Composite Real-Time Image Processing for Railways Track Profile Measurement

Cesare Alippi[1], Ettore Casagrande[2], Fabio Scotti[1], Vincenzo Piuri[1]

[1]Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy
[2]Tecnogamma, Italy

Abstract

Checking railway status is critical to guarantee high operating safety, proper maintenance schedule, low maintenance and operating costs. This operation consists of the analysis of the rail profile and level as well as overall geometry and ondulation. Traditional detection systems are based on mechanical devices in contact with the track. Innovative approaches are based on laser scanning and image analysis. This paper presents an efficient composite technique for track profile extraction with real-time image processing. High throughput is obtained by algorithmic pre-filtering to restrict the image area containing the track profile, while high accuracy is achieved by neural reconstruction of the profile itself.

1. Introduction

Safety in railways and tramways is one of the key issues of public transportation companies. The state of the tracks is relevant in this perspective, in particular when high-speed trains are envisioned. Frequent monitoring of the tracks is therefore critical to plan proper and cost-effective maintenance. Detection of wear and deformation of tracks at an early stage allows for better scheduling of the maintenance, avoiding the need of immediate action when dangerous conditions are observed. Advance maintenance planning reduces also costs since the limited human and equipment resources can be better used. Besides, accurate maintenance decreases the acoustic pollution due to bad coupling between wheel and track: this is relevant especially within the town borders.

To detect the track profile by means of tactile techniques mechanical devices in contact with the track are traditionally used. The main characteristics of the profile are observed indirectly through the analysis of the position of suited leverages. This approach has several practical problems. Accuracy and completeness of profile reconstruction may be not accurate since the contact point between the mechanical sensor and the track may be large and restricted to a specific area of the track. Besides, components are subject to wear and difficulties in passing joint point areas. On the other hand, acceptable operation quality needs dedicated rail carriages running at low speed. The above characteristics induce high costs of acquisition and operation of the track monitoring system. Moreover, only part of the track parameters (namely, geometry and ondulation) can be observed, while the most important ones (namely, profile and level of the wear of the track) cannot be measured.

An experimental system available on the market was realized by using laser technologies to replace the tactile sensors. Relevance of this technique is wear avoidance due to lack of physical contact between moving components. This solution detects only the track surface in 8 points with an insufficient accuracy for early global detection of incipient deformations. Besides, detection of complex profiles (e.g., grooved track) is not allowed and a mechanical truing system is required to align lasers.

We considered an innovative approach based on image analysis and processing to reconstruct the whole track profile. The image is generated by lighting the track with a laser beam and acquired by a CCD camera. Since no contact between the monitoring system and the tracks is required, no wear occurs and the speed of the rail carriage can be higher. Carriage speed is limited only by the real-time processing ability of the monitoring system. Due to the amount of information to be processed, a high-performance architecture is needed for real-time analysis since it is not possible to store all images and process them off line. Pipelining and parallelism allow higher performance when very high operating speed is required. Differences between reconstructed and reference profiles point out the track deformations. High-level image analysis avoids for continuous and accurate alignment of the monitoring system with the track since the image processing method can be design to be self-aligning. Some companies performed partial experiments similar to ours, but none was reported to be satisfactory.

In this paper we present the image processing system and the composite technique for real-time profile analysis. The composite approach consists of an algorithmic pre-processing to identify the strip - in the whole image - in which the track profile lays and a neural processing for fine profile reconstruction within such a strip. The system was tested on still images.

0-7803-5276-9/99/$10.00 © 1999 IEEE

2. The composite detection system

The detection system consists of a laser source, whose beam is collimated by a suited optic lens into a light plane, two 512x512-pixel CCD cameras for complete optimum observation of the track, a digital processing system per camera, and a supervision system (Fig. 1). Each digital processing system performs real-time profile filtering and extraction by using a composite approach from images of the corresponding CCD camera. The supervision system collects the partial views of the track, reconstructs the whole profile, identifies and stores the deformed profiles. Real-time operation is needed since 200 track sections per second must be captured and processed to guarantee a sufficient accuracy of deformation localization. The detection system is conceived for on-board operation on a regular train, even if – at the moment - was tested with still carriages only.

