# Characterizing the performance of automatic road detection using error propagation

S. Gautama * W. Goeman J. D'Haeyer W. Philips

*TELIN, Ghent University, St.Pietersnieuwstraat 41, B-9000 Gent, Belgium*

**Abstract**

A methodology is introduced to predict the performance of automatic road detection using image examples of typical road types. In contrast to previous work on road detection, the focus is on characterizing the detection performance to achieve reliable performance measures of the detection. It is studied how noise, like road markings, shadows, trees and buildings, influences the detection of the road. This noise is modeled using second-order statistics and its effects are calculated using error propagation on the detection equations. The method predicts the performance in terms of detection rate and gives the optimal parameter set that is needed for this detection. Experiments have been conducted on a set of images of typical roads in very-high-resolution satellite images.

*Key words:* road detection, performance characterization, error propagation

* Corresponding author.
*Email address:* sidharta.gautama@ugent.be (S. Gautama).

## 1 Introduction

Although digital maps are available for many areas, producers of geospatial data continuously need to keep this data accurate, up-to-date and complete. Not only do they need to control the quality of their data, the high demands of current applications puts pressure to deliver products of increasingly higher spatial resolution. Upgrading the spatial accuracy of existing data however is a labour intensive and expensive process. In this paper, we focus on the problem of quality assessment of digital maps and more specifically digital road maps. We investigate automatic road detection in very-high-resolution (VHR) satellite images, which forms the basis for automated quality assessment and quality control in road GIS databases.

The proposed system for quality assessment is based on object based spatial registration, where detected objects in the image are registered and compared to corresponding objects in the GIS data (cfr. [1]). The system consists out of two stages: 1) a low-level feature detection process, which extracts roads and junctions using an improved ridge detector, and 2) a high-level matching process, which uses graph matching to find correspondences between the detected image information and the road data (cfr.fig.1). The graph matching process is driven by the spatial relations between the objects and takes into account different errors that can occur (e.g. spatial inaccuracy, data inconsistencies between image and GIS data). The matched objects can be used to calculate a rubbersheeting transformation between image and GIS data that is able to compensate the local distortions that can occur between the datasets. Additionally the object-to-object mapping is useful to define measures of change between datasets.

Much of the performance of the system depends on the quality of the first stage, i.e. the extraction of road information from the images. Although much effort has been spent on designing algorithms for road detection (e.g. [2]), this complex problem is still not fully solved. In contrast with previous efforts, which are aimed at improving the detection, we focus on characterizing the performance of the detection. There are several reasons for this:

(1) The industry needs to be able to assess the benefits that can be gained with image based road detection for its specific needs. An analysis tool should be ready to assess the expected quality of road detection for a given dataset to be able to decide if the use of this technology can decrease operational costs. Often the results published in the literature are based on a small number of images and are as such insufficient to predict the performance on larger datasets or data of different type (e.g. a different resolution or region). Based on example images that describe typical roads in the region, a tool is needed that predicts the expected detection rate.

(2) An important difficulty in using a detection algorithm is that a number of parameters need to be determined by the operator. The choice of the parameter set has a big impact on the detection performance and is dependent on the image content. Since the operator is not necessarily an image processing expert, a tool should be available to estimate the "optimal" parameter set given example images.

(3) An object based quality assessment system naturally depends on the quality of the detection in the image. Decisions need to be taken based on the information gathered by the detection, so one needs to be able to estimate the reliability of the information. The high-level matching process, which compares image information with the GIS data, needs information

about the reliability of object detection in terms of true and false positive detection. The reason is that the matching process needs to differentiate between inconsistencies due to change and inconsistencies due to outliers (i.e. false positive detection). Knowledge about the expected number of outliers is used to make this distinction.

Although numerous papers exist on road detection, the focus is always on improving or extending the detection operators. In this work, the emphasis is not so much on precise localization but on robust detection. Therefore we have studied how noise, like activity on the road, shadows, trees and buildings next to the road, influences the detection of the road. This noise is modeled using second-order statistics and its effects are calculated using error propagation on the detection equations [5]. Error propagation allow us to predict the detection performance for a road in a given image, based on simple image statistics. In addition, the optimal parameter set can be determined that gives the best detection, taking into account the type of road, image noise and structural noise. Prediction of the performance can be done in an interactive manner, where the user shows typical road examples and gets in return the estimated performance on his dataset. On the other hand, it can be used during the detection, where in addition to the detected road pixels information is given on the reliability.

