THESE

pour l'obtention du grade de

DOCTEUR de l'UNIVERSITE
DE TECHNOLOGIE DE TROYES

Spécialité : OPTIMISATION ET SURETE DES SYSTEMES

présentée et soutenue par

Karim TOUT

le 6 avril 2018

Automatic vision system for surface inspection and monitoring: Application to wheel inspection

JURY

Rapporteurs:	
M. Jean-Michel MOREL	Professeur des universités, ENS Paris-Saclay
M. Ali MOHAMMAD-DJAFARI	Directeur de recherche du CNRS, L2S, Supélec Paris
Examinateurs:	
Mme. Christine FERNANDEZ-MALOIGNE	Professeur des universités, Univ. de Poitiers
M. Bachar EL HASSAN	Professeur, Univ. Libanaise
M. Florent RETRAINT	Enseignant-chercheur, UTT (Directeur de thèse)
M. Rémi COGRANNE	Maître de conférences, UTT (Directeur de thèse)


هذا من فضل ربي

Acknowledgments

I wish to express my gratitude to all those who helped and encouraged me during my doctoral studies.

First of all, I would like to express my deepest gratitude to my supervisors, Florent RETRAINT and Rémi COGRANNE, for their patience, motivation, and immense knowledge. I thank them for their highly professional guidance, unlimited support and unceasing encouragement. Their expertise in the domain as well as their valuable remarks have contributed the most part to the success of my PhD thesis. I highly value the friendly yet professional environment that they created during these last years of my PhD. I am extremely thankful and indebted to them for that.

Likewise, I would like to express my deepest gratitude to Henry BERNARD, Hugues DUGRES, Xavier GAUCHER, Jean-Yves JONET, Alexis BECK, Elouma EL AWN and all the personals at Mefro Wheels France, especially the QSE department, for the great collaboration during this PhD. Since day one, they supported me and granted me their full confidence, which led to the success of this PhD. I highly value their work ethics and the professional environment in their company. I learned a lot from them, especially in terms of teamwork.

I would like to express my special thanks to Mr. Jean-Michel MOREL and Mr. Ali MOHAMMAD-DJAFARI for accepting to review my PhD thesis. I would also like to thank Ms. Christine FERNANDEZ-MALOIGNE and Mr. Bachar EL HASSAN for agreeing to examine this thesis. Their valuable remarks and hard questions helped improve the quality of this thesis.

Thank you to my PhD student fellows of LM2S, Azzam, Kim, Canh, Hoai, and to all my Lebanese friends inside and outside UTT, especially my childhood friends Hussein and Samer. They deserve my sincerest thanks. Their friendship and assistance have meant more to me than I could ever express.

Last but not least, I need to dedicate this PhD thesis to my family. My parents Marwan and Mageda, my brothers Issam and Hassan, and my sister Hala, thank you for supporting me and encouraging me. Thank you for believing in me. Thank you for asking me every single day about my progress. And thank you for being there for me anytime.

Résumé

L'inspection visuelle des produits industriels a toujours été l'une des applications les plus reconnues du contrôle de qualité. Cette inspection reste en grande partie un processus manuel mené par des opérateurs et ceci rend l'opération peu fiable. Par conséquent, il est nécessaire d'automatiser cette inspection pour une meilleure efficacité. L'objectif principal de cette thèse est de concevoir un système d'inspection visuelle automatique pour l'inspection et la surveillance de la surface du produit. L'application spécifique de l'inspection de roues est considérée pour étudier la conception et l'installation du système d'imagerie. Ensuite, deux méthodes d'inspection sont développées: une méthode de détection des défauts à la surface du produit et une méthode de détection d'un changement brusque dans les paramètres du processus d'inspection non stationnaire. Parce que dans un contexte industriel, il est nécessaire de contrôler le taux de fausses alarmes, les deux méthodes proposées s'inscrivent dans le cadre de la théorie de la décision statistique. Un modèle paramétrique des observations est développé. Les paramètres du modèle sont estimés afin de concevoir un test statistique dont les performances sont analytiquement connues. Enfin, l'impact de la dégradation de l'éclairage sur la performance de détection des défauts est étudié afin de prédire les besoins de maintenance du système d'imagerie. Des résultats numériques sur un grand nombre d'images réelles mettent en évidence la pertinence de l'approche proposée.

Abstract

Visual inspection of finished products has always been one of the basic and most recognized applications of quality control in any industry. This inspection remains largely a manual process conducted by operators, and thus faces considerable limitations that make it unreliable. Therefore, it is necessary to automatize this inspection for better efficiency. The main goal of this thesis is to design an automatic visual inspection system for surface inspection and monitoring. The specific application of wheel inspection is considered to study the design and installation setup of the imaging system. Then, two inspection methods are developed: a defect detection method on product surface and a change-point detection method in the parameters of the non-stationary inspection process. Because in an industrial context it is necessary to control the false alarm rate, the two proposed methods are cast into the framework of hypothesis testing theory. A parametric approach is proposed to model the non-anomalous part of the observations. The model parameters are estimated to design a statistical test whose performances are analytically known. Finally, the impact of illumination degradation on the defect detection performance is studied in order to predict the maintenance needs of the imaging system. Numerical results on a large set of real images highlight the relevance of the proposed approach.

Table of Contents

1 Introduction **1**
- 1.1 General context . . . 1
- 1.2 Main objectives . . . 2
- 1.3 Organization of this manuscript . . . 3

2 Quality control **7**
- 2.1 Automotive industry . . . 7
  - 2.1.1 Evolution and challenges . . . 7
  - 2.1.2 Wheel manufacturing process . . . 8
- 2.2 Quality control . . . 10
  - 2.2.1 Quality control for wheels . . . 11
  - 2.2.2 Wheel surface inspection . . . 13
- 2.3 Non-Destructive Testing for surface inspection . . . 17
  - 2.3.1 Penetrant Testing . . . 20
  - 2.3.2 Magnetic Particle Testing . . . 22
  - 2.3.3 X-Ray Radiographic Testing . . . 24
- 2.4 Visual Testing . . . 26
  - 2.4.1 Automated visual inspection . . . 27
- 2.5 Conclusion . . . 28

3 Automated visual inspection system **29**
- 3.1 Inspection requirements and conditions . . . 30
- 3.2 Camera selection . . . 32
  - 3.2.1 Area scan or line scan . . . 33
  - 3.2.2 Monochrome or color camera . . . 35
  - 3.2.3 Sensor type . . . 37
  - 3.2.4 Resolution . . . 39
  - 3.2.5 Pixel size and sensor size . . . 41
  - 3.2.6 Interface . . . 42
- 3.3 Lens selection . . . 44
  - 3.3.1 Lens mount . . . 44
  - 3.3.2 Image circle and sensor size . . . 45
  - 3.3.3 Resolution and pixel size . . . 45
  - 3.3.4 Focal length . . . 46
  - 3.3.5 Aperture and depth of field . . . 47
- 3.4 Lighting system . . . 49
  - 3.4.1 Lighting technique . . . 50
  - 3.4.2 Lighting source . . . 52
- 3.5 Processing platform . . . 54
- 3.6 System design and installation . . . 56

3.7 Conclusion . . . 60

4 Defect detection method 61
4.1 Problem formulation . . . 64
4.2 Adaptive parametric linear model . . . 65
4.2.1 Background model . . . 65
4.2.2 Parameter tuning of background model . . . 68
4.2.3 Noise model . . . 70
4.3 Statistical detection of anomalies . . . 72
4.3.1 Statistical test . . . 72
4.3.2 Anomaly Detectability . . . 74
4.4 Wheel Inspection Characteristics . . . 75
4.4.1 Region of Interest Extraction . . . 76
4.4.2 Data Preparation . . . 78
4.5 Experiments and results . . . 80
4.5.1 Common core of all experiments . . . 80
4.5.2 Improvement of Detection Accuracy Due to the Model Adaptivity . . . 81
4.5.3 Improvement of Detection Accuracy Due to Heteroscedastic Noise Model . . . 82
4.5.4 Comparison Between Empirical and Theoretically Established Results . . . 85
4.5.5 Comparison With the State-of-the-art . . . 86
4.5.6 Real defects . . . 90
4.6 Conclusion . . . 94

5 Wheels coating monitoring method 95
5.1 Change-point detection problem statement . . . 97
5.1.1 CUSUM procedure . . . 97
5.1.2 Difficulties of non-stationarity and criterion of optimality . . . 98
5.2 Proposed change-point detection method . . . 99
5.2.1 Process modeling . . . 100
5.2.2 2FLW-SEQ procedure . . . 101
5.3 2FLW-SEQ performances . . . 102
5.3.1 Minimizing the probability of missed detection . . . 103
5.3.2 The worst-case probability of false alarm . . . 105
5.4 Paint coating intensity variation . . . 107
5.4.1 Painting process . . . 107
5.4.2 Topcoat monitoring . . . 108
5.5 Experiments and results . . . 109
5.5.1 Parameters tuning . . . 110
5.5.2 Maximal detection delay and detection performances . . . 112
5.5.3 Efficiency of the bounds and detectability performance . . . 113
5.5.4 Real case scenario of change-point . . . 114

5.5.5 Process modeling and detection criteria . . . 116
5.6 Conclusion . . . 119

6 AVI system maintenance 121
6.1 Illumination degradation . . . 121
6.2 Effects on the defect detection method . . . 122
6.3 Detection of light threshold . . . 126
6.3.1 Non-centrality parameter and pixel intensity . . . 126
6.3.2 Procedure to detect the lowest acceptable illumination level . . . 127
6.4 Experiments and results . . . 129
6.4.1 Common core of all experiments . . . 129
6.4.2 Accuracy of pixel intensity estimation . . . 130
6.4.3 Precision of the lowest acceptable pixel intensity . . . 132
6.5 Conclusion . . . 134

7 General conclusion and perspectives 137
7.1 Summary of the presented work . . . 137
7.2 Perspectives . . . 140

8 French Summary 143
8.1 Introduction . . . 143
8.2 Système d'inspection visuelle automatisé . . . 146
8.3 Méthode de détection de défauts . . . 148
8.3.1 Méthodologie . . . 150
8.3.2 Simulations . . . 154
8.4 Méthode de surveillance du revêtement des roues . . . 155
8.4.1 Méthodologie . . . 157
8.4.2 Simulations . . . 159
8.5 Maintenance du système vision . . . 161
8.5.1 Méthodologie . . . 162
8.5.2 Simulations . . . 163
8.6 Conclusion . . . 164

Bibliography 169

List of Figures

2.1 Image of the wheel and its two main components 9
2.2 Illustration of the disc stamping process 9
2.3 Illustration of the rim-forming process 10
2.4 Illustration of the two general cases of local appearance defects . . . 14
2.5 An example of a dimensional defect on the surface of the wheel . . . 16
2.6 General concept of NDT techniques 18
2.7 Principle of the penetrant testing technique 20
2.8 Principle of the magnetic particle testing technique 23

3.1 An example of a wheel to be inspected with an illustration of its design 49
3.2 Back lighting 51
3.3 Bright field lighting 51
3.4 Dark field lighting 51
3.5 Diffuse lighting 52
3.6 Illustration of the AVI prototype installation 58
3.7 Two images of the same wheel, one acquired with camera $n°1$ (a) while the other with one of the side cameras (b). 59

4.1 Illustration of the noise model showing, for several images, pixels' variance σ^2 as a function of their expectation μ 71
4.2 Illustration, in R^3 of observations along with their projections on the parity space $\mathbf{R}(\mathbf{H}_k)^{\perp}$ and a surface of constant power for which "anomaly-to-noise ratio" ϱ_k, see (4.18) is equal. 75
4.3 Description of the different elements of a wheel 76
4.4 Illustration of steps for pilot hole detection 77
4.5 Key element localization for two different types of wheels 78
4.6 Circular galbe 79
4.7 Typical example of the unfolded galbe of a wheel. 80
4.8 Example of a typical defect that it is wished to detect 81
4.9 Real power curves with $\ell = 3$ PCs and $\ell = 0$ PCs function of the defect intensity with a fixed defect radius $= 5$ 82
4.10 Real power curves with $\ell = 3$ PCs and $\ell = 0$ PCs function of the defect radius with a fixed defect intensity $= 150$ 83
4.11 Block 1 and Block 2 83
4.12 Empirical distributions of normalized residuals norm over block 1 and block 2 using the heteroscedastic noise model 84
4.13 Empirical distributions of normalized residuals norm over block 1 and block 2 using the AWGN model 85
4.14 Empirical and theoretical distributions of normalized residuals norm for images with and without defect 86

4.15 Real and Theoretical ROC Curves 87
4.16 ROC curves of the four detection methods 88
4.17 Detection probability function of the defect radius with a fixed defect intensity = 150 89
4.18 Images of wheels with various defects along with their corresponding model image and residual image 93

5.1 A typical example of variation of wheel images mean value 109
5.2 Empirical and theoretical cumulative distributions of the normalized residuals r_n with two different values of the first window size L with polynomial degrees $q-1=2$. 111
5.3 Empirical false alarm probability $\alpha(R)$ and detection power $\beta(M)$ over a run length $R=5000$ for 3 different values of M, plotted as a function of the decision threshold τ. 112
5.4 Empirical and theoretical false alarm probability $\alpha(R)$ over three different values of run length R, plotted as a function of the decision threshold τ. 113
5.5 Empirical and theoretical detection power $\beta(M)$ for 2 different values of M and 2 different values of false alarm rate $\alpha(R)$, plotted as a function of the change amplitude Δ. 114
5.6 Real example of mean value variation with a change-point at index 2434. 115
5.7 Result of the proposed 2FLW-SEQ detection method with $M=5$. . 115
5.8 Result of the proposed 2FLW-SEQ detection method with $M=3$. . 116
5.9 Empirical ROC curves for the proposed 2FLW-SEQ method and the CUSUM method with and without the polynomial model, computed over a run length $R=5000$, with a maximal detection delay $M=5$ and change amplitude $\Delta=60$. 117
5.10 Average detection delay as a function of the average run length to false alarm for the proposed 2FLW-SEQ method and the CUSUM method with polynomial model with a maximal detection delay $M=5$ and change amplitude $\Delta=60$. 117
5.11 Empirical detection power $\beta(M)$ as a function of the average run length to false alarm for the proposed 2FLW-SEQ method and the CUSUM method with polynomial model with a maximal detection delay $M=5$ and change amplitude $\Delta=40$. 118

6.1 Flowchart of the proposed procedure to detect the lowest acceptable illumination level 128
6.2 Images of the same galbe zone acquired at three different illumination levels 130
6.3 Estimation of the mean pixel expectation $\overline{\boldsymbol{\mu}}_i$ with the model (6.17), using the non-centrality parameter of image i, and the non-centrality parameter and mean pixel expectation of only the first image. 131

6.4 Empirical and theoretical detection power, along with the estimation of pixel expectation with the model (6.17), once using only the first image and then using images acquired periodically. 133
6.5 Estimation error for different reference images. 134

8.1 Installation du système vision 149
8.2 Deux images de la même roue, une acquise avec la caméra $n°1$, et l'autre avec la caméra $n°2$ 150
8.3 Distributions empirique et théorique de la norme des résidus normalisés pour des images avec et sans défauts 154
8.4 Courbes COR réelle et théorique 155
8.5 Images de roues contenant différents défauts, avec l'image du modèle de l'arrière-plan, et l'image résiduelle 156
8.6 Probabilité de fausse alarme $\alpha(R)$ et puissance de détection $\beta(M)$ avec $R = 5000$ pour 3 valeurs différentes de M en fonction du seuil de décision τ. 160
8.7 Exemple réel de la variation de la valeur moyenne avec un changement brusque à l'indice 2434. 160
8.8 Résultat de la méthode proposée 2FLW-SEQ avec $M = 5$ 161
8.9 Procédure proposée pour détecter valeur minimale acceptable d'éclairage 162
8.10 Puissance de détection empirique et théorique, avec les estimations de l'intensité des pixels à l'aide du modèle (8.29) en utilisant la première image ou des images reçues périodiquement. 164

List of Tables

2.1 List of the main potential defects with their characteristics 15

3.1 Hardware specifications 57

4.1 Proposed Adaptive Model Parameters 91

5.1 The empirical detection power $\beta(M)$ and the Hellinger distance H_D for different values of L and q. 111

8.1 Spécification du système vision 148

List of Abbreviations

Acronym	**W**hat (it) **S**tands **F**or
A/D	Analog-to-Digital.
ADD	Average Detection Delay.
AIA	Global Association for Vision Information.
ARIMA	Autoregressive Integrated Moving Average.
ARLFA	Average Run Length to False Alarm.
AVI	Automated Visual Inspection.
AWGN	Additive White Gaussian Noise.
CAD	Computer-Aided Design.
CCD	Charge-Coupled Device.
CDF	Cumulative Distribution Function.
CFAR	Constant False Alarm Rate.
CHT	Circular Hough Transform.
CMOS	Complementary Metal-Oxide Semiconductor.
CPU	Central Processing Unit.
CRF	Camera Response Function.
CT	Computed Tomography.
CUSUM	Cumulative Sum.
DR	Digital Radiography.
DT	Destructive Testing.
FMEA	Failure Mode and Effects Analysis.
FPGA	Field-Programmable Gate Array.
GPU	Graphics Processing Unit.
i.i.d	independent and identically distributed.
I/O	Input Output.
ISO	International Organization for Standardization.
LED	Light-Emitting Diode.
LS	Least-Square.
ML	Maximum Likelihood.
MT	Magnetic particle Testing.
NDE	Non-Destructive Evaluation.
NDT	Non-Destructive Testing.
PCA	Principal Component Analysis.
PCB	Printed Circuit Board.
PDF	Probability Density Function.
PHOT	Phase Only Transform.
PPM	Products Per Million.
PT	Penetrant Testing.
RGB	Red Green Blue.
RMSE	Root Mean Square Error.

ROC	Receiver Operating Characteristic.
ROI	Region Of Interest.
RT	Radiographic Testing.
SNR	Signal-to-Noise Ratio.
TQC	Total Quality Control.
UBCP	Uniformly Best Constant Power.
UMP	Uniformly Most Powerful.
VT	Visual Testing.
WLS	Weighted Least-Square.
2FLW-SEQ	Two Fixed-Length Windows Sequential Method.

Glossary of Notations

Notation	Definition
M	Number of lines in the inspected image.
N	Number of columns in the inspected image.
w	width of block in image $\mathbf{Z}$.
h	height of block in image $\mathbf{Z}$.
k	Index of block in image $\mathbf{Z}$.
z	Pixel value.
μ	Expectation value.
ξ	Noise value.
θ	impact of the anomaly on pixel's expectation.
σ	Standard deviation.
τ	Decision threshold.
δ	Statistical test.
α_0	False alarm probability.
α_1	Missed-detection probability.
$\beta(\delta)$	Power of the test δ.
ℓ	Number of added principal components.
r	Residuals.
ϱ	Non-centrality parameter.
(a, b)	Heteroscedastic noise model parameters.
$\widetilde{\boldsymbol{\mu}}_k^{\text{ls}}$	Least-square estimation of expectations in block k.
$\widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}}$	Estimation of the noise variance in block k obtained from $\widetilde{\boldsymbol{\mu}}_k^{\text{ls}}$.
$\widetilde{\boldsymbol{\mu}}_k$	Estimation of expectations in block k.
$\widetilde{\boldsymbol{\Sigma}}_k$	Estimation of noise variance in block k.
$\chi^2_\Upsilon(\varrho)$	The non-central χ-squared distribution with degrees of freedom Υ and a non-central parameter ϱ.
$F_{\chi^2_\Upsilon}(x, \varrho)$	The non-central χ-squared cumulative distribution function with degrees of freedom Υ and a non-central parameter ϱ.
$F^{-1}_{\chi^2_\Upsilon}(x, \varrho)$	Inverse of the non-central χ-squared cumulative distribution function with degrees of freedom Υ and a non-central parameter ϱ.
$\boldsymbol{S}$	Family of surfaces defined by $\boldsymbol{S} = \left\{ \boldsymbol{S_c} : \left\| \widetilde{\boldsymbol{\Sigma}}_k^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}_k} \boldsymbol{\theta}_k \right\|_2^2 = c^2 \right\}$.
x_i	Observation of index i.
N	Index of the last acquired observation.
Θ	Parameter space $\theta \in \Theta$.
p_θ	Probability density function associated to distribution $\mathcal{P}\theta$.
L	Length of the first window of the sequential method.
M	Length of the second window of the sequential method, and the maximal detection delay.

R	Run length.
K_M	Vector of size L that describes the change in mean value.
Δ	Amplitude of the change in mean value.
$\widetilde{S}^N_{N-L+1}$	Results of the sequential method.
Φ	Standard normal cumulative distribution function.
T	Stopping time.
v	Change point index.
$\widetilde{\beta}$	Lower bound of the detection power β.
$\widetilde{\alpha}$	Upper bound of the false alarm probability α.
H_D	Hellinger distance.

$\mathbb{E}$	Mathematical expectation.
$\mathbb{P}[E]$	Probability that an event E occurs.
$\mathbb{R}$	Set of real numbers.
$\mathbb{Z}$	Set of integer numbers.
$\mathbb{P}_{md}$	Worst-case probability of missed detection.
$\mathbb{P}_{fa}$	Worst-case probability of false alarm.

$\mathcal{Z}$	Set of pixel indices.
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2.
$\mathcal{H}_0$, $\mathcal{H}_1$	Null hypothesis and alternative hypothesis.
$\mathcal{P}\theta$	Probability distribution with parameter θ.

$\mathbf{Z}$	Inspected raw image.
$\mathbf{H}$	Parametric model of expectations.
$\mathbf{d}$	Contributions of basis vectors of model $\mathbf{H}$.
$\mathbf{P}$	Principal components added to the model.
$\mathbf{R}(\mathbf{H})$	Subspace spanned by the columns of $\mathbf{H}$.
$\mathbf{R}(\mathbf{H})^\perp$	Orthogonal complement of $\mathbf{R}(\mathbf{H})$, or parity space.
$\mathbf{P}^\perp_{\mathbf{H}}$	Linear projection onto $\mathbf{R}(\mathbf{H})^\perp$.
$\mathbf{I}_n$	Identity matrix of size $n \times n$.
$\mathbf{W}$	Rejection matrix of model $\mathbf{H}$.

Chapter 1

Introduction

1.1 General context

In the present highly competitive world class manufacturing scenario, where manufacturers in almost every industry sector find themselves competing with companies from every part of the world, customer satisfaction is a key element for survival and success. One critical determinant of customer satisfaction is the quality of the product. Guaranteeing the quality of produced products has always been a necessity to please the customers. Hence, many manufacturers tend to review the quality of their products not from their standpoint, but rather from the perspective of the customer.

Nowadays, customers are better informed and more attentive to what they perceive to be the quality of a product. While the ability of the product to perform its expected functions remains the main concern for the customer, its visual perception is also very important [1]. Manufacturers that reached high capabilities on the technical and functional aspects of production are now differentiated by their control of the perception of their product, specifically the visual perception of its surface. This lead to the necessity to carry out a visual inspection to ensure that each manufactured product meets the expected visual characteristics.

Visual inspection of finished products has always been one of the basic and most recognized applications of quality control in any industry. It is performed at the end of the production line, when all the manufacturing steps are completed. Its importance comes from the fact that it is the last step of inspection, and the last resort to stop a defective product before it reaches the customer. This inspection remains largely a (fully or partially) manual process conducted by operators, whose main role is to inspect each and every manufactured product. Unfortunately, this manual process faces considerable limitations which weaken it efficiency and make it unreliable for inspection.

While human inspection benefits from some advantages, mainly its great flexibility with regard to the various types and shapes of inspected products, it faces, on the other hand, many major disadvantages. First, the variability of decision for an inspector over time (factors related to fatigue and motivation) and the variability of decision between different operators for the evaluation of the same product. A second disadvantage for human inspection is the repeatability of decision of the same inspector when evaluating the same product various times.

These disadvantages result in an uncertainty and a lack of precision during the inspection.

To overcome these difficulties, automated inspection has proven to be the best alternative for industries to rely on. The share of variability and subjectivity in decision-making is eliminated with an automated surface inspection. When testing the same product multiple times with an automated inspection system, the result will be the same every time.

1.2 Main objectives

An automatic visual inspection (AVI) system is composed of two main subsystems. First, the image acquisition subsystem which is hardware based. It has the role of transforming the optical scene into an array of numerical data received by the processing platform. Second, the image processing subsystem which is software based. It mainly consists of image processing methods that are developed to analyze the acquired data and give the final inspection result.

The first objective of this work consists in designing the image acquisition subsystem. Indeed, the general requirements are that any AVI system must be fast, cost effective and reliable. However, to be able to put together an efficient AVI system, a more in depth description of the inspection requirements is necessary. It is absolutely essential for the designed system to be tailored to the specific demands of the application. In this respect, the complete design process must be driven by the application requirements and conditions. Therefore, it is important to specify the problem in a totally clear and precise way.

A second objective consists in developing a defect detection method for surface inspection. The method has to be general enough for the inspection of a wide range of complex surfaces, allowing it to be used in various quality inspection domains. In addition, in an industrial context, it is necessary to control the false alarm rate to prevent unnecessary stoppage in production, while ensuring the highest detection of defects. With these constraints in mind, it is best to resort to hypothesis testing theory to design a test whose statistical performances are analytically known. Furthermore, the method could make use of prior statistical information about the inspected surface to make the inspection more insensitive to variations in acquisition conditions.

A third objective consists in monitoring online a non-stationary process to detect abrupt changes in the process mean value. In our case, this non-stationary process results from the variation of the paint quantity on the inspected wheels surface, where the abrupt change corresponds to a sudden lack of paint. Since the monitored process is non-stationary, in other words, naturally changes over

time, the proposed method must be able to distinguish those "regular" process changes from abrupt changes resulting from potential failures. In addition, since this method aims at being applied for industrial processes, it is required to detect the change within a given maximal detection delay and to control the false alarm probability over a fixed run length. Hence, the proposed method falls within the category of online sequential detection method, operating, however, under a non-classical optimization criteria that is suitable for industrial applications.

A final objective consists in predicting the maintenance needs of the AVI system, required due to the degradation of illumination over time. On one hand this will enable to maintain the system in a well-functioning state, and, on the other hand, to prevent excessive stoppage in production for unnecessary maintenance operations. The state of the system is best judged according to the performance of the inspection tasks. Studying the impact of the degradation of illumination on the performance of the detection method will enable to estimate the illumination level below which the inspection is no longer efficient.

1.3 Organization of this manuscript

To address the objectives defined above, the overall structure of this thesis consists of the five following chapters:

- chapter 2 highlights the importance of quality control for automotive industries, and describes the techniques that are applied to inspect the manufactured wheels, specifically the inspection of their surface. This inspection is still largely a manual process that faces considerable limitations. With the goal to automatize this inspection process, this chapter then reviews the most widely used NDT techniques dedicated for surface inspection, investigating their suitability for the specific task of wheel surface inspection.

- chapter 3 presents in detail the complete procedure to design an AVI system for the real-time surface inspection of finished wheels. More precisely, this chapter focuses on the design of the imaging system, and its installation setup. First, the requirements and conditions under which the inspection must be performed are defined. Some of these requirements are related to the inspected wheel, while others are imposed by industrial constraints. Next, this chapter provides a general overview on all the characteristics of each key element of the AVI system. This overview serves as a brief reference to design a complete imaging system from scratch. Then, based on the predefined inspection requirements, a detailed discussion on the appropriate choice of each key element is presented. Finally, the installation setup of the AVI system is described.

- chapter 4 studies the defect detection problem on surfaces inspected using an imaging system. To control the false alarm rate, the proposed method relies on

hypothesis testing theory to design a test whose statistical performances are analytically known. The anomalous-free content of the image that represents the inspected surface acts here as a nuisance parameter as it has no interest for defect detection, while it must be carefully taken into account as it may hide potential defects. Hence, it is proposed to design an adaptive model of the imaged surface. This model allows to address the inspection of a wide range of objects, and it can be further used for any application that requires a model of the nuisance parameters.
Then, since no prior information about the occurrence of the defects is available, the problem is that of a statistical test between composite hypotheses. In such a case, a Uniformly Best Constant Power (UBCP) test is proposed, which is based on the rejection of the nuisance parameters. Finally, to take in consideration the noise corrupting the image, an accurate and realistic noise model is adopted. This will allow to establish with highest precision the theoretical statistical properties of the proposed test.
The proposed detection method is then applied for wheels surface inspection. Due to the nature of the wheels, the different elements are analyzed separately. Numerical results on a large set of real images show both the accuracy of the proposed adaptive model and the sharpness of the ensuing statistical test.

- chapter 5 addresses the problem of monitoring online a non-stationary process to detect abrupt changes in the process mean value. It starts by briefly recalling the well-known cumulative sum (CUSUM) procedure, and then states the problem of change-point detection for a non-stationary process emphasizing on the main difficulties and limitations of the CUSUM in this context. The main particularity of the problem addressed in this chapter is that the distribution parameters of a non-stationary process may "naturally" change over time, which follows that the detection hypotheses are composite. A usual solution is to use a generalized likelihood ratio that consists in substituting these unknown parameters by their estimations using the maximum likelihood estimation. Hence, at first, the model used to deal with observations' non-stationarity is presented. Second, the ensuing statistical test is detailed. Then, to comply with requirements on low false alarm probability and highest change-point detection performance under a maximal delay constraint, the performance of the proposed method is studied.
  The proposed sequential method is then applied for the problem of paint coating intensity variation on produced wheels. Numerical results obtained on a wide range of real data are presented, and the sharpness of the theoretical performance for the proposed method is studied.

- chapter 6 investigates the problem of predicting the maintenance needs of the AVI system. Since the imaging system consists of hardware components, the aging of such components could reduce the inspection performance, and even-

tually cause a failure. This chapter focuses on the degradation of illumination caused by the deterioration of LED modules over time. This degradation will have a negative impact on the detection method proposed in chapter 4. Hence, the approach proposed in this chapter is to study the impact of the degradation of illumination on the performance of the detection method. Modeling this impact will allow to estimate the illumination level below which the system is no longer reliable. Finally, a hardware solution is proposed to detect the moment when this illumination level is reached.

- chapter 7 concludes this thesis and presents some perspectives of future works.

Chapter 2

Quality control

2.1 Automotive industry

2.1.1 Evolution and challenges

The automotive industry has been undergoing a period of growth due to the ever-increasing global demand for automobiles. On average, the production of automobiles has been growing by 2.2% every year since 1975 [2], thus making the automotive industry one of the world's most important economic sectors by revenue.

On the other hand, the automotive industry has entered the phase of "globalization". The globalization of the market, the development of transport and communication means that today, a company has less and less geographical privileges in the market [2]. It must therefore be highly competitive in order to conquer distant markets and face greater competition.

With the ongoing competitive pressures, and in order to survive and stand out from the competition, the only alternative for automotive industries, including wheels manufacturing industries, is to grant lot of importance to the continuous improvement of their product, along with an increasing response to customer needs. What does the customer need ? "*Consumers tend to go for best quality products at nominal rate*s [3]". During the economic crisis of 2008, the automotive industry was among the sectors that were hit the most [4]. However, this crisis only lasted two years, after which the automotive industry recovered. This was due to many factors including the investment in improving the quality of the products. Indeed, developing high quality products will have a positive effect on different aspects of a company, and is essential for profits as well as for the survival in a competitive environment. The benefits of high-quality products extends even to other aspects of the industry:

- **Business aspect:**

  In order to benefit from a good competitiveness, the company can improve its quality / price ratio: at equal selling price it must increase the quality level, or at the same quality, the company must reduce the selling price, which reduces the costs of poor quality, improves its working methods and simplifies the design of its products.

  On the other hand, consumers are now more demanding in terms of quality. The certification of the quality assurance system applied by the company has

become today a significant commercial argument that reassures the potential customer on the level of quality of the company. It is more and more demanded by customers.

- **Financial aspect:**

  Quality costs are expensive. The term "Quality costs" has been extensively used to describe the costs of poor quality. A non-conformity is a waste not only of raw materials but also of labor working time and energy. Investing in quality control management has become a necessity for all industries. In his book [5], which is considered by many to be the main reference for quality control, Juran explained how and why investing in quality will eventually lead to both reductions in costs and increases in sales revenue.

- **Technical aspect:**

  Controlling the quality in an industry will not only prevent defective products to reach the customer, but will also help to characterize the defect types and their causes. This will lead to a better understanding of production techniques, and thus improvement in the manufacturing processes: lowering the costs of poor quality necessarily requires a review and improvement of manufacturing and production techniques.

2.1.2 Wheel manufacturing process

The wheel is an essential part of any vehicle. It is considered to be a security component that has no lifespan, which requires the manufacturers to pay particular attention to certain aspects of the product. In order to ensure its safety, the wheel must have a correct geometry allowing the vehicle to be stable, and must be robust enough to withstand the weight of the vehicle under all circumstances and to participate in the reduction of fuel consumption. On the other hand, the wheel is also considered to be an aesthetic component. Its overall appearance, such as its color, gloss, and texture, greatly affects the perception of the product to the customer. Thus, the image of the wheel has a direct impact on the image of its manufacturer.

Manufactured wheels for passenger cars and commercial vehicles usually consist of two parts: rim and disc [6]. The rim is the part of the wheel on which the tire is mounted and supported: international norms regulate most of its geometrical characteristics to guarantee functionality of pieces developed by different wheel and tire manufacturers. The disc is the part of the wheel which is the supporting member between the vehicle axle and the rim. Except for its central part, which is imposed by the customer, the disc profile is less standardized than the rim since it needs to be designed by taking into account several variable interfaces (brake calipers, trims, ...) and performance requirements (fatigue resistance, bolt hole resistance, ...)



Figure 2.1: Image of the wheel and its two main components

which are specific for each wheel design. Figure 2.1 shows a CAD illustration of the wheel and its two main components.

The manufacturing process of the wheel can be divided into three steps:

- **Disc stamping process:**

  The disc begins as a flat blank sheet of predefined size. It then follows a series of several stamping operations in order to reach its the final shape. Each disc is designed to satisfy specific customer's requirements, strongly influencing its geometry. Taking into account the main characteristics of wheel discs, you can consider the stamping process as a sequence of progressive deforming phases covering drawing, forming, flanging, cutting and coining operations. Figure 2.2 illustrates the disc stamping process.

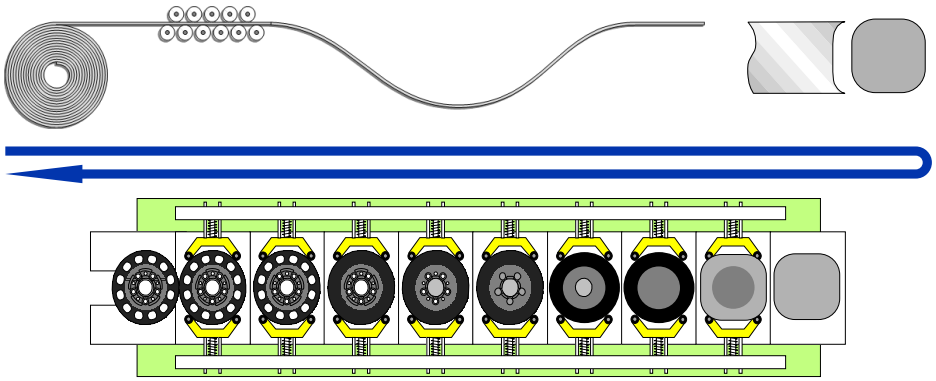


Figure 2.2: Illustration of the disc stamping process

- **Rim forming process:**

  A rectangular blank sheet of steel coil is cut with a predefined length to obtain the desired rim size. The starting blank sheet is curved to create a cylindrical shape where the edges are welded and the final welded surface is refined and

recalibrated to ensure the "roundness" of the piece. In order to obtain the desired rim profile, the previous cylindrical piece is deformed through multiple progressive rolling operations and then calibrated. A specific zone of the rim is locally deformed to create an opportune flat area; the valve hole is cut (centered on the flat surface) and coined to avoid burrs all around. Figure 2.3 shows the different steps of the rim-forming process.

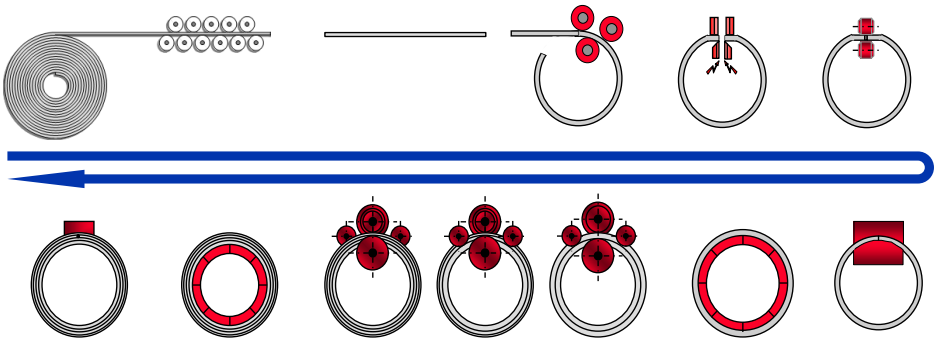


Figure 2.3: Illustration of the rim-forming process

- **Wheel mounting and painting:**

  When the disc and the rim are available, the wheel manufacturing process ends with the mounting phase, when the disc is pushed into the rim creating a forced fitting joint. Then the two components are additionally linked using different welding techniques. At the end, the wheel is painted through electro-coating (black cataphoresis); other colored paintings can be added, if requested, to improve the usual aspect of the final product.

2.2 Quality control

The aspect of quality control has been evolving since the beginning of time. Even the definition of the term "quality" has been extensively debated and modified to better follow this evolution. Its most recent definition by the International Organization for Standardization (ISO) is the "*degree to which a set of inherent characteristics of an object fulfills requirements* [7]".

Guaranteeing the quality of manufactured products has always been a necessity to please the customers. In the early days, and before the Industrial Revolution, the quality control was restricted to the people who made the product. They dealt directly with their customers and were responsible for the quality of their products. However, as global demand kept increasing, manufacturing practices and techniques evolved accordingly. This Industrial Revolution motivated a more general view of quality, where the concept of interested parties extends beyond a focus solely on the customer. In fact, many other parties of interest had to be considered, within

and outside the company, as they are also affected by, or affect, the quality of the product [8]. These parties include the employees, the suppliers, the investors, etc. and many others.

As a result, the concept of quality control evolved from being exclusively the responsibility of the inspectors, to a company-wide agenda where all the company shared responsibility on quality-related issues. This concept has been defined as the Total Quality Control (TQC) and was introduced by A. V. Feigenbaum in his book [9] "*Total Quality Control*".

In view of this evolving concept of quality, and with the considerable development in technologies, the quality control process has become a continuous process, starting from the conception / creation of the product, and engaging along the whole manufacturing process. Therefore, it is necessary to distinguish between two main categories of quality control: process control that aims to ensure the proper functioning and stability of all the processes and systems involved in the manufacturing task, and product control that is related to the quality of the product itself. This work only focuses on the product quality control task.

There are two main types of product control covering the majority of quality problems industries may encounter:

- **Technical control:**

  This type of control involves all the technical aspects of the product, i.e. all the details that ensure the good functioning of the product. It is related to technical defects that usually occur during the manufacturing process. Most of these defects are associated with the dimensions of the product, and can be sometimes located on its surface.

- **Appearance control:**

  This kind of control is related to appearance defects, such as scratches, marks, cracks, etc., that do not have a direct impact on the proper functioning of the product, but rather are associated with its aesthetics. These defects are usually referred to as accidental defects. They may occur during manufacturing or during the product handling. Customers are strict about this type of defect, as it directly affects the brand. These defects are located on the surface of the product, the one visible for the customer.

2.2.1 Quality control for wheels

With the continuous demand to produce high quality products, wheel industries have been making advancement in their quality control techniques. Finding the defects and their potential causes was the first step, which was followed by updating and improving the manufacturing processes to avoid such defects. As a

consequence, the customer returns for technical defects on the wheels, which are process related, declined significantly to a new lowest level. However, on the other hand, the customer returns for appearance defects were not significantly reduced. In fact, the major factor that created this difference in customer returns between the two types of defects is their inspection process.

Technical defects in manufactured wheels are measurable defects, which means that the effect of the defect can be measured and compared to a reference value for decision. They can be either dimensional defects that portray a geometrical design error, or mechanical defects related to the mechanical features of the wheel. In both cases, theses defects are mostly caused by a machine malfunctioning or a faulty process. Hence, defects of this type are repetitive and consistent, as they affect almost all of the production.

To control dimensional defects, many methods have been employed to monitor the dimensions of the different parts of the wheel, such as the height and thickness of the rim or the radius of the valve hole, for example. Some of these methods are conducted offline on a sample of the production using suitable calipers or micrometers. Many other industries use online automated dimension measurement techniques to test the totality of the production. These techniques are sometimes based on image and signal processing methods, or mechanical methods, and they include many technologies as electronic and pneumatic gauges, optical and imaging sensory, etc. [10].

On the other hand, the quality control process to check for mechanical defects in wheels is a well-defined process that has been regulated and standardized. All wheel manufacturers follow the same standard methods to test their products for such defects. The ISO standard [11] entitled "*Road vehicles – Passenger car wheels for road use – Test methods*" describes two test methods to evaluate fatigue strength characteristics of wheels used on passenger cars. These inspection techniques fall into the category of destructive testing (DT), even called mechanical testing, in which case a product is forced to reach its failing point by application of various load factors [12, 13]. These techniques are essentially applied offline on a small sample of the production, and when a defect is detected, the production stops in order to diagnose and find the faulty process or machine.

As for the appearance defects, the automatic inspection is more complicated. Appearance defects mostly occur because of a random error during product handling or due to a process related fault. These defects usually affect some products in a random manner, without consistency, which means that it is mandatory to inspect every single product for defects. Therefore, the testing process has to be online and fast thus limiting the potential testing techniques that can be used to detect such defects. In fact, in such case, only non-destructive testing can be used, where the future usefulness of the inspected wheel is not affected and its component materials

remain intact. Another downside to the detection of such defects is that no prior information, such as the size or shape, of the potential defect is available for the inspection process, as there is no agreed standard for appearance defect types in the wheel industry. Till today, this inspection relies largely on manual human control.

2.2.2 Wheel surface inspection

One of the basic and most recognized applications of quality control in any industry is the inspection of the finished product. It is performed at the end of the production line, when all the manufacturing steps are completed. Its importance comes from the fact that it is the last step of inspection, and the last resort to stop a defective product before it reaches the customer. For wheel manufacturing, the last step of production is the painting, or more precisely the coating process. Several coats of paints are spread over the surface of the wheel to give it its final aesthetic appearance. This appearance (color, gloss, texture, etc.) of the coated surface of a wheel greatly affects the perception on its quality for customers. Hence, an inspection procedure for this surface is required to ensure the quality of the wheel before delivery.

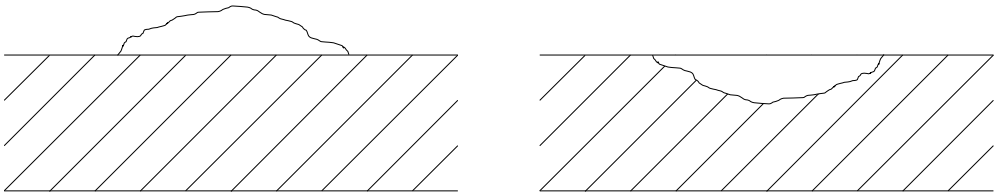
2.2.2.1 Surface defects

Given a set of product specifications, any deviation from what is standard, or normal, to the product is considered an anomaly. Furthermore, if the anomaly surpasses certain acceptance limits, which are usually defined by the customer, it is then referred to as a defect. One of the most important factors to judge the quality of products produced by an industry is the degree of variability between the manufactured products of the same type. During the day, a wheel industry may produce over 20 000 wheels of different types and designs. The heterogeneity between different products of the same type can be considered as nonconformity if it surpasses a certain level.
To more precisely feature the types of defects that may occur, it is necessary to categorize appearance defects into two different types:

- **Local appearance defects:**

  Local defects are defined as a local heterogeneity or distortion from the reference texture of a surface. They represent a sudden variation on a limited area of the inspected surface. Such defects stand out on the surface of the wheel and are the most frequent type of defects detected during the final inspection.

  These defects can take different shapes and sizes whether they are scratches, marks, geometrical deformation, etc. But they can all be classified under two general types of defects. Figure 2.4 represents a depiction of these two classes on a cross-section illustration of a wheel surface. The first class of defects, shown in figure 2.4(a), include all the defects that create a spot or a bump over the surface of the wheel, while the second class of defects, shown in figure 2.4(b), include the ones that break the surface of the wheel and penetrate



(a) (b)

Figure 2.4: Illustration of the two general cases of local appearance defects

it with a small depth. This classification is very important for choosing the most relevant non-destructive inspection method to be used.

Table 2.1 highlights the most frequent types of local appearance defects that can be found on the surface of the wheel. For each type of defects, table 2.1 provides common information defined according to the customer criteria or acquired during the inspection process. These information include the usual cause of the defect, its acceptance limits in terms of average size in millimeters, its severity using as reference the Failure Mode and Effects Analysis manual (FMEA) [14], its frequency of appearance in terms of products per million (ppm), and some real case examples of such defect.

- **Global appearance defects:**

  Global defects affect the entire surface of the wheel in a uniform manner. In this case, it is difficult to judge the wheel by itself as the defect is only visible if compared to previous manufactured wheels. Detecting this heterogeneity between different products of the same type is a challenging task. In fact, no product is perfect, and two products of the same type are never exactly the same, hence judging the nonconformity of a product due to a global defect is not always straightforward.

  Almost all of these defects are caused by the final manufacturing process, the painting process, and more precisely the topcoat layer. This is the last layer of coating applied over the surface of the wheel and it maintains the aesthetic appearance characteristics of the wheel surface. Any increase or decrease in the amount of paint used for this layer will lead to an excess or shortage of color on the wheel surface, thus generating a global defect. Therefore, it is necessary for the inspection method to be able to examine the color of the wheel and its intensity.

Additionally, the surface of the wheel is not only linked to an aesthetic aspect. Of course, during the inspection of finished products that will be delivered to the customers, the main focus is set on its aesthetics. The surface of the wheel is the one visible to the customer, thus on which his judgment of conformity will be based.

Table 2.1: List of the main potential defects with their characteristics

Type	Cause	Acceptance Limits	Severity (FMEA)	Frequency (ppm)	Examples
Smear	Industrial environment	Area $\leq 60\ mm^2$	2	$\sim 3\,000$	
Imprint	Stamping process	Area $\leq 30\ mm^2$	4	$\sim 2\,000$	
Painting Drops	Painting process	Area $\leq 20\ mm^2$	7	$\sim 5\,000$	
Black Dots	Cataphoresis process	Area $\leq 30\ mm^2$	3	$\sim 2\,000$	
Scratch	Product handling	Width $\leq 2\ mm$ Length $\leq 15\ mm$	7	$\sim 1\,500$	

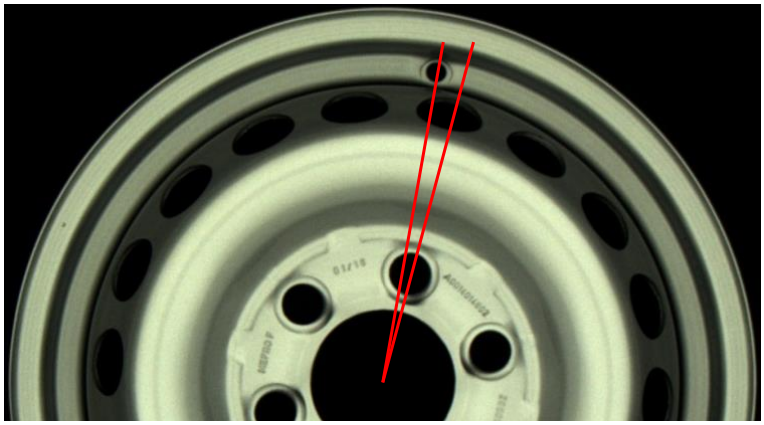


Figure 2.5: An example of a dimensional defect on the surface of the wheel

However, the surface of the wheel also contains multiple key elements that are essential for the proper functioning of the wheel. Those key elements are designed with respect to precise geometrical forms and measurements that have to be checked during the inspection process. Any error in these measurements is to be considered as a critical defect that leads to the immediate rejection of the wheel. Figure 2.5 shows an example of a dimensional defect visible on the surface of the wheel, where the valve hole is not in its correct position, since it must be aligned with one of the countersinks. This is due to an error during the mounting phase, when the rim and the disc are joined together.

This type of defect is classified under technical defects, or, more precisely, as measurement defects, and ideally must be inspected and rejected during the manufacturing process. Despite that, a non-negligible number of defective products can reach the final stage of inspection. Therefore, the final quality control has to be capable of detecting such defects.

As a result, we will define the term **surface defects** that serves as a more general class of defects that includes all the appearance defects present on the surface of the wheel, either local or global defects, and some technical defects that are related to the measurements of the key elements mentioned above.

2.2.2.2 Current setup: Manual visual inspection

At the end of the wheel's manufacturing cycle, the quality of the wheel surface is very often inspected and evaluated by human inspectors. Their main role is to inspect each and every product manufactured and to decide whether the final product is acceptable or not for delivery to the customer. A single operator might have to inspect thousands of manufactured products during the day. During the product inspection process, inspectors are should have a list of potential defects, which includes specified details about geometrical features, severity measure, and the acceptance limits of a number of types of defects. It usually also contains

illustrations and photos to provide assistance to the understanding of each defect. However, these acceptance limits are not usually the same for all produced wheels. A wheel industry will have many different types and designs of wheels to adapt to the demand of different customers. And every customer may have his own specifications and acceptance limits that have to be respected. This makes the inspection phase a complicated process for inspectors.

The main advantage of human evaluation is its great flexibility with regard to the various types and shapes of appearance defects that can be present on the surface of the wheel. However, human inspection faces major disadvantages. First the variability of decision for an inspector over time (factors related to fatigue and motivation) and the variability of decision between different operators for the evaluation of the same wheel. Another factor that is used to study the efficiency of human inspection is the repeatability of decision of the same inspector when evaluating the same product various times [15, 16]. These disadvantages result in an uncertainty and a lack of precision for the detection of certain appearance defects. Indeed, evaluating the aesthetics of a product is very subjective, and can differentiate from an operator to another. As for the inspection of dimensional defects, this task cannot be done by human inspection. Dimensional defects are usually very small (few millimeters), thus unnoticeable and undetectable by the human eye within few seconds.

To overcome these difficulties, automated inspection has proven to be the best alternative for industries to rely on [17]. The share of variability and subjectivity in decision-making is eliminated with an automated surface inspection. When testing the same product multiple times with an automated inspection system, the result will be the same every time. Furthermore, the higher precision and resolution that an automated system can reach will enable it to perform better for detecting dimensional defects.

2.3 Non-Destructive Testing for surface inspection

Non-destructive testing (NDT), also referred to as non-destructive evaluation (NDE), is defined as the class of methods used to inspect, examine, and evaluate the condition of a part, a material, or a system for non-conformity, or differences in characteristics without destroying its usability and future usefulness. The goal of NDT techniques in therefore to detect any defects that may affect the availability, safety of use, or more generally, the conformity of a product to fulfill its intended use [18, 19]. In addition, NDT techniques can be used for dimensional inspection, as to measure the thickness of a part or a coating, or even to determine the physical properties of a material such as, for example, its electrical conductivity. These techniques play an important role in the quality control process, not only the quality of the final product, but also the quality of the assembly parts as well as

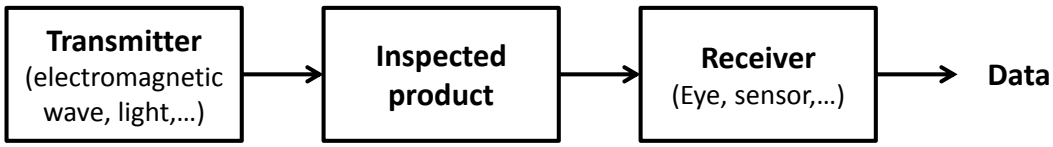


Figure 2.6: General concept of NDT techniques

the initial raw materials.

Since the beginning of the 20th century, the development and application of NDT techniques in different industrial sectors have been growing rapidly. It has now become an essential part of every industry at all stages of the production process. New techniques are being regularly introduced to respond to various industrial problems in various sectors, including nuclear, aerospace, civil engineering and automotive industries. The speed and reliability of the employed techniques have been essential for reducing maintenance costs and optimizing the lifetime of installations. The innovation in NDT was made possible by the development of electronics, sensors and especially information technology and computer science tools. These tools have improved the application of NDT by introducing the possibility to model the physical phenomena on which the different inspection methods are based. This modeling allowed the design of sensors optimized to answer specific industrial problems. In addition, signal and image processing techniques, and data mining, played a very important role in the development of the NDT.

Various NDT techniques have been developed, each one having its own advantages and limitations making it more or less appropriate for a given application. In general, they all consist in exciting the inspected product by a suitable physical signal and collecting the response data either physically or through a sensor. Then, processing the received data will help characterize the condition of the product and more particularly the presence of defects. Figure 2.6 illustrates this general concept.

In the literature, NDT methods are usually classified into two families according to whether they favor the detection of surface defects or internal defects [20, 21]. This is the most important factor to respect when choosing the appropriate method to perform the intended inspection task. However, this choice heavily depends on other important criteria as well. Some of them are physical constraints such as the nature of the inspected parts (material, shape, etc.) and the environment in which the inspection is to take place (thermal, chemical, pressure, radioactivity, etc.). And other related to the cost of the proposed method (material, labor, time consuming, etc.).

For the specific application of wheel surface inspection studied in this work, the

current setup employed by wheel industries is the visual inspection, or as referred to in the NDT community, the visual testing (VT). This choice is rational as the VT is simple and practical for surface inspection applications. However, it is important to review other NDT techniques to investigate their ability to perform better in the case study presented in this work. It is normal to focus only on NDT methods that can be automated, and that are used for surface defect detection. We will present the most widely used methods in this category. But first, in order to choose the most suitable among these methods to perform the task, we will start by listing the relevant choice criteria for our case, according to which each method will be judged. Some of these criteria are imposed by the specific characteristics of wheel surface inspection, while others are defined by the industry.

- **Inspection tasks:** What is the inspection process searching for ? In our case, the detection of surface defects is the main task. A detailed description of the variety of surface defects has been presented in section 2.2.2.1, where it has been shown that the inspection system must be able not only to detect various local appearance defects, but also deal with global defects, and perform measurements of certain geometrical forms.

- **Ease of automation:** As discussed earlier, the automation of the inspection procedure is essential to improve the reliability of the inspection. Therefore, an important requirement for the inspection system is the ability to automatize the entire procedure (all the steps of the inspection). For some NDT techniques, the automation can be very costly and may require large machinery and installations, which is not preferable for the industries, as they usually tend to reduce the costs and may have limited space for installation.

- **Product handling:** Products that reach this step of inspection are finished products that will be delivered to customers, hence must be handled accordingly during the inspection. Preferably, a contactless NDT technique is to be used to guarantee that the product will not be altered in any way. Additionally, some NDT techniques involve a pretreatment step before the inspection. Most of these treatments of the surface of the wheel may weaken its integrity and diminish its aesthetic appearance.

- **Record keeping:** The main purpose of the inspection procedure is to minimize the customer returns for defective products. When a customer returns with a complaint about a defective product that he received, the industry has to be able to properly respond to his complaint and to verify the cause of this mistake. As a result, it is necessary to keep a record for all the products that have been inspected and delivered to customers. This computerized record must contain a maximum quantity of information to describe the product (wheel type, wheel color, conformity, etc.) and ideally enable a second

evaluation.

- **Industrial requirements:** From an industrial point of view, many other aspects of the inspection system have to be taken in consideration, mainly the cost of inspection and the environmental impact. First, the cost of inspection does not only refer to the first installation of the machinery, but also include the continuous costs of maintaining the inspection process. Some NDT techniques require a continuous supply of resources (chemical materials, water supply, etc.) in order to perform the inspection. On the other hand, industries are required to follow certain environmental regulations throughout the production. Many NDT techniques come with warnings regarding to their environmental impact and workplace hygiene (precautions concerning the risks of fire, explosion, pollution of water. etc.). These precautions have to be fully respected at all time.

2.3.1 Penetrant Testing

Penetrant testing (PT) is one of the oldest and most widely used NDT methods for the detection of surface defects in solid materials. Its popularity can be attributed to the fact that it can be applied to virtually any material, magnetic or non-magnetic, provided that its surface is not extremely rough or porous. The variety of materials that are commonly inspected using this method extends to steel [22], ceramic [23], aluminum [24], carbon fiber/epoxy composite [25], polymer [26], and even rocks [27]. It is becoming more and more important in many industrial sections, including the automotive industry [28], and more recently the aerospace industry [29–31].

There exist several PT techniques applied for different applications. However, the common fundamental steps of all these procedures are illustrated in figure 2.7.

Choosing the suitable technique for PT is a crucial step. In fact, the PT procedure can vary from an application to another depending on various factors including the size and the material of the inspected product, its surface roughness and condition, the type and size of discontinuities that are expected, and the environmental conditions under which the inspection is performed. In a general manner, the PT procedure consists of successive steps summarized as follows [32]:

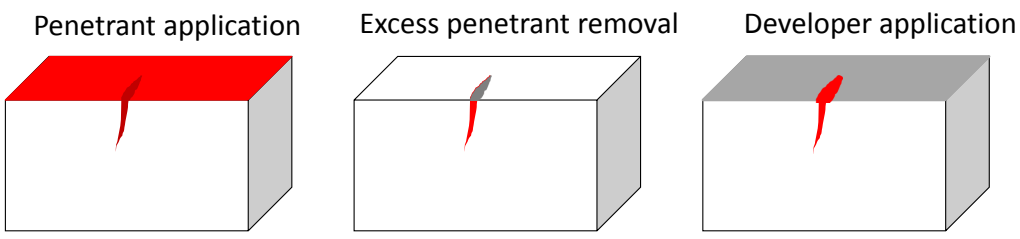


Figure 2.7: Principle of the penetrant testing technique

- **Pre-cleaning:** It is vital that the product to be inspected is perfectly clean. Its surface must be dry and free from any dust, rust, grease, oil, painting, water or any other pollution. Otherwise, their presence may restrict the entry of penetrant and can very easily lead to serious discontinuities being falsely detected. Pre-cleaning methods that can be used to clean the surface vary depending on the material and what is required to remove. Some of these methods are mechanical, such as brushing or abrasive blasting, but they are not considered as an adequate solution, as they might cause small surface cracks to be closed. An alternative more suitable solution is to apply chemical methods that are often done using cleaning solvents that allow a quick drying.