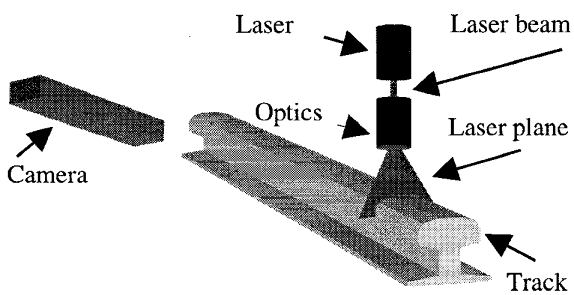

Figure 1: The detection system.

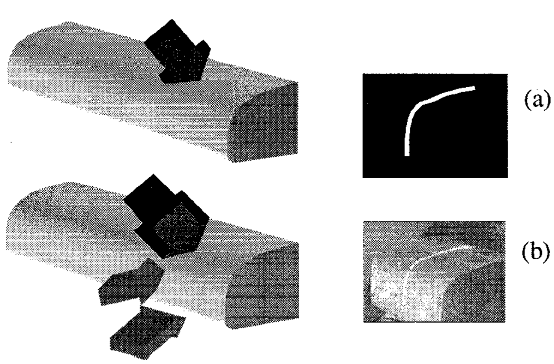

Figure 2: The CCD images: (a) ideal, (b) real.

Identification of the profile is made complex by the presence of noise and environmental disturbances that modify the ideal profile reflection to be observed by the camera (Fig. 2a). Real images (e.g., Fig. 2b and 3) are affected by environmental light, multiple reflections, track oxidation, greasy track, speckle effect due to track roughness, non-infinitesimal thickness of the laser plane scanning the track, optic aberrations, CCD sensor saturation, and image distortions due to vibrations. Note that the roughly approximate position of the profile in the image is known a priori since the laser and the cameras are still with the rail carriage. Besides, the profile is approximately laying in a linear direction, i.e., cutting the image in stripes only one point of the profile belongs to each stripe. This characteristic allows for parallel processing since each stripe can be analyzed independently to reach 10ms image processing time without affecting the profile accuracy.

In each column of the image localizing the position of the track profile means to find the position of the maximum laser reflection intensity. In the ideal case the intensity distribution along the column is Gaussian. Localizing the maximum implies therefore detect the position of the expected Gaussian profile with the maximum likelihood.

To tackle this application, we tested both traditional filtering techniques with minimum-square approximation and neural network techniques. In the first case results were quite poor due to the inability of capturing all non-linearities and distorsions. In the second one the number of pixels to be processed in each column and the variety of the possible maximum light profile position leaded to large inaccurate networks, that are also difficult to train.

It is worth noting that highly approximate localization of the area of interest in each image is quite trivial for the human observer, even without experience (see Fig. 2 and 3). Track profile localization does not need to take into account all details in the whole column, but only the area around the maximum lighting. Experiments have shown that no information out of a 40-pixel strip centered approximately on the maximum lighting is necessary for accurate reconstruction of the track profile. Besides, this area of interest corresponds approximately to the zone around the highest-intensity Gaussian profile in the column. Such area can be easily found by identifying the maximum correlation of the light profile with the Gaussian reference: correlation can be effectively used. Finer localization of the maximum must deals with all non-linearities presented above, which are difficult to be captured algorithmically while are easy described by examples. In the literature, neural networks were proved effective for this kind of tasks.

Our approach is therefore *composite* since exploits the best features and performance of both of these techniques within their individual application limits. Algorithmic filtering by correlation is used to center the attention on the 40-pixel strip (Fig. 4a), while the neural network perform the fine track profile localization at subpixel accuracy (Fig. 4b).

