

A Statistical Approach to Industrial Anomaly Detection

Zhaohui Sun, Robert Kaucic, Paulo Mendoca, Ali Can
Visualization and Computer Vision Lab, GE Global Research
One Research Circle, Niskayuna, New York 12309, USA

Abstract

In this paper, we present a statistical approach to anomaly detection and monitoring through image analysis, and its application in non-destructive evaluation. A non-parametric statistical model is created by Parzen window density estimation at each pixel location, based on the observations of a number of defect-free images and the derived low level features. A test image is compared against the learned statistics and pixels not fitting the model are called out as defects, which include material changes and manufacturing defects. Image normalization and registration are carried out to factor out the image-to-image variations of appearance change and spatial misalignment, and anomaly detection on multiple views are combined for overall detection. The assisted defect recognition and monitoring can dramatically increase screening efficiency and consistency, thus improving quality control.

Keywords: Assisted defect recognition, image analysis, industrial inspection, non-destructive evaluation, statistical modeling, Parzen windowing, quality assurance, X-ray imaging.

1 Introduction

Non-destructive evaluation (NDE) of industrial parts is essential to manufacturing productivity and quality control. The internal structures of the manufacturing parts are probed by various imaging modalities, including X-ray, ultrasound, infrared, and eddy current, and digital images are captured and analyzed to identify potential defects. Assisted defect recognition (ADR) aims at automatically screening the images, calling out defects, and prioritizing the ones needing visual inspection. ADR can be applied to a wide variety of applications, including the turbine blades of jet engines, ships, and power generators. The goal is to achieve performance equivalent to or better than human operators, with much higher efficiency and consistency, thus enhancing quality control.

Assisted defect recognition has been under study for decades, and a number of systems have been reduced to practice in specialized domains [6, 1, 3, 5, 4, 2]. Among the different approaches, one [8] runs detection on 3D models after reconstruction from 2D images on CT data or from CAD models. Another approach [7] runs detection on 2D images directly. Image-based approaches use templates, that include geometry (e.g., shape and size), features (e.g., intensity, texture, color), and local statistics [9] to locate abnormalities. One particular interesting method, proposed in [7], uses learned bi-modal statistics for exhaustive defect detection.

In this paper, we propose a non-parametric statistical approach to assisted defect recognition following the Bayes rule, where the likelihood is approximated by a Gaussian mixture through

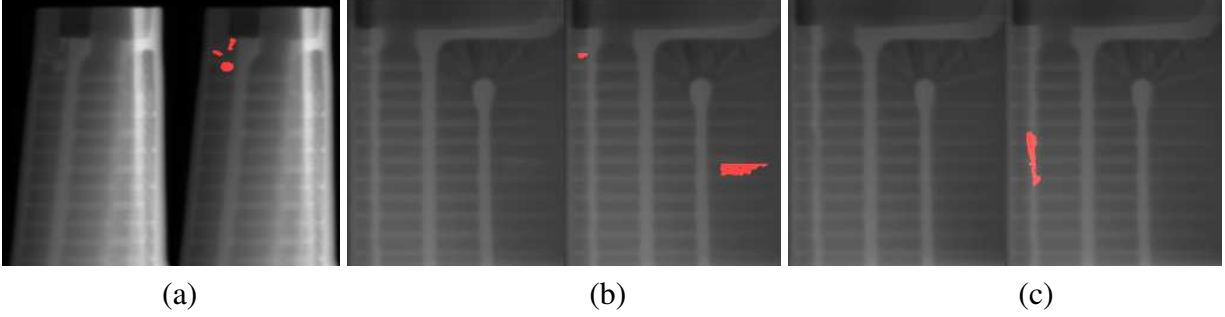


Figure 1: Defects of shot in (a), excess material in (b), and negative material in (c) in aircraft engine blade.

Parzen windowing and prior domain knowledge is used for anomaly detection. Images are spatially aligned by minimizing disjoint information [10] and spatially variant normalization is employed to factor out the image-to-image and part-to-part variations. The approach is validated on thousands of X-ray turbine blade images through ROC analysis. It can be easily adapted to different imaging modalities and applications.