## 2 Ridge detection

Lines in an image can be seen as narrow valleys or ridges in the intensity surface if one views the image as a terrain model. In [3], different approaches to line detection are reviewed. In this work, line detection is performed based

on polynomial interpolation to determine pixels belonging to road structures in the image, the "facet model" [4]. This is a standard method for ridge detection. The image is regarded as a function $I(i, j)$. Lines are detected as ridges and ravines in this function by locally approximating the image function by its second order Taylor polynomial. The polynomial is used to approximate first and second order derivatives of the image function in each pixel. The direction of the line can be determined from the Hessian matrix of the Taylor polynomial. The gradient and curvature information in each pixel are used to classify a pixel in a number of topological classes based on their sign or magnitude. Line points are mainly characterized by a high second directional derivative, i.e. a high curvature perpendicular to the line direction.

The calculation of the partial derivatives can be done in various ways. The facet model determines a least squares fit of a polynomial $F$ to the image data $I$ over a window of size $N = w^2$ with window size $w$. The origin is chosen in the central pixel of the window. The value of the polynomial $F$ in pixel $(i, j)$ is given by:

$$\begin{aligned} F(i, j, \vec{\theta}) &= a_1 + a_2 i + a_3 j + a_4 i^2 + a_5 ij + a_6 j^2 \\ &= \vec{m}^T \vec{\theta} \\ \text{with } \vec{m} &= \begin{bmatrix} 1 \; i \; j \; i^2 \; ij \; j^2 \end{bmatrix}^T \\ \vec{\theta} &= \begin{bmatrix} a_1 \; \cdots \; a_6 \end{bmatrix}^T \end{aligned} \tag{1}$$

The facet model searches the least-squares solution $\vec{\theta}$, given the image data $\vec{x}$

containing the intensity value $I(i,j)$ in each pixel $(i,j)$:

$$\begin{aligned}
&\arg\min_{\vec{\theta}} r(\vec{\theta}) \text{ with } r(\vec{\theta}) = \left\| M\vec{\theta} - \vec{x} \right\|^2 \\
&M = \begin{bmatrix} 1 & i_1 & j_1 & i_1^2 & i_1 j_1 & j_1^2 \\ \vdots & & & & & \vdots \\ 1 & i_N & j_N & i_N^2 & i_N j_N & j_N^2 \end{bmatrix} \in \mathfrak{N}^{N \times 6} \\
&\vec{x} = \begin{bmatrix} I(i_1, j_1) & \ldots & I(i_N, j_N) \end{bmatrix}^T \in \mathfrak{N}^{N \times 1}
\end{aligned} \tag{2}$$

This leads to the linear system $M^T M\vec{\theta} = M^T\vec{x}$ with the solution $\vec{\theta}_0$ given by $\vec{\theta}_0 = (M^T M)^{-1} M^T \vec{x}$.

The matrix $M$ is independent of the position of the window within the image, meaning that the calculation of $(M^T M)^{-1} M^T$ needs to be performed only once for the processing of an image with a fixed window size $w$. On the basis of the parameters $\vec{\theta}$ of the interpolated surface $F$, the gradient and Hessian in a certain pixel can be calculated. In our model we are only interested in the gradient and the Hessian in the central pixel of the window (i.e. $i = j = 0$):

$$\begin{aligned}
gradient(I) &= \left[ \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right]^T = \begin{bmatrix} a_2 + 2a_4 i + a_5 j \\ a_3 + a_5 i + 2a_6 j \end{bmatrix} \\
Hessian(I) &= \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2a_4 & a_5 \\ a_5 & 2a_6 \end{bmatrix}
\end{aligned} \tag{3}$$

## 3 Error Analysis

We wish to give a more quantitative analysis of the performance of ridge detection. More specifically, we wish to be able to predict the performance of the detector for a given dataset and the according parameter set that gives optimal results. For this, we analyse the influence that perturbations on the intensity values have on the estimation of the parameters by using error propagation [5]. Additive random perturbations are assumed on the input $\vec{x}$ and the perturbations are described by the covariance matrix $\Sigma_{\vec{x}}$. The propagation of the error on the input $\vec{x}$ to the estimated parameters $\vec{\theta}$ is given by:

$$\begin{aligned} \hat{\vec{x}} = \vec{x} + \triangle\vec{x}, \quad \hat{\vec{\theta}} = \vec{\theta} + \triangle\vec{\theta} \\ \Sigma_{\triangle\vec{\theta}} = (M^T M)^{-1} M^T \, \Sigma_{\triangle\vec{x}} \, M (M^T M)^{-1} \end{aligned} \tag{4}$$

