ROBUSTNESS OF DIFFERENT FEATURES FOR ONE-CLASS CLASSIFICATION AND ANOMALY DETECTION IN WIRE ROPES

Esther-Sabrina Platzer, Joachim Denzler, Herbert Süße
Chair for Computer Vision, Friedrich Schiller University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany
platzer,denzler@informatik.uni-jena.de, herbert.suesse@informatik.uni-jena.de

Josef Nägele and Karl-Heinz Wehking
Insitute for Mechanical Handling and Logistics, University of Stuttgart, Holzgartenstrae 15B, 70174 Stuttgart, Germany
naegele,wehking@ift.uni-stuttgart.de

Keywords: Anomaly detection, novelty detection, one-class classification, linear prediction, local binary pattern.

Abstract: Automatic visual inspection of wire ropes is an important but challenging task. Anomalies in wire ropes usually are unobtrusive and their detection is a difficult job. Certainly, a reliable anomaly detection is essential to assure the safety of the ropes. A one-class classification approach for the automatic detection of anomalies in wire ropes is presented. Different well-established features from the field of textural defect detection are compared to context-sensitive features extracted by linear prediction. They are used to learn a Gaussian mixture model which represents the faultless rope structure. Outliers are regarded as anomaly. To evaluate the robustness of the method, a training set containing intentionally added, defective samples is used. The generalization ability of the learned model, which is important for practical life, is exploited by testing the model on different data sets from identically constructed ropes. All experiments were performed on real-life rope data. The results prove a high generalization ability, as well as a good robustness to outliers in the training set. The presented approach can exclude up to 90 percent of the rope as faultless without missing one single defect.

1 INTRODUCTION

Wire ropes are used in many fields of logistics. They are deployed as load cable for bridges, elevators and ropeways. This implies a high strain by external powers every day. Unfortunately, this can lead to structural anomalies or even defects in the rope formation. A defective rope bears a high risk for human life. This motivates the strict rules summarized in the European norm (EN 12927-7, 2004), which instruct a regular inspection of wire ropes .

Risky defects, prominent in wire ropes, are small wire fractions, missing wires, and damaged rope material due to lightening strokes. Furthermore, structural anomalies caused by interweavement of the rope ends or a reduced stress are also in the focus of interest. In figure 1, two exemplary defects are marked in the rope. Visual inspection of wire ropes is a difficult and dangerous task. Besides, the inspection speed is quite high (on average 0.5 meters/second) which makes it a hard effort, to concentrate on the passing rope without missing small defective rope regions.

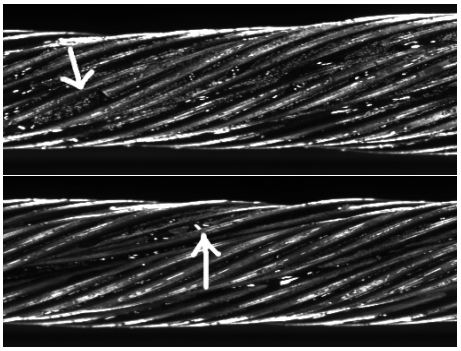

Figure 1: Rope defects: in the upper image you can see a wire fraction and in the image below a wire is missing.

A prototypic acquisition system was developed to

This manuscript is originally published in *VISAPP 2009: Proceedings of the 4th Int. Conf. on Computer Vision Theory and Applications, 2009, pp 171–178.*
www.visapp.org/VISAPP2009

overcome these limitations (Moll, 2003). Four line cameras record the passing rope and yield four different views. By this, the rope can be inspected in the office without time pressure. The images in figure 1 were acquired with this system.

Defects and anomalies in wire ropes are unimposing and small. The image quality is deranged by mud, powder, grease or water and the lighting conditions change frequently. Therefore, the choice of features for the detection task is important. Recent approaches for defect or anomaly detection focus on fault detection in material-surfaces. In (Platzer et al., 2008) we introduced a one-class classification approach for anomaly detection in wire ropes using linear prediction (LP) coefficients as features and a Gaussian mixture for model learning. The former work is extended in this paper by two main aspects: the performance of LP features is compared to that of other well-established features from the field of textural defect detection, and the robustness to outliers in the training set as well as the generalization ability of the presented approach are carefully evaluated. The last point is of particular interest for the practical relevance of the method.