The rest of the paper is organized as follows. Section 2 formulates the problem and general approach. Section 3 addresses image pre-processing, including image formation, normalization, and spatial registration. The non-parametric statistical approach based on the Bayesian rule is studied in Section 4. The approach is validated in Section 5, and the paper is concluded in Section 6.

2 The Problem and Approach

2.1 Problem Formulation

The anomaly detection task is to identify defect indications on the testing parts through non-destructive imaging and image analysis. Given the intensity/texture patterns from a large number of defect-free parts, how well does the current observation fit those patterns, what is the chance of testing parts having defect pixels, and how does the operating point balance the detection rate and false alarm rate based on the specifications? The goal is to increase quality control by driving inspection efficiency and consistency.

Given a set of defect-free images $\{I_n\}_{n=1}^N$, defect types $\{L_j\}_{j=1}^J$, and a test image $I(u, v)$, we would like to build a non-parametric statistical model $\mathcal{M}(u, v)$ and a labeling function

$$f : I(u, v) \Rightarrow L(u, v), \quad (1)$$

which maps image observation I to defect labels L at each pixel location (u, v) . In other words, the mapping function makes a decision of defect indication and defect type in the test image based on the statistical model \mathcal{M} built from the defect-free images.

In the following, we use X-ray images of aircraft engine blades to illustrate the ideas and validate the algorithms. Typical defects in engine blades include excess material (e.g. shots), negative material (e.g. casting defects), scarfs, over drills, and mis-drills. Examples of defects in X-ray images from a leading edge view and flat view are highlighted in Fig. 1, including shots, excess material, and negative material. We develop algorithms to automatically screen images

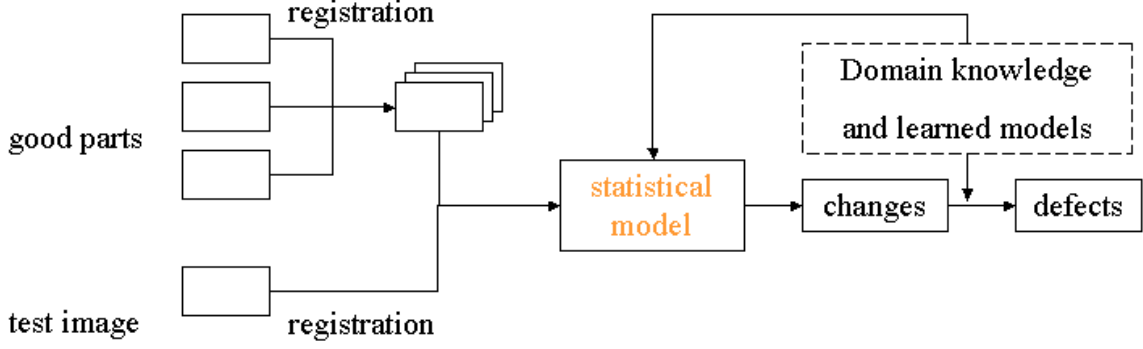


Figure 2: The statistical approach to anomaly detection.

for such defects, filter out the defect-free images, and prioritize the images which need visual inspection.

Part defects manifest themselves as image changes under appropriate imaging modalities, but changes in the image do not necessarily correspond to manufacturing defects. The part-to-part variation and image-to-image variation pose great challenge to automatic defect recognition systems. There are a number of image variation sources, including 1) Image misalignment, due to part placement, pose changes, and geometric distortion in the imaging process; 2) Image appearance change, due to variations in imaging source and detector; 3) Part variations within normal specification range; and 4) Part changes out of specification. Of all the potential changes, only the changes out of specification are of concern for defect recognition. How to factor out the irrelevant image changes and identify the defect indications with consistent performance becomes the biggest challenge. From a technical point of view, this is related to expressive feature extraction, feature selection, pattern classification, and performance validation.