The matrix $(M^T M)^{-1} M^T$ can be worked out to a closed expression. Due to the linear system, the error propagation of $\vec{x}$ on $\vec{\theta}$ is only dependent on the window size $w$ and not on the input intensity values (and consequently independent of the observed perturbed image structure). For uncorrelated noise with variance $\sigma^2$, Eq.(4) simplifies to $\Sigma_{\triangle\vec{\theta}} = \sigma^2 (M^T M)^{-1}$ since in this case $\Sigma_{\triangle\vec{x}} = \sigma^2 I$. However, in this work, the general case is followed.

The parameter covariance matrix $\Sigma_{\triangle\vec{\theta}}$ allows to estimate the variance on the detected gradient and curvature of the ridge detector. The gradient magnitude and eigenvalues of the Hessian are given by

$$\begin{aligned} G \equiv \|\vec{g}\| = \sqrt{a_2^2 + a_3^2} \\ \lambda_{1,2} = a_4 + a_6 \pm \sqrt{a_4^2 + a_5^2 + a_6^2 - 2a_4 a_6} \end{aligned} \tag{5}$$

The variance on these measurements up to first order is given by

$$\begin{aligned} \sigma_G^2 &= \left(\frac{\partial G}{\partial \vec{\theta}}\right)^T \Sigma_{\vec{\theta}} \left(\frac{\partial G}{\partial \vec{\theta}}\right) \\ \sigma_{\lambda_1}^2 &= \left(\frac{\partial \lambda_1}{\partial \vec{\theta}}\right)^T \Sigma_{\vec{\theta}} \left(\frac{\partial \lambda_1}{\partial \vec{\theta}}\right) \end{aligned} \tag{6}$$

In combination with Eq.(4), this gives the relation between the perturbations on the image data and the perturbations on the detection measurements $G$ and $\lambda_1$. This relation allows to estimate the expected performance of the detector for a given parameter set. The basic ridge detector operates by setting a threshold $t_1$ on the first eigenvalue $\lambda_1$ of the Hessian. Pixels that exceed this threshold are selected as ridge pixels. To predict the performance of the detector, the eigenvalue $\lambda_1$ is regarded as a stochastic variable using a Gaussian distribution with variance $\sigma_{\lambda_1}$. The probability $p(\lambda_1 > t_1)$ expresses when the curvature $\lambda_1$ of a pixel exceeds $t_1$. This distribution is given by the survival function (i.e. the complement of the cumulative density function) of the Gaussian:

$$\begin{aligned} p\left(\lambda_1 > t_1\right) &= \frac{1}{\sigma_{\lambda_1}\sqrt{2\pi}} \int_{t_1}^{\infty} e^{-\frac{(\lambda_1 - \overline{\lambda_1})^2}{2\sigma_{\lambda_1}^2}} d\lambda_1 \\ &= \tfrac{1}{2}\left[1 - erf\left(\frac{\lambda_1 - \overline{\lambda_1}}{\sigma_{\lambda_1}\sqrt{2}}\right)\right] \end{aligned} \tag{7}$$

Given the survival function, calculated with the appropiate statistics for the road, a prediction of the detection performance can be made in terms of true positive and false negative detection. In addition, if statistics for the noise structures in the surroundings of a road (e.g. buildings, trees) are measured, a prediction of the false positive and true negative detection can be made.

## 4 Experimental results

In this section, the derived expressions are applied to predict the performance of the detection. For these experiments, an IKONOS panchromatic satellite image has been used, acquired over the city of Ghent, Belgium. The image is a standard GEO product with a one-meter pixel resolution. A number of small image subsets have been selected containing straight examples of typical roads. These subsets have been manually rotated until each road is positioned parallel to the vertical axis. Figure 2 shows the selected image subsets referenced $R1$ to $R5$.

