AUTOMATED RAILWAY SIGNS DETECTION. PRELIMINARY RESULTS

Marco Guerrieri [1], Giuseppe Parla [2]

[1,2] *Polytechnic School, University of Palermo*
90100, Palermo, Italy
[1] *marco.guerrieri@unipa.it*
[2] *giuseppe.parla@unipa.it*

Nowadays safety in railways is mostly achieved by automated system technologies such as ERTMS/ETCS. Nevertheless, on local railways (suburban and regional lines) several tasks still depend on the choices and actions of a human crew. With the aim to improve safety in such type of railways, this research proposes a system for the automatic detection and recognition of railway signs by means of the digital image processing technique. First field applications, carried out on the Italian railway network, show that the proposed system is very accurate (the percentage of correctly detected railway signs is about 97%), even at high train speeds.

Keywords: Railway safety, digital image processing technique, railway sign detection and recognition

1. Introduction

The railway safety in Europe, evaluated in terms of fatal train collisions and derailments per billion train-kilometres, has gradually improved since 1990, with a reduction in the accident rate of 4% per year (ERA, 2016).

The estimated average number of fatal train collisions and derailments per billion train-kilometres was about 4.8 in 1990 and 1.1 in 2015 (ERA, 2016).

Despite the encouraging trend, in 2014 the European member states reported a total of 2,076 accidents resulting in 1,054 killed and 819 seriously injured casualties.

These data show an increase of 5% in the number of "significant accidents" (Directive 2004/49/EC) compared to 2013. The fatality risk for an average train passenger is about 0.12 fatalities per billion kilometres (ERA, 2016).

Even with this positive datum, railway safety is a relevant issue especially for the suburban and regional lines that are not equipped with modern safety technologies, such as the European Rail Traffic Management System/European Train Control System (ERTMS/ETCS).

Human factors play a key role in the above-mentioned types of railways and, although indirectly, they can affect railway capacity and delays (Nagy and Csiszár, 2015). For example, the fixed block should have a minimum length (L) greater than the stopping sight distance (s) (Sostarics and Zobory, 1981) that is a function of the train driver perception-reaction time:

$$L \geq s(v_{max}) = t_p \cdot v_{max} + \frac{v_{max}^2}{2 \cdot a}, \qquad (1)$$

in which: v_{max} is the maximum train speed in the line under analysis, t_p is the train driver perception-reaction time and "a" is the train deceleration.

Traditionally, in Italy the following values have been adopted for a conventional railway network: $v_{max} = 160$ km/h, $a = 0.82$ m/s^2 $t_p = 3$ s.

Assuming a reasonable margin of safety, from Eq. (1) we obtain L = 1,350 m.

Under the hypothesis of traffic flow stationarity, the railway line capacity both in case of fixed block system C_f and moving block system C_m can be calculated as follows (Delfino and Galaverna, 2003):

$$C_f = \frac{3600}{\Delta t} = \frac{3600 \cdot v}{L + L* + l_t + f} , \qquad (2)$$

$$C_m = \frac{3600 \cdot v}{1.5 \cdot (t_p \cdot v + \frac{v^2}{2 \cdot a}) + l_t + f} , \qquad (3)$$

with:

- ΔT = minimum headway between trains [s];
- v = train speed [m/s];
- $L*$ = L for $V \le 160$ km/h ($L*$ = 1,350 m);
- $L*$ = 2L for 160 km/h < $V \le 200$ km/h ($L*$ = 2,700 m);
- $L*$=4L for 200 km/h < $V \le 250$ km/h ($L*$ = 5,400 m);
- L_t = train length [m];
- f = margin of safety [m].