Features based on local binary patterns (LBP) were first introduced by Ojala (Ojala et al., 1996) for texture classification. Recently they were used for defect detection in fabrics (Tajeripour et al., 2008) and for real-time surface inspection (Mäenpää et al., 2003). Textural features, extracted from co-occurrence matrices, were proposed by Harlick in the early 70's (Harlick et al., 1973) and are frequently used for texture description (Chen et al., 1998). (Iivarinen, 2000) compares two histogram-based methods for surface defect detection using LBP and co-occurrence matrices. (Rautkorpi and Iivarinen, 2005) used shaped-based co-occurrence matrices for the classification of metal surface defects. (Vartiainen et al., 2008) focus on the detection of irregularities in regular, periodic patterns. They separate the image data in a regular and an irregular part. Based on the resulting irregularities, we compute local histograms, which serve as features. Another important category of features for texture analysis and textural defect detection are wavelet-based features. (Kumar and Pang, 2002) for example use Gabor features for the detection of defects in textured material. However, the computation of these features requires large filter banks and high computational costs. Due to the huge size of rope data sets (20-30 GB) the time-consuming computation of Gabor features seems to be not the best choice. In (Varma and Zisserman, 2003) the authors state, that similar results to that obtained by the usage of wavelet features can be resolved with help of joint neighborhood distributions and less computational effort.

The one-class classification strategy proposed in (Platzer et al., 2008) was chosen due to a lack of defective training samples for supervised classification. In contrast, it is no problem to design a huge sample set of faultless training samples. With this faultfree training set a model of the intact rope structure can be learned and in the detection step outliers with regard to this model are classified as defect. However, the only available ground truth information about this training data is the labeling of the human expert. In the following, there remains a small uncertainty of underdiagnosed defects in the training set. For this reason the robustness of the proposed method to outliers in the training set is evaluated. Results obtained by learning from a faultless training set are compared to those, obtained by learning from a training set with intentionally added, faulty samples. The generalization ability of a learned model is a further important point, especially for the practical relevance of the presented method. There exist only a limited number of different construction types for wire ropes. The differences between them are mainly a different number of wires and strands, different thickness of single wires, the length of twist and the diameter. If just one model for every possible rope type would have to be learned in advance, this would save a lot of computational effort. However, the rope data from different ropes differs significantly due to the changing acquisition conditions and a different mounting of the ropes. Nevertheless, it is desirable to have just one model for every construction type and to overcome the challenges of a changing acquisition environment. Therefore, the generalization ability of the learned models is evaluated by learning and testing on different rope data from nearly identical constructed ropes.

The paper is structured as follows: in section 2 we briefly review the feature extraction using linear prediction and give a description of the used textural features and their extraction. The one-class classification of wire rope data is shortly summarized in section 3. Experiments, revealing the usability and robustness of our approach, have been performed on real-life rope data and results are presented in section 4. A conclusion and a discussion about future work is given in section 5.

2 FEATURE EXTRACTION

In this section the different features are briefly reviewed. Their extraction from the underlying rope data is described, as it differs for the LP features in

contrast to the remaining features. In the following, a short motivation for the different features used in this context, is given.

Local binary patterns (LBP) code the local graylevel-structure of a pixel neighborhood. Histograms based on the resulting codes lead to a local feature distribution. Since local binary patterns incorporate contextual information from a local neighborhood, a comparison of their performance with that of the LP features is of particular interest.

(Harlick et al., 1973) introduced a set of 14 different textural features computed from co-occurrence matrices. They reveal the spatial distribution of graylevels an though seem to be an interesting choice for structures with a certain regularity, like wire ropes.

The detection of irregularities, proposed in (Vartianinen et al., 2008) focuses on anomalies in regular, periodic patterns. Since the structure of wire ropes is not perfectly periodic, but offers some regular periodicities, we used the detected irregularities for the computation of local, histogram-based features.

2.1 Linear Prediction Based Features

Linear prediction can be seen as one key technique in speech recognition (Rabiner and Juang, 1993). It is used to compute parameters determining the spectral characteristics of the underlying speech signal.