### *4.1 Verification of the derived expressions*

The position $x_{road}$ of the central axis of the road is used to measure the covariance matrix $\Sigma_{\vec{x}}$ and the mean curvature $\overline{\lambda_1}$, by sampling $\vec{x}$ and $\lambda_1$ for a given window size $w$ along the axis. This position $x_{road}$ is determined manually by examining the mean curvature profile of the image subset. Based on the measurements, $\Sigma_{\vec{\theta}}$ and $\sigma^2_{\lambda_1}$ are calculated using Eq.(4) and Eq.(6). This allows us to predict the survival function $p(\lambda_1 > t_1)$ for a given threshold $t_1$ on the curvature $\lambda_1$.

Fig. 3 shows an example plot of the survival function for image subset $R_5$ using a window size $w = 11$. The solid curve shows the empirical distribution measured over the central axis of the road. The dashed curve shows the predicted distribution using the measured $\Sigma_{\vec{x}}$ and $\overline{\lambda_1}$. In this example, the two distributions show a good resemblance. To quantify the goodness-of-fit we measured the discrepancy between the distributions for each threshold sample point.

Fig. 4 summarizes this discrepancy for the five image subsets for different window sizes $w$. The bars in this plot show the mean discrepancy between the empirical and estimated distribution, averaged out over a threshold interval $[0, t_f]$. The upper bound $t_f$ of this interval is defined by $p(\lambda_1 > t_f) = 0.01$. The error bars on the plot in Fig. 4 show the minimum and maximum discrepancy for each sample. Fig. 3 shows the location of the maximum discrepancy for road type $R_5$ with window $w = 11$.

Fig. 4 shows a mean discrepancy of less than 5% for the higher window sizes ($w > 11$). The larger mean discrepancy for small windows occur when the gaussian model is less adequate to describe the statistics of the road. Small local artifacts on the road (shadow, road markings, cars), which are smoothed out with larger windows, could in this case be better modeled using a mixture of normal distrbutions.

### *4.2 Optimal performance*

The distribution that is measured in Fig. 3 only gives information about the useful signal, i.e. the road that needs to be detected. This gives an upper bound to the threshold $t_1$. A lower bound is defined by the properties of the noise, i.e. undesired structures in the image which are falsely detected. If the threshold is chosen too low, the detection of these false negatives will increase. The performance of the detection for each threshold can be summarized in a Receiver Operating Characteristic (ROC) curve. Sensitivity and specificity are

defined as follows:

$$\begin{aligned} sensitivity &= \frac{TP}{TP+FN} = \frac{detected\,road\,pixels}{road\,length} \\ specificity &= \frac{TN}{TN+FP} = 1 - \frac{detected\,noise\,pixels}{noise\,length} \end{aligned} \tag{8}$$

where $\{TP, TN, FN, FP\}$ stands for true positive etc. True positives and false negatives are measured on the road axis $x = x_{road}$ and are normalized using the total road length in the image. True negatives and false positives are measured along a vertical axis $x = x_{noise}$ representative for the noisy structures in the vicinity of the road. The position of this axis is determined automatically by taking the position of the second largest peak in the mean curvature profile of the road and its vicinity. Normalisation is performed using the total length of the axis, which in this case equals the total road length. Fig. 8a illustrates the position $x = x_{road}$ and $x = x_{noise}$ for road type $R_1$.

The reason for this approach is that it can be meaningful to characterize false detections in the direct surroundings of a road. Since for some applications a rough registration between image and GIS data is known, regions-of-interest can be defined where roads are expected in the image. Road detection in this case should then be aimed at distinguishing the useful signal from noise in the immediate surroundings. In our example we measure the noise statistics of the structures beside the road and use this to model the expected falsely detected noise pixels using Eq.(7). In this case $\Sigma_{\vec{x}}$ and $\overline{\lambda_1}$ are measured for $x = x_{noise}$ for the same window size $w$ as the signal.

Using the ROC curve, the optimal performance for this curve is defined as the point on the curve closest to the upper right corner $(1, 1)$. The upper right corner is equivalent to a road which is completely detected with no detection of noisy structures (i.e. perfect detection). The point on the curve closest to

perfect detection defines the optimal threshold $t_1$ for the given window size. In Fig. 5, the empirical and predicted ROC curves for road type $R_1$ are plotted for window size $w = 11$ and $w = 13$. The full curve shows the empirical plot, where detected and falsely detected road pixels have been measured on the road axis and the position $x_{noise}$. The dashed curve shows the predicted plot using the covariance matrices measured in these positions. The sample point closest to the upper right corner defines the optimal threshold for each window size. The difference between the empirical and predicted ROC curve for a given window size is mainly due to a less accurate estimation of the true noise survival function. Fig. 7 shows both survival functions for road type $R_1$ for $w = 11$. The figure shows a good correspondence between the empirical and predicted survival for $x = x_{road}$ and a less accurate estimation for $x = x_{noise}$. This is because the noise introduced by the building structures are less well captured in a single gaussian distribution and should be described by a mixture model. On the road itself, where the noise structures are small compared to the window size, the estimation is more accurate.