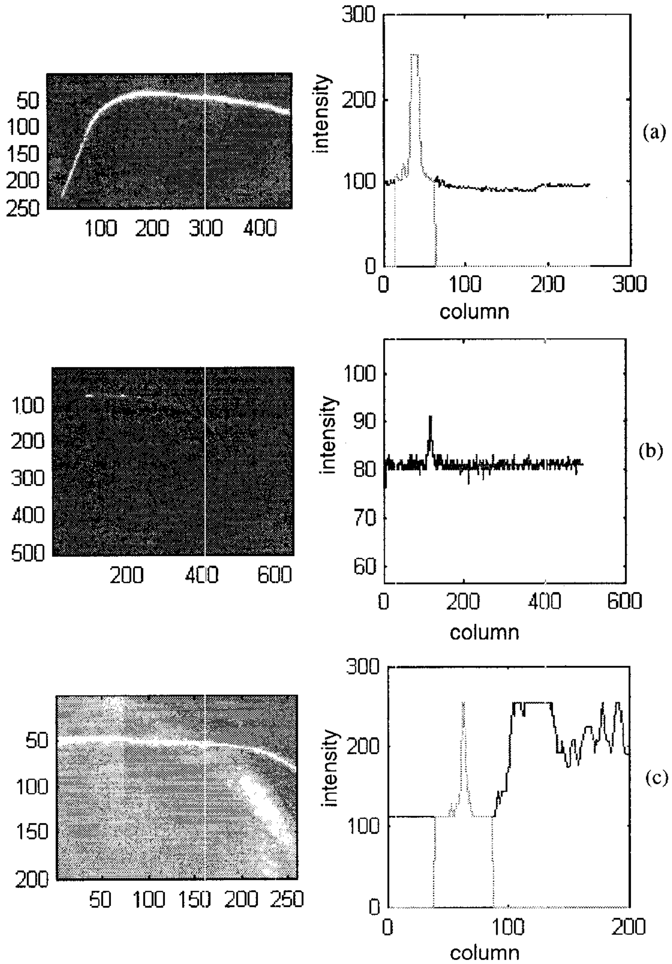

Figure 3: Typical disturbances: (a) saturation, (b) low intensity, (c) environmental reflections; the observed images are on the left, typical light intensities along image columns are on the right.

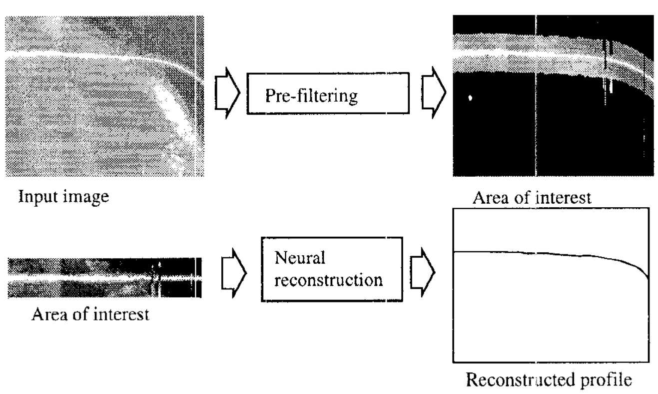

Figure 4: The composite processing approach: (a) algorithmic filtering, (b) neural processing.

3. The algorithmic pre-filtering

Identification of the region containing the track profile drawn by laser reflection within an image column is obtained by convolving the pixel intensity with a Gaussian distribution [1]. The convolution is repeated by positioning the maximum value of the Gaussian profile in each pixel of the column. The maximum value of the convolution corresponds to the position in which the light intensity is more similar to the expected Gaussian distribution. To reduce the computational complexity the convolution can be performed every few pixel positions instead of every pixel.

