A Generic Framework for Defect Detection on Vessel Structures based on Image Saliency

Francisco Bonnin-Pascual and Alberto Ortiz
Department of Mathematics and Computer Science
University of the Balearic Islands
07122 Palma de Mallorca, Spain
{xisco.bonnin, alberto.ortiz}@uib.es

***Abstract*—Seagoing vessels have to undergo regular visual inspections in order to detect defects such as cracks and corrosion before they result into catastrophic consequences. These inspections are currently performed manually by ship surveyors at a great cost, so that any level of assistance during the inspection process would significatively decrease the inspection cost. In this paper, we describe a novel framework for visually detecting the aforementioned defects. This framework is generic and flexible in the sense that it can be easily configured to compute the features that perform better for the inspection at hand. In this regard, inspired by the idea of conspicuity, this work considers contrast and symmetry as features for detecting defects on vessel structures. A combination operator is additionally tested in order to merge the information provided by these features and improve the detection performance. Experimental results for the different configurations of the detection framework show better classification results than state of the art methods.**

I. Introduction

Vessels are nowadays one of the most cost effective ways to transport goods around the world. Despite the efforts to avoid maritime accidents, these still occur, and, from time to time, have catastrophic consequences in environmental, human and/or economic terms. Structural failures, such as cracks and corrosion, are the main cause of these accidents.

Classification Societies impose extensive inspection schemes in order to ensure the structural integrity of vessels. To carry out this task, the vessel has to be emptied and situated in a dockyard where high scaffoldings are installed to allow the human inspectors to access the highest parts of the vessel structure. This process can mean the visual assessment of more than 600,000 m^2 of steel. Besides, the surveys are on many occasions performed in hazardous environments for human operation. Moreover, total expenses involved by the infrastructure can reach up to one million dollars for certain sorts of vessels. Therefore, it is clear that any level of automation of the inspection process that can lead to a reduction of the inspection time/costs and an increase in the safety of the operation is fully justified.

The European project INCASS has among their goals the development of a micro-aerial vehicle fitted with cameras, to be used for the visual inspection of vessels [1]. The collected images are intended to be processed afterwards to detect the defective areas. Regarding the latter, this paper presents a novel approach for automatic detection of defects in images taken from vessel structures. A framework is proposed as a generic classifier that can be configured to make use of different features, potentially leading to different defect detectors each. Furthermore, the framework foresees the combination of the respective feature responses in order to enhance the overall output quality. The conspicuousness of defects in general, together with the kind of defects that can be expected in metallic surfaces (i.e. cracks and corrosion), have guided the feature selection process.

Literature contains just a few contributions regarding defect detection over vessel structures. For example, [2] presents a method to identify structural anomalies over image reconstructions of underwater ship hulls. Restricting to those contributions which use just visual sensors, [3] and [4] present detectors of cracks and corrosion for vessel structures. The drawback of these methods is that they require a previous training stage (e.g. to learn which is the color of corrosion) or tuning their working parameters (e.g. to know how thin and elongated must be a dark collection of pixels to be considered as a crack), whose value is typically related with the distance from which the images have been taken. The use of such previous stages is common among many contributions for defect detection in regular surfaces (see for example [5], [6]).

To the best of our knowledge, only one method has been published for generic defect detection in vessel structure images [7]. This approach makes use of a Bayesian framework to compute the probability for every pixel of corresponding to some kind of defective situation. This probability is based on the information learned in a previous training stage.

II. A Flexible Framework for Defect Detection

The design of the defect detector has been oriented towards a flexible framework which allows an easy integration of different features and their combinations. To attain this level of flexibility, we considered that the framework must cover the following aspects: (1) it should allow computing one or more features that are potentially useful to discriminate between defective and non-defective areas; (2) final features response should not depend on scale; (3) one or more combination operators should be available to merge the information provided by the features and try to find the combination (if any) that improves the classification performance; and (4) related to the previous point, one o more normalization operators should be available to adapt all the features responses to a certain range, in order to prevent loosing information when combining them.