To determine the optimal window size and associated threshold $(w_{opt}, t_{opt})$ for a certain road type, the ROC curves for different window sizes are calculated. The curve which passes closest to the upper right corner defines the parameter set $(w_{opt}, t_{opt})$. We have performed this procedure for road type $R_1$ for window sizes $w \in [7, 19]$ and $t_1 \in [0, 4]$. In Fig. 6, the distance of each sample point on the ROC curves to the upper right corner is plotted in a contour plot for both empirical and estimated data. The structure of both contour plots is comparable with the location of optimal performance being situated in the same area in parameter space. The exact minimum is different for empirical and predicted data, being resp. $\{w_{opt} = 13, t_{opt} = 0.75\}$ and $\{w_{opt} = 11, t_{opt} =$

0.8} (cfr.Table 1). However, looking at the actual ROC curves in Fig. 5, we see that for the predicted curves the difference between $w = 11$ and $w = 13$ is very small.

More important than the exact location of the optimum however is the relevance of the predicted performance. Table 1 summarizes the predicted and empirical true positive and false negative detection for given parameter sets, as well as the Euclidean distance to the point of perfect detection. We see that for the optimal estimated parameter set $(11, 0.8)$, there is a good correspondence for the true positive detection and a slight underestimation for the false negative detection. Also compared to the optimal empirical parameter set $(13, 0.75)$, the difference is small. Fig. 8 shows both the optimal predicted and optimal empirical detection result in overlay on the image.

In a last test, the results obtained for road type $R_1$ are compared with the detection over a larger region containing similar roads (cfr. Fig. 9). The methodology to determine the detection rate is slightly different because it was more difficult to locate the precise road axes. Instead two regions have been defined in the image: one region containing the roads and one region containing the road surroundings up to 20 meter at the left and right of a road. These regions are used to distinguish between the useful signal and the noise. To normalize the true and false positive detection rate, the area of each region is used instead of road and noise length in Eq.(8). The detection performance in the full image using the optimal predicted parameter set is given in Table 1. We see that the predicted performance overestimates the detection rate over the full image. In the case of the false negatives this is expected, since by taking only the statistics over $x_{noise}$, i.e. the location of the most detected pixels beside the road, this amounts to a worst-case estimation. In the case of the true

positives, the overestimation is mainly due to the fact that the image subset $R_1$ is actually a well behaved example of the road (note that the empirical TP rate is also 90% compared to 76% for the full image). The full image contains at the bottom a number of roads with trees, which leads to a bad detection in this part. Additionally, junctions are not detected by the ridge detector since the line model does not hold anymore in these cases. Junctions are not covered in our model. When we neglect the roads containing trees in the image, the TP rate increases to 82%, which is closer to the predicted 89%.

## 5 Conclusion

In this paper, an analysis of the performance of road detection is introduced. A methodology has been established where, based on the image statistics of a road and its immediate surroundings, the performance of the detection can be predicted as well as the optimal parameter set which is needed for the detection of this type of road. The performance is characterized in function of the detection rate. Experiments have been conducted on a set of images of typical roads in an IKONOS satellite image. Results show a mean discrepancy between the predicted and empirical survival function of less than 5% for larger window sizes, measured on the central axis of the road. By using the image statistics of the road axis and the road vicinity, ROC curves can be calculated to determine the optimal parameter set. The optimization function, which is used to determine the parameters, uses the euclidean distance to perfect detection. The function shows a similar shape for the predicted and empirical ROC curves. Differences are mainly due to the modeling of the noise structures like buildings, which in some cases is less accurate. The normal distribution,

which is used in this paper, could be extended to a mixture modeling to capture the image content more accurately.

## References


[1] S.Gautama and A.Borghgraef, "Detecting change in road networks using continuous relaxation labeling" in Proc. ISPRS Workshop "High Resolution Mapping from Space 2003". Hannover. 6 pag. Oct 2003.

[2] J.Mena, "State of the art on automatic road extraction for GIS update: a novel classification", Pattern Recognition Letters 24:3037–3058, 2003.