Accuracy of the identification of the area of interest may be reduced by the presence of noise in the input image. Since the laser reflection has usually intensity definitely greater than noise, the Gaussian profile is likely understandable. Problems are actually due to modification in the Gaussian amplitude (related inaccuracies in the lens and focusing as well as to the real thickness of the laser plane), to track reflectivity variations leading to CCD saturation, and to system oscillations due to rail carriage motion. External sources of errors are the possible reflections of environmental lights. In Fig. 5, some typical light distributions are shown.

The column analysis may have difficulties to discriminate a Gaussian distribution from a saturation border. An effective solution consists of applying the convolution to the derivative both of the light intensity distribution and the Gaussian function. Even in the presence of strong external light leading to saturation, this approach maximizes the correlation in correspondence of the Gaussian laser reflection only (Fig. 6). This approach is successful also if the Gaussian profile is very near to the saturation region.

4. The neural profile reconstruction

The fine profile reconstruction is obtained by a more accurate analysis of the area of interest identified by the algorithmic pre-filtering. In each image column the neural approach identifies the position of the maximum of the Gaussian distribution in the 40-pixel strip by minimizing the difference between the theoretical Gaussian and the actual profiles. Separating the pre-filtering phase from the fine neural positioning allows for separating the accuracy of reconstruction in the neural

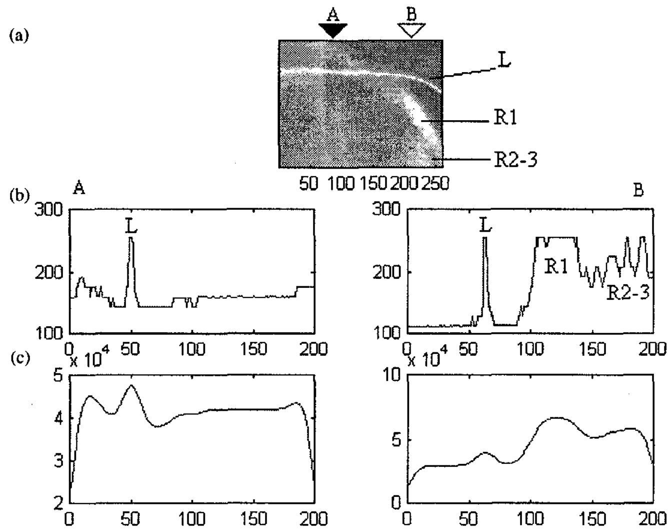

Figure 5: The algorithmic pre-filtering: (a) the image, (b) the intensity distribution in columns A (typical) and B (with saturation), (c) the convolution values along these columns.

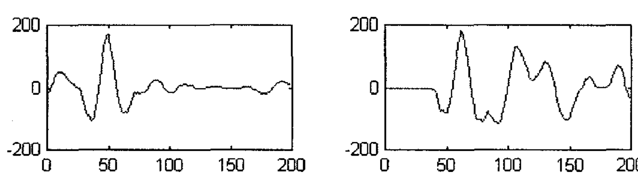

Figure 6: The derivative pre-filtering: the convolution values along the columns.

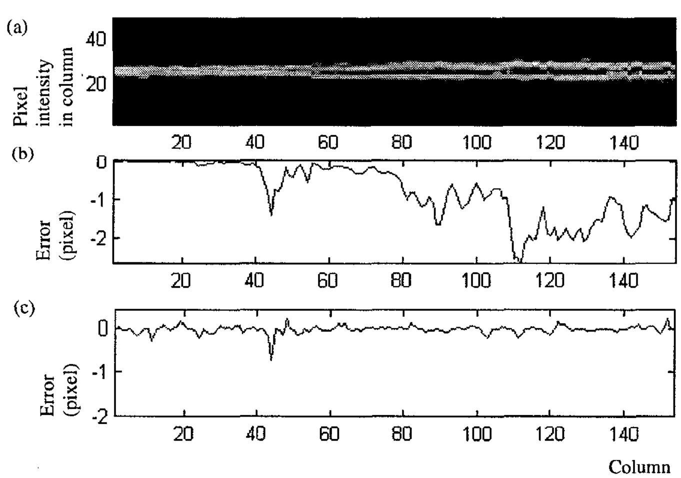

Figure 7: Learning error: (a) the training image, (b) non-uniform sampling, (c) uniform sampling.