2.2 The Approach

The proposed approach, as illustrated in Fig. 2, includes a modeling phase and a much faster evaluation phase, connected by the statistical model. The model is defined as

$$\mathcal{M} = (p_j(u, v), p_j^\alpha(u, v), I_T(u, v), I_0(u, v), Q_j, S_j, \sigma_j), \quad (2)$$

where $p_j(u, v)$ is the defect probability at pixel (u, v) , $p_j^\alpha(u, v)$ is the defect prior at pixel (u, v) based on the domain knowledge, I_T is a template image used for spatial alignment, I_0 is the baseline image for appearance normalization, j is the defect index, Q_j is the probability threshold separating normal from abnormal variations, S_j is the minimum defect size, and σ_j is the standard deviation of the Gaussian kernel. We use likelihood p_j and prior p_j^α for statistical analysis, template image I_T and baseline image I_0 for spatial and appearance normalization, and (Q_j, S_j, σ_j) as an operating point for defect detection.

The approach starts with building a statistical model \mathcal{M} from a large number of defect-free parts. A template image I_T is chosen from the good parts. The remaining images are spatially aligned with the template image by pairwise image registration. The baseline image I_0 is derived from the spatially aligned images, and the image intensities are normalized to compensate the variation in appearance. Low-level image features are extracted, and a non-parametric statistical

distribution $p_j(u, v)$ is created for each feature at pixel (u, v) . A connected component algorithm is carried out for the pixels having probabilities over threshold Q_j , and only regions larger than size S_j are kept for further analysis. This region analysis can alleviate incidental detection due to noisy image data, image misalignment, and image appearance change. The parameters of (Q_j, S_j, σ_j) are derived from ROC analysis based on a large set of defect-free images and a relatively smaller set of labeled defect images.

At the evaluation and screening phase, a test image I is aligned with the template image I_T , and normalized with respect to the baseline image I_0 . Low-level features are extracted and compared with the learned statistical model. Pixels are called out as defects if the probability is over threshold Q_j , the defect area size is larger than S_j , and the prior domain constraints are met. The process is repeated for all defect types until all potential defects are evaluated. Domain knowledge in the form of a prior distribution $p_j^\alpha(u, v)$ and learned statistics are further employed to confirm if image changes correspond to defect indications. When multiple images are taken at different viewpoints, the detection results are combined to make a final call.

The approach has a few advantages. 1) The statistics-based approach can be formulated in the Bayesian framework and therefore builds upon a sound foundation; 2) The image-based approach is separated from image formation and the application, and therefore very generic. It models the statistical distribution of the parts, not the physics behind image formation. The model requires no defect parts, which are often difficult to obtain. Therefore it is easy to customize to different applications by providing a large number of defect-free images and few, or no, labeled defect images; 3) It can be applied to different imaging modalities (e.g. X-ray, infrared, ultrasonic, eddy current, etc.) and various industrial parts (e.g. turbine blades, automotive parts, etc.).

3 Image Processing

Image processing is carried out to factor out irrelevant image-to-image changes, such as spatial misalignment and appearance changes. Minor pose variations in the part positioning and attenuation differences due to source and detector gain variations are accounted for by registering each part to a reference blade and normalizing the image using a spatially varying median filter.

Given two images A and B , the registration task can be formulated as

$$T^* = \arg \min_T \mathcal{D}(\mathbf{A}(\mathbf{s}), \mathbf{B}(T(\mathbf{s}))), \quad (3)$$

i.e., finding the optimal geometric transform T^* that brings the moving image \mathbf{B} into spatial alignment with the fixed image \mathbf{A} , where \mathbf{s} is a pixel site on the images. The geometric transform can be represented by different models, including translation, rotation, scaling, shear, affine, perspective, piecewise affine, and deformable.

A generic image registration method has been developed by minimization of disjoint information [10] in the ITK (Insight Toolkit) framework. It captures various geometric transforms (translation, rigid, similarity, and affine) and metrics (disjoint information, mutual information, normalized correlation, and mean squared error) in the same framework. The transform parameters and optimization step sizes can be initialized as input parameters, and a region of interest (ROI) can also be specified to limit the spatial region used for registration. The checkerboard patterns in Fig. 3(a,b) demonstrate the spatial alignment before and after registration, where adjacent blocks are taken from the reference image and the test images, respectively.

