

Real time classification of rail defects

M. Nitti, C. Mandriota, E. Stella, A. Distante

Istituto Elaborazione Segnali ed Immagini – C.N.R., Bari, Italy

Abstract

In the last years the detection and classification of surface defects of material is assuming great importance. Visual inspection can help to increase the product quality and, in particular context, the maintenance of products. The railway infrastructure is a particular field in which the periodical surface inspection of rolling plane can help an operator to prevent critical situation. A defect on rolling surface appears generally as a grey level variation useful for its classification. Main idea is to utilize the image processing to help a human operator in the detection of defects on the rolling surface. The prototype realised uses two Dalsa line scanner camera SP-12 to acquire the left and right rail image with a sampling rate of 2 mm per line. An encoder connected to the axel box with 2 mm resolution generates the line acquisition trigger for the cameras. The left and right images are processed to extract the rolling surface strip by image and to classify defects. We use neural network to tracking the rolling surface in the image. This method is able to track the rail also in the switch, cross level, and so on. The detection of defect uses a gradient oriented approach to emphasis the image regions with grey level variation. Four directions 0° , 45° , 90° and 135° will be considered as defect principal direction. The union of all four normalised histograms is used as input sample for a neural network classifier. A test phase has been performed on a real trolley.

1 Introduction

The maintenance of railway infrastructure can be done using different systems based on x-Ray, ultrasound or with image processing as we have done.

Actually the system based on image processing has some limitations such as rail surface localization and the algorithms used to detect the defect. So that the light and camera positioning are important to normalize the rail image.

The camera and light have been positioned in a way to obtain a non saturation rail image with the rail surface that is a part of railway. Furthermore the image contains the sleepers and all attach part of rail to the sleepers.

Developing a system based on image processing requires the solution of different tasks summarized as follows:

- Railway tracking to extract the rolling plane surface
- Detection of defects

The solution of these tasks resolves only the algorithmic approach to the problem. These algorithms, finally, will be translate in a programming language and tested to measure the obtained performance.

In this work we describe the solutions used for each tasks and the development of prototype used to test the methods listed above.

2 Rail tracking

Founding the exact rail position in the image is the first job that the system must do. The precision obtained is very important as the subsequent will work with rolling plane extracted in the image.

A rail tracking system must consider several factors: the rail can appear with different forms (UIC 50, UIC 60 and so on); the rail illumination can change; the defects present on the rail surface can modify the rail geometry; the principal rail must be followed also in presence of switch .

To satisfy all requirements we have tested different approaches:

- Singular Value Decomposition with threshold
- Singular Value Decomposition with neural network classifier
- Gradient approach with neural network to implement tracking with close loop ring.

All approaches have provided good precision, but generally the Singular Value Decomposition methods require much CPU time. The approach based on gradient has provided good performance requiring few CPU time. The main idea is based on close loop ring system.

Let $I(x, y)$ be a grey level image of railway of dimension $[N, M]$ containing a rail oriented in the $\vec{Y}$ direction. Let I_w be a sub-window with dimension $[N_w, M_w]$ placed on the image and containing the rail. The M_w dimension must be three times the dimension of the rail surface along $\vec{X}$ (e.g. N_w).

Convolving I_w with the mask $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$ we obtain the gradient image $G_{\vec{X}} = I_w * \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$ and calculating the mean value along the $\vec{Y}$ direction

we obtain a vector $M(x) = \text{mean}(G_X(x, y)), x = 0, \dots, M_W$. In the figure 1 we show an example of the processing described above.

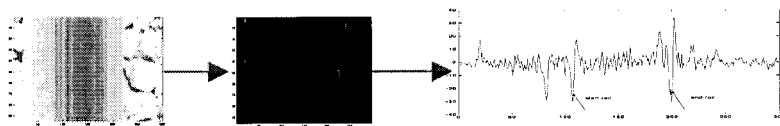


Figure 1: From left to right image: rail image, gradient image, vector $M(x)$

Vector $M(x)$ provides a synthetic representation of rail position using only M_W elements. A neural network K is created with M_W input neurons, few hidden neurons and one output neuron with linear transfer functions has been created. It has been trained with the well known back-propagation algorithm with momentum. The network is trained to give 0 if the two peaks representing the contour of rail are located exactly to the middle of $M(x)$; -1 if they are located on the right side of $M(x)$ middle and 1 if they are located to the left side of $M(x)$ middle. The network output provides a method to control the exact localization of rail in the middle of $M(x)$, therefore for centering the sub-image I_W on the rail head is sufficient to control that the network output is next to zero.

