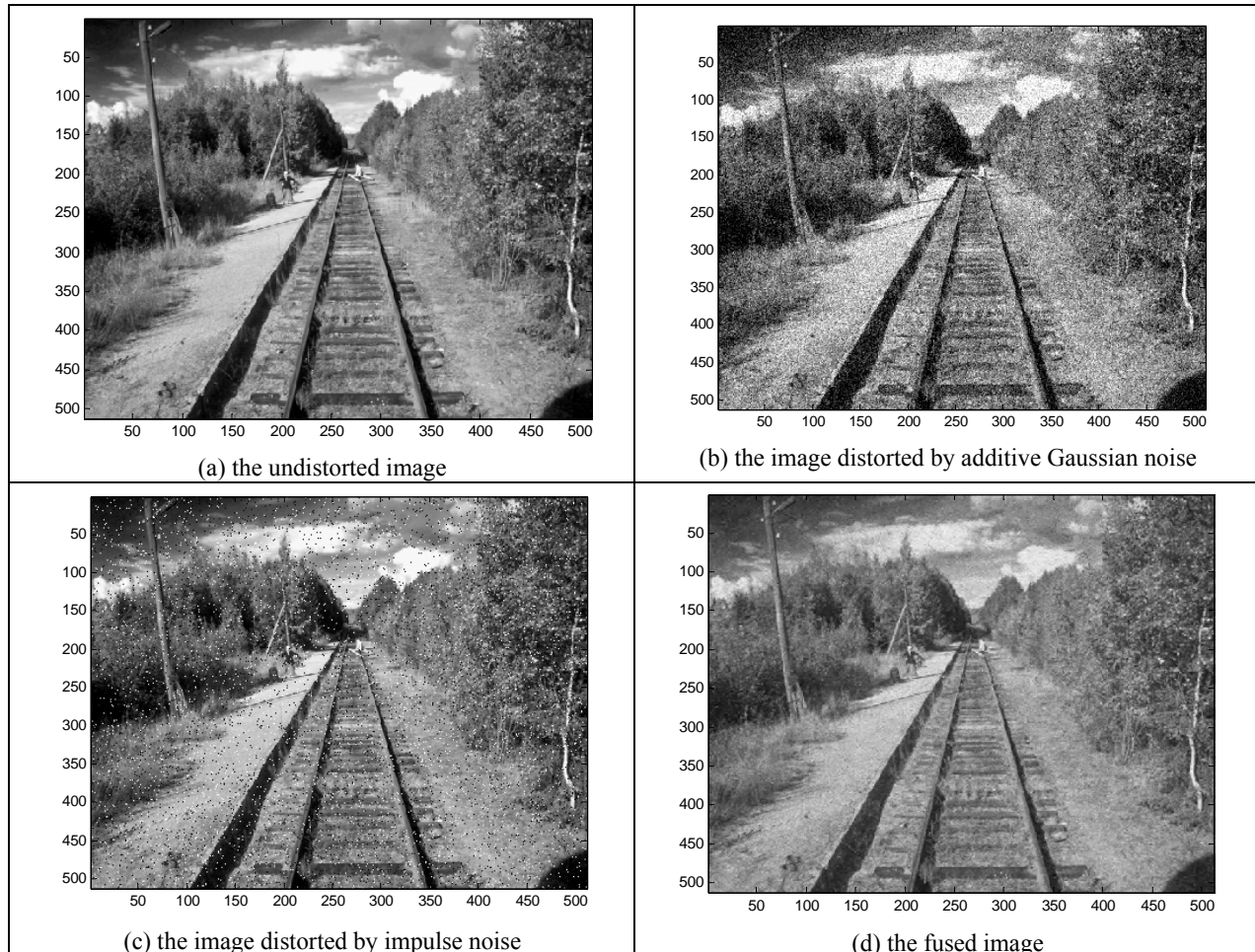




image quality in comparing with the pre-filtering and well-known fusion schemes.

Figure-1 contains the results of simulation modeling for the typical railroad scene. A person crossing

the rails is practically invisible for a locomotive driver under additive Gaussian distortions (Figure-1, b). Also, it is difficult to make a decision in the case of impulse noise shown on Figure-1, c.

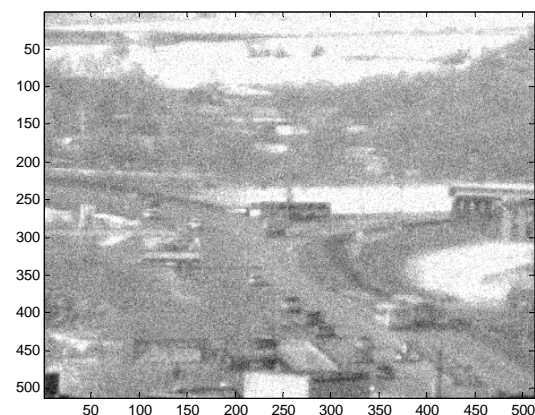


**Figure-1.** The results of modeling: (a) the undistorted image, (b) the image distorted by additive Gaussian noise; (c) the image distorted by impulse noise; (d) the fused image.

The situation is much worse for the algorithms which make decisions regarding train speed decreasing and/or stoppage automatically. The fused image allows both the locomotive driver and embedded alarm software to make the right decision.

Figures 2, 3, 4 demonstrate the effectiveness of the suggested fusion algorithms when the normalized cross-correlation between video and infrared images is about 0.5, i.e. there are two different images of the one and the same scene. It happens when cameras or sensors operating in different spectral subbands have different spatial resolutions.