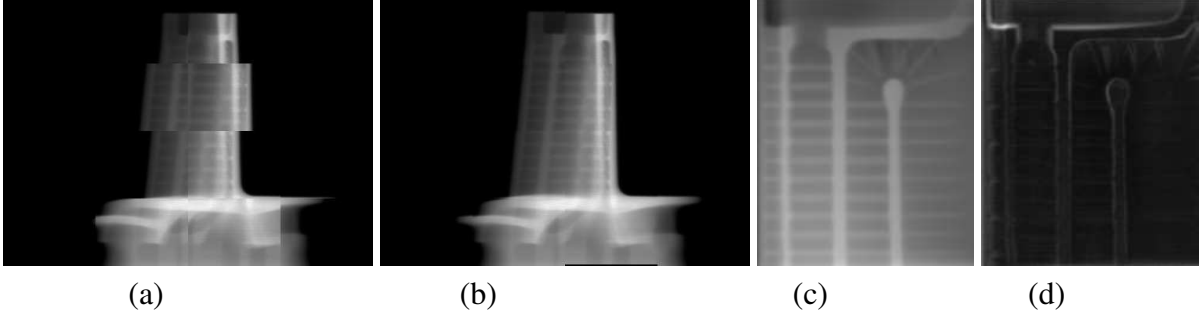


Figure 3: Checkerboard patterns before and after image registration in (a,b). Median and standard deviation of the defect-free images after normalization in (c,d).

The X-ray image is mainly used for qualitative evaluation. It is common to observe global appearance shift from image to image. This variation is compensated by applying spatially varying normalization. The images are spatially filtered using a 2-D median filter with a large radius, removing all the fine structures and details. In the modeling phase, the filtered intensity code values at pixel (u, v) from all the defect-free images are further filtered by a 1D median filter, generating a baseline image I_0 . In the evaluation phase, a test image I is spatially median filtered as I' , and compensated by the difference between I' and the baseline image I_0 , i.e.,

$$I'(T(u, v)) \Leftarrow I(T(u, v)) - F(I(T(u, v))) + I_0(u, v), \quad (4)$$

where (u, v) is the pixel location, T is the geometric transform to warp the test image I to the common coordinate system, F is the low pass median filtering, and I_0 is the baseline image. Examples of the median and standard deviation images after spatial alignment and normalization are shown in Fig. 3(a). The standard deviation is scaled up for better visualization. After compensation, the large uncertainty is mainly due to the part-to-part variation and registration error.

Figs. 3(c,d) give an example of the spatial and appearance normalization. The average of all defect-free images from the flat view is shown in (c). By checking the sharp edges, we see the registration achieves enough accuracy. Any misalignment will blur out the fine structures. The part-to-part appearance variation is shown in the standard deviation map in (d).

4 Statistical Modeling

Let y denote the state of image pixels being normal or abnormal and let x denote the image observation, in terms of the low-level image features, after spatial alignment and appearance normalization. The probability density function (PDF) of a test image being normal or abnormal given the extracted image feature can be written as

$$p(y|x) \propto p(x|y) \cdot p(y), \quad (5)$$

based on the Bayes rule, where the probability $p(y|x)$ is proportional to the likelihood $p(x|y)$ and prior $p(y)$. The details of these terms are discussed in the following.