- **Penetrant application:** The penetrant can be applied to the surface of the inspected product in virtually any effective manner. That includes spraying, brushing, electrostatic application, dipping the part into the penetrant, immersion, or just pouring it on the surface. It is important to ensure that the area of interest is completely covered with the penetrant solution. The main concern is to ensure that the penetrant solution does not dry during the penetration time, referred to as dwell time, which is the period of time from when the penetrant is applied to the surface until it is cleaned out. Typical dwell times are between five and sixty minutes, and are usually defined by the penetrant supplier, and may vary according to the specification and type of discontinuity being sought.

- **Excess penetrant removal:** This is considered to be the most delicate step of the PT procedure, because the excess penetrant must be carefully removed from the inspected surface while removing as little penetrant as possible from discontinuities. If this step goes wrong, discontinuities will not be properly detectable afterwards, and the process must start over. In general, there are three different techniques for excess surface penetrant removal: direct rinsing with water, cleaning with a solvent, or post-treatment with an emulsifier and then rinsed with water. The choice of the appropriate removal method depends on the type of penetrant that has been used.

- **Developer application:** A uniform and thin layer of the developer is then applied to the entire inspected surface. It acts as a blotting agent, drawing penetrant out of the discontinuities to the surface where it will be spread over a larger area, thus more visible. The developer must be given sufficient time to draw the entrapped penetrant from the discontinuity out to the inspected surface. In some case, this time may reach as long as sixty minutes. Depending on the developer type, it may be applied either by dusting, dipping or spraying. There are four main types of developer: dry developer, water-based developer, solvent-based developer, and the rarely used film type developer.

- **Inspection:** After the development time, the surface is ready to be inspected. In an automatic setup, this is done using a camera, with an appropriate lighting that depends on the penetrant dye type that has been used. Penetrants based on a fluorescent dye are mostly used as they are considered to be more sensitive than other types of dye. In that case, the inspection has to be carried out in a dark booth, with ultraviolet lighting, where the penetrant will have a bright glowing yellow/green color.

- **Post-cleaning:** After the inspection, the final step in this process is to thoroughly remove all traces of any remaining penetrant and developer from the surface of the inspected product prior to it being delivered to the customer.

It is clear that PT is substantially more sensitive than a purely visual inspection, mainly because of the spreading and glowing characteristics of the dye. However, it suffers some major drawbacks in general, and is especially not applicable in the application studied in this work.
First, the limitations of PT are mainly related to the inspected product itself. The shape of the inspected surface has to be simple to be able to perform a PT [33,34]. This is not the case for wheel inspection as the surface of the wheel has a complex geometrical design, with multiple gaps and holes in which the penetrant can get trapped. In addition, PT is not preferable for the inspection of coated surfaces for fear of damaging the paint by the use of chemical products.
Secondly, several limitations of PT are related to its process. It is relatively slow, with many environmental risks, and can hardly be crried out in a fully automatic manner. In the last decade, constant efforts have been dedicated to automate the whole procedure [31,35]. An important advantage of this automation is the decline in environmental issues that comes with the PT. However, it is not possible to eliminate all the potential hazards due to the chemicals involved in the process [36]. As a consequence, many regulations and standards are imposed on industries that use PT. In addition, the PT procedure requires a continuous supply of materials and chemicals, such as the dye or the developer, which implies continuous costs for the industry to maintain the inspection process.
Finally, a crucial downside of PT that makes it inapplicable for wheel inspection is the fact that it can only locate discontinuities that are open to the surface (cracks). All the defects that do not break the surface of the wheel cannot be detected, such as bumps, stamps, or imprints.

2.3.2 Magnetic Particle Testing

Magnetic particle testing (MT) is a relatively simple NDT technique that is primarily used for detecting discontinuities located at or near the surface of magnetic materials. MT is governed by the laws of magnetism, thus restricted to the inspection of materials that can support magnetic flow, or ferromagnetic materials, such as iron, nickel, cobalt, etc. It is widely used in many industries

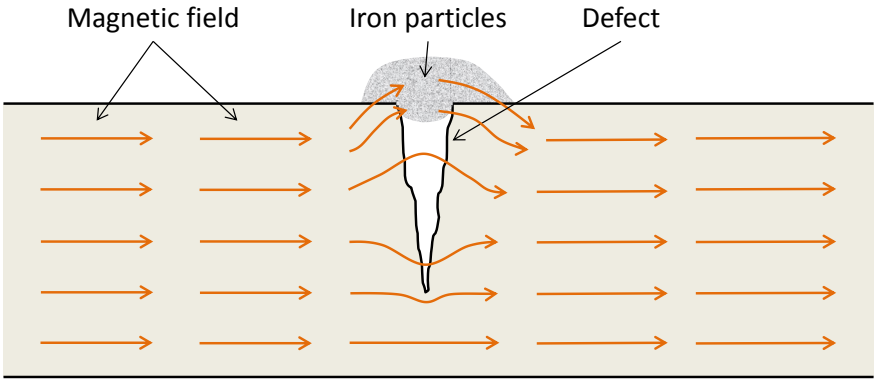


Figure 2.8: Principle of the magnetic particle testing technique

including the automotive industry [37,38].

The basic principle of MT is the disturbance of magnetic fields by surface defects in magnetized ferromagnetic materials, and is illustrated in figure 2.8. When a material is magnetized, the lines of force in its internal field will tend to distribute themselves evenly through the material, provided that the material is homogeneous. The presence of a discontinuity presents an interruption to the field, thus an increase in reluctance. The electric flow prefers the path of least reluctance and will therefore redistribute themselves in the material by bending around the discontinuity. The defect will cause some of the lines of magnetic force to depart from the surface and thus to create a magnetic leakage field. In order to produce a leakage field, a discontinuity must interrupt the field usually considered to be within 45° to the perpendicular. If, for example, a crack is oriented perpendicular to the flux lines the effect of disturbance of the magnetic field is at a maximum compared with other orientations. It follows that in order to detect a discontinuity with any orientation, the part must be magnetized in at least two orthogonal directions [39].

The MT procedure consists of three main steps. The first step is to magnetize the product under inspection. Many magnetization techniques can be used, including magnetic flow, current flow, coil, and central conductor technique. Several factors will dictate which magnetization technique is the most relevant for a particular test. The primary factor is the most likely direction of possible discontinuities. It is important to understand that the orientation of defects relative to the magnetic line of field determines if the defect can or cannot be detected. The second step is to apply iron particles, either in a dry or wet suspended form, on the magnetized surface. These particles are attracted to the area of flux leakage, creating a visible indication of the flaw. The magnetic leakage field will hold the particles in a ridge on top of the crack. By taking advantage of this effect, an accumulation of magnetic particle forms, which is much wider than the crack itself, thus turning an otherwise invisible crack into a visible one. In either case, dry

or wet particles, the choice has to be made between color contrast or fluorescent particles techniques. Most of the MT techniques prefer the use of fluorescent particles, due to the increased sensitivity provided by the higher contrast ratio. Finally, the inspection task is carried out, usually with a camera, to decide on the conformity of the product, after which a demagnetization is usually needed [19].

The MT technique has been proved to be one of the simplest and most sensitive methods in NDT. Furthermore, and unlike some other NDT techniques, MT has shown a very high potential to be automated and integrated in the production line [37, 38, 40]. With the aid of digital image processing, the procedure becomes fully automatic [41]. However, it has some major limitations that restrict its use in many industries. Probably the single largest limitation to MT is the fact that it is applicable only to ferromagnetic materials, and cannot be used to inspect nonferrous materials, such as aluminum. As a result, in the wheel industry for example, only steel wheels manufacturers are able to use the MT technique. In addition, some MT techniques, more precisely the direct magnetization techniques, could face some difficulties with painted products, as the paint coating would reduce the electrical contact.
Another limitation for the MT technique is its weak flexibility to the inspected product geometry. When inspecting products with complex shapes, the flux flow through the product is not uniform, and difficult to predict. This becomes even harder in case where the product design contains abrupt changes such as holes or keyways. This will probably cause flux leakage and thus produce non-relevant indications that make the interpretation difficult.
A final downside that should be noted is the health, safety and environmental impact of the MT method. More details can be found in [42, 43].

2.3.3 X-Ray Radiographic Testing

It is believed that the radiographic testing (RT) was the basis and the start point of the NDT terminology, and it all started with the discovery of X-ray in 1895 [44]. It involves exposing the product to be inspected to penetrating radiation (X-ray or gamma radiation) so that the radiation passes through the product into a detector medium placed on the opposite side of that product [45]. The choice of the proper radiation for the intended application is mainly related to the ability of the radiation to pass through the material. Hence the choice depends heavily on the material characteristics and thickness. For wheel inspection, the usual choice in the industry is the X-ray radiation. In this case, the classical detector is an X-ray film. However in recent years the transition has been made to electronic sensors. The radiation passing through the inspected object and into the detector defines the test results. Darker areas in the received image means that more radiations have passed through the product, and lighter areas means that less radiations have penetrated. If there is a void or defect in the part, more radiation passes through, causing a darker image on the detector. Another aspect of the RT is that it can

be used for volumetric inspection of industrial products. In fact, an object which has a high density, i.e. a thicker object, will absorb more radiation, thus produces lighter areas on the received image, and, on the other hand, if an object has a low density, it will absorb less radiation producing darker areas on the received image.

NDT with X-rays, known as X-ray testing, can be considered as a visual testing technique. The distinction, however, is made in the NDT community due to the fact that X-ray testing uses radiation as a transmitter. It is widely used in many applications such as inspection of automotive parts, including wheels [46–48], inspection of blades in aircraft and turbines [49] quality control of welds [50, 51], cargo and containers inspection [52,53], baggage screening [54], and quality control of electronic circuits [55] among others [56].

In order to achieve efficient and effective X-ray testing, automated and semi-automated systems are being developed to execute this task. Compared to manual X-ray testing, automated systems offer the advantages of objectivity and reproducibility for every test. In recent years, the most widely used X-ray imaging systems employed for NDT in industries are digital radiography (DR) and 3-dimensional computed tomography (CT). Instead of the traditional radiographic film, these techniques use electronic sensors that allow the production of digital X-ray images by converting the X-ray energy into an electrical signal [57]. The most common sensors are amorphous silicon or selenium flat-panel detectors which are based on a charged-coupled device (CCD) camera. These panels are simple to use, and relatively fast to support a real-time inspection. The X-ray images are captured by the flat panels and then recorded digitally, which will allow for them to be processed at any time [58]. In fact, the 3D CT technique is simply an extension of the DR. It provides a cross-section image of the inspected product so that each object is clearly separated from any others. This section-by-section reconstruction makes it possible to achieve a complete exploration of the inspected product, but is, however, very time consuming, and thus cannot be applied in some industries.

X-ray testing can be difficult to use on components with irregular surface areas since those irregularities can influence the thickness of the materials that are being evaluated. Despite that, many solutions using X-ray testing for the inspection of wheels can be found in the market. Although they can detect surface defects and cracks in certain orientations, they are mainly used to find subsurface volumetric defects. Continuous development is done to offer fully automatic X-ray inspection systems for wheel inspection based on image processing methods. Most of these systems use multiple view inspection techniques to overcome the wheel irregularities, and to detect defects with different orientations [47, 48]. In such systems, several images of the wheel with different views are acquired in predefined precise positions. The number of views required for the inspection depends of the design and size of the wheel, and the test specifications. Detailed description of such systems can be found in [59–61].

Nevertheless, X-ray testing is not convenient for the inspection of the wheel aesthetics. This is mainly due to the specific characteristics of the aesthetic defects that are present on the wheel surface. Such defects, like painting drops and smear marks, have a negligible or nonexistent impact on the X-ray radiation passing through, compared to the impact of the wheel material. As detailed earlier, an important factor for the X-ray detection is the density of the inspected object. For aesthetic defects, their density compared to the density of the wheel metal is insignificant. Thus the X-ray radiation will be greatly affected by a slight metal irregularity on the wheel surface, but not sufficiently affected by the presence of an aesthetic defect, such as a complete lack of coating. As a result, the contrast introduced by the presence of such defects in the X-ray image of the wheel will not be relevant.

2.4 Visual Testing

Visual testing (VT) was the first NDT technique used in industry, and is by far the most commonly used test method. In fact, because most NDT techniques require a final visual inspection phase, VT is inherent in most of theses techniques, as the ones described in the previous section, but was the last method to be formally acknowledged as an independent NDT method [19]. As its name indicates, VT involves the visual observation to examine and evaluate the surface condition of a product, from discontinuities, such as corrosion or cracks, to the coating condition of painted surfaces. It can be applied successfully to virtually anything from man-made products and structures to organic matter. This last feature enable the VT to be used in almost any industry where a final inspection is required.

In its basic definition, VT is simply the inspection of a surface with eyesight. However, in the case of NDT, it has a much broader meaning. It became a regulated inspection technique, which involves a training program for the operators, and requires several additional components to properly perform the inspection task. Probably the most important one is the lighting system, as in all cases of surface inspection, the lighting conditions are essential for the reliability of the visual control. In general, one can differentiate between two main types of VT inspections [62].
The basic type is the direct visual testing, where the optical path between the inspected surface and the operator's eye is not broken, hence it is only available when the inspected product is directly accessible by the operator. For such inspection scheme, it is essential to guarantee optimal inspection conditions, such as environmental and lighting conditions, for the inspector to be able to detect the smallest anomalies comfortably. In addition, when inspecting a variety of surfaces, operators can benefit from some simple visual aids, such as magnifiers or portable

measuring devices, to enhance their views and enable a better inspection.
On the other hand, the second type of VT is the remote visual testing. It is characterized by an interrupted optical path between the operator's eye and the inspected product. This type of inspection is usually used to access ares on the product surface that the operator cannot reach, or when the inspected product is located in hazardous locations for the human inspectors, such as high pressured or high radiation rooms. Hence, it is applied whenever the eye cannot obtain a direct, unobstructed view of the inspected surface without the use of another optical instrument or device. Such devices vary from telescopes, borescopes, microscopes, and most recently digital cameras. In general, a monitor is used to display the acquired image for the operator to perform the inspection task. It is also possible to digitally record the image for future inspection.

Eventually, it can be seen that the VT is rather a simple technique that enjoys many advantages. It usually does not require expensive equipment, and can be used to examine virtually any component anywhere on the surface. However, it faces some major drawbacks, mostly related to the human factor as discussed in section 2.2.2.2. In many cases, especially in harsh environments, the inspection tends to be tedious or difficult, even for the best trained operators. In addition, human operators are slow compared to modern high-speed production rates. To overcome these drawbacks, automated visual inspection has proven to be the best alternative for industries to rely on [17,63], in which case the inspection system can acquire image data, process and analyze this data and make an evaluation automatically.

2.4.1 Automated visual inspection

It has always been a dream for manufacturing industries to build a machine that is able to "see" and describe what it saw. Such a machine does not necessarily attempt to emulate the human vision system, but rather comply with various industrial requirements, such as the high speed, low cost and high reliability. This concept has been known in the industrial world as machine vision [64].

Machine vision traditionally refers to the use of computer vision in an industrial or practical application or process where it is necessary to execute a certain function or outcome based on the image analysis done by the vision system. More simply, the automatic extraction of information from digital images for process control or inspection of manufactured products. Over the last years, machine vision has played an essential role in controlled industrial applications. It provides innovative solutions in the direction of industrial automation. A variety of industrial applications have benefited from machine vision systems to improve their manufacturing processes, including among others, electronics component manufacturing [65, 66], fabric inspection [67], food sorting [68–70], assembly systems [71], integrated circuits inspection [72], etc. . . .

In general, machine vision applications can be classified into four main categories [73]: process control, parts identification, robotic guidance, and automated visual inspection (AVI). Visual inspection is the oldest category of machine vision applications, and it presently represents the widest variety of installed industrial systems. AVI systems are typically more repeatable than human inspectors, and they can often inspect products at a far higher rate. They are typically used to evaluate the aspect of finished products, while providing numerical data about the detected defects features. In addition, they are used to perform dimensional measurements to verify the integrity of the inspected product. In the literature, AVI systems can be further classified according to the type of inspection performed, such as inspection of dimensional quality, surface quality, structural quality, or operational quality. Most of the time, AVI systems are designed to perform two or more types of quality inspection at the same time. The survey in [74] provides several examples of industrial vision systems under the previous classification. Other reviews can be found in the literature that specialize in a single application field, such as [75] for automatic PCB inspection, or [76] for automatic fruit harvesting.

2.5 Conclusion

The quality of manufactured products has always been a major concern for many industries. With recent technological advancements, new quality control techniques are being developed to tackle this concern. Accordingly, wheel manufacturing industries have managed to reduce the customer returns for technical defects, but were not able to achieve the same for appearance defects. This is due to the fact that the quality inspection of appearance defects on the wheel surface is still a manual process conducted by operators. This manual process faces considerable limitations which weakens it efficiency and make it unreliable for inspection. Hence the necessity to find new inspection methods to meet the inspection requirements of the industry.

This chapter reviews the most widely used NDT techniques for surface inspection. The advantages and disadvantages of each technique are then evaluated based on the pre-defined requirements of the inspection system. It is concluded that the automated visual inspection technique is the most appropriate for our case, as it meets all the requirements. In the following chapters, it is proposed to design a complete AVI system, based on the practical requirements of the present industrial context, to perform the wheel surface inspection task.

Chapter 3

Automated visual inspection system

Whether it was the inspection of fruits and vegetables [77], or the inspection of steel surfaces [78], the overall configuration of an AVI system is always the same. A lighting system generates light in a specific manner to illuminate the inspected object with the aim to improve the quality of the acquired images. The illuminated scene is projected onto the sensor of a digital camera, which transforms the light into electrical signals, to create a digital image. A lens is typically added to the camera to define the field of view and the area of the scene to capture. The resulting digital image is then sent to a processing platform that controls the whole AVI system. With the help of image processing techniques, the processing platform analyzes and processes the acquired images to give the final inspection information about the conformity of the inspected object.

So as it can be seen, an AVI system is composed of two main subsystems.
First, the image acquisition subsystem which is hardware based. It has the role of transforming the optical scene into an array of numerical data received by the processing platform. In total, the four key elements that can always be found in this subsystem are the camera, the lens, the lighting system, and the processing platform. In many cases, several of each component may be needed depending on the specific application.
Second, the image processing subsystem which is software based. It mainly consists of image processing methods that are developed to analyze the acquired data and give the final inspection result.

It is worth noting that a complete AVI system also incorporates multiple software solutions which have the role to control the different elements of the system, including the cameras and the lighting system.

In this chapter, we will only focus on the image acquisition subsystem. More precisely, this chapter will discuss in detail the challenging task of designing an AVI system for the real-time surface inspection of finished wheels.

To this purpose, we will adopt a work plan that is typical for the design of such systems. The first step is to define and well-understand the system and inspection requirements, in agreement with the constraints imposed by the industry. Then start with a conceptual design that suggests some first ideas on how to meet the pre-

defined requirements, to finally outline the basis of the whole system. Afterwards, it is possible to choose the suitable hardware for the AVI system, i.e. the cameras, the lenses, the lighting system, and the processing platform. Finally, the prototype of the complete system is built based on the conceptual design, during which possible adjustments are probably required to obtain satisfactory results. When the final installation setup is decided and installed, the software based inspection process can start.

3.1 Inspection requirements and conditions

When designing an AVI system, the first thing to do is to ask the question: What exactly does the system need to achieve, and under which conditions ?

Indeed, the general requirements are that any AVI system must be fast, cost effective and reliable. However, to be able to put together an efficient AVI system, a more in depth description of the inspection requirements is necessary. It is absolutely essential for the designed system to be tailored to the specific demands of the application. In this respect, the complete design process must be driven by the application requirements and conditions. Therefore, it is important to specify the problem as clearly and precisely as possible.

Real-time surface inspection of finished wheels faces a number of challenges. Some are imposed by the specific characteristics of wheel surface inspection, while others are imposed by the industry. In what follows, we will present the major inspection requirements and conditions that will have the utmost effect on the design of our inspection system. In addition, we will suggest some first possible solutions to some of these problems, which will help put forward the conceptual design of the system. Of course, the final system could be very different from that first envisaged.

1. **Wheel surface complexity:** The surface of the wheel is not a flat surface. It rather has a complex geometrical design, with irregularities in different directions. Those irregularities may sometimes hide defects when inspecting the wheel from a single direction. In addition, multiple elements are present on the wheel surface that have to be identified and inspected. A good example of such elements is the stamping on the wheel surface, which usually contains various information on the manufactured wheel. Hence, the imaging system should ensure the visibility of these elements regardless their location on the wheel surface.
   Therefore, a probable solution is to use multiple cameras to take images of the wheel from different viewpoints simultaneously. These viewpoints have to be chosen in a way to allow a total inspection of the complete wheel surface, despite its complexity. Furthermore, a bright lighting source could be required to properly illuminate the wheel, ensuring the good visibility of all the elements present on its surface.

2. **Wheel surface texture:** At the final stage of the manufacturing process, all

the wheels are painted through electro-coating that gives the wheel a black color. Afterwards, an important amount of the finished wheels are finally coated with a glossy paint layer, with different colors, to improve the style rendition of the wheel surface. This paint topcoat usually renders the wheel surface more reflective than regular metal surfaces. As a result, any direct lighting on the wheel surface will generate some light reflection artifacts in the wheel image that could result in false alarms during the detection procedure.
Therefore, it is important to design a lighting system that is suitable for glossy products, to avoid generating light reflections on the wheel surface. Ideally, for any type of application, a uniform illumination of the inspected product is always favored. In addition, due to the variability in topcoat colors, the lighting system should enable a correct recreation of real colors in the resulting wheel image.

3. **Variability in wheel designs:** During the day, a wheel industry may produce over 20 000 wheels with a large variety in types and designs. This variety could reach up to 100 different designs of manufactured wheels. Not only the design of the wheel surface that varies, but more importantly the size (height and width) of the wheels may largely differ. This parameter plays a major role in defining the field of view of the imaging system. In our case, the width of the wheels ranges from 13 inches diameter, to 17 inches diameter for the biggest wheel design.
Accordingly, the AVI system should be designed in a way that enables the inspection of the whole surface regardless of the wheel design. The installation of the system should especially be able to adapt to the wheel width and height. A possible solution could be to design an adjustable installation setup of the imaging system over the conveyor belt. This will enable to modify the distance and angle between the cameras and the inspected wheel according to the wheel design.

4. **Variability in defect types:** There is a large variety of defect types that are located on the wheel surface. This variety of defects is usually studied by the size of the defect, its shape, and its probable cause. Some of these defects can be identified as scratches, marks, geometrical deformation, etc.... However, several new defect types can appear on the wheel surface upon any modification in the production process. Hence, the AVI system should ensure the detection of surface defects regardless of their type. The only requirement that is defined is that in this work, the smallest defects that are wished to be detected have a size of about 2 mm in a single direction.
This last requirement is very important for the design of the imaging system. In fact, for a specific detail to be available for inspection, the resolution of the camera must be better than the size of the detail itself. Thus, the resolution will be defined partially in correspondence with the order of magnitude of the smallest defects that are intended to be detected on the wheel surface. On the other hand, from a software point of view, the defect detection method to develop should be

general enough to be able to adapt to the variability, not only in defects, but also in wheel surface textures.

5. **Productivity conditions:** In order to increase the productivity and maintain it above a certain limit imposed by the industry, the wheels do not stop under the AVI system, but are rather moving at a constant speed of about 1m/s. This is a major challenge to address when designing the imaging system. Indeed, a high quality inspection is not possible without a high quality image. Any type of blurring or distortion in the acquired image will definitely decrease its quality, thus the overall inspection efficiency. Additionally, when capturing the images of the different wheels, they should ideally be at the same position every time. Otherwise, the different images will not have similar profiles during their inspection, which makes the detection method much more complex.
   Consequently, the AVI system must be able to capture the image of the moving wheel at the precise moment when it passes through. Furthermore, it is essential to use cameras that are adapted to capture images of moving objects without any distortion in the resulting image.

6. **Installation conditions:** As the imaging system has to be installed over the conveyor belt, the space is restricted. Accordingly, in our case, the allowed distance between the camera and the inspected wheel ranges between 550 mm and 600 mm.
   Therefore it is important to keep in mind these conditions while designing the installation setup of the imaging system. On the other hand, the choice of some hardware will also be affected by these conditions. For example, the lenses for the cameras have to be suitable with the dimensions of the wheel, the inspection distance, and the camera characteristics.

3.2 Camera selection

The first practical decision to make on the way to design an AVI system is the choice of the suitable camera. And the best way to start is to ask the question: What should the system be able to "see". This question does not only concern the shape and size of the inspected object, but also the level of detail required to perform a proper inspection. Following the inspection requirements presented in section 3.1, it is clear that many parameters will have an impact on the final decision. Correctly identifying these parameters will help define the necessary specifications that the camera should have in order to deliver precisely what is required.

In general, cameras that are designed to perform quality inspection have high specifications compared to regular cameras. Still, with this high performance, such cameras are usually small in size to be able to fit in any type of installation. In most cases, they are additionally customized to endure harsh conditions that potentially exist in industrial environments. The term "industrial camera" is formally used to

refer to such type of cameras. In fact, the main benefit of industrial cameras is that they enable the user to receive the acquired images as uncompressed "raw" data that are not processed in any way. This type of image is generally preferred during the image processing stage, as it guarantees that no information is lost during the image acquisition. This will result in a more accurate inspection of the captured scene.

In recent years, a new type of cameras, referred to as smart cameras, has been growing in popularity. The smart camera is defined as a complete or nearly complete vision system, contained in the camera body itself. Not only does it capture images, but also uses computer vision technologies to process the image. Although this might seem like a huge advantage over standard industrial cameras, smart cameras are not designed for general purpose but rather for particular applications [79]. Therefore, they are usually not used in AVI systems, as they are limited in capabilities and do not offer the flexibility that a complete vision system could offer.

In what follows, we will discuss the camera selection process, exploring in detail various factors that play essential roles in the final decision.

3.2.1 Area scan or line scan

When designing a machine vision system, more precisely an AVI system, we are confronted with seemingly endless options. Perhaps the most critical decision is the choice of the scan technique. There are two subdivisions of scan techniques within the world of industrial cameras: line scan cameras and area scan cameras. While both perform a similar role, these technologies are miles apart in how an image is captured. The question of whether an area or line scan camera should be used is a question of the specific application in which they will be used.

An image is composed of a specific number of lines. Line scan cameras use one sensor comprised of just 1, 2 or 3 lines of pixels. As the object moves past the camera, or the camera moves over the inspected object, the image data is captured line by line, with the individual lines then reconstructed into an entire image during the processing stage.
Since line scan systems require parts in motion to build the image, they are often well-suited for products in continuous motion, or high-speed applications. They are thus best employed to inspect rolled or sheet objects, mainly fabric inspection. In [80], a review of automated fabric defect detection systems showed that most of the current setups for such systems are based on a line scan camera, in order to ensure a high-speed inspection. Many examples of such systems have been reported in the literature [81–84] Additionally, in [81], the authors described the advantages of using the line scan camera over an area scan camera for fabric inspection. They concluded that line scan cameras should be used in the analysis of moving fabric, while area scan cameras should only be used in the analysis of static fabric, to avoid

the blurring effect in the acquired images. Indeed, line scan cameras can obtain images from the fabric surface area at high speeds in the form of lines. However, they must be synchronized to the moving fabric by means of an encoder to obtain the true movement direction of the manufactured fabric. This is sometimes a complex task to assure an accurate image-line triggering for line scan cameras.
Other applications that use line scan cameras to inspect sheet objects include conveyor belt inspection [85–87] and steel surface inspection [88]. In all these previous applications, the camera is usually fixed, and the object is moving across it to capture the whole image. In other applications that adopt the line scan technique, the inspected object is the one stationary, and the camera is moving across it. Such applications include street inspection tasks, and the evaluation of the concrete in tunnels and bridges, where the camera is attached to a vehicle that moves over the inspected area. For example, in [89] and [90], a line scan camera was used to design a vehicular inspection system for detecting respectively road pavement defects, and concrete cracks in a tunnel. The common features that all these cases share is that the inspected surface cannot fit in a single image.

In contrast, area scan cameras, also known as matrix cameras, are equipped with a rectangular sensor featuring numerous lines of pixels that are exposed at the same time. The image data is thus recorded in one single step, and is also processed in the same way. They serve indeed a more general purpose, hence are used in the majority of machine vision systems. However, they are best suited towards applications where the object is stationary, even if only momentarily, or moving at a slow pace. They are usually the most suitable techniques when inspecting individual objects, where a single image is sufficient to capture the whole area of interest. Uninterrupted capture of continuous materials by an area scan camera can be complicated and time consuming, as it is achieved only by capturing overlapping images, and thus requires software to crop each individual image, eliminate distortion and assemble the images in the correct sequence.

In certain applications, line scan systems have specific advantages over area scan systems. For example, inspecting round or cylindrical parts may require multiple area scan cameras to cover the entire part surface. This can be done with only one line scan camera. Rotating the part in front of a single line scan camera captures the entire surface by unwrapping the image. Another advantage of line scan cameras is that they can fit more easily into tight spaces for instances when the camera must peek through rollers on a conveyor to view the bottom of a part. In addition, line scan cameras can build continuous images not limited to a specific vertical resolution. This allows for much higher resolutions than with area scan cameras, where the achievable resolution is limited and predefined. Indeed, line scan cameras resolution is only specified in the horizontal axis since the achievable resolution in the vertical direction will depend on the intended object inspection.

However, line scan cameras have many disadvantages compared to area scan

cameras, and which favored the decision to use this later one in many applications, and in this particular work. First and foremost, the inspected wheels arrive individually one by one, where the inspection of each wheel has to be done fast before the arrival of the next one. Hence the benefit of using a line scan camera for the inspection of continuous objects is not applicable. In such a case, it is preferable to use an area scan camera, where a single image is sufficient to capture the whole area of interest. This will help to save up the time required to build up the wheel image from multiple line scans, and the use of an external hardware to perform it. Secondly, area scan cameras can be more easily focused and adjusted to the specific scene compared to line scan cameras. They offer an easier setup and alignment, and are usually much cheaper for similar specifications. Finally, an inspection system based on a line scan camera is sometimes difficult to install depending on the inspection settings. The line rate of the camera must be synchronized to the speed of the moving object. If this is not the case, and the line rate is fixed while the object speed varies (due to varying conveyor speeds), the object image on the monitor or in image memory is elongated or compressed. Hence an additional hardware, usually an incremental encoder, is required to provide the adjustment of the line rate to match the current speed of the material under inspection.

3.2.2 Monochrome or color camera

On a digital camera sensor, there are millions of pixels. Each pixel has a tiny light cavity or "photosite" to record an image. When the camera's shutter button is pressed and the exposure begins, each of these photosites is uncovered to collect photons and store those as an electrical signal. Once the exposure finishes, the camera closes each of these photosites, and then tries to assess how many photons fell into each cavity by measuring the strength of the electrical signal. The signals are then converted into digital values, with a precision that is determined by the bit depth. However, each small cavity cannot distinguish how much of each color has fallen in, so the sensor can only record gray scale images. This procedure briefly describes the functioning of a monochrome camera. Indeed, a monochrome camera cannot reproduce a color image by itself, but it is possible by the use of an additional hardware. Characteristics regrading the color will appear on the black and white images of a monochrome camera through the use of colored illumination. However, this solution is only limited to a single color. This is also the case when using a particular color filter installed on the camera lens. Only the color of this particular color filter could be distinguished in the resulting black and white image.

To capture proper color images, a filter has to be placed over each cavity that permits only particular colors of light. Most of the current digital cameras can only capture one of three primary colors in each cavity, and so they discard roughly two third of the incoming light. As a result, the camera has to approximate the other two primary colors in order to have full color at every pixel. The most common type of color filter array is called a "Bayer filter". It consists of three colors of small

filter: red, green and blue. There are twice as many green filters in a Bayer array to accurately reflect the way the human eye sees color; it is more sensitive to green light. The array only passes the intensity of one of the three colors in each cavity to the sensor. Hence, each pixel represents the information of a single color. This is referred to as the RAW information. In order to create the final color image, the two remaining colors at each pixel are estimated (interpolated) using the eight surrounding pixels. This technique is called Bayer demosaicing or Bayer to RGB Conversion, and it is usually executed in the camera itself. It is the process of translating this Bayer array of primary colors into a final image which contains full-color information at each pixel.

Unlike with color, monochrome sensors also do not require demosaicing to create the final image; the values recorded at each photosite effectively just become the values at each pixel. As a result, monochrome sensors are able to achieve a slightly higher resolution. One should also consider whether the quality of a monochrome sensor outweighs the flexibility of a color sensor. One can always convert color into monochrome afterwards, for example. Furthermore, with color capture, any arbitrary color filter can be applied in post-production to customize the monochrome conversion, whereas with monochrome capture, the effects of a lens-mounted color filter are irreversible. Therefore, a monochrome camera should only be selected if color is not relevant, since the color filter on color cameras inherently makes them less sensitive.
For the specific application presented in this work, color images are essential. Manufactured wheels are always painted with a homogeneous color, but it is not limited to a single color. Indeed, most of the wheels will have either a black color topcoat, a white color topcoat, or a gray color topcoat. A monochrome image can be suitable for use with all these colors as they present a variation in the gray level. However, the wheel could be coated with other colors as well. For example, spare wheels are sometimes painted with a red color topcoat to separate them. Distinguishing between a spare wheel and a regular wheel having the same shape and design cannot be possible unless the color characteristic of the wheel is provided.
An additional reason to use a color camera is simply because operators do often find it easier to look at images in color. Although it is proposed to design a fully automatic inspection system that does not require human intervention, but it is wished to create a computerized record of each inspected wheel. When a customer returns with a complaint about a defective product that he received, the industry has to be able to properly respond to his complaint and to verify the cause of this mistake. As a result, it is necessary to keep a record for all the products that have been inspected and delivered to customers. This computerized record must contain a maximum quantity of information to describe the product (wheel type, wheel color, conformity, etc.) and ideally enables a second evaluation. To this purpose, it is important to provide a color image that describes the real aspect of the wheel.

3.2.3 Sensor type

The next important step involves picking a suitable sensor, built either around charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) image sensor technology. Both have the same task of transforming light (photons) into electrical signals (electrons). This information is, however, transmitted by both sensor types using different ways and means and the design of each is also fundamentally different.

In CCD sensors the charges of the light-sensitive pixels are shifted and converted into signals. The charges of the pixels, which are created by exposure to a semiconductor, are transported to a central A/D converter with the support of many very small shifting operations (vertical and horizontal shift registers), similar to that of a "bucket chain". The transfer of the charges is forced with the support of electrical fields, which are created by electrodes in the sensors.

In contrast, in the CMOS sensors, a capacitor as a charge storage is put in parallel to each individual pixel. This capacitor is charged with the exposure of each pixel by its photoelectric current. The voltage created in the capacitor is proportional to the brightness and the exposure time. Differently than CCDs, the electrons captured in the capacitors by the exposure of the sensor to light are not shifted to a single output amplifier but are transformed into a measurable voltage directly at the source by means of each pixel's own associated electronic circuit. This voltage can then be made available to the analog signal processor.

Unlike the CMOS sensor, the pixels on a CCD sensor use the entire sensor surface to capture the light, with no conversion electronics placed on the sensor's surface. This leaves more space for pixels on the surface, which in turn means more light is captured. This leads to the CCD sensor being more light-sensitive than CMOS sensor technology. On the other side, for CMOS sensors, by using additional electronic circuits per pixel, each pixel can be addressed, without the charge having to be shifted as with CCDs. This results in the image information being able to be read much more quickly than with CCD sensors and that artifacts due to overexposure such as, for example, blooming and smearing, occur far less frequently or not at all.

In light of the many technological improvements in recent years, the trend on the sensor market is increasingly pointed toward CMOS technology. Indeed, in the last few years, modern CMOS image sensors have emerged as a strong alternative to CCDs. In [91], the authors reviewed the advantages and disadvantages of the classical CMOS sensor compared to the CCD sensor. The main advantages were centered on their low power consumption [92], lower cost, high-speed imaging, and the ability to avoid blooming and smearing effects, which are typical problems of CDD technology [93]. While the main disadvantages of CMOS sensor compared

to the CCD sensor are the sensitivity, the generated noises, especially under low illumination, and the lower image quality. Then they presented the most important advances in the field of CMOS image sensors to overcome these disadvantages, focusing on fields such as sensitivity, low noise, low power consumption, low voltage operation, high-speed imaging and good dynamic range. They concluded that CMOS imagers are competitive with CCDs in many application areas, especially industrial applications, due to their low cost, low power consumption, integration capability, etc... Nevertheless, they demonstrated that CCD technology will continue as predominant in high performance systems, such as medical imaging, astronomy, etc... because of its better image quality.

Another important factor to consider when choosing the suitable sensor technology for the intended application is the shutter type. The shutter protects the sensor within the camera against incoming light, opening only at the moment of exposure. The selected shutter or exposure time provides the right "dose" of light and determines how long the shutter remains open. If the exposure period is too short, then the images end up underexposed; if it is too long, then the photos are overexposed. There are two shutter type options, either global shutter or rolling shutter. The difference between the two shutter variants is in the way they handle exposure to light.
The global shutter approach opens to allow the light to strike the entire sensor surface all at once. Depending on the frame rate, a moving object is thus exposed in a rapid succession. Hence they are suitable for applications that include moving objects.
On the other hand, rolling shutter exposes the image line-by-line, where the exposure time does not begin and end simultaneously, but rather for each individual row respectively. Depending on which exposure time is selected and the speed of the imaged object, distortions can occur when photographing objects moving during the exposure process, making this shutter type unsuitable for some applications. These distortions occur if the object or camera continues moving during the row-by-row exposure. As the image data is gathered, the exposed rows are reconstituted in the same sequence into a complete image. The sequential exposure of the individual rows is also visible in the way the distortions are formed in the reconstituted image. This is known as the rolling shutter effect.

For the wheel inspection task presented in this work, and in order to increase the productivity, it is proposed to take a photo of each wheel while it is passing under the imaging system. Stopping the wheel under the camera to take its photo, even briefly, decreases the number of inspected wheels per day below the limit imposed by the industry. Although there is no need to abandon the possibility of a rolling shutter just because the objects are moving, it is preferable to use the global shutter technique to make sure to avoid this rolling shutter effect. This is generally the case for the inspection of moving objects, and is especially true in an industrial scheme where the inspection specifications, more precisely the conveyor speed, may

increase over time for productivity reasons. As a result, the performances of a rolling shutter camera may no longer be sufficient to avoid the distortions.

Traditionally, CCD sensors have been the only ones to function with global shutters, while CMOS sensors only offered the ability to use the rolling shutter. However, as CMOS technology begins to surpass CCD, there are now high quality global shutter CMOS sensors available, where both shutter types are supported. At the end of this discussion, it is decided to consider CMOS sensors for our application, since the trend on the sensor market is increasingly pointed toward CMOS technology.

3.2.4 Resolution

Even after the previous steps, the choice of the most suitable camera is far from being over. Taking a look at the vision systems market, it can be seen that providers offer a very wide variety of cameras to choose from. Each provider then tries to keep ahead of the competition by providing a set of useful features that help to improve image quality, assess image data more effectively or control processes with greater precision. Such features are mainly software based, and include adjustable parameters to control the gain, the ISO, the exposure time, the Gamma correction, among many others. However, improving the image by software remains restricted to the camera's hardware capabilities. Hence the choice of the camera has first to be based on its physical features, or specifications, more precisely its resolution, its pixel size and its sensor size. These three factors exhibit a relationship that must be exploited depending on the needs of the application.

In practice, resolution describes a measurement of the smallest possible distance between two lines or points such that they can still be perceived as separate from one another within the image. In a camera spec sheet, the resolution refers solely to the total number of pixels on the sensor. This information by itself is not sufficient to determine the correct choice of resolution for the application. Hence, another definition can be found in industrial image processing, where the resolution is stated in terms of how much of an area or which dimensions of the object being inspected are depicted on a single pixel. For a specific detail to be available for inspection, the resolution must be better than the size of the detail itself. Thus, the resolution will be defined partially in correspondence with the order of magnitude of the smallest defects that are intended to be detected on the wheel surface. As mentioned in the system requirements (in section 3.1), the smallest defects that are wished to be detected have a size of about 2 mm in a single direction.

At this point, there is no general rule that defines the number of pixels necessary to represent the smallest defect, to be able to distinguish it in the image. This number heavily depends on the wheel texture, or the background of the area surrounding the defect. If the background is very smooth, one can consider that a

small number of pixels is sufficient to perceive the defect as an anomaly. As the texture is getting rougher, the needed number of pixels to represent the smallest possible defect increases. In this work, the vision system is intended to inspect several different designs of wheels, with different textures. In this case, it is mandatory to consider the roughest wheel texture to define the necessary number of pixels for the smallest defect.
An additional consideration to take into account when making the choice is the quantity of information needed to be gathered about the defect. In case where the defect is only depicted with a limited number of pixels, it is usually difficult to extract its geometrical characteristics as its length and width, or even its shape. These information are important to be able to classify the defect and judge its severity. In this work, the goal of the inspection system is not only limited to the detection of the defect, but also to characterize this defect according to various variables, such as its size and shape, in order to create a record of all the possible defects with their category.
With these considerations in mind, the number of pixels per defect is set to a large value of 8 pixels in a single direction.

Another factor that plays a role in calculating the proper resolution is the field of view, in other words, the dimensions of the area to be inspected and its potential displacement amplitude. It is calculated by summing the maximum dimensions of the area to be inspected (length or width, depending on the case), the range of motion, and a safety factor of a few millimeters.
In our case, for the wheel inspection application, the dimensions of the wheels defer from one design to another. Hence it is necessary to consider the biggest wheel to define our field of view. The dimensions of the biggest manufactured wheel in our case is of 17 inches diameter, which converts to 431.8 millimeters, or to round it up to 432 mm.
The second parameter is the range of motion of the wheel under the imaging system. Because the wheel is set on a conveyor belt that passes under the camera, the displacement of the wheel in the image could occur in two directions. The first is when the camera takes the picture a little bit late or early, hence the wheel will be shifted in the image in the direction of movement of the wheel. The second possible displacement will occur in the opposite direction, in case where the wheel is not perfectly centered on the conveyor belt. Therefore, the range of motion parameter will be considered as the maximum displacement in the two directions. In the direction of movement of the conveyor belt, it is proposed to install a retroreflective sensor to signal the presence of the wheel exactly under the camera. Hence, the position of the wheel in this direction will be precise, and the range of motion will be limited to about 2 mm. In the opposite direction, it is proposed to center the wheel by installing adjustable guide rollers on each side of the conveyor belt. The gap between the guide rollers on both sides can be adjusted according to the wheel size being manufactured. However, the range of possible gaps cannot precisely cover all the possible dimensions of wheels. Hence, for some wheels, the gap is

set to the nearest possible value to ensure a good wheel positioning. Even so, a non-negligible displacement could occur and could span up to 15 mm. As a result, the range of motion parameter is set to its maximum value of 15 mm, which adds up with the wheel dimensions to a total of 447 mm. We finally add a safety factor of a few millimeters to guarantee that the whole wheel is present in the image.
However, when designing the inspection system, it is important to anticipate some possible modifications in the products. For example, if the industry introduces an 18 inches wheel in the future, the inspection system should be able to adapt to this new feature without much adjustment. Hence, it is proposed to add an offset of 50 mm around the largest current wheel to ensure the possibility of inspecting wheels with a diameter up to 2 inches more, without any adjustment to the system. Consequently, the field of view for our application can be set to 500 mm.

Finally, with all the information above, it is possible to determine the minimal camera resolution required to respect the application requirements. This is done using the following equation:

$$res = \frac{(field\ of\ view \times number\ of\ pixels\ for\ smallest\ defect)}{smallest\ defect\ size} \tag{3.1}$$

where res is the minimal resolution in each direction. Then, applying the calculations will lead to $res = \frac{500 \times 8}{2} = 2000$ pixels. Among the available industrial cameras in the market, the most suitable camera resolution that we can find is a 4 MP camera with 2046×2046 pixels in the horizontal and vertical directions respectively. This is more than sufficient for the intended application, and it ensures that the system will be able to perform even if the system specifications evolve slightly.

Then, knowing the resolution of the camera, it is necessary to calculate the exposure time. As previously mentioned in section 3.1, the wheels do not stop under the imaging system, but they are in continuous movement with a constant speed of about $1m/s$. In order to eliminate the blurring in the acquired image, the wheel must not move more than one pixel during the exposure time. To ensure that with a camera of 2046 pixels in one direction, and a field of view of 500 mm, the exposure time must be set to a value not more than $\frac{1}{4000} = 245\ \mu s$. Recent industrial cameras can reach values of exposure time as low as 10 μs for some models. Keep in mind that with such a small value of exposure time, the lesser quantity of light accessing the camera sensor has to be compensated with a bright light source.

3.2.5 Pixel size and sensor size

Indeed, a higher number of pixels corresponds to an increase in the camera's resolution, i.e. the camera's ability to resolve fine details of an object and produce clear and sharp images. Larger pixels will allow the camera to have a higher

sensitivity to light and a higher dynamic range, which represents the camera's ability to reproduce the brightest and darkest portions of the image and how many variations in between. Light is the signal used by the sensor to generate and process the image data. Since the area of each pixel is increased, more photons can be collected over the surface of each pixel, rendering the camera more sensitive to light. Additionally, the greater the available surface, the better the Signal-to-Noise Ratio (SNR), especially for large pixels measuring 3.5 μm or greater. A higher SNR translates into better image quality. However, having larger pixels also means that a higher percentage of the sensor's light-sensitive area is used in each pixel, allowing for the conversion of more photons to electrons for each picture element. A trade-off therefore exists between resolution and sensitivity. Larger pixels will allow the camera to be more light-sensitive whereas smaller pixels over the same area will allow the camera to see finer details and smaller objects.
Therefore the necessity of a larger sensor, onto which more pixels can fit, which produces a higher resolution. The real benefit here is that the individual pixels can still be large enough to ensure a good SNR, unlike on smaller sensors where there is less space available and thus smaller pixels must be used. Hence, larger sensors allow for higher pixel density and larger pixel size. So the question becomes, why not just use the largest sensor possible ? The main disadvantage of having a larger sensor is that they are much more expensive, so they are not always beneficial. Another disadvantage is that larger sensors induce bigger cameras, which is not always convenient in an industrial installation, where the space may be limited.

Classic machine vision cameras have varyingly large sensors, with standard sizes. Industrial cameras usually use 1/3 inches sensors in case of resolutions of 640×480 pixels, cameras with 1280×1024 pixels mainly 1/2 inches. The quite popular camera resolution of 1600×1200 pixels often uses a somewhat larger sensor with 1/1.8 inches with the same pixel size. For our application, it is proposed to use a non-standard camera resolution of 4 MP with 2046×2046 pixels. With this resolution in mind, the choice of the sensor and pixel sizes will be limited to the availability in the market. The most suitable camera that can be found has a sensor size of $11.3\ mm \times 11.3\ mm$ and a pixel size of $5.5\ \mu m \times 5.5\ \mu m$. The performance of the camera with these dimensions is considered to be typical in machine vision applications, with a dynamic range of 58.7 dB and an SNR of 40.8 dB, according to the camera's technical sheet.

3.2.6 Interface

The final step on the path to find the proper camera for the wheels inspection application is selecting the suitable interface. The interface serves as the liaison between the camera and the computing platform, forwarding image data from the camera sensor to the components that process the images. As with many of the criteria for camera selection, there is no single best option interface, but rather the selection has to be done in accordance with the most appropriate interface to the application. Many factors are to be considered. However, the two most important

considerations in choosing the most suitable interface are the maximum bandwidth, or data rate, that can be achieved and the maximum theoretical cable length that can be achieved. A third factor can also be considered, as it is of importance in the inspection application presented in this work, which is the possibility of installing a multi-camera system.

Several interfaces are available to design a vision system. The Global Association for Vision Information (AIA) documents all the camera-to-computer interface standards available in the industry, along with detailed specifications on the hardware. It covers a variety of old-to-new interfaces, including IEEE 1394, Camera Link, CoaXPress, GigE Vision and USB Vision [64]. The latter two interfaces are the most recent, and have been adopted by the machine vision industry as the primary interfaces for imaging. This is due to their superiority in addressing modern requirements for data transmission, along with their lower cost as they do not require a frame grabber to function. The frame grabber is a piece of hardware that is added to the computing platform and which is responsible for capturing the streaming control and image data to and from the camera. In many cases this component is more expensive than the camera itself. Hence, we will only discuss the choice between these two technologies.

The GigE Vision standard is a widely adopted camera interface developed using the Ethernet communication standard (IEEE 802.3). It is the fastest growing interface on digital cameras used for industrial image processing. This popularity is the result of its data rate being able to attain more than 100 MB/s without the use of a frame grabber, and with a cable distance which can span up to 100 meters with a single cable. The interface's Ethernet framework enables the creation of camera networks, which makes GigE Vision the ideal interface for multi-camera applications. Multi-port GigE cards allow operating up to four GigE Vision cameras with full bandwidth using a single PC add-in slot. This allows users to simultaneously trigger cameras in multi-camera applications over the existing Ethernet cable, without need for additional I/O trigger cables, and in a very precise manner with a latency as low as a few microseconds.

When it comes to USB, one of the most popular computer interfaces in the world, a significant evolution introduced a substantial increase in speed in 2008. Going from USB 2.0 at 60 MB/s to 10 times the speed with USB 3.0 at 625 MB/s, USB 3.0 became an attractive interface for vision systems for its speed and plug-and-play nature. The official USB3 Vision standard defines the necessary elements that make USB 3.0 industry-compliant for vision applications. It takes in connectors and cable characteristics. It also defines the communication between a USB3 Vision device and USB3 Vision-compliant software. As with GigE Vision, this standard is an important step in allowing customers to rely on one interface choice over a long period of time. Just as the GigE technology, USB3 Vision cameras can be connected to a processing unit without additional hardware.

However, USB3 cable lengths are not as long as is possible with GigE, and can usually span up to only 5 meters in length.

In a wheel industry, there are usually multiple production lines that separate different types of wheels and ensure a higher productivity. At the end of each line, the inspection task is performed. In this work, it is wished to design a vision system that enables the inspection of all the wheels ate the end of 2 production lines. Hence, it is necessary to install an imaging system, which includes among others the camera and lighting, at the end of each line, along with a single shared operation room, where the computing platforms are located. This will allow a single operator to monitor the automatic inspection process, and to intervene in case of a system malfunctioning. In this case, a greater cable length is mandatory, which will favor the use of the GigE cameras. In addition, as previously mentioned, GigE Vision is the ideal interface for multi-camera applications, such as the one intended to design in this work.
Therefore, it is decided to consider the GigE Vision interface in this work.

3.3 Lens selection

The lens is at the front of the optical system. It captures the image and delivers it to the image sensor in the camera. Indeed, choosing the appropriate lens for the vision system is a vital part. In order to get a good and sharp image, you do not only need a good camera, but also the right lens for this specific camera, and for the application as a whole.

Many factors play a role when selecting the lens for the application. Some of these factors are directly related to the camera specifications, such as the camera resolution and its sensor size, while other factors depend on the imaging system, ranging from the installation constraints to the lighting conditions. In what follows, we will discuss some of the main factors that helped defining the proper lens for our application.

3.3.1 Lens mount

The search for a lens begins by looking at the type of lens mount that the camera has. This is the interface between the camera body and the lens. Lens mounts come in standard sizes and are labeled according to the screw-threaded type of the camera body.
Dozens of different mounts are available in the market. However, for machine vision cameras, or industrial cameras, C-mount is the most widely available type of lens mount. In fact, the type of lens mount to be used depends on the camera's sensor size. A C-mount is appropriate for a maximum sensor diagonal of 20 mm, which corresponds to 1.5 inches. For larger sensors, an F-mount is generally used, though this type is seldom used in industrial applications. On

the other hand, for smaller sensors, CS-mount and S-mount lenses are more suitable.

The camera that we chose to use in this work has a sensor size of 11.3 $mm \times$ 11.3 mm, which corresponds to a 15.98mm diagonal. Hence, a C-mount lens is convenient as it conforms with the camera housing type, and is suitable to the sensor size.

3.3.2 Image circle and sensor size

After determining the right mount, it is needed to determine the right image circle size, because C-Mount lenses have different image circle sizes depending on the model. Similar to the camera's sensor size, the image circle diameter of a lens is given in inches in its data sheet. Ideally, we are supposed to choose the same value for these two parameters. However, this is not always possible, especially for non-standard sensor sizes, as the one considered in this work, where it is difficult to find a lens that has the same image circle diameter. Hence, the rule to follow is that the image projected by the lens must cover an area equal or slightly larger than the camera sensor in order to fully utilize the sensor, while minimizing the costs.
In fact, using a lens with a small image circle diameter on a camera with a larger sensor size will cause some pixels to be wasted on the sensor, as the image circle will be too small and not cover the whole sensor surface. On the opposite, if the lens has a larger image circle diameter than the camera's sensor size, we will get an image without any problems. However, in this case, a large portion of the image circle is not used, which is a waste of money. Indeed, it does not matter how large the lens is, the size of the image is always determined by the sensor size. But the larger the lens, the more expensive it is. Thus, for a smaller sensor, it is always better to use an appropriately smaller lens.

With these considerations in mind, it is proposed to use a lens with an image circle diameter of 1 inch to be suitable for our camera's image sensor size.

3.3.3 Resolution and pixel size

A camera with a high pixel number in its data sheet implies high resolution; but this is only partially true when considered in combination with the lens. A lens needs to be optimally matched to the pixel number irrespective of how many pixels the camera and sensor can offer. A high resolution image can only be created if a high resolution lens is used.
The resolution of a lens is given in line pairs per millimeter (lp/mm) and specifies how many lines on a millimeter appear as separate from one another. Using a pattern of black and white lines, a measurement is made of how close the lines can be to each other while still being able to detect the boundaries between them. The more line pairs that appear as differentiated, the better the resolution of the

lens. This resolution has also a direct impact on the contrast level of the image. In fact, the boundaries between black and white lines are not perfectly sharp. The transition from black to white, or inversely, will always produce a blur into gray tones. The lower the lens resolution, the less sharp the transition, hence the lower the effective contrast. This could be prevented by always choosing a lens with a resolution higher than the camera's resolution. This will ensure a noticeably higher contrast even if the fine structuring detected by the lens is lost to a certain degree on the camera sensor. Once again, this comes on the expense of a higher lens cost.

For the considered camera in this work, its resolution and the physical measurements of its sensor result in a value of about 182 pixels/mm, which translates to a total of 91 lp/mm. Checking the available lens resolutions in the market, the choice has been set on a standard lens resolution of 120 lp/mm, which is more than enough for our camera.

3.3.4 Focal length

The next important consideration when selecting a lens is its focal length. A lens with a focal length approximately equal to the diagonal size of the sensor format reproduces a perspective that generally appears normal to the human eye, i.e. with the same field of view on objects with similar sizes. Lenses with shorter than normal focal lengths, also called "wide angle" lenses, can capture a larger field of view. Lenses with longer than normal focal lengths, or "telephoto" lenses, capture a smaller field of view. Therefore, when considering focal length, it is necessary to consider the sensor size, the field of view wished to capture, and approximately how far from the inspected object the lens is located, also known as the "working distance".

Many lens vendors provide lens selection calculators on their websites that produce a recommended focal length based on the approximate form of the focal length equation. Using this equation is straightforward when the image sensor size is known. Indeed, the focal length f is given by:

$$f = \frac{(sensor\ size \times working\ distance)}{field\ of\ view} \tag{3.2}$$

It would be impossible to provide a lens for each focal length there is, so there exist some standard values that all manufacturers provide. These are called "fixed focal length" lenses. The most commonly found lenses fall between 8 mm and 25 mm. When the calculated focal length has a non-standard value, two options exist: either choose the first lower standard focal length, where the camera views a larger area, or choose the first higher standard focal length, where cropping of the sides may occur. The general tendency in such a case would be to select the lower value option, as the field of view is generally important to maintain for an application. Another solution exists that consists in using a "zoom" lens, as they

conveniently provide a wide range of focal lengths, and thus field of view, in a single lens. However, for machine vision applications, this type of lens is usually avoided because it comes with some costs. One of them being the fact the optics are not as optimized as you would find in a fixed focal length lens, and so image quality is generally not as good. For instance, optical aberrations may occur, in which case either the light from any point of the image does not focus to a single point at the focal plane, or in other cases the focus will be different for different wavelengths of light.

In our application, all the required parameters to calculate the focal length have been previously defined, except the working distance. In machine vision applications, this parameter is usually controlled by the available space for the installation. This is also the case for our application. As the imaging system has to be installed over the conveyor belt, the space is restricted. Accordingly, the allowed distance between the camera and the inspected wheel ranges between 550 mm and 600 mm. It has been defined in the system requirements. Hence, to find the lens with the proper focal length, we must calculate the focal length required by the system on the two edges of the working distance. This will lead to the first focal length of $f_1 = \frac{11,3 \times 550}{500} = 12.43mm$ and the second focal length of $f_2 = \frac{11,3 \times 600}{500} = 13.56mm$. Following the discussion above, it is better to choose a focal length that is near the value of f_1 to avoid losing the defined field of view. In the market, C-mount lenses with a focal length of 12.5 mm and a resolution of 120 lp/mm exist, hence the choice has been set on this value of focal length. Although this focal length is a little bit higher than the value of f_1, but the resulting cropped surface of the image is much lower than the offset that we included in the field of view. Thus the wheel is guaranteed to be visible in its totality in the acquired image.

3.3.5 Aperture and depth of field

The aperture is the opening at the rear of the lens that determines how much light travels through the lens and falls on the image sensor. The size of the aperture's opening is measured in f-stops. They give the ratio of the focal length over the aperture diameter. The f-stops work as inverse values, such that a small f-number corresponds to a larger or wider aperture size, which results in more light coming through the lens. Conversely a large f-number results in a smaller or narrower aperture size and therefore less amount of light reaches the camera sensor.

