Combining Structure and Appearance for Anomaly Detection in Wire Ropes

Esther-Sabrina Wacker[1] and Joachim Denzler[1]

Chair for Computer Vision, Friedrich Schiller University of Jena
Ernst-Abbe-Platz 2, 07743 Jena, Germany,
{esther.wacker,joachim.denzler}@uni-jena.de,

Abstract. We present a new approach for anomaly detection in the context of visual surface inspection. In contrast to existing, purely appearance-based approaches, we explicitly integrate information about the object geometry. The method is tested using the example of wire rope inspection as this is a very challenging problem.
A perfectly regular 3d model of the rope is aligned with a sequence of 2d rope images to establish a direct connection between object geometry and observed rope appearance. The surface appearance can be physically explained by the rendering equation. Without a need for knowledge about the illumination setting or the reflectance properties of the material we are able to sample the rendering equation. This results in a probabilistic appearance model. The density serves as description for normal surface variations and allows a robust localization of rope surface defects.
We evaluate our approach on real-world data from real ropeways. The accuracy of our approach is comparable to that of a human expert and outperforms all other existing approaches. It has an accuracy of 95% and a low false-alarm-rate of 1.5%, whereupon no single defect is missed.

Keywords: anomaly detection, image-based analysis, surface inspection

1 Introduction

Automatic surface inspection is a research area of rising interest. It is an important problem as the inspection task is an exhausting and monotonous work for a human with high quality claims on the other hand. In addition, surface analysis in general is a difficult problem, as the visual appearance of surfaces is highly subjected to various kinds of noise and changing lighting conditions.

A good example for such a task is the visual inspection of wire ropes. This is a very important problem, as damaged ropes pose a risk for the human life. Furthermore, the long, heavy ropes cannot be unmounted, are often contaminated

This manuscript is originally published in *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science, Volume 6855, Springer, 2011, pp 163–170.* The final publication is available at
http://link.springer.com/chapter/10.1007%2F978-3-642-23678-5_18

with *e.g* mud or oil and their material is highly reflective. In consequence, the surface appearance of an intact rope exhibits various characteristics. In contrast, defects in the surface structure are often very small and inconspicuous. Some examples for typical surface defects are displayed in the upper images of Fig. 5. Due to the high intra-class variability and the poor inter-class separability, a discrimination between defect and normal appearance variation is a difficult problem. Furthermore, a common problem of visual inspection tasks is the limited amount of available defective samples which hinders a supervised learning. For this reason, anomaly detection techniques [1,4], also known as one-class classification [9] have been used in the past for defect detection in material surfaces [8,11]. In general, these approaches are highly dependent on their choice of features used to represent the intact class. Platzer *et al* [7] compared the performance of different textural features for the problem of defect detection in wire rope surfaces. Their results underline the importance of context information for the problem of surface defect detection, especially with respect to the complex structure of wire ropes. In [6] Platzer *et al* focused on contextual anomaly detection by modeling the intact class with help of Hidden Markov Models. Haase *et al* [2] diagnosed contextual anomalies in the rope surface with help of an autoregressive model which predicts the intact surface appearance given its neighborhood. Nevertheless, no approach achieves the accuracy of a human inspector.

We state that the main reason for this is the lack of *geometrical* context in these purely appearance-based approaches. Therefore, we present a model-based approach for visual surface inspection. By fusing a geometrical structure model with a statistical appearance model we achieve a much better discrimination between a real defect and normal appearance variations. In a first step the model geometry is estimated in an image-based manner with help of a perfectly regular 3d rope model introduced recently by Wacker and Denzler [10]. In contrast to our work, they used this model to monitor important rope parameters but they did not address the problem of rope surface defect detection. We introduce a statistical appearance model which is linked to the geometric constraints implied by the rope structure. This allows a description of the surface appearance dependent on the position in the rope. Our method is data-driven and purely image-based. Moreover, we have no need for calibration information with respect to camera positions or the illumination setting.

The remainder of this paper is structured as follows: in section 2 the 3d model and the geometry estimation are summarized. Section 3 explains how this structural model can be linked to an statistical appearance model based on the *rendering equation*, which gives a physical explanation for light transport. Finally, section 4 turns to the problem of anomaly detection for defect analysis. A special focus will be laid on a validation strategy, which normalizes the learned appearance model with respect to small inaccuracies, which result from the geometry estimation step. Our experimental evaluation on real-world rope data is provided in section 5. Finally, conclusions are given in section 6.