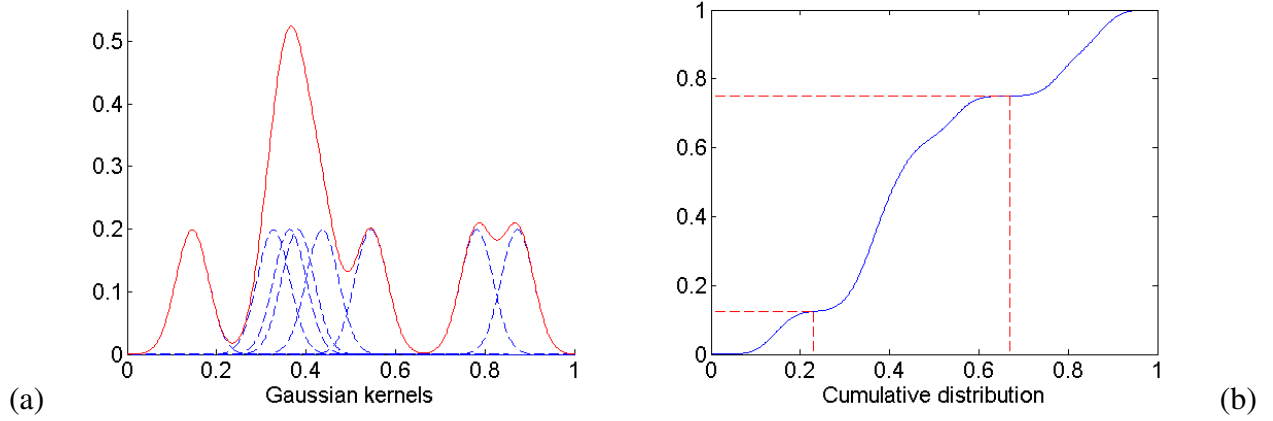


Figure 4: (a) Probability density function as a Gaussian mixture by Parzen window density approximation. (b) Cumulative density function.

4.1 The Likelihood

The likelihood $p(x|y)$ specifies the probability density function of image features x given image state y . When an image is represented by a set of conditionally independent low-level features (x^1, \dots, x^K) , such as image intensity and texture, the likelihood can be decomposed as

$$p(x|y) = p(x^1, \dots, x^K|y) = \prod_{k=1}^K p(x^k|y), \quad (6)$$

and the log likelihood becomes a summation

$$\log p(x|y) = \sum_{k=1}^K \log p(x^k|y). \quad (7)$$

In the following, we use only one image feature, the image intensity, and drop the feature index k for simplicity.

Given a total of N defect-free images, at each pixel location (u, v) , the probability distribution $p(x(u, v)|y(u, v))$ can be approximated by a mixture of local kernels [9]. When the kernel is chosen as a Gaussian distribution $G(\mu, \sigma)$ with fixed standard deviation σ , the PDF can be approximated by a Gaussian mixture through Parzen windowing,

$$p(x|y) = \frac{1}{N} \sum_{n=1}^N G(x|y; x_n, \sigma) = \frac{1}{\sqrt{2\pi}\sigma N} \sum_{n=1}^N e^{-\frac{(x-x_n)^2}{2\sigma^2}} \quad (8)$$

where x_n is the feature (e.g. intensity code value) of image n at pixel (u, v) after pre-processing.

As an example shown in Fig. 4(a), intensity code values x_n are taken from eight images at the same pixel location (u, v) . A Gaussian kernel $G(x_n, \sigma)$ is fit at each point, shown as the dashed blue lines. The summation of the Gaussian mixture yields the PDF in solid red. The idea is to take observation from a large number of pre-processed defect-free images and approximate the PDF at each pixel location by a Gaussian mixture. Anything out of normal range is considered as the potential defect. It is expected the observations have little uncertainty in the defect-free

regions and large uncertainty along structure boundary due to spatial misalignment and part-to-part variation, as shown in Fig. 3(a). Therefore, the statistical approach can effectively factor out process noise and irrelevant image changes for defect recognition. Checking Fig. 4(a), we see most of the samples concentrate in the range of $[0.22, 0.68]$. Samples with PDF over 0.68 or below 0.22 potentially are drawn from regions containing positive or negative materials.

The Gaussian kernel is effective for positive and negative material detection, by studying the tails at both ends. It is also possible to derive one-sided kernels for positive or negative detection, such as the one-sided Gaussian kernel,

$$G'(x; \mu, \sigma) = \begin{cases} 2G(x; \mu, \sigma) & \text{if } x \geq \mu \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

It can further reduce the false alarm rate by narrowing down the defect types.