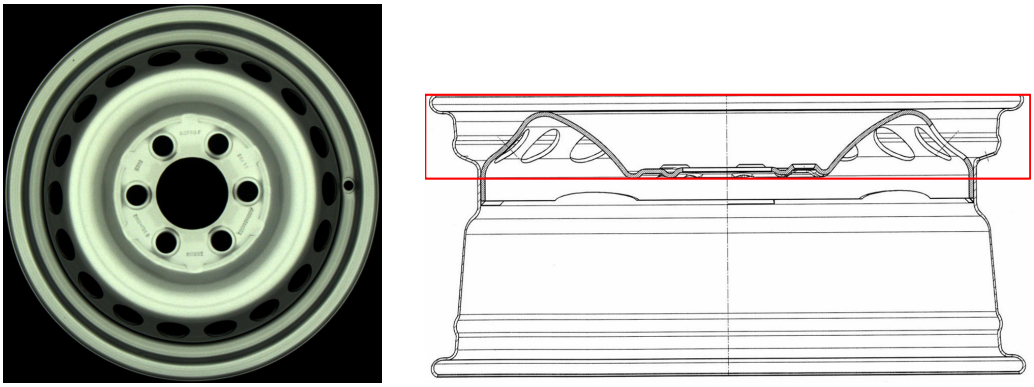
The value of the aperture of the lens is also the main parameter determining the depth of field. Other parameters also affect the depth of field, such as the object distance and the focal length. However, these parameters are usually fixed in the application. Hence, modifying the aperture is the easiest and most often utilized means to adjust it. The depth of field can be defined as the front-to-back zone of a picture in which the image is razor sharp. As soon as an object falls out of this

range, it begins to lose focus at an accelerating degree the farther out of the zone it falls; e.g. closer to the lens or deeper into the background. With any depth of field zone, or focus zone, there is a plane of optimum focus in which the object is most sharp. A shallow depth of field means that the area in focus is small, and a deep depth of field means that the area in focus is large.
The relationship between the aperture and the depth of field is easy: Large aperture means a small f-number which results in a shallow (small) depth of field. Vice versa, small aperture means a larger f-number which results in a deeper (larger) depth of field.

Choosing the proper aperture for an inspection application depends heavily on the inspected object shape. In the case where the inspection surface of the object is in a single plane, i.e. at the same distance from the camera, the plane of optimum focus could be set on this plane. As for the depth of field, there are no constraints on its size because the region of interest will always be in focus whether a large or small depth of field has been considered. Therefore, it is preferable to use a large aperture, i.e. a small depth of field, which enables a greater amount of light to reach the camera sensor, thus less additional lighting equipment to be used. This will help reduce the costs of the overall imaging system.

However, for the wheel inspection application studied in this work, this is not the case. The surface of the wheel to be inspected has a complex geometrical design, with elements on different distance levels from the camera. Figure 3.1(a) represents an image of the surface of a wheel intended to be inspected. The largest part of the wheel surface is the disc, while only the upper part of the rim surface is visible for inspection, which is the part surrounding the disc. Figure 3.1(b) shows a CAD illustration of the side view cross section of the same wheel. The red rectangle in the image locates the whole surface that is visible to the camera. It can be seen that this visible surface consists of several layers, starting from the rim sides, which are the highest points of the wheel, and going through the different layers of the disc until reaching its center, which is the lowest visible point that the camera can see.

To properly perform the inspection of the wheel surface, it is necessary that the whole zone located in the red rectangle to be in focus. The aperture should then be set to a value small enough to enable the depth of field to be larger than the width of the red rectangle. In addition, due to the complex shape of the wheel surface, multiple viewpoints are necessary to enable the inspection of its whole surface. Therefore, it is proposed in this work to use multiple cameras located around the wheel. These cameras will always be directed towards the wheel center, creating a hemisphere above the wheel surface, with the ability to modify the angle between each camera and the horizontal. The smaller the angle, the longer the wheel surface will appear to the camera, thus the width of the red rectangle will get larger. Then, as the angle is approaching zero, the depth of field will become the whole width of



(a) Example of a wheel (b) Side view cross section of the wheel

Figure 3.1: An example of a wheel to be inspected with an illustration of its design

the wheel surface. Generally speaking, for the proposed wheel inspection scheme, the larger the depth of field, the better for the inspection, considering that the focus will be higher in the depth of field zone, at its maximum at the plane of optimal focus, and will decrease going further from it. But this will come at a cost, because as the aperture is getting smaller, the light reaching the camera sensor is reduced, hence the need to compensate for the lack of light with a sufficient light source.

For machine vision applications, lenses with a manual iris are generally used to control the aperture level. These lenses are usually designed with two manual rings that can be adjusted: one is used to set the aperture level, and the other is used to select the plane of optimal focus. The maximum apertures are often between f1.2 and f1.8 for the considered focal length, and they can decrease to reach the level C "close", which refers to the iris being completely closed. For our application, it is sufficient to choose a lens with an aperture ranging between f1.4 and f22.

3.4 Lighting system

Without a doubt, an AVI system relies on quality images to ensure quality inspection results. High quality images enable the system to accurately interpret the information extracted from the object under inspection, resulting in reliable, and reproducible system performance. One of the most important factors that influence the quality of the acquired image in any AVI system is the lighting configuration. Indeed, images are created by analyzing the reflected light from the inspected object, not by analyzing the object itself. Therefore, the AVI system lighting is the most important aspect to consider after a camera-lens selection.

Designing a lighting system is not an easy task. Unlike the camera-lens selection that can be accomplished by following standard mathematical rules, when it

comes to the lighting system, there are very few standard rules to follow. Only one rule is for sure: a well-designed lighting system must enhance the features being inspected, while minimizing contrast of the surroundings. Many lighting system manufacturers provide their own guide to help design the suitable lighting configuration for the intended application. These guides usually consist of general practical steps for the designer to follow, along with a catalog containing a detailed description of the most popular lighting techniques. They mostly represent a series of experiments designed to understand the interactions between the different light sources and the inspected object, by taking into account numerous factors, including but not limited to, the size of the inspected object, its geometrical shape, its surface features, and other important inspection requirements. Then, bearing in mind the results of these experiments, the designer will be able to choose one or a few possible lighting techniques from the provided catalog. As a result, a prototype can be designed based on the specifications detailed in the catalog, in order to acquire some samples of real images. These images will finally help adjust the prototype until satisfactory results are obtained.

In this work, it is proposed to follow the guide proposed in [94] as the main reference, and to use the catalog in [95] to choose the suitable lighting technique. This catalog includes a list of 129 different lighting techniques, with their full description and specifications, which are necessary information to properly design the prototype. Then, after discussing the choice of the lighting technique, it is possible to decide on the type of light source that is most appropriate.

3.4.1 Lighting technique

The goal of lighting, as outlined in two of the most prominent machine vision lighting references [96] and [97], is to provide a consistent lighting environment that is immune to external and ambient lighting changes, reduce the amount of glare in the image, and most importantly, increase the contrast in the image such that the features being inspected are enhanced, while obscuring those that are not. Indeed, high contrast features simplify the integration of the inspection methods and improve the overall reliability of the system.
In addition, another important factor other than the contrast, is the uniformity of light on the inspected surface. In any AVI application, uniform lighting on the inspected object is crucial. Non-uniform illumination by an external light source can be significantly inefficient, forcing the use of additional image processing algorithms that are slower, more expensive and less reliable for inspection. This is especially true for real-time applications as the one presented in this work. In such cases, time is a major constraint. For that reason, it is best to have a well-designed lighting system that ensures a uniform illumination rather than to invest important processor cycles into making the inspection algorithm robust to light irregularities on the inspected surface.

Therefore, images with poor contrast and uneven illumination require more

effort from the system and increase processing time. Hence the importance of properly choosing the lighting technique that is best for the intended application.

In general, there exist a very large variety of lighting techniques in machine vision applications. The catalog in [95] for example, contains a list of 129 different lighting techniques, each with a different design. A practical way to classify these different lighting techniques, is according to the location of the light source relative to the inspected object and camera [94]. In total, they can be classified into four main categories:

- **Back lighting:**

Back lighting locates the source of light behind the inspected object toward the camera. This type of lighting is mainly used to inspect the outline shape of the object for dimensional measurements. It captures the inspected object as a dark silhouette against a bright background, creating a clear image of the outline of the inspected object, without any reproduction of the surface features.

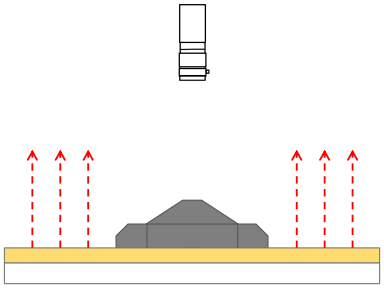


Figure 3.2: Back lighting

- **Bright field lighting:**

Bright field lighting, also called directional lighting, locates the source of light on the axis of the camera to about ±45° off the axis. This type of lighting is the most commonly used as it creates the best reproduction of what a person would see. However, it faces major drawbacks for inspection tasks, as it is limited to flat surfaces, and could generate "hotspot" reflections on reflective objects.

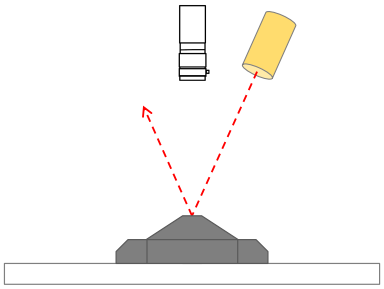


Figure 3.3: Bright field lighting

- **Dark field lighting:**

Dark field lighting is also considered as a directional lighting, but with the light source positioned at angles of 45° or less from the horizontal. Contrary to bright field lighting, this type of lighting highlights the edges of the inspected object, while flat surfaces remain dark. Hence, it cannot be used for inspection applications where the inspected surface presents several irregularities in different directions.

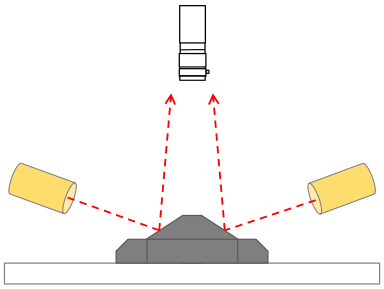


Figure 3.4: Dark field lighting

- **Diffuse lighting:**

Diffuse lighting is most commonly used on shiny and reflective objects where even, but multi-directional light is needed. It produces the most uniform illumination with the most realistic reproduction of the inspected object. There are multiple types of diffused lighting, some combining bright field and dark field lighting techniques to create a very even illumination of the inspected surface while eliminating reflections.

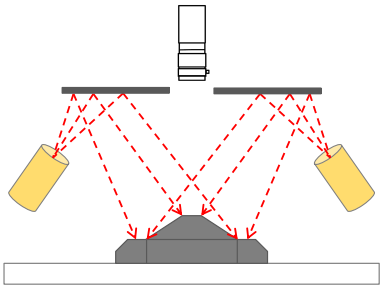


Figure 3.5: Diffuse lighting

To choose the most suitable lighting technique for our application, let us go back to the requirements defined in section 3.1. Here the focus is only on the characteristics of the inspected object itself, the surface of the wheel. Two main characteristics of the wheel will have major impact on the choice of the lighting technique. The first one is the shape of the inspected wheel surface. As mentioned previously, the surface of the wheel is not flat. It rather has a complex geometrical design, with irregularities in different directions. This can be clearly seen in Figure 3.1(b), which shows a CAD illustration of the side view cross section of a standard wheel. The second characteristic is the texture of the inspected wheel surface. Indeed, most of the finished wheels are coated with a glossy paint layer that renders the wheel surface more reflective.

With these characteristics in mind, and following the previous discussion on the different lighting techniques, it is obvious that the diffuse lighting technique is the most appropriate one for our application. It ensures a uniform illumination on the complex surface of the inspected wheel, with minimal light reflection artifacts in the acquired image which could result in false alarms during the detection procedure.

3.4.2 Lighting source

Several types of light sources are commonly used in machine vision. Quartz-halogen lamps, Fluorescent lamps, and Light-Emitting Diodes (LEDs) are by far the most widely used lighting sources in visual inspection systems [94]. In what follows, it is proposed to compare these three main types of light source in order to choose the most appropriate one for our application.

When an electric current passes through a thin tungsten wire, its temperature rises releasing thermally generated photons. This is the basic concept behind classical incandescent lamps. Then, an enclosing glass bulb is necessary to prevent air from reaching the hot filament, which would otherwise be quickly destroyed by oxidation. However, the evaporated tungsten gradually condenses on the inner surface of the glass envelope, reducing light transmission. Therefore, a major advancement

was the development of quartz-halogen lamps, for which the tungsten filament is sealed into a small quartz envelope filled with a halogen gas. When tungsten evaporates from the filament, it combines with the gas to form a tungsten halide which does not react with the quartz envelope, ensuring that it never reaches the outside glass envelope. Although this resulted in an improved life span and greater efficiency over classical incandescent lamps, this was not enough to maintain their popularity for AVI lighting systems [98].

The main drawback for quartz-halogen lamps is their limited life span. They can normally be expected to operate for about 2 000 hours [98], which is far shorter than the life expectancy of fluorescent and LED lamps. This means that a higher maintenance rate for the AVI system is required, which is not cost-effective.
Another important factor that affects the performance of quartz-halogen lamps is the reality that they change color when intensity changes. This is a major issue when the AVI system requires different light intensity levels for the inspection of different product types. For our particular case, manufactured wheels come in different colors that each requires a different brightness level. Indeed, to ensure the same inspection efficiency between various products, black coated wheels have to be much more illuminated than white coated wheels. But if by adjusting the light intensity from a wheel type to another, the color of the emitted light varies, then the inspection method must be able to adapt to those changes. This might be a complex, and more importantly a time consuming task. Hence, it is better to avoid such a problem in the first place by using a light source that does not have the same issue.

Moving on to fluorescent lamps. Each fluorescent lamp has two electrodes placed at the opposite ends of a glass tube filled with low-pressure mixture of ionized gaseous mercury and an inert gas. When current flows through the gas between the electrodes, the gas is ionized and emits ultraviolet radiation. The inside of the tube is coated with varying blends of metallic and rare-earth phosphor salts; substances that absorb ultraviolet radiation and fluoresce (re-radiate the energy as visible light) [99].

Despite the higher initial cost of a fluorescent lamp, compared to a quartz-halogen lamp, its lower energy consumption over its working life makes it cheaper. This life span, while higher than the one of quartz-halogen lamps, it is still far shorter than the life expectancy of LED lamps. They generally need to be replaced every six months to ensure a regular and steady brightness.
In addition, the brightness of fluorescent lamps, or more precisely their light output, has very low intensity, and is quite inconsistent compared to quartz-halogen lamps and LED lamps. This is mainly due to the unstable nature of the gas used in such lamps. As a result, fluorescent lamps cannot ensure a uniform lighting over the whole illuminated surface, which renders the inspection a difficult task.

Finally, let us review the features of LED which make it the ideal light source for the majority of AVI applications. LED is a semiconductor light source that is based on diodes which emit light when they are activated. When an LED module is

switched on, electrons are able to recombine with electron holes within the device, releasing energy in the form of photons [99]. LEDs are highly efficient and are electrically safe, since they only require low operating voltages, generating little unwanted heat.

The main features that contributed most in the increasing significance of LEDs for AVI systems are related to their light output performance. LEDs offer high intensity with very even, stable output with no flicker, producing a very bright light. This results in a clean, noise-free inspection of images with the highest possible accuracy.
In addition to providing better inspection performance and being incredibly reliable, LEDs require very little maintenance, as their life span could reach up to 100 000 hours. More importantly, LED modules do not die instantly; instead their light output slowly degrades over time [100]. This ensures the well-functioning of the AVI system for a long period of time, and helps avoid a sudden failure in illumination that could cause the system to break.
A final important feature of LEDs is that they are physically small as individual chips. These individual LED chips can then be packaged and mounted directly onto a rigid board, or flexible circuit, or even assembled into a 3D structure. Hence, it is possible to build light sources with a very wide range of shapes, to go along with the many possible complex designs of the lighting system. A popular packaging of LED chips is in square modules, so that they could generate a narrow beam of light that can be directed as needed, rather than spreading unwanted light in all directions.

For all these reasons, in this work, it is proposed to design a diffused lighting system that is based on multiple white LED modules. These modules are square shaped, and are linked together to generate the necessary amount of light required to ensure a quality inspection of all types of wheels. The whole lighting system is then connected to the power supply through a driver circuit, which enables to control the light output.

3.5 Processing platform

The last step on the road to design an AVI system is the choice of the processing platform. Recent years have been witnessing a fierce market competition to develop high-performance computing platforms. In general, these platforms can be classified into three categories [101]: Multi-core CPU, GPU, and FPGA.

The first, multi-core general-purpose microprocessors, or multi-core CPUs, integrate a few cores (from two to ten) on the same integrated circuit chip in an effort to speed up execution of computationally intensive tasks. The second, general-purpose graphics processing units, or GPUs, consist of a large number of cores (as many as several hundred), and are specifically oriented to maximizing execution throughput for parallel applications. The third, field-programmable gate arrays (FPGAs),

are becoming important, especially when higher performance/power computation ratios are desired.

Many researches have been dedicated to compare these three platforms by evaluating their performances in different sectors [102–106]. The obvious conclusion one gets after surveying the existing literature is that there is no clear winner for all problems [106]. Every platform has its own specific advantages that make it optimal for specific applications.

Modern GPUs are designed as programmable processors employing a large number of processor cores. At first glance, it seems that they are superior in performance than multi-core CPUs, which contain less number of cores. However, this is not always the case. GPUs are usually used not as standalone systems, but rather as computing co-processors along with a main CPU. This is because they are not optimized to deal with sequential processing. GPUs are also weak for non-parallel tasks since they do not have basic features like branch prediction. In fact, GPUs are especially well-suited to address problems that can be expressed as data-parallel computations in which the same program is executed on many data elements in parallel, by mapping data elements to parallel processing threads [105]. However, they face some major limitations when it comes to processing large amounts of data. Indeed, the GPU has its own high-speed memory to perform the processing tasks, but it is usually limited in space. Hence the need to transfer data between the CPU's memory and GPU's memory, in some cases multiple times. This is a major disadvantage of the GPU as these transfers are slow, and introduce undesirable delays in the program. Therefore, to ideally exploit the full performance of a GPU, a trade-off has to be respected between data transfers and processing parallelism. This cannot be generalized and has to be studied case by case.

On the other hand, FPGAs present a different case. An FPGA chip contains arrays of configurable logic blocks, programmable input/output blocks and a hierarchy of re-configurable routing resources that allow the blocks to be connected together. Logic blocks can be configured to perform different combinations of combinational and sequential logic functions. This is the major advantage of FPGAs over CPUs and GPUs, their re-configurability [105]. The function of an FPGA chip can be redefined, at the hardware level, to perform any type of processing to fulfill the requirements of the developers. This means that the system can be optimized to perform certain tasks with the highest possible performance. In contrast, if the FPGA has not been configured to perform the specific task, the performance will drastically drop. Hence, with each new method to be added to the system, the FPGA has to be reconfigured to meet the new requirements. However, reconfiguring an FPGA is a difficult and time-consuming process. FPGAs can only be reconfigured through a totally new design cycle where the chip's gates will be reallocated to a new architecture. In certain cases, new external block elements

have to be added to the FPGA chip to perform the required task. This whole process could take weeks to several months with each new added method [107]. Therefore, it is obvious that using FPGAs is not ideal at the early stages of designing a new system. During this stage of prototyping, the inspection tasks and the methods behind them will pass through various modifications and changes before figuring out the final inspection strategy. Hence, at first, the processing platform has to be highly flexible. Then later on, after deciding on the methods to perform the inspection, FPGAs will potentially be the ideal platform to use.

In conclusion, at this stage, it is best to adopt multi-core CPUs as the main processing platform, as they are the most flexible solution and the easiest to program, while providing high performances [107]. A GPU can then be added as a co-processor to perform the highly parallel tasks if necessary.

In this work, it is proposed to use a high-speed workstation with a Dual Intel Xeon Processor, each with 10 cores, 25 MB cache and a clock rate of 3.1 GHz. The system has a total of 32 GB of RAM. In addition, we added a 4 GB NVIDIA Quadro GPU. The image processing tasks are performed using the OpenCV library implemented using the C++ programming language, while the user interface is developed in Delphi.

3.6 System design and installation

In this section, it is proposed to describe the design of the AVI system, and its installation setup. While this section is located after the hardware selection, it is important to note that in reality, the conceptual design of the whole system must be done earlier, just after defining the system requirements. During this first process of design, the goal is to study the requirements imposed by the specific application, and then describe a model of a system that can meet them. Indeed, the concept design is the starting point in which the outlines of the different functions of the system are set. Then, based on this first design, the selection of the corresponding hardware can be done.

Before proceeding with the installation setup description, let us first review the different elements that have been selected to build the imaging system. Table 3.1 summarizes the results of the selection process detailed along this chapter. This table lists the main specifications of the four main elements of the AVI system: the camera, the lens, the lighting system, and the processing platform.

In this work, it is proposed to install a total of four cameras to properly conduct the inspection of the wheel surface. The first camera, referred to as camera $n°1$, is installed directly above the conveyor belt, over which the inspected wheels will be located. On the other hand, the three additional side cameras, referred to as

Table 3.1: Hardware specifications

Camera	Area scan Color camera CMOS sensor Global shutter Resolution: 2046×2046 (4 MP) Sensor size: $11.3mm \times 11.3mm$ Pixel size: $5.5\mu m \times 5.5\mu m$ GigE interface
Lens	C-mount lens Image circle diameter: 1 inch Resolution: 120 lp/mm Fixed focal length: 12.5 mm Aperture: f1.4 - f22
Lighting system	Diffused lighting technique White LED modules Controlled voltage output
Processing platform	Dual Intel Xeon processor 20 cores in total Cache: 25 MB Clock rate: 3.1 GHz Ram: 32 GB 4 GB NVIDIA Quadro GPU Image processing: C++ / OpenCV User interface: Delphi

camera $n°2$, camera $n°3$, and camera $n°4$, are installed around the inspected wheel at an angle of 120° apart. Each of these side cameras is installed on a curved track with a slider that enables to modify the angle between the camera and the wheel, which goes from 30° to 75°, while always maintaining the same distance between them. In addition, it is possible to modify this distance between each of the cameras and the inspected wheel. Hence, all the cameras can be positioned to adapt to the inspected wheel type. To acquire the images, a single trigger is used to control the four cameras in order to record the four images at exactly the same moment.

The main concept behind this design is to distribute the cameras in a hemispherical way around the inspected wheel, enabling a full inspection of its surface from all sides. Then, the diffused lighting system will be installed at the center of this hemisphere, directly above the conveyor belt. An illustration of the described

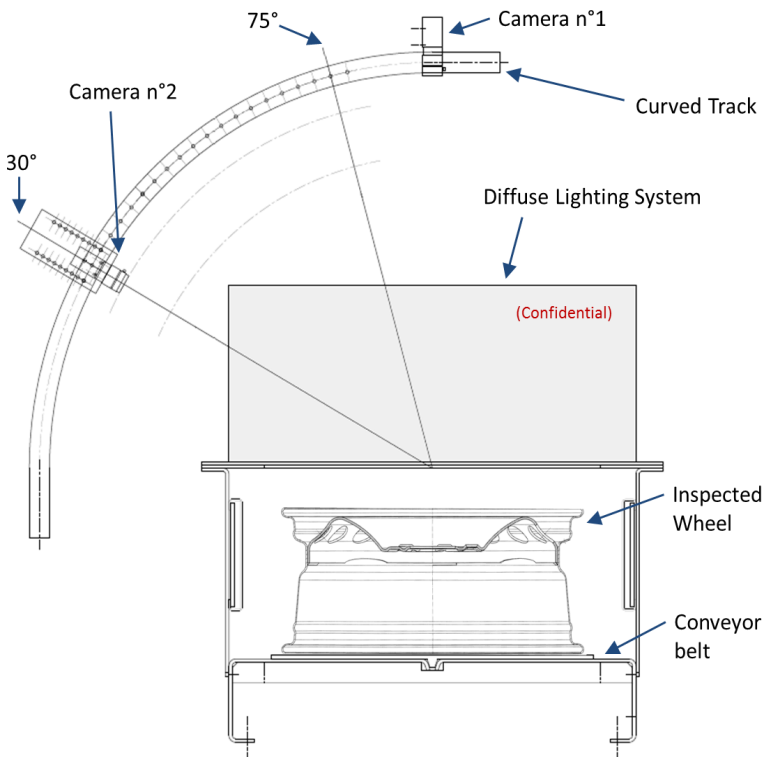


Figure 3.6: Illustration of the AVI prototype installation

installation setup is depicted in Figure 3.6. It only shows camera $n°1$ and camera $n°2$, as the installation of cameras $n°3$ and $n°4$ will be similar to that of camera $n°2$, but at different angles.

As for the processing platform, it is not depicted in Figure 3.6 as it is not installed near the imaging system. In fact, as mentioned in section 3.2.6, the processing platform is installed in an operation room far from the imaging system itself, encouraging the use of the GigE interface for the cameras. Indeed, this type of installation has many advantages. First, it enables the operators to have a single shared operation room to monitor different processes around the factory from one location, hence allowing for a faster intervention in case of a system malfunctioning. Second, the processing platform is usually designed to perform in certain conditions, at certain temperatures, in order to maintain the highest performances. Some of these conditions may not always be guaranteed in an industrial installation. In such a case, a possible solution is to use computer chassis that are customized to adapt to harsh industrial conditions. This solution, however, is not always efficient.

Afterwards, the first step after installing the AVI system is to calibrate the cameras to capture real colors. The calibration was performed using a ColorChecker [108]. It consists of a cardboard-framed arrangement of 24 squares of



(a) Image acquired using camera $n°1$. (b) Image acquired using a side camera.

Figure 3.7: Two images of the same wheel, one acquired with camera $n°1$ (a) while the other with one of the side cameras (b).

painted samples formulated to emulate common natural colors. It is used to assess the color rendering accuracy of a camera in order to calibrate it. Calibration is performed by capturing the image of the ColorChecker, then adjusting the appropriate RGB settings in the cameras' software so that the color reading matched with the corresponding provided reference. The resulting RGB settings have been used in capturing all the images presented in this work.

Figure 3.7 shows two images of the same wheel. The image on the left is acquired using camera $n°1$, while the image on the right was shot with one of the side cameras. As it can be seen, the side camera enables to better view some regions of the wheel surface that are not visible for camera $n°1$, allowing for their inspection. Indeed, the defect detection procedure for the two images is the same, but on different regions of the wheel.

Finally, the proposed wheel inspection procedure is divided into two main steps: pre-processing and defect detection. In summary, the pre-processing step involves the identification of the inspected wheel, the detection and inspection of its key elements, along with several small inspection tasks, to make sure that the inspected wheel correctly match with its CAD model and that it complies with geometrical requirements. A final procedure is to prepare the wheel for the defect detection step by splitting it into different circular zones and then unfolding these zones to create rectangular matrices on which the detection is performed. Meanwhile, the defect detection step consists of two main methods: a defect detection method for the detection of local defects present on the wheel surface, and a sequential detection method to monitor the wheels' topcoat intensity. Both the detection method and the sequential method are not limited to wheel surface inspection. Indeed, the detection

method is general enough to be applied for the detection of defects in any case where no prior information about the defects is available. As for the sequential method, it can also be applied for the monitoring of various industrial and non-industrial processes. These two methods, along with the pre-processing procedure, will be detailed in the following chapters.

3.7 Conclusion

This chapter discusses in detail the complete procedure to design an AVI system for the real-time surface inspection of finished wheels. The first step is to define the requirements and conditions under which the inspection must be performed. Some of these requirements are related to the inspected wheel, while others are imposed by the industry. Next, some possible solutions are proposed to meet each of the requirements, with the goal to outline the initial design of the whole system. Then, a detailed discussion on the appropriate choice of each key element of the AVI system is presented. These key elements are the cameras, the lenses, the lighting system, and the processing platform. Finally, this chapter provides a full description of the AVI system design and its installation setup.

The focus in this chapter was on the hardware part of the AVI system. More precisely the design of the imaging system, and its installation setup. In the following chapters, we will present the methods that have been developed to perform the inspection tasks.

Chapter 4

Defect detection method

This chapter is inspired by our paper: "Karim Tout, Rémi Cogranne, Florent Retraint, Statistical decision methods in the presence of linear nuisance parameters and despite imaging system heteroscedastic noise: Application to wheel surface inspection, In Signal Processing, Volume 144, Pages 430–443, 2018"

Following the design and installation of the imaging subsystem, which was discussed in detail in chapter 3, it is time to address the inspection tasks carried out by the AVI system. As seen in chapter 2, two categories of defects have to be detected on the wheel surface: local defects and global defects. Hence, the inspection tasks consists of two main methods: a defect detection method for the detection of local defects present on the wheel surface, and a sequential detection method to monitor the wheels' topcoat intensity, hence addressing the detection of global defects. While the imaging subsystem has been designed for the specific application of real-time surface inspection of finished wheels, both the defect detection method and the sequential method are not limited to wheel surface inspection, but are rather suitable for various other applications. This chapter will solely focus on the development of the defect detection method for the inspection of local defects, while the detection of global defects will be addressed in the next chapter (chapter 5).

Local defects are defined as a local heterogeneity or distortion from the reference texture of a surface. They represent a sudden variation on a limited area of the inspected surface. They stand out on the surface of the wheel and are the most frequent type of defects detected during the final inspection. These defects can take different shapes and sizes whether they are scratches, marks, geometrical deformation, etc. Therefore, it is intended to design detection method that is general enough to be applied for the detection of defects in any case where no prior information about the defects is available.

Prior methods for surface anomaly detection based upon images captured with an AVI system can be divided into three categories [109–111]: 1) Generic methods that are highly flexible as they are not based on any prior knowledge on inspected objects [112], 2) Specific methods that are based on ground truth or examples of a reference [111], and 3) Methods based on computer vision and image processing, see [113, Chap. 15] , that usually require prior information on the non-anomalous object.

The first category includes anomaly detection methods that do not require any prior model of object's structure. Different types of filters [114, 115], such as median filter or Gaussian filter, morphological operations [116], and histogram equalization [117], have been all applied for noise reduction, image enhancement, with the aim to improve the contrast between the anomaly and the non-anomalous background. These tools, followed by pattern recognition methods [118, 119] or thresholding operations [120, 121], illustrate the core architecture of this type of methods. First-order gradient filter followed by thresholding is one of the most commonly used approaches in this category [115, 121, 122]. More recently, state-of-the-art image processing methods, such as multi-resolution representation [123], sparse dictionary learning [124] and variational methods [125], have all been applied for automatic anomaly detection. Similarly, classification methods have been used for automatic recognition of anomalies, mainly with the help of supervised machine learning [120, 121, 126–128]. These methods consist of separating the inspected image into regions of distinct statistical behavior, based on the assumption that common properties can define all kinds of anomalies and distinguish them from any non-anomalous background. The existence of such properties is doubtful in practice and these methods are thus often sensitive to the object and anomaly geometry and to the presence of noise.

In the second category, the detection methods are based on a ground truth: a reference image of the non-anomalous background used as a model [111, 129]. The detection is thus straightforward as it is usually based on mere differences between the reference and the inspected image. If a significant difference is identified, the inspected image is classified as defective [117, 130]. Usually, the reference image is created by averaging multiple anomaly-free images [131]. In an alternative approach, the reference image is estimated from the inspected image using a filter consisting of several masks [132]. This approach is efficient but is very sensitive to experimental conditions, such as object position, illumination and projection angles. Moreover an accurate ground truth may be difficult to obtain in practice.

Finally, methods from the third category make use of prior statistical information on the non-anomalous object. Two main approaches have been proposed to introduce statistical prior knowledge: Bayesian and non-Bayesian approaches. The Bayesian statistical approach allows the design of efficient and rather simple methods for anomaly detection. However those methods require 1) that the anomaly occurs with known prior probability and, 2) that the non-anomalous object is also random with known a priori distribution. Those requirements limit the application of Bayesian methods.
For a more detailed review on methods for automatic defects detection, the reader is referred to [111, 112, 133, 134].

In the anomaly detection problem considered in the present work, the non-anomalous background of the inspected surface has no interest in the detection

process while it may hide the anomalies and, hence, may prevent their detection. In addition, we do not always have prior distributions to model both the inspected surface and the occurrence of anomalies. In such situations, it is more convenient to represent the expected non-anomalous background as a linear combination of basis functions, and to consider non-Bayesian hypothesis testing methods for anomaly detection.

As a result, it proposed to design a detection method that belongs to the third category, and, specifically, to non-Bayesian approaches. It is based on an adaptive model of the non-anomalous part of the inspected surface, also referred to as the "background". This original adaptive model is interesting as it allows the inspection of a large range of surfaces, with different geometries, without prior information or Computer-Aided Design (CAD) models of the inspected object; this extends the application of the proposed methodology to various quality inspection domains. In addition, the use of this model with an heteroscedastic statistical noise model of digital images prevents the need to calibrate the imaging system. Eventually, the proposed model is accurate enough to allow the detection of small defects that are hardly visible by naked eyes. The proposed method is then applied for wheels surface inspection to detect "appearance defects" that are located on the surface of the wheel. Indeed the surface of the wheel is rather complex to inspect and requires an accurate model, while as in most of industrial applications, the large number of wheels produced every day requires mastering precisely the properties of the statistical test. Hence, this specific application studied in this work allows challenging the efficiency of the proposed methodology.

Prior works [135,136] also rely on a similar approach for hidden data detection; fundamental differences, however, are that in the case studied in this work, no information of the potential anomaly (shape, size, position, etc.) is available and that the adaptive model is much more accurate which allows its use in a much wider range of applications.

The main contributions of the proposed method are the following:

1. An adaptive statistical model is proposed to represent the imaged surface. This model only requires knowledge of inspected objects geometry making, thus, the anomaly detection system is fully automatic and applicable to a wide range of surfaces.

2. The proposed model is accurate, to ensure high detection performance, and computationally simple, for real-time applications.

3. The heteroscedastic noise model is used to accurately describe the noise properties in raw images. Accordingly, for other types of images, the heteroscedastic model can be replaced with the appropriate model without having any effect on the detection accuracy.

4. The statistical properties of the method are explicitly provided. The detection threshold only depends on the false-alarm probability. Consequently, an operator can, for instance, prescribe a false-alarm probability easily and can know which type of anomalies can be detected with which probability.

4.1 Problem formulation

Let $\mathbf{Z} = \{z_{m,n}\}$ denote the noisy image, of the inspected surface, of size $M \times N$, where $(m,n) \in \mathcal{Z} = (\{1,\dots,M\} \times \{1,\dots,N\})$. During acquisition, each pixel is corrupted with various noises that change its value from the one expected upon counted photons on the camera sensor. Therefore, each pixel value $z_{m,n}$ at location (m,n) can be represented as:

$$z_{m,n} = \mu_{m,n} + \xi_{m,n} \tag{4.1}$$

where $\mu_{m,n}$ is the expectation of pixel $z_{m,n}$, or the noise-free value, and $\xi_{m,n}$ represents the noises corrupting the pixel at this location. It is usually assumed that all the noises corrupting the pixel value can be modeled as a Gaussian random variable [137]. As a consequence, the statistical distribution of value for the pixel at location (m,n) is given by:

$$z_{m,n} \sim \mathcal{N}\left(\mu_{m,n}, \sigma^2_{m,n}\right) \tag{4.2}$$

where $\sigma^2_{m,n}$ is the noise variance. This representation of a pixel is considered when no anomaly is present in the inspected surface. On the contrary, when an anomaly is present in the inspected surface, the expected value of the pixel is affected. Hence, $z_{m,n}$ can be written as:

$$z_{m,n} = \mu_{m,n} + \theta_{m,n} + \xi_{m,n} \tag{4.3}$$

where $\theta_{m,n}$ is the impact of the anomaly on pixel's expectation. In fact, the anomaly affects a limited area of the image, therefore $\theta_{m,n}$ is equal to zero except for a few pixels in which the anomaly is located. Then, when an anomaly is present, the model of the pixel at location (m,n) becomes:

$$z_{m,n} \sim \mathcal{N}\left(\mu_{m,n} + \theta_{m,n}, \sigma^2_{m,n}\right) \tag{4.4}$$

When inspecting an image of a surface with the goal of detecting an anomaly, two situations may occur $\mathcal{H}_0 = \{\text{there is no anomaly}\}$ and $\mathcal{H}_1 = \{\text{there is an anomaly}\}$. From equations (4.2) and (4.4), anomaly detection problem can be represented as a decision between the two following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \left\{\mathbf{z}_{m,n} \sim \mathcal{N}(\mu_{m,n}, \sigma^2_{m,n}),\ \forall (m,n) \in \mathcal{Z}\right\} \\ \mathcal{H}_1 : \left\{\mathbf{z}_{m,n} \sim \mathcal{N}(\mu_{m,n} + \theta_{m,n}, \sigma^2_{m,n}),\ \forall (m,n) \in \mathcal{Z}\right\}, \end{cases} \tag{4.5}$$

with $\theta_{m,n} \neq 0$ for some (m,n).

Formally, a statistical test δ is a mapping $\delta : \mathbb{R}^{M.N} \mapsto \{\mathcal{H}_0; \mathcal{H}_1\}$. When testing two hypotheses, the ultimate goal is to design a Uniformly Most Powerful (UMP) test, which maximizes the power function and satisfies a prescribed constraint on false-alarm probability regardless the anomaly. Let us denote $\mathbb{P}_{\mathcal{H}_i}$ the probability under hypothesis $\mathcal{H}_i$ with $i = \{1, 2\}$. The false alarm probability of a test is defined as:

$$\alpha_0(\delta) = \mathbb{P}_{\mathcal{H}_0}(\delta(\mathbf{Z}) = \mathcal{H}_1)$$

Conversely, the power of a test δ is defined as:

$$\beta(\delta) = \mathbb{P}_{\mathcal{H}_1}(\delta(\mathbf{Z}) = \mathcal{H}_1),$$

which also corresponds to $1 - \alpha_1(\delta)$ where $\alpha_1(\delta)$ is the missed-detection probability.

However, resolving such a problem is not straightforward due to various difficulties. In practice, the main difficulty is the presence of unknown nuisance parameters, in the definition of hypotheses, that have no interest for the anomaly detection problem. These nuisance parameters are the pixels' expectation $\mu_{m,n}$ that describe the background, or the non-anomalous part of the inspected surface. Though this nuisance parameter is not related to the detection problem, it must be carefully taken into account, through the design of a model that accurately describes this background, such that the nuisance parameter does not prevent the detection of anomalies. The model has to be designed with the highest accuracy for distinguishing anomalies from background and, hence, to enhance the detection of the former. However, due to the diversity of inspected surfaces in various applications, this becomes a complex task.

In addition, imaging devices are characterized by an heteroscedastic noise model that makes pixels' variance $\sigma^2_{m,n}$ a function of expectations $\mu_{m,n}$ [137]. This property greatly complexifies the removal of the nuisance parameter as well as the design of the ensuing optimal statistical test.

Another difficulty resides in the fact that in this work, it is considered that no information of the potential anomaly (shape, size, position, etc.) is available. In general, this is the case for many applications. Consequently, as the anomaly cannot be modeled, it is necessary for the background model to be accurate enough to describe the non-anomalous part of the inspected surface, while at the same time avoid the subtraction of the anomaly.

4.2 Adaptive parametric linear model

4.2.1 Background model

For each inspected surface, an original adaptive model is applied to subtract the anomaly-free background. In fact the background, the anomalous-free content of

the image that represents the inspected surface, acts here as a nuisance as it has no interest for anomaly detection while it must be carefully taken into account. It is proposed to use a parametric linear model to represent the background. Such a model has, indeed, indisputable advantages: it is simple and, hence, usually computationally efficient and can be used within hypothesis testing theory to design the anomaly detection method.

Again, let $\mathbf{Z} = \{z_{m,n}\}$ denote the noisy image of size $M \times N$. The inspected area corresponding to image $\mathbf{Z}$ is split into non-overlapping small blocks of size w and h (for width and height respectively). Let us also denote $\mathbf{z}_k$ the k-th block of the inspected image $\mathbf{Z}$; though this block can be seen as a matrix of pixels, it is represented as a vector for the application of the proposed method, typically the pixels are read lexicographically.

The linear parametric model that has been adopted is based on the following model for the block $\mathbf{z}_k$, when no anomaly is present:

$$\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}\mathbf{d}_k, \mathbf{\Sigma}_k). \tag{4.6}$$

Here $\mathcal{N}$ represents the Gaussian distribution thus the model (4.6) belongs to the very usual one that represents a block $\mathbf{z}_k$ as a sum of non-anomalous content corrupted with additive Gaussian noise. However, it is proposed in this work to use an original model for the content, non-anomalous part, and a more realistic model for the noise than the Additive White Gaussian Noise (AWGN) that models all the pixels as realization of independent and identically distributed (i.i.d) random variables. The model of the noise is presented in detail in the next subsection.

The linear parametric model is an obvious model (4.6) to represent the content. It essentially consists in representing all the pixels of the block $\mathbf{z}_k$ as a weighted sum of basis vectors that represent the columns of the matrix $\mathbf{H}$. The weight of this sum represents the vector of parameters $\mathbf{d}_k$. The model of $\mathbf{H}$ is based on the following two dimensional algebraic polynomial:

$$f(x, y) = \sum_{i=0}^{d_x} \sum_{j=0}^{d_y} c_{i,j} x^i y^j \tag{4.7}$$

with d_x and d_y the degrees of the polynomial on x and y respectively.

Over the block $\mathbf{z}_k$ the (discrete) coordinates can also be put into vector form, denoted as $\mathbf{x}$ and $\mathbf{y}$. The coefficients $c_{i,j}$ of the polynomial (4.7) can also be put into a vector $\mathbf{c}_k$ of size $(d_x + 1) \times (d_y + 1)$. Denoting as a matrix $\mathbf{F}$ the polynomial model:

$$\mathbf{F} = (\mathbf{1}, \mathbf{x}, \mathbf{y}, \mathbf{x}\mathbf{y}, \ldots \mathbf{x}^{d_x}\mathbf{y}^{d_y}),$$

the model of the background (4.7) can be written as:

$$\mathbb{E}[\mathbf{z}_k] = \mathbf{F}\mathbf{c}_k, \tag{4.8}$$

where $\mathbb{E}$ represents the expectation.

The model (4.6) - (4.8) is simple and efficient enough for several applications, see [138–140] for examples in modeling of Internet traffic and image processing. However, for other applications, as the one of wheel inspection presented in this work, the non-anomalous background is much too complex to be represented with a simple model that remains the same for all the blocks. In fact, a trade-off must be found to keep the degrees of the polynomial (4.7) as small as possible, for improving ensuing detection performance, while representing with the highest accuracy the content, to enhance the detection of anomalies within the residual noise.

This trade-off leads us to the design of an adaptive model, for which the matrix $\mathbf{F}$ does not remain the same but, instead, changes to take into account the specificity of each block. To this purpose, the proposed method actually exploits the shape of the inspected surface to represent the pixels that share similar profiles. The design of an adaptive linear model based on this idea is done using the Principal Component Analysis (PCA). PCA is a powerful tool to identify patterns in data, and highlight their similarities. In fact, the first principal components retain most of the variation present in the data, which can be added to the model (4.8) to approximate better the background.

Indeed, multiple methods for dimensionality reduction, other than the PCA, can be found in the literature. Probably the most robust ones are sparse dictionary learning methods that proved their efficiency, especially in image modeling [141]. However, considering the PCA for our proposed model can be justified by many reasons. First, our proposed model is mainly based on a polynomial model which can be designed to ensure the orthogonality between the model and the defect. Adding the principal components to the model will not affect this orthogonality, which is crucial for the defect detectability as will be discussed later in section 4.3.2. On the other hand, sparse dictionary learning methods are much more robust than the PCA, and thus may incorporate a large portion of the defect in the background model. Second, it is proposed in this work to design a background model accurate enough to represent a wide range of surfaces without prior training step. The model is then computed for each inspected surface on the image of the very same surface itself. To this purpose, the PCA can be applied as it can represent the observations without requiring a prior learning dictionary. A final reason can be mentioned is that the PCA provides good results for a rather low computational complexity, compared to other methods in the same category.

Similarly to the model (4.8) one can approximate the pixels' value from the block $\mathbf{z}_k$ as:

$$\mathbb{E}\left[\mathbf{z}_k\right] \approx \mathbf{P}_k\, \boldsymbol{\alpha}_k. \tag{4.9}$$

where $\boldsymbol{\alpha}_k \in \mathbb{R}^{\ell}$ is the vector of weights for the different Principal Components and the size ℓ the number of selected principal components. The Principal components

that are added within matrix $\mathbf{P}_k$ are extracted from the inspected image itself. The index k, here, emphasizes that for different blocks the part of the principal components differs.

With the addition of the adaptive part due to the first few components, the proposed model for representing the background, that is the expectation of the block $\mathbf{z}_k$ (4.6), can be written as :

$$\mathbb{E}[\mathbf{z}_k] = \mathbf{H}_k \, \mathbf{d}_k. \tag{4.10}$$

where the matrix $\mathbf{H}_k$ is made of the polynomial basis vector(4.8), from the matrix $\mathbf{F}$ and the first principal components (4.9), from the matrix $\mathbf{P}_k$. The matrix $\mathbf{H}_k$ is thus given by the concatenation of those matrices:

$$\mathbf{H}_k = (\mathbf{F}|\mathbf{P}_k)\,.$$

Similarly the weighting $\mathbf{d}_k$ vector represents the contribution of those different basis vectors.

4.2.2 Parameter tuning of background model

The range of applications for the proposed parametric model is not only limited to surface inspection, but can rather be used for any application that requires a model of the nuisance parameters for their rejection [142]. In particular, parametric models that are based on a polynomial model have been extensively used in image processing applications, such as image compression [143], or image coding [144]. Furthermore, the concept of adding an adaptive part to the model, in order to aid the polynomial part, has also proven its efficiency. It was mainly used in image processing applications to help model the discontinuities and edges in the image, whether for radiographic image inspection [110], or even to detect hidden data in images [136, 138].

As for the particular domain of surface inspection, the proposed adaptive model can be efficiently applied on partially smooth surfaces, with low textures. In practice, the polynomial part is primarily efficient to accurately represent homogeneous surfaces, or smooth surfaces, with little to no texture. Then, adding the principal components will offer a higher flexibility, and will enhance the performances of the model to handle minor surface complexities. Hence, the proposed model can be used for the inspection of a variety of surfaces, including steel surface [112], ceramic tiles' surface [145], glass surface [132], among others that have mostly a low textured surface.

Consequently, the texture of the inspected surface will have a major role in defining $\mathbf{H}_k$. In total, 5 parameters have to be properly tuned to accurately model the background, while ensuring an efficient detection of defects. The choice of these

parameters essentially depends, on the one hand, on the level of complexity of the background and its overall shape, and, on the other hand, on both type and size of potential defects.

First, the degrees of the polynomial d_x and d_y have to be large enough to accurately model the background. Depending on the level of complexity of the background in each direction, d_x and d_y might be defined differently. The more the complexity in one direction, the larger the value of polynomial degree in that same direction. Furthermore, it is proposed to add an adaptive part to the model, i.e. the principal components, to better approximate the background. This adaptive part has a role to identify common features in the background, or patterns, and model them using the PCA. In the present study, it is proposed to apply the PCA in a single direction, which represents the direction of the main pattern in the background, but indeed it can be applied on various directions simultaneously. Only the first ℓ principal components are added to the model. This number will increase with the complexity of the pattern, which will be defined mainly by the inspected surface shape. As a result, to more accurately model the background, it is preferable to have large values for the three parameters d_x, d_y and ℓ.

However, having larger values for these parameters may result in a large part of defects being modeled within the background. Therefore, when subtracting the background from the original image, a large part of potential defects will also be subtracted. This would reduce the level of detectability of those defects.

Second, the width w and height h of each block mainly depend on the potential defects size in the inspected area. If the defect affects the majority of pixels inside the block, the estimate of the linear model parameters $\mathbf{d}_k$ will be significantly impacted by the presence of the defect. Consequently, a large portion of the defect will be removed with the background subtraction, thus reducing the level of detectability. Therefore, it is important to define the size of the block according to the potential defects size, in a way to ensure that the majority of pixels in the block belong to the background. It is important to note that in most cases, the defect surface is more textured than the background itself. Hence, even if the defect occupies the majority of the block, it may always be detectable to a certain degree. This is due to the fact that the parametric model is designed to represent the background, therefore the more textured surface of the defect will not be well modeled and accordingly a portion of it will remain after the background rejection.

On the other hand, the size of the block has to remain reasonably small such that the parametric model may be able to accurately model the background. Larger blocks include more background data, and thus may require higher polynomial degrees and more principal components to enable the good modeling of the background.

Section 4.5.6 further discusses the choice of all those parameters, in the case of wheel surface inspection, and presents the methodology used to select the most

suitable values upon experimental data.

4.2.3 Noise model

The proposed method relies on the image of the inspected surface. However, any image is corrupted during its acquisition by various sources of noise. A usual model of noise corrupting raw images (that are not processed for quality enhancement, compression, etc.) can be obtained by considering the shot noise separately, due to Poissonian process of photo-counting, and the various read-out noises. In fact the former noise has the specificity that its variance depends on the expected number of counted photons. While on the opposite the latter noise has a variance that depends on experimental setup (such as temperature, exposure time, etc.) that is constant for all the pixels.

It is usually assumed that the number of photons counted over each pixel is high, so that the Poissonian process can be approximated as a Gaussian distribution, and that the read-out noise can also be modeled as a Gaussian random-variable [137,146]. Hence, this gives a model for all the noises corrupting the pixel at location (m, n) that can be written as follows:

$$z_{m,n} \sim \mathcal{N}\left(\mu_{m,n}, \sigma^2_{m,n}\right) \tag{4.11}$$

where $\mu_{m,n}$ is the expectation of pixel $z_{m,n}$, which represents its noise-free value, and the variance of all the noises is given by:

$$\sigma^2_{m,n} = a\,\mu_{m,n} + b. \tag{4.12}$$

These parameters (a, b) of the heteroscedastic noise model remain the same for all the pixels. Beside they depend on several acquisition parameters, hence parameters (a, b) are also constant for all the images taken with the same acquisition settings.

The model of the noise (4.11) - (4.12) is well known for being more accurate than the usual AWGN model for raw images and allows us to take into account the variance of each pixel in the ensuing statistical test, to improve its accuracy [146–148].
Additionally, it is important to note that it is possible to use any other type of images rather than the raw type, provided that the appropriate noise model for that type of images is used [149, 150]. In fact, for many applications, the model of noise corrupting the acquired image may be more complex, in which cases the noise characterization becomes a major problem. Many researches dealt with such cases, for instance by providing flexible approaches to modeling complex noise based on a robust version of the PCA [151, 152]. In all cases, replacing the noise model will only affect the normalization factor in the ensuing statistical test. However, in the present application, obtaining raw images is simple and ensures to keep as much information on the inspected surface as possible; there is, hence, no

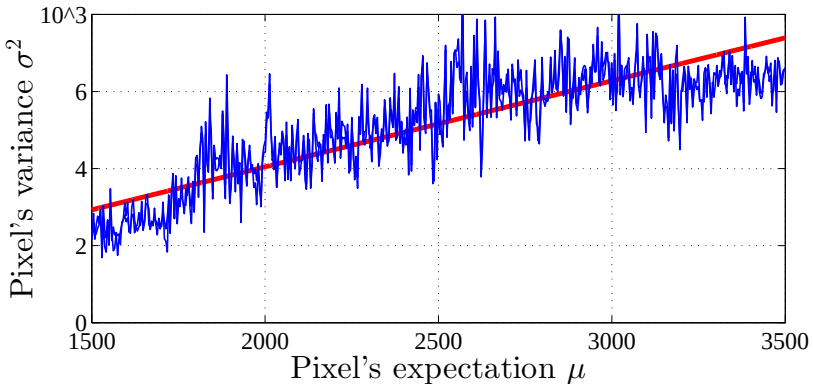


Figure 4.1: Illustration of the noise model showing, for several images, pixels' variance σ^2 as a function of their expectation μ

need to use a more sophisticated model for the noise corrupting such type of images.

An example of the relationship between pixels' expectation and variance is illustrated in Figure 4.1. This figure shows the variance of pixels as a function of estimated expectation from the same pixels along with the estimated noise variance estimated from the model (4.12). Those estimates have been obtained from a few RAW images.

In any industrial installation, it is supposed that the camera does not change so it can be calibrated easily. It is thus assumed that the noise model parameters (a, b) are known. The only parameter that has to be estimated, given an image of a surface under inspection, is the expectation of each pixel. However one can note that the noise corrupting the raw images cannot be modeled with i.i.d random variables, the celebrated Maximum Likelihood Estimation (ML) does not coincide with the Least-Square (LS). To tackle this estimation problem without applying a time-consuming optimization algorithm a two-step approach is proposed.
A first estimation is obtained applying the mere LS:

$$\widetilde{\boldsymbol{\mu}}_k^{\text{ls}} = \mathbf{H}_k \left(\mathbf{H}_k^{\text{T}} \mathbf{H}_k\right)^{-1} \mathbf{H}_k^{\text{T}} \mathbf{z}_k.$$

Then, a rough estimation of the noise variance is obtained from $\widetilde{\boldsymbol{\mu}}_k^{\text{ls}}$ by:

$$\widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}} = \mathbf{I}_{w \times h} \times (a\, \widetilde{\boldsymbol{\mu}}_k^{\text{ls}} + b),$$

where $\mathbf{I}_{w \times h}$ denotes identity matrix of size $w \times h$. This rough estimation of the covariance is thus reused to update the estimation of the expectation using the well-known Weighted Least-Square (WLS) given by:

$$\begin{cases} \widetilde{\boldsymbol{\mu}}_k = \mathbf{H}_k \left(\mathbf{H}_k^{\text{T}} \widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}^{-1}} \mathbf{H}_k \right)^{-1} \mathbf{H}_k^{\text{T}} \widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}^{-1}} \mathbf{z}_k, \\ \widetilde{\boldsymbol{\Sigma}}_k = \mathbf{I}_{w \times h} \times (a\, \widetilde{\boldsymbol{\mu}}_k + b). \end{cases} \tag{4.13}$$

It is, of course, possible to continue this procedure. It has been observed that this two-step method is a good trade-off between accuracy and computational time.

4.3 Statistical detection of anomalies

When inspecting an image of a surface with the goal of detecting an anomaly, two situations may occur $\mathcal{H}_0 = \{\text{there is no anomaly}\}$ and $\mathcal{H}_1 = \{\text{there is an anomaly}\}$. As described above, see Eq. (4.6), when there is no anomaly, any block of the image can be modeled as $\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k, \mathbf{\Sigma}_k)$. On the opposite, when an anomaly is present on the surface, any block can be modeled as $\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k + \boldsymbol{\theta}_k, \mathbf{\Sigma}_k)$. Here $\boldsymbol{\theta}_k$ represents the impact of the anomaly on pixels expectation. As described above, the anomaly affects a limited area of the image, therefore $\boldsymbol{\theta}_k$ is equal to zero except in a few blocks on which the anomaly is located.
Note that in this study, we consider that the presence of the anomaly has no effect on the variance.

Hence the goal of the studied anomaly detection problem is to decide between these two following composite hypotheses:

$$\begin{cases} \mathcal{H}_0 : \{\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k, \mathbf{\Sigma}_k)\,,\ \forall k \in \{1,\dots,K\}\} \\ \mathcal{H}_1 : \{\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k + \boldsymbol{\theta}_k, \mathbf{\Sigma}_k)\,,\ \forall k \in \{1,\dots,K\}\}\,, \end{cases} \tag{4.14}$$

with of course, $\boldsymbol{\theta}_k \neq 0$ for some k.

4.3.1 Statistical test

Formally, a statistical test δ is a mapping $\delta : \mathbb{R}^{w.h} \mapsto \{\mathcal{H}_0; \mathcal{H}_1\}$. The false alarm probability of a test is defined as:

$$\alpha_0(\delta) = \mathbb{P}_{\mathcal{H}_0}(\delta(\mathbf{Z}) = \mathcal{H}_1)$$

where $\mathbb{P}_{\mathcal{H}_0}$ denotes the probability under hypothesis $\mathcal{H}_0$. Conversely, the power of a test δ is defined as:

$$\beta(\delta; \boldsymbol{\theta}_k) = \mathbb{P}_{\mathcal{H}_1}(\delta(\mathbf{Z}) = \mathcal{H}_1),$$

which also corresponds to $1 - \alpha_1(\delta; \boldsymbol{\theta}_k)$ where α_1 is the missed-detection probability. We note that the power function $\beta(\delta; \boldsymbol{\theta}_k)$ depends, of course, on the anomaly $\boldsymbol{\theta}_k$.

When testing composite hypotheses, the ultimate goal is to design a Uniformly Most Powerful (UMP) test, which maximizes the power function and satisfies a prescribed constraint on false-alarm probability regardless the anomaly. However, such a test seldom exists. Hence, it is proposed to remove the nuisance parameters $\mathbf{H}_k\mathbf{d}_k$ and to design a Uniformly Best Constant Power (UBCP) test. Indeed, the expectation under $\mathcal{H}_0$ given by $\mathbf{H}_k\mathbf{d}_k$ has no interest for the testing problem (4.14) but must be taken into account.

To remove the nuisance parameters [153], the idea is to project the observations $\mathbf{z}_k$ onto the orthogonal complement of the subspace spanned by the columns of $\mathbf{H}_k$. This is achieved by using the projector:

$$\mathbf{P}^{\perp}_{\mathbf{H}_k} = \mathbf{I}_{w\times h} - \left(\mathbf{H}_k \left(\mathbf{H}_k^{\mathrm{T}} \widetilde{\mathbf{\Sigma}}_k^{-1} \mathbf{H}_k\right)^{-1} \mathbf{H}_k^{\mathrm{T}}\right) \widetilde{\mathbf{\Sigma}}_k^{-1}, \tag{4.15}$$

where the estimated covariance $\widetilde{\mathbf{\Sigma}}_k$ is given using the estimated expectation $\widetilde{\boldsymbol{\mu}}_k$ (4.13). One can note that the projection of observations $\mathbf{z}_k$ onto $\mathbf{P}^{\perp}_{\mathbf{H}_k}$ corresponds to subtracting from the observation the estimated expectation $\widetilde{\boldsymbol{\mu}}_k$. However, because the variance is not constant over all the pixels, it is necessary to normalize the "residuals" by dividing each residual by its standard deviation. Those normalized residuals can be written as follows:

$$\mathbf{r}_k = \widetilde{\mathbf{\Sigma}}_k^{-1/2} \left(\mathbf{P}^{\perp}_{\mathbf{H}_k} \mathbf{z}_k\right). \tag{4.16}$$

where $\mathbf{A}^{-1/2}$ represents the "square root" of the matrix $\mathbf{A}$ defined such that $\left(\mathbf{A}^{-1/2} \times \mathbf{A}^{-1/2}\right)^{-1} = \mathbf{A}$.

It is then easy to establish [110,153,154] that the norm of the normalized "residuals" $\mathbf{r}_k$ follows the distribution

$$\|\mathbf{r}_k\|_2^2 \sim \begin{cases} \chi^2_{\Upsilon}(0)\,, \forall k \in \{1,\ldots,K\} & \text{under } \mathcal{H}_0 \\ \chi^2_{\Upsilon}\left(\varrho_k\right)\,, \forall k \in \{1,\ldots,K\} & \text{under } \mathcal{H}_1, \end{cases} \tag{4.17}$$

where $\chi^2_{\Upsilon}\left(\varrho_k\right)$ denotes the non-central χ-squared distribution with $\Upsilon = w \times h - p$ degrees of freedom, here p denotes the number of columns of $\mathbf{H}_k$, and the non-central parameter ϱ_k under hypothesis $\mathcal{H}_1$ is given by :

$$\varrho_k = \left\| \widetilde{\mathbf{\Sigma}}_k^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}_k} \boldsymbol{\theta}_k \right\|_2^2. \tag{4.18}$$

Here ϱ_k denotes the "anomaly-to-noise" ratio [110] and is essential to define how detectable the anomaly is.