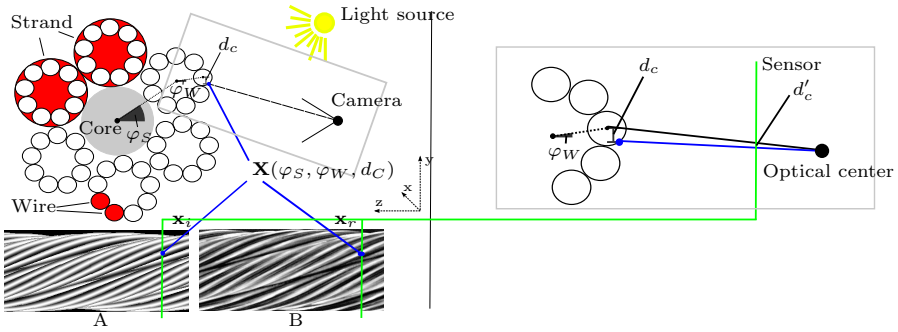

Fig. 1: Scenario sketch: given the point correspondence of a rope pixel $\mathbf{x}_r$ in the real rope image (B) and a rope pixel $\mathbf{x}_i$ in the aligned artificial model projection (A) a 3d surface point $\mathbf{X}$ of the rope can be parametrized by the two phase angles φ_S, φ_W and the 2d distance d'_c of $\mathbf{x}_i$ to its corresponding projected wire centerline. d'_c results from a 1:1 mapping of the unknown 3d distance d_c.

2 Geometric Rope Model

To estimate the rope geometry from 2d rope images, we use the framework described recently by Wacker and Denzler [10]. Their approach focuses on the image-based monitoring of important rope parameters and is not suitable for the automatic detection of surface defects.

A rope has a hierarchical structure composed of J strands $\mathbf{S}_j$ which comprise I wires $\mathbf{W}_i$. A wire centerline $\mathbf{W}_{i,j}$ of wire i in strand j for the time step t can be described by a sum of two parametrized helices:

$$\mathbf{W}_{i,j}(\mathbf{p},t) = \underbrace{\begin{pmatrix} t \\ r_S \sin(\varphi_S(\mathbf{p},t)) \\ -r_S \cos(\varphi_S(\mathbf{p},t)) \end{pmatrix}}_{\mathbf{S}_j} + \underbrace{\begin{pmatrix} 0 \\ r_W \sin(\varphi_W(\mathbf{p},t)) \\ -r_W \cos(\varphi_W(\mathbf{p},t)) \end{pmatrix}}_{\mathbf{W}_i}. \tag{1}$$

$\mathbf{p}$ is a vector of free model parameters and $\varphi_S(\mathbf{p},t)$, $\varphi_W(\mathbf{p},t)$ are the phase angles of the helices which are dependent on the model parametrization. The cross section through this model for one time step is shown in the top of Fig. 1.

By means of analysis-by-synthesis this parametric model is aligned with the digitally acquired 2d rope images. For that purpose an artificial 2d projection of the 3d rope model is computed. Real rope images and the artificial projections are then registered by optimizing the free model parameters in a non-linear fashion and these steps are repeated until convergence. We obtain a correspondence between a pixel $\mathbf{x}_i$ in the artificial projection and a pixel $\mathbf{x}_r$ in the real image.

In contrast to [10] we use this correspondence to form a parametric description of each surface point $\mathbf{X}$ in the rope. Fig. 1 clarifies that every 3d surface point can be described by the two phase angles φ_S and φ_W of the corresponding

wire centerline and the 3d distance d_c to this surface point (time is neglected):

$$\mathbf{X}(\varphi_S(\mathbf{p}), \varphi_W(\mathbf{p}), d_c) = \mathbf{W}_{i,j}(\mathbf{p}) + \underbrace{\begin{pmatrix} 0 \\ d_c \\ -\sqrt{0.5\varnothing_W^2 - d_c^2} \end{pmatrix}}_{\mathbf{n}'} \tag{2}$$

Here $\varnothing_W$ is the known diameter of the wires and $\mathbf{n}'$ points into the direction of the surface normal. As the rope model reveals no volumetric information d_c is unknown, but there exists a 1:1 mapping to the measurable 2d distance d'_c of an image pixel $\mathbf{x}_i$ to its corresponding projected wire centerline. Therefore, we will use the parametric description $\theta = (\varphi_S(\mathbf{p}), \varphi_W(\mathbf{p}), d'_c)$ to characterize a surface point in the rope and to build a combined model for structure and appearance.