To test the method we have extracted manually by a long image strip the exact rail position. This vector samples the rail position every 95 lines considering a frame of 512x95 pixel. The training sample has been created extracting sequentially the patches and processing it in accordance with the following flow diagram:

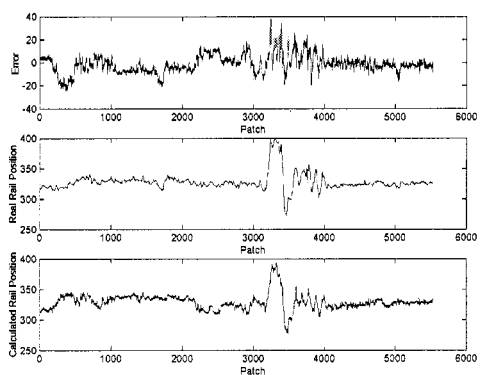
1. Calculate vector $M(x)$ and add the network output to the current rail position until the network output is near zero.
2. If the rail position founded is located on the left/right of real rail position extract some random vectors $M(x)$ from the left/right of rail to reinforce the network (The $M(x)$ with rail in central position with output network zero is always present).
3. Retrain the network. Set the learning rate to 0.001, the momentum to 0.2 and execute 1000 epochs with goal set to 0.001 MSE.
4. Go to step 1 until there is a rail image.

The first test has been done by using a new patch of 512x95 pixels. The centre of rail is located manually and the tracking starts. In each step the image is shifted of one pixel to left and the network output is used to locate the new rail position. In this way for each rail position a location error is calculated. We have obtained a mean error location of -2.1448 pixels with a variance of 0.3676 pixels.

The second test consists in tracking the same railway used to learn the network. The mean error obtained is -2.31 pixel for a variance of 7.8 pixel. The

visual check of methods has proved a good tracking of rail providing always a window of 95x95 pixels on the rail head.

The graphs of error, real rail position and calculated rail position are shown in figure 2. We can observe some situations where the error is too high as well as the normal situations where the error is near to zero. We must consider that



the railway used to test the system presents some anomalies. The important thing is that the window centred on the rolling plane must be always on the rolling plane. This is a crucial aspects for solving the defect recognition problem checking the patch with superimposed window centred on the rolling plane. We have obtained good performance.

Figure 2: Error, real rail position and calculated position in a test

3 Detection of defects

Rail defects have different forms and can be classified by operator specialized in different way. In Italy they are classified in relation to defect localization and type of defect [1]. In the following sections we consider only the defect present

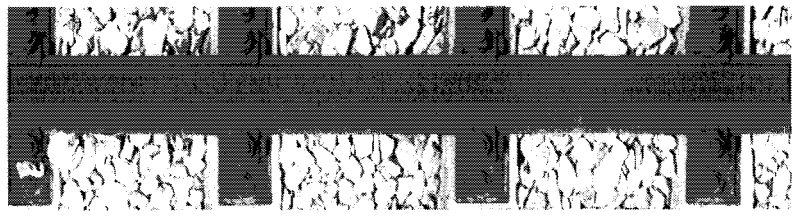
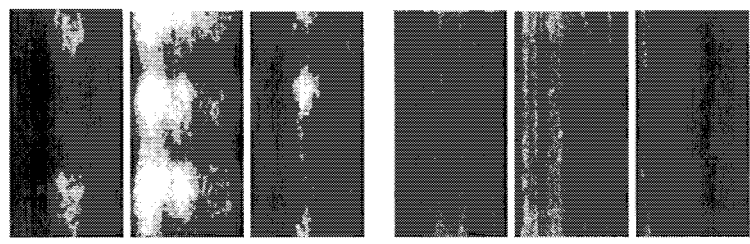


Figure 3: Rail Image of 512x2048 pixels

on the rolling surface, in particular on the rail head. This defect can present a privilege (as cracking, welding, shelling) or none (corrugation, blob, spot).



Corrugation Rail

Good Rail

Figure 4: Example of defect and no defect (128x64 pixels)

Figure 3 shows a rail image extracted from a long strip image of 512 pixel width acquired with the system described below. Figure 4 shows a set of corrugation and good rail images.

To perform the defect recognizing task a bank of Gabor filters has been used to emphasize the defect frequency. Moreover a network able to separate the “good rail” class from “defected rail” class has been considered.

In spatial domain 2D complex Gabor filter is given by[2]:

$$f(x, y) = \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right\} \cdot \cos(2\pi\mu_0(x \cos \theta + y \sin \theta))$$

where μ_0 represents the filter frequency and $\sigma_x = \sigma_y$ the amplitude of Gaussian envelope along the x and y directions. The Gabor filters can be used to emphasize all frequency components having θ directions with frequency μ_0 and pass band that is function of σ_x and σ_y .