Based on the residuals $\mathbf{r}_k$ and their distribution, see Eq. (4.17), the UBCP test can be written as follows

$$\delta = \begin{cases} \mathcal{H}_0 & \text{if } \|\mathbf{r}_k\|_2^2 \leq \tau \\ \mathcal{H}_1 & \text{if } \|\mathbf{r}_k\|_2^2 > \tau, \end{cases} \tag{4.19}$$

where, in order to guarantee the false-alarm probability α_0, the decision threshold τ is set as follows:

$$\tau = F^{-1}_{\chi^2_{\Upsilon}}(1-\alpha_0; 0) \tag{4.20}$$

where $F_{\chi^2_{\Upsilon}}(x, \varrho_k)$ and $F^{-1}_{\chi^2_{\Upsilon}}(x, \varrho_k)$ resp. represent the non-central χ^2 cumulative distribution function and its inverse with non-centrality parameter ϱ_k.
Similarly the power function of the test is given by:

$$\beta(\delta, \boldsymbol{\theta}_k) = F_{\chi^2_{\Upsilon}}(\tau, \varrho_k). \tag{4.21}$$

One can note from the previous results, Eq. (4.20)-(4.21) two important things. First of all, the threshold τ only depends on the false-alarm probability α_0 and is thus constant for all the blocks. Second, the detectability of the anomaly only depends on the non-centrality parameter ϱ_k (4.18). More precisely, Eq. (4.18) shows that ϱ_k is defined as the part of the anomaly $\boldsymbol{\theta}_k$ that lies in the orthogonal complement of the subspace spanned by $\mathbf{H}_k$. Hence, an anomaly $\boldsymbol{\theta}_k$ is detectable if and only if $\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k \neq 0$.

4.3.2 Anomaly Detectability

To better understand the effects of the rejection of nuisance parameters on the defect detectability, let us consider the simple case where the observations $\mathbf{z}_k \in \mathbb{R}^3$ and $\mathbf{d}_k \in \mathbb{R}$, shown in figure 4.2. In this case, the nuisance parameter is scalar ($\text{rank}(\mathbf{H}_k) = 1$), and the column space of $\mathbf{H}_k$ is a vector $\mathbf{R}(\mathbf{H}_k)$. Its orthogonal complement $\mathbf{R}(\mathbf{H}_k)^{\perp}$, also referred to as the parity space, is then a plane orthogonal to the vector spanned by $\mathbf{H}_k$, and is depicted in blue in figure 4.2. When projecting the anomalous observations $\mathbf{z}_k$ onto the parity space, the nuisance parameters will be rejected, and only the projection of the anomaly $\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k$ will remain in the residuals. Consequently, the detectability of the anomaly $\boldsymbol{\theta}_k$ will mainly depend on how much of the anomaly is present in the parity space, depicted by the value $\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k$.

Then, in this work, it is proposed to normalize the remaining residuals by taking into consideration the noise corrupting the image. To this purpose, the more realistic heteroscedastic model of the noise has been used, allowing to establish with highest precision the theoretical statistical properties of the ensuing test. This noise model represents the variance of the noise corrupting the image as a linear combination of the pixel's expectation $\boldsymbol{\mu}_k$. Hence, because the variance is not constant over all the pixels, it is necessary to normalize the residuals by dividing each residual by its standard deviation. This procedure is illustrated in figure 4.2. Let us consider two different anomalies $\boldsymbol{\theta}_k$ and $\boldsymbol{\theta}'_k$ that have different projections onto the parity space $\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k \neq \mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}'_k$. If these two anomalies belong to the same surface $\left\{ \boldsymbol{S_c} : \left\| \widetilde{\boldsymbol{\Sigma}}_k^{-1/2}\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k \right\|_2^2 = c^2 \right\}$ with c a positive constant, this means that they have the same value of non-centrality parameter ϱ_k after normalization, thus have the same level of detectability. Indeed, the surface $\boldsymbol{S_c}$ is defined by the heteroscedastic noise model, and thus it is shaped like a cone with the radius increasing with $\boldsymbol{\mu}_k$.

We have observed that the proposed adaptive model is very efficient in the sense that it represents the background accurately (non-anomalous part of the inspected wheel) while preserving the vast majority of the anomaly within the orthogonal complement.

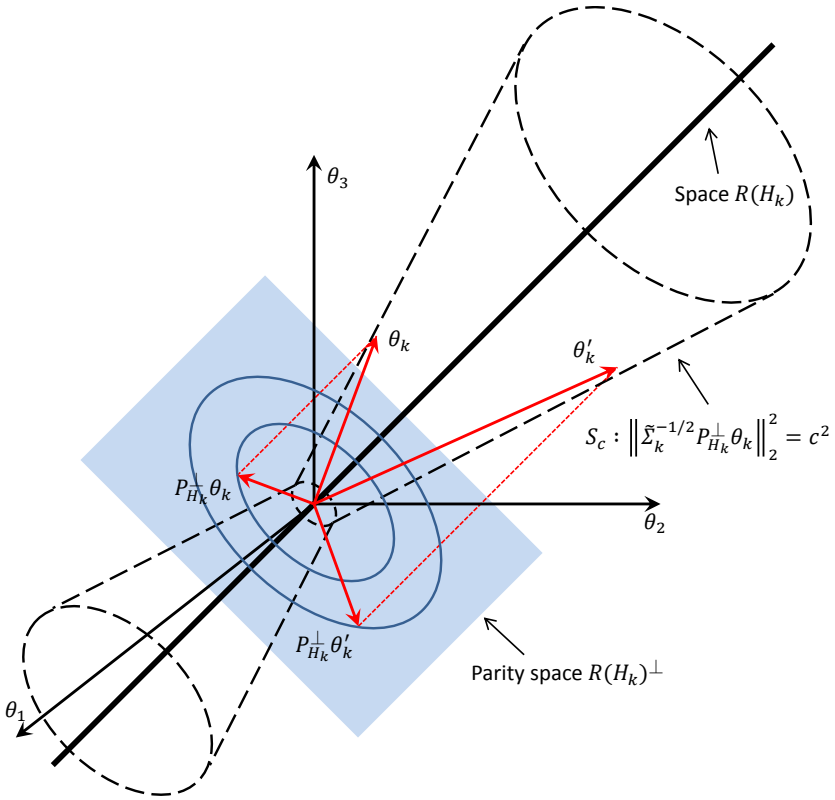
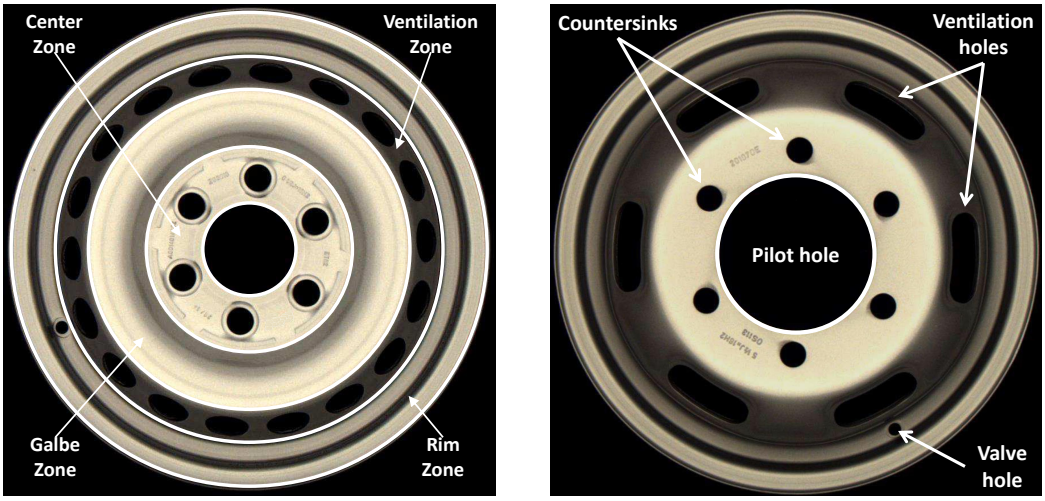


Figure 4.2: Illustration, in R^3 of observations along with their projections on the parity space $\mathbf{R}(\mathbf{H}_k)^{\perp}$ and a surface of constant power for which "anomaly-to-noise ratio" ϱ_k, see (4.18) is equal.

4.4 Wheel Inspection Characteristics

The use of AVI systems has been extending to reach various applications. For wheel surface inspection, the detection of "appearance defects" is one of the most challenging tasks that has not been studied yet. "Appearance defects", such as scratches or painting drops, do not have a direct impact on the proper functioning of the wheel, but rather are associated with the aesthetics of the product. These defects are located on the upper surface of the wheel, the one visible to the client. There is a large variety of "appearance defects" that can be classified according to their shape, size, and location on the wheels' surface.

As for the wheels, there are plenty of wheel designs with different shapes and sizes, but they all have some common features and essential elements that characterize a wheel. Based on those elements, it is possible and necessary to define the regions of interest (ROI) on the surface of the wheel, in order to carry out properly the wheel inspection.



(a) Wheel parts (b) Key elements

Figure 4.3: Description of the different elements of a wheel

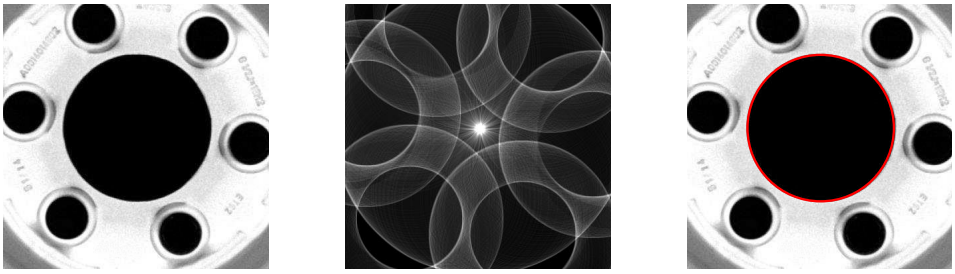
4.4.1 Region of Interest Extraction

The face of the wheel is a complicated surface to inspect. Each wheel is designed with specific parameters that define its form and geometry. These parameters can be used to split the wheel into different parts (zones), on which the detection method will be applied, see Figure 4.3(a). Multiple elements can be found in these zones that have to be identified in order not to consider their presence as a defect and, on the opposite, to detect anomalies on those elements. The key elements are the pilot hole, the countersinks, the valve hole and the ventilation holes, see Figure 4.3(b).

In brief, countersinks are used to mount the wheel on a vehicle, the valve hole is needed to mount the tire valve and the ventilation holes are important to cool the brakes (discs and pads). Those elements must be detected and localized prior to the detection of anomalies. Hence, let us first briefly describe how those key elements are detected. Indeed, the position of those elements will be used to perform a geometrical readjustment, which can be considered as a self-calibration.

It is important to note that the parameters describing the geometry of the wheel, and especially the key elements mentioned above, are known because the design of the wheels currently manufactured is also known. The knowledge of those parameters, such as the wheel radius and pilot hole radius, for instance, is useful as it allows to reduce the search area.

First of all, it is needed to detect the center of the wheel that coincides with the center of the pilot hole. The detection of the pilot hole and the localization of its center is carried out using the Circular Hough Transform (CHT). This is one of the most robust and commonly used methods for circular shape detection [155]. In addition, the prior knowledge of the radius makes the CHT computationally very efficient as the only two unknown parameters are the coordinates (x_0, y_0) of the



(a) Wheel center (b) CHT result (c) Hole Detection

Figure 4.4: Illustration of steps for pilot hole detection

center of the pilot hole. An example of the application of CHT for the pilot hole detection and localization of its center is presented in Figure 4.4.

Once the Pilot hole is located, the countersinks and the valve hole can be detected. Once again, knowing the distance from the wheel center to the countersinks and the valve hole helps reduce the search area along with the computational complexity. The countersinks and the valve hole are also detected using the CHT. The detection of countersinks only adds a simple additional check on the angle between the detected circles to make sure that all the countersinks are detected at the right position. Then, because on any wheel the valve hole is either in front of a countersink, Figure 4.3(a), or between two countersinks, Figure 4.3(b), the detection of the valve hole is done on a small number of areas that correspond to the known distance from the pilot hole center.

The last and most complicated step is the detection of the ventilation holes. Those elements may exhibit a wide range of different designs for aesthetic reasons but also in order to reduce the weight of the wheel with minimal structural impact. This high variety in shape of the ventilation holes explains the choice of the active contour models, or Snakes, for their detection [156, 157]. Such models have been extensively used in image segmentation in order to detect complex geometrical forms.

The final result of the detection procedure to locate all those elements for two different types of wheels is shown in Figure 4.5.

Once the pilot hole, countersinks, valve hole and ventilation holes are detected and located, one can split the wheel within several areas, see Figure 4.3(a), which are all inspected separately. Note that, though not very original, the detection of those elements has to be extremely robust as any error will lead to a false alarm of the anomaly detection method. To give an idea, the number of wheels produced over one year is about 4 million wheels, and yet, not a single error has been observed in those key elements location.

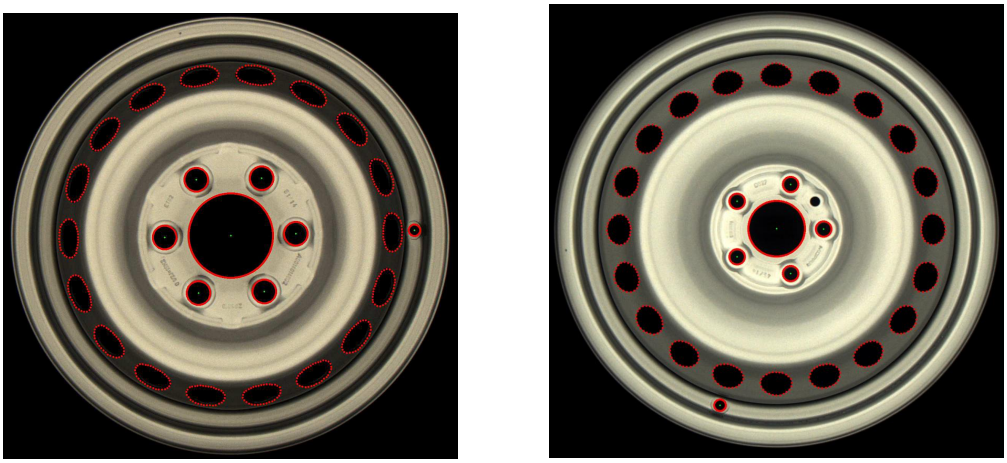


Figure 4.5: Key element localization for two different types of wheels

4.4.2 Data Preparation

AVI systems for anomaly detection have been widely used for their efficiency and unbiased results compared to human inspection. But, on the other hand, one great drawback in certain cases is that such systems might be time-consuming. Thus the urge to find new solutions that are fast and reliable.

In fact, in the industrial world, we call the production line, the set of sequential operations put together to produce an end product that is suitable for consumption. In order to maintain the synchronization of the production process, certain constraints have to be respected along the production line. One of the most important constraints to be respected is the pre-defined delay for each operation. The anomaly detection process is one of the main operations on the production line, essential to ensure the safety and well-being of the end product. Thus the urge to find new solutions that are fast and not time-consuming.

A first solution might be the use of simple and basic processing methods for anomaly detection, such as the gradient filtering for instance. However, such methods usually go short in terms of detection accuracy and precision, thus making them insufficient for the client requirements. Another solution is to use multitasking, or parallel processing. In fact, this approach highly depends on the inspected product, together with other important factors.

For wheel inspection, the fact that all the different zones are totally independent from one another, allows us to split the wheel, and apply the detection process in parallel on each zone. Another important advantage in splitting the wheel is that each individual zone has specific characteristics and features, thus specific model parameters. As we can see, all the different zones of the wheel described in figure 4.3(a) have a circular form. To apply appropriately the detection process, a

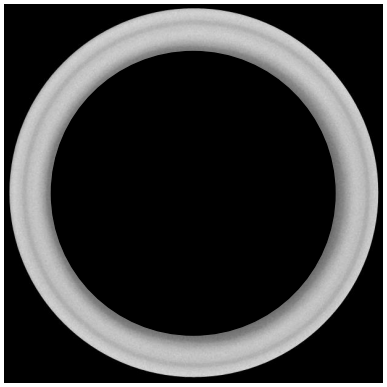


Figure 4.6: Circular galbe

rectangular image is required. Thus, we proceed by transforming the circular form into a rectangular one.

To explain the data preparation procedure, the galbe zone is considered as an example (see Figure 4.3(a)).

First, we start by isolating the galbe zone from the rest of the wheel, as illustrated in figure 4.6.

The second step is to unfold the circular area to create a rectangular one. Let us denote $(\mathbf{r_1}, \mathbf{r_2})$ the lowest and highest radius of the galbe area respectively. Normally, a circle of radius $\mathbf{r}$ has a perimeter of $2\boldsymbol{\pi}\mathbf{r}$. As $\mathbf{r_2} > \mathbf{r_1}$, hence $2\boldsymbol{\pi}\mathbf{r_2} > 2\boldsymbol{\pi}\mathbf{r_1}$, the circular area will turn into an isosceles trapezoid with a length ratio of $\mathbf{r_2}/\mathbf{r_1}$. In fact, to get a rectangle, either you define the length as $2\boldsymbol{\pi}\mathbf{r_1}$, thus you lose some pixels, or you define it as $2\boldsymbol{\pi}\mathbf{r_2}$ and you add the missing pixels to complete the form. For the purpose of not losing any information, it is clearly best to define the length of the resulting matrix as $2\boldsymbol{\pi}\mathbf{r_2}$ and add the missing pixels by duplicating its neighbors. It is worth noting that the number of missing pixels is proportional to the ratio $\mathbf{r_2}/\mathbf{r_1}$, which is in our case less than 1.5 for all the different types of wheels.

As a result, to unfold the circular galbe properly, the resulting rectangular matrix must be of height $\mathbf{r_2} - \mathbf{r_1} + 1$ and of length $2\pi\mathbf{r_2}$. In fact, the unfolding procedure can be represented as a transition from the Polar coordinate system to the Cartesian coordinate system. Let us denote $(\mathbf{x_c}, \mathbf{y_c})$ the coordinates of the center of the wheel. Each pixel of the galbe in the Polar coordinate system is represented by its radius $\mathbf{r}$ and its angle $\boldsymbol{\theta}$. The transition system can be represented as :

$$\begin{cases} \mathbf{x} = \mathbf{x_c} + \mathbf{r} * cos(\boldsymbol{\theta}) \\ \mathbf{y} = \mathbf{y_c} + \mathbf{r} * sin(\boldsymbol{\theta}) \end{cases} \tag{4.22}$$

where $(\mathbf{x}, \mathbf{y})$ are the Cartesian coordinates of the correspondent Polar coordinates $(\mathbf{r}, \boldsymbol{\theta})$.
Figure 4.7 represents the resulting unfolded galbe to be inspected.

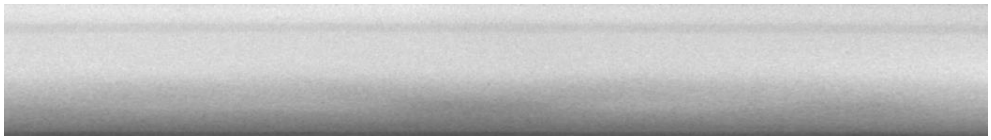


Figure 4.7: Typical example of the unfolded galbe of a wheel.

4.5 Experiments and results

4.5.1 Common core of all experiments

All the images used in this work are raw images that are made of 2046×2046 pixels of 12 bits depth; In what follows, most of the experiments presented only use the red channel for simplicity and clarity.

Regarding to the adaptive part of the background model, as mentioned in section 4.2.1, it is beneficial to use the shape of the inspected surface to better design this part of the model. Due to the circular shape of the wheel, we applied the PCA by considering the columns of the unfolded image as different observations. Hence, the columns of $\mathbf{P}_k$ in equation (4.9) represent the part of the first principal components, computed for each inspected wheel on the image on the very same wheel itself, that corresponds to the location along the rows of the extracted block $\mathbf{z}_k$ and reshaped such that it is constant along the rows.

Then, in order to estimate accurately the heteroscedastic noise model parameters (a, b), a batch of raw images of test pattern has been used. As mentioned previously, these parameters (a, b) only depend on acquisition parameters. All the images of the different types of wheels are acquired using the same camera settings; the parameters (a, b) are, thus, constant for all the images. Using the method proposed in [147] for the estimation of parameters (a, b) from the noise model resulted in $a = 2.23$ and $b = -420$.

The rest of this section is divided into four parts. In the first part, the advantages of adding the adaptive part, based on the PCA, in the proposed model of the background are proven. The second part investigates the advantages of using the heteroscedastic noise model rather than the usual AWGN model. The third part is aimed at studying the performance and accuracy of the proposed adaptive model of the background. And finally, the fourth and last part compares the performance obtained using the proposed detection method with performance from other recently proposed surface defect detection methods. Note that the first two parts of this section can also be considered as a comparison between the proposed method and other parametric methods that neither contain adaptive parts within linear model, nor take into account accurate imaging system heteroscedastic noise model.

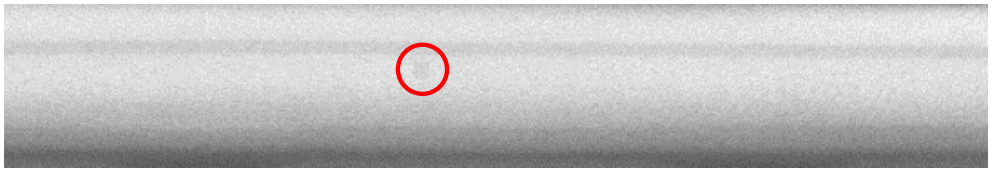


Figure 4.8: Example of a typical defect that it is wished to detect

4.5.2 Improvement of Detection Accuracy Due to the Model Adaptivity

Figure 4.8 presents an example of the galbe zone, with a typical defect that represents the lower limit of the detection criteria, above which the defect is intended to be detected. It has a circular form which can be considered to simulate various types of real defects. Note that the defect is highlighted with a red circle as it is rather difficult to see from naked eyes. This defect is used in the first and third parts of the experiments.

In the first part of the experiments, the goal is to investigate the advantages of adding the adaptive part, in other words, the Principal Components, to the proposed background model. To this purpose, it is needed to compare the performance of the proposed model in two different scenarios, where in the first scenario the background model consists of the polynomial part and the adaptive part, and in the second scenario the background model only consists of the polynomial part.

The defect in Figure 4.8, has been used to perform a Monte-Carlo simulation on 3 000 images. Because it is hardly possible to obtain many images with similar defects, we picked randomly a set of 3 000 non-anomalous images on which the defect has been superimposed. In fact, it is rather difficult to obtain images with defects, while images of wheels without any defect are easy to obtain. The proposed model parameters have to be adjusted in a way to highlight the effect of the adaptive part. As the Principal Components are used to assist the polynomial function to better model the complexity of the wheel along the rows, it is then possible to increase the block height and study the performance. Therefore the block is set to a size of $h = 40$ (height) and $w = 40$ (width).

Figure 4.9 and Figure 4.10 compare the mean power value of the statistical test performed on the 3 000 images, in the two-case scenarios mentioned above, for different values of the defect intensity and defect radius respectively. In Figure 4.9 the defect radius is set to 5, and in Figure 4.10 the defect intensity is set to 150. In both cases, the false-alarm probability is set to 0.01.

In the two figures, the blue plots have been obtained using the proposed adaptive model with the degrees of the polynomial set to $d_y = 5$ (along the height) and $d_x = 2$ (along the width) and the number of Principal Components added to this model is $\ell = 3$ giving us a total of $p = 21$ parameters, which is very small compared

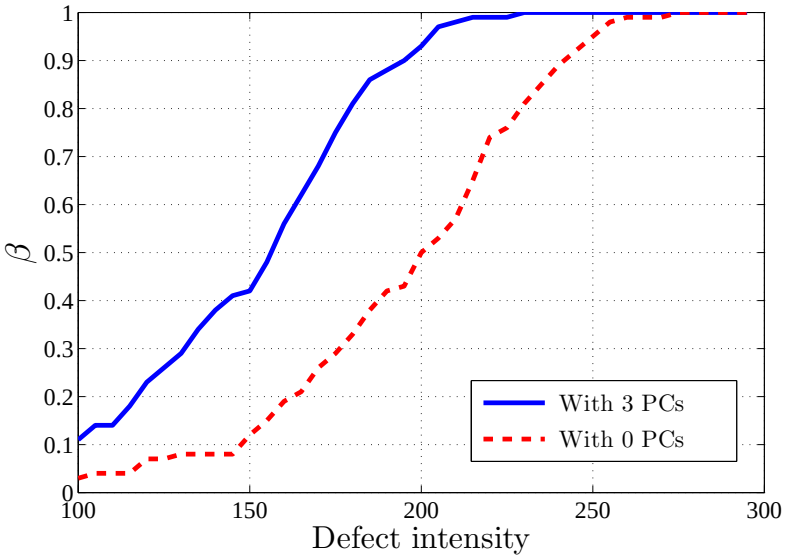


Figure 4.9: Real power curves with $\ell = 3$ PCs and $\ell = 0$ PCs function of the defect intensity with a fixed defect radius = 5

to the number of pixels (1600). As for the red plots, they have been obtained using the same polynomial degrees, but without any Principal Components added.

As it can be clearly seen in the two Figures, the model with Principal Components added outperforms the model without any Principal Components added in terms of detection power. For a defect radius of 5 and a defect intensity of 150, the detection power values can be read from both figures. If no Principal Components are added, the detection power only reaches a value of 0.12. By adding 3 Principal Components to the model, the detection power increases to a value of 0.42, which is almost 4 times higher than the previous case. One can also notice from Figures 4.9 and 4.10 that the two plots converge to a nearby value of 1 with the increase of the defect intensity or radius.

4.5.3 Improvement of Detection Accuracy Due to Heteroscedastic Noise Model

In this section it is aimed to highlight the advantages of using the heteroscedastic noise model rather than the usual AWGN model. As mentioned in subsection 4.2.3, the heteroscedastic noise model expresses the relationship between pixels variance σ^2 and their expectation μ as a linear polynomial. As for the AWGN model, the variance σ^2 is considered constant all over the image, independent from pixels expectation.

To investigate the choice of the noise model, it is necessary to study the effect of pixels' expectation on the normalization process. In fact, the noise model that

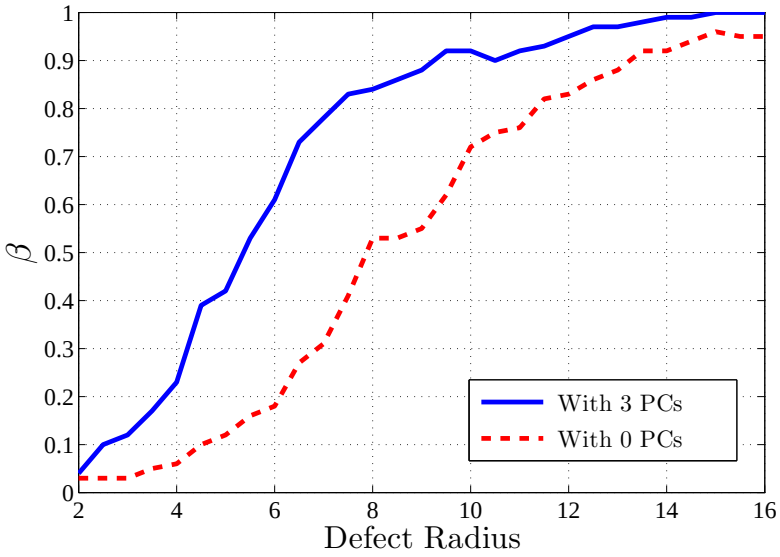


Figure 4.10: Real power curves with $\ell = 3$ PCs and $\ell = 0$ PCs function of the defect radius with a fixed defect intensity = 150

one uses defines the covariance matrix which has a primary role to normalize the residuals norm. From (4.17), the empirical distribution of the normalized residuals norm must follow a central χ-squared distribution with $\Upsilon = w \times h - p$ degree of freedom. Any inaccurate normalization will have a direct effect on the centrality of the non-anomalous normalized residuals norm.

Figure 4.11 represents two blocks within two different regions of the galbe zone. As one can notice, the block 1 is located on a bright region of the galbe zone, with high pixels' expectation values. As for the block 2, it is located on a darker region of the galbe zone, which means lower values of pixels' expectation. It is important to note that Block 1 and Block 2 do not contain any type of defect.
Those two blocks are used to perform a Monte-Carlo simulation on 3 000 non-anomalous images using the proposed adaptive model, the first time with the heteroscedastic noise model, and the second time with the usual AWGN model.

Figure 4.12 represents the empirical distributions of normalized residuals norm $\|\mathbf{r}_k\|_2^2$ for Block 1 and Block 2 using the heteroscedastic noise model. The two



Figure 4.11: Block 1 and Block 2

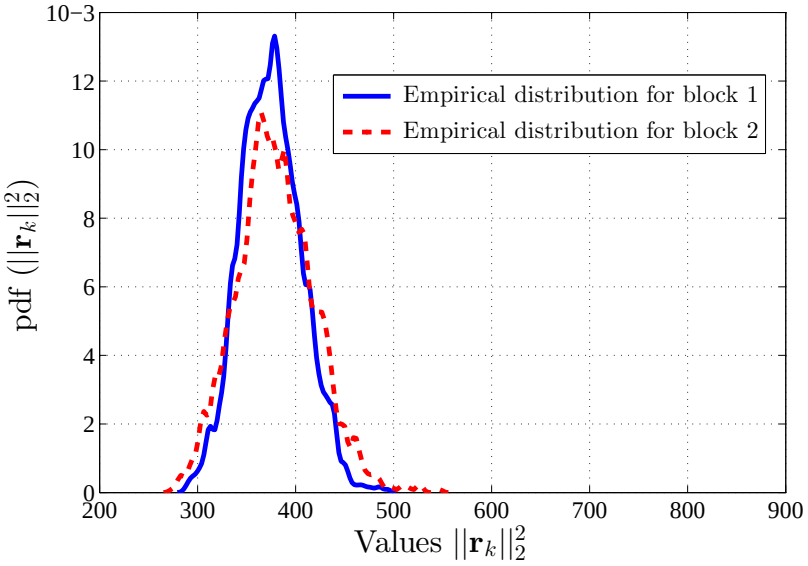


Figure 4.12: Empirical distributions of normalized residuals norm over block 1 and block 2 using the heteroscedastic noise model

distributions are centered around the same value of $w \times h - p = 20 \times 20 - 21$, which indicates that the normalization is adapted to the pixels expectation variation.

On the other hand, Figure 4.13 represents the empirical distributions of the normalized residuals norm $\|\mathbf{r}_k\|_2^2$ for Block 1 and Block 2 using the AWGN model. We observe that the empirical distribution for Block 1, which has high values of pixels expectation, is shifted to the right, while the empirical distribution for Block 2, which has low values of pixels expectation, is shifted to the left.

This result can be explained by the model from Equation (4.12). By using the AWGN model, the normalization of the residuals norm is independent from pixels' expectation values. As a consequence, the variance is considered constant all over the image. In reality, the relationship between pixels' expectation and variance is a linear polynomial (4.12) with a positive slope value (see Figure 4.1), which means that the pixel variance increases with its expectation. For Block 1, having higher pixels' expectation values in reference to the average expectation value, implies higher variance values, which also indicates the need for a higher normalization factor. As for Block 2, it is the other way around; Lower pixels expected values in reference to the average expectation value, implies lower variance values, which also indicates the need for a lower normalization factor.

These results show the efficiency of the heteroscedastic noise model to achieve the appropriate normalization of the residuals norm. In conclusion, one can clearly state that, in case of dealing with raw images, the AWGN model is not efficient, and that the use of the heteroscedastic noise model is crucial to achieve the perfect normalization results.

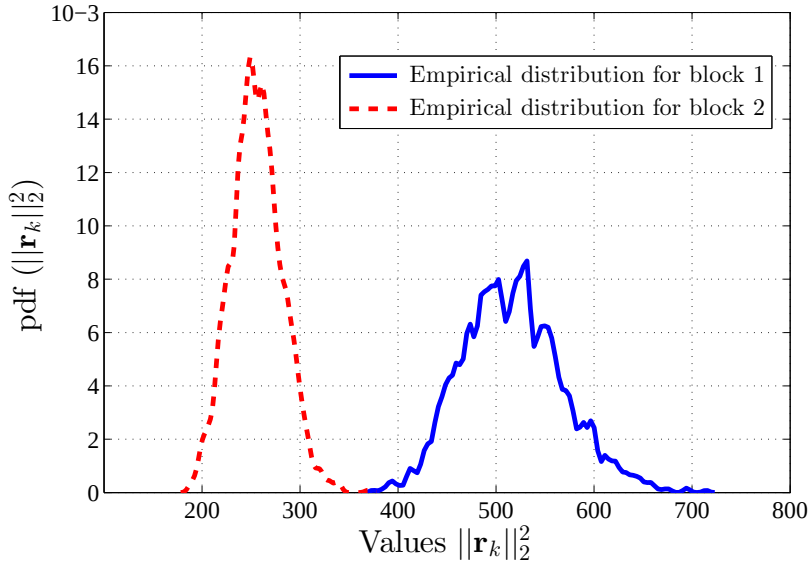


Figure 4.13: Empirical distributions of normalized residuals norm over block 1 and block 2 using the AWGN model

4.5.4 Comparison Between Empirical and Theoretically Established Results

In the third part of the experiments, it is wished to show the relevance of the proposed statistical test and the accuracy of the theoretical results. Once more a Monte-Carlo simulation has been performed on 3 000 images with and without the presence of the defect shown in Figure 4.8. This defect has an intensity of 200 and a radius of 5.

Figure 4.14 presents the empirical distribution of normalized residuals norm $\|\mathbf{r}_k\|_2^2$ on which the proposed test is based, see Eq. (4.19). It has been obtained using the proposed adaptive model over blocks with size $h = 20$ (height) and $w = 20$ (width). Note that for this experiment, the size of the block has been reduced to more precisely model the background. The degrees of the polynomial used are $d_y = 5$ (along the height) and $d_x = 2$ (along the width) and the number of Principal Components added to this model is $\ell = 3$.

The gap between the empirical distribution under $\mathcal{H}_0$ and the empirical distribution under $\mathcal{H}_1$ is due to the non-central parameter ϱ_k under hypothesis $\mathcal{H}_1$ (4.18). As mentioned above, the defect used in this simulation represents the lower limit of the detection criteria, above which the defect is intended to be detected. With this information in mind, Figure 4.14 shows that the detection of the defect is possible with certain classification error.

Besides, Figure 4.14 also compares the empirical distribution under $\mathcal{H}_0$ with the theoretical one (4.17). A small discrepancy can be observed between the empirical and the theoretical distributions.

The effect of the discrepancy can be illustrated in Figure 4.15 which represents

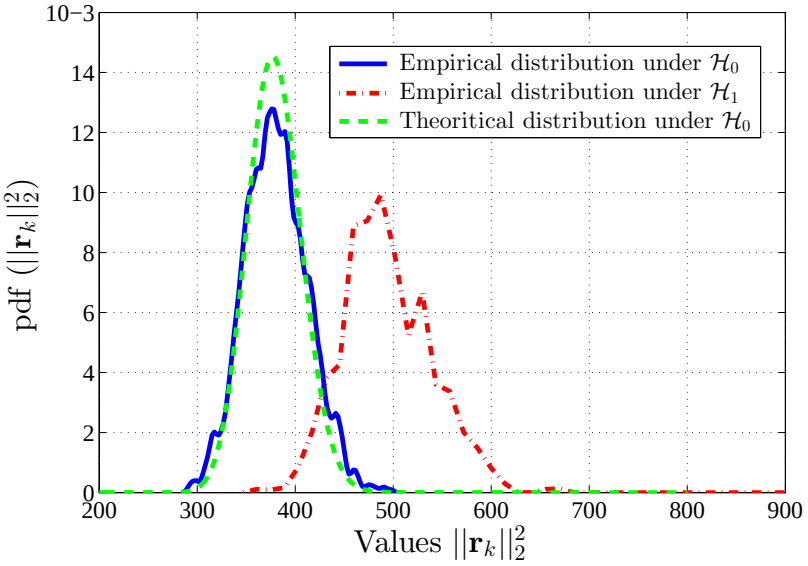


Figure 4.14: Empirical and theoretical distributions of normalized residuals norm for images with and without defect

the real and theoretical ROC curves. As one can notice, the performance of the proposed model is slightly lower than it is expected theoretically.

This can be explained by the two following facts. First, the estimation of pixels expectation is not perfect and has itself a non-negligible variance, which is not yet taken into account in the proposed test. Second, the proposed adaptive model, though efficient, is not perfect and, hence, maybe sometimes unable to describe the background with highest accuracy, putting part of the non-anomalous background among the residuals.

4.5.5 Comparison With the State-of-the-art

Finally, in this last section of numerical results, it is wished to compare the performance of the proposed method with methods recently proposed in the literature for surface defect detection. To this intent, Monte-Carlo simulations have been performed on 3 000 images with and without the presence of a defect that has the same form as the one shown in Figure 4.8, but with an intensity of 300 and a radius of 5.

The first simulation has been performed using the proposed adaptive model over blocks with size $h = 20$ (height) and $w = 20$ (width), with the degrees of the polynomial set to $d_y = 5$ and $d_x = 2$ and the number of Principal Components added to this model set to $\ell = 3$.

The second simulation has been performed using the detection method detailed in [158]. It consists of two phases: 1) a global estimation based on the Phase Only

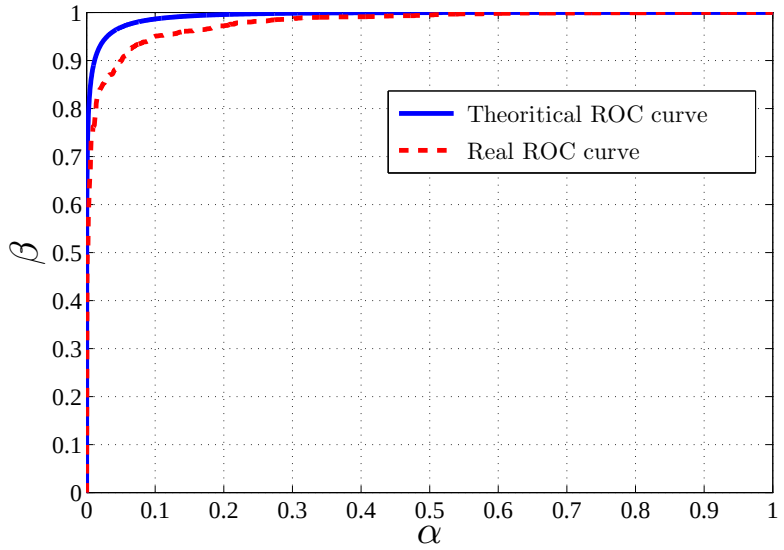


Figure 4.15: Real and Theoretical ROC Curves

Transform (PHOT) method, which considers the defect as an abrupt change in the image regularity, thus removes the regularity by normalizing the image's Fourier Transform by its magnitude. 2) Then a local refinement procedure which locally refines the estimated region based on the distributions of pixel intensities derived from defect and defect-free regions. Two parameters have to be tuned correctly depending on the input image texture and defect size: the Mahalanobis distance which is the threshold value for detection, and the size of the squared patch used in the local refinement step. They are set to 4.0 and 5×5 respectively, as suggested in the paper [158].

The third simulation uses the detection method introduced in [159]. They propose a regularity measure for defect detection in non-textured and homogeneously textured surfaces based on PCA. The method consists of a small neighborhood window that slides over the inspected image and for each window the regularity measure is then derived from the PCA. Again, two parameters have to be properly selected, as they have major effects on the detection performance. First, a control constant K which defines the threshold value for detection. A small value of K will generate false alarms, whereas a large value of K will result in high miss-detection rate. In order to choose the proper value of K, the paper proposes to select the minimum K value that generates no false alarms when applied to a defect-free training sample. For the proposed performance test, the proper value of K is 3. The second parameter to be tuned is the sliding window size. It should be large enough to contain the entire defect area, however, if it is excessively large it may smooth out the defective area and result in miss-detection. The defect implemented in the performance test has a radius of 5, thus the selected window size is 15×15.

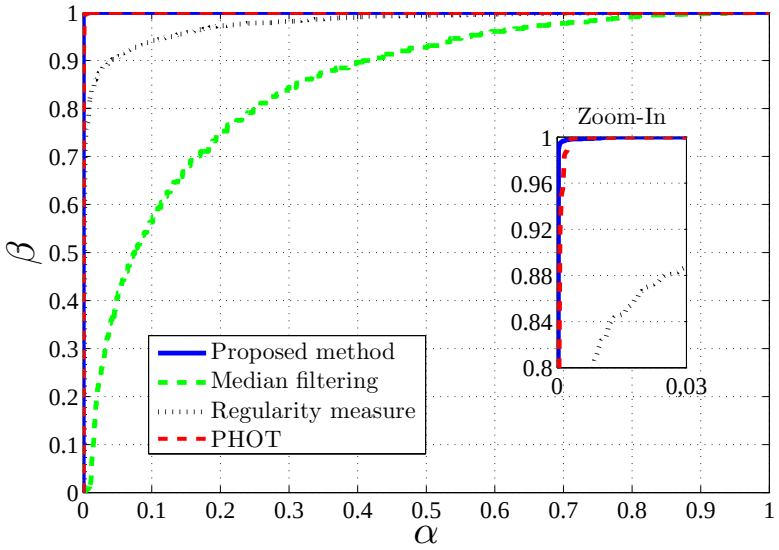


Figure 4.16: ROC curves of the four detection methods

Finally, the proposed adaptive method is compared to the detection method presented in [160] which consists of four subsystems: sensing, detection, classification and post-processing. Only the detection step is for interest in this study. It is based on a foreground extraction step using a median filter, and then a multi-zone detection technique, where each image is divided into multiple overlapped squared areas that undergo a thresholding procedure. The most crucial parameter that achieves an efficient defect detection performance is the median filter size. After many tests on defective and defect-free images, it is set to 13×13.

Figure 4.16 presents the corresponding ROC curves of the four mentioned detection methods.

The proposed adaptive method outperforms the detection method based on the regularity measure [159] and the one based on the median filtering [160]. However, compared to the method based on PHOT [158], the ROC curves seem to be indistinguishable. For a better look, a zoomed-in portion of the plot is illustrated on the same figure, with a false-alarm probability ranging between 0 and 0.03, and a detection probability between 0.8 and 1. This subplot shows that the proposed adaptive method performs slightly better than the one based on PHOT. Still, this comparison cannot be considered as conclusive.

Therefore, to present a more consistent comparison, it is wished to analyze the effect of the defect radius on the performances of the proposed adaptive method and the one based on PHOT. To this purpose, a similar Monte-Carlo simulation as the one seen in the first part of the experiments is performed, to represent the detection probability value, function of the defect radius, with a defect intensity of 150, and a false-alarm probability of 0.01. The degree of the polynomial used for the proposed adaptive model is $d_y = 5$ and $d_x = 2$ and the number of Principal

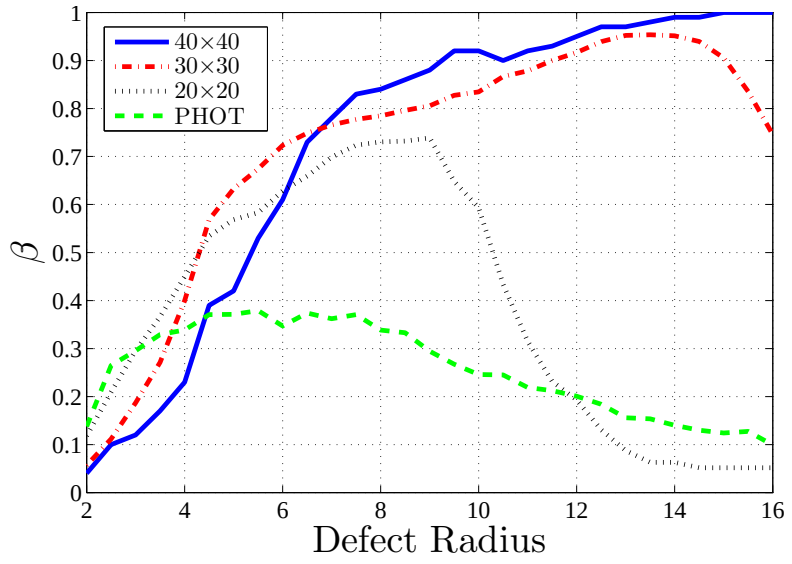


Figure 4.17: Detection probability function of the defect radius with a fixed defect intensity = 150

Components added is $\ell = 3$. Additionally, another goal of this simulation is to show the effect of the block size used in the proposed adaptive model on the detection performance. Hence, three different block sizes have been considered in this simulation.

Figure 4.17 illustrates the simulation results.

The green plot, which corresponds to the method based on PHOT, starts off with the highest detection probability for very small defects with a radius less than 3. It reaches its peak for defects with a radius ranging between 5 and 7. Then the performance starts to decline with the size of the defect. That can be explained in the fact that its detection phase is based on a global method, the PHOT, which must be applied on the whole image at once in order to properly remove the image regularity. Hence, a small defect is seen as an abrupt irregularity while a larger defect becomes slowly considered as a part of the image regularity, and thus partially removed.

As for the proposed adaptive method, it can be noticed that the performance highly depends on the considered block size. For small defects, using a smaller block size leads to a higher detection probability. For a defect radius ranging between 2 and 4, the proposed adaptive model with a block size of 20×20 is the best choice. For a defect radius ranging between 4 and 7, a block size of 30×30 is more appropriate. And for larger defects, a larger block size of 40×40 becomes the superior choice. That is due to the fact that our method considers that the block is a background image containing a defect, thus the defect occupies a smaller space in the block than the background itself. For a specific block size, the performance of the proposed method continues to rise, with the slope becoming smaller as the

defect is becoming larger, till it reaches a certain point when the defect occupies the totality of the block at which the performance drastically declines. This sudden decline in the performance can be observed in Figure 4.17 for a block size of 20×20 at a defect radius of 9, and for a block size of 30×30 at a defect radius of 14. Therefore, depending on the inspected surface, and on the range on the size of potential defects, one can choose the appropriate block size that gives the best performance.

Note that in chapter 3, the number of pixels for the smallest defect intended to be detected on the wheel surface was set to a value of 8 pixels in a single direction, i.e. a defect radius of 4 pixels. Therefore, since all the potential defects will have a radius above 4 pixels, it can be concluded from the comparison in Figure 4.17 that the proposed adaptive method outperforms the method based on PHOT.

4.5.6 Real defects

In this section, it is wished to present some results by applying the proposed adaptive model on wheel images containing real defects located on the surface of the wheel.

For each example image with defects, the data preparation procedure presented in Section 4.4.2 has been applied to obtain the unfolded image of the specific zone of the wheel that contains the defect. Then, the proposed adaptive model has been used to model the unfolded image background and create a model image that ideally does not contain any trace of the defect. Finally, the results are illustrated by a residual image, resulting from the subtraction of the model image from the unfolded image.

Choosing the appropriate parameters for the proposed adaptive model is vital to achieve the best detection results. Each zone of the wheel has its unique features and characteristics, according to which the choice of the proposed adaptive model parameters highly depends. These features and characteristics are not only related to the background homogeneity or pattern, but also to the defect size and shape which dependent much on the specific zone on which the defect is located.

The example images used in this experiment represent numerous types of wheels, with various designs and dimensions, containing defects that are located on various zones of the wheel, as the galbe zone, the rim zone and the ventilation zone. Table 4.1 lists the proposed adaptive model parameters for each zone of the wheel. The parameters described are the block width w and the block height h, the degrees of the polynomial d_x and d_y, and the number of added Principal Components ℓ.
A detailed, but rather general, discussion on the choice of these parameters has been done in Section 4.2.2. In the following, a more in depth explanation of the parameters choice for each zone will be provided, based on the zone texture and

Table 4.1: Proposed Adaptive Model Parameters

Zone	w	h	d_x	d_y	ℓ
Galbe Zone	50	25	2	5	3
Rim Zone	40	15	2	5	5
Ventilation Zone	30	20	2	5	3

potential defect size.

First, it is obvious that, due to the circularity of wheels, in all the different zones pixels share similar values along the horizontal direction, after unfolding process. This characteristic justifies the choice of d_x to a relatively low value of 2 for all the different zones. Along the vertical direction, the texture is much more complex with multiple light reflection artifacts, thus the necessity of higher polynomial degree $d_y = 5$ and the assistance of the Principal Components $\ell = 3$. The Rim zone presents a special case as it contains sharper edges along the rows which explains the use of a higher number of Principal Components ($\ell = 5$) to better model the background.

Regarding the block size, the choice of its width w and height h is mainly related to the defect size, although its height must be always maintained to a low value to help model the complex background along the rows. Briefly speaking, the galbe zone is the largest zone of the wheel, and defects located in this area are usually bigger in size, thus the choice of $w = 50$ and $h = 25$. Defects located on the rim zone, which is the boundary zone of the wheel, are usually medium sized scratches caused by the mishandling of the product, hence the choice of $w = 40$ and a lower value of $h = 15$ due to the special case mentioned above. Finally, defects located on the ventilation zone are usually of small size as they are trapped between the ventilation holes, which explains the choice of $w = 30$ and $h = 20$.

Following this discussion, it is possible to understand the variability of these parameters for the different zones of the wheel. However, to properly determine the exact values of the model parameters for each zone, it has to be performed by simulations. To this purpose, multiple simulations, with different values of model parameters, have been performed to ensure the best detection efficiency. For each simulation, a different combination of model parameters has been used, with the goal to maximize the detection power β for a fixed value of false alarm α. These simulations were performed in two steps; The first step was defining the size of the block w and h according to the potential defects sizes that could be present on each zone of the wheel. Then, in the next step, multiple combinations of the model parameters d_x, d_y and ℓ were tested to determine the ones that maximize the detection power β for a fixed value of false alarm α with the pre-defined block

size.

Figure 4.18 presents, for a small set of example images, its corresponding unfolded image, model image and residual image, resulting from the use of the proposed adaptive model with the proper parameters corresponding to the zone on which the defect is located.

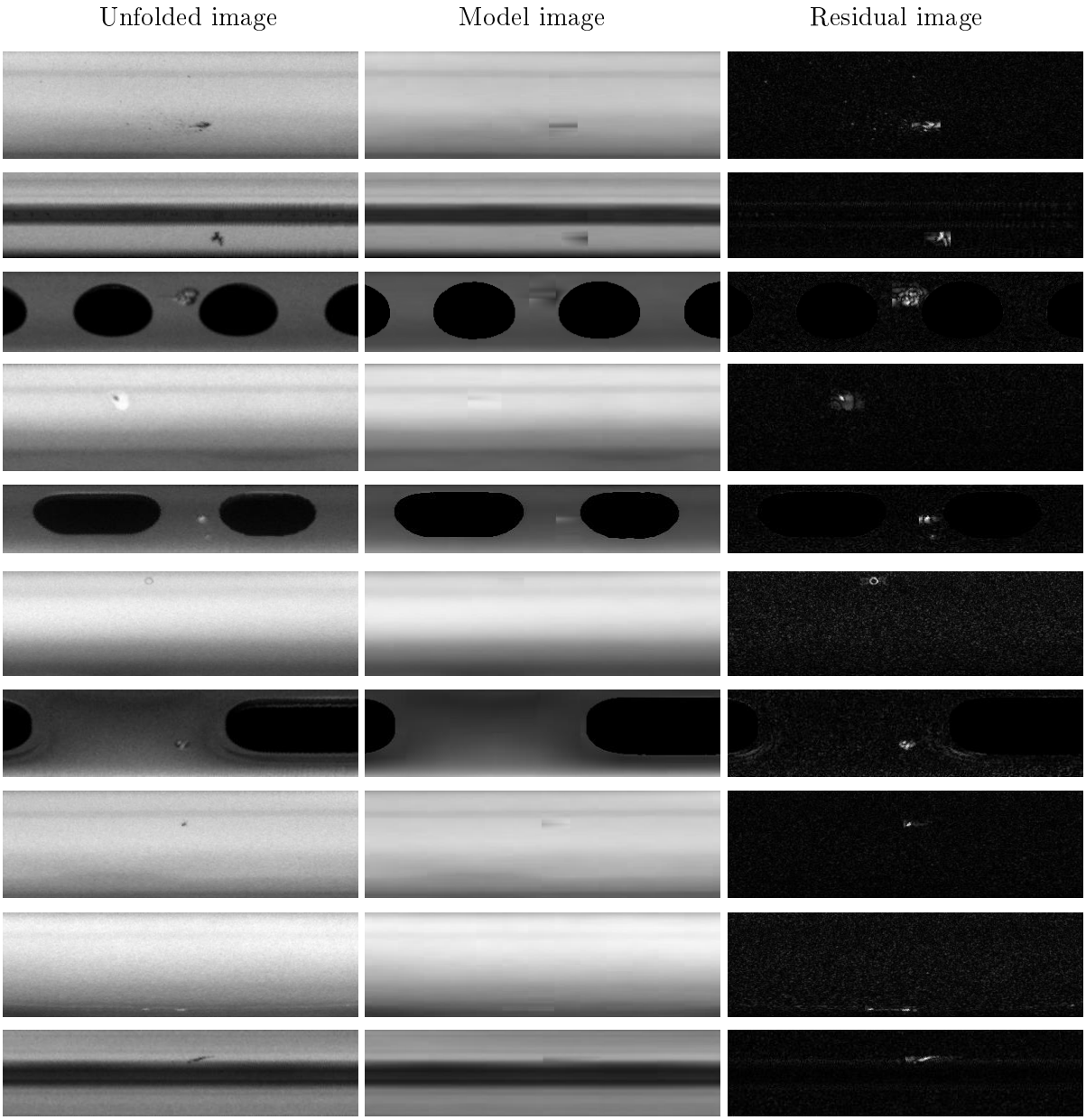


Figure 4.18: Images of wheels with various defects along with their corresponding model image and residual image

4.6 Conclusion

This chapter proposes a novel method for fully automatic anomaly detection on objects inspected using an imaging system. In order to address the inspection of a wide range of objects and to allow the detection of any anomaly, an original adaptive linear parametric model is proposed; The great flexibility of this adaptive model offers highest accuracy for a wide range of complex surfaces while preserving detection of small defects. In addition, because the proposed original model remains linear it allows the application of the hypothesis testing theory to design a test whose statistical performances are analytically known. Another important novelty of this method is that it takes into account the specific heteroscedastic noise of imaging systems. Indeed, in such systems, the noise level depends on the pixels' intensity which should be carefully taken into account for providing the proposed test with statistical properties.

The proposed detection method is then applied for the wheel surface inspection problem studied in this work. Due to the nature of the wheels, the different elements are analyzed separately. Numerical results on a large set of real images show both the accuracy of the proposed adaptive model and the sharpness of the ensuing statistical test. Furthermore, experimental results on wheel images containing real defects located on the surface of the wheel show the relevance in practice of the proposed adaptive model. The influence of the model parameters will be studied on a wide range of wheels to find the most accurate ones and the variance of the estimations will be taken into account to be able to establish with accuracy the performance of the proposed test.

Chapter 5

Wheels coating monitoring method

This chapter is inspired by our paper: "Karim Tout, Florent Retraint, Rémi Cogranne, Non-stationary process monitoring for change-point detection with known accuracy: Application to wheels coating inspection, in IEEE Access, 2018"

After addressing the problem of local defect detection in chapter 4, the remaining inspection task revolves around the detection of global defects. Such defects affect the entire surface of the wheel in a uniform manner. In this case, it is difficult to judge the wheel by itself as the defect is only visible if compared to previous manufactured wheels. Global defects are mainly related to the topcoat layer. This is the last layer of coating applied over the surface of the wheel and it maintains the aesthetic appearance characteristics of the wheel surface. Any increase or decrease in the amount of paint used for this layer will lead to an excess or shortage of color on the wheel surface, thus generating a global defect. Therefore, it is wished to develop a method to monitor the wheels' topcoat intensity over time, in order to detect an abrupt change that corresponds to a sudden lack of paint.

In recent years, the change-point detection topic has been receiving increasing attention in various domains. It addresses the problem of detecting the point or multiple points at which a "significant change" occurs in a time series. These points are referred to as change points. The change-point detection process must be able to distinguish between a "significant change" indicating an abnormal event, and an "insignificant change" due to noise and that indicates a predicted or a normal behavior of data. Distinguishing change-points from spurious noise is very important in order to keep a false alarm rate. However surprisingly, sequential methods are hardly provided with established, or bounded, false-alarm probability and power functions.

In general, change-point detection methods can be classified into "posteriori" and "sequential" methods. The choice of the appropriate class of methods depends heavily on the application.

"Posteriori" methods, also referred to as offline or retrospective methods, are considered in many applications, such as climate change study [161], biological applications [162, 163], econometric applications [164], and utility change in social media [165], to cite few topics. Such methods can only be applied after all

the data, or observations, are received. Then, the objective is to detect all the change-points available in the data, along with estimating their locations. In applications for which these types of methods are used, the goal is usually to analyze time series and not to take immediate action after detecting the change points.

On the opposite, many other applications analyze data in real time with the goal to take an immediate response as soon as a change in the data is detected, as it can reveal a system failure which must be handled. In such cases, a real-time data acquisition and analysis processes are required in order to raise an alarm as soon as a change-point is detected. Such problems fall within the scope of "Sequential" methods, also referred to as online or real-time methods, in which it is assumed that the data are received sequentially, and that until a change-point is detected the process is allowed to continue. Contrariwise, when the data change, typically revealing a failure or a change in the underlying process, the aim is to detect the change-point with a minimal delay, in order to take the relevant actions, while also preserving a low false-alarm.
Obviously, minimizing the detection delay and the false-alarm rate are contradictory goals. Sequential methods have been especially attracting attention from the industrial world, in which the term control chart is widely used, for quality control applications [166–168]. Industries have been pushing to produce higher quality and innovative products, which requires more and more manufacturing processes, while on the opposite, they are also required to reduce costs and production time. Hence, early fault detection for these industries is crucial to minimize downtime, reduce the product losses, and thus reduce manufacturing costs.

"Sequential" change-point detection methods can be further categorized into "parametric" and "non-parametric" methods.