4.2 The Prior

The prior $P(y)$ captures the domain and prior knowledge of the industrial parts and imaging process. A carefully designed prior can dramatically improve detection performance.

A simple way to enforce prior knowledge is to use an ROI mask. It is a labeled mask image, indicating the probability of defect appearing in certain image regions. For example, excess material tends to appear in the cavities and chambers, not on the thick walls. It can also be used to exclude certain regions, such as the background region and other regions not of interest. Such a mask image can be manually drawn, outlined from a CAD model. It can also be computed from image content. For example, the probability of having excess material is inversely proportional to the material thickness, i.e. the intensity code value. At last, the prior can be derived from the neighboring pixel sites using the Markov random field.

4.3 Anomaly Detection

Having computed the CDF from likelihood and prior in (5), we can evaluate the cumulative density function (CDF)

$$P(y|x) = \int_{-\infty}^y p(t|x)dt, \quad (10)$$

and use the CDF to derive the detection function f for anomaly detection.

For excess and negative material detection, the detection function is defined as

$$f(x) = \begin{cases} 1 & \text{if } P(y|x) > Q^1 \\ -1 & \text{if } P(y|x) < Q^2 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

When the $CDF(x)$ is greater than the threshold Q^1 or below Q^2 , the pixel is called out as a potential defect. For the PDF example in Fig. 4(a), the corresponding CDF is shown in Fig. 4(b). Pixels with a CDF over $Q^1 = 0.68$ are called out as potential excess material and pixels with a CDF below $Q^2 = 0.22$ are called out as potential negative material.

Connected component analysis is carried out on the potential defect pixels, and only regions with certain sizes are confirmed as defects. This can alleviate the incidental detection due to noise and drive down the false alarm rate.

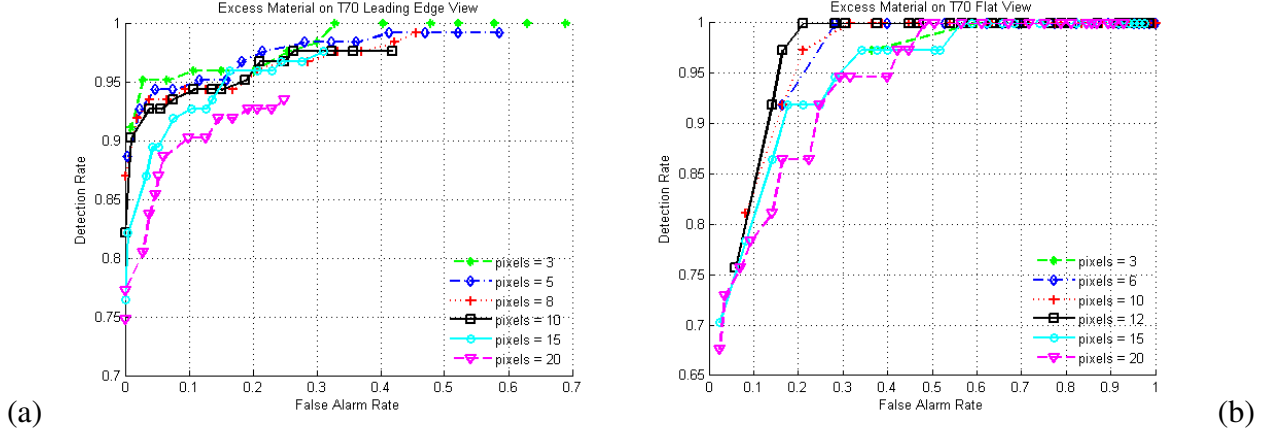


Figure 5: ROC analysis for (a) the leading edge view and (b) the flat view.

5 Experimental Results

The proposed statistical approach to industrial inspection has been applied to ADR of X-ray turbine blades used in aircraft engines and validated on several blade views and thousands of test images. ROC analysis has been carried out to pick the optimal operating point for different study cases.