The behavior of the underlying signal is modelled by forecasting the value $x(t)$ of the signal x at time t by a linear combination of the p past values $x(t-k)$ with $k = 1,\ldots,p$, where p is the order of the autoregressive process. The prediction $\hat{x}(t)$ of a 1-D signal can be written as

$$\hat{x}(t) = -\sum_{k=1}^{p} \alpha_k x(t-k) \tag{1}$$

with the following prediction error

$$e(t) = x(t) - \hat{x}(t) = x(t) + \sum_{k=1}^{p} \alpha_k x(t-k). \tag{2}$$

This motivates the choice of linear prediction for feature extraction. For the prediction of the actual value the contextual information of the past values is used and is implicitly incorporated in the resulting feature.

Based on a least-squares formulation, the optimal parameters $\vec{\alpha} = (1, \alpha_1 \ldots \alpha_p)$ can be obtained by solving the normal equations (Makhoul, 1975). The optimal coefficients are derived by use of the autocorrelation method and the Levinson-Durbin recursion (Makhoul, 1975; Rabiner and Juang, 1993). Free parameters of this method are framesize and the order of the process. In experiments the optimal framesize

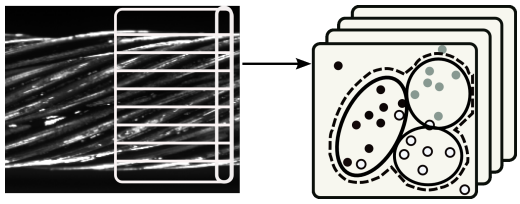

Figure 2: Multichannel version of the classification model. For every channel (horizontal white boxes) a feature is extracted and examined in a separate feature space. The vertical white box marks the signal values, which are actually predicted.

was found to be 20 camera lines with an incremental overlap of 10 lines. Best results were achieved for order $p = 8$.

Rope data, obtained from the acquisition system, can be seen as a sequence of 2-d images. Thus, with 1-d linear prediction it is not possible to analyze the 2-d signal. To overcome this, the rope data is considered as a multichannel time series. The signal $\vec{x}$ consists of c channels $\vec{x} = (\vec{x}_1\ \vec{x}_2 \ldots \vec{x}_c)^T$ and every channel represents a 1-dimensional time series $\vec{x}_i(t) = (x_i(1),\ldots,x_i(t))$. For every channel i of this signal an individual 1-d linear prediction is performed, leading to the estimate $\hat{x}_i(t)$, the squared prediction error for the whole frame, and the coefficient vector $\vec{\alpha}_i$. These components are used as corresponding feature for the actual frame and the channel i. Best results were obtained with a combined feature vector, including prediction coefficients and the squared error. In the training step, a separate model for every channel is learned. This is schematically depicted in figure 2. By this, the different appearance of the rope at different positions in the images is taken into consideration.

2.2 Local Binary Pattern

For the local binary pattern (LBP) a texture region is seen as a joint distribution of $P+1$ pixel-graylevels in a predefined neighborhood. Often a circular neighborhood with radius R and P equally spaced samples is chosen. The center pixel grayvalue g_c serves as threshold for the binarization of the neighborhood pixels g_p, $p = 1,\ldots,P$. The local binary pattern operator can be summarized as follows:

$$LBP_{P,R}(g_c) = \sum_{p=1}^{P} s(g_p - g_c) 2^{p-1} \tag{3}$$

$$\text{with } s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0\,. \end{cases} \tag{4}$$

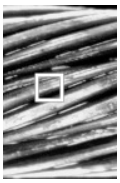

Figure 3: Detection window with 20×20 pixels in relation to the wire rope image data.