3 Combined Model for Structure and Appearance

The rendering equation is a physical model describing the observed radiance at a surface point of an geometric object. It was first introduced by Kajiya [3] in 1986 and is an integral equation describing the propagation of light. One of the most common formulations of the rendering equation is:

$$L_O(\mathbf{X}, \omega_o) = L_E(\mathbf{X}, \omega_o) + \int_\Omega f_r(\mathbf{X}, \omega_i, \omega_o)\ L_I(\mathbf{X}, \omega_i)\ (\omega_i \cdot \mathbf{n}) d\omega_i. \tag{3}$$

The radiance which can be observed at a surface point $\mathbf{X}$ depends on the viewing direction ω_o, the emitted amount of light L_E and the reflected radiance which results from the incoming radiance L_I, the bidirectional reflectance distribution function f_r of the surface point and the inner product of surface normal $\mathbf{n}$ and the inward direction ω_i integrated over the hemisphere Ω.

Usually, in visual inspection scenarios we have neither calibration information nor knowledge about the illumination setting so that ω_o and ω_i are unknown. However, the relation between camera, object and position of the light source(s) typically stays fixed. This implies that the viewing direction and the incident angle of the incoming light depend only on the parametrization θ of the surface point $\mathbf{X}$, which we derived in section 2. Fig. 1 clarifies this scenario. In this case, the rendering equation can be re-parametrized and the emitting term L_E can be neglected for non-emitting objects like the rope:

$$\tilde{L}_O(\theta) = L_O(\mathbf{X}(\theta)) = \int_\Omega f_r(\mathbf{X}(\theta), \omega_i)\ L_I(\mathbf{X}(\theta), \omega_i)(\omega_i \cdot \mathbf{n}(\mathbf{X}(\theta)))d\omega_i, \tag{4}$$

Now, we are able to sample the observed irradiance L_O at a surface point $\mathbf{X}$ of the rope only dependent on its parametrization θ without additional knowledge about the camera position or the illumination setting. As our goal is the estimation of a *representative* surface appearance model including normal appearance variations, we exploit the periodic structure of a rope to obtain several

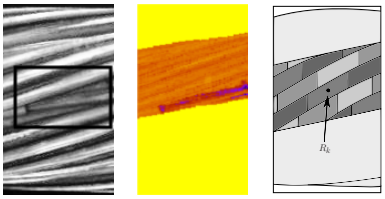

Fig. 2: Original rope image with defect (left), corresponding probability map (middle) for the strand with the defect and sketch of the rope regions (right).

samples for the same surface point. We consider a whole sequence of rope images which are aligned with the rope model for this purpose.

The appearance model is learned from an images of an intact rope. We are interested in the likelihood of observing a gray value g_r at the position $\mathbf{x}_r$ in the real rope image given its corresponding 3d surface point $\mathbf{X}(\theta)$. This can be formulated as a density estimation problem. We estimate the joint distribution $p(g_r, \theta)$ for any parametrization θ and its corresponding observed gray values g_r in a non-parametric manner. To obtain a dense representation we apply a 4d Parzen estimator. This density constitutes a combined model for appearance and structure, which allows to describe the normal surface appearance of each surface point subjected to the underlying rope geometry.

4 Defect Analysis

Once having learned the rope surface appearance model, the defect diagnosis can be treated as anomaly detection problem. Again, the input rope images must be aligned with the rope model to obtain the parametrization θ of each surface point. Subsequently, the appearance representation is extracted from the density $p(g_r, \theta)$ as a function of the position in the rope. A probability map can be computed which contains the likelihood of observing gray value g_r for a pixel $\mathbf{x}_r$ in the real rope image given its corresponding parametrization θ

$$p(g_r \mid \theta) = \frac{p(g_r, \theta)}{p(\theta)}. \tag{5}$$

Fig. 2 shows a real rope image including a typical defect on the left and its corresponding probability map for the strand of interest in the middle. The darker the color in the probability map, the smaller the obtained likelihood.

Nevertheless, an alignment of a rigid rope model with the flexible structure of a real rope leads to systematic registration inaccuracies which arise mainly in the border areas between two strands. In these regions a robust estimation of the appearance model is hindered. Hence, we normalize the appearance model with respect to these stability variations.