Figure 1 shows the relationships capacity-speed for the fixed and moving block systems respectively. The curve $C_m = C_m(v)$ is the so-called Lehner curve (Papacostas and Prevedouro, 2001; Vuchic, 2007).
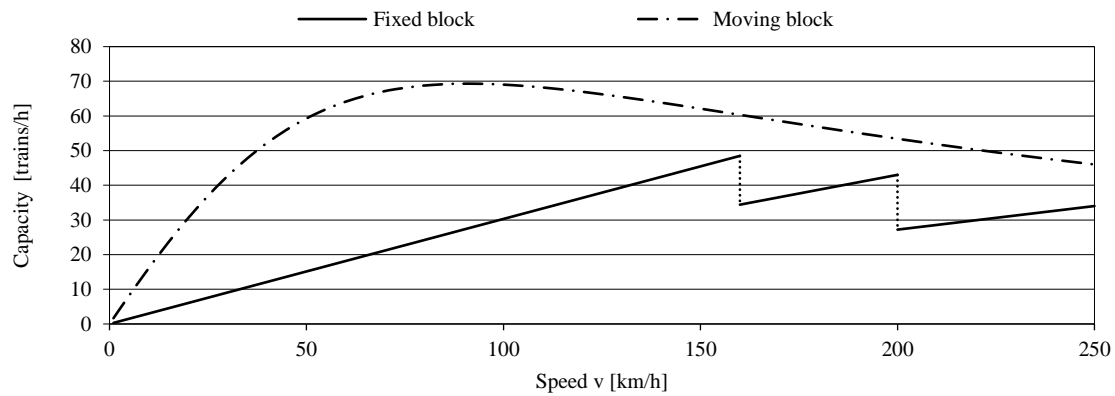

Figure 1. Railway line capacity (l_t = 400 m, f = 200 m, t_p= 3s, a = 0.8 m/s^2, k = 1.5)

Some researchers have shown that the amount of information on what action a train driver should take, can be increased by means of a device that pre-detects the colour of traffic signs (Marmo *et al.*, 2006). This approach increases safety in case of driver inattention and prevents train collisions (Marmo *et al.*, 2006).

In this regard a special framework for railway infrastructure detection based on computer vision was implemented on a diagnostic train (Marmo *et al.*, 2006; Nassu and Ukai, 2010; Lee *et al.*, 2014).

This research proposes a novel system for the automatic detection and recognition of traffic signs by means of the digital image processing technique (Moutakki *et al.*, 2018), but with a very simplified algorithm.

The first tests were conducted in the railway "Caltanissetta-Palermo" in Italy. To this end, some of the most relevant Italian railways signs were taken into account (RFI, n.d.). However, the proposed methodology can be easily extended to traffic signs of other countries.

2. A Novel Approach for Sign Recognition and Detection

The first phase of the proposed methodology is devoted to detect sign edges. The Canny algorithm (Freeman, 1974; Otsu, 1979) is then used to reduce the image noise (edge detection). The Canny algorithm analyses the behaviour of the gradient operator applied to a noisy edge. The high quality of the results from the Canny method can be justified by the fact that the method utilises two thresholds, one for detecting the sharpest edges, the other for detecting the weakest ones.

The algorithm works separately on the three chromatic layers of the image under analysis (see Fig. 2) and, only later, the processing results are reunited into one Boolean image (Fig. 3). By considering two threshold values s_1 and s_2 ($s_1>s_2$), the algorithm finds that the pixel placed at position (i,j), on the generic chromatic plane p of image I(i,j,p), is certainly a Canny edge pixel if the intensity value of pixel I(i,j,p) appears to be superior to the highest threshold value s_1. While the lowest value s_2 designates the limit under which the pixel is not a Canny edge (Guerrieri *et al.*, 2012).

For each point {K} of the determined edges, suitable convolution masks and congruence imposed on the chromatic intensities will help to:

- determine the gradient direction;
- create, for the eight pixels adjacent to each edge point, two brighter {C} and less bright {D} sets; an example of the pixel matching procedure to such sets is given by pixel intensities in the image shown in Figure 4;
- eliminate any pixel belonging to {C} and {D} at the same time;
- impose congruence of the result on the previous step with Canny edge {K}.

So, a slight edge thickening is obtained, thus closing (by connectivity) and removing any discontinuities (see Fig. 5). In addition, the algorithm is capable to discriminate some clearer pixels near the track edge, which do not belong to its micro wrinkles (Guerrieri *et al*., 2012).