Transforming the binary vector into a decimal number (3) results in a pixel label, based on the neighborhood information. (Ojala et al., 2000) developed a rotational invariant and uniform extension of the local binary pattern. For the anomaly detection task there is no need for rotational invariance due to the constant rope orientation. The uniformity of the pattern is defined based on the number of 0/1 transitions U in the binary vector. The resulting LBP code is computed as follows:

$$LBP^{u}_{P,R}(g_c) = \begin{cases} \sum_{p=1}^{P} s(g_p - g_c), & U \leq 2 \\ P+1, & \text{otherwise} . \end{cases} \quad (5)$$

A histogram with a predefined number of bins is built from the underlying code distribution and serves as feature. The optimal parameters P and R and the number of quantization levels for the local histograms were determined in extensive experiments. We found the optimal parameter setting to be $R = 1$, $P = 8$ with 16 quantization levels for the histogram. As already mentioned, defects usually have just a small elongation. Hence, the histogram computation resulting in the feature vector is done for a small detection window (20×20 pixels), which moves over the underlying frame of rope data. To get an impression of the window size in relation to the rope data, a detection window is visualized in figure 3. By this, more than one feature is obtained for every frame.

2.3 Co-occurrence Features

Features for texture classification based on co-occurrence matrices were first introduced by (Harlick et al., 1973). A co-occurrence matrix is defined with respect to a certain displacement vector $\vec{d} = (d_x, d_y)$ and results in the joint distribution of co-occurring grayvalues. The relative frequency p_{ij}, which defines the co-occurrence of two neighboring grayvalues (with respect to $\vec{d}$) i and j, is defined as

$$p_{ij}(\vec{d}) = \lambda | \{(x,y) : I(x,y) = i, I(x+d_x, y+d_y) = j\} | \quad (6)$$

with $i, j \in \{0 \ldots G-1\}$ and G the number of gray levels. I represents an image of size $M \times N$ and λ is a normalization factor such that $\sum_{ij} p_{ij}(\vec{d}) = 1$.

Harlick introduced 14 different textural features (Harlick et al., 1973). Experiments for the determination of the most discriminative ones were performed. As the information theoretic texture features named difference entropy, information measure one, information measure two and the maximum correlation coefficient lead to the best results, a combination of these four features is used. Furthermore, a displacement vector of 2 pixels length with an angle of 90 degrees has led to the best results in our tests with varying parameters. As co-occurrence matrices lead to a global representation of the underlying texture, they are usually computed for a local region of interest. For the detection of small, regional anomalies in the rope structure this is important, as the small defect will not be recognized with global features. Again, a detection window of 20×20 pixels was used for the computation of the co-occurrence matrices and the following feature computation.

2.4 Features Based on Pattern Irregularity

(Vartianinen et al., 2008) describe an approach for the detection of irregularities in regular patterns based on the Fourier transform. Regular patterns result in distinct frequency peaks in the Fourier domain. By filtering out these frequencies and transforming the data back to the spatial domain, a perfectly regular pattern can be obtained. On the other hand, it is possible to substract this regular part from a unit function in the frequency domain, which results in the irregular part of the pattern:

$$\begin{aligned} I(x,y) &= \mathfrak{F}^{-1}(I(u,v)) && (7) \\ &= \mathfrak{F}^{-1}((\mathbf{1} - \mathfrak{M}(u,v) + \mathfrak{M}(u,v))I(u,v)) && (8) \\ &= \underbrace{\mathfrak{F}^{-1}(\mathfrak{M}(u,v)I(u,v))}_{\text{regular part}} && (9) \\ &\quad + \underbrace{\mathfrak{F}^{-1}((\mathbf{1} - \mathfrak{M}(u,v))I(u,v))}_{\text{irregular part}} \end{aligned}$$

$I(x,y)$ is the input image, $I(u,v)$ is the Fourier transformed image, $\mathfrak{F}$ is the Fourier transform and $\mathfrak{M}(u,v)$ is the filter function in the frequency domain. $\mathbf{1}$ represents the unit function. Without prior knowledge about the pattern structure a reasonable filter function is self-filtering (Bailey, 1993). Filtering is performed with the magnitude of the Fourier spectrum $\mathfrak{M}(u,v) = |I(u,v)|$. As the rope consists of regular structures, filtering is done with regard to the irregular part of the data. For the computation of the local histograms again a detection window of size 20×20

is used. Experimental evaluation of the used quantization levels for histogram computation showed, that a histogram with 16 bins performed the best.