This generic framework has been designed as a modular pipeline which involves different stages that can be configured (or even removed) depending on our needs, so that different configurations result into different defect detectors (see Fig. 1). Within the framework, each feature is intended to be computed as a different thread and the information that they all provide can be finally combined to make up the detection output.

The framework consists of the following stages:

- *Pre-feature computation*. The first stage provides the information necessary to compute the features. From an input color image one can obtain, for example, the gray-scale image, the saturation image (from HSV color space), etc. These images are called *pre-feature maps*.
- *Scale-space generation*. This stage scales the pre-feature maps using a range of scale factors to obtain a collection of pyramids. The computation of each pyramid level can include filtering the input map using some kind of kernel. One can compute, for example, a Gaussian pyramid which progressively low-pass filters and sub-samples the pre-feature map, an oriented Gabor pyramid for a preferred orientation θ, etc.
- *Feature computation*. This is the core stage within the pipeline. It is in charge of computing the value of the features for all the pixels of the image. Since this can be fed with one or more multi-scale pyramids, each feature can be computed combining the information provided at different scales. The outputs are called *feature maps*.
- *Normalization*. It normalizes the different feature maps to the same range of values to enable their combination.
- *Combination operator*. The last stage combines the normalized feature maps in order to obtain a single map, which is called the *defect map*. The mean and the median operators are some examples of simple combinations.

The resultant defect map is a single-channel map where defective areas are supposed to be labelled with higher values.

III. Detecting Defects on Vessel Structures

Vessel structures consist of large surfaces that usually present a regular texture. When these surfaces are inspected from a certain distance, a defect represents a discontinuity that alters the regularity of the texture. Based on that, texture-related features seem to be a good option to differentiate between defective and non-defective areas.

Furthermore, defects can also be considered as rare phenomena that may appear on such regular surfaces. Since they are rare, defects potentially attract the visual attention of the surveyor during a visual inspection process. Following these ideas, we propose to use texture-based features typically used in cognitive models to predict human eye fixations.

Among them, we focus on those which can be evaluated through a saliency map. This consists in a topographic map that represents the conspicuousness of the different areas of the input image [8]. This is typically shown as a gray-scale image where locations with higher conspicuity values are closer to white and less salient areas are closer to black. Notice that this representation fits with our definition of defect map.

Taking all these considerations into account, contrast and symmetry have been selected as the features for detecting defects on vessel structures. Three different versions of the defect detector have been developed: using only contrast, using only symmetry, and using a combination of both features.

A. The Contrast-based Defect Detector

Local contrast in intensity, color and orientation is typically used in computation models of attention, where these features are usually combined into a single saliency map (e.g. [9], [10]).

The contrast-based saliency model by Itti et al. [11] has been used as source of inspiration to design a contrast-based defect detector. For its implementation, each one of the stages of the generic pipeline has been particularized as follows:

- *Pre-feature computation*. Five pre-feature maps are computed: (a) an intensity map using $I = (r+g+b)/3$ with r, g and b being the red, green and blue channels of the input image; (b) a red channel map using $R = r - (g+b)/2$; (c) a green channel map using $G = g - (r+b)/2$; (d) a blue channel map by means of $B = b-(r-g)/2$; and (e) a yellow channel map using $Y = (r+g)/2-|r-g|/2-b$ (negative values are set to zero).
- *Scale-space generation*. Nine pyramids are generated: five Gaussian pyramids for the pre-feature maps (I, R, G, B and Y) plus four Gabor pyramids computed from the intensity map for orientations $\theta \in \{0°, 45°, 90°, 135°\}$. All the pyramids are computed to contain seven scales ranging from 1:1 (scale one) to 1:64 (scale seven).
- *Feature computation*. The contrast level in intensity, color and orientation is found as indicated in [11]. This process consists in computing center-surround differences between fine and coarse scales from the pyramids; that is, it computes the difference between each pixel of a fine (or center) scale c and its corresponding pixel in a coarse (or surrounding) scale s. In our implementation $c \in \{1, 2, 3\}$ and $s = c + \delta$, with $\delta \in \{3, 4\}$.
  Furthermore, a normalization operator $N(.)$ is used prior to combine the across-scale differences into three conspicuity maps – $\overline{I}$ for intensity, $\overline{C}$ for color and $\overline{O}$ for orientation – and also, before combining these maps to obtain the final saliency map as $S = 1/3(N(\overline{I})+N(\overline{C})+N(\overline{O}))$.