Figure 2. Original image

Figure 3. Binary image obtained from Canny algorithm

Figure 4. Matching to sets {C} and {D}

Figure 5. Binary image {K} without discontinuities

For each chromatic plane of the analysed image, a moving average filter Med is applied; its kernel is:

$$\text{Med} = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}. \tag{4}$$

The algorithm requires an iterative approach: the image at the iteration step n+1 derives from the image convolved with the moving average filter (4) at step n:

$$I(i,j,p,n+1) = I(i,j,p,n) \times \text{Med}, \tag{5}$$

in which (i,j) are the pixel position indices within the original image I, p indicates the chromatic plane and "×" represents the convolution operator. The moving average filter (4) could eliminate the information on the edge details as shown in Figure 6.

Figure 6. Post-processing image implementing a moving average filter

Therefore, the use of another operator emphasizing the track edge is essential. For the purpose of the research, a range filter was iteratively applied as follows:

$$R(i,j,p,n+1) = \max[R(i\text{-}1,j\text{-}1,p,n),...,R(i,j,p,n),...,R(i+1,j+1,p,n)] \\ - \min[R(i\text{-}1,j\text{-}1,p,n),...,R(i,j,p,n),...,R(i+1,j+1,p,n)] \qquad . \tag{6}$$

At first, obviously, image R(i,j,p)≡I(i,j,p). Generally speaking, the range filter uses by default a 3-by-3 neighbourhood but in some cases a neighbourhood with different shapes and sizes can be taken into account. Consequently, for each iteration step n and chromatic plane p, the suggested methodology is able to reduce

pixel intensities, which mark the edge of the detected railway sign. In the light of the above reason, it is essential to introduce a new matrix, denoted with Z, which is initially empty and of equal size to I.

$$Z(i,j,p,n+1) = Z(i,j,p,n) + |I(i,j,p,n-1) \times Med - R(i,j,p,n)| . \tag{7}$$

At first, obviously, image Z(i,j,p)≡I(i,j,p). In this process a congruence condition, indicated here with ψ_c , needs to be imposed. In fact, it is necessary to correctly attribute every generic pixel at position (i,j) to a colorimetrically homogeneous area. Thus, a new matrix G, and therefore a new image, initially empty and of equal size with respect to image I, should be introduced:

$$G(i,j,p,n_f) = \psi_c(Z(i,j,p,n_f)), \tag{8}$$

in which n_f is the last of the n steps. Figure 8 shows image $G(i,j,p,n_f)$.

In order to simulate the way in which a train driver perceives different colours (Bruno *et al.*, 2012; Vitabile *et al.*, 2001), the segmentation is carried out in the HSV colour space, which is a three-dimensional space that can be visualised as a single cone, based on a cylindrical-coordinate representation of points in an RGB colour model: hue, saturation (y axis), and brightness value (x axis), as depicted in Figure 7.
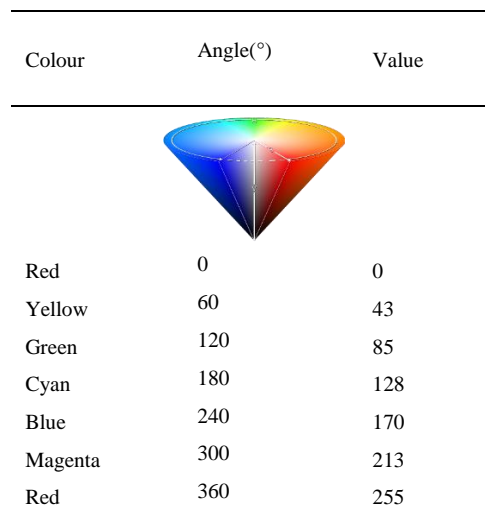
Colour	Angle(°)	Value
Red	0	0
Yellow	60	43
Green	120	85
Cyan	180	128
Blue	240	170
Magenta	300	213
Red	360	255