3 ONE-CLASS CLASSIFICATION

In order to exclude as many rope meters as possible from further inspection, the theory of one-class classification seems to be a good choice. A separation between faultless and faulty samples is required. In this case, the faultless samples represent the target class ω_T and the defects are considered as outliers ω_O. As it is no problem to construct a large training set of defect-free feature samples, a representation for this target density $p(\vec{x} \mid \omega_T)$ has to be found without any knowledge about the outlier density $p(\vec{x} \mid \omega_O)$ (Tax, 2001). Here, $\vec{x}$ is the feature vector.

For one-class classification problems the false negative rate (FNR) is the only rate which can be measured directly from the training data. The false positive rate (FPR) is the most important measure for defect detection, but cannot be obtained without a sample set containing a sufficient number of defective samples. In case of a uniform distributed outlier density, however, a minimization of the FNR in combination with a minimization of the descriptive volume of the target density $p(\vec{x} \mid \omega_T)$ results in a minimization of the FNR and FPR (Tax, 2001).

There exist many different methods for one-class classification (also called novelty detection) (Hodge and Austin, 2004; Markou and Singh, 2003). In (Platzer et al., 2008) two approaches, namely the K-means clustering and a Gaussian mixture model (GMM), were compared. In contrast to our former work, where the training sample set contained only faultless samples, the learning step is now modified. Model learning is performed on a sample set with intentionally included samples from defective rope regions. The aim is to evaluate the robustness of the method against outliers in the training set. This would reduce the need of a human inspector, who determines an optimal, faultfree rope region for model learning.

3.1 Decision Making

For a classification into target class and outliers, a threshold is defined on the density. This threshold is based on the mean and the minimal probability reached in the training. It is stated, that an optimal threshold should be within the range of mean and minimum probability. As the training samples are all considered as defect-free, the minimum probability gives the lower bound for the likelihood of faultless samples. To account for possible outliers, the threshold is varied in this range and the evaluation is done by means of receiver operating characteristics (ROC). Since anomaly detection is a security relevant application, it is important not to miss any defect. As a consequence, the optimal threshold maximizes the TPR (number of samples correctly classified to the target class) while keeping the FPR zero.

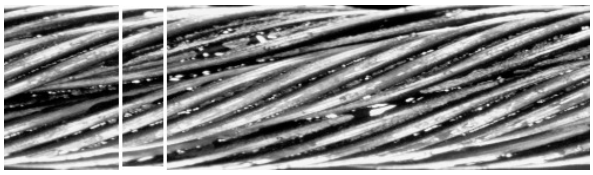

Figure 4: In this rope region a wire is missing. The window marks the frame, which was detected as an outlier by the described system.

Due to feature extraction with a detection window (or in case of the LP features based on one single channel), a rope frame consists of more than one feature. Accordingly, the decision for the overall frame is based on the decisions for the single windows/channels. In case of feature extraction by a detection window, the frame is classified as outlier, if one of the corresponding windows is rejected as outlier. For the channel-based LP features a further preprocessing is necessary. As the channels have no spatial extension like the detection window, one single channel is prone to noise. Therefore, a local channel-neighborhood consisting of 15 channels is scanned for potential outliers and only if the number of channel-votes exceeds five channels, the whole frame is rejected as defect.

Since a defect usually lasts over several frames, the whole defect is regarded as detected, if one frame in this range is rejected as outlier. Consequently, defects are detected but not localized at the moment. Figure 4 displays one defect, detected by the described system. The borders of the frame, detected as outlier are depicted by the window.

4 EXPERIMENTS AND RESULTS

In the following section experiments and their outcomes are presented. All experiments were performed on authentic rope data, acquired from real ropeways. In the generalization experiment (subsection 4.3) the data used for model learning was acquired in a controlled environment, but testing was again performed on data from a real ropeway. Model

learning was done with a Gaussian mixture composed of five mixture components and rope data belonging to one of the four views. Testing was performed on all four views and the resulting ROC curves were averaged over the different views. Interference between the views was not yet considered. The length of the used rope regions in all experiments is given by the number of camera lines, followed by the corresponding length in meters put into brackets. Learning on 20.000 camera lines (2m rope) of one view takes between 25 seconds and one minute on a Intel Pentium 4 with 3,4GHz, according to the choice of features. Surely, the LP model learning needs the most time due to the separate computation of one model for every channel. A testrun on 600.000 (60m rope) lines takes approximately between 14 (Irregularity) and 25 minutes (LP) per view, which gives an average detection speed of 14 till 23 seconds per meter of rope (10.000 camera lines). Again the computation time depends on the choice of features.