On the one hand, "non-parametric" or data-driven methods have the advantage not to require any assumptions or any model on the data. They are based on statistical methods, especially supervised or non-supervised learning, to build detection rules based on large set of observations. Such decision rules are then applied to new data. While not requiring a model to describe the observations, those methods may, however, be limited, typically when the manufacturing process can largely change, and they are hardly provided with known statistical performances.

On the other hand, "parametric" methods are used when a sufficient information on the monitoring process is available such that a statistical model of the observations can be designed. In other words, this approach requires that some distributional knowledge of the data is available and employed into the detection scheme. A common limitation of such methods is that they rely on pre-specified parametric models that are based on a priori information about the form of the data distribution, which is not always available.

The studied process in this work is the variation of the paint quantity on the

inspected wheels surface. This process is shown to be non-stationary in the mean with a constant variance. Therefore, it is proposed to design a parametric sequential method with the goal to monitor a non-stationary process in real time in order to detect an abrupt change in its mean. This abrupt change results from a spray gun nozzle getting partially blocked, which will be translated into a sudden change in the paint intensity caused by the lack of paint on the wheel. In an industrial situation, it is required to detect the change within a given maximal detection delay (number of observations after the change) and it is wished to control the false-alarm probability over a fixed run length. In this operational context, a two fixed-length windows sequential method (2FLW-SEQ) based on the well-known CUSUM procedure is proposed for which the statistical performances are bounded.

The main contributions of the proposed method are the following:

1. A two fixed-length windows sequential method (2FLW-SEQ) is proposed for monitoring a non-stationary process in real time. The first window is considered to deal with the non-stationarity of the process, while the second window is the one used for the sequential detection procedure.

2. The proposed sequential procedure operates under the non-classical criteria of minimizing the worst-case probability of missed detection under the constraint of a maximal detection delay, while controlling the false alarm probability for a given number of observations.

3. A statistical study of the proposed method is established that allows to lower bound the detection power as a function of the maximal allowed detection delay, and enables to upper bound the false alarm probability for a given number of observations.

4. The proposed context enables the user to prescribe a maximal detection delay and a false alarm probability for a given number of observations, and can know which change amplitudes can be detected with guaranteed minimal probability.

5.1 Change-point detection problem statement

This section formally states the usual problem of abrupt change-point detection and recalls the well-known CUSUM method before highlighting the main particularity of the problem addressed in this work.

5.1.1 CUSUM procedure

The sequential change-point detection problem can be formulated as follows. Let us consider $\{x_n\}_{n\geq 1}$ a sequence of independent and identically distributed ($i.i.d$) observations that are acquired sequentially. At the beginning, the sequence is considered in a normal state, and the observations follow a probability distribution $\mathcal{P}_{\theta_0}$. Then,

at an unknown point $v \geq 0$ (the change-point), the sequence reaches an abnormal state, in which the observations follow a different probability distribution $\mathcal{P}_{\theta_1}$. The problem formulation can be rewritten as follows:

$$x_n \sim \begin{cases} \mathcal{P}_{\theta_0} & if \quad 1 \leq n \leq v, \\ \mathcal{P}_{\theta_1} & if \quad n \geq v+1, \end{cases} \tag{5.1}$$

The sequential change-point detection consists in detecting the change-point v as soon as it occurs, while at the same time preserving a low false alarm rate.

For the online continuous inspection, for each new observation received, a decision rule is computed to test between the two following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \{\theta = \theta_0\}, \\ \mathcal{H}_1 : \{\theta = \theta_1\}, \end{cases} \tag{5.2}$$

As long as the test (also called the stopping rule) fails to reject $\mathcal{H}_0$, the data acquisition continues.

When the observations x_i are statistically independent, a usual approach to decide between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ is to use the cumulative sum (CUSUM) procedure which can be defined, for observations up to N as follows [169]:

$$\delta_N = \begin{cases} 0 & if \quad S_1^N = \max(S_1^{N-1} + s_N - \lambda; 0) < \tau, \\ 1 & if \quad S_1^N = \max(S_1^{N-1} + s_N - \lambda; 0) \geq \tau, \end{cases} \tag{5.3}$$

where λ is a constant that avoids spurious false-alarms, τ is a conveniently predefined threshold and, for initialization, $S_1^0 = 0$. Though the decision statistics s_N and the constant λ were not defined in [169], the logarithm of the well-known likelihood ratio is commonly used:

$$s_n = \log\left(\frac{p_{\theta_1}(x_n)}{p_{\theta_0}(x_n)}\right), \tag{5.4}$$

where p_{θ_0} and p_{θ_1} are the probability density functions (PDF) under hypotheses associated with distributions $\mathcal{P}_{\theta_0}$ and $\mathcal{P}_{\theta_1}$ respectively, which are assumed to be known, and the constant λ is usually the average of the expected values $\lambda = 1/2\,(\mathbb{E}_{\mathcal{H}_0}[s] + \mathbb{E}_{\mathcal{H}_1}[s])$.

5.1.2 Difficulties of non-stationarity and criterion of optimality

The studied process in this work, which is the wheels coating variation, is non-stationary in the mean. As a consequence, the problem of detecting abrupt change in an i.i.d random sequence is not relevant anymore because (1) the hypotheses are composite, that is they are characterized by a set of possible parameters Θ_0 and Θ_1 and (2) for observation x_n the PDFs p_{θ_0} and p_{θ_1} are unknown.

In fact, when monitoring a non-stationarity process whose distribution parameters may "naturally" change over time, the change-point detection problem as stated in (5.1)–(5.2) is no longer relevant.

Indeed, since under the hypothesis $\mathcal{H}_0$ the distribution parameter θ_0 may change within the set Θ_0, the hypotheses are defined by:

$$\begin{cases} \mathcal{H}_0 : \{\theta \in \Theta_0\}, \\ \mathcal{H}_1 : \{\theta \in \Theta_1\}, \end{cases} \tag{5.5}$$

and one should instead consider the following sequential test problem:

$$x_n \sim \begin{cases} F_{\theta_{0,n}}\,,\ \theta_{0,n} \in \Theta_0 & if \quad 1 \leq n \leq v, \\ F_{\theta_{1,n}}\,,\ \theta_{1,n} \in \Theta_1 & if \quad n \geq v+1, \end{cases} \tag{5.6}$$

The main issue to tackle those scientific difficulties is to have an accurate model of Θ_0 and Θ_1; in other words, to be able to model with enough accuracy the set of "regular" changes in the process from the abrupt changes that reveals a malfunctioning.

Regarding the scientific difficulties, when the distribution parameters $\theta_{0,n}$ and $\theta_{1,n}$ are unknown, in such a context the likelihood ratio (5.4) cannot be calculated for a given observation x_n. A usual solution that is adopted in this work is to use a generalized likelihood ratio that consists in substituting unknown parameters $\theta_{0,n}$ and $\theta_{1,n}$ by their estimations using the maximum likelihood estimation.

The second main challenge, in our industrial context, is the introduction of an unusual criterion of optimality. Indeed the CUSUM has been shown to be asymptotically optimal with respect to the criterion that consists in minimizing the average worst case detection, see [170–172] for details on the so-called Lorden's criterion and CUSUM optimality.

However, a minimal average delay is not equivalent to a maximal detection accuracy for a given detection delay. Focusing on a practical industrial context, the proposed method aims at maximizing the probability of change-point detection for a fixed maximal delay; this is justified for cost-reduction purposes as the change point corresponds in practice to a malfunction in a production process.

5.2 Proposed change-point detection method

As discussed in section 5.1.2, it is wished to design a change point detection method in the case of a non-stationary process with a constraint on the maximal detection delay. This section first presents how to deal with the process non-stationarity that represents a nuisance parameter; then the novel two fixed-length window sequential method (2FLW-SEQ) is presented that fits with the constraint on the detection

delay, and rejects online the nuisance parameter generated by the process non-stationarity.

5.2.1 Process modeling

Let us consider a sliding window of size L. After the first L observations, for each new received data x_N, the window slides by one point to contain the observations from x_{N-L+1} to x_N. Let $Y_N = (x_{N-L+1}, \ldots, x_{N-1}, x_N)^T$ denotes this window after the reception of observation x_N. The vector Y_N is modeled with the following normal distribution:

$$\mathbf{Y}_N \sim \mathcal{N}(\mu_N, \sigma^2 \mathbf{I}_L), \tag{5.7}$$

where μ_N is the expectation in this window, $\mathbf{I}_L$ is the identity matrix of size L, and σ^2 is the variance which is assumed constant for all windows $Y_N, \forall N \geq L$.

A linear parametric model is proposed to represent the expectation μ_N. It essentially consists in representing all the observations in the window Y_N as a weighted sum of q basis vectors that represent the columns of a matrix $\mathbf{H}$ of size $L \times q$. The weight of this sum represents the vector of q parameters $\mathbf{d}_N$.
Hence, the expectation μ_N can be written as:

$$\mu_N = \mathbf{H}\mathbf{d}_N. \tag{5.8}$$

The model of $\mathbf{H}$ is based on the following algebraic polynomial:

$$h(x) = \sum_{j=0}^{q-1} d_j x^j, \tag{5.9}$$

with $q-1$ the degree of the algebraic polynomial.

The use of a linear parametric model in statistical testing theory has been widely exploited [142, 173], especially the case of a polynomial model [110, 135, 174]. One can also note that we have used such approach of polynomial image modeling for the defect detection method on the wheel surface detailed in the previous chapter.
However, here it is used in a simplistic manner within a sequential detection method to remove the possible slight "natural" intensity changes that are not abnormal and should thus be removed.

It follows from Eqs. (5.7) and (5.8) that in the absence of any anomaly, the vector of observations Y_N is modeled by:

$$\mathbf{Y}_N \sim \mathcal{N}(\mathbf{H}\mathbf{d}_N, \boldsymbol{\sigma}^2 \mathbf{I}_L). \tag{5.10}$$

On the opposite, when a defect happens in the process, a change occurs in the mean value which will affect all the observations after the change-point. Consequently, when the change occurs, the observations Y_N can be modeled as:

$$\mathbf{Y}_N \sim \mathcal{N}(\mathbf{H}\mathbf{d}_N + \Delta\mathbf{K}_M, \boldsymbol{\sigma}^2 \mathbf{I}_L), \tag{5.11}$$

where the sudden shift in the mean value is described by the vector K_M, of size L, containing $L-M$ zeros before the change occurs and minus ones M times after, and the constant $\Delta > 0$ represents the amplitude of the change. Here, M is the number of maximal acceptable observations with defects. For example, the change vector $K_1 = (0, 0, ..., 0, -1)$ describes a change that only affects the last observation in the window of size L.

It is important to note that the "acceptable" variation of mean value, modeled by $\mathbf{H}\mathbf{d}_N$, is a nuisance parameter as it is of no use for the considered detection problem. To deal with this nuisance parameter, it is proposed to use the maximum likelihood (ML) estimation method to perform a rejection of this nuisance parameter as follows:

$$\mathbf{r}_N = \frac{1}{\sigma}\mathbf{W}\mathbf{Y}_N. \tag{5.12}$$

Here $\mathbf{W}$ is the orthogonal projection of size $L-q \times L$, onto the null space of $\mathbf{H}$, whose columns correspond to the eigenvectors of the matrix $\mathbf{I}_L - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}$ associated with eigenvalues equal to 1. The vector $\mathbf{r}_N$ represents the projection of the observations onto the null space of $\mathbf{H}$.

5.2.2 2FLW-SEQ procedure

Among others, the matrix $\mathbf{W}$ has the following useful properties: $\mathbf{W}\mathbf{W}^T = \mathbf{I}_{L-q}$ and $\mathbf{W}\mathbf{H} = 0$; it thus follows from Eqs. (5.10)-(5.12), that the residuals $\mathbf{r}_N$ can be modeled under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ by the following statistical distribution:

$$\begin{cases} \mathcal{H}_0 : \{\mathbf{r}_N \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{L-q})\} \\ \mathcal{H}_1 : \left\{\mathbf{r}_N \sim \mathcal{N}\left(\frac{\Delta}{\sigma}\boldsymbol{\theta}_M, \mathbf{I}_{L-q}\right)\right\}, \end{cases} \tag{5.13}$$

where $\boldsymbol{\theta}_M$ represents the shift of expectation, due to the process failure, projected onto the null space of $\mathbf{H}$: $\boldsymbol{\theta}_M = \mathbf{W}\mathbf{K}_M$.

From the definition of the hypotheses in Eq. (5.13), after the rejection of the nuisance parameter $\mathbf{H}\mathbf{d}_N$, it is obvious that the considered detection problem essentially consists in the detection of a specific signal in noise.

Due to the industrial aspect of our operational context, it is proposed to use a sequential method with a fixed window of length M which also corresponds to a pre-defined fixed maximal detection delay.

Similar approaches have been studied in the context of sequential detection in [175, 176]. They proposed to use the well-known match space detection which is given in our case by:

$$\delta_N = \begin{cases} 0 & if \ \ \widetilde{S}^N_{N-L+1} = \boldsymbol{\theta}_M^T\mathbf{r}_N < \tau \\ 1 & if \ \ \widetilde{S}^N_{N-L+1} = \boldsymbol{\theta}_M^T\mathbf{r}_N \geq \tau. \end{cases} \tag{5.14}$$

From Eq. (5.13) it is straightforward to establish the statistical distribution of results $\widetilde{S}_{N-L+1}^{N}$ of the proposed 2FLW-SEQ:

$$\begin{cases} \mathcal{H}_0 : \left\{ \widetilde{S}_{N-L+1}^{N} \sim \mathcal{N}(\mathbf{0}, \|\boldsymbol{\theta}_M\|_2^2) \right\} \\ \mathcal{H}_1 : \left\{ \widetilde{S}_{N-L+1}^{N} \sim \mathcal{N}\left(\frac{\Delta}{\sigma} \|\boldsymbol{\theta}_M\|_2^2, \|\boldsymbol{\theta}_M\|_2^2 \right) \right\}. \end{cases} \tag{5.15}$$

which can be normalized, for the sake of clarity, as follows

$$\begin{cases} \mathcal{H}_0 : \left\{ \dfrac{\widetilde{S}_{N-L+1}^{N}}{\|\boldsymbol{\theta}_M\|_2} \sim \mathcal{N}(\mathbf{0}, 1) \right\} \\ \mathcal{H}_1 : \left\{ \dfrac{\widetilde{S}_{N-L+1}^{N}}{\|\boldsymbol{\theta}_M\|_2} \sim \mathcal{N}(\dfrac{\Delta}{\sigma} \|\boldsymbol{\theta}_M\|_2, 1) \right\}. \end{cases} \tag{5.16}$$

It is important to note that the choice of the first window size L and the polynomial degrees $q-1$ is crucial and essentially depends on the observations.

First, L must be much greater than M in order to avoid any significant impact of the abrupt change on the estimate of the linear model parameters $\mathbf{d}_N$. On the opposite, L must remain reasonably small such that the linear model will well model the observations' expectation and to ensure that the residuals $\mathbf{r}_N$ follow a standard normal distribution under $\mathcal{H}_0$.

As for q, it is the opposite scenario. Indeed, high polynomial degrees may lead to the shift being eliminated with the projection (5.12), and thus removed from the residuals. On the other hand, very small polynomial degrees may not be sufficient to properly model the process, and thus putting parts of the healthy observations among the residuals, and probably losing the standard normal distribution under $\mathcal{H}_0$.

5.3 2FLW-SEQ performances

In this section, the statistical properties of the proposed sequential method are studied in terms of probability of false alarm for a given run-length (number of observations under $\mathcal{H}_0$) and probability of change-point detection under the maximal delay constraint.

A sequential change-point detection procedure stops as soon as its decision rule δ_n becomes 1. Then, the stopping time T is defined as the smallest observation index n for which $\delta_n = 1$. A correct change detection consists in stopping the sequential procedure after the change has occurred, which means $T \geq v$ where v is the change point index. A false alarm is raised in case where $T < v$, i.e. the process has been stopped before the change occurred.

A usual criteria for a sequential procedure is to detect the change as soon as it occurs, thus minimizing the detection delay $T - v$. Many criteria have been used to investigate the optimality of change point detection algorithms concerning the detection delay, as the "mean delay", the "conditional mean delay", the "worst mean delay", etc.... [169, 170, 177, 178]. In that context, the CUSUM algorithm has been proven to be optimal in [170–172].

However, in the proposed detection scheme, the goal is to fix a detection delay after which the change detection is considered too late. In fact, minimizing the detection delay does not necessarily lead to a higher detection power, or to a small probability of missed detection. Therefore, the aim of the proposed sequential method is to minimize the worst-case probability of missed detection under constraint on the worst-case probability of false alarm for a given run length.

5.3.1 Minimizing the probability of missed detection

The stopping time of the classical CUSUM procedure is given by:

$$T_c = \inf_{n \geq 1} \{n : \max_{1 \leq k \leq n} \widetilde{S}_k^n \geq \tau\} \tag{5.17}$$

In this context, the CUSUM procedure takes into account all previous observations. However, for the proposed sequential method, after collecting the first L observations, the stopping time can be defined as:

$$T_{2FLW} = \inf_{n \geq L} \{n : \widetilde{S}_{n-L+1}^n \geq \tau\} \tag{5.18}$$

The probability of missed detection can be considered as the probability that the detection delay is higher than the acceptable one defined as M, knowing that the detection is made after the change has occurred with $T \geq v$.
Then, to the purpose of minimizing the probability of missed detection, the following criteria can be applied:

$$\mathbb{P}_{md}(M) = \sup_{v \geq L} \mathbb{P}(T - v + 1 > M \mid T \geq v) \tag{5.19}$$

where $\mathbb{P}_{md}(M)$ is the worst-case probability of missed detection. Minimizing this probability will lead to maximizing the detection probability denoted as $\beta(M) = 1 - \mathbb{P}_{md}(M)$.
Eq. (5.19) can be developed to:

$$\mathbb{P}_{md}(M) = \sup_{v \geq L} \frac{\mathbb{P}\left(\bigcap_{n=L}^{M+v-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right)}{\mathbb{P}\left(\bigcap_{n=L}^{v-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right)} \tag{5.20}$$

It is complicated to calculate the exact value of $\mathbb{P}_{md}(M)$, instead it is proposed to calculate an upper bound. It can be seen that:

$$\begin{aligned}&\mathbb{P}\left(\bigcap_{n=L}^{M+v-1}\left\{\widetilde{S}_{n-L+1}^{n}<\tau\right\}\right)\leq\\&\mathbb{P}\left(\left\{\bigcap_{n=L}^{v-1}\left\{\widetilde{S}_{n-L+1}^{n}<\tau\right\}\right\}\bigcap\left\{\widetilde{S}_{M+v-L}^{M+v-1}<\tau\right\}\right)\end{aligned}\tag{5.21}$$

Note that in Eq. (5.21), the two events have common observations of indexes $(M+v-L,...,v-1)$. In order to calculate the result $\widetilde{S}_{M+v-L}^{M+v-1}$, observations of indexes $(M+v-L,...,M+v-1)$ have been projected onto the null space of the model matrix $\mathbf{H}$, and then the resulting residuals have been multiplied by $\boldsymbol{\theta}_M$ which represents the shift of expectation, due to the process failure, projected onto the null space of $\mathbf{H}$. Because all the common observations are healthy observations, as they are acquired before the change v, their effect is neglected when multiplied by $\boldsymbol{\theta}_M$. Following that, the two events can be considered as independent, and Eq. (5.21) can be written as:

$$\begin{aligned}&\mathbb{P}\left(\bigcap_{n=L}^{M+v-1}\left\{\widetilde{S}_{n-L+1}^{n}<\tau\right\}\right)\leq\\&\mathbb{P}\left(\left\{\bigcap_{n=L}^{v-1}\left\{\widetilde{S}_{n-L+1}^{n}<\tau\right\}\right\}\right)\cdot\mathbb{P}\left(\left\{\widetilde{S}_{M+v-L}^{M+v-1}<\tau\right\}\right)\end{aligned}\tag{5.22}$$

Then, from Eq. (5.20), we get:

$$\mathbb{P}_{md}(M)\leq\mathbb{P}\left(\left\{\widetilde{S}_{M+v-L}^{M+v-1}<\tau\right\}\right)=\mathbb{P}_{\mathcal{H}_1}\left(\left\{\widetilde{S}_1^L<\tau\right\}\right)\tag{5.23}$$

where $\mathbb{P}_{\mathcal{H}_1}$ is the probability under $\mathcal{H}_1$. Based on (5.15), under $\mathcal{H}_1$, the result $\widetilde{S}_1^L$ is a Gaussian random variable with mean $\frac{\Delta}{\sigma}\|\boldsymbol{\theta}_M\|_2^2$ and variance $\|\boldsymbol{\theta}_M\|_2^2$.
As a result, the worst-case probability of missed detection can be upper bounded as:

$$\mathbb{P}_{md}(M)\leq\Phi\left(\frac{\tau}{\|\boldsymbol{\theta}_M\|_2}-\frac{\Delta}{\sigma}\|\boldsymbol{\theta}_M\|_2\right)\tag{5.24}$$

with Φ the standard normal cumulative distribution function.

Finally, the power function $\beta(M)$ of the proposed test (5.15), that is the probability of detecting a failure after at most M observations, is bounded by:

$$\beta(M)\geq 1-\Phi\left(\frac{\tau}{\|\boldsymbol{\theta}_M\|_2}-\frac{\Delta}{\sigma}\|\boldsymbol{\theta}_M\|_2\right).\tag{5.25}$$

In what follows, this lower bound will be referred to as $\widetilde{\beta}(M)$.

5.3.2 The worst-case probability of false alarm

For a given run length R and at a given time ℓ, the false alarm probability is given by:

$$\mathbb{P}_0(\ell \leq T \leq \ell + R) \tag{5.26}$$

Hence, the worst-case probability of false alarm for all $\ell \geq L$ can be defined as:

$$\mathbb{P}_{fa}(R) = \sup_{\ell \geq L} \mathbb{P}_0(\ell \leq T \leq \ell + R) \tag{5.27}$$

The calculation of the exact value of $\mathbb{P}_{fa}(R)$ is absurd, instead it is proposed to calculate an upper bound only. In this way, it is possible to guarantee a false alarm rate lower than that bound for all $\ell \geq L$.

The calculation will be done in two steps. First, the proof that the worst-case probability of false alarm is indeed the probability of false alarm at the starting point L. And then, the second step is to determine the upper bound.

First, let us start the proof of the following equality:

$$\mathbb{P}_{fa}(R) = \sup_{\ell \geq L} \mathbb{P}_0(\ell \leq T \leq \ell + R) = \mathbb{P}_0(L \leq T \leq L + R) \tag{5.28}$$

Let us denote $U_\ell = \mathbb{P}_0(T = \ell)$ for all $\ell \geq L$. For the first point L, it can be clearly seen that:

$$U_L = \mathbb{P}_0(\widetilde{S}_1^L \geq \tau) \tag{5.29}$$

and that:

$$\begin{aligned} U_{L+1} &= \mathbb{P}_0\left(\left\{\widetilde{S}_1^L < \tau\right\} \bigcap \{\widetilde{S}_2^{L+1} \geq \tau\}\right) \\ &\leq \mathbb{P}_0\left(\left\{\widetilde{S}_2^{L+1} \geq \tau\right\}\right) \end{aligned} \tag{5.30}$$

As all the observations of indexes $(1, ..., L+1)$ follow the same distribution under $\mathcal{H}_0$, then the inequality in Eq. (5.30) can be rewritten as:

$$U_{L+1} \leq \mathbb{P}_0\left(\left\{\widetilde{S}_1^L \geq \tau\right\}\right) = U_L \tag{5.31}$$

In a similar manner, for $\ell > L$, we can verify that:

$$U_\ell = \mathbb{P}_0\left(\bigcap_{n=L}^{\ell-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\} \bigcap \left\{\widetilde{S}_{\ell-L+1}^\ell \geq \tau\right\}\right) \tag{5.32}$$

and that:

$$\begin{aligned} U_{\ell+1} &= \mathbb{P}_0\left(\bigcap_{n=L}^{\ell} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\} \bigcap \left\{\widetilde{S}_{\ell-L+2}^{\ell+1} \geq \tau\right\}\right) \\ &\leq \mathbb{P}_0\left(\bigcap_{n=L+1}^{\ell} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\} \bigcap \left\{\widetilde{S}_{\ell-L+2}^{\ell+1} \geq \tau\right\}\right) \\ &\leq \mathbb{P}_0\left(\bigcap_{n=L}^{\ell-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\} \bigcap \left\{\widetilde{S}_{\ell-L+1}^\ell \geq \tau\right\}\right) = U_\ell \end{aligned} \tag{5.33}$$

Therefore, it is concluded that $(U_\ell)_{\ell \geq L}$ is a decreasing sequence.
Now let us define $V_\ell = \mathbb{P}_0(\ell \leq T \leq \ell + R)$ for all $\ell \geq L$. It can be seen that:

$$V_\ell = \sum_{n=\ell}^{\ell+R-1} \mathbb{P}_0(T = n) = \sum_{n=\ell}^{\ell+R-1} U_n \tag{5.34}$$

Then:

$$V_\ell - V_{\ell+1} = \sum_{n=\ell}^{\ell+R-1} U_n - \sum_{n=\ell+1}^{\ell+R} U_n = U_\ell - U_{\ell+R} \geq 0 \tag{5.35}$$

Consequently, $(V_\ell)_{\ell \geq L}$ is also a decreasing sequence. As a result, the equality in Eq. (5.28) is proven to be correct:

$$\sup_{\ell \geq L} V_\ell = V_L = \mathbb{P}_0(L \leq T \leq L + R) = \mathbb{P}_{fa}(R) \tag{5.36}$$

The second step consists in calculating the upper bound of V_L. From Eq. (5.29), U_L can be rewritten as:

$$U_L = 1 - \mathbb{P}_0(\widetilde{S}_1^L < \tau) \tag{5.37}$$

Similarly for all $\ell > L$, Eq. (5.32) can be rewritten as:

$$\begin{aligned} U_\ell &= \mathbb{P}_0\left(\bigcap_{n=L}^{\ell-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \\ &\quad - \mathbb{P}_0\left(\bigcap_{n=L}^{\ell-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\} \bigcap \left\{\widetilde{S}_{\ell-L+1}^\ell < \tau\right\}\right) \\ &= \mathbb{P}_0\left(\bigcap_{n=L}^{\ell-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \\ &\quad - \mathbb{P}_0\left(\bigcap_{n=L}^{\ell} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \end{aligned} \tag{5.38}$$

It follows from Eqs. (5.37), (5.38), and (5.34), that the worst-case probability of false detection V_L is:

$$V_L = 1 - \mathbb{P}_0\left(\bigcap_{n=L}^{L+R-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \tag{5.39}$$

For any two positive integers $n \neq n'$, it is proven that the covariance of the two Gaussian variables $\widetilde{S}_{n-L+1}^n$ and $\widetilde{S}_{n'-L+1}^{n'}$ is non-negative $cov\left(\widetilde{S}_{n-L+1}^n, \widetilde{S}_{n'-L+1}^{n'}\right) \geq 0$. As a consequence, one can immediately get:

$$\mathbb{P}_0\left(\bigcap_{n=L}^{L+R-1} \left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \geq \prod_{n=L}^{L+R-1} \mathbb{P}_0\left(\left\{\widetilde{S}_{n-L+1}^n < \tau\right\}\right) \tag{5.40}$$

Thus, V_L is upper bounded by:

$$V_L \leq 1 - \prod_{n=L}^{L+R-1} \mathbb{P}_0 \left(\left\{ \widetilde{S}_{n-L+1}^n < \tau \right\} \right) \tag{5.41}$$

Finally, based on (5.15), under $\mathcal{H}_0$, the results $\widetilde{S}_{n-L+1}^n \; \forall n \geq L$ are Gaussian random variables with zero mean and variance $\|\boldsymbol{\theta}_M\|_2^2$. As a result, the probability of having a false alarm $\alpha(R)$ after R observations is bounded by:

$$\alpha(R) \leq 1 - \Phi \left(\frac{\tau}{\|\boldsymbol{\theta}_M\|_2} \right)^R, \tag{5.42}$$

In what follows, this upper bound will be referred to as $\widetilde{\alpha}(R)$.

Equations (5.25) and (5.42) emphasize the main advantages of the proposed approach.

First, the statistical performance of the proposed test is bounded. The false alarm probability $\alpha(R)$ is upper bounded which will enable to calculate a detection threshold τ using a pre-defined false alarm rate knowing that the application is guaranteed not to exceed. On the other hand, the detection power $\beta(M)$ of the test is lower bounded which will allow to guarantee, for a pre-defined false alarm rate, a minimal detection power that the application will not decrease bellow.

Second, the false alarm probability $\alpha(R)$ only depends on the prescribed run-length R and the maximal acceptable detection delay M.

Last, the power function (5.25) shows that the accuracy of the proposed method essentially depends on the "change-to-noise ratio" Δ/σ, along with the maximal acceptable detection delay M.

5.4 Paint coating intensity variation

5.4.1 Painting process

Wheel paint has two purposes; to protect the underlying metal from the harsh environment to which it is exposed, and most importantly to improve the look of the wheel. Modern wheel coating methods consist of five main steps, starting with the pre-treatment which removes and cleans excess metal to form a smooth surface structure, and ending with the topcoats which provide the surface characteristics including color, appearance, gloss, smoothness, and weather resistance [179]. The focus in this chapter is on the topcoats as they are the only visible layer.

Wheel topcoats are usually composed of several layers of paint coatings, with a precise thickness, spread on the whole surface of the wheel one after another [180]. They are generally applied in the form of liquid or powder using spray atomizers, also called spray gun nozzles [180]. The appearance (color, gloss, texture, etc...) of a coated surface greatly affects the perception on the product quality.

In fact, every wheel manufacturer has a list of client requirements that defines every detail concerning the final product, including a "top-coat requirements" list that contains specifications about the color, the gloss level, and many other aspects of the topcoats. Given this set of specifications, any significant deviation from what is standard or normal to the product is considered an anomaly that has to be correctly detected.

However, it is important to note that in this context a defective process will not only affect one wheel, but all of the following products. Therefore, a fast and accurate detection of any anomaly, as soon as it appears, is necessary in order to reduce the number of defective products, thus reducing the loss. Moreover, the deviation that is considered as anomalous is hardly distinguishable from other normal deviations, and hence may remain unnoticed by visual inspection.

All those points lead us to the necessity of an automatic inspection system that monitors the variations of the topcoat intensity, and signals the change-point with minimal delay time. The detection process has to be fast and sufficiently efficient in order to distinguish between a normal state and the anomalous state.

Technically speaking, many factors influence the quality of the coating, thus its appearance, such as temperature, paint viscosity, solvents, etc. [179, 180] .... Specifically for liquid painting, as time goes by, the viscosity of the paint in the paint bath decreases (the paint becomes more pasty) since the solvents are evaporating over time. This process may be faster or slower depending on the neighboring temperature [179, 180]. To rectify the effects of this process, usually the operators tend to increase the paint/airflow on the spray gun nozzle. These variations in the topcoats remain in the acceptable zone in accordance with the technical requirements.

This paper focuses on a usual problem, that is when the spray gun nozzle partially clogs, or gets blocked, which will be translated in a sudden change in the intensity of the topcoats.

5.4.2 Topcoat monitoring

To the purpose of monitoring the variation of paint coating intensity on produced wheels, it is wished to consider a block containing s pixels in the image of the wheel, over which the mean value of all pixels is computed. The considered window maintains the same size and position on the surface of the wheel for all images. Then, for one image of a wheel, let $Z = \{z_w\}_{w=1}^{s}$ denote the window containing s pixels and $m = s^{-1} \sum_{w=1}^{s} z_w$ the mean value of pixels' intensity. Note that the behavior of the observations is independent from the window position on the surface of the wheel.

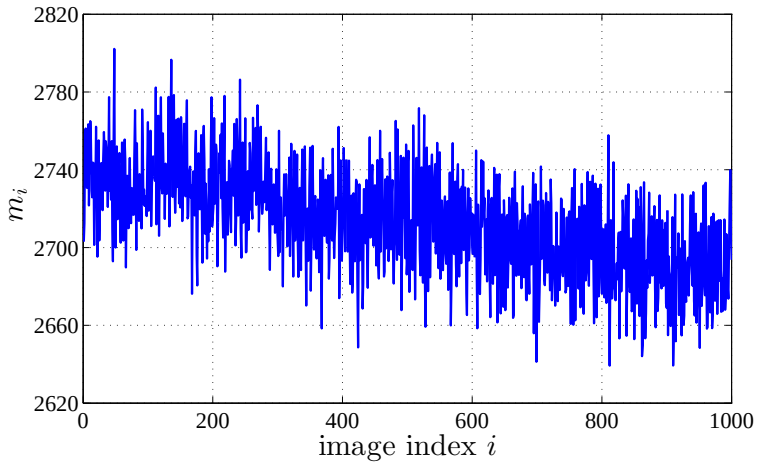


Figure 5.1: A typical example of variation of wheel images mean value

The variation of the mean value m from a wheel image to another describes the variation of the topcoat intensity. Indeed, the mean value is a sufficient parameter to detect coating failure as the change in pixel values that it causes affects the whole surface of the wheel. Figure 5.1 shows an example of series of mean pixels' value m_i for 1 000 images of consecutive wheels without change points, with i the image index. The observed variation in the mean values is considered to be normal, and it is due to the reasons detailed previously. It is shown that the mean value of observations m_i evolves smoothly.

Note that, the window Z has always the same position from the center of the wheel, but not exactly the same position on the wheel image. In fact, the wheels are not perfectly centered under the imaging system, which means that from an image to another, the position of the wheel may differ by few pixels. In addition, the illumination system is not ideal, meaning that the distribution of light over the whole wheel surface is not perfectly uniform, hence some locations on the wheel are slightly more or less illuminated than others.

Therefore, it is concluded that the variance of the variable m_i is only related to the imaging system which is not modified during the acquisition, thus it remains constant for all observations, whether before or after the change. Based on these factors, and based on the behavior of the variable m_i observed in Figure 5.1, the process can be considered as a non-stationary process in the mean, with a constant variance over all the observations.

5.5 Experiments and results

In this section, five types of results are presented. First, the proper choice of the first window length L and the degree of the polynomial $q-1$ is discussed with simulation results. The second experiment aims to study the effect of the second window length M on the performances of the proposed test. In the third part,

it is wished to examine the efficiency of the bounds calculated in subsection 5.3 and to study the detectability of the proposed test function of the abrupt change amplitude, given a set of requirements. Next, the fourth experiment is a study of a real case scenario with a real change point in the observations. Finally, a performance comparison is conducted to highlight the advantages of modeling the observations and examine the difference in the detection criteria between our approach and the CUSUM method.

To conduct these experiments, a data base of 500 000 successive healthy images has been acquired. The acquired images are made of 2046×2046 pixels of 12 bits depth. The procedure described in section 5.4 has been applied to obtain the observations m_i with $i = \{1, 2, ..., 500\,000\}$. The observed standard deviation, related to the imaging system, is $\sigma = 22$. As supposed in section 5.4, this parameter is assumed to be constant during the monitoring process. However, the variance can be changed with the acquisition conditions, for instance, with the illumination intensity.

To deal with the problem of imaging acquisition system drift, the variance is periodically computed (typically at the beginning of each week).

5.5.1 Parameters tuning

First, let us start by discussing the choice of the first window length L, and the degree of the polynomial $q-1$. In fact, as mentioned in subsection 5.2.2, the choice of parameters L and q has an important role, on the one hand, to increase the detection performances of the test, and, on the other hand, to correctly model the paint coat intensity process as a Gaussian process. Hence, multiple Monte-Carlo simulations, with different values of L and q, have been performed to correctly tune these parameters to ensure the best performances.

Two important factors are directly affected by the change in these parameters, those are the detection power $\beta(M)$ and the accuracy of the standard Gaussian distribution model for residuals' r_N distribution under $\mathcal{H}_0$. This accuracy can be expressed using the Hellinger distance H_D between the empirical residuals r_N and the theoretical standard Gaussian distribution.

Table 5.1 contains the calculated values of $\beta(M)$ and H_D for values of L ranging from 50 to 1 000, and values of $q-1$ ranging from 2 to 10. These results have been obtained for a maximal detection delay $M = 5$, over a run length of $R = 5\,000$ which represents about half of a day production, and for a pre-defined value of false alarm rate $\alpha(R) = 10^{-2}$.

It can be observed from Table 5.1 that for a certain polynomial degree, increasing L will lead to a better detection performance as $\beta(M)$ increases, however, the Hellinger distance H_D increases alongside which indicates a decrease in accuracy. For large values of L, as $L = 1\,000$, small values of polynomial degree are not even sufficient to correctly represent the observations under $\mathcal{H}_0$, which can be seen by

Table 5.1: The empirical detection power $\beta(M)$ and the Hellinger distance H_D for different values of L and q.

		First window size L											
		50		100		150		200		250		1 000	
		$\beta(M)$	H_D	$\beta(M)$	H_D	$\beta(M)$	H_D	$\beta(M)$	H_D	$\beta(M)$	H_D	$\beta(M)$	H_D
Degrees of the polynomial $q-1$	2	0.4005	0.0405	0.8079	0.0426	0.8481	0.0433	0.9351	0.0443	0.8741	0.0451	0.5387	0.0564
	3	0.0559	0.0387	0.5135	0.0417	0.7549	0.0423	0.8020	0.0438	0.8775	0.0444	0.7049	0.0563
	4	0.0455	0.0373	0.2813	0.0409	0.4185	0.0415	0.6594	0.0429	0.8052	0.0438	0.8416	0.0463
	5	0.0441	0.0354	0.0601	0.0407	0.2378	0.0410	0.3493	0.0422	0.5582	0.0434	0.8706	0.0381
	10	0.0040	0.0264	0.0217	0.0364	0.0523	0.0395	0.0425	0.0414	0.0524	0.0420	0.6525	0.0350

the increase in H_D and the decrease in $\beta(M)$. Thus it is necessary to increase the polynomial degree just to correctly model the observations.

On the other side, for a certain value of L, increasing q will lead to an increase in the accuracy, in favor of a decreasing performance. For large values of q, as $q = 10$, the accuracy increases significantly, however, the test performance is low. To choose the optimal values of L and q, it is important to have the maximal detection power $\beta(M)$ alongside a sufficient accuracy so that the empirical performance matches at best the theoretical performance study. For the values $L = 200$ and $q - 1 = 2$, we have the best detection power $\beta(M) = 0.9351$.

Then, to better understand the relation between the Hellinger distance and the accuracy, figure 5.2 represents a comparison between the theoretical standard Gaussian cumulative distribution function (cdf) and the empirical cumulative distribution

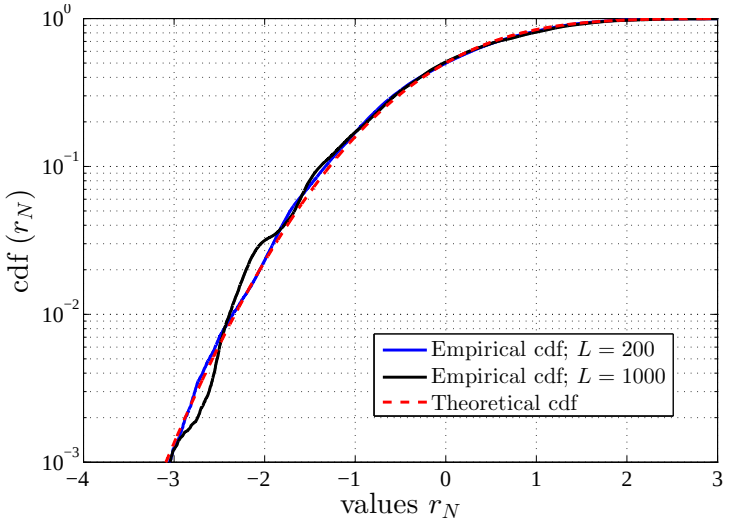


Figure 5.2: Empirical and theoretical cumulative distributions of the normalized residuals r_n with two different values of the first window size L with polynomial degrees $q - 1 = 2$.

functions of the residuals r_N for $L = 200$ and $L = 1\,000$, and with $q - 1 = 2$.

It can be seen that for $L = 200$, the empirical distribution is accurate enough compared to the theoretical distribution, and that moving from $H_D = 0.0443$ for $L = 200$ to the highest distance values $H_D = 0.0564$ for $L = 1\,000$ will only have a small effect on the accuracy of the distribution under $\mathcal{H}_0$. Therefore, the choice of the parameters can be made on the basis of the highest detection power $\beta(M)$ for a Hellinger distance H_D lower than a certain value after which the accuracy is considered no longer acceptable.

As a result, the correct choice of the parameters in our application is $L = 200$ and $q - 1 = 2$, which will be considered in all following experiments.

5.5.2 Maximal detection delay and detection performances

Secondly, it is proposed to study the effect of the second window length M on the detection performances. The same data base has been used to perform a Monte-Carlo simulation, for which a simulated shift of amplitude $\Delta = 60$ has been superimposed on some of the observations.

Figure 5.3 represents the empirical false alarm probability $\alpha(R)$ and detection power $\beta(M)$ over a run length $R = 5\,000$ for 3 different values of the maximal allowed detection delay $M = \{1, 3, 5\}$, as a function of the decision threshold τ. It can be observed that when M increases, $\|\boldsymbol{\theta}_M\|_2$ increases, which affects both the false alarm rate $\alpha(R)$ and the detection power $\beta(M)$, as seen in (5.15). However, the increase rate of $\beta(M)$ is larger than the one of $\alpha(R)$. Hence, the shift between the detection power and the false alarm probability becomes larger which implies a better detection performance, but at a larger delay M.

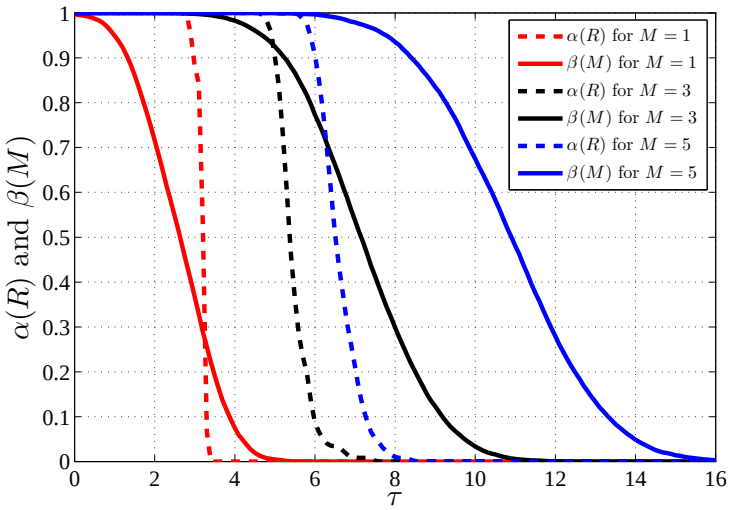


Figure 5.3: Empirical false alarm probability $\alpha(R)$ and detection power $\beta(M)$ over a run length $R = 5000$ for 3 different values of M, plotted as a function of the decision threshold τ.

As a result, it can be seen that the choice of M essentially depends on the application requirements. Depending on the application, this test allows to either increase the detection performance at a cost of a larger detection delay, or decrease the detection delay at a cost of a lower detection performance.

5.5.3 Efficiency of the bounds and detectability performance

In the third part of the experiments, and because one of the main assets of the proposed change-point detection method is that its performance properties are bounded, it is wished first to examine the efficiency of the bounds calculated for $\alpha(R)$ and $\beta(M)$.

Figure 5.4 shows the empirical false alarm probability $\alpha(R)$ and its theoretical upper bound $\widetilde{\alpha}(R)$ for three different values of the run length $R = \{50, 500, 5\,000\}$ as a function of the detection threshold τ. The maximal delay for detection is set to $M = 5$.
It can be observed that the upper bound is accurate and relatively tight. However, as the run-length increases, one can notice that the upper bound is gradually losing its accuracy for smaller values of false alarm. At $\alpha(R) = 10^{-2}$, the distance between the empirical threshold and the theoretical one obtained by the upper bound is 0.7 for $R = 50$, but it increases to 1.2 for $R = 5\,000$. This is due mainly to the fact that the observations are not totally independent.

In fact, the calculation of the upper bound of the false alarm probability is based on the inequality in equation (5.40) which is greatly affected by the independence of the observations. When R increases, the number of events in equation (5.40) increases, resulting in an increase in the difference between the probability of their

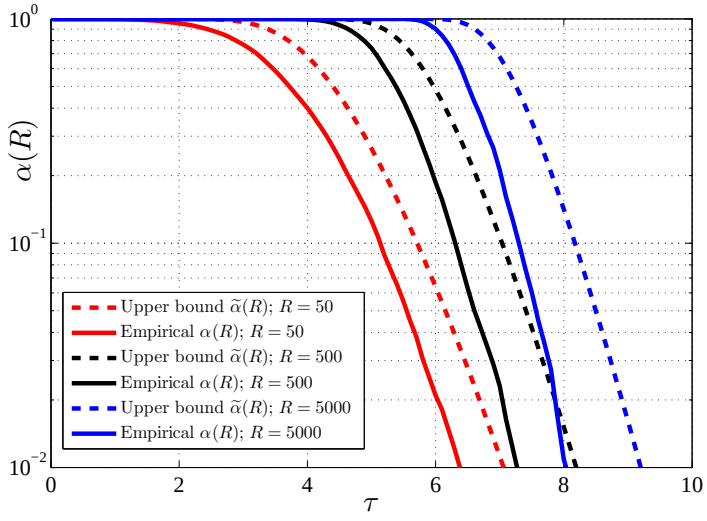


Figure 5.4: Empirical and theoretical false alarm probability $\alpha(R)$ over three different values of run length R, plotted as a function of the decision threshold τ.

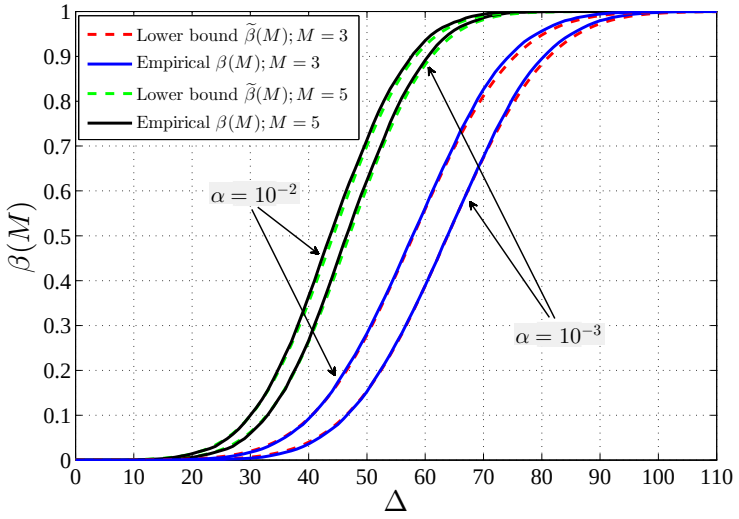


Figure 5.5: Empirical and theoretical detection power $\beta(M)$ for 2 different values of M and 2 different values of false alarm rate $\alpha(R)$, plotted as a function of the change amplitude Δ.

intersection (the first term (5.40)) and the product of their individual probabilities (the second term in (5.40)). As a result, the sharpness of the upper bound for the false alarm probability decreases.
In addition, a second factor can be the fact that the data base used to perform these experiments is rather small to be generally accurate in the empirical results for large values of run length as $R = 5\,000$.

Then, in order to test the detectability of the proposed test and the sharpness of the detection power lower bound, figure 5.5 presents the empirical detection probability $\beta(M)$ and its theoretical lower bound $\widetilde{\beta}(M)$ for two different values of $M = \{3, 5\}$ and two different false alarm rates $\alpha(R) = \{10^{-2}, 10^{-3}\}$ over a run length $R = 5\,000$, as a function of the change amplitude Δ.
First, it can be seen that the theoretical lower bound is precise and really tight for the different parameter values. Second, for a fixed value of the false alarm rate, when M increases, the detection power $\beta(M)$ increases accordingly.
This result confirms the one obtained in the second experiment in Figure 5.3.

5.5.4 Real case scenario of change-point

Next, it is wished to exemplify the efficiency of the proposed 2FLW-SEQ sequential detection method on a real case scenario with a real change-point in the observations.

Figure 5.6 portrays a real case of observations when the spray gun nozzle got partially clogged. As a consequence, a sudden shift in the observations of amplitude $\Delta = 55$ can be seen at exactly the image index 2434. The blue plot represents the real observations, while the red plot represents the expectation values (5.8)

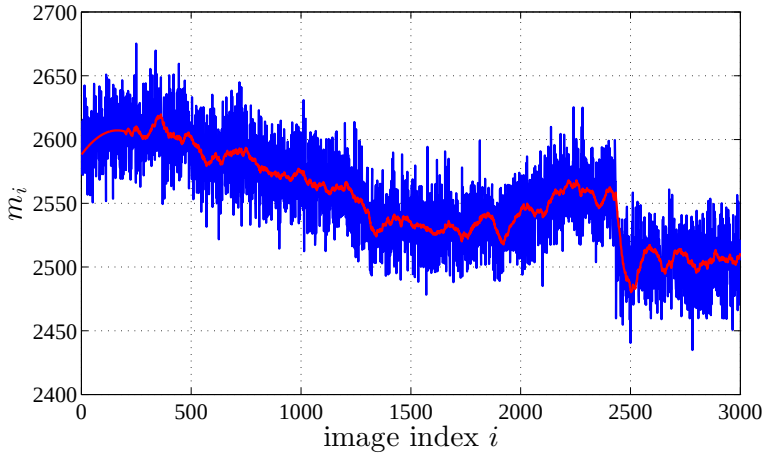


Figure 5.6: Real example of mean value variation with a change-point at index 2434.

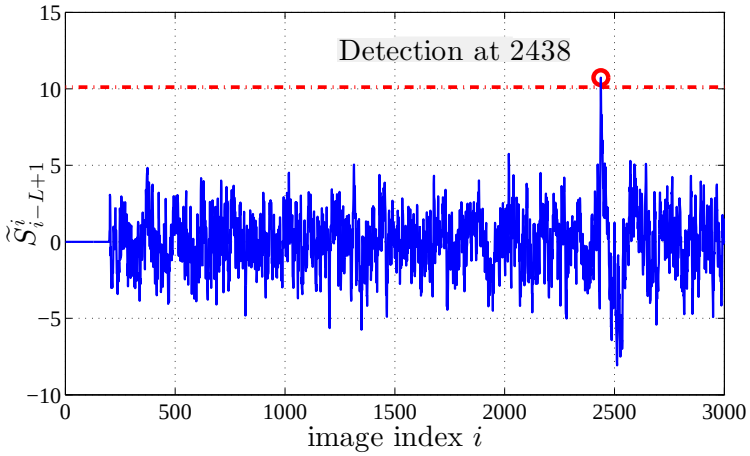


Figure 5.7: Result of the proposed 2FLW-SEQ detection method with $M = 5$.

estimated using the polynomial model over a window of size $L = 200$ and a degree of $q - 1 = 2$. Because it is aimed to be as close as possible from the real practical requirement that corresponds to the specific application of paint coat monitoring, the false alarm rate is set to $\alpha(R) = 10^{-3}$ over a run length $R = 5\,000$. This will result in a detection threshold of $\tau = 10.12$ for $M = 5$ and a threshold of $\tau = 8.2$ for $M = 3$.

Figure 5.7 illustrates the result of the proposed 2FLW-SEQ method with $M = 5$. It can be seen that the change point is detected at the index 2438 which means a delay of exactly 5 defective wheels.

Then, the same experiment has been performed for a maximal allowed delay of $M = 3$ where the detection power is much lower than the previous case of $M = 5$, as seen in Figure 5.5. Figure 5.8 illustrates the corresponding result where it can be seen that the change-point has been missed.

Note that, usually when the change is detected, the sequential process stops.

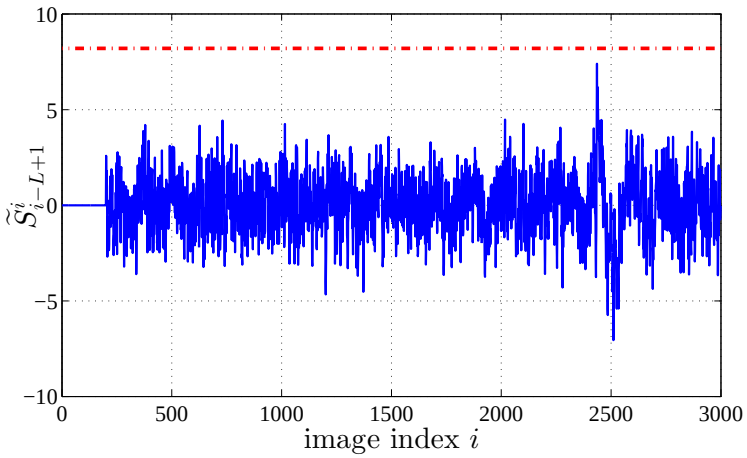


Figure 5.8: Result of the proposed 2FLW-SEQ detection method with $M = 3$.

However to better illustrate the results of the test, the sequential procedure was allowed to continue. It is shown in Figure 5.6 that after the change occurs, the observations return to a state similar to the one just before the change occurred. Then, just after the change, the sequential procedure will re-operate under the hypothesis $\mathcal{H}_0$, and the results $\widetilde{S}^i_{i-L+1}$ will return to have a Gaussian distribution with zero mean and a variance $\|\boldsymbol{\theta}_M\|_2^2$, as it can be seen in Figures 5.7 and 5.8.

5.5.5 Process modeling and detection criteria

Last, but not least, the first goal is to investigate the advantages of modeling the paint coat intensity process to deal with its non-stationarity. To this purpose, it is proposed to compare the performance of the original 2FLW-SEQ method presented in this chapter with a classical sequential detection method, more precisely the well-known CUSUM, in two different scenarios.

In the first scenario, the polynomial model is used to represent the expectation of the last L observations, while in the second scenario only the mean value of the last L observations is considered. In addition, the proposed 2FLW-SEQ is included in the comparison in order to show its efficiency. Note that when the polynomial model is used, the optimal parameters obtained from the first part of the experiments are considered, i.e. $L = 200$ and $q - 1 = 2$. However, for the CUSUM without a model, multiple simulations with different values of L have been conducted and lead to the choice of $L = 20$ which is the best in terms of detection power.

Figure 5.9 presents the empirical ROC curves for the proposed 2FLW-SEQ method and the CUSUM method with and without the polynomial model, computed over a run length $R = 5000$, with a maximal detection delay $M = 5$ and change amplitude $\Delta = 60$. It can be seen that using the polynomial model actually improves the performance of the CUSUM method. Figure 5.9 also shows that the proposed 2FLW-SEQ method outperforms the CUSUM method even when using

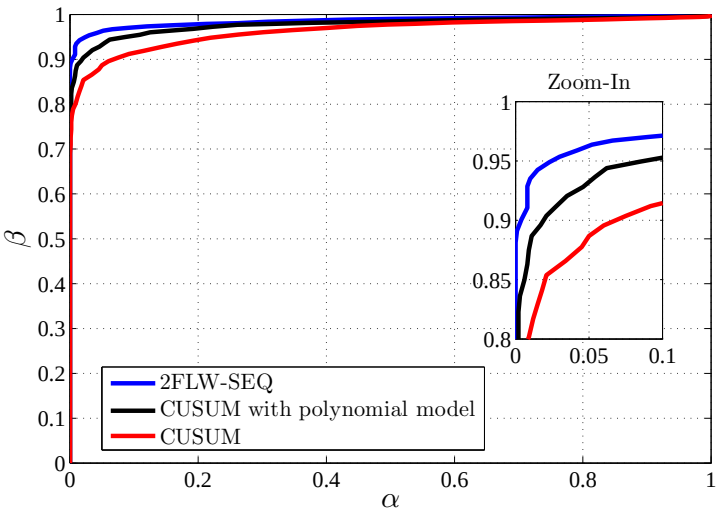


Figure 5.9: Empirical ROC curves for the proposed 2FLW-SEQ method and the CUSUM method with and without the polynomial model, computed over a run length $R = 5000$, with a maximal detection delay $M = 5$ and change amplitude $\Delta = 60$.

the polynomial model.

Indeed, the CUSUM method has proven many times to be optimal as mentioned in section 5.3, however, this optimality is related to the average detection delay. To better understand the difference in the detection criteria under which each of the proposed 2FLW-SEQ method and the CUSUM method operates, two sets of

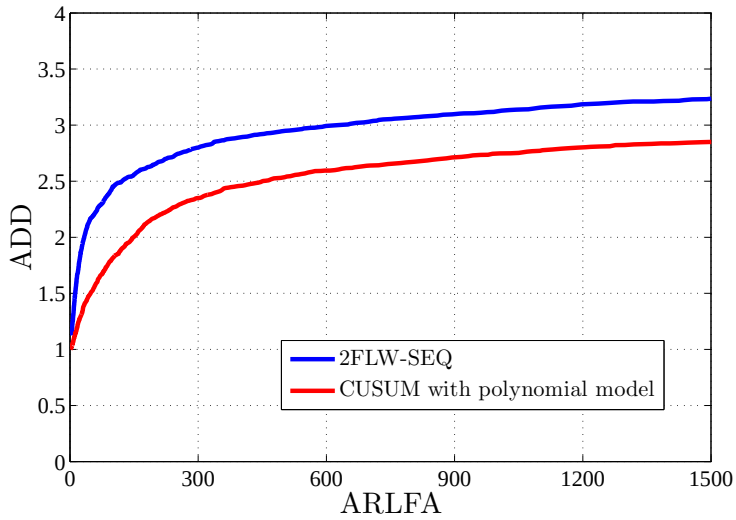


Figure 5.10: Average detection delay as a function of the average run length to false alarm for the proposed 2FLW-SEQ method and the CUSUM method with polynomial model with a maximal detection delay $M = 5$ and change amplitude $\Delta = 60$.

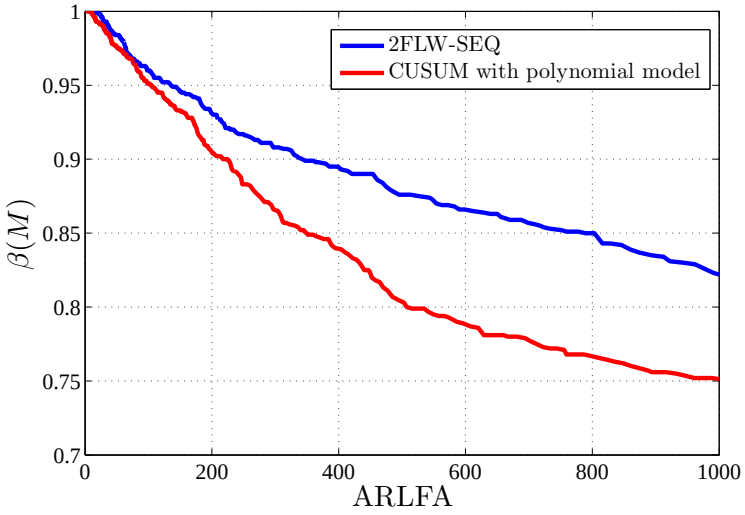


Figure 5.11: Empirical detection power $\beta(M)$ as a function of the average run length to false alarm for the proposed 2FLW-SEQ method and the CUSUM method with polynomial model with a maximal detection delay $M = 5$ and change amplitude $\Delta = 40$.

simulations are conducted.