Figure 7. HSV colour space

Figure 8. Post processing image after the implementation of median filter, range operator and congruence conditions

In accordance to Vitabile *et al*. (2001) and De La Escalera *et al*. (2003) in a true perspective plan railway signs can be detected by distinguishing colours with different shades in the image {G} and using only information on chrominance H (cfr. Fig. 9) and saturation S (Fig. 10). By means of the following

specialised algorithm image {G}, in the original RGB colour space, has been converted into HSV colour space:

$$G_{\{R;G;B\}} \xrightarrow{\quad RGB \quad to \quad HSV \quad} G_{\{H;S;V\}} \tag{9}$$

Figure 9. Chromatic plane H of the image of Fig. 8

In order to obtain a more accurate selection of the areas containing the objects of interest (Fig. 9 and Fig. 10), the saturation plane {S} and the chromaticity plane {H} have been binarised using Otsu's algorithm (Otsu, 1979), obtaining the threshold matrices {H_{TR}} and {S_{TR}}, in which $_{TR}$ is Otsu's threshold level. The overall results – with no objects except the railway signs of interest - are gathered in a single Boolean matrix B_{wu}, as follows (see Fig. 11 (a)):

$$B_{wu}(i,j) = \{H_{TR}(i,j) \cup S_{TR}(i,j)\} \tag{10}$$

Figure 10. Chromatic plane S of the image of Fig. 8
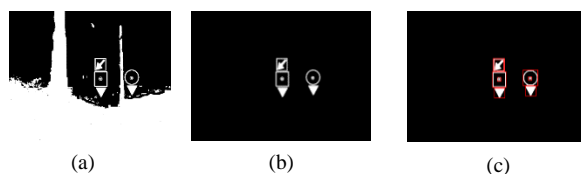

| (a) | (b) | (c) |

Figure 11. (a) Matrix union; (b) Binary morphology operations;
(c) Bounding box – Matrix B_{wf}

The position of railway signs in the bounding box can be detected with the Freeman chain code method (Freeman, 1961; Freeman, 1974; Guerrieri, *et al.*, 2012; Guerrieri, *et al.*, 2013) (see Fig. 11(b)), whose main result is matrix $B_{wf.}$ (Fig.11(c)).

The methodology that allows detecting (Fig.11(c)) and tracking a predefined railway sign in a partial area of one scene (see Fig. 12) is the so-called "object recognition" and is based on the template correlation.

Therefore, the correlation between all the binary objects detected in the partial area of image $\{B_{wf}\}$ (see Fig. 12) and the template of real signs catalogued and previously saved in memory needs to be evaluated. An example of some Italian railway signs is given in Figure 13.
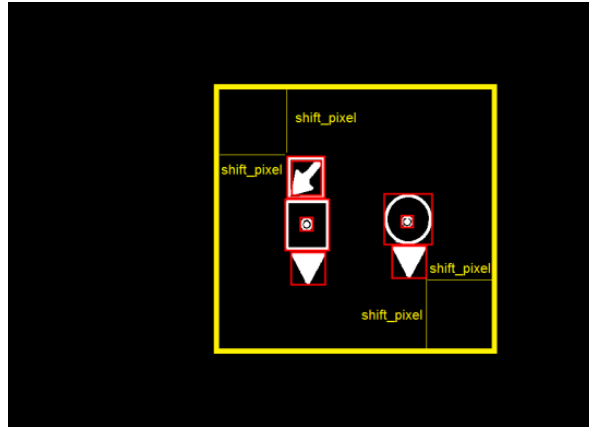

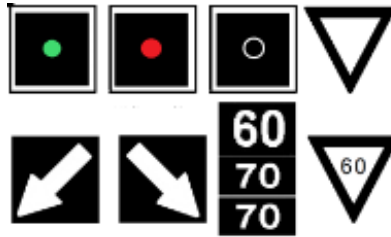*Figure 12.* Area of search in image B_{wf}

Figure 13. Example of template for comparison and correlation between the railway signs

In other words, the implemented processes consist in searching for the correlation between the template of a real sign and its image in the sequences of the video-recorded images. This search ends once the position with the highest correlation coefficient between the two (Fig. 13 and Fig. 14) is found, and if this value is higher than a prefixed value of correlation coefficient c*. More precisely: "when moving the template along the rows and columns of the area to be searched G, a coefficient of similarity between the two images is calculated, at each position" (see Fig. 15 and Fig. 16 concerning another image of the video-recorded images).