4.1 Comparison of Features

To compare the performance of the different features, model learning was done for every feature on the same training set, consisting of 20.000 lines (2m rope) of defect-free rope data from a real ropeway. Experiments were performed on a connected region of rope data, containing 600.000 camera lines (60m rope) and covering all known defects in the rope. The receiver operating characteristics in figure 5 point out, that the LP features outperform all other features. However, for all used features respectable results were obtained. Features based on detected irregularities perform the worst. A lot of noise is contained in the rope raw data and the structure is not perfectly regular, so that a certain amount of irregularities are detected in every frame. This results in a less discriminative behavior. Context-sensitive features like the LP features and the LBP features perform the best. Although their maximum TPRs are not the maximum ones obtained for an FPR of zero, their overall characteristics show a more robust behavior. The reason for the decreased true positive rate in both cases is one single, easily missed defect in one single view. This becomes also clear from table 1, where the maximum TPR for every view and feature, reached for a FPR of zero is listed. The underdiagnosed errors in view 1 for the LP features and view 3 for the LBP features were manually inspected. It was discovered, that these defects are present in more than one view and were correctly detected in the remaining views. It can be educed from this, that interference between the different views would eliminate this misclassifications. These results reveal the importance of context-sensitive features for this challenging task.

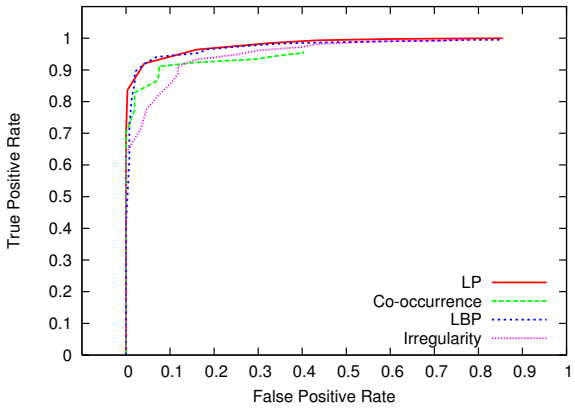

Figure 5: Comparison of the ROC curves for the different choice of features.

Table 1: Comparison of the maximum TPR for a FPR of zero for all features and all views.

Featuretype	view1	view2	view3	view4
LP	0.96	0.93	0.94	0.62
Co-occurrence	0.78	0.77	0.77	0.88
LBP	0.90	0.94	0.39	0.72
Irregularity	0.62	0.89	0.71	0.78

4.2 Robustness to Outliers

To evaluate the robustness to outliers in the training set, model learning was performed on a training set containing 200.000 (20m rope) lines of rope data. For learning, the view containing the most defects (9 defects) was chosen. Testing was performed on the remaining three views, also containing each at least seven defects. For comparison, the same experiment was performed with a model, learned from 200.000 defect-free camera lines (20m rope). The resulting ROC curves are compared in figure 6. The ROC curve in figure 6(a) gives the averaged ROC for the model, learned on defect-free training data. Figure 6(b) visualizes the results obtained with a model, learned from a training set including outliers. Obviously, the method is robust to few outliers in the training set, as the results differ not significantly from each other. The LP features show the most robust behavior to outliers in the training set1. As in the first experiment, one defect in one single view was missed at an early stage by the LBP, which results in a decreased maximum TPR for this feature. The size of the training

set was increased in the experiment, to incorporate as many defects as possible.

4.3 Generalization Ability

For the evaluation of the generalization ability, learning was performed on a real, faultless rope, acquired in a controlled environment. Testing on the other hand was performed on different rope data from a real ropeway containing defective regions. Both ropes belong to the same construction type and they only differ in their diameter by 10 pixels. In figure 7 the results are depicted by the corresponding ROC curves, averaged over all views. Figure 7(a) is generated by learning and testing on the same rope from the ropeway and figure 7(b) shows the result for learning in the controlled setup and testing on real-life rope data. In both cases the size of the learning set was 20.000 camera lines (2m rope). The LP features obviously show the best generalization performance, whereas the generalization performance of the LBP is the worst. Co-occurrence features as well as irregularity-based features show a surprising good generalization performance. In general, these results indicate a quite good generalization ability of the overall approach.