[3] C.Steger, "An unbiased detector of curvilinear structures", IEEE Transactions on Pattern Analysis Machine Intelligence 20(2):113–125, 1998.

[4] R.Haralick, L.Watson and T.Laffey, "The topographic primal sketch", International Journal of Robotics Research 2(1):50–72, 1983.

[5] R.Haralick, "Propagating Covariance in Computer Vision" International Journal on Pattern Recognition and Artificial Intelligence 10(5):561–572, 1996.

| | $(w, t_1)$ | TP | FN | distance to perfect detection |
|---|---|---|---|---|
| predicted (optimal) | (11, 0.8) | 89% | 33% | 0.35 |
| empirical | (11, 0.8) | 90% | 40% | 0.41 |
| empirical (optimal) | (13, 0.75) | 85% | 36% | 0.39 |
| full image | (11, 0.8) | 76% | 19% | |
| full image no trees | (11, 0.8) | 82% | 19% | |

Table 1

Comparison of empirical and predicted performance

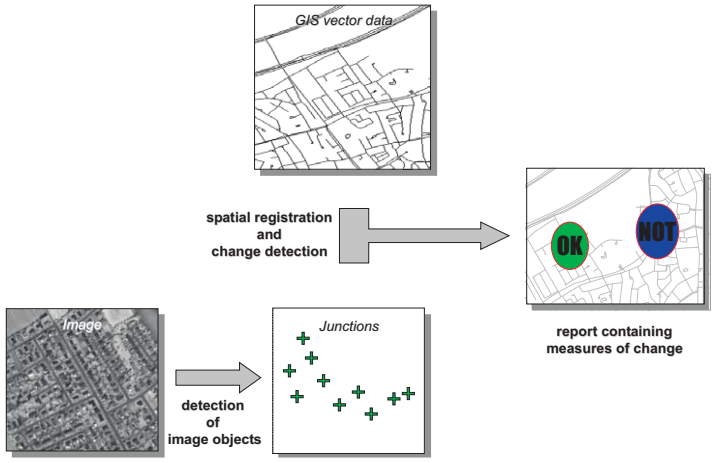
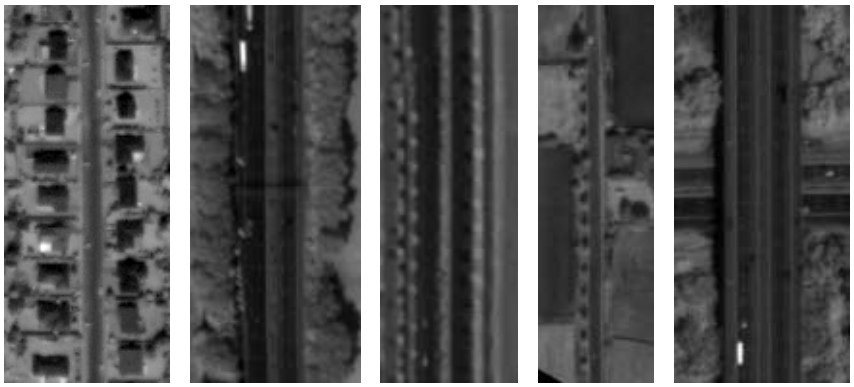



Fig. 1. Overview of the system for automated quality assessment of GIS data using VHR images



(a) R1 (b) R2 (c) R3 (d) R4 (e) R5

Fig. 2. IKONOS image subsets showing typical roads in the region of Ghent, Belgium. (a)-(e) Road type $R_1$-$R_5$. For road type $R_2$, $R_3$ and $R_5$ the right lane is selected for the analysis.

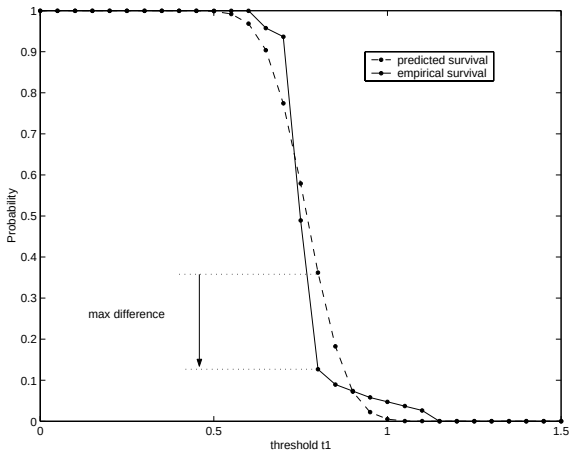



Fig. 3. Plot of the survival in function of the threshold $t_1$ on the curvature $\lambda_1$, for road type R5 with $w = 11$.