Normalization and combination stages are not employed for this version of the detector.

B. The Symmetry-based Defect Detector

A saliency model based on the Gestalt principle of symmetry was presented in [12]. In their paper, they discuss local symmetry as a measure of saliency and investigate its role in visual attention. The results suggest that symmetry is a salient structural feature for humans, as well as the suitability of their method for predicting human eye fixations in complex photographic images. Furthermore, their results show that,

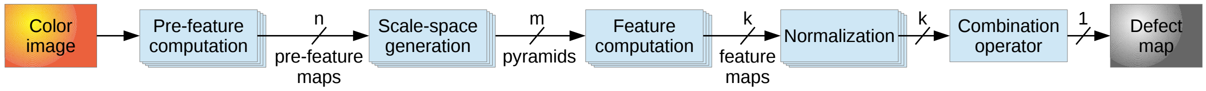

Figure 1. Generic framework for defect detection.

on many occasions, their symmetry operators outperform the contrast-saliency model by Itti et al. [11].

For all these reasons, symmetry is the second feature that has been selected for this study. To implement the symmetry-based defect detector, each stage of the generic framework has been particularized as follows:

- *Pre-feature computation*. It just computes an intensity map as $I = (r + g + b)/3$.
- *Scale-space generation*. This stage computes a simple sub-sample pyramid with five scales, ranging from 1:1 (scale one) to 1:16 (scale five).
- *Feature computation*. The symmetry value is computed for each level of the pyramid using the isotropic operator [12]. To obtain the final symmetry map, the five responses (one per pyramid level) are normalized using the normalization operator $N(.)$ and finally added together across-scale into an scale 1:1 map.

Normalization and combination stages are not employed for this version of the detector.

C. Combination of Contrast and Symmetry

In order to deeper explore the possibilities of the selected features, the generic framework has been configured once more to combine the information that they convey. To implement this version of the defect detector, contrast and symmetry maps are computed as indicated in the previous sections. The normalization stage is now configured to implement the normalization operator $N(.)$ [11], which promotes the areas labelled as potentially defective by any of the feature maps. To merge the normalized maps, we propose the linear combination: $OR = 1/2(\overline{Co} + \overline{Sy})$. This operator allows any defective point in any of the maps to be promoted so that it stands out in the final defective map. This version of the defect detector will be referred to as the *OR* configuration.

Figure 3 shows the generic pipeline configured to implement the three versions of our defect detector.

IV. Assessment of the Defect Detector

In this study, we have used a dataset comprising 73 images of vessel structures including defective areas (cracks, coating breakdown and different kinds of corrosion). The images have been collected at different distances and at different lighting conditions. This dataset is available online (http://dmi.uib.es/~xbonnin/resources) and also includes the ground truth, consisting in black and white images where defects are labelled in white (see Fig. 4:B).

In a first kind of experiments, we evaluated the performance of the proposed defect detectors. Figure 4 presents some examples of defect maps provided for the different cases, namely, the contrast-based detector, the symmetry-based detector and the version which combines both features.

At first sight, it can be observed that all the defect detectors tend to label as defective the areas that are indicated as defective in the ground truth image. This suggests that the different detectors can attain good classification rates.

In order to perform a quantitative evaluation, the True Positive Rate (TPR), or sensitivity, and the False Positive Rate (FPR), or fall-out, have been computed for the three defect detectors. To this end, the defect maps were thresholded for different values of a threshold τ to obtain the corresponding ROC curves. Furthermore, the values for the Area Under the Curve (AUC) have been calculated for all the ROC curves. These results are presented in Fig. 2 using dashed lines.