computation from the accuracy in windowing the area of interest. The profile is in fact reconstructed by adding the very accurate distance (measured with sub-pixel accuracy) of the maximum value from the bottom of the analyzed column in the strip to the distance (measured in pixels) of the bottom itself from the base of the whole image. The overall accuracy is therefore related only to the accuracy of the neural reconstruction. The accuracy of the pre-filtering (typically about 2 pixels in our approach) is useful to center the Gaussian profile approximately in the middle of the strip so that the neural reconstruction can focus its abilities mainly on the central area of the strip to achieve very high accuracy efficiently.

The neural network that was shown effective for the envisioned application is the Radial Basis Function (RBF) network [2]. This kind of networks is suited for interpolating multi-variable functions. A RBF network has a Three-layered feed-forward topology. Input neurons are used to distribute the input values to all subsequent neurons. Each hidden neuron generates its output by applying a radial function (typically a Gaussian function) to the difference between the input vector and the centers' vector. The output neuron computes the weighted sum of the hidden neurons' outputs, possibly with a threshold. The number of hidden neurons and the centers can be determined from the analysis of the data available for training. A minimum-square algorithm is used to identify the weights.

In our application the input layer is composed of 40 inputs corresponding to the 40 pixels of the image strip. Each input value is the intensity of the light collected by the corresponding CCD pixel. Experimentally 6 neurons were shown sufficient for the hidden layer to achieve the desired accuracy.

To achieve an error goal equal to 10^{-3}, the training set was typically composed by 50 input vectors. The training set must contain enough examples to allow the network to capture the desired behavior. We do not need to use all columns in an image since the light reflection change gradually along the truck profile. The generalization ability allows the network to operate correctly even for reflections never previously seen. Conversely, if we sample the columns only in one part of the image, reconstruction will be accurate only in the portion learnt during training and may become

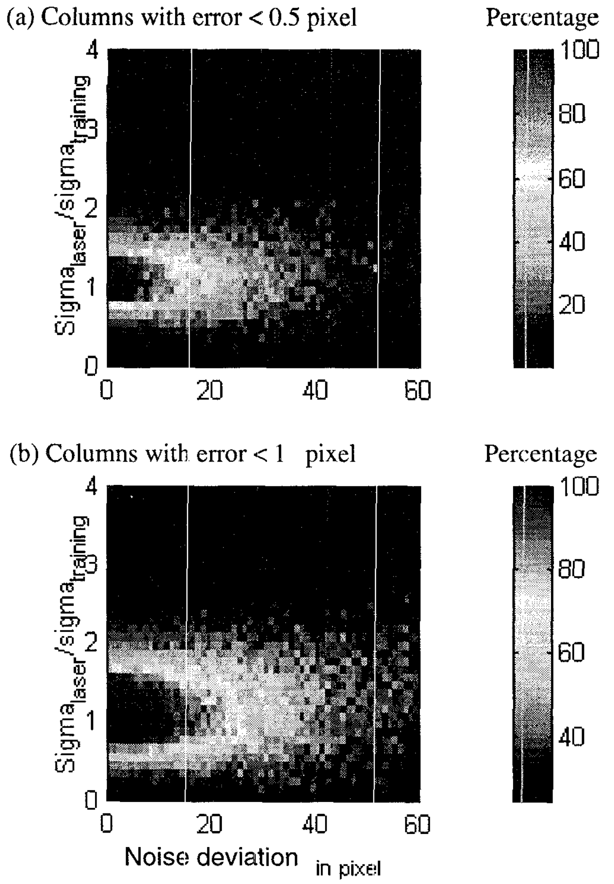

Figure 8: The percentage of successful profile identification with an error less than (a) 0.5 and (b) 1 pixel. The ratio $\sigma_{laser}/\sigma_{training}$ measures the variation of the light width in the real case with respect to the training examples.