Figure 5.10 represents the average detection delay (ADD) as a function of the average run length to false alarm (ARLFA) for the proposed 2FLW-SEQ method, with a maximal detection delay set to $M = 5$, and for the CUSUM method with polynomial model with $q - 1 = 2$ and change amplitude $\Delta = 60$. It can be seen that in this context, the CUSUM has proven to be optimal and, hence, outperforms the proposed 2FLW-SEQ method.

On the other hand, as noted in the section 5.3, the aim of the proposed 2FLW-SEQ method is to minimize the worst-case probability of missed detection under constraint on the worst-case probability of false alarm for a given run length. To highlight this criteria, figure 5.11 represents the empirical detection power $\beta(M)$ as a function of the ARLFA for the same sequential methods and with a change amplitude $\Delta = 40$. The smaller value of the change amplitude is considered to emphasize better the difference.

Obviously, figure 5.11 shows that, in this context, with the increasing values of the ARLFA, the proposed 2FLW-SEQ method outperforms the CUSUM method in terms of detection power. Indeed, it is well known that minimizing the average detection delay does not necessarily lead to a higher detection power under a given maximal delay, or to a small probability of missed detection.

5.6 Conclusion

This chapter addresses the problem of monitoring online a non-stationary process to detect abrupt changes in the process mean value. Two main challenges are addressed: First, the monitored process is non-stationary ; i.e. naturally changes over time and it is necessary to distinguish those"regular" process changes from abrupt changes resulting from potential failures. Second, this method aims at being applied for industrial processes where the performance of the detection method must be accurately controlled. A novel sequential method, based on two fixed-length windows, is proposed to detect abrupt changes with guaranteed accuracy while dealing with non-stationary process. The first window is used for estimating the non-stationary process parameters while the second window is used to execute the detection. A study on the performances of the proposed method provides analytical expressions of the test statistical properties. This allows to bound the false alarm probability for a given number of observations while maximizing the detection power as a function of a given detection delay.

For the studied process of wheels coating monitoring, the mean value of pixels from all wheel images are used to measure the coating intensity. Numerical results on a large set of images show the accuracy of the proposed model, the efficiency of the proposed detection method, and the sharpness of the statistical performances theoretically established.

Chapter 6

AVI system maintenance

Likewise all measurement systems, an AVI system essentially consists of hardware solutions that require maintenance. The aging of such components could reduce the performance of the system, and eventually cause a failure. For AVI systems, the illumination degradation over time is a major concern [181]. The light emission of light sources is based on chemical compounds that have characteristic aging behavior in their substantial structure [182]. Indeed, since light is the signal that is used by the camera sensor to generate the processing data, the illumination of the imaged scene has a major impact on the performance of the inspection system.

Usually, detection methods that are based on a reference model of the inspected surface are very sensitive to changes in illumination. In such methods, the reference model is derived from calibration data collected under certain imaging conditions, and thus can be regarded as "valid" only if the conditions, under which the calibration data were collected, do not change during normal use of the model. Alternatively, a major advantage of our detection method proposed in chapter 4 is that it does not use a general reference model that should adapt to changes in the imaging system, but rather creates a unique background model for each inspected surface derived from its image. Despite that, the detectability of anomalies still depends on the "anomaly-to-noise" ratio, which will also be impacted by the change in illumination.

It is then important to study the aging behavior of the light source to determine the moment at which the illumination level is no more sufficient for the system to execute its intended functions. However, depending on the kind of light source and on its working conditions, this aging is not always easy to model. An alternative approach proposed in this chapter is to study the impact of the degradation of illumination on the performance of the detection method. This will allow to estimate the illumination level under which the system is no longer reliable. Finally, a hardware solution is proposed to detect the moment when this illumination level is reached.

6.1 Illumination degradation

LEDs benefit from two major attributes that have contributed in the growing interest in their use; the potential for long life and the reduced energy consumption. Unlike conventional light sources, LED modules do not die instantly, instead

their light output slowly degrades over time [100]. This feature ensures a proper functioning of the AVI system for a long period of time, and helps avoid a sudden failure in illumination. However, even if the LED is technically operating and producing light, after a certain amount of time, the light produced by the LED might become insufficient for the intended application. This amount of time is usually referred to as the lifetime of the LED, rather than the time needed for its complete failure.

Many LED manufacturers provide an estimated degradation rate of their products. However, this is rather a vague information of how long the illumination device will work, as it is usually measured under optimal conditions. In reality, the degradation rate of a LED module, which in a way defines its lifetime, is heavily affected by various factors, most importantly its functioning temperature. In fact, the degradation rate of many light sources, including LEDs, is sped up by temperature [100], especially under long-term operation. Many efforts have been made to study the degradation rate of LEDs [183–187], taking into account some of the factors that contribute to this degradation. The term that has been adopted for such studies is "lumen maintenance", which is how the intensity of emitted light tends to diminish over time. The results are generally presented as predictions of the gradual output degradation as a function of time, for different case studies each with different set of conditions. The precision of these predictions is however doubtful as many factors are usually not taken into account during the study, either because of unknown failure modes and mechanisms, or lack of field data [183, 185]. Hence, basing the maintenance of the AVI system on such predictions in not efficient. More importantly, in almost all of these studies, the conditions under which the LED is operating are known and fully measured. This is usually not the case for industrial applications, where the contributing factors for LED degradation are unknown and even unpredictable. As an example, the lack of continuous cleaning of the LED modules may be an additional cause of light degradation, as the light emitted from the LEDs is also attenuated by the industrial dust that covers the LED modules.

Therefore, an alternative approach, proposed in this chapter, is to monitor the illumination from the AVI system itself and to signal the point at which the level of illumination is no further sufficient.

6.2 Effects on the defect detection method

Light is the signal that is used by the camera sensor to generate the data of the imaged scene. The sensor of the camera, whether it was a CCD sensor or a CMOS sensor, records the image by using photosensitive elements that convert light energy to electrical energy [188]. Each pixel has a tiny light cavity or "photosite" that collects photons and store those as an electrical signal. This signal is then

quantified as a digital value, which represents the pixel intensity. The number of photons collected by each photosite will then mainly define, among other factors, the corresponding pixel intensity. However, not all camera sensors respond in a similar manner to the incoming light. Each camera model has a specific response relating the scene radiance to the image intensity. This response is mapped into a function, known as the camera response function (CRF) [189].

Regardless of the camera model, the CRF is always a monotonically increasing function [189–191]. This characteristic indicates that when the scene radiance increases, the acquired image intensity will increase. Consequently, when the scene radiance decreases, the acquired image intensity will decrease. Hence, due to the light source output degradation, the intensity of the acquired images from the AVI system will monotonically decrease over time. This in turn will affect the performance of the defect detection method proposed in chapter 4.

Based on Eq. (4.16), after the rejection of the nuisance parameters, the norm of the normalized residuals $\mathbf{r}$ in one block of the image follows the distribution:

$$\|\mathbf{r}\|_2^2 \sim \begin{cases} \chi^2_\Upsilon(0) & \text{under } \mathcal{H}_0 \\ \chi^2_\Upsilon(\boldsymbol{\varrho}) & \text{under } \mathcal{H}_1, \end{cases} \tag{6.1}$$

with the non-central parameter ϱ under hypothesis $\mathcal{H}_1$ is given by :

$$\boldsymbol{\varrho} = \left\| \widetilde{\boldsymbol{\Sigma}}^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}} \boldsymbol{\theta} \right\|_2^2. \tag{6.2}$$

Then, based on the residuals $\mathbf{r}$ and their distribution (Eq. (6.1)), a UBCP test is designed as follows

$$\delta = \begin{cases} \mathcal{H}_0 & \text{if } \|\mathbf{r}\|_2^2 \leq \boldsymbol{\tau} \\ \mathcal{H}_1 & \text{if } \|\mathbf{r}\|_2^2 > \boldsymbol{\tau}, \end{cases} \tag{6.3}$$

where, in order to guarantee the false-alarm probability α_0, the decision threshold $\boldsymbol{\tau}$ is set as follows:

$$\boldsymbol{\tau} = F^{-1}_{\chi^2_\Upsilon}(1 - \alpha_0; 0) \tag{6.4}$$

Similarly the power function of the test is given by:

$$\beta(\delta, \boldsymbol{\theta}) = F_{\chi^2_\Upsilon}(\boldsymbol{\tau}, \boldsymbol{\varrho}). \tag{6.5}$$

First, from Eq. (6.4), it can be seen that the threshold $\boldsymbol{\tau}$ is independent from the pixel intensities $\boldsymbol{\mu}$. In fact, it only depends on the false-alarm probability α_0, and thus remains constant even if $\boldsymbol{\mu}$ decreased over time.

Then, from Eq. (6.5), and with a constant threshold $\boldsymbol{\tau}$, it can be seen that the detectability of the anomaly only depends on the non-centrality parameter $\boldsymbol{\varrho}$ (6.2). This non-centrality parameter is essentially the "anomaly-to-noise" ratio that will define how detectable the anomaly is, thus the performance of the detection method.

When the illumination decreases, the anomaly $\boldsymbol{\theta}$ also decreases as the whole image intensity will decrease. However, this is also true for the noise. Indeed, one main contribution of the proposed detection method is to consider the more realistic heteroscedastic model of the noise. This will allow to establish with highest precision the theoretical statistical properties of the ensuing test. This noise model represents the variance of the noise corrupting the image as a linear combination of the pixel's expectation $\boldsymbol{\mu}$ given by

$$\sigma^2 = a\,\boldsymbol{\mu} + b. \tag{6.6}$$

The heteroscedastic noise model parameters (a, b) only depend on acquisition parameters. Therefore, as the camera settings are not modified over time, these parameters remain constant even with the degradation of illumination.

Then, since both the anomaly and the noise decrease with the decrease in illumination, it is not obvious if the non-centrality parameter $\boldsymbol{\varrho}$ will eventually decrease or not.

To demonstrate that the non-centrality parameter $\boldsymbol{\varrho}$ will decrease with the degradation of illumination, let us denote $(\boldsymbol{\mu_1}, \boldsymbol{\theta_1}, \boldsymbol{\varrho_1})$ the background intensity, the anomaly intensity, and the non-centrality parameter, respectively, for a block of the image acquired from an imaged scene, and $(\boldsymbol{\mu_2}, \boldsymbol{\theta_2}, \boldsymbol{\varrho_2})$ the ones for the same block of the image acquired from the same scene but with a decrease in illumination (i.e. smaller light luminance). The block is of size w and h for width and height respectively. Due to the monotonic characteristic of the CRF, it is possible to deduce that $\boldsymbol{\mu_2} < \boldsymbol{\mu_1}$ and that $\boldsymbol{\theta_2} < \boldsymbol{\theta_1}$. Then, since the acquired images are in raw format, the CRF is assumed to be linear [192, 193]. This is indeed true regardless of the camera model, considering that all camera sensors are designed to produce electrical signals that are linearly related to the scene radiance. For most cameras, though, non-linearities are later purposely introduced in the camera's electronics to mimic the response of the human visual system, to account for non-linearities in display systems, or to create a variety of aesthetic effects. The most recognized non-linearity is the gamma correction [194, 195]. As a result, this linearity for raw images induces that both the background and the anomaly will respond in a similar manner to the decrease in illumination.
Hence, let us consider that $\boldsymbol{\mu_2} = K\boldsymbol{\mu_1}$ and that $\boldsymbol{\theta_2} = K\boldsymbol{\theta_1}$ where $K < 1$ is a constant that corresponds to the decrease ratio of pixel intensity (or light illumination under linear CRF). This equality, in turn, highlights the fact that for a same anomaly in the scene, the "anomaly-to-background" ratio will remain constant with the decrease in illumination.

Then, let us start the proof of the following inequality:

$$\boldsymbol{\varrho_2} < \boldsymbol{\varrho_1} \tag{6.7}$$

This inequality can be written as:

$$\left\| \widetilde{\boldsymbol{\Sigma}}_2^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}_2} \boldsymbol{\theta_2} \right\|_2^2 < \left\| \widetilde{\boldsymbol{\Sigma}}_1^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}_1} \boldsymbol{\theta_1} \right\|_2^2 \tag{6.8}$$

Considering that this inequality is true, and by replacing $\boldsymbol{\theta_2}$ by $K\boldsymbol{\theta_1}$ and $\boldsymbol{\mu_2}$ by $K\boldsymbol{\mu_1}$, it follows that:

$$\sum_{i=1}^{w\times h} \frac{K^2 \left(\mathbf{P}^{\perp}_{\mathbf{H}_2}\boldsymbol{\theta_1}\right)_i^2}{aK\boldsymbol{\mu_{1}}_i + b} < \sum_{i=1}^{w\times h} \frac{\left(\mathbf{P}^{\perp}_{\mathbf{H}_1}\boldsymbol{\theta_1}\right)_i^2}{a\boldsymbol{\mu_{1}}_i + b} \tag{6.9}$$

with V_i is the i-th element of the vector V. From Eq. (4.15), and after a brief calculation, it is possible to assume that the projection matrix in both cases will be the same $\mathbf{P}^{\perp}_{\mathbf{H}_2} = \mathbf{P}^{\perp}_{\mathbf{H}_1}$. Then, the inequality in Eq. (6.9) can be rewritten as:

$$\sum_{i=1}^{w\times h} \left(\mathbf{P}^{\perp}_{\mathbf{H}_1}\boldsymbol{\theta_1}\right)_i^2 \left(\frac{K^2}{aK\boldsymbol{\mu_{1}}_i + b} - \frac{1}{a\boldsymbol{\mu_{1}}_i + b}\right) < 0 \tag{6.10}$$

Indeed, the inequality in Eq. (6.10) is true only if:

$$K^2(a\boldsymbol{\mu_{1}}_i + b) - K(a\boldsymbol{\mu_{1}}_i) - b < 0 \tag{6.11}$$

Solving this quadratic inequality, we get that the inequality in Eq. (6.7) is true if:

$$\begin{cases} 0 < K < 1 & \text{for } b \geq 0 \\ \dfrac{-b}{a\boldsymbol{\mu_{1}}_i + b} < K < 1 & \text{for } b < 0 \text{ and } \boldsymbol{\mu_{1}}_i > \dfrac{-2b}{a} \end{cases} \tag{6.12}$$

In general, the parameter b of the heteroscedastic noise model is a positive value. However, in some digital imaging sensors, the collected charge is added to a base value, referred to as the pedestal value, to constitute an offset-from-zero of the output pixels [137]. This offset could then generate a negative value of the parameter b. Indeed, with $b < 0$, the smallest value that the pixel intensity $\boldsymbol{\mu}$ can have is $-b/a$, as it cannot have a negative value.

As we can see from Eq. (6.12), for $b < 0$, an additional condition is set on the pixel intensity values of the background $\boldsymbol{\mu_{1}}_i$. In this work, and for any inspection system in general, this condition on $\boldsymbol{\mu_{1}}_i$ is always ensured, since the usual pixel intensities of the inspected surface are much higher than the value $-2b\ /\ a$. This continues to be true even with the decrease in illumination as the detection performances will largely drop before reaching such small values of pixel intensity. Consequently, the variable K is also lower bounded, as the same condition applies for $\boldsymbol{\mu_2}$. This explains the lower term of the inequality in Eq. (6.12) for $b < 0$.

In conclusion, and with the above conditions on $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$, Eq. (6.12) indicates that if $K < 1$, the inequality in Eq. (6.7) is true. In other words, when the background intensity decreases, the non-centrality parameter will also decrease. Consequently, the statistical performance of the detection method will therefore decrease.

6.3 Detection of light threshold

The inspection performance is the most significant and most consistent factor to evaluate the working state of an AVI system. For many applications, this performance should not decrease under a certain limit, which is usually predefined according to the nature of the inspected products. Hence, the necessity to monitor the variation of the inspection performance over time, and signal the moment at which it is no longer sufficient. At this point, a maintenance operation is required to make the necessary adjustments to the system for it to regain its performance in the accepted zone.

As seen in section 6.2, the degradation of illumination will have a negative impact on the performance of the detection method proposed in chapter 4. Hence, in this section, it is proposed to model this impact in a way that will enable to determine the illumination level below which the detection performance is no further satisfactory. Then, to detect the moment when the AVI system reaches this illumination level, a simple hardware solution is proposed.

6.3.1 Non-centrality parameter and pixel intensity

Let us suppose the presence of an anomaly on the inspected surface. This anomaly has a known non-centrality parameter $\boldsymbol{\varrho_1}$ for a certain background intensity $\boldsymbol{\mu_1}$. Then, after an unknown change in the illumination that could be either an increase or a decrease, this same anomaly has a new non-centrality parameter $\boldsymbol{\varrho_2}$ that is known. Hence the goal is to determine the change ratio in illumination that has occurred. In other words, it is proposed to determine $\boldsymbol{\mu_2}$, knowing the values of $\boldsymbol{\mu_1}$, $\boldsymbol{\varrho_1}$, and $\boldsymbol{\varrho_2}$.

Therefore, let us consider that $\boldsymbol{\varrho_2} = R\boldsymbol{\varrho_1}$ where R is a constant that represents the change ratio in the non-centrality parameter. And let us suppose that $\boldsymbol{\mu_2} = K\boldsymbol{\mu_1}$ where K is unknown. Due to the linearity of the CRF, and for a same anomaly in the scene, the "anomaly-to-background" ratio will remain constant with the change in illumination. Hence, $\boldsymbol{\theta_2} = K\boldsymbol{\theta_1}$. Then, we get:

$$\left\|\widetilde{\boldsymbol{\Sigma}}_2^{-1/2}\mathbf{P}_{\mathbf{H}_2}^{\perp}\boldsymbol{\theta_2}\right\|_2^2 = R\left\|\widetilde{\boldsymbol{\Sigma}}_1^{-1/2}\mathbf{P}_{\mathbf{H}_1}^{\perp}\boldsymbol{\theta_1}\right\|_2^2 \tag{6.13}$$

By replacing $\boldsymbol{\theta_2}$ by $K\boldsymbol{\theta_1}$ and $\boldsymbol{\mu_2}$ by $K\boldsymbol{\mu_1}$, it follows that:

$$\sum_{i=1}^{w\times h}\frac{K^2\left(\mathbf{P}_{\mathbf{H}_2}^{\perp}\boldsymbol{\theta_1}\right)_i^2}{aK\boldsymbol{\mu_1}_i + b} = R\sum_{i=1}^{w\times h}\frac{\left(\mathbf{P}_{\mathbf{H}_1}^{\perp}\boldsymbol{\theta_1}\right)_i^2}{a\boldsymbol{\mu_1}_i + b} \tag{6.14}$$

with V_i is the i-th element of the vector V. Again, the projection matrix in both cases will be the same. Then, the equality in Eq. (6.14) can be rewritten as:

$$\sum_{i=1}^{w\times h}\left(\mathbf{P}_{\mathbf{H}_1}^{\perp}\boldsymbol{\theta_1}\right)_i^2\left(\frac{K^2}{aK\boldsymbol{\mu_1}_i + b} - \frac{R}{a\boldsymbol{\mu_1}_i + b}\right) = 0 \tag{6.15}$$

Indeed, the equality in Eq. (6.15) is true only if:

$$K^2(a\boldsymbol{\mu_{1}}_i + b) - K(Ra\boldsymbol{\mu_{1}}_i) - Rb = 0 \tag{6.16}$$

The only solution for the above quadratic equation in Eq. (6.16), which respects the conditions on the various parameters, is:

$$K = \frac{Ra\boldsymbol{\mu_{1}}_i + \sqrt{(Ra\boldsymbol{\mu_{1}}_i)^2 + 4Rb\,(a\boldsymbol{\mu_{1}}_i + b)}}{2\,(a\boldsymbol{\mu_{1}}_i + b)} \tag{6.17}$$

As a result, given the same anomaly on the inspected surface, Eq. (6.17) will allow to determine the change ratio in pixel intensity, function of the change ratio in the non-centrality parameter R along with the initial background intensity $\boldsymbol{\mu_1}$. Eq. (6.17) also verify that if $R = 1$, then $K = 1$.

6.3.2 Procedure to detect the lowest acceptable illumination level

Due to the degradation of illumination over time, the performance of the detection method proposed in chapter 4 will decrease, because it has been designed to match a prescribed false alarm probability, so called CFAR (constant false alarm rate) detector. This performance corresponds to the detection power $\boldsymbol{\beta}$ of the statistical test in Eq. (6.3), for a constant false alarm probability $\boldsymbol{\alpha_0}$. Let us denote $\boldsymbol{\beta_t}$ the value of the lowest acceptable detection power that the system is allowed to reach. This detection power represents the level of detectability of a particular anomaly present on the inspected surface. From Eq. (6.5), it can be seen that the detectability of the anomaly depends on its non-centrality parameter $\boldsymbol{\varrho}$ (6.2). The higher $\boldsymbol{\varrho}$ is, the better the performance of the test to detect the corresponding anomaly.
As the goal of the detection method is to detect all anomalies present on the inspected surface, it is then necessary to consider the smallest anomaly intended to be detected to define the value of $\boldsymbol{\beta_t}$. Note that the term "smallest" does not refer to the size nor to the shape of the anomaly, but rather means the anomaly that has the smallest non-centrality parameter $\boldsymbol{\varrho}$ for the same background intensity.
Then, the goal is to detect the luminance or pixel intensity value at which the proposed defect detection method will have a statistical performance that reaches this value of $\boldsymbol{\beta_t}$, in order to start a maintenance operation to regain the performance.

To this purpose, the proposed procedure can be divided into two main parts. The first part consists in determining the pixel intensity of the background below which the detection performance is no further satisfactory. Then, the second part consists in monitoring the AVI system to detect the moment at which this pixel intensity is reached. The complete procedure is illustrated in Figure 6.1.

The first step is for the user to define $\boldsymbol{\beta_t}$ for a constant false alarm probability $\boldsymbol{\alpha_0}$. Using Eq. (6.4), the detection threshold $\boldsymbol{\tau}$ of the test can be determined as it

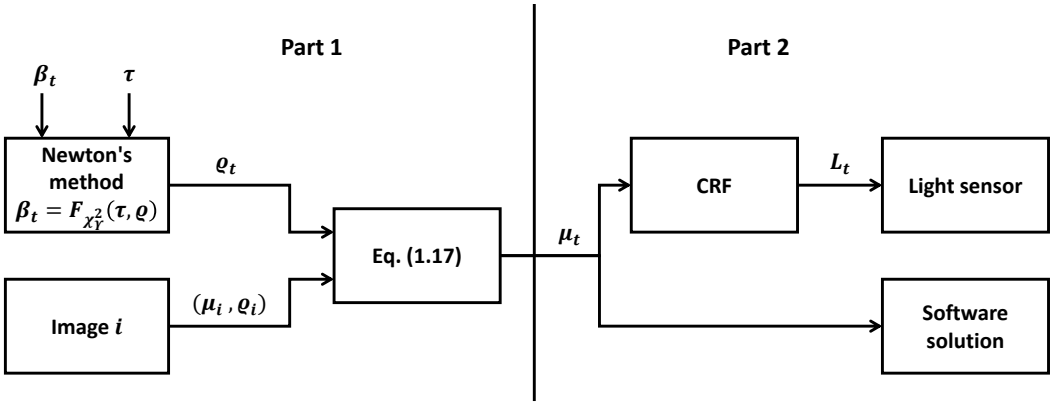


Figure 6.1: Flowchart of the proposed procedure to detect the lowest acceptable illumination level

only depends on the predefined false alarm probability $\boldsymbol{\alpha_0}$.
The next step is to calculate the non-centrality parameter $\boldsymbol{\varrho_t}$ that results in the value of $\boldsymbol{\beta_t}$, with the detection threshold $\boldsymbol{\tau}$. To calculate $\boldsymbol{\varrho_t}$, it is proposed to use Eq. (6.5), and to apply Newton's method [196] to solve the following equation:

$$\boldsymbol{\beta_t} = F_{\chi^2_\Upsilon}(\boldsymbol{\tau}, \boldsymbol{\varrho_t}) \tag{6.18}$$

The resulting value of $\boldsymbol{\varrho_t}$ represents the value of the non-centrality parameter caused by the presence of the smallest anomaly intended to be detected on the inspected surface, at the lowest allowed illumination level.

Indeed, during this procedure, it is supposed that the current illumination level of the system has not yet reached its lowest allowed value, but is rather higher. Thus, if we consider a block on a an image acquired at the current illumination level, its pixel intensity will be higher than the one of this same block if the image was acquired at the lowest allowed illumination level. In other words, if we denote $\boldsymbol{\mu_i}$ the pixel intensity of the background at the current illumination level, and $\boldsymbol{\mu_t}$ the pixel intensity of the background at the lowest allowed illumination level, we have $\boldsymbol{\mu_i} > \boldsymbol{\mu_t}$. Since $\boldsymbol{\mu_i}$ is known, the goal here is to calculate the decrease ratio in pixel intensity that will allow to determine $\boldsymbol{\mu_t}$.
To this purpose, it is proposed to use Eq. (6.2) to determine the value of the non-centrality parameter $\boldsymbol{\varrho_i}$ at the current image of index i, that corresponds to the same smallest anomaly for which the non-centrality parameter at the lowest allowed illumination level is $\boldsymbol{\varrho_t}$. Then, having $\boldsymbol{\mu_i}$ and $\boldsymbol{\varrho_i}$ for image index i, and $\boldsymbol{\varrho_t}$ at the lowest allowed illumination level, it is possible to determine the value of $\boldsymbol{\mu_t}$ using Eq. (6.17). Note that, the heteroscedastic noise model parameters (a, b) are known, and they only depend on acquisition parameters. Therefore, as the camera settings are not modified over time, these parameters remain constant even with the degradation of illumination.

Finally, the second part of the procedure consists in monitoring the AVI system to detect the moment at which this pixel intensity $\boldsymbol{\mu_t}$ is reached. This pixel

intensity is related to the scene radiance through the CRF. Hence, monitoring the status of the AVI system can either be done by monitoring the pixel intensity of the acquired images (software-based solution), or by monitoring the light intensity of the lighting system (hardware-based solution). In this work, it is proposed to adopt the hardware-based solution by installing a light sensor under the lighting system. This sensor will help monitor the change in illumination levels over time in order to signal the moment when it reaches its lowest allowed value, denoted $\boldsymbol{L_t}$. The value of $\boldsymbol{L_t}$, which will be expressed in lumen (lm), can be determined from $\boldsymbol{\mu_t}$ using the CRF. However, even if the CRF of the camera is unknown or inaccessible, the illumination level $\boldsymbol{L_t}$ can still be calculated. Indeed, since the CRF is linear for raw images that are used in this work, the decrease ratio in the illumination is similar to the decrease ratio in pixel intensity. Therefore, since the decrease ratio in pixel intensity has been calculated, the value of $\boldsymbol{L_t}$ can be simply determined by capturing the illumination level $\boldsymbol{L_i}$ under which the image of index i was acquired.

As a result, the light sensor will signal the moment when the illumination level reaches $\boldsymbol{L_t}$, indicating the need for a maintenance operation. This operation aims to regain a higher level of illumination, either by cleaning the light source modules or by increasing the power supply of the illumination system.

6.4 Experiments and results

6.4.1 Common core of all experiments

Once again, all the images that are used in this chapter are raw images that are made of 2046×2046 pixels of 12 bits depth; Note that for clarity and simplicity, in the following experiments, only the red channel has been used.

To conduct the experiments, it was proposed to simulate the effects of the illumination degradation over time by controlling the power supply for the LEDs. Note that this procedure does not aim to simulate the model of degradation of LEDs, but rather to acquire images at different levels of illumination. Hence, the power supply for the LEDs has been decreased step by step to create 60 different levels of illumination. For each level, a total of 500 images of the same wheel have been acquired. As a result, this data base consists of 60 sets of images, each containing 500 images acquired at a different illumination level, to make a total of 30 000 images.

Then, an anomaly has been superimposed on each image, which represents the smallest anomaly intended to be detected. As discussed in section 6.2, the linearity of the CRF for raw images induces that both the background and the anomaly will respond in a similar manner to the decrease in illumination. Hence, in order to guarantee a constant "anomaly-to-background" ratio, the amplitude of the anomaly has been adjusted for each level of illumination. More precisely, for each illumination level, the amplitude of the superimposed anomaly is multiplied by the ratio of illumination decrease. This anomaly will be used in all the experiments.

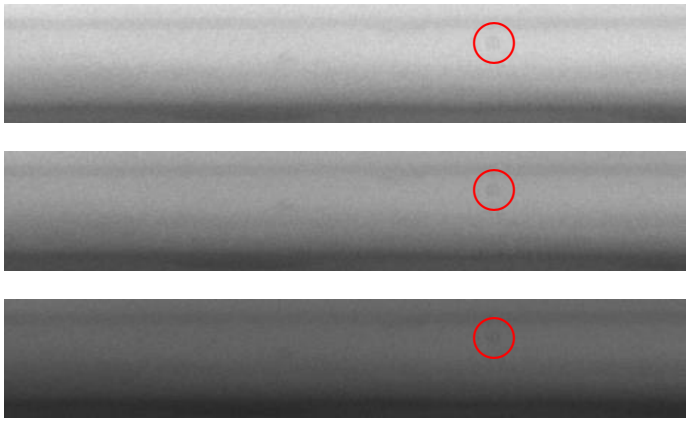


Figure 6.2: Images of the same galbe zone acquired at three different illumination levels

Figure 6.2 represents three images of the same galbe zone acquired at three different levels of illumination, with the inserted anomaly highlighted with a red circle. The first image has the index $i = 1$ and thus belongs to the set number 1 which has the brightest illumination. The second image has the index $i = 15\,000$ and thus belongs to the set number 30 with mid-range illumination. And finally, the third image has the index $i = 30\,000$ and thus belongs to the set number 60 which has the darkest illumination.

Finally, since all the images for these experiments were acquired using the same camera settings, the heteroscedastic noise model parameters (a, b) remain constant. It is therefore proposed to reuse the prior-work in [147] to estimate once the parameters (a, b). Such an estimation resulted in the values $a = 1.08$ and $b = -550.3$.

The rest of this section is divided into two parts. In the first part, the accuracy of the model proposed in section 6.3.1 is studied. Then, the second part investigates the ability of the procedure detailed in section 6.3.2 to correctly estimate the pixel intensity of the background that corresponds to a predefined lowest detection power.

6.4.2 Accuracy of pixel intensity estimation

Let us consider a block on the image of the galbe zone that contains the anomaly. This block is set to a size of $h = 20$ (height) and $w = 20$ (width). And let us denote $\boldsymbol{\mu_i}$ the background of this block for the image of index i. This background resulted from the application of the adaptive model proposed in section 4.2.1, where the degrees of the polynomial used are $d_y = 5$ (along the height) and $d_x = 2$ (along the width), and the number of Principal Components added to this model is $\ell = 3$. For simplicity, it is proposed to use the mean value of each background $\boldsymbol{\mu_i}$ to represent

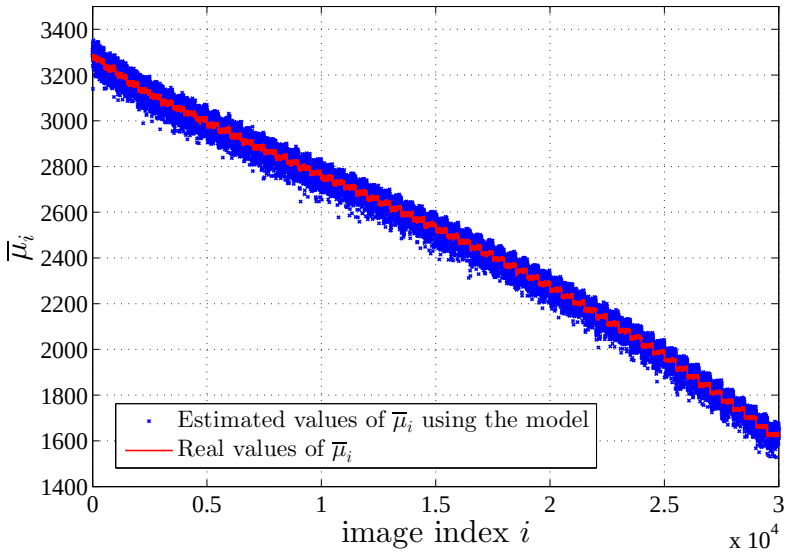


Figure 6.3: Estimation of the mean pixel expectation $\overline{\boldsymbol{\mu}}_i$ with the model (6.17), using the non-centrality parameter of image i, and the non-centrality parameter and mean pixel expectation of only the first image.

its pixel intensity, which will be denoted $\overline{\boldsymbol{\mu}}_i$.

The red plot in Figure 6.3 represents the values of $\overline{\boldsymbol{\mu}}_i$ for $i = \{1, \ldots, 30\,000\}$. Each set of successive 500 images have approximately the same value of $\overline{\boldsymbol{\mu}}_i$, as they are acquired under the same conditions of illumination. With each decrease in illumination, the value of $\overline{\boldsymbol{\mu}}_i$ drops to a new level that is lower than the previous one.

Let us now suppose that only $\overline{\boldsymbol{\mu}}_1$ is known, while $\overline{\boldsymbol{\mu}}_{i>1}$ are all unknown. Hence, it is wished to use Eq. (6.17) to estimate the values of $\overline{\boldsymbol{\mu}}_{i>1}$ using only the background of the first image $\overline{\boldsymbol{\mu}}_1$ as the reference. To do so, it is necessary to determine the change ratio in the non-centrality parameter between the first image of index $i = 1$, and all the other images for $i > 1$. Therefore, it is proposed to use Eq. (6.2) to calculate the values of $\boldsymbol{\varrho_i}$ for $i = \{1, \ldots, 30\,000\}$, where $\boldsymbol{\varrho_i}$ represents the non-centrality parameter corresponding to the anomaly in the image of index i.

Figure 6.3 illustrates the simulation results. The red plot represents the real values of $\overline{\boldsymbol{\mu}}_i$ calculated using the adaptive model of the background, while the blue plot represents the estimated values of $\overline{\boldsymbol{\mu}}_i$ with the model in Eq. (6.17) using the non-centrality parameter of image i, and the non-centrality parameter and mean pixel expectation of only the first image of index $i = 1$.

It can be seen that the estimated values of $\overline{\boldsymbol{\mu}}_i$ correctly follow their real values, with a total root mean square error of $RMSE = 29.99$ on 12 bits images. In perspective, for a normalized image, this $RMSE$ is equal to 7.3×10^{-3} which

indicates the accuracy of the estimation.
While it seems that the calculations to determine the model in Eq. (6.17) are straightforward and thus the estimation must be exact, this is not the case. Indeed, many factors will have an impact on the accuracy of the estimation. Mainly, Eq. (6.17) is based on the assumption that the projection matrices for two different levels of illumination will be the same, which is not precise. While the difference will be very slight between the two matrices, they are not exactly equal. However, even with this assumption, Eq. (6.17) provides an accurate estimation of the change ratio in background intensity.

In conclusion, the main result that can be derived from this simulation is that Eq. (6.17) will allow to estimate the background intensity for any level of illumination, given only the background intensity of one single image, along with the change ratio in non-centrality parameter. The accuracy of this estimation will be further highlighted in the following experiments.

6.4.3 Precision of the lowest acceptable pixel intensity

The goal of this chapter is to detect the illumination level that corresponds to a predefined lowest acceptable detection power. Because this procedure heavily depends on an accurate estimation of the pixel intensity that corresponds to this detection power, it is proposed in this second part of the simulations to study the accuracy to assign for each pixel intensity, a corresponding detection power. In other words, for the same smallest anomaly intended to be detected, each illumination level will be assigned a corresponding detection power.

Figure 6.4 illustrates the results of the simulation. First, the blue and red plots represent respectively the empirical and theoretical detection power for the defect shown in Figure 6.2, function of the mean background pixel intensity $\overline{\boldsymbol{\mu}}$. It can be seen that the detection power was at its highest for the images with the brightest illumination, and then starts to decrease with the degradation of illumination. A non-negligible fluctuation in the blue plot can also be seen. This can be explained by the fact that the empirical detection power for each illumination level is estimated on 500 images only, which is rather small to be accurate in the empirical results. Another contributing factor to this fluctuation is that, the proposed adaptive model that is used to represent the background, though efficient, is not perfect and, hence, maybe sometimes unable to describe the background with highest accuracy.

Then, the black and green plots both represent the estimated values of $\overline{\boldsymbol{\mu}}$ for different values of $\boldsymbol{\beta}$, using the procedure described in section 6.3.2. For each value of $\boldsymbol{\beta}$, the corresponding non-centrality parameter $\boldsymbol{\varrho}$ is calculated by applying Newton's method on Eq. (6.18). Next, for the black plot, the corresponding background intensity $\overline{\boldsymbol{\mu}}$ is calculated using Eq. (6.17) given only the background intensity of the first image, along with the change ratio in non-centrality parameter.

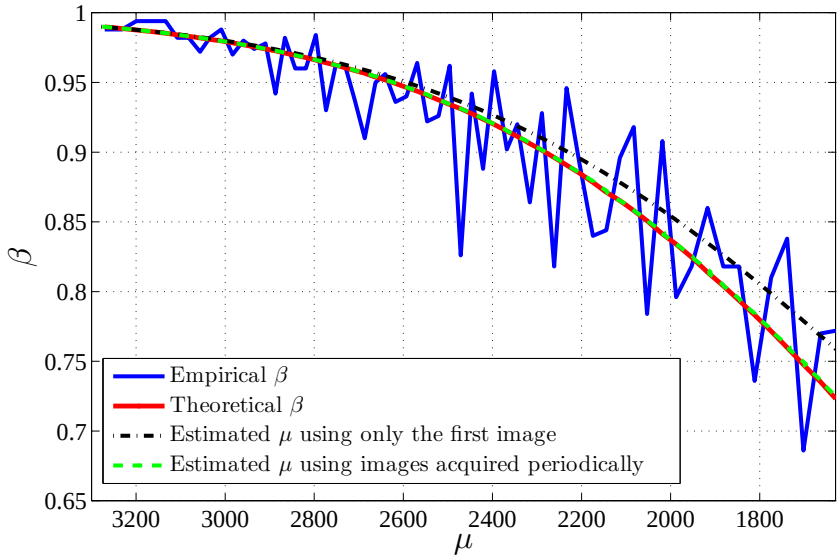


Figure 6.4: Empirical and theoretical detection power, along with the estimation of pixel expectation with the model (6.17), once using only the first image and then using images acquired periodically.

As for the green plot, a more realistic scenario is considered, where the image that is used to estimate the value of $\overline{\boldsymbol{\mu}}$ is updated periodically during time. It can be seen that in both cases, for high levels of illumination, the estimation of $\overline{\boldsymbol{\mu}}$ is very accurate when compared to the theoretical power curve (red plot). However, with the illumination degradation, the estimates obtained from a single image begins to go further away from the theoretical values of $\overline{\boldsymbol{\mu}}$, while the ones obtained from images acquired periodically maintain their accuracy. It can then be concluded that the estimation of pixels intensity will be more accurate if it is closer to the pixels intensity of the reference image.

To better understand the previous results, and to validate the previous conclusion, let us fix the value of the lowest acceptable detection power to $\boldsymbol{\beta_t} = 0.85$. Theoretically, this value of $\boldsymbol{\beta_t}$ corresponds to a mean background intensity of $\overline{\boldsymbol{\mu}}_t = 2050$. Then, let us denote $\widetilde{\overline{\boldsymbol{\mu}}}_t$ the estimated value of the mean background intensity for the same $\boldsymbol{\beta_t} = 0.85$ using the procedure described in section 6.3.2. This estimation is performed multiple times, by changing the reference mean background intensity, denoted $\overline{\boldsymbol{\mu}}_i$, to the one that corresponds to each illumination level. Figure 6.5 shows the variation of the estimation error at each illumination level, where the reference image is changed. It can be seen that the estimation error is at its lowest value when the reference image has a background intensity of $\overline{\boldsymbol{\mu}}_i = \overline{\boldsymbol{\mu}}_t$. However, when the reference background intensity $\overline{\boldsymbol{\mu}}_i$ is getting farther from $\overline{\boldsymbol{\mu}}_t$, the estimation error starts to increase. Indeed, as seen in the first part of the simulations in section 6.4.2, the estimation of the change ratio in background intensity, though efficient, is not perfect. Hence, even a slight error in the estimation will be further aggravated when

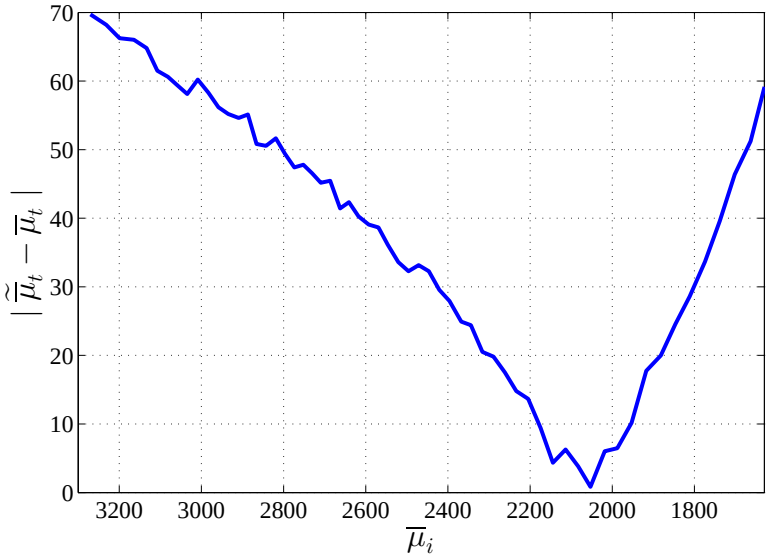


Figure 6.5: Estimation error for different reference images.

the ratio of change between the two illumination levels is more important.

In conclusion, these simulations show that periodically updating the estimation of the lowest allowed illumination level over time will improve its accuracy.
Note that in this simulation, between the first illumination level, which represents the current status of the system, and the lowest allowed illumination level for $\boldsymbol{\beta_t} = 0.85$, the pixel intensity decreases to about 60% of its initial value. Due to the linearity of the CRF for raw images, this means that the illumination level of the LEDs has also decreased to about 60% of its initial intensity. Depending on the LED quality and its working conditions, the time period necessary for the LED light output to decline to 60% of its initial flux usually reaches thousands of hours [100]. However, this time period is further decreased in an industrial environment, since the AVI system will get very dusty way before thousands of hours of use.

6.5 Conclusion

This chapter addressed the problem of predicting the maintenance needs of the AVI system due to illumination degradation over time. Given the same defect on the inspected surface, a model is created that allows to determine the change ratio in pixel intensity, function of the change ratio in the non-centrality parameter along with the initial background intensity. Then, starting from a predefined lowest acceptable detection power, the complete procedure to determine the corresponding lowest acceptable illumination level is described. The final step consists in monitoring the change in illumination levels over time using a light sensor, in order to signal the moment when it reaches its lowest allowed value. Numerical results on a large set of real images show the accuracy of the calculated model to determine the change ratio in pixel intensity. It is seen that this accuracy is improved if the

change ratio is small. It is then concluded that periodically updating the estimation of the lowest allowed illumination level over time will improve its accuracy.

Chapter 7

General conclusion and perspectives

7.1 Summary of the presented work

The main objective of this thesis is to design a fully automatic system for the inspection of wheels' surface. Following the review on the most widely used NDT techniques for surface inspection, it is concluded that the automated visual inspection technique is the most appropriate for this specific inspection operation. Hence, at first, this thesis addressed the problem of designing and installing an imaging system, given a set of requirements and constraints. A practical work plan has been adopted. The first consists in defining the requirements and conditions under which the inspection must be performed. Some of these requirements are related to the inspected wheel, while others are imposed by the industry. Alongside, some possible solutions are proposed to meet each of the requirements, with the goal to outline the initial design of the whole system. Next, a general overview on all the characteristics of each key element of the AVI system is provided. This overview serves as a brief reference to design a complete imaging system from scratch. Then, based on the predefined inspection requirements, a detailed discussion on the appropriate choice of each key element of the AVI system is presented. These key elements are the cameras, the lenses, the lighting system, and the processing platform. Finally, a full description of the AVI system design and its installation setup is presented. This system consists of a total of four cameras installed over the production line of a wheel industry. These cameras are distributed in a hemispherical way around the inspected wheel, enabling a full inspection of its surface from all sides.

Afterwards, the first inspection task addressed in this thesis is the defect detection procedure. The proposed approach consists in modeling the non-anomalous background of the inspected surface using a parametric model. An original adaptive linear model is designed that consists of two parts: A fixed part based on a two-dimensional algebraic polynomial, which is primarily efficient to accurately represent homogeneous surfaces, or smooth surfaces, with little to no texture. Then, an adaptive part based on the PCA, which will offer a higher flexibility, and will enhance the performances of the model to handle minor surface complexities. Hence, this model can be used for the inspection of a variety of surfaces that have mostly a low textured surface. Since this model makes uses of prior physical information

about the inspected surface, a detailed discussion on tuning its parameters is also provided.

From a statistical point of view, this background is considered as a nuisance parameter that has to be removed. Indeed, this background has no interest in the detection process while it may hide the defects and, hence, may prevent their detection. Therefore, a statistical test based on the rejection of this nuisance parameter is proposed. This test exploits the linearity of the background model to perform the rejection through a linear projection. The remaining residuals are then normalized using the specific heteroscedastic noise of the imaging system. This noise model is well known for being more accurate than the usual AWGN model for raw images and allows to take into account the variance of each pixel in the ensuing statistical test, to improve its accuracy. The statistical performance of the test is then established. More precisely, the test is UBCP, with a maximal and constant detection power on a local surface defined by the anomaly-to-noise ratio. Through a discussion on the defect detectability, it is seen that this ratio mainly depends on how much of the anomaly still remains after the background rejection. Furthermore, since the noise corrupting the image depends on the pixels' expectation, the anomaly-to-noise ratio is also affected by the background intensity. This last feature will be further exploited in chapter 6.

The proposed detection method is then applied for the wheel surface inspection problem studied in this work. Due to the nature of the wheels, the different elements are analyzed separately. Several simulations on a large data set of real images have been performed. First, the advantages of adding the adaptive part, based on the PCA, in the proposed model of the background are shown. Second, the advantages of using the heteroscedastic noise model rather than the usual AWGN model are highlighted. The third part of the simulations shows both the accuracy of the proposed adaptive model and the sharpness of the ensuing statistical test. Then, the fourth part compares the performance obtained using the proposed detection method with performance from other recently proposed surface defect detection methods. Finally, some results are presented by applying the proposed adaptive model on wheel images containing real defects located on the surface of the wheel.

The second inspection task addressed in this thesis aims at monitoring online a non-stationary process to detect abrupt changes in the process mean value. In our case, this non-stationary process results from the variation of the paint quantity on the inspected wheels surface, where the abrupt change corresponds to a sudden lack of paint. The proposed approach consists in modeling the acceptable changes in the non-stationary process to be able to distinguish them from the abrupt changes that reveal a malfunctioning. A polynomial model is considered and is applied on a fixed length window containing the last set of acquired observations. Indeed, this acceptable variation of mean value, modeled using the polynomial model, is a nuisance parameter as it is of no use for the considered detection problem. Hence, the maximum likelihood (ML) estimation method was considered to perform a rejection of this nuisance parameter through a linear projection. Then, a sequential method

with a fixed length window is proposed. This second window also corresponds to a predefined fixed maximal detection delay under which the detection is performed. In summary, the proposed sequential method is based on two fixed-length windows, where the first window is used for estimating the non-stationary process parameters while the second window is used to execute the detection. Finally, to comply with requirements on low false alarm probability and highest change-point detection performance under a maximal delay constraint, the performance of the proposed method is studied. This study shows that the statistical performance of the proposed test is bounded. The false alarm probability is upper bounded which will enable to calculate a detection threshold using a predefined false alarm rate knowing that the application is guaranteed not to exceed. On the other hand, the detection power of the test is lower bounded which will allow to guarantee, for a predefined false alarm rate, a minimal detection power that the application will not decrease bellow.

The proposed sequential method is then applied for the wheel coating monitoring problem studied in this work. The mean value of pixels from all wheel images are used to measure the coating intensity. To properly study the performance of the proposed method, the proper choice of the first window length and the degree of the polynomial is first discussed. Several numerical results on a large set of images are presented using the proper parameters. These results show the accuracy of the proposed model, the efficiency of the proposed detection method, and the sharpness of the statistical performances theoretically established. In addition, the effect of the second window length on the performances of the proposed test is studied. It was shown that, depending on the application, the proposed test allows to either increase the detection performance at a cost of a larger detection delay, or decrease the detection delay at a cost of a lower detection performance. Finally, a performance comparison is conducted to highlight the advantages of modeling the observations and examine the difference in the detection criteria between our approach and the CUSUM method. It was concluded that the CUSUM outperforms the proposed sequential method in terms of average detection delay (ADD) as a function of the average run length to false alarm (ARLFA). On the other hand, the proposed sequential method outperforms the CUSUM in terms of detection power as a function of the average run length to false alarm (ARLFA).

The final problem addressed in this thesis is the prediction of maintenance needs of the AVI system due to LED degradation over time. The proposed approach is to study the impact of the degradation of illumination on the performance of the detection method. Two characteristics of the camera response function CRF are exploited. First, since the CRF is a monotonically increasing function, the degradation in illumination induces a decrease in the background intensity of the inspected image. Secondly, since the CRF is linear for raw images, a same defect on the inspected surface will maintain a constant anomaly-to-background ratio during the illumination degradation. Hence, given the same defect on the inspected surface, these two features are used to create a model, which allows to determine the change

ratio in pixel intensity, function of the change ratio in the non-centrality parameter along with the initial background intensity. Then, starting from a predefined lowest acceptable detection power, the complete procedure to determine the corresponding lowest acceptable illumination level is described. The final step consists in monitoring the change in illumination levels over time using a light sensor, in order to signal the moment when it reaches its lowest allowed value.

Numerical results on a large set of real images show the accuracy of the calculated model to determine the change ratio in pixel intensity. It is seen that this accuracy is improved if the change ratio is small. It is then concluded that periodically updating the estimation of the lowest allowed illumination level over time will improve its accuracy.

7.2 Perspectives

Following this brief summary of the presented work in this thesis, it is proposed to present some perspectives that could extend the discussion on some subjects and offer new potential research ideas.

- In chapter 3, during the selection of the main processing platform, it was noted that the choice between the multi-core CPU, the GPU, and the FPGA is not obvious since each has its own specific advantages that make it optimal for specific applications. Hence, for the prototype, it was decided to adopt multi-core CPUs as they are the most flexible solution and the easiest to program, while providing high performances. However, now that the inspection methods are decided, it is interesting to test each individual platform, and compare their performance in terms of processing time, in order to decide on the fastest solution.

- Even so four cameras have been installed, all the simulations used only images acquired from the camera directly above the inspected wheel. It is then expected to use the side cameras to improve the inspection of the whole wheel surface. Indeed, some parts of the wheel surface, such as the ventilation zone, are better illuminated with the side cameras, which would improve the detectability of defects located on those parts as discussed in chapter 6. Therefore, a complete calibration of the multi-camera system would be of importance to properly distribute the inspection tasks on the different cameras.

- A polynomial model, aided with PCA, has been considered to represent the background of the inspected image. This model proved its efficiency for wheel surface inspection, as it maintained most of the defect within the residuals. It would be interesting to try other types of parametric models that are more robust, and study the detectability of defects in such cases.

- The statistical test designed in chapter 4 for defect detection could be further extended to take into account the variance of the estimation of pixels expectation. This could potentially improve the precision of the detection, and provide a better control of the false alarm rate.

- The research on surface defect detection methods is expanding. Deep learning is a new emerging research domain that showed some potentials for defect detection on surfaces with different types of textures [197]. It is then important to study its detection capabilities, and investigate its efficiency for industrial applications, especially its sensibility to variations in the inspection system.

- In chapter 5, modeling the non-stationary process over the first window was conducted using a polynomial model. Replacing this model with any other type of parametric models is feasible, and will not alter the detection procedure performed on the second window. Hence, another interesting perspective is to test the accuracy of other types of parametric models to model the non-stationary process, such as the autoregressive integrated moving average (ARIMA) for example, and to compare its accuracy with the one of the proposed polynomial model.

- A final perspective is related to the problem of monitoring the mean value of the background presented in chapter 6. It consists in monitoring online the mean value of a Gaussian random sequence to detect the moment it exceeds a prescribed value. The main challenge to design a statistical test is that the variance is an affine function of the mean value.

Chapter 8

French Summary

8.1 Introduction

Dans l'hypothèse où les fabricants de presque tous les secteurs d'activité se retrouvent en concurrence avec des entreprises de toutes les régions du monde, la satisfaction du client est un élément clé de survie et de succès. Un déterminant essentiel de la satisfaction du client est la qualité du produit. Garantir la qualité des produits fabriqués a toujours été une nécessité pour satisfaire les clients. Par conséquent, de nombreux fabricants ont tendance à examiner la qualité de leurs produits non pas de leur point de vue, mais plutôt du point de vue du client.

De nos jours, les clients sont mieux informés et plus attentifs à ce qu'ils perçoivent comme étant la qualité d'un produit. Alors que la capacité du produit à remplir ses fonctions attendues reste la principale préoccupation du client, sa perception visuelle est également très importante. Les fabricants qui ont atteint des capacités élevées sur les aspects techniques et fonctionnels de la production se différencient par le contrôle de la perception de leur produit, en particulier la perception visuelle de sa surface. Cela a conduit la nécessité d'effectuer une inspection visuelle pour s'assurer que chaque produit fabriqué répond aux caractéristiques visuelles attendues.

L'inspection visuelle des produits finis a toujours été l'une des applications les plus reconnues du contrôle de qualité de n'importe quelle industrie. Il est effectué à la fin de la chaîne de production, lorsque toutes les étapes de fabrication sont terminées. Son importance vient du fait que c'est la dernière étape de l'inspection, et le dernier recours pour arrêter un produit défectueux avant qu'il n'atteigne le client. Cette inspection reste en grande partie un processus manuel (entièrement ou partiellement) conduit par les opérateurs, dont le rôle principal est d'inspecter chaque produit manufacturé. Malheureusement, ce processus manuel est confronté à des limites considérables qui le rendent peu fiable pour l'inspection.

Alors que l'inspection humaine bénéficie de certains avantages, principalement sa grande flexibilité en ce qui concerne les différents types et formes de produits inspectés, elle présente en revanche de nombreux inconvénients majeurs. D'abord, la variabilité de la décision pour un inspecteur dans le temps (facteurs liés à la fatigue et à la motivation) et la variabilité de la décision entre différents opérateurs pour l'évaluation du même produit. Un deuxième inconvénient pour l'inspection humaine est la répétabilité de la décision du même inspecteur lorsqu'il évalue le même produit plusieurs fois. Ces inconvénients entraînent une incertitude et un manque de précision lors de l'inspection.

Pour surmonter ces difficultés, l'inspection automatisée s'est avérée être la

meilleure alternative sur laquelle les industries peuvent compter. La part de la variabilité et de la subjectivité dans la prise de décision est éliminée grâce à une inspection de surface automatisée. Lorsqu'un même produit est inspecté plusieurs fois avec un système d'inspection automatisé, le résultat est, a priori, le même à chaque fois.

Un système d'inspection visuelle automatique (AVI) est composé de deux principaux sous-systèmes. Tout d'abord, le sous-système d'acquisition d'image qui est basé sur le matériel. Il a pour rôle de transformer la scène optique en un ensemble de données numériques reçues par la plateforme de traitement. Deuxièmement, le sous-système de traitement de l'information qui est basé principalement sur des méthodes de traitement d'images développées pour analyser les données acquises et donner le résultat final de l'inspection.

Le premier objectif de ce travail consiste à concevoir le sous-système d'acquisition d'images. En effet, les exigences générales sont que tout système AVI doit être rapide, rentable et fiable. Cependant, pour être en mesure de mettre en place un système AVI efficace, une description plus détaillée des exigences d'inspection est nécessaire. Il est absolument essentiel que le système conçu soit adapté aux exigences spécifiques de l'application.

Un deuxième objectif consiste à développer une méthode de détection de défauts pour l'inspection de surface. La méthode doit être suffisamment générale pour l'inspection d'une large gamme de surfaces, ce qui permet de l'utiliser dans divers domaines d'inspection de la qualité. De plus, dans un contexte industriel, il est nécessaire de contrôler le taux de fausses alarmes pour éviter un arrêt inutile de la production, tout en assurant la meilleure détection des défauts. Compte tenu de ces contraintes, il est préférable de recourir à la théorie des tests d'hypothèses pour concevoir un test dont les performances statistiques sont analytiquement connues. En outre, la méthode peut utiliser des informations statistiques préalables sur la surface inspectée pour rendre l'inspection insensible aux variations des conditions d'acquisition.

Un troisième objectif consiste à surveiller en ligne un processus non stationnaire pour détecter des changements brusques dans sa valeur moyenne. Dans notre cas, ce processus non stationnaire résulte de la variation de la quantité de peinture sur la surface des roues inspectées, où le changement brusque correspond à un manque soudain de peinture. Puisque le processus surveillé est non stationnaire, c'est-à-dire qu'il change naturellement avec le temps, la méthode proposée doit être capable de distinguer les changements de processus "réguliers" des changements brusques résultant de défaillances potentielles. De plus, étant donné que cette méthode vise à être appliquée à des processus industriels, il est nécessaire de détecter le changement dans un délai maximal de détection donné et de contrôler la probabilité de fausse alarme sur un nombre d'échantillons donné. Ainsi, la méthode proposée s'inscrit dans la catégorie de méthodes de détection séquentielle, fonctionnant selon un critère d'optimisation non classique adapté aux applications industrielles.

Un dernier objectif consiste à prévoir les besoins de maintenance du système AVI, requis en raison de la dégradation de l'éclairage dans le temps. D'une part,

cela permettra de maintenir le système dans un état de fonctionnement satisfaisant et, d'autre part, d'éviter un arrêt excessif de la production pour des opérations de maintenance inutiles. L'étude de l'impact de la dégradation de l'éclairage sur la performance de la méthode de détection permettra d'estimer le niveau d'éclairage au-dessous duquel l'inspection n'est plus efficace.