5 DISCUSSION AND OUTLOOK

We presented important and meaningful extensions of our former work. The one-class classification approach for anomaly detection in wire ropes, presented in (Platzer et al., 2008), was extended by three well-established features from the field of textural defect detection, and their performances were compared. The results emphasize the necessity of context-sensitive features for this challenging task. Although all compared features result in a good performance, the overall behavior of the context-sensitive features is better. With the presented approach 70 till 90 percent of the rope can be excluded from a reinspection by a human expert and only a region of 10 cm around a detected defect has to be re-examined again. Experiments emphasizing the robustness and generalization ability of the approach were presented. They pointed out that a perfect, faultless training set is not essential for model learning. Especially from the practical point of view, this is an important improvement, precisely because one cannot assure a completely defect-free training set. With regard to the generalization ability, the learned model was evaluated on real-life rope data of different, individual ropes, which belong to the same construction type. Results indicate a good generalization ability of learned models with respect to ropes, which are constructed in the same way. In the context of practical applicability this is a remarkable improvement, as it is a difficult and time consuming task to learn an individual model of the respective rope previous to every detection run.

In future, the focus will be turned to the defect localization. Using not only context-sensitive features, but also incorporating them into a context-based classification could lead to a further improvement of the system. Interference between the different views, acquired by the prototype system is also an interesting point under investigation. For both problems, hidden Markov models seem to be a good framework.

ACKNOWLEDGMENTS

This research is supported by the German Research Foundation (DFG) within the particular projects DE 735/6-1 and WE 2187/15-1.

REFERENCES

Bailey, D. (1993). Frequency domain self-filtering for pattern detection. In *Proceedings of the first New Zealand Conference on Image and Vision Computing*, pages 237–243.

Chen, C. H., Pau, L. F., and Wang, P. S. P., editors (1998). *The Handbook of Pattern Recognition and Computer Vision (2nd Edition)*. World Scientific Publishing Co.

EN 12927-7 (2004). Safety requirments for cableways installations designed to carry persons. ropes. inspection, repair and maintenance. European Norm: EN 12927-7:2004.

Harlick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621.

Hodge, V. J. and Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2):85–126.

Iivarinen, J. (2000). Surface Defect Detection with Histogram-Based Texture Features. volume 4197 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 140–145.

Kumar, A. and Pang, K. H. (2002). Defect Detection in Textured Materials Using Gabor Filters. *IEEE Transactions on Industry Applications*, 38(2):425–440.

Mäenpää, T., Turtinen, M., and Pietikäinen, M. (2003). Real-time surface inspection by texture. *Real-Time Imaging*, 9(5):289–296.

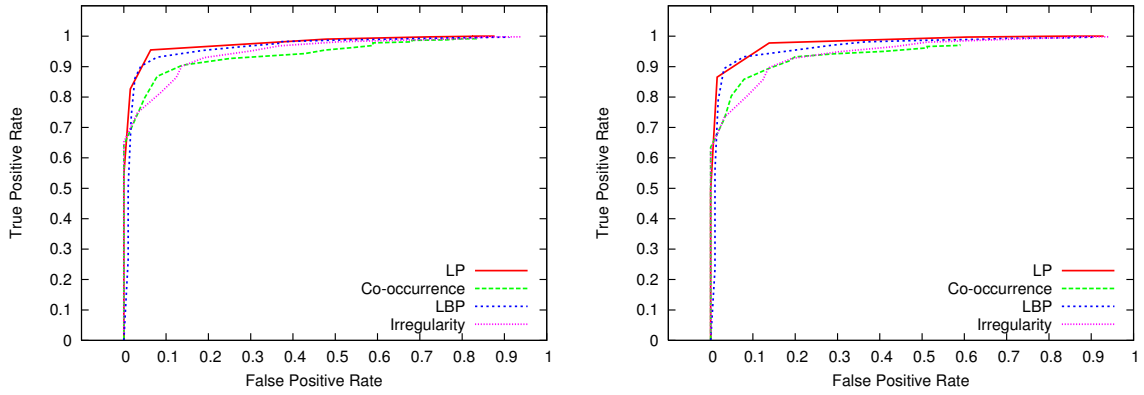

Figure 6: Comparison of ROC curves for learning with a defect-free training set (a) and a training set including defects (b).