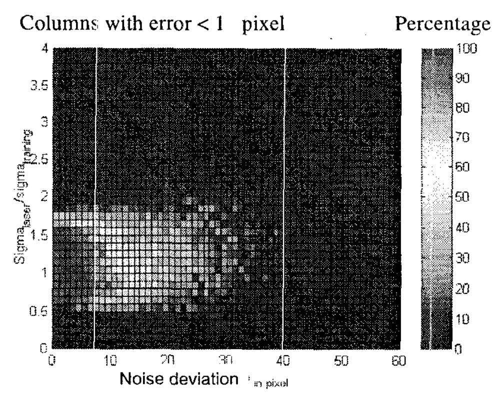

Figure 9: Hardening the training set for laser reflection width.

very poor elsewhere (Figs. 7a and 7b). If samples are uniformly taken along the profile, the network achieves a uniform high accuracy (Fig. 7c): the standard deviation is 0.1 pixel.

The robustness of the neural reconstruction with respect to variation of the reflection is critical for high accuracy. For example, as shown in Fig. 7, 40% variation of the reflection width makes the accuracy jump from 0.1 to over 2 pixels. To create a robust network, we studied the effects of varying the application parameters (e.g., shape of the reflection, saturation index, and noise) on synthetic images. Fig. 8 reports the percentage ability of successful identification of the profile with the specified accuracy. In the absence of noise, the deformation of the light width that the network is able to tolerate even if not present in the training set ranges between 0.8 and 1.4, i.e., the network tolerates better width enlargement. The behavior is similar also in the presence of noise. To balance this behavior, we suggest to introduce some reflections with width slightly smaller (e.g., 0.9) than the expected one in the training set to enhance the generalization ability symmetrically.

Adding noise to the training set decreases the generalization ability of the network. The network recognizes well only the vectors very similar to the ones used in training. To deal with noise and accuracy contemporaneously, the training set must be created in a different way.

To harden the network with respect to variation of the reflection shape, for each expected position of the profile we include several sample vectors corresponding to Gaussian profiles having different width. Since the non-saturated laser reflection is 5-pixel width, we added reflections 4- and 6-pixel wide. To avoid unnecessary generation of large networks, the training set is created without changing the total number of vectors: we uniformly extract samples from the set composed by profiles having all considered widths. Fig. 9 shows that this approach enlarges the correct recognition region: from 4 to 6 pixels the correct recognition is now at least 90% even in the presence of small noise. The network is also able to identify profiles with deformation index equal to 1.5, i.e., a Gaussian reflection profile 7.5-pixel wide even if it was never seen. This behavior holds for a noise deviation up to 4 pixels, which is relevant even in the real images. Higher generalization ability and noise immunity without increasing (sometimes even

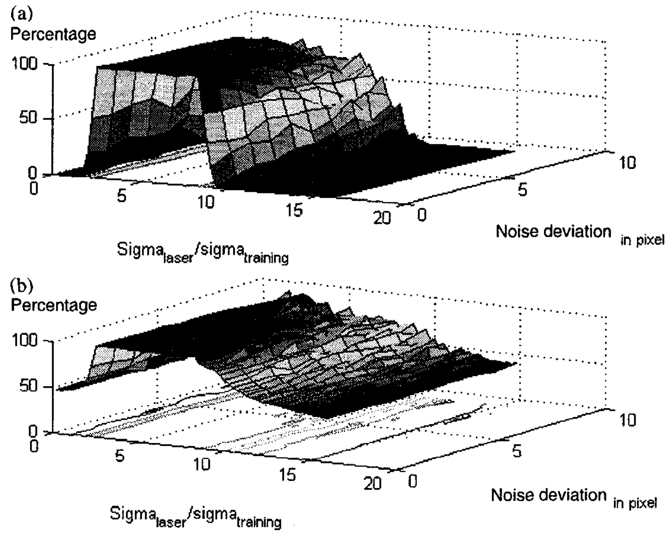

Figure 10: Successful identification with error < 0.5 pixel by selecting widths every (a) 0.5 and (b) 0.1 pixels in the range from 4 to 5 pixels.