For our experiments we have chosen four Gabor filters with $\sigma_x = \sigma_y = 2.0$, $\mu_0 = \sqrt{2}/8$ and we have fixed orientation θ equal to 0, $\pi/4$, $\pi/2$ and $3\pi/4$ as can be seen in figure 5.

We define A as the set of rail surface images and I as any image of A that can contain defect or no. The image I is convolved using the four Gabor filters obtaining four different images $I_0, I_{\pi/4}, I_{\pi/2}, I_{3\pi/4}$ where:

$$I_0 = I * \text{KerneK}Gabor\left(\sqrt{2}/8, 0, 2.0\right)$$

$$I_{\pi/4} = I * \text{KerneK}Gabor\left(\sqrt{2}/8, \pi/4, 2.0\right)$$

$$I_{\pi/2} = I * \text{KerneK}Gabor\left(\sqrt{2}/8, \pi/2, 2.0\right)$$

$$I_{3\pi/4} = I * \text{KerneK}Gabor\left(\sqrt{2}/8, 3\pi/4, 2.0\right)$$

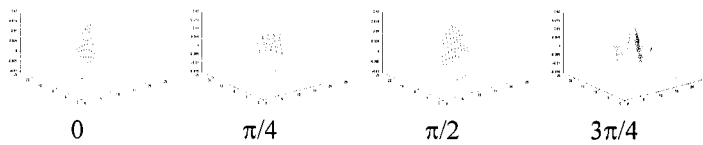


Figure 5: Gabor Filters example

Each filtered image is scaled between [0 1] considering Min_ImageGabor and Max_ImageGabor equal to:

$$\text{Min_ImageGabor} = \min_{I \in A} \left(I_0, I_{\pi/4}, I_{\pi/2}, I_{3\pi/4} \right)$$

and

$$\text{Max_ImageGabor} = \max_{I \in A} \left(I_0, I_{\pi/4}, I_{\pi/2}, I_{3\pi/4} \right)$$

so scaled image is given by:

$$I_x^{\text{scaled}} = (I_x - \text{Min_ImageGabor}) / (\text{Max_ImageGabor} - \text{Min_ImageGabor})$$

For each I_x^{scaled} a vector histogram extracted using by K points uniformly distributed between 0 and 1. The union of histograms forms the feature vector of 4*K elements used for the classification task.

A Self Organized Map network [3] is used to classify the feature vector using 400 rail surface image of 128x64 pixels with 200 images representing a defect and 200 images representing a good rail. K is equal to 30 and the feature vector has 120 components.

We have tested different SOM [4] networks changing the number of hidden nodes in order to choose the best network. We have obtained the best performance with a hidden grid SOM of (5x14) as can be seen in table 1.

Nodes in the Hidden layer	5 x 2)	5 x 4)	5 x 6)	5 x 8)	5 x 10)	5 x 12)	5 x 14)	5 x 16)	5 x 18)	5 x 20)
Misclassification errors	9	1	5	5	4	5	2	1	3	3

Table 1 - relationship between grid dimension and misclassification error

Table 2 shows a grid of (5 x 8) where white box refers to good rail cluster and black box to a corrugation rail cluster. Grey box indicates a misclassification box. The number written in the box represents the number of time that the node has won over 400 test images. A white text indicates winning corrugation image and black text indicates winning good rail. Only 4 misclassification errors over 400 samples have been obtained.

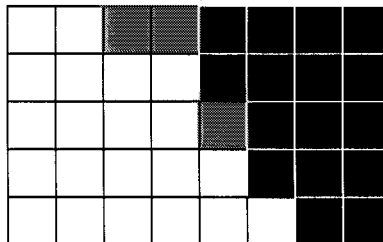


Table 2 – In white box good rail; in black box corrugation rail; in gray box misclassification

We have tested the performance using “Leave One Out Procedure” [6]. Leave One Out Error is obtained by training the SOM network using 399 elements and calculating the network response on the other one. The test provides a generalization error of 2 % (8 errors / 400 image). Different tests have shown the ambiguity of some images.

4 Prototype development

The prototype development has required the following project steps:

- camera definition – type of camera, lens and position
- frame grabber
- light definition – type of light and position
- mechanical parts
- Software production

In the following section we describe how we define each single part of the prototype.