The correlation coefficient c is evaluated as follows (Lewis, 1995; Bruno *et al.*, 2012):

$$c(i,j) = \frac{\sum_{i,j}\left[G(i,j)-\overline{G}_{u,v}\right]\left[T_p(i-u,j-v)-\overline{T}_p\right]}{\sqrt{\sum_{i,j}\left[G(i,j)-\overline{G}_{u,v}\right]^2 \sum_{i,j}\left[T_p(i-u,j-v)-\overline{T}_p\right]^2}} \geq c^*, \tag{11}$$

where G (see Eq. (8) is the image and the sum is over i; j under the window containing feature T_p positioned at (u; v) and $\overline{T}_p$ is the mean of the feature (Template) and $\overline{G}_{u,v}$ is the mean of G(i; j) in the region under the feature. We refer to (11) as *normalized cross-correlation* (Lewis, 1995).

Finally, the proposed procedure introduces suitable threshold values on the size of objects before they are detected in the tracking phase (Guerrieri *et al.*, 2012; Bruno *et al.*, 2012).
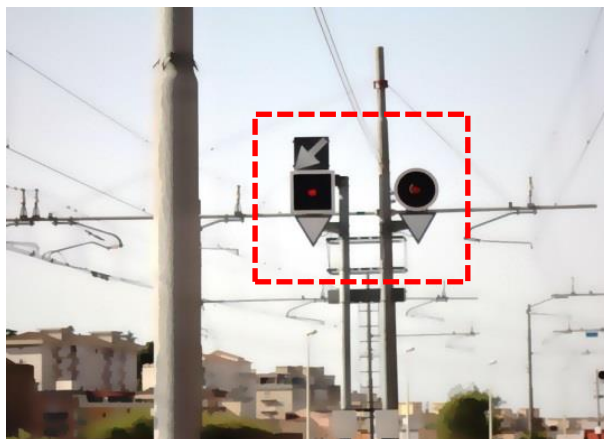
Figure 14. Search area of the railway signs

Figure 15. Tracking phase

Figure 16. Signs classification phase

3. Preliminary Results

The proposed methodology was applied in a case study in order to obtain its preliminary validation on the basis of experimental data. First analysis concerned 400 signs of the line "Caltanissetta-Palermo" (Italy) - that is a strategic suburban railway - near the railway station (within a distance of about 75 km).

Scenes were collected on a regional train, equipped with a high-resolution CCD camera (model "Sony FDR AX 100").

In this first experiment we considered only the railway sections with a train speed less than about 80 km/h and good weather conditions. The percentage of correctly detected railway signs was found to be about 97% (Table 1) in case of good weather conditions.

Table 1. Details of the results obtained

Caltanissetta-Palermo railway	
Total signs	400
Visible signs	390
Detected signs	378
Detected	96.92%
Occluded or damage signs	10

4. Conclusion

The railway safety is a relevant issue for those suburban and regional lines that are not equipped with modern safety technologies, because several tasks still depend on the choices and actions of a human crew. In addition, although indirectly, human factors can affect the capacity of the railway transportation systems. In order to increase the amount of information on what action a train driver has to be taken, this research implemented a novel methodology for railway sign detection. The novel methodology allows an automatic sign detection based on the digital image processing technique and relatively simple algorithms.

In order to obtain a validation based on experimental data, the proposed automated railway sign detection was applied to a case study of a suburban railway ("Caltanissetta-Palermo" in Italy). Actually, only the railway sections with train speeds less than about 80 km/h were considered.