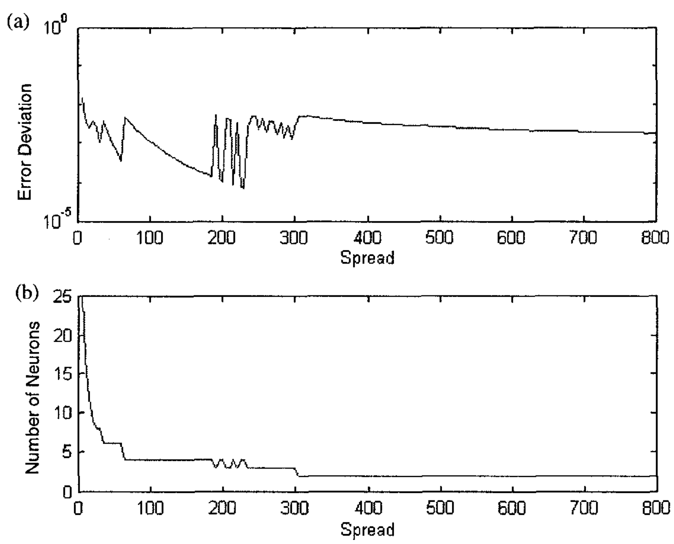

Figure 11: Generalization ability vs. spread: (a) the error deviation, (b) the number of neurons required to perform identification.

decreasing) the number of neurons can be achieved with the approach described above by increasing the number of different widths. The effectiveness of this constructive technique for the training set can be also observed in Fig. 10 as higher insensitivity to laser width variation and to noise is concerned.

The generalization ability and the accuracy are also affected by the spread of the radial function, i.e., of the width at half height. In Fig. 11 the error deviation and the number of neurons are reported for different values of spread, in the case of 40 noise-free validation vectors applied at the end of learning. The first relative minimum value of the error deviation is at 180, while the absolute minimum is at 230. Even if they are not so different for noise-free input vectors, the error deviation obtained in the case of noisy inputs having 5-pixel deviation becomes 0.198 and 0.204 pixel, respectively. Therefore, the maximum generalization does not coincides with the absolute minimum error deviation vs. the spread. We experimentally observed that the optimum value corresponds to the minimum reached without discontinuities (i.e., 180 in the example).

5. Conclusions

An innovative approach to track profile measurement is presented. A real-time image-processing-based technique was adopted to reconstruct and measure the profile by analyzing a laser-scanned CCD-camera image. A prototype of the detection technology was tested for more than one year in the Milan underground, while the reconstruction technique was verified on simulated and real images. An accuracy of the same magnitude of the track roughness was achieved with a still monitoring system: typical resolution is 25μm. Similar results are also expected for system on board of a moving rail carriage. Simulations were performed to mimic light reflections and the damped small low-frequency oscillations that are typical of the moving carriages: results are still attractive and show the efficiency and the effectiveness of the proposed approach. A more extensive on-field experimentation is required to verify the real effects and interference of external light conditions, reflection, and vibrations as well as to certify the accuracy of the measurement system.

Acknowledgements

Authors thank Azienda Trasporti Milanesi SpA for having allowed in-field experiments in the Milan underground.

References

[1] A.V. Oppenheim, *Signals and Systems*, Prentice-Hall, 1990

[2] J.A. Anderson, *An Introduction to Neural Networks*, The MIT Press Cambridge, 1995