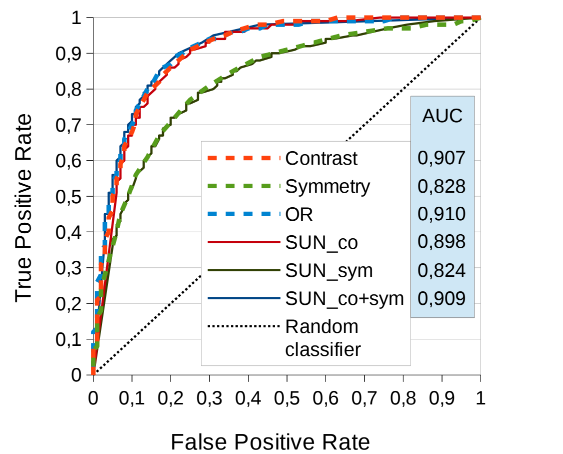

Figure 2. ROC curves and AUC values for the three versions of our defect detector (dashed lines) and the SUN-based classifiers (continuous lines).

Comparing the different ROC curves and AUC values, some interesting results can be stated: (1) the three new defect detectors present good performances during the classification task, with ROC curves relatively close to the (0,1) corner, and AUC values above 0.8; (2) contrast performs better than symmetry for the dataset employed in this study, what suggests that contrast provides more information to discriminate between defective and non-defective areas in vessel structures; (3) the detector which combines both features provides slightly better results than the version based only on contrast (i.e. symmetry provides complementary information).

In a second kind of experiments, the performances of the defect detectors presented in this paper have been compared with the one presented in [7]. This algorithm combines contrast and symmetry information through the Bayesian framework SUN [13] to provide a saliency value for every pixel in the image.

Notice that, despite both approaches use contrast and symmetry, the SUN-based detector requires from a training stage to estimate the probability distributions, and its feature combination is performed within a probabilistic formulation.

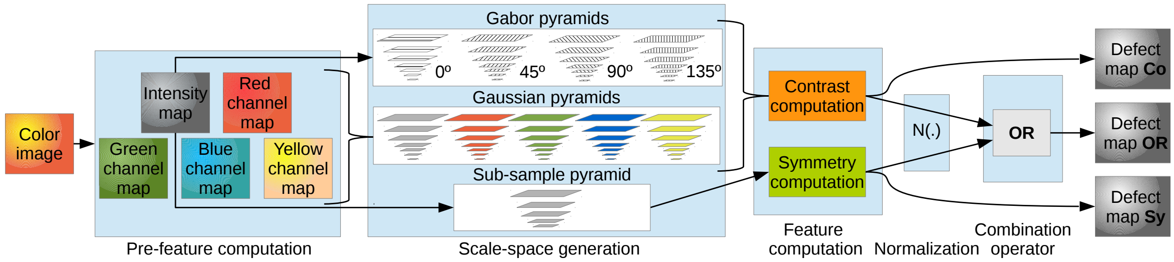

Figure 3. The three versions of the defect detector implemented using our generic framework.

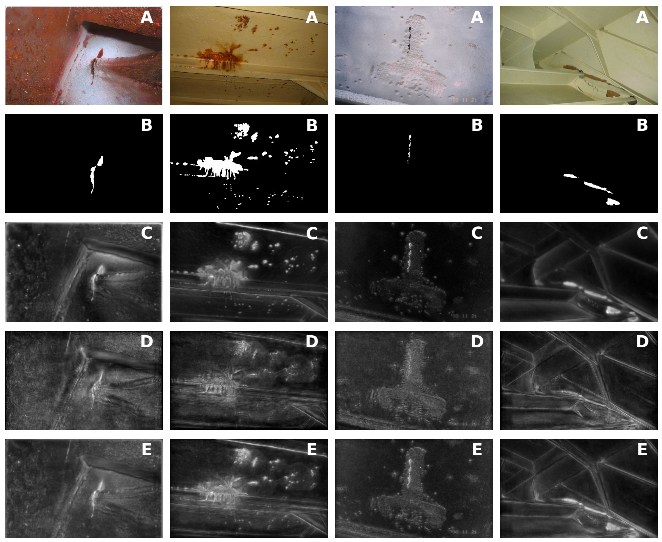

Figure 4. Test images (A) with their associated ground truth (B) and defect maps obtained using contrast (C), symmetry (D) and both features combined through the OR operator (E).