Pour répondre aux objectifs définis ci-dessus, la structure globale de ce chapitre se compose des cinq sections suivantes :

- Section 8.2 présente la procédure de conception du système AVI pour l'inspection de surface des roues en temps réel. Plus précisément, ce chapitre se concentre sur la conception du système d'imagerie et son installation. Premièrement, les exigences et les conditions dans lesquelles l'inspection doit être effectuée sont définies. Certaines de ces exigences sont liées à la roue inspectée, tandis que d'autres sont imposées par des contraintes industrielles. Ensuite, en fonction des exigences d'inspection prédéfinies, le choix approprié de chaque élément clé du système vision est présenté. Enfin, la configuration de l'installation du système AVI est décrite.
- Section 8.3 étudie le problème de détection de défauts sur des surfaces inspectées à l'aide d'un système de vision. Pour contrôler le taux de fausses alarmes, la méthode proposée s'appuie sur la théorie des tests d'hypothèses pour concevoir un test dont les performances statistiques sont analytiquement connues. Le contenu conforme de l'image (sans défauts) qui représente la surface inspectée agit ici comme un paramètre de nuisance. Comme il n'a aucun intérêt pour la détection des défauts, alors il doit être soigneusement pris en compte, car il pourra potentiellement masquer les défauts. Par conséquent, il est proposé de développer un modèle adaptatif de la surface imagée. Ce modèle permet d'aborder l'inspection d'une large gamme d'objets, et il peut être utilisé pour toute application nécessitant un modèle des paramètres de nuisance.

  Puis, comme aucune information préalable sur l'occurrence des défauts n'est disponible, le problème est celui d'un test statistique entre hypothèses composites. Dans un tel cas, un test uniformément le plus puissant avec puissance constante (UPPC) est proposé, basé sur la réjection des paramètres de nuisance. Enfin, pour prendre en compte le bruit présent dans l'image, un modèle de bruit précis et réaliste est adopté. Cela permet d'établir avec une grande précision les propriétés statistiques théoriques du test proposé.

  La méthode de détection proposée est ensuite appliquée pour l'inspection de la surface des roues. Les résultats numériques sur un grand nombre d'images réelles montrent à la fois la précision du modèle adaptatif proposé et l'efficacité du test statistique qui en découle.
- Section 8.4 traite le problème de la surveillance en ligne d'un processus non stationnaire pour détecter des changements brusques dans sa valeur moyenne. La principale particularité du problème abordé dans ce chapitre est que les

paramètres de la distribution d'un processus non stationnaire peuvent changer naturellement dans le temps, ce qui fait que les hypothèses de détection sont composites. Une solution habituelle consiste à utiliser un rapport de vraisemblance généralisé qui consiste à substituer ces paramètres inconnus par leurs estimations en utilisant l'estimation du maximum de vraisemblance. Ainsi, dans un premier temps, le modèle utilisé pour traiter la non-stationnarité des observations est présenté. Deuxièmement, le test statistique qui suit est détaillé. Ensuite, pour respecter les exigences relatives à une faible probabilité de fausse alarme et, aux performances de détection du point de changement les plus élevées et une contrainte sur le délai maximal de détection, les performances de la méthode proposée sont étudiées.

La méthode séquentielle proposée est ensuite appliquée au problème de la variation de l'intensité du revêtement de peinture des roues produites. Les résultats numériques obtenus sur un grand nombre d'images réelles sont présentés et la précision de la performance théorique de la méthode proposée est étudiée.

- Section 8.5 étudie le problème de la prévision des besoins de maintenance du système AVI. Étant donné que le système d'imagerie est composé de composants électroniques, le vieillissement de ces composants peut réduire les performances d'inspection et éventuellement provoquer une défaillance. Ce chapitre se concentre sur la dégradation de l'éclairage causée par la détérioration des modules LED dans le temps. Cette dégradation a un impact négatif sur la méthode de détection proposée dans la section 8.3. Ainsi, l'approche proposée dans ce chapitre est l'étude de l'impact de la dégradation de l'éclairage sur la performance de la méthode de détection. La modélisation de cet impact permettra d'estimer le niveau d'éclairage au-dessous duquel le système n'est plus fiable. Enfin, une solution industrielle est proposée pour détecter le moment où ce niveau d'éclairage est atteint.
- Section 8.6 conclut le chapitre par une brève synthèse des travaux et présente quelques perspectives.

8.2 Système d'inspection visuelle automatisé

La configuration globale d'un système AVI est toujours la même. Un système d'éclairage génère de la lumière d'une manière spécifique pour éclairer l'objet inspecté dans le but d'améliorer la qualité des images acquises. La scène éclairée est projetée sur le capteur d'une caméra numérique, qui transforme la lumière en signaux électriques, pour finalement créer une image numérique. Une lentille est généralement ajoutée à la caméra pour définir le champ de vision et la zone de la scène à capturer. L'image numérique résultante est ensuite envoyée à une plateforme de traitement qui contrôle l'ensemble du système AVI. Avec la prise en charge des techniques de traitement d'image, la plateforme de traitement analyse et traite les images acquises pour donner les informations d'inspection finales sur la

conformité de l'objet inspecté.

Dans cette section, nous nous concentrerons uniquement sur la conception du sous-système d'acquisition d'images. L'inspection de la surface des roues en temps réel fait face à un certain nombre de défis. Certains sont imposés par les caractéristiques spécifiques de l'inspection de la surface des roues tandis que d'autres sont imposées par le cadre industriel :

- La surface de la roue a une forme géométrique complexe, avec des irrégularités dans des directions différentes. Ces irrégularités peuvent parfois masquer des défauts lors de l'inspection de la roue à partir d'une seule direction.
- La variabilité dans la conception des roues et les différents types de défauts sont des défis majeurs à relever. Le système d'acquisition d'images AVI doit être conçu de manière à permettre l'inspection de toute la surface, quelle que soit la forme de la roue. Par ailleurs, la méthode de détection utilisée au sein de l'AVI doit également pouvoir s'adapter à cette variabilité.
- La plupart des roues sont recouvertes par une couche de peinture brillante qui rend la surface de la roue plus réfléchissante. Cela rend difficile la conception d'un système d'éclairage approprié qui minimise les éventuelles réflexions de la lumière. Ce phénomène peut entraîner des fausses alarmes pendant la procédure de détection.
- Afin d'augmenter la productivité et la maintenir au-dessus d'une certaine limite imposée par l'industrie, les roues ne s'arrêtent pas sous le système AVI, mais plutôt se déplacent à une vitesse constante d'environ un mètre par seconde. Par conséquent, le système AVI doit être capable de capturer l'image de la roue en mouvement sans aucune distorsion ou flou pouvant diminuer la qualité de l'image et donc l'efficacité de l'inspection.

En fonction de ces exigences d'inspection prédéfinies, le choix approprié de chaque élément clé du système de vision peut être effectué. Ces éléments clés sont la caméra, l'objectif, le système d'éclairage et la plateforme de traitement de données. Le tableau 8.1 liste les caractéristiques de chacun de ces éléments clés.

Le système d'imagerie est installé sur la ligne de production d'une industrie de fabrication de roues. La figure 8.1 montre une illustration de la configuration de l'installation. La caméra $n°1$ est installée directement au-dessus de la roue contrôlée, tandis que les caméras $n°2-3-4$ sont installées sur le côté selon un angle de 120°. Seule la caméra $n°2$ est représentée dans la figure 8.1. Elle est installée sur une piste circulaire avec un curseur qui permet de modifier l'angle entre la caméra et la roue, qui va de 30° à 75°, ainsi que la distance à la roue. Par conséquent, la caméra peut être positionnée pour s'adapter au type de roue inspectée. Un seul déclencheur est utilisé pour contrôler les 4 caméras pour enregistrer les 4 images au même moment.

Enfin, la figure 8.2 montre deux images de la même roue. L'image de gauche est acquise en utilisant la caméra $n°1$, tandis que l'image de droite a été prise avec la

Table 8.1: Spécification du système vision

Caméra	Caméra matricielle couleur Capteur CMOS Résolution : 2046×2046 (4 MP) Taille du capteur : $11.3mm \times 11.3mm$ Taille du pixel : $5.5\mu m \times 5.5\mu m$
Objectif	Résolution : 120 lp/mm Longueur focale fixe : 12.5 mm Ouverture : f1.4 - f22
Système d'éclairage	Technique d'éclairage diffusée Modules LED blancs
Plateforme de traitement	Dual Intel Xeon processeur 3.1 GHz 20 cores en total Ram : 32 GB 4 GB NVIDIA Quadro GPU Traitement d'image : C++ / OpenCV

caméra $n°2$. Comme on peut le voir, la caméra $n°2$ permet de mieux voir certaines régions de la surface de la roue qui ne sont pas visibles par la caméra $n°1$, permettant ainsi leur inspection. En effet, la procédure de détection des défauts pour les deux images est la même, mais est effectuée sur différentes régions de la roue.

8.3 Méthode de détection de défauts

Les méthodes de détection de défauts à partir d'images peuvent être divisées en trois catégories [109–111] :

- Les méthodes génériques très flexibles ne reposant sur aucune connaissance préalable de l'objet inspecté. Ces méthodes sont généralement basées sur des outils de traitement d'images et l'amélioration des défauts (amélioration du contraste, détection des contours, etc...) et la reconnaissance de formes. Ils consistent à séparer l'image inspectée en régions de comportement statistique distinct, en partant de l'hypothèse que des propriétés communes peuvent définir toutes sortes d'anomalies et les distinguer de l'arrière-plan conforme. L'existence de telles propriétés n'est pas évidente en pratique et ces méthodes sont souvent sensibles à la géométrie de l'objet et de l'anomalie et, à la présence de bruit.
- Les méthodes spécifiques basées sur une référence [111]. La détection est alors simplement basée sur les différences observées entre la référence et l'image inspectée. Cette approche est efficace, mais aussi très sensible aux conditions expérimentales, telles que la position de l'objet, l'illumination, etc.... De plus, une référence peut être difficile à obtenir dans la pratique.

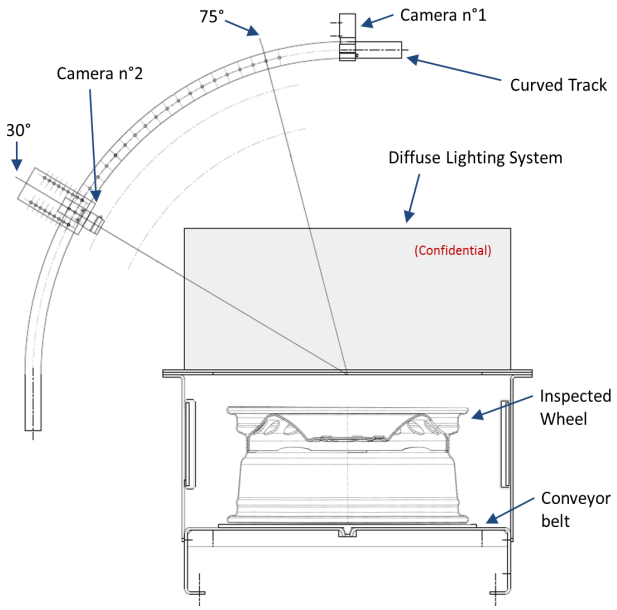


Figure 8.1: Installation du système vision

- Les méthodes basées sur la vision par ordinateur et le traitement d'images [110, 111], qui nécessitent habituellement des informations préalables sur l'objet inspecté. Deux approches principales ont été proposées pour introduire des connaissances statistiques a priori: les approches bayésiennes et non bayésiennes. L'approche statistique bayésienne permet la conception de méthodes simples et efficaces de détection d'anomalies. Cependant, ces méthodes requièrent une connaissance a priori sur les défauts recherchés. Cette exigence limite l'application des méthodes bayésiennes.

Dans le problème de détection d'anomalie considéré, l'arrière-plan conforme de la surface inspectée n'a aucun intérêt dans le processus de détection. De ce fait, il est plus pratique de représenter l'arrière-plan conforme par un modèle paramétrique et d'utiliser une méthode non bayésienne pour la détection d'anomalies.

Les principales contributions de la méthode proposée sont les suivantes :

- Un modèle statistique adaptatif est proposé pour représenter la surface imagée. Ce modèle nécessite seulement la connaissance de la géométrie des objets inspectés. Ainsi, le système de détection des anomalies est entièrement automatique et applicable à une large gamme de surfaces.
- Le modèle proposé assure des performances de détection élevées et est simple en termes de calcul pour les applications en temps réel.
- Le modèle de bruit hétéroscédastique est utilisé pour décrire avec précision les propriétés de bruit dans les images raw. En conséquence, pour d'autres types d'images, le modèle hétéroscédastique peut être remplacé par le modèle



(a) Image acquise avec la caméra $n°1$



(b) Image acquise avec la caméra $n°1$

Figure 8.2: Deux images de la même roue, une acquise avec la caméra $n°1$, et l'autre avec la caméra $n°2$

approprié sans effet majeur sur la précision de la détection.

- Les propriétés statistiques de la méthode sont explicitement fournies. Le seuil de détection ne dépend que de la probabilité de fausse alarme. Par conséquent, un opérateur peut, par exemple, prescrire facilement une probabilité de fausse alarme et savoir quel type d'anomalies peut être détecté avec quelle probabilité.

8.3.1 Méthodologie

Soit $\mathbf{Z} = \{z_{m,n}\}$ l'image bruitée, de la surface inspectée, de taille $M \times N$, où $(m,n) \in \mathcal{Z} = (\{1,\dots,M\} \times \{1,\dots,N\})$. Au cours de l'acquisition, chaque pixel est corrompu par divers bruits qui changent sa valeur. Par conséquent, chaque valeur de pixel $z_{m,n}$ à l'emplacement (m,n) peut être représentée comme suit :

$$z_{m,n} = \mu_{m,n} + \xi_{m,n} \tag{8.1}$$

où $\mu_{m,n}$ est l'espérance du pixel $z_{m,n}$, et $\xi_{m,n}$ représente les bruits qui corrompent le pixel à cet endroit. On suppose généralement que tous les bruits qui modifient la valeur du pixel peuvent être modélisés comme une variable aléatoire gaussienne [137]. En conséquence, la distribution statistique de la valeur du pixel à l'emplacement (m,n) est donnée par :

$$z_{m,n} \sim \mathcal{N}\left(\mu_{m,n}, \sigma^2_{m,n}\right) \tag{8.2}$$

où $\sigma^2_{m,n}$ est la variance du bruit. Cette représentation d'un pixel est considérée lorsqu'aucune anomalie n'est présente sur la surface inspectée. Au contraire, lorsqu'une anomalie est présente sur la surface inspectée, la valeur attendue du pixel est affectée. Par conséquent, $z_{m,n}$ peut s'écrire :

$$z_{m,n} = \mu_{m,n} + \theta_{m,n} + \xi_{m,n} \tag{8.3}$$

où $\theta_{m,n}$ est l'impact de l'anomalie sur l'espérance des pixels. En fait, l'anomalie affecte une zone limitée de l'image, donc $\theta_{m,n}$ est égale à zéro à l'exception de

quelques pixels dans lesquels se trouve l'anomalie. Par suite, lorsqu'une anomalie est présente, le modèle du pixel à l'emplacement (mn) devient :

$$z_{m,n} \sim \mathcal{N}\left(\mu_{m,n} + \theta_{m,n}, \sigma^2_{m,n}\right) \tag{8.4}$$

Lors de l'inspection d'une image, deux situations peuvent survenir: $\mathcal{H}_0 = \{$il n'y a pas d'anomalie $\}$ et $mathcal H_1 = \{$il y a une anomalie $\}$.
A partir des équations (8.2) et (8.4), le problème de détection d'anomalie peut être représenté comme une décision entre les deux hypothèses suivantes :

$$\begin{cases} \mathcal{H}_0 : \left\{\mathbf{z}_{m,n} \sim \mathcal{N}(\mu_{m,n}, \sigma^2_{m,n})\,, \forall (m,n) \in \mathcal{Z}\right\} \\ \mathcal{H}_1 : \left\{\mathbf{z}_{m,n} \sim \mathcal{N}(\mu_{m,n} + \theta_{m,n}, \sigma^2_{m,n})\,, \forall (m,n) \in \mathcal{Z}\right\}, \end{cases} \tag{8.5}$$

avec $\theta_{m,n} \neq 0$ pour quelques (m,n).

En pratique, la difficulté principale est la présence de paramètres de nuisance inconnus, dans la définition des hypothèses, qui n'ont aucun intérêt pour le problème de détection d'anomalies. Ces paramètres de nuisance sont l'espérance des pixels $\mu_{m,n}$ qui décrivent la surface inspectée. Bien que ce paramètre de nuisance ne soit pas lié au problème de détection, il doit être soigneusement pris en compte, à travers la conception d'un modèle qui décrit précisément cet élément, de sorte que le paramètre de nuisance n'empêche pas la détection d'anomalies.

Dans ce but, pour chaque surface inspectée, un modèle adaptatif original est proposé pour soustraire l'arrière-plan conforme. La zone inspectée correspondant à l'image $\mathbf{Z}$ est divisée en petits blocs de taille w et h (respectivement pour la largeur et la hauteur). Notons aussi $\mathbf{z}_k$ le bloc numéro k de l'image inspectée $\mathbf{Z}$.
Le modèle paramétrique adaptatif proposé consiste essentiellement à représenter tous les pixels du bloc $\mathbf{z}_k$ comme une somme pondérée de vecteurs de base représentant les colonnes de la matrice $\mathbf{H}_k$. Le poids de cette somme représente le vecteur des paramètres $\mathbf{d}_k$. Le modèle proposé se compose de deux parties distinctes. Une première partie fixe basée sur un polynôme algébrique bidimensionnel de degrés d_x et d_y sur x et y respectivement. Une deuxième partie adaptative basée sur une analyse en composantes principales, qui change d'un bloc à un autre pour prendre en compte la spécificité de chaque bloc.

Par conséquent, le modèle proposé pour représenter l'arrière-plan, c'est-à-dire l'espérance du bloc $\mathbf{z}_k$, peut être écrit comme :

$$\mathbb{E}\left[\mathbf{z}_k\right] = \mathbf{H}_k\, \mathbf{d}_k. \tag{8.6}$$

Ensuite, pour la modélisation du bruit, il est proposé d'utiliser le modèle hétéroscédastique. En fait, un modèle habituel du bruit corrompant les images raw peut être obtenu en considérant séparément le bruit de grenaille, dû au processus de photo-comptage Poissonien, et les divers bruits électroniques. Alors que la variance du processus Poissonien est proportionnelle à son espérance, les bruits électroniques sont modélisés comme une variable aléatoire gaussienne à variance

constante [137, 146]. En raison du très grand nombre de photons comptés, le processus Poissonien peut être approché par une distribution gaussienne [137, 146]. Par conséquent, la valeur du pixel à l'emplacement (m, n) peut être modélisée comme une variable aléatoire gaussienne dont la variance est donnée par :

$$\sigma^2_{m,n} = a\,\mu_{m,n} + b. \tag{8.7}$$

Ces paramètres (a, b) du modèle de bruit hétéroscédastique restent les mêmes pour tous les pixels. En outre, ils dépendent de plusieurs paramètres d'acquisition, donc les paramètres (a, b) sont également constants pour toutes les images prises avec les mêmes paramètres d'acquisition.

Dans toute installation industrielle, il est supposé que la caméra ne change pas donc le système peut être calibré facilement. On suppose donc que les paramètres du modèle de bruit (a, b) sont connus. Le seul paramètre qui doit être estimé, étant donné l'image d'une surface à inspecter, est l'espérance de chaque pixel. Une première estimation est obtenue en utilisant la méthode des moindres carrés :

$$\widetilde{\boldsymbol{\mu}}_k^{\text{ls}} = \mathbf{H}_k \left(\mathbf{H}_k^{\text{T}}\mathbf{H}_k\right)^{-1} \mathbf{H}_k^{\text{T}}\mathbf{z}_k.$$

Ensuite, une estimation de la variance du bruit est obtenue à partir de $\widetilde{\boldsymbol{\mu}}_k^{\text{ls}}$ par :

$$\widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}} = \mathbf{I}_{w\times h} \times (a\,\widetilde{\boldsymbol{\mu}}_k^{\text{ls}} + b),$$

où $\mathbf{I}_{w\times h}$ dénote une matrice d'identité de taille $w \times h$. Cette estimation approximative de la covariance est donc réutilisée pour mettre à jour l'estimation de l'espérance en utilisant la méthode des moindres carrés pondérés :

$$\begin{cases} \widetilde{\boldsymbol{\mu}}_k = \mathbf{H}_k \left(\mathbf{H}_k^{\text{T}}\widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}^{-1}}\mathbf{H}_k\right)^{-1} \mathbf{H}_k^{\text{T}}\widetilde{\boldsymbol{\Sigma}}_k^{\text{ls}^{-1}}\mathbf{z}_k, \\ \widetilde{\boldsymbol{\Sigma}}_k = \mathbf{I}_{w\times h} \times (a\,\widetilde{\boldsymbol{\mu}}_k + b). \end{cases} \tag{8.8}$$

Par conséquent, l'objectif du problème de détection d'anomalies étudié est de décider entre les deux hypothèses composites suivantes :

$$\begin{cases} \mathcal{H}_0 : \{\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k, \boldsymbol{\Sigma}_k)\,,\ \forall k \in \{1, \ldots, K\}\} \\ \mathcal{H}_1 : \{\mathbf{z}_k \sim \mathcal{N}(\mathbf{H}_k\mathbf{d}_k + \boldsymbol{\theta}_k, \boldsymbol{\Sigma}_k)\,,\ \forall k \in \{1, \ldots, K\}\}\,, \end{cases} \tag{8.9}$$

avec $\boldsymbol{\theta}_k \neq 0$ pour quelques k.
Lors du test d'hypothèses composites, le but ultime est de concevoir un test UPP (Uniformément le Plus Puissant,) qui maximise la fonction de puissance et satisfait une contrainte prescrite sur la probabilité de fausse alarme, quelle que soit l'anomalie. Cependant, un tel test existe rarement. Par conséquent, il est proposé d'appliquer le principe d'invariance pour supprimer les paramètres de nuisance $\mathbf{H}_k\mathbf{d}_k$ et de concevoir un test Uniformément le Plus Puissant avec puissance Constante (UPPC). En effet, l'espérance sous $\mathcal{H}_0$ donnée par $\mathbf{H}_k\mathbf{d}_k$ n'a aucun intérêt

pour le problème de test (8.9), mais doit être prise en compte. Pour supprimer les paramètres de nuisance, l'idée est de projeter les observations $\mathbf{z}_k$ sur le complément orthogonal du sous-espace engendré par les colonnes de $\mathbf{H}_k$. Ceci est réalisé en utilisant le projecteur :

$$\mathbf{P}^{\perp}_{\mathbf{H}_k} = \mathbf{I}_{w\times h} - \left(\mathbf{H}_k \left(\mathbf{H}_k^{\mathrm{T}} \widetilde{\mathbf{\Sigma}}_k^{-1} \mathbf{H}_k\right)^{-1} \mathbf{H}_k^{\mathrm{T}}\right) \widetilde{\mathbf{\Sigma}}_k^{-1}, \tag{8.10}$$

où la covariance estimée $\widetilde{\mathbf{\Sigma}}_k$ est donnée en utilisant l'espérance estimée $\widetilde{\boldsymbol{\mu}}_k$ (8.8). Cependant, comme la variance n'est pas constante sur tous les pixels, il est nécessaire de normaliser les résidus en divisant chaque résidu par son écart-type. Ces résidus normalisés peuvent s'écrire comme suit :

$$\mathbf{r}_k = \widetilde{\mathbf{\Sigma}}_k^{-1/2} \left(\mathbf{P}^{\perp}_{\mathbf{H}_k} \mathbf{z}_k\right). \tag{8.11}$$

où $\mathbf{A}^{-1/2}$ représente la "racine carrée" de la matrice $\mathbf{A}$ définie telle que $\left(\mathbf{A}^{-1/2} \times \mathbf{A}^{-1/2}\right)^{-1} = \mathbf{A}$.
Il est alors facile d'établir [110, 153, 154] que la norme des résidus normalisés $\mathbf{r}_k$ suit la distribution :

$$\|\mathbf{r}_k\|_2^2 \sim \begin{cases} \chi^2_{\Upsilon}(0)\,, \forall k \in \{1,\ldots,K\} & \text{under } \mathcal{H}_0 \\ \chi^2_{\Upsilon}\left(\varrho_k\right)\,, \forall k \in \{1,\ldots,K\} & \text{under } \mathcal{H}_1, \end{cases} \tag{8.12}$$

où $\chi^2_{\Upsilon}\left(\varrho_k\right)$ représente la distribution χ-carré non-centrale avec $\Upsilon = w \times h - p$ degrés de liberté. Ici, p représente le nombre de colonnes de $\mathbf{H}_k$ et le paramètre de non-centralité ϱ_k sous l'hypothèse $\mathcal{H}_1$ est donné par :

$$\varrho_k = \left\| \widetilde{\mathbf{\Sigma}}_k^{-1/2} \mathbf{P}^{\perp}_{\mathbf{H}_k} \boldsymbol{\theta}_k \right\|_2^2. \tag{8.13}$$

Ici, ϱ_k dénote le rapport "défaut-sur-bruit" [110] et est essentiel pour définir la détectabilité de l'anomalie.
Suivant les résidus $\mathbf{r}_k$ et de leur distribution, voir Eq. (8.12), le test UPPC peut être écrit comme suit :

$$\delta = \begin{cases} \mathcal{H}_0 & \text{if } \|\mathbf{r}_k\|_2^2 \leq \tau \\ \mathcal{H}_1 & \text{if } \|\mathbf{r}_k\|_2^2 > \tau, \end{cases} \tag{8.14}$$

où afin de garantir la probabilité de fausse alarme α_0, le seuil de décision τ est défini par :

$$\tau = F^{-1}_{\chi^2_{\Upsilon}}(1-\alpha_0; 0) \tag{8.15}$$

où $F_{\chi^2_{\Upsilon}}(x, \varrho_k)$ et $F^{-1}_{\chi^2_{\Upsilon}}(x, \varrho_k)$ représente respectivement la fonction de distribution cumulative χ^2 et son inverse avec le paramètre de non-centralité ϱ_k.
De même, la fonction de puissance du test est donnée par :

$$\beta(\delta, \boldsymbol{\theta}_k) = F_{\chi^2_{\Upsilon}}(\tau, \varrho_k). \tag{8.16}$$

On peut noter à partir des résultats précédents, Eq. (8.15)-(8.16), deux choses importantes. Tout d'abord, le seuil τ ne dépend que de la probabilité de fausse alarme α_0 et est donc constant pour tous les blocs. Deuxièmement, la détectabilité de l'anomalie dépend uniquement du paramètre de non-centralité ϱ_k (8.13). Plus précisément, Eq. (8.13) montre que ϱ_k est défini comme la partie de l'anomalie $\boldsymbol{\theta}_k$ qui se trouve dans le complément orthogonal du sous-espace engendré par $\mathbf{H}_k$. Par conséquent, une anomalie $\boldsymbol{\theta}_k$ est détectable si et seulement si $\mathbf{P}^{\perp}_{\mathbf{H}_k}\boldsymbol{\theta}_k \neq 0$.

8.3.2 Simulations

Il est souhaité de montrer la pertinence du test statistique proposé et l'exactitude des résultats théoriques. Une simulation Monte-Carlo a été réalisée sur 3 000 images de roues. Comme il est très difficile d'obtenir beaucoup d'images avec des défauts similaires, nous avons pris au hasard un ensemble de 3 000 images conformes sur lesquelles un défaut a été superposé. Ce défaut représente la limite inférieure des critères de détection, au-dessus de laquelle le défaut est destiné à être détecté.
la figure 8.3 représente la distribution empirique de la norme des résidus normalisés $\|\mathbf{r}_k\|_2^2$ sur laquelle repose le test proposé, voir Eq. (8.14). Elle a été obtenue en utilisant le modèle adaptatif proposé sur des blocs de taille $h = 20$ (hauteur) et $w = 20$ (largeur). Les degrés du polynôme utilisé sont $d_y = 5$ (le long de la hauteur) et $d_x = 2$ (le long de la largeur) et le nombre de composantes principales ajoutées à ce modèle est de 3. L'écart entre la distribution empirique sous $\mathcal{H}_0$ et

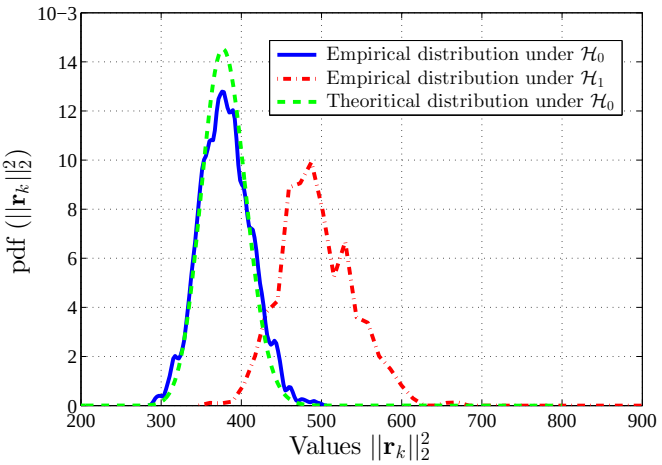


Figure 8.3: Distributions empirique et théorique de la norme des résidus normalisés pour des images avec et sans défauts

la distribution empirique sous $\mathcal{H}_1$ est dû au paramètre de non-centralité ϱ_k sous l'hypothèse $\mathcal{H}_1$ (8.13). Cet écart montre que la détection du défaut est possible avec certaines erreurs de classification.
En outre, la figure 8.3 compare aussi la distribution empirique sous $\mathcal{H}_0$ avec la distribution théorique (8.12). Une petite différence peut être observée entre les distributions empiriques et théoriques. L'effet de cette différence peut être illustré

dans la figure 8.4 qui représente les courbes COR réelles et théoriques. Comme on peut le constater, la performance du modèle proposé est légèrement inférieure à ce que l'on peut attendre théoriquement. Cela peut être expliqué par les deux

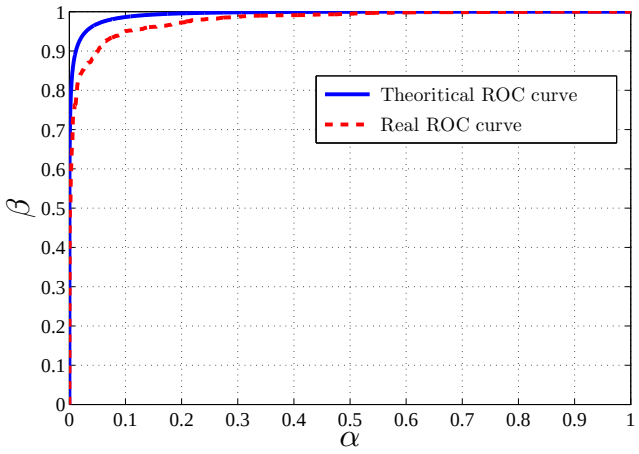


Figure 8.4: Courbes COR réelle et théorique

faits suivants. Premièrement, l'estimation de l'espérance des pixels n'est pas parfaite et possède elle-même une variance non négligeable, qui n'a pas été prise en compte dans le test proposé. Deuxièmement, le modèle adaptatif proposé, malgré son efficacité, n'est pas parfait et par conséquent, peut parfois ne pas décrire l'arrière-plan de l'image avec une grande précision en plaçant une partie de l'arrière-plan conforme parmi les résidus.

Finalement, la figure 8.5 présente, pour un petit ensemble d'images, l'image réelle, l'image du modèle de l'arrière-plan et l'image des résidus correspondante. On peut voir que le modèle proposé est relativement précis et conserve la plus grande partie du défaut dans l'image des résidus.

8.4 Méthode de surveillance du revêtement des roues

Les méthodes séquentielles de détection des points de changement peuvent être classées en deux classes : les méthodes paramétriques et les méthodes non paramétriques.
D'une part, les méthodes non paramétriques ont l'avantage de ne nécessiter ni hypothèse ni aucun modèle sur les observations. Ils sont basés sur des méthodes statistiques pour construire des règles de détection basées sur un grand nombre d'observations. Des règles de décision sont ensuite appliquées aux nouvelles observations. Bien que ces méthodes ne nécessitent pas de modèle pour décrire les observations, elles peuvent cependant être limitées, typiquement lorsque le processus de fabrication change légèrement et qu'elles sont rarement munies de performances statistiques connues.
D'un autre côté, les méthodes paramétriques sont utilisées lorsqu'une information

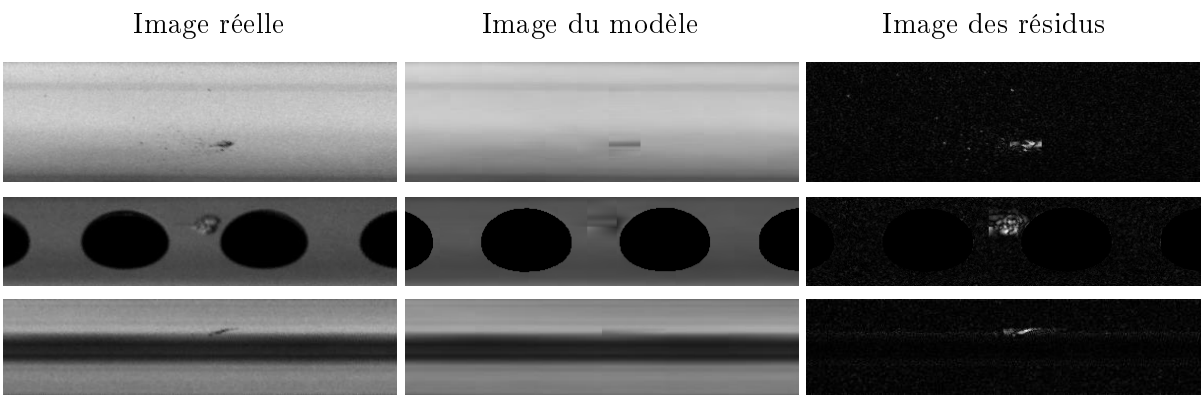


Figure 8.5: Images de roues contenant différents défauts, avec l'image du modèle de l'arrière-plan, et l'image résiduelle

suffisante sur le processus de surveillance est disponible, de sorte qu'un modèle statistique des observations peut être conçu. En d'autres termes, cette approche nécessite que certaines connaissances distributionnelles des données soient disponibles et utilisées dans le schéma de détection. Une limitation commune de ces méthodes est qu'elles s'appuient sur des modèles paramétriques prédéfinis basés sur des informations à priori sur la distribution des données.

Dans ce travail, il est proposé de concevoir une méthode séquentielle paramétrique dans le but de surveiller un processus non stationnaire en temps réel afin de détecter un changement brusque dans sa moyenne. Dans une situation industrielle, il est nécessaire de détecter le changement dans un délai maximal de détection donné (nombre d'observations après le changement) et il est souhaitable de contrôler la probabilité de fausse alarme sur une longueur de parcours fixe. Dans ce contexte opérationnel, une méthode séquentielle avec deux fenêtres de longueurs fixes (2FLW-SEQ) basée sur la méthode du CUSUM est proposée. Cette méthode séquentielle est ensuite appliquée pour la surveillance du revêtement des roues. En effet, quand un pistolet de pulvérisation se bloque partiellement, cela se traduit par un changement soudain de l'intensité de l'image causée par un manque de peinture sur la roue.
Les principales contributions de la méthode proposée sont les suivantes :

- Une méthode séquentielle avec deux fenêtres de longueurs fixes (2FLW-SEQ) est proposée pour surveiller un processus non stationnaire en temps réel. La première fenêtre est considérée pour modéliser la non-stationnarité du processus tandis que la seconde fenêtre est utilisée pour la procédure de détection séquentielle.
- La procédure séquentielle proposée fonctionne selon les critères non classiques de la minimisation de la probabilité de fausse détection sous la contrainte d'un délai de détection maximal tout en contrôlant la probabilité de fausse alarme pour un nombre donné d'observations.

- Une étude statistique de la méthode proposée permet de trouver une borne inférieure de la puissance de détection en fonction du délai maximal de détection autorisé et une borne supérieure de la probabilité de fausse alarme pour un nombre donné d'observations.
- Le contexte proposé permet à l'utilisateur de prescrire un délai de détection maximal et une probabilité de fausse alarme pour un nombre donné d'observations et peut savoir quelles amplitudes de changement peuvent être détectées avec une probabilité minimale garantie.

8.4.1 Méthodologie

Considérons une fenêtre glissante de taille L. Après les L premières observations, pour chaque nouvelle observation reçue x_N la fenêtre contient les observations de x_{N-L+1} à x_N. $Y_N = (x_{N-L+1}, \ldots, x_{N-1}, x_N)^T$ représente la fenêtre après l'ajout de l'observation x_N. Le vecteur Y_N est modélisé par la distribution normale suivante :

$$\mathbf{Y}_N \sim \mathcal{N}(\mu_N, \sigma^2 \mathbf{I}_L), \tag{8.17}$$

où μ_N est l'espérance dans cette fenêtre, $\mathbf{I}_L$ est la matrice d'identité de taille L, et σ^2 est la variance supposée constante pour tous $Y_N, \forall N \geq L$.
Un modèle paramétrique linéaire est proposé pour représenter l'espérance μ_N. Cela consiste essentiellement à représenter toutes les observations de la fenêtre Y_N comme une somme pondérée de q vecteurs de base qui représentent les colonnes d'une matrice $\mathbf{H}$ de taille $L \times q$. Le poids de cette somme représente le vecteur de q paramètres $\mathbf{d}_N$. Par conséquent, l'espérance μ_N peut s'écrire comme :

$$\mu_N = \mathbf{H}\mathbf{d}_N. \tag{8.18}$$

Le modèle $\mathbf{H}$ est basé sur un polynôme de degré $q-1$.
Il résulte des équations (8.17) et (8.18) qu'en absence de défaut, le vecteur d'observations Y_N est modélisé par :

$$\mathbf{Y}_N \sim \mathcal{N}(\mathbf{H}\mathbf{d}_N, \boldsymbol{\sigma}^2 \mathbf{I}_L). \tag{8.19}$$

Au contraire, lorsqu'un défaut survient dans le processus, il se produit un changement dans la valeur moyenne qui affectera toutes les observations après le point de changement. Par conséquent, lorsque le changement se produit, les observations Y_N peuvent être modélisées comme suit :

$$\mathbf{Y}_N \sim \mathcal{N}(\mathbf{H}\mathbf{d}_N + a\mathbf{K}_M, \boldsymbol{\sigma}^2 \mathbf{I}_L), \tag{8.20}$$

où le changement brusque de la valeur moyenne est décrit par le vecteur K_M, de taille L, contenant $(L-M)$ zéros avant le changement et (M) fois -1 après, et la constante $a > 0$ représente l'amplitude du changement. Ici, M est le nombre d'observations maximales acceptables après le changement. Par exemple, le vecteur de changement $K_1 = (0, 0, ..., 0, -1)$ décrit un changement qui n'affecte que la dernière observation

dans la fenêtre de taille L.
Il est important de noter que la variation acceptable de la valeur moyenne, modélisée par $\mathbf{Hd}_N$, est un paramètre de nuisance. En effet, il n'a pas d'intérêt pour le problème de détection considéré. Pour traiter ce paramètre de nuisance, il est proposé d'utiliser la méthode d'estimation du maximum de vraisemblance (ML) pour effectuer un rejet de ce paramètre de nuisance :

$$\mathbf{r}_N = \frac{1}{\sigma}\mathbf{W}\mathbf{Y}_N. \tag{8.21}$$

Ici $\mathbf{W}$ est la projection orthogonale de taille $L-q \times L$, sur l'espace nul de $\mathbf{H}$, dont les colonnes correspondent aux vecteurs propres de la matrice $\mathbf{I}_L - \mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}$ associés à des valeurs propres égales à 1. Le vecteur $\mathbf{r}_N$ représente la projection des observations sur l'espace nul de $\mathbf{H}$.
Parmi d'autres, la matrice $\mathbf{W}$ a les propriétés utiles suivantes: $\mathbf{W}\mathbf{W}^T = \mathbf{I}_{L-q}$; il résulte donc des équations (8.19)-(8.21) que les résidus $\mathbf{r}_N$ peuvent être modélisés sous les hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ par la distribution statistique suivante :

$$\begin{cases} \mathcal{H}_0 : \{\mathbf{r}_N \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{L-q})\} \\ \mathcal{H}_1 : \left\{\mathbf{r}_N \sim \mathcal{N}\left(\frac{a}{\sigma}\boldsymbol{\theta}_M, \mathbf{I}_{L-q}\right)\right\}, \end{cases} \tag{8.22}$$

où $\boldsymbol{\theta}_M$ représente le décalage d'espérance, dû au changement brusque, projeté sur l'espace nul de $\mathbf{H}$: $\boldsymbol{\theta}_M = \mathbf{W}\mathbf{K}_M$.
Il est proposé d'utiliser une méthode séquentielle avec une fenêtre fixe de longueur M qui correspond également à un délai de détection fixe prédéfini. Des approches similaires ont été étudiées dans le contexte de la détection séquentielle dans [175,176]. Ils ont proposé d'utiliser le "match space detection" qui est donné dans notre cas par :

$$\delta_N = \begin{cases} 0 & if \quad \widetilde{S}^N_{N-L+1} = \boldsymbol{\theta}_M^T \mathbf{r}_N < \tau \\ 1 & if \quad \widetilde{S}^N_{N-L+1} = \boldsymbol{\theta}_M^T \mathbf{r}_N \geq \tau. \end{cases} \tag{8.23}$$

De l'équation (8.22) il est facile d'établir la distribution statistique des résultats $\widetilde{S}^N_{N-L+1}$ du 2FLW-SEQ proposé :

$$\begin{cases} \mathcal{H}_0 : \left\{\widetilde{S}^N_{N-L+1} \sim \mathcal{N}(\mathbf{0}, \|\boldsymbol{\theta}_M\|_2^2)\right\} \\ \mathcal{H}_1 : \left\{\widetilde{S}^N_{N-L+1} \sim \mathcal{N}\left(\frac{a}{\sigma}\|\boldsymbol{\theta}_M\|_2^2, \|\boldsymbol{\theta}_M\|_2^2\right)\right\}. \end{cases} \tag{8.24}$$

Enfin, la puissance de détection $\beta(M)$ du test proposé (8.24), c'est-à-dire la probabilité de détecter un changement après au plus M observations, est bornée par :

$$\beta(M) \geq 1 - \Phi\left(\frac{\tau}{\|\boldsymbol{\theta}_M\|_2} - \frac{a}{\sigma}\|\boldsymbol{\theta}_M\|_2\right). \tag{8.25}$$

avec Φ la fonction de répartition de la loi normale centrée réduite. Et la probabilité d'avoir une fausse alarme $\alpha(R)$ après R observations est bornée par :

$$\alpha(R) \leq 1 - \Phi\left(\frac{\tau}{\|\boldsymbol{\theta}_M\|_2}\right)^R, \tag{8.26}$$

Les équations (8.25) et (8.26) soulignent les principaux avantages de l'approche proposée. Premièrement, la performance statistique du test proposé est bornée. La probabilité de fausse alarme $\alpha(R)$ est majorée ce qui permet de calculer un seuil de détection τ en utilisant un taux de fausse alarme prédéfini sachant que l'application est garantie de ne pas le dépasser. D'autre part, la puissance de détection $\beta(M)$ du test est minorée, ce qui permet de garantir, pour un taux de fausse alarme prédéfini, une puissance de détection minimale. Deuxièmement, la probabilité de fausse alarme $\alpha(R)$ ne dépend que de la longueur de parcours prescrite R et du délai maximal de détection acceptable M. Enfin, la fonction de puissance (8.25) montre que la précision de la méthode proposée dépend essentiellement du rapport "changement sur bruit" a/σ, ainsi que du délai maximal de détection acceptable M.

8.4.2 Simulations

Dans le but de surveiller la variation de l'intensité du revêtement de peinture sur les roues produites, on considère un bloc contenant s pixels dans l'image de la roue sur lequel la valeur moyenne de tous les pixels est calculée. La fenêtre considérée conserve la même taille et la même position sur la surface de la roue pour toutes les images. Ensuite, pour une image d'une roue, soit $Z = \{z_w\}_{w=1}^{s}$ dénotent la fenêtre contenant s pixels et $m = s^{-1}\sum_{w=1}^{s} z_w$ la valeur moyenne de l'intensité des pixels. Pour effectuer les simulations, une base de données de 500 000 images conformes et successives a été acquise. La procédure décrite ci-dessus a été appliquée pour obtenir les observations m_i avec $i = \{1, 2, ..., 500\,000\}$. L'écart-type observé, lié au système d'imagerie, est $\sigma = 22$. Ce paramètre est supposé être constant pendant le processus de surveillance.

Tout d'abord, il est proposé d'étudier l'effet de la seconde fenêtre M sur les performances de la détection. La base de données a été utilisée pour effectuer une simulation Monte-Carlo, pour laquelle un décalage simulé d'amplitude $a = 60$ a été superposé à certaines observations. La figure 8.6 représente la probabilité de fausse alarme empirique $\alpha(R)$ et la puissance de détection $\beta(M)$ avec $R = 5\,000$ pour 3 valeurs différentes du délai maximal de détection autorisé $M = \{1, 3, 5\}$, en fonction du seuil de décision τ. On peut observer que lorsque M augmente, $\|\boldsymbol{\theta}_M\|_2$ augmente, ce qui affecte à la fois le taux de fausses alarmes $\alpha(R)$ et la puissance de détection $\beta(M)$, comme vu dans l'équation (8.24). Cependant, le taux d'augmentation de $\beta(M)$ est plus grand que celui de $\alpha(R)$. Par conséquent, le décalage entre la puissance de détection et la probabilité de fausse alarme augmente, ce qui implique une meilleure performance de détection, mais avec un délai M plus important. On peut alors conclure que le choix de M dépend essentiellement des exigences de l'application. En fonction de l'application, ce test permet soit d'augmenter les performances de détection au prix d'un délai de détection plus important, soit de diminuer le délai de détection au prix d'une performance de détection plus faible.

Ensuite, il est souhaité d'illustrer l'efficacité de la méthode de détection séquen-

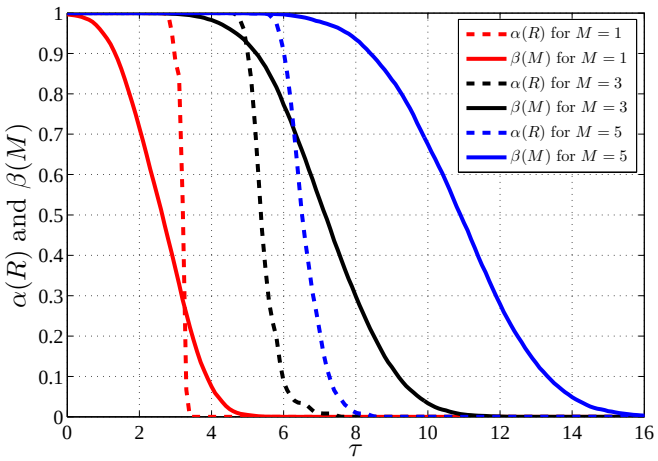


Figure 8.6: Probabilité de fausse alarme $\alpha(R)$ et puissance de détection $\beta(M)$ avec $R = 5000$ pour 3 valeurs différentes de M en fonction du seuil de décision τ.

tielle 2FLW-SEQ proposée sur un scénario réel avec un changement réel dans les observations. La figure 8.7 représente un cas réel d'observations lorsque le pistolet de pulvérisation est partiellement bouché. En conséquence, un changement brusque dans les observations d'une amplitude $a = 55$ peut être vu exactement à l'indice de l'image 2434. La courbe bleue représente les observations réelles, tandis que la courbe rouge représente les valeurs d'espérance (8.18) estimées en utilisant le modèle polynomial sur une fenêtre de taille $L = 200$ et un degré de $q - 1 = 2$. Le taux de fausse alarme est fixé à $\alpha(R = 10^{-3}$ avec $R = 5\,000$. Cela entraînera un seuil de détection de $\tau = 10.12$ pour $M = 5$. Figure 8.8 illustre le résultat de la méthode

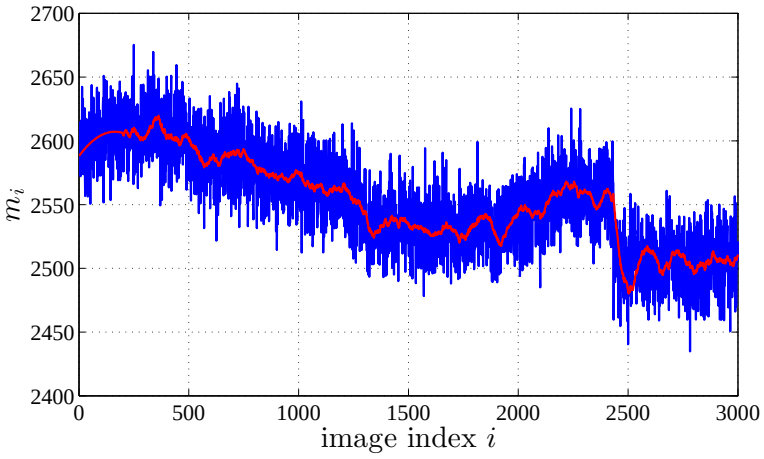


Figure 8.7: Exemple réel de la variation de la valeur moyenne avec un changement brusque à l'indice 2434.

proposée 2FLW-SEQ avec $M = 5$. On peut voir que le point de changement est détecté à l'index 2438 ce qui signifie un délai exact de 5 roues défectueuses.

Il faut noter que, généralement lorsque le changement est détecté, le processus

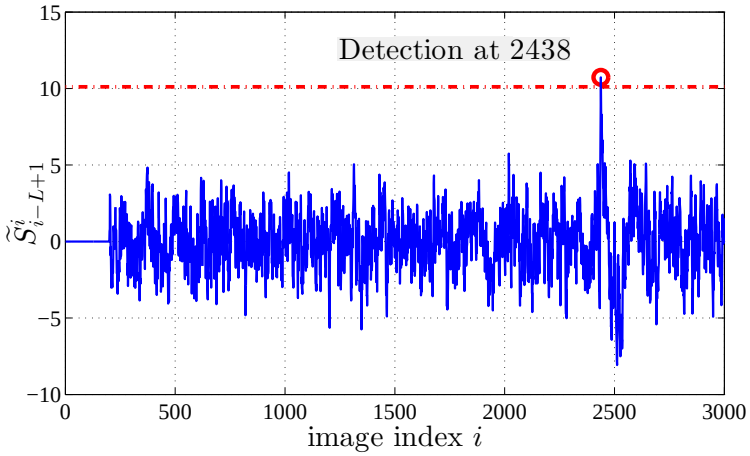


Figure 8.8: Résultat de la méthode proposée 2FLW-SEQ avec $M = 5$

séquentiel s'arrête. Cependant, pour mieux illustrer les résultats du test, la procédure séquentielle a été autorisée à continuer. Il est montré sur la figure 8.7 qu'après le changement, les observations retournent à un état similaire à celui juste avant le changement. Alors, juste après le changement, la procédure séquentielle va réopérer sous l'hypothèse $\mathcal{H}_0$, et les résultats $\widetilde{S}^i_{i-L+1}$ reviendront à une distribution gaussienne avec une moyenne nulle et une variance $\|\boldsymbol{\theta}_M\|_2^2$, comme on peut le voir dans la figure 8.8.

8.5 Maintenance du système vision

Le vieillissement des composants électroniques du système vision pourrait réduire les performances du système et éventuellement provoquer une défaillance. Pour de tels systèmes, la dégradation de l'éclairage dans le temps est une préoccupation majeure [181]. En fait, les sources lumineuses sont basées sur des composés chimiques qui ont un comportement de vieillissement caractéristique [182]. Par conséquent, comme la lumière est le signal utilisé par le capteur de la caméra pour générer les données, l'éclairage de la scène imagée a un impact majeur sur la performance du système d'inspection. Il est alors important d'étudier le comportement du vieillissement de la source lumineuse pour déterminer le moment où le niveau d'éclairage n'est plus suffisant pour que le système puisse exécuter ses fonctions prévues. Cependant, en fonction du type de source lumineuse et de ses conditions de travail, ce vieillissement n'est pas toujours facile à modéliser. Une approche alternative proposée dans ce travail est d'étudier l'impact de la dégradation de l'éclairage sur la performance de la méthode de détection. Cela permettra d'estimer le niveau d'éclairage au-dessous duquel le système n'est plus fiable. Enfin, une solution industrielle est proposée pour détecter le moment où ce niveau d'éclairage est atteint.

8.5.1 Méthodologie

La procédure proposée peut être divisée en deux parties principales. La première partie consiste à déterminer l'intensité de pixel de l'arrière-plan en dessous de laquelle la performance de détection n'est plus satisfaisante. Ensuite, la seconde partie consiste à surveiller le système AVI pour détecter le moment où cette intensité de pixel est atteinte. La procédure complète est illustrée dans la figure 8.9.

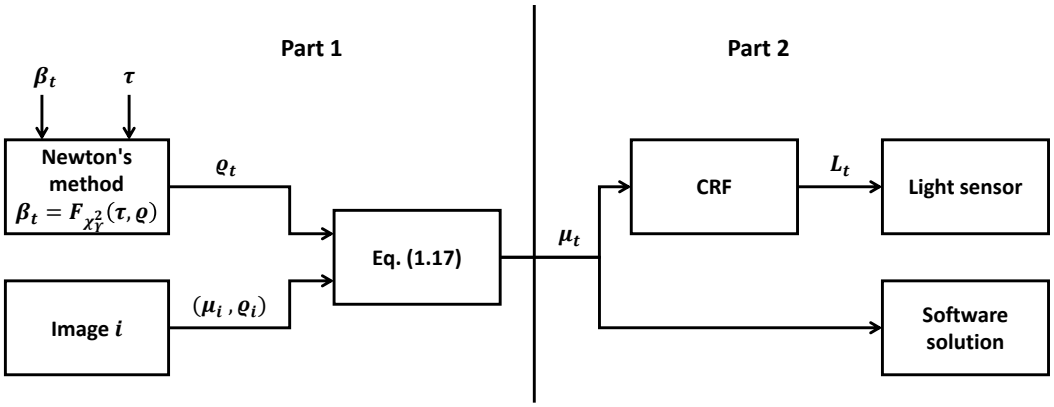


Figure 8.9: Procédure proposée pour détecter valeur minimale acceptable d'éclairage

La première étape consiste à définir la puissance minimale acceptable $\boldsymbol{\beta_t}$ pour une probabilité de fausse alarme constante $\boldsymbol{\alpha_0}$. En utilisant l'équation (8.15), le seuil de détection $\boldsymbol{\tau}$ du test peut être déterminé, car il ne dépend que de la probabilité de fausse alarme prédéfinie $\boldsymbol{\alpha_0}$.
L'étape suivante consiste à calculer le paramètre de non-centralité $\boldsymbol{\varrho_t}$ qui correspond à la valeur $\boldsymbol{\beta_t}$, avec le seuil de détection $\boldsymbol{\tau}$. Pour calculer $\boldsymbol{\varrho_t}$, il est proposé d'utiliser l'équation (8.16), et d'appliquer la méthode de Newton [196] pour résoudre l'équation suivante :

$$\boldsymbol{\beta_t} = F_{\chi^2_\Upsilon}(\boldsymbol{\tau}, \boldsymbol{\varrho_t}) \tag{8.27}$$

La valeur résultante de $\boldsymbol{\varrho_t}$ représente la valeur du paramètre de non-centralité causée par la présence de la plus petite anomalie destinée à être détectée avec un niveau d'éclairage minimal acceptable.
En effet, lors de cette procédure, il est supposé que le niveau d'éclairage du système n'a pas encore atteint sa valeur minimale acceptable, mais est supérieur. Ainsi, si nous notons $\boldsymbol{\mu_i}$ l'intensité du pixel de l'arrière-plan au niveau d'éclairage actuel, et $\boldsymbol{\mu_t}$ l'intensité du pixel au niveau d'éclairage minimal acceptable, nous avons $\boldsymbol{\mu_i} > \boldsymbol{\mu_t}$.
Puisque $\boldsymbol{\mu_i}$ est connu, le but ici est de calculer le taux de diminution en intensité de pixel qui permettra de déterminer $\boldsymbol{\mu_t}$.
Pour cela, il est proposé d'utiliser l'équation (8.13) pour déterminer la valeur du paramètre de non-centralité $\boldsymbol{\varrho_i}$, qui correspond à la même anomalie pour laquelle le paramètre de non-centralité $\boldsymbol{\varrho_t}$ est déterminé. Alors, ayant $\boldsymbol{\mu_i}$ et $boldsymbol\varrho_i$ pour l'image actuelle, et $\boldsymbol{\varrho_t}$ au niveau d'éclairage minimal acceptable, il est possible de déterminer la valeur de $\boldsymbol{\mu_t}$ comme suit.

Considérons que $\boldsymbol{\varrho_t} = R\boldsymbol{\varrho_i}$ et que $\boldsymbol{\mu_t} = K\boldsymbol{\mu_i}$, où R est une constante connue et K une constante inconnue. En raison de la linéarité de la CRF, et pour une même anomalie dans la scène, on a $\boldsymbol{\theta_t} = K\boldsymbol{\theta_i}$. Ensuite, nous obtenons :

$$\left\|\widetilde{\boldsymbol{\Sigma}}_2^{-1/2}\mathbf{P}_{\mathbf{H}_2}^{\perp}\boldsymbol{\theta_2}\right\|_2^2 = R\left\|\widetilde{\boldsymbol{\Sigma}}_1^{-1/2}\mathbf{P}_{\mathbf{H}_1}^{\perp}\boldsymbol{\theta_1}\right\|_2^2 \tag{8.28}$$

En remplaçant $\boldsymbol{\theta_t}$ par $K\boldsymbol{\theta_i}$ et $\boldsymbol{\mu_t}$ par $K\boldsymbol{\mu_i}$, et avec un petit calcul, on obtient :

$$K = \frac{Ra\boldsymbol{\mu_i} + \sqrt{(Ra\boldsymbol{\mu_i})^2 + 4Rb\,(a\boldsymbol{\mu_i} + b)}}{2\,(a\boldsymbol{\mu_i} + b)} \tag{8.29}$$

où (a, b) sont les paramètres du modèle hétéroscédastique du bruit.

Enfin, la seconde partie de la procédure consiste à surveiller le système AVI pour détecter le moment où cette intensité de pixel $\boldsymbol{\mu_t}$ est atteinte. Cette intensité de pixel est liée à la luminosité de la scène à travers la fonction de réponse de la caméra (CRF). Ainsi, dans ce travail, il est proposé d'installer un capteur de lumière sous le système d'éclairage. Ce capteur aidera à surveiller le changement du niveau d'éclairage dans le temps afin de signaler le moment où il atteint sa valeur minimale acceptable, notée $\boldsymbol{L_t}$.