For the leading edge view images shown in Fig. 3 and Fig. 1(a), 81 defect-free images are used to build the statistical model, and a total of 525 images are tested, including 382 defect-free images and 143 images with shots, excess material, and negative material. Each test image is spatially aligned with the template image, normalized in appearance, and compared against the learned statistics at each pixel location. We sweep the probability threshold Q (e.g. from 0.94 to 1 at a step size of 0.005) with fixed defect size S . Defection indications and types are called out at each step of Q . Compared with the ground truth, the results can be classified as true positive, true negative, false positive, and false negative. Evaluation of the detection rate and false alarm rate yields one point on the ROC curve. The ROC for excess material is shown in Fig. 5(a). Sweeping the full range of Q yields a specific ROC curve, and change of size threshold $S = 3, 5, 8, 10, 15, 20$ creates a set of ROC curves. From them, we can pick the best operating point to balance the detection and false alarm rates. The learned statistical model is applied to the test data and the results are reported. The overall performance is 95% detection rate and 8% false alarm rate. The detection includes defects due to material gain and loss. It takes about 5.5 seconds to review an image on a personal computer. The same approach has been applied to the flat view for validation. The ROC curves for positive material are shown in Fig. 5(b). Overall, it achieves 97.3% detection rate at 16.47% false alarm rate. The performance is not as good as that on leading edge view due to some very subtle linear structure defects. It is expected the results can be further improved by exploring the correlation between different views. The false alarm rate can also be improved by classification in a post-processing step.

Building the statistical model takes most of the computation, as it involves labeling the defect images, processing a large number of images, and running the ROC analysis to pick the best operating point. However, it is only done once per part. Testing and screening is very fast, normally within seconds. At current stage, the algorithms can filter out the defect-free images and prioritize the ones need visual inspection, thus increasing throughput and efficiency.

6 Conclusions

We have presented a non-parametric statistical approach to anomaly detection of industrial parts by non-destructive imaging and image analysis. The approach factors out the image-to-image appearance change, spatial alignment, and normal statistical variation, and calls out defects not fitting the learned statistics and prior knowledge. The algorithm has been validated on thousands of X-ray turbine blade images through ROC analysis. The statistical approach is very generic and can be easily customized to various imaging modalities and applications.

References

- [1] H. Boerner and H. Strecker. Automated x-ray inspection of aluminum castings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(1):79–91, 1988.
- [2] L. Fillatre, I. Nikiforov, and F. Retraint. A simple algorithm for defect detection from a few radiographies. *Journal of Computers*, 2(6):26–34, 2007.
- [3] D. Mery and D. Filbert. Automated flaw detection in aluminum castings based on the tracking of potential defects in a radiosopic image sequence. *IEEE Transactions on Robotics and Automation*, 18(6):890–901, 2002.
- [4] D. Mery, T. Jaeger, and D. Filbert. A review of methods for automated recognition of casting defects. *Insight*, 44(7):428–436, 2002.
- [5] J. Mundy, A. Noble, C. Marinos, V.D. Nguyen, A. Heller, J. Farley, and A.T. Tran. An object-oriented approach to template guided visual inspection. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 386–392, Champaign, IL, 1992.
- [6] T. S. Newman and A. K. Jain. A survey of automated visual inspection. *Computer Vision and Image Understanding*, 61(2):231–262, 1995.
- [7] Van-Duc Nguyen, Alison Noble, Joseph Mundy, John Janning, and Joseph Ross. Exhaustive detection of manufacturing flaws as abnormalities. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 945–952, Santa Barbara, CA, 1998.
- [8] J.A. Noble, R. Gupta, J. Mundy, A. Schmitz, and R.I. Hartley. High precision x-ray stereo for automated 3D CAD-based inspection. *IEEE Transactions on Robotics and Automation*, 14(2):292–302, 1998.
- [9] Chris Stauffer and Eric Grimson. Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):747–757, 2000.
- [10] Zhaohui Sun and Anthony Hoogs. Image comparison by compound disjoint information. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 857–862, New York, NY, June 2006.