The preliminary results obtained show that more than 97% of the railway signs were actually detected via image processing analysis. The proposed methodology is still going to improve its performance in other common cases of very noisy images due to rain or fog and those in which the train speed is higher than 80 km/h.

References

1. Bruno, L., Parla, G., Celauro, C. (2012) Improved Traffic Signal Detection and Classification via Image Processing Algorithms. *Procedia - Social and Behavioral Sciences* (53), 811–821. https://doi.org/10.1016/j.sbspro.2012.09.930
2. De La Escalera, A., Armingol, J. M. A., Mata, M. (2003) Traffic sign recognition and analysis for intelligent vehicles. *Image and vision computing,* 21, 247–258. - https://doi.org/10.1016/S0262-8856(02)00156-7
3. Delfino, A., Galaverna, M. (2003) Mobile and fixed block: Carrying capacity analysis. *Ingegneria Ferroviaria* 6: 555-565.
4. Directive 2004/49/EC, Commission Directive 2009/149/EC and Regulation (EC) No 91/2003.
1. ERA. Railway Safety Performance in the European Union. Biennial Report. European Union Agency for railway, 2016.
2. Freeman, H. (1961) On the encoding of arbitrary geometric configurations. *Ire Trans. Electron. Comput.* EC-10, 260-268.
3. Freeman, H. (1974) Computer Processing of Line Drawing Images. *ACM Computing Surveys.* 6(1), 57-97. https://doi.org/10.1109/TEC.1961.5219197
4. Guerrieri, M., Parla, G., Corriere, F. (2013) A new methodology to estimate deformations of longitudinal safety barrier. *ARPN Journal of Engineering and Applied Sciences*, 8(9), 763-769.
5. Guerrieri, M., Parla, G., Ticali, D. (2012) A theoretical and experimental approach to reconstructing the transverse profile of worn-out rails. *Ingegneria Ferroviaria* 67(1), 23-37.
6. Lee, D., Yoon, H., Chung, M., Kim, J. (2014) Robust Sign Recognition System at Subway Stations Using Verification Knowledge. *ETRI Journal*, 36(5), 696-703. - https://doi.org/10.4218/etrij.14.2214.0007
7. Lewis, J.P. (1995) Fast Template Matching. Vision Interface. *Canadian Image Processing and Pattern Recognition Society*, 15-19, pp. 120-123.

8. Marmo, R., Lombardi, L., Gagliardi, N. (2006) Railway Sign Detection and Classification. *Proceedings of the IEEE ITSC 2006*, Toronto, Canada, September 17-20, 1358-1363.
9. Moutakki, Z., Mohamed Ouloul, I., Afdel, K., Amghar, A. (2018) Real-time system based on feature extraction for vehicle detection and classification. *Transport and Telecommunication*, 19(2), 93–102.
10. Nagy, E., Csiszár, C. (2015) Analysis of delay causes in Railway passenger transportation. *Periodica Polytechnica Transportation Engineering,* 43(2), 73-80. https://doi.org/10.3311/PPtr.7539
11. Nassu, B.T., Ukai, M. (2010) Automatic Recognition of Railway Signs Using SIFT Features. *IEEE Intelligent Vehicles Symposium*, University of California, San Diego, CA, USA, June 21-24, 348-354. https://doi.org/10.1109/IVS.2010.5548127
12. Otsu, N. (1979) A threshold selection method from grey-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62 - 66.
13. Papacostas, C.S., Prevedouro, P.D. (2001) *Transportation Engineering and Planning* (3rd Edition), Prentice Hall.
14. Sostarics, G., Zobory, I. (1981) Railway brake-gear design method by means of probability theory. *Periodica Polytechnica Transportation Engineering*, 9(2), 119-131.
15. Vitabile, S., Pollacia, G., Pilato, G., Sorbello, F. (2001) Road signs recognition using a dynamic pixel aggregation technique in the HSV color space. *Proceedings of the 11th International Conference on image analysis and processing*, Palermo, Italy, 572 - 577.
16. Vuchic, V.R. (2007) *Urban Transit Systems and Technology*, John Wiley & Sons.