8.5.2 Simulations

Pour effectuer les simulations, il a été proposé de simuler les effets de la dégradation de l'éclairage dans le temps en contrôlant l'alimentation des LED. Notez que cette procédure ne vise pas à simuler le modèle de dégradation des LED, mais plutôt à acquérir des images à différents niveaux d'illumination. Par conséquent, l'alimentation des LED a été réduite progressivement pour créer 60 niveaux d'éclairage différents. Pour chaque niveau, 500 images de la même roue ont été acquises. En conséquence, la nouvelle base de données se compose de 60 séries d'images, contenant chacune 500 images acquises à un niveau d'éclairage unique, pour avoir un total de 30 000 images. Ensuite, un défaut a été superposé sur chaque image qui représente le plus petit défaut destiné à être détecté. L'amplitude de ce défaut a été ajustée pour chaque niveau d'éclairage pour respecter la linéarité du CRF.
La figure 8.10 illustre les résultats de la simulation. Premièrement, les courbes bleue et rouge représentent respectivement la puissance de détection empirique et théorique en fonction de l'intensité moyenne des pixels de l'arrière-plan $\overline{\boldsymbol{\mu}}$. On peut voir que la puissance de détection est la plus élevée pour les images avec l'éclairage le plus élevé, puis commence à diminuer avec la dégradation de l'éclairage.

Ensuite, les courbes noire et verte représentent les valeurs estimées de $\overline{\boldsymbol{\mu}}$ pour différentes valeurs de $\boldsymbol{\beta}$ en utilisant la procédure décrite dans la section 8.5.1. Pour la courbe noire, l'intensité de l'arrière-plan $\overline{\boldsymbol{\mu}}$ est calculée en utilisant l'équation (8.29) étant donné seulement l'intensité de l'arrière-plan de la première image avec le rapport de changement dans le paramètre de non-centralité. Alors que pour la courbe verte, un scénario plus réaliste est considéré où l'image qui est utilisée pour estimer

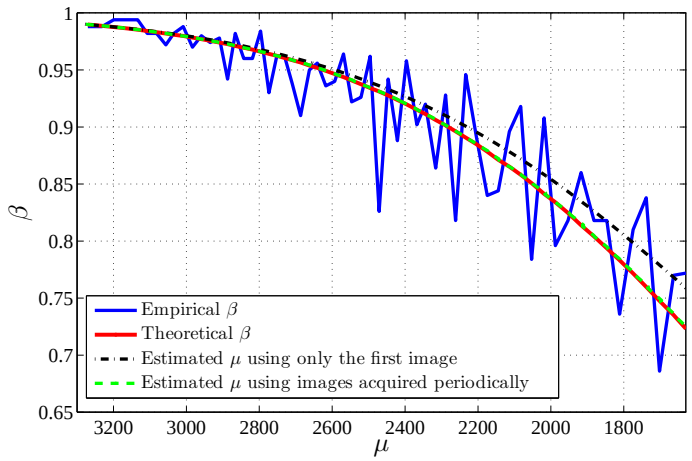


Figure 8.10: Puissance de détection empirique et théorique, avec les estimations de l'intensité des pixels à l'aide du modèle (8.29) en utilisant la première image ou des images reçues périodiquement.

la valeur de $\overline{\boldsymbol{\mu}}$ est mise à jour périodiquement au cours du temps. On peut voir que dans les deux cas, pour des niveaux d'éclairage élevés, l'estimation de $\overline{\boldsymbol{\mu}}$ est très précise par rapport à la courbe de puissance théorique (courbe rouge). Cependant, avec la dégradation de l'éclairage, les estimations obtenues à partir d'une seule image commencent à s'éloigner des valeurs théoriques de $\overline{\boldsymbol{\mu}}$ tandis que celles obtenues à partir d'images acquises périodiquement maintiennent leur précision.
En conclusion, ces simulations montrent que la mise à jour périodique de l'estimation du niveau d'éclairage minimal acceptable dans le temps améliore sa précision.

8.6 Conclusion

L'objectif principal de cette thèse est de concevoir un système entièrement automatique pour l'inspection de la surface des roues. Dans un premier temps, cette thèse a abordé le problème de la conception et de l'installation d'un système de vision compte tenu d'un ensemble d'exigences et de contraintes. La première étape consiste à définir les exigences et les conditions dans lesquelles l'inspection doit être effectuée. Certaines de ces exigences sont liées à la roue inspectée, tandis que d'autres sont imposées par le cadre industriel. Ensuite, en fonction des exigences d'inspection prédéfinies, le choix approprié de chaque élément clé du système AVI est effectué. Ces éléments clés sont les caméras, les objectifs, le système d'éclairage et la plateforme de traitement. Enfin, une description complète de la conception du système AVI et de son installation est présentée. Ce système contient quatre caméras installées sur la ligne de production de roues. Ces caméras sont réparties de manière hémisphérique autour de la roue inspectée permettant une inspection complète de sa surface.

Ensuite, la première tâche d'inspection abordée dans cette thèse est la procédure

de détection des défauts. L'approche proposée consiste à modéliser l'arrière-plan de la surface inspectée à l'aide d'un modèle paramétrique. Un modèle linéaire adaptatif original est proposé composé de deux parties : une partie fixe basée sur un polynôme algébrique bidimensionnel qui est principalement efficace pour représenter avec précision des surfaces homogènes ou des surfaces lisses avec peu ou sans texture. Ensuite, une partie adaptative basée sur les composantes principales qui offre une plus grande flexibilité et améliore les performances du modèle pour gérer des complexités potentielles de surface. Par conséquent, ce modèle peut être utilisé pour l'inspection d'une variété de surfaces qui ont principalement une faible texture. D'un point de vue statistique, cet arrière-plan est considéré comme un paramètre de nuisance qui doit être supprimé. En effet, il n'a aucun intérêt dans la détection et par conséquent, empêcher leur détection. Par conséquent, un test statistique basé sur le rejet de ce paramètre de nuisance est proposé. Ce test exploite la linéarité du modèle adaptatif pour effectuer le rejet à travers une projection linéaire. Les résidus restants sont ensuite normalisés en utilisant le bruit hétéroscédastique spécifique du système d'imagerie. Ce modèle de bruit est bien connu pour être plus précis que le modèle de bruit blanc gaussien habituel considéré pour les images raw et permet de prendre en compte la variance de chaque pixel dans le test statistique. Les performances statistiques du test sont ensuite établies. Plus précisément, le test est UPPC avec une puissance de détection maximale et constante sur une surface locale définie par le rapport "défaut sur bruit".
La méthode de détection proposée est ensuite appliquée au problème d'inspection de la surface de la roue étudié dans ce travail. En raison de la nature des roues, les différents éléments sont analysés séparément. Plusieurs simulations sur un grand ensemble de données d'images réelles ont été effectuées. Ces simulations montrent à la fois la précision du modèle adaptatif proposé et la netteté du test statistique qui s'ensuit. Enfin, certains résultats sont présentés en appliquant le modèle adaptatif proposé sur des images de roues contenant des défauts réels situés sur la surface de la roue.

La deuxième tâche d'inspection abordée dans cette thèse vise à surveiller en ligne un processus non stationnaire pour détecter des changements brusques dans sa valeur moyenne. Dans notre cas, ce processus non stationnaire résulte de la variation de la quantité de peinture à la surface des roues inspectées où le changement brusque correspond à un manque brusque de peinture. L'approche proposée consiste à modéliser les changements acceptables dans le processus non stationnaire pour pouvoir les distinguer des changements brusques qui révèlent un dysfonctionnement. Un modèle polynomial est considéré et appliqué sur une fenêtre de longueur fixe contenant l'ensemble des dernières observations acquises. En effet, cette variation acceptable de la valeur moyenne, modélisée à l'aide du modèle polynomial est un paramètre de nuisance, car elle est inutile pour le problème de détection considéré. Par conséquent, la méthode d'estimation du maximum de vraisemblance (ML) a été considérée pour effectuer un rejet de ce paramètre de nuisance à travers une projection linéaire. Ensuite, une méthode séquentielle avec une fenêtre de longueur fixe est proposée. Cette seconde fenêtre correspond également à un délai de détection

fixe prédéfini sous lequel la détection est effectuée. En résumé, la méthode séquentielle proposée est basée sur deux fenêtres de longueur fixe, où la première fenêtre est utilisée pour estimer les paramètres du processus non stationnaires tandis que la seconde fenêtre est utilisée pour exécuter la détection. Enfin, les performances de la méthode proposée sont étudiées. Cette étude montre que la performance statistique du test proposé est bornée. La probabilité de fausse alarme possède une borne supérieure, ce qui permet de calculer un seuil de détection en utilisant un taux de fausse alarme prédéfini sachant que l'application est garantie pour ne pas le dépasser. D'autre part, la puissance de détection du test possède une borne inférieure, ce qui permet de garantir, pour un taux de fausse alarme prédéfini, une puissance de détection minimale que l'application ne va pas diminuer au-dessous.
La méthode séquentielle proposée est ensuite appliquée au problème de surveillance du revêtement des roues étudié dans ce travail. La valeur moyenne des pixels de toutes les images de roues est utilisée pour mesurer l'intensité du revêtement. Plusieurs résultats numériques sur un grand nombre d'images sont présentés. Ces résultats montrent la précision du modèle proposé et l'efficacité de la méthode de détection proposée. De plus, l'effet de la seconde fenêtre sur les performances du test proposé est étudié. Il a été montré que, selon l'application, le test proposé permet soit d'augmenter les performances de détection au prix d'un délai de détection plus important, soit de diminuer le délai de détection au prix d'une performance de détection plus faible.

Le dernier problème abordé dans cette thèse est la prédiction des besoins de maintenance du système AVI en raison de la dégradation des LED dans le temps. L'approche proposée consiste à étudier l'impact de la dégradation de l'éclairage sur la performance de la méthode de détection. À partir d'une puissance de détection minimale acceptable prédéfinie, la procédure complète pour déterminer le niveau d'éclairage minimal acceptable est décrite. Étant donné le même défaut sur la surface inspectée, cette procédure utilise un modèle qui permet de déterminer le taux de changement en intensité de pixel fonction du taux de changement dans le paramètre de non-centralité. L'étape finale consiste à surveiller la variation du niveau d'éclairage dans le temps à l'aide d'un capteur de lumière afin de signaler le moment où il atteint sa valeur minimale acceptable.
Des résultats numériques sur un grand nombre d'images réelles montrent la précision du modèle calculé pour déterminer le taux de changement en intensité de pixel. On voit que cette précision est améliorée si le taux de changement est faible. Il est alors conclu que la mise à jour périodique de l'estimation du niveau d'éclairage minimal acceptable dans le temps améliore sa précision.

Suite à cette brève synthèse des travaux présentés dans cette thèse, il est proposé de présenter quelques perspectives qui pourraient enrichir la discussion sur certains sujets et offrir de nouvelles pistes de recherche potentielles :

- Toutes les simulations présentées dans ce travail ont utilisé uniquement des images acquises à partir de la caméra directement au-dessus de la roue inspectée.

Il est alors prévu d'utiliser les caméras latérales pour améliorer l'inspection de toute la surface de la roue. En effet, certaines zones de la surface de la roue, telles que la zone contenant les trous de ventilation, sont mieux éclairées avec les caméras latérales, ce qui améliorerait la détectabilité des défauts situés sur ces zones comme discuté dans la section 8.5.

- Un modèle polynomial, accompagné des composantes principales, a été considéré pour modéliser l'arrière-plan de l'image inspectée. Ce modèle a prouvé son efficacité pour l'inspection de la surface des roues. Il serait intéressant de tester d'autres types de modèles paramétriques et d'étudier la détectabilité des défauts dans de tels cas.
- Le test statistique conçu dans la section 8.3 pour la détection des défauts pourrait être étendu pour prendre en compte plusieurs paramètres. Tout d'abord, la variance de l'estimation de l'espérance de pixels, et en second lieu, l'impact de la présence du défaut sur la variance. Cela pourrait potentiellement améliorer la précision de la détection, et fournir un meilleur contrôle du taux de fausses alarmes.
- Dans la section 8.4, la modélisation du processus non stationnaire sur la première fenêtre a été réalisée à l'aide d'un modèle polynomial. Remplacer ce modèle par un autre type de modèle paramétrique est réalisable et ne modifiera pas la procédure de détection effectuée sur la seconde fenêtre. Par conséquent, une autre perspective intéressante est de tester la précision d'autres types de modèles paramétriques pour modéliser le processus non stationnaire et de comparer sa précision avec celle du modèle polynomial proposé.
- Une dernière perspective est liée au problème de la surveillance de la valeur moyenne de l'arrière-plan présenté dans la section 8.5. Ce problème consiste à surveiller en ligne la valeur moyenne d'une séquence aléatoire gaussienne pour détecter le moment où elle dépasse une valeur prescrite. Le principal défi pour concevoir un test statistique dans notre cas est que la variance est une fonction affine de la valeur moyenne.

Bibliography

[1] M. E. H. Creusen and J. P. L. Schoormans, "The different roles of product appearance in consumer choice*," *Journal of Product Innovation Management*, vol. 22, no. 1, pp. 63–81, 2005. [Online]. Available: http://dx.doi.org/10.1111/j.0737-6782.2005.00103.x (Cited on page 1.)

[2] M. Holweg, *The Evolution of Competition in the Automotive Industry*. London: Springer London, 2008, pp. 13–34. [Online]. Available: http://dx.doi.org/10.1007/978-1-84800-225-8_2 (Cited on page 7.)

[3] R. Bingham and Z. Zhang, *The Economies Of Central City Neighborhoods*. Avalon Publishing, 2001. [Online]. Available: https://books.google.fr/books?id=T3RPAAAAMAAJ (Cited on page 7.)

[4] X. Bai, "The effects of the 2007-2009 economic crisis on global automobile industry," Ph.D. dissertation, Applied Economics Thesis at Department of Economics and Finance, State University of New York College at Buffalo. (Cited on page 7.)

[5] J. Juran and A. Godfrey, *Juran's Quality Handbook*, ser. JURAN'S QUALITY HANDBOOK. McGraw Hill, 1999. [Online]. Available: https://books.google.fr/books?id=beVTAAAAMAAJ (Cited on page 8.)

[6] "Wheels and rims for pneumatic tyres – Vocabulary, designation and marking," International Organization for Standardization, Standard, Nov. 2004. (Cited on page 8.)

[7] "Quality management systems – Fundamentals and vocabulary," International Organization for Standardization, Standard, Sep. 2015. (Cited on page 10.)

[8] D. Rollinson, *Organisational Behaviour and Analysis: An Integrated Approach*. FT Prentice Hall, 2008. [Online]. Available: https://books.google.fr/books?id=BzJtQr3Rdp0C (Cited on page 11.)

[9] A. Feigenbaum, *Total Quality Control*, ser. Industrial engineering series. McGraw-Hill, 1991. [Online]. Available: https://books.google.fr/books?id=B9FTAAAAMAAJ (Cited on page 11.)

[10] M. Curtis and F. Farago, *Handbook of Dimensional Measurement*. Industrial Press, 2007. [Online]. Available: https://books.google.fr/books?id=ssTZAAAAMAAJ (Cited on page 12.)

[11] "Road vehicles – Passenger car wheels for road use – Test methods," International Organization for Standardization, Standard, Jun. 2015. (Cited on page 12.)

[12] P. R. RAJU, B. SATYANARAYANA, K. RAMJI, and K. S. BABU, "Evaluation of fatigue life of aluminium alloy wheels under bending loads," *Fatigue and Fracture of Engineering Materials and Structures*, vol. 32, no. 2, pp. 119–126, 2009. [Online]. Available: http://dx.doi.org/10.1111/j.1460-2695.2008.01316.x (Cited on page 12.)

[13] L. Wang, Y. Chen, C. Wang, and Q. Wang, "Fatigue life analysis of aluminum wheels by simulation of rotary fatigue test," *Strojni?ki vestnik - Journal of Mechanical Engineering*, vol. 57, no. 1, 2011. (Cited on page 12.)

[14] D. Stamatis, *Failure Mode and Effect Analysis: FMEA from Theory to Execution*. ASQ Quality Press, 2003. [Online]. Available: https://books.google.fr/books?id=TTxI8jbTkVwC (Cited on page 14.)

[15] A. S. Guerra, M. Pillet, and J.-L. Maire, "Control of variability for man measurement," in *12th IMEKO TC1-TC7 joint Symposium on Man, Science and Measurement*, Annecy, France, Sep. 2008, p. nc. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00344007 (Cited on page 17.)

[16] J.-L. Maire, M. Pillet, and N. Baudet, "Gage r2&e2: an effective tool to improve the visual control of products," *International Journal of Quality & Reliability Management*, vol. 30, no. 2, pp. 161–176, 2013. [Online]. Available: https://doi.org/10.1108/02656711311293571 (Cited on page 17.)

[17] P. Kopardekar, A. Mital, and S. Anand, "Manual, hybrid and automated inspection literature and current research," *Integrated Manufacturing Systems*, vol. 4, no. 1, pp. 18–29, 1993. [Online]. Available: https://doi.org/10.1108/09576069310023838 (Cited on pages 17 and 27.)

[18] G. Workman, P. Moore, and A. S. for Nondestructive Testing, *Nondestructive Testing Overview*, ser. Nondestructive testing handbook. American Society for Nondestructive Testing, 2012. [Online]. Available: https://books.google.fr/books?id=eBRrMAEACAAJ (Cited on page 17.)

[19] C. Hellier, *Handbook of Nondestructive Evaluation, Second Edition*, ser. Mechanical Engineering. McGraw-Hill Education, 2012. [Online]. Available: https://books.google.fr/books?id=TjKNY_gCmIkC (Cited on pages 17, 24 and 26.)

[20] J. DUMONT-FILLON, "Contrôle non destructif (cnd)," *Techniques de l'ingénieur CND : méthodes globales et volumiques*, vol. base documentaire : TIB585DUO., no. ref. article : r1400, 2017, fre. [Online]. Available: http://www.techniques-ingenieur.fr/base-documentaire/mesures-analyses-th1/cnd-methodes-globales-et-volumiques-42585210/controle-non-destructif-cnd-r1400/ (Cited on page 18.)

[21] "Non-destructive testing: A guidebook for industrial management and quality control personnel," International Atomic Energy Agency (IAEA), Training Course, Jan. 1999. (Cited on page 18.)

[22] B. Mehta, "Surface and subsurface crack analysis in carbon steel samples using magnetic particle testing and liquid penetrant testing," *Materials Evaluation*, vol. 73, no. 4, 2015. (Cited on page 20.)

[23] H. Fischer, F. Karaca, and R. Marx, "Detection of microscopic cracks in dental ceramic materials by fluorescent penetrant method," *Journal of Biomedical Materials Research*, vol. 61, no. 1, pp. 153–158, 2002. [Online]. Available: http://dx.doi.org/10.1002/jbm.10148 (Cited on page 20.)

[24] S. Ranganayakulu, A. Kucheludu, R. Gowtham, and B. Ramesh Kumar, "Inspection on aluminum plates by implementation of ndt techniques," *J. Pure Appl. Ultrson*, vol. 37, pp. 57–61, 2015. (Cited on page 20.)

[25] K. Imieli?ska, M. Castaings, R. Wojtyra, J. Haras, E. L. Clezio, and B. Hosten, "Air-coupled ultrasonic c-scan technique in impact response testing of carbon fibre and hybrid: glass, carbon and kevlar/epoxy composites," *Journal of Materials Processing Technology*, vol. 157, pp. 513 – 522, 2004, achievements in Mechanical and Materials Engineering Conference. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924013604011355 (Cited on page 20.)

[26] F. Awaja, S. Zhang, M. Tripathi, A. Nikiforov, and N. Pugno, "Cracks, microcracks and fracture in polymer structures: Formation, detection, autonomic repair," *Progress in Materials Science*, vol. 83, pp. 536 – 573, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0079642516300366 (Cited on page 20.)

[27] R. Gardner and H. Pincus, "Fluorescent dye penetrants applied to rock fractures," *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, vol. 5, no. 2, pp. 155 – 156, 1968. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0148906268900314 (Cited on page 20.)

[28] K. Abend, "Fully automated dye-penetrant inspection of automotive parts," *Computer Standards and Interfaces*, vol. 21, no. 2, p. 157, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092054899992144X (Cited on page 20.)

[29] X. Guirong, G. Xuesong, Q. Yuliang, and G. Yan, "Analysis and innovation for penetrant testing for airplane parts," *Procedia Engineering*, vol. 99, pp. 1438 – 1442, 2015, 2014 Asia-Pacific International Symposium on Aerospace Technology, APISAT2014 September 24-26, 2014 Shanghai,

China. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877705814037989 (Cited on page 20.)

[30] J. Zheng, W. F. Xie, M. Viens, L. Birglen, and I. Mantegh, "Design of an advanced automatic inspection system for aircraft parts based on fluorescent penetrant inspection analysis," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 57, no. 1, pp. 18–34, 2015. [Online]. Available: http://www.ingentaconnect.com/content/bindt/insight/2015/00000057/00000001/art00005 (Cited on page 20.)

[31] T. Vetterlein, M. Wagener, H. Rongen, and C. Sampson, "Automated dye penetrant systems with process control and documentation in the aerospace industry," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 48, no. 3, pp. 171–173, 2006. [Online]. Available: http://www.ingentaconnect.com/content/bindt/insight/2006/00000048/00000003/art00009 (Cited on pages 20 and 22.)

[32] N. Tracy, P. Moore, and A. S. for Nondestructive Testing, *Liquid Penetrant Testing*, ser. Nondestructive testing handbook. American Society for Nondestructive Testing, 1999. [Online]. Available: https://books.google.fr/books?id=DeoEAAAACAAJ (Cited on page 20.)

[33] M. Schröder, C. Biedermann, and R. Vilbrandt, "On the applicability of dye penetrant tests on vacuum components: Allowed or forbidden?" *Fusion Engineering and Design*, vol. 88, no. 9, pp. 1947 – 1950, 2013, proceedings of the 27th Symposium On Fusion Technology (SOFT-27); Liège, Belgium, September 24-28, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0920379612005352 (Cited on page 22.)

[34] N. P. Migoun and N. V. Delenkovsky, "The ways of penetrant testing applicability for rough surfaces," in *Proceedings of 17th World Conf. on NDT, Shanghai, China*, 2008, pp. 25–28. (Cited on page 22.)

[35] D. Popescu, F. Anania, C. Cotet, and C. Amza, "Fully automated liquid penetrant inspection line simulation model for increasing productivity," *International Journal Simulation Modelling*, vol. 12, no. 2, pp. 82–93, 2013. (Cited on page 22.)

[36] B. L. Luk and A. H. S. Chan, "Ergonomics, safety and health issues in magnetic particles inspection and dye penetrant inspection," in *IMECS*, 2007. (Cited on page 22.)

[37] T. VETTERLEIN, "Application of magnetic particle inspection in the field of the automotive industry," in *Abstracts of 17th World Conference on Non-Destructive Testing*, 2008. (Cited on pages 23 and 24.)

[38] M. Maass, W. A. K. Deutsch, and F. Bartholomai, "State of the art mt and ut test stations in the german automotive industry," in *Fall Conference & Quality Testing Show 2012*, 2012, pp. 2–6. (Cited on pages 23 and 24.)

[39] U. Ewert, G.-R. Jaenisch, K. Osterloh, U. Zscherpel, C. Bathias, M. Hentschel, A. Erhard, J. Goebbels, H. Hanselka, J. Nuffer, and W. Daum, *Performance Control and Condition Monitoring*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 831–912. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-30300-8_16 (Cited on page 23.)

[40] T. Nishimine, O. Tsuyama, T. Tanaka, and H. Fujiwara, "Automatic magnetic particle testing system for square billets," in *Industry Applications Conference, 1995. Thirtieth IAS Annual Meeting, IAS '95., Conference Record of the 1995 IEEE*, vol. 2, Oct 1995, pp. 1585–1590 vol.2. (Cited on page 24.)

[41] J. Luo, Z. Tian, and J. Yang, "Fluorescent magnetic particle inspection device based on digital image processing," in *Proceeding of the 11th World Congress on Intelligent Control and Automation*, June 2014, pp. 5677–5681. (Cited on page 24.)

[42] D. Lovejoy, *Health and safety in magnetic particle inspection*. Dordrecht: Springer Netherlands, 1993, pp. 409–425. [Online]. Available: http://dx.doi.org/10.1007/978-94-011-1536-0_15 (Cited on page 24.)

[43] ——, *Environmental impact of magnetic particle inspection*. Dordrecht: Springer Netherlands, 1993, pp. 427–432. [Online]. Available: http://dx.doi.org/10.1007/978-94-011-1536-0_16 (Cited on page 24.)

[44] W. C. Roentgen, "Ueber eine neue art von strahlung., aus den sitzungsberichten der wuerzburger physik," *Medic. Gesellschaft, Wuerzburg*, 1895. (Cited on page 24.)

[45] R. Bossi, F. Iddings, and G. Wheeler, *Radiographic testing*, ser. Nondestructive testing handbook. American Society for Nondestructive Testing, 2002. [Online]. Available: https://books.google.fr/books?id=0MLZAAAAMAAJ (Cited on page 24.)

[46] H. Boerner and H. Strecker, "Automated x-ray inspection of aluminum castings," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 1, pp. 79–91, Jan 1988. (Cited on page 25.)

[47] L. Pizarro, D. Mery, R. Delpiano, and M. Carrasco, "Robust automated multiple view inspection," *Pattern Analysis and Applications*, vol. 11, no. 1, pp. 21–32, Jan 2008. [Online]. Available: http://dx.doi.org/10.1007/s10044-007-0075-9 (Cited on page 25.)

[48] M. Carrasco and D. Mery, "Automatic multiple view inspection using geometrical tracking and feature analysis in aluminum wheels," *Machine*

Vision and Applications, vol. 22, no. 1, pp. 157–170, Jan 2011. [Online]. Available: http://dx.doi.org/10.1007/s00138-010-0255-2 (Cited on page 25.)

[49] V. D. Nguyen, A. Noble, J. Mundy, J. Janning, and J. Ross, "Exhaustive detection of manufacturing flaws as abnormalities," in *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)*, Jun 1998, pp. 945–952. (Cited on page 25.)

[50] J. Shao, H. Shi, D. Du, L. Wang, and H. Cao, "Automatic weld defect detection in real-time x-ray images based on support vector machine," in *2011 4th International Congress on Image and Signal Processing*, vol. 4, Oct 2011, pp. 1842–1846. (Cited on page 25.)

[51] M. Malarvel, G. Sethumadhavan, P. C. R. Bhagi, S. Kar, T. Saravanan, and A. Krishnan, "Anisotropic diffusion based denoising on x-radiography images to detect weld defects," *Digital Signal Processing*, vol. 68, pp. 112 – 126, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200417301148 (Cited on page 25.)

[52] G. Chen, J. Turner, D. Nisius, K. Holt, and A. Brooks, "Linatron mi6, the x-ray source for cargo inspection," *Physics Procedia*, vol. 66, pp. 68 – 74, 2015, the 23rd International Conference on the Application of Accelerators in Research and Industry - CAARI 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1875389215001649 (Cited on page 25.)

[53] M. Abdolshah, M. Teimouri, and R. Rahmani, "Classification of x-ray images of shipping containers," *Expert Systems with Applications*, vol. 77, pp. 57 – 65, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417417300362 (Cited on page 25.)

[54] A. Mazoochi, F. Rahmani, F. A. Davani, and R. Ghaderi, "A novel numerical method to eliminate thickness effect in dual energy x-ray imaging used in baggage inspection," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 763, pp. 538 – 542, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900214008328 (Cited on page 25.)

[55] Y. Wang, M. Wang, and Z. Zhang, "Microfocus x-ray printed circuit board inspection system," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 17, pp. 4929 – 4931, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030402614004756 (Cited on page 25.)

[56] D. Mery, "X-ray testing by computer vision," in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2013, pp. 360–367. (Cited on page 25.)

[57] J. Rowlands, "The physics of computed radiography," *Physics in medicine and biology*, vol. 47, no. 23, p. R123, 2002. (Cited on page 25.)

[58] D. Mery, *X-ray Image Processing*. Cham: Springer International Publishing, 2015, pp. 109–148. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-20747-6_4 (Cited on page 25.)

[59] G. Theis and T. Kahrs, "Fully automatic x-ray inspection of aluminium wheels," in *8th European Conference on Non-Destructive Testing (ECNDT 2002)*, 2002. (Cited on page 25.)

[60] F. Herold, S. Frantz, K. Bavendiek, and R. Grigat, "Building blocks of third-generation automatic defect recognition system," in *9th European Conference on NDT*, 2006. (Cited on page 25.)

[61] F. HEROLD, "Automatic wheel inspection using building blocks," in *9th European Conference on NDT*, 2006. (Cited on page 25.)

[62] P. Moore and A. S. for Nondestructive Testing, *Visual Testing*, ser. Nondestructive testing handbook. American Society for Nondestructive Testing, 2010. (Cited on page 26.)

[63] P. Kopardekar, A. Mital, and S. Anand, "Manual, hybrid and automated inspection literature and current research," *Integrated Manufacturing Systems*, vol. 4, no. 1, pp. 18–29, 1993. (Cited on page 27.)

[64] B. G. Batchelor, "Selecting cameras for machine vision," in *Machine Vision Handbook*. London: Springer London, 2012, pp. 477–506. [Online]. Available: https://doi.org/10.1007/978-1-84996-169-1_11 (Cited on pages 27 and 43.)

[65] J. L. C. Sanz and D. Petkovic, "Machine vision algorithms for automated inspection thin-film disk heads," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 830–848, Nov 1988. (Cited on page 27.)

[66] B.-J. You, Y. S. Oh, and Z. Bien, "A vision system for an automatic assembly machine of electronic components," *IEEE Transactions on Industrial Electronics*, vol. 37, no. 5, pp. 349–357, Oct 1990. (Cited on page 27.)

[67] K. Hanbay, M. F. Talu, and Ömer Faruk Özgüven, "Fabric defect detection systems and methods – a systematic literature review," *Optik - International Journal for Light and Electron Optics*, vol. 127, no. 24, pp. 11 960 – 11 973, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030402616311366 (Cited on page 27.)

[68] G. ElMasry, S. Cubero, E. Moltó, and J. Blasco, "In-line sorting of irregular potatoes by using automated computer-based machine vision system," *Journal of Food Engineering*, vol. 112, no. 1, pp. 60 – 68, 2012. [Online]. Available:

http://www.sciencedirect.com/science/article/pii/S0260877412001690 (Cited on page 27.)

[69] M. Sofu, O. Er, M. Kayacan, and B. Ceti?li, "Design of an automatic apple sorting system using machine vision," *Computers and Electronics in Agriculture*, vol. 127, pp. 395 – 405, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169916304513 (Cited on page 27.)

[70] E. Davies, "6 - computer vision for automatic sorting in the food industry," in *Computer Vision Technology in the Food and Beverage Industries*, ser. Woodhead Publishing Series in Food Science, Technology and Nutrition, D.-W. Sun, Ed. Woodhead Publishing, 2012, pp. 150 – 180. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780857090362500065 (Cited on page 27.)

[71] P. Nerakae, P. Uangpairoj, and K. Chamniprasart, "Using machine vision for flexible automatic assembly system," *Procedia Computer Science*, vol. 96, pp. 428 – 435, 2016, knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 20th International Conference KES-2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050916318798 (Cited on page 27.)

[72] W.-Y. Wu, M.-J. J. Wang, and C.-M. Liu, "Automated inspection of printed circuit boards through machine vision," *Computers in Industry*, vol. 28, no. 2, pp. 103 – 111, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0166361595000631 (Cited on page 27.)

[73] H. Golnabi and A. Asadpour, "Design and application of industrial machine vision systems," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 6, pp. 630 – 637, 2007, 16th International Conference on Flexible Automation and Intelligent Manufacturing. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0736584507000233 (Cited on page 28.)

[74] E. N. Malamas, E. G. Petrakis, M. Zervakis, L. Petit, and J.-D. Legat, "A survey on industrial vision systems, applications and tools," *Image and Vision Computing*, vol. 21, no. 2, pp. 171 – 188, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S026288560200152X (Cited on page 28.)

[75] M. Moganti, F. Ercal, C. H. Dagli, and S. Tsunekawa, "Automatic pcb inspection algorithms: A survey," *Computer Vision and Image Understanding*, vol. 63, no. 2, pp. 287 – 313, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S107731429690020X (Cited on page 28.)

[76] A. Jiménez, A. Jain, R. Ceres, and J. Pons, "Automatic fruit recognition: a survey and new results using range/attenuation images," *Pattern Recognition*, vol. 32, no. 10, pp. 1719 – 1736, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320398001708 (Cited on page 28.)

[77] B. Zhang, W. Huang, J. Li, C. Zhao, S. Fan, J. Wu, and C. Liu, "Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review," *Food Research International*, vol. 62, no. Supplement C, pp. 326 – 343, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0963996914001707 (Cited on page 29.)

[78] N. Neogi, D. K. Mohanta, and P. K. Dutta, "Review of vision-based steel surface inspection systems," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, p. 50, Nov 2014. [Online]. Available: https://doi.org/10.1186/1687-5281-2014-50 (Cited on page 29.)

[79] Y. Shi and F. D. Real, *Smart Cameras: Fundamentals and Classification*. Boston, MA: Springer US, 2010, pp. 19–34. [Online]. Available: https://doi.org/10.1007/978-1-4419-0953-4_2 (Cited on page 33.)

[80] H. Y. Ngan, G. K. Pang, and N. H. Yung, "Automated fabric defect detection – a review," *Image and Vision Computing*, vol. 29, no. 7, pp. 442–458, 2011. (Cited on page 33.)

[81] K. Hanbay, M. F. Talu, and Ö. F. Özgüven, "Fabric defect detection systems and methods – a systematic literature review," *Optik-International Journal for Light and Electron Optics*, vol. 127, no. 24, pp. 11 960–11 973, 2016. (Cited on page 33.)

[82] L. Norton-Wayne, M. Bradshaw, and A. Jewell, "Machine vision inspection of web textile fabric," in *BMVC92*. Springer, 1992, pp. 217–226. (Cited on page 33.)

[83] M. Bradshaw, "The application of machine vision to the automated inspection of knitted fabrics," *Mechatronics*, vol. 5, no. 2-3, pp. 233–243, 1995. (Cited on page 33.)

[84] C.-S. Cho, B.-M. Chung, and M.-J. Park, "Development of real-time vision-based fabric inspection system," *IEEE Transactions on Industrial Electronics*, vol. 52, no. 4, pp. 1073–1079, 2005. (Cited on page 33.)

[85] Y. Yang, C. Miao, X. Li, and X. Mei, "On-line conveyor belts inspection based on machine vision," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 19, pp. 5803–5807, 2014. (Cited on page 34.)

[86] C. C. Fromme, D. J. Stager, T. E. Pilarski, B. Bancroft, and T. E. Hegadorn, "Conveyor belt inspection system and method," Jan. 24 2006, uS Patent 6,988,610. (Cited on page 34.)

[87] T. Wei, "Research on detection technology for cracks of coal conveyor belt," *North University of China, Taiyuan*, pp. 48–49, 2010. (Cited on page 34.)

[88] N. Neogi, D. K. Mohanta, and P. K. Dutta, "Review of vision-based steel surface inspection systems," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, p. 50, 2014. (Cited on page 34.)

[89] T. S. Nguyen, M. Avila, and S. Begot, "Automatic detection and classification of defect on road pavement using anisotropy measure," in *2009 17th European Signal Processing Conference*, Aug 2009, pp. 617–621. (Cited on page 34.)

[90] S.-N. Yu, J.-H. Jang, and C.-S. Han, "Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel," *Automation in Construction*, vol. 16, no. 3, pp. 255–261, 2007. (Cited on page 34.)

[91] M. Bigas, E. Cabruja, J. Forest, and J. Salvi, "Review of cmos image sensors," *Microelectronics Journal*, vol. 37, no. 5, pp. 433 – 451, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0026269205002764 (Cited on page 37.)

[92] S. Kempainen, "Cmos image sensors: eclipsing ccds in visual information?" *EDN*, vol. 42, no. 21, pp. 101–120, 1997. (Cited on page 37.)

[93] M. Hillebrand, N. Stevanovic, B. J. Hosticka, J. E. S. Conde, A. Teuner, and M. Schwarz, "High speed camera system using a cmos image sensor," in *Proceedings of the IEEE Intelligent Vehicles Symposium 2000 (Cat. No.00TH8511)*, 2000, pp. 656–661. (Cited on page 37.)

[94] N. Instruments, "A practical guide to machine vision lighting," 2017, [Online; accessed 9-October-2017]. [Online]. Available: http://www.ni.com/white-paper/6901/en/ (Cited on pages 50, 51 and 52.)

[95] B. G. Batchelor, "Lighting-viewing methods," in *Machine Vision Handbook*. London: Springer London, 2012, pp. 1345–1560. (Cited on pages 50 and 51.)

[96] K. G. Harding, "Machine vision – lighting." in *Encyclopedia of Optical Engineering*. CRC Press, 2003, pp. 1227–1336. (Cited on page 50.)

[97] B. G. Batchelor, "Lighting-viewing methods," in *Machine Vision Handbook*. London: Springer London, 2012, pp. 319–327. (Cited on page 50.)

[98] ——, "Illumination sources," in *Machine Vision Handbook*. London: Springer London, 2012, pp. 283–317. (Cited on page 53.)

[99] T. Bergin, J. Cusack, and K. DeSmet, "Advantages of led lighting in vision inspection systems," 2010, [Online; accessed 7-November-2017]. [Online]. Available: http://www.quadtechworld.com/downloads/whitepapers/wp022010.pdf (Cited on pages 53 and 54.)

[100] N. Narendran and Y. Gu, "Life of led-based white light sources," *J. Display Technol.*, vol. 1, no. 1, p. 167, Sep 2005. (Cited on pages 54, 122 and 134.)

[101] P. D. Michailidis and K. G. Margaritis, "Scientific computations on multi-core systems using different programming frameworks," *Applied Numerical Mathematics*, vol. 104, no. Supplement C, pp. 62 – 80, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016892741400213X (Cited on page 54.)

[102] M. Birk, M. Zapf, M. Balzer, N. Ruiter, and J. Becker, "A comprehensive comparison of gpu- and fpga-based acceleration of reflection image reconstruction for 3d ultrasound computer tomography," *Journal of Real-Time Image Processing*, vol. 9, no. 1, pp. 159–170, Mar 2014. [Online]. Available: https://doi.org/10.1007/s11554-012-0267-4 (Cited on page 55.)

[103] J. A. Gomez-Pulido, M. A. Vega-Rodriguez, J. M. Sanchez-Perez, S. Priem-Mendes, and V. Carreira, "Accelerating floating-point fitness functions in evolutionary algorithms: a fpga-cpu-gpu performance comparison," *Genetic Programming and Evolvable Machines*, vol. 12, no. 4, pp. 403–427, Dec 2011. [Online]. Available: https://doi.org/10.1007/s10710-011-9137-2 (Cited on page 55.)

[104] S. Kestur, J. D. Davis, and O. Williams, "Blas comparison on fpga, cpu and gpu," in *2010 IEEE Computer Society Annual Symposium on VLSI*, July 2010, pp. 288–293. (Cited on page 55.)

[105] D. Zou, Y. Dou, and F. Xia, "Optimization schemes and performance evaluation of smith?waterman algorithm on cpu, gpu and fpga," *Concurrency and Computation: Practice and Experience*, vol. 24, no. 14, pp. 1625–1644, 2012. [Online]. Available: http://dx.doi.org/10.1002/cpe.1913 (Cited on page 55.)

[106] C. Grozea, Z. Bankovic, and P. Laskov, *FPGA vs. Multi-core CPUs vs. GPUs: Hands-On Experience with a Sorting Application.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 105–117. [Online]. Available: https://doi.org/10.1007/978-3-642-16233-6_12 (Cited on page 55.)

[107] T. Instruments, "Multicore socs stay a step ahead of soc fpgas," 2016, [Online; accessed 11-November-2017]. [Online]. Available: http://www.ti.com/lit/wp/spry296/spry296.pdf (Cited on page 56.)

[108] D. Pascale, "Rgb coordinates of the macbeth colorchecker," *The BabelColor Company*, pp. 1–16, 2006. (Cited on page 58.)

[109] A. Kumar, "Computer-vision-based fabric defect detection: A survey," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 1, pp. 348–363, Jan 2008. (Cited on pages 61 and 148.)

[110] R. Cogranne and F. Retraint, "Statistical detection of defects in radiographic images using an adaptive parametric model," *Signal Processing*, vol. 96, no. Part B, pp. 173 – 189, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168413003599 (Cited on pages 61, 68, 73, 100, 148, 149 and 153.)

[111] D. Mery, T. Jaeger, and D. Filbert, "A review of methods for automated recognition of casting defects," *INSIGHT-WIGSTON THEN NORTHAMPTON-*, vol. 44, no. 7, pp. 428–436, 2002. (Cited on pages 61, 62, 148 and 149.)

[112] N. Neogi, D. K. Mohanta, and P. K. Dutta, "Review of vision-based steel surface inspection systems," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, p. 50, Nov 2014. (Cited on pages 61, 62 and 68.)

[113] H. Barrett and K. Myers, *Foundations of image science*, ser. Wiley series in pure and applied optics. Wiley-Interscience, 2004. [Online]. Available: https://books.google.fr/books?id=pHBTAAAAMAAJ (Cited on page 61.)

[114] Y. Zhang, Z. Lu, and J. Li, *Fabric Defect Detection and Classification Using Gabor Filters and Gaussian Mixture Model.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 635–644. (Cited on page 62.)

[115] W. bin Li, C. hou Lu, and J. chuan Zhang, "A local annular contrast based real-time inspection algorithm for steel bar surface defects," *Applied Surface Science*, vol. 258, no. 16, pp. 6080 – 6086, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169433212004217 (Cited on page 62.)

[116] Alaknanda, R. Anand, and P. Kumar, "Flaw detection in radiographic weld images using morphological approach," *NDT & E International*, vol. 39, no. 1, pp. 29 – 33, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0963869505000794 (Cited on page 62.)

[117] J.-B. Martens, "Adaptive contrast enhancement through residue-image processing," *Signal Processing*, vol. 44, no. 1, pp. 1 – 18, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0165168495000112 (Cited on page 62.)

[118] H. Boerner and H. Strecker, "Automated x-ray inspection of aluminum castings," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 1, pp. 79–91, Jan 1988. (Cited on page 62.)

[119] T. Liao and J. Ni, "An automated radiographic ndt system for weld inspection: Part i – weld extraction," *NDT & E International*, vol. 29, no. 3, pp. 157 – 162, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0963869596000096 (Cited on page 62.)

[120] H. Jia, Y. L. Murphey, J. Shi, and T.-S. Chang, "An intelligent real-time vision system for surface defect detection," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 3, Aug 2004, pp. 239–242 Vol.3. (Cited on page 62.)

[121] D. Naso, B. Turchiano, and P. Pantaleo, "A fuzzy-logic based optical sensor for online weld defect-detection," *IEEE Transactions on Industrial Informatics*, vol. 1, no. 4, pp. 259–273, Nov 2005. (Cited on page 62.)

[122] J. Blackledge and D. Dubovitskiy, "A surface inspection machine vision system that includes fractal texture analysis," vol. 3, no. 2, pp. 76–89, 2008. (Cited on page 62.)

[123] R. N. Strickland and H. I. Hahn, "Wavelet transform methods for object detection and recovery," *IEEE Transactions on Image Processing*, vol. 6, no. 5, pp. 724–735, May 1997. (Cited on page 62.)

[124] H. Y. Liao and G. Sapiro, "Sparse representations for limited data tomography," in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, May 2008, pp. 1375–1378. (Cited on page 62.)

[125] E. Konuko?lu and B. Acar, "Hdf: Heat diffusion fields for polyp detection in ct colonography," *Signal Processing*, vol. 87, no. 10, pp. 2407 – 2416, 2007, special Section: Total Least Squares and Errors-in-Variables Modeling. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168407001247 (Cited on page 62.)

[126] P. Campadelli, E. Casiraghi, and D. Artioli, "A fully automated method for lung nodule detection from postero-anterior chest radiographs," *IEEE Transactions on Medical Imaging*, vol. 25, no. 12, pp. 1588–1603, Dec 2006. (Cited on page 62.)

[127] G. Wang and T. Liao, "Automatic identification of different types of welding defects in radiographic images," *NDT & E International*, vol. 35, no. 8, pp. 519 – 528, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0963869502000257 (Cited on page 62.)

[128] R. R. da Silva, L. P. Calôba, M. H. Siqueira, and J. M. Rebello, "Pattern recognition of weld defects detected by radiographic test," *NDT & E International*, vol. 37, no. 6, pp. 461 – 470, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0963869503001786 (Cited on page 62.)

[129] H. Strecker, "A local feature method for the detection of flaws in automated x-ray inspection of castings," *Signal Processing*, vol. 5, no. 5, pp. 423 – 431, 1983. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0165168483900051 (Cited on page 62.)

[130] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. (Cited on page 62.)

[131] M. Haindl and S. Mikeš, *Model-Based Texture Segmentation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 306–313. (Cited on page 62.)

[132] D. Mery and O. Medina, *Automated Visual Inspection of Glass Bottles Using Adapted Median Filtering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 818–825. (Cited on pages 62 and 68.)

[133] C. H. Chen, *Handbook of Pattern Recognition and Computer Vision*, 4th ed. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 2010. (Cited on page 62.)

[134] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 196 – 210, 2015, infrastructure Computer Vision. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1474034615000208 (Cited on page 62.)

[135] R. Cogranne and F. Retraint, "An asymptotically uniformly most powerful test for lsb matching detection," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 464–476, March 2013. (Cited on pages 63 and 100.)

[136] R. Cogranne, C. Zitzmann, F. Retraint, I. V. Nikiforov, P. Cornu, and L. Fillatre, "A local adaptive model of natural images for almost optimal detection of hidden data," *Signal Processing*, vol. 100, no. Supplement C, pp. 169 – 185, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168414000450 (Cited on pages 63 and 68.)

[137] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1737–1754, Oct 2008. (Cited on pages 64, 65, 70, 125, 150 and 152.)

[138] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, Feb 2016. (Cited on pages 67 and 68.)

[139] H. Yin, C. Lin, B. Sebastien, B. Li, and G. Min, "Network traffic prediction based on a new time series model," *International Journal of Communication Systems*, vol. 18, no. 8, pp. 711–729, 2005. [Online]. Available: http://dx.doi.org/10.1002/dac.721 (Cited on page 67.)

[140] T. N. Nguyen, R. Cogranne, G. Doyen, and F. Retraint, "Detection of interest flooding attacks in named data networking using hypothesis testing," in *2015 IEEE International Workshop on Information Forensics and Security (WIFS)*, Nov 2015, pp. 1–6. (Cited on page 67.)

[141] I. Tosic and P. Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, March 2011. (Cited on page 67.)

[142] M. Basseville and I. Nikiforov, "Fault isolation for diagnosis: Nuisance rejection and multiple hypotheses testing," *Annual Reviews in Control*, vol. 26, no. 2, pp. 189 – 202, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1367578802000299 (Cited on pages 68 and 100.)

[143] R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, "Rate-distortion optimized tree-structured compression algorithms for piecewise polynomial images," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 343–359, March 2005. (Cited on page 68.)

[144] R. Kazinnik, S. Dekel, and N. Dyn, "Low bit-rate image coding using adaptive geometric piecewise polynomial approximation," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2225–2233, Sept 2007. (Cited on page 68.)

[145] S. H. Hanzaei, A. Afshar, and F. Barazandeh, "Automatic detection and classification of the ceramic tiles? surface defects," *Pattern Recognition*, vol. 66, no. Supplement C, pp. 174 – 189, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S003132031630379X (Cited on page 68.)

[146] T. H. Thai, R. Cogranne, and F. Retraint, "Statistical model of quantized dct coefficients: Application in the steganalysis of jsteg algorithm," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 1980–1993, May 2014. (Cited on pages 70 and 152.)

[147] ——, "Camera model identification based on the heteroscedastic noise model," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 250–263, Jan 2014. (Cited on pages 70, 80 and 130.)

[148] T. H. Thai, F. Retraint, and R. Cogranne, "Statistical detection of data hidden in least significant bits of clipped images," *Signal Processing*, vol. 98, no. Supplement C, pp. 263 – 274, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168413004635 (Cited on page 70.)

[149] ——, "Generalized signal-dependent noise model and parameter estimation for natural images," *Signal Processing*, vol. 114, no. Supplement C, pp. 164 – 170, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168415000869 (Cited on page 70.)

[150] ——, "Camera model identification based on the generalized noise model in natural images," *Digital Signal Processing*, vol. 48, no. Supplement C, pp. 285 – 297, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200415003012 (Cited on page 70.)

[151] Y. Wang, C. Xu, C. Xu, and D. Tao, "Beyond rpca: Flattening complex noise in the frequency domain," 2017. (Cited on page 70.)

[152] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang, "Robust principal component analysis with complex noise," in *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ser. ICML'14. JMLR.org, 2014, pp. II–55–II–63. [Online]. Available: http://dl.acm.org/citation.cfm?id=3044805.3044899 (Cited on page 70.)

[153] M. Fouladirad, L. Freitag, and I. Nikiforov, "Optimal fault detection with nuisance parameters and a general covariance matrix," *International Journal of Adaptive Control and Signal Processing*, vol. 22, no. 5, pp. 431–439, 2008. [Online]. Available: http://dx.doi.org/10.1002/acs.976 (Cited on pages 73 and 153.)

[154] R. Cogranne and F. Retraint, "A new tomography model for almost optimal detection of anomalies," in *2013 IEEE International Conference on Image Processing*, Sept 2013, pp. 1461–1465. (Cited on pages 73 and 153.)

[155] J. Illingworth and J. Kittler, "A survey of the hough transform," *Computer Vision, Graphics, and Image Processing*, vol. 44, no. 1, pp. 87 – 116, 1988. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0734189X88800331 (Cited on page 76.)

[156] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, Jan 1988. [Online]. Available: https://doi.org/10.1007/BF00133570 (Cited on page 77.)

[157] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2029–2039, Nov 2008. (Cited on page 77.)

[158] J. Choi and C. Kim, "Unsupervised detection of surface defects: A two-step approach," in *2012 19th IEEE International Conference on Image Processing*, Sept 2012, pp. 1037–1040. (Cited on pages 86, 87 and 88.)

[159] D.-M. Tsai, M.-C. Chen, W.-C. Li, and W.-Y. Chiu, "A fast regularity measure for surface defect detection," *Machine Vision and Applications*, vol. 23, no. 5, pp. 869–886, Sep 2012. [Online]. Available: https://doi.org/10.1007/s00138-011-0403-3 (Cited on pages 87 and 88.)

[160] F. G. Bulnes, D. F. García, F. J. de la Calle, R. Usamentiaga, and J. Molleda, "A non-invasive technique for online defect detection on steel strip surfaces," *Journal of Nondestructive Evaluation*, vol. 35, no. 4, p. 54, Sep 2016. [Online]. Available: https://doi.org/10.1007/s10921-016-0370-8 (Cited on page 88.)

[161] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, "A review and comparison of changepoint detection techniques for climate data," *Journal of Applied Meteorology and Climatology*, vol. 46, no. 6, pp. 900–915, 2007. (Cited on page 95.)

[162] G. E. Evans, G. Y. Sofronov, J. M. Keith, and D. P. Kroese, "Estimating change-points in biological sequences via the cross-entropy method," *Annals of Operations Research*, vol. 189, no. 1, pp. 155–165, Sep 2011. [Online]. Available: https://doi.org/10.1007/s10479-010-0687-0 (Cited on page 95.)

[163] T. Polushina and G. Sofronov, "A cross-entropy method for change-point detection in four-letter dna sequences," in *2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Oct 2016, pp. 1–6. (Cited on page 95.)

[164] J. Chen and A. K. Gupta, "Testing and locating variance changepoints with application to stock prices," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 739–747, 1997. (Cited on page 95.)

[165] A. Aprem and V. Krishnamurthy, "Utility change point detection in online social media: A revealed preference framework," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1869–1880, April 2017. (Cited on page 95.)

[166] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2000. (Cited on page 96.)

[167] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, ser. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 2014. [Online]. Available: https://books.google.fr/books?id=zhsbBAAAQBAJ (Cited on page 96.)

[168] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press Cambridge, 2009, vol. 40. (Cited on page 96.)

[169] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954. [Online]. Available: http://www.jstor.org/stable/2333009 (Cited on pages 98 and 103.)

[170] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, 1971. [Online]. Available: http://www.jstor.org/stable/2240115 (Cited on pages 99 and 103.)

[171] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *The Annals of Statistics*, vol. 14, no. 4, pp. 1379–1387, 1986. [Online]. Available: http://www.jstor.org/stable/2241476 (Cited on pages 99 and 103.)

[172] Y. Ritov, "Decision theoretic optimality of the cusum procedure," *The Annals of Statistics*, vol. 18, no. 3, pp. 1464–1469, 1990. [Online]. Available: http://www.jstor.org/stable/2242064 (Cited on pages 99 and 103.)

[173] C. R. Rao, C. R. Rao, M. Statistiker, C. R. Rao, and C. R. Rao, *Linear statistical inference and its applications*. Wiley New York, 1973, vol. 2. (Cited on page 100.)

[174] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, Feb 2016. (Cited on page 100.)

[175] R. Cogranne, "A sequential method for online steganalysis," in *2015 IEEE International Workshop on Information Forensics and Security (WIFS)*, Nov 2015, pp. 1–6. (Cited on pages 101 and 158.)

[176] B. K. Guépié, L. Fillatre, and I. Nikiforov, "Sequential detection of transient changes," *Sequential Analysis*, vol. 31, no. 4, pp. 528–547, 2012. (Cited on pages 101 and 158.)

[177] A. Shiryaev, "The problem of the most rapid detection of a disturbance in a stationary process," in *Soviet Math. Dokl*, vol. 2, no. 795-799, 1961. (Cited on page 103.)

[178] K. W. Kemp, "The average run length of the cumulative sum chart when a v-mask is used," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 23, no. 1, pp. 149–153, 1961. [Online]. Available: http://www.jstor.org/stable/2983850 (Cited on page 103.)

[179] N. K. Akafuah, S. Poozesh, A. Salaimeh, G. Patrick, K. Lawler, and K. Saito, "Evolution of the automotive body coating process – a review," *Coatings*, vol. 6, no. 2, 2016. [Online]. Available: http://www.mdpi.com/2079-6412/6/2/24 (Cited on pages 107 and 108.)

[180] N. K. Akafuah, *Automotive Paint Spray Characterization and Visualization*. Dordrecht: Springer Netherlands, 2013, pp. 121–165. [Online]. Available: https://doi.org/10.1007/978-94-007-5095-1_5 (Cited on pages 107 and 108.)

[181] M. Ottavian, M. Barolo, and S. García-Muñoz, "Maintenance of machine vision systems for product quality assessment. part i. addressing changes in lighting conditions," *Industrial & Engineering Chemistry Research*, vol. 52, no. 35, pp. 12 309–12 318, 2013. [Online]. Available: http://dx.doi.org/10.1021/ie303295t (Cited on pages 121 and 161.)

[182] I. Jahr, *Lighting in Machine Vision*. Wiley-VCH Verlag GmbH & Co. KGaA, 2007, pp. 73–203. [Online]. Available: http://dx.doi.org/10.1002/9783527610136.ch3 (Cited on pages 121 and 161.)

[183] M.-H. Chang, D. Das, P. Varde, and M. Pecht, "Light emitting diodes reliability review," *Microelectronics Reliability*, vol. 52, no. 5, pp. 762 – 782, 2012, reliability of High-Power LED Packaging and Assembly. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0026271411003283 (Cited on page 122.)

[184] J. Fan, K. C. Yung, and M. Pecht, "Lifetime estimation of high-power white led using degradation-data-driven method," *IEEE Transactions on Device and Materials Reliability*, vol. 12, no. 2, pp. 470–477, June 2012. (Cited on page 122.)

[185] B. Hamon and W. van Driel, "Led degradation: From component to system," *Microelectronics Reliability*, vol. 64, no. Supplement C, pp. 599 – 604, 2016, proceedings of the 27th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0026271416301585 (Cited on page 122.)

[186] W. van Driel, M. Schuld, B. Jacobs, F. Commissaris, J. van der Eyden, and B. Hamon, "Lumen maintenance predictions for led packages," *Microelectronics Reliability*, vol. 62, no. Supplement C, pp. 39 – 44, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002627141630052X (Cited on page 122.)

[187] S. H. Park and J. H. Kim, "Lifetime estimation of led lamp using gamma process model," *Microelectronics Reliability*, vol. 57, no. Supplement C, pp. 71 – 78, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0026271415302547 (Cited on page 122.)

[188] R. Ramanath, W. E. Snyder, Y. Yoo, and M. S. Drew, "Color image processing pipeline," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 34–43, Jan 2005. (Cited on page 122.)

[189] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 1, 1999, p. 380 Vol. 1. (Cited on page 123.)

[190] M. D. Grossberg and S. K. Nayar, "Determining the camera response from images: what is knowable?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1455–1467, Nov 2003. (Cited on page 123.)

[191] ——, "Modeling the space of camera response functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1272–1282, Oct 2004. (Cited on page 123.)

[192] Y. W. Tai, X. Chen, S. Kim, S. J. Kim, F. Li, J. Yang, J. Yu, Y. Matsushita, and M. S. Brown, "Nonlinear camera response functions and image deblurring: Theoretical analysis and practice," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2498–2512, Oct 2013. (Cited on page 124.)

[193] G. Joffre, W. Puech, F. Comby, and J. Joffre, "High dynamic range images from digital cameras raw data," in *ACM SIGGRAPH 2005 Posters*, ser. SIGGRAPH '05. New York, NY, USA: ACM, 2005. [Online]. Available: http://doi.acm.org/10.1145/1186954.1187036 (Cited on page 124.)

[194] H. Farid, "Blind inverse gamma correction," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1428–1433, Oct 2001. (Cited on page 124.)

[195] Y. Tsin, V. Ramesh, and T. Kanade, "Statistical calibration of ccd imaging process," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 1, 2001, pp. 480–487 vol.1. (Cited on page 124.)

[196] C. Ding, "An efficient algorithm for computing the noncentrality parameters of chi-squared tests," *Communications in Statistics - Simulation and Computation*, vol. 23, no. 3, pp. 861–870, 1994. [Online]. Available: https://doi.org/10.1080/03610919408813204 (Cited on pages 128 and 162.)

[197] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," *Journal of Manufacturing Systems*, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0278612518300037 (Cited on page 141.)