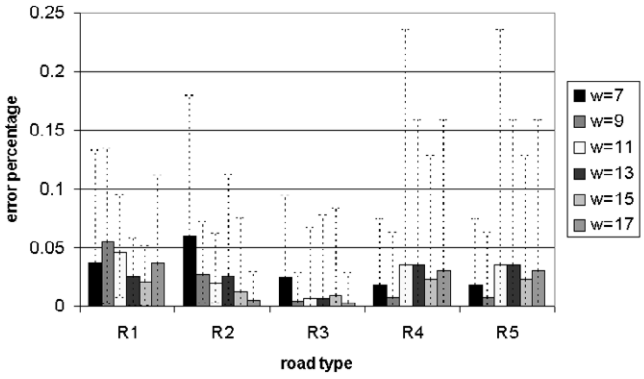



Fig. 4. Overview of the mean and maximum discrepancy between the empirical and predicted survival for different window sizes $w$ and road types

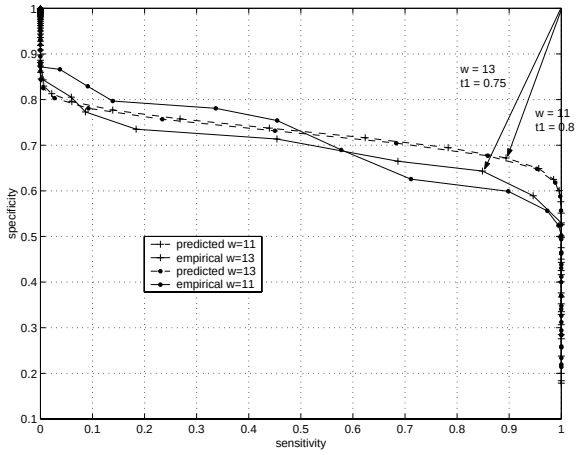
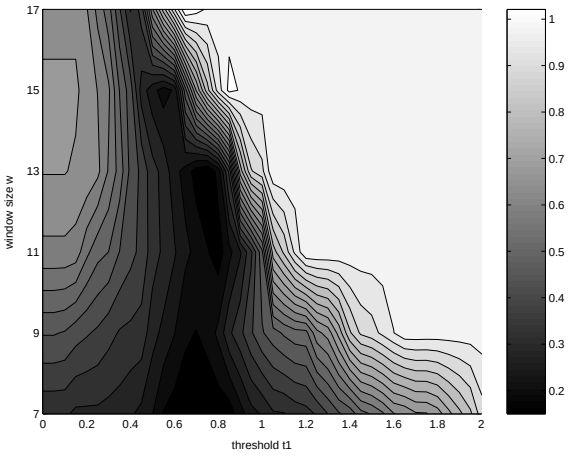
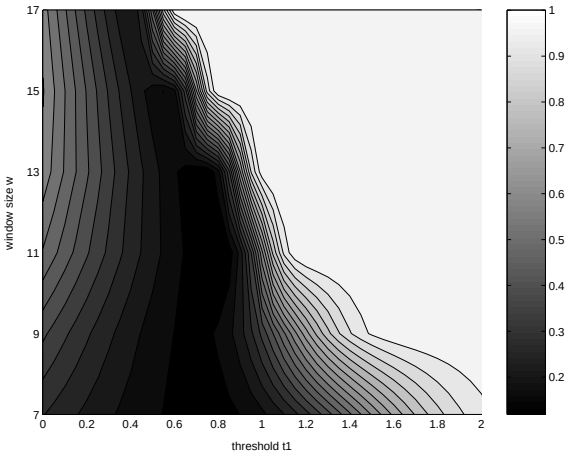



Fig. 5. Empirical and predicted ROC curves for road type $R_1$ for window size $w = 11$ and $w = 13$. The empirical and predicted point of optimal performance are marked by the arrows.




(a) empirical




(b) predicted

Fig. 6. Contour plot of road type $R_1$ showing the distance in the ROC curve of every sample point to perfect detection.