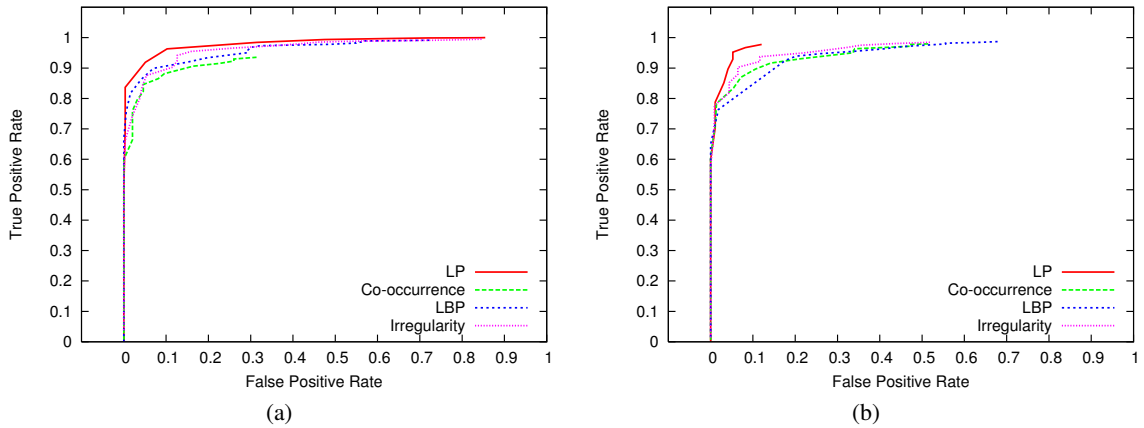

Figure 7: Comparison of the ROC curves for learning and testing on the same data set (a) and learning and testing on two different, identically constructed rope data sets (b).

Makhoul, J. (1975). Linear Prediction: A Tutorial Review. *Proceedings of the IEEE*, 63(4):561– 580.

Markou, M. and Singh, S. (2003). Novelty detection: a review - part 1: statistical approaches. *Signal Processing*, 83(12):2481 – 2497.

Moll, D. (2003). Innovative procedure for visual rope inspection. *Lift Report*, 29(3):10–14.

Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59.

Ojala, T., Pietikäinen, M., and Mäenpää, T. (2000). Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 404–420.

Platzer, E.-S., Denzler, J., Süße, H., Nägele, J., and Wehking, K.-H. (2008). Challenging anomaly detection in wire ropes using linear prediction combined with one-class classification. In *Proceedings of the 13th International Fall Workshop Vision, Modeling and Visualization*, pages 343–352.

Rabiner, L. and Juang, B.-H. (1993). *Fundementals of speech recognition*. Prentice Hall PTR.

Rautkorpi, R. and Iivarinen, J. (2005). Shape-Based Co-occurrence Matrices for Defect Classification. In *Proceedings of the 14th Scandinaian Converence on Image Analysis (SCIA)*, pages 588–597.

Tajeripour, F., Kabir, E., and Sheikhi, A. (2008). Fabric Defect Detection Using Modified Local Binary Patterns. *EURASIP Journal on Advances in Signal Processing*, 8(1):1–12.

Tax, D. M. J. (2001). *One-Class classification: Concept-learning in the absence of counter-examples*. Phd thesis, Delft University of Technology.

Varma, M. and Zisserman, A. (2003). Texture Classification: Are Filter Banks Necessary? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 691–698.

Vartianinen, J., Sadovnikov, A., Kamarainen, J.-K., Lensu, L., and Kälviäinen, H. (2008). Detection of irregularities in regular patterns. *Machine Vision and Applications*, 19(4):249–259.