4.1 Camera and frame buffer definition,

We have selected a DALSA line scanner camera model SP-12 to acquire the rail image considering the rail movement with respect to the train.. The resolution is 2048 pixel/line and at each trigger signal pulse the camera acquires an image line. A VLDS bus channel at 30 MHz provides the communication channel to the frame grabber. It is possible to set the gain (1x, 2x, 4x and 8x) and the integration time. The camera requires a F-mount lens, but can be modified for a C-mounts lens. The pixel dimension is $14\mu\text{m} \times 14\mu\text{m}$ for a total CCD lengths of 28.672 mm. The maximum line rate for this camera is 14100 line/sec and believe 52800 line/sec with 512 pixel resolution. The lens used has a focal distance of 6 mm positioning the camera at 860 mm over the rolling plane. Consequently the pixel resolution on the image plane of 2x2 mm. With this resolution and 512 pixel/line camera the maximum train speed is 380 Km/h. The maximum integration time is in relation to the maximum line scan frequency. With a maximum of 14100 line/sec, the integration time can vary in a range of $[0,70 \mu\text{s}]$, but considering the pixel resolution imposed, the time integration range is between $[0,20 \mu\text{s}]$ with a maximum train movement of 0.5611 mm during the exposition time.

The DALSA camera requires a trigger pulse to acquire a line. An optical encoder mounted on the axel box generates a trigger pulse every 2 mm. Considering a wheel diameter of 860 mm, we obtain that is necessary an encoder with 1350 pulse per revolution.

The trigger signal is connected to the frame grabber that controls the camera. The frame grabber chosen is the ITI PC-DIG that can receive directly the trigger signal pulse from the wheel encoder and generates consequently the EXSYNK signals to DALSA SP-12. All transferred cycle is synchronized by MCLK of PC-DIG clock at 30 MHz and the STROBE signal of DALSA has the

same phase of MCLK. For more information about the PC-DIG see the ITI PC-DIG manual.

4.2 Light definition

The light system is used to guarantee a uniform and invariant illumination level in relation to the particular problem. The strip of rail interested to the line scanner acquisition has the dimension of 1024 x 2 mm, but considering the rail-bogie-coach movement we can increase the dimension to 1024 x 100 mm. A time invariant lumen level is required, and this constraint limits the light source to DC power supply. Using the AC power supply, for example 220 VAC 50 Hz we observe the flicker phenomena. Furthermore, the lumen emitted from the object observed must be sufficient to impress the CCD. The resistance to the shock and vibration of lamps must be adequate to the application.

We have tried some solution using DC power supply lamps for automotive and films application discarding the situation where the lumen level was insufficient or where the bulbs had poor resistance to vibrations.

After these tests we have chosen a commercial lamp of OSRAM model 41850 FL. It is characterized by 12 V DC power supply; 100 W power consumption; an angle beam of 24° degrees and 54100 total lux distributed on a surface of 200 mm diameter with a distance from lamps of 500 mm.

The configuration proposed uses eight lamps in series realizing a light strip. Each lamp has a diameter of 130 mm for a total length of 1040 mm that is sufficient to illuminate the image. The row of lamps is positioned at a distance of 390 mm from the rolling plane with an angle of 30~35 degrees in the camera direction to reduce the light reflection into the camera. In this situation the light surface has a diameter of 200 mm with overlap. The test done has provided good results.