One can see from the figures and Table-1 that the proposed method provides better image enhancement both visually and numerically under the PSNR and SSIM criteria.



**Figure-2.** The noised video image.

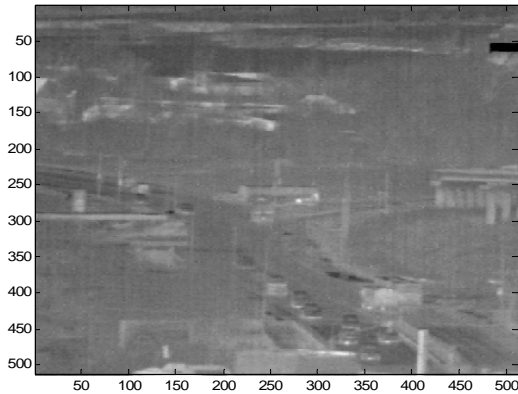


Figure-3. The infrared image.



Figure-4. The fused image.

Table-1. Results of comparison.

Image	PSNR, dB	SSIM
Corrupted by additive Gaussian noise with $\sigma_z^2=34$	18.1	0.7045
Corrupted by multiplicative noise with exponential pdf, $\sigma_z^2=56$	16.4	0.4215
Fused by averaging	17.8	0.5633
Fused by maximum-selection fusion rule	18.5	0.5547
Fused by maximum-selection fusion rule with preliminary hard thresholding	19.4	0.7322
Fused by the selected suggested fusion rule	26.5	0.8945

## CONCLUSIONS

Aggregation of the proposed fusion methods within the locomotive hardware is suggested to increase the effectiveness of the following processing algorithms for object detection and recognition.

The choice of the fusion algorithm depends on the current conditions of viewing. Experimental results have shown that the first fusion scheme works perfectly well under daylight, and when the distance between the head of the train and the barrier is relatively large, for example, when the train moves on direct part (without turns) of the railroad path. Otherwise, the second fusion algorithm works better under bad weather conditions. At the current moment, the suitable method is being explored to automatically differ situations when and which of the proposed fusion algorithms to apply.

## ACKNOWLEDGEMENTS

The work was carried out with financial support of the Ministry of Education and Science of the Russian Federation in the framework of the Agreement № 14.574.21.0047, 19.06.2014.

## REFERENCES

- [1] J. Berger, R. Coifman and M. Goldberg. 1994. Removing noise from music using local trigonometric bases and wavelet packets. J. Audio Eng. Sci. 42(10): 808-818.
- [2] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez. 2005. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. IEEE Transactions on Geoscience and Remote Sensing, 43 (10): 2376-2385.
- [3] S. Mallat, Z. Zhang. 1993. Matching pursuits with time-frequency dictionaries. IEEE Transactions on Signal Processing. 41(12): 3397-3415.
- [4] R. A. DeVore. 1998. Nonlinear approximation. Acta Numerica. pp. 51-150.
- [5] O. Rockinger, T. Fechner. 1998. Pixel-level image fusion: The case of image sequences. SPIE Proceeding. 3374: 378-388.



- [6] Image Fusion: Algorithms and Applications. 2008. Ed. Tania Stathaki. Academic Press. p. 519.
- [7] H. Li, B. S. Manjunath, S. K. Mitra. 1995. Multisensor Image Fusion using the Wavelet Transform, Graphical Models and Image Processing. 57. pp. 235-245.
- [8] P. Burt, R. Kolczynski. 1993. Enhanced image capture through fusion. Proceedings of 4<sup>th</sup> Intl. Conference on Computer Vision, Berlin. pp. 173-182.
- [9] J. Shapiro. 1993. Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans. on Signal Processing. 41: 3445-3462.
- [10] A. Said and W. A. Pearlman. 1996. A new fast and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans Circ. and Syst. Video Tech. vol. 6.
- [11] Cheng Peng; Chan A. 2001. Speckle noise removal in SAR image based on SOT structure in wavelet domain. IEEE Transactions on Geoscience and Remote Sensing. 7, pp. 3039-3041.
- [12] Bekhtin Y. S., Bryantsev A. A. 2011. Matching Pursuit of Coherent Structures during Wavelet-based Fusion of Noisy Multispectral Images. Proceedings of 11<sup>th</sup> International Conference on Pattern Recognition and Information Processing (PRIP'2011), May 18-20, Minsk, Belarus.
- [13] Bekhtin Y. S., Bryantsev A. A., Malebo D. P. 2013. Wavelet-based fusion of noisy multispectral images using Spatial Oriented Trees, Proceedings of 2nd Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 15-20 June, pp. 113-116, DOI 10.1109/MECO.2013.6601332.
- [14] A. Ostroukh, V. Nikonov, I. Ivanova, T. Morozova, V. Strakhov. 2014. Distributed System of Real Time Head Gesture Recognition in Development of Contactless Interfaces. Middle East Journal of Scientific Research. 20 (12): 2177-2183. DOI: 10.5829/idosi.mejsr.2014.20.12.21105.
- [15] Ostroukh, V. Nikonov, I. Ivanova, T. Morozova, K. Sumkin, D. Akimov. 2014. Development of Contactless Integrated Interface of Complex Production Lines. Journal of Artificial Intelligence (JAI). 7(1): 1-12. DOI: 10.3923/jai.2014.1.12.
- [16] A. V. Ostroukh, S.A. Vasuhova, M.N. Krasnyanskiy, A. Samarutunga. 2011. Study the prospects and problems of integration of human-computer: artificial intelligence, robotics, technological singularity and virtual reality. Prospects of Science. No 4 (19): 109-112.