Different regions in the rope can be encoded with help of the two phase angles φ_S, φ_W of the 3d model. This allows a separation into K discrete region classes

R_k as sketched in the right hand side of Fig. 2. In order to increase the robustness of the appearance model with respect to systematic registration inaccuracies, we normalize the expectation of all rope regions. Hence, we compute the average likelihood $\overline{p}(R_k)$ for each rope region R_k and all N_k rope pixels belonging to R_k:

$$\overline{p}(R_k) = \frac{1}{N_k} \sum_{n=1}^{N_k} p(g_r^n \mid \theta^n). \tag{6}$$

This average is used to obtain a normalized likelihood according to (5):

$$\tilde{p}(g_r \mid \theta) = p(g_r \mid \theta) \, \frac{1}{\epsilon + \overline{p}(R_k)}. \tag{7}$$

$\epsilon > 0$ is a stabilization factor. The validation compensates for a systematic problem caused by the alignment of a rigid model with flexible real-world data. Thus, the normalization is data-independent and can be performed on the training set.

Finally, the resulting probability map for the input rope image including the normalized likelihoods $\tilde{p}(g_r \mid \theta)$ is filtered along the wire course. To transfer this soft classification result into a hard discrimination between suspicious changes and normal variations in the rope surface, a thresholding operation can be used.

5 Experiments

We evaluate our approach on real-world data taken from real ropeways under realistic acquisition conditions. Our data set comprises 400 meters of rope in total which corresponds to 7.7 GB of data. It was carefully selected by a human expert to ensure, that a maximum amount of appearance variations and surface defects are contained. The used system [5] operates with four line cameras, which are equally placed around the rope. A concatenation of the four individual 1d measurements results in four different 2d image sequences which are referenced as `view 1 - 4` from now on. Thus the amount of rope meters is quadrupled and the set of natural variations which occur during the acquisition process is augmented. The reference labeling is also provided by a human expert. The appearance model is trained on 5 m of rope which are known to be defect free. The remaining 395 m were used for testing.

5.1 Overall Performance

In order to evaluate the overall performance of our approach, we compute Receiver Operating Characteristic (ROC) curves for each sequence. The results can be seen in Fig. 3. The Area Under the Curve (AUC) value for each curve is given in the legend. The True Positive Rate (TPR) represents the total area of recovered defects and the False Positive Rate (FPR) relates to the false alarm rate (both measured in camera lines). As it is not sufficient to measure the error just as a function of the total length of detected anomalies we furthermore introduce

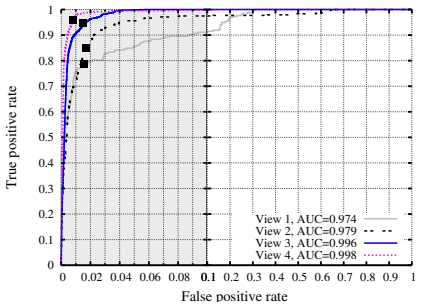

Fig. 3: Model-based approach: ROC curves and the 50% recovery of each defect marked by black squares.

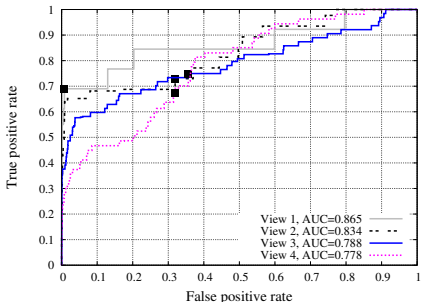

Fig. 4: Comparison to HMM-approach: ROC curves and 50% recovery of each defect marked by black squares.

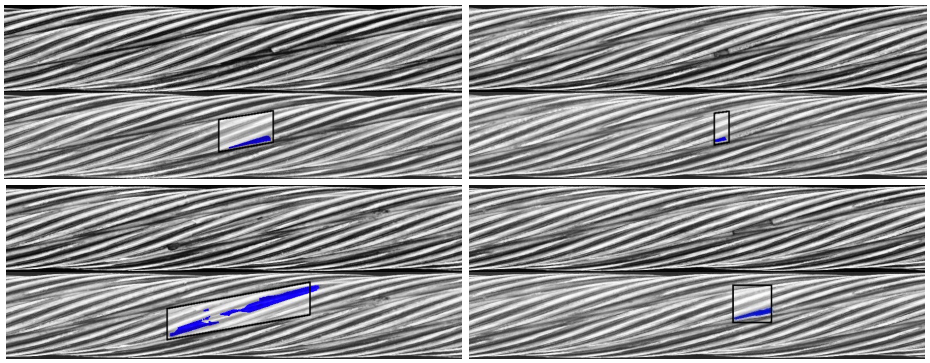

Fig. 5: Recovered defects: original rope image (upper image in each group) and result with recovered defect (blue) and ground truth labeling (black box).