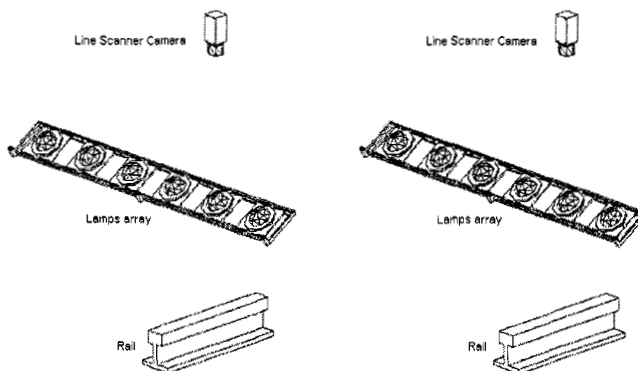


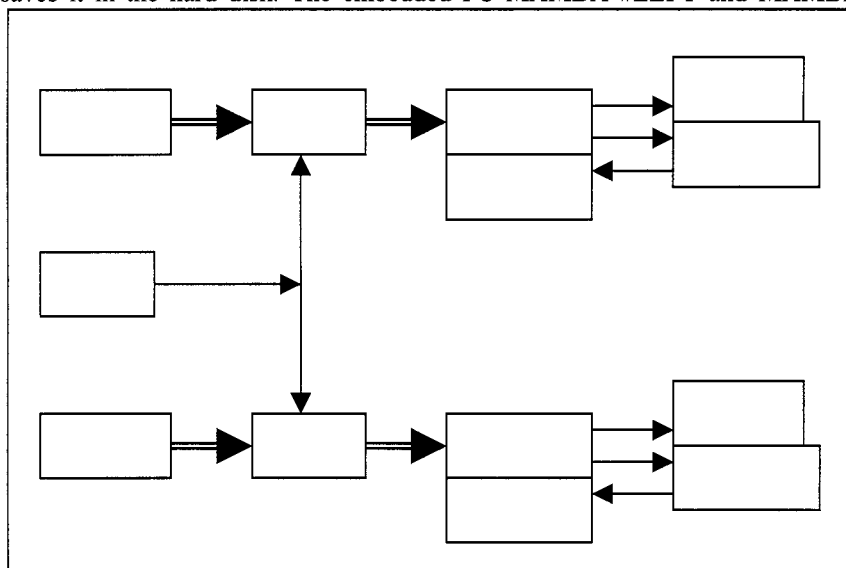
Figure 6: left and right equipment

4.3 Hardware architecture

The hardware architecture provides the necessary equipment to perform the acquisition and processing of rail image. We analyse the hardware architecture for only one rail. For the second rail it is the same.

The block diagram of the system is described below.

The encoder mounted on the axel box provides the trigger TTL signal to each PC-DIG card to acquire a line from the line scanner camera left and right. In this way the perfect synchronization between the left and right frame is ensured. The image is transferred via DMA to PC #LEFT and PC #RIGHT that saves it in the hard disk. The embedded PC MAMBA #LEFT and MAMBA



#RIGHT are used in the off-line phase to process the image contained in the hard disks. The results are display through the monitors. The MAMBA architecture used is the MAMBA 100 with PENTIUM II 1 GHz.

4.4 Software architecture

The development of software architecture is made in Visual C++ version 6.00. We have used the Application Program Interface for PC-DIG called IFC version 5.2 and for MAMBA 100 called SAPERA. The program realizes two principal tasks: an acquisition task and a processing task. Only the PC is used in the acquisition task. The images provided by the PC-DIG are recorded in the hard disk in a raw format. This format consists in a binary sequence of 512 pixel per each line. The maximum number of line per file is 1331200. This number depends from the total number of bytes recorded for a CD-ROM. When a raw file has a dimension of 650 MB, the software closes it and opens another raw file. In this way it is always possible to export the images acquired on CDs.

In the off-line processing, the software loads from the hard disk all data necessary to execute a reliable tracking and to detect the rail defect. It reads the image file, localize the rail and detects the presence of a defect. The test effectuated have provided good results.

5 Future work

This works is a part of a project oriented to obtain a prototype able to detect autonomously the rail defect located on the head of rail. The training set used in this paper is only a little part of all possible defects. We are working on a new approach to create an exhaustive training set. The best solution is to use an unsupervised learning network able to classify an infinite training set where the occurrence of class samples is different. For example on a railway of 100 Km, we hope that only 1% of rail presents defect. During this continuous learning phase, a human operator reinterprets the result obtained by the Network by using his own experience to link the class with the rail quality.

Our future work will explore adaptive unsupervised learning as Multiple Self-Organizing Maps to obtain better results.

Acknowledgements

This work was supported by district Puglia with *PROGETTO POP PUGLIA 1994-1999*, Sottoprogramma 7, Misura 7.4.1 and title “*Sistema di riconoscimento real-time dei difetti del binario mediante elaborazione delle immagini*”

References

- [1] M. Nitti, E. Stella – “*Sistema di Riconoscimento Real Time dei Difetti del Binario Mediante Elaborazione delle Immagini*”, Progetto Pop Puglia 1994-1999, 24/07/2000, prot. 160/2000.
- [2] Thomas P. Weldon, William E. Higgins, Dennis F. Dunn – “*Efficient Gabor Filter Design For Texture Segmentation*”, Pattern Recognition, Vol. 29, No. 12, pp.2005 – 2015, 1996
- [3] C. Distante, P. Siciliano, K. C. Persaud – “*Dynamic Cluster Recognition with multiple Self-Organizing Maps*”
- [4] M. Nitti, C.Mandriota, E. Stella, C. Distante – “*Rail defect classification by adaptive self organized map*” – SPIE International Symposium on Intelligent Systems and Advanced Manufacturing – Novembre 2001 – Boston – USA [4572-58]