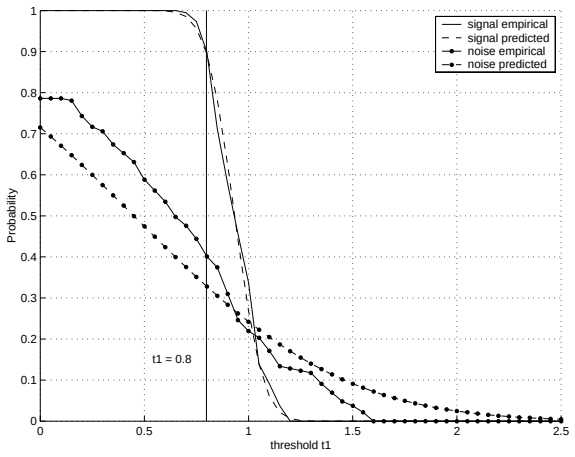
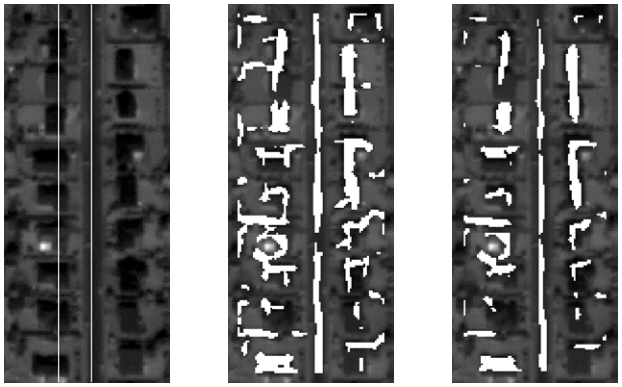
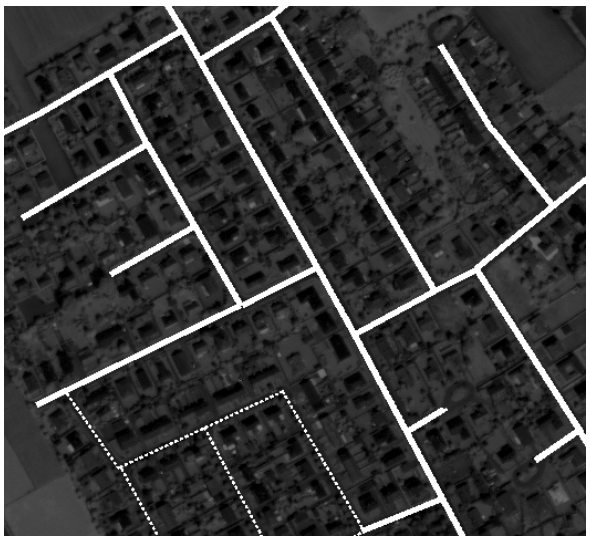



Fig. 7. Empirical and predicted survival function for road type $R_1$ for $x = x_{road}$ and $x = x_{noise}$.

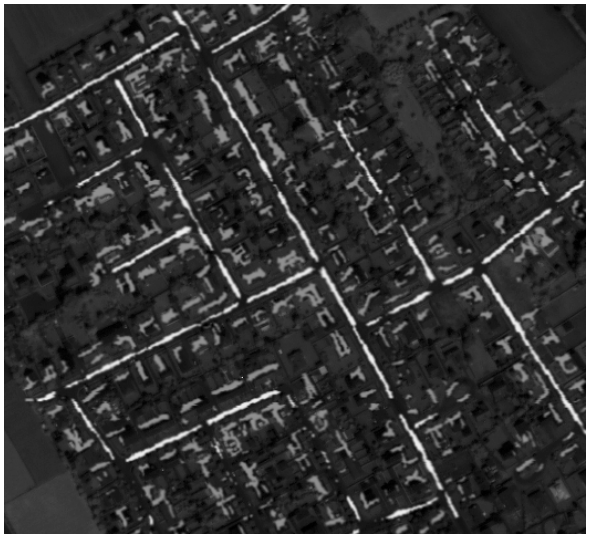


(a) (b) (c)

Fig. 8. Detected result on road type $R_1$ using optimal parameter settings. The original image shows the position $x = x_{road}$ and $x = x_{noise}$. The associated detection rate is summarized in Table 1. (a) original (b) predicted (c) empirical



(a) ground truth defining the road region (roads containing trees are marked with a dotted line)



(b) detection result (white: true positives, light gray: false positives

Fig. 9. Overall detection result using predicted optimal parameter settings.