the 50% recovery case. The black squares on each curve mark the recognition rates, which can be achieved if *every* known defect is recognized to at least 50% of its extent. Note, that these rates are bounded to the most inconspicuous defects in the sequence and the overall recognition rate is significantly higher than 50% in all cases. Keep in mind, that for the application it is not important to recover 100% of the defect area. But, it is crucial to recover *every single* defect to at least a certain extent while minimizing the FPR. In Fig. 5 some of our detection results are displayed. These results underline the high accuracy of the presented approach. As in most security relevant applications, the final decision must be made by a human expert who needs an image context of around 5 cm around each system alarm to judge weather it is a critical anomaly or a false alarm. With a false alarm rate of 1.5% for the 50% defect recovery case, a human expert would have to re-inspect only 103 m of the rope instead of 395 m.

5.2 Comparison to other Rope Defect Detection Approaches

We compare our results to the one obtained with the Hidden-Markov model (HMM) approach of Platzer *et al* [6] which leads to the best published results so far with regard to an individual analysis of each camera view.

Fig. 4 shows the ROC curves obtained on the same dataset with the HMM approach. Again the AUC values for each curve are given and the black squares mark the recognition rates obtained for the 50% recovery case of each defect.

It is obvious that our approach outperforms the HMM-based strategy. Particularly, in case of views 2–4 the HMM approach fails with an unfeasible high false alarm rate if the request is a detection of *every* single defect to at least 50%. But for a security-relevant task this claim is essential and this is not guaranteed by the existing approaches.

6 Summary and Conclusions

We presented a new approach for anomaly detection in wire ropes. The combination of a statistical appearance model with a parametric description of the object geometry leads to a position-dependent appearance representation. This combination allows a clearly enhanced discrimination between normal appearance variations and suspicious anomalies. One open question is the automatic determination of an optimal threshold. At the moment, the optimal threshold is evaluated with ROC curves, which always require a labeled data set.

Our results obtained on real-world rope data are very accurate and comparable to those of a human expert. We achieve low false alarm rates of 1.5% while fulfilling the claim that *every* single defect is recovered to a certain extent. This outperforms all existing approaches for automatic rope inspection and marks a clear improvement with respect to the practical applicability. Furthermore, our approach allows a precise localization of the defects.

References

1. Chandola, V., Banerjee, A., Kumar, V.: Anomaly Detection: A Survey. ACM Computing Surveys 41(3), 1–58 (2009)
2. Haase, D., Wacker, E.-S., Schukat-Talamazzini, E.-G., Denzler, J.: Analysis of Structural Dependencies for the Automatic Visual Inspection of Wire Ropes. In: VMV 2010: Vision, Modeling & Visualization. pp. 49–56 (2010)
3. Kajiya, J.T.: The rendering equation. ACM SIGGRAPH Computer Graphics 20(4), 143 – 150 (1986)
4. Markou, M., Singh, S.: Novelty detection: a review - part 1: statistical approaches. Signal Processing 83(12), 2481 – 2497 (2003)
5. Moll, D.: Innovative procedure for visual rope inspection. Lift Report 29(3), 10–14 (2003)
6. Platzer, E.-S., Nägele, J., Wehking, K.-H., Denzler, J.: HMM-Based Defect Localization in Wire Ropes - A New Approach to Unusual Subsequence Recognition. In: Proceedings of the 31st Annual Symposium of the German Association for Pattern Recognition (DAGM). pp. 442–451 (2009)

7. Platzer, E.-S., Süße, H., Nägele, J., Wehking, K.-H., Denzler, J.: On the Suitability of Different Features for Anomaly Detection in Wire Ropes. In: Computer Vision, Imaging and Computer Graphics: Theory and Applications. pp. 296–308. Springer (2010)
8. Tajeripour, F., Kabir, E., Sheikhi, A.: Fabric Defect Detection Using Modified Local Binary Patterns. EURASIP Journal on Advances in Signal Processing 8(1), 12 pages (2008)
9. Tax, D.M.J.: One-class classification - Concept-learning in the absence of counter-examples. Phd thesis, Technische Universität Delft (2001)
10. Wacker, E.-S., Denzler, J.: An Analysis-by-Synthesis Approach to Rope Condition Monitoring. In: Proceedings of the 6th International Symposium on Visual Computing (ISVC). LNCS, vol. 6454, pp. 459–468. Springer (2010)
11. Xie, X.: A Review of Recent Advances in Surface Defect Detection using Texture analysis Techniques. Electronic Letters on Computer Vision and Image Analysis 7(3), 1–22 (2008)