To perform the assessment, three different configurations of the SUN-based detector have been considered: using only contrast, using only symmetry and using both features. The three configuration have been evaluated through Leave-One-Out-Cross-Validation [14] and their corresponding ROC curves and AUC values have been computed and included in Fig. 2.

As can be observed, the results obtained with the defect detection framework presented in this paper are slightly better than the ones obtained using the SUN framework. These results are even better taking into account that our detectors do not rely on knowledge learned during a previous learning stage, as SUN-based detectors do.

V. Conclusions

A novel algorithm for defect detection on vessel structures has been presented. It has been devised as a generic framework that can be configured to select the features (and the way to combine them) that provide a more successful classification of the defective and non-defective areas. The selection of the features for our particular problem has been inspired by the idea of conspicuity and taking into account the kind of defects that appear in the metallic structures of vessels.

The different configurations of the defect detector have provided good classification performances. In comparison with previous methods, the presented framework does not require from tuning a large set of working parameters nor performing a previous training stage. Furthermore, the use of pyramids allows the algorithm to successfully detect the detective areas in pictures taken from different distances.

As future work, the defect detector is planned to be assessed with images collected using robotic devices (e.g. [1]).

Acknowledgment

This work is partially supported by Govern the les Illes Balears project nr. AAEE50/2015, and by project INCASS. The latter has received research funding from the EU FP7 under GA 605200. This publication reflects only the authors views and the European Union is not liable for any use that may be made of the information contained therein.

References

[1] F. Bonnin-Pascual, A. Ortiz, E. Garcia-Fidalgo, and J. P. Company, "A Micro-Aerial Platform for Vessel Visual Inspection based on Supervised Autonomy," in *IROS*, 2015, pp. 46–52.

[2] P. Ozog and R. M. Eustice, "Identifying Structural Anomalies in Image Reconstructions of Underwater Ship Hulls," in *OCEANS*, 2015.

[3] F. Bonnin-Pascual, "Detection of Cracks and Corrosion for Automated Vessels Visual Inspection," Master's thesis, UIB, 2010. [Online]. Available: {http://dmi.uib.es/~xbonnin/mthesis.pdf}

[4] F. Bonnin-Pascual and A. Ortiz, "Corrosion Detection for Automated Visual Inspection," in *Developments in Corrosion Protection*, D. M. Aliofkhazraei, Ed. InTech, 2014, ch. 25, pp. 619–632.

[5] M. Niederhöfer and V. Lohweg, "Application-based Approach for Automatic Texture Defect Recognition on Synthetic Surfaces," in *ETFA*, 2008, pp. 229–232.

[6] M. R. Jahanshahi and S. F. Masri, "Effect of Color Space, Color Channels, and Sub-Image Block Size on the Performance of Wavelet-Based Texture Analysis Algorithms: An Application to Corrosion Detection on Steel Structures," in *Computing in Civil Eng.*, 2013, pp. 685–692.

[7] F. Bonnin-Pascual and A. Ortiz, "A Probabilistic Approach for Defect Detection Based on Saliency Mechanisms," in *ETFA*, 2014.

[8] C. Koch and S. Ullman, "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry," *Human Neurobiology*, vol. 4, no. 4, pp. 219–227, 1985.

[9] A. Borji, M. N. Ahmadabadi, and B. N. Araabi, "Cost-sensitive Learning of Top-Down Modulation for Attentional Control," *MVA*, vol. 22, no. 1, pp. 61–76, 2011.

[10] L. Zhang, B. Qiu, X. Yu, and J. Xu, "Multi-scale Hybrid Saliency Analysis for Region of Interest Detection in Very High Resolution Remote Sensing Images," *IVC*, vol. 35, pp. 1–13, 2015.

[11] L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," *PAMI*, vol. 20, no. 11, pp. 1254–1259, 1998.

[12] G. Kootstra and L. Schomaker, "Prediction of Human Eye Fixations Using Symmetry," in *Conf. of the Cognitive Sci. Soc.*, 2009, pp. 56–61.

[13] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian Framework for Saliency Using Natural Statistics," *Jour. of Vision*, vol. 8, no. 7, pp. 1–20, 2008.

[14] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Wiley